Las Integrales de Riemann, Lebesgue

y

Henstock-Kurzweil

Wilman Brito

Dedicatoria

A mis Alejandros:
Sebastian y *Rubén*.

Prólogo

Con el transcurrir del tiempo, un hermoso jardín de integrales se ha ido sembrando paulatinamente en la amplia y fértil tierra de las matemáticas. Esta increíble y rica diversidad de integrales no cesa. Un poco más de 100 integrales han crecido, hasta el momento, en ciertas parcelas de ese hermoso terreno, cada una de ellas con uno o varios nombres que las identifican. Mencionemos, por ejemplo, las integrales de Newton, Cauchy, Riemann, Darboux, Harnack, Cauchy-Riemann, Lebesgue, Stieltjes, Riemann-Stieltjes, Lebesgue-Stieltjes, Denjoy, Perron, Henstock-Kurzweil, McShane, C-integral, Pfeffer, BV-integral, Haar, Radon, Daniell, Burkill, Itô, Hellinger, Kolmogoroff, Khinchin, Bochner, Dunford, Pettis, Bartle, Gelfand, Wiener, Feynman, Birkhoff, Dinculeanu, Dobrakov, Dunford-Pettis, Gelfand-Pettis, Bartle-Dunford, Henstock-Kurzweil-Dunford, Henstock-Kurzweil-Pettis, Daniell-Bourbaki, Denjoy-McShane, Denjoy-Bochner, Denjoy-Pettis, etc. Unas tienen un carácter histórico como las de Newton y Cauchy, algunas fueron absorbidas por otras que son más nuevas y mejores, otras son maneras o formas equivalentes de expresar una misma integral, etc. ¿Por qué tantas integrales? Pues bien, existen, fundamentalmente, varios aspectos a considerar: algunas de ellas nacieron de un problema particular, otras surgieron de una necesidad de generalizar una integral a contextos más amplios o abstractos, algunas otras deben su existencia a un análisis profundo y exhaustivo de una integral modificando algún aspecto de la misma, etc. Por ejemplo, la integral de Riemann es una generalización de la de Cauchy y nace de un problema concreto que Riemann quería resolver, pero es equivalente a la de Darboux; la integral de Lebesgue generaliza a la de Riemann pues surge de las deficiencias y limitaciones que esta última posee, pero es equivalente a las integrales de Daniel y de McShane. Las integrales de Denjoy, Perron y Henstock-Kurzweil se crean a partir de la necesidad de resolver el problema de las primitivas propuesto por Newton-Leibniz. Posteriormente se demostró que ellas tres eran equivalentes y contienen, en su interior, a la integral de Lebesgue. A su vez, la integral de McShane, cuya definición es muy similar a la de Henstock-Kurzweil, se obtiene por medio de una interrogante y que resulta ser equivalente a la de Lebesgue. Las integrales de Bochner, Dunford y Pettis fueron diseñadas para trabajar en espacios de Banach, mientras que la de Haar se desarrolla en grupos topológicos localmente compactos. Por otro lado, la de Itô se genera a partir de los procesos estocásticos asociados a movimientos Brownianos, etc. Los libros de Frank E. Burk, "**A Garden of Integrales**" [28], Stefan Schwabik y Ye Guoju, "**Topics in Banach Space Integration**", [119], Ivan N. Pesin, "**Classical and Modern Integration Theories**" [109], el artículo de T. H. Hildebrandt "*Integration in Abstract Spaces*, etc., poseen abundante y buena información de algunas de ellas.

En estas notas desarrollaremos sólo tres de estas integrales, las que aparecen en el título. La integral de Riemann, que se mantiene y continúa enseñándose a pesar de los años transcurridos desde su creación, será breve. En su presentación sólo mostraremos alguna de sus propiedades elementales pasando por el formidable Teorema Fundamental de Vitali-Lebesgue que describe cómo son, en realidad, las funciones Riemann integrables, pero haciendo mucho énfasis en las deficiencias que posee dicha integral con el sólo propósito de justificar el por qué Lebesgue construyó su integral. Por otro lado, la integral de Lebesgue se desarrolla más ampliamente que las otras dos por las siguientes razones: primero, debemos fabricar gran parte del aparataje de la Teoría de la Medida y de las Funciones Medibles que son necesarias para la construcción de dicha integral, este enfoque consume una gran parte del texto; luego mostramos sus poderosos teoremas de convergencia y finalmente extendemos brevemente esa noción de integral a un contexto totalmente abstracto. Todo ello conduce a la creación de una integral que, aunque es bastante complicada en su construcción, es totalmente superior a la integral de Riemann en todo sentido. Su amplío abanico de aplicaciones justifican, con creces, ese gran esfuerzo en su construcción. Existen otras integrales equivalentes a la integral de Lebesgue que evitan el uso de la Teoría de la Medida tales como las integrales de Daniell, de Mikusiński y la de McShane. Sin embargo, es importante aclararlo, la Teoría de la Medida Exterior que conlleva a la noción de Conjunto Medible y luego a la de Medida es importante en sí misma y no sólo por el hecho de servir como un puente en la construcción de la integral de Lebesgue. A pesar de las inmensas bondades que posee la integral de Lebesgue, ella no está exenta de sus propias deficiencias: por ejemplo, la integral de Lebesgue no es capaz de integrar a todas las funciones derivadas, en otras palabras, ella no satisface el Teorema Fundamental del Cálculo en toda su generalidad; tampoco puede integrar funciones que poseen en algún punto de su dominio una "fuerte oscilación". Integrales de la forma $\int_0^\infty \operatorname{sen}(x^2)\,dx$ son elusivas a la integral de Lebesgue, etc. En cambio, con la integral de Henstock-Kurzweil, que es una extensión propia de la integral de Lebesgue, se subsanan algunas de esas deficiencias. Por ejemplo, el Teorema Fundamental del Cálculo para la integral de Henstock-Kurzweil se cumple para toda función derivada, cualquier función Lebesgue integral es Henstock-Kurzweil integrable con integrales iguales y también posee los poderosos teoremas de convergencia válidos para la integral de Lebesgue. Además, otro punto a su favor que es muy importante, es que dicha integral no usa la Teoría de la Medida en su construcción y muy poco de ella en su desarrollo posterior. Cuando intentamos comparar las integrales de Lebesgue con la integral de Henstock-Kurzweil, tenemos que admitir que hay bondades en ambos lados y, por supuestos, sus respectivas deficiencias. Por ejemplo, la integral de Henstock-Kurzweil no es una integral absoluta, es decir, si f es Henstock-Kurzweil integrable, entonces no es cierto, en general, que su valor absoluto $|f|$ sea Henstock-Kurzweil integrable cosa que si ocurre con la integral de Lebesgue. Este hecho impide que se pueda desarrollar una teoría $\mathcal{HK}_p([a,b])$ similar a la teoría de los espacios $\mathrm{L}_p([a,b])$ para $p \in [1,+\infty]$. Sin embargo, a pesar de no ser la integral de Henstock-Kurzweil una integral absoluta este hecho no es, en lo absoluto, una calamidad, sino más bien, constituye, en ciertos aspectos, una enorme fortaleza de dicha integral. Por otro lado, no existe una extensión canónica de la integral de Henstock-Kurzweil a espacios abstractos, de modo que la busqueda de una integral perfecta aun continua.

Ahora detallaremos brevemente cómo hemos organizado el contenido de estas extensas notas. Los capítulos que van del 1 al 4 constituyen los recordatorios básicos que necesitaremos para los restantes capítulos: casi todas las demostraciones de los resultados requeridos en el desarrollo de estas notas se encuentran en dichos recordatorios de modo que el lector no tendrá que dejar su lectura para ir a la busqueda de otro libro para consultar la demostración de un resultado en particular. Sin embargo, el lector está en pleno derecho de saltarse esos capítulos si considera que tales conocimientos no le son ajenos y comenzar desde el capítulo 5 para volver la mirada hacia atrás cada vez que sea necesario recordar un enunciado y(o) su prueba de un resultado particular. Los capítulos que comienzan desde 5 hasta el 10 tratan sobre la medida y la integral de Lebesgue en $\mathbb{R}$. Aunque la integral de Lebesgue abarca más de la mitad del libro, queda por fuera, sin embargo, un inmenso caudal de conocimientos relativos a tal integral. Libros tales como [111, 14, 60, 106] etc. tratan muchos otros tópicos que no mencionamos ni consideramos en estas notas. Los capítulos 11 y 12 constituyen ciertas consideraciones abstractas de la medida e integral de Lebesgue. El capítulo 13 es interesante ya que desarrolla algunos importantes teoremas sobre la convergencia de medidas. Los resultados demostrados en este capítulo no son utilizados en estas notas por lo que el lector, sino está interesados en ellos en ese momento, puede evitarlos en una primera lectura. Finalmente, el capítulo 14 es una incursión a una integral (o varias integrales) que es fantástica por todos lados: contiene a la integral de Lebesgue, integra cualquier derivada y no usa la teoría de la medida de Lebesgue en su construcción. Ella es la integral de Henstock-Kurzweil. Una casi febril investigación se ha desarrollado en los últimos tiempos en torno a esta integral y sus posibles generalizaciones. Aquí nos restringimos al estudio de funciones integrables según Henstock-Kurzweil cuyo dominio es un intervalo $[a, b] \subseteq \mathbb{R}$. Los libros de R. A. Gordon [66], D. S. Kurz y C. W. Swartz [84], C. W. Swartz [125], R. G. Bartle [38], A. G. Das [40] y muchos otros introducen al lector al estudio de esta tres integrales haciendo menor énfasis en las dos primeras, pero dedicándole más espacio a la integral de Henstock-Kurzweil o integral de Henstock como la llama Gordon. El capítulo finaliza con una breve exposición de las definiciones de las integrales de Denjoy y Perron que son equivalentes a la integral de Henstock-Kurzweil y la integral distribucional de Denjoy que es más general que las anteriores.

A modo de Advertencia: Tratándose de una versión preliminar, este trabajo contendrá, con toda seguridad, un montón de errores, omisiones, demostraciones medio sospechosas, otras incompletas, insuficiencia de ejercicios y algunas otras cosas indeseables. Por tal motivo el lector, en plena posesión de sus facultades, si acepta leer cualquier parte del mismo y se tropieza con algunos de los errores u omisiones que se encuentran en él, se compromete a reportarlo a mi persona y también puede sugerir, si lo cree necesario, otros caminos y vías más adecuadas para una mejor y futura presentación.

wbrito@ula.ve

britow78@gmail.com

Índice general

Prólogo v

1. Preliminares 1
1.1. **Un poco de Teoría de Conjuntos** 1
1.1.1. Conjuntos 2
1.1.2. Funciones 6
1.1.3. Familias Indexadas. Productos Cartesianos 9
1.2. **Conjuntos Numerables y otros más Numerosos** 10
1.2.1. Ejemplos de Conjuntos Numerables 18
1.2.2. El Teorema de Cantor y Conjuntos No-numerables 41
1.2.3. Ejemplos de Conjuntos No-numerables 44
1.2.4. Un Juego y la No-numerabilidad de $\mathbb{R}$. 46
1.3. **El Axioma de Elección y sus Aliados** 47
1.3.1. El Axioma de Elección 47
1.3.2. El Lema de Zorn 49
1.3.3. Principio del Buen-Orden 53
1.3.4. Números Ordinales 57
1.3.5. Números Cardinales 67
1.3.6. $\aleph_1$ y el Primer Ordinal No-numerable 68
1.3.7. La Aritmética de los Cardinales 69
1.3.8. La Cardinalidad de $\mathbb{R}$ y otros Conjuntos Similares 73
1.3.9. La Hipótesis del Continuo 76
1.3.10. Conjunto de Bernstein 78
1.4. **Problemas** 80

2. Los Números Reales 83
2.1. **Algunas Propiedades de los Números Reales** 83
2.1.1. Principio de Arquímedes 83
2.1.2. Conjuntos Acotados 85
2.1.3. Límites 91
2.1.4. El Teorema de Bolzano-Weierstrass 92
2.1.5. Los Números Reales Extendidos 95

2.1.6. Limites Superior e Inferior de una Sucesión . . . 98
2.1.7. Límites Superior e Inferior de Conjuntos . . . 105
2.1.8. Series Absolutamente Convergentes y Familias Sumables . . . 107
2.1.9. Caracterizando Series Absolutamente Convergentes . . . 109
2.1.10. Familias Sumables . . . 114
2.2. **Espacios Topológicos** . . . 119
2.2.1. Espacios Métricos . . . 126
2.2.2. El Teorema de Categoría de Baire . . . 135
2.2.3. Espacios Normados . . . 144
2.2.4. La Topología Producto . . . 146
2.2.5. El Espacio de Baire . . . 150
2.3. **Problemas** . . . 152

3. Funciones Continuas **155**
3.1. **Propiedades Básicas** . . . 155
3.1.1. Funciones Continuas con Soportes Compactos . . . 157
3.1.2. Más sobre Funciones Continuas . . . 161
3.1.3. Oscilación y Discontinuidad de una Función . . . 167
3.1.4. Convergencia de Sucesiones de Funciones . . . 175
3.1.5. Una Función Continua Nunca Diferenciable . . . 179
3.1.6. Funciones Semicontinuas . . . 183
3.1.7. Convergencia Puntual en $\mathrm{Sc}([a,b])$ . . . 187
3.1.8. Funciones Acotadas en $\mathrm{Sc}([a,b])$ . . . 189
3.2. **Problemas** . . . 191

4. Desigualdades de Hölder y Minkowski en $\mathbb{R}^n$ **193**
4.1. **Convexidad** . . . 193
4.1.1. Las Desigualdades AM-GM . . . 195
4.1.2. Las Desigualdades de Hölder y Minkowski . . . 199
4.2. **Problemas** . . . 204

5. El Conjunto de Cantor y su Media Hermana **209**
5.1. **Representaciones Ternarias y Binarias** . . . 209
5.1.1. Representaciones Ternarias . . . 210
5.1.2. Representaciones Binarias . . . 213
5.1.3. El Conjunto Ternario de Cantor . . . 214
5.1.4. Propiedades del Conjunto Ternario de Cantor . . . 218
5.1.5. Conjuntos Tipo-Cantor de Medida Cero . . . 226
5.1.6. Conjuntos Tipo-Cantor de Medida Positiva . . . 229
5.1.7. La Función de Cantor . . . 232
5.2. **Problemas** . . . 236

6. La Medida de Lebesgue en $\mathbb{R}$ **239**
6.1. **Introducción** . . . 239
6.2. **La Medida Exterior de Lebesgue** . . . 240
6.2.1. Condiciones bajo la cual μ^* es σ-aditiva . . . 250
6.2.2. Conjuntos de Contenido Cero . . . 253
6.3. **La Medida de Lebesgue** . . . 255

6.3.1. La σ-álgebra de Lebesgue . . . 255
6.3.2. La σ-álgebra de Borel . . . 263
6.3.3. La Cardinalidad de la σ-álgebra de Borel . . . 268
6.3.4. Conjuntos Analíticos . . . 271
6.3.5. La Medida de Lebesgue . . . 279
6.3.6. El Lema de Borel-Cantelli . . . 287
6.3.7. Criterios de Medibilidad . . . 293
6.3.8. Medida Interior . . . 300
6.4. **Conjuntos Medibles con Propiedades Especiales** . . . 304
6.5. **Conjuntos no-medibles** . . . 309
6.5.1. Conjunto de Vitali . . . 310
6.5.2. Conjunto no-medible de un Grupo Aditivo . . . 318
6.5.3. Conjunto Saturado no-medible . . . 320
6.5.4. Conjunto de Bernstein . . . 323
6.5.5. Conjunto de Sierpiński . . . 324
6.5.6. Ultrafiltros no-medibles . . . 326
6.5.7. Conjunto de Lévy . . . 332
6.6. **Notas Breves sobre El Problema de la Medida** . . . 344
6.6.1. El Problema de la Medida de Lebesgue y El Axioma de Elección . . . 344
6.6.2. El Problema de la Medida de Lebesgue y la Hipótesis del Continuo . . . 346
6.6.3. El Problema de la Medida de Lebesgue y la Aditividad Finita . . . 350
6.6.4. El Problema de la Medida de Lebesgue y el Axioma de Determinación . . . 352
6.7. **Ejemplos y Contraejemplos Usando la Función de Cantor** . . . 354
6.8. **Ejercicios Resueltos** . . . 358
6.9. **Problemas** . . . 364

7. Funciones Medibles **371**
7.1. **Definición** . . . 371
7.2. **Propiedades Básicas** . . . 377
7.2.1. Aproximación de Funciones Medibles . . . 382
7.2.2. Los Teoremas de Severini-Egoroff y de Lusin . . . 386
7.2.3. Convergencia en Medida . . . 394
7.3. **Ejercicios Resueltos** . . . 398
7.4. **Problemas** . . . 403

8. La Integral de Riemann **405**
8.1. **Introducción** . . . 405
8.2. **La Integral de Newton** . . . 407
8.3. **Construcción de la Integral de Riemann** . . . 409
8.3.1. La Integral de Darboux - Su Construcción . . . 411
8.3.2. Equivalencia de las Integrales de Riemann y Darboux . . . 416
8.3.3. El Teorema de Vitali-Lebesgue . . . 418
8.3.4. Consecuencias del Teorema de Vitali-Lebesgue . . . 423
8.3.5. Propiedades Básicas de la Integral de Riemann . . . 426
8.3.6. El Teorema Fundamental del Cálculo . . . 429
8.3.7. Limitaciones y Deficiencias de la Integral de Riemann . . . 436
8.4. **Ejercicios Resueltos** . . . 447

8.5. **Problemas** . . . 451

9. Diferenciación y un Teorema de Lebesgue **453**
9.1. **Funciones Absolutamente Continuas y de Variación Acotada** . . . 453
9.1.1. Funciones Lipschitz y la condición (N) de Lusin . . . 454
9.1.2. Funciones de Variación Acotada . . . 461
9.1.3. Cubrimientos de Vitali . . . 472
9.1.4. El Teorema de Densidad de Lebesgue . . . 478
9.1.5. El Teorema de Diferenciabilidad de Lebesgue . . . 483
9.1.6. Funciones Absolutamente Continuas . . . 498
9.1.7. El Teorema de Banach-Zarecki . . . 512
9.2. **Ejercicios Resueltos** . . . 517
9.3. **Problemas** . . . 523

10. La Integral de Lebesgue **527**
10.1. **Construcción de la Integral de Lebesgue** . . . 528
10.1.1. Propiedades de la Integral de Lebesgue . . . 532
10.1.2. Los Poderosos Teoremas de Convergencias . . . 540
10.1.3. El Teorema Fundamental del Cálculo de Lebesgue . . . 545
10.1.4. La Integral de Lebesgue via Funciones Simples . . . 546
10.1.5. Integrales Dependiendo de un Parámetro . . . 568
10.1.6. La Integral de Lebesgue sin Medida . . . 571
10.2. **El Espacio** $\mathrm{L}_1(X,\mu)$ . . . 576
10.2.1. Densidad en el espacio $\mathrm{L}_1([a,b])$ . . . 586
10.2.2. El Lema de Riemann-Lebesgue . . . 590
10.2.3. La Completación del Espacio $\mathcal{R}([a,b])$ . . . 592
10.2.4. Conjuntos Uniformemente Integrables . . . 593
10.2.5. Los Teoremas de Convergencia de Vitali . . . 601
10.2.6. El TFC para la Integral de Lebesgue . . . 608
10.2.7. El Teorema de Vitali-Carathéodory . . . 619
10.2.8. Regla de la Cadena e Integración por Partes . . . 624
10.2.9. Cambio de Variable para la Integral de Lebesgue . . . 625
10.2.10.Puntos de Lebesgue de una Función Integrable . . . 628
10.3. **Los Espacios** $\mathrm{L}_p(X,\mu),\ 1 < p \leq +\infty$ . . . 630
10.3.1. Las Desigualdades de Hölder y Minkowski en $\mathrm{L}_p(X,\mu)$ . . . 630
10.3.2. Convergencia Fuerte y Débil en L_p, $1 < p < +\infty$ . . . 635
10.3.3. La inclusión $\mathrm{L}_q \subseteq \mathrm{L}_p$ para $1 \leq p < q$ . . . 640
10.3.4. Conjuntos Uniformemente Integrables en L_p para $p > 1$ . . . 643
10.3.5. Densidad en los Espacios $\mathrm{L}_p(X,\mu)$ . . . 645
10.3.6. Separabilidad de los Espacios $\mathrm{L}_p(\mathbb{R},\mu),\ p \in [1,+\infty)$ . . . 646
10.3.7. El Espacio $\mathrm{L}_\infty(X,\mu)$ . . . 647
10.3.8. Convolución en $\mathrm{L}_p(\mathbb{R},\mu)$ . . . 655
10.4. **Ejercicios Resueltos** . . . 665
10.5. **Problemas** . . . 674

11. Medida e Integración Abstracta **679**
11.1. **Espacios de Medidas** 679
11.1.1. Medidas sin Átomos 686
11.1.2. Completación de una Medida 689
11.1.3. El Teorema de Extensión de Carathéodory 692
11.1.4. La Medida de Lebesgue-Stieltjes en $\mathbb{R}$ 700
11.1.5. Funciones de Variación Acotada sobre $\mathbb{R}$ 709
11.1.6. Funciones Medibles 710
11.1.7. Funciones Integrables 712
11.2. **Medida Producto y el Teorema de Fubini** 715
11.2.1. Clases Monótonas 715
11.2.2. Medida Producto y el Teorema de Fubini 718
11.3. **La Medida de Lebesgue en $\mathbb{R}^n$** 734
11.3.1. Cambio de Variable: caso lineal 747
11.3.2. Cambio de Variable: caso no-lineal 753
11.3.3. El Teorema de Sard 765
11.4. **Medida de Borel sobre el Espacio de Cantor $2^{\mathbb{N}}$** 767
11.4.1. σ-álgebra Producto 768
11.4.2. Una Métrica sobre $2^{\mathbb{N}}$ 771
11.4.3. Una Medida sobre $2^{\mathbb{N}}$ 772
11.5. **Ejercicios Resueltos** 774
11.6. **Problemas** 778

12. El Teorema de Radon-Nikodým **781**
12.1. **Medidas con Signos e Integración** 782
12.1.1. El Teorema de Drewnowski 796
12.1.2. Integración Respecto a una Medida con Signo 798
12.1.3. El Teorema de Radon-Nikodým 800
12.2. **Aplicaciones del Teorema de Radon-Nikodým** 810
12.2.1. Una Identidad en $L_1(X, \mathcal{M}, \nu)$ 810
12.2.2. El Teorema de Descomposición de Lebesgue 810
12.2.3. El Teorema de Representación de Riesz - El Dual de $L_p(X, \nu)$, $1 \leq p < +\infty$ 811
12.2.4. Existencia de la Esperanza Condicional 816
12.2.5. Unicidad de la Medida de Lebesgue 819
12.3. **Diferenciación de Medidas** 820
12.3.1. La Función Maximal de Hardy-Littlewood 823
12.3.2. El Teorema de Hardy-Littlewood 824
12.3.3. El Teorema de Diferenciabilidad de Lebesgue 826
12.4. **Problemas** 829

13. Convergencia en ca$(\mathcal{M}, \mathbb{R})$ **833**
13.1. **Convergencia según Antosik-Mikusiński** 834
13.2. **Medidas Uniformemente σ-aditivas** 837
13.3. **Los Teoremas de Nikodým y el de Vitali-Hahn-Saks** 844

14. La Integral de Henstock-Kurzweil **849**
14.1. **Construcción de la Integral de Henstock-Kurzweil** 850
14.1.1. El Teorema Fundamental del Cálculo 858
14.1.2. Propiedades Básicas 863
14.1.3. El Lema de Saks-Henstock 868
14.1.4. El Segundo Teorema Fundamental del Cálculo 875
14.1.5. Integrabilidad Absoluta 876
14.1.6. La clase $\mathcal{L}_{\mathrm{HK}}([a,b])$ 881
14.1.7. Comparando Integrabilidad: Lebesgue y Henstock-Kurzweil 883
14.1.8. Los Teoremas de Convergencia en $\mathcal{HK}([a,b])$ 888
14.1.9. La Norma de Alexiewicz 896
14.2. **La Integral de McShane** 897
14.3. **La C-integral** 904
14.4. **Las Integrales de Denjoy, Perron y Distribucional** 908
14.4.1. La Integral de Denjoy 908
14.4.2. La Integral de Perron 910
14.4.3. La Integral de Denjoy Distribucional 912
14.5. **Problemas** 917

Bibliografía **918**

Índice Alfabético **925**

CAPÍTULO 1

Preliminares

Este capítulo describe algunas de las herramientas básicas que se van a requerir en el transcurso de estas notas. Hemos tratado de incluir la casi totalidad de los resultados que se requieren para desarrollar las tres integrales que aparecen en el título, lo cual ha permitido que la longitud de estas notas sea extremadamente larga, pero garantizándole al lector que no recurrirá a otro texto para la demostración de los resultados utilizados.

1.1. Un poco de Teoría de Conjuntos

En esta sección revisaremos de manera sucinta algunas nociones básicas de la Teoría de Conjuntos la cual constituye la base de las matemáticas modernas. El padre fundador de tan fascinante teoría fue Georg Ferdinand Ludwig Philipp Cantor (1845-1918). A partir de 1874 y por más de 30 años Cantor desarrolla de manera intuitiva una teoría general de conjuntos haciendo énfasis en los conjuntos que poseen infinitos miembros. Con la noción de conjunto infinito, Cantor logra sacar de la oscuridad dicho concepto llevándolo a extremos inconcebibles: crea una jerarquía infinita y creciente de infinitos. Muchas de sus ideas chocaron con una resistencia férrea de parte de prominentes matemáticos como Leopold Kronecker (1823-1891) quien afirmaba:

> "*No se qué predomina en la teoría de Cantor: filosofía o teología, pero de lo que sí estoy seguro es que allí no hay matemática.*"

A pesar de las críticas recibidas sobre su incipiente teoría, Cantor logra obtener el respaldo de muchos matemáticos, en especial de uno de los más brillante, prolífico y respetado del momento: David Hilbert (1862-1943) quien afirmó, de modo premonitorio, lo siguiente:

> "*Del Paraíso creado por Cantor para nosotros, nadie podrá expulsarnos.*"

Esta afirmación es compartida por Paul Cohen quien afirma:

> "*Todos coinciden, aun si se cree o no que la Teoría de Conjuntos se refiere a una realidad existente, en que hay una belleza en su sencillez y en su ámbito de aplicación*"

Más aun, en su libro Naive Set Theory, Paul R. Halmos hace notar que:

> "*Los matemáticos están de acuerdo en que cada uno de ellos debe saber algo de Teoría de Conjuntos; el desacuerdo comienza al tratar de decidir qué tanto es algo*"

Por otro lado, Edwin Hewitt y Karl Stromber afirman, al comienzo del Capítulo 1 en su libro Real and Abstract Analysis:

> "*Desde el punto de vista de un lógico, las matemáticas son la Teoría de Conjuntos y sus consecuencias. Para el analista, los conjuntos y conceptos definidos inmediatamente a partir de ellos son herramientas esenciales, y la manipulación de conjuntos es una operación que debe llevar a cabo continuamente*."

1.1.1. Conjuntos

Para hacer matemáticas superiores se requiere de una Teoría de Conjuntos robusta, práctica y conveniente. Dos de los más importantes sistemas de axiomas con los cuales se pueden crear tal Teoría de Conjuntos y que han permitido desarrollar casi toda la matemática existente hasta el presente son: la que se basa en la Axiomática de Zermelo-Fraenkel (**ZF**) cuyos creadores fueron Ernst Zermelo (1871-1953) y Abraham Fraenkel (1891-1965). Por lo general, a tal axiomática se le añade un axioma adicional conocido como el Axioma de Elección (**ZFC**), y la otra es la Teoría de Conjuntos sustentada sobre la Axiomática de Zermelo-Fraenkel-von Neumann-Bernays-Gödel (**ZFNBG**). En la primera axiomática, los conceptos primitivos corresponden a las ideas intuitivas de "*conjunto*" y "*pertenencia*", mientras que en la segunda se parte de las nociones de "*clase*" y "*pertenencia*". En ésta última teoría un conjunto es, por definición, una clase la cual es un miembro de alguna otra clase, pero donde existen clases que no son conjuntos (véase, [55], [108]). Una buena justificación para optar por cualquiera de las dos axiomatizaciones es que las Teorías de Conjuntos que se construye con ellas permiten un desarrollo adecuado del sistema de los números reales, incluyendo sus operaciones aritméticas así como las demostraciones de sus propiedades. También el Análisis, la Topología, el Álgebra y, en general, casi todas las otras ramas de la matemática han podido ser desarrolladas gracias a dichas teorías. Puesto que la existencia de clases que no son conjuntos sólo aparece una sola vez en estas notas, hemos optado por adoptar la Teoría de Conjuntos que se construye con el sistema (**ZFC**). En esta sección no describiremos explícitamente la totalidad de los Axiomas de Zermelo-Fraenkel, tan sólo nos ocuparemos de formular ciertas definiciones y operaciones usuales de conjuntos con las que trabajaremos y generar algunas de sus consecuencias. Referencias donde se pueden estudiar tales axiomas y muchas de sus consecuencias son, por ejemplo, [127], [55], [108], [70], [74], etc.

Comúnmente, un **conjunto** se describe como una colección (o reunión, o agrupación, etc) de objetos de cualquier naturaleza llamados los **elementos** o **miembros** del conjunto pero evitando definir lo que es una colección o lo que es un objeto con el sólo propósito de no incurrir en un circulo vicioso. Por tal motivo, los términos "conjunto" y "elemento" permanecerán sin ser definidos y serán aceptados como entidades fundamentales confiando en que el lector posee una noción, o sentimiento intuitivo, de lo que es un "conjunto" y lo que es "elemento de un conjunto". Los elementos que pertenecen o forman parte de un conjunto particular, digamos X, serán denotados por el símbolo "$x \in X$" que se lee: "x es un **elemento** o **miembro** de X", o también se dirá que "x **pertenece** a X." Análogamente, el enunciado "$x \notin X$" significa que "x **no pertenece** a X", o bien que "x **no es un miembro o elemento** de X".

En general, usaremos letras minúsculas tales como $a, b, c, \ldots, x, y, z$ para indicar los miembros o elementos de un conjunto, y letras mayúsculas A, B, C, $\ldots$, X, Y, Z, $\mathfrak{A}, \mathfrak{B}$, $\ldots$, $\mathcal{A}, \mathcal{B}, \mathcal{C}$, etc., para designar conjuntos. Si los elementos de un conjunto son a su vez conjuntos (los cuales serán representados por letras mayúsculas), entonces dicho conjunto será llamado una **familia**, o una

colección de conjuntos e indicado con una letra tipo gótica $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \ldots$, o tipo caligrafía $\mathcal{A}, \mathcal{B}, \mathcal{C}, \ldots$ Como siempre, usaremos el símbolo $\mathbb{N}$ para denotar el conjunto de los números naturales, esto es, $\mathbb{N} = \{1, 2, 3, \ldots\}$, mientras que $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{I}$ y $\mathbb{R}$ representan, respectivamente, el conjunto de los números enteros, los números racionales, los números irracionales y los números reales.

Una de las ideas básicas de conjuntos es la siguiente.

Definición 1.1.1. *Sean A y B conjuntos. Diremos que A es un* ***subconjunto*** *de B si todo elemento de A pertenece al conjunto B.*

Escribiremos $A \subseteq B$ o $A \supseteq B$ para denotar que A es un subconjunto de B. En ocasiones, en lugar de decir que "A es un *subconjunto de* B", diremos que "A está **incluido en** B". La negación de $A \subseteq B$, en notación $A \nsubseteq B$, y que se expresa diciendo que A **no es un subconjunto** de B, significa que *existe al menos un elemento de A que no es miembro de B*.

Un método usual de obtener subconjuntos de un conjunto dado es el siguiente: se parte de un *conjunto* X y se considera una propiedad $\mathbf{P}(x)$ referente a los elementos $x \in X$ la cual puede o no ser cierta para algunos de sus miembros. En este sentido, cualquier conjunto de la forma

$$A = \{x \in X : \mathbf{P}(x) \text{ es verdadera}\} \tag{1}$$

define un subconjunto de X. ¿Qué ocurre si a la propiedad $\mathbf{P}$ no se le impone ningún tipo de limitaciones? Por ejemplo, suponga que aceptamos la siguiente "*idea ingenua*":

Axioma de Abstracción. *Dada* ***cualquier propiedad*** P *existe un conjunto cuyos elementos son aquellos que poseen la propiedad dada. De modo más formal,*

$$(\exists\, X)(\forall\, x)[x \in X \Leftrightarrow \mathbf{P}(x)].$$

Una consecuencia lógica que se deriva de la aceptación del Axioma de Abstracción es la existencia del *conjunto de todos los conjuntos*. En efecto, basta considerar la propiedad $\mathbf{P}(x)$ como la afirmación: "*x es un conjunto*" para obtener tal conjunto. Denotemos por $\mathfrak{U}$ la colección de todos los conjuntos. Lo que Russell demostró, con un argumento enteramente elemental, es que $\mathfrak{U}$, como conjunto, no existe, originándose con ello la así llamada "Paradoja de Russell". Pero, ¿qué es una paradoja? Pues bien, una paradoja implica, a menudo, un argumento muy convincente que conduce a una conclusión errónea que parece correcta, o a una conclusión correcta que parece incorrecta o sorprendente. En términos sencillos, una *paradoja* es un razonamiento en doble sentido: supone la verdad de algo y concluye su falsedad. Similarmente, si supone su falsedad entonces se llega a su verdad. Entre 1893 y 1903, Friedrich Ludwig Gottlob Frege (1884-1925) en un intento por axiomatizar la incipiente teoría de conjuntos de Cantor, también llamada la teoría de conjuntos "ingenua", incluyó entre sus axiomas el Axioma de Abstracción y es aquí donde aparece Bertrand Arthur William Russell (1872-1970). Russell razonó del modo siguiente: *dado cualquier conjunto X y cualquier objeto x, las reglas de la lógica dictan que $x \in X$ o $x \notin X$*. En particular, un conjunto X, o es miembro de sí mismo, o no lo es. Russell entonces considera la colección $\mathcal{R}$ constituida por los conjuntos que no son miembros de si mismo, es decir, $\mathcal{R} = \{X \in \mathfrak{U} : X \notin X\}$. Puesto que $\mathfrak{U}$ es, por el Axioma de Abstracción, un conjunto, resulta que $\mathcal{R}$ también es un conjunto lo que genera la siguiente contradicción:

$$\mathcal{R} \in \mathcal{R} \quad \Leftrightarrow \quad \mathcal{R} \notin \mathcal{R}.$$

Por esto,

Paradoja de Russell. *La colección $\mathfrak{U}$ **no es un conjunto**.*

La conclusión fundamental que se extrae del resultado anterior es la siguiente: la no aceptación del Axioma de Abstracción impide la construcción de colecciones tan grandes como $\mathfrak{U}$, $\mathcal{R}$ y muchas otras. Puesto que la Teoría de Conjuntos basada en la axiomática de (**ZFC**) prescinde del Axioma de Abstracción colecciones gigantesca como las de Russell están prohibidas en esta Teoría de Conjuntos pues ellas no son conjuntos, por lo que:

Hecho Universal: *expresiones del tipo $X \in X$ **no son aceptadas** cualquiera sea el conjunto X.*

Como suele suceder en muchas partes de las matemáticas, existen convenciones que resultan ser muy adecuadas. Por ejemplo, en la Teoría de Conjuntos, postular la existencia de un conjunto que no posee elementos es una de ellas. A tal conjunto se le llama el **conjunto vacío** y denotado por $\emptyset$. El conjunto vacío está caracterizado por la siguiente propiedad: "$x \in \emptyset$" *nunca se satisface, cualquiera sea x*. Es importante destacar que, una vez admitido la existencia del conjunto vacío, siempre se cumple que $\varnothing \subseteq X$, para cualquier conjunto X. En efecto, suponer que $\varnothing \nsubseteq X$ significa que existe algún $x \in \varnothing$ tal que $x \notin X$, pero como $x \in \varnothing$ nunca se satisface, entonces ello obliga a sentenciar que $\varnothing \subseteq X$. De esto último se deduce que el conjunto vacío es único.

Definición 1.1.2. *Dos conjuntos A y B son **iguales**, en notación, $A = B$, si ocurre que $A \subseteq B$ y $B \subseteq A$. Si la relación $A = B$ no se cumple, entonces diremos que A y B son **distintos** y lo denotaremos por $A \neq B$.*

La notación "$A \subsetneqq B$" significa que $A \subseteq B$ pero $A \neq B$, que se expresa diciendo que A es un **subconjunto propio** de B.

Definición 1.1.3. *Dado un conjunto X, indicaremos por $\mathscr{P}(X)$ al **conjunto potencia** o **de las partes de** X, es decir,*

$$\mathscr{P}(X) = \{A : A \subseteq X\}.$$

Por ejemplo,

$$\mathscr{P}(\varnothing) = \{\varnothing\}, \quad \mathscr{P}(\{\varnothing\}) = \{\varnothing, \{\varnothing\}\}, \text{ etc.}$$

En general, si $X \subseteq Z$, entonces $\mathscr{P}(X) \subseteq \mathscr{P}(Z)$.

Definición 1.1.4. *Dados los conjuntos A y B, la **unión** e **intersección** de ambos conjuntos, denotados por $A \cup B$ y $A \cap B$ respectivamente, se definen como:*

$$A \cup B = \{x : x \in A \ \text{ ó } \ x \in B\} \qquad y \qquad A \cap B = \{x : x \in A \ \ y \ \ x \in B\}.$$

En el caso particular en que $A \cap B = \varnothing$, entonces se dice que A y B son conjuntos **disjuntos** o **ajenos**. Se sigue inmediatamente de la definición anterior que las operaciones de unión e intersección son conmutativas, esto es, $A \cup B = B \cup A$ y $A \cap B = B \cap A$. Además,

$$A \cap B \subseteq A \subseteq A \cup B \qquad \text{y} \qquad A \cap B \subseteq B \subseteq A \cup B.$$

La unión e intersección de conjuntos se distribuyen según las siguientes igualdades:

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \qquad \text{y} \qquad A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

Más aun, la siguiente es una caracterización de $A \subseteq B$ en términos de la unión y la intersección.

$$A \subseteq B \quad \Leftrightarrow \quad A = A \cap B \quad \Leftrightarrow \quad B = A \cup B.$$

Definición 1.1.5. *Dados los conjuntos A y B, la **diferencia** $A \setminus B$ es el conjunto formado por todos los elementos de A que no son miembros de B, esto es,*

$$A \setminus B = \{x : x \in A \quad y \quad x \notin B\}.$$

Es importante observar lo siguiente referente a la diferencia de conjuntos:

(a) $A \setminus B = \varnothing$ si, y sólo si, $A \subseteq B$.

(b) $A \setminus B = A \setminus (A \cap B)$ y $A \setminus \varnothing = A$.

(c) $A \cap B = A \setminus (A \setminus B)$.

(d) Si $A \setminus B \subseteq A$, entonces $A \cap B = \varnothing$.

En el caso particular en que X es un conjunto fijo y $A \subseteq X$, entonces a $X \setminus A$ se le llama el **complemento** de A (relativo a X) y se denota por A^c. Observe que si X es un conjunto y $A, B \subseteq X$, entonces $A \setminus B = A \cap B^c$.

Definición 1.1.6. *La **diferencia simétrica** de los conjuntos A y B se expresa en la forma*

$$\begin{aligned} A \bigtriangleup B &= (A \setminus B) \cup (B \setminus A) \\ &= (A \cup B) \setminus (A \cap B). \end{aligned}$$

Algunas de las propiedades que son válidas con esta operación de conjuntos son las siguientes: si A, B, C, D son conjuntos arbitrarios, entonces

(a_1) $A \bigtriangleup B = B \bigtriangleup A$

(b_1) $A \bigtriangleup \varnothing = A$

(c_1) $A \bigtriangleup A = \varnothing$.

(d_1) $A \bigtriangleup B = A^c \bigtriangleup B^c$.

(e_1) $(A \bigtriangleup B) \bigtriangleup C = A \bigtriangleup (B \bigtriangleup C)$.

(f_1) $A \cap (B \bigtriangleup C) = (A \cap B) \bigtriangleup (A \cap C)$.

(g_1) $(A \cup B \cup C) \setminus (A \cap B \cap C) = (A \bigtriangleup B) \cup (B \bigtriangleup C)$.

(h_1) $(A \cup B) \bigtriangleup (C \cup D) \subseteq (A \bigtriangleup C) \cup (B \bigtriangleup D)$.

Puesto que no existe ninguna limitación para restringirnos a dos conjuntos en las definiciones de unión e intersección, podemos considerar uniones e intersecciones arbitrarias de conjuntos.

Definición 1.1.7. *Sea $\mathscr{A}$ una familia arbitraria de conjuntos. Definimos la **unión** e **intersección**, respectivamente, de dicha familia como*

$$\bigcup \mathscr{A} = \bigcup_{A \in \mathscr{A}} A = \{x : x \in A \text{ para algún } A \in \mathscr{A}\}$$

y

$$\bigcap \mathscr{A} = \bigcap_{A \in \mathscr{A}} A = \{x : x \in A \text{ para todo } A \in \mathscr{A}\},$$

Si $\mathscr{A}$ es una familia numerable, digamos $\mathscr{A} = \{A_1, A_2, \ldots\}$, entonces, en lugar de escribir $\bigcup_{A\in\mathscr{A}} A$, usaremos la notación $\bigcup_{n=1}^{\infty} A_n$. Lo mismo se hará con la intersección, es decir, escribiremos $\bigcap_{n=1}^{\infty} A_n$ en lugar de $\bigcap_{A\in\mathscr{A}} A$. Más aun, si $\mathscr{A} = \{A_{mn} : m, n = 1, 2, \ldots\}$, entonces las notaciones

$$\bigcup_{m,n=1}^{\infty} A_{mn} = \bigcup_{m=1}^{\infty}\bigcup_{n=1}^{\infty} A_{mn} \qquad \text{y} \qquad \bigcap_{m,n=1}^{\infty} A_{mn} = \bigcap_{m=1}^{\infty}\bigcap_{n=1}^{\infty} A_{mn}$$

se usarán frecuentemente. Como antes, si ocurre que $A \cap B = \varnothing$ para todo par de conjuntos A, B en $\mathscr{A}$, entonces diremos que $\mathscr{A}$ es una **familia disjunta** o que los conjuntos de $\mathscr{A}$ son **disjuntos dos a dos**.

Suponga que X es un conjunto no vacío y que $\mathscr{A}$ es una familia de subconjuntos de X. Si $X = \bigcup_{A\in\mathscr{A}} A$, entonces diremos que $\mathscr{A}$ es un **cubrimiento** de X. Si, además, la familia $\mathscr{A}$ es disjunta, entonces se dice que $\mathscr{A}$ es una **partición** de X o que X es una **unión disjunta** de $\mathscr{A}$.

A diferencia de los elementos de la unión y de la intersección, los del producto cartesiano son de naturaleza distinta a los elementos de A y de B.

Definición 1.1.8. *Sean X, Y conjuntos no vacíos. El **producto cartesiano** $X \times Y$ se define por*

$$X \times Y = \{(x, y) : x \in X,\, y \in Y\}.$$

Recuerde que todo *par ordenado* (x, y) se define como $(x, y) = \{\{a\}, \{a, b\}\}$. De esto e sigue $(x, y) \subseteq \mathscr{P}(\{x, y\})$. Algunas propiedades importantes sobre familias de conjuntos y que se usan frecuentemente son las siguientes. Sean $\mathscr{A}, \mathscr{B}$ familias de conjuntos. Entonces se verifica que:

$$\Big(\bigcup_{A\in\mathscr{A}} A\Big) \cap \Big(\bigcup_{B\in\mathscr{B}} B\Big) = \bigcup_{(A,B)\in\mathscr{A}\times\mathscr{B}} (A \cap B)$$

y

$$\Big(\bigcap_{A\in\mathscr{A}} A\Big) \cup \Big(\bigcap_{B\in\mathscr{B}} B\Big) = \bigcap_{(A,B)\in\mathscr{A}\times\mathscr{B}} (A \cup B).$$

También se cumplen las **Leyes de Morgan**: si X es un conjunto no vacío y $\mathscr{A} \subseteq \mathscr{P}(X)$, entonces

$$X \setminus \bigcup_{A\in\mathscr{A}} A = \bigcap_{A\in\mathscr{A}} (X \setminus A) \qquad \text{y} \qquad X \setminus \bigcap_{A\in\mathscr{A}} A = \bigcup_{A\in\mathscr{A}} (X \setminus A).$$

lo que comúnmente se escribe como

$$\Big(\bigcup_{A\in\mathscr{A}} A\Big)^c = \bigcap_{A\in\mathscr{A}} A^c \qquad \text{y} \qquad \Big(\bigcap_{A\in\mathscr{A}} A\Big)^c = \bigcup_{A\in\mathscr{A}} A^c$$

1.1.2. Funciones

Sean X, Y conjuntos no vacíos. Una **relación** de X en Y es cualquier subconjunto $\mathcal{R}$ de $X \times Y$. En lo que sigue, cualquier elemento (x, y) de $\mathcal{R}$ se indicará por el símbolo $x\,\mathcal{R}\,y$. Si $X = Y$, entonces a la relación $\mathcal{R}$ se le llama **relación binaria**.

Definición 1.1.9. *Una **función**, o **aplicación**, de X en Y es una relación f de X en Y que posee la siguiente propiedad adicional: si $(x,y) \in f$ y $(x,z) \in f$, entonces $y = z$.*

Siguiendo la tradición, a la función f de la definición anterior la denotaremos, en lo sucesivo, por el símbolo $f : X \to Y$. Así, toda función $f : X \to Y$ asigna a cada uno de los elementos $x \in X$ un único $y \in Y$ al que designaremos por $f(x)$. Al conjunto X se le llama el **dominio** de la función f, mientras que a Y se le llama el **contradominio** de f.

Definición 1.1.10. *Dos funciones $f : X \to Y$ y $g : X' \to Y'$ son **iguales** si $X = X'$, $Y = Y'$ y $f(x) = g(x)$ para todo $x \in X$.*

Una función $f : X \to Y$ se llama **inyectiva** o **uno a uno** si dados $x, y \in X$ arbitrarios, la igualdad $f(x) = f(y)$ implica que $x = y$. La función f se dice que es **sobreyectiva**, o simplemente **sobre**, si $Y = f(X)$, es decir, si para cada $y \in Y$ existe un $x \in X$ tal que $y = f(x)$. Si f es tanto inyectiva así como también sobreyectiva, entonces la diremos que es **biyectiva**. Para una función $f : X \to Y$, el conjunto

$$\mathrm{Gra}(f) \;=\; \big\{(x, f(x)) \in X \times Y : \, x \in X\big\}$$

es llamado el **gráfico** de f. Si $f : X \to Y$ es una función y $A \subseteq X$, entonces la **imagen** de A por f es el conjunto

$$f(A) \;=\; \big\{f(x) \in Y : \, x \in A\big\}.$$

Por otro lado, si $B \subseteq Y$, la **imagen inversa** de B por f es el conjunto

$$f^{-1}(B) \;=\; \big\{x \in X : \, f(x) \in B\big\}.$$

Es fácil ver que si $\mathcal{A} \subseteq \mathscr{P}(X)$, entonces

$$f\Big(\bigcup_{A \in \mathcal{A}} A\Big) \;=\; \bigcup_{A \in \mathcal{A}} f(A) \qquad \text{y} \qquad f\Big(\bigcap_{A \in \mathcal{A}} A\Big) \;\subseteq\; \bigcap_{A \in \mathcal{A}} f(A).$$

Observe que la última inclusión puede ser propia. En efecto, si existen elementos $x, y \in X$ con $x \neq y$ para los cuales $f(x) = f(y)$, entonces tomando $A = \{x\}$ y $B = \{y\}$, se tiene que $A \cap B = \varnothing$, de donde $f(A \cap B) = \varnothing$, mientras que $f(A) \cap f(B) = \{f(x)\}$. La construcción de este ejemplo sólo es posible si nuestra función f no es inyectiva, de modo que *si f es inyectiva, entonces*

$$f\Big(\bigcap_{A \in \mathcal{A}} A\Big) \;=\; \bigcap_{A \in \mathcal{A}} f(A).$$

Para la imagen inversa se cumple que si $\mathcal{B} \subseteq \mathscr{P}(Y)$, entonces

$$f^{-1}\Big(\bigcup_{B \in \mathcal{B}} B\Big) \;=\; \bigcup_{B \in \mathcal{B}} f^{-1}(B) \qquad \text{y} \qquad f^{-1}\Big(\bigcap_{B \in \mathcal{B}} B\Big) \;=\; \bigcap_{B \in \mathcal{B}} f^{-1}(B).$$

Si $B \subseteq Y$, también es válida la siguiente igualdad:

$$f^{-1}(Y \setminus B) \;=\; X \setminus f^{-1}(B).$$

Más aun, dado $A \subseteq X$, se tiene que

$$A \subseteq f^{-1}\big(f(A)\big),$$

mientras que si $B \subseteq Y$, entonces
$$f\big(f^{-1}(B)\big) \subseteq B.$$
Ya hemos visto que $A \subseteq f^{-1}\big(f(A)\big)$. ¿Bajo qué condiciones $f^{-1}(f(A)) = A$? Para que ocurra la igualdad $f^{-1}(f(A)) = A$ cualquiera que sea $A \subseteq X$, es necesario y suficiente que f sea *inyectiva*. Similarmente, f es *sobreyectiva* si, y sólo si, $f(f^{-1}(B)) = B$ para todo $B \subseteq Y$.

Si $f : X \to Y$ y $g : Y \to Z$ son funciones, entonces podemos definir la **función compuesta** $g \circ f : X \to Z$ como $(g \circ f)(x) = g(f(x))$ para todo $x \in X$. Sea A un subconjunto de X. La aplicación $i_A : A \to X$, definida por $i(x) = x$ para todo $x \in A$, se llama la **aplicación inclusión** de A en X. En el caso particular cuando $A = X$, la aplicación inclusión de X en X, se llama la **función identidad** y será indicada por $\mathrm{Id} : X \to X$. Cada función biyectiva $f : A \to B$ da origen a otra función biyectiva, llamada la **inversa** de f y denotada por $f^{-1} : B \to A$ tal que $f \circ f^{-1} = f^{-1} \circ f = \mathrm{Id}$.

Sean $f : X \to Y$ una función y A un subconjunto no vacío de X. La **restricción** de f al subconjunto A es la aplicación $f|_A : A \to Y$ definida por $(f|_A)(x) = f(x)$ para todo $x \in A$. Nótese que $f|_A = f \circ i_A$, donde i_A es la inclusión de A en X. Por otro lado, dada una función $g : A \to Y$, toda aplicación $f : X \to Y$ tal que $g = f|_A$ se llama una **extensión** de g al conjunto X. La función $\chi_A : X \to \mathbb{R}$ definida por
$$\chi_A(x) = \begin{cases} 1 & \text{si } x \in A, \\ 0 & \text{si } x \notin A \end{cases}$$
se le denomina la función **característica** de A. En el caso particular cuando $A = \mathbb{Q}$, a $\chi_{\mathbb{Q}}$ se le llama la **función de Dirichlet**.

Sea $f : X \to \mathbb{R}$ una función. Diremos que f es **no-negativa** sobre X si $f(x) \geq 0$ para todo $x \in X$. Similarmente, decir que f es **no-positiva** sobre X significa que $f(x) \leq 0$ para todo $x \in X$. De modo más general, si $f, g : X \to \mathbb{R}$ son funciones, entonces $f \leq g$ sobre X, significa que $(g - f)(x) \geq 0$ para todo $x \in X$.

Si $f : X \to \mathbb{R}$ es una función, entonces f se puede escribir en la forma $f = f^+ - f^-$, donde
$$f^+(x) = \begin{cases} f(x) & \text{si } f(x) \geq 0 \\ 0 & \text{si } f(x) < 0 \end{cases} \qquad \text{y} \qquad f^-(x) = \begin{cases} -f(x) & \text{si } f(x) \leq 0 \\ 0 & \text{si } f(x) > 0 \end{cases}$$
Observe que tanto f^+ así como f^- son ambas no-negativas. A f^+ y f^- se les llaman la **parte positiva** y la **parte negativa** de f, respectivamente. El **valor absoluto** de f se define entonces como $|f| = f^+ + f^-$. Finalmente, la **proyección** de $X \times Y$ sobre X es la aplicación $\mathrm{pr}_X : X \times Y \to X$ definida por $\mathrm{pr}_X(x, y) = x$ para todo $(x, y) \in X \times Y$. Similarmente, la **proyección** de $X \times Y$ sobre Y es la aplicación $\mathrm{pr}_Y : X \times Y \to Y$ definida por $\mathrm{pr}_Y(x, y) = y$ para todo $(x, y) \in X \times Y$.

Definición 1.1.11. *Sea X un conjunto no vacío. Una* ***relación de equivalencia*** *sobre X es una relación binaria $\mathcal{R}$ sobre dicho conjunto que es*

(a) ***reflexiva****: $x \,\mathcal{R}\, x$ para todo $x \in X$,*

(b) ***simétrica****: $x \,\mathcal{R}\, y \Rightarrow y \,\mathcal{R}\, x$, para todo $x, y \in X$, y*

(c) ***transitiva****: $x \,\mathcal{R}\, y$ y $y \,\mathcal{R}\, z \Rightarrow x \,\mathcal{R}\, z$, para todo $x, y, z \in X$.*

Frecuentemente usaremos el símbolo $\sim$ en lugar $\mathcal{R}$. En consecuencia, escribiremos $x \sim y$ en lugar de $x \mathcal{R} y$ y diremos que x y y son $\sim$**equivalentes**. La **clase de equivalencia de** x **módulo** $\sim$ es el conjunto

$$C_x \;=\; \{y \in X \,:\, x \sim y\}.$$

Observe que cualesquiera sean $x, y \in X$, se verifica que $C_x = C_y$ o bien $C_x \cap C_y = \varnothing$. Más aun, puesto que $x \in C_x$ para todo $x \in X$, resulta que las clases de equivalencias forman una partición de X, es decir, $X = \bigcup_{x \in X} C_x$. Al conjunto

$$X/\!\sim \;=\; \{C_x : x \in X\},$$

se le llama el **cociente** de X por la relación $\sim$.

La función $Q : X \to X/\!\sim$ definida por $Q(x) = C_x$ para cada $x \in X$, es claramente sobreyectiva y se le llama la **aplicación cociente** o **canónica** sobre X.

1.1.3. Familias Indexadas. Productos Cartesianos

Todo conjunto no vacío X puede ser considerado como una familia indexada por sus propios elementos, es decir, $X = \{z_x : x \in X\}$, donde $z_x = x$ para cada $x \in X$. Con frecuencia, resulta más práctico y útil, asignarle a cada elemento $x \in A$ una etiqueta distinta. Un modo de hacer esto es el siguiente: se considera un conjunto no vacío I (cuyos elementos llamaremos **índices**) de modo que exista una aplicación biyectiva $x(\cdot) : I \to X$. La imagen de cada elemento $\alpha \in I$ por medio de $x(\cdot)$, es decir, $x(\alpha)$, se denotará por x_α y entonces el conjunto X se identificará con $\{x_\alpha : \alpha \in I\}$, al que denotaremos por el símbolo $(x_\alpha)_{\alpha \in I}$ y se dirá que X está **indexado por el conjunto** I. Cuando I es un **conjunto dirigido**, es decir, cuando sobre I existe una relación $\preceq$ entre sus elementos verificando la propiedad: cualesquiera sean $\alpha, \beta \in I$, existe un $\xi \in I$ tal que

$$\alpha \;\preceq \xi \qquad\qquad \text{y} \qquad\qquad \beta \;\preceq \xi,$$

entonces diremos que $(x_\alpha)_{\alpha \in I}$ es una **red** en X. Cuando $I = \mathbb{N}$, entonces diremos que $(x_n)_{n=1}^{\infty}$ es una **sucesión** en X.

En general, si $\mathcal{A}$ es una familia de conjuntos y si suponemos que I es un conjunto no vacío y $x(\cdot) : I \to \mathcal{A}$ es una aplicación biyectiva, entonces la colección $\mathcal{A}$ se identifica con la **familia de conjuntos** $\{A_\alpha : \alpha \in I\}$, lo que frecuentemente escribiremos como $\mathcal{A} = (A_\alpha)_{\alpha \in I}$. En este caso, la unión de los elementos de la familia $\mathcal{A}$ se escribirá como $\bigcup_{\alpha \in I} A_\alpha$ en lugar de $\bigcup_{A \in \mathcal{A}} A$, y lo mismo para la intersección. Si $I = \mathbb{N}$, entonces a la familia $\mathcal{A} = (A_n)_{n=1}^{\infty}$ se le llama una **sucesión de conjuntos**. Si $\mathcal{A} = (A_\alpha)_{\alpha \in I}$ es una familia de conjuntos donde I es un conjunto dirigido, entonces diremos que $(A_\alpha)_{\alpha \in I}$ es una **red de conjuntos**. La red $(A_\alpha)_{\alpha \in I}$ se llama **creciente** (respectivamente, **decreciente**) si $A_\alpha \subseteq A_\beta$ (respectivamente, $A_\alpha \supseteq A_\beta$) siempre que $\alpha \preceq \beta$. Cuando $I = \mathbb{N}$ entonces hablaremos de una sucesión creciente o decreciente de conjuntos. Si las inclusiones son todas estrictas, entonces diremos que la sucesión es **estrictamente creciente** (respectivamente, **estrictamente decreciente**).

Definición 1.1.12. *Sea $(A_\alpha)_{\alpha \in I}$ una familia cualquiera de conjuntos. Se define el* ***producto cartesiano*** *de esta familia como el conjunto de todas las funciones $x(\cdot)$ que tienen dominio I tal que $x(\alpha) = x_\alpha \in A_\alpha$ para cada $\alpha \in I$, es to es,*

$$\prod_{\alpha \in I} A_\alpha \;=\; \left\{ x(\cdot) : I \to \bigcup_{\alpha \in I} A_\alpha \;\middle|\; x(\alpha) = x_\alpha \in A_\alpha \;\; \textit{para cada } \alpha \in I \right\}.$$

Según lo expresado anteriormente, podemos también escribir

$$\prod_{\alpha \in I} A_\alpha = \Big\{ (x_\alpha)_{\alpha \in I} \, : \, x_\alpha \in A_\alpha \ \text{ para cada } \alpha \in I \Big\}.$$

Si cada conjunto A_α es no vacío, entonces toda función $x \in \prod_{\alpha \in I} A_\alpha$ es llamada una **función de elección** para la familia $(A_\alpha)_{\alpha \in I}$. Si ocurre que todos los A_α son iguales, digamos, $A_\alpha = A$ para todo $\alpha \in I$, entonces el producto cartesiano $\prod_{\alpha \in I} A_\alpha$ se denotará brevemente por A^I. En el caso particular en que $I = \{1, \ldots, n\}$ para algún $n \in \mathbb{N}$, escribiremos A^n en lugar de A^I. En general, escribiremos $\prod_{n=1}^{\infty} A_n$ como sinónimo de $\prod_{n \in \mathbb{N}} A_n$. El conjunto $\mathbb{K}^n$ es llamado el **espacio Euclidiano de dimensión** n **(o** n**-dimensional)**. Observe que si X es un conjunto arbitrario, entonces $\mathbb{R}^X$ constituye el conjunto de todas las funciones $f : X \to \mathbb{R}$. De interés es el producto cartesiano $\prod_{n=1}^{\infty} A_n$ donde $A_n = \{0, 1\}$ para todo entero $n \geq 1$. A éste producto lo denotaremos por $2^{\mathbb{N}}$, el cual consiste de todas las sucesiones $(x_n)_{n=1}^{\infty}$ donde cada $x_n \in \{0, 1\}$. Finalmente, para cada $\beta \in I$ se considera la aplicación $p_\beta : \prod_{\alpha \in I} A_\alpha \to A_\beta$ definida por $p_\beta((x_\alpha)_{\alpha \in I}) = x_\beta$. A p_β se llama la β-ésima **proyección**. Claramente p_β es una aplicación sobreyectiva.

Si $(A_\alpha)_{\alpha \in I}$ y $(B_\beta)_{\beta \in J}$ son familias de conjuntos, entonces el producto de sus uniones e intersecciones satisfacen:

$$\Big(\bigcup_{\alpha \in I} A_\alpha \Big) \times \Big(\bigcup_{\beta \in J} B_\beta \Big) = \bigcup_{(\alpha, \beta) \in I \times J} A_\alpha \times B_\beta,$$

$$\Big(\bigcap_{\alpha \in I} A_\alpha \Big) \times \Big(\bigcap_{\beta \in J} B_\beta \Big) = \bigcap_{(\alpha, \beta) \in I \times J} A_\alpha \times B_\beta.$$

1.2. Conjuntos Numerables y otros más Numerosos

En esta sección introduciremos un método para “comparar” el “*número*” de elementos que poseen dos conjuntos. Esto se hará a través de la noción de *cardinalidad*. Posteriormente, si los conjuntos poseen un cierto “orden”, entonces, además de comparar el número de elementos que ellos poseen, también estaremos interesados en preservar la “posición” que ellos ocupan en cada conjunto. Comenzaremos con la noción de “igual de números de elementos” entre dos conjuntos.

En el año 1874 Cantor demostró que existía una correspondencia uno-a-uno entre $\mathbb{N}$ y el conjunto de los números algebraicos (en realidad fue Dedekind quien lo hizo). Posteriormente, demuestra que no existe correspondencia uno-a-uno entre $\mathbb{N}$ y el conjunto de los números reales. Estos hechos le permitió considerar la existencia de una correspondencia uno-a-uno como un criterio para comparar el tamaño de dos conjuntos.

Definición 1.2.1. *Dos conjuntos A y B se dice que son* ***equipotentes****, o* ***biyectables****, si existe una función biyectiva $f : X \to Y$.*

Escribiremos $A \approx B$ para abreviar la expresión “ A y B son equipotentes”. Esta relación, evidentemente, nos muestra que los conjuntos A y B poseen el “*mismo número de elementos*”. Esta idea permite que intentemos asignar a cualquier conjunto A un objeto de la Teoría de Conjuntos, al que llamaremos **número cardinal** y denotado por $\operatorname{card}(A)$, de modo que

$$X \approx A \quad \Leftrightarrow \quad \operatorname{card}(X) = \operatorname{card}(A).$$

Una motivación para esto es observar, usando la noción de conjuntos equipotentes, que:

(a) $A \approx A$ para cualquier conjunto A,

(b) si $A \approx B$, entonces $B \approx A$ y

(c) si $A \approx B$ y $B \approx C$, entonces $A \approx C$

Nótese que la relación $\approx$ se comporta como una *relación de equivalencia* sobre la colección $\mathfrak{U}$ de todos los conjuntos. Sin embargo, como $\mathfrak{U}$ no es un conjunto y puesto que la definición de relación de equivalencia, Definición 1.1.11, se formuló sólo para conjuntos, tropezamos con un serio problema: $\approx$ *no es una relación de equivalencia sobre* $\mathfrak{U}$. ¿Cómo resolver este impasse? Uno puede intentar manejar esta situación definiendo el cardinal del conjunto X del modo siguiente:

$$\mathrm{card}(X) \;=\; \mathcal{C}_X$$

donde $\mathcal{C}_X = \{A \in \mathfrak{U} : A \approx X\}$. Observe que con esta definición

$$X \approx A \quad \Leftrightarrow \quad \mathrm{card}(X) \;=\; \mathrm{card}(A).$$

Sin embargo, como estamos trabajando en la Teoría de Conjuntos basada en la axiomática de **ZFC**, resulta que tal definición no es apta desde nuestro punto de vista ya que $\mathrm{card}(X)$ debe ser un conjunto y no tenemos certeza de que $\mathcal{C}_X$ lo sea. ¿Cómo definir, entonces, la cardinalidad de un conjunto en nuestra teoría? Pues bien, para dar una definición precisa de cardinalidad debemos apoyarnos en el Axioma de Elección y la Teoría de Ordinales que desarrollaremos brevemente en el próximo capítulo. Sin embargo, para ciertos tipos de conjuntos podemos aproximarnos a una tal definición.

Definición 1.2.2 (Bolzano). *Diremos que un conjunto A es **finito** si ocurre que $A = \varnothing$, o existe un $n \in \mathbb{N}$ tal que $A \approx \{1, 2, \ldots, n\}$. En este caso se dice que A tiene n-**elementos** y escribiremos $\mathrm{card}(A) = n$.*

Es importante destacar que en base a esta definición se tiene que: si A y B *son conjuntos finitos*, entonces:

$$A \approx B \quad \Leftrightarrow \quad \mathrm{card}(A) \;=\; \mathrm{card}(B)$$

Nótese que si A y B son conjuntos finitos con $\mathrm{card}(A) = m$, $\mathrm{card}(B) = n$ y, además, $m \neq n$, entonces $A \not\approx B$. Esta observación nos dice, en particular, que *un conjunto finito no puede ser equipotente a ningún subconjunto propio de sí mismo*.

Definición 1.2.3 (Bolzano). *Un conjunto A se llama **infinito** si él no es finito. Un conjunto infinito A se dice que es **numerable** si $A \approx \mathbb{N}$, en caso contrario diremos que A es **no-numerable**. La expresión "A es **a lo más numerable**" significa que A, o es finito, o es infinito numerable.*

Fijemos un conjunto A y sea

$$\mathcal{B}\mathrm{i}(\mathbb{N}, A) \;=\; \{f : \mathbb{N} \to A \mid f \text{ es biyectiva}\}.$$

Con la notación anterior, nuestra definición de conjunto numerable se puede expresar de esta forma:

*A es **numerable** si, y sólo si, $\mathcal{B}\mathrm{i}(\mathbb{N}, A) \neq \varnothing$.*

Es importante observar que si A es numerable, entonces siempre podemos hacer una lista infinita de sus elementos y escribir a A, por ejemplo, como $A = \{a_1, a_2, a_3, \ldots\}$. En efecto, basta elegir cualquier función $f \in \mathcal{B}\mathrm{i}(A, \mathbb{N})$ y luego definir $a_n = f(n) \in A$ para cada $n \in \mathbb{N}$. Nótese que al ser f inyectiva, todos los a_n so distintos dos a dos y puesto que ella también es sobreyectiva, cada elemento de A es de la forma a_n para un único $n \in \mathbb{N}$. En este caso diremos que $\{a_1, a_2, a_3, \ldots\}$ es una ***enumeración*** de A. Por supuesto, *A puede ser enumerado de muchas formas diferentes* pues ello depende de la elección de la función f en $\mathcal{B}\mathrm{i}(A, \mathbb{N})$. Por otro lado, decir que *un conjunto A es **no-numerable** significa que A es infinito y no existe ninguna biyección de A en $\mathbb{N}$*, lo que también se puede expresar en la forma:

*A es **no-numerable** si, y sólo si, A es infinito y $\mathcal{B}\mathrm{i}(\mathbb{N}, A) = \varnothing$.*

Un principio que es fundamental en matemáticas es el siguiente:

Definición 1.2.4 (Principio del Buen-Orden). *Si A es cualquier subconjunto **no vacío** de $\mathbb{N}$, entonces A posee un **primer elemento**, esto es, existe un $n_0 \in A$ tal que $n_0 \leq n$ para todo $n \in A$.*

El Principio del Buen-Orden es el responsable del siguiente hecho: cualquier subconjunto de $\mathbb{N}$ es a lo más numerable.

Teorema 1.2.5. *Si A es un subconjunto **no vacío** de $\mathbb{N}$, entonces A es **a lo más numerable**. En particular, si A es infinito, entonces él se puede representar por medio de una sucesión **estrictamente creciente**, es decir, $A = \{m_n : n \in \mathbb{N}\}$, donde*

$$m_1 < m_2 < m_3 < \cdots < m_n < \cdots$$

Prueba. Si A es finito, la conclusión es obvia. Suponga entonces que A es infinito. Puesto que A es no vacío, el Principio del Buen-Orden garantiza que A posee un primer elemento, es decir, existe un $m_1 \in A$, tal que

$$m_1 \leq a \quad \text{para todo } a \in A.$$

Ahora bien, como A es infinito, el conjunto $A_1 = A \setminus \{m_1\}$ es no vacío y, de nuevo, por el Principio del Buen-Orden, existe un $m_2 \in A_1$ tal que

$$m_2 \leq a \quad \text{para todo } a \in A_1.$$

Por supuesto, como $m_1 \notin A_1$, resulta que

$$m_1 < m_2.$$

Sea $A_2 = A_1 \setminus \{m_2\} = A \setminus \{m_1, m_2\}$. Entonces A_2 es no vacío y se repite, como antes, el procedimiento anterior para hallar un $m_3 \in A_2$ tal que

$$m_1 < m_2 < m_3.$$

En definitiva, teniendo en cuenta que A es infinito, podemos continuar indefinidamente con el argumento antes descrito para concluir que el conjunto A se puede escribir en la forma $A = \{m_n : n \in \mathbb{N}\}$, donde

$$m_1 < m_2 < m_3 < \cdots < m_n < \cdots$$

La aplicación $f : \mathbb{N} \to A$ definida por $f(n) = m_n$ es claramente biyectiva y termina la prueba.■

Algunas consecuencias inmediatas que se obtienen de la definición de conjuntos infinitos son las siguientes: Sean A y B conjuntos **no-vacíos**.

(Nu_1) *Si A es **a lo más numerable** (respectivamente, **no-numerable**) y B es **equipotente** con A, entonces B es **a lo más numerable** (respectivamente, **no-numerable**)*. Esto sigue del hecho de que la composición de dos funciones biyectivas es biyectiva.

(Nu_2) *Si A es **numerable** y $B \subseteq A$, entonces B es **a lo más numerable***. En efecto, la numerabilidad de A nos garantiza la existencia de una función biyectiva $f : A \to \mathbb{N}$. El conjunto $A' = f(B) \subseteq \mathbb{N}$ es, por el Teorema 1.2.5, a lo más numerable y, en consecuencia, la función $g : B \to A'$ definida por $g(b) = f(b)$ para todo $b \in B$ es biyectiva. El resultado ahora sigue de (Nu_1).

(Nu_3) *Si A es **no-numerable** y $A \subseteq B$, entonces B es **no-numerable***. Es consecuencia de (Nu_2).

El siguiente resultado nos provee de una caracterización muy útil de los conjuntos numerables.

Teorema 1.2.6. *Sea A un **conjunto infinito**. Las siguientes condiciones son equivalentes:*

(1) *A es **numerable**.*

(2) *Existe una **función sobreyectiva** $f : \mathbb{N} \to A$.*

(3) *Existe una **función inyectiva** $g : A \to \mathbb{N}$.*

Prueba. $(1) \Rightarrow (2)$. Suponga que A es numerable y escoja una función $f \in \mathcal{B}i(\mathbb{N}, A)$. Puesto que f es sobreyectiva, (2) sigue.

$(2) \Rightarrow (3)$. Suponga que $f : \mathbb{N} \to A$ es una función sobreyectiva y observe que $f^{-1}(\{a\}) \neq \varnothing$ para todo $a \in A$. Puesto que $f^{-1}(\{a\}) \subseteq \mathbb{N}$ el Principio del Buen-Orden nos garantiza que $\min f^{-1}(\{a\})$ existe y es único. Esto permite definir la función $g : A \to \mathbb{N}$ por

$$g(a) \;=\; \min f^{-1}(\{a\}) \quad \text{para cada} \quad a \in A.$$

Veamos ahora que g es inyectiva. En efecto, suponga que $a, a' \in A$ con $a \neq a'$. Entonces $f^{-1}(\{a\}) \cap f^{-1}(\{a'\}) = \varnothing$, lo cual implica que $\min f^{-1}(\{a\}) \neq \min f^{-1}(\{a'\})$. Por esto $g(a) \neq g(a')$ y g es inyectiva.

$(3) \Rightarrow (1)$. Suponga que $g : A \to \mathbb{N}$ es una función inyectiva. Puesto que A es infinito y $g(A) \subseteq \mathbb{N}$, resulta que $g(A)$ también es numerable. Sea $h : g(A) \to \mathbb{N}$ una biyección. Teniendo en cuenta que $g : A \to g(A)$ es una biyección, se tiene que $f = h \circ g$ es una biyección y termina la prueba. ■

Siguiendo la tradición, usaremos el símbolo $\aleph_0$ para designar el *número de elementos* de $\mathbb{N}$, es decir, escribiremos

$$\operatorname{card}(\mathbb{N}) \;=\; \aleph_0.$$

La **cardinalidad** de un conjunto numerable se puede definir sin ambigüedad del modo siguiente:

Definición 1.2.7. *Si A es cualquier **conjunto numerable**, definimos la **cardinalidad** de A por*

$$\operatorname{card}(A) \;=\; \aleph_0$$

Observe que, similar al caso finito, si A y B *son conjuntos numerables*, entonces

$$A \approx B \quad \Leftrightarrow \quad \mathrm{card}(A) \;=\; \mathrm{card}(B).$$

¿Cómo definir $\mathrm{card}(A)$ si A es no-numerable? Pues bien, es aquí donde se presenta el meollo del asunto. A diferencia de la familia de los conjuntos numerables en donde existe un único objeto, $\aleph_0$, que los identifica a todos, en el caso de la colección de los conjuntos no-numerables no hay tal objeto. En realidad, existe una cantidad infinita de "objetos" en orden "estrictamente creciente" para conjuntos no-numerables. Sin embargo, la relación anterior permite que podamos introducir una definición (incompleta) de cardinalidad del modo siguiente:

Definición 1.2.8 (incompleta). *A cada conjunto A se le asigna un único objeto, al que llamaremos la **cardinalidad** de A, $\mathrm{card}(A)$, tal que:*

$$A \approx B \quad \Leftrightarrow \quad \mathrm{card}(A) \;=\; \mathrm{card}(B).$$

El problema con esta definición es que no se especifica qué cosa es $\mathrm{card}(A)$ o cómo se escoge y, por lo tanto, se le considera incompleta. Más adelante veremos, cuando hallamos introducido la noción de *número cardinal*, que ésta definición incompleta es, en realidad, una buena definición, es decir, usando el Axioma de Elección y la Teoría de Conjuntos Bien-Ordenados se demuestra que existe una operación que es compatible con la relación $\approx$.

Teorema 1.2.9. *Sean A y B conjuntos no vacíos con A **finito** y B **numerable**. Entonces $A \cup B$ es numerable.*

Prueba. Es suficiente suponer que A y B son disjuntos. Siendo A finito, existe un $n \in \mathbb{N}$ y una función biyectiva $f : N_n \to A$, donde $N_n = \{1, 2, \ldots, n\}$. Similarmente, existe una función biyectiva $g : \mathbb{N} \to B$. Considere la función

$$h : \mathbb{N} \setminus N_n \to \mathbb{N}$$

definida por

$$h(i) \;=\; i - n, \quad \text{para todo } \; i \in \mathbb{N} \setminus N_n.$$

Claramente h es biyectiva. Finalmente, la función $\varphi : \mathbb{N} \to A \cup B$ definida por

$$\varphi\big|_{N_n} \;=\; f, \qquad \varphi\big|_{\mathbb{N} \setminus N_n} \;=\; g \circ h$$

es biyectiva y termina la prueba. ■

Uno de los resultados fundamentales acerca de la noción de conjuntos numerables es el siguiente el cual se puede demostrar por del medio del genial Método de la Diagonal de Cantor.

Teorema 1.2.10. *Sea $(A_n)_{n=1}^{\infty}$ una colección infinita numerable y **disjunta** de conjuntos **a lo más numerables**. Entonces $\bigcup_{n=1}^{\infty} A_n$ es **numerable**.*

Prueba. En vista del Teorema 1.2.9 es suficiente suponer que cada A_n es infinito numerable. Puesto que cada conjunto A_n es numerable, podemos hacer una lista de sus elementos, por ejemplo, del modo siguiente:

$$A_n \;=\; \{a_{nj} \,:\, j = 1, 2, \ldots\}.$$

Disponga ahora todos los elementos de cada uno de los conjuntos A_n en el siguiente arreglo matricial infinito (los puntos suspensivos indican que las sucesiones se prolongan indefinidamente a la derecha y hacia abajo):

$$\begin{array}{lcccc}
A_1 & a_{11} & a_{12} & a_{13} & \cdots \\
A_2 & a_{21} & a_{22} & a_{23} & \cdots \\
A_3 & a_{31} & a_{32} & a_{33} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots \\
A_n & a_{n1} & a_{n2} & a_{n3} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{array}$$

y haga una lista de ellos siguiendo las diagonales sucesivas come se muestra en la figura adjunta:

$$\begin{tikzcd}
a_{11} \arrow[r] & a_{12} \arrow[dl] & a_{13} \arrow[r] & a_{14} \arrow[dl] & \cdots \\
a_{21} \arrow[d] & a_{22} \arrow[ur] & a_{23} \arrow[dl] & a_{24} & \cdots \\
a_{31} \arrow[ur] & a_{32} \arrow[dl] & a_{33} & a_{34} & \cdots \\
a_{41} \arrow[d] & a_{42} & a_{43} & a_{44} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{tikzcd}$$

Ahora bien, como la colección $(A_n)_{n=1}^{\infty}$ es disjunta, resulta que los a_{ij} son todos distintos entre sí, lo cual permite que se pueda establecer una correspondencia biunívoca entre $\mathbb{N}$ y $\bigcup_{n=1}^{\infty} A_n$ tal como se muestra en la figura adjunta.

$$\begin{array}{ccccccccccc}
a_{11} & a_{12} & a_{21} & a_{31} & a_{22} & a_{13} & a_{14} & a_{23} & a_{32} & a_{41} & \cdots \\
\uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \vdots \\
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & \cdots
\end{array}$$

Esto establece la numerabilidad de $\bigcup_{n=1}^{\infty} A_n$ y termina la prueba. ■

La demostración del resultado anterior posee una pequeña sutileza que con mucha frecuencia pasa desapercibida. En efecto, aunque dicha prueba pareciera, en una primera mirada, ser constructiva, ella no lo es. La razón es la siguiente: considere la colección

$$\mathcal{B} = \{\mathcal{B}\mathrm{i}(\mathbb{N}, A_n) : n \in \mathbb{N}\}.$$

Como cada conjunto $\mathcal{B}\mathrm{i}(\mathbb{N}, A_n)$ es no vacío, podemos seleccionar una función f_n en él para producir la enumeración de A_n, es decir, como la elección es arbitraria, esto nos indica que existen muchas maneras de enumerar los elementos de A_n. Una vez que los conjuntos A_n han sido expresados en la forma $A_n = \{a_{nj} : j = 1, 2, \dots\}$, el resto de la prueba es, efectivamente, constructiva. Sin embargo, como acabamos de ver, para poder elegir una función en cada uno de los conjuntos $\mathcal{B}\mathrm{i}(\mathbb{N}, A_n)$ debemos hacer uso del enigmático y necesario Axioma de Elección, en su versión numerable, que discutiremos un poco más adelante.

Corolario 1.2.11. *Sea $(A_n)_{n=1}^{\infty}$ una familia infinita numerable de **conjuntos numerables**. Entonces para cualquier entero $k \geq 2$,*

$$\bigcup_{n=1}^{\infty} A_n \qquad y \qquad \prod_{n=1}^{k} A_n$$

*son **numerables**. Más aun, si A es cualquier **conjunto numerable**, entonces*

$$\mathcal{P}_{\text{fin}}(A) \ = \ \{F \subseteq A \ : \ \operatorname{card}(F) \text{ es } \textbf{\textit{finito}}\}$$

*también es **numerable**.*

Prueba. Si la sucesión $(A_n)_{n=1}^{\infty}$ es disjunta, el resultado sigue del teorema anterior. Suponga entonces que la sucesión $(A_n)_{n=1}^{\infty}$ no es disjunta y considere la sucesión $(B_n)_{n=1}^{\infty}$ definida por:

$$B_1 \ = \ A_1 \qquad \text{y} \qquad B_n \ = \ A_n \setminus (A_1 \cup \cdots \cup A_{n-1}) \quad \text{para } n \geq 2.$$

Es claro que dicha sucesión posee las siguientes propiedades: (i) cada $B_n \subseteq A_n$ y, en consecuencia, es a lo más numerable, (ii) la familia $(B_n)_{n=1}^{\infty}$ es disjunta y (iii) $\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$. Se sigue del Teorema 1.2.10 que $\bigcup_{n=1}^{\infty} A_n$ es numerable.

Para demostrar $\prod_{n=1}^{k} A_n)$ es numerable, es suficiente comprobar que $\mathbb{N} \times \mathbb{N}$ es numerable. Veamos esto. Para cada $n \in \mathbb{N}$ escribamos $A_n = \{(m, n) : m \in \mathbb{N}\}$. Puesto que los conjuntos A_n son infinitos y disjuntos dos a dos, el Teorema 1.2.10 nos garantiza que $\bigcup_{n=1}^{\infty} A_n = \mathbb{N} \times \mathbb{N}$ es numerable.

Finalmente, para verificar que $\mathcal{P}_{\text{fin}}(A)$ es numerable, suponga que A es un conjunto numerable y sea $A = \{a_1, a_2, \ldots\}$ una enumeración de A. Para cada $n \in \mathbb{N}$, considere el conjunto

$$A_n \ = \ \mathscr{P}(\{a_1, \ldots, a_n\}) \ = \ \{F \ : \ F \subseteq \{a_1, \ldots, a_n\}\}.$$

Claramente A_n es finito y, por consiguiente, $\bigcup_{n=1}^{\infty} A_n$ es numerable. Observe ahora que

$$\bigcup_{n=1}^{\infty} A_n \ = \ \mathcal{P}_{\text{fin}}(A)$$

y termina la prueba. ∎

Tal vez el lector sienta curiosidad en saber por qué, en el Corolario 1.2.11 (b), no se incluyó la afirmación: *el producto numerable de conjuntos numerables es numerable*. La razón es simple: tal afirmación es falsa. Por ejemplo, si tomamos $A_n = \{0, 1\}$ para todo $n \in \mathbb{N}$, entonces, como veremos más adelante, $\prod_{n=1}^{\infty} A_n = 2^{\mathbb{N}}$ es no-numerable.

Un hecho fundamental que se deriva de la parte (a) del Corolario 1.2.11 es el siguiente principio conocido con el nombre de:

> **Principio del Palomar $\mathfrak{c}$-Infinito (PP$_{\mathfrak{c}}$).** *Sea A un **conjunto no-numerable**. Si todos los elementos de A se distribuyen en una colección a lo más numerable $(A_n)_{n=1}^{\infty}$ de conjuntos, entonces existe al menos un conjunto, digamos A_{n_0}, que contiene una cantidad **infinita no-numerable** de elementos.*

Si en el principio anterior el conjunto A es infinito numerable y si consideramos sólo una colección finita $\{A_1, A_2, \ldots, A_n\}$ de conjuntos, se tiene esta otra versión infinita del Principio del Palomar.

Principio del Palomar $\aleph_0$-Infinito $(\mathbf{PP}_{\aleph_0})$. *Sea A un **conjunto infinito numerable**. Si todos los elementos de A se distribuyen en una colección finita $\{A_1, A_2, \ldots, A_n\}$ de conjuntos, entonces al menos uno de los conjuntos, digamos A_{i_0}, debe contener una cantidad **infinita numerable** de elementos*.

Otro hecho interesante que poseen los conjuntos infinitos numerables es que ellos pueden ser particionados en infinitos conjuntos cada uno de los cuales es infinito numerable. Esta afirmación es suficiente demostrarla para $\mathbb{N}_0$.

Teorema 1.2.12 (Partición de $\mathbb{N}$). *Existe una sucesión $(A_n)_{n=0}^{\infty}$ de subconjuntos de $\mathbb{N}_0$ tal que:*

(a) cada A_n es infinito,

(b) $A_m \cap A_n = \varnothing$ si $m \neq n$ y

(c) $\mathbb{N}_0 = \bigcup_{n=0}^{\infty} A_n$.

Prueba. Considere, como se muestra en la figura, el siguiente arreglo matricial de $\mathbb{N}_0$ siguiendo las diagonales:

$$\begin{array}{lcccccccc}
 & 0 & 2 & 5 & 9 & 14 & 20 & \cdots & A_0 \\
1^{\text{a}}\text{ - diagonal} & 1 & 4 & 8 & 13 & 19 & \cdots & & A_1 \\
2^{\text{a}}\text{ - diagonal} & 3 & 7 & 12 & 18 & \cdots & & & A_2 \\
3^{\text{a}}\text{ - diagonal} & 6 & 11 & 17 & \cdots & & & & A_3 \\
4^{\text{a}}\text{ - diagonal} & 10 & 16 & \cdots & & & & & A_4 \\
5^{\text{a}}\text{ - diagonal} & 15 & \cdots & & & & & & A_5 \\
 & \vdots & & & & & & & \vdots
\end{array}$$

Si $f(n,i)$ denota la entrada sobre la n-ésima fila y la i-ésima columna, entonces es fácil demostrar que

$$f(n,i) \;=\; \frac{(n+i)(n+i+1)}{2} \;+\; i$$

para todo $n, i \in \mathbb{N}_0$. En efecto, observe que para cada $k \geq 1$, existen $k+1$ elementos en la k-ésima diagonal. De aquí se sigue que, para cada $m \geq 1$, el número total de elementos que se encuentran por arriba de la m-ésima diagonal es:

$$\sum_{k=0}^{m-1}(k+1) \;=\; \sum_{k=1}^{m} k \;=\; \frac{m(m+1)}{2}.$$

Por consiguiente, si denotamos por $f(m,0)$ la entrada inicial sobre la m-ésima diagonal, tendremos que

$$f(m,0) \;=\; \frac{m(m+1)}{2},$$

la cual también se cumple si $m = 0$. Sean ahora $n, i \in \mathbb{N}_0$ con $n + i > 0$. Entonces $f(n,i)$ está en la $(n+i)$-diagonal y es el i-ésimo elemento en dicha diagonal comenzando desde $f(n+i, 0)$. De allí que, tomando $m = n + i$ en la igualdad anterior resulta que

$$f(n,i) \;=\; f(n+i,0) \;+\; i \;=\; \frac{(n+i)(n+i+1)}{2} \;+\; i.$$

La sucesión $(A_n)_{n=0}^{\infty}$ definida por $A_n = \{f(n,i) : i \in \mathbb{N}_0\}$ para cada $n \in \mathbb{N}_0$, cumple con las propiedades deseadas. ■

Ya hemos visto que $2\mathbb{N} \approx \mathbb{N}$, es decir, $\mathbb{N}$ *contiene un subconjunto propio equipotente con él*. Similarmente, $\mathbb{N} \approx \mathrm{Pri}(\mathbb{N})$, donde $\mathrm{Pri}(\mathbb{N})$ es el conjunto (infinito) de todos los números primos. Tal vez una de las características más sobresaliente que definen a los conjuntos infinitos y que ningún conjunto finito la posee viene dado por el siguiente:

Teorema 1.2.13 (Dedekind). *Un conjunto X es **infinito** precisamente cuando X es **equipotente a un subconjunto propio** de sí mismo.*

Prueba. Si X es infinito numerable, entonces (Nu_2) nos dice que cualquier subconjunto infinito de X es numerable y, en consecuencia, equipotente a X. Suponga ahora que X es no-numerable y sea $a_1 \in X$. Como X es infinito, $X \setminus \{a_1\}$ es no vacío. Seleccione $a_2 \in X \setminus \{a_1\}$. En general, sea $n \in \mathbb{N}$ con $n \geq 1$ y suponga que los términos $a_1, \ldots, a_{n-1}$ han sido escogidos. De nuevo, como $X \setminus \{a_1, \ldots, a_{n-1}\}$ es infinito, elija un a_n en dicho conjunto. De este modo se construye un subconjunto infinito numerable $A = \{a_1, a_2, \ldots, a_n, \ldots\}$ de X. Puesto que X es no-numerable, se sigue del Teorema 1.2.9 que $X \setminus A$ es infinito. Sea $f : X \to X$ definida por

$$f(x) \;=\; \begin{cases} x & \text{si } x \notin A \\ a_{n+1} & \text{si } x = a_n,\; n = 1, 2, \ldots. \end{cases}$$

Claramente f es inyectiva. Si tomamos $Y = f(X)$, resulta que $Y = X \setminus \{a_1\}$ es un subconjunto propio de X y se cumple que $Y \approx X$. Fin de la prueba. ■

Del resultado anterior se concluye que:

Corolario 1.2.14. *Si X es un **conjunto infinito**, entonces X contiene una copia de $\mathbb{N}$, es decir, existe un subconjunto A de X tal que $A \approx \mathbb{N}$.*

1.2.1. Ejemplos de Conjuntos Numerables

Nuestra definición de conjunto numerable establece la existencia de una correspondencia biunívoca entre el conjunto numerable, digamos A, y $\mathbb{N}$. Esta definición, por supuesto, está atada a un problema de existencia y, por consiguiente, no siempre es fácil determinar una tal biyección aun estando en conocimiento de que nuestro conjunto es numerable. El siguiente ejemplo expone, de manera contundente, esta situación.

(N_0) $\mathrm{Pri}(\mathbb{N})$, el **conjunto de todos los números primos**, es **numerable**. Existen muchas y variadas maneras de demostrar este resultado. La siguiente es la siempre elegante, hermosa, simple y viejita prueba dada por el propio Euclides. Suponga, para generar una contradicción, que la totalidad de los números primos es finito, digamos

$$\mathrm{Pri}(\mathbb{N}) \;=\; \{p_1,\, p_2, \ldots,\, p_n\}$$

y considere el número natural $q = p_1 p_2 \cdots p_n + 1$. Claramente $q > p_i$ para todo $i = 1, \ldots, n$ por lo que $q \notin \operatorname{Pri}(\mathbb{N})$. Veamos de inmediato que esto conduce a una contradicción. En efecto, como q no es un número primo, resulta, por el Teorema Fundamental de la Aritmética, que él es divisible por algún p_i y como $p_1 p_2 \cdots p_n$ también es divisible por p_i, tenemos que $1 = q - p_1 p_2 \cdots p_n$ es divisible por $p_i > 1$ lo cual es imposible. Por consiguiente, nuestro punto de partida de que $\operatorname{Pri}(\mathbb{N})$ era finito es falsa y, por lo tanto, él es infinito numerable. ■

Por siglos el hombre ha intentado, sin éxito (hasta ahora), conocer de algún medio o fórmula que le permita generar todos los números primos. En el momento en que el hombre esté en posesión de un tal mecanismo, muchos de los problemas difíciles que aun permanecen confusos y sin resolver en el ámbito de la Teoría de Números se podrán aclarar y solucionar (y, por supuesto, sus tarjetas de créditos y sus cuentas bancarias estarían en peligro). Por tal razón, el ejemplo anterior nos revela que aunque sepamos que un determinado conjunto es numerable, en nuestro caso $\operatorname{Pri}(\mathbb{N})$, puede resultar una tarea ardua y tremendamente difícil exhibir una función biyectiva explícita entre dicho conjunto y $\mathbb{N}$.

(N_1) $\mathbb{N}_0$ **es numerable**. La aplicación $f : \mathbb{N} \to \mathbb{N}_0$ definida por

$$f(n) = n - 1 \quad \text{para todo } n \in \mathbb{N},$$

es biyectiva. Esto sigue también del Teorema 1.2.9.

(N_2) $2\mathbb{N} = \{2, 4, 6, \ldots, 2n, \ldots\}$ **es numerable**. La función $f : \mathbb{N} \to 2\mathbb{N}$ dada por $f(n) = 2n$ para todo $n \in \mathbb{N}$ es claramente una biyección. Similarmente, el conjunto $\mathbb{N}_2 = \{1, 4, 9, \ldots, n^2, \ldots\}$ es numerable.

(N_3) $\mathbb{Z}$ **es numerable**. En efecto, la función $f : \mathbb{Z} \to \mathbb{N}$ definida por

$$f(n) = \begin{cases} 2n & \text{si } n = 1, 2, \ldots \\ 1 - 2n & \text{si } n = 0, -1, -2, \ldots. \end{cases}$$

es biyectiva.

(N_4) $\mathbb{N}^n$ y $\mathbb{Z}^n$ **son numerables** para cada $n \in \mathbb{N}$. Esto es consecuencia de (b) del Corolario 1.2.11. Por lo tanto,

$$\operatorname{card}(\mathbb{N}^n) = \operatorname{card}(\mathbb{Z}^n) = \aleph_0$$

para cualquier $n \in \mathbb{N}$.

(N_5) $\mathbb{Q}$ **es numerable**. Recordemos que

$$\mathbb{Q} = \left\{ \frac{m}{n} : m, n \in \mathbb{Z},\ n \neq 0 \right\}.$$

Tal vez una de las primeras sorpresas acerca de la numerabilidad de un conjunto lo constituye, sin duda alguna, el conjunto $\mathbb{Q}$ de los números racionales. Este conjunto, como sabemos, tiene una sorprendente e increíble propiedad: *entre dos números racionales distintos, no importa cuan cercano estén, existe entre ellos una cantidad infinita numerable de racionales distintos*. Este hecho pudiera hacer pensar que $\mathbb{Q}$ es "más numeroso" que $\mathbb{N}$: *entre dos números naturales distintos habitan, a lo sumo, sólo una cantidad finita de números naturales, pero infinitos racionales*. Y, sin embargo, como veremos de inmediato, ambos conjuntos $\mathbb{N}$ y $\mathbb{Q}$ son equipotentes.

Una manera simple de demostrar la numerabilidad de $\mathbb{Q}$, usando el Teorema 1.2.6, es considerar la aplicación $f : \mathbb{Z} \times \mathbb{N} \to \mathbb{Q}$ definida por

$$f(m, n) = \frac{m}{n} \quad \text{para todo } m \in \mathbb{Z},\ n \in \mathbb{N}$$

la cual es claramente sobreyectiva. En efecto, como $\mathbb{Z} \times \mathbb{N}$ es numerable, escoja una función biyectiva $g : \mathbb{N} \to \mathbb{Z} \times \mathbb{N}$. Se sigue entonces que la composición $f \circ g : \mathbb{N} \to \mathbb{Q}$ es sobreyectiva y entonces el Teorema 1.2.6 termina la prueba.

Existen, por supuesto, muchas otras formas distintas de demostrar que $\mathbb{Q}$ es numerable. En los siguientes ejemplos veremos algunas maneras diferentes y, a veces sorprendentes, de contar a $\mathbb{Q}$. También, el artículo de David M. Bradley, *Counting the Positive Rationals: A Brief Survey*, disponible en Internet, contiene otros ejemplos que ilustran la numerabilidad de $\mathbb{Q}$.

(N_5^0) **Numerabilidad de $\mathbb{Q}$ según Cantor** (1873). Puesto que $\mathbb{Q} = \mathbb{Q}^- \cup \{0\} \cup \mathbb{Q}^+$, donde $\mathbb{Q}^+ = \{q \in \mathbb{Q} : q > 0\}$ y $\mathbb{Q}^- = \{q \in \mathbb{Q} : q < 0\}$, es suficiente demostrar que $\mathbb{Q}^+$ es numerable. Es importante destacar que, para contar a $\mathbb{Q}^+$, sus fracciones deben estar escritas en forma irreducibles, esto quiere decir que, si $m/n \in \mathbb{Q}^+$, entonces m y n son primos relativos. Denote por $\mathbb{Q}^+_{\text{irre}}$ las fracciones de $\mathbb{Q}^+$ que son irreducibles. Sea $A_1 = \mathbb{N}$ y para cada entero $n \geq 2$, considere el conjunto

$$A_n = \left\{ \frac{m}{n} \in \mathbb{Q}^+_{\text{irre}} \,:\, m \in \mathbb{N} \right\} \setminus \bigcup_{k=1}^{n-1} A_k.$$

Como los conjuntos A_n son disjuntos dos a dos y cada uno de ellos es numerable, se sigue entonces del Teorema 1.2.10 que $\bigcup_{n=1}^{\infty} A_n = \mathbb{Q}^+_{\text{irre}}$ es numerable. Esta es la demostración clásica de que $\mathbb{Q}^+$ es numerable dada por G. Cantor en 1873. En efecto, Cantor dispuso los elementos de cada conjunto $B_n = \{m/n : m \in \mathbb{N}\}$, con $n \in \mathbb{N}$, en un arreglo matricial infinito como se sugiere en el gráfico adjunto y luego seleccionó las fracciones que allí aparecen, siguiendo las diagonales, pero teniendo la precaución de omitir todos aquellas fracciones que ya habían sido encontradas en las diagonales anteriores; esto, por supuesto, no es otra cosa que la definición de los conjuntos A_n:

$$\begin{array}{cccccccc}
B_1 & \frac{1}{1} & \frac{2}{1} & \frac{3}{1} & \frac{4}{1} & \frac{5}{1} & \cdots & A_1 \\
B_2 & \frac{1}{2} & \frac{2}{2} & \frac{3}{2} & \frac{4}{2} & \frac{5}{2} & \cdots & A_2 \\
B_3 & \frac{1}{3} & \frac{2}{3} & \frac{3}{3} & \frac{4}{3} & \frac{5}{3} & \cdots & A_3 \\
B_4 & \frac{1}{4} & \frac{2}{4} & \frac{3}{4} & \frac{4}{4} & \frac{5}{4} & \cdots & A_4 \\
B_5 & \frac{1}{5} & \frac{2}{5} & \frac{3}{5} & \frac{4}{5} & \frac{5}{5} & \cdots & A_5 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots
\end{array}$$

$$\begin{array}{cccccccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & \cdots \\
\updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \cdots \\
\frac{1}{1} & \frac{2}{1} & \frac{1}{2} & \frac{1}{3} & \frac{3}{1} & \frac{4}{1} & \frac{3}{2} & \frac{2}{3} & \frac{1}{4} & \frac{1}{5} & \frac{5}{1} & \cdots
\end{array}$$

De esta forma se concluye que $\mathbb{Q}^+$ es numerable. Aunque tal razonamiento no genera duda alguna, sin embargo, no es claro determinar, con exactitud, cuál es el número natural asignado a cada fracción. Por ejemplo, ¿a cuál número natural corresponde la fracción $123/123321123111$? Otra pregunta difícil de responder es: ¿cuál es el número racional que sigue, usando el argumento de la diagonal de Cantor, inmediatamente después de uno dado? Por esto, el procedimiento sugerido por Cantor, si bien "cuenta a todos los números racionales positivos" no permite, de manera explícita, establecer cuál es número natural que corresponde a cada fracción en $\mathbb{Q}^+$.

Se sigue de (b) del Corolario 1.2.11 que $\mathbb{Q}^n$ es **numerable** para todo entero $n \geq 1$.

(N_5^1) **Numerabilidad de $\mathbb{Q}$ según Lauwerier** (1991). La idea de Hans Lauwerier consiste en hacer una lista de los números racionales en $(0,1]$ comenzando con la fracción $1/1$ y considerando, para cada $n \in \mathbb{N}$ con $n \geq 2$, todas las fracciones m/n escritas en forma irreducibles que satisfacen la condición $1 \leq m < n$. Colóquelas en orden creciente según su denominador, es decir:

$$\frac{1}{1}, \frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{3}{4}, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, \frac{1}{6}, \frac{5}{6}, \frac{1}{7}, \frac{2}{7}, \frac{3}{7}, \frac{4}{7}, \frac{5}{7}, \frac{6}{7}, \ldots$$

Para completar la lista de los racionales en $\mathbb{Q}^+_{\text{irre}}$, intercale, entre dos fracciones consecutivas, todos los recíprocos de las fracciones aparecidas en la lista anterior comenzando con la fracción $1/2$:

$$\frac{1}{1}, \frac{2}{1}, \frac{1}{2}, \frac{3}{1}, \frac{1}{3}, \frac{3}{2}, \frac{2}{3}, \frac{4}{1}, \frac{1}{4}, \frac{4}{3}, \frac{3}{4}, \frac{5}{1}, \frac{1}{5}, \frac{5}{2}, \frac{2}{5}, \frac{5}{3}, \frac{3}{5}, \ldots$$

De esta forma se cuentan "todos" los racionales en $\mathbb{Q}^+_{\text{irre}}$. Otra manera muy similar de contar a $\mathbb{Q}^+$ es agrupar sus fracciones según la **suma** de su numerador y su denominador: primero la fracción cuya suma es 2, entonces se agrupan aquellas cuya suma es igual a 3, y así sucesivamente.

$$\underbrace{\frac{1}{1}}_{2};\ \underbrace{\frac{1}{2}, \frac{2}{1}}_{3};\ \underbrace{\frac{1}{3}, \frac{3}{1}}_{4};\ \underbrace{\frac{1}{4}, \frac{2}{3}, \frac{3}{2}, \frac{4}{1}}_{5};\ \underbrace{\frac{1}{5}, \frac{5}{1}}_{6};\ \underbrace{\frac{1}{6}, \frac{2}{5}, \frac{3}{4}, \frac{4}{3}, \frac{5}{2}, \frac{6}{1}}_{7};\ \ldots$$

Observe el gran parecido de la sucesión de Lauwerier con la sucesión de Cantor dispuesta siguiendo las diagonales.

(N_5^2) **Numerabilidad de $\mathbb{Q}$ según S. Abbott**. Una idea muy similar a la Lauwerier es la dada por S. Abbott (véase, por ejemplo, [1], pág. 24). Para cada $n \in \mathbb{N}$, considere el siguiente conjunto:

$$A_n = \left\{ \pm\frac{p}{q} \in \mathbb{Q}_{\text{irre}} \,:\, p, q \in \mathbb{N}_0,\ q \neq 0,\ p + q = n \right\}.$$

Una muestra de los elementos de algunos de los conjuntos A_n sigue a continuación:

$$A_1 = \left\{\frac{0}{1}\right\}, \quad A_2 = \left\{\frac{1}{1}, -\frac{1}{1}\right\}, \quad A_3 = \left\{\frac{1}{2}, -\frac{1}{2}, \frac{2}{1}, -\frac{2}{1}\right\}, \quad A_4 = \left\{\frac{1}{3}, -\frac{1}{3}, \frac{3}{1}, -\frac{3}{1}\right\}.$$

Lo importante es observar que:

(i) cada conjunto A_n es finito.

(ii) $A_m \cap A_n = \varnothing$ si $m \neq n$ y

(iii) cada número racional escrito en forma irreducible aparece una, y sola una vez, en exactamente uno de los conjuntos A_n.

Resulta de lo anterior que $\bigcup_{n=1}^{\infty} A_n = \mathbb{Q}$ es numerable. Una lista de los elementos de $\mathbb{Q}$ es mostrada abajo, aunque, como en el caso de la matriz de Cantor, puede ser una tarea ardua dar explícitamente una fórmula para dicha correspondencia.

$$\begin{array}{cccccccccccc}
\frac{0}{1} & \frac{1}{1} & -\frac{1}{1} & \frac{1}{2} & -\frac{1}{2} & \frac{2}{1} & -\frac{2}{1} & \frac{1}{3} & -\frac{1}{3} & \frac{3}{1} & -\frac{3}{1} & \cdots \\
\uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \cdots \\
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & \cdots
\end{array}$$

(N_5^3) **Numerabilidad de $\mathbb{Q}$ según Sagher** (1989). Tal vez un modo más práctico de contar los números racionales sea considerar el siguiente argumento debido a Yoram Sagher, (Amer. Math. Monthly, 96(9), 1989). Sea $m/n \in \mathbb{Q}_{\text{irre}}$ y suponga que m y n son enteros positivos. El Teorema Fundamental de la Aritmética nos muestra que podemos expresar, de modo único, a m y n en la forma:

$$m = p_1^{i_1} \cdots p_k^{i_k} \qquad \text{y} \qquad n = q_1^{j_1} \cdots q_l^{j_l},$$

donde los $p_1, \ldots, p_k, q_1, \ldots, q_l$ son números primos y los $i_1, \ldots, i_k, j_1, \ldots, j_l \in \{1, \ldots, 9\}$. Observe que debido a que la fracción m/n está en forma irreducible, se tiene que $p_i \neq q_j$ para todo i, j. La función de Sagher $S : \mathbb{Q} \to \mathbb{N}$ que cuenta los racionales es la siguiente: $S(1) = 1$ y

$$S\left(\frac{m}{n}\right) = p_1^{2i_1} \cdots p_k^{2i_k} \cdot q_1^{2j_1 - 1} \cdots q_l^{2j_l - 1}.$$

Es fácil establecer que S es inyectiva y sobre. Por ejemplo,

$$S(1/10^8) = 10^{15}, \qquad S(22/7) = 11^2 \cdot 2^2 \cdot 7^1 = 3388, \qquad S(3/5) = 3^2 \cdot 5^1 = 45.$$

Por otro lado, para determinar cuál es número racional asociado a un número natural N, expréselo, como antes, en la forma $N = p_1^{i_1} \cdots p_k^{i_k}$. Ahora separe los números primos con potencias pares e impares, esto es,

$$N = \left(p_{\alpha_1}^{2j_1} \cdots p_{\alpha_n}^{2j_n}\right)\left(p_{\beta_1}^{2k_l - 1} \cdots p_{\beta_s}^{2k_s - 1}\right)$$

y entonces defina

$$S^{-1}(N) = \frac{p_{\alpha_1}^{j_1} \cdots p_{\alpha_n}^{j_n}}{p_{\beta_1}^{k_l} \cdots p_{\beta_s}^{k_s}}.$$

Por ejemplo: si $N = 2^7 \cdot 3^4 \cdot 17 \cdot 7^2 \cdot 5$, entonces sepárelos como $N = 3^4 \cdot 7^2 \cdot 2^7 \cdot 5 \cdot 17$ y ahora defina

$$S^{-1}(N) = S^{-1}(3^4 \cdot 7^2 \cdot 2^7 \cdot 5 \cdot 17) = \frac{3^2 \cdot 7^1}{2^4 \cdot 5^1 \cdot 17^1}.$$

Aunque la descomposición de un número entero positivo en factores primos puede ser una tarea agotadora si dicho número es muy grande, no deja de ser interesante la función de Sagher, la cual resulta ser más atractiva que el procedimiento usado por Cantor a la hora de contar a los racionales.

(N_5^4) **Numerabilidad de $\mathbb{Q}$ usando fracciones continuas**. La Teoría de las Fracciones Continuas añade otra forma de contar los racionales. En efecto, es un hecho bien conocido que cualquier número racional p/q admite una *única representación decimal como una fracción continua finita* de la forma:

$$\frac{p}{q} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots \, a_{n-2} + \cfrac{1}{a_{n-1} + \cfrac{1}{a_n + 1}}}}},$$

donde $a_1, a_2, \ldots, a_n \in \mathbb{N}$ y a_0 es un entero no-negativo. De allí que si definimos

$$b_k = a_0 + a_1 + \cdots + a_k, \quad \text{para} \quad 0 \leq k \leq n,$$

entonces la aplicación $f : \mathbb{Q} \to \mathbb{N}$ dada por

$$f\left(\frac{p}{q}\right) = \sum_{k=0}^{n} 2^{b_k}$$

constituye una biyección entre $\mathbb{Q}$ y $\mathbb{N}$.

(N_5^5) **Numerabilidad de $\mathbb{Q}$ según Kantrowitz** (2000). Un principio general descubierto por Robert Kantrowitz en el año 2000 (Math. Mag. 73(2000), 40-42) establece que:

> *Todas las* ***palabras de longitud finita*** *que se pueden formar a partir de un alfabeto finito es* ***numerable****.*

Un ***alfabeto finito*** no es otra cosa que conjunto finito cuyos elementos, a los que llamaremos *letras*, cumplen con la siguiente regla: *se permite construir cualquier* ***palabra finita*** *con las letras del alfabeto colocándolas una delante de la otra*. El número de letras que contiene cada palabra finita se llama la *longitud* de dicha palabra (las letras que aparecen en cada palabra pueden repetirse tantas veces como se quiera).

Suponga que $\mathcal{A}$ es un *alfabeto finito* con $r \in \mathbb{N}$ letras y sea $\mathcal{P}\mathrm{al}(\mathcal{A})$ el conjunto de todas las "palabras" de longitud finita que se pueden formar con el alfabeto $\mathcal{A}$. Para cada $n \in \mathbb{N}$, sea E_n el conjunto de todas las palabras de longitud n. Entonces E_n es un conjunto finito con exactamente r^n palabras, de modo que existe una función inyectiva f_1 de E_1 sobre el segmento inicial $\{1, 2, \ldots, r\}$, una función inyectiva f_2 de E_2 sobre el siguiente segmento $\{1+r, 2+r, \ldots, r+r^2\}$ y así, sucesivamente. Observe que, para cada $n \geq 2$, la función f_n aplica el conjunto E_n sobre el segmento

$$\{1 + r + r^2 + \cdots + r^{n-1},\ 2 + r + r^2 + \cdots + r^{n-1}, \ldots, r + r + r^2 + \cdots + r^n\}$$

de longitud r^n. De esto se sigue que la colección $(E_n)_{n=1}^{\infty}$ es disjunta y entonces, por el Teorema 1.2.10, el conjunto $\bigcup_{n=1}^{\infty} E_n = \mathcal{P}\mathrm{al}(\mathcal{A})$ es numerable.

Ejemplos:

(a) Si $\mathcal{A} = \{0,1,2,3,4,5,6,7,8,9,/,-\}$, entonces $\mathcal{P}\mathrm{al}(\mathcal{A})$ es numerable. Observe que si m/n se interpreta como el cociente de números naturales y $-$ como el signo menos, entonces $\mathbb{Q}$ se puede identificar con un subconjunto de $\mathcal{P}\mathrm{al}(\mathcal{A})$ y, por lo tanto, $\mathbb{Q}$ es numerable.

(b) Sea $\mathcal{B} = \{0,1,2,3,4,5,6,7,8,9,x,+,-,(,),\wedge,/\}$ y denote por $\mathbb{Q}[x]$ el conjunto formado por todos los polinomios en la variable x con coeficientes racionales. Puesto que $\mathcal{P}\mathrm{al}(\mathcal{B})$ es numerable y $\mathbb{Q}[x]$ es un subconjunto de $\mathcal{P}\mathrm{al}(\mathcal{B})$, resulta que $\mathbb{Q}[x]$ es numerable. Por ejemplo, el polinomio $x^3 - \frac{4}{5}x^2 + 2$ puede ser representado por la palabra $x \wedge 3 - (4/5)x \wedge 2 + 2$.

(N_5^6) **Numerabilidad de $\mathbb{Q}$ según Calkin-Wilf** (2000). Una manera muy elegante de contar los elementos de $\mathbb{Q}^+_{\mathrm{irre}}$ fue la que propusieron Neil Calkin y Herbert Wilf en el año 2000 (véase, por ejemplo, [2], p. 104-108) que consiste de la siguiente sucesión, respetando el orden en que están dispuestos sus términos:

$$\mathbf{CW}\ :\ \frac{1}{1}, \frac{1}{2}, \frac{2}{1}, \frac{1}{3}, \frac{3}{2}, \frac{2}{3}, \frac{3}{1}, \frac{1}{4}, \frac{4}{3}, \frac{3}{5}, \frac{5}{2}, \frac{2}{5}, \frac{5}{3}, \frac{3}{4}, \frac{4}{1}, \ldots$$

Lo primero que llama la atención en esa lista es que:

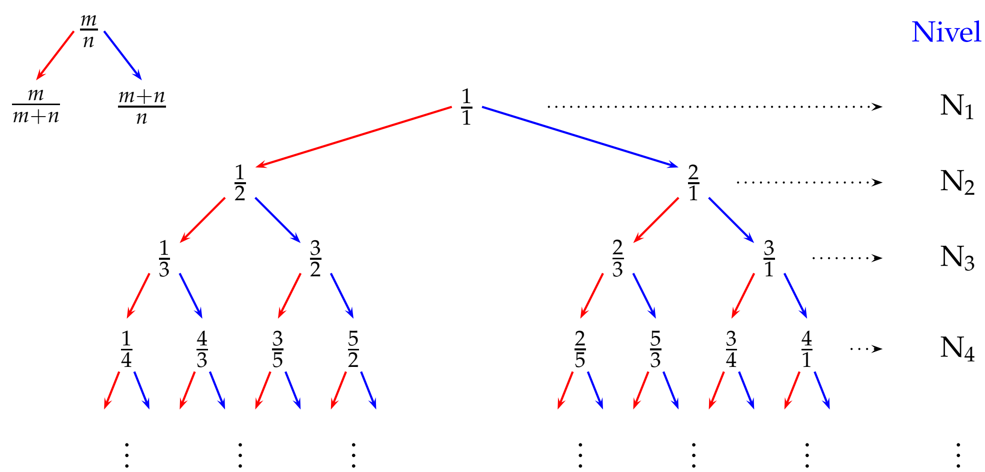
(CW_0) *El denominador de cada fracción en la lista* **CW** *es el numerador de la siguiente fracción*. Esto significa que el n-ésimo número racional en la lista **CW** se puede expresar en la forma $b(n)/b(n+1)$, $n = 0,1,2,\ldots$, donde b es una exquisita y elegante función de $\mathbb{N}$ en $\mathbb{N}$ que cuenta, para cada $n \in \mathbb{N}$, el número de formas, o maneras, de escribir el entero n como suma de potencias de 2 pero imponiendo la siguiente restricción: *cada potencia debe ser usada a lo sumo dos veces*. Por ejemplo, partiendo de $b(0) = 1$, tenemos que

$$\begin{array}{lll} b(1) = 1 & \text{pues} & 1 = 2^0, \\ b(2) = 2 & \text{ya que} & 2 = 2^1 = 2^0 + 2^0, \\ b(3) = 1 & \text{pues} & 3 = 2^1 + 2^0, \\ b(4) = 3 & \text{ya que} & 4 = 2^2 = 2^1 + 2^1 = 2^1 + 2^0 + 2^0, \\ b(5) = 2 & \text{pues} & 5 = 2^2 + 2^0 = 2^1 + 2^1 + 2^0, \\ \quad\vdots & & \end{array}$$

Pero, ¿cómo se construye tan extraña y afortunada lista? La forma más simple de construir la lista **CW** es a través del así llamado **árbol de Calkin-Wilf**, que consiste en aplicar las siguientes dos reglas sencillas:

(CW_1) $\frac{1}{1}$ *es el tope del árbol*.

(CW_2) *Cada vértice (o nodo) es un número racional* $\frac{m}{n}$ *que posee dos ramas: la* ***rama izquierda*** *que se define por* $\frac{m}{m+n}$, *mientras que la* ***rama derecha*** *viene dada por* $\frac{m+n}{n}$.

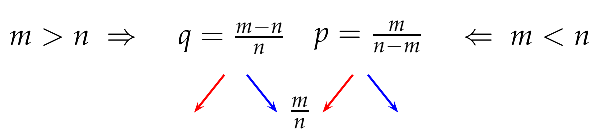


Árbol de Calkin-Wilf

Observe que si $x = m/n$, entonces sus ramas izquierda y derecha se pueden escribir, respectivamente, en la forma:

$$\frac{m}{m+n} = \frac{\frac{m}{n}}{\frac{m}{n}+\frac{n}{n}} = \frac{x}{x+1} \qquad \text{y} \qquad \frac{m+n}{n} = \frac{m}{n}+1 = x+1.$$

Nótese también que si m/n es la **rama izquierda** de algún vértice p, entonces $m < n$ y $p = \frac{m}{n-m}$. Similarmente, si m/n es la **rama derecha** del vértice q, entonces $m > n$ y $q = \frac{m-n}{n}$.



La lista **CW** se genera, siguiendo el árbol de Calkin-Wilf, del modo siguiente: coloque la fracción del nivel 1 en el primer lugar y continúe colocando, de izquierda a derecha, las fracciones de los niveles siguientes separadas por comas.

Algunas de las formidables e increíbles propiedades que posee el árbol de Calkin-Wilf son mostradas a continuación.

(CW_3) *Todas las fracciones que aparecen en el árbol de Calkin-Wilf son **irreducibles**, es decir, pertenecen a* $\mathbb{Q}^+_{\text{irre}}$.

Prueba. Es claro que $\frac{1}{1}$, la primera fracción en el tope del árbol, está en forma irreducible. Suponga, para generar una contradicción, que alguna fracción en el árbol, digamos m/n, no está en forma irreducible. Esto significa que existe algún entero $k > 1$ tal que

$$m = km' \qquad \text{y} \qquad n = kn'$$

con $m', n' \in \mathbb{N}$. Para fijar idea, suponga también que m/n aparece en el nivel N_j y que todas las fracciones por encima de dicho nivel están en forma reducida. Por supuesto, la fracción m/n proviene de un vértice p/q que está en el nivel N_{j-1} y, por consiguiente, está, por hipótesis, en forma reducida. Si m/n es la rama izquierda de p/q, entonces

$$\frac{p}{q} = \frac{m}{n-m} = \frac{km'}{k(n'-m')},$$

contrario a nuestra suposición. De forma similar, si m/n es la rama derecha de p/q, entonces

$$\frac{p}{q} = \frac{m-n}{n} = \frac{k(m'-n')}{kn'},$$

obteniéndose de nuevo una contradicción. Esto prueba nuestra afirmación. ∎

Otra forma de demostrar lo anterior es observar que si m y n son primos relativos, también lo son m y $m+n$, así como también, n y $m+n$.

(CW$_4$) ***Cualquier fracción*** *en* $\mathbb{Q}^+_{\text{irre}}$ *aparece en el árbol, es decir, pertenece a* **CW**.

Prueba. Suponga que la conclusión es falsa, es decir, que existe al menos una fracción irreducible en $\mathbb{Q}^+$ que no está en **CW**. Denote por R el conjunto de todas las fracciones irreducible en $\mathbb{Q}^+$ que no aparecen en el árbol. Nuestra hipótesis nos dice que $R \neq \varnothing$. De todas las fracciones que están en R hay, al menos una, cuyo denominador es el más pequeño: denótelo por n y reúna a todas esas fracciones en un conjunto R_n. De éste último conjunto de fracciones existe al menos una cuyo numerador es el más pequeño: designe a esa fracción por m/n. Observe que:

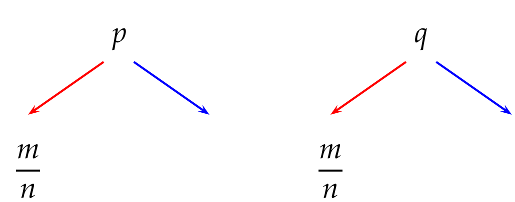
(i) Si $m > n$, entonces $(m-n)/n \in R$. En efecto, si $(m-n)/n \notin R$, entonces dicho elemento estaría en **CW** y, en consecuencia, sus dos ramas $(m-n)/m$ y m/n también estarían en **CW** lo cual es imposible por la elección de m/n. Ahora bien, el hecho de que $(m-n)/n \in R$ y $m-n < m$ nos revela la existencia de un elemento en R cuyo numerador es menor que m, lo que de nuevo conduce a una contradicción por la elección de m/n.

(ii) Similarmente, si $m < n$, entonces $m/(n-m) \in R$ lo que conduce, razonando como en el caso anterior, a una contradicción. La conclusión es clara: $R = \varnothing$ y con esto finaliza la prueba de (CW$_4$). ∎

(CW$_5$) *Cada fracción en el árbol aparece una, y sola una, vez*.

Prueba. Observe, en primer lugar, que $1/1$ aparece una sola vez en el árbol ya que cada rama izquierda, digamos $m/(m+n)$ es menor que 1 y cada rama derecha $(p+q)/n$ es mayor que 1 cualquiera sea el vértice m/n. Suponga ahora que existe al menos una fracción en el árbol que aparece, por lo menos, dos veces y sea R el conjunto de tales fracciones. Como antes, sea R_1 el subconjunto de R formado por las fracciones que tienen el denominador más pequeño y, de estos últimos, sea m/n el elemento de R_1 que posee el numerador más pequeño.

(i) Si $m < n$, entonces m/n es la rama izquierda de al menos dos vértices, digamos p y q.



De aquí se sigue que

$$p = \frac{m}{n-m} = q,$$

y, por lo tanto, la fracción $m/(n-m)$ se repite más de una vez. Esto prueba que $m/(n-m) \in R$ y como su denominador es menor que n se obtiene entonces una contradicción con la elección de m/n.

(ii) Si $m > n$, entonces un razonamiento similar al anterior conduce de nuevo a una contradicción. Por consiguiente, $R = \varnothing$ y termina la prueba. ■

Las afirmaciones (CW_3), (CW_4) y (CW_5) se combinan para obtener una biyección entre $\mathbb{N}$ y $\mathbf{CW} = \mathbb{Q}^+$ como se muestra en la figura adjunta:

$$\begin{array}{ccccccccccc}
1 & \frac{1}{2} & 2 & \frac{1}{3} & \frac{3}{2} & \frac{2}{3} & 3 & \frac{1}{4} & \frac{4}{3} & \frac{3}{5} & \cdots \\
\uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \cdots \\
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & \cdots
\end{array}$$

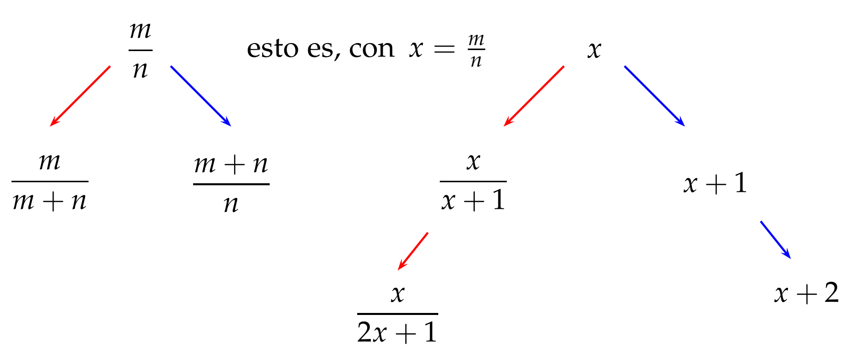
El lector tiene todo el derecho a preguntarse: bueno, y *¿cuál es, explícitamente, la biyección entre* $\mathbb{N}$ *y* $\mathbb{Q}^+$*?* La respuesta es fácil si podemos determinar, cada vez que elijamos una fracción en el árbol, cuál es la fracción que le sigue. He aquí el modo de obtenerla.

(CW_6) *Para cada* x *en el árbol, el siguiente elemento o sucesor viene dado por*

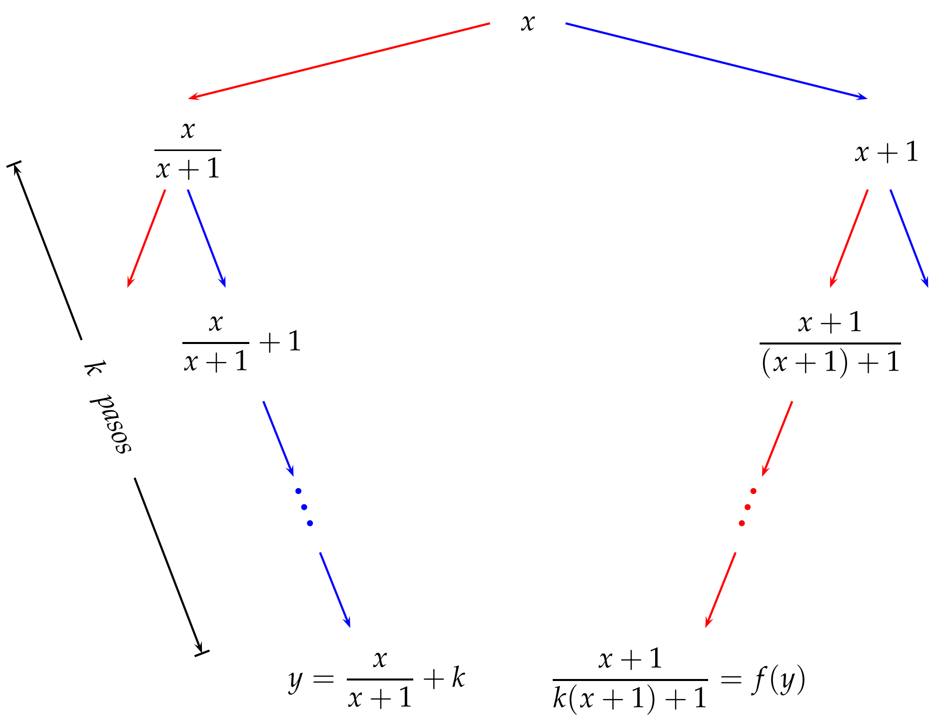
$$f(x) = \frac{1}{2[x]+1-x},$$

donde $[x]$ *es la parte entera de* x.

Prueba. Fijemos una fracción $x = m/n$ del árbol **CW** con sus dos ramas $\frac{x}{x+1}$ y $x+1$ como se muestra en el gráfico adjunto.



De éste diagrama se puede deducir, sin dificultad, que la rama derecha de x es $x+1$ y que la rama derecha de $x+1$ es $x+2$, es decir, la segunda rama derecha de x es $x+2$ y, en general, la k-ésima rama derecha de x es $x+k$. Similarmente, la rama izquierda de x es $\frac{x}{x+1}$ y la rama izquierda de $\frac{x}{x+1}$ es $\frac{x}{2x+1}$ y, en general, la k-ésima rama izquierda de x es $\frac{x}{kx+1}$.



Veamos ahora cómo obtenemos, dada una fracción en el árbol de Calkin-Wilf, su sucesora inmediata. Suponga que nuestra fracción es la rama izquierda de $x = m/n$, es decir, la fracción $m/(m+n)$ la cual podemos reescribir en la forma

$$y \;=\; \frac{x}{x+1}.$$

Resulta entonces que su sucesora es la fracción

$$x+1 \;=\; \frac{1}{1-y},$$

lo que nos proporciona una fórmula simple para hallar el siguiente elemento en el árbol siempre que y sea la rama izquierda del vértice $x = m/n$. Observe que, en este caso, $0 \le y < 1$ y, por consiguiente, $[y] = 0$, de modo que su sucesor $\frac{1}{1-y}$ viene dado por

$$f(y) \;=\; \frac{1}{2[y]+1-y} \;=\; \frac{1}{1-y}.$$

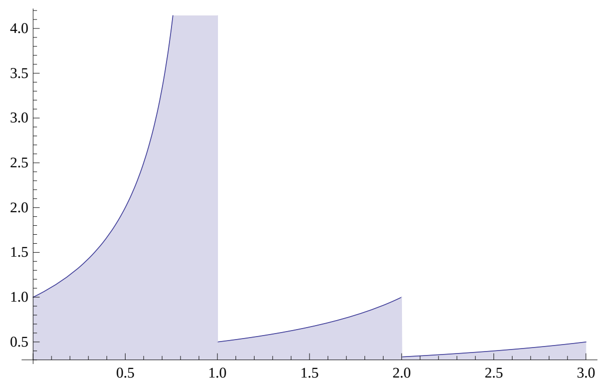
Suponga ahora que y es alguna rama derecha en el árbol y queremos hallar su sucesora inmediata $f(y)$ en dicho árbol. Para determinar $f(y)$ debemos, en primer lugar, hallar el **primer vértice** del árbol, digamos x, desde cuyas dos ramas y después de k pasos tomando las ramas derechas de $x/(x+1)$ y las ramas izquierdas de $x+1$, respectivamente, conducen a y y $f(y)$ (véase el gráfico anterior). Se tiene entonces que

$$y \;=\; \frac{x}{x+1}+k \;=\; \frac{m}{m+n}+k,$$

de donde se sigue que $[y] = k$ ya que $m/(m+n) < 1$ y, por lo tanto, $\dfrac{x}{x+1} = y - [y]$. De esto resulta que

$$f(y) \;=\; \frac{x+1}{k(x+1)+1} \;=\; \frac{1}{k+\frac{1}{x+1}} \;=\; \frac{1}{[y]+1-\frac{x}{x+1}} \;=\; \frac{1}{2[y]-y+1}$$

la cual representa la hermosa y sencilla fórmula descubierta por Moshe Newman en el 2003, a propósito de un problema formulado por Donald E. Knuth en el 2001.



Definiendo $f^n = f(f^{n-1})$ para cada $n \geq 2$, donde $f^1 = f$, se tiene entonces que

$$\begin{array}{lllll} f(0) = 1, & f^2(0) = \dfrac{1}{2}, & f^3(0) = 2, & f^4(0) = \dfrac{1}{3}, & f^5(0) = \dfrac{3}{2}, \\ f^6(0) = \dfrac{2}{3}, & f^7(0) = 3, & f^8(0) = \dfrac{1}{4}, & f^9(0) = \dfrac{4}{3}, & f^{10}(0) = \dfrac{5}{3}, \end{array}$$

por lo que la función $\mathrm{CW} : \mathbb{N} \to \mathbb{Q}^+$ definida por

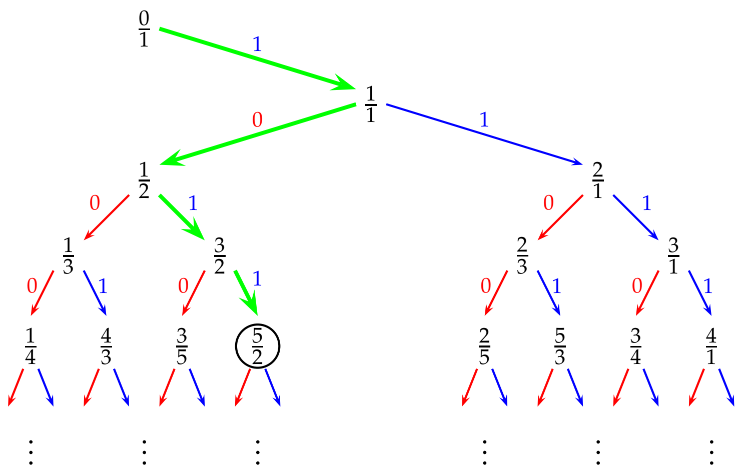
$$\mathrm{CW}(n) \;=\; f^n(0), \quad \text{para cada } n \in \mathbb{N}$$

es la biyección buscada. Observe que determinar el número racional asociado a cada n puede tomar muchísimos pasos (si n es muy grande) ya que la función CW es una relación de recurrencia. ■

(CW_7) *Otra propiedad notable del árbol de Calkin-Wilf*. Otra forma de hallar la n-ésima fracción en la sucesión **CW** es proceder del modo siguiente: exprese a n en su forma binaria, digamos

$$n \;=\; b_k 2^k + b_{k-1}2^{k-1} + \cdots + b_1 2 + b_0 \;=\; \left(b_k b_{k-1} \ldots b_1 b_0\right)_2$$

y entonces siga el camino de las ramas en el árbol de Calkin-Wilf determinado por los dígitos $b_k, \ldots, b_0$ pero comenzando con la fracción $\frac{0}{1}$ como se muestra en el gráfico adjunto. Aquí, $b_i = 1$ indica que "*debes tomar la rama derecha*", mientras que $b_i = 0$ significa que "*debes tomar la rama* *izquierda*" en cada vértice.



Por ejemplo, si $n = 11$, entonces como $11 = 1 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2 + 1 \cdot 2^0 = (1011)_2$ resulta que la fracción que corresponde a $n = 11$ es $\frac{5}{2}$ (véase el gráfico anterior). ■

(N_5^7) **Numerabilidad de $\mathbb{Q}$ según Farey**. Otro interesante y fascinante modo de contar los racionales que no se repiten lo constituye la inigualable sucesión de Farey. El conjunto de las ***fracciones, o sucesión, de Farey de orden*** **n**, denotado por $\mathcal{F}_n$, es el conjunto de todas las fracciones irreducibles en el intervalo cerrado $[0, 1]$, colocadas en *orden creciente*, cuyos denominadores no exceden a n. Así,

$$\frac{p}{q} \in \mathcal{F}_n \qquad \Leftrightarrow \qquad \frac{p}{q} \in \mathbb{Q}_{\text{irre}}, \quad 0 \leq p < q \leq n.$$

Por ejemplo,

$$\begin{aligned}
\mathcal{F}_1 &= \left\{\frac{\mathbf{0}}{\mathbf{1}}, \frac{\mathbf{1}}{\mathbf{1}}\right\} \\
\mathcal{F}_2 &= \left\{\frac{\mathbf{0}}{\mathbf{1}}, \frac{\mathbf{1}}{\mathbf{2}}, \frac{\mathbf{1}}{\mathbf{1}}\right\} \\
\mathcal{F}_3 &= \left\{\frac{\mathbf{0}}{\mathbf{1}}, \frac{\mathbf{1}}{\mathbf{3}}, \frac{\mathbf{1}}{\mathbf{2}}, \frac{\mathbf{2}}{\mathbf{3}}, \frac{\mathbf{1}}{\mathbf{1}}\right\} \\
\mathcal{F}_4 &= \left\{\frac{\mathbf{0}}{\mathbf{1}}, \frac{\mathbf{1}}{\mathbf{4}}, \frac{\mathbf{1}}{\mathbf{3}}, \frac{\mathbf{1}}{\mathbf{2}}, \frac{\mathbf{2}}{\mathbf{3}}, \frac{\mathbf{3}}{\mathbf{4}}, \frac{\mathbf{1}}{\mathbf{1}}\right\}
\end{aligned}$$

$$\begin{aligned}
\mathcal{F}_5 &= \left\{\frac{\mathbf{0}}{\mathbf{1}}, \frac{\mathbf{1}}{\mathbf{5}}, \frac{\mathbf{1}}{\mathbf{4}}, \frac{\mathbf{1}}{\mathbf{3}}, \frac{\mathbf{2}}{\mathbf{5}}, \frac{\mathbf{1}}{\mathbf{2}}, \frac{\mathbf{3}}{\mathbf{5}}, \frac{\mathbf{2}}{\mathbf{3}}, \frac{\mathbf{3}}{\mathbf{4}}, \frac{\mathbf{4}}{\mathbf{5}}, \frac{\mathbf{1}}{\mathbf{1}}\right\} \\
\mathcal{F}_6 &= \left\{\frac{\mathbf{0}}{\mathbf{1}}, \frac{\mathbf{1}}{\mathbf{6}}, \frac{\mathbf{1}}{\mathbf{5}}, \frac{\mathbf{1}}{\mathbf{4}}, \frac{\mathbf{1}}{\mathbf{3}}, \frac{\mathbf{2}}{\mathbf{5}}, \frac{\mathbf{1}}{\mathbf{2}}, \frac{\mathbf{3}}{\mathbf{5}}, \frac{\mathbf{2}}{\mathbf{3}}, \frac{\mathbf{3}}{\mathbf{4}}, \frac{\mathbf{4}}{\mathbf{5}}, \frac{\mathbf{5}}{\mathbf{6}}, \frac{\mathbf{1}}{\mathbf{1}}\right\} \\
&\vdots
\end{aligned}$$

Observe que:

(i) *$\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}_n$ para cualquier $n \geq 1$.*

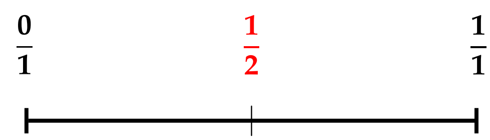
Una de las cosas curiosas que poseen cada uno de los conjuntos $\mathcal{F}_n$ con $n \geq 2$ y que fue descubierta por John Farey en 1816 es que tal conjunto se construye a partir de $\mathcal{F}_{n-1}$ copiando, en primer lugar, todos sus elementos en $\mathcal{F}_n$ y luego insertando las **medianas** entre fracciones consecutivas de $\mathcal{F}_{n-1}$ pero solamente cuando los denominadores de tales medianas no excedan a n. Pero, ¿qué son las medianas? El detalle en la construcción de los conjuntos $\mathcal{F}_n$ consiste en una forma muy peculiar de sumar fracciones, llamada la mediana de dos fracciones y definida del modo siguiente:

Definición 1.2.15. *La **mediana** de dos fracciones p/q y p'/q' se define como*

$$\frac{p}{q} \oplus_F \frac{p'}{q'} = \frac{p+p'}{q+q'}.$$

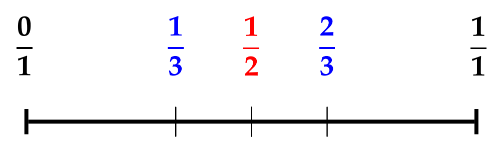
Pues bien, partiendo de las dos fracciones $\frac{0}{1}$ y $\frac{1}{1}$ se generan, con la ayuda de las medianas, los conjuntos $\mathcal{F}_n$ para $n \geq 2$. En efecto, $\mathcal{F}_2$ se construye a partir de $\mathcal{F}_1$ insertando, entre las dos fracciones de $\mathcal{F}_1$, la fracción $\frac{1}{2}$ la cual se obtiene "sumando" las dos fracciones de $\mathcal{F}_1$ al estilo Farey, es decir,

$$\frac{0}{1} \oplus_F \frac{1}{1} = \frac{1}{2}$$



$\mathcal{F}_3$ se construye de $\mathcal{F}_2$ insertando entre ellos las medianas de los elementos de $\mathcal{F}_2$, es decir,

$$\frac{0}{1} \oplus_F \frac{1}{2} = \frac{1}{3}, \qquad \frac{1}{2} \oplus_F \frac{1}{1} = \frac{2}{3}$$



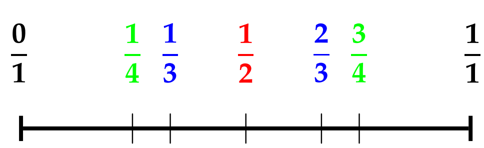
Un procedimiento similar se hace para construir $\mathcal{F}_4$ pero aquí hay que tener la precaución de evitar incorporar medianas cuyos denominadores excedan a 4. Por ejemplo, las medianas

$$\frac{1}{3} \oplus_F \frac{1}{2} = \frac{2}{5} \qquad \text{y} \qquad \frac{1}{2} \oplus_F \frac{2}{3} = \frac{3}{5}$$

no pueden ser incluidas en $\mathcal{F}_4$ ya que su denominador excede a 4. Recuerde que, por definición, los denominadores de las fracciones de $\mathcal{F}_4$ no deben ser mayores que 4, por lo que las únicas medianas permitidas en $\mathcal{F}_4$ son

$$\frac{0}{1} \oplus_F \frac{1}{3} = \frac{1}{4} \qquad \text{y} \qquad \frac{2}{3} \oplus_F \frac{1}{1} = \frac{3}{4}$$

Luego, las fracciones que viven en $\mathcal{F}_4$ son



De esta forma se construye, para cada $n \geq 2$, la sucesión de Farey $\mathcal{F}_n$ a partir de $\mathcal{F}_{n-1}$.

Es importante tener en cuenta el siguiente hecho respecto a las medianas.

(1) Si p/q y p'/q' son dos fracciones arbitrarias con $0 \leq \frac{p}{q} < \frac{p'}{q'}$, entonces

$$\frac{p}{q} < \frac{p+p'}{q+q'} < \frac{p'}{q'}.$$

Prueba. El hecho de que $\frac{p}{q} < \frac{p'}{q'}$ nos garantiza que $p'q - pq > 0$ y, por lo tanto,

$$\frac{p+p'}{q+q'} - \frac{p}{q} = \frac{p'q-pq}{q(p+q)} > 0 \qquad \text{y} \qquad \frac{p'}{q'} - \frac{p+p'}{q+q'} = \frac{p'q-pq}{q'(p+q)} > 0.$$

Fin de la prueba. ■

Esta propiedad era muy bien conocida por Arquímedes y también por algunos geómetras de la India, lo que hace presumir que tal vez otras personas la conociesen. Además, ella fue utilizada por Nicolás Chuquet en el año de 1484 para aproximar a $\sqrt{n}$ cuando $n \leq 14$.

Observe que, por definición, los racionales que aparecen en cualquier sucesión de Farey de orden n están dispuestos en orden creciente y que, por lo tanto, *no es posible que una misma fracción en $\mathcal{F}_n$ aparezca en dos lugares diferentes*. Por consiguiente:

Corolario 1.2.16. *Cada **fracción irreducible** en $[0,1]$ aparece una, y solo una, vez en $\bigcup_{n=1}^{\infty} \mathcal{F}_n$, es decir,*

$$\mathbb{Q}_{\text{irre}} \cap [0,1] = \bigcup_{n=1}^{\infty} \mathcal{F}_n.$$

En particular, $\mathbb{Q}_{\text{irre}} \cap [0,1]$ es numerable.

Prueba. En efecto, si m/n es una fracción irreducible en $(0,1)$, entonces claramente $m/n \in \mathcal{F}_n$ y el hecho de que ella aparezca una sola vez es consecuencia de la observación anterior. Finalmente, puesto que $\bigcup_{n=1}^{\infty} \mathcal{F}_n$ es numerable, entonces $\mathbb{Q}_{[0,1]} = \mathbb{Q} \cap [0,1]$ es numerable. ■

Otra de las propiedades fundamentales que posee cada sucesión de Farey es la siguiente:

(2) Si p/q y p'/q' son dos términos consecutivos en $\mathcal{F}_n$, entonces

$$|p'q - pq| = 1.$$

Prueba. La demostración la haremos por inducción sobre n. La conclusión se cumple trivialmente si $n = 1$. Suponga ahora que el resultado es cierto para cualquier par de fracciones consecutivas en $\mathcal{F}_n$ y probemos que el resultado también se cumple para $\mathcal{F}_{n+1}$. Suponga que $\frac{p}{q}$ y $\frac{p'}{q'}$ son dos fracciones consecutivas en $\mathcal{F}_n$ y escribamos

$$\mathcal{F}_n = \left\{ \ldots, \frac{p}{q}, \frac{p'}{q'}, \ldots \right\}.$$

Por hipótesis, $|pq' - p'q| = 1$. Sólamente existen dos posibilidades: que $q + q' \leq n+1$, o bien, que $q + q' > n+1$.

(a) si $q+q' \le n+1$, entonces

$$\mathcal{F}_{n+1} \;=\; \left\{\dots, \frac{p}{q}, \frac{p+p'}{q+q'}, \frac{p'}{q'}, \dots\right\}$$

y se cumple que $|p(q+q') - q(p+p')| = |pq' - p'q| = 1$. Similarmente,

$$|p'(q+q') - q'(p+p')| \;=\; |pq' - p'q| = 1.$$

(b) si $q+q' > n+1$, entonces $\frac{p+p'}{q+q'} \notin \mathcal{F}_{n+1}$ y, por lo tanto,

$$\mathcal{F}_{n+1} \;=\; \left\{\dots, \frac{p}{q}, \frac{p'}{q'}, \dots\right\}.$$

Esto nos dice que $|pq' - p'q| = 1$ y el resultado es válido para $\mathcal{F}_{n+1}$. Fin de la prueba. ■

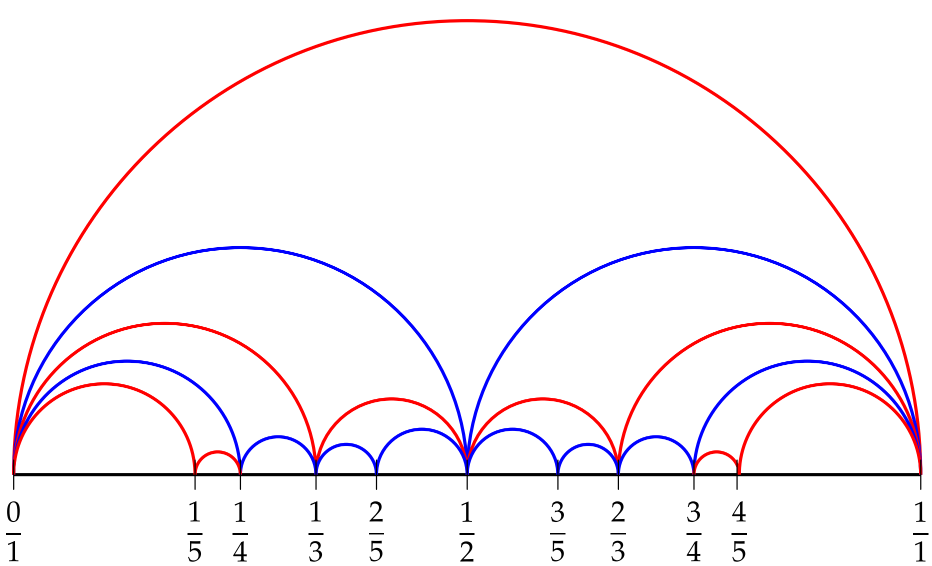
(3) *Toda fracción de Farey es* ***simétrica respecto a la fracción*** $1/2$ *en el siguiente sentido: cada fracción* $\frac{p}{q} \in \mathcal{F}_n$ *con* $n > 2$ *y* $\frac{p}{q} \ne \frac{1}{2}$, *posee su simétrica* $1 - \frac{p}{q} \in \mathcal{F}_n$ *y ambas son equidistantes de* $1/2$.

Suponga que $a_{n,k}$ denota el término de $\mathcal{F}_n$ que ocupa el lugar k contado de izquierda a derecha, entonces, por ejemplo:

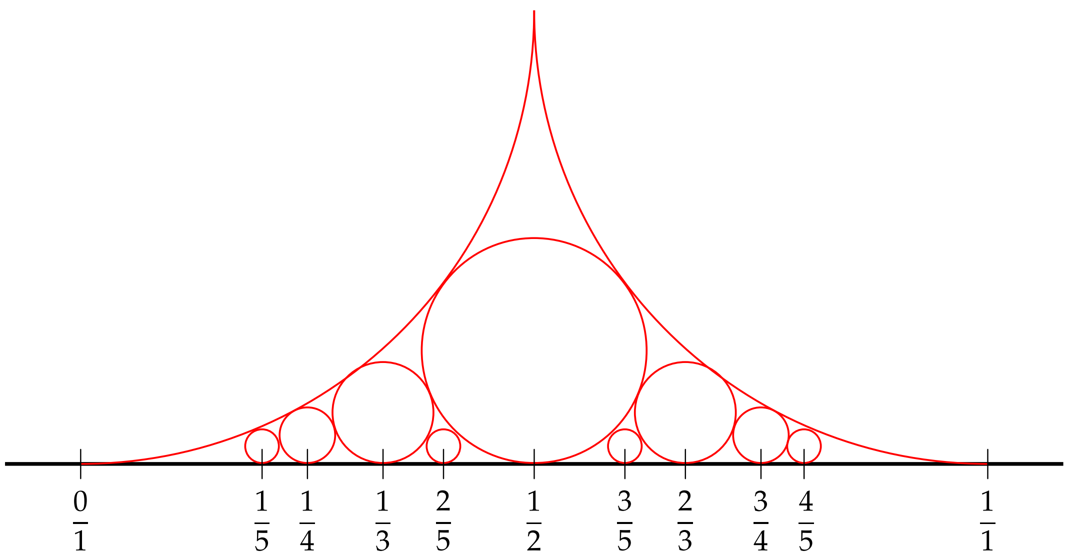
$$a_{4,3} \;=\; \frac{1}{3}, \qquad a_{4,5} \;=\; \frac{2}{3} \;=\; 1 - \frac{1}{3}, \qquad a_{6,6} \;=\; \frac{2}{5}, \qquad a_{6,8} \;=\; \frac{3}{5} \;=\; 1 - \frac{2}{5}$$

y así con todos los elementos de $\mathcal{F}_n$.

Existen hermosas y variadas formas de hacer diagramas con las sucesiones de Farey. Por ejemplo, comenzando con $\mathcal{F}_1$, dibuje un semicírculo de diámetro 1 pasando por los puntos $(0,0)$ y $(1,0)$. Luego, cuando se haya construido $\mathcal{F}_2$, dibuje dos semicírculos tangentes, ambos de diámetro $1/2$, el primero pasando por los puntos $(0,0)$ y $(1/2,0)$, mientras que el segundo semicírculo pase por $(1/2,0)$ y $(1,0)$. Continúe construyendo 4, 8, etc. semicírculos pasando por cada par de puntos consecutuvos de $\mathcal{F}_3$, $\mathcal{F}_4$, etc. como se muestra en el gráfico adjunto.



Las sucesiones de Farey han sido objeto de intensos estudios. Muchas son sus propiedades y aplicaciones (véase, por ejemplo, [68]). El siguiente es otro caso curioso de lo que se puede hacer con ellas, la cual fue encontrada en 1938 por Lester R. Ford. En plano y sobre el eje X dibuje, por cada fracción $p/q \in \mathcal{F}_n$, un círculo tangente al punto $(p/q, 0)$ y de radio $1/q^2$. Resulta que estos círculos, llamados **círculos de Ford**, tienen la propiedad de que cualesquiera sean las fracciones consecutivas p/q y p'/q', los círculos construidos sobre ellos son tangentes (véase la figura adjunta).



Nota Curiosa. Un comentario final respecto a las fracciones de Farey. En Londres existió, desde 1704 hasta 1841, una revista llamada "*The Ladies Diary*", también conocida como "*The Woman Almanache*", que se dedicaba a mostrar calendarios, acertijos, problemas matemáticos y otros "Entretenimientos Particulares, Peculiarmente Adaptados para el Uso y Diversión de la Mujer-Bella". En la edición de 1747 apareció el siguiente problema propuesto por J. May de Amsterdam:

> *"Se requiere encontrar (por un teorema general) el número de fracciones irreducibles, cada una menor que la unidad, tal que el denominador más grande sea menor que* 100*".*

Este problema permaneció sin resolverse por espacio de 4 años hasta que, en 1751, un caballero de nombre R. Flitcon dio una solución satisfactoria: *existen 3003 fracciones irreducibles*, aunque no suministró una lista de ellas y, menos aun, una fórmula explícita para hallarlos. Correspondió al matemático francés Charles Haros crear una tabla de tales fracciones, la cual fue publicada en el año de 1802 en el "*Journal de l'Ecole Polytechique*" que contenía, además, sus respectivas aproximaciones decimales. Para lograr su objetivo Haros tuvo que usar las medianas para su construcción. Posteriormente, un inglés de nombre Henry Goodwyn, se ocupó de confeccionar su propia tabla de fracciones irreducibles cuyos denominadores comenzaban desde 1 y terminaban en 1024, la cual publicó en la "Royal Society" el 25 de Abril de 1816. No había transcurrido un mes desde su publicación cuando un geólogo inglés, John Farey, publicó una nota en "*The Philosophical Magazine and Journal*" titulada: "Sobre una curiosa Propiedad de las Fracciones Vulgares", donde se percataba sobre el uso de las medianas para construir nuevas fracciones irreducible pero sin ofrecer ninguna prueba de ello. Por suerte, la nota de Farey fue publicada en la revista francesa "*Bulletin de la Societé Philomatique*" y el curioso Cauchy la leyó y demostró que el método de las medianas sugerido por Farey era correcto y, desde entonces, tal sucesión lleva su nombre.

(N_5^8) **Numerabilidad de $\mathbb{Q}$ según Stern-Brocot**. Existe otro árbol, tan genial, fascinante y hermoso como el árbol de Calkin-Wilf, que permite *no sólo contar, sino también ordenar de menor a mayor* en una sucesión creciente, todos los racionales en $\mathbb{Q}^+_{\text{irre}}$, es decir, cada nivel del árbol, llamémoslo $\mathcal{SB}_n$, contiene $2^n - 1$ elementos de $\mathbb{Q}^+_{\text{irre}}$ tal que, para cada $n \in \mathbb{N}$ los elementos de $\mathcal{SB}_n$ están dispuestos en orden creciente:

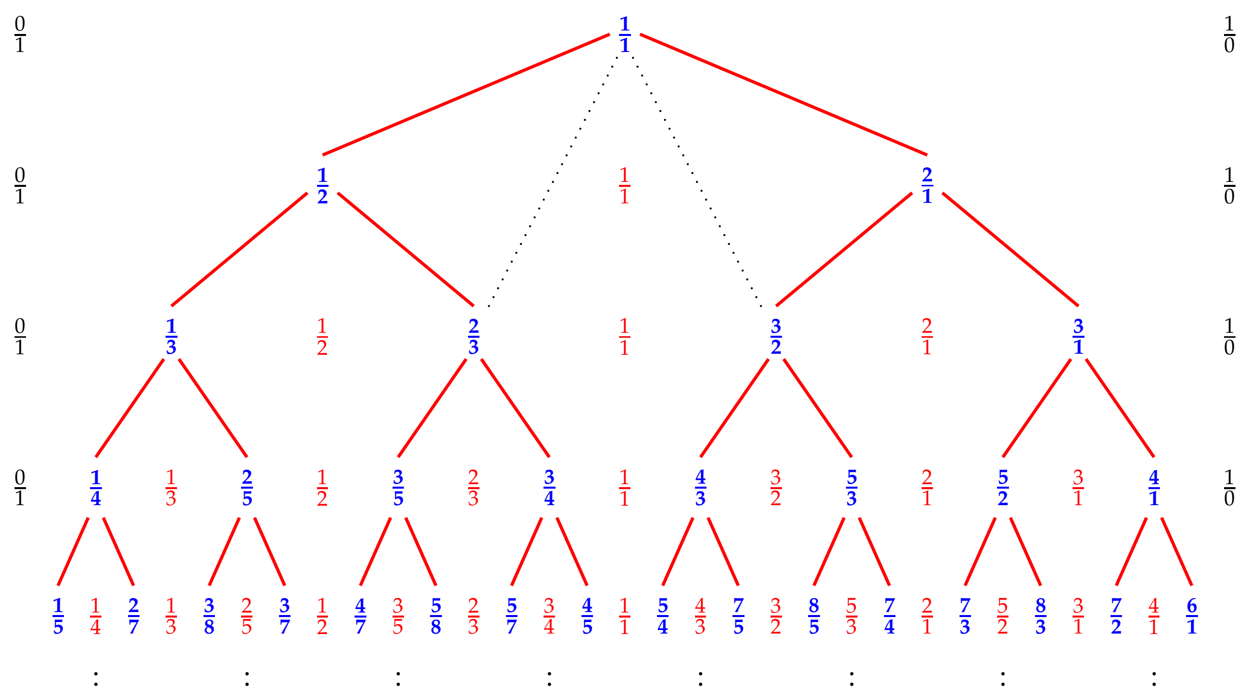
$$r_1 < r_2 < \cdots < r_{2^n-1},$$

$$\mathcal{SB}_1 \subseteq \mathcal{SB}_2 \subseteq \cdots \subseteq \mathcal{SB}_n \subseteq \cdots \qquad \text{y} \qquad \bigcup_{n=1}^{\infty} \mathcal{SB}_n = \mathbb{Q}^+_{\text{irre}}$$

Este árbol fue descrito por primera vez por el matemático alemán Moritz Stern en 1858 y tres años después, e independientemente, por el relojero francés Achielle Brocot. Las reglas para la construcción del árbol de Stern-Brocot son similares a la construcción de las sucesiones de Farey de orden n pero con una excepción: la suma de los denominadores de las medianas no requieren que estén acotadas por n. Dichas reglas se pueden expresar del modo siguiente:

(SB_1) *Se comienza con las fracciones* $\dfrac{0}{1}$ *y* $\dfrac{1}{0}$.

(SB_2) *Se inserta, tantas veces como sea necesaria, la mediana entre dos fracciones consecutivas.*



Árbol de Stern-Brocot

Pues bien, partiendo de las dos fracciones $\frac{0}{1}$ y $\frac{1}{0}$ se generan, con la ayuda de las medianas, las siguientes fracciones: en el primer paso se obtiene la mediana $\frac{1}{1} = \frac{0}{1} + \frac{1}{0}$ la cual se añade a las dos fracciones anteriores para formar:

$$\frac{0}{1}, \ \mathbf{\frac{1}{1}}, \ \frac{1}{0}. \tag{1}$$

En el siguiente paso añaden, al conjunto anterior, las medianas $\frac{1}{2}$ y $\frac{2}{1}$ obteniéndose:

$$\frac{0}{1}, \mathbf{\frac{1}{2}}, \mathbf{\frac{1}{1}}, \mathbf{\frac{2}{1}}, \frac{1}{0}. \tag{2}$$

En el próximo paso se añaden a las fracciones que aparecen en (2) las cuatro nuevas medianas obtenidas de éste: $\frac{1}{3}, \frac{2}{3}, \frac{3}{2}$ y $\frac{3}{1}$, esto es:

$$\frac{0}{1}, \mathbf{\frac{1}{3}}, \mathbf{\frac{1}{2}}, \mathbf{\frac{2}{3}}, \mathbf{\frac{1}{1}}, \mathbf{\frac{3}{2}}, \mathbf{\frac{2}{1}}, \mathbf{\frac{3}{1}}, \frac{1}{0},$$

y se continúa ad infinitum. El ***árbol de Stern-Brocot*** *es el que se construye sólo con las nuevas medianas* tal como se muestra en el gráfico anterior.

Si por cada $n \in \mathbb{N}$, denotamos por $\mathcal{SB}_n$ el conjunto de todas las fracciones de $\mathcal{SB}_{n-1}$ añadiéndole las medianas de cada dos fracciones consecutivas de $\mathcal{SB}_{n-1}$, se tiene que:

$$\begin{aligned}
\mathcal{SB}_0 &= \left\{\mathbf{\frac{0}{1}}, \mathbf{\frac{1}{0}}\right\}\\
\mathcal{SB}_1 &= \left\{\mathbf{\frac{0}{1}}, \mathbf{\frac{1}{1}}, \mathbf{\frac{1}{0}}\right\}\\
\mathcal{SB}_2 &= \left\{\mathbf{\frac{0}{1}}, \mathbf{\frac{1}{2}}, \mathbf{\frac{1}{1}}, \mathbf{\frac{2}{1}}, \mathbf{\frac{1}{0}}\right\}\\
\mathcal{SB}_3 &= \left\{\mathbf{\frac{0}{1}}, \mathbf{\frac{1}{3}}, \mathbf{\frac{1}{2}}, \mathbf{\frac{2}{3}}, \mathbf{\frac{1}{1}}, \mathbf{\frac{3}{2}}, \mathbf{\frac{2}{1}}, \mathbf{\frac{3}{1}}, \mathbf{\frac{1}{0}}\right\}\\
\mathcal{SB}_4 &= \left\{\mathbf{\frac{0}{1}}, \mathbf{\frac{1}{4}}, \mathbf{\frac{1}{3}}, \mathbf{\frac{2}{5}}, \mathbf{\frac{1}{2}}, \mathbf{\frac{3}{5}}, \mathbf{\frac{2}{3}}, \mathbf{\frac{3}{4}}, \mathbf{\frac{1}{1}}, \mathbf{\frac{4}{3}}, \mathbf{\frac{3}{2}}, \mathbf{\frac{5}{3}}, \mathbf{\frac{2}{1}}, \mathbf{\frac{5}{2}}, \mathbf{\frac{3}{1}}, \mathbf{\frac{4}{1}}, \mathbf{\frac{1}{0}}\right\}\\
&\ \vdots
\end{aligned}$$

Observe que cada conjunto $\mathcal{SB}_n$ contiene a todos los vértices de los niveles anteriores del árbol. Con estas definiciones es fácil probar las siguientes conclusiones:

(a) $\mathcal{SB}_{n-1} \subseteq \mathcal{SB}_n$ para todo $n \geq 1$.

(b) $\mathcal{F}_n \subseteq \mathcal{SB}_n$ para todo $n \geq 1$.

(c) $p/q \in \mathcal{SB}_n \quad \Leftrightarrow \quad q/p \in \mathcal{SB}_n$.

(d) $\operatorname{card}(\mathcal{SB}_n) = 2^n + 1$ para todo $n \geq 1$.

(e) Cada conjunto $\mathcal{SB}_n$ está ordenado de menor a mayor y todas sus elementos son fracciones irreducibles.

Nota: Observe que (c) indica una *simetría inversa* respecto a la fracción $\frac{1}{1}$, es decir, la fracción p/q y su inversa q/p son ambas equidistante de $\frac{1}{1}$, lo cual significa que el número de fracciones que existen entre cada una de ellas y $1/1$ es el mismo.

Algunas de las propiedades notables de los conjuntos $\mathcal{SB}_n$, árbol de Stern-Brocot, y similares a las sucesiones de Farey, son las siguientes:

(SB$_3$) Si m/n y m'/n' son dos términos consecutivos en cualquier conjunto $\mathcal{SB}_k$, entonces

$$m'n - mn' = 1 \tag{SB}$$

y, por lo tanto,

$$\frac{m'}{n'} - \frac{m}{n} = \frac{1}{n \cdot n'}.$$

Prueba. La relación es cierta para el caso de las fracciones $\frac{0}{1}$, $\frac{1}{0}$ ya que $1 \cdot 1 - 0 \cdot 0 = 1$. Observe ahora que cuando se inserta una nueva mediana $(m+m')/(n+n')$ entre dos fracciones consecutivas m/n y m'/n', entonces los casos que se deben chequear son:

$$(m+m')n - m(n+n') = 1 \qquad \text{y} \qquad m'(n+n') - (m+m')n' = 1$$

y ambas ecuaciones son equivalentes a la condición original $m'n - mn' = 1$. De allí que (SB) es invariante en cualquier conjunto $\mathcal{SB}_k$. ■

(SB$_4$) *Cada nueva mediana insertada en cada paso en la construcción del árbol de Stern-Brocot está en forma reducida*.

Prueba. Suponga que alguna mediana del árbol, digamos $(m+m')/(n+n')$, no está en forma reducida. Esto significa que existe un $k > 1$ tal que

$$\frac{m+m'}{n+n'} = \frac{kp}{kq}$$

donde p y q son primos relativos. Puesto que

$$\frac{m}{n} < \frac{m+m'}{n+n'} = \frac{kp}{kq},$$

se sigue de (SB$_3$) que

$$1 = (kp) \cdot n - m \cdot (kq) = k \cdot (pn - mq)$$

y como $k > 1$, resulta que $pn - mq = 1/k$ no es un entero, lo cual es absurdo. ■

(SB$_5$) Si m/n, m''/n'' y m'/n' son tres términos consecutivos en cualquier conjunto $\mathcal{SB}_k$, entonces

$$\frac{m''}{n''} = \frac{m+m'}{n+n'}.$$

Prueba. Esto es así por definición. ■

(SB$_6$) *Todas las fracciones construidas en el k-ésimo paso están ordenadas de menor a mayor*.

Prueba. Sigue de lo anterior. ■

(SB$_7$) *Ninguna fracción del árbol de Stern-Brocot aparece más de una vez*.

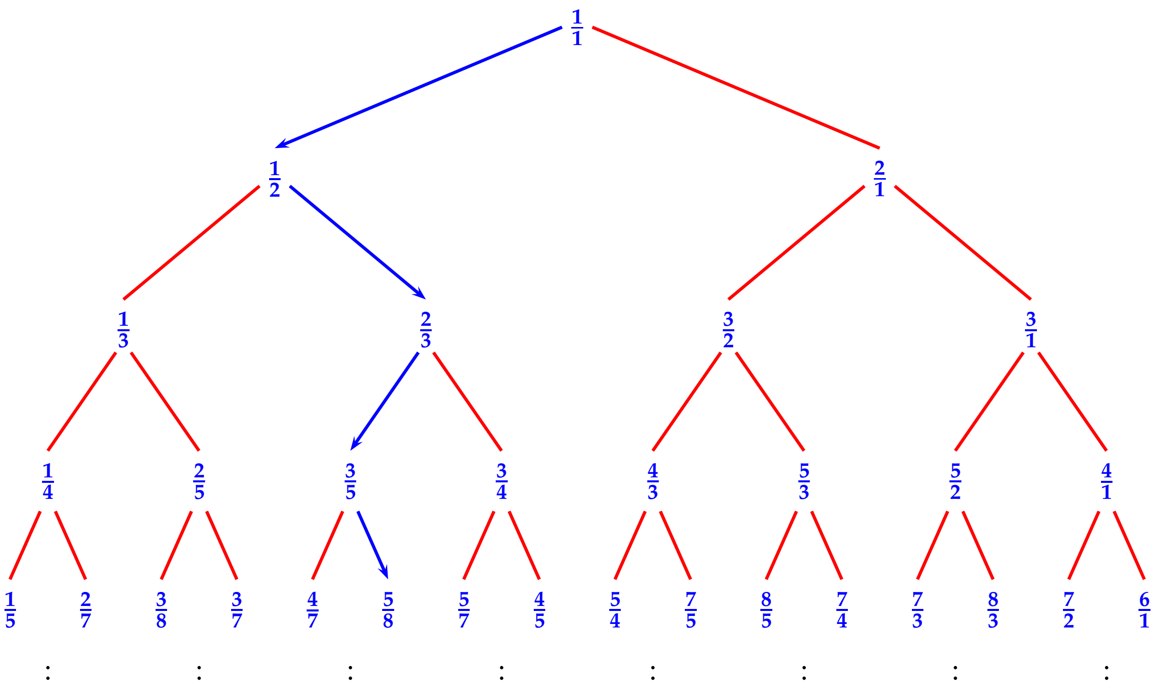
Prueba. Es consecuencia de (SB$_6$). ■

(SB$_8$) *Cada fracción en $\mathbb{Q}^+_{\text{irre}}$ aparece en el árbol de Stern-Brocot*.

Prueba. Sea $m/n \in \mathbb{Q}^+_{\text{irre}}$. Por simetría es suficiente demostrar que si $m/n \leq 1$, entonces $m/n \in \mathcal{SB}_n$. Pero si $m/n \leq 1$, se sigue del Corolario 1.2.16 que $m/n \in \mathcal{F}_n \subseteq \mathcal{SB}_n$ y termina la prueba. ■

Esto último nos indica que existe una biyección entre $\mathbb{Q}^+_{\text{irre}}$ y las fracciones del árbol de Stern-Brocot y, por lo tanto, $\mathbb{Q}^+_{\text{irre}}$ es numerable.

Son muchas las propiedades que posee el árbol de Stern-Brocot. Queremos finalizar este pequeño análisis con la siguiente curiosidad: si se sigue en zig-zag la rama infinita del árbol de Stern-Brocot comenzando con la fracción $\frac{1}{1}$ como se indica en el gráfico,



es decir,

$$\frac{1}{1} \;\to\; \frac{1}{2} \;\to\; \frac{2}{3} \;\to\; \frac{3}{5} \;\to\; \frac{5}{8} \;\to\; \frac{8}{13} \;\to\; \frac{13}{21} \;\to\cdots$$

resulta que los numeradores y denominadores de esas fracciones no son otra cosa que los exquisitos e increíbles números de Fibonacci f_n los cuales, como se sabe, se definen para cada $n \geq 3$ como:

$$f_n \;=\; f_{n-1} + f_{n-2}, \qquad f_1 \;=\; 1, \quad f_2 \;=\; 1.$$

Por ejemplo, los 15 primeros números de Fibonacci son:

$$1, \quad 1, \quad 2, \quad 3, \quad 5, \quad 8, \quad 13, \quad 21, \quad 34, \quad 55, \quad 89, \quad 144, \quad 233, \quad 377, \quad 610, \ldots$$

Dos hechos que son fascinantes y a la vez sorprendentes de tales números son los siguientes:

(1º) Para cada entero $n \geq 1$, f_n se puede representar en la forma:

$$f_n \;=\; \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right],$$

donde $\phi = (1+\sqrt{5})(2)$ y $\phi' = (1-\sqrt{5})/2$ son los números de oro.

(2º) ¿Qué relación existe entre los números de Fibonacci y el triángulo de Pascal? Recordemos que una representación del triángulo de Pascal toma, por lo general, la forma siguiente:

$$\begin{array}{c}
1\\
1 \quad 1\\
1 \quad 2 \quad 1\\
1 \quad 3 \quad 3 \quad 1\\
1 \quad 4 \quad 6 \quad 4 \quad 1\\
1 \quad 5 \quad 10 \quad 10 \quad 5 \quad 1\\
1 \quad 6 \quad 15 \quad 20 \quad 15 \quad 6 \quad 1\\
1 \quad 7 \quad 21 \quad 35 \quad 35 \quad 21 \quad 7 \quad 1\\
1 \quad 8 \quad 28 \quad 56 \quad 70 \quad 56 \quad 28 \quad 8 \quad 1\\
\vdots \qquad\qquad \vdots \qquad\qquad \vdots
\end{array}$$

Si en la representación anterior desplazamos sus elementos tal cual se muestra en la siguiente gráfica y consideramos las suma de las diagonales mostradas en el mismo, se obtienen los susodichos números de Fibonacci.

$$\begin{array}{lllllllll}
1 & 1 & 2 & 3 & 5 & 8 & 13 & 21 & 34\\
1 & 1 & & & & & & & \\
1 & 2 & 1 & & & & & & \\
1 & 3 & 3 & 1 & & & & & \\
1 & 4 & 6 & 4 & 1 & & & & \\
1 & 5 & 10 & 10 & 5 & 1 & & & \\
1 & 6 & 15 & 20 & 15 & 6 & 1 & & \\
1 & 7 & 21 & 35 & 35 & 21 & 7 & 1 & \\
1 & 8 & 28 & 56 & 70 & 56 & 28 & 8 & 1\\
\vdots & & \vdots & & \vdots & & \vdots & &
\end{array}$$

Pues bien, la sucesión de fracciones de Farey en zig-zag es $(f_{n+1}/f_{n+2})_{n=0}^{\infty}$ la cual converge al fabuloso **número de oro**

$$-\phi' \;=\; \frac{\sqrt{5}-1}{2} \;\approx\; 0{,}6180339887\ldots$$

cuando $n \to \infty$.

La siguiente pregunta es natural: *Dado cualquier número irrational* α*, ¿existe alguna rama infinita del árbol de Stern-Brocot que converge a* α*?*

(N_6) $\mathcal{A}_{\text{lg}}(\mathbb{R})$, el **conjunto de los números algebraicos**, es **numerable** (Dedekind).

Denotemos por $\mathbb{Z}[x]$ el conjunto de todos los polinomios con coeficientes enteros, es decir, $p \in \mathbb{Z}[x]$ si

$$p(x) \;=\; a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

donde los coeficientes $a_0, a_1, \ldots, a_n$ son números enteros y no todos son ceros. Un número real x_0 se dice **algebraico** si es raíz de algún polinomio $p \in \mathbb{Z}[x]$, es decir, si existe $p \in \mathbb{Z}[x]$ tal que $p(x_0) = 0$. Por ejemplo, cualquier número racional p/q es algebraico ya que él satisface la ecuación $p(x) = qx - p = 0$. También ocurre que muchos números irracionales son algebraicos. Por ejemplo, $\sqrt{2}$ es algebraico ya que el polinomio $p(x) = x^2 - 2$ satisface $p(\sqrt{2}) = 0$.

Cualquier número real que no es algebraico se llama **trascendente**, es decir, un número trascendente es aquel que no satisface ninguna ecuación algebraica con coeficientes enteros. La primera prueba de la existencia de números trascendentes fue dada por Joseph Liouville quien, en 1844, descubrió una clase muy extensa de tales números. Por ejemplo, todos lo números de la forma

$$1 + \frac{1}{n} + \frac{1}{n^2} + \frac{1}{n^6} + \frac{1}{n^{24}} + \cdots + \frac{1}{n^{k!}} + \cdots$$

son trascendentes, donde n es cualquier número entero mayor que 1. Aunque este descubrimiento de Liouville genera muchos números trascendentes, sigue siendo un reto difícil para un matemático demostrar que un sospechoso particular es o no trascendente. Por tal razón, cuando Charles Hermite demostró, en 1873, que *e es trascendente*, los matemáticos no dejaron de asombrarse ante la belleza y sencillez de la prueba. Nueve años más tarde del descubrimiento de Hermite, en 1882, Ferdinand Lindemann demostró que π pertenecía al mismo clan. Una prueba de la existencia de números trascendentes fue dada a conocer por G. Cantor sin dar ningún ejemplo concreto de uno de ellos. Hasta el día de hoy constituye la prueba más elemental de la existencia de tales números. Un hecho curioso es el siguiente: algún tiempo antes de que Cantor demostrara que $\mathbb{R}$ era no-numerable, ya él había probado la numerabilidad de $\mathbb{Q}$, él le preguntó a Dedekind si podía probar la no-numerabilidad de $\mathbb{R}$. Dedekind trabajó un tiempo en ese problema pero no logró demostrarlo. Sin embargo, Dedekind le ofreció a Cantor la demostración de que el conjuntos de los números algebraicos es numerable.

Prueba de que $\mathcal{A}_{\mathrm{lg}}(\mathbb{R})$ es numerable. La idea de la demostración es considerar, para cada $n \in \mathbb{N}$, el conjunto $\mathbb{Z}_n[x]$ de todos los polinomios en $\mathbb{Z}[x]$ de grado n, esto es,

$$\mathbb{Z}_n[x] = \left\{ \sum_{j=0}^{n} a_j x^j \ : \ a_0, a_1, \ldots, a_n \in \mathbb{Z},\ a_n \neq 0 \right\}.$$

Veamos que dicho conjunto es numerable. En efecto, considere la aplicación $f : \mathbb{Z}_n[x] \to \mathbb{Z}^{n+1}$ definida por

$$f\left(\sum_{j=0}^{n} a_j x^j \right) = (a_0, a_1, \ldots, a_n).$$

Claramente f_n es inyectiva. Además, como $\mathbb{Z}^{n+1}$ es numerable, existe una función biyectiva $\Phi : \mathbb{Z}^{n+1} \to \mathbb{N}$. Definamos ahora $\varphi : \mathbb{Z}_n[x] \to \mathbb{N}$ por $\varphi(p) = (\Phi \circ f)(p)$ para cualquier $p \in \mathbb{Z}_n[x]$. Puesto que la composición de dos funciones inyectivas es inyectiva, resulta del Teorema 1.2.6 que $\mathbb{Z}_n[x]$ es numerable. Se sigue del Corolario 1.2.11 que

$$\mathbb{Z}[x] = \bigcup_{n=1}^{\infty} \mathbb{Z}_n[x]$$

es numerable. Finalmente, observe que $\mathcal{A}_{\mathrm{lg}}(\mathbb{R})$ se puede escribir como sigue:

$$\mathcal{A}_{\mathrm{lg}}(\mathbb{R}) = \bigcup_{p \in \mathbb{Z}[x]} Z(p),$$

donde $Z(p) = \{x \in \mathbb{R} : p(x) = 0\}$. Puesto que cada polinomio $p \in \mathbb{Z}[x]$ posee a lo sumo un número finito raíces, resulta que $\mathcal{A}_{\mathrm{lg}}(\mathbb{R})$ es numerable por ser unión numerable de conjuntos finitos. Esto termina la prueba. ■

1.2.2. El Teorema de Cantor y Conjuntos No-numerables

Hasta este momento hemos podido constatar la existencia de una cantidad inmensa de conjuntos infinitos numerables pero no así la existencia de conjuntos no-numerables. Por lo tanto, la noción general de infinito sólo tendrá relevancia si logramos probar que *no todos los conjuntos infinitos son equipotentes a* $\mathbb{N}$, es decir, si se logra demostrar que existen conjuntos no-numerables. En esta corta sección demostraremos, usando un poderoso resultado de Cantor, la existencia de conjuntos no-numerables. Es importante destacar que demostrar que un cierto candidato, digamos B, es no-numerable siguiendo nuestra definición puede resultar una tarea difícil pues ello conlleva a verificar, en primer lugar, que él es infinito y, segundo, que *ninguna función* $f : B \to \mathbb{N}$ *puede ser biyectiva*. Por fortuna, existen otras formas distintas e interesantes de verificar la no-numerabilidad de un conjunto: por ejemplo, uno de ellos consiste en utilizar el método de prueba por contradicción, es decir, suponer que alguna extraña función $f : B \to \mathbb{N}$ es biyectiva y generar una contradicción. Una segunda opción es saber que un determinado conjunto, digamos A, es no-numerable y entonces encontrar alguna función biyectiva $f : B \to A$. Un tercer modo de verificar la no-numerabilidad de un conjunto es utilizando el así llamado "**Método de la Diagonal de Cantor**", el cual fue ideado por G. Cantor para demostrar que $\mathbb{R}$ es un conjunto no-numerable, etc.

Recordemos que si X y Y son conjuntos no vacíos, entonces Y^X *consiste del conjunto de todas las funciones* $f : X \to Y$. En particular, si $Y = \{0, 1\}$, escribiremos 2^X en lugar de $\{0, 1\}^X$. Un resultado sencillo pero que es de suma importancia es el siguiente:

Teorema 1.2.17 (Cantor). *Sea X un conjunto no vacío. Entonces $\mathscr{P}(X)$ y 2^X son equipotentes; es decir,*

$$\operatorname{card}\big(\mathscr{P}(X)\big) \;=\; \operatorname{card}\big(2^X\big).$$

Prueba. La aplicación $\varphi : \mathscr{P}(X) \to 2^X$ definida por

$$\varphi(A) \;=\; \chi_A$$

para todo $A \in \mathscr{P}(X)$ es biyectiva. En efecto, sean $A, B \in \mathscr{P}(X)$ con $A \neq B$. Esto significa que existe un $x \in X$ tal que $x \in A \setminus B$, o $x \in B \setminus A$. Sin perder generalidad suponga que $x \in A \setminus B$. Entonces $x \in A$ y $x \notin B$ de modo que

$$\varphi(A)(x) \;=\; \chi_A(x) \;=\; 1 \qquad\qquad \text{y} \qquad\qquad \varphi(B)(x) \;=\; \chi_B(x) \;=\; 0$$

y, en consecuencia, $\varphi(A) \neq \varphi(B)$. Esto prueba que φ es inyectiva. Para demostrar la sobreyectividad, tome cualquier $f \in 2^X$ y considere el conjunto

$$A \;=\; \big\{x \in X \,:\, f(x) = 1\big\}.$$

Es claro que $\varphi(A) = f$ y termina la prueba. ■

De la misma forma en que la igualdad de dos conjuntos A y B se puede descomponer en dos inclusiones: $A \subseteq B$ y $B \subseteq A$, la noción de cardinalidad, a pesar de no haber sido definida en forma precisa, también se puede descomponer en dos desigualdades y ser usadas para comparar conjuntos.

Definición 1.2.18. *Sean A y B conjuntos no vacíos. Definimos*

$$\operatorname{card}(A) \leq \operatorname{card}(B)$$

*si existe una **aplicación inyectiva** $f : A \to B$.*

Escribiremos

$$\operatorname{card}(A) < \operatorname{card}(B)$$

para indicar que $\operatorname{card}(A) \leq \operatorname{card}(B)$ pero $\operatorname{card}(A) \neq \operatorname{card}(B)$. Observe que esto último significa que: *ninguna función inyectiva* $f : A \to B$ *puede ser sobreyectiva*. Una manera de ilustrar estos hechos es observando, por ejemplo, que:

(a) $\operatorname{card}(A) < \operatorname{card}(B)$ *si* A *es* ***finito*** *y* B *es* ***infinito***.

(b) $\aleph_0 \leq \operatorname{card}(X)$ si X *es cualquier* ***conjunto infinito*** (Corolario 1.2.14), y

(c) $\aleph_0 < \operatorname{card}(X)$ si X *es cualquier conjunto* ***infinito no-numerable***.

Otro resultado extraordinario, llamado el Teorema de Cantor por E. Zermelo pero cuya prueba fue dada por primera vez por G. Hessenberg y que posee consecuencias profundas, establece que *no existe sobreyección entre un conjunto* X *y su potencia* $\mathscr{P}(X)$, o dicho de otro modo: *en cualquier conjunto no vacío* X *siempre existen más subconjuntos que elementos*.

Teorema 1.2.19 (Teorema de Cantor). *Cualquier conjunto arbitrario no vacío* X *es* ***equipotente a un subconjunto propio*** *de* $\mathscr{P}(X)$*, pero no es equipotente a* $\mathscr{P}(X)$*; es decir,*

$$\operatorname{card}(X) < \operatorname{card}(\mathscr{P}(X)).$$

Prueba. Que X no es equipotente a $\mathscr{P}(X)$ significa que: *ninguna función* $f : X \to \mathscr{P}(X)$ *puede ser sobreyectiva*. Para ver esto, suponga que $f : X \to \mathscr{P}(X)$ es una función sobreyectiva. Para cada $x \in X$, $f(x)$ es un subconjunto de X que puede o no contener a x. Considere entonces el conjunto $F = \{x \in X : x \notin f(x)\}$. Afirmamos que no existe $x \in X$ tal que $F = f(x)$. En efecto, como estamos asumiendo que f es sobreyectiva, existe algún $x_0 \in X$ para el cual $F = f(x_0)$. Observe, sin embargo, que: $x_0 \in F$ si, y sólo si, $x_0 \notin f(x_0) = F$. Esta contradicción establece que f no puede ser sobreyectiva y, en consecuencia, X y $\mathscr{P}(X)$ no son equipotentes. Más aun, si definimos $g : X \to \mathscr{P}(X)$ por $g(x) = \{x\}$ para todo $x \in X$, resulta que g es inyectiva, lo cual prueba que X es equipotente a un subconjunto propio de $\mathscr{P}(X)$ y concluye la prueba. ■

Una consecuencia inmediata del Teorema de Cantor es que:

Corolario 1.2.20. $\mathscr{P}(\mathbb{N})$ ***es no-numerable***.

Prueba. Si $\mathscr{P}(\mathbb{N})$ fuese numerable, entonces el Teorema 1.2.6 nos garantizaría la existencia de una aplicación inyectiva de $\mathbb{N}$ sobre $\mathscr{P}(\mathbb{N})$ lo que estaría en contradicción con el Teorema de Cantor. ■

Otra consecuencia del Teorema de Cantor es impedir la existencia del conjunto de todos los conjuntos. En efecto, suponga que tal conjunto existe y llamémoslo $\mathfrak{U}$. Por definición, todo subconjunto de $\mathfrak{U}$ es asimismo un elemento de $\mathfrak{U}$ y, en consecuencia, $\mathscr{P}(\mathfrak{U})$ es un subconjunto de $\mathfrak{U}$, es decir,

$$\mathscr{P}(\mathfrak{U}) \subseteq \mathfrak{U}$$

lo cual implica que $\operatorname{card}(\mathscr{P}(\mathfrak{U})) = \operatorname{card}(2^{\mathfrak{U}}) \leq \operatorname{card}(\mathfrak{U})$. Pero entonces, esto contradice el Teorema de Cantor el cual afirma que $\operatorname{card}(\mathfrak{U}) < \operatorname{card}(2^{\mathfrak{U}})$. Por esto, $\mathfrak{U}$ no existe como conjunto. Similarmente, el Teorema de Cantor impide la formación, como conjunto, de todos los números cardinales, de todos los conjuntos que son equipotentes a un conjunto dado, etc.

Recordemos que cuando intentamos definir la cardinalidad de un conjunto no-numerable tropezábamos con el hecho de que no existía un objeto que los identificase a todos, sino que existía una cantidad "*infinita de infinitos*" en orden creciente. Pues bien, estamos listo para mostrar tal colección. Una de las consecuencias extraordinarias del Teorema 1.2.19 se evidencia en el siguiente hecho:

$$\aleph_0 \ < \ \operatorname{card}(\mathscr{P}(\mathbb{N})) \ < \ \operatorname{card}(\mathscr{P}(\mathscr{P}(\mathbb{N}))) \ < \ \cdots,$$

es decir, existe una **jerarquía infinita**, estrictamente creciente, de **conjuntos infinitos**.

Teorema 1.2.21. *Sean A, B y C conjuntos arbitrarios. Entonces:*

(a) $\operatorname{card}(A) \leq \operatorname{card}(A)$.

(b) $\operatorname{card}(A) \leq \operatorname{card}(B)$ *y* $\operatorname{card}(B) \leq \operatorname{card}(C) \ \Rightarrow \ \operatorname{card}(A) \leq \operatorname{card}(C)$.

(c) $A \subseteq B \ \Rightarrow \ \operatorname{card}(A) \leq \operatorname{card}(B)$.

Observe que el resultado anterior establece que $\leq$ es *casi* una relación de orden entre conjuntos. Que ello se comporte en realidad como una relación de orden, se debe a un extraordinario resultado debido a Cantor, Bernstein y Schröeder.

Teorema 1.2.22 (Cantor-Bernstein-Schröeder). *Sean A y B conjuntos arbitrarios y suponga que*

$$\operatorname{card}(A) \ \leq \ \operatorname{card}(B) \qquad \text{y} \qquad \operatorname{card}(B) \ \leq \ \operatorname{card}(A).$$

Entonces $\operatorname{card}(A) = \operatorname{card}(B)$.

Prueba. La demostración la haremos en dos pasos. Primero demostraremos que: *si X, A' y B' son conjuntos arbitrarios con*

$$X \subseteq B' \subseteq A' \qquad \text{y} \qquad \operatorname{card}(X) \ = \ \operatorname{card}(A'),$$

entonces $\operatorname{card}(B') = \operatorname{card}(A')$. En efecto, sea $f : A' \to X$ una función biyectiva y defina las sucesiones de conjuntos $(A_n)_{n=0}^{\infty}$ y $(B_n)_{n=0}^{\infty}$ del modo siguiente:

$$A_0 \ = \ A' \qquad \text{y} \qquad B_0 \ = \ B'$$

y para cada $n \in \mathbb{N}_0$, sean

$$A_{n+1} \ = \ f(A_n), \qquad B_{n+1} \ = \ f(B_n). \tag{CBS}$$

Puesto que $A_0 \supseteq B_0 \supseteq X$, se sigue de (CBS) e inducción que $A_n \supseteq B_n \supseteq A_{n+1}$ para todo $n \in \mathbb{N}_0$. Considere ahora el conjunto $C_n = A_n \setminus B_n$ para cada entero $n \geq 0$ y sean

$$C \ = \ \bigcup_{n=0}^{\infty} C_n \qquad \text{y} \qquad D \ = \ A' \setminus C.$$

Por (CBS) y el hecho de que f es inyectiva, se tiene que $f(C_n) = C_{n+1}$ y, por lo tanto,

$$f(C) \ = \ \bigcup_{n=1}^{\infty} C_n.$$

Finalmente, defina $g : A' \to B'$ por

$$g(x) \ = \ \begin{cases} f(x) & \text{si } x \in C \\ x & \text{si } x \in D. \end{cases}$$

Puesto que $g|_C$ y $g|_D$ son funciones inyectivas y sus imágenes son conjuntos disjuntos, se deduce que g es una función inyectiva de A' sobre $f(C) \cup D = B'$.

El segundo paso es suponer que $\text{card}(A) \leq \text{card}(B)$ y $\text{card}(B) \leq \text{card}(A)$. Esto significa que existen funciones inyectivas $f : A \to B$ y $g : B \to A$ y, por consiguiente, la función $g \circ f : A \to A$ es inyectiva y $\text{card}(A) = \text{card}(g(f(A)))$. Más aun, como $g(f(A)) \subseteq g(B) \subseteq A$, se sigue de la primera parte que $\text{card}(g(B)) = \text{card}(A)$. Por otro lado, puesto que $\text{card}(B) = \text{card}(g(B))$ se concluye entonces que $\text{card}(A) = \text{card}(B)$ y termina la prueba. ■

Otra manera de formular el Teorema de Cantor-Bernstein-Schröeder es del modo siguiente: *Si $f : A \to B$ y $g : B \to A$ son* ***funciones inyectivas****, entonces existe una* ***biyección*** *de A en B.*

1.2.3. Ejemplos de Conjuntos No-numerables

En lo que sigue mostraremos algunos ejemplos concretos de conjuntos que son no-numerables. En el transcurso de estas notas aparecerán otros ejemplos interesantes de tales conjuntos. Comenzaremos demostrando un hecho que es fundamental en matemáticas:

Teorema 1.2.23 (Cantor). *$\mathbb{R}$ es* ***no-numerable****.*

Prueba. Para demostrar la no-numerabilidad de $\mathbb{R}$ vamos a utilizar el Método de la Diagonal de Cantor. Puesto que cualquier subconjunto infinito de un conjunto numerable es numerable, si podemos encontrar un subconjunto de $\mathbb{R}$ que es no-numerable, entonces $\mathbb{R}$ también será no-numerable. El subconjunto de $\mathbb{R}$ que queremos analizar es el intervalo $(0, 1)$. Suponga entonces que $(0, 1)$ es numerable y sea $\{x_1, x_2, x_3, \ldots\}$ una lista de **todos** los elementos de $(0, 1)$. Escriba cada elemento $x_n \in (0, 1)$ usando su representación decimal, esto es,

$$x_n = 0.x_{n1}x_{n2}x_{n3}\ldots$$

donde cada $x_{ni} \in \{0, 1, 2, \ldots, 9\}$ para todo n y todo i en $\mathbb{N}$. Antes de continuar, observe que la representación decimal de cada número en $(0, 1)$ no es única. Por ejemplo,

$$0.539999\ldots \qquad \text{y} \qquad 0.540000\ldots$$

representan el mismo número. Por esta razón, convenimos en no usar la version que termina en 9's. Disponga ahora de todos los elementos en $(0, 1)$ en el siguiente arreglo:

$$\begin{aligned}
x_1 &= 0.\mathbf{x_{11}}\, x_{12}\, x_{13} \ldots\, x_{1n} \ldots \\
x_2 &= 0.x_{21}\, \mathbf{x_{22}}\, x_{23} \ldots\, x_{2n} \ldots \\
x_3 &= 0.x_{31}\, x_{32}\, \mathbf{x_{33}} \ldots\, x_{3n} \ldots \\
&\vdots \qquad\qquad \ddots \\
x_n &= 0.x_{n1}\, x_{n2}\, x_{n3} \ldots \mathbf{x_{nn}} \ldots \\
&\vdots \qquad\qquad\qquad \vdots
\end{aligned}$$

Ahora viene lo realmente genial: vamos a construir, usando los elementos de la "diagonal" en el arreglo anterior, un elemento $b \in (0, 1)$ que no aparece en la lista $\{x_1, x_2, x_3, \ldots\}$, es decir, distinto de todos los x_n. En efecto, si para cada $n \in \mathbb{N}$, definimos

$$b_n = \begin{cases} x_{nn} - 1 & \text{si} \quad x_{nn} \neq 0 \\ 1 & \text{si} \quad x_{nn} = 0, \end{cases}$$

entonces $b = 0.b_1b_2b_3\ldots$ es claramente un número en $(0,1)$ que difiere de cada x_n precisamente en la cifra decimal x_{nn} y, por lo tanto, b no es igual a ningún x_n. Esto, por supuesto, es contrario a la suposición de que la sucesión $x_1, x_2, x_3, \ldots$ comprendía **a todos los números reales** en $(0,1)$. Esta contradicción establece que $(0,1)$ no es numerable y, en consecuencia, tampoco lo es $\mathbb{R}$. ■

Más adelante veremos otros modos de demostrar que $\mathbb{R}$ es no-numerable. La cardinalidad de $\mathbb{R}$ será denotada por

$$\operatorname{card}(\mathbb{R}) \;=\; \mathfrak{c}.$$

Otros Ejemplos.

(NN_0) Cualquier intervalo no-degenerado $I \subseteq \mathbb{R}$ es no-numerable.

(NN_1) $\mathbb{I} = \mathbb{R} \setminus \mathbb{Q}$, el conjunto de los **números irracionales**, es no-numerable.

(NN_2) $\mathbb{R}^n$ es no-numerable para cada $n \in \mathbb{N}$. Un modo fácil de ver esto es pensar a $\mathbb{R}$ como canónicamente sumergido en $\mathbb{R}^n$, es decir, identificando a $\mathbb{R}$ con el subespacio $\mathbb{R} \times \{0\} \times \cdots \times \{0\}$ de $\mathbb{R}^n$ donde a cada punto $x \in \mathbb{R}$ se le asigna el vector $(x, 0, \ldots, 0) \in \mathbb{R}^n$. De este modo podemos suponer que $\mathbb{R} \subseteq \mathbb{R}^n$ para cada $n \geq 1$. Como $\mathbb{R}$ es no-numerable, se sigue entonces de (Nu_3) que $\mathbb{R}^n$ es no-numerable. Este ejemplo fue descubierto por Cantor en 1877 cuando $n = 2$, quien en una carta dirigida a Dedekind le decía: "Lo veo, pero no puedo creerlo", aparentemente asombrado por el hecho de que $\mathbb{R}^2$ fuese tan numeroso como $\mathbb{R}$.

(NN_3) $\mathbb{N}^{\mathbb{N}}$, la familia de todos los **subconjuntos infinitos** de $\mathbb{N}$, es no-numerable. En efecto, si escribimos $\mathbb{N}^{\infty} = \mathbb{N}^{\mathbb{N}}$, resulta que

$$\mathscr{P}(\mathbb{N}) \;=\; \mathbb{N}^{<\infty} \cup \mathbb{N}^{\infty},$$

donde $\mathbb{N}^{<\infty} = \mathcal{P}_{\mathrm{fin}}(\mathbb{N})$. Como $\mathbb{N}^{<\infty}$ es numerable, Corolario 1.2.11, y $\mathscr{P}(\mathbb{N})$ es no-numerable, se concluye que $\mathbb{N}^{\mathbb{N}}$ es no-numerable.
Otra notación frecuentemente usada para el conjunto $\mathbb{N}^{\mathbb{N}}$ es $\mathbb{N}^{\omega}$.

(NN_4) $\mathcal{T}_{\mathrm{ras}}(\mathbb{R})$, el conjunto de todos los **números trascendentes**, es no-numerable. Esto es consecuencia de los siguientes hechos:

$$\mathbb{R} \;=\; \mathcal{A}_{\mathrm{lg}}(\mathbb{R}) \cup \mathcal{T}_{\mathrm{ras}}(\mathbb{R}),$$

la no-numerabilidad de $\mathbb{R}$ y la numerabilidad de $\mathcal{A}_{\mathrm{lg}}(\mathbb{R})$. En efecto, si $\mathcal{T}_{\mathrm{ras}}(\mathbb{R})$ fuese numerable, entonces $\mathcal{A}_{\mathrm{lg}}(\mathbb{R}) \cup \mathcal{T}_{\mathrm{ras}}(\mathbb{R})$ también lo sería y, en consecuencia, $\mathbb{R}$ sería numerable. Esta contradicción establece que $\mathcal{T}_{\mathrm{ras}}(\mathbb{R})$ es no-numerable.

(NN_5) $\prod_{n=1}^{\infty}\{0,1\} \;=\; 2^{\mathbb{N}}$, el conjunto de todas las sucesiones de 0's y 1's, es no-numerable. La demostración es muy similar a la ofrecida por Cantor para probar que $\mathbb{R}$ es no-numerable (usando el Método de la Diagonal) y, en consecuencia, se omite.

(NN_6) $\operatorname{card}\big(\mathbb{R}^{[0,1]}\big)$ es no-numerable. Más aun, $\operatorname{card}\big(\mathbb{R}^{[0,1]}\big) > \mathfrak{c}$. En efecto, en primer lugar, observe que el conjunto $\mathcal{C}$ de todas las funciones constantes definidas sobre $[0,1]$ posee cardinalidad $\mathfrak{c}$ y como $\mathcal{C} \subseteq \mathbb{R}^{[0,1]}$ resulta que $\operatorname{card}\big(\mathbb{R}^{[0,1]}\big) \geq \mathfrak{c}$. Esto prueba la no-numerabilidad de $\operatorname{card}\big(\mathbb{R}^{[0,1]}\big)$. Para verificar la segunda afirmación, suponga que $\operatorname{card}\big(\mathbb{R}^{[0,1]}\big) = \mathfrak{c}$ y sea $\Phi : [0,1] \to \mathbb{R}^{[0,1]}$ una biyección. Nótese que, para cada $x \in [0,1]$, $\Phi(x) \in \mathbb{R}^{[0,1]}$ de modo que, si hacemos $\Phi(x) = \Phi_x$, entonces $\mathbb{R}^{[0,1]} = \{\Phi_x : x \in [0,1]\}$. Identifique cada Φ_x con el conjunto de sus imágenes, esto es, $\Phi_x = \{\Phi_x(y) \in \mathbb{R} : y \in [0,1]\}$. Afirmamos que existe una función

$g \in \mathbb{R}^{[0,1]}$ tal que $\Phi_x \neq g$ para todo $x \in [0,1]$. En efecto, considere la función $g : [0,1] \to \mathbb{R}$ definida por

$$g(x) \;=\; \Phi_x(x) + 1 \quad \text{para todo} \quad x \in [0,1].$$

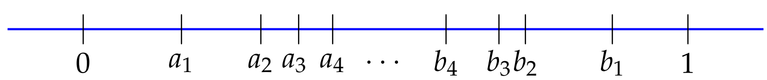
Esta función g difiere de cualquier $\Phi_x \in \mathbb{R}^{[0,1]}$ al menos en el punto $y = x$. Por esto, Φ no puede ser biyectiva y entonces $\operatorname{card}\big(\mathbb{R}^{[0,1]}\big) > \mathfrak{c}$.

1.2.4. Un Juego y la No-numerabilidad de $\mathbb{R}$.

El siguiente juego es otra forma elegante de presentar la no-numerabilidad de $\mathbb{R}$. Dos jugadores, denotados por α y β, juegan el siguiente juego infinito en $\mathbb{R}$. Se fija un subconjunto $X \subseteq [0,1]$ y el juego lo comienza el jugador α eligiendo un número real $a_1 \in (0,1)$. Enseguida β responde seleccionando un número $b_1 \in (a_1, 1)$. El turno ahora es de α quien escoge un número $a_2 \in (a_1, b_1)$. La respuesta de β consiste en elegir un punto $b_2 \in (a_2, b_1)$ y se continua ad infinitum. En consecuencia, se obtienen dos sucesiones, la del jugador α, $(a_n)_{n=1}^{\infty}$, que es estrictamente creciente, mientras que la de β, $(b_n)_{n=1}^{\infty}$, es estrictamente decreciente relacionadas por las siguientes desigualdades:

$$a_{n-1} \;<\; a_n \;<\; b_{n-1} \qquad\qquad \text{y} \qquad\qquad a_n \;<\; b_n \;<\; b_{n-1}$$

para todo $n \geq 1$, donde hemos puesto $a_0 = 0$ y $b_0 = 1$.



Puesto que toda sucesión monótona acotada converge, véase el Teorema 2.1.23, página 93, resulta que $\lim_{n\to\infty} a_n = a \in (0,1)$. Se declara **ganador** al jugador α si $a \in X$, en caso contrario el jugador β es el ganador.

Lema 1.2.24. *Sea X un subconjunto no vacío de $\mathbb{R}$. Si X es **numerable**, entonces β posee una **estrategia ganadora**.*

Prueba. La conclusión es inmediata si $X = \varnothing$. Suponga entonces que X es infinito numerable y sea $\{x_1, x_2, \ldots\}$ una enumeración de X. Considere la siguiente estrategia del jugador β: *en el n-ésimo movimiento, β selecciona $b_n = x_n$ si tal movimiento es legal, en caso contrario él elige como b_n cualquier número real permitido según las reglas establecidas*. De esto resulta que, por cada $n \geq 1$, se tiene que $x_n \leq a_n$ o $x_n \geq b_n$. Puesto que $a_n < a < b_n$ para todo $n \geq 1$, concluimos que $a \notin X$. Esto, por supuesto, significa que β **siempre gana** con esa estrategia. ■

Si tomamos $X = [0,1]$ en lema anterior, entonces claramente α gana el juego sin importar como juegue su oponente. De esto se obtiene que:

Corolario 1.2.25. *El intervalo $[0,1]$ es **no-numerable**. En particular, $\mathbb{R}$ es **no-numerable**.*

Nota Adicional 1.2.1 En el Corolario 1.2.11 (a), vimos que la unión numerable de conjuntos numerables sigue siendo numerable. ¿Qué ocurre si, en lugar de considerar una unión numerable de conjuntos numerables, suponemos una colección no-numerable de conjuntos numerables? ¿Cómo es su unión? ¿Será dicha unión no-numerable? El siguiente ejemplo nos muestra que la respuesta puede, en general, ser falsa.

Existe una colección ***no-numerable*** *de conjuntos, digamos* $(A_\alpha)_{\alpha\in\mathbb{R}}$, *tal que*

(1) *cada* A_α *es infinito numerable*,

(2) $A_\alpha \subsetneq A_\beta$ *si* $\alpha < \beta$ *y*

(3) $\displaystyle\bigcup_{\alpha\in\mathbb{R}} A_\alpha$ *es numerable.*

En efecto, para cada $\alpha \in \mathbb{R}$, considere el conjunto $A_\alpha = \{q \in \mathbb{Q} : q \leq \alpha\}$. Puesto que $\mathbb{Q}$ es numerable, resulta que cada A_α es infinito numerable y, por supuesto, su unión $\bigcup_{\alpha\in\mathbb{R}} A_\alpha = \mathbb{Q}$ es numerable. Además, si $\alpha < \beta$, entonces para cualquier $q \in A_\alpha$, se tiene que $q \leq \alpha < \beta$ y, por consiguiente, $q \in A_\beta$. Esto prueba que $A_\alpha \subseteq A_\beta$. Más aun, por la densidad de lo números racionales, existe $q_0 \in \mathbb{Q}$ tal que $\alpha < q_0 < \beta$, de modo que $q_0 \in A_\beta$, pero $q_0 \notin A_\alpha$. ■

Otro hecho interesante de la no-numerabilidad es que ella puede ser utilizada, en combinación con el Teorema del Valor Medio, para dar una prueba del siguiente resultado de Volterra.

Teorema de Volterra. *No existe ninguna* ***función continua*** $f : \mathbb{R} \to \mathbb{R}$ *tal que*

$$f(\mathbb{Q}) \subseteq \mathbb{I} \qquad \text{y} \qquad f(\mathbb{I}) \subseteq \mathbb{Q}. \tag{$*$}$$

Prueba. Suponga lo contrario, es decir, que cualquier función continua $f : \mathbb{R} \to \mathbb{R}$ satisface $(*)$. Fijemos una tal f. Por definición se tiene que f toma al menos dos valores, uno racional y otro irracional. Sean entonces $p \in \mathbb{Q}$ y $s \in \mathbb{I}$ tales que $f(p) = \alpha \in \mathbb{I}$ y $f(s) = \beta \in \mathbb{Q}$. Asuma que $\alpha < \beta$. Puesto que f es continua, la Propiedad del Valor Intermedio nos dice que f asume todos los valores en el intervalo cerrado $[\alpha, \beta]$, es decir, $[\alpha, \beta] \subseteq f(\mathbb{R})$. Por otro lado, como $f(\mathbb{I}) \subseteq \mathbb{Q}$, resulta que $f(\mathbb{I})$ es numerable y, por supuesto, $f(\mathbb{Q})$ también lo es. Por consiguiente, $f(\mathbb{R})$ es numerable lo cual es imposible ya que $[\alpha, \beta]$ es no-numerable y $[\alpha, \beta] \subseteq f(\mathbb{R})$. ■

1.3. El Axioma de Elección y sus Aliados

1.3.1. El Axioma de Elección

Suponga que $\mathcal{A} = \{A_\alpha : \alpha \in J\}$ es una familia arbitraria de conjuntos cada uno de los cuales es no vacío. ¿Existe algún procedimiento que permita construir un nuevo conjunto eligiendo uno, y sólo un punto, en cada uno de los conjuntos A_α de $\mathcal{A}$? El Axioma de Elección, o Axioma de Zermelo como también se le conoce, es un axioma de la teoría de conjuntos que postula la existencia de un tal conjunto pero sin dar ninguna indicación de cómo se hace tal elección.

Axioma de Elección (AC). *Dada cualquier colección arbitraria* $(X_\alpha)_{\alpha\in J}$ *de conjuntos no vacíos, siempre se puede elegir, de cada uno de los conjuntos* X_α, *uno, y sólo un miembro* $x_\alpha \in X_\alpha$ *para construir un nuevo conjunto* X *con tales elementos*.

El Axioma de Elección fue propuesto por primera vez por Ernst Zermelo en 1904 quien lo utilizó precisamente para demostrar que *todo conjunto puede ser bien-ordenado*. De inmediato provocó reacciones: resultó, para algunos matemáticos, un plato un tanto difícil de digerir y, por supuesto, generó mucha controversia. La no existencia de un "procedimiento" o "descripción explícita" para elegir los puntos de cada uno de los conjuntos en una *colección arbitraria* de conjuntos es lo que produce la controversia. Como veremos un poco más adelante, existen colecciones de conjuntos que son increíblemente enormes y, por lo tanto, no existe, en términos generales, un procedimiento para escoger un elemento en cada uno de los conjuntos de esas colecciones. A pesar de la ausencia de un tal *mecanismo de selección*, existen situaciones donde es posible hacerlo sin requerir el Axioma de Elección. En efecto, si nuestra familia de conjuntos consta sólo de un número finito de conjuntos, entonces la elección de un punto en cada uno de los conjuntos de la familia se puede llevar a cabo sin invocar el Axioma de Elección. También, si $(A_\alpha)_{\alpha\in J}$ es una familia arbitraria de subconjuntos de $\mathbb{N}$, entonces el Principio del Buen-Orden puede ser invocado para hacer tal elección: de cada conjunto A_α seleccione su *primer elemento*. Por supuesto, este caso tampoco requiere el uso del Axioma de Elección.

El Axioma de Elección siempre fue, desde sus inicios, un axioma polémico. Paul Bernays (1888-1977) y Adolf Abraham Fraenkel (1891-1965) afirman de él lo siguiente:

> El Axioma de Elección (junto con la Hipótesis del Continuo) es probablemente el más interesante y más discutido axioma en matemáticas después del Axioma de las Paralelas de Euclides.

En 1938 esa controversia fue profundamente iluminada por el lógico-matemático Kurt Gödel quien demostró que si la Teoría de Conjuntos construida con el sistema **ZF** es consistente, entonces también lo es la Teoría de Conjuntos construida con el sistema **ZFC**. Posteriormente, en 1963, Paul Cohen cierra el ciclo al demostrar que si al sistema **ZF** se le añade la negación del Axioma de Elección, la nueva Teoría de Conjuntos que se edifica con ella también es consistente. Dicho de otro modo, el Axioma de Elección es independiente de los axiomas de Zermelo-Fraenkel, lo que permite concluir que ni la verdad, ni la falsedad de dicho axioma pueden ser demostrados en **ZF**. Su aceptación, en términos generales, se sustenta sobre el hecho de que dicho axioma es tremendamente útil. Muchos resultados importantes y fundamentales en Análisis Real, Análisis Funcional, Álgebra, Topología, etc. sólo se pueden demostrar si se acepta, sin limitaciones, el Axioma de Elección. Una muestra de ello se puede ver, por ejemplo, en el libro de H. Herrlich: **Axiom of Choice** [71].

Es importante destacar que el Axioma de Elección lo que realmente afirma es simplemente la *existencia de una función de lección* y, en consecuencia, se puede formular ligeramente diferente, aunque equivalente, del modo siguiente:

Axioma de Elección (AC). *Si $(X_\alpha)_{\alpha\in J}$ es una familia de conjuntos tal que X_α es **no vacío** para todo $\alpha \in J$, entonces el producto cartesiano $\Pi_{\alpha\in J}X_\alpha$ es no vacío, es decir, existe al menos una función de elección para la familia $(X_\alpha)_{\alpha\in J}$.*

En el siguiente resultado se muestra un "procedimiento" explícito para construir una función de elección de una familia de conjuntos sin usar el Axioma de Elección.

Teorema 1.3.1. *Sea X un conjunto no vacío y suponga que $f : \mathbb{N} \to X$ es una función **sobreyectiva**. Entonces existe una función **inyectiva** $g : X \to \mathbb{N}$ tal que $f(g(x)) = x$ para todo $x \in X$.*

Prueba. El requerimiento de que $f(g(x)) = x$ significa que $g(x)$ debe ser un elemento del conjunto $f^{-1}(\{x\}) = \{n \in \mathbb{N} : f(n) = x\}$. Ahora bien, la sobreyectividad de f nos garantiza que el conjunto $f^{-1}(\{x\})$ es no-vacío para cada $x \in X$. Defina $g : X \to \mathbb{N}$ demandando que

$$g(x) = \min f^{-1}(\{x\})$$

para cada $x \in X$. Puesto que el mínimo de cada subconjunto no-vacío de $\mathbb{N}$ existe (Principio del Buen-Orden) y es único, resulta que g está bien definida, es claramente inyectiva y se cumple que $f(g(x)) = x$ para cualquier $x \in X$. ■

¿Qué ocurre si, en el teorema anterior, consideramos $\mathbb{N} = X$ y una función sobreyectiva arbitraria $f : X \to X$? ¿Se puede, en este caso, determinar una función inyectiva $g : X \to X$ tal que $f(g(x)) = x$ para todo $x \in X$? La respuesta es que en presencia del Axioma de Elección una tal g siempre existe. En efecto, como antes, la sobreyectividad de f nos muestra que el conjunto $f^{-1}(\{x\}) \neq \varnothing$ para cada $x \in X$. Sea $A_x = f^{-1}(\{x\})$ para $x \in X$ y observe que $X = \bigcup_{x \in X} A_x$. El Axioma de Elección nos dice que podemos elegir, de cada conjunto A_x, un único punto a_x. La función $g : X \to X$ definida $g(x) = a_x$ para cada $x \in X$ posee las propiedades requeridas.

Otro resultado de cierta utilidad que también es consecuencia del Axioma de Elección es el siguiente.

Teorema 1.3.2. *Sean X un conjunto arbitrario no vacío y sea $f : X \to X$ una función. Entonces, para cada conjunto $A \subseteq X$ se cumple que*

$$\operatorname{card}(f(A)) \leq \operatorname{card}(A).$$

Prueba. Para cada $y \in f(A)$, sea $X_y = f^{-1}(\{y\})$. Puesto que $X_y \neq \varnothing$ para cada elemento $y \in f(A)$ y como $X_{y_1} \cap X_{y_2} = \varnothing$ si $y_1 \neq y_2$, el Axioma de Elección nos garantiza la existencia de una función de elección $g \in \prod_{y \in f(A)} X_y$. Veamos que g es inyectiva. En efecto, en primer lugar, observe que $g : f(A) \to A$. Sean $y_1, y_2 \in f(A)$ con $y_1 \neq y_2$. Entonces $g(y_1) \in X_{y_1}$, $g(y_2) \in X_{y_2}$ y como ambos conjuntos son disjuntos, resulta que $g(y_1) \neq g(y_2)$. Esto prueba que g es inyectiva y, por lo tanto, $\operatorname{card}(f(A)) \leq \operatorname{card}(A)$. ■

Nota Adicional 1.3.2 Existen formas más débiles del Axioma de Elección que han sido propuestos para excluir ciertos resultados que pueden ser catalogados como extraños tal como la Paradoja de Banach-Tarski. Entre esas forman están, por ejemplo, el Axioma de Elección Numerable, $\mathbf{AC}_\omega$, el cual establece que:

Axioma de Elección Numerable ($\mathbf{AC}_\omega$). *Si $(X_n)_{n=1}^{\infty}$ es una sucesión de conjuntos no vacíos, entonces $\prod_{n=1}^{\infty} X_n$ es **no vacío**.*

Otro axioma es el Axioma de Elección Múltiple, **DC**, que reza lo siguiente:

Axioma de Elección Múltiple (**DC**). *Sea R una relación binaria sobre un conjunto no vacío X con la propiedad de que para cada $x \in X$ exsite un $y \in X$ de modo que $(x, y) \in R$. Entonces existe una sucesión $(x_n)_{n=0}^{\infty}$ en X tal que $(x_n, x_{n+1}) \in R$ para todo $n = 0, 1, 2, \ldots$*

Es fácil establecer que $\mathbf{AC} \Rightarrow \mathbf{DC} \Rightarrow \mathbf{AC}_\omega$ y que ninguna de las implicaciones opuestas son válidas.

1.3.2. El Lema de Zorn

Entre las numerosas y variadas formas equivalentes de pensar el Axioma de Elección se encuentra el así llamado Lema de Zorn, un resultado formulado por Max Zorn (1906-1993) en 1935 y que resulta ser extremadamente útil en varias ramas del quehacer matemático. Por ejemplo, el Lema de Zorn es fundamental para demostrar resultados importantes tales como: el Teorema de

Hahn-Banach, el Teorema de Krein-Milman, el Teorema del Ultrafiltro, la prueba de la existencia de una base de Hamel en cualquier espacio vectorial no trivial, etc.

Recordemos que una relación binaria sobre un conjunto X no es otra cosa que cualquier subconjunto $\mathcal{R}$ de $X \times X$.

Definición 1.3.3. *Una relación binaria $\mathcal{R}$ sobre un conjunto X se dice que es un* ***orden parcial*** *si ella es*

(a) ***reflexiva***: $(x, x) \in \mathcal{R}$ para todo $x \in X$,

(b) ***antisimétrica***: si (x, y) y (y, x) están en $\mathcal{R}$, entonces $x = y$,

(c) ***transitiva***: si (x, y) y (y, z) están en $\mathcal{R}$, entonces $(x, z) \in \mathcal{R}$.

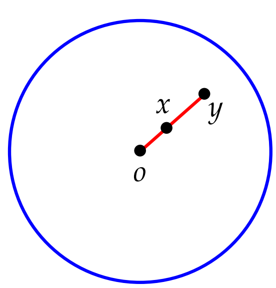
En lugar de $(x, y) \in \mathcal{R}$ también se acostumbra a escribir $x\mathcal{R}y$. En lo que sigue escribiremos, en lugar de $\mathcal{R}$, el símbolo $\preceq$ para denotar un orden parcial sobre X. En este caso, la expresión $(x, y) \in \mathcal{R}$ se escribe como $x \preceq y$. Un conjunto X equipado con un orden parcial $\preceq$ es llamado un **conjunto parcialmente ordenado** y denotado por $(X, \preceq)$. Dos elementos x, y en un conjunto parcialmente ordenado se dicen que son **comparables** si $x \preceq y$ o $y \preceq x$. Un conjunto parcialmente ordenado en el cual cualquier par de elementos son comparables es llamado un **conjunto totalmente (o linealmente) ordenado** y a dicho orden se le denomina un **orden lineal** o **total**. Una **cadena** en un conjunto parcialmente ordenado es un subconjunto que está totalmente ordenado. En un conjunto parcialmente ordenado $(X, \preceq)$ la relación $x \prec y$ significa que $x \preceq y$ pero $x \neq y$. Con frecuencia escribiremos $y \succeq x$ (respectivamente, $y \succ x$) como sinónimo de $x \preceq y$ (respectivamente, $x \prec y$).

Sea $(X, \preceq)$ un conjunto parcialmente ordenado y sea $A \subseteq X$. Un elemento $x \in X$ es una **cota superior** de A si $a \preceq x$ para todo $a \in A$. Si x_0 es una cota superior de A y si cualquier otra cota superior x de A satisface $x_0 \preceq x$, entonces se dice que x_0 es el **supremo** de A. En este caso escribiremos $x_0 = \sup A$. Si, además, $x_0 \in A$, entonces se dice que x_0 es el **máximo** o el elemento **más grande** de A. Por otro lado, un elemento $x_0 \in X$ se dice que es un elemento **maximal** en X si no existe $y \in X$ para el cual $x_0 \prec y$; es decir, si existe un elemento $x \in X$ que satisface $x_0 \preceq x$, entonces $x = x_0$.

Observe que un elemento maximal no tiene porque ser, en el orden dado, el más grande de todos. Por ejemplo, sea $X = \{x \in \mathbb{R}^2 : \| x \|_2 \leq 1\}$, donde $\| \cdot \|_2$ es la norma euclidiana:

$$\| x \|_2 = \sqrt{x_1^2 + x_2^2} \quad \text{para cada } x = (x_1, x_2) \in \mathbb{R}^2.$$

Sobre X defina el siguiente orden parcial $\preceq$: si $x, y \in X$, $x \preceq y$ *si, y sólo si*, $x \in I_y$, donde I_y es el segmento radial que va desde el origen al punto y.



Es claro que cualquier par de vectores $x, y \in X$ no son comparables si ellos están sobre segmentos radiales distintos. De esto se sigue que cualquier vector $v \in \{x \in \mathbb{R}^2 : \| x \|_2 = 1\}$ es maximal pero no es un máximo.

Las nociones de ínfimo, mínimo y minimal se definen de modo enteramente similar. La demostración del próximo resultado se puede ver, por ejemplo, en [71].

Lema 1.3.4 (Lema de Zorn). *Sea $(X, \preceq)$ un conjunto parcialmente ordenado. Si* ***cualquier cadena*** *en X posee una* ***cota superior****, entonces X posee un* ***elemento maximal****.*

Con mucha frecuencia el Lema de Zorn se utiliza cuando $\mathfrak{F}$ es una familia de subconjuntos de un conjunto dado X ordenado por la relación de inclusión $\subseteq$ con la propiedad de que cualquier cadena $\mathcal{C} \subseteq \mathfrak{F}$, su unión $\bigcup \mathcal{C}$, también esté en $\mathfrak{F}$. En este caso, $\bigcup \mathcal{C}$ es una cota superior para $\mathcal{C}$ con respecto a $\subseteq$. En este caso particular, el Lema de Zorn se expresa del modo siguiente:

Corolario 1.3.5 (Principio Maximal de Hausdorff). *Sea $\mathfrak{F}$ una familia de subconjuntos no vacíos de un conjunto no vacío X. Suponga que los elementos de $\mathfrak{F}$ están ordenados por la relación de inclusión $\subseteq$ y que para cualquier cadena $\mathcal{C} \subseteq \mathfrak{F}$, se cumple que su unión $\bigcup \mathcal{C}$ también está en $\mathfrak{F}$. Entonces $\mathfrak{F}$ posee un elemento maximal.*

Una de las aplicaciones clásicas del Lema de Zorn es la demostración de la existencia de una base de Hamel en cualquier espacio vectorial no trivial. Recordemos que si X es un espacio vectorial sobre un cuerpo $\mathbb{F}$, entonces cualquier expresión de la forma $a_1x_1 + \cdots + a_nx_n$, donde $x_j \in X$ y $a_j \in \mathbb{F}$ para todo $j = 1, \ldots, n$ es llamada un **combinación lineal** de los elementos $x_1, \ldots, x_n$. Un conjunto de vectores, digamos $\{x_1, \ldots, x_n\}$, se dicen que son **linealmente dependientes** si existen escalares no todos nulos $a_1, \ldots, a_n$ en $\mathbb{F}$ tal que $a_1x_1 + \cdots + a_nx_n = 0$. En caso contrario se dice que los vectores $\{x_1, \ldots, x_n\}$ son **linealmente independientes**, es decir, si la única solución de la ecuación $a_1x_1 + \cdots + a_nx_n = 0$ es la nula, esto es, $a_1 = \cdots = a_n = 0$. De modo más general, diremos que un subconjunto no vacío $\mathcal{B}$ de X es **linealmente independiente** si cualquier subconjunto no vacío y finito de $\mathcal{B}$ es linealmente independiente. Se dice también que $\mathcal{B}$ **genera** a X si cualquier vector en X es combinación lineal de algún subconjunto finito de $\mathcal{B}$. En este caso, escribiremos $[\mathcal{B}] = X$.

Definición 1.3.6. *Sea X un* ***espacio vectorial no trivial*** *sobre un cuerpo $\mathbb{F}$. Un subconjunto no vacío $\mathcal{B}$ de X es un* ***base de Hamel*** *de X si $\mathcal{B}$ es linealmente independiente y genera a X.*

Observe que si $\mathcal{B}$ es una base de Hamel de X, entonces el vector $0 \notin \mathcal{B}$.

Teorema 1.3.7 (Base de Hamel). *Si X es un* ***espacio vectorial no trivial*** *sobre un cuerpo $\mathbb{F}$, entonces X posee una* ***base de Hamel****.*

Prueba. Considere la familia

$$\mathfrak{F} = \big\{F \subseteq X : F \text{ es linealmente independiente}\big\}.$$

Puesto que $X \neq \{0\}$, cualquier conjunto $F = \{a\}$, con $a \neq 0$, es linealmente independiente. Luego, $\mathfrak{F} \neq \varnothing$. Definamos el siguiente orden parcial sobre $\mathfrak{F}$: si $F, G \in \mathfrak{F}$, declaramos que

$$F \preceq G \qquad \text{si, y sólo si} \qquad F \subseteq G. \tag{1}$$

Suponga ahora que $\mathcal{C}$ es una cadena arbitraria en $\mathfrak{F}$ y sea $F_0 = \bigcup \mathcal{C}$. Veamos que $F_0 \in \mathfrak{F}$. En efecto, sea A un subconjunto finito de F_0, digamos $A = \{x_1, \ldots, x_n\}$. Entonces existen conjuntos $C_1, \ldots, C_n$ en $\mathcal{C}$ tales que $x_i \in C_i$ para $i = 1, \ldots, n$, y como $\mathcal{C}$ es una cadena, existe un i_0 en $\{1, \ldots, n\}$ tal que $C_i \subseteq C_{i_0}$ para $i = 1, \ldots, n$. Esto prueba que $A = \{x_1, \ldots, x_n\} \subseteq C_{i_0}$ y como

C_{i_0} es linealmente independiente, se sigue que A también es linealmente independiente y, por lo tanto, $F_0 \in \mathfrak{F}$. Un llamado al Lema de Zorn (o al Principio Maximal de Hausdorff) nos revela que $\mathfrak{F}$ posee un elemento maximal, digamos $\mathcal{B} \in \mathfrak{F}$. Queda por demostrar que $\mathcal{B}$ genera a X. En efecto, suponga por un momento que $\mathcal{B}$ no genera a X. Esto significa que algún $x_0 \in X$ no se puede representar como una combinación lineal de elementos de $\mathcal{B}$. Si ahora definimos $\mathcal{B}_0 = \{x_0\} \cup \mathcal{B}$, resulta que hemos encontrado un nuevo conjunto que es linealmente independiente y, además, contiene propiamente a $\mathcal{B}$, violando la maximalidad de $\mathcal{B}$. Esto termina la prueba. ■

Fijemos un conjunto linealmente independiente $F \subseteq X$, donde X un espacio vectorial no trivial sobre $\mathbb{F}$. Un argumento enteramente similar al resultado anterior, pero ahora trabajando con la familia

$$\mathfrak{F}_F = \{L \subseteq X : L \text{ es linealmente independiente y } F \subseteq L\},$$

conduce a la existencia de una base de Hamel $\mathcal{B}$ con $F \subseteq \mathcal{B}$.

Corolario 1.3.8. *Sea X un **espacio vectorial no trivial** sobre un cuerpo $\mathbb{F}$. Si $F \subseteq X$ es **linealmente independiente**, entonces existe una base de Hamel $\mathcal{B}$ en X tal que $F \subseteq \mathcal{B}$.*

Prueba. Considere, como en la prueba del resultado anterior, la familia

$$\mathfrak{F}_F = \{L \subseteq X : L \text{ es linealmente independiente y } F \subseteq L\}.$$

Observe que $\mathfrak{F}_F \neq \varnothing$ ya que $F \in \mathfrak{F}_F$. El mismo argumento dado en la prueba del Teorema 1.3.7 muestra que cualquier cadena $\mathcal{C}$ en $\mathfrak{F}_F$, en el orden dado por (1), posee una cota superior. Por el Lema de Zorn existe un elemento maximal en $\mathfrak{F}_F$, es decir, un conjunto linealmente independiente en X, digamos $\mathcal{B}$, el caul contiene a F y es maximal con respecto al orden establecido en $\mathfrak{F}$. Falta por demostrar que $\mathcal{B}$ genera a X, sin embargo, el mismo procedimiento utilizado en la prueba del resultado anterior nos conduce a que $\mathcal{B}$ genera a X y termina la prueba. ■

Finalizamos esta sección con otro resultado útil sobre la existencia de bases de Hamel que usa de nuevo el Lema de Zorn.

Corolario 1.3.9. *Sea X un **espacio vectorial no trivial** sobre un cuerpo $\mathbb{F}$. Si $F \subseteq X$ **genera a** X, entonces existe una base de Hamel $\mathcal{B}$ en X tal que $\mathcal{B} \subseteq F$.*

Prueba. Sea

$$\mathfrak{G} = \{G \subseteq X : G \text{ es linealmente independiente y } G \subseteq F\}.$$

Puesto que F genera a X, resulta que $F \neq \{0\}$. Seleccionando cualquier $x \in F$ con $x \neq 0$ se tiene que $\{x\} \in \mathfrak{G}$ lo cual prueba que $\mathfrak{G}$ es no vacío. Usando de nuevo el orden dado por (1) en $\mathfrak{G}$, resulta que cada cadena en $\mathfrak{G}$ posee una cota superior, por lo que una nueva aplicación del Lema de Zorn nos garantiza la existencia de un elemento maximal $\mathcal{B}$ en $\mathfrak{G}$. Por supuesto, $\mathcal{B} \subseteq F$ y como $\mathcal{B}$ es linealmente independiente, entonces un argumento enteramente similar a la demostración del resultado anterior conduce a que $\mathcal{B}$ genera a X y termina la prueba. ■

1.3.3. Principio del Buen-Orden

Entre los conjuntos infinitos, el conjunto de los números naturales, con su orden natural, $(\mathbb{N}, \leq)$ es un conjunto que disfruta de la siguiente propiedad, conocida como el **Principio del Buen-Orden**: *cualquier subconjunto no vacío de* $\mathbb{N}$ *contiene un primer elemento*, es decir, el elemento más pequeño (o mínimo) del subconjunto. Si pudiéramos extender dicho principio a cualquier conjunto no-numerable para algún tipo de orden, abrigaríamos la esperanza de poder trabajar con cualquier conjunto con tal orden del mismo modo conque trabajamos con $\mathbb{N}$ y, por supuesto, eso nos conduciría a extender nuestra manera tradicional de contar más allá de los naturales y, también, dispondríamos de una extensión del proceso de inducción matemática en cualquier conjunto. Por tales motivos, el principio del buen-orden es una propiedad que pudiéramos pensar como *altamente deseada*. Éste principio fue introducido por G. Cantor para desarrollar algunos de sus resultados sobre la teoría de los subconjuntos infinitos de $\mathbb{R}$.

Sea $(X, \preceq)$ un conjunto parcialmente ordenado y suponga que A es un subconjunto no vacío de X. Recordemos que un elemento $a_0 \in A$ se dice que es el **primer elemento**, o el **elemento mínimo** o **más pequeño** en A, si $a_0 \preceq a$ para todo $a \in A$. El primer elemento de un conjunto A, si existe, es único.

Definición 1.3.10. *Sea $(X, \preceq)$ un conjunto parcialmente ordenado. Diremos que $\preceq$ es un* ***buen-orden*** *en (o sobre) X, o que $(X, \preceq)$ es un conjunto* ***bien-ordenado****, si cualquier subconjunto no vacío A de X posee un primer elemento.*

De la definición anterior se pueden derivar al menos tres consecuencias importantes:

(a) ***Todo buen-orden*** *$\preceq$ sobre un conjunto no vacío X automáticamente convierte a dicho conjunto en un* ***conjunto totalmente ordenado***. En efecto, si $x, y \in X$, entonces el conjunto $A = \{x, y\}$ posee, por ser $\preceq$ un buen-orden sobre X, un primer elemento, es decir, o bien $x \preceq y$, o bien $y \preceq x$. Por esta razón, *siempre supondremos que el orden en un conjunto bien-ordenado es total*.

(b) *Si $(X, \preceq)$ es un* ***conjunto infinito bien-ordenado****, entonces sus elementos se pueden etiquetar en orden creciente*. En efecto, sea x_0 es el primer elemento de X. Como X es infinito, $X \setminus \{x_0\}$ es no vacío. Sea x_1 el primer elemento de $X \setminus \{x_0\}$ y observe que $x_0 \prec x_1$. En general, si para cada $n \in \mathbb{N}$, x_n es el primer elemento de $X \setminus \{x_0, \ldots, x_{n-1}\}$, entonces

$$x_0 \prec x_1 \prec x_2 \prec \cdots \prec x_n$$

y continúe.

(c) *En* ***conjuntos bien-ordenados*** *no existen* ***sucesiones estrictamente decrecientes***. Para ver esto, sea $(X, \preceq)$ un conjunto bien-ordenado y suponga que existe una sucesión $(x_n)_{n=0}^{\infty}$ en X tal que $x_0 \succ x_1 \succ x_2 \succ \cdots$. Entonces el subconjunto no vacío $A = \{x_0, x_1, x_2, \ldots\}$ de X no tendría primer elemento lo cual daría lugar a una contradicción.

El orden lexicográfico es un ejemplo de un buen-orden en el producto cartesiano de dos conjuntos bien-ordenados. Recordemos su definición.

Definición 1.3.11 (El Orden Lexicográfico). *Sean $(A, \preccurlyeq_A)$ y $(B, \preccurlyeq_B)$ dos conjuntos parcialmente ordenados. El* ***orden lexicográfico****, también conocido como el* ***orden del diccionario****, es una relación de orden $\preceq$ definida sobre el producto cartesiano $A \times B$ del modo siguiente: para todo $(a, b), (a', b') \in A \times B$,*

$$(a, b) \preceq (a', b') \quad \Leftrightarrow \quad \begin{cases} a \preccurlyeq_A a', \ \text{o} \\ a = a' \ \wedge \ b \preccurlyeq_B b' \end{cases}$$

Nótese que la regla que define a $\preceq$ es la misma regla que se utiliza para ordenar las palabras en cualquier diccionario. De allí su nombre.

Ejemplo 1.3.1. *Sean* $(A, \preccurlyeq_A)$ *y* $(B, \preccurlyeq_B)$ *conjuntos bien-ordenados. Si el producto cartesiano* $A \times B$ *está provisto del orden lexicográfico* $\preceq$*, entonces* $(A \times B, \preceq)$ *es un conjunto bien-ordenado.*

Prueba. Sea X un subconjunto no vacío de $A \times B$. Observe que

$$X_1 \;=\; \big\{a \in A \,:\, (a,b) \in X\big\}$$

es un subconjunto no vacío de A y, en consecuencia, como $(A, \preccurlyeq_A)$ está bien-ordenado él posee un primer elemento, llamémoslo a_0. De modo similar, el conjunto

$$X_2 \;=\; \big\{b \in B \,:\, (a_0,b) \in X\big\}$$

posee, en B, un primer elemento, digamos b_0. Resulta claro, por la definición del orden lexicográfico, que (a_0, b_0) es el primer elemento de X y, por lo tanto, $A \times B$ con el orden lexicográfico $\preceq$ es un conjunto bien-ordenado. ■

Por ejemplo, el orden lexicográfico en $\mathbb{N} \times \mathbb{N}$ viene dado por:

$$\begin{array}{ccccccccccc}
 & (0,0) & < & (0,1) & < & (0,2) & < & (0,3) & < & \cdots \\
< & (1,0) & < & (1,1) & < & (1,2) & < & (1,3) & < & \cdots \\
< & (2,0) & < & (2,1) & < & (2,2) & < & (2,3) & < & \cdots \\
< & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots
\end{array}$$

En el siguiente resultado se muestra que cualquier conjunto admite un buen-orden. Este resultado, al que llamaremos Principio del Buen-Orden, fue formulado por G. Cantor pero demostrado por E. Zermelo en 1904 haciendo uso del Axioma de Elección. La prueba que se muestra a continuación se basa en una de las formas equivalentes de dicho axioma: el Lema de Zorn.

Teorema 1.3.12 (Principio del Buen-Orden). *Si X es cualquier **conjunto infinito**, entonces X puede ser **bien-ordenado**.*

Prueba. Sea X un conjunto infinito y considere la familia

$$\mathfrak{F} \;=\; \Big\{(A, \preccurlyeq_A) \,:\, A \subseteq X \quad \text{y} \quad \preccurlyeq_A \ \text{ es un buen-orden sobre } A\Big\}.$$

Puesto que cualquier conjunto finito está bien-ordenado por cualquier orden lineal, resulta que $\mathfrak{F} \neq \varnothing$. Sobre $\mathfrak{F}$ se define el orden parcial $\precsim$ declarando que: $(A, \preccurlyeq_A) \precsim (B, \preccurlyeq_B)$ si, y sólo si,

(1) $A \subseteq B$,

(2) $\preccurlyeq_A$ y $\preccurlyeq_B$ coinciden sobre A y,

(3) si $x \in B \setminus A$, entonces $a \preccurlyeq_B x$ para todo $a \in A$.

Sea ahora $\mathcal{C}$ una cadena en $\mathfrak{F}$ y definamos $C = \bigcup\{A : (A, \preccurlyeq_A) \in \mathcal{C}\}$. Sobre C conviene definir el siguiente orden: $x \preccurlyeq_C y$ si, y sólo, $x \preccurlyeq_A y$ para algún conjunto $(A, \preccurlyeq_A) \in \mathcal{C}$ tal que $x, y \in A$. Observe que un tal conjunto A siempre existe. En efecto, sean $x, y \in C$. Entonces $x \in A'$, $y \in A''$ para ciertos $A', A'' \in \mathcal{C}$. Como $\mathcal{C}$ es una cadena, entonces $A' \subseteq A''$ o $A'' \subseteq A'$. Sea A

el conjunto de $\mathcal{C}$ que contiene al otro. Entonces $x, y \in A$ y, por lo tanto, $x \preccurlyeq_A y$. Claramente $\preccurlyeq_{\mathcal{C}}$ está bien definido y constituye un buen-orden sobre C. Por esto, $(C, \preccurlyeq_C) \in \mathfrak{F}$ y es claro que $(C, \preccurlyeq_C)$ es una cota superior para $\mathcal{C}$. Se sigue del Lema de Zorn que el conjunto $\mathfrak{F}$ posee un elemento maximal, digamos $(A_0, \preccurlyeq)$. Afirmamos que $A_0 = X$. En efecto, suponga por un momento que $A_0 \neq X$ y sea x cualquier elemento en $X \setminus A_0$. Ordene el conjunto $B_0 = A_0 \cup \{x\}$ con el mismo orden que posee A_0 estipulando, además, que $a \preccurlyeq x$ para todo $a \in A_0$. Entonces $(B_0, \preccurlyeq)$ es un elemento de $\mathfrak{F}$ tal que $(A_0, \preccurlyeq) \precnsim (B_0, \preccurlyeq)$, lo que evidentemente contradice la maximalidad de $(A_0, \preccurlyeq)$. Por esto $A_0 = X$ y $\preccurlyeq$ es un buen-orden sobre X. ■

Nota Adicional 1.3.3 Es un hecho ya establecido que en *la Teoría de Conjuntos* **ZF** *el* Axioma de Elección*, el* Lema de Zorn *y el* Principio del Buen-Orden *son todos equivalentes entre sí* (véase, por ejemplo, [74]). Esto significa que en la Teoría de Conjuntos **ZFC**, el Lema de Zorn y el Principio del Buen-Orden siempre se cumplen. En tal sentido, y en lo que sigue, *asumiremos que nuestra Teoría de Conjuntos es la que se obtiene del sistema* **ZFC** *por lo que el* Axioma de Elección *será usado libremente en esta notas*.

Definición 1.3.13. *Sea $(X, \preceq)$ un conjunto bien-ordenado. Para cada $x \in X$, al subconjunto de X,*

$$\mathrm{Seg}(x) = \{a \in X : a \prec x\},$$

se le llama el ***segmento inicial*** *determinado por x.*

Por ejemplo, la definición moderna de los números naturales $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$ usa tales segmentos en la siguiente forma:

$$0 = \mathrm{Seg}(0) = \varnothing, \quad 1 = \mathrm{Seg}(1) = \{0\}, \quad 2 = \mathrm{Seg}(2) = \{0, 1\},$$

y, en general,

$$n = \mathrm{Seg}(n) = \{0, 1, 2, \ldots, n-1\}.$$

Un par de observaciones son pertinentes referentes a los conjuntos $\mathrm{Seg}(x)$ de la definición anterior.

(a) $x \notin \mathrm{Seg}(x)$ para todo $x \in X$.

(b) Cada par de segmentos iniciales en X son $\subseteq$-comparables. En efecto, sean $x, y \in X$. Sin $x = y$ no hay nada que probar. Suponga que $x \neq y$. Por ser $\preceq$ un orden total, se tiene que $x \prec y$ o $y \prec x$ y, por consiguiente,

$$\mathrm{Seg}(x) \subseteq \mathrm{Seg}(y) \qquad \text{o} \qquad \mathrm{Seg}(y) \subseteq \mathrm{Seg}(x).$$

Si $(X, \preceq)$ y $(Y, \preceq')$ son conjuntos bien-ordenados, escribiremos, si fuese necesario, $\mathrm{Seg}_X(x)$ y $\mathrm{Seg}_Y(y)$ para denotar segmentos iniciales de X y de Y respectivamente. En general, si $A \subseteq X$, entonces definimos

$$\mathrm{Seg}_A(x) = \mathrm{Seg}(x) \cap A = \{a \in A : a \prec x\}.$$

La existencia de conjuntos bien-ordenados permite obtener la siguiente versión del principio de inducción en $\mathbb{N}$, llamado el Principio de Inducción Transfinita.

Teorema 1.3.14 (Inducción Transfinita). *Sea $(X, \preceq)$ un **conjunto bien-ordenado**. Si A es un subconjunto no vacío de X verificando las siguientes dos condiciones:*

*(a) el **primer elemento** de X pertenece a A, y*

*(b) para cada $x \in X$, la condición: "$\mathrm{Seg}(x) \subseteq A$ **implica que** $x \in A$,"*

entonces $A = X$.

Prueba. Suponga que $A \neq X$. El subconjunto $B = X \setminus A$ de X es no vacío y, gracias al hecho de X está bien-ordenado, B posee un primer elemento, llamémoslo x_0. Por (a) tenemos que x_0 no es el primer elemento de X. Ahora bien, por nuestra elección de x_0 vemos que si $y \prec x_0$, entonces $y \in A$, lo cual significa que $\mathrm{Seg}(x_0) \subseteq A$. Un llamado a la condición (b) nos revela $x_0 \in A$, lo que contradice el hecho de que $x_0 \notin A$. Por esto, $A = X$. ∎

Hemos visto que los elementos de todo conjunto bien-ordenado se pueden disponer en orden creciente. La siguiente definición, la cual es fundamental para describir lo que llamaremos número ordinal, permite identificar conjuntos bien-ordenados respetando, además, el orden en el cual sus elementos están dispuestos.

Definición 1.3.15. *Sean $(X, \preceq)$ y $(X', \preceq')$ dos conjuntos bien-ordenados. Una función $f : X \to X'$ se dice que es un **orden-isomorfismo** si f es biyectiva y preserva el orden, es decir, $f(x) \preceq' f(y)$ siempre que $x \preceq y$. En este caso diremos que los conjuntos X y Y son **orden-isomorfos** y lo escribiremos como $f : (X, \preceq) \cong_0 (X', \preceq')$, o simplemente como $(X, \preceq) \cong_0 (X', \preceq')$.*

Observe que si f es un orden-isomorfismo entre dos conjuntos bien-ordenados $(X, \preceq)$ y $(X', \preceq')$, entonces f envía el primer elemento de X en el primer elemento de X', el segundo elemento de X en el segundo elemento de X', y así sucesivamente. Similarmente, por ser f^{-1} también un orden-isomorfismo, él envía el primer elemento de X' en el primer elemento de X, el segundo elemento de X' en el segundo elemento de X, etc. Por esta razón, la afirmación $(X, \preceq) \cong_0 (X', \preceq')$ significa que X y X' son, esencialmente, lo mismo, o como se dice frecuentemente, X y X' poseen el *mismo tipo de orden*. Por supuesto, si $f : X \to Y$ y $g : Y \to Z$ son orden-isomorfismos, también lo es $g \circ f$.

Algunas de las propiedades de los conjuntos bien-ordenados se presentan a continuación.

Teorema 1.3.16. *Sean $(X, \preceq)$ y $(X', \preceq')$ **conjuntos bien-ordenados**. Si $f : (X, \preceq) \to (X', \preceq')$ es un orden-isomorfismo, entonces se cumple que*

$$f\big(\mathrm{Seg}_X(a)\big) = \mathrm{Seg}_{X'}\big(f(a)\big).$$

para todo $a \in X$.

Prueba. Sea $a \in X$. Si $y \in f(\mathrm{Seg}_X(a))$, entonces existe algún $x \in \mathrm{Seg}_X(a)$ tal que $y = f(x)$. Puesto $x \in \mathrm{Seg}_X(a)$, resulta que $x \prec a$ y como f preserva el orden, entonces $y = f(x) \prec f(a)$, es decir, $y \in \mathrm{Seg}_{X'}(f(a))$. Esto prueba que $f(\mathrm{Seg}_X(a)) \subseteq \mathrm{Seg}_{X'}(f(a))$. Para demostrar la otra inclusión, suponga que $y \in \mathrm{Seg}_{X'}(f(a))$. Puesto que $\mathrm{Seg}_{X'}(f(a)) \subseteq X' = f(X)$, resulta que existe algún $x \in X$ tal que $y = f(x)$. Veamos que $x \in \mathrm{Seg}_X(a)$. En efecto, admitir que $x \notin \mathrm{Seg}_X(a)$, significa que $a \preceq x$ y, por lo tanto, $f(a) \preceq' f(x) = y$ lo que contradice el hecho de que $y \prec' f(a)$. Por esto $\mathrm{Seg}_{X'}(f(a)) \subseteq \mathrm{Seg}_X(a))$ y termina la prueba. ∎

Un hecho simple, pero fundamental, es el siguiente.

Teorema 1.3.17. *Sean $(X, \preceq)$ un **conjunto bien-ordenado**, $Y \subseteq X$ y $f : X \to Y$ un orden-isomorfismo. Entonces $x \preceq f(x)$ para todo $x \in X$.*

Prueba. Sea $A = \{x \in X : f(x) \prec x\}$. Vamos a demostrar que $A = \varnothing$. Suponga, para generar una contradicción, que $A \neq \varnothing$. Entonces A posee un primer elemento al que denotaremos por x_0. Puesto que $x_0 \in A$ se sigue que $f(x_0) \prec x_0$. Sea $x_1 = f(x_0)$. Teniendo en cuenta que $x_1 \prec x_0$ resulta, aplicando f, que $f(x_1) \prec f(x_0) = x_1$ lo cual demuestra que $x_1 \in A$ y, en consecuencia, como x_0 es el elemento más pequeño de A se tiene que $x_0 \prec x_1$. Esta contradicción establece que $A = \varnothing$ y concluye la prueba. ■

Hemos convenido en definir que dos conjuntos bien-ordenados son orden-isomorfos si existe al menos una función biyectiva entre ellos que preserva el orden. En el siguiente resultado se establece que, en realidad, existe un único orden-isomorfismo entre ellos.

Teorema 1.3.18. *Sean $(X, \preceq)$ y $(X', \preceq')$ **conjuntos bien-ordenados**. Si $f : X \to X'$ es un orden-isomorfismo, entonces f es **único**.*

Prueba. Suponga que $g : X \cong_0 X'$ es otro orden-isomorfismo y considere, en primer lugar, la aplicación $h : X \to X$ definida por $h = f^{-1} \circ g$. Es fácil establecer que h es un orden-isomorfismo. Se sigue del Teorema 1.3.17 que $x \preceq h(x)$ para todo $x \in X$. Si ahora aplicamos f a lo anterior, vemos que $f(x) \preceq' f(h(x)) = g(x)$. Similarmente, definiendo $h' = g^{-1} \circ f$, resulta, como antes, que h' es un orden-isomorfismo verificando que $x \preceq h'(x)$ para todo $x \in X$. De esto se sigue que $g(x) \preceq' g(h'(x)) = f(x)$ y, por lo tanto, $f = g$. Fin de la prueba. ■

Otra consecuencia del Teorema 1.3.17 es que ningún segmento inicial de un conjunto bien-ordenado $(X, \preceq)$ puede ser orden-isomorfo a dicho conjunto.

Corolario 1.3.19. *Sea $(X, \preceq)$ un **conjunto bien-ordenado**. Entonces para cualquier $x \in X$, el segmento inicial $\mathrm{Seg}(x)$ no es orden-isomorfo a X.*

Prueba. Suponga $f : (X, \preceq) \cong_0 (\mathrm{Seg}(a), \preceq)$ es un orden-isomorfismo para algún $a \in X$. Por el Teorema 1.3.17 sabemos que $x \preceq f(x)$ para todo $x \in X$. En particular, $a \preceq f(a)$. Por otro lado, como f es sobreyectiva, entonces $f(X) = \mathrm{Seg}(a)$, de donde resulta que $f(a) \in \mathrm{Seg}(a)$ y, por lo tanto, $f(a) \prec a$. Esta contradicción termina la prueba. ■

1.3.4. Números Ordinales

El objetivo fundamental de esta sección es la de introducir un criterio que asigne a cada conjunto bien-ordenado X un único objeto de la Teoría de Conjuntos, denotado por $\mathrm{ord}(X)$, de modo que

$$X \cong_0 Y \quad \Leftrightarrow \quad \mathrm{ord}(X) \;=\; \mathrm{ord}(Y). \tag{$*$}$$

Tal vez un buen indicio para aproximarnos a ese objeto es el siguiente argumento. Suponga que $(X, \preceq)$ es un conjunto bien-ordenado y considere la siguiente colección de conjuntos

$$\mathrm{Seg}(X) \;=\; \big\{\mathrm{Seg}(x) : x \in X\big\}.$$

Puesto que X es un conjunto bien-ordenado, cualquier par de elementos $x, y \in X$ son comparables y, por lo tanto, también lo son los segmentos $\mathrm{Seg}(x)$ y $\mathrm{Seg}(y)$. Afirmamos que $(\mathrm{Seg}(X), \subseteq)$

es un conjunto bien-ordenado. Para ver esto, sea S un subconjunto no-vacío de Seg(X). Esto significa que

$$S = \{\mathrm{Seg}(x) : x \in A\}$$

para algún subconjunto no vacío A de X y como X es un conjunto bien-ordenado resulta que A posee un primer elemento, llamémoslo x_0. Entonces $\mathrm{Seg}(x_0)$ es el menor elemento de S y termina la prueba de nuestra afirmación. Veamos que $(X, \preceq)$ y $(\mathrm{Seg}(X), \subseteq)$ son orden-isomorfos. En efecto, la función

$$f_* : (X, \preceq) \to (\mathrm{Seg}(X), \subseteq),$$

definida por

$$f_*(x) = \mathrm{Seg}(x) \quad \text{para todo } x \in X \tag{2}$$

es biyectiva y, por supuesto, preserva el orden. Se sigue del Teorema 1.3.18 que f_* es el único orden-isomorfismo que cumple (2). Podemos proponer tomar $\mathrm{ord}(X) = f_*(X)$. En el transcurso de esta sección veremos una manera precisa de formalizar esta idea.

Siguiendo a John von Neumann (1903-1957) podemos extender la definición de número natural a cualquier conjunto bien-definido del modo siguiente.

Definición 1.3.20. *Un **número ordinal** es un conjunto bien-ordenado $(X, \preceq)$ tal que $\mathrm{Seg}(x) = x$ para todo $x \in X$.*

Usualmente usaremos el término "ordinal" en lugar de "número ordinal". Los *ordinales finitos* son esencialmente los números naturales (junto con el 0) que se construyen partiendo del conjunto vacío $\varnothing$ por una aplicación repetida de la "operación sucesor", esto es,

$$\begin{aligned}
0 &= \varnothing = \mathrm{Seg}(0), \\
1 &= \{0\} = \{\varnothing\} = \mathrm{Seg}(1), \\
2 &= \{0,1\} = \{\varnothing, \{\varnothing\}\} = \mathrm{Seg}(2), \\
3 &= \{0,1,2\} = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\} = \mathrm{Seg}(3), \\
&\vdots \\
n &= \{0,1,2,\ldots,n-1\} = \mathrm{Seg}(n), \\
&\vdots \\
\omega_0 &= \{0,1,2,\ldots\} = \mathrm{Seg}(\mathbb{N}_0) \\
\omega_0 + 1 &= \{0,1,2,\ldots,\omega_0\} \\
&\vdots
\end{aligned}$$

Estos números, los ordinales, son usados para etiquetar los pasos de cualquier proceso inductivo transfinito.

Ya hemos visto que si $(X, \preceq)$ y $(Y, \preceq')$ son conjuntos bien-ordenados y si ellos son orden-isomorfos, entonces dicho isomorfismo es único. Lo que resulta interesante es que si ellos también son ordinales, entonces $X = Y$ como se muestra en el siguiente:

Teorema 1.3.21. *Si $(X, \preceq)$ y $(X', \preceq')$ son **números ordinales** tales que $(X, \preceq) \cong_0 (X', \preceq')$, entonces $X = X'$.*

Prueba. Sea $f : (X, \preceq) \cong_0 (X', \preceq')$ el único orden-isomorfismo dado por el Teorema 1.3.18. Vamos a demostrar que $f = \mathrm{Id}$. Para ver esto, considere el conjunto

$$A = \{x \in X : f(x) \neq x\}$$

y veamos que $A = \varnothing$. En efecto, suponga por un momento que $A \neq \varnothing$ y sea a el primer elemento de A. Sea $x \in \mathrm{Seg}_X(a)$. Entonces $x \prec a$ y como $x \notin A$ resulta que $f(x) = x$, de donde se sigue que $x = f(x) \prec' f(a)$. Esto prueba que $\mathrm{Seg}_X(a) \subseteq \mathrm{Seg}_{X'}(f(a))$. Recíprocamente, sea $y \in \mathrm{Seg}_{X'}(f(a))$. Como $\mathrm{Seg}_{X'}(f(a)) = f(\mathrm{Seg}_X(a))$, existe un $x \in \mathrm{Seg}_X(a)$ tal que $y = f(x)$. Ya que $x \in \mathrm{Seg}_X(a)$, se sigue de la primera parte que $x = f(x) = y$ y, así, $y \in \mathrm{Seg}_X(a)$, es decir, $\mathrm{Seg}_{X'}(f(a)) \subseteq \mathrm{Seg}_X(a)$ y, por lo tanto,

$$\mathrm{Seg}_X(a) = \mathrm{Seg}_{X'}(f(a)).$$

Sin embargo, como X y Y son ordinales, resulta que $a = \mathrm{Seg}_X(a) = \mathrm{Seg}_{X'}(f(a)) = f(a)$, contrario al hecho de que $a \in A$. Esta contradicción establece que $A = \varnothing$ y termina la prueba.■

Corolario 1.3.22. *Si $(X, \preceq)$ es un **conjunto bien-ordenado** y $(X', \preceq')$ es un **ordinal** con $X \cong_0 X'$, entonces $X = X'$.*

Prueba. En vista al resultado anterior es suficiente demostrar que $(X, \preceq)$ es un ordinal. Sea $f : (X, \preceq) \cong_0 (Y, \preceq')$ el único orden-isomorfismo entre ellos y sea $x \in X$. Entonces

$$\begin{aligned} x &= f^{-1}(f(x)) \\ &= f^{-1}(\mathrm{Seg}_{X'}(f(x))) \quad \text{por ser } X' \text{ un ordinal} \\ &= \mathrm{Seg}_X(f^{-1}(f(x))) \quad \text{por el Teorema 1.3.16} \\ &= \mathrm{Seg}_X(x), \end{aligned}$$

lo cual finaliza la prueba. ■

Dos observaciones importantes podemos derivar de la noción de número ordinal.

(1) *Si $(X, \preceq)$ es un número ordinal, entonces se cumple que*:

$$x \in X \quad \Rightarrow \quad x \subseteq X.$$

Esto sigue directamente del hecho de que $x = \mathrm{Seg}(x) \subseteq X$.

(2) *La relación de orden de cualquier número ordinal se puede sustituir por la relación de subconjunto.* Para ver esto, sea $(X, \preceq)$ un número ordinal y sean $x, y \in X$. Entonces

$$x \preceq y \quad \Leftrightarrow \quad \mathrm{Seg}(x) \subseteq \mathrm{Seg}(y) \quad \Leftrightarrow \quad x \subseteq y.$$

La primera equivalencia siempre se cumple en cualquier conjunto bien-ordenado, mientras que la segunda $x = \mathrm{Seg}(x) \subseteq \mathrm{Seg}(y) = y$ es válida por ser X un número ordinal. Por consiguiente,

$$(X, \preceq) \text{ es un número ordinal} \Leftrightarrow \begin{cases} (X, \subseteq) \text{ es un conjunto bien-ordenado} \quad \text{y} \\ \\ x = \{a \in X : a \subseteq x\} \text{ para todo } x \in X. \end{cases} \tag{BO}$$

Por esta razón, cuando escribamos: "sea X un número ordinal" sin mencionar el buen-orden, siempre se podrá suponer, cuando convenga, que el orden de éste es $\subseteq$.

Para alcanzar el objetivo propuesto al comienzo de esta sección, debemos comprobar que *cualquier conjunto bien-ordenado es orden-isomorfo a un único número ordinal*. Esto, sin embargo, tomará un poquito de tiempo. Comencemos.

Teorema 1.3.23. *Sea $(X, \preceq)$ un **número ordinal**. Si $a \in X$, entonces $a = \mathrm{Seg}(a)$ es un **número ordinal**.*

Prueba. Fijemos $a \in X$. Puesto que todo subconjunto de un conjunto bien-ordenado hereda dicha propiedad, resulta que $(\mathrm{Seg}_X(a), \preceq)$ está bien-ordenado. Pongamos $Y = \mathrm{Seg}_X(a)$ y veamos que $\mathrm{Seg}_Y(b) = b$ para cualquier $b \in Y$. En efecto, si $b \in Y$, entonces $b \prec a$ y, por lo tanto, $\mathrm{Seg}_X(b) \subseteq \mathrm{Seg}_X(a)$. De esto se sigue que

$$
\begin{aligned}
z \in \mathrm{Seg}_Y(b) \quad &\Leftrightarrow \quad z \in Y \ \wedge \ z \in \mathrm{Seg}_X(b) \\
&\Leftrightarrow \quad z \in \mathrm{Seg}_X(a) \ \wedge \ z \prec b \\
&\Leftrightarrow \quad z \prec a \ \wedge \ z \prec b \\
&\Leftrightarrow \quad z \in \mathrm{Seg}_X(a) \cap \mathrm{Seg}_X(b) \ = \ \mathrm{Seg}_X(b) \ = \ b,
\end{aligned}
$$

donde la última igualdad es válida por el hecho de ser X un número ordinal. Esto nos muestra que $\mathrm{Seg}_X(a)$ es un número ordinal y termina la prueba. ■

El siguiente resultado estable que cualquier subconjunto de un número ordinal el cual también es un número ordinal, es un segmento inicial.

Teorema 1.3.24. *Sea $(X, \preceq)$ un **número ordinal**. Si $Y \subsetneqq X$ es un **número ordinal**, entonces $Y = \mathrm{Seg}(a)$ para algún $a \in X$.*

Prueba. Sea a el primer elemento de $X \setminus Y$. Entonces $\mathrm{Seg}_X(a) \subseteq Y$. Para demostrar que $Y \subseteq \mathrm{Seg}_X(a)$, seleccione un elemento arbitrario $b \in Y$. Como Y y X son ambos ordinales, resulta que

$$\mathrm{Seg}_Y(b) \ = \ b \ = \ \mathrm{Seg}_X(b).$$

De aquí se deduce que $b \preceq a$. En efecto, si fuera $a \prec b$, entonces $a \in \mathrm{Seg}_X(b)$ y por lo anterior tendríamos que $a \in \mathrm{Seg}_Y(b)$ y, en consecuencia, $a \in Y$ lo cual es imposible. Así, $b \preceq a$. Más aun, como $b \in Y$ resulta que $b \neq a$ y, por consiguiente, $b \prec a$. Esto último nos muestra que $b \in \mathrm{Seg}_X(a)$ y, por lo tanto, $Y \subseteq \mathrm{Seg}_X(a)$. Así $Y = \mathrm{Seg}_X(a)$ y finaliza la prueba. ■

En lo que sigue, el símbolo **Ord** *se usará para denotar la colección de todos los **números ordinales***. Del último resultado se observa que: si $X, Y \in \mathbf{Ord}$ con $X \neq Y$, entonces

$$
\begin{aligned}
X \subseteq Y \quad &\Leftrightarrow \quad X \ = \ \mathrm{Seg}_X(a) \quad \text{para algún } a \in Y \\
&\Leftrightarrow \quad X \ = \ a \quad \text{puesto que } \mathrm{Seg}_X(a) = a \\
&\Leftrightarrow \quad X \ \in \ Y \quad \text{puesto que } a \in Y
\end{aligned}
$$

Por consiguiente,

Hecho 1. *El orden de subconjunto $\subseteq$ y el orden de pertenencia $\in$ considerados sobre* **Ord** *son equivalentes* y, por lo tanto, la clase **Ord** está bien-ordenada por $\subseteq$, o equivalentemente, por $\in$. Esta observación, combinada con (BO) de la página 59, permite la siguiente formulación equivalente de número ordinal.

Definición 1.3.25. *Un conjunto bien-ordenado X es un **número ordinal**, que denotaremos por* $\operatorname{ord}(X)$, *si, y sólo si,*

(a) *X es **transitivo**, esto es,* $(x \in X \Rightarrow x \subseteq X)$, *y*

(b) *X está bien-ordenado por $\in$.*

De la definición anterior resulta claro que: *si X y X' son conjuntos bien-ordenados, entonces*

$$X \cong_0 X' \quad \Leftrightarrow \quad \operatorname{ord}(X) = \operatorname{ord}(X').$$

Observe que si tenemos dos números ordinales, digamos $(X, \preceq)$ y $(Y, \preceq')$, y si $X \cap Y \neq \varnothing$, entonces $\preceq$ y $\preceq'$ coinciden sobre $X \cap Y$ y, en consecuencia, $X \cap Y$ está bien-ordenado (con uno cualquiera de los dos ordenes anteriores). Más aun, se tiene que:

Teorema 1.3.26. *Si $(X, \preceq)$ y $(Y, \preceq')$ son **números ordinales**, entonces $X \cap Y$ es un **número ordinal**.*

Prueba. Claramente $X \cap Y$ está bien-ordenado. Sea $a \in X \cap Y$. Por definición, $\operatorname{Seg}_X(a) = a = \operatorname{Seg}_Y(a)$ y en consecuencia,

$$a = \operatorname{Seg}_X(a) \cap \operatorname{Seg}_Y(a) = \{z \in X \cap Y : z \prec a\} = \operatorname{Seg}_{X \cap Y}(a).$$

La prueba es completa. ■

Si bien es cierto que ningún conjunto bien-ordenado es orden-isomorfo a ninguno de sus segmentos iniciales resulta, sin embargo, que siempre es cierto, para cualquier par de números ordinales, lo siguiente.

Teorema 1.3.27. *Sean $(X, \preceq)$ y $(Y, \preceq')$ **números ordinales**. Entonces se cumple una, y sólo una, de las siguientes afirmaciones:*

(a) $X = Y$.

(b) $X \cong_0 \operatorname{Seg}_Y(y)$ para algún $y \in Y$.

(c) $Y \cong_0 \operatorname{Seg}_X(x)$ para algún $x \in X$.

Prueba. En primer lugar observe que (a) y (b) no pueden ocurrir al mismo tiempo. En efecto, si lo anterior fuese cierto, entonces $Y = X \cong_0 \operatorname{Seg}_Y(y)$ para algún $y \in Y$ lo cual es imposible gracias al Corolario 1.3.19. El mismo argumento nos revela que tampoco puede ser cierto que (a) y (c) se cumplan a la vez. Veamos que tampoco (b) y (c) son posibles. Para ver esto, suponga lo contrario y considere los orden-isomorfismos $f : X \cong_0 \operatorname{Seg}_Y(y)$ para algún $y \in Y$ y $g : Y \cong_0 \operatorname{Seg}_X(x)$ para algún $x \in X$. Entonces la aplicación $g \circ f : X \to \operatorname{Seg}_X(x)$ es un orden-isomorfismo, lo que de nuevo contradice el Corolario 1.3.19.

Suponga ahora que $X \neq Y$ y veamos que (b) o (c) se cumplen. Puesto que $X \neq Y$, entonces $X \subseteq Y$ o $Y \subseteq X$ y el resultado sigue del Teorema 1.3.24. Suponga, para construir una contradicción, que ni (b) ni (c) se cumplen, es decir, que $X \ncong_0 \operatorname{Seg}_Y(b)$ para todo $b \in Y$ y también que $Y \ncong_0 \operatorname{Seg}_X(a)$ para todo $a \in X$. Puesto que $X \cap Y \subseteq X$ y $X \cap Y \subseteq Y$ se sigue del

Teorema 1.3.26 que $X \cap Y$ es un número ordinal y, así, gracias al Teorema 1.3.24, $X \cap Y = \mathrm{Seg}_X(a)$ para algún $a \in X$ y, similarmente, $X \cap Y = \mathrm{Seg}_Y(b)$ para algún $b \in Y$. Usando ahora el hecho de que X y Y son números ordinales, resulta que

$$a = \mathrm{Seg}_X(a) = X \cap Y = \mathrm{Seg}_Y(b) = b.$$

Afirmamos que $a \in \mathrm{Seg}_X(a)$. En efecto, como $a \in X$ y $a = b \in Y$ resulta que $a \in X \cap Y = \mathrm{Seg}_X(a)$, es decir, $a \in \mathrm{Seg}_X(a)$ lo cual, como sabemos, es imposible. Esta contradicción establece el resultado. ∎

En lo sucesivo, un número ordinal será denotado, siguiendo la tradición, por una de las letras griegas α, β, etc. El Teorema 1.3.27 permite *comparar* cualquier par de números ordinales del modo siguiente: si α y β son números ordinales y si definimos

$$\alpha \leq \beta \qquad \text{si, y sólo si,} \qquad \alpha \in \beta \quad \text{ó} \quad \alpha = \beta,$$

resulta que:

Definición 1.3.28 (Tricotomía de Ordinales). *Para cualesquiera dos números ordinales α y β, se cumple una, y sólo una, de las siguiente tres posibilidades: $\alpha < \beta$, $\alpha = \beta$ ó $\beta < \alpha$.*

La relación de orden $\leq$ que acabamos de definir sobre la colección **Ord** de los números ordinales la llamaremos el **orden canónico** de los números ordinales.

Similar a la definición de sucesión en un conjunto X, ahora generalizamos dicha noción a conjunto de índices que son números ordinales.

Definición 1.3.29. *Sea X un conjunto no-vacío y θ un ordinal infinito. Cualquier aplicación $L : \mathrm{Seg}(\theta) \to X$ se llama una* ***sucesión transfinita*** *en X.*

Identificaremos, como siempre, la aplicación L de la definición anterior con el conjunto de sus imágenes $\{x_\alpha : \alpha < \theta\}$, donde $x_\alpha = L(\alpha)$ para cualquier $\alpha < \theta$. Siguiendo la tradición, escribiremos $(x_\alpha)_{\alpha<\theta}$ en lugar de $\{x_\alpha : \alpha < \theta\}$. Si en lugar de puntos se considera una colección $\mathcal{A}$ de subconjuntos de X, entonces estaremos hablando de una **sucesión transfinita de conjuntos** que denotaremos por $(A_\alpha)_{\alpha<\theta}$. La sucesión transfinita de conjuntos $(A_\alpha)_{\alpha<\theta}$ se dice

(a) **no decreciente** si $F_{\alpha'} \subseteq F_\alpha$ siempre que $\alpha' < \alpha < \theta$ y

(b) **no creciente** si $F_{\alpha'} \supseteq F_\alpha$ siempre que $\alpha' < \alpha < \theta$.

Sea $(x_\alpha)_{\alpha<\theta}$ una sucesión transfinita en un conjunto X. Diremos que la sucesión $(x_\alpha)_{\alpha<\theta}$ es **estrictamente creciente** si $x_\alpha < x_\upsilon$ siempre que $\alpha < \upsilon < \theta$. Si ocurre que $x_\alpha > x_\upsilon$ para cualquier $\alpha < \upsilon < \theta$, entonces diremos que la sucesión $(x_\alpha)_{\alpha<\theta}$ es **estrictamente decreciente**.

Teorema 1.3.30. *Sea $\mathcal{A}$ un conjunto arbitrario de* ***números ordinales****. Entonces $(\mathcal{A}, \leq)$ es un* ***conjunto bien-ordenado****.*

Prueba. Suponga que $(\mathcal{A}, \leq)$ no está bien-ordenado. Se sigue de la observación (c) de la página 53, que $\mathcal{A}$ contiene una sucesión $(\alpha_n)_{n=0}^\infty$ estrictamente decreciente. Por definición, esto significa que,

$$\alpha_0 \ni \alpha_1 \ni \cdots \ni \alpha_n \ni \cdots$$

lo cual es equivalente, gracias al **Hecho 1**, a que

$$\alpha_0 \supset \alpha_1 \supset \cdots \supset \alpha_n \supset \cdots$$

Esto último nos revela que $(\alpha_n)_{n=1}^{\infty}$ es una sucesión estrictamente decreciente contenida en el conjunto bien-ordenado α_0 lo cual es imposible. ■

Corolario 1.3.31. *Sea $\mathcal{A}$ un conjunto de **números ordinales**. Las siguientes condiciones son equivalentes:*

(1) *$\mathcal{A}$ es un **número ordinal**.*

(2) *Para cada $\alpha \in A$, si β es un **número ordinal** con $\beta < \alpha$, entonces $\beta \in \mathcal{A}$.*

Prueba. $(1) \Rightarrow (2)$ sigue de la definición de número ordinal. Para ver que $(2) \Rightarrow (1)$, observe, en primer lugar, que gracias al Teorema 1.3.30, $\mathcal{A}$ es un conjunto bien-ordenado. Sólo resta demostrar que $\alpha = \mathrm{Seg}(\alpha)$ para cada $\alpha \in A$. Veamos esto. Sea $\alpha \in \mathcal{A}$. Como α es un número ordinal, también lo es $\mathrm{Seg}(\alpha)$, de modo que si $\beta \in \mathrm{Seg}(\alpha)$, entonces el Teorema 1.3.23 no dice que β es un número ordinal satisfaciendo $\beta < \alpha$ y, por lo tanto, $\beta \in \alpha$. Esto prueba que $\mathrm{Seg}(\alpha) \subseteq \alpha$. Para demostrar la otra inclusión, es decir, $\alpha \subseteq \mathrm{Seg}(\alpha)$, sea $\beta \in \alpha$. Por el Teorema 1.3.23, β es un número ordinal tal que, por definición, $\beta < \alpha$. Nuestra hipótesis ahora nos garantiza que $\beta \in \mathcal{A}$ y, por lo tanto, $\beta \in \mathrm{Seg}(\alpha)$. Con esto se concluye que $\alpha = \mathrm{Seg}(\alpha)$ y, por consiguiente, $\mathcal{A}$ es un número ordinal. ■

El siguiente resultado es uno de los más importantes en la obtención de números ordinales.

Teorema 1.3.32. *Sea $\mathcal{A}$ un conjunto de **números ordinales**. Entonces **existe un número ordinal** α_0 con las siguientes propiedades:*

(a) $\alpha \leq \alpha_0$ para todo $\alpha \in \mathcal{A}$, y

(b) si β es otro número ordinal tal que $\alpha \leq \beta$ para todo $\alpha \in \mathcal{A}$, entonces $\alpha_0 \leq \beta$.

Prueba. Cada elemento de $\mathcal{A}$ es un conjunto de ordinales que verifica la condición del Corolario 1.3.31. De allí que

$$\alpha_0 = \bigcup_{\alpha \in \mathcal{A}} \alpha$$

es un conjunto de ordinales y como tal satisface la condición del Corolario 1.3.31. De allí que α_0 es un número ordinal. Por otro lado, para cada $\alpha \in \mathcal{A}$, se tiene que $\alpha \subseteq \alpha_0$, lo cual es equivalente a decir que $\alpha \in \alpha_0$ y, así, por definición, $\alpha \leq \alpha_0$. Para terminar la prueba, suponga que β es un número ordinal tal que $\alpha \leq \beta$ para todo $\alpha \in \mathcal{A}$. Entonces $\alpha \subseteq \beta$ para todo $\alpha \in \mathcal{A}$, de donde se sigue que $\alpha_0 \subseteq \beta$ y, por lo tanto, $\alpha_0 \leq \beta$. ■

El número ordinal α_0 del resultado anterior se llama el **supremo** de $\mathcal{A}$ y será denotado, en lo sucesivo, por $\alpha_0 = \sup \mathcal{A}$. Así,

$$\alpha_0 = \sup A = \bigcup_{\alpha \in \mathcal{A}} \alpha$$

y α_0 no es, necesariamente, un elemento de $\mathcal{A}$. Por ejemplo, si

$$\alpha_0 = \sup \mathbb{N} = \{1, 2, \ldots\}$$

resulta que $\alpha_0 \notin \mathbb{N}$.

La siguiente definición permitirá la construcción de una colección indescriptiblemente inmensa de ordinales, todos con la misma cardinalidad.

Definición 1.3.33. *Sea α un ordinal. Llamaremos **sucesor inmediato** de α al número ordinal $\alpha \cup \{\alpha\}$ al que denotaremos por α^+ o $\alpha + 1$. Si α y β son ordinales y $\beta^+ = \alpha$, entonces diremos que β es el **predecesor inmediato** de α. Un ordinal sin un predecesor inmediato es llamado un **ordinal límite**, esto es, α es un ordinal límite si*

$$\alpha = \sup\{\beta : \beta < \alpha\}.$$

Nótese que, para cualquier ordinal α siempre existe α^+ y se cumple que $\alpha < \alpha^+$. Sin embargo, no todo ordinal posee un predecesor inmediato. Por lo tanto, un ordinal α es un ordinal límite si α no es sucesor de ningún otro ordinal.

A partir de este momento, usaremos el símbolo ω para designar *el **ordinal de los números naturales***, esto es,

$$\omega = \operatorname{ord}(\mathbb{N}).$$

Observe que ω es el *primer ordinal infinito numerable* y, en consecuencia, el *primer ordinal límite*. Usando la definición anterior podemos comenzar a producir ordinales infinitos. En efecto, el siguiente número ordinal después ω es ω^+ y se puede continuar generando ordinales numerables del modo siguiente:

$$\omega^+ = \omega + 1, \quad (\omega+1)^+ = \omega + 2, \quad (\omega+2)^+ = \omega + 3, \cdots$$

En esta escala, después de $\omega, \omega+1, \omega+2, \ldots$ viene

$$\omega + \omega = \omega 2 = \{1, 2, \ldots, \omega, \omega+1, \omega+2, \ldots\}.$$

Similarmente, después de $\omega 2, \omega 2 + 1, \omega 2 + 2, \ldots$ se consigue

$$\omega 2 + \omega = \omega 3 = \{1, 2, \ldots, \omega, \omega+1, \omega+2, \ldots, \omega 2, \omega 2+1, \omega 2+2, \ldots\}.$$

Más allá de $\omega 2, \omega 3, \ldots$ está ω^2 y después de $\omega^2, \omega^3, \ldots$ se obtiene ω^ω. Si se continúa indefinidamente con este proceso se logra construir una gigantesca cantidad de números ordinales

$$\begin{aligned}&\omega, \ldots, \omega 2, \ldots, \omega 3, \ldots, \omega^2, \ldots, \omega^2+\omega, \ldots, \omega^2+\omega 2, \ldots, \omega^3, \ldots,\\ &\omega^\omega, \ldots, \omega^{\omega^\omega}, \ldots\end{aligned}$$

Esto ordinales son sólo una brevísima muestra de los así llamados *ordinales numerables*. Es costumbre definir

$$\epsilon = \sup\left\{\omega, \omega^\omega, \omega^{\omega^\omega}, \ldots\right\}$$

y, en consecuencia, formar $\epsilon + 1$, $\epsilon + \omega$, ϵ^ϵ, $\epsilon^{\epsilon^\epsilon}$, etc. Es importante destacar que la lista de los ordinales nunca termina y que ninguno de los ordinales: ω, $\omega 2$, $\ldots$, ω^2, $\ldots$, ω^ω, $\ldots, \epsilon, \ldots$ posee un predecesor inmediato. Cada uno de ellos es, por supuesto, un ordinal límite.

Si ahora consideramos la colección $\mathbf{Ord}(\omega)$ *de todos los **ordinales numerables***, entonces el Teorema 1.3.32 permite definir el ordinal

$$\omega_1 = \sup \mathbf{Ord}(\omega)$$

Nótese que ω_1 es no-numerable. En efecto, si ω_1 fuese numerable, entonces $\omega_1 \in \omega_1$ lo cual está prohibido en **ZFC**. De hecho, ω_1 es *el primer ordinal no-numerable* ya que si $\alpha < \omega_1$, entonces $\alpha \in \omega_1$ y, en consecuencia, α es numerable.

De este modo, los ordinales numerables ofrecen un camino distinto para producir conjuntos no-numerables. Por supuesto, podemos, como antes, continuar definiendo

$$\omega_1 + 1,\ \omega_1 + 2,\ \ldots,\ \omega_1 + \omega,\ \ldots,\ \omega_1 + \omega 2,\ \ldots$$

Los dos ejemplos siguientes permiten visualizar, a vuelo de pájaros, que todos los ordinales construidos anteriormente, con excepción de ω_1 y sus seguidores, son numerables.

Ejemplo 1.3.2. *Si α es un ordinal con $\alpha \geq \omega$, entonces la aplicación $f : \alpha \to \alpha^+$ definida por*

$$\begin{aligned} f(1) &= \alpha, \\ f(n+1) &= n, \quad \textit{para } n < \omega \\ f(\beta) &= \beta, \quad \textit{si } \omega \leq \beta < \alpha \end{aligned}$$

es una biyección. Por inducción, se comprueba que $\omega, \omega+1, \omega+2, \ldots$ todos son numerables.

Ejemplo 1.3.3. *Similarmente, la aplicación $g : \omega \to \omega 2$ dada por*

$$f(n) = \begin{cases} n/2 & \textit{si } n \textit{ es par} \\ \omega + (n-1)/2 & \textit{si } n \textit{ es impar} \end{cases}$$

también es una biyección. Por lo tanto, usando ejemplo anterior, se prueba que $\omega 2, \omega 2 + 1, \omega 2 + 2, \ldots$ son numerables. Continúe.

La importancia del siguiente resultado reside, precisamente, en impedir que la inmensa colección **Ord** sea un conjunto en **ZFC**.

Teorema 1.3.34. *Sea $\mathcal{A}$ un conjunto de **números ordinales**. Entonces existe un $\beta_0 \in \mathbf{Ord} \setminus \mathcal{A}$ tal que $\alpha < \beta_0$ para todo $\alpha \in \mathcal{A}$.*

Prueba. Por el Teorema 1.3.32 existe el número ordinal $\alpha_0 = \sup \mathcal{A}$. Es claro que el número ordinal $\beta_0 = \alpha_0^+$ satisface la conclusión del teorema. ■

Del resultado anterior se concluye que **Ord** *no es un conjunto*. En efecto, suponer que **Ord** es un conjunto implicaría, gracias al Teorema 1.3.34, que existe un ordinal fuera de **Ord** lo cual es absurdo.

Ya vimos que en **ZFC** la colección $\mathcal{U}$ de todos los conjuntos no es un conjunto. Sin embargo, se puede probar que $\mathcal{U}$ es una *jerarquía acumulativa*, es decir, cualquier conjunto pertenece a algún V_α para algún ordinal α, donde los V_α son definidos como sigue:

$$\begin{aligned} V_0 &= \varnothing. \\ V_{\alpha+1} &= \mathscr{P}(V_\alpha). \\ V_\lambda &= \bigcup_{\alpha<\lambda} V_\alpha, \quad \text{si } \lambda \text{ es un ordinal límite.} \end{aligned}$$

$$\text{Entonces} \quad \mathcal{U} = \bigcup_{\alpha \in \mathbf{Ord}} V_\alpha.$$

Uno de los resultados fundamentales que conduce a definir con claridad la noción de número cardinal es el que se muestra en el próximo teorema. Sin embargo, su demostración se sustenta sobre un axioma llamado el Axioma de Sustitución o Esquema de Reemplazo que reza así:

> **Axioma de Sustitución**. *Sea* $\mathbf{P}(u,v)$ *una propiedad tal que para cada objeto* u *existe un único objeto* v *para el cual* $\mathbf{P}(u,v)$ *se satisface. Entonces, para cada conjunto* A*, existe un conjunto* B *con la siguiente propiedad: para cada* $a \in A$*, existe* $b \in B$ *para el cual* $\mathbf{P}(a,b)$ *se cumple.*

Sea $\mathbf{F}$ la operación definida por la propiedad $\mathbf{P}$, esto es, $\mathbf{F}(a)$ es el único b para el cual $\mathbf{P}(a,b)$ se cumple. Entonces el Axioma de Sustitución puede ser establecido como sigue:

Para cualquier conjunto A*, existe un conjunto* B *tal que: para todo* $a \in A$*,* $\mathbf{F}(a) \in B$.

Por supuesto, B puede contener elementos que no son de la forma $\mathbf{F}(a)$ para algún $a \in A$. Sin embargo, siempre se puede sustituir el conjunto B por el conjunto

$$\{b \in B : \mathbf{P}(a,b) \text{ para algún } a \in A\}$$

al que llamaremos la **imagen** de A por $\mathbf{F}$ y denotado por $\mathbf{F}(A) = \{\mathbf{F}(a) : a \in A\}$. En resumen, el Axioma de Sustitución permite, dado el conjunto A, construir una función $\mathbf{F} : A \to B$ tal que $\mathrm{Dom}(\mathbf{F}) = A$, $\mathrm{Im}(\mathbf{F}) = \mathbf{F}(A)$ y $\mathbf{F}(a)$ sea el único $b \in B$ para el cual $\mathbf{P}(a,b)$ se satisface.

Teorema 1.3.35. *Sea* $(X, \preceq)$ *un* ***conjunto bien-ordenado****. Entonces,* X *es* ***orden-isomórfico*** *a único número ordinal.*

Prueba. La prueba la haremos en dos pasos. El primero consiste en demostrar que: *para cada* $x \in X$*, existe un ordinal* α *tal que* $\mathrm{Seg}(x)$ *es orden-isomorfo a* α. Para ver esto, considere el conjunto

$$A = \{x \in X : \mathrm{Seg}(x) \cong_0 \alpha \text{ para algún } \alpha \in \mathbf{Ord}\}$$

Vamos a demostrar, usando el Principio de Inducción Transfinita, que $A = X$. Observe que si X es vacío, entonces $A = \varnothing$ y no hay nada que probar. Suponga que $X \neq \varnothing$. Queremos demostrar que:

(i) el primer elemento de X está en A y

(ii) para todo $x \in X$, si $\mathrm{Seg}(x) \subseteq A$, entonces $x \in A$.

Sea x_0 el primer elemento de X. Puesto que $\mathrm{Seg}(x_0) = \varnothing$ y $\varnothing$ es el ordinal 0, resulta que $x_0 \in A$. Esto prueba (i). Para demostrar (ii), suponga que $x \in X$ es tal que $\mathrm{Seg}(x) \subseteq A$, donde $x \neq x_0$. Elija cualquier $x' \in \mathrm{Seg}(x)$ y observe que como $x' \in A$, entonces existe un ordinal β para el cual $\mathrm{Seg}(x') \cong_0 \beta$. El Teorema 1.3.21 nos garantiza entonces que β es único para cada $x' \in \mathrm{Seg}(x)$. Lo anterior permite que podamos considerar el conjunto $\mathrm{Seg}(x)$ y la propiedad

$$\mathbf{P}(x,\beta) : \text{ si } x' \prec x, \text{ entonces } \beta \text{ es el único número ordinal tal que } \beta \cong_0 \mathrm{Seg}(x').$$

Por el Axioma de Sustitución, existe un conjunto B y una función $f : \mathrm{Seg}(x) \to B$ tal que $f(x')$ es el único ordinal β orden-isomorfo a $\mathrm{Seg}(x')$ para cada $x' \in \mathrm{Seg}(x)$. Uno puede suponer que

$$B = \{\beta \in \mathbf{Ord} : \beta \cong_0 \mathrm{Seg}(x') \text{ para algún } x' \in \mathrm{Seg}(x)\}.$$

Resulta del Corolario 1.3.31 que $\alpha = f(\mathrm{Seg}(x)) = B$ es un número ordinal y como f es estrictamente creciente, se concluye que f es un orden-isomorfismo, esto es, $\mathrm{Seg}(x) \cong_0 \alpha$. Esto prueba $x \in A$ y así, gracias al Principio de Inducción Transfinita, resulta que $A = X$.

El segundo paso consiste en aplicar, de nuevo, el Axioma de Sustitución al conjunto X y a la propiedad (justificada en el paso anterior):

$$\mathbf{P}(x, \beta): \ \beta \text{ es un número ordinal tal que } \beta \cong_0 \mathrm{Seg}(x)$$

para cada $x \in X$. Entonces existe un conjunto B' y una función $g : X \to B'$ tal que $g(x)$ es el único número ordinal β orden-isomorfo al segmento inicial $\mathrm{Seg}(x)$. Un argumento similar al desarrollado en el paso anterior nos revela que X es orden-isomorfo al número ordinal $\alpha = f(X)$. El hecho de que X es orden-isomorfo a un único ordinal sigue del Teorema 1.3.21. Esto termina la prueba. ■

1.3.5. Números Cardinales

Una vez definido el número ordinal de cualquier conjunto bien-ordenado, estamos en condiciones de terminar nuestra tarea: definir el número cardinal de cualquier conjunto X, en realidad, de cualquier conjunto bien-ordenado.

Definición 1.3.36. *Un número ordinal α es llamado un **ordinal inicial** si α no es equipotente a ningún ordinal $\beta < \alpha$.*

Por ejemplo, cualquier número natural es un ordinal inicial. ω es un ordinal inicial ya que él no es equipotente a ningún número natural. Sin embargo, $\omega + 1$ no es un ordinal inicial por el simple hecho de que $\mathrm{card}(\omega) = \mathrm{card}(\omega + 1)$, pero $\omega < \omega + 1$. De modo similar, ninguno de los ordinales $\omega + 2, \omega + 3, \ldots, \omega + \omega, \ldots$ son iniciales.

Que a todo conjunto se le puede asignar un único número cardinal es consecuencia del siguiente resultado.

Teorema 1.3.37. *Cualquier conjunto no vacío X es **equipotente** a un **único ordinal inicial**.*

Prueba. Sea X un conjunto no vacío. El Principio del Buen-orden nos garantiza la existencia de un buen-orden sobre X. Por el Teorema 1.3.35 existe un único ordinal α que es orden-isomorfo a X. Sea

$$F_\alpha = \{\beta \in \mathbf{Ord} : \beta \approx \alpha\}.$$

Entonces F_α es un conjunto bien-ordenado y, por lo tanto, el más pequeño de los elementos de F_α es un ordinal inicial equipotente a X. ■

Lo que acabamos de probar permite justificar, de modo simple, la siguiente definición.

Definición 1.3.38. *Sea X un conjunto. El **número cardinal**, o **cardinal** de X, denotado por $\mathrm{card}(X)$, es el único ordinal inicial equipotente a X.*

Nótese que $\mathrm{card}(X)$ es el ordinal más pequeño que es equipotente con X. Dicho de otro modo, *κ es un número cardinal si él no es equipotente a ningún ordinal $\alpha < \kappa$.* Por ejemplo, el ordinal más pequeño que es equipotente con $\mathbb{N}$ es ω. Por consiguiente, $\mathrm{card}(\omega)$ es un número cardinal al que llamaremos **el cardinal de los números naturales** y que denotaremos por $\aleph_0$. Así,

$$\aleph_0 = \mathrm{card}(\omega).$$

Es interesante observar que $\mathbb{N}$ puede ser pensado de dos maneras: como ω, el *primer ordinal infinito* y también como $\aleph_0$, *el primer cardinal infinito*.

Aunque estamos usando el sistema de los números ordinales para "medir" conjuntos, no todos ellos sirven para ese propósito. Por ejemplo, ningún *ordinal numerable* $\alpha > \omega$ puede ser usado para definir un número cardinal puesto que, en este caso, $\operatorname{card}(\omega) = \operatorname{card}(\alpha)$ y entonces se puede determinar una biyección entre ambos conjuntos.

Por supuesto, la definición de número cardinal que acabamos de suministrar cumple la receta: *si* X, Y *son conjuntos, entonces*

$$X \approx Y \quad \Leftrightarrow \quad \operatorname{card}(X) \;=\; \operatorname{card}(Y).$$

Claramente cualquier número natural es un número cardinal finito.

1.3.6. $\aleph_1$ y el Primer Ordinal No-numerable

Este es otro modo de visualizar el primer ordinal no-numerable. Considere la colección $F_{\mathfrak{c}}$ de todos los números ordinales no-numerables orden-isomorfos a $\mathbb{R}$; es decir,

$$F_{\mathfrak{c}} \;=\; \big\{\beta \in \mathbf{Ord} \;:\; \beta \approx \operatorname{ord}(\mathbb{R})\big\}.$$

Por lo visto anteriormente, Teorema 1.3.30, $F_{\mathfrak{c}}$ es un conjunto bien-ordenado y, en consecuencia, posee un primer elemento al que denotaremos por ω_1. Este número ordinal ω_1 posee las siguientes propiedades:

(1) ω_1 *es el* ***primer ordinal no-numerable***. Por consiguiente, como ω_1 es un ordinal inicial, cualquier ordinal $\beta < \omega_1$ es un *ordinal numerable* y, en consecuencia, ω_1 es el *conjunto de todos los ordinales numerables*:

$$\omega_1 \;=\; \big\{\beta \in \mathbf{Ord} \;:\; \beta \approx \operatorname{ord}(\mathbb{N})\big\} \;=\; \sup\{\alpha \in \mathbf{Ord} \;:\; \alpha < \omega_1\}.$$

(2) ***Ninguna sucesión*** *en* ω_1 *alcanza a* ω_1, esto significa que si $(\alpha_n)_{n=1}^{\infty}$ es cualquier sucesión en ω_1, entonces existe un ordinal $\alpha \in \omega_1$ tal que $\alpha_n \leq \alpha$ para cualquier entero $n \geq 1$. En efecto, basta tomar $\alpha = \sup\{\alpha_n : n \in \mathbb{N}\}$ y observar que $\alpha \in \omega_1$.

Definamos ahora

$$\aleph_1 \;=\; \operatorname{card}(\omega_1).$$

Observe que $\aleph_1$ es *el número cardinal no-numerable más pequeño mayor que* $\aleph_0$. Un paso más. Sea

$$\omega_2 \;=\; \big\{\beta \in \mathbf{Ord} \;:\; \beta \approx \omega_1\big\} \;=\; \sup\{\alpha \in \mathbf{Ord} \;:\; \omega_1 \leq \alpha < \omega_2\},$$

y defina

$$\aleph_2 \;=\; \operatorname{card}(\omega_2).$$

En general, si para cada $n \in \mathbb{N}$, el número cardinal $\aleph_n$ ha sido obtenido, entonces definimos

$$\aleph_{n+1} \;=\; \operatorname{card}(\omega_{n+1}),$$

donde ω_{n+1} es el conjunto de todos los números ordinales cuya cardinalidad es $\aleph_n$. De nuevo, $\aleph_{n+1}$ *es el número cardinal más pequeño mayor que* $\aleph_n$. Para ordinales transfinitos, comenzando con ω, definimos

$$\aleph_\omega \;=\; \operatorname{card}\bigg(\bigcup_{n=1}^{\infty} \omega_n\bigg) \;=\; \operatorname{card}(\sup\{\omega_n : n \in \mathbb{N}\}),$$

y se continúa la construcción de cada $\aleph_\alpha$ para todo número ordinal $\alpha > \omega$. Cada cardinal $\aleph_\alpha$, donde α es un ordinal arbitrario, será llamado un **alef**. Nótese que si $\alpha \geq 1$, entonces $\aleph_\alpha$ es no-numerable.

Observe que todo **cardinal infinito** posee un cardinal sucesor inmediato, es decir, $\aleph_\beta$ es un cardinal sucesor inmediato si y sólamente si el índice β es un ordinal sucesor inmediato. Un cardinal que no es un cardinal sucesor inmediato es llamado un cardinal límite. Ejemplos de cardinales límites son $\aleph_0$, $\aleph_\omega$, $\aleph_{\omega 2}$, $\aleph_{\omega_1}$. Así, la sucesión transfinita de cardinales infinitos es estrictamente creciente:

$$\aleph_0 < \aleph_1 < \aleph_2 < \cdots < \aleph_\omega < \cdots < \aleph_{\omega_1} < \cdots < \aleph_{\omega_\omega} < \cdots$$

Estos cardinales infinitos eran para Cantor algo sagrado: constituían los escalones que conducen a Ω^*, el infinito absoluto, el trono de Dios. Nótese que todo número cardinal es un número natural, o un alef.

Lo antes expuesto nos permite ahora establecer una definición formal de los términos "finito", "numerable" y "no-numerable":

(a) Un conjunto es **finito** si su cardinalidad es menor que $\aleph_0$.

(b) Un conjunto es **numerable** si su cardinalidad es igual a $\aleph_0$.

(c) Un conjunto es **no-numerable** si su cardinalidad es mayor o igual que $\aleph_1$.

Puesto que la cardinalidad de cualquier conjunto está en una de las categorías anteriores, la **Ley de Tricotomía de los Cardinales** establece que cualquier par de cardinales, digamos $\mathfrak{a}$ y $\mathfrak{b}$, son comparables, esto es,

$$\mathfrak{a} < \mathfrak{b}, \qquad \mathfrak{a} = \mathfrak{b} \qquad \text{o} \qquad \mathfrak{a} > \mathfrak{b}.$$

Una de las sorpresas de este principio ocurrió cuando Friedrich Hartogs demostró en el año 1915 que la *Ley de Tricotomía implica el Axioma de Elección*.

Nota Adicional 1.3.4 La siguiente observación será usada frecuentemente. *Si un conjunto X tiene* ***cardinalidad*** $\aleph_1$ *entonces*:

(1) X está *bien-ordenado* y

(2) *Existe una biyección entre ω_1 y X.*

En consecuencia, cada elemento de X puede ser identificado con un único elemento de ω_1 y viceversa. Por esta razón, siempre podemos *representar* a X en la forma:

$$X = \{x_\alpha : \alpha < \omega_1\} = (x_\alpha)_{\alpha < \omega_1}.$$

1.3.7. La Aritmética de los Cardinales

En esta sección definiremos las operaciones aritméticas (suma, multiplicación y exponenciación) de números cardinales y algunas de sus propiedades, pero no así la de números ordinales las cuales difiere sustancialmente de la aritmética cardinal para conjuntos infinitos. Por ejemplo, 2^ω es un ordinal numerable, mientras que $2^{\aleph_0}$ es un cardinal no-numerable.

La aritmética de los cardinales finitos es sencilla en el sentido de que si A y B son conjuntos finitos y *disjuntos* tales que $\operatorname{card}(A) = m$ y $\operatorname{card}(B) = n$, entonces

$$m + n = \operatorname{card}(A \cup B) \qquad \text{y} \qquad m \cdot n = \operatorname{card}(A \times B).$$

Más aun, las operaciones de resta, división y exponenciación se practican de la forma habitual. Por otro lado, la aritmética de los cardinales transfinitos, aunque se basa sobre ideas similares a la del caso finito, es bien distinta a ésta ya que, por ejemplo, no incluye operaciones como la resta y la división, aunque sí la exponenciación. Sin embargo, no es imprescindible en su definición que los conjuntos sean disjuntos.

Definición 1.3.39. *Sean A y B conjuntos arbitrarios. Si* $\operatorname{card}(A) = \mathfrak{a}$ *y* $\operatorname{card}(B) = \mathfrak{b}$, *definimos*

$$\mathfrak{a} + \mathfrak{b} \;=\; \operatorname{card}\big((A \times \{0\}) \cup (B \times \{1\})\big) \qquad y \qquad \mathfrak{a} \cdot \mathfrak{b} \;=\; \operatorname{card}(A \times B).$$

La legitimidad de las operaciones anteriores se sustentan sobre el hecho de que $\mathfrak{a}+\mathfrak{b}$ y $\mathfrak{a}\cdot\mathfrak{b}$ no dependen de la elección de los conjuntos A y B. Se sigue de la definición anterior que

$$\aleph_0 + \aleph_0 \;=\; \aleph_0 \qquad \text{y} \qquad \aleph_0 \cdot \aleph_0 \;=\; \aleph_0.$$

ya que $(\mathbb{N} \times \{0\}) \cup (\mathbb{N} \times \{1\}) \approx \mathbb{N}$ y $\mathbb{N} \times \mathbb{N} \approx \mathbb{N}$. También

$$2^{\aleph_0} + 2^{\aleph_0} \;=\; 2^{\aleph_0} \qquad \text{y} \qquad 2^{\aleph_0} \cdot 2^{\aleph_0} \;=\; 2^{\aleph_0}.$$

por un argumento similar. Véase también el próximo teorema. Por ejemplo, si $\mathfrak{J}_{\mathbb{Q}}$ constituye *la colección de todos los intervalos abiertos con extremos racionales,* entonces

$$\operatorname{card}(\mathfrak{J}_{\mathbb{Q}}) \;=\; \aleph_0.$$

En efecto, existen $\aleph_0$ intervalos abiertos cuyo extremo inferior es un número racional. Una vez fijado un intervalo abierto cuyo extremo izquierdo es un número racional, existen $\aleph_0$ posibilidades de que su extremo derecho sea un racional. Por lo tanto,

$$\operatorname{card}(\mathfrak{J}_{\mathbb{Q}}) \;=\; \aleph_0 \cdot \aleph_0 \;=\; \aleph_0.$$

Las operaciones de suma y producto de números cardinales son siempre conmutativas, asociativas y distributivas, esto es,

$$\mathfrak{a}+\mathfrak{b} \;=\; \mathfrak{b}+\mathfrak{a} \qquad\qquad \mathfrak{a}\cdot\mathfrak{b} \;=\; \mathfrak{b}\cdot\mathfrak{a}$$
$$\mathfrak{a}+(\mathfrak{b}+\mathfrak{d}) \;=\; (\mathfrak{a}+\mathfrak{b})+\mathfrak{d} \qquad\qquad \mathfrak{a}\cdot(\mathfrak{b}\cdot\mathfrak{d}) \;=\; (\mathfrak{a}\cdot\mathfrak{b})\cdot\mathfrak{d}$$
$$\mathfrak{a}\cdot(\mathfrak{b}+\mathfrak{d}) \;=\; \mathfrak{a}\cdot\mathfrak{b}+\mathfrak{a}\cdot\mathfrak{d}$$

Otro hecho que resulta contrario a la lógica de la aritmética de los números cardinales finitos es el siguiente:

Teorema 1.3.40. *Sean $\mathfrak{a}$ y $\mathfrak{b}$ números cardinales. Si $\mathfrak{b}$ es* ***infinito*** *y $\mathfrak{a} < \mathfrak{b}$, entonces*

$$\mathfrak{a} + \mathfrak{b} \;=\; \mathfrak{b} \qquad y \qquad \mathfrak{a}\cdot\mathfrak{b} \;=\; \mathfrak{b}.$$

Prueba. Sean A y B conjuntos tales que $\operatorname{card}(A) = \mathfrak{a}$ y $\operatorname{card}(B) = \mathfrak{b}$. Sin perder generalidad asumiremos que $A \cap B = \varnothing$. En este caso, $\operatorname{card}\big((A \times \{0\}) \cup (B \times \{1\})\big) = \operatorname{card}(A \cup B)$. Como $\mathfrak{a} < \mathfrak{b}$, existe una función inyectiva $f : A \to B$ que no es sobreyectiva. Si ahora definimos $F : A \to f(A)$ por

$$F(x) \;=\; f(x) \quad \text{para todo } x \in A,$$

resulta que F es biyectiva, $\operatorname{card}(A) = \operatorname{card}(F(A))$, $F(A) \cup B = B$ y

$$\mathfrak{a} + \mathfrak{b} \;=\; \operatorname{card}(A \cup B) \;=\; \operatorname{card}(F(A) \cup B) \;=\; \operatorname{card}(B) \;=\; \mathfrak{b}.$$

La prueba de que $\mathfrak{a} \cdot \mathfrak{b} = \mathfrak{b}$ es similar y se invita al lector a realizarla. ■

Del resultado anterior se tiene que si $\mathfrak{a}, \mathfrak{b}$ son cardinales infinitos, entonces se cumple que

$$\mathfrak{a} + \mathfrak{b} \;=\; \mathfrak{a} \cdot \mathfrak{b} \;=\; \text{máx}\{\mathfrak{a},\, \mathfrak{b}\}.$$

En particular, cualesquiera sean $\alpha, \beta \in \mathbf{Ord}$,

$$\aleph_\alpha + \aleph_\beta \;=\; \aleph_\alpha \cdot \aleph_\beta \;=\; \text{máx}\{\aleph_\alpha,\, \aleph_\beta\}.$$

(1) Puesto que $\aleph_0 < 2^{\aleph_0}$, resulta de lo anterior que

$$\aleph_0 + 2^{\aleph_0} \;=\; 2^{\aleph_0} \qquad\qquad \text{y} \qquad\qquad \aleph_0 \cdot 2^{\aleph_0} \;=\; 2^{\aleph_0}.$$

(2) Similarmente, $n + \aleph_0 = \aleph_0$ y $n \cdot \aleph_0 = \aleph_0$ para cualquier $n \in \mathbb{N}$.

En general, si $(\kappa_\alpha)_{\alpha<\theta}$ es una sucesión de números cardinales, su **suma cardinal**, $\sum_{\alpha<\theta} \kappa_\alpha$, se define como

$$\sum_{\alpha<\theta} \kappa_\alpha \;=\; \operatorname{card}\Big(\bigcup_{\alpha<\theta} A_\alpha\Big),$$

donde $\{A_\alpha : \alpha < \theta\}$ es cualquier colección disjunta de conjuntos con $\operatorname{card}(A_\alpha) = \kappa_\alpha$ para todo $\alpha < \theta$. Una definición similar sirve para el producto cardinal.

Para definir la exponenciación de números cardinales, recordemos que para cualquier conjunto X, $\operatorname{card}(\mathscr{P}(X)) = \operatorname{card}\big(2^X\big)$. Usando esta información podemos convenir en definir:

Definición 1.3.41. *Si* $\operatorname{card}(A) = \mathfrak{a}$ *y* $\operatorname{card}(B) = \mathfrak{b}$, *definimos*

$$\mathfrak{a}^{\mathfrak{b}} \;=\; \operatorname{card}(A^B).$$

Como antes, se puede demostrar que esta definición de exponenciación no depende sobre la elección de los conjuntos A y B. Mientras que la suma y multiplicación de números cardinales son relativamente simples, debemos advertir que la evaluación de la exponenciación es más complicada. Por ejemplo, por el Teorema de Cantor, $2^{\aleph_\alpha} > \aleph_\alpha$, en otras palabras,

$$2^{\aleph_\alpha} \;\geq\; \aleph_{\alpha+1}.$$

Pero no hay mucho que pueda demostrarse de $2^{\aleph_\alpha}$, excepto la siguiente propiedad:

$$2^{\aleph_\alpha} \;\leq\; 2^{\aleph_\beta} \qquad \text{siempre que } \alpha \leq \beta.$$

De la definición anterior se puede demostrar, sin mucha dificultad, que:

(a) $\mathfrak{a} \leq \mathfrak{a}^{\mathfrak{b}}$ si $\mathfrak{b} > 0$.

(b) $\mathfrak{a} \leq \mathfrak{b}^{\mathfrak{a}}$ si $\mathfrak{b} > 1$.

(c) Si $\mathfrak{a}_1 \leq \mathfrak{a}_2$ y $\mathfrak{b}_1 \leq \mathfrak{b}_2$, entonces $\mathfrak{a}_1^{\mathfrak{b}_1} \leq \mathfrak{a}_2^{\mathfrak{b}_2}$.

Más propiedades de la exponenciación viene dada por el siguiente resultado.

Teorema 1.3.42. *Sean* $\mathfrak{a}, \mathfrak{b}$ *y* γ *números cardinales. Entonces,*

(1) $\mathfrak{a}^{\mathfrak{b}+\mathfrak{d}} = \mathfrak{a}^{\mathfrak{b}} \cdot \mathfrak{a}^{\mathfrak{d}}$.

(2) $\left(\mathfrak{a}^{\mathfrak{b}}\right)^{\mathfrak{d}} = \mathfrak{a}^{\mathfrak{b}\cdot\mathfrak{d}}$.

(3) $(\mathfrak{a} \cdot \mathfrak{b})^{\mathfrak{d}} = \mathfrak{a}^{\gamma} \cdot \mathfrak{b}^{\mathfrak{d}}$.

Prueba. Sean A, B y C conjuntos tales que $\mathfrak{a} = \mathrm{card}(A)$, $\mathfrak{b} = \mathrm{card}(B)$ y $\mathfrak{d} = \mathrm{card}(C)$. Para demostrar (1), suponga que B y C son disjuntos. Vamos a construir una función inyectiva F de $A^B \times A^C$ sobre $A^{B\cup C}$. Si $(f,g) \in A^B \times A^C$, defina $F(f,g) = h$, donde $h : B \cup C \to A$ viene dada por

$$h(x) = \begin{cases} f(x) & \text{si } x \in B \\ g(x) & \text{si } x \in C. \end{cases}$$

Es claro que $\varphi \in A^{B\cup C}$ y que F es inyectiva. Por supuesto, dada cualquier función $h \in A^{B\cup C}$, las funciones $f = h|_B$ y $g = h|_C$ satisfacen $F(f,g) = h$ y, por lo tanto, F es sobreyectiva.

Para demostrar (2) vamos a considerar la función $F : A^{B\times C} \to (A^B)^C$ definida por $F(f) = g$, donde cada $f : B \times C \to A$ y $g : C \to A^B$ se define, para todo $c \in C$ y todo $b \in B$, por $g(c)(b) = f(b,c)$. Observe que como $g(c) \in A^B$, la expresión $g(c)(b) = f(b,c)$ siempre tiene sentido. Dejamos a cargo del lector la tarea de comprobar que F es biyectiva y también la prueba de (3). ■

Por ejemplo,

$$2^{2^{\aleph_0}} \cdot 2^{2^{\aleph_0}} = 2^{2^{\aleph_0}+2^{\aleph_0}} = 2^{2^{\aleph_0}}.$$

Corolario 1.3.43. *Si* $2 \leq \mathfrak{b} \leq 2^{\mathfrak{a}}$ *y* $\mathfrak{a}$ *es* ***infinito****, entonces* $\mathfrak{b}^{\mathfrak{a}} = 2^{\mathfrak{a}}$.

Prueba. De los resultados anteriores se tiene que

$$2^{\mathfrak{a}} \leq \mathfrak{b}^{\mathfrak{a}} \leq \left(2^{\mathfrak{a}}\right)^{\mathfrak{a}} = 2^{\mathfrak{a}\cdot\mathfrak{a}} = 2^{\mathfrak{a}}.$$

y, por lo tanto, $\mathfrak{b}^{\mathfrak{a}} = 2^{\mathfrak{a}}$. ■

En particular, tomando $\mathfrak{a} = \mathfrak{b} = \aleph_0$ en el resultado anterior, resulta que

$$\aleph_0^{\aleph_0} = 2^{\aleph_0}.$$

Nota Adicional 1.3.5 Fijemos un cardinal infinito κ. Si para cada $\lambda < \kappa$ y cada $\alpha < \lambda$ existe un conjunto A_α con $\mathrm{card}(A_\alpha) \leq \kappa$, entonces

$$\mathrm{card}\Big(\bigcup_{\alpha<\lambda} A_\alpha\Big) \leq \kappa.$$

En efecto, $\mathrm{card}(\bigcup_{\alpha<\lambda} A_\alpha) \leq \kappa \cdot \lambda = \kappa \cdot \kappa = \kappa$.

Sea θ un ordinal límite. Si $(\alpha_\beta)_{\beta<\theta}$ es una sucesión estrictamente creciente de ordinales, definimos

$$\alpha = \sup\{\alpha_\beta : \beta < \theta\}.$$

A α lo llamaremos el **límite** de la sucesión creciente y lo denotaremos por $\alpha = \lim\limits_{\beta\to\theta} \alpha_\beta$.

Un cardinal infinito $\aleph_\alpha$ es un **límite fuerte** si $2^{\aleph_\beta} < \aleph_\alpha$ para todo $\beta < \alpha$.

Definición 1.3.44. *Un cardinal infinito κ se llama **singular** si existe una sucesión estrictamente creciente $(\alpha_\beta)_{\beta<\theta}$ de ordinales, donde θ es un ordinal límite, tal que:*

(a) $\theta < \kappa$,

(b) $\alpha_\beta < \kappa$ *para todo* $\beta < \theta$ *y* $\kappa = \lim_{\beta\to\theta} \alpha_\beta$.

*Un cardinal infinito que no es singular se llama **regular**.*

Por ejemplo, $\aleph_\omega$ es un cardinal singular, pues ω es un ordinal límite,

$$\aleph_\omega = \lim_{n\to\omega} \aleph_n,$$

$\omega < \aleph_\omega$ y $\aleph_n < \aleph_{n+1}$ para cada $n < \omega$. Similarmente, los cardinales $\aleph_{\omega+\omega}$, $\aleph_{\omega\cdot\omega}$, $\aleph_{\omega_1}$ son singulares. Se puede demostrar que existen cardinales singulares arbitrariamente grandes. Por ejemplo, fijemos un cardinal infinito, digamos $\aleph_\alpha$. Entonces

$$\aleph_{\alpha+\omega}\ 0\ \lim_{n\to\omega} \aleph_{\alpha+n}$$

es un cardinal singular mayor que $\aleph_\alpha$.

Definición 1.3.45. *Un cardinal no-numerable κ se llama **inaccesible** si él es un cardinal límite y a la vez regular.*

Otra manera de expresar lo anterior es como sigue: κ es un cardinal inaccesible si:

(1) $\kappa > \aleph_0$,

(2) para todo cardinal $\mathfrak{a} < \kappa$, se tiene que $2^{\mathfrak{a}} < \kappa$, y

(3) para cualquier conjunto de índices I con $\operatorname{card}(I) < \kappa$ y cualquier colección $\{\mathfrak{a}_i : i \in I\}$ de cardinales tal que $\mathfrak{a}_i < \kappa$ para todo $i \in I$, se cumple que $\sup\{\mathfrak{a}_i : i \in I\} < \kappa$.

En otras palabras, un cardinal κ es *inaccesible* si no puede alcanzar por "debajo"; es decir, si no puede obtenerse a partir de las operaciones básicas entre cardinales, utilizando únicamente cardinales menores que él mismo. Los cardinales inaccesibles forman una cofradía de los así llamados **los grandes cardinales** los cuales no pueden demostrarse, en **ZFC**, su consistencia y, en consecuencia, deben asumirse como un axioma adicional.

1.3.8. La Cardinalidad de $\mathbb{R}$ y otros Conjuntos Similares

.

En esta parte analizaremos la cardinalidad del conjunto $\mathbb{R}$ de los números reales y estudiaremos algunos otros ejemplos de conjuntos que poseen esa cardinalidad. Ya hemos visto que $\mathbb{R}$ es un conjunto no-numerable y habíamos asignado el símbolo $\mathfrak{c}$ al número de sus elementos. Ahora es el momento de darle sentido a ese símbolo. De la definición de exponenciación de números cardinales y el Teorema de Cantor, Teorema 1.2.17, se tiene que

$$\operatorname{card}(2^{\mathbb{N}}) = 2^{\aleph_0} = \operatorname{card}(\mathscr{P}(\mathbb{N})).$$

Veamos ahora que:

Teorema 1.3.46. $\mathrm{card}(\mathbb{R}) = 2^{\aleph_0}$.

Prueba. Puesto que toda sucesión de números racionales es una aplicación $s : \mathbb{N} \to \mathbb{Q}$, resulta que su gráfico es un subconjunto de $\mathbb{N} \times \mathbb{Q}$ por lo que el conjunto $\mathcal{S}(\mathbb{Q})$, formado por todas las sucesiones de números racionales, está contenido en $\mathscr{P}(\mathbb{N} \times \mathbb{Q})$. Más aun, como cada número real x es el límite de una sucesión perteneciente a $\mathcal{S}(\mathbb{Q})$ se tiene, usando el Teorema de Cantor, Teorema 1.2.17, y las propiedades de la exponenciación de cardinales, que

$$\mathrm{card}(\mathbb{R}) \le \mathrm{card}(\mathcal{S}(\mathbb{Q})) \le \mathrm{card}\big(\mathscr{P}(\mathbb{N} \times \mathbb{Q})\big) = 2^{\mathrm{card}(\mathbb{N}\times\mathbb{Q})} = 2^{\aleph_0 \cdot \aleph_0} = 2^{\aleph_0}.$$

Esto prueba que $\mathrm{card}(\mathbb{R}) \le 2^{\aleph_0}$. Para demostrar la otra desigualdad, es decir, $2^{\aleph_0} \le \mathrm{card}(\mathbb{R})$ considere la función $f : 2^{\mathbb{N}} \to \mathbb{R}$ definida por

$$f\big((a_n)_{n=1}^{\infty}\big) = 0.a_1 a_2 a_3 \ldots$$

donde $a_n \in \{0, 1\}$ para todo $n \ge 1$. Es fácil ver que f es inyectiva por lo que

$$2^{\aleph_0} = \mathrm{card}\big(2^{\mathbb{N}}\big) \le \mathrm{card}\big(\mathbb{R}\big).$$

La prueba es completa. ■

Por el resultado anterior tenemos que $\mathfrak{c} = 2^{\aleph_0}$ y, por lo tanto, haciendo uso del Teorema de Cantor, Teorema 1.2.19, se tiene que

$$\aleph_0 < 2^{\aleph_0} < 2^{2^{\aleph_0}} < 2^{2^{2^{\aleph_0}}} < \cdots$$

En el siguiente resultado se muestran algunos conjuntos cuya cardinalidad es la del continuo.

Teorema 1.3.47. (1) $\mathrm{card}(\mathbb{C}) = 2^{\aleph_0}$, *donde* $\mathbb{C}$ *es el conjunto de los* ***números complejos***.

(2) $\mathrm{card}\big(\mathbb{R}^n\big) = 2^{\aleph_0}$ *para cada* $n \in \mathbb{N}$.

(3) $\mathrm{card}\big(\mathbb{N}^{\mathbb{N}}\big) = 2^{\aleph_0}$.

(4) $\mathrm{card}\big(\mathbb{R}^{\mathbb{N}}\big) = 2^{\aleph_0}$.

(5) $\mathrm{card}\big(C(\mathbb{R})\big) = 2^{\aleph_0}$, *donde* $C(\mathbb{R})$ *consiste de todas las* ***funciones continuas*** *de* $\mathbb{R}$ *en* $\mathbb{R}$.

(6) $\mathrm{card}\big(\mathcal{O}_{\mathbb{R}}\big) = 2^{\aleph_0}$, *donde* $\mathcal{O}_{\mathbb{R}}$ *es la colección de todos los* ***subconjuntos abiertos*** *de* $\mathbb{R}$.

Prueba. (1) Puesto que la aplicación $\varphi : \mathbb{R} \times \mathbb{R} \to \mathbb{C}$ dada por $\varphi(a, b) = a + ib$ es biyectiva, resulta que

$$\mathrm{card}(\mathbb{C}) = \mathrm{card}(\mathbb{R} \times \mathbb{R}) = 2^{\aleph_0} \cdot 2^{\aleph_0} = 2^{\aleph_0 + \aleph_0} = 2^{\aleph_0}.$$

(2) Esto sigue de lo siguiente:

$$\mathrm{card}(\mathbb{R}^n) = \big(2^{\aleph_0}\big)^n = 2^{n \cdot \aleph_0} = 2^{\aleph_0}.$$

(3) El conjunto $\mathbb{N}^{\infty}$ tiene cardinalidad $\aleph_0^{\aleph_0}$. Se sigue del Corolario 1.3.43 que

$$\mathrm{card}\big(\mathbb{N}^{\mathbb{N}}\big) = \aleph_0^{\aleph_0} = 2^{\aleph_0}.$$

(4) Esto es consecuencia de (2) del Teorema 1.3.42:

$$\operatorname{card}\big(\mathbb{R}^{\mathbb{N}}\big) \;=\; \big(2^{\aleph_0}\big)^{\aleph_0} \;=\; 2^{\aleph_0\cdot\aleph_0} \;=\; 2^{\aleph_0}.$$

(5) Usaremos el siguiente hecho: *Toda **función continua** $f:\mathbb{R}\to\mathbb{R}$ queda completamente determinada por sus valores en* $\mathbb{Q}$. En particular, si dos funciones continuas $f,g:\mathbb{R}\to\mathbb{R}$ son tales que $f(r)=g(r)$ para todo $r\in\mathbb{Q}$, entonces $f=g$. Considere ahora la función $\Phi: C(\mathbb{R})\to\mathbb{R}^{\mathbb{Q}}$ definida por $\Phi(f)=f|_{\mathbb{Q}}$. De la observación anterior resulta que Φ es inyectiva y, por lo tanto,

$$\operatorname{card}(C(\mathbb{R})) \;\le\; \operatorname{card}\big(\mathbb{R}^{\mathbb{Q}}\big) \;=\; \big(2^{\aleph_0}\big)^{\aleph_0} \;=\; 2^{\aleph_0}.$$

Por otro lado, como $C(\mathbb{R})$ contiene a todas la funciones constantes, se tiene que

$$2^{\aleph_0} \;\le\; \operatorname{card}(C(\mathbb{R}))$$

y termina la prueba de (5).

(6) En primer lugar, observe que la función $f:\mathbb{R}\to\mathcal{O}_{\mathbb{R}}$ dada por $f(a)=(a,+\infty)$ es inyectiva, de donde se sigue que

$$2^{\aleph_0} \;=\; \operatorname{card}(\mathbb{R}) \;\le\; \operatorname{card}\big(\mathcal{O}_{\mathbb{R}}\big).$$

Para demostrar la otra desigualdad, nótese que todo conjunto abierto G es la unión de los intervalos abiertos con extremos racionales contenidos en G y como $\mathcal{I}_{\mathbb{Q}}$, el conjunto de todos los intervalos abiertos con extremos racionales, es numerable resulta que existen a lo sumo $2^{\aleph_0}$ de tales uniones. Por esto

$$\operatorname{card}\big(\mathcal{O}_{\mathbb{R}}\big) \;\le\; 2^{\aleph_0}$$

y termina la prueba. ■

Del Teorema 1.3.47 (6) se deduce que:

(7) La cardinalidad de $\mathcal{F}$, *la colección de todos los **subconjuntos cerrados** de* $\mathbb{R}$, es $2^{\aleph_0}$.

Esto sigue simplemente del hecho de que existe una biyección entre $\mathcal{O}_{\mathbb{R}}$ y $\mathcal{F}$ dada por $f(G)=G^c$ para todo $G\in\mathcal{O}_{\mathbb{R}}$. En particular, la colección $\mathcal{F}_{\mathfrak{c}}$ *de todos los elementos de* $\mathcal{F}$ *que son **no-numerables*** posee cardinalidad $2^{\aleph_0}$. En efecto, como todo intervalo cerrado no degenerado es no-numerable, entonces

$$2^{\aleph_0} \;=\; \operatorname{card}\big(\{[x,1]\,:\,x\in\mathbb{R},\ x<1\}\big) \;\le\; \operatorname{card}(\mathcal{F}_{\mathfrak{c}}) \;\le\; \operatorname{card}(\mathcal{F}) \;=\; 2^{\aleph_0}.$$

Si denotamos por $\mathcal{F}_\sigma$ *la familia de todas las **uniones numerables** de **conjuntos cerrados*** y por $\mathcal{G}_\delta$ *la familia de todas las **intersecciones numerables** de **conjuntos abiertos***, resulta que

$$\operatorname{card}(\mathcal{G}_\delta) \;=\; 2^{\aleph_0} \;=\; \operatorname{card}(\mathcal{F}_\sigma).$$

(8) La cardinalidad de $\mathcal{P}$, *la colección de todos los **subconjuntos perfectos** de* $\mathbb{R}$, es $2^{\aleph_0}$.

Recordemos que un conjunto $P\subseteq\mathbb{R}$ es ***perfecto*** si él cerrado y todos sus puntos son puntos de acumulación. Puesto que todo intervalo cerrado no degenarado es un conjunto perfecto (y no-numerable) resulta que

$$2^{\aleph_0} \;=\; \operatorname{card}\big(\{[x,1]\,:\,x\in\mathbb{R},\ x<1\}\big) \;\le\; \operatorname{card}(\mathcal{P}) \;\le\; \operatorname{card}(\mathcal{F}_{\mathfrak{c}}) \;\le\; \operatorname{card}(\mathcal{F}) \;=\; 2^{\aleph_0}.$$

En la sección sobre cardinalidad habíamos informado que se podía demostrar que $\mathbb{I}$, el conjunto de todos los números irracionales, tiene la misma cardinalidad que la de $\mathbb{R}$. Aquí está una prueba de ese hecho.

Teorema 1.3.48. *Si A es un **subconjunto numerable** de $\mathbb{R}$, entonces*

$$\operatorname{card}(\mathbb{R} \setminus A) = 2^{\aleph_0}.$$

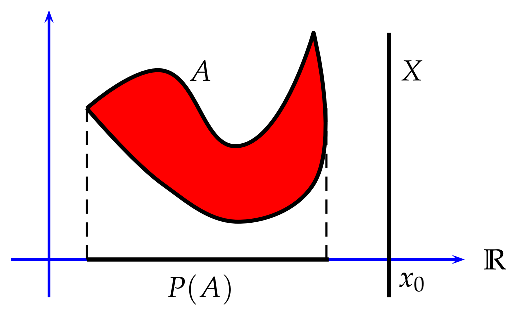
Prueba. Puesto que $\operatorname{card}(\mathbb{R} \times \mathbb{R}) = \operatorname{card}(\mathbb{R})$, podemos asumir que $A \subseteq \mathbb{R} \times \mathbb{R}$. En este caso, vamos a demostrar que

$$\operatorname{card}\big((\mathbb{R} \times \mathbb{R}) \setminus A\big) = 2^{\aleph_0}.$$

Sea $P : A \to \mathbb{R}$ la proyección sobre la primera coordenada, esto es, $P(x, y) = x$ para todo $(x, y) \in A$. Puesto que P es inyectiva, resulta que

$$\aleph_0 = \operatorname{card}(A) < \operatorname{card}(\mathbb{R}) = 2^{\aleph_0}$$

y, por lo tanto, $\mathbb{R} \setminus P(A) \neq \varnothing$.



Sea $x_0 \in \mathbb{R} \setminus P(A)$. Claramente el conjunto $X = \{x_0\} \times \mathbb{R}$ es disjunto al conjunto A y, en consecuencia, $X \subseteq (\mathbb{R} \times \mathbb{R}) \setminus A$. Puesto que $\operatorname{card}(X) = \operatorname{card}(\mathbb{R}) = 2^{\aleph_0}$ se tiene que

$$\operatorname{card}(\mathbb{R} \times \mathbb{R} \setminus A) \geq 2^{\aleph_0}.$$

Esto termina la prueba. ■

Una consecuencia inmediata del resultado anterior es que:

$$\operatorname{card}(\mathbb{I}) = \operatorname{card}(\mathbb{R} \setminus \mathbb{Q}) = 2^{\aleph_0} = \operatorname{card}(\mathbb{R}).$$

1.3.9. La Hipótesis del Continuo

Con el descubrimiento de la existencia de infinitos niveles de infinitos: $\aleph_0, \aleph_1, \aleph_2, \ldots$ y que $\aleph_0 < \mathfrak{c}$, surgen dos preguntas fundamentales:

(1) *¿Es $\mathfrak{c}$ el infinito no-numerable más pequeño*?

(2) *¿Son el Método de la Diagonal de Cantor y los números ordinales los únicos modos de probar la existencia de conjuntos no-numerables*?

La respuesta a la primera pregunta persiste, hasta el sol de hoy, aun sin poder ni demostrarse, ni refutarse, conociéndose dicho problema con el nombre de "la Hipótesis del Continuo". Sin embrago, como ya vimos anteriormente, el Método de la Diagonal y los números ordinales, no son los únicos modos de probar la existencia de conjuntos no-numerables: el Teorema de Cantor, la medida exterior de Lebesgue combinado con el Axioma de Elección, constituyen, entre otros, medios para obtenerlos.

La primera formulación de Cantor de la Hipótesis del Continuo, conocida también como *Hipótesis del Continuo Débil*, ocurrió en el año 1878 y es la siguiente: estando en conocimiento de que $\aleph_0 < 2^{\aleph_0}$,

$$\text{¿existirá algún } \textbf{conjunto infinito } A \subseteq \mathbb{R} \text{ tal que } \aleph_0 < \operatorname{card}(A) < \mathfrak{c}?$$

La Hipótesis del Continuo Débil afirma que un tal conjunto A no existe, en otras palabras: *cualquier subconjunto infinito de números reales, o es numerable, o posee la cardinalidad del continuo.*

Hipótesis del Continuo Débil (CH). *Si A es un **subconjunto infinito** de $\mathbb{R}$, entonces,*

$$\operatorname{card}(A) = \aleph_0 \qquad \text{o} \qquad \operatorname{card}(A) = \mathfrak{c}.$$

Algunos años más tarde, después de desarrollar su teoría de los números ordinales transfinitos, Cantor presenta esta otra forma de la Hipótesis del Continuo:

El conjunto $\mathbb{R}$ tiene la cardinalidad de todos los números ordinales numerables.

Finalmente, en 1890, después de desarrollar la noción de exponenciación entre números cardinales y su simbolismo: los alephs, y de demostrar que que $\aleph_0 < \aleph_1 \leq 2^{\aleph_0} = \mathfrak{c}$, él presenta su tercera y última forma de la Hipótesis del Continuo:

Hipótesis del Continuo (CH). $2^{\aleph_0} = \aleph_1$.

Es decir, en la sucesión transfinita de cardinales no-numerables $\aleph_1, \aleph_2 \ldots$, la Hipótesis del Continuo afirma que $2^{\aleph_0} = \aleph_1$, es decir, $\mathfrak{c} = 2^{\aleph_0}$ es el *primer cardinal no numerable*.

$$\begin{array}{cccccccccccc}
0 & 1 & \cdots & \omega & \omega+1 & \cdots & \omega+\omega & \cdots & \omega_1 & \cdots & \omega_2 & \cdots \\
 & & & \| & & & & \textbf{CH} & \| & & \| & \\
 & & & \aleph_0 & & & & & 2^{\aleph_0} = \aleph_1 & & \aleph_2 &
\end{array}$$

Trivialmente, la Hipótesis del Continuo implica la Hipótesis del Continuo Débil. La diferencia entre ambas formas es relevante al Axioma de Elección: *la Hipótesis del Continuo implica que el conjunto de los números reales puede ser bien-ordenado, mientras que la Hipótesis del Continuo Débil no la implica*. De hecho, la Hipótesis del Continuo es equivalente a la Hipótesis del Continuo Débil más la aceptación de la existencia de un buen-orden sobre $\mathbb{R}$. Para efectos prácticos usaremos la expresión "Hipótesis del Continuo" para referirnos a cualquiera de sus versiones.

Existe una versión generalizada de la Hipótesis del Continuo expresada del modo siguiente:

Hipótesis del Continuo Generalizada (GCH). $2^{\aleph_\alpha} = \aleph_{\alpha+1}$ para cualquier ordinal α.

La forma en que fue presentada la Hipótesis del Continuo Generalizada, supone ya al Axioma de Elección, pero puede ser presentada sin hacer uso de ese axioma: *Para cada cardinal infinito $\mathfrak{m}$, se tiene que*

$$\text{si } \mathfrak{m} < \mathfrak{a} < 2^{\mathfrak{m}} \quad \text{entonces} \quad \mathfrak{m} = \mathfrak{a} \quad \text{o} \quad \mathfrak{a} = 2^{\mathfrak{m}}.$$

Una virtud de la Hipótesis del Continuo Generalizada, la cual hace la hace interesante, es que ella da una solución completa al problema de computar $\mathfrak{a}^{\mathfrak{m}}$ para cardinales infinitos simplificando notablemente la exponenciación de los números cardinales. Por ejemplo, bajo la **GCH**, si $\mathfrak{a} \leq \mathfrak{m}$, entonces $\mathfrak{a}^{\mathfrak{m}} = \mathfrak{m}^{+}$.

1.3.10. Conjunto de Bernstein

El Principio del Buen-Orden en compañía de la Hipótesis del Continuo permiten construir ciertos subconjuntos extraños en $\mathbb{R}$. Por ejemplo, existe un subconjunto de $\mathbb{R}$ que tiene la rara particularidad que tanto él, así como su complemento, intersectan a cualquier subconjunto cerrado no-numerable de $\mathbb{R}$. Tal conjunto se conoce con el nombre de **Monstruo de Bernstein** o, simplemente, **conjunto de Bernstein** y su demostración se apoya en la Hipótesis del Continuo e Inducción Transfinita.

Teorema 1.3.49 (Conjunto de Bernstein). *En* **ZFC + CH** *existe un subconjunto no vacío* $\mathbf{B}$ *de* $\mathbb{R}$ *tal que*

$$F \cap \mathbf{B} \neq \varnothing \qquad y \qquad F \cap (\mathbb{R} \setminus \mathbf{B}) \neq \varnothing$$

para cualquier ***subconjunto cerrado no-numerable*** F *de* $\mathbb{R}$.

Prueba. Considere la familia $\mathcal{F}_{\mathfrak{c}}$ formada por todos los subconjuntos cerrados no-numerables de $\mathbb{R}$. Ya hemos visto que $\operatorname{card}(\mathcal{F}_{\mathfrak{c}}) = 2^{\aleph_0}$ y como estamos asumiendo la Hipótesis del Continuo, resulta que $2^{\aleph_0} = \aleph_1$. Esto nos indica que podemos indexar a dicho conjunto con los números ordinales menores que ω_1, esto es, $\mathcal{F}_{\mathfrak{c}} = \{F_\alpha : \alpha < \omega_1\}$. Usaremos inducción transfinita para construir a $\mathbf{B}$. Para ello es necesario, invocando el Principio del Buen-Orden, asumir que $\mathbb{R}$ está bien-ordenado y, por consiguiente, que cada F_α también lo está. Sean p_1, q_1 los primeros dos elementos de F_1. Fijemos un ordinal arbitrario α con $1 < \alpha < \omega_1$ y suponga que p_β, q_β han sido construidos para todo $\beta < \alpha$. Puesto que $\bigcup_{\beta<\alpha}\{p_\beta, q_\beta\}$ es numerable y F_α es no-numerable, resulta que el conjunto $K_\alpha := F_\alpha \setminus \bigcup_{\beta<\alpha}\{p_\beta, q_\beta\}$ es no vacío y, por lo tanto, bien-ordenado. Seleccionemos ahora los dos primeros elementos de K_α, digamos p_α, q_α. El Principio de Inducción Transfinita nos permite construir el conjunto $\{p_\alpha, q_\alpha : 1 \leq \alpha < \omega_1\}$. Definamos ahora $\mathbf{B} = \{p_\alpha : \alpha < \omega_1\}$. Puesto que $p_\alpha \in \mathbf{B} \cap F_\alpha$ y $q_\alpha \in (\mathbb{R} \setminus \mathbf{B}) \cap F_\alpha$ para cualquier $\alpha < \omega_1$ y como cualquier conjunto cerrado no-numerable es un F_α para algún α, resulta que el conjunto $\mathbf{B}$ tiene las propiedades deseadas. ∎

Observe que si $\mathbf{B}$ es conjunto de Bernstein también lo es $\mathbb{R} \setminus \mathbf{B}$. Más aun,

Corolario 1.3.50. *Si* $\mathbf{B}$ *es* ***conjunto de Bernstein****, entonces* ***cualquier conjunto cerrado*** $F \subseteq \mathbf{B}$ *es, necesariamente,* ***numerable****.*

Prueba. Si $F \subseteq \mathbf{B}$ fuese no-numerable, entonces él debería contener puntos de $\mathbb{R} \setminus \mathbf{B}$ lo que resultaría imposible. Por esto, F es numerable. ∎

En particular, $\mathbf{B}$ *no contiene a ningún conjunto perfecto* pues tales conjuntos son cerrados y no-numerables (véase el Ejemplo **TBC** (4) en la página 139).

Nota Adicional 1.3.6 Similar al Axioma de Elección, la Hipótesis del Continuo es otro axioma muy popular, importante y de gran impacto pues varios resultados fundamentales en Matemáticas sólo pueden ser demostrados si se acepta dicho axioma. Tal vez su popularidad se deba a estos dos factores: el primero es la sencillez de su formulación, y el segundo por ser el primero de los 23 problemas propuestos por David Hilbert en el famoso congreso de matemáticas celebrado en París en el año 1900. Aun que esto último le confirió un privilegio excepcional, dicha hipótesis también ha resultado ser muy polémica: *¿por qué* $\mathfrak{c}$ *debe ser el primer cardinal no numerable*? La Hipótesis del Continuo no puede ser demostrada que es verdadera ni tampoco que es falsa en el sistema **ZFC**. Por consiguiente, cualquier matemático tiene total y absoluta libertad de tomarla o dejarla, es decir, aceptarla como un axioma y añadirselo al sistema **ZFC** para formar el nuevo sistema **ZFC** + **CH**, o rechazarla añadiendo su negación a **ZFC** para obtener **ZFC** + $\neg$**CH**. Lo importante de esos dos sistemas es que en 1938 Kurt Gödel demostró, asumiendo la consistencia de **ZFC**, la consistencia del sistema **ZFC** + **CH** y años más tarde Paul Cohen, en 1963, demostró la consistencia del sistema **ZFC** + $\neg$**CH** siempre que **ZFC** sea consistente. Sin embargo, ambos creían que dicha hipótesis era falsa. Si bien es cierto que los trabajos de Gödel y Cohen permiten cierta tranquilidad en cuanto a que una contradicción en uno de los dos sistemas produce de inmediato una contradicción en el otro, resulta que en definitiva la Hipótesis del Continuo aun sigue siendo una materia pendiente.

¿Cuál es la relación entre **GCH** *y* **AC**? Se puede demostrar que **AC** *no implica* la **CH** ni la **GCH**; sin embargo, vale la siguiente implicación: **GCH** $\Rightarrow$ **AC** la cual fue demostrada por primera vez por W. Sierpinski, aunque algunos años antes A. Tarski había anunciado la misma implicación pero sin dar ningún indicio de su prueba (véase, por ejemplo, [62] para una demostración de este hecho).

En 1883 Cantor introdujo la noción de conjunto bien-ordenado. Maravillado por las propiedades de tales conjuntos y, especialmente, por los "Alephs", Cantor los concibe como una "*ley del pensamiento*", formulando su famoso Principio del Buen-Orden: *Todo conjunto puede ser bien-ordenado*. Después de un tiempo y mucha reflexión Cantor cambia de opinión y cree que dicho principio requiere de una demostración. Un pedacito de historia después del Congreso en Paris en 1900 es la siguiente: en el Tercer Congreso de Matemáticas celebrado en Heidelberg en el año 1904, el matemático Gyula König quien también firmaba como Julius König, ofreció una prueba, en presencia de Hilbert y el propio Cantor, que la Hipótesis del Continuo es falsa, es decir, que $\mathfrak{c}$ no es un $\aleph$. En consecuencia, añadió König, la afirmación de Cantor de que cualquier conjunto puede ser bien-ordenado también es falsa ya que $\mathbb{R}$ no puede ser bien-ordenado. Cantor encontró devastador el argumento de König pues utilizaba sólo herramientas cantorianas y no le fue posible encontrar ningún error en esa demostración. Sin embargo, al día siguiente, según cuenta G. Kowaleswski, E. Zermelo encontró un error en la demostración, aunque Ivor Grattan-Guinness y W. Purkert afirman que fue Hausdorff, un tiempo después, quien encontró las fallas en la prueba ofrecida por König. En cualquier caso, la conclusión es inequívoca: existía un error en la demostración de König. Un mes después del Congreso de Heidelberg, Zermelo, quien también creía en la afirmación de Cantor, escribió una carta a Hilbert, editor de la revista Mathematische Annalen para ese momento, en la cual demostraba el Teorema del Buen-Orden. Zermelo basó su demostración sobre un nuevo postulado que, de inmediato, causó gran furor: el enigmático Axioma de Elección o Axioma de Zermelo como también se le conoce. La demostración de Zermelo generó una febril controversia en los siguientes cuatro años involucrando a matemáticos de Alemania, Inglaterra, Francia, Holanda, Hungría, Italia y los Estados Unidos de América. Hilbert, sin embargo, le escribe a Borel pidiéndole que preparara una respuesta a la prueba de Zermelo. La respuesta de Borel, la cual rechazaba la demostración de Zermelo, fue terminada el 1 de Diciembre de 1904. En ella Borel afirma lo siguiente: *Lo que Zermelo realmente muestra es que las siguientes dos proposiciones son equivalentes*:

(A) El Teorema del Buen-Orden

(B) El Axioma de Elección

e insiste que Zermelo no ha demostrado que la equivalencia de (A) *y* (B) *provee una solución general al problema* (A).

1.4. Problemas

(1) Sean A, B, C, D conjuntos arbitrarios. Pruebe que

(a_1) $A \triangle B = B \triangle A$

(b_1) $A \triangle \varnothing = A$

(c_1) $A \triangle A = \varnothing$.

(d_1) $A \triangle B = A^c \triangle B^c$.

(e_1) $(A \triangle B) \triangle C = A \triangle (B \triangle C)$.

(f_1) $A \cap (B \triangle C) = (A \cap B) \triangle (A \cap C)$.

(g_1) $(A \cup B \cup C) \setminus (A \cap B \cap C) = (A \triangle B) \cup (B \triangle C)$.

(h_1) $(A \cup B) \triangle (C \cup D) \subseteq (A \triangle C) \cup (B \triangle D)$.

(2) Sea $\mathscr{A}$ una familia arbitraria de conjuntos. Pruebe que

$$\left(\bigcup_{A \in \mathscr{A}} A\right)^c = \bigcap_{A \in \mathscr{A}} A^c \qquad \text{y} \qquad \left(\bigcap_{A \in \mathscr{A}} A\right)^c = \bigcup_{A \in \mathscr{A}} A^c$$

(3) Sean $(A_\alpha)_{\alpha \in I}$ y $(B_\beta)_{\beta \in J}$ familias de conjuntos. Demuestre que

$$\left(\bigcup_{\alpha \in I} A_\alpha\right) \times \left(\bigcup_{\beta \in J} B_\beta\right) = \bigcup_{(\alpha,\beta) \in I \times J} A_\alpha \times B_\beta,$$

$$\left(\bigcap_{\alpha \in I} A_\alpha\right) \times \left(\bigcap_{\beta \in J} B_\beta\right) = \bigcap_{(\alpha,\beta) \in I \times J} A_\alpha \times B_\beta.$$

(4) Sean $f : X \to Y$ una función y $\mathcal{A}$ una familia arbitraria de subconjuntos de X. Pruebe que

$$f\left(\bigcup_{A \in \mathcal{A}} A\right) = \bigcup_{A \in \mathcal{A}} f(A) \qquad \text{y} \qquad f\left(\bigcap_{A \in \mathcal{A}} A\right) \subseteq \bigcap_{A \in \mathcal{A}} f(A).$$

Demuestre que si f es **inyectiva**, entonces

$$f\left(\bigcap_{A \in \mathcal{A}} A\right) = \bigcap_{A \in \mathcal{A}} f(A).$$

(5) Sean $f : X \to Y$ una función. Pruebe que para cualquier conjunto $A \subseteq X$, $f^{-1}(f(A)) = A$ si, y sólo f es inyectiva. Similarmente, f es sobreyectiva si, y sólo si, $f(f^{-1}(B)) = B$ para todo $B \subseteq Y$.

(6) Sea A un conjunto numerable. Pruebe que la aplicación $\varphi : \mathcal{P}_{\text{fin}}(A) \to \mathbb{N}$ definida por

$$\varphi(\{a_{j_1}, \ldots, a_{j_n}\}) = p_1^{j_1} p_2^{j_2} \cdots p_n^{j_n}$$

es biyectiva, donde $\{a_{j_1}, \ldots, a_{j_n}\} \in \mathcal{P}_{\text{fin}}(A)$ y los números $p_1, \ldots, p_n$ son los primeros n números primos.

(7) Muestre que la aplicación $\varphi : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ dada por

$$\varphi(m, n) = 2^m 3^n$$

es inyectiva. Concluya, usando el Teorema 1.2.6, que $\mathbb{N} \times \mathbb{N}$ es numerable. La misma conclusión se obtiene si definimos $\varphi_1(m, n) = 2^{m-1}(2n - 1)$.

(8) Sea $A \subseteq \mathbb{R}^n$ un conjunto numerable. Pruebe que cada uno de los siguientes conjuntos

$$x + A = \{x + a : a \in A\} \qquad \text{y} \qquad \lambda \cdot A = \{\lambda \cdot a : a \in A\}$$

es numerable para cada $x \in \mathbb{R}^n$ y cada $\lambda \in \mathbb{R}$.

(9) Sea $\mathcal{I}$ una familia disjunta de intervalos no-degenerados, es decir, ningún punto $x = [x, x]$ pertenece a la familia. Demuestre que $\mathcal{I}$ es a lo más numerable.

(10) Pruebe que $[0, 1] \cup S$ es no-numerable para cualquier subconjunto $S \subseteq \mathbb{R}$.

(11) Pruebe que el conjunto

$$\mathcal{E} = \Big\{m^n : m \in \mathbb{N},\ n \in \mathbb{Z}\Big\}$$

es numerable.

(12) Sea X un subconjunto numerable de $\mathbb{R}$. ¿Existe algún $a \in \mathbb{R}$ tal que

$$(a + X) \cap X = \varnothing?$$

(13) Sea X un conjunto infinito. Si A es cualquier subconjunto finito de X, pruebe que X y $X \setminus A$ son equipotentes y, en consecuencia, $\operatorname{card}(X) = \operatorname{card}(X \setminus A)$.

(14) Demuestre que si X es un conjunto no-numerable y A es un conjunto a lo más numerable, entonces $\operatorname{card}(X) = \operatorname{card}(X \setminus A)$.

(15) Pruebe que el conjunto $\mathcal{P}\mathrm{ol}(\mathbb{R})$ de todos los polinomios reales con coeficientes enteros es numerable.

(16) Pruebe que $\prod_{n=1}^{\infty}\{0, 1\} = 2^{\aleph_0}$ es no-numerable.

(17) Demuestre que, si para cada $n \in \mathbb{N}$, $\operatorname{card}(A_n) = \mathfrak{c}$, entonces $\operatorname{card}\big(\prod_{n=1}^{\infty} A_n\big) = \mathfrak{c}$.

(18) Sea $A \subseteq \mathbb{R}$ un conjunto no-numerable. Pruebe que existe un $x \in \mathbb{R}$ tal que $A \cap (x, +\infty)$ y $A \cap (-\infty, x)$ son ambos no-numerables.

(19) *Dos conjuntos A y B se dice que son **casi-disjuntos** si $A \cap B$ es finito*. Pruebe que si X es un conjunto numerable, entonces la familia $\mathcal{C}_{\text{disj}}(X)$, *formada por todos los subconjuntos de X que son dos a dos casi-disjuntos*, es no-numerable.

(20) **El Hotel de Hilbert**. *Ya hemos podido constatar que con los conjuntos infinitos se está permito hacer ciertas "operaciones" que no son posibles hacerlas con ningún conjunto finito*. Veamos este otro ejemplo. En una cierta ciudad nombrada $\aleph_\alpha$, existe un famoso hotel llamado el **Hotel de Hilbert** que dispone de infinitas habitaciones numeradas en el orden habitual: $1, 2, 3, \ldots$ Cierta noche llega a la ciudad un turista y se dirige raudo al Hotel de Hilbert en busca de una habitación. Al llegar solicita al recepcionista (que no era matemático) una habitación y éste inmediatamente le responde: - *Lo sentimos mucho caballero, el hotel está "totalmente lleno"*. Cantor, quien para ese momento era el gerente del hotel, al oír al recepcionista le dice al turista: *Disculpe ud. al caballero, de inmediato lo ubicaremos en una habitación*. El recepcionista, sorprendido e incrédulo por la respuesta de su jefe, se pregunta a asimismo: ¿y cómo diablos le va a dar una habitación a ese caballero si todas las habitaciones están ocupadas? A lo mejor echa a la calle a alguien. Cantor toma el micrófono de la recepción y gentilmente le informa a sus huéspedes, pidiendo las disculpas de rigor, que deben abandonar su actual habitación y alojarse en la siguiente. De este modo, el huésped que ocupaba la habitación número 1 se muda a la 2, el que ocupaba la habitación número 2 se mueve a la número 3 y así sucesivamente. Completada la orden queda la habitación número 1 desocupada y Cantor, esbozando una picara sonrisa, le hace entrega de dicha habitación al turista quien muy complacido por la forma en que Cantor había resuelto el problema le extiende la mano en agradecimiento. No habían transcurrido 5 minutos cuando por el lobby del hotel aparecen 30 nuevos turistas solicitando cada uno de ellos una habitación y el recepcionista vuelve a responder: *el hotel está lleno, pero permítanme llamar a mi jefe que él es matemático y a lo mejor sabe cómo darle habitaciones a todos ustedes*. Así fue, llegó Cantor y de nuevo resolvió el problema sin "echar a nadie a la calle" y dándole habitación a cada uno de los 30 turistas. Pues bien, el problema no para allí. Una hora después aparecieron infinitos turistas (por supuesto, en una cantidad numerable) solicitando habitaciones. De nuevo hace su aparición Cantor y le da habitaciones a todos los recién llegados sin dejar sin habitación a los anteriores inquilinos. ¿Sabe ud. cómo lo hizo Cantor?

(21) Demuestre que si α y β son cardinales, entonces:

(*a*) $\alpha \leq \alpha^\beta$ si $\beta > 0$.

(*b*) $\alpha \leq \beta^\alpha$ si $\beta > 1$.

(*c*) Si $\alpha_1 \leq \alpha_2$ y $\beta_1 \leq \beta_2$, entonces $\alpha_1^{\beta_1} \leq \alpha_2^{\beta_2}$.

(22) Demuestre que en **ZF** el Axioma de Elección, el Lema de Zorn y el Principio del Buen-Orden son equivalentes.

(23) Use el Principio del Buen-Orden para dar una prueba del Teorema Fundamental de la Aritmética.

(24) Demuestre que el Axioma de Elección no implica la Hipótesis del Continuo.

CAPÍTULO 2

Los Números Reales

2.1. Algunas Propiedades de los Números Reales

Confiamos en que el lector posee cierta experiencia con el sistema de los números reales $\mathbb{R}$ por lo que no le dedicaremos tiempo a su construcción. En particular, asumiremos familiaridad con el conjunto de todos los números enteros,

$$\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\},$$

con el conjunto de todos los números naturales, también llamado, enteros positivos,

$$\mathbb{N} = \{1, 2, 3, \dots\},$$

con el conjunto de todos los números racionales

$$\mathbb{Q} = \{p/q \,:\, p \in \mathbb{Z},\, q \in \mathbb{N}\},$$

y su complemento, $\mathbb{I} = \mathbb{R} \setminus \mathbb{Q}$, el conjunto de todos los números irracionales. En lo que sigue escribiremos $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. También asumiremos el conocimiento de las propiedades del orden en $\mathbb{R}$.

2.1.1. Principio de Arquímedes

Recordemos que, dado cualquier número real x, llamamos la **parte entera** de x al único entero $[x]$ definido por:

$$[x] = \text{máx}\,\{\mathbb{Z} \cap (-\infty, x]\} = \text{máx}\,\{z \in \mathbb{Z} \,:\, z \leq x\},$$

es decir, $[x]$ es *el mayor entero menor o igual a x*. Observe que $[x]$ queda completamente determinado por las dos propiedades siguientes:

$$[x] \in \mathbb{Z} \qquad \text{y} \qquad x - 1 < [x] \leq x, \tag{1}$$

o de modo equivalente,

$$[x] \in \mathbb{Z} \qquad \text{y} \qquad [x] \leq x < [x] + 1.$$

Fijemos un $x \in \mathbb{R}$ y considere cualquier $n \in \mathbb{N}$. Se sigue de (1) que $nx - 1 < [nx] \leq nx$ y así,

$$x - \frac{1}{n} < \frac{[nx]}{n} \leq x.$$

De esto último, tomando límite cuando $n \to \infty$, se tiene que

$$x = \lim_{n\to\infty} \frac{[nx]}{n}.$$

Puesto que x fue elegido de modo arbitrario y la sucesión $([nx]/n)_{n=1}^{\infty}$ es una sucesión de números racionales convergiendo a x, resulta que $\mathbb{Q}$ es *denso* en $\mathbb{R}$.

Teorema 2.1.1 (Principio de Arquímedes). *Si a y b son* ***números reales positivos****, entonces existe un* ***entero positivo*** *n tal que $na > b$.*

Prueba. Puesto que $b/a > 0$, resulta que $1 + b/a > 1$ y, por consiguiente,

$$n = \left[1 + \frac{b}{a}\right]$$

es un entero ≥ 1. Se sigue de la definición de parte entera que

$$n > \left(1 + \frac{b}{a}\right) - 1 = \frac{b}{a}$$

y puesto que $a > 0$, resulta entonces que $na > b$ y termina la prueba. ■

Cuando a y b son **enteros positivos**, el Principio del Buen-Orden también puede ser usado para dar otra demostración del Principio de Arquímedes. En efecto, suponga que la conclusión es falsa. Entonces $na \leq b$ para todo $n \in \mathbb{N}$. Considere el conjunto

$$S = \{b - na : n \in \mathbb{N}\}.$$

Resulta que S consiste sólo de enteros positivos y, en consecuencia, por el Principio del Buen-Orden él posee un primer elemento el cual denotaremos por k. De esto se sigue que existe un único $m \in \mathbb{N}$ tal que

$$k = b - ma \qquad \text{y} \qquad b - ma < b - na \quad \text{para todo } n \neq m.$$

Puesto que $b - (m+1)a \in S$, se tiene que

$$b - (m+1)a = (b - ma) - a < b - ma$$

lo cual contradice la elección de $b - ma$. Esta contradicción establece el resultado.

Corolario 2.1.2. *Si $\varepsilon > 0$, entonces existe un* ***entero positivo*** *n tal que $1/n < \varepsilon$.*

Prueba. Basta tomar $a = \varepsilon$ y $b = 1$ en el teorema anterior. ■

Teorema 2.1.3. *Sean $x, y \in \mathbb{R}$. Si para cada $\varepsilon > 0$ ocurre que $x < y + \varepsilon$, entonces $x \leq y$.*

Prueba. Si suponemos que $x > y$, entonces el número $\varepsilon = x - y > 0$ debería satisfacer la desigualdad $x < y + \varepsilon = y + (x - y) = x$, lo que resulta ser un disparate. Por esto, $x \leq y$. ■

2.1.2. Conjuntos Acotados

Sean $a, b \in \mathbb{R}$ con $a < b$. Recordemos que un **intervalo** con extremos a y b es cualquier conjunto de la forma

$$\begin{aligned}(a,b) &= \{x \in \mathbb{R} : a < x < b\},\\ [a,b] &= \{x \in \mathbb{R} : a \leq x \leq b\},\\ (a,b] &= \{x \in \mathbb{R} : a < x \leq b\},\\ [a,b) &= \{x \in \mathbb{R} : a \leq x < b\}.\end{aligned}$$

A tales intervalos los llamaremos intervalos de **longitud finita** o, simplemente, **intervalos finitos**. Algunas veces también los denominaremos **intervalos acotados**. Por otro lado, los intervalos **no acotados** o de **longitud infinita** se definen como

$$\begin{aligned}(a,+\infty) &= \{x \in \mathbb{R} : a < x\},\\ [a,+\infty) &= \{x \in \mathbb{R} : a \leq x\},\\ (-\infty,b] &= \{x \in \mathbb{R} : x \leq b\},\\ (-\infty,b) &= \{x \in \mathbb{R} : x < b\}.\end{aligned}$$

También se acostumbra a escribir $\mathbb{R} = (-\infty, +\infty)$. Convenimos en definir $(a,a) = (a,a] = [a,a) = \varnothing$ para cualquier $a \in \mathbb{R}$. Un intervalo finito y cerrado se llama **degenerado** si es de la forma $[a,a] = \{a\}$ para algún $a \in \mathbb{R}$. Un intervalo finito I con extremos a y b es **no-degenerado** si $a < b$ o $b < a$. La **longitud** de cualquier intervalo acotado I con extremos a y b, se define como

$$\ell(I) = b - a.$$

Si I no es acotado pondremos $\ell(I) = +\infty$. Suponga que $G = I_1 \cup \cdots \cup I_n$ es la unión de n intervalos acotados disjuntos dos a dos, entonces convenimos en definir la ***longitud de*** G como:

$$\ell(G) = \ell(I_1) + \cdots + \ell(I_n).$$

Si algún $\ell(I_i) = +\infty$, entonces escribiremos $\ell(G) = +\infty$.

Definición 2.1.4. *Un subconjunto A de $\mathbb{R}$ se dice que está **acotado superiormente** si existe una constante M tal que $A \subseteq (-\infty, M]$, es decir, $a \leq M$ para todo $a \in A$.*

Cualquier número real M para el cual $a \leq M$ para todo $a \in A$, se llama una **cota superior** de A. Observe que el número M puede o no estar en A. Similarmente, si existe un número N tal que $N \leq a$ para todo $a \in A$, entonces diremos que A está **acotado inferiormente**. A un tal N se le llama una **cota inferior** de A. Si el conjunto A está acotado tanto inferiormente así como superiormente, entonces diremos que A está **acotado**. En este caso, existen $m, M \in \mathbb{R}$ tal que $A \subseteq [m, M]$. Si A no está acotado superiormente (respectivamente, no está acotado inferiormente), escribiremos $\sup A = +\infty$ (respectivamente, $\inf A = -\infty$).

Sea X un conjunto no vacío y sea $f : X \to \mathbb{R}$ una función. Diremos que f está **acotada** sobre X si el conjunto $f(X)$ está acotado. Observe que si f está acotada sobre X, entonces, por lo dicho anteriormente, existen $m, M \in \mathbb{R}$ tales que $f(X) \subseteq [m, M]$, esto es,

$$m \leq f(x) \leq M \quad \text{para todo } x \in X.$$

Equivalentemente, decir que f está acotada sobre X significa que existe una constante $K > 0$ tal que $|f(x)| \leq K$ para todo $x \in X$. En lo que sigue, el símbolo $\mathcal{B}_\infty(X)$ lo usaremos para denotar *el conjunto de todas las funciones* $f : X \to \mathbb{R}$ *que están* ***acotadas*** *sobre* X; esto es:

$$\mathcal{B}_\infty(X) \;=\; \big\{f \in \mathbb{R}^X \;:\; f \text{ está } \textbf{acotada} \text{ sobre } X\big\}.$$

Es fácil ver que si $f, g \in \mathcal{B}_\infty(X)$ y $a \in \mathbb{R}$, entonces $f + g$ y $a \cdot f$ permanecen en $\mathcal{B}_\infty(X)$; es decir, $\mathcal{B}_\infty(X)$ es un espacio vectorial sobre $\mathbb{R}$.

Definición 2.1.5. *Sea A un subconjunto no vacío de $\mathbb{R}$. Un número real a_0 se dice que es el* ***supremo*** *de A, que escribiremos $a_0 = \sup A$, si las siguientes dos condiciones se cumplen:*

$(\sup_1)$ *$x \leq a_0$ para todo $x \in A$, y*

$(\sup_2)$ *si $M \in \mathbb{R}$ es tal que $x \leq M$ para todo $x \in A$, entonces $a_0 \leq M$.*

Observe que si $a_0 = \sup A$, entonces la condición $(\sup_1)$ de la definición anterior dice que A *está acotado superiormente*, mientras que la condición $(\sup_2)$ establece que a_0 es la *menor de todas las cotas superiores* de A. Si el número $a_0 = \sup A$ pertenece al conjunto A, entonces diremos que a_0 es el **máximo**, el **mayor** o el **más grande** elemento de A.

La notación

$$\sup_{x \in A} x \;=\; \sup\big\{x \;:\; x \in A\big\} \;=\; \sup A$$

se usará indistintamente en estas notas. Por ejemplo, si $A_n = \{x_n, x_{n+1}, \ldots\}$, escribiremos

$$\sup_{j \geq n} x_j \;=\; \sup A_n.$$

Una manera sencilla de identificar el supremo de un conjunto (cuando éste existe) es por medio del siguiente resultado.

Teorema 2.1.6. *Sea A un subconjunto de $\mathbb{R}$* ***no vacío*** *y* ***acotado superiormente***. *Entonces $a_0 = \sup A$ si, y sólo si,*

(a) $x \leq a_0$ para todo $x \in A$, y

(b) dado $\varepsilon > 0$, existe $x \in A$ tal que $a_0 - \varepsilon < x$.

Prueba. Suponga que $a_0 = \sup A$ es finito. Entonces, por la definición de supremo, se sigue que (a) se cumple. Para demostrar (b), asuma que la conclusión es falsa, es decir, que $a_0 - \varepsilon \geq x$ para todo $x \in A$. Esto último significa que $a_0 - \varepsilon$ es una cota superior y se sigue, usando $(\sup_2)$, que $a_0 < a_0 - \varepsilon$, lo cual es imposible. Por esto, (b) también se cumple.

Recíprocamente, suponga que (a) y (b) se satisfacen y sea a_1 una cota superior de A, esto es, $x \leq a_1$ para todo $x \in A$. Vamos a demostrar que $a_0 \leq a_1$. En efecto, dado $\varepsilon > 0$, usemos (b) para obtener un $x \in A$ tal que $a_0 - \varepsilon < x$. Ahora bien, como $x \leq a_1$, resulta que $a_0 < a_1 + \varepsilon$ y, en consecuencia, por el Teorema 2.1.3 tenemos que $a_0 \leq a_1$. ■

El **ínfimo** de un subconjunto A no vacío y acotado inferiormente de $\mathbb{R}$, $\inf A$, se define de manera similar a la del supremo, es decir, $b_0 = \inf A \in \mathbb{R}$ significa que se cumplen las siguientes condiciones:

(Inf$_1$) $b_0 \leq x$ para todo $x \in A$, es decir, *A está acotado inferiormente*, y

(Inf$_2$) si $b \in \mathbb{R}$ es tal que $b \leq x$ para todo $x \in A$, entonces $b \leq b_0$. Esto último significa que b_0 es la *mayor de todas las cotas inferiores* de A.

Si $b_0 = \text{ínf}\, A$ pertenece al conjunto A, entonces diremos que b_0 es el **mínimo**, el **menor** o el **más pequeño** elemento de A. Similar al caso del supremo también escribiremos, en algunos casos, $\text{ínf}_{a \in A}\, a$ en lugar de $\text{ínf}\, A$.

Teorema 2.1.7. *Sea A un subconjunto de $\mathbb{R}$* ***no vacío*** *y* ***acotado inferiormente****. Entonces $b_0 = \text{ínf}\, A$ si, y sólo si,*

(a) $b_0 \leq x$ para todo $x \in A$, y

(b) dado $\varepsilon > 0$, existe $x \in A$ tal que $x < b_0 + \varepsilon$.

Prueba. Es similar a la del Teorema 2.1.6 y, por consiguiente, se omite. ■

De modo similar al supremo de un conjunto, *si $b_0 = \text{ínf}\, A$, entonces* ***existe una sucesión*** *$(x_n)_{n=1}^{\infty}$ en A tal que*

$$b_0 \;=\; \lim_{n\to\infty} x_n. \tag{3}$$

Se sigue de la definición que, *si A y B son subconjuntos no vacíos y acotados de $\mathbb{R}$, entonces*

$$A \subseteq B \quad \Rightarrow \quad -\infty < \text{ínf}\, B \leq \text{ínf}\, A \leq \sup A \leq \sup B < +\infty.$$

Definición 2.1.8. *Sean A y B subconjuntos no vacíos de $\mathbb{R}$. Diremos que ellos que están* ***ordenadamente separados*** *si*

$$a \leq b \quad \text{para todo } a \in A \text{ y todo } b \in B.$$

Por ejemplo, si $\alpha \leq \beta$, entonces $A = (-\infty, \alpha)$ y $B = (\beta, +\infty)$ están ordenadamente separados. Un resultado que es fundamental, entre otras cosas, para definir la integral de Riemann por medio de las sumas de Darboux es el siguiente.

Teorema 2.1.9. *Sean A y B subconjuntos* ***no vacíos*** *y* ***acotados*** *de $\mathbb{R}$* ***ordenadamente separados****. Entonces*

(a) $\sup A \leq \text{ínf}\, B$.

(b) $\sup A = \text{ínf}\, B$ si, y sólo si, para cada $\varepsilon > 0$, existen $a_\varepsilon \in A$ y $b_\varepsilon \in B$ tales que

$$b_\varepsilon - a_\varepsilon < \varepsilon.$$

Prueba. (a) Fijemos $a \in A$. Entonces, por la hipótesis, $a \leq b$ para todo $b \in B$ lo cual significa que a es una cota inferior para B y, así, $a \leq \text{ínf}\, B$ para cualquier $a \in A$. Esto, por supuesto, significa que $\text{ínf}\, B$ es una cota superior para A y, en consecuencia, $\sup A \leq \text{ínf}\, B$.

(b) Suponga que $\sup A = \text{ínf}\, B$ y sea $\varepsilon > 0$. Puesto que $\sup A - \varepsilon/2 < \sup A$, se sigue de las propiedades del supremo que existe un $a_\varepsilon \in A$ tal que $\sup A - \varepsilon/2 < a_\varepsilon \leq \sup A$. Similarmente, de las propiedades del ínfimo, podemos elegir un $b_\varepsilon \in B$ tal que $\text{ínf}\, B \leq b_\varepsilon < \text{ínf}\, B + \varepsilon/2$. Finalmente, si definimos $\delta = \text{ínf}\, B = \sup A$, tendremos que

$$\delta \geq a_\varepsilon > \delta - \frac{\varepsilon}{2} \qquad \text{y} \qquad \delta \leq b_\varepsilon < \delta + \frac{\varepsilon}{2},$$

de donde se sigue que $0 < b_\varepsilon - a_\varepsilon < (\delta + \varepsilon/2) - (\delta - \varepsilon/2) = \varepsilon$.

Sea $\varepsilon > 0$ y suponga que existen $a_\varepsilon \in A$ y $b_\varepsilon \in B$ tal que $b_\varepsilon - a_\varepsilon < \varepsilon$. Puesto que $b_\varepsilon \geq \text{ínf}\, B$ y $a_\varepsilon \leq \sup A$, resulta que

$$\text{ínf}\, B \;-\; \sup A \;\leq\; b_\varepsilon \;-\; a_\varepsilon \;<\; \varepsilon.$$

Como $\varepsilon > 0$ es arbitrario, resulta que $\text{ínf}\, B \leq \sup A$ y así, gracias a la primera parte, se concluye que $\sup A = \text{ínf}\, B$ y termina la prueba. ■

Definición 2.1.10. *Sean ahora A y B subconjuntos no vacíos de $\mathbb{R}$ y definamos la **suma** y el **producto algebraico** de A y B como:*

$$\begin{aligned} A \,+\, B &= \{a + b \,:\, a \in A,\; b \in B\} \\ AB &= \{ab \,:\, a \in A,\; b \in B\} \end{aligned}$$

Si $A = \{\lambda\}$ para algún $\lambda \in \mathbb{R}$, escribiremos λB en lugar de $\{\lambda\}B$ y $-B$ en lugar de $(-1)B$. Similarmente, escribiremos $A - B$ en lugar de $A + (-B)$. Observe que por la conmutatividad de la suma y el producto en $\mathbb{R}$, se tiene que

$$A + B \;=\; B + A \qquad \text{y} \qquad AB \;=\; BA.$$

Teorema 2.1.11. *Sea $A \subseteq \mathbb{R}$ **no vacío** y **acotado**. Entonces*

$$\sup(-A) \;=\; -\,\text{ínf}\, A \qquad y \qquad \text{ínf}(-A) \;=\; -\sup A.$$

Prueba. Suponga que $a_0 = \text{ínf}\, A$ y sea $\varepsilon > 0$ elegido arbitrariamente. Escojamos un $a \in A$ tal que $a < a_0 + \varepsilon$. Entonces $-a > -a_0 - \varepsilon$ y puesto que $-a \in -A$, resulta que $\sup(-A) \geq -a > -a_0 - \varepsilon$. Esto prueba que $-\sup(-A) < a_0 + \varepsilon$ para cada $\varepsilon > 0$, de donde se sigue, invocando al Teorema 2.1.3, que $-\sup(-A) \leq \text{ínf}\, A$, esto es,

$$\sup(-A) \;\geq\; -\,\text{ínf}\, A.$$

Por otro lado, $a \in -A$ implica que $-a \in A$ y, en consecuencia, $-a \geq \text{ínf}\, A$, es decir, $a \leq -\,\text{ínf}\, A$, de donde se deduce que

$$\sup(-A) \;\leq\; -\,\text{ínf}\, A.$$

La demostración de la segunda afirmación es inmediata si se intercambia A por $-A$ en la prueba de la igualdad anterior. ■

Diremos que $A \subseteq \mathbb{R}$ es un **conjunto no-negativo** si $a \geq 0$ para todo $a \in A$.

Teorema 2.1.12. *Si A y B subconjuntos **no vacíos** y **acotados superiormente** de $\mathbb{R}$, entonces $A + B$ y AB también son **acotados superiormente** y se cumple que:*

(1_1) $\sup(A + B) \;=\; \sup A + \sup B$.

(2_2) $\sup(AB) \;=\; (\sup A)(\sup B)$ *si A y B son no-negativos. En particular,*

$$\sup(\lambda A) \;=\; \lambda \sup A.$$

siempre que $\lambda \in \mathbb{R}^+$.

Prueba. Sean $a_0 = \sup A$ y $b_0 = \sup B$ y notemos que, por definición,

$$a \leq a_0 \quad \text{para todo } a \in A \qquad \text{y} \qquad b \leq b_0 \quad \text{para todo } b \in B. \tag{2.1.1}$$

(a) De las dos desigualdades anteriores se sigue que

$$a + b \leq a_0 + b_0$$

para todo $a \in A$ y todo $b \in B$ lo cual implica que $a_0 + b_0$ es una *cota superior* de $A + B$. Por consiguiente,

$$\sup(A+B) \leq a_0 + b_0 = \sup A + \sup B.$$

Fijemos $\varepsilon > 0$ y escojamos, haciendo uso del Teorema 2.1.6, un $a \in A$ y un $b \in B$ de modo tal que

$$a_0 - \frac{\varepsilon}{2} < a \qquad \text{y} \qquad b_0 - \frac{\varepsilon}{2} < b.$$

Entonces $a + b \in A + B$ y se cumple que

$$a_0 + b_0 - \varepsilon = \Big(a_0 - \frac{\varepsilon}{2}\Big) + \Big(b_0 - \frac{\varepsilon}{2}\Big) < a + b.$$

Un nuevo llamado al Teorema 2.1.6 nos revela que $\sup(A+B) = \sup A + \sup B$.

(b) Si $a_0 b_0 = 0$, entonces entonces uno de ellos es cero. Suponga que $a_0 = 0$. Como A es no-negativo resulta que $0 \leq a \leq a_0$ para todo $a \in A$ y, en consecuencia, $A = \{0\}$. Por esto, $AB = \{0\}$ y, así, $\sup(AB) = 0 = (\sup A)(\sup B)$. Consideremos ahora el caso cuando $a_0 b_0 > 0$. Por (2.1.1) tenemos que $ab \leq a_0 b_0$ para todo $a \in A$ y todo $b \in B$ de modo que $a_0 b_0$ es una *cota superior* de AB. Fijemos $\varepsilon > 0$ y escojamos un $\delta > 0$ que satisfaga $0 < \delta < \varepsilon/(a_0 + b_0)$. Invoquemos de nuevo al Teorema 2.1.6 para escoger un $a \in A$ y un $b \in B$ que cumplan

$$a_0 - \delta < a \qquad \text{y} \qquad b_0 - \delta < b.$$

Ahora,

$$ab > (a_0 - \delta)(b_0 - \delta) = a_0 b_0 - (a_0 + b_0)\delta + \delta^2 > a_0 b_0 - \varepsilon$$

de donde se sigue, gracias al Teorema 2.1.6, que $\sup(AB) = (\sup A)(\sup B)$. En el caso particular cuando $B = \{\lambda\}$, se tiene que $\sup(\lambda A) = \lambda \sup A$. ■

Observe que si A y B son subconjuntos de $\mathbb{R}$ acotados inferiormente, entonces (imitando la prueba del resultado anterior) se cumple que

$$\text{ínf}(A+B) = \text{ínf } A + \text{ínf } B$$

y, en consecuencia, si A y B son *subconjuntos acotados* de $\mathbb{R}$, entonces $A + B$ también es *acotado*.

Es importante destacar que si $A \subseteq \mathbb{R}$ es acotado, entonces por lo que acabamos de ver, el conjunto $A - A$ también lo es y $0 \in A - A$. Más aun, $A - A$ es un **conjunto simétrico** en el sentido de que

$$z \in A - A \quad \Leftrightarrow \quad -z \in A - A.$$

En efecto, $x - y \in A - A$ si, y sólo si, $-(x-y) = y - x \in A - A$ cualesquiera sean $x, y \in A$. De esto se sigue que $x - y \in A - A$ si, y sólo si, $|x - y| \in A - A$ y, por lo tanto,

$$\sup(A - A) = \sup\{|x-y| : x, y \in A\}.$$

Corolario 2.1.13. *Si $A \subseteq \mathbb{R}$ es* ***no vacío*** *y* ***acotado****, entonces*

$$\sup A \;-\; \inf A \;=\; \sup\big\{|x-y| \,:\, x, y \in A\big\}.$$

Prueba. Por la observación anterior, el Teorema 2.1.11 y el Teorema 2.1.12, tenemos que

$$\begin{aligned}\sup\big\{|x-y| \,:\, x, y \in A\big\} &= \sup(A-A)\\ &= \sup\big(A+(-A)\big)\\ &= \sup A \,+\, \sup(-A)\\ &= \sup A \,-\, \inf A.\end{aligned}$$

La prueba es completa. ■

El siguiente axioma, conocido como el **Axioma de Completitud** de $\mathbb{R}$, o también como el **Axioma del Supremo**, es el que permite que $\mathbb{R}$ sea, realmente, un conjunto extraordinariamente especial.

Definición 2.1.14 (Axioma del Supremo). Cualquier subconjunto **no vacío** A de $\mathbb{R}$ **acotado superiormente** posee un supremo.

Por ejemplo, si consideramos el conjunto $A = \{x \in \mathbb{R} : x^2 < 2\}$, entonces $A \neq \varnothing$ ya que, por ejemplo, $1 \in A$. Además, A está acotado superiormente por 2. Se sigue del Axioma del Supremo que $a_0 = \sup A$ existe. Es un ejercicio cotidiano verificar que $a_0 = \sqrt{2}$. Sin embargo, si reemplazamos a $\mathbb{R}$ por $\mathbb{Q}$, el Axioma del Supremo no es válido. En efecto, si tomamos $A = \{x \in \mathbb{Q} : x^2 < 2\}$, resulta que A es no vacío y acotado superiormente, pero $\sup A$ no existe en $\mathbb{Q}$.

Puesto que la relación $\sup(-A) = -\inf A$ se cumple para cualquier conjunto no vacío $A \subseteq \mathbb{R}$, y ya que el conjunto A está acotado inferiormente si, y sólo si, $-A$ está acotado superiormente, se sigue entonces del Axioma del Supremo que:

Corolario 2.1.15 (Axioma del Ínfimo). *Cualquier subconjunto* ***no vacío*** *A de $\mathbb{R}$* ***acotado inferiormente*** *posee un ínfimo.*

Una bonita aplicación del Axioma del Supremo la constituye la siguiente caracterización de los intervalos de $\mathbb{R}$.

Lema 2.1.16. *Sea $I \subseteq \mathbb{R}$. Si para cada $x, y \in I$ con $x < y$, ocurre que el intervalo abierto $(x, y) \subseteq I$, entonces I es un* ***intervalo****.*

Prueba. Suponga, en primer lugar, que I es acotado y sean

$$a \;=\; \inf I \qquad\qquad \text{y} \qquad\qquad b \;=\; \sup I.$$

Entonces $I \subseteq [a, b]$. Vamos a demostrar que $(a, b) \subseteq I$. En efecto, sea $c \in (a, b)$. Puesto que $\inf I = a < c$, se sigue del Teorema 2.1.7 que existe un $x \in I$ tal que $a \leq x < c$. Similarmente, como $c < b = \sup I$, el Teorema 2.1.6 nos revela la existencia de un $y \in I$ tal que $c < y \leq b$. La combinación de los dos hechos anteriores nos dice que $x, y \in I$ con $x < y$ y, así, invocando nuestra hipótesis tenemos que $c \in (x, y) \subseteq I$. Esto prueba que $(a, b) \subseteq I \subseteq [a, b]$ y, entonces I es un intervalo.

Los otros casos, por ejemplo, si I está acotado inferiormente pero no superiormente o viceversa, o si I no está acotado, se dejan como ejercicio al lector. ■

Como una consecuencia del resultado anterior se tiene que

Corolario 2.1.17. *Si $(I_\alpha)_{\alpha\in D}$ es una **familia de intervalos** de $\mathbb{R}$ y si existe al menos un $a \in \mathbb{R}$ tal que $a \in \bigcap_{\alpha\in D} I_\alpha$, entonces $I = \bigcup_{\alpha\in D} I_\alpha$ es un **intervalo**.*

Prueba. Sean $x, y \in I$ arbitrarios con $x < y$. De acuerdo al Lema 2.1.16 es suficiente demostrar que $(x, y) \subseteq I$. Ahora bien, como $x, y \in I$, existen $\alpha, \beta \in D$ tales que $x \in I_\alpha$ y $y \in I_\beta$. Consideremos los siguientes casos:

(*a*) Si $y \leq a$, entonces $(x, y) \subseteq (x, a) \subseteq I_\alpha \subseteq I$.

(*b*) Si $a \leq x$, entonces $(x, y) \subseteq (a, y) \subseteq I_\beta \subseteq I$.

(*c*) Si $x < a < y$, entonces $(x, y) \subseteq (x, a] \cup [a, y) \subseteq I_\alpha \cup I_\beta \subseteq I$.

En cualquier caso $(x, y) \subseteq I$ y termina la prueba. ■

2.1.3. Límites

La noción de límite es *la noción básica* del Análisis. Es esa noción la que separa el Análisis del Álgebra. Intuitivamente, decir que una sucesión $(x_n)_{n=1}^\infty$ de números reales converge a un límite L significa que *eventualmente* casi-todos los términos de la sucesión están, tanto como se desee, muy próximos al número L.

Definición 2.1.18. *Un número real L es el **límite** de una sucesión de números reales $(x_n)_{n=1}^\infty$ si, para cada $\varepsilon > 0$, existe un entero $N \geq 1$ tal que*

$$\left|x_n - L\right| < \varepsilon \quad \text{para todo } n \geq N.$$

*En este caso se dice que la sucesión **converge** a L y se escribe $\lim\limits_{n\to\infty} x_n = L$.*

Una sucesión $(x_n)_{n=1}^\infty$ se dice que **diverge** si ella no converge. Es conocido y fácil de establecer que el límite de una sucesión es único.

Teorema 2.1.19. *Sea $(x_n)_{n=1}^\infty$ una sucesión de números reales. Si $(x_n)_{n=1}^\infty$ **converge**, entonces el conjunto $\{x_n : n \in \mathbb{N}\}$ es **acotado**.*

Prueba. Sea $L = \lim_{n\to\infty} x_n$. Tomando $\varepsilon = 1$, resulta de la definición de límite, que existe un $N \in \mathbb{N}$ tal que $|x_n - L| < 1$ para todo $n \geq N$. En otras palabras,

$$L - 1 < x_n < L + 1 \quad \text{para todo } n \geq N.$$

Si ahora definimos $M = \max\{x_1, \ldots, x_{N-1}, L+1\}$ y $m = \min\{x_1, \ldots, x_{N-1}, L-1\}$, resultará que

$$m \leq x_n \leq M \quad \text{para todo } n \in \mathbb{N}.$$

Fin de la prueba. ■

Las propiedades algebraicas básicas de las sucesiones de números reales son las siguientes:

Teorema 2.1.20. *Sean* $(x_n)_{n=1}^{\infty}$ *y* $(y_n)_{n=1}^{\infty}$ *sucesiones en* $\mathbb{R}$ *tales que* $\lim_{n\to\infty} x_n = L$ *y* $\lim_{n\to\infty} y_n = M$. *Entonces*

(a) $\lim_{n\to\infty} (x_n + y_n) = L + M$,

(b) $\lim_{n\to\infty} \alpha x_n = \alpha L$ *para cualquier* $\alpha \in \mathbb{R}$,

(c) $\lim_{n\to\infty} x_n y_n = LM$, *y*

(d) $\lim_{n\to\infty} \dfrac{x_n}{y_n} = \dfrac{L}{M}$ *siempre que* $M \neq 0$.

Prueba. Los detalles se dejan a cargo del lector. ■

Teorema 2.1.21. *Sea* A *un subconjunto de* $\mathbb{R}$ ***no vacío*** *y* ***acotado superiormente***. *Si* $a_0 = \sup A$, *entonces* ***existe una sucesión*** $(x_n)_{n=1}^{\infty}$ *en* A *tal que*

$$a_0 \;=\; \lim_{n\to\infty} x_n.$$

Prueba. Por cada entero $n \geq 1$, seleccione, usando el Teorema 2.1.6, un punto $x_n \in A$ tal que $a_0 - 1/n < x_n$. De esto se sigue que

$$a_0 - \frac{1}{n} \;<\; x_n \;\leq\; a_0 \;<\; a_0 + \frac{1}{n}$$

y, así, $|x_n - a_0| < 1/n$ para todo $n \geq 1$. Tomando límite cuando $n \to \infty$ queda establecido el resultado. ■

2.1.4. El Teorema de Bolzano-Weierstrass

Recordemos que una sucesión $(x_n)_{n=1}^{\infty}$ en $\mathbb{R}$ está **acotada** si el conjunto $\{x_1, x_2, \ldots\}$ está acotado, es decir, si existe una constante $M > 0$ tal que $|x_n| \leq M$ para todo $n \geq 1$.

Definición 2.1.22. *Una sucesión* $(x_n)_{n=1}^{\infty}$ *se dice que es* ***monótona creciente***, *o simplemente,* ***creciente***, *si*

$$x_1 \;\leq\; x_2 \;\leq\; \cdots \;\leq\; x_n \;\leq\; \cdots$$

y es ***monótona decreciente***, *o* ***decreciente***, *si*

$$x_1 \;\geq\; x_2 \;\geq\; \cdots \;\geq\; x_n \;\geq\; \cdots$$

Si las desigualdades son estrictas en ambos casos, en la definición anterior, entonces diremos que la sucesión es **estrictamente creciente** y **estrictamente decreciente**, respectivamente. En ocasiones se usará la expresión $(x_n)_{n=1}^{\infty}$ es **monótona no-decreciente** en lugar de monótona creciente. Similarmente, $(x_n)_{n=1}^{\infty}$ es **monótona no-creciente** se usa como sinónimo de monótona decreciente. Una sucesión se dice que es **monótona** si ella es monótona creciente o monótona decreciente.

Uno de los resultados fundamentales que garantiza la convergencia de una sucesión acotada es el siguiente.

Teorema 2.1.23. *Sea $(x_n)_{n=1}^{\infty}$ una **sucesión acotada** en $\mathbb{R}$. Si $(x_n)_{n=1}^{\infty}$ es **monótona**, entonces ella **converge**. En particular,*

(*i*) $\lim\limits_{n\to\infty} x_n = \sup\limits_{n\geq 1} x_n$ *si* $(x_n)_{n=1}^{\infty}$ *es **creciente**.*

(*ii*) $\lim\limits_{n\to\infty} x_n = \inf\limits_{n\geq 1} x_n$ *si* $(x_n)_{n=1}^{\infty}$ *es **decreciente**.*

Prueba. (*i*). Suponga que $(x_n)_{n=1}^{\infty}$ es acotada y monótona creciente. Sea $a = \sup A$, donde hemos puesto $A = \{x_n : n \in \mathbb{N}\}$. Dado $\varepsilon > 0$, por el Teorema 2.1.6, existe un $x_{n_0} \in A$ tal que $a - \varepsilon < x_{n_0}$. Ahora bien, como nuestra sucesión es creciente, resulta que $x_{n_0} \leq x_n$ para todo $n \geq n_0$ y entonces, haciendo uso del hecho de que $x_n \leq a = \sup A$ para todo $n \in \mathbb{N}$, podemos concluir que

$$a - \varepsilon \ < \ x_{n_0} \ \leq \ x_n \ \leq \ a \ < \ a + \varepsilon \quad \text{para todo} \quad n \geq n_0,$$

es decir, $|x_n - a| < \varepsilon$ para todo $n \geq n_0$. Esto prueba que la sucesión $(x_n)_{n=1}^{\infty}$ converge a a y, por lo tanto,

$$\lim_{n\to\infty} x_n \ = \ a \ = \ \sup_{n\in\mathbb{N}} x_n.$$

Si la sucesión $(x_n)_{n=1}^{\infty}$ es acotada y monótona decreciente, entonces $(-x_n)_{n=1}^{\infty}$ es una sucesión monótona creciente y se sigue de lo anterior y el Teorema 2.1.11 que

$$-\lim_{n\to\infty} x_n \ = \ \lim_{n\to\infty}(-x_n) \ = \ \sup_{n\in\mathbb{N}}(-x_n) \ = \ -\inf_{n\in\mathbb{N}} x_n.$$

y termina la prueba. ■

Si $(x_n)_{n=1}^{\infty}$ es una sucesión en $\mathbb{R}$ y si $(n_k)_{k=1}^{\infty}$ es una *sucesión estrictamente creciente* de enteros positivos

$$n_1 \ < \ n_2 \ < \ \cdots \ < \ n_k \ < \ \cdots$$

entonces la sucesión $(x_{n_k})_{k=1}^{\infty}$ se llama una **subsucesión** de $(x_n)_{n=1}^{\infty}$. Resulta interesante observar que:

(*a*) Si $(n_k)_{k=1}^{\infty}$ es una sucesión estrictamente creciente de enteros positivos, entonces $n_k \geq k$ para todo $k \geq 1$.

(*b*) Si $(x_n)_{n=1}^{\infty}$ es una sucesión monótona acotada, entonces cualquier subsucesión de ella hereda esa propiedad.

(*c*) Si $(x_n)_{n=1}^{\infty}$ es una sucesión que converge a un punto $a \in \mathbb{R}$, entonces cualquier subsucesión de ella también converge y lo hace hacia el mismo punto a.

Teorema 2.1.24. *Cualquier **sucesión** $(x_n)_{n=1}^{\infty}$ en $\mathbb{R}$ posee una **subsucesión monótona**.*

Prueba. Considere el siguiente conjunto

$$D \ = \ \{n \in \mathbb{N} \ : \ x_n \ \leq \ x_k \ \text{ para todo } k \geq n\}.$$

Por supuesto, sólo existen dos posibilidades para el conjunto D: que sea finito o infinito.

(a) Si D es finito, sea $n_0 = \max D$. Observe que si n es cualquier número natural tal que $n > n_0$, entonces $n \notin D$ y, en consecuencia, por la definición de D, existe un $k_n \geq n$ tal que $x_n > x_{k_n}$. Esto es,

$$\textit{para cada } n > n_0, \textit{ existe } k_n \geq n \textit{ tal que } x_n > x_{k_n}. \qquad (*)$$

Vamos a usar lo anterior para obtener de $(x_n)_{n=1}^{\infty}$ una subsucesión decreciente. En efecto, como $n_0 + 1 \notin D$, existe, por $(*)$, un $n_1 > n_0 + 1$ tal que $x_{n_0+1} > x_{n_1}$. De nuevo, como $n_1 > n_0$, existe un $n_2 > n_1$ tal que $x_{n_1} > x_{n_2}$. Suponga que hemos construido $n_1 < n_2 < \cdots < n_k$ tal que $x_{n_j} > x_{n_{j+1}}$ para todo $j < k$. Una vez más, como $n_k \notin D$, existe, por $(*)$, un $n_{k+1} > n_k$ tal que $x_{n_k} > x_{n_{k+1}}$. Es claro que $(n_k)_{k=1}^{\infty}$ es una sucesión estrictamente creciente y que $(x_{n_k})_{k=1}^{\infty}$ es una subsucesión decreciente de $(x_n)_{n=1}^{\infty}$.

(b) Suponga que D es infinito. Se sigue del Teorema 1.2.5 que D se puede representar por una sucesión estrictamente creciente, digamos,

$$n_1 < n_2 < \cdots < n_k < \cdots$$

Es claro, por la definición de D, que $(x_{n_k})_{k=1}^{\infty}$ resulta ser una subsucesión creciente de $(x_n)_{n=1}^{\infty}$. La prueba es completa ■

El siguiente resultado es poderosamente importante ya que establece una condición suficiente para la existencia de subsucesiones convergentes.

Teorema 2.1.25 (Bolzano-Weierstrass). *Si $(x_n)_{n=1}^{\infty}$ es una* ***sucesión acotada*** *en $\mathbb{R}$, entonces ella posee una* ***subsucesión convergente****.*

Prueba. Usemos el Teorema 2.1.24 para seleccionar una subsucesión monótona $(x_{n_k})_{k=1}^{\infty}$, de la sucesión acotada $(x_n)_{n=1}^{\infty}$. Como $(x_{n_k})_{k=1}^{\infty}$ también es acotada, necesariamente ella converge gracias al Teorema 2.1.23. ■

Definición 2.1.26. *Una sucesión $(x_n)_{n=1}^{\infty}$ en $\mathbb{R}$ se dice que es una* ***sucesión de Cauchy*** *si, para cada $\varepsilon > 0$, existe un $N \in \mathbb{N}$ tal que*

$$|x_n - x_m| < \varepsilon \quad \textit{para todo } m, n \geq N.$$

Teorema 2.1.27. *Si $(x_n)_{n=1}^{\infty}$ es una* ***sucesión de Cauchy*** *en $\mathbb{R}$, entonces ella es* ***acotada****.*

Prueba. Para $\varepsilon = 1$, existe un $N \in \mathbb{N}$ tal que $|x_n - x_m| < 1$ para todo $m, n \geq N$. En particular,

$$|x_n - x_N| < 1 \quad \text{para todo } n \geq N.$$

de donde se obtiene que

$$|x_n| < 1 + |x_N| \quad \text{para todo } n \geq N.$$

Finalmente, si hacemos $M = \max\{|x_1|, |x_2|, \ldots, |x_{N-1}|, 1 + |x_N|\}$ vemos que

$$|x_n| \leq M \quad \text{para todo} \quad n \in \mathbb{N}$$

lo cual prueba que la sucesión $(x_n)_{n=1}^{\infty}$ es acotada. ■

Otro hecho interesante es el siguiente:

Teorema 2.1.28. *Si $(x_n)_{n=1}^{\infty}$ es una* ***sucesión de Cauchy*** *la cual posee una subsucesión $(x_{n_k})_{k=1}^{\infty}$ que converge a algún punto $x \in \mathbb{R}$, entonces $x_n \to x$.*

Prueba. Sea $\varepsilon > 0$ y seleccione un $k_0 \in \mathbb{N}$, usando el hecho de que $x_{n_k} \to x$, tal que

$$|x_{n_k} - x| \ < \ \varepsilon/2 \quad \text{para todo } k \geq k_0. \tag{1}$$

Por otro lado, como $(x_n)_{n=1}^{\infty}$ es de Cauchy, existe un $N \in \mathbb{N}$ tal que

$$|x_n - x_m| \ < \ \varepsilon/2 \quad \text{para todo } m, n \geq N. \tag{2}$$

Sea $N_1 = \max\{N, k_0\}$. Si $k \geq N_1$, entonces $k \geq k_0$ y (1) implica que

$$|x_{n_k} - x| \ < \ \varepsilon/2.$$

También, como $n_k > k \geq N_1 \geq N$, entonces (2) implica que

$$|x_k - x_{n_k}| \ < \ \varepsilon/2.$$

De estas dos últimas desigualdades se sigue que, para todo $k \geq N_1$,

$$|x_k - x| \ \leq \ |x_k - x_{n_k}| + |x_{n_k} - x| \ < \ \varepsilon.$$

Esto termina la prueba. ■

Teorema 2.1.29 ($\mathbb{R}$ es completo). *Una sucesión $(x_n)_{n=1}^{\infty}$ en $\mathbb{R}$ es* ***convergente*** *si, y sólo si, ella es de* ***Cauchy****.*

Prueba. Si $(x_n)_{n=1}^{\infty}$ es convergente, entonces claramente es de Cauchy. Recíprocamente, suponga que $(x_n)_{n=1}^{\infty}$ es de Cauchy. Entonces ella es acotada y se sigue del Teorema de Bolzano-Weierstrass que existe una subsucesión de $(x_n)_{n=1}^{\infty}$ que converge a un punto $x \in \mathbb{R}$. Por el Teorema 2.1.28, $x_n \to x$ y termina la prueba. ■

El resultado anterior es interesante por dos razones: la primera es que en $\mathbb{R}$ las sucesiones de Cauchy y las sucesiones convergentes son indistinguibles. La segunda es que cualquier sucesión de Cauchy converge sin mencionar cuál es su límite, un hecho que puede resultar muy útil en algunas circunstancias. También es importante tener presente que *si $(x_n)_{n=1}^{\infty}$ es de Cauchy, entonces se puede determinar la existencia una subsucesión estrictamente creciente de enteros positivos, digamos $(n_k)_{k=1}^{\infty}$, tal que*

$$\left|x_{n_k} - x_{n_{k+1}}\right| \ < \ 2^{-k} \quad \text{para todo } k \in \mathbb{N}.$$

Se invita al lector a llenar los detalles de esa afirmación.

2.1.5. Los Números Reales Extendidos

El Sistema de los **Números Reales Extendidos**, al que también se le llama **Recta Extendida**, se define como

$$\overline{\mathbb{R}} \ = \ \mathbb{R} \cup \{-\infty, +\infty\},$$

donde $-\infty$ y $+\infty$ son dos objetos que no pertenecen a $\mathbb{R}$. Un elemento $a \in \overline{\mathbb{R}}$ se dice que es **finito** si $a \in \mathbb{R}$, en caso contrario se llama **infinito**.

La naturaleza de los objetos $-\infty$ y $+\infty$ es totalmente irrelevante. Lo que es importante, en realidad, es cómo dichos objetos se integran con los números reales a través de las operaciones algebraicas usuales y las relaciones de orden. Comencemos por extender el orden de $\mathbb{R}$ a $\overline{\mathbb{R}}$ declarando que:

$$-\infty < +\infty \qquad \text{y} \qquad -\infty < x < +\infty \quad \text{para todo } x \in \mathbb{R}.$$

Dados $x, y \in \overline{\mathbb{R}}$, escribiremos $x < y$ y diremos que x ***es menor que*** y, cuando se satisface una de las siguientes condiciones:

(a) $x, y \in \mathbb{R}$ y $x < y$ en el orden usual de $\mathbb{R}$,

(b) $x \neq +\infty$ y $y = +\infty$,

(c) $x = -\infty$ y $y \neq -\infty$.

Escribiremos $x > y$ cuando $y < x$. El símbolo $x \leq y$, que se lee x ***es menor o igual que*** y, significa que $x < y$ o bien $x = y$. Similarmente, escribiremos $x \geq y$ cuando $y \leq x$.

La relación binaria $<$ que acabamos de definir sobre $\overline{\mathbb{R}}$ es, en realidad, una relación de orden total, la cual permite introducir, de la manera usual, la noción de intervalos. Por ejemplo, para cualesquiera $a, b \in \overline{\mathbb{R}}$ con $a < b$,

$$[a, b] = \{x \in \overline{\mathbb{R}} : a \leq x \leq b\}.$$

De modo similar se definen los intervalos $[a, b)$, $(a, b]$ y (a, b). En ocasiones escribiremos $\overline{\mathbb{R}} = [-\infty, +\infty]$. Podemos introducir una topología sobre $\overline{\mathbb{R}}$ declarando que: G es **abierto** en $\overline{\mathbb{R}}$ si, y sólo si, G es abierto en $\mathbb{R}$ o G es de la forma

$$(a, +\infty] \qquad \text{o} \qquad [-\infty, a)$$

para algún $a \in \mathbb{R}$.

¿Cómo se han de sumar y multiplicar los símbolos $-\infty$ y $+\infty$ con los restantes elementos de $\overline{\mathbb{R}}$? Ellos se harán de acuerdo a las siguientes reglas:

(1) Para todo $x \in \mathbb{R}$,

$$x + (+\infty) = (+\infty) + x = +\infty \qquad \text{y} \qquad x + (-\infty) = (-\infty) + x = -\infty$$

(2) Si $x \in \mathbb{R}$ con $x > 0$, entonces

$$x \cdot (+\infty) = (+\infty) \cdot x = +\infty \qquad \text{y} \qquad x \cdot (-\infty) = (-\infty) \cdot x = -\infty.$$

(3) Si $x \in \mathbb{R}$ con $x < 0$, entonces

$$x \cdot (+\infty) = (+\infty) \cdot x = -\infty \qquad \text{y} \qquad x \cdot (-\infty) = (-\infty) \cdot x = +\infty.$$

En particular,

$$(-1)(+\infty) = -(+\infty) = -\infty \qquad \text{y} \qquad (-1)(-\infty) = -(-\infty) = +\infty.$$

El convenimiento más importante relativo a la aritmética de $\overline{\mathbb{R}}$ es el siguiente:

(4) $0\cdot(+\infty) \;=\; (+\infty)\cdot 0 \;=\; 0\cdot(-\infty) \;=\; (-\infty)\cdot 0 \;=\; 0.$

(5) Finalmente,

$$(+\infty)+(+\infty) \;=\; +\infty \;=\; (-\infty)\cdot(-\infty)$$

y

$$(-\infty)+(-\infty) \;=\; -\infty \;=\; (+\infty)\cdot(-\infty).$$

Precaución: La relación $a+(-a)=0$, la cual es válida para todo $a\in\mathbb{R}$, no se extiende a los elementos $-\infty$ y $+\infty$ de $\overline{\mathbb{R}}$, es decir, declaramos que las expresiones

$$(+\infty)+(-\infty) \qquad \text{y} \qquad (-\infty)+(+\infty),$$

no tienen significado en $\overline{\mathbb{R}}$ y, por consiguiente, **se prohíben**. En efecto, cualquier "valor" que se pretenda asignarle a tales expresiones conlleva a obtener resultados absurdos. Por ejemplo, si intentamos definir $\infty-\infty=0$, entonces, usando (1), tendríamos que $1+\infty=-1+\infty$ de donde se obtendría que $1=-1$. Otro aspecto que se pierde en $\overline{\mathbb{R}}$ es que el producto de dos límites de sucesiones no es necesariamente igual al límite del producto. Por ejemplo, si tomamos $a_n=n$ y $b_n=1/n$ para todo $n\in\mathbb{N}$, entonces $\lim_{n\to\infty}a_n=+\infty$ y $\lim_{n\to\infty}b_n=0$. Sin embargo,

$$0 \;=\; \Big(\lim_{n\to\infty}a_n\Big)\cdot\Big(\lim_{n\to\infty}b_n\Big) \;\neq\; \lim_{n\to\infty}a_n\cdot b_n \;=\; 1.$$

Similarmente, la expresión $x/0$ con $x\in\overline{\mathbb{R}}$, carece de significado en $\overline{\mathbb{R}}$. Sin embargo, la expresión $x/\pm\infty$ siempre tiene sentido para cualquier $x\in\mathbb{R}$ y su valor es cero, es decir, $x/\pm\infty=0$.

El siguiente resultado es una generalización en $\overline{\mathbb{R}}$ del Teorema 2.1.3.

Teorema 2.1.30. *Sean $a,b\in\overline{\mathbb{R}}$. Si para cada $c\in\mathbb{R}$ con $c>b$ ocurre que $a<c$, entonces $a\leq b$.*

Prueba. Suponga, para obtener una contradicción, que $a>b$. Entonces $a\neq-\infty$, y $b\neq+\infty$. Escojamos ahora un número real c tal que $a>c>b$. Como $c>b$, nuestra hipótesis nos dice que $a<c$, lo cual contradice nuestra elección de c. Por esto, $a\leq b$ y termina la prueba. ■

Recordemos que si A es un **subconjunto no vacío y acotado** de $\mathbb{R}$, hemos acordado en definir el $\sup A$ como la menor de las cotas superiores de A y a $\inf A$ como la mayor de las cotas inferiores de A. Si A **no está acotado**, extendemos esta noción declarando que:

$$\inf A \;=\; -\infty \quad \textit{si } A \textbf{ no está acotado inferiormente},$$

$$\sup A \;=\; +\infty \quad \textit{si } A \textbf{ no está acotado superiormente}.$$

Por consiguiente, se cumple que

$$\inf A \;\leq\; \sup A \qquad \text{si, y sólo si} \qquad A\neq\varnothing. \tag{2.1.2}$$

¿Qué ocurre cuando $A=\varnothing$? Hemos visto que si A y B son subconjuntos arbitrarios no vacíos de $\mathbb{R}$ con $A\subseteq B$, entonces

$$\inf B \;\leq\; \inf A \;\leq\; \sup A \;\leq\; \sup B.$$

Puesto que $\varnothing \subseteq A$ es una relación válida para cualquier subconjunto A de $\mathbb{R}$, si queremos que exista concordancia con lo antes expuesto y que las desigualdades anteriores sigan siendo válidas *cualesquiera sean los conjuntos* A y B con $A \subseteq B$, entonces se debe postular que:

$$\text{ínf}\,\varnothing \;=\; +\infty \qquad\qquad \text{y} \qquad\qquad \sup \varnothing \;=\; -\infty.$$

Establecidas estas consideraciones vemos que $\text{ínf}\,A$ y $\sup A$ **siempre existen como elementos de** $\overline{\mathbb{R}}$ cualquiera sea el conjunto $A \subseteq \mathbb{R}$. En particular,

$$\sup(-A) \;=\; -\,\text{ínf}\,A \qquad\qquad \text{y} \qquad\qquad \text{ínf}(-A) \;=\; -\sup A.$$

2.1.6. Limites Superior e Inferior de una Sucesión

Cuando una sucesión $(x_n)_{n=1}^{\infty}$ de números reales *converge*, se pueden recopilar algunas *buenas* propiedades de dicha sucesión. Por ejemplo, podemos tener información de cuál es su límite, también sabemos que ella es acotada, que cualquiera de sus subsucesiones converge y todas lo hacen hacia el mismo punto, etc. Si *la sucesión no converge* parte de la información anterior se pierde, pero puede ser de interés explorar el comportamiento de sus subsucesiones. Para analizar este caso, fijemos una sucesión $(x_n)_{n=1}^{\infty}$ en $\mathbb{R}$ y consideremos el siguiente conjunto:

$$\mathfrak{L}^*(x_n) \;=\; \Big\{ z \in \overline{\mathbb{R}} \,:\, z = \lim_{k\to\infty} x_{n_k}, \text{ para alguna subsucesión } (x_{n_k})_{k=1}^{\infty} \text{ de } (x_n)_{n=1}^{\infty} \Big\}$$

Observe que nuestro conjunto $\mathfrak{L}^*(x_n)$ consiste de todos los puntos de acumulación de la sucesión $(x_n)_{n=1}^{\infty}$. Ahora analizaremos en profundidad esta situación.

(1) *Si* $(x_n)_{n=1}^{\infty}$ *está acotada, pero no converge*, el Teorema de Bolzano-Weierstrass nos garantiza la existencia de al menos una subsucesión de $(x_n)_{n=1}^{\infty}$ que converge a algún número real, de donde resulta que $\mathfrak{L}^*(x_n)$ es un conjunto *no vacío* y, por supuesto, *acotado*. Esto último permite asegurar que

$$m \;=\; \text{ínf}\,\mathfrak{L}^*(x_n) \qquad\qquad \text{y} \qquad\qquad M \;=\; \sup \mathfrak{L}^*(x_n)$$

siempre existen en $\mathbb{R}$. Por supuesto, los números m y M representan el ***menor*** y el ***mayor***, respectivamente, de los puntos límites de la sucesión $(x_n)_{n=1}^{\infty}$. Esos números son llamados el ***límite inferior*** y el ***límite superior***, respectivamente, de la sucesión $(x_n)_{n=1}^{\infty}$. Observe que, en este caso, $m < M$ y

$$\{m, M\} \;\subseteq\; \mathfrak{L}^*(x_n) \;\subseteq\; [m, M].$$

Aunque la sucesión $(x_n)_{n=1}^{\infty}$ no converge, los números $\text{ínf}\,\mathfrak{L}(x_n)$ y $\sup \mathfrak{L}(x_n)$, como veremos un poco más abajo, permiten obtener información sobre la "*distribución*" de los puntos tanto de la sucesión $(x_n)_{n=1}^{\infty}$ así como de sus subsucesiones convergentes.

(2) *¿Qué ocurre si la sucesión* $(x_n)_{n=1}^{\infty}$ *no está acotada*? Por ejemplo, si $(x_n)_{n=1}^{\infty}$ no está acotada superiormente, resulta del Teorema 2.1.24 que ella posee una subsucesión creciente, digamos $(x_{n_k})_{k=1}^{\infty}$, tal que

$$\lim_{k\to\infty} x_{n_k} \;=\; \sup\{x_{n_k} : k \in \mathbb{N}\} \;=\; +\infty \in \mathfrak{L}^*(x_n).$$

Esto prueba que $\sup \mathfrak{L}^*(x_n) \;=\; +\infty$. Similarmente, si $(x_n)_{n=1}^{\infty}$ no está acotada inferiormente, entonces $-\infty \in \mathfrak{L}^*(x_n)$ y, así, $\text{ínf}\,\mathfrak{L}^*(x_n) \;=\; -\infty$. Todo lo anterior garantiza, por supuesto, que

$\mathfrak{L}^*(x_n)$ *siempre es un subconjunto* ***no vacío*** *de* $\overline{\mathbb{R}}$ independientemente si la sucesión $(x_n)_{n=1}^{\infty}$ es o no acotada y, por lo tanto,

$$\sup \mathfrak{L}^*(x_n) \qquad \text{y} \qquad \text{ínf}\, \mathfrak{L}^*(x_n)$$

existen en $\overline{\mathbb{R}}$.

Definición 2.1.31. *Sea $(x_n)_{n=1}^{\infty}$ una sucesión en $\mathbb{R}$. Definimos el* ***límite superior*** *y el* ***límite inferior*** *de la sucesión $(x_n)_{n=1}^{\infty}$, respectivamente, como*

$$\limsup_{n\to\infty} x_n = \sup \mathfrak{L}^*(x_n) \qquad y \qquad \liminf_{n\to\infty} x_n = \inf \mathfrak{L}^*(x_n).$$

Algunas veces escribiremos $\limsup x_n$ en lugar de $\limsup_{n\to\infty} x_n$. Igual consideración haremos para $\liminf_{n\to\infty} x_n$. La notación

$$\varliminf_{n\to\infty} x_n \qquad \text{y} \qquad \varlimsup_{n\to\infty} x_n$$

también se usa con mucha frecuencia para denotar, respectivamente, el límite inferior y el límite superior de la sucesión $(x_n)_{n=1}^{\infty}$. Observe que, por definición, existe al menos una subsucesión $(x_{n_k})_{k=1}^{\infty}$ de $(x_n)_{n=1}^{\infty}$ tal que

$$\limsup_{n\to\infty} x_n = \lim_{k\to\infty} x_{n_k}.$$

Lo mismo se cumple para el límite inferior.

Puesto que $\mathfrak{L}^*(x_n)$ es no vacío para cualquier sucesión $(x_n)_{n=1}^{\infty}$ en $\mathbb{R}$, resulta de (2.1.2) que

$$\liminf_{n\to\infty} x_n = \inf \mathfrak{L}^*(x_n) \leq \sup \mathfrak{L}^*(x_n) = \limsup_{n\to\infty} x_n. \tag{1}$$

Más aun, si la sucesión $(x_n)_{n=1}^{\infty}$ converge a un punto z, entonces también converge cualquiera de sus subsucesiones y, por supuesto, lo hace hacia el mismo punto z. Por esto,

$$\begin{aligned} \lim_{n\to\infty} x_n = z \in \overline{\mathbb{R}} \quad &\Leftrightarrow \quad \mathfrak{L}^*(x_n) = \{z\} \\ &\Leftrightarrow \quad \liminf_{n\to\infty} x_n = z = \limsup_{n\to\infty} x_n. \end{aligned}$$

Dicho de otra manera:

Corolario 2.1.32. *Una sucesión $(x_n)_{n=1}^{\infty}$ en $\mathbb{R}$* ***converge*** *si, y sólo si,*

$$\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n = \lim_{n\to\infty} x_n \in \mathbb{R}.$$

Observe que gracias a (1) y la última igualdad del corolario anterior podemos concluir que:

Corolario 2.1.33. *Una sucesión $(x_n)_{n=1}^{\infty}$ en $\mathbb{R}$* ***converge*** *si, y sólo si $\liminf\limits_{n\to\infty} x_n$ y $\limsup\limits_{n\to\infty} x_n$ existen (en $\mathbb{R}$) y*

$$\limsup_{n\to\infty} x_n \leq \liminf_{n\to\infty} x_n. \tag{2}$$

Dos de las propiedades importantes que relacionan directamente los límites inferior y superior, el cual es consecuencia del Teorema 2.1.11, son las siguientes:

$$\limsup_{n\to\infty} (-x_n) = -\liminf_{n\to\infty} x_n$$

y

$$\liminf_{n\to\infty} (-x_n) = -\limsup_{n\to\infty} x_n.$$

Si bien la definición de los límites superior e inferior es impecable desde el punto de vista teórico, ellas no son muy adecuadas a la hora calcular los límites inferiores y superiores de una sucesión en particular, pues habría que determinar, en primer lugar, *todos* los puntos límites de dicha sucesión, una tarea que pudiera ser un tanto difícil, y luego elegir el ínfimo y el supremo de tales puntos límites. La caracterización dada en el próximo teorema es, en muchos aspectos, más practica y, por supuesto, más útil.

Suponga de nuevo que $(x_n)_{n=1}^{\infty}$ es una sucesión en $\mathbb{R}$ y defina, para cada entero $n \geq 1$, los siguientes números reales extendidos:

$$m_n = \inf\{x_n, x_{n+1}, \dots\} \qquad \text{y} \qquad M_n = \sup\{x_n, x_{n+1}, \dots\}.$$

Observe que:

(ι_1) $m_n \leq x_j \leq M_n$ para todo $j \geq n \in \mathbb{N}$,

(ι_2) $m_1 \leq m_2 \leq \cdots \leq m_n \leq \cdots$ y

(ι_3) $M_1 \geq M_2 \geq \cdots \geq M_n \geq \cdots$.

Puesto que las sucesiones $(m_n)_{n=1}^{\infty}$ y $\left(M_n\right)_{n=1}^{\infty}$ son ambas monótonas en $\overline{\mathbb{R}}$, resulta que los límites

$$\lim_{n\to\infty} M_n = \inf_{n\geq 1} M_n = \inf_{n\geq 1} \sup_{k\geq n} x_n$$

y

$$\lim_{n\to\infty} m_n = \sup_{n\geq 1} m_n = \sup_{n\geq 1} \inf_{k\geq n} x_n.$$

siempre existen en $\overline{\mathbb{R}}$.

Teorema 2.1.34. *Sea $(x_n)_{n=1}^{\infty}$ una sucesión en $\mathbb{R}$. Entonces*

$$\limsup_{n\to\infty} x_n = \inf_{n\geq 1} \sup_{k\geq n} x_n \qquad \text{y} \qquad \liminf_{n\to\infty} x_n = \sup_{n\geq 1} \inf_{k\geq n} x_n.$$

Prueba. Sea $M = \lim_{n\to\infty} M_n$, el cual, como acabamos de ver, siempre existe en $\overline{\mathbb{R}}$. Para demostrar que $\sup \mathfrak{L}^*(x_n) = M$ tomemos cualquier $z \in \mathfrak{L}^*(x_n)$. Por definición, existe una subsucesión $(x_{n_j})_{j=1}^{\infty}$ de $(x_n)_{n=1}^{\infty}$ tal que $z = \lim_{j\to\infty} x_{n_j}$ y entonces, por (ι_1), $x_{n_j} \leq M_{n_j}$ para todo $j \geq 1$, de donde se sigue que

$$z = \lim_{j\to\infty} x_{n_j} \leq \lim_{j\to\infty} M_{n_j} = M.$$

Esta desigualdad nos revela que M es una cota superior de $\mathfrak{L}^*(x_n)$ y, por lo tanto,

$$\sup \mathfrak{L}^*(x_n) \leq M. \tag{2.1.3}$$

Para probar la otra desigualdad suponga, en primer lugar, que $M_1 = +\infty$. Esto significa que la sucesión $(x_n)_{n=1}^{\infty}$ no está acotada superiormente y, por consiguiente, $+\infty = \sup \mathfrak{L}^*(x_n)$. Se sigue de (2.1.3) que $M = +\infty$ y, así,

$$\sup \mathfrak{L}^*(x_n) \;=\; M \;=\; +\infty.$$

Suponga entonces que $M_1 < +\infty$. Por la propiedad del supremo, existe un x_{n_1} tal que

$$M_1 - 1 \;<\; x_{n_1} \;\leq\; M_1.$$

Considere ahora $M_{n_1+1} = \sup\{x_{n_1+1}, x_{n_1+2}, \ldots\}$ y note que como $M_{n_1+1} \leq M_1$, entonces $M_{n_1+1} < +\infty$. Usemos de nuevo la propiedad del supremo para hallar un x_{n_2} tal que

$$M_{n_1+1} - \frac{1}{2} \;<\; x_{n_2} \;\leq\; M_{n_1+1}.$$

Observe que $n_2 > n_1$. Una vez más. Considere $M_{n_2+1} = \sup\{x_{n_2+1}, x_{n_2+2}, \ldots\}$ y, como antes, apliquemos la propiedad del supremo para hallar un x_{n_3} tal que

$$M_{n_2+1} - \frac{1}{3} \;<\; x_{n_3} \;\leq\; M_{n_2+1}.$$

Si se continúa indefinidamente con este procedimiento, se logra obtener una subsucesión $(x_{n_j})_{j=1}^{\infty}$ de $(x_n)_{n=1}^{\infty}$ tal que, para todo $j \in \mathbb{N}$,

$$M_{n_j+1} - \frac{1}{1+j} \;<\; x_{n_{j+1}} \;\leq\; M_{n_j+1}.$$

Por el Teorema del Sandwich para límites se obtiene que

$$\lim_{j\to\infty} x_{n_j} \;=\; \lim_{j\to\infty} M_{n_j+1} \;=\; M \;\in\; \mathfrak{L}^*(x_n).$$

De esto se concluye que $M \leq \sup \mathfrak{L}^*(x_n)$ y, por lo tanto, usando (2.1.3), vemos que $M = \sup \mathfrak{L}^*(x_n)$. El caso del límite inferior es similar y se deja a cargo del lector. ■

La manera de expresar a los límites inferior y superior dada en el resultado anterior permite asignarle una interpretación geométrica que resulta ser muy conveniente.

Teorema 2.1.35. *Sea $(x_n)_{n=1}^{\infty}$ una sucesión en $\mathbb{R}$. Entonces*

$$\limsup_{n\to\infty} x_n \;=\; a_0 \in \mathbb{R}$$

si, y sólo si, para cada $\varepsilon > 0$:

(a) existe un $N \in \mathbb{N}$ tal que $x_n < a_0 + \varepsilon$ para todo $n \geq N$, y

(b) $\{n \in \mathbb{N} : a_0 - \varepsilon < x_n\}$ es infinito.

Prueba. Suponga que $\limsup_{n\to\infty} x_n = a_0 \in \mathbb{R}$ y sea $\varepsilon > 0$ elegido arbitrariamente. Veamos que (a) y (b) se cumplen. En efecto, puesto que $a_0 = \lim_{n\to\infty} M_n$, donde $M_n = \sup\{x_n, x_{n+1}, \ldots\}$ para todo $n \in \mathbb{N}$, entonces existe un entero $N \in \mathbb{N}$ tal que

$$|M_n - a_0| \;<\; \varepsilon \quad \text{para todo } n \geq N,$$

es decir, $a_0 - \varepsilon < M_n < a_0 + \varepsilon$ para todo $n \geq N$, de donde se sigue que

$$x_n \ \leq \ \sup\{x_n, x_{n+1}, \dots\} \ = \ M_n \ < \ a_0 \ + \ \varepsilon \quad \text{para todo } n \geq N.$$

Esto prueba que

$$x_n \ < \ a_0 \ + \ \varepsilon \quad \text{para todo } n \geq N$$

y (a) se cumple. Para demostrar (b) recordemos que, por el Teorema 2.1.34, $a_0 = \operatorname{ínf}_{n\geq 1} \sup_{k\geq n} x_k$ de donde resulta, por la definición de ínfimo, que

$$a_0 \ \leq \ \sup\{x_k, x_{k+1}, \dots\} \ = \ M_k \qquad \text{para todo } k \in \mathbb{N}. \tag{2.1.4}$$

En particular, $a_0 \leq M_1$. Ahora bien, usando las propiedades del supremo, existe un entero n_1 tal que $x_{n_1} > M_1 - \varepsilon$, y, por consiguiente,

$$x_{n_1} \ > \ a_0 - \varepsilon.$$

Tomando $k = n_1$ en (2.1.4), existe un $n_2 > n_1$ tal que

$$x_{n_1} \ > \ M_{n_1} - \varepsilon \ > \ a_0 - \varepsilon.$$

Procediendo indefinidamente con el mecanismo anterior, se obtiene una subsucesión $(x_{n_k})_{k=1}^{\infty}$ de $(x_n)_{n=1}^{\infty}$ tal que

$$x_{n_k} \ > \ a_0 \ - \ \varepsilon \qquad \text{para todo } k \in \mathbb{N}.$$

Puesto que $n_k > k$ para todo $k \in \mathbb{N}$ resulta de lo anterior que

$$\{n \in \mathbb{N} \, : \, a_0 - \varepsilon \ < \ x_n\}$$

es infinito y así (b) también se cumple. Es claro que (a) y (b) implican que $\limsup\limits_{n\to\infty} x_n = a_0$. ■

Observe que la condición (a) del resultado anterior establece que *todos los términos* de la sucesión $(x_n)_{n=1}^{\infty}$, salvo un número infinito ellos, están a la izquierda de $a_0 + \varepsilon$, mientras que la condición (b) expresa que *siempre hay una infinidad de términos* de la sucesión a la derecha de $a_0 - \varepsilon$.

Un argumento enteramente similar a la demostración del teorema anterior permite obtener el siguiente resultado para el límite inferior.

Teorema 2.1.36. *Sea $(x_n)_{n=1}^{\infty}$ una sucesión en $\mathbb{R}$. Entonces $\liminf\limits_{n\to\infty} x_n = b_0 \in \mathbb{R}$ si, y sólo si, para cada $\varepsilon > 0$,*

(a) existe un $N \in \mathbb{N}$ tal que $b_0 - \varepsilon < x_n$ para todo $n \geq N$, y

(b) $\{n \in \mathbb{N} \, : \, x_n < b_0 + \varepsilon\}$ es infinito.

Prueba. Se omite por ser similar al resultado anterior. ■

Cuando el límite inferior, respectivamente, el límite superior es infinito se obtiene la siguiente caracterización.

Teorema 2.1.37. *Sea* $(x_n)_{n=1}^{\infty}$ *una sucesión en* $\mathbb{R}$. *Son equivalentes:*

(1) $\liminf\limits_{n\to\infty} x_n = +\infty$.

(2) *Para cada* $M > 0$, *existe un* $N \in \mathbb{N}$ *tal que* $x_n \geq M$ *para todo* $n \geq N$.

Prueba. Sea $M > 0$. Entonces

$$\begin{aligned}
\liminf_{n\to\infty} x_n = +\infty \ \Leftrightarrow\ \sup_{n\geq 1} m_n &= +\infty \\
&\Leftrightarrow \text{ existe un } \ N \in \mathbb{N} \ \text{ tal que } m_N \geq M \\
&\Leftrightarrow \text{ existe un } \ N \in \mathbb{N} \ \text{ tal que } \inf\{x_N, x_{N+1}, \ldots\} \geq M \\
&\Leftrightarrow \text{ existe un } \ N \in \mathbb{N} \ \text{ tal que } x_n \geq M \text{ para todo } n \geq N.
\end{aligned}$$

Teorema 2.1.38. *Sea* $(x_n)_{n=1}^{\infty}$ *una sucesión en* $\mathbb{R}$. *Son equivalentes:*

(1) $\limsup\limits_{n\to\infty} x_n = -\infty$.

(2) *Para cada* $M > 0$, *existe un* $N \in \mathbb{N}$ *tal que* $x_n \leq M$ *para todo* $n \geq N$.

Prueba. Es similar al caso anterior y se omite. ■

Ejemplo 2.1.1. *Si* $x_n = n$ *para todo* $n \in \mathbb{N}$, *entonces* $m_n = n$ *y* $M_n = +\infty$, *por lo que*

$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty} m_n = +\infty \quad y \quad \limsup_{n\to\infty} x_n = \lim_{n\to\infty} M_n = +\infty.$$

Ejemplo 2.1.2. *Si* $x_n = (-1)^n$ *para todo* $n \in \mathbb{N}$, *entonces*

$$\liminf_{n\to\infty} (-1)^n = -1 \qquad y \qquad \limsup_{n\to\infty} (-1)^n = 1.$$

Ejemplo 2.1.3. *Si* $x_n = (-1)^n/n$ *para todo* $n \in \mathbb{N}$, *entonces*

$$\liminf_{n\to\infty} \frac{(-1)^n}{n} = 0 = \limsup_{n\to\infty} \frac{(-1)^n}{n}.$$

Teorema 2.1.39. *Sean* $(x_n)_{n=1}^{\infty}$ y $(y_n)_{n=1}^{\infty}$ *sucesiones en* $\mathbb{R}$. *Entonces se cumple que*

(a) $\limsup\limits_{n\to\infty} (x_n + y_n) \leq \limsup\limits_{n\to\infty} x_n + \limsup\limits_{n\to\infty} y_n$.

(b) $\liminf\limits_{n\to\infty} (x_n + y_n) \geq \liminf\limits_{n\to\infty} x_n + \liminf\limits_{n\to\infty} y_n$.

(c) $\limsup\limits_{n\to\infty} x_n \leq \limsup\limits_{n\to\infty} y_n$ *y* $\liminf\limits_{n\to\infty} x_n \leq \liminf\limits_{n\to\infty} y_n$ *si* $x_n \leq y_n$ $\forall n \geq 1$.

Prueba. Sólo demostraremos (a), pues la prueba de (b) es muy similar y (c) es inmediata. Observe, en primer lugar, que si uno de los límites superiores de la derecha en (a) es infinito, entonces la igualdad se cumple trivialmente. Suponga entonces que el lado derecho es finito y definamos, para cada $n \in \mathbb{N}$,

$$M_n = \sup\{x_n, x_{n+1}, \dots\} \qquad \text{y} \qquad M'_n = \sup\{y_n, y_{n+1}, \dots\}$$

Entonces, para todo $k \geq n$,

$$x_k \leq M_n \qquad \text{y} \qquad y_k \leq M'_n$$

y, en consecuencia,

$$x_k + y_k \leq M_n + M'_n$$

para todo $k \geq n$. Esto prueba que $M_n + M'_n$ es una cota superior para el conjunto

$$A_n = \{x_n + y_n,\, x_{n+1} + y_{n+1}, \dots\},$$

y, por lo tanto,

$$\sup A_n = \sup\{x_n + y_n,\, x_{n+1} + y_{n+1}, \dots\} \leq M_n + M'_n.$$

Como la sucesión $(\sup A_n)_{n=1}^{\infty}$ es decreciente, resulta que

$$\begin{aligned}
\limsup_{n\to\infty}\,(x_n + y_n) &= \inf_{n\geq 1}\sup_{n\geq k} A_n \\
&\leq \lim_{n\to\infty}\left(M_n + M'_n\right) \\
&= \lim_{n\to\infty} M_n + \lim_{n\to\infty} M'_n \\
&= \limsup_{n\to\infty}\, x_n + \limsup_{n\to\infty}\, y_n.
\end{aligned}$$

Para demostrar (c) observe que si $x_n \leq y_n$ para todo $n \geq 1$, entonces

$$M_n \leq M'_n \qquad \text{y} \qquad m_n \leq m'_n$$

donde $m_n = \inf\{x_n, x_{n+1}, \dots\}$ y $m'_n = \{y_n, y_{n+1}, \dots\}$. Tomando el límite cuando $n \to \infty$ en las desigualdades anteriores se obtienen los resultados deseados. ■

Nota Adicional 2.1.1 La relación de orden impuesta sobre $\overline{\mathbb{R}}$ hace que $-\infty$ sea el elemento más pequeño y $+\infty$ el elemento más grande. Esta relación de orden es compatible con la topología puesto que los conjuntos abiertos son uniones de intervalos. Si se considera la aplicación $\varphi = \overline{\mathbb{R}} \to [-1,1]$ definida por

$$\varphi(x) = \begin{cases} \dfrac{x}{\sqrt{1+x^2}} & \text{si } x \in \mathbb{R} \\ \pm 1 & \text{si } x = \pm\infty \end{cases}$$

entonces es fácil ver que φ es un homeomorfismo que es compatible con la relación de orden, es decir, para todo $x, y \in \overline{\mathbb{R}}$,

$$x < y \quad \Rightarrow \quad \varphi(x) < \varphi(y). \tag{1}$$

De esta información se sigue que si $(x_n)_{n=1}^{\infty}$ es una sucesión monótona en $\overline{\mathbb{R}}$, entonces ella converge. En efecto, de (1) se sigue que $(\varphi(x_n))_{n=1}^{\infty}$ es una sucesión monótona en $[-1,1]$ y, en consecuencia, converge. Puesto que φ^{-1} es continua, resulta que $(x_n)_{n=1}^{\infty}$ converge en $\overline{\mathbb{R}}$. Estas consideraciones permiten que la noción de limite superior y limite inferior puedan ser tratados en $[-1,1]$ en lugar de $\overline{\mathbb{R}}$.

2.1.7. Límites Superior e Inferior de Conjuntos

Las nociones de límites superiores e inferiores de conjuntos es importante en la Teoría de la Medida. En lo que sigue X denotará un subconjunto arbitrario no vacío.

Definición 2.1.40. *Sea X un conjunto no vacío. Dada una sucesión* $(A_n)_{n=1}^{\infty}$ *de subconjuntos de X, definimos las sucesiones* $(\underline{A}_n)_{n=1}^{\infty}$ *y* $(\overline{A}_n)_{n=1}^{\infty}$ *por*

$$\underline{A}_n = \bigcap_{k=n}^{\infty} A_k \qquad y \qquad \overline{A}_n = \bigcup_{k=n}^{\infty} A_k$$

para $n = 1, 2, \ldots$ *Los conjuntos*

$$\liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} \underline{A}_n = \bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} A_k \qquad y \qquad \limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \overline{A}_n = \bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k$$

los llamaremos, respectivamente, el ***límite inferior*** *y el* ***límite superior*** *de la sucesión* $(A_n)_{n=1}^{\infty}$.

La notación

$$\underline{\lim_{n\to\infty}}\, A_n \qquad \text{y} \qquad \overline{\lim_{n\to\infty}}\, A_n$$

se usan frecuentemente para denotar a los límites inferior y superior, respectivamente, de la sucesión $(A_n)_{n=1}^{\infty}$. En ocasiones, escribiremos $\liminf A_n$ en lugar de $\liminf_{n\to\infty} A_n$ y similarmente para el límite superior. De las Leyes de Morgan se sigue que

$$\limsup_{n\to\infty} A_n^c = \big(\liminf_{n\to\infty} A_n\big)^c \qquad \text{y} \qquad \liminf_{n\to\infty} A_n^c = \big(\limsup_{n\to\infty} A_n\big)^c.$$

La siguiente caracterización de los conjuntos $\liminf A_n$ y $\limsup A_n$ resulta conveniente para ciertos propósitos.

Teorema 2.1.41. *Sea* $(A_n)_{n=1}^{\infty}$ *una sucesión de subconjuntos de X. Entonces*

(a) $x \in \liminf_{n\to\infty} A_n$ *si, y sólo si, existe un* $N \in \mathbb{N}$ *tal que, para todo* $n \geq N$, $x \in A_n$, *es decir,*

$$\liminf_{n\to\infty} A_n = \big\{x : x \in A_n \text{ excepto para un número finito de } n\text{'s}\big\}.$$

(b) $x \in \limsup_{n\to\infty} A_n$ *si, y sólo si, para cada* $N \in \mathbb{N}$ *existe* $n \geq N$ *tal que* $x \in A_n$, *es decir,*

$$\limsup_{n\to\infty} A_n = \big\{x : x \in A_n \text{ para infinitos } n\text{'s}\big\}.$$

Prueba. (a) Sea $x \in \liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} \underline{A}_n$. Entonces existe un $N \in \mathbb{N}$ tal que $x \in \underline{A}_N = \bigcap_{n=N}^{\infty} A_n$. Por esto, $x \in A_n$ para todo $n \geq N$. Recíprocamente, sean $x \in X$ y $N \in \mathbb{N}$ de modo que $x \in A_n$ para todo $n \geq N$. Entonces

$$x \in \bigcap_{n=N}^{\infty} A_n = \underline{A}_N \subseteq \liminf_{n\to\infty} A_n.$$

(b) Es similar a la anterior y, por lo tanto, se deja a cargo del lector. ■

Algunas de las propiedades de los conjuntos $\liminf_{n\to\infty} A_n$ y $\limsup_{n\to\infty} A_n$ se detallan a continuación.

Teorema 2.1.42. *Sea* $(A_n)_{n=1}^{\infty}$ *una sucesión de subconjuntos de* X*. Entonces*

(a) $(\underline{A}_n)_{n=1}^{\infty}$ *es* ***creciente*** *y* $(\overline{A}_n)_{n=1}^{\infty}$ *es* ***decreciente***.

(b) $\underline{A}_n \subseteq \overline{A}_m$ *para cualesquiera* $m, n \in \mathbb{N}$.

(c) $\liminf_{n\to\infty} A_n \subseteq \limsup_{n\to\infty} A_n$.

Prueba. (a) $(\underline{A}_n)_{n=1}^{\infty}$ es monótona creciente pues, para cualquier $n \in \mathbb{N}$,

$$\underline{A}_n = \bigcap_{k=n}^{\infty} A_k = A_n \cap \bigcap_{k=n+1}^{\infty} A_k \subseteq \bigcap_{k=n+1}^{\infty} A_k = \underline{A}_{n+1}.$$

Similarmente, para cualquier $n \in \mathbb{N}$,

$$\overline{A}_n = \bigcup_{k=n}^{\infty} A_k = A_n \cup \bigcup_{k=n+1}^{\infty} A_k \supseteq \bigcup_{k=n+1}^{\infty} A_k = \overline{A}_{n+1}$$

lo cual prueba que $(\overline{A}_n)_{n=1}^{\infty}$ es monótona decreciente.

(b) Sean $m, n \in \mathbb{N}$ y suponga que $n < m$. Entonces, por la primera parte,

$$\underline{A}_n \subseteq \underline{A}_m \subseteq A_m \subseteq \bigcup_{k=m}^{\infty} A_k = \overline{A}_m.$$

(c) Sea $n \in \mathbb{N}$. Por (2) tenemos que $\underline{A}_n \subseteq \overline{A}_m$ para $m = 1, 2, \ldots$ y, en consecuencia,

$$\overline{A}_n \subseteq \bigcap_{m=1}^{\infty} \overline{A}_m = \limsup_{n\to\infty} A_n,$$

Puesto que esta última desigualdad es válida para todo $n \in \mathbb{N}$, resulta que

$$\liminf_{n\to\infty} A = \bigcup_{n=1}^{\infty} \overline{A}_n \subseteq \limsup_{n\to\infty} A_n.$$

■

Definición 2.1.43. *Sea* $(A_n)_{n=1}^{\infty}$ *una sucesión arbitraria de subconjuntos de* X*. Si*

$$\limsup_{n\to\infty} A_n = \liminf_{n\to\infty} A_n,$$

entonces a éste conjunto lo denotaremos por $\lim_{n\to\infty} A_n$ *y lo llamaremos el* ***límite*** *de la sucesión* $(A_n)_{n=1}^{\infty}$.

Cuando la sucesión de conjuntos $(A_n)_{n=1}^{\infty}$ es monótona, los límites superiores e inferiores adoptan formas muy sencillas tal como lo muestra el siguiente resultado.

Teorema 2.1.44. *Sea $(A_n)_{n=1}^{\infty}$ una **sucesión monótona** de subconjuntos de X. Si $(A_n)_{n=1}^{\infty}$ es **decreciente** (respectivamente, **creciente**), entonces*

$$\lim_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} A_n \qquad \Big(\text{respectivamente}, \lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n\Big).$$

Prueba. Suponga que $(A_n)_{n=1}^{\infty}$ es decreciente. Entonces $A_{n-1}\cap A_n = A_n$ para todo $n \geq 2$ y, por consiguiente,

$$\underline{A}_n = A_n \cap \bigcap_{k=n+1}^{\infty} A_k = (A_{n-1}\cap A_n)\cap \bigcap_{k=n+1}^{\infty} A_k = \bigcap_{k=n-1}^{\infty} A_k = \cdots = \bigcap_{k=1}^{\infty} A_k.$$

De aquí se sigue que

$$\liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} \underline{A}_n = \bigcap_{k=1}^{\infty} A_k.$$

De modo similar, usando de nuevo el hecho de que $(A_n)_{n=1}^{\infty}$ es monótona decreciente, resulta que

$$\overline{A}_n = \bigcup_{k=n}^{\infty} A_k = A_n$$

y, en consecuencia,

$$\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \overline{A}_n = \bigcap_{n=1}^{\infty} A_n.$$

Esto termina la prueba de la primera parte. La prueba de la segunda parte es similar a la primera y, por lo tanto, se omite. ■

2.1.8. Series Absolutamente Convergentes y Familias Sumables

Sea $(x_n)_{n=1}^{\infty}$ una sucesión de números reales y para cada entero $n \geq 1$ considere la **suma parcial** $s_n = \sum_{i=1}^{n} x_i$. A la expresión

$$\sum_{n=1}^{\infty} x_n = \lim_{n\to\infty} s_n \tag{α_1}$$

se le llama la **serie asociada a la sucesión** $(x_n)_{n=1}^{\infty}$. Observe que dicho límite puede o no existir. Si él existe, diremos entonces que la **serie converge**, lo cual escribiremos como $\sum_{n=1}^{\infty} x_n < \infty$, o también como $\sum_{n=1}^{\infty} x_n = a$ para algún $a \in \mathbb{R}$. Por ejemplo, si los términos de la sucesión $(x_n)_{n=1}^{\infty}$ son **no-negativos** y si la sucesión $(s_n)_{n=1}^{\infty}$ está **acotada superiormente**, resulta del Teorema 2.1.23, que ella converge ya que es una sucesión creciente. En este caso se tiene que

$$\sum_{n=1}^{\infty} x_n := \lim_{n\to\infty} \sum_{i=1}^{n} x_i = \sup\left\{\sum_{i=1}^{n} x_i \,:\, n \in \mathbb{N}\right\}.$$

Si el límite en (α_1) no existe, se dirá entonces que la **serie diverge**.

Teorema 2.1.45. *Si la serie* $\sum_{n=1}^{\infty} x_n$ ***converge****, entonces* $\lim_{n\to\infty} x_n = 0$.

Prueba. Sea $a = \sum_{n=1}^{\infty} x_n$. si $(s_n)_{n=1}^{\infty}$ es la sucesión de las sumas parciales de $(x_n)_{n=1}^{\infty}$, entonces $x_n = s_n - s_{n-1}$ para todo $n \geq 2$ y por lo tanto,

$$\lim_{n\to\infty} x_n \;=\; \lim_{n\to\infty} s_n \;-\; \lim_{n\to\infty} s_{n-1} \;=\; a - a \;=\; 0.$$

La prueba es completa. ∎

Es fácil establece que si $\sum_{n=1}^{\infty} x_n$ y $\sum_{n=1}^{\infty} y_n$ son series convergentes, entonces:

(a) $\sum_{n=1}^{\infty}(x_n + y_n)$ converge y

$$\sum_{n=1}^{\infty}(x_n + y_n) \;=\; \sum_{n=1}^{\infty} x_n + \sum_{n=1}^{\infty} y_n.$$

(b) $\sum_{n=1}^{\infty} a\, x_n$ converge para cualquier $a \in \mathbb{R}$ y

$$\sum_{n=1}^{\infty} a\, x_n \;=\; a \sum_{n=1}^{\infty} x_n.$$

(c) Si $x_n \leq y_n$ para todo $n \in \mathbb{N}$, entonces

$$\sum_{n=1}^{\infty} x_n \;\leq\; \sum_{n=1}^{\infty} y_n.$$

El siguiente criterio, conocido como el Criterio de Cauchy para Series, es una elegante, útil y práctica forma de caracterizar serie s convergentes a través de la rigurosa forma del $\varepsilon - N$.

Teorema 2.1.46 (Criterio de Cauchy para Series). *Las siguientes condiciones son equivalentes para una serie* $\sum_{n=1}^{\infty} x_n$.

(1) *La* ***serie converge****.*

(2) *Para cada* $\varepsilon > 0$, *existe un* $N \in \mathbb{N}$ *tal que*

$$\left| \sum_{k=n+1}^{\infty} x_k \right| \;<\; \varepsilon \quad \textit{para todo } n \geq N.$$

(3) *Para cada* $\varepsilon > 0$, *existe un* $N \in \mathbb{N}$ *tal que*

$$\left| \sum_{k=n+1}^{m} x_k \right| \;<\; \varepsilon \quad \textit{para todo } m > n \geq N.$$

Prueba. $(1) \Rightarrow (2)$. Suponga que la serie converge, digamos $\sum_{n=1}^{\infty} x_n = a$ para algún $a \in \mathbb{R}$ y sea $(s_n)_{n=1}^{\infty}$ la sucesión de las sumas parciales de $(x_n)_{n=1}^{\infty}$. Entonces, para cada $\varepsilon > 0$, existe un entero positivo N tal que

$$|a - s_n| < \varepsilon \quad \text{para todo } n \geq N.$$

Puesto que

$$a - s_n = \lim_{m\to\infty} s_m - s_n = \lim_{m\to\infty} \sum_{k=n+1}^{m} x_k = \sum_{k=n+1}^{\infty} x_k$$

se obtiene que $(1) \Rightarrow (2)$.
$(2) \Rightarrow (3)$. Si (2) se cumple, entonces existe un $N \in \mathbb{N}$ tal que

$$\left|\sum_{k=n+1}^{\infty} x_k\right| < \varepsilon \quad \text{para todo } n \geq N.$$

Si $m > n \geq N$, entonces

$$\left|\sum_{k=n+1}^{m} x_k\right| \leq \left|\left|\sum_{k=n+1}^{\infty} x_k\right| - \left|\sum_{k=m+1}^{\infty} x_k\right|\right| \leq \max\left\{\left|\sum_{k=n+1}^{\infty} x_k\right|, \left|\sum_{k=m+1}^{\infty} x_k\right|\right\} < \varepsilon$$

y (3) se cumple.
$(3) \Rightarrow (1)$. Suponga que (3) es verdadero. Entonces

$$|s_m - s_n| = \left|\sum_{k=n+1}^{m} x_k\right| < \varepsilon \quad \text{para todo } m > n \geq N$$

lo cual prueba que la sucesión $(s_n)_{n=1}^{\infty}$ es de Cauchy en $\mathbb{R}$. La completitud de $\mathbb{R}$ nos garantiza que la serie converge. ■

2.1.9. Caracterizando Series Absolutamente Convergentes

Definición 2.1.47. *Sea $(x_n)_{n=1}^{\infty}$ una sucesión en $\mathbb{R}$. La serie $\sum_{n=1}^{\infty} x_n$ se dice que es* ***absolutamente convergente*** *si la serie $\sum_{n=1}^{\infty} |x_n|$ converge.*

¿Cuál es la relación entre una serie absolutamente convergente y la serie original? El siguiente resultado, válido en cualquier espacio de Banach, establece que toda serie absolutamente convergente, converge.

Teorema 2.1.48. *Sea $(x_n)_{n=1}^{\infty}$ una sucesión en $\mathbb{R}$. Si $\sum_{n=1}^{\infty} |x_n|$* ***converge****, entonces también* ***converge*** *la serie $\sum_{n=1}^{\infty} x_n$.*

Prueba. Suponga que $\sum_{n=1}^{\infty} |x_n|$ converge. Puesto que

$$0 \leq x_n + |x_n| \leq 2|x_n|$$

se sigue que la serie $\sum_{n=1}^{\infty}(x_n + |x_n|)$ converge y, en consecuencia, por la propiedad de la suma de dos series, se tiene que

$$\sum_{n=1}^{\infty} x_n = \sum_{n=1}^{\infty}(x_n + |x_n|) - \sum_{n=1}^{\infty} |x_n|$$

converge. ■

El recíproco del resultado anterior no es, en general, válido. Por ejemplo, es bien conocido que la serie $\sum_{n=1}^{\infty}(-1)^n/n$ converge, pero ella no es absolutamente convergente.

Definición 2.1.49. *Un **reordenamiento** de una serie $\sum_{n=1}^{\infty} x_n$ es otra serie de la forma $\sum_{n=1}^{\infty} x_{\pi(n)}$, donde $\pi : \mathbb{N} \to \mathbb{N}$ es cualquier biyección, es decir, una permutación de $\mathbb{N}$.*

Observe que un reordenamiento de $\sum_{n=1}^{\infty} x_n$ es otra serie con los mismos términos pero dispuesto en un orden diferente. Un concepto que es de importancia fundamental en la Teoría de Series y donde el orden de los términos no altera la convergencia de la serie es el siguiente.

Definición 2.1.50. *Una serie $\sum_{n=1}^{\infty} x_n$ se dice que es **incondicionalmente convergente** si para cualquier permutación $\pi : \mathbb{N} \to \mathbb{N}$, la serie $\sum_{n=1}^{\infty} x_{\pi(n)}$ converge y*

$$\sum_{n=1}^{\infty} x_n = \sum_{n=1}^{\infty} x_{\pi(n)}.$$

Es decir, una serie $\sum_{n=1}^{\infty} x_n$ es incondicionalmente convergente si no importa cómo reordenes los términos de la sucesión $(x_n)_{n=1}^{\infty}$ siempre obtienes una serie convergente que, además, nunca cambia su valor. Observe que, tomando la permutación identidad, vemos que $\sum_{n=1}^{\infty} x_n$ converge. Uno de los ejemplos más simple de una serie incondicionalmente convergente se obtiene cuando se considera cualquier serie convergente $\sum_{n=1}^{\infty} x_n$ cuyos términos son no-negativos. Éste hecho es consecuencia de un resultado sorprendente demostrado por Weierstrass y Riemann. La implicación (1) implica (2) en el siguiente resultado se debe a Weierstrass y la otra a Riemann.

Teorema 2.1.51 (Weierstrass-Riemann). *Sea $(x_n)_{n=1}^{\infty}$ una **sucesión de números reales**. Las siguientes condiciones son equivalentes:*

*(1) $\sum_{n=1}^{\infty} x_n$ es **absolutamente convergente**.*

*(2) $\sum_{n=1}^{\infty} x_n$ es **incondicionalmente convergente**.*

Prueba. (1) ⇒ (2). Suponga que $\sum_{n=1}^{\infty} |x_n| < \infty$ y sea π una permutación de $\mathbb{N}$. Hagamos

$$M = \sum_{n=1}^{\infty} |x_n| \qquad \text{y} \qquad m_n = \sup\{\pi(k) : 1 \leq k \leq n\}, \; n = 1, 2, \ldots$$

Para cada $n \in \mathbb{N}$ se verifica:

$$S_n = \sum_{k=1}^{n} |x_{\pi(k)}| \leq \sum_{k=1}^{m_n} \leq M$$

Este hecho nos muestra que la sucesión de términos no-negativos $(S_n)_{n=1}^{\infty}$ es creciente y acotada y, por lo tanto, converge. Además,

$$\sum_{n=1}^{\infty} |x_{\pi(n)}| \leq M = \sum_{n=1}^{\infty} |x_n|.$$

Si se considera la permutación inversa π^{-1} de π e invirtiendo los papeles de las series $\sum_{n=1}^{\infty}|x_n|$ y $\sum_{n=1}^{\infty}|x_{\pi(n)}|$ se obtiene la desigualdad:

$$\sum_{n=1}^{\infty}|x_n| \;\leq\; \sum_{n=1}^{\infty}|x_{\pi(n)}|,$$

de donde

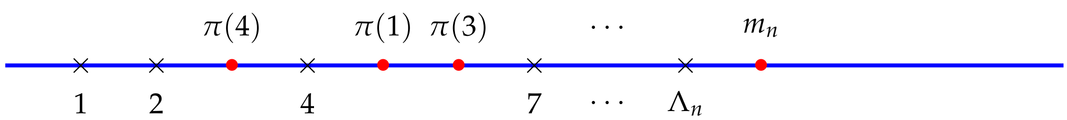
$$\sum_{n=1}^{\infty}|x_n| \;=\; \sum_{n=1}^{\infty}|x_{\pi(n)}| \;=\; M. \qquad (\alpha_2)$$

Falta verificar que

$$\sum_{n=1}^{\infty}x_n \;=\; \sum_{n=1}^{\infty}x_{\pi(n)}.$$

Para demostrar esto, considere, para cada $n \in \mathbb{N}$, el conjunto

$$\Lambda_n \;=\; \{p \in \mathbb{N} \,:\, p < m_n,\; p \notin \{\pi(1), \ldots, \pi(n)\}\}.$$



Observe que

$$\left|\sum_{k=1}^{m_n}x_k - \sum_{k=1}^{n}x_{\pi(k)}\right| \;=\; \left|\sum_{k\in\Lambda_n}x_k\right| \;\leq\; \sum_{k\in\Lambda_n}|x_k|$$

y

$$\sum_{k\in\Lambda_n}|x_k| \;=\; \sum_{k=1}^{m_n}|x_k| - \sum_{k=1}^{n}|x_{\pi(k)}|.$$

Se sigue entonces de (α_2) que $\sum_{k=1}^{m_n}|x_k|$ y $\sum_{k=1}^{n}|x_{\pi(k)}|$ tienen el mismo límite M cuando $n \to +\infty$. Luego, $\sum_{k=1}^{m_n}x_k - \sum_{k=1}^{n}x_{\pi(k)}$ tiende a cero y, se verifica que

$$\sum_{n=1}^{\infty}x_n \;=\; \sum_{n=1}^{\infty}x_{\pi(n)}.$$

$(2) \Rightarrow (1)$. Sea $\sum_{n=1}^{\infty}x_n$ una serie incondicionalmente convergente y suponga, para generar una contradicción, que ella no es absolutamente convergente, es decir, $\sum_{n=1}^{\infty}|x_n| = +\infty$. Lo que vamos a hacer es construir una permutación π sobre $\mathbb{N}$ tal que la serie $\sum_{n=1}^{\infty}x_{\pi(n)}$ diverja, lográndose, de este modo, una contradicción.

Para cada $n \in \mathbb{N}$, sean

$$p_n = \frac{|x_n| + x_n}{2} \qquad \text{y} \qquad q_n = \frac{|x_n| - x_n}{2}.$$

Observe que, para cualquier $n \in \mathbb{N}$,

$$p_n \geq 0, \quad q_n \geq 0, \quad p_n - q_n = x_n \quad \text{y} \quad p_n + q_n = |x_n|. \tag{α_3}$$

Sin perder generalidad, asumiremos que $p_n, q_n > 0$ para todo $n \geq 1$. Siendo $(p_n)_{n=1}^{\infty}$ y $(q_n)_{n=1}^{\infty}$ subsucesiones de $(x_n)_{n=1}^{\infty}$, existen sucesiones estrictamente crecientes en $\mathbb{N}$:

$$i_1 < i_2 < \cdots < i_n < \cdots \qquad \text{y} \qquad j_1 < j_2 < \cdots < j_n < \cdots$$

tales que $p_n = x_{i_n}$ y $q_n = x_{j_n}$ para todo $n \geq 1$. Observe que gracias a (α_3) se tiene que

$$\mathbb{N} = \{i_n : n \in \mathbb{N}\} \cup \{j_n : n \in \mathbb{N}\} \qquad \text{y} \qquad \{i_n : n \in \mathbb{N}\} \cap \{j_n : n \in \mathbb{N}\} = \varnothing.$$

La construcción de π será efectuada en dos pasos, el primero de los cuales es el siguiente:

(a_1) *Las series* $\displaystyle\sum_{n=1}^{\infty} p_n$ *y* $\displaystyle\sum_{n=1}^{\infty} q_n$ *son* ***ambas divergentes***.

En efecto, si ambas series fueran convergentes, entonces $\sum_{n=1}^{\infty} p_n + \sum_{n=1}^{\infty} q_n = \sum_{n=1}^{\infty} |x_n|$ sería convergente, lo que resulta contrario a nuestra suposición. Suponga ahora que una de las series converge y, para fijar idea, suponga que $\sum_{n=1}^{\infty} q_n < +\infty$. Puesto que la serie $\sum_{n=1}^{\infty} x_n$ converge, por hipótesis, resulta que $\sum_{n=1}^{\infty} x_n + \sum_{n=1}^{\infty} q_n$ también converge. Sin embargo, por la primera parte sabemos que $\sum_{n=1}^{\infty} p_n$ diverge y entonces se obtiene la siguiente contradicción

$$+\infty > \sum_{n=1}^{\infty} x_n + \sum_{n=1}^{\infty} q_n = \sum_{n=1}^{\infty} p_n = +\infty.$$

(a_2) *Existen subsucesiones estrictamente crecientes* $(m_n)_{n=1}^{\infty}$ *y* $(k_n)_{n=1}^{\infty}$ *de* $\mathbb{N}$ *tales que la serie*

$$p_1 + \cdots + p_{m_1} - \big(q_1 + \cdots + q_{k_1}\big) + p_{m_1+1} + \cdots + p_{m_2} - \big(q_{k_1+1} + \cdots + q_{k_2}\big) + \cdots \tag{$*$}$$

no converge. Ésta serie, como se demuestra a continuación, resulta ser un reordenamiento de $\sum_{n=1}^{\infty} x_n$ lo que producirá la contradicción que andamos buscando.

Fijemos un par de números reales α, β con $1 < \alpha < \beta$ y escoja sucesiones en $\mathbb{R}$, digamos $(\alpha_n)_{n=1}^{\infty}$ y $(\beta_n)_{n=1}^{\infty}$ con $\alpha_n < \beta_n$, $n = 1, 2, \ldots$, tales que

$$\lim_{n\to\infty} \alpha_n = \alpha \qquad \text{y} \qquad \lim_{n\to\infty} \beta_n = \beta.$$

Para demostrar (a_2) usaremos, en primer lugar, el hecho de que $\sum_{n=1}^{\infty} p_n = +\infty$, para elegir el entero positivo más pequeño, llamémoslo m_1, tal que

$$p_1 + \cdots + p_{m_1} > \beta_1 \qquad \text{y} \qquad p_1 + \cdots + p_{m_1-1} \leq \beta_1$$

De esto se sigue que

$$\beta_1 < p_1 + \cdots + p_{m_1} = \big(p_1 + \cdots + p_{m_1-1}\big) + p_{m_1} \leq \beta_1 + p_{m_1}.$$

Definiendo $S_1^p = p_1 + \cdots + p_{m_1}$, resulta que

$$\big|S_1^p - \beta_1\big| \leq p_{m_1}.$$

Vamos a definir los primeros m_1 valores de nuestra permutación π declarando que:

$$\pi(1) = i_1, \quad \pi(2) = i_2, \quad \dots, \quad \pi(m_1) = i_{m_1}.$$

Similarmente, siendo divergente la serie $\sum_{n=1}^{\infty} q_n$, existe un $k_1 \in \mathbb{N}$, que de nuevo elegiremos como el entero positivo más pequeño, de modo tal que

$$q_1 + \cdots + q_{k_1} > p_1 + \cdots + p_{m_1} - \alpha_1,$$

y

$$q_1 + \cdots + q_{k_1 - 1} \leq p_1 + \cdots + p_{m_1} - \alpha_1.$$

De allí que

$$\alpha_1 - q_{k_1} \leq p_1 + \cdots + p_{m_1} - \big(q_1 + \cdots + q_{k_1 - 1}\big) - q_{k_1} < \alpha_1$$

y, por lo tanto,

$$\big|S_1^q - \alpha_1\big| \leq q_{k_1},$$

donde hemos puesto $S_1^q = p_1 + \cdots + p_{m_1} - \big(q_1 + \cdots + q_{k_1}\big)$. Los siguientes valores de π, comenzando desde $m_1 + 1$ hasta k_1, se obtienen definiendo

$$\pi(m_1 + 1) = j_1, \quad \pi(m_1 + 2) = j_2, \ \dots, \ \pi(k_1) = j_{k_1}.$$

Usando de nuevo el hecho de que las series $\sum_{n=1}^{\infty} p_n$ y $\sum_{n=1}^{\infty} q_n$ son divergentes podemos, como antes, escoger los enteros positivos más pequeños, digamos m_2 y k_2, con $k_2 > k_1$ y $m_2 > m_1$, tales que

$$p_1 + \cdots + p_{m_1} + p_{m_1+1} + \cdots + p_{m_2} > q_1 + \cdots + q_{k_1} + \beta_2,$$

y

$$q_1 + \cdots + q_{k_1} + q_{k_1+1} + \cdots + q_{k_2} > p_1 + \cdots + p_{m_1} + p_{m_1+1} + \cdots + p_{m_2} - \alpha_2.$$

De esto se desprende, por las elecciones de m_2 y k_2, que si definimos

$$S_2^p = p_1 + \cdots + p_{m_1} - \big(q_1 + \cdots + q_{k_1}\big) + p_{m_1+1} + \cdots + p_{m_2} > \beta_2$$

y

$$S_2^q = p_1 + \cdots + p_{m_1} - \big(q_1 + \cdots + q_{k_1}\big) + p_{m_1+1} + \cdots + p_{m_2} - \big(q_{k_1+1} + \cdots + q_{k_2}\big) < \alpha_2,$$

entonces,

$$\big|S_2^p - \beta_2\big| \leq p_{m_2} \qquad \text{y} \qquad \big|S_2^q - \alpha_2\big| \leq q_{k_2}.$$

Hagamos

$$\pi(k_1 + 1) = i_{m_1+1}, \ \dots, \ \pi(m_2) = i_{m_2}, \quad \pi(m_2 + 1) = j_{k_1+1}, \ \dots, \ \pi(k_2) = j_{k_2}.$$

El procedimiento anterior, que se puede llevar a cabo indefinidamente gracias al hecho (a_1), culmina con la obtención de las dos sucesiones $(m_n)_{n=1}^{\infty}$ y $(k_n)_{n=1}^{\infty}$ de $\mathbb{N}$ tal que la serie $(*)$ es, por construcción, un reordenamiento de $\sum_{n=1}^{\infty} x_n$. Veamos ahora que dicha serie diverge. En efecto, para cada $n \in \mathbb{N}$, sean

$$\begin{aligned}
S_n^p &= p_1 + \cdots + p_{m_1} - \big(q_1 + \cdots + q_{k_1}\big) + p_{m_1+1} + \cdots + p_{m_2} - \big(q_{k_1+1} + \cdots + q_{k_2}\big) + \cdots + \\
&\quad\ p_{m_{n-1}+1} + \cdots + p_{m_n}, \\
S_n^q &= p_1 + \cdots + p_{m_1} - \big(q_1 + \cdots + q_{k_1}\big) + p_{m_1+1} + \cdots + p_{m_2} - \big(q_{k_1+1} + \cdots + q_{k_2}\big) + \cdots + \\
&\quad\ p_{m_{n-1}+1} + \cdots + p_{m_n} - \big(q_{k_{n-1}+1} + \cdots + q_{k_n}\big).
\end{aligned}$$

las sumas parciales de la serie $(*)$. Resulta entonces, por el procedimiento antes descrito, que para cada $n \in \mathbb{N}$,

$$\left|S_n^p - \beta_n\right| \leq p_{m_n} \qquad \text{y} \qquad \left|S_n^q - \alpha_n\right| \leq q_{k_n}.$$

Ahora bien, como la serie $\sum_{n=1}^{\infty} x_n$ es convergente, entonces $\lim_{n\to\infty} x_n = 0$, de donde se sigue que $\lim_{n\to\infty} p_n = 0 = \lim_{n\to\infty} q_n$. Por esto,

$$\lim_{n\to\infty} S_n^p = \beta \qquad \text{y} \qquad \lim_{n\to\infty} S_n^q = \alpha.$$

Puesto que $\alpha \neq \beta$, se concluye que la serie $(*)$ no converge contradiciendo, como habíamos afirmado, nuestra hipótesis. La prueba es completa. ■

2.1.10. Familias Sumables

Del Teorema de Weierstrass-Riemann se concluye que el orden de los términos de una serie que no es absolutamente convergente afecta tanto a su convergencia así como a su suma. En vista de esto, es razonable formularse la siguiente pregunta: ¿es posible tener una definición alternativa para la suma de una serie donde el orden de los términos no sea importante? La respuesta, la cual es afirmativa, se desarrolla a través de una teoría de sumas sin orden.

Siguiendo la idea de la convergencia de series, intentaremos darle significado a expresiones del tipo $\sum_{i\in I} x_i$, donde $(x_i)_{i\in I}$ es una familia arbitraria no-numerable de números reales. Para alcanzar ese objetivo aprovecharemos la existencia del supremo en $\overline{\mathbb{R}}$ de cualquier subconjunto A de $\mathbb{R}$.

En lo que sigue, supondremos que I es un conjunto infinito, el cual puede ser numerable o no-numerable y, como antes, denotemos por $\mathcal{P}_{\text{fin}}(I)$ la colección de todas las partes finitas de I, esto es,

$$\mathcal{P}_{\text{fin}}(I) = \big\{F \subseteq I : F \text{ es finito}\big\}.$$

Suponga, en primer lugar, que $(x_i)_{i\in I}$ es una familia de *números reales no-negativos*. Para cada $F \in \mathcal{P}_{\text{fin}}(I)$, defina

$$s(F) = \begin{cases} \displaystyle\sum_{i\in F} x_i & \text{si } F \neq \varnothing \\ 0 & \text{si } F = \varnothing. \end{cases}$$

Observe que la aplicación $s : \mathcal{P}_{\text{fin}}(I) \to [0, +\infty]$ es creciente en el siguiente sentido: si $F, G \in \mathcal{P}_{\text{fin}}(I)$, entonces

$$F \subseteq G \quad \Rightarrow \quad s(F) \leq s(G).$$

Más aun, si $F \cap G = \varnothing$, entonces $s(F \cup G) = s(F) + s(G)$.

Definición 2.1.52. *Sea $(x_i)_{i\in I}$ una familia de números reales no-negativos. Si el conjunto de números no-negativos $\{s(F) : F \in \mathcal{P}_{\text{fin}}(I)\}$ está* ***acotado superiormente****, entonces diremos que la familia $(x_i)_{i\in I}$ es* ***sumable*** *y escribiremos*

$$\sum_{i\in I} x_i = \sup\big\{s(F) : F \in \mathcal{P}_{\text{fin}}(I)\big\}.$$

Si la familia $\{s(F) \ : \ F \in \mathcal{P}_{\text{fin}}(I)\}$, en la definición anterior, no está acotada superiormente, pondremos

$$\sum_{i \in I} x_i \ = \ +\infty.$$

En el siguiente resultado veremos que cuando $I = \mathbb{N}$ las nociones de serie convergente y de familia sumable coinciden.

Teorema 2.1.53. *Sea $(x_n)_{n \in \mathbb{N}}$ una sucesión de números reales **no-negativos**. Son equivalentes:*

*(1) La familia $(x_n)_{n \in \mathbb{N}}$ es **sumable**.*

*(2) La serie $\sum_{n=1}^{\infty} x_n$ **converge**.*

Prueba. $(1) \Rightarrow (2)$. Suponga que $(x_n)_{n \in \mathbb{N}}$ es sumable y sea M tal que

$$\sum_{n \in \mathbb{N}} x_n \ = \ \sup\{s(F) \ : \ F \in \mathcal{P}_{\text{fin}}(\mathbb{N})\} \ = \ M.$$

Claramente la sucesión $\left(\sum_{i=1}^{n} x_n\right)_{n=1}^{\infty}$ es creciente y acotada por M. Se sigue del Teorema 2.1.23 que ella converge. Esto prueba (2).

$(2) \Rightarrow (1)$ Suponga que $\sum_{n=1}^{\infty} x_n$ converge y sea $\sum_{n=1}^{\infty} x_n = M$ para algún $M \geq 0$. Considere cualquier $F \in \mathcal{P}_{\text{fin}}(\mathbb{N})$ y sea $n_F = \max\{n : n \in F\}$. Tenemos entonces que

$$F \subseteq \{1, 2, \ldots, n_F\}$$

y, por lo tanto, se verifica que

$$\sum_{i \in F} x_i \ \leq \ \sum_{i=1}^{n_F} x_i \ \leq \ M.$$

De esto se sigue que $\sup\left\{s(F) \ : \ F \in \mathcal{P}_{\text{fin}}(\mathbb{N})\right\} \leq M$ por lo que la familia $(x_n)_{n \in \mathbb{N}}$ es sumable. La prueba es completa. ■

Para poder extender la noción de familia sumable a cualquier colección $(x_i)_{i \in I}$ de números reales, no necesariamente no-negativos, debemos verificar lo siguiente.

Corolario 2.1.54. *Sea $(x_i)_{i \in I}$ una familia **sumable** de números reales no-negativos. Entonces, para cada $\varepsilon > 0$, existe un conjunto $F_\varepsilon \in \mathcal{P}_{\text{fin}}(I)$ con la siguiente propiedad:*

$$\left| x \ - \ \sum_{i \in F} x_i \right| \ < \ \varepsilon$$

para cualquier conjunto finito $F \supseteq F_\varepsilon$.

Prueba. Pongamos $x = \sum_{i \in I} x_i$ y sea $\varepsilon > 0$. De las propiedades del supremo, existe un conjunto finito $F_\varepsilon \subseteq I$ tal que

$$x \ - \ \varepsilon \ = \ \sum_{i \in I} x_i \ - \ \varepsilon \ < \ \sum_{i \in F_\varepsilon} x_i.$$

Ahora bien, en virtud de que la aplicación s es creciente, se tiene que $s(F_\varepsilon) \subseteq s(F)$ para cualquier $F \in \mathcal{P}_{\text{fin}}(I)$ con $F \supseteq F_\varepsilon$ y, por consiguiente,

$$x \ - \ \varepsilon \ < \ \sum_{i \in F_\varepsilon} x_i \ \leq \ \sum_{i \in F} x_i \ < \ x \ + \ \varepsilon.$$

De esto se sigue que

$$\left| x - \sum_{i\in F} x_i \right| < \varepsilon \tag{1}$$

para cualquier conjunto finito $F \supseteq F_\varepsilon$. Fin de la prueba. ■

El siguiente resultado expresa una condición similar a la Condición de Cauchy para series convergentes.

Teorema 2.1.55 (Criterio de Cauchy). *Sea $(x_i)_{i\in I}$ una familia **sumable** de números reales no-negativos. Entonces, para cada $\varepsilon > 0$, existe un conjunto $F_\varepsilon \in \mathcal{P}_{\rm fin}(I)$ con la siguiente propiedad: para cualquier $G \in \mathcal{P}_{\rm fin}(I)$ con $G \cap F_\varepsilon = \varnothing$ se cumple que*

$$\sum_{i\in G} x_i < \varepsilon.$$

Prueba. Sea $\varepsilon > 0$ y use el corolario anterior para hallar un conjunto $F_\varepsilon \in \mathcal{P}_{\rm fin}(I)$ tal que

$$\left| x - \sum_{i\in F} x_i \right| < \varepsilon$$

para cualquier conjunto finito $F \supseteq F_\varepsilon$, donde $x = \sum_{i\in I} x_i$. Seleccione ahora cualquier conjunto $G \in \mathcal{P}_{\rm fin}(I)$ con $G \cap F_\varepsilon = \varnothing$. Entonces $F = G \cup F_\varepsilon \supseteq F_\varepsilon$ y, por lo anterior,

$$\begin{aligned}
\sum_{i\in G} x_i &= \sum_{i\in G\cup F_\varepsilon} x_i - \sum_{i\in F_\varepsilon} x_i \\
&= \sum_{i\in G\cup F_\varepsilon} x_i - x - \left(\sum_{i\in F_\varepsilon} x_i - x \right) \\
&\leq \left| \sum_{i\in G\cup F_\varepsilon} x_i - x \right| + \left| \sum_{i\in F_\varepsilon} x_i - x \right| \\
&< \varepsilon + \varepsilon = 2\varepsilon.
\end{aligned}$$

La prueba es completa. ■

Un aspecto importante que hay que destacar del Criterio de Cauchy anterior y que era de esperarse es que casi-todos los términos de una familia sumable son ceros, en otras palabras:

Corolario 2.1.56. *Si $(x_i)_{i\in I}$ una **familia sumable** de números reales **no-negativos**, entonces el conjunto*

$$I^* = \{ i \in I : x_i > 0 \}$$

*es **a lo más numerable**.*

Prueba. Si el conjunto I es numerable, no hay nada que demostrar. Suponga entonces que I es no-numerable y por cada $n \in \mathbb{N}$, considere el conjunto

$$I_n = \left\{ i \in I : x_i \geq \frac{1}{n} \right\}.$$

Afirmamos que $I^* = \bigcup_{n=1}^{\infty} I_n$. Para ver esto, observe, en primer lugar, que $\bigcup_{n=1}^{\infty} I_n \subseteq I^*$. Para demostrar la otra inclusión, sea $i \in I^*$. Entonces $x_i > 0$ y se sigue del Principio de Arquímedes que existe un entero $m \geq 1$ tal que $x_i > 1/m$ y, por lo tanto, $i \in I_m \subseteq \bigcup_{n=1}^{\infty} I_n$. Veamos ahora que cada uno de los conjuntos I_n es finito. En efecto, como la familia $(x_i)_{i\in I}$ es sumable, por el Criterio de Cauchy, Teorema 2.1.55, resulta que, por cada entero $n \geq 1$, existe un $F_n \in \mathcal{P}_{\text{fin}}(I)$ tal que

$$i \notin F_n \quad \Rightarrow \quad x_i < \frac{1}{n}.$$

(Aquí hemos tomado $G = \{i\}$ de modo que $G \cap F_n = \varnothing$). Por consiguiente, $I_n \subseteq F_n$ para $n = 1, 2, \ldots$. Esto prueba que I_n es finito y, en consecuencia, I^* es a lo más numerable. ■

Estamos listo para dar, de forma general, una definición de suma para cualquier familia $(x_i)_{i\in I}$ de números reales no necesariamente no-negativos.

Definición 2.1.57. *Una familia $(x_i)_{i\in I}$ de números reales se llama **sumable** si, dado $\varepsilon > 0$, existe un conjunto $F_\varepsilon \in \mathcal{P}_{\text{fin}}(I)$ con la siguiente propiedad: para cualquier conjunto finito $F \supseteq F_\varepsilon$*

$$\left| x - \sum_{i\in F} x_i \right| < \varepsilon.$$

Como antes, escribiremos $x = \sum_{i\in I} x_i$ si la familia $(x_i)_{i\in I}$ es sumable. Una pequeña aclaratoria es necesaria con respecto a la definición de una familia sumable $(x_i)_{i\in\mathbb{N}}$, cuando los x_i son números reales arbitrarios. En este caso, no es verdad que la noción de serie convergente coincide con la de familia sumable. En la literatura sobre series, la noción de familia sumable es equivalente a la de serie incondicionalmente convergente, un concepto que como ya hemos visto resulta ser más fuerte que la convergencia usual de series.

Finalizamos esta sección con esta otra observación. Suponga que $(x_i)_{i\in I}$ es una familia arbitraria de números reales tal que el conjunto

$$\left\{ \sum_{i\in F} x_i \;:\; F \in \mathcal{P}_{\text{fin}}(I) \right\}$$

está acotado, digamos, por una constante $M > 0$. Por cada $F \in \mathcal{P}_{\text{fin}}(I)$, considere los subconjuntos

$$F^+ = \{i \in F \,:\, x_i \geq 0\} \qquad \text{y} \qquad F^- = \{i \in F \,:\, x_i < 0\}.$$

Entonces

$$\sum_{i\in F} |x_i| = \sum_{i\in F^+} x_i - \sum_{i\in F^-} x_i = \left| \sum_{i\in F^+} x_i \right| + \left| \sum_{i\in F^-} x_i \right| \leq 2M.$$

Esto prueba que el conjunto de números no-negativos

$$\left\{ \sum_{i\in F} |x_i| \;:\; F \in \mathcal{P}_{\text{fin}}(I) \right\}$$

está acotado y, en consecuencia, por la primera parte, la familia $(|x_i|)_{i\in I}$ es sumable. En este caso definimos

$$\sum_{i\in I} |x_i| = \sup\left\{ \sum_{i\in F} |x_i| \;:\; F \in \mathcal{P}_{\text{fin}}(I) \right\}.$$

La noción de familia sumable se puede llevar a cabo para sucesiones dobles en $\overline{\mathbb{R}}$ del modo siguiente: suponga que para cada $(m,n) \in \mathbb{N} \times \mathbb{N}$, $x_{(m,n)} \in \overline{\mathbb{R}}^+$ y defina

$$S = \sup \left\{ \sum_{(m,n)\in F} x_{(m,n)} : F \in \mathcal{P}_{\text{fin}}(\mathbb{N} \times \mathbb{N}) \right\}.$$

Teorema 2.1.58. *Con la notación anterior, se tiene que*

$$\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} x_{(m,n)} = S = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} x_{(m,n)}.$$

Si, además, $\varphi : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ *es una aplicación biyectiva, entonces*

$$S = \sum_{k=1}^{\infty} x_{\varphi(k)}.$$

Prueba. Sean $M, N \in \mathbb{N}$ arbitrarios y observe que

$$\sum_{m=1}^{M} \sum_{n=1}^{N} x_{(m,n)} = \sum_{n=1}^{N} \sum_{m=1}^{M} x_{(m,n)} \leq S.$$

De esto se sigue que

$$\sum_{m=1}^{M} \sum_{n=1}^{\infty} x_{(m,n)} = \lim_{N\to\infty} \sum_{m=1}^{M} \sum_{n=1}^{N} x_{(m,n)} \leq S$$

y similarmente,

$$\sum_{n=1}^{N} \sum_{m=1}^{\infty} x_{(m,n)} = \lim_{M\to\infty} \sum_{n=1}^{N} \sum_{m=1}^{M} x_{(m,n)} \leq S.$$

Tomando límite sobre M y N, respectivamente, en las expresiones anteriores se obtiene que

$$\begin{aligned} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} x_{(m,n)} &= \lim_{N\to\infty} \lim_{N\to\infty} \sum_{m=1}^{M} \sum_{n=1}^{N} x_{(m,n)} \\ &= \lim_{M\to\infty} \lim_{M\to\infty} \sum_{n=1}^{N} \sum_{m=1}^{M} x_{(m,n)} = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} x_{(m,n)} \leq S. \end{aligned}$$

Para demostrar la otra desigualdad, tome $F \in \mathcal{P}_{\text{fin}}(\mathbb{N} \times \mathbb{N})$. Entonces

$$\sum_{(m,n)\in F} x_{(m,n)} \leq \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} x_{(m,n)}$$

y, por lo tanto,

$$S \leq \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} x_{(m,n)}.$$

Esto termina la prueba. ■

En la práctica se suele sustituir el término $x_{(m,n)}$ por $x_{m,n}$ para cada $m, n \in \mathbb{N}$.

2.2. Espacios Topológicos

Los intervalos abiertos son las piezas fundamentales en el desarrollo de un nuevo estudio de las propiedades de $\mathbb{R}$ llamado la *estructura topológica* de $\mathbb{R}$. Por ejemplo, $\mathbb{R}$ disfruta de una importantísima propiedad denominada **propiedad de Hausdorff** la cual establece que: *para cualesquiera dos números* $x, y \in \mathbb{R}$ *con* $x \neq y$, *existen intervalos abiertos* G_x *y* G_y *tales que*

$$x \in G_x, \qquad y \in G_y \qquad\qquad \text{y} \qquad\qquad G_x \cap G_y \;=\; \varnothing.$$

En efecto, basta tomar $r = |x - y|/2$ y definir

$$G_x \;=\; (x - r,\, x + r) \qquad\qquad \text{y} \qquad\qquad G_y \;=\; (y - r,\, y + r)$$

para comprobar que $G_x \cap G_y = \varnothing$. En este caso también se dice que $\mathbb{R}$ es un **espacio de Hausdorff**.

Definición 2.2.1. *Un conjunto* $G \subseteq \mathbb{R}$ *es* ***abierto*** *si, para cada* $x \in G$, *existe un* $\delta > 0$ *tal que*

$$(x - \delta,\, x + \delta) \;\subseteq\; G.$$

Denotemos por $\mathcal{O}_{\mathbb{R}}$ *la colección de todos los* ***subconjuntos abiertos*** *de* $\mathbb{R}$. Es un hecho ya establecido que la familia $\mathcal{G}$ cumple con las siguientes propiedades:

(a) $\varnothing, \mathbb{R} \in \mathcal{O}_{\mathbb{R}}$,

(b) si $\{G_\alpha : \alpha \in J\}$ es una subcolección de $\mathcal{O}$, entonces $\bigcup_{\alpha \in J} G_\alpha \in \mathcal{O}_{\mathbb{R}}$, y

(c) si $\{G_1, \ldots, G_n\}$ es cualquier subcolección finita de $\mathcal{O}_{\mathbb{R}}$, entonces $\bigcap_{i=1}^{n} G_i \in \mathcal{O}_{\mathbb{R}}$.

En general, se tiene el siguiente concepto fundamental.

Definición 2.2.2. *Sea X un conjunto no vacío y suponga que $\mathcal{T}$ es una colección no vacía de subconjuntos de X. Diremos que $\mathcal{T}$ es una* ***topología*** *sobre X siempre que dicha familia cumpla con las siguientes propiedades:*

(a_0) *$\varnothing$, $X \in \mathcal{T}$,*

(b_0) *si $\{G_\alpha : \alpha \in J\}$ es cualquier colección de elementos de $\mathcal{T}$, entonces $\bigcup_{\alpha \in J} G_\alpha \in \mathcal{T}$, y*

(c_0) *si para cualquier $n \in \mathbb{N}$, $G_1, \ldots, G_n \in \mathcal{T}$, entonces $\bigcap_{i=1}^{n} G_i \in \mathcal{T}$.*

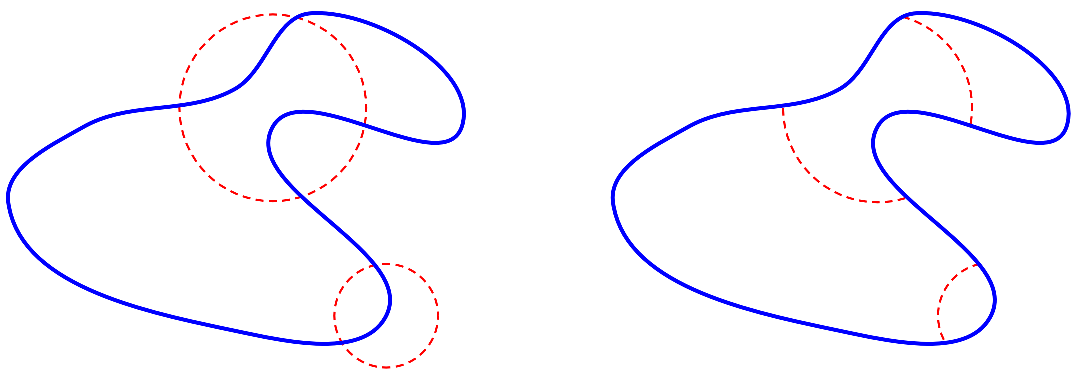
Un **espacio topológico** es un par $(X, \mathcal{T})$, donde X es un conjunto no vacío y $\mathcal{T}$ es una topología sobre X. Los elementos de cualquier topología $\mathcal{T}$ sobre un conjunto X son llamados **conjuntos** $\mathcal{T}$**-abiertos** o simplemente **abiertos**. Una subcolección $\mathcal{B} \subseteq \mathcal{T}$ se llama una **base** de $\mathcal{T}$ si para cualquier conjunto abierto $U \subseteq X$, existe una familia $(V_\alpha)_{\alpha \in I}$ incluida en $\mathcal{B}$ tal que $U \subseteq \bigcup_{\alpha \in I} V_\alpha$. Un espacio topológico $(X, \mathcal{T})$ se dice que es 2º **numerable** si existe una base de $\mathcal{T}$ que es numerable. *Todos los espacios topológicos considerados en estas notas se supondrán ser* ***Hausdorff***, es decir, cualesquiera sean $x, y \in X$ con $x \neq y$, existen abiertos disjuntos U, V en X tales que $x \in U$ y $y \in V$.

Observe que $\mathscr{P}(X)$ es claramente una topología sobre X, llamada la **topología discreta**. De igual forma, $\{\varnothing, X\}$ también es otra topología sobre X conocida como la **topología trivial**. Ambas topología son extremas, es decir, cualquier otra topología $\mathcal{J}$ sobre X verifica que $\{\varnothing, X\} \subseteq \mathcal{J} \subseteq \mathscr{P}(X)$.

Sea $(X, \mathcal{T})$ un espacio topológico. Cualquier subconjunto no vacío Y de X puede ser considerado en sí mismo como un espacio topológico definiendo la topología $\mathcal{T}_Y$ sobre Y del modo siguiente: $\mathcal{T}_Y := \{G \cap Y : G \in \mathcal{T}\}$, esto es, si $V \subseteq Y$, entonces

$$V \text{ es abierto en } Y \text{ si, y sólo si, existe } G \in \mathcal{T} \text{ tal que } V = G \cap Y.$$

En este caso se dice que $(Y, \mathcal{T}_Y)$ es un **subespacio** de $(X, \mathcal{T})$ y a $\mathcal{T}_Y$ se le llama la **topología inducida** por $\mathcal{T}$.



Por ejemplo, si $Y = [a, b]$, entonces para cada $x \in [a, b]$ y $\delta > 0$,

$$V(x, \delta) = [a, b] \cap (x - \delta, x + \delta)$$

es abierto en $[a, b]$. En particular, si δ es tal que $a + \delta < b$, entonces $V(a, \delta) = [a, a + \delta)$ es abierto en $[a, b]$.

Sea $(X, \mathcal{T})$ un espacio topológico y sea $x \in X$. Un subconjunto V de X se llama un **entorno abierto**, o simplemente, un **entorno** del punto x, si existe un conjunto abierto G tal que $x \in G \subseteq V$. Un subconjunto F de X se dice que es un **conjunto cerrado** si $X \setminus F$ es un conjunto abierto. Se sigue de las propiedades de los conjuntos abiertos y las leyes de Morgan que:

(a_0') $\varnothing$ y X son conjuntos cerrados,

(b_0') si $\{F_\alpha : \alpha \in J\}$ es cualquier colección de subconjuntos cerrados de X, entonces $\bigcap_{\alpha \in J} F_\alpha$ es cerrado, y

(c_0') si para cualquier $k \in \mathbb{N}$, $F_1, \ldots, F_k$ son subconjuntos cerrados de X, entonces $\bigcup_{i=1}^{k} F_i$ también es cerrado.

Sea F un subconjunto no vacío de X. Diremos que F es un **conjunto** F_σ, o algunas veces, un conjunto de tipo F_σ, si existe una sucesión $(F_n)_{n=1}^{\infty}$ de subconjuntos cerrados en X tal que

$$F = \bigcup_{n=1}^{\infty} F_n.$$

La familia de todos los ***subconjuntos*** F_σ *será denotada por* $\mathcal{F}_\sigma$. El complemento de un conjunto F_σ lo llamaremos un **conjunto** G_δ, o simplemente, un conjunto de tipo G_δ, es decir, un conjunto G es un G_δ si existe una sucesión $(G_n)_{n=1}^\infty$ de subconjuntos abiertos en X tal que

$$G = \bigcap_{n=1}^{\infty} G_n.$$

La familia de todos los ***subconjuntos*** G_δ *será denotada por* $\mathcal{G}_\delta$. Un conjunto que simultáneamente se puede representar tanto como un F_σ así como un G_δ será llamado un **conjunto ambiguo**, es decir, A es ambiguo si $A \in \mathcal{F}_\sigma \cap \mathcal{G}_\delta$.

Ejemplo 2.2.1. (1) *$\mathbb{Q}$ es un F_σ, mientras que los números irracionales $\mathbb{R} \setminus \mathbb{Q}$, es un G_δ-denso.* ¿Es $\mathbb{Q}$ un conjunto ambiguo? Más adelante veremos, como consecuencia del Teorema de Categoría de Baire, que a $\mathbb{Q}$ le está negada la posibilidad de poder expresarse como un G_δ.

(2) *Si $F \subseteq \mathbb{R}$ es* ***cerrado****, entonces F es un G_δ. En particular, F es ambiguo.*

Prueba. Para cada $n \in \mathbb{N}$, sea

$$G_n = \bigcup_{x\in F} (x - 1/n,\, x + 1/n).$$

Como cada G_n es abierto y $F \subseteq G_n$ para todo n, resulta que $F \subseteq \bigcap_{n=1}^\infty G_n$. Para demostrar la otra inclusión, tomemos $y \in \bigcap_{n=1}^\infty G_n$. Entonces $y \in G_n$ para todo $n \in \mathbb{N}$. Fijado un n, existe un $x \in F$ tal que $y \in (x - 1/n,\, x + 1/n)$ lo cual dice que $y \in \overline{F}$ y como F es cerrado, entonces $y \in F = \overline{F}$. Esto prueba que $\bigcap_{n=1}^\infty G_n \subseteq F$ y termina la demostración de (2). ■

(3) *Cualquier* ***conjunto abierto*** *G en $\mathbb{R}$ es un F_σ.* En particular, G es ambiguo pues trivialmente es un G_δ.

Observe que para todo $a, b \in \mathbb{R}$ con $a < b$,

$$[a, b) = \bigcap_{n=1}^{\infty} (a - 1/n,\, b) \in \mathcal{G}_\delta \qquad y \qquad (a, b) = \bigcup_{n=1}^{\infty} [a + 1/n,\, b) \in \mathcal{F}_\sigma.$$

Sea $E \subseteq X$ y considere la colección

$$\mathcal{C}_E = \{T \subseteq X : E \subseteq T,\ T \text{ es cerrado}\}$$

Puesto X es cerrado, resulta que $\mathcal{C}_E \neq \varnothing$. Se sigue de (b_0') que el conjunto

$$\overline{F} = \bigcap_{T \in \mathcal{C}_E} T$$

es cerrado y, además, es el *conjunto cerrado más pequeño que contiene a* E. A $\overline{E}$ se le llama la clausura de E. Cualquier punto $x \in \overline{E}$ es llamado un **punto de clausura** de E. Es fácil ver que si $E \subseteq F \subseteq X$, entonces $\overline{E} \subseteq \overline{F}$.

Teorema 2.2.3. *$x \in \overline{E}$ si, y sólo si, $G \cap E \neq \varnothing$ para cualquier conjunto abierto G conteniendo a x.*

Prueba. Suponga que $x \in \overline{E}$ y sea G un conjunto abierto conteniendo a x. Veamos que la igualdad $G \cap E = \varnothing$ no puede ocurrir. Aceptando que $G \cap E = \varnothing$ y definiendo $F_0 = X \setminus G$, tendremos que F_0 es un conjunto cerrado conteniendo a E pero no a x, lo cual es imposible pues $x \in \overline{E} = \bigcap_{F \in \mathcal{F}} F \subseteq F_0$. Recíprocamente, suponga que $G \cap E \neq \varnothing$ para todo conjunto abierto G conteniendo a x pero que $x \notin \overline{E}$. Entonces $G = X \setminus \overline{E}$ es un conjunto abierto conteniendo a x pero disjunto de E. Esta contradicción establece el resultado. ■

Sea E un subconjunto de X. *La unión de todos los conjuntos abiertos contenidos en* E es llamado el **interior** de E y denotado por $\operatorname{int}(E)$. Observe que si E no contiene ningún subconjunto abierto no vacío, entonces $\operatorname{int}(E) = \varnothing$. Se sigue de nuestra definición que si $\operatorname{int}(E) \neq \varnothing$, entonces $\operatorname{int}(E)$ es *el conjunto abierto más grande contenido* en E. Por consiguiente, $x \in \operatorname{int}(E)$ si, y sólo si, $G \subseteq E$ para algún conjunto abierto G conteniendo a x. De esto se sigue que si $E \subseteq X$, entonces:

|| ► $$(\overline{E})^c = \operatorname{int}(E^c).$$

En efecto,

$$\begin{aligned} x \in (\overline{E})^c &\Leftrightarrow x \notin \overline{E} \\ &\Leftrightarrow G \cap E = \varnothing \text{ para algún conjunto abierto } G \text{ conteniendo a } x \\ &\Leftrightarrow G \subseteq E^c \text{ para algún conjunto abierto } G \text{ conteniendo a } x \\ &\Leftrightarrow x \in \operatorname{int}(E^c). \end{aligned}$$

Uno de los resultados importantes que caracteriza a los subconjuntos abiertos de $\mathbb{R}$, el cual fue demostrado por G. Cantor y que usaremos frecuentemente, es el siguiente.

Teorema 2.2.4 (Cantor: Caracterización de abiertos en $\mathbb{R}$). *Sea $G \subseteq \mathbb{R}$ **abierto no vacío**. Entonces existe una sucesión $(I_n)_{n=1}^{\infty}$ de **intervalos abiertos** y **disjuntos dos a dos** tal que $G = \bigcup_{n=1}^{\infty} I_n$.*

Prueba. Sea $x \in G$. Puesto G es abierto, existe un intervalo abierto J_x conteniendo a x tal que $J_x \subseteq G$. Considere ahora la familia

$$\mathcal{G}_x = \{J \subseteq G : J \text{ es un intervalo abierto conteniendo a } x\}.$$

Por la observación anterior el conjunto $\mathcal{G}_x$ es no vacío y, además, $x \in \bigcap_{J \in \mathcal{G}_x} J$. Se sigue del Corolario 2.1.17 que $I_x = \bigcup_{J \in \mathcal{G}_x} J$ es un intervalo que es abierto (por ser unión de conjuntos abiertos) y que, por definición, es *el más grande* de los intervalos abiertos que están contenidos en G y que contienen a x. Es claro que

$$G = \bigcup_{x \in G} I_x.$$

Veamos ahora que la familia $\mathfrak{G} = \{I_x : x \in G\}$ es disjunta. En efecto, sean $x, y \in G$ y suponga que $I_x \neq I_y$. Si fuese $I_x \cap I_y \neq \varnothing$, entonces $J_0 = I_x \cup I_y$ sería un intervalo abierto contenido en G y conteniendo a x ya que contiene a I_x, de donde resultaría que J_0 es más grande que I_x. Esto, por supuesto, contradice la definición de I_x y, en consecuencia, $I_x \cap I_y = \varnothing$ si $x \neq y$.

Si se selecciona uno, y sólo un número racional r_{I_x} en cada uno de los intervalos $I_x \in \mathfrak{G}$, tendremos que ningún otro intervalo abierto perteneciente a la familia $\mathfrak{G}$ y distinto de I_x puede contener a r_{I_x} y, por consiguiente, la aplicación $\varphi : \mathfrak{G} \to \mathbb{Q}$ dada por $\varphi(I_x) = r_{I_x}$ para todo $x \in G$ es inyectiva. Como $\mathbb{Q}$ es numerable, resulta entonces que $\mathfrak{G}$ también es numerable.

Finalmente, si $(I_n)_{n=1}^{\infty}$ es una enumeración de $\mathfrak{G}$, entonces se tendrá que $G = \bigcup_{n=1}^{\infty} I_n$ y de este modo termina la prueba. ■

Del último parágrafo en la demostración del resultado anterior se concluye que:

Corolario 2.2.5. *Sea $(G_\alpha)_{\alpha\in D}$ una colección de **intervalos abiertos**. Si $(G_\alpha)_{\alpha\in D}$ es **disjunta**, entonces D es a lo más numerable.*

Cada uno de los intervalos abiertos I_n que conforman a G en el Teorema 2.2.4 se llama una **componente conexa abierta**, o simplemente, una **componente conexa**, de G. En general, si K es un subconjunto cerrado de $[a,b]$, la sucesión de intervalos abiertos y disjuntos dos a dos que, según el Teorema 2.2.4, particionan a $G = [a,b] \setminus K$ se llaman **contiguos** o **adyacentes** a K.

Definición 2.2.6. *Un espacio topológico (X,τ) se llama **conexo** si no se puede representar como la unión de dos conjuntos abiertos no vacíos. Un subconjunto Y de X es conexo si él, como subespacio, es conexo.*

Es fácil ver que que un espacio topológico (X,τ) es conexo si, y sólo si, los únicos subconjuntos de X que son ambos abiertos y cerrados son $\varnothing$ y X. En efecto, suponga que $G \subseteq X$ es abierto y cerrado. Si $G \neq \varnothing$, entonces por ser G cerrado, resulta que $X \setminus G$ es un abierto no vacío tal que $X = G \cup (X \setminus G)$ contradiciendo el hecho de que X es conexo. La otra implicación es, por supuesto, inmediata.

Si $x \in X$, entonces la colección C_x constituida por todos los subconjuntos conexos de X que contienen a x se llama la **componente conexa** de x. Es fácil establecer que la familia de todas la componentes conexas de X, $\mathcal{C} = \{C_x : x \in X\}$, constituyen una partición de X, es decir, $\mathcal{C}$ es una familia disjunta cuya unión es X.

En el siguiente resultado se establece de modo definitivo cuáles son los subconjuntos de $\mathbb{R}$ que son conexos.

Teorema 2.2.7. *Un conjunto no vacío $S \subseteq \mathbb{R}$ es **conexo** si, y sólo si, S es un **intervalo** o consta de **un único punto**.*

Prueba. Suponga que S es conexo. Si S consta de único punto, no hay nada que probar. Asuma entonces que S posee más de un punto y no es un intervalo. Esto significa, según el Lema 2.1.16, que existen $x, y \in S$ con $x < y$ tal que el intervalo cerrado $[x,y]$ no está contenido en S, es decir, existe $z \in (x,y)$ tal que $z \notin S$. Si ahora consideramos los conjuntos

$$U = (-\infty, z) \cap S \qquad \text{y} \qquad V = (z, +\infty) \cap S$$

resulta que U y V son subconjuntos abiertos disjuntos incluidos en S tal que $S = U \cup V$. Esto, por supuesto, contradice el hecho de que S es conexo. La otra implicación es inmediata. ■

Definición 2.2.8. *Sea (X,τ) un espacio topológico y sea $D \subseteq X$. Se dice que D es **totalmente disconexo** si sus únicos subconjuntos conexos son los puntos.*

La definición anterior tiene una interpretación geométrica muy interesante: significa que entre dos puntos cualesquiera $x, y \in D$, siempre hay puntos que no pertenecen a D. Cuando $D = X$, decir que X es totalmente disconexo es equivalente a afirmar que cada conjunto de τ es simultáneamente abierto y cerrado. (Verifique esto!)

Uno puede afinar un poco más la conclusión del Teorema 2.2.4 para derivar un resultado similar pero sin exigir que la sucesión de intervalos $(I_n)_{n=1}^{\infty}$ sea disjunta. Recordemos antes que si I es un intervalo acotado con extremos a y b, $a < b$, el **interior** de I es el intervalo abierto (a, b).

Definición 2.2.9. *Dos intervalos I y J se dicen que son **casi-disjuntos**, o **no-superpuestos**, si sus interiores son disjuntos. Una colección $\mathfrak{F}$ de intervalos se llama **casi-disjunta** o **no-superpuesta**, si sus elementos son dos a dos casi-disjuntos.*

Las expresiones: **contiguos**, o **adyacentes**, se usan como sinónimos de casi-disjuntos o no-superpuestos. Observe que todo *intervalo* se puede dividir en una cantidad finita de subintervalos casi-disjuntos. En efecto, sean $a_1, \ldots, a_n$ puntos interiores en un intervalo I tal que $a_1 < a_2 < \cdots < a_n$. Si I está acotado, digamos $[a, b]$, entonces los intervalos $I_1 = [a, a_1]$, $I_2 = [a_1, a_2], \ldots, I_n = [a_n, b]$ son casi-disjuntos. Los otros casos son similares. El siguiente es una generalización del Teorema 2.2.4.

Lema 2.2.10 (Otra caracterización de abiertos). *Sea G un subconjunto **abierto no vacío** de $\mathbb{R}$. Entonces existe una familia **numerable casi-disjunta** $(K_n)_{n=1}^{\infty}$ de intervalos cerrados y acotados tal que*

$$G = \bigcup_{n=1}^{\infty} K_n.$$

Prueba. Para cada $n \in \mathbb{N}$, considere la colección $\mathcal{R}_n$ de los *intervalos diádicos* de longitud $1/2^n$ definida por

$$\mathcal{R}_n = \left\{ \left[\frac{k}{2^n}, \frac{k+1}{2^n} \right] : k \in \mathbb{Z} \right\}.$$

Las siguientes propiedades son fáciles de establecer:

(a) Cada colección $\mathcal{R}_n$ es casi-disjunta, $n \geq 1$.

(b) $\mathbb{R} = \bigcup_{K \in \mathcal{R}_n} K$ para cada $n \in \mathbb{N}$.

(c) Para cualesquiera $n, m \in \mathbb{N}$ con $n \geq m$, todo intervalo perteneciente a $\mathcal{R}_n$ está contenido en algún intervalo perteneciente a $\mathcal{R}_m$.

(d) Todo intervalo perteneciente a $\mathcal{R}_n$ tiene longitud $\frac{1}{2^n}$.

Nuestro objetivo es construir, inductivamente, una sucesión de familias muy especiales, digamos $(\mathcal{R}'_n)_{n=1}^{\infty}$, de modo que $\mathcal{R}'_n$ sea un subconjunto de $\mathcal{R}_n$, para cada $n \in \mathbb{N}$. Sea $\mathcal{R}'_1$ el conjunto de todos los intervalos $K \in \mathcal{R}_1$ tales que $K \subseteq G$. Suponga que hemos construido las familias $\mathcal{R}'_1, \mathcal{R}'_2, \ldots, \mathcal{R}'_n$ y defina $\mathcal{R}'_{n+1}$ como el conjunto de todos los intervalos pertenecientes a $\mathcal{R}_{n+1}$ que están contenidos en G y que tienen interior disjunto con todos los interiores de los intervalos pertenecientes a $\bigcup_{k=1}^{n} \mathcal{R}'_k$. Finalmente, definamos

$$\mathcal{R} = \bigcup_{n=1}^{\infty} \mathcal{R}'_n.$$

Observe que cada uno de los conjuntos $\mathcal{R}'_n$ es a lo más numerables y, en consecuencia, $\mathcal{R}$ *es numerable*. Suponga ahora que $K_1, K_2 \in \mathcal{R}$ son intervalos distintos, digamos $K_1 \in \mathcal{R}'_n$ y $K_2 \in \mathcal{R}'_m$

con $n \geq m$. Si $n > m$ entonces, por construcción, el interior de K_1 es disjunto del interior de cualquier intervalo perteneciente a $\bigcup_{k=1}^{n-1} \mathcal{R}'_k$, en particular, el interior de K_1 es disjunto del interior de K_2. Si $n = m$, se sigue de la propiedad (a), que los interiores de K_1 y K_2 son disjuntos. Esto prueba que *la colección numerable $\mathcal{R}$ es casi-disjunta*. Verifiquemos ahora que

$$G = \bigcup_{K \in \mathcal{R}} K.$$

Es claro, por construcción, que $\bigcup_{K \in \mathcal{R}} K \subseteq G$. Para demostrar la otra inclusión, sea $x \in G$. Como G es abierto, existe un $n \geq 1$ tal que todo intervalo abierto con centro en x y longitud $1/2^n$ está contenido en G. Se sigue de las propiedades (b) y (d) que existe un $K \in \mathcal{R}_n$ tal que $x \in K$ y, además, $K \subseteq G$. Puede ocurrir sólo una de las dos posibilidades siguientes: (1) que $K \in \mathcal{R}'_n \subseteq \mathcal{R}$ y, en consecuencia, $G \subseteq \bigcup_{K \in \mathcal{R}} K$; o, (2) que $K \notin \mathcal{R}'_n$. En este caso, existe un $m < n$ y un $K_1 \in \mathcal{R}'_m$ tal que los interiores de K y K_1 se intersectan. En vista de la propiedad (c), podemos hallar un intervalo $K_2 \in \mathcal{R}'_m$ conteniendo a K de modo que $K_1, K_2 \in \mathcal{R}'_m$ y, por consiguiente, sus interiores se intercepten. Esto último lo que indica es que $K_1 = K_2$, gracias a la propiedad (a) y, por lo tanto, $x \in K \subseteq K_1 = K_2 \in \mathcal{R}$. En cualquier caso, $x \in \bigcup_{K \in \mathcal{R}} K$ lo cual da por finalizada la prueba. ■

Teorema 2.2.11. *Sea E un subconjunto no vacío de $\mathbb{R}$. Son equivalentes:*

(a) $x \in \overline{E}$.

(b) Existe una sucesión $(x_n)_{n=1}^{\infty}$ en E tal que $\lim_{n\to\infty} x_n = x$.

Prueba. $(a) \Rightarrow (b)$ Sea $x \in \overline{E}$. Si $E_n = (x - 1/n, x + 1/n) \cap E$ para cada $n \in \mathbb{N}$, entonces $E_n \neq \varnothing$, y así, seleccionando un punto x_n de cada E_n, vemos que $\lim_{n\to\infty} x_n = x$.

$(b) \Rightarrow (a)$ Suponga que (b) se cumple y sea G un conjunto abierto conteniendo a x. Escojamos un $\varepsilon > 0$ de modo tal que $(x - \varepsilon, x + \varepsilon) \subseteq G$. Por otro lado, como $\lim_{n\to\infty} x_n = x$, para el ε seleccionado, existe un $N \in \mathbb{N}$ tal que $|x_n - x| < \varepsilon$ para todo $n \geq N$, es decir, $x_n \in (x - \varepsilon, x + \varepsilon)$ para todo $n \geq N$. Por esto, $x_n \in G \cap E$ para todo $n \geq N$, lo cual demuestra que $G \cap E \neq \varnothing$ y termina la prueba. ■

La siguiente definición es extremadamente útil e importante en muchas ramas del árbol de las matemáticas.

Definición 2.2.12. *Sea (X, τ) un espacio topológico y sea $D \subseteq X$. Diremos que D es **denso** en X si $X = \overline{D}$.*

Ejemplo 2.2.2. (1) *$\mathbb{Q}$ y $\mathbb{R} \setminus \mathbb{Q}$ son los ejemplos clásicos de subconjuntos densos en $\mathbb{R}$.*

(2) *Sea D el conjunto de los **racionales diádicos** en $\mathbb{R}$, es decir, si denotamos por*

$$D_n = \left\{ \frac{k}{2^n} \,:\, k \in \mathbb{Z} \right\},$$

*para cada $n \in \mathbb{N}$, entonces $D = \bigcup_{n=1}^{\infty} D_n$ es **denso** en $\mathbb{R}$.*

En efecto, sea $x \in \mathbb{R}$ y sea $\varepsilon > 0$. Sin perder generalidad, asumiremos que $x \in (0, 1]$. Use el Principio de Arquímedes para hallar un $n \in \mathbb{N}$ tal que $1/2^n < \varepsilon$. Para el n hallado, divida el intervalo $(0,1]$ en 2^n partes iguales y observe que $x \in ((k-1)/2^n, k/2^n]$ para algún entero $k = 1, \ldots, 2^n$. Resulta de esto que $|x - k/2^n| \leq 1/2^n < \varepsilon$ y concluye la demostración.

(3) *Sean D_1, D_2 subconjuntos de un espacio topológico (X, τ) con $D_1 \subseteq D_2$. Si D_1 es* ***denso*** *en X, entonces D_2 también lo es y $\overline{D}_1 = \overline{D}_2 = X$.*

Sea E un subconjunto no vacío de $\mathbb{R}$. Diremos que $x_0 \in \mathbb{R}$ es un **punto límite** de E si existe una sucesión $(x_n)_{n=1}^{\infty}$ en E que converge a x_0. En vista del Teorema 2.2.11, *los puntos límites y los puntos de clausura son exactamente los mismos*. Por otro lado, $x_0 \in \mathbb{R}$ es un **punto de acumulación** de E si $G \cap (E \setminus \{x_0\}) \neq \varnothing$ para todo conjunto abierto G conteniendo al punto x_0 o, de forma equivalente, si $\operatorname{card}(G \cap E) > 1$ para todo entorno G de x_0. Denotaremos por E' el conjunto de todos los puntos de acumulación de E. Observe que todo punto de acumulación es un punto límite. Es fácil establecer que si x_0 es un punto de acumulación de E, entonces existe una sucesión $(x_n)_{n=1}^{\infty}$ en E que converge a x_0 tal que $x_m \neq x_n$ para todo $m \neq n$. Por consiguiente, la escasa diferencia entre punto límite y punto de acumulación es que la existencia de una sucesión convergente puede, en el primer caso, ser la sucesión constante, mientras que en el segundo caso *siempre* se puede elegir una sucesión convergente cuyos términos sean distintos dos a dos. Un punto $x_0 \in E$ se dice que es **aislado** en E si $x_0 \notin E'$, es decir, si existe un entorno G de x_0 tal que $G \cap E = \{x_0\}$.

Corolario 2.2.13. *Sea $E \subseteq \mathbb{R}$. Son equivalentes:*

(1) *E es cerrado.*

(2) *$E' \subseteq E$.*

Prueba. (1) $\Rightarrow$ (2). Sea $x \in E'$. En particular, x es un punto límite de E, de donde se sigue, usando el Teorema 2.2.11, que $x \in E = \overline{E}$, pues E es cerrado.

(2) $\Rightarrow$ (1). Sea $x \in \overline{E}$ y suponga que $x \notin E$. Como $x \in \overline{E}$, se tiene que $G \cap E \neq \varnothing$ para cualquier entorno G de x. Pero como $x \notin E$, resulta que $G \cap (E \setminus \{x\}) \neq \varnothing$, es decir, $x \in E'$ y, así, por hipótesis, $x \in E$. ■

2.2.1. Espacios Métricos

Sea X un conjunto no vacío. Una **métrica** sobre X es una aplicación $d : X \times X \to [0, +\infty)$ que cumple con las siguientes propiedades:

(1) $0 \leq d(x,y) < +\infty$ para todo $x, y \in \mathbb{R}$.

(2) $d(x,y) = 0$ si, y sólo si, $x = y$.

(3) $d(x,y) = d(y,x)$ para todo $x, y \in \mathbb{R}$.

(4) $d(x,y) \leq d(x,z) + d(z,y)$ para todo $x, y, z \in \mathbb{R}$.

Al par (X, d) se le llama un **espacio métrico**. En ocasiones, en lugar del símbolo $d(x,y)$ escribiremos $\operatorname{dist}(x,y)$.

Una sucesión $(x_n)_{n=1}^{\infty}$ en el espacio métrico (X, d) se dice que **converge** a un punto x si

$$\lim_{n\to\infty} d(x_n, x) \;=\; 0.$$

La sucesión $(x_n)_{n=1}^{\infty}$ se dice que es de **Cauchy** si, dado $\varepsilon > 0$, existe un $N \in \mathbb{N}$ tal que $d(x_m, x_n) < \varepsilon$ para todo $m, n \geq N$. Diremos que (X, d) es un **espacio métrico completo** si toda sucesión de Cauchy en (X, d) converge.

Dos hechos importantes que hay que destacar en relación a las sucesiones de Cauchy, y cuyas pruebas se dejan a cargo del lector, son los siguientes:

(1) *Si $(x_n)_{n=1}^{\infty}$ es una sucesión de Cauchy en X, entonces se puede determinar la existencia una subsucesión $(n_k)_{k=1}^{\infty}$ de enteros positivos tal que $d(x_{n_k}, x_{n_{k+1}}) < 2^{-k}$ para todo $k \in \mathbb{N}$.*

(2) *Si $(x_n)_{n=1}^{\infty}$ es una sucesión de Cauchy y $(x_{n_j})_{j=1}^{\infty}$ es una subsucesión de $(x_n)_{n=1}^{\infty}$ que converge a algún punto $x \in X$, entonces la sucesión $(x_n)_{n=1}^{\infty}$ también converge y lo hace hacia el mismo punto x.*

Sea (X, d) un espacio métrico. Para cada $x \in X$ y cada $r > 0$ los conjuntos

$$\begin{aligned} U(x,r) &= \{y \in X : d(x,y) < r\}, \\ B(x,r) &= \{y \in X : d(x,y) \leq r\}, \\ S(x,r) &= \{y \in X : d(x,y) = r\} \end{aligned}$$

se llaman, respectivamente, una **bola abierta**, una **bola cerrada** y la **esfera** con centro en x y radio $r > 0$. Un conjunto $G \subseteq X$ se dice que es **abierto** en X si para cada $x \in G$, existe un $r > 0$ tal que la bola abierta $U(x, r) \subseteq G$. La colección de todos los subconjuntos abiertos de X será denotada por $\mathcal{J}_d$ y llamada la **d-topología** de X o la **topología generada** por d. Una subcolección $\mathcal{B}$ de $\mathcal{J}_d$ se dice que es una **base** para X si todo conjunto abierto en X es la unión de alguna subcolección de miembros de $\mathcal{B}$. Si la colección $\mathcal{B}$ es numerable, entonces diremos que X posee una **base numerable**.

Sea $A \subseteq X$. Es un ejercicio sencillo establecer que:

$$x \in \overline{A} \text{ si, y sólo si, existe una sucesión } (a_n)_{n=1}^{\infty} \text{ en } A \text{ tal que } \lim_{n\to\infty} d(a_n, x) = 0.$$

Definición 2.2.14. *Un espacio métrico (X, d) se dice que es **separable** si existe un conjunto $D \subseteq X$ que es numerable y denso en X.*

Teorema 2.2.15. *Todo **espacio métrico separable** (X, d) posee una **base numerable**.*

Prueba. Para ver esto, sea $D = \{x_n : n \in \mathbb{N}\}$ un conjunto numerable y denso de X, y considere la colección $\mathcal{B} = \{U(x_n, q) : n \in \mathbb{N}, q \in \mathbb{Q}\}$. Claramente $\mathcal{B}$ es numerable, de modo que sólo resta verificar que es una base para X. En efecto, sea G un subconjunto abierto de X y sea $y \in G$. Por definición, existe un $\varepsilon > 0$ tal que $U(y, \varepsilon) \subseteq G$. Puesto que D es denso en X, podemos elegir un $x_m^y \in D$ de modo que $0 < d(y, x_m^y) < \varepsilon/2$. Usemos ahora el hecho de $\mathbb{Q}$ es denso en $\mathbb{R}$ para hallar un $q_y \in \mathbb{Q}$ tal que $d(y, x_m^y) < q_y < \varepsilon/2$. De esto se sigue

$$y \in U(x_m^y, q_y) \subseteq U(y, \varepsilon) \subseteq G$$

y, en consecuencia, $G \subseteq \bigcup_{y \in G} U(x_m^y, q_y)$. Esto prueba que $\mathcal{B}$ es una base numerable para X. ■

Un espacio métrico (X, d) se dice que es **Polaco** si él es completo y separable. En general, cualquier espacio topológico (X, τ) para el cual exista una métrica completa d tal que la d-topología $\mathcal{J}_d$ coincida con τ es llamado un espacio Polaco. Por ejemplo, cualquier espacio compacto metrizable es un espacio Polaco. Una de las propiedades fundamentales de los espacios Polacos es que cualquier subespacio abierto, así como cada subespacio cerrado, en un espacio Polaco X siguen siendo Polacos.

Definición 2.2.16. *Sean A y B subconjuntos de (X, d). La **distancia** entre A y B se define por*

$$\operatorname{dist}(A, B) \;=\; \operatorname{ínf}\big\{d(a, b)\,:\, a \in A,\ b \in B\big\}.$$

Observe que si $A \cap B \neq \varnothing$, entonces $\operatorname{dist}(A, B) = 0$, es decir, la condición $\operatorname{dist}(A, B) = 0$ no implica, en general, que $A = B$. Más adelante veremos bajo qué condiciones sobre los conjuntos A y B se obtiene que $\operatorname{dist}(A, B) = 0$.

En particular, si $x \in X$ y A es un subconjunto no vacío de X, la **distancia** entre x y A, se define como

$$\operatorname{dist}(x, A) \;=\; \operatorname{dist}(\{x\}, A) \;=\; \operatorname{ínf}\,\big\{d(x, a)\,:\, a \in A\big\}.$$

Es una tarea fácil establecer las siguientes propiedades:

(a) $\operatorname{dist}(x, A) \;=\; \operatorname{dist}(x, \overline{A})$,

(b) $\operatorname{dist}(x, A) \;=\; 0$ si, y sólo si, $x \in \overline{A}$, y

(c) $\big|\operatorname{dist}(x, A) - \operatorname{dist}(y, A)\big| \;\leq\; d(x, y)$ cualesquiera sean $x, y \in X$.

Un conjunto $A \subseteq X$ se dice que es **acotado** si existe una bola cerrada que lo contiene. Si A es acotado, el **diámetro** de A, $\operatorname{diam}(A)$, se define mediante el número

$$\operatorname{diam}(A) \;=\; \sup\,\big\{d(a, b)\,:\, a, b \in A\big\}.$$

Pondremos $\operatorname{diam}(A) = \infty$, si el conjunto A no es acotado. Observe que si A es un subconjunto acotado de $\mathbb{R}$, entonces gracias al Corolario 2.1.13 se tiene que

$$\operatorname{diam}(A) \;=\; \sup A - \operatorname{ínf} A.$$

El siguiente resultado es una pieza fundamental para la demostración de varios resultados importantes en Análisis, entre ellos, el irrenunciable Teorema de Categoría de Baire.

Teorema 2.2.17 (Encaje de Cantor). *Sea $(F_n)_{n=1}^{\infty}$ una sucesión de **subconjuntos cerrados** no vacíos en un **espacio métrico completo** (X, d) tal que*

(a) $F_1 \supseteq F_2 \supseteq \cdots \supseteq F_n \supseteq \cdots$, y

(b) $\lim\limits_{n\to\infty} \operatorname{diam}(F_n) = 0$.

Entonces existe un único $x_0 \in X$ tal que $\bigcap_{n=1}^{\infty} F_n = \{x_0\}$.

Prueba. Por cada $n \in \mathbb{N}$, seleccionemos un único $x_n \in F_n$. Afirmamos que la sucesión $(x_n)_{n=1}^{\infty}$ es de Cauchy en X. En efecto, sea $\varepsilon > 0$ y usemos el hecho de $\lim_{n\to\infty} \operatorname{diam}(F_n) = 0$ para elegir un entero $N > 0$ tal que $\operatorname{diam}(F_N) < \varepsilon$. Como la sucesión $(F_n)_{n=1}^{\infty}$ es decreciente, se sigue que si $m, n \geq N$, entonces $d(x_n, x_m) < \varepsilon$. En efecto, como $x_n \in F_n \subseteq F_N$ y también $x_m \in F_m \subseteq F_N$, resulta que $d(x_n, x_m) \leq \operatorname{diam}(F_N) < \varepsilon$. Por esto $(x_n)_{n=1}^{\infty}$ es de Cauchy y por la completitud de X ella converge a un $x_0 \in X$. Puesto que todos los términos de la sucesión $(x_n)_{n=1}^{\infty}$, salvo un número finito de ellos se quedan dentro de F_k para todo $k \in \mathbb{N}$, resulta que $x_0 \in \overline{F_k} = F_k$ para todo $k \in \mathbb{N}$. Por esto, $x_0 \in \bigcap_{k=1}^{\infty} F_k$. Para demostrar la otra inclusión, observe que como $\bigcap_{n=1}^{\infty} F_n \subseteq F_m$ para todo $m \in \mathbb{N}$, entonces la existencia de algún $y \in \bigcap_{n=1}^{\infty} F_n$ nos indicaría que $x_0, y \in F_m$ y, por consiguiente, $d(x_0, y) \leq \operatorname{diam}(F_m) \to 0$ cuando $m \to \infty$. Esto prueba que $y = x_0$ y termina la demostración. ■

Cantor demostró, como ya vimos, que $\mathbb{R}$ es no-numerable usando su incuestionable Método de la Diagonal. Otra forma de ver eso es aplicando el Teorema de Encaje de Cantor.

Corolario 2.2.18 (Cantor). *$\mathbb{R}$ es **no-numerable**.*

Prueba. Es suficiente demostrar, por (Nu_3), página 13, que $[0,1]$ es no-numerable. De nuevo, para fabricar una contradicción, suponga que $[0,1]$ es numerable y sea

$$x_1,\ x_2,\ \ldots,\ x_n,\ \ldots$$

una lista de *todos* los números en $[0,1]$. Nuestra tarea será construir una sucesión decreciente de intervalos cerrados $(I_n)_{n=1}^{\infty}$ en $[0,1]$ tal que

$$x_n \notin I_n \qquad \text{y} \qquad \ell(I_n) < 1/3^n \tag{1}$$

para todo $n \geq 1$. En efecto, divida $[0,1]$ en tres intervalos cerrados de igual longitud, es decir, $[0,1/3]$, $[1/3,2/3]$ y $[2/3,1]$. Es claro que x_1 no puede estar en los 3 intervalos. Sea I_1 uno de esos intervalos para el cual $x_1 \notin I_1$. Proceda, como antes, dividiendo el intervalo I_1 en tres subintervalos cerrados cada uno de longitud $1/3^2$. Por supuesto, uno de esos intervalos, al que denotaremos por I_2, no contiene a x_2. Si se continúa indefinidamente con este procedimiento se obtiene una sucesión decrecientes de intervalos compactos satisfaciendo (1). Del Teorema de Encaje de Cantor se sigue la existencia un único $x_0 \in [0,1]$ tal que

$$\bigcap_{n=1}^{\infty} I_n \;=\; \{x_0\}.$$

Puesto que $x_0 \in I_n$ para todo $n \geq 1$, se sigue de nuestra construcción que $x_0 \neq x_n$ para todo $n \geq 1$ obteniéndose, de este modo, una contradicción pues hemos encontrado un número en $[0,1]$, el susodicho x_0, que no está en lista. Por esto, $[0,1]$ es no-numerable. ■

Nota Adicional 2.2.2 El siguiente resultado es otra versión del Teorema de Encaje de Cantor en $\mathbb{R}$ que no requiere que los conjuntos sean cerrados:

Teorema 2.2.19 (Intervalos Encajados de Cantor). *Sea $(I_n)_{n=1}^{\infty}$ una sucesión de intervalos en $\mathbb{R}$. Suponga que cada I_n es **acotado** con extremos a_n y b_n y que:*

(a) $I_1 \supseteq I_2 \supseteq \cdots \supseteq I_n \supseteq \cdots$ y

(b) $\lim\limits_{n\to\infty} \ell(I_n) = 0$.

*Entonces existe un **único** $x_0 \in \mathbb{R}$ tal que $x_0 \in \bigcap_{n=1}^{\infty} \overline{I}_n$.*

Nótese que no se asume que los intervalos I_n sean cerrados, aunque la conclusión es que $a_n \leq x_0 \leq b_n$ para todo $n \geq 1$. Observe que esto último no significa que $x_0 \in I_n$ para todo $n \geq 1$ como se puede ver, por ejemplo, si elegimos $I_n = (0, 1/n)$ para todo $n \geq 1$. Por supuesto, la prueba es idéntica al Teorema 2.2.17 tomando $F_n = \overline{I}_n$ para cada $n \geq 1$.

Definición 2.2.20. *Un punto $x \in \mathbb{R}$ se dice que es un **punto de condensación** de un conjunto $E \subseteq \mathbb{R}$, si la intersección de E con cada entorno abierto V de x es un conjunto no-numerable o, de modo equivalente, si $\operatorname{card}(E \cap V) = \mathfrak{c}$ para todo entorno abierto V de x.*

Si denotamos por $P_{\mathfrak{c}}$ el conjunto de los puntos de condensación de E, resulta que

$$P_{\mathfrak{c}} \;\subseteq\; E' \;\subseteq\; \overline{E}$$

Teorema 2.2.21. *Si E es un **subconjunto no-numerable** de $\mathbb{R}$, entonces E contiene al menos un punto de condensación.*

Prueba. Puesto que $\mathbb{R} = \bigcup_{n=1}^{\infty}[-n, n]$, entonces

$$E = E \cap \mathbb{R} = \bigcup_{n=1}^{\infty} \big(E \cap [-n, n]\big)$$

y como E es no-numerable, debe existir algún N tal que $E \cap [-N, N]$ es no-numerable. Divida el intervalo $[-N, N]$ en dos subintervalos cerrados no-superpuestos de igual longitud. Uno de ellos, que denotaremos por F_1, tiene la propiedad de que $F_1 \cap E$ es no-numerable. De nuevo, divida el intervalo F_1 en dos subintervalos cerrados no-superpuestos de igual longitud y, como antes, la intersección de uno de ellos con E es no-numerable. Llamémoslo F_2. Si continuamos indefinidamente con este procedimiento obtenemos una sucesión $\big(F_n\big)_{n=1}^{\infty}$ de intervalos cerrados con las siguientes propiedades:

(1) $F_1 \supseteq F_2 \supseteq \cdots \supseteq F_n \supseteq \cdots$,

(2) $\lim\limits_{n\to\infty} \operatorname{diam}(F_n) = 0$ y

(3) $F_n \cap E$ es no-numerable para todo $n \geq 1$.

Se sigue del Teorema de Encaje de Cantor que $\bigcap_{n=1}^{\infty} F_n = \{x_0\}$. Claramente $x_0 \in E$ y como $\lim_{n\to\infty} \operatorname{diam}(F_n) = 0$, resulta que cualquier entorno V de x_0 contiene a uno de los F_m para un cierto m y, por consiguiente, $V \cap E \supseteq F_m \cap E$ es un conjunto no-numerable. Esto nos revela que x_0 es un punto de condensación de E y termina la prueba. ■

Definición 2.2.22. *Un subconjunto P de $\mathbb{R}$ se llama **perfecto** si él cerrado y todos sus puntos son puntos de acumulación.*

El siguiente resultado establece que todo subconjunto cerrado no-numerable de $\mathbb{R}$ se puede particionar en dos conjuntos, uno de ellos siendo perfecto y el otro a lo más numerable.

Teorema 2.2.23 (Cantor-Bendixson). *Sea E un subconjunto **cerrado no-numerable** de $\mathbb{R}$. Entonces E se puede escribir en la forma $E = P \cup N$, donde P es **perfecto** y N es **a lo más numerable**.*

Prueba. Defina $P = \big\{x \in E : x \text{ es punto de condensación de } E\big\}$ y sea $N = E \setminus P$. Observe que

$$x \in N \Leftrightarrow \text{existe un entorno } V_x \text{ de } x \text{ tal que } \operatorname{card}(V_x \cap E) \leq \aleph_0.$$

Veamos que N es a lo más numerable. En efecto, sea $\mathcal{V} = (V_n)_{n=1}^{\infty}$ una enumeración de todos los intervalos abiertos con extremos racionales. Sabemos que $\mathcal{V}$ es numerable. Para cada $x \in N$, seleccionemos un miembro $V_x \in \mathcal{V}$ tal que $x \in V_x$ y $V_x \cap E$ es a lo más numerable. Resulta entonces que la subcolección

$$\mathcal{V}_* = \big\{V_x \in \mathcal{V} : x \in V_x,\ \ V_x \cap E \text{ es a lo más numerable}\big\}$$

es numerable y se cumple que

$$N = \bigcup_{x\in N} V_x \cap E \subseteq \bigcup_{V_x \in \mathcal{V}_*} V_x \cap E,$$

es decir, N está incluido en una unión numerable de conjuntos a lo más numerables y, por tanto, N es a lo más numerable.

P es cerrado. Para demostrar esta afirmación observe que por el Teorema 2.2.21, $P \neq \varnothing$. Sea $x \in \overline{P}$ y sea G_x un entorno de x. Entonces $G_x \cap P \neq \varnothing$. Sea $z \in G_x \cap P$. Como G_x un entorno de $z \in P$, él contiene, por definición, una cantidad no-numerable de puntos de E, lo cual significa que $x \in P$ y, así, P es cerrado.

Veamos ahora que $P' = P$. En efecto, puesto que todo punto de condensación de E es claramente un punto de acumulación, tenemos que $P \subseteq P'$. Por otro lado, como P es cerrado, se sigue del Corolario 2.2.13 que $P' \subseteq P$. Esto prueba que P es perfecto y que $E = P \cup N$. ■

En el siguiente resultado se establece que uniones de familias arbitrarias de conjuntos abiertos en $\mathbb{R}$ siempre se pueden reducir a uniones numerables. Ese resultado constituye, por supuesto, una buena economía y revela, una vez más, el gran poder de los conjuntos abiertos.

Teorema 2.2.24 (Lindelöf). *Si $\mathcal{V} = \{V_\alpha : \alpha \in D\}$ es cualquier **colección de subconjuntos abiertos no vacíos** de $\mathbb{R}$, entonces existe una subcolección numerable $\{V_n \in \mathcal{V} : n \in \mathbb{N}\}$ de $\mathcal{V}$ tal que*

$$\bigcup_{\alpha \in D} V_\alpha = \bigcup_{n=1}^{\infty} V_n.$$

Prueba. Sea $\mathcal{J}$ la familia de todos los intervalos abiertos no-degenerados cuyos extremos son números racionales. Puesto que $\mathbb{Q}$ es numerable, resulta que la colección $\mathcal{J}$ también es numerable. Sea $(J_n)_{n=1}^{\infty}$ una enumeración de $\mathcal{J}$. Fijemos $\alpha \in D$. Como V_α es abierto, para cada $x \in V_\alpha$ podemos determinar un intervalo $J_{n_x} \in \mathcal{J}$ tal que $x \in J_{n_x} \subseteq V_\alpha$. Es claro que

$$\bigcup_{\alpha \in D} V_\alpha = \bigcup_{n_x \in \mathbb{N}} J_{n_x}.$$

Finalmente, por cada intervalo J_{n_x}, seleccionemos el correspondiente conjunto V_{n_x} tal que $J_{n_x} \subseteq V_{n_x}$. Entonces

$$\bigcup_{\alpha \in D} V_\alpha = \bigcup_{n_x \in \mathbb{N}} J_{n_x} \subseteq \bigcup_{n_x \in \mathbb{N}} V_{n_x} \subseteq \bigcup_{\alpha \in D} V_\alpha$$

y termina la prueba. ■

Una de las nociones topológicas más importantes y que se usa frecuentemente es la de compacidad. Sea K un subconjunto de un espacio topológico $(X, \mathcal{T})$.

Definición 2.2.25. *Una colección $\mathcal{V} = \{V_\alpha : \alpha \in D\}$ de subconjuntos de X se dice que es un **cubrimiento abierto** de K si cada V_α es un conjunto abierto y $K \subseteq \bigcup_{\alpha \in I} V_\alpha$.*

Si D_0 es un subconjunto de D y si la subcolección $\mathcal{V}_0 = \{V_\alpha : \alpha \in D_0\}$ cubre a K, entonces decimos que $\mathcal{V}_0$ es un **subcubrimiento** de K. Si D_0 es finito, diremos que $\mathcal{V}_0$ es un **subcubrimiento finito** de K. Los conjuntos compactos tiene la virtud de reducir a cubrimientos finitos cualquier cubrimiento infinito.

Definición 2.2.26. *Un subconjunto K de un espacio topológico $(X, \mathcal{T})$ se dice que es **compacto** si cualquier cubrimiento abierto $\mathcal{V} = \{V_\alpha : \alpha \in D\}$ de K se reduce a un subcubrimiento finito, es decir, existen $V_{\alpha_1}, \dots, V_{\alpha_n}$ en $\mathcal{V}$ tal que $K \subseteq \bigcup_{k=1}^{n} V_{\alpha_k}$.*

Lema 2.2.27. *Sea $(K_n)_{n=1}^{\infty}$ una sucesión **subconjuntos compactos** de un espacio topológico $(X, \mathcal{T})$. Entonces:*

(a) $K_1 \cup \cdots \cup K_n$ es compacto para cualquier $n \in \mathbb{N}$.

(b) $\bigcap_{n=1}^{\infty} K_n$ es compacto.

Prueba. (a). Sea $\mathcal{V} = \{V_\alpha : \alpha \in D\}$ un cubrimiento abierto de $K_1 \cup \cdots \cup K_n$. Como $\mathcal{V}$ es un cubrimiento abierto de cada K_j existen $V^j_{\alpha_1}, \ldots, V^j_{\alpha_{n_j}}$ en $\mathcal{V}$ tal que $K_j \subseteq \bigcup_{i=1}^{n_j} V^j_{\alpha_i}$. Luego

$$K_1 \cup \cdots \cup K_n \subseteq \bigcup_{j=1}^{n} \bigcup_{i=1}^{n_j} V^j_{\alpha_i}$$

y termina la prueba de (a). El lector puede consultar la demostración del Teorema 2.2.34, página 134, para una prueba sencilla de la parte (b). ■

Teorema 2.2.28. *Todo intervalo $[a, b]$, con $a, b \in \mathbb{R}$, $a < b$, es **compacto**.*

Prueba. Suponga que $[a, b]$ no es compacto. Esto significa que existe un cubrimiento abierto de $[a, b]$, digamos $\mathcal{V} = \{V_\alpha : \alpha \in D\}$, con la propiedad de que *ninguna subfamilia finita de $\mathcal{V}$ cubre a* $[a, b]$. Vamos ahora, con esa hipótesis, a fabricar una contradicción. Comencemos. Dividamos el intervalo $I_1 = [a, b]$ en dos subintervalos de igual longitud. Al menos uno de esos dos intervalos, que denotaremos por I_2, no puede ser cubierto por una subfamilia finita de $\mathcal{V}$. Dividamos de nuevo el intervalo I_2 en dos subintervalos de igual longitud y, como antes, uno de ellos, que denotaremos por I_3, no puede ser cubierto por una subfamilia finita de $\mathcal{V}$. Si repetimos indefinidamente el procedimiento anterior se obtiene una sucesión $(I_n)_{n=1}^{\infty}$ de intervalos cerrados con las siguientes propiedades:

(1) $I_1 \supseteq I_2 \supseteq \cdots \supseteq I_n \supseteq \cdots$,

(2) $\lim\limits_{n\to\infty} \operatorname{diam}(I_n) = \lim\limits_{n\to\infty} \dfrac{b-a}{2^n} = 0$, y

(3) *ninguna subfamilia finita de $\mathcal{V}$ cubre a ninguno de los I_n, $n = 1, 2, \ldots$*

El Teorema de Encaje de Cantor, Teorema 2.2.17, nos garantiza entonces la existencia de un único $x_0 \in \mathbb{R}$ tal que

$$\bigcap_{n=1}^{\infty} I_n = \{x_0\}.$$

Como $x_0 \in I_1 = [a, b] \subseteq \bigcup_{\alpha \in D} V_\alpha$, existe un $\alpha \in D$ tal que $x_0 \in V_\alpha$. Siendo V_α un conjunto abierto, existe un intervalo abierto $(x_0 - \varepsilon, x_0 + \varepsilon) \subseteq V_\alpha$ para algún $\varepsilon > 0$. Por otro lado, usando (2), podemos hallar un $N \in \mathbb{N}$ de modo tal que $(b-a)/N < \varepsilon$ y, en consecuencia, el intervalo $I_N \subseteq (x_0 - \varepsilon, x_0 + \varepsilon) \subseteq V_\alpha$. Esto último es incompatible con nuestra construcción de los intervalos I_n, ya que por (3), ninguno de tales intervalos puede ser cubierto por una subfamilia finita de $\mathcal{V}$. Esta contradicción establece que $[a, b]$ es compacto y termina la prueba. ■

Recordemos que si $K \subseteq \mathbb{R}$, entonces un subconjunto $V \subseteq K$ *es abierto en* K si, y sólo si, V es de la forma $V = G \cap K$ para algún conjunto abierto $G \subseteq \mathbb{R}$.

Teorema 2.2.29. *Sea $K \subseteq \mathbb{R}$ **compacto**. Entonces*

*(1) K es **cerrado** y **acotado**.*

*(2) Si $F \subseteq K$ es **cerrado**, entonces F es **compacto**.*

Prueba. (1). Puesto que $\mathbb{R} = \bigcup_{n=1}^{\infty}(-n, n)$, resulta entonces que

$$K = K \cap \mathbb{R} = \bigcup_{n=1}^{\infty} K \cap (-n, n)$$

y por lo dicho anteriormente, la colección $\{K \cap (-n, n) : n = 1, 2, \ldots\}$ es un cubrimiento abierto del compacto K y, por consiguiente, existen $n_1, \ldots, n_k$ en $\mathbb{N}$ tal que $K = \bigcup_{j=1}^{k} K \cap (-n_j, n_j)$. Si ahora definimos $N = \max\{n_1, \ldots, n_k\}$, vemos que $K \subseteq (-N, N)$ y, por lo tanto, K es acotado. Probemos ahora que K es cerrado. Para este fin, fijemos un $x_0 \in \mathbb{R} \setminus K$. Para cada $x \in K$, usemos el hecho de que $\mathbb{R}$ es Hausdorff para hallar conjuntos abiertos y disjuntos V_x y $V_x(x_0)$ de x y x_0 respectivamente. Observe que $\mathcal{V} = \{V_x : x \in K\}$ es un cubrimiento abierto de K con la propiedad de que $x_0 \notin V_x$ para todo $x \in K$. La compacidad de K permite reducir el cubrimiento anterior a un subcubrimiento finito, digamos, $\{V_{x_1}, \ldots, V_{x_n}\}$. Claramente $G = \bigcup_{i=1}^{n} V_{x_i}$ es abierto con $x_0 \in G \subseteq \mathbb{R} \setminus K$. Esto prueba que $\mathbb{R} \setminus K$ es abierto, y así, K es cerrado.

(2). Suponga que F es un subconjunto cerrado de K. Sea $\mathcal{V}$ un cubrimiento abierto de F. Como F es cerrado, entonces $\mathbb{R} \setminus K$ es abierto y, en consecuencia, $\mathcal{V} \cup (\mathbb{R} \setminus K)$ es un cubrimiento abierto de K. Por compacidad, existen $V_1, \ldots, V_n$ en $\mathcal{V}$ tal que $K \subseteq V_1 \cup \cdots \cup V_n \cup (\mathbb{R} \setminus K)$. Claramente $\{V_1, \ldots, V_n\}$ es un cubrimiento de F. ■

Teorema 2.2.30 (Heine-Borel). *Un subconjunto no vacío $K \subseteq \mathbb{R}$ es **compacto** si, y sólo si, K es **cerrado** y **acotado**.*

Prueba. Si K es compacto, entonces el Teorema 2.2.29 nos dice que K es cerrado y acotado. Recíprocamente, suponga que K es cerrado y acotado. Sean $a = \inf K$ y $b = \sup K$. Puesto que $K \subseteq [a, b]$ es, por hipótesis, un subconjunto cerrado del compacto $[a, b]$ entonces, por una nueva aplicación del Teorema 2.2.29, tenemos que K es compacto. ■

Una consecuencia de los Teoremas de Heine-Borel y de Bolzano-Weierstrass es el siguiente resultado que establece que *sucesiones son suficientes para caracterizar compacidad*, un resultado que también es válido para cualquier compacto viviendo en un espacio métrico arbitrario.

Teorema 2.2.31 (Weierstrass). *Sea K un subconjunto no vacío de $\mathbb{R}$. Las siguientes condiciones son equivalentes:*

*(1) K es **compacto**.*

*(2) K es **secuencialmente compacto**, es decir, cualquier sucesión en K posee una subsucesión que converge a algún punto de K.*

Prueba. (1) $\Rightarrow$ (2). Suponga que K es compacto. Por el Teorema de Heine-Borel, K es cerrado y acotado. Sea $(x_n)_{n=1}^{\infty}$ una sucesión en K. Como K es acotado, también lo es la sucesión $(x_n)_{n=1}^{\infty}$, y entonces, por el Teorema de Bolzano-Weierstrass, ella posee una subsucesión convergente a un punto de K pues K es cerrado.

(2) $\Rightarrow$ (1). Suponga que (2) se cumple. Veamos que K es cerrado y acotado. Para demostrar que K es cerrado, tomemos $x \in \overline{K}$. Entonces existe una sucesión $(x_n)_{n=1}^{\infty}$ en K tal que $x_n \to x$. Por hipótesis, existe una subsucesión de $(x_n)_{n=1}^{\infty}$ que converge a un punto $y \in K$ y, por supuesto, $x = y \in K$. Esto prueba que K es cerrado. Para verificar que K es acotado, suponga que no lo es. Esto significa que, para cada $n \in \mathbb{N}$, existe un $x_n \in K$ tal que $|x_n| > n$. Por hipótesis, existe una una subsucesión $(x_{n_k})_{k=1}^{\infty}$ de $(x_n)_{n=1}^{\infty}$ que converge a un punto $x_0 \in K$. Aquí viene la contradicción: como toda sucesión convergente es acotada, existe una constante $M > 0$ tal que $|x_{n_k}| \leq M$ para todo $k \in \mathbb{N}$, sin embargo, por construcción $|x_{n_k}| \geq n_k > k$ para todo $k \in \mathbb{N}$. Esto revela que $k \leq M$ para todo $k \in \mathbb{N}$, es decir, la sucesión de todos los números naturales está acotada lo cual es imposible. Por esto K es acotado y termina la prueba. ■

Las aplicaciones del Teorema de Weierstrass son amplias. He aquí una de ellas que usaremos un poco más adelante.

Teorema 2.2.32. *Si A es **cerrado**, B es **compacto** y $A \cap B = \varnothing$, entonces $\operatorname{dist}(A, B) > 0$.*

Prueba. Suponga que $\operatorname{dist}(A, B) = 0$. Entonces, por definición, para cada $n \in \mathbb{N}$, existen $a_n \in A$ y $b_n \in B$ tales que $|a_n - b_n| < 1/n$. Ahora bien, como B es compacto se sigue del Teorema de Weierstrass que existe una subsucesión $(b_{n_k})_{k=1}^{\infty}$ de $(b_n)_{n=1}^{\infty}$ tal que $b_{n_k} \to b \in B$. Por supuesto, la correspondiente subsucesión $(a_{n_k})_{k=1}^{\infty}$ de $(a_n)_{n=1}^{\infty}$ también converge y lo hace hacia b. Más aun, como A es cerrado, resulta que $b \in A$, de donde se obtiene la incuestionable contradicción $b \in A \cap B = \varnothing$. ■

Del Teorema 2.2.29 también se deduce que si $(K_\alpha)_{\alpha \in I}$ es una familia arbitraria de subconjuntos compactos, entonces $\bigcap_{\alpha \in I} K_\alpha$ es compacto, aunque dicha intersección puede ser vacía. Sin embargo, si la familia $(K_\alpha)_{\alpha \in I}$ es numerable, digamos $(K_n)_{n=1}^{\infty}$, y además decreciente, entonces $\bigcap_{n=1}^{\infty} K_n \neq \varnothing$. Lo anterior permite justificar la siguiente definición.

Definición 2.2.33. *Una familia $(F_\alpha)_{\alpha \in I}$ de subconjuntos de un espacio topológico $(X\,\mathcal{T})$ se dice que tiene la **propiedad de intersección finita** (**PIF**) si, para cada subconjunto finito $F \subseteq I$, se cumple que $\bigcap_{\alpha \in F} F_\alpha \neq \varnothing$.*

Una de las tantas caracterizaciones hermosas que poseen los espacios compactos de Hausdorff es la siguiente:

Teorema 2.2.34. *Sea K un subconjunto no vacío de un espacio topológico $(X, \mathcal{T})$. Son equivalentes:*

*(1) K es **compacto**.*

*(2) Para cualquier familia $(K_\alpha)_{\alpha \in D}$ de subconjuntos **cerrados** de K con la **propiedad de intersección finita** se cumple que*

$$\bigcap_{\alpha \in D} K_\alpha \neq \varnothing.$$

Prueba. (1) $\Rightarrow$ (2). Suponga que K es compacto y sea $(K_\alpha)_{\alpha \in D}$ una familia de subconjuntos cerrados de K con la propiedad de intersección finita. Si ocurriera que $\bigcap_{\alpha \in D} K_\alpha = \varnothing$, entonces, tomando $V_\alpha = \mathbb{R} \setminus K_\alpha$ para cada $\alpha \in D$, resultaría que la familia $(V_\alpha)_{\alpha \in D}$ sería un cubrimiento abierto de K del que, por la compacidad de K, se podría extraer un subcubrimiento finito, digamos $V_{\alpha_1}, \ldots, V_{\alpha_n}$. De esto se seguiría que $\bigcap_{k=1}^{n} K_{\alpha_k} = \varnothing$ lo cual es una contradicción. La otra implicación es más sencilla de probar y se deja a cargo del lector. ■

Una consecuencia inmediata del resultado anterior es el siguiente.

Corolario 2.2.35. *Si $(K_\alpha)_{\alpha\in I}$ es una familia **decreciente** de subconjuntos **compactos no vacíos** de un espacio topológico $(X, \mathcal{T})$, entonces $\bigcap_{\alpha\in I} K_\alpha$ es compacto y no vacío.*

Prueba. Puesto que cualquier intersección finita de los K_α es no vacía, se sigue del Teorema 2.2.34 que $K := \bigcap_{\alpha\in I} K_\alpha$ es no vacío y, además, como K es cerrado (por ser intersección de conjuntos cerrados, Teorema 2.2.29) y contenido en el compacto K_1 tenemos, por una nueva aplicación del Teorema 2.2.29, que K es compacto. ■

2.2.2. El Teorema de Categoría de Baire

Ya hemos visto, un resultado de George Cantor, que $\mathbb{R}$ es un conjunto no-numerable, lo que también es equivalente a afirmar que *ninguna colección numerable de puntos de* $\mathbb{R}$ puede ponerse en correspondencia uno a uno con todos los elementos de $\mathbb{R}$. René Baire extiende el resultado de Cantor al demostrar que *ninguna colección numerable de conjuntos de primera categoría* puede cubrir a $\mathbb{R}$. Esto se obtendrá como consecuencia de un extraordinario resultado de mayor alcance que el de Cantor llamado el Teorema de Categoría de Baire. Recordemos que conjunto D de un espacio métrico (X, d) es **denso** en X si $\overline{D} = X$. Notemos que ello significa que el conjunto cerrado más pequeño que contiene a D es precisamente X. En general, sean A y B subconjuntos de X y suponga que $A \subseteq B$. Diremos que A **denso en** B si $B \subseteq \overline{A}$. Una condición equivalente a la definición de densidad que no hace referencia a ningún punto del espacio y que se usa frecuentemente es el siguiente:

Teorema 2.2.36. *Sean (X, d) un espacio métrico y D un subconjunto no vacío de X. Entonces, D es **denso** en X si, y sólo si, para cada subconjunto abierto no vacío G de X, $G \cap D \neq \varnothing$.*

Prueba. Supongamos que D es denso en X y sea G un subconjunto abierto no vacío de X. Si se diera el caso que $G \cap D = \varnothing$, entonces $F = X \setminus G$ sería un conjunto cerrado conteniendo a D y, en consecuencia, $\overline{D} \subseteq F \neq X$, lo que contradice la densidad de D.

Recíprocamente, suponga que $G \cap D \neq \varnothing$ para cada subconjunto abierto no vacío G de X. De ocurrir $\overline{D} \neq X$, entonces $G := X \setminus \overline{D}$ sería un conjunto abierto no vacío que satisface $G \cap D = \varnothing$. Esta contradicción establece que $\overline{D} = X$. ■

Puesto que todo conjunto abierto no vacío G se puede expresar como una unión numerable y disjunta de intervalos abiertos no vacíos, resulta que para demostrar que un conjunto $D \subseteq \mathbb{R}$ es denso en $\mathbb{R}$ sólo basta con verificar que $D \cap I \neq \varnothing$ para cualquier intervalo abierto I.

El siguiente principio es fundamental en matemáticas.

‖ ▶ **Principio del Palomar o de Dirichlet.** *Si n palomas se distribuyen en m palomares y si $n > m$, entonces al menos uno de los palomares debe contener más de una paloma.*

Teorema 2.2.37 (Kronecker). *Sea $\xi \in \mathbb{R}$ un **número irracional**. Entonces el conjunto*

$$D_\xi = \{x \in \mathbb{R} : x = p + q\xi, \quad p, q \in \mathbb{Z}\}$$

*es **numerable** y **denso** en $\mathbb{R}$.*

Prueba. Claramente la aplicación $f : D_\xi \to \mathbb{N} \times \mathbb{N}$ definida por $f(p + q\xi) = (p, q)$ es inyectiva y como $\mathbb{N} \times \mathbb{N}$ es un conjunto numerable, se sigue del Teorema 1.2.6, página 13, que D_ξ es numerable.

Para demostrar la densidad de D_ξ en $\mathbb{R}$, sea I un intervalo abierto arbitrario y veamos que $D_\xi \cap I \neq \varnothing$. Lo primero que vamos a hacer es demostrar la siguiente:

Afirmación: *Para cada $n \in \mathbb{N}$, existe un único $p_n \in \mathbb{Z}$ tal que*

$$p_n + n\xi \in (0,1).$$

En efecto, sea $m = [n\xi]$ la parte entera de $n\xi$. Entonces $m \leq n\xi < m+1$ y, así, $0 \leq -m + n\xi < 1$. Ahora bien, como $n \geq 1$ y ξ es irracional, resulta que $-m + n\xi > 0$ y, en consecuencia, si hacemos $p_n = -m$, entonces $p_n + n\xi \in (0,1)$ y nuestra afirmación queda demostrada.

Puesto que I es un intervalo no vacío, su longitud $\ell(I) > 0$ y, por consiguiente, existe un $k \in \mathbb{N}$ tal que $\ell(I) > 1/k$. Dividamos el intervalo $(0,1)$ en k partes iguales y disjuntas, digamos $J_1, \ldots, J_k$. Escojamos ahora $k+1$ enteros distintos $n_1, \ldots, n_{k+1}$ y hagamos uso de nuestra afirmación para seleccionar enteros $p_{n_1}, \ldots, p_{n_{k+1}}$ de modo que $x_j = p_{n_j} + n_j\xi \in (0,1)$ para $j = 1, \ldots, k+1$. Puesto que ξ es irracional, todo los x_j's son distintos y entonces, por el Principio del Palomar, dos de tales puntos, digamos x_i y x_l pertenecen a uno, y sólo uno, de los intervalos anteriores al que denotaremos por J_{j_0} y, en consecuencia, $|x_i - x_l| < 1/k$. Esto, por supuesto, nos indica que podemos elegir un entero m tal que

$$m(x_i - x_l) \in I. \tag{1}$$

Por otro lado, como $x_j = p_{n_j} + n_j\xi$ para $j = 1, \ldots, k+1$, resulta entonces que

$$m(x_i - x_l) = m(p_{n_i} - p_{n_l}) + m(n_i - n_l)\xi \in D_\xi. \tag{2}$$

De (1) y (2) se sigue que $m(x_i - x_l) \in D_\xi \cap I$ y, por lo tanto, $D_\xi \cap I \neq \varnothing$. La prueba es completa. ■

Nota Adicional 2.2.3 Con un argumento similar al anterior, se prueba que el conjunto

$$D_\xi^2 = \{x \in \mathbb{R} : x = 2p + q\xi, \ \ p, q \in \mathbb{Z}\}$$

también es denso en $\mathbb{R}$. En efecto, la demostración es idéntica a la dada en el Teorema de Kronecker, sólo se reemplaza el intervalo $(0,1)$ por el intervalo $(0,2)$ en la Afirmación.

Observe, además, que cada conjunto D_ξ con $\xi \in \mathbb{R} \setminus \mathbb{Q}$ es un **subgrupo** (aditivo) de $\mathbb{R}$ y la familia $\mathcal{D}_\xi = \{x + D_\xi : x \in \mathbb{R}\}$ es una partición (disjunta) de $\mathbb{R}$. Para ver esto último, considere la siguiente relación sobre $\mathbb{R}$:

$$x \sim y \quad \Leftrightarrow \quad x - y \in D_\xi.$$

Es fácil establecer que $\sim$ es una relación de equivalencia sobre $\mathbb{R}$ por lo que la colección $\mathcal{E} = \{x + D_\xi : x \in \mathbb{R}\}$ formada por todas las clases de equivalencias determinadas por $\sim$ es una partición de $\mathbb{R}$. Estos conjuntos pueden ser usados para justificar, vía el Axioma de Elección, la existencia de conjuntos no medibles según Lebesgue.

Definición 2.2.38. *Sean (X,d) un espacio métrico y $E \subseteq X$. Diremos que E es **nunca-denso** en X, si $\operatorname{int}(\overline{E}) = \varnothing$. Si ocurre que $\operatorname{int}(\overline{E}) \neq \varnothing$, entonces se dice que E es **denso en alguna parte** de $\mathbb{R}$.*

El término **conjunto raro** se usa, con cierta frecuencia, como un sinónimo de conjunto nunca-denso. Por supuesto, E es nunca-denso si, y sólo si, $\overline{E}$ es nunca-denso. Observe que un conjunto nunca-denso no puede ser un entorno de ninguno de sus puntos y, por consiguiente,

*E es **nunca-denso** si, y sólo si, para cada conjunto abierto no vacío $G \subseteq X$, existe un conjunto abierto no vacío $J \subseteq G$ tal que $J \cap E = \varnothing$.*

Prueba. En efecto, suponga que existe un subconjunto abierto no vacío G de X con la propiedad de que cualquier subconjunto abierto no vacío de G intersecta a E. Esto, por supuesto, significa que $\overline{E}$ contiene a G lo que es imposible por ser E nunca-denso. ■

Teorema 2.2.39. *Sea F un subconjunto **cerrado** de un espacio métrico (X, d). Para cualquier $B \subseteq F$ se cumple que*

$$F \setminus \operatorname{int}(B) = \overline{F \setminus B}. \tag{2.2.1}$$

*En particular, $\operatorname{int}(B) = \varnothing$ si, y sólo si, $F \setminus B$ es **denso** en F.*

Prueba. Puesto que $\operatorname{int}(B) \subseteq B$, resulta entonces que $F \setminus B \subseteq F \setminus \operatorname{int}(B)$ y, como además, $F \setminus \operatorname{int}(B)$ es cerrado, se concluye que

$$\overline{F \setminus B} \subseteq F \setminus \operatorname{int}(B).$$

Por otro lado, suponga que $x \in F$ con $x \notin \overline{F \setminus B}$. Entonces existe un entorno G_x de x tal que $G_x \cap (F \setminus B) = \varnothing$. Esto, por supuesto, significa que $x \in G_x \subseteq B$, lo cual quiere decir que $x \in \operatorname{int}(B)$ y, en consecuencia, $x \notin F \setminus \operatorname{int}(B)$. Hemos demostrado que $F \setminus \operatorname{int}(B) \subseteq \overline{F \setminus B}$ y termina la prueba. ■

Si en el resultado anterior tomamos $F = X$ y $B \subseteq X$, entonces

$$B \text{ es } \textbf{nunca-denso} \quad \Leftrightarrow \quad X \setminus \overline{B} \text{ es } \textbf{denso} \text{ en } X.$$

Definición 2.2.40. *Sea A un subconjunto de un espacio métrico (X, d). Diremos que*

*(a) A es de **primera categoría** en X si existe una sucesión $(A_n)_{n=1}^{\infty}$ de subconjuntos nunca-densos de X tal que $A = \bigcup_{n=1}^{\infty} A_n$.*

*(b) A es de **segunda categoría** en X si él no es de primera categoría.*

Es claro que si $(B_n)_{n=1}^{\infty}$ es una sucesión de conjuntos de primera categoría, entonces $\bigcup_{n=1}^{\infty} A_n$ es de primera categoría. Es bien conocido que $\mathbb{Q}$ y $\mathbb{R} \setminus \mathbb{Q}$ son ambos densos en $\mathbb{R}$. Más aun, para cada número irracional ξ, el conjunto $\xi \cdot \mathbb{Q} = \{\xi \cdot r : r \in \mathbb{Q}\}$ es denso en $\mathbb{R}$. Sin embargo, si D es cualquier subconjunto a lo más numerable de números irracionales, entonces existen intersecciones de la forma

$$\bigcap_{\xi \in D} \xi \cdot \mathbb{Q},$$

que pueden ser vacías. ¿Cuándo se puede garantizar que cualquier intersección numerable de conjuntos densos sea siempre no vacía? La respuesta viene dada por uno de los resultados más prolíficos del análisis : el Teorema de Categoría de Baire (véase, por ejemplo, el libro [23] donde se muestran sorprendentes y casi inimaginables aplicaciones de dicho teorema).

Teorema 2.2.41 (Categoría de Baire). *Sea (X, d) un **espacio métrico completo**. Si $(G_n)_{n=1}^{\infty}$ es una sucesión de subconjuntos no vacíos, **abiertos** y **densos** en X. Entonces $G = \bigcap_{n=1}^{\infty} G_n$ es un G_δ-**denso** en X.*

Prueba. Sea V un conjunto abierto no vacío de X. Queremos demostrar que

$$V \cap \bigcap_{n=1}^{\infty} G_n \neq \varnothing.$$

Como G_1 es abierto y denso en X, se tiene que $V \cap G_1$ es abierto y no vacío. Escoja un punto $x_1 \in V \cap G_1$ y seleccione un conjunto abierto $J_1 \subseteq V \cap G_1$ con las siguientes propiedades:

$$x_1 \in J_1, \qquad \operatorname{diam}(J_1) < \frac{1}{2}, \qquad \overline{J_1} \subseteq V \cap G_1.$$

Como J_1 es un conjunto abierto no vacío y G_2 es abierto y denso en X, resulta que $J_1 \cap G_2$ es abierto y no vacío. Sea $x_2 \in J_1 \cap G_2$ y escojamos, de nuevo, un conjunto abierto $J_2 \subseteq J_1 \cap G_2$ para el cual:

$$x_2 \in J_2, \qquad \operatorname{diam}(J_2) < \frac{1}{2^2}, \qquad \overline{J_2} \subseteq J_1 \cap G_2.$$

Teniendo en cuenta, una vez más, que $J_2 \cap G_3$ es un abierto no vacío, podemos repetir el proceso anterior y continuarlo indefinidamente para obtener una sucesión de intervalos cerrados $(\overline{J_n})_{n=1}^{\infty}$ verificando las siguientes propiedades:

1) $\overline{J_1} \supseteq \overline{J_2} \supseteq \cdots \supseteq \overline{J_n} \supseteq \cdots$,

2) $\lim\limits_{n\to\infty} \operatorname{diam}(\overline{J_n}) = \lim\limits_{n\to\infty} \operatorname{diam}(J_n) = 0$, y

3) $\overline{J_n} \subseteq J_{n-1} \cap G_n$ para todo $n \geq 1$, donde hemos puesto $J_0 = V$.

Podemos entonces invocar al Teorema de Encaje de Cantor para concluir que

$$\varnothing \neq \bigcap_{n=1}^{\infty} \overline{J_n} \subseteq \bigcap_{n=1}^{\infty} \big(J_{n-1} \cap G_n\big) \subseteq V \cap \bigcap_{n=1}^{\infty} G_n.$$

Esto termina la prueba. ■

Varias consecuencias inmediatas se derivan del Teorema de Categoría de Baire.

TCB (0) $\mathbb{R}$ *es **no-numerable**.*

Suponga que $\mathbb{R}$ es numerable. Entonces existe una sucesión $(x_n)_{n=1}^{\infty}$ en $\mathbb{R}$ tal que $\mathbb{R} = \bigcup_{n=1}^{\infty}\{x_n\}$. Para cada $n \in \mathbb{N}$, el conjunto $G_n = \mathbb{R} \setminus \{x_n\}$ es abierto y denso en $\mathbb{R}$, por lo que el Teorema de Categoría de Baire nos garantiza que $\bigcap_{n=1}^{\infty} G_n$ es denso en $\mathbb{R}$, lo cual es imposible ya que

$$\bigcap_{n=1}^{\infty} G_n = \mathbb{R} \setminus \bigcup_{n=1}^{\infty}\{x_n\} = \varnothing.$$

TCB (1) *Si (X, d) es un **espacio métrico completo**, entonces él es de **segunda categoría**.*

Suponga, para construir una contradicción, que X es de primera categoría. Entonces existe una sucesión $(A_n)_{n=1}^{\infty}$ de subconjuntos nunca-densos tal que $X = \bigcup_{n=1}^{\infty} A_n$. Sin perder generalidad, podemos suponer que cada A_n es cerrado. Por el Teorema 2.2.39 cada conjunto $G_n = X \setminus A_n$ es abierto y denso en X y, por consiguiente, gracias al

Teorema de Categoría de Baire, $\bigcap_{n=1}^{\infty} G_n$ es denso en X, lo que constituye, por supuesto, una contradicción ya que

$$\varnothing \neq \bigcap_{n=1}^{\infty} G_n = X \setminus \bigcup_{n=1}^{\infty} A_n = \varnothing.$$

Esto prueba que X es de segunda categoría. ■

TCB (2) *Si $F \subseteq X$ es **cerrado**, entonces F satisface la conclusión del Teorema de Categoría de Baire, es decir, si $(G_n)_{n=1}^{\infty}$ es una sucesión de subconjuntos abiertos y densos en F, entonces $\bigcap_{n=1}^{\infty} G_n$ es denso en F.*

La prueba procede de manera idéntica a la del Teorema de Categoría de Baire y, por lo tanto, se omite.

TCB (3) *Sea $(F_n)_{n\in\mathbb{N}}$ una sucesión de **subconjuntos cerrados** en un **espacio métrico completo** (X,d) tal que $X = \bigcup_{n=1}^{\infty} F_n$. Entonces*

$$V = \bigcup_{n=1}^{\infty} \operatorname{int}(F_n)$$

*es **abierto** y **denso** en X.*

Prueba. Observemos que V es abierto por ser unión de conjuntos abiertos. Para cada $n \in \mathbb{N}$, sea $V_n = \operatorname{int}(F_n)$ y definamos $G_n = V_n \cup (X \setminus F_n)$. Fijemos $n \in \mathbb{N}$. Nuestro objetivo es probar que G_n es abierto y denso en X. En efecto, G_n es abierto por ser unión de dos conjuntos abiertos. Para ver que G_n es denso en X, tomemos cualquier subconjunto abierto no vacío G de X. Ocurre entonces que o bien $G \subseteq F_n$, en cuyo caso

$$G \subseteq V_n \subseteq G_n, \quad \text{es decir,} \quad G \cap G_n \neq \varnothing,$$

o bien $G \nsubseteq F_n$, de donde se obtiene que $G \cap (X \setminus F_n) \neq \varnothing$, y por consiguiente $G \cap G_n \neq \varnothing$. Esto prueba G_n es denso en X. Por el Teorema de Categoría de Baire, $E = \bigcap_{n=1}^{\infty} G_n$ es denso en X. Afirmamos que $E \subseteq V$. En efecto, suponga que $x \in E \subseteq X = \bigcup_{n=1}^{\infty} F_n$. Entonces existe algún $n_0 \in \mathbb{N}$ tal que $x \in F_{n_0}$. Por otro lado, $x \in G_{n_0}$ pues $x \in E = \bigcap_{n=1}^{\infty} G_n$, y así, $x \in G_{n_0} \cap F_{n_0} = V_{n_0} \subseteq V$; es decir, $x \in V$ y por lo tanto $E \subseteq V$. De aquí se sigue que V es denso en X y concluye la prueba. ■

TCB (4) *Si $F \subseteq \mathbb{R}$ es **perfecto**, es decir, **cerrado y sin puntos aislados**, entonces F es **no-numerable**.*

En efecto, suponga que F es numerable, digamos, $F = \{x_1, x_2, \ldots\}$. Para cada $n \geq 1$, considere el conjunto $G_n = \mathbb{R} \setminus \{x_n\}$. Como $\mathbb{R}$ no posee puntos aislados, cada conjunto $\{x_n\}$ es cerrado en $\mathbb{R}$, de donde resulta que $(G_n)_{n=1}^{\infty}$ es una sucesión de subconjuntos abiertos y densos en $\mathbb{R}$. Se sigue del Teorema de Categoría de Baire que $\bigcap_{n=1}^{\infty} G_n$ es denso en $\mathbb{R}$. Ahora bien, como F es un subespacio de $\mathbb{R}$, tenemos que $F \cap G_n$ es abierto en F y, además, como ningún punto de F es aislado, resulta que $F \cap G_n$ también es denso en F. Se sigue de TCB (2) que $\bigcap_{n=1}^{\infty}(F \cap G_n)$ es denso en F. En particular, $\bigcap_{n=1}^{\infty}(F \cap G_n) \neq \varnothing$, lo cual es imposible ya que

$$\varnothing \neq \bigcap_{n=1}^{\infty}(F \cap G_n) = F \cap \bigcap_{n=1}^{\infty} G_n = F \cap (\mathbb{R} \setminus F) = \varnothing.$$

Esto termina la prueba. ■

TCB (5) *Si $(G_n)_{n=1}^{\infty}$ es una sucesión de subconjuntos G_δ-**densos** en un **espacio métrico completo** (X,d), entonces $\bigcap_{n=1}^{\infty} G_n$ también es un G_δ-**denso**.*

En efecto, para cada $n \in \mathbb{N}$ seleccionemos una sucesión $(G_{nj})_{j=1}^{\infty}$ de subconjuntos abiertos tal que $G_n = \bigcap_{j=1}^{\infty} G_{nj}$. Teniendo en cuenta que $G_n \subseteq G_{nj}$ para todo $j \geq 1$, entonces la densidad de cada G_n garantiza la de cada G_{nj}. Puesto $\{G_{nj} : n, j \in \mathbb{N}\}$ es una colección numerable de subconjuntos abiertos y densos en X, el Teorema de Categoría de Baire nos revela que

$$\bigcap_{n=1}^{\infty} G_n = \bigcap_{n=1}^{\infty} \bigcap_{j=1}^{\infty} G_{nj}$$

es un G_δ-denso.

TCB (6) *Si $G \subseteq \mathbb{R}$ es un G_δ-**denso**, entonces G es **no-numerable**.*

Escriba $G = \bigcap_{n=1}^{\infty} G_n$ donde cada G_n es abierto y, por supuesto, denso ya que G lo es. Suponga ahora que $G = \{x_1, x_2, \ldots\}$ es numerable y, como en TCB (4), defina $H_n = \mathbb{R} \setminus \{x_n\}$ para todo $n \geq 1$. Entonces $(H_n)_{n=1}^{\infty}$ es una sucesión de subconjuntos y abiertos densos en $\mathbb{R}$ y, gracias al Teorema de Categoría de Baire, se tiene que $H = \bigcap_{n=1}^{\infty} H_n$ es un G_δ-denso. Finalmente, usando TBC (5) se obtiene la siguiente contradicción:

$$\varnothing = G \cap (\mathbb{R} \setminus G) = \left(\bigcap_{n=1}^{\infty} G_n\right) \cap \left(\bigcap_{n=1}^{\infty} H_n\right) \neq \varnothing.$$

Este hecho, combinado con TCB (5), es el que permite considerar a los conjuntos G_δ-densos como "*más numerosos*" de los que son simplemente densos. Por tal razón, *una propiedad* **P**(*x*) *la cual se cumple para todos los puntos de un conjunto G_δ-denso, se llama* ***abundante, genérica*** *o* ***típica*** y a un tal conjunto, un **conjunto abundante**.

Una consecuencia inmediata de TCB (6) es el siguiente:

Corolario 2.2.42. *$\mathbb{Q}$ no es un conjunto de tipo G_δ.*

Esto último permite construir subconjuntos de $\mathbb{R}$ que no son ni de tipo de G_δ así como tampoco de tipo F_σ. En efecto, para cada $a \in \mathbb{R}$, pongamos $G_a = \mathbb{Q} \cap (-\infty, a)$ y $F_a = (\mathbb{R} \setminus \mathbb{Q}) \cap [a, +\infty)$. Se sigue de lo anterior que el conjunto $H_a = G_a \cup F_a$ no es de tipo G_δ, ni de tipo F_σ.

Ya hemos tenido oportunidad de mostrar funciones que son continuas en los irracionales pero discontinuas en los racionales. El siguiente resultado establece, como consecuencia del resultado anterior, que lo contrario nunca es posible.

TCB (7) ***No existe*** *ninguna función $f : \mathbb{R} \to \mathbb{R}$ que* ***sea continua en los racionales*** *y* ***discontinua en los irracionales****.*

Prueba. Supongamos que una tal f existe. Por el Teorema 3.1.25 sabemos que $\mathrm{PC}(f)$ es un G_δ, mientras que, por hipótesis, $\mathrm{PC}(f) = \mathbb{Q}$. La combinación de estos hechos nos dice que $\mathbb{Q}$ es un G_δ-denso lo cual contradice el resultado anterior. ■

TCB (8) *Existencia de* ***conjuntos de Lusin***.

Bajo el imperio de la Hipótesis de Continuo se tiene que *cualquier conjunto* $A \subseteq \mathbb{R}$ *con* $\mathrm{card}(A) < \mathfrak{c}$ *es de primera categoría*. En efecto, como $\mathrm{card}(A) < \mathfrak{c}$, la Hipótesis de Continuo nos asegura que A es a lo más numerable y, por lo tanto, podemos escribir a A como una sucesión (finita o infinita), digamos $A = \{a_n : n \in F \subseteq \mathbb{N}\}$. Puesto que $A_n = \{a_n\}$ es nunca-denso para cada $n \in F$, resulta que $A = \bigcup_{n\in F} A_n$ es de primera categoría.

Usando ideas similares a las utilizadas en la construcción del conjunto de Bernstein, se puede construir otro conjunto extraño denominado "conjunto de Lusin" donde todos los subconjuntos de primera categoría que viven en él son necesariamente numerables. Como antes, ello es posible si se acepta la Hipótesis del Continuo más la existencia de un buen-orden en $\mathbb{R}$, es decir, el Axioma de Elección.

Teorema 2.2.43 (Conjunto de Lusin). *Asumiendo la* ***Hipótesis del Continuo****, existe un subconjunto no vacío* $\mathbb{L}$ *de* $\mathbb{R}$ *con las siguientes propiedades:*

(a) $\mathrm{card}(\mathbb{L}) = \aleph_1$ *y*

(b) $\mathrm{card}(\mathbb{L} \cap A) \leq \aleph_0$ *para cualquier conjunto* $A \subseteq \mathbb{R}$ *de* ***primera categoría***.

Prueba. Sea $\mathcal{F}_{\mathfrak{c}}$ la familia de todos los subconjuntos cerrados de $\mathbb{R}$. Por el Teorema 1.3.47 (8), página 74, sabemos que $\mathrm{card}(\mathcal{F}_{\mathfrak{c}}) = \aleph_1 = \mathfrak{c}$ (recuerde que estamos asumiendo la Hipótesis de Continuo). Denotemos por $\mathcal{F}_{\mathrm{nd}}$ los miembros de $\mathcal{F}_{\mathfrak{c}}$ que son nunca-densos. Puesto que $\{x\} \in \mathcal{F}_{\mathrm{nd}}$ para cada $x \in \mathbb{R}$, entonces $\mathrm{card}(\mathcal{F}_{\mathrm{nd}}) \geq \mathfrak{c}$ y, así,

$$\mathfrak{c} \;\leq\; \mathrm{card}(\mathcal{F}_{\mathrm{nd}}) \;\leq\; \mathrm{card}(\mathcal{F}_{\mathfrak{c}}) \;=\; \mathfrak{c}.$$

Este hecho permite que podamos enumerar a $\mathcal{F}_{\mathrm{nd}}$ por medio de una sucesión transfinita, digamos, $\mathcal{F}_{\mathrm{nd}} = \{F_\alpha : \alpha < \omega_1\}$. Para construir nuestro conjunto $\mathbb{L}$ observe que como $F_\alpha \neq \mathbb{R}$ para todo $\alpha < \omega_1$, podemos comenzar eligiendo un $x_1 \in \mathbb{R}$ tal que $x_1 \notin F_1 \in \mathcal{F}_{\mathrm{nd}}$. Fijemos un $\xi < \omega_1$ y supongamos que la familia $(x_\alpha)_{\alpha<\xi}$ ha sido construida. Consideremos el conjunto

$$X_\xi \;=\; \Big(\bigcup_{\alpha<\xi} F_\alpha\Big) \cup \{x_\alpha : \alpha < \xi\}$$

y observemos que como ξ es numerable, X_ξ es de primera categoría en $\mathbb{R}$. Por el Teorema de Categoría de Baire, $X_\xi \neq \mathbb{R}$. Si ahora elegimos un $x_\xi \in \mathbb{R} \setminus X_\xi$, entonces el Principio de Inducción Transfinita da por finalizada la construcción de $(x_\xi)_{\xi<\omega_1}$. Definamos ahora

$$\mathbb{L} = \{x_\xi : \xi < \omega_1\}.$$

Afirmamos que $\mathbb{L}$ posee las propiedades requeridas. En efecto:

(a) $\mathbb{L}$ *es no numerable*. Considere la aplicación $f : \omega_1 \to \mathbb{L}$ dada por $f(\alpha) = x_\alpha$. Veamos que ella es biyectiva. En efecto, es claro que f es sobreyectiva. Para ver que ella también es inyectiva, suponga que $\alpha < \xi < \omega_1$. Entonces, por construcción, se tiene $x_\alpha \neq x_\xi$. Esto prueba que $\mathrm{card}(\mathbb{L}) = \mathfrak{c} = \aleph_1$.

(b) $\mathbb{L}$ *satisface* (b). Sea $A = \bigcup_{n=1}^{\infty} A_n$ un conjunto de primera categoría en $\mathbb{R}$, donde cada A_n es cerrado y nunca-denso. Puesto que $\mathcal{F}_{\mathrm{nd}}$ es la familia de todos los conjuntos cerrados nunca densos de $\mathbb{R}$, resulta que para cada $n \in \mathbb{N}$, existe un $F_{\alpha_n} \in \mathcal{F}_{\mathrm{nd}}$ tal que $A_n = F_{\alpha_n}$. Por construcción, sabemos que

$$\mathbb{L} \cap A_n = \mathbb{L} \cap F_{\alpha_n} \subseteq \{x_\beta : \beta < \alpha\}$$

y como éste último conjunto es numerable, entonces $\mathbb{L} \cap A = \bigcup_{n=1}^{\infty} \mathbb{L} \cap A_n$ es numerable.

Esto termina la prueba. ■

Al conjunto $\mathbb{L}$ del resultado anterior se le llama un ***conjunto de Lusin***. Observe que, por construcción, *todo subconjunto no-numerable incluido en* $\mathbb{L}$ *es de segunda categoría. En particular,* $\mathbb{L}$ *es de* ***segunda categoría***.

Una nota interesante, que puede ser de interes al lector, es siguiente resultado:

Si existe un ***conjunto de Lusin*** *y cualquier subconjunto de* $\mathbb{R}$ *de* ***cardinalidad menor*** *que* $\mathfrak{c}$ *es de* ***primera categoría***, *entonces la* ***Hipótesis del Continuo*** *es* ***verdadera***.

Definición 2.2.44. *Una función* $f : [a, b] \to \mathbb{R}$ *se dice que es* ***puntualmente discontinua*** *si* $\mathrm{PC}(f)$ *es denso en* $[a, b]$.

Corolario 2.2.45. *Si* $f : [a, b] \to \mathbb{R}$ *es una función* ***puntualmente discontinua***, *entonces* $\mathrm{PC}(f)$ *es un* G_δ-***denso*** *en* $[a, b]$.

Prueba. Sólo resta demostrar que $\mathrm{PC}(f)$ es un G_δ. Esto, sin embargo, es consecuencia del Teorema 3.1.25, página 168. ■

¿Qué tan complicado es el conjunto de los puntos de continuidad de una función arbitraria $f : [a, b] \to \mathbb{R}$? ¿En qué otros casos es $\mathrm{PC}(f)$ un G_δ-denso? El siguiente resultado establece la existencia de una amplia categoría de funciones, cada una de las cuales posee abundantes puntos de continuidad, es decir, sus puntos de continuidad forman un G_δ-denso en $[a, b]$.

Teorema 2.2.46 (Baire). *Sea* $f : [a, b] \to \mathbb{R}$ *una función arbitraria. Suponga que existe una sucesión de funciones en* $C([a, b])$, *digamos* $(f_n)_{n \in \mathbb{N}}$, *tal que* $f(x) = \lim_{n \to \infty} f_n(x)$ *existe para cada* $x \in [a, b]$. *Entonces* $\mathrm{PC}(f)$ *es un* G_δ-***denso*** *en* $[a, b]$. *En particular,* $\mathrm{PC}(f)$ *es* ***no-numerable***.

Prueba. $\mathrm{PC}(f)$ es un G_δ gracias al Teorema 3.1.25. Veamos que él también es denso en $[a, b]$. Para $k, m, n \in \mathbb{N}$, definamos

$$F_{kmn} = \left\{ x \in [a, b] : |f_m(x) - f_n(x)| \leq \frac{1}{k} \right\}.$$

Siendo f_m, f_n funciones continuas, también lo es $|f_m - f_n|$ y, en consecuencia, cada F_{kmn} es cerrado en $[a, b]$. Fijando $k, m \in \mathbb{N}$ y definiendo

$$F_{km} = \bigcap_{n=m}^{\infty} F_{kmn},$$

resulta que cada F_{km} también es cerrado en $[a,b]$. Afirmamos que $[a,b] = \bigcup_{m=1}^{\infty} F_{km}$ para cada $k \in \mathbb{N}$. En efecto, sea $k \in \mathbb{N}$ y sea $x \in [a,b]$. Puesto que $\lim_{n\to\infty} f_n(x) = f(x)$, existe un $m \in \mathbb{N}$ tal que $\left|f_n(x) - f(x)\right| \leq 1/(2k)$ para todo $n \geq m$. Ahora bien, si $n \geq m$, entonces

$$\left|f_m(x) - f_n(x)\right| \;\leq\; \left|f_m(x) - f(x)\right| \;+\; \left|f_n(x) - f(x)\right| \;\leq\; \frac{1}{2k} \;+\; \frac{1}{2k} \;=\; \frac{1}{k}$$

lo que prueba que $x \in F_{kmn}$ para todo $n \geq m$ y, por lo tanto,

$$x \in \bigcap_{n=m}^{\infty} F_{kmn} \;=\; F_{km} \;\subseteq\; \bigcup_{i=1}^{\infty} F_{ki}$$

Esto prueba nuestra afirmación. Por TCB (3), cada $G_k = \bigcup_{m=1}^{\infty} \operatorname{int}(F_{km})$ es abierto y denso en $[a,b]$ y, en consecuencia, por el Teorema de Categoría de Baire, el conjunto

$$E \;=\; \bigcap_{k=1}^{\infty} G_k$$

es denso en $[a,b]$.

Nos proponemos demostrar que $E \subseteq \mathrm{PC}(f)$. En efecto, sean $x_0 \in E$ y $\varepsilon > 0$. Elijamos un $k \in \mathbb{N}$ de modo tal que $\frac{1}{k} < \frac{\varepsilon}{3}$. Como $x_0 \in G_k = \bigcup_{m=1}^{\infty} \operatorname{int}(F_{km})$, existe un $m \in \mathbb{N}$ tal que $x_0 \in \operatorname{int}(F_{km})$. Siendo $\operatorname{int}(F_{km})$ abierto, podemos elegir un $\delta_0 > 0$ de modo que intervalo abierto $(x_0 - \delta_0, x_0 + \delta_0) \subseteq \operatorname{int}(F_{km})$. Observemos ahora que si $|x - x_0| < \delta_0$, entonces $x \in \operatorname{int}(F_{km}) \subseteq F_{km}$ y así, $\left|f_m(x) - f_n(x)\right| \leq 1/k$ para cualquier $n \geq m$. Usando la continuidad del valor absoluto, se obtiene que

$$\left|f_m(x) - f(x)\right| \;=\; \left|f_m(x) - \lim_{n\to\infty} f_n(x)\right| \;=\; \lim_{n\to\infty} \left|f_m(x) - f_n(x)\right| \;\leq\; \frac{1}{k}.$$

En particular, $\left|f_m(x_0) - f(x_0)\right| \leq 1/k$. Por otro lado, como f_m continua en x_0, existe un δ_1 tal que $\Big|f_m(x) - f_m(x_0)\Big| \leq \varepsilon/3$ siempre que $x \in (x_0 - \delta_1, x_0 + \delta_1)$. Sea $\delta = \min\{\delta_0, \delta_1\}$. Si $|x - x_0| < \delta$, entonces

$$\begin{aligned}
\left|f(x) - f(x_0)\right| \;&\leq\; \left|f(x) - f_m(x)\right| \;+\; \left|f_m(x) - f_m(x_0)\right| \;+\; \left|f_m(x_0) - f(x_0)\right| \\
&\leq\; \frac{1}{k} \;+\; \frac{\varepsilon}{3} \;+\; \frac{1}{k} \;\leq\; \frac{\varepsilon}{3} \;+\; \frac{\varepsilon}{3} \;+\; \frac{\varepsilon}{3} \;=\; \varepsilon
\end{aligned}$$

Esto prueba que f es continua en x_0 y, en consecuencia, $E \subseteq \mathrm{PC}(f)$. De aquí se sigue que $\mathrm{PC}(f)$ es un G_δ-denso en $[a,b]$. La última parte es consecuencia de TCB (6). ■

El símbolo $\mathcal{B}_1([a,b])$ denotará *el conjunto de todas las funciones* $f : [a,b] \to \mathbb{R}$ *que son límites puntuales de sucesiones de funciones continuas*, esto es, $f \in \mathcal{B}_1([a,b])$ si, y sólo si, existe una sucesión $(f_n)_{n=1}^{\infty}$ en $C([a,b])$ tal que

$$f(x) \;=\; \lim_{n\to\infty} f_n(x)$$

para cada $x \in [a,b]$. Cada elemento de $\mathcal{B}_1([a,b])$ se llamará, en lo sucesivo, una **función de la primera clase de Baire**. Resulta claro que

$$C([a,b]) \;\subseteq\; \mathcal{B}_1([a,b]).$$

Se sigue del resultado anterior que para cada $f \in \mathcal{B}_1([a,b])$, el conjunto $\mathrm{PC}(f)$ es un G_δ-denso en $[a,b]$.

Corolario 2.2.47. *Si $f : [a, b] \to \mathbb{R}$ es una función **continua** tal que $f'(x)$ existe para todo $x \in [a, b]$, entonces $f' \in \mathcal{B}_1([a, b])$.*

Prueba. Extienda f continuamente al intervalo $[a, b+1]$ poniendo $f(x) = f(b)$ para todo x en $[b, b+1]$. Si para cada $n \in \mathbb{N}$, definimos $g_n : [a, b] \to \mathbb{R}$ por

$$g_n(x) \;=\; n\big[f(x + 1/n) - f(x)\big],$$

resulta que cada g_n es continua y se cumple que, para cada $x \in [a, b]$,

$$f'(x) \;=\; \lim_{n\to\infty} \frac{f(x+1/n) - f(x)}{1/n} \;=\; \lim_{n\to\infty} g_n(x).$$

Esto prueba que $f' \in \mathcal{B}_1([a, b])$. ■

2.2.3. Espacios Normados

Sea X un espacio vectorial sobre $\mathbb{R}$. Una **norma** sobre X es una aplicación $\|\cdot\| : X \to \mathbb{R}$ que cumple con las siguientes propiedades:

(1) $\|x\| \geq 0$ para todo $x \in X$,

(2) $\|x\| = 0$ si y, sólo si, $x = 0$,

(3) $\|a \cdot x\| = a \cdot \|x\|$ para todo $x \in X$ y todo $a \in \mathbb{R}$, y

(4) $\|x + y\| \leq \|x\| + \|y\|$ para todo $x, y \in X$. (Desigualdad Triangular)

El par $(X, \|\cdot\|)$ es llamado un **espacio normado**. Se sigue de (4) e inducción que

$$\left\| \sum_{k=1}^{n} x_k \right\| \;\leq\; \sum_{k=1}^{n} \|x_k\|$$

cualesquiera sean $x_1, \ldots, x_n$ en X. Más aun,

$$\Big| \|x\| - \|y\| \Big| \;\leq\; \|x - y\| \tag{1a}$$

para todo $x, y \in X$. Observe que gracias a las propiedades (3) y (4) cualquier norma es una función convexa. Es importante también destacar que toda norma sobre un espacio vectorial da origen a una métrica natural definida por

$$d(x, y) \;=\; \|x - y\|, \quad x, y \in X, \tag{2a}$$

y, en consecuencia, todo espacio normado $(X, \|\cdot\|)$ siempre será pensado como un espacio métrico bajo dicha métrica.

Definición 2.2.48. *Dos normas $\|\cdot\|$ y $\|\cdot\|_*$ sobre X son **equivalentes** si existen constantes a, $b > 0$ tales que*

$$a \cdot \|x\| \;\leq\; \|x\|_* \;\leq\; b \cdot \|x\| \quad \textit{para todo } x \in X.$$

Observe que afirmar que $\|\cdot\|$ y $\|\cdot\|_*$ son equivalentes significa que las topologías inducidas por ambas normas coinciden sobre X, es decir, la aplicación identidad $\mathrm{id} : (X, \|\cdot\|) \to (X, \|\cdot\|_*)$ es un homeomorfismo.

Sea $(X, \|\cdot\|_X)$ un espacio normado. Recordemos que una aplicación $f : X \to \mathbb{R}$ se llama un **funcional lineal** si $f(ax+by) = af(x) + bf(y)$ para todo $x, y \in X$ y todo $a, b \in \mathbb{R}$. El **dual** de X, denotado por X^*, consiste de *todos los funcionales lineales continuos* $x^* : X \to \mathbb{R}$. Es claro que, con las operaciones usuales, X^* es un espacio vectorial sobre $\mathbb{R}$. Recordemos que un funcional lineal $x^* : X \to \mathbb{R}$ *es continuo si, y sólo si, él es acotado*, es decir, existe una constante $C > 0$ tal que $|x^*(x)| \leq C\|x\|$ para todo $x \in X$. En este caso, se define el número $\|x^*\|$ por

$$\|x^*\| \;=\; \inf\big\{C > 0 \,:\, |x^*(x)| \leq C\|x\|\big\}.$$

Es fácil verificar $\|x^*\|_{X^*}$ es una norma sobre X^* llamada la **norma dual**. Más aun,

$$\|x^*\| \;=\; \sup\big\{|x^*(x)| \,:\, \|x\| \leq 1\big\}$$

de donde se sigue

$$|x^*(x)| \;\leq\; \|x^*\|\,\|x\| \quad \text{para todo } x \in X.$$

Con esta información se prueba si dificultad que $(X^*, \|\cdot\|)$ es siempre un espacio de Banach independientemente de si X es o no de Banach.

Sea $(Y, \|\cdot\|_Y)$ otro espacio de Banach. Diremos que Y es el **dual** de X, denotado por $Y = X^*$, si existe una aplicación lineal isométrica $T : X^* \to Y$ tal que $T(X^*) = Y$, esto significa que T es lineal, continua, biyectiva satisfaciendo

$$\|T(x^*)\| \;=\; \|x^*\| \quad \text{para todo } x^* \in X^*.$$

Teorema 2.2.49 (Acotación Uniforme). *Sea $(X, \|\cdot\|)$ un **espacio de Banach** y suponga que $\mathcal{F}$ es una familia no vacía incluida en X^* que está **puntualmente acotada** en el siguiente sentido: para cada $x \in X$, existe una constante $M_x \geq 0$ tal que*

$$\sup_{x^* \in \mathcal{F}} |x^*(x)| \;\leq\; M_x.$$

Entonces existe una constante $M \geq 0$ tal que

$$\sup_{x^* \in \mathcal{F}} \|x^*\| \;\leq\; M.$$

Prueba. Por cada entero $k \geq 1$ y cada $x^* \in \mathcal{F}$, considere el conjunto

$$E_{k,x^*} \;=\; \big\{x \in X \,:\, |x^*(x)| \leq k\big\}.$$

Pongamos ahora,

$$E_k \;=\; \bigcap_{x^* \in \mathcal{F}} E_{k,x^*}.$$

Observe que $X = \bigcup_{k=1}^{\infty} E_k$. Sólo se necesita verificar que $X \subseteq \bigcup_{k=1}^{\infty} E_k$. En efecto, tome cualquier $x \in X$ y usemos nuestra hipótesis para hallar una constante $M_x \geq 0$ tal que $|x^*(x)| \leq M_x$ para todo $x^* \in \mathcal{F}$. Si ahora elegimos cualquier $k \geq M_x$, entonces se cumple que $x \in E_k$ y, por lo tanto, $X \subseteq \bigcup_{k=1}^{\infty} E_k$. Por otro lado, como cada $x^* \in \mathcal{F}$ es una aplicación continua, resulta que

E_{k,x^*} es cerrado y, en consecuencia, E_k también lo es. Se sigue del Teorema de Categoría de Baire, véase **TCB** (3), que el conjunto

$$V = \bigcup_{k=1}^{\infty} \operatorname{int}(E_k)$$

es abierto y denso en X. De aquí se sigue la existencia de algún $k_0 \in \mathbb{N}$ tal que $\operatorname{int}(E_{k_0}) \neq \varnothing$. En particular, existe algún $x_0 \in \operatorname{int}(E_{k_0})$ y un cierto $\delta > 0$ tal que la bola abierta $U(x_0, \delta) \subseteq \operatorname{int}(E_{k_0})$. Resulta claro que

$$\sup_{\substack{x^* \in \mathcal{F} \\ x \in U(x_0,\delta)}} |x^*(x)| \leq k_0.$$

Finalmente, sea $x \in X$ con $x \neq 0$ y defina $y = \frac{\delta}{2} \| x \|^{-1} x$. Puesto que $\| y \| < \delta$, entonces $y + x_0 \in U(x_0, \delta)$ y se sigue de la desigualdad anterior que

$$\begin{aligned} |x^*(x)| &\leq \left| x^* \left(\frac{2 \| x \|}{\delta} y \right) \right| \\ &= \frac{2 \| x \|}{\delta} |x^*(y)| \\ &\leq \frac{2 \| x \|}{\delta} \big(|x^*(y + x_0)| + |x^*(x_0)| \big) \\ &\leq \frac{2 \| x \|}{\delta} 2k_0 \\ &= \frac{4k_0}{\delta} \| x \| \end{aligned}$$

para cualquier $x^* \in \mathcal{F}$. Por supuesto, como dicha desigualdad se cumple trivialmente si $x = 0$, resulta que tomando $M = 4k_0/\delta$ se obtiene el resultado deseado; es decir,

$$\sup_{x^* \in \mathcal{F}} \| x^* \| \leq M.$$

Fin de la prueba. ■

2.2.4. La Topología Producto

Sea X un conjunto no vacío y sea $\mathcal{E} \subseteq \mathscr{P}(X)$. Afirmamos que existe una **única topología** sobre X, la cual denotaremos por $\tau_X(\mathcal{E})$, con las siguientes propiedades:

(a) $\mathcal{E} \subseteq \tau_X(\mathcal{E})$, y

(b) $\tau_X(\mathcal{E})$ es la *topología más pequeña* conteniendo a $\mathcal{E}$, es decir, si $\mathcal{J}$ es otra topología sobre X con $\mathcal{E} \subseteq \mathcal{J}$, entonces $\tau_X(\mathcal{E}) \subseteq \mathcal{J}$.

La prueba de la existencia y unicidad de $\tau_X(\mathcal{E})$ es simple. En efecto, considere la familia $\mathfrak{T}(\mathcal{E})$ formada por todas las topologías sobre X que contienen a $\mathcal{E}$. Observe que $\mathfrak{T}(\mathcal{E})$ es no vacía ya que $\mathscr{P}(X) \in \mathfrak{T}(\mathcal{E})$ y entonces defina

$$\tau_X(\mathcal{E}) = \bigcap_{\mathcal{J} \in \mathfrak{T}(\mathcal{E})} \mathcal{J}.$$

Es un ejercicio sencillo establecer que $\tau_X(\mathcal{E})$ es una topología sobre X que satisface (a) y (b). Por supuesto, ella es única. A $\tau_X(\mathcal{E})$ se le llama la **topología generada** por $\mathcal{E}$.

El siguiente mecanismo, basado en la construcción anterior, es un proceso natural para generar topologías sobre un conjunto dado. Comencemos con un conjunto no vacío X, el cual puede o no estar provisto de alguna topología, y sea $(Y, \mathcal{J}_Y)$ un espacio topológico arbitrario. Se considera una función $f : X \to (Y, \mathcal{J}_Y)$ y lo que se desea es construir la topología más pequeña sobre X, digamos $\mathcal{J}_X$, que permita que f sea $\mathcal{J}_X - \mathcal{J}_Y$ continua, es decir, $f^{-1}(G) \in \mathcal{J}_X$ para cualquier $G \in \mathcal{J}_Y$. Hay evidencia de la existencia de al menos una topología que hace que f sea $\mathcal{J}_X - \mathcal{J}_Y$ continua. En efecto, si tomamos a X con la topología discreta, $\mathscr{P}(X)$, resulta entonces claro que f es $\mathscr{P}(X) - \mathcal{J}_Y$ continua. La elección de la topología discreta no siempre es la más afortunada por lo que es deseable poder contar con otra topología, distinta a la discreta, que nos resuelva el problema. Consideremos entonces

$$\mathcal{J}_X = \left\{ f^{-1}(G) : G \in \mathcal{J}_Y \right\}.$$

Es fácil establecer que $\mathcal{J}_X$ es la topología más pequeña sobre X que hace que f sea $\mathcal{J}_X - \mathcal{J}_Y$ continua.

Generalicemos el argumento anterior en el siguiente sentido: tomemos el conjunto X y en lugar de tener un único espacio topológico $(Y, \mathcal{J}_Y)$ y una única función $f : X \to Y$, elijamos una familia de espacios topológicos, digamos $(Y_\alpha, \mathcal{J}_\alpha)_{\alpha \in I}$, y una familia de funciones $f_\alpha : X \to Y_\alpha$ con $\alpha \in I$, y como antes, queremos construir la topología más pequeña sobre X que permita que *todas las funciones* f_α sean continuas. He aquí el procedimiento. Para cada $\alpha \in I$, sea

$$\tau_\alpha = \left\{ f_\alpha^{-1}(G) : G \in \mathcal{J}_\alpha \right\}$$

y pongamos

$$\mathcal{E} = \bigcup_{\alpha \in I} \tau_\alpha.$$

Observe que por ser cada uno de los conjuntos τ_α una topología sobre X (la cual garantiza que la respectiva función f_α sea continua, pero no las demás), resulta que $\varnothing, X \in \mathcal{E}$. En general, $\mathcal{E}$ no es una topología sobre X, pero sí lo es $\tau_X(\mathcal{E})$, la topología generada por $\mathcal{E}$. A esta topología se le llama la **topología inicial o proyectiva** sobre X asociada a la familia $(f_\alpha)_{\alpha \in I}$. Es claro que $\tau_X(\mathcal{E})$ es la topología más pequeña sobre X bajo la cual cada f_α es $\tau_X(\mathcal{E}) - \mathcal{J}_\alpha$ continua.

Denote por $\mathcal{E}_{\text{fin}}$ *la colección de las intersecciones de familias finitas de* $\mathcal{E}$, esto es, $A \in \mathcal{E}_{\text{fin}}$ si, y sólo si, existe $F \in \mathcal{P}_{\text{fin}}(I)$ y conjuntos $A_\alpha \in \mathcal{E}$ con $\alpha \in F$ tal que

$$A = \bigcap_{\alpha \in F} A_\alpha.$$

Si ahora se considera $\mathcal{E}_\sigma$ *la colección de todas las uniones de familias arbitrarias de* $\mathcal{E}_{\text{fin}}$, entonces

$$\tau_X(\mathcal{E}) = \mathcal{E}_\sigma. \tag{1}$$

En primer lugar, observe que, si $A, B \in \mathcal{E}_{\text{fin}}$, entonces $A \cap B \in \mathcal{E}_{\text{fin}}$. Más aun, $\mathcal{E}_\sigma$ contiene a $\mathcal{E}$. Para verificar la igualdad (1), sólo debemos demostrar que $\mathcal{E}_\sigma$ es una topología sobre X. Claramente, $\varnothing, X \in \mathcal{E}_\sigma$. Sean ahora $U, V \in \mathcal{E}_\sigma$. Entonces $U = \bigcup_{\alpha \in F} A_\alpha$ y $V = \bigcup_{\beta \in G} B_\beta$ donde $(A_\alpha)_{\alpha \in F}$ y $(B_\beta)_{\beta \in G}$ son familias de elementos de $\mathcal{E}_{\text{fin}}$. Se deduce que

$$U \cap V = \left(\bigcup_{\alpha \in F} A_\alpha \right) \cap \left(\bigcup_{\beta \in G} B_\beta \right) = \bigcup_{(\alpha, \beta) \in F \times G} (A_\alpha \cap B_\beta)$$

y puesto que $A_\alpha \cap B_\beta \in \mathcal{E}_{\text{fin}}$ para cualquier $(\alpha, \beta) \in F \times G$, resulta que $U \cap V \in \mathcal{E}_\sigma$. Por inducción se sigue que cualquier intersección finita de elementos de $\mathcal{E}_\sigma$ permanece dentro de $\mathcal{E}_\sigma$.

Suponga ahora que $(U_\alpha)_{\alpha\in F}$ es una familia de elementos de $\mathcal{E}_\sigma$. Entonces, U_α es la unión de una familia de elementos de $\mathcal{E}_{\text{fin}}$ y, en consecuencia, $\bigcup_{\alpha\in F} U_\alpha$ es la unión de una familia de elementos de $\mathcal{E}_{\text{fin}}$. Esto prueba que $\bigcup_{\alpha\in F} U_\alpha \in \mathcal{E}_\sigma$ y, así, $\mathcal{E}_\sigma$ es una topología sobre X. Como $\mathcal{E}_\sigma$ que contiene a $\mathcal{E}$ se tiene que $\tau(\mathcal{E}) \subseteq \mathcal{E}_\sigma$. Por otro lado, puesto que $\tau(\mathcal{E})$ es una topología conteniendo a $\mathcal{E}$ se tiene que $\mathcal{E}_{\text{fin}} \subseteq \tau(\mathcal{E})$ y, por consiguiente, $\mathcal{E}_\sigma \subseteq \tau(\mathcal{E})$. Esto termina la prueba de (1).

Como una aplicación de lo que acabamos de ver, mostraremos de inmediato cómo se obtiene la **topología producto** de una familia de espacios topológicos. Suponga que $(X_\alpha, \mathfrak{J}_\alpha)_{\alpha\in I}$ es una familia de espacios topológicos y sea
$$X = \prod_{\alpha\in I} X_\alpha$$
su producto cartesiano. Para cada $\beta \in I$, sea
$$p_\beta : \prod_{\alpha\in I} X_\alpha \to X_\beta$$
la **proyección** de X sobre X_β, esto es, $p_\beta((x_\alpha)_{\alpha\in I}) = x_\beta$ para todo $(x_\alpha)_{\alpha\in I} \in X$. La topología inicial sobre X asociada a la familia $(p_\alpha)_{\alpha\in I}$ la llamaremos la **topología producto** sobre X y la denotaremos, en lo sucesivo, por $\mathfrak{T}_p$. Por consiguiente, *$\mathfrak{T}_p$ es la topología más pequeña sobre X que hace que cada proyección p_α sea continua.*

Usando la igualdad dada en (1) sabemos que $\mathfrak{T}_p = \mathcal{E}_\sigma$, donde
$$\mathcal{E} = \bigcup_{\alpha\in I} p_\alpha^{-1}(\mathfrak{J}_\alpha),$$
y $p_\alpha^{-1}(\mathfrak{J}_\alpha) = \{p_\alpha^{-1}(U) : U \in \mathfrak{J}_\alpha\}$. Observe que, para cada $U \in \mathfrak{J}_\alpha$,
$$p_\beta^{-1}(U) = \prod_{\alpha\in I} V_\alpha, \qquad \text{donde} \qquad V_\alpha = \begin{cases} U & \text{si } \alpha = \beta \\ X_\alpha & \text{si } \alpha \neq \beta. \end{cases}$$
Recordemos que $\mathcal{E}_{\text{fin}}$ es la familia formada por todas las intersecciones de colecciones finitas de elementos de $\mathcal{E}$, es decir, $A \in \mathcal{E}_{\text{fin}}$ si, y sólo si, existe un conjunto finito $F \subseteq I$ y conjuntos abiertos $U_\alpha \subseteq X_\alpha$ con $\alpha \in F$ tal que
$$A = \bigcap_{\alpha\in F} p_\alpha^{-1}(U_\alpha) = \prod_{\beta\in I} V_\beta, \qquad \text{donde} \qquad V_\beta = \begin{cases} U_\alpha & \text{si } \alpha \in F \\ X_\alpha & \text{si } \alpha \notin F. \end{cases}$$
Finalmente, $\mathcal{E}_\sigma$ es la colección que se obtiene formando todas las uniones de familias arbitrarias de conjuntos pertenecientes a $\mathcal{E}_{\text{fin}}$, es decir, $V \in \mathcal{E}_\sigma$ si, y sólo si, existe una familia $(W_\beta)_{\beta\in F}$ incluida en $\mathcal{E}_{\text{fin}}$ tal que
$$V = \bigcup_{\beta\in F} W_\beta.$$
Observe que $\mathcal{E}_{\text{fin}}$ constituye *una base para la topología producto*. Sus elementos serán llamados los ***conjuntos básicos*** de $\mathfrak{T}_p$.

Por definición, cada proyección $p_\alpha : X \to X_\alpha$ es una función continua cuando X está dotado de la topología producto. Lo que resulta también interesante es que p_α es una **aplicación abierta**, esto es, *transforma conjuntos abiertos en conjuntos abiertos*. Para ver esto, fijemos un $\alpha \in I$ y sea $U \in \mathcal{E}$. Entonces existe un $\beta \in I$ y un conjunto $U_\beta \in \mathcal{J}_\beta$ tal que $U = p_\beta^{-1}(U_\beta)$. Puesto que p_α es una aplicación sobreyectiva, se tiene que

$$p_\alpha(U) = V_\alpha \in \mathcal{J}_\alpha \quad \text{donde} \quad V_\alpha = \begin{cases} U_\alpha & \text{si } \alpha = \beta \\ X_\alpha & \text{si } \alpha \neq \beta. \end{cases}$$

Si $U \in \mathcal{E}_{\text{fin}}$, entonces $U = \bigcap_{\beta \in F} p_\alpha^{-1}(U_\beta)$ donde $F = \{\beta \in I : U_\beta \neq X_\beta\}$ es finito. Se sigue de lo anterior que $p_\alpha(U)$ es abierto en X_α. En el caso general, puesto que todo abierto $U \subseteq X$ es unión de conjuntos básicos, esto es, $U = \bigcup_{\beta \in F} W_\beta$ donde $(W_\beta)_{\beta \in F}$ es alguna colección de conjuntos básicos, se sigue que

$$p_\alpha(U) = \bigcup_{\beta \in F} p_\alpha(W_\beta) \in \mathcal{J}_\alpha.$$

En conclusión, cada proyección es una función ***continua, abierta*** y ***sobreyectiva***.

Uno de los resultados fundamentales en Análisis lo constituye el Teorema de Tychonoff el cual afirma que el producto arbitrario de espacios compactos, provisto de la topología producto, es compacto. La demostración que a continuación ofrecemos se apoya en el Lema de Zorn. Sin embargo, también se pueden dar demostraciones de dicho resultado usando el Principio del Buen Orden, redes, ultrafiltros, etc.

Teorema 2.2.50 (Tychonoff). *Sea $(X_\alpha, \mathcal{J}_\alpha)_{\alpha \in I}$ una familia de **espacios topológicos compactos**. Entonces el producto $(\prod_{\alpha \in I} X_\alpha, \mathfrak{T}_p)$ es compacto.*

Prueba. Sea $X = \prod_{\alpha \in I} X_\alpha$ dotado de la topología producto y sea $\mathcal{A}$ una colección de subconjuntos de X con la propiedad de intersección finita. Vamos a demostrar que $\bigcap_{A \in \mathcal{A}} \overline{A} \neq \varnothing$. Para alcanzar ese objetivo usaremos el Lema de Zorn argumentando del modo siguiente: considere la colección $\mathfrak{D}_\mathcal{A}$ de todas las familias $\mathcal{D}$ de subconjuntos de X tales $\mathcal{A} \subseteq \mathcal{D}$ y $\mathcal{D}$ tiene la propiedad de intersección finita. Dotemos a $\mathfrak{D}_\mathcal{A}$ del orden parcial estricto $\subsetneqq$ y sea $\mathcal{C}$ una cadena en $\mathfrak{D}_\mathcal{A}$. Puesto que $\bigcup_{\mathcal{B} \in \mathcal{C}} \mathcal{B}$ es claramente una cota superior para los elementos de $\mathcal{C}$, el Lema de Zorn nos garantiza la existencia de un elemento maximal $\mathcal{D}_0 \in \mathfrak{D}_\mathcal{A}$. Será suficiente demostrar que

$$\bigcap_{D \in \mathcal{D}_0} \overline{D} \neq \varnothing.$$

Para verificar esto, tomemos un $\alpha \in I$ arbitrario y considere la colección

$$\mathcal{F}_\alpha = \{p_\alpha(D) : D \in \mathcal{D}_0\}.$$

de subconjuntos de X_α. Esta colección tiene la propiedad de intersección finita ya que $\mathcal{D}_0$ la tiene. Puesto que X_α es compacto, el Teorema 2.2.34 nos revela que $\bigcap_{D \in \mathcal{D}_0} \overline{p_\alpha(D)} \neq \varnothing$. Seleccione un punto $x_\alpha \in X_\alpha$ tal que

$$x_\alpha \in \bigcap_{D \in \mathcal{D}_0} \overline{p_\alpha(D)}.$$

Sea $x = (x_\alpha)_{\alpha\in I} \in X$ y veamos que $x \in \overline{D}$ para todo $D \in \mathcal{D}_0$. Observe que ello es equivalente a demostrar que si U es un subconjunto abierto arbitrario de X conteniendo a x, entonces $U \cap D \neq \varnothing$. En efecto, sea U_α un abierto en X_α con $x_\alpha \in U_\alpha$. Como $x_\alpha \in \overline{p_\alpha(D)}$, resulta que U_α intersecta a $p_\alpha(D)$ en algún punto $p_\alpha(y)$ para algún $y \in D$. De esto se sigue que $y \in p_\alpha^{-1}(U_\alpha) \cap D$ y entonces $V \cap D \neq \varnothing$ para cualquier $V \in \mathcal{E}_{\text{fin}}$ que contenga a x. Finalmente, como U es unión de una colección de conjuntos pertenecientes a $\mathcal{E}_{\text{fin}}$ el resultado sigue. Esto prueba que $\bigcap_{D\in\mathcal{D}_0} \overline{D} \neq \varnothing$ y termina la prueba. ■

En particular, si dotamos al conjunto $\{0,1\}$ de la topología discreta y hacemos $X_n = \{0,1\}$ para todo $n \in \mathbb{N}$, resulta que él es compacto y, entonces, $\prod_{n=1}^\infty X_n = 2^{\mathbb{N}}$ es compacto con la topología producto.

Si para cada $n \in \mathbb{N}$, el espacio topológico $(X_n, \mathfrak{J}_n)$ es 2º numerable, entonces $(\prod_{n=1}^\infty X_n, \mathfrak{T}_p)$ también es 2º numerable. En efecto, suponga que $\mathcal{B}_n$ es una base numerable de X_n para cada $n \geq 1$. Para cada conjunto finito $F \subseteq \mathbb{N}$, sea $U_F = \prod_{n=1}^\infty U_n$, donde $U_n \in \mathcal{B}_n$ para cada $n \in F$, mientras que $U_n = X_n$ si $n \notin F$. Si se considera la colección $\mathcal{B} \subseteq \mathfrak{T}_p$ formada por todos los productos de la forma U_F con $F \subseteq \mathbb{N}$ finito, entonces $\mathcal{B}$ es una base numerable para $\mathfrak{T}_p$.

2.2.5. El Espacio de Baire

El objetivo de esta sección es estudiar brevemente un espacio Polaco de particular importancia en la Teoría Descriptiva de Conjuntos, específicamente en el estudio de los conjuntos analíticos y que denotaremos por $\mathcal{N}$.

En lo que sigue escribiremos

$$\mathbb{N}^{<\infty} = \bigcup_{k=1}^\infty \mathbb{N}^k,$$

es decir, $\mathbb{N}^{<\infty}$ consiste de todas las **sucesiones finitas** de números naturales. Otras notaciones relevantes que usaremos son las siguientes: si $s \in \mathbb{N}^{<\infty}$, el símbolo $\ell(s)$ será usado para denotar la longitud de s, es decir, su número de elementos. Por ejemplo, si $s = (s_1, \ldots, s_k)$, entonces $\ell(s) = k$. Para cada par de elementos $s, t \in \mathbb{N}^{<\infty}$, digamos $s = (s_1, \ldots, s_k)$ y $t = (t_1, \ldots, t_m)$, convenimos en escribir $s \leq t$ para indicar que s es un segmento inicial de t, es decir, $t = (s_1, \ldots, s_k, t_{k+1}, \ldots, t_m)$. En este caso diremos que t es una **extensión** de s. Por ejemplo,

$$(4,7,3) \leq (4,7,3,9,2,1) \quad \text{pero} \quad (4,7,9) \nleq (4,7,3,9,2,1).$$

Escribiremos $s < t$, si $s \leq t$ pero $s \neq t$.

Para describir $\mathcal{N}$, debemos comenzar con el conjunto $\mathbb{N}$ al cual dotaremos de su topología discreta $\mathscr{P}(\mathbb{N})$. Para cada $n \in \mathbb{N}$, pongamos $\mathbb{N}_n = \mathbb{N}$ y considere el espacio producto $\prod_{n=1}^\infty \mathbb{N}_n = \mathbb{N}^{\mathbb{N}}$. Si dotamos a $\mathbb{N}^{\mathbb{N}}$ con la topología producto $\mathfrak{T}_p$, entonces al espacio topológico $(\mathbb{N}^{\mathbb{N}}, \mathfrak{T}_p)$ lo denotaremos brevemente por $\mathcal{N}$ y lo llamaremos el **espacio de Baire**. Puesto que $\mathbb{N}$ posee la topología discreta, todos sus subconjuntos son abiertos. En particular, para cada $m \in \mathbb{N}$, el conjunto $\{m\}$ es abierto y

$$p_m^{-1}(\{m\}) = \mathbb{N}_1 \times \cdots \mathbb{N}_{m-1} \times \{m\} \times \mathbb{N}_{m+1} \times \cdots$$

Como antes, sea

$$\mathcal{E} = \bigcup_{n=1}^\infty p_n^{-1}(\mathscr{P}(\mathbb{N})).$$

Sabemos que una base para la topología $\mathfrak{T}_p$ viene dada por la colección $\mathcal{E}_{\text{fin}}$ formada por todas las intersecciones de colecciones finitas de elementos de $\mathcal{E}$. Sin embargo, si se considera la colección $\mathcal{B}_p = \{\mathcal{O}_s : s \in \mathbb{N}^{<\infty}\} \subseteq \mathfrak{T}_p$, donde para cada $s = (s_1, \ldots, s_k) \in \mathbb{N}^{<\infty}$,

$$
\begin{aligned}
\mathcal{O}_s &= \prod_{i=1}^{k} p_{s_i}^{-1}(\{n_i\}) \\
&= \Big\{(n_k)_{k=1}^{\infty} \in \mathbb{N}^{\mathbb{N}} \,:\, n_1 = s_1, \ldots, n_k = s_k\Big\}.
\end{aligned}
$$

entonces $\mathcal{B}_p$ también es una base para $\mathfrak{T}_p$. En este sentido, un conjunto $U \subseteq \mathbb{N}^{\mathbb{N}}$ es $\mathfrak{T}_p$-**abierto**, o simplemente **abierto**, si para cada $x = (n_k)_{k=1}^{\infty} \in U$ existe un $k \in \mathbb{N}$ tal que $\mathcal{O}_{(n_1,\ldots,n_k)} \subseteq U$. Esto nos indica que los conjuntos abiertos de $\mathcal{N}$ son muy simples: ellos son exactamente los subconjuntos de $\mathbb{N}^{\mathbb{N}}$ que comienzan con conjunto finito prescrito de números naturales. Dicho de otra manera,

*$U \subseteq \mathbb{N}^{\mathbb{N}}$ es **abierto** si, y sólo si, existe un conjunto $S \subseteq \mathbb{N}^{<\infty}$ tal que $U = \bigcup_{s\in S} \mathcal{O}_s$.*

Observe que si $s, t \in \mathbb{N}^{<\infty}$, entonces

$$
s \leq t \quad \Rightarrow \quad \mathcal{O}_s \supseteq \mathcal{O}_t.
$$

Por otro lado, si no ocurre que $s \leq t$ ni tampoco que $t \leq s$, entonces debe existir algún $i \in \mathbb{N}$ para el cual $i < \text{mín}\{\ell(s), \ell(t)\}$ y $n_i \neq m_i$. En este caso, $\mathcal{O}_s \cap \mathcal{O}_t = \varnothing$. En tal situación escribiremos $s \perp t$.

Recordemos que un espacio topológico $(X, \mathcal{T})$ es **totalmente disconexo** si éste tiene una base cuyos elementos son abiertos y cerrados a la vez.

Teorema 2.2.51. *$\mathcal{N}$ es **totalmente disconexo**.*

Prueba. Veamos que cada conjunto en $\mathcal{B}_p$ es abierto y cerrado a la vez. Fijemos un elemento arbitrario $s = (s_1, \ldots, s_k) \in \mathbb{N}^{<\infty}$ y sea $\mathcal{O}_s \in \mathcal{B}_p$. Puesto que $\mathcal{O}_s$ es ciertamente abierto, sólo basta verificar que él es cerrado, pero esto es consecuencia inmediata de la igualdad

$$
\mathcal{N} \setminus \mathcal{O}_s = \bigcup \{\mathcal{O}_t \,:\, t \perp s\}.
$$

La prueba es completa. ■

Recordemos que un punto $x = (n_k)_{k=1}^{\infty} \in \mathcal{N}$ es un *punto límite* de $A \subseteq \mathcal{N}$, si $\mathcal{O}_s \cap A \neq \varnothing$ para todo abierto básico $\mathcal{O}_s \subseteq \mathcal{N}$ para el cual $x \in \mathcal{O}_s$. Por supuesto, esto significa que: para cada $k \in \mathbb{N}$, existe algún $y = (m_k)_{k=1}^{\infty} \in A$ con $x \neq y$ y $(n_1, \ldots, n_k) = (m_1, \ldots, m_k)$. También recordemos que un conjunto es cerrado si contiene a todos sus puntos límites.

Es fácil establecer que la aplicación $d : \mathbb{N}^{\mathbb{N}} \times \mathbb{N}^{\mathbb{N}} \to [0, +\infty)$ dada por

$$
d(x, y) = \begin{cases} 0 & \text{si} \quad x = y \\ 2^{-\ell(x,y)} & \text{si} \quad x \neq y, \quad \text{y} \quad \ell(x, y) = \text{ínf}\{n \in \mathbb{N} \,:\, x_n \neq y_n\}, \end{cases}
$$

siendo $x = (x_n)_{n=1}^{\infty}$ y $y = (y_n)_{n=1}^{\infty}$ elementos arbitrarios de $\mathbb{N}^{\mathbb{N}}$, es una **métrica** sobre $\mathbb{N}^{\mathbb{N}}$ que genera la topología producto. Más aun, $(\mathcal{N}, d)$ resulta ser un *espacio métrico completo* y *separable* el cual es de importancia fundamental en el estudio y desarrollo de los espacio Polacos y los conjuntos analíticos. En efecto, el conjunto

$$\mathcal{D} \;=\; \Big\{(n_k)_{k=1}^{\infty} \in \mathcal{N} \,:\, (n_k)_{k=1}^{\infty} \text{ es eventualmente } 0\Big\}$$

es subconjunto numerable y denso de $\mathcal{N}$. Para ver esto, recuerde que $(n_k)_{k=1}^{\infty} \in \mathcal{N}$ es **eventualmente constante** si existe un $N \in \mathbb{N}$ y un entero $m \geq 0$ tal que $n_k = m$ para todo $k \geq N$. Claramente $\mathcal{D}$ es numerable, de modo que sólo resta ver que él es denso. Pero ser denso en $\mathcal{N}$ significa que $\mathcal{D} \cap \mathcal{O}_s \neq \varnothing$ para cualquier $s \in \mathbb{N}^{<\infty}$ y, obviamente, si $s = (n_1, \ldots, n_k) \in \mathbb{N}^{<\infty}$, entonces $(n_1, \ldots, n_k, 0, 0, \ldots) \in \mathcal{D} \cap \mathcal{O}_s$. En particular, $\mathcal{N}$ es un espacio 2º **numerable**.

2.3. Problemas

(1) Sea A un conjunto acotado inferiormente y sea

$$B \;=\; \big\{b \in \mathbb{R} \,:\, b \text{ es una cota inferior de } A\big\}.$$

Demuestre que B está acotado superiormente y que $\sup B = \inf A$.

(2) Si $\sup A < \sup B$, demuestre que existe al menos un $b \in B$ tal que $\sup A < b$.

(3) Sean A y B subconjuntos acotados de $\mathbb{R}$. Pruebe que $A \cup B$ y $A \cap B$ están acotados y

$$\inf(A \cup B) \;=\; \min\{\inf A, \inf B\} \qquad \text{y} \qquad \sup(A \cup B) \;=\; \max\{\sup A, \sup B\}.$$

(4) Sea $(A_\alpha)_{\alpha \in I}$ una familia arbitraria de subconjuntos no vacíos de $\mathbb{R}$.

(a) Si para cada $\alpha \in I$, A_α está acotado superiormente y el conjunto $S_1 = \{\sup A_\alpha \,:\, \alpha \in I\}$ está acotado superiormente, entonces $\bigcup_{\alpha \in I} A_\alpha$ está acotado superiormente y

$$\sup\Big(\bigcup_{\alpha \in I} A_\alpha\Big) \;=\; \sup_{\alpha \in I}\big(\sup A_\alpha\big).$$

¿Se obtiene la misma conclusión si la hipótesis "S_1 está acotado superiormente" no se verifica?

(b) Si para cada $\alpha \in I$, A_α está acotado inferiormente y el conjunto $S_2 = \{\inf A_\alpha \,:\, \alpha \in I\}$ está acotado inferiormente, entonces $\bigcup_{\alpha \in I} A_\alpha$ está acotado inferiormente y

$$\inf\Big(\bigcup_{\alpha \in I} A_\alpha\Big) \;=\; \inf_{\alpha \in I}\big(\inf A_\alpha\big).$$

¿Se obtiene la misma conclusión si la hipótesis "S_2 está acotado inferiormente" no se verifica?

(5) Sea $(x_{mn})_{m,n=1}^{\infty}$ una sucesión doble en $\mathbb{R}$ tal que $\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}|x_{mn}| < +\infty$. Demuestre que para cualquier biyección $\sigma : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ se cumple que

$$\sum_{m=1}^{\infty}\Big(\sum_{n=1}^{\infty} x_{mn}\Big) \;=\; \sum_{n=1}^{\infty}\Big(\sum_{m=1}^{\infty} x_{mn}\Big) \;=\; \sum_{k=1}^{\infty} x_{\sigma(k)}.$$

(6) Sea $(x_n)_{n=1}^{\infty}$ una sucesión en $\mathbb{R}$ y sea $a \in \mathbb{R}$. Pruebe que:

   (a) Si $a < \varliminf_{n\to\infty} x_n$, entonces existe un $N \in \mathbb{N}$ tal que $a < x_n$ para todo $n \geq N$.

   (a') Si $\varliminf_{n\to\infty} x_n < a$, entonces $\{n : x_n < a\}$ es infinito.

   (b) Si existe un $N \in \mathbb{N}$ tal que $a < x_n$ para todo $n \geq N$, entonces $a \leq \varliminf_{n\to\infty} x_n$.

   (b') Si $\{n : x_n < a\}$ es infinito, entonces $\varliminf_{n\to\infty} x_n \leq a$.

(7) Sea $(x_n)_{n=1}^{\infty}$ una sucesión en $\mathbb{R}$ y sea $b \in \mathbb{R}$. Pruebe que:

   (a) Si $b > \varlimsup_{n\to\infty} x_n$, entonces existe un $N \in \mathbb{N}$ tal que $b > x_n$ para todo $n \geq N$.

   (a') Si $\varlimsup_{n\to\infty} x_n > b$, entonces $\{n : x_n > b\}$ es infinito.

   (b) Si existe un $N \in \mathbb{N}$ tal que $b > x_n$ para todo $n \geq N$, entonces $b \geq \varlimsup_{n\to\infty} x_n$.

   (b') Si $\{n : x_n > b\}$ es infinito, entonces $\varlimsup_{n\to\infty} x_n \geq b$.

(8) Sean $(x_n)_{n=1}^{\infty}$ y $(y_n)_{n=1}^{\infty}$ dos sucesiones en $\mathbb{R}$. Pruebe que:

   (a) Si $(y_n)_{n=1}^{\infty}$ converge, entonces

$$\varlimsup_{n\to\infty} (x_n + y_n) \;=\; \varlimsup_{n\to\infty} x_n + \varlimsup_{n\to\infty} y_n$$

   y

$$\varliminf_{n\to\infty} (x_n + y_n) \;=\; \varliminf_{n\to\infty} x_n + \varliminf_{n\to\infty} y_n.$$

   (b) Si las sucesiones $(x_n)_{n=1}^{\infty}$ y $(y_n)_{n=1}^{\infty}$ son no negativas, entonces

$$\varlimsup_{n\to\infty} (x_n y_n) \;\leq\; \Big(\varlimsup_{n\to\infty} x_n\Big)\Big(\varlimsup_{n\to\infty} y_n\Big)$$

   y

$$\varliminf_{n\to\infty} (x_n y_n) \;\geq\; \Big(\varliminf_{n\to\infty} x_n\Big)\Big(\varliminf_{n\to\infty} y_n\Big)$$

   (c) Si las sucesiones $(x_n)_{n=1}^{\infty}$ y $(y_n)_{n=1}^{\infty}$ son no negativas y $(y_n)_{n=1}^{\infty}$ converge, entonces

$$\varlimsup_{n\to\infty} (x_n y_n) \;=\; \Big(\varlimsup_{n\to\infty} x_n\Big)\Big(\varlimsup_{n\to\infty} y_n\Big)$$

   y

$$\varliminf_{n\to\infty} (x_n y_n) \;=\; \Big(\varliminf_{n\to\infty} x_n\Big)\Big(\varliminf_{n\to\infty} y_n\Big)$$

(9) Sea $(x_n)_{n=1}^{\infty}$ una sucesión y, para cada $n \geq 1$, defina

$$\sigma_n \;=\; \frac{x_1 + x_2 + \cdots + x_n}{n}.$$

   Pruebe que:

   (a) $\varliminf_{n\to\infty} x_n \leq \varliminf_{n\to\infty} \sigma_n \leq \varlimsup_{n\to\infty} \sigma_n \leq \varlimsup_{n\to\infty} x_n$.

   (b) Si $x_n \to x$, entonces $\sigma_n \to x$.

(10) Sean $\left(A_n\right)_{n=1}^{\infty}$ y $\left(B_n\right)_{n=1}^{\infty}$ sucesiones de subconjuntos de $\mathbb{R}$ y suponga que

$$A = \varliminf_{n\to\infty} A_n \qquad \text{y} \qquad B = \varlimsup_{n\to\infty} A_n.$$

Pruebe que

(a) $\displaystyle \varliminf_{n\to\infty} A_n = \left\{x \in \mathbb{R} : \sum_{n=1}^{\infty} \chi_{A_n^c}(x) < +\infty\right\},$

$\displaystyle \varlimsup_{n\to\infty} A_n = \left\{x \in \mathbb{R} : \sum_{n=1}^{\infty} \chi_{A_n}(x) = +\infty\right\}.$

(b) $\displaystyle \left(\varlimsup_{n\to\infty} A_n\right)^c = \varliminf_{n\to\infty} A_n^c.$

(c) $\displaystyle \chi_A = \varliminf_{n\to\infty} \chi_{A_n}$ y $\displaystyle \chi_B = \varlimsup_{n\to\infty} \chi_{A_n}.$

(d) $\displaystyle \varliminf_{n\to\infty} A_n \cap \varlimsup_{n\to\infty} B_n \subseteq \varlimsup_{n\to\infty} \left(A_n \cap B_n\right).$

(e) $\displaystyle \varlimsup_{n\to\infty} A_n \setminus \varliminf_{n\to\infty} A_n = \varlimsup_{n\to\infty} \left(A_n \triangle A_{n+1}\right).$

(f) Si $\lim\limits_{n\to\infty} A_n = A'$ y $\lim\limits_{n\to\infty} B_n = B'$, entonces

$$\lim_{n\to\infty} \left(A_n \times B_n\right) = A' \times B'.$$

(g) Sea $\left(F_n\right)_{n=1}^{\infty}$ una sucesión disjunta de subconjuntos cerrados de un espacio métrico completo (X, d). Demuestre que $\bigcup_{n=1}^{\infty} F_n$ es cerrado.

CAPÍTULO 3

Funciones Continuas

3.1. Propiedades Básicas

La noción de continuidad, como sabemos, es una de las más importantes y de mayor utilidad en matemáticas. En lo que sigue, supondremos que (X,d) es un espacio métrico o, más general, que (X,τ) es un espacio topológico de Hausdorff.

Definición 3.1.1. *Una función $f:(X,\tau)\to\mathbb{R}$ es **continua en un punto** $x_0\in X$ si, dado $\varepsilon>0$, existe un entorno abierto V_{x_0} de x_0 tal que $|f(x)-f(x_0)|<\varepsilon$ para cualquier $x\in V_{x_0}$.*

La función f se dice **continua en** X, o **sobre** X, si f es continua en todos los puntos de X. Si (X,d) es un espacio métrico, entonces podemos reemplazar el entorno abierto V_{x_0} de x_0 por alguna bola abierta $B(x_0,\delta)$. Una simple y útil caracterización de continuidad en términos de convergencia de sucesiones es el siguiente.

Teorema 3.1.2. *Sea $f:(X,d)\to\mathbb{R}$ una función y sea $x_0\in X$. Las siguientes condiciones son equivalentes:*

(1) *f es **continua en** x_0.*

(2) *Si $(x_n)_{n=1}^{\infty}$ es cualquier sucesión en X tal que $\lim_{n\to\infty} x_n = x_0$, entonces*

$$\lim_{n\to\infty} f(x_n) \;=\; f(x_0).$$

Prueba. Se deja a cargo del lector.

Otro aspecto importante respecto a la definición de continuidad en un punto tiene que ver con la noción de límites laterales. Suponga que $X=[a,b]$. Una función $f:[a,b]\to\mathbb{R}$ es **continua** en $x_0\in[a,b]$ significa que

$$\lim_{x\to x_0} f(x) \;=\; f(x_0)\in\mathbb{R},$$

en otras palabras,

(1) $f(x_0)$ existe,

(2) $\lim_{x\to x_0} f(x)$ existe, y

(3) $\lim_{x\to x_0} f(x) = f(x_0)$.

Recuerde que $\lim\limits_{x\to x_0} f(x)$ existe significa que $f(x_0^-) = f(x_0^+)$, donde

$$f(x_0^-) = \lim_{x\to x_0^-} f(x) \qquad \text{y} \qquad f(x_0^+) = \lim_{x\to x_0^+} f(x)$$

Para la continuidad en los puntos extremos, debe ocurrir que $f(a^+) = f(a)$ y $f(b^-) = f(b)$.

El símbolo $C(X)$ lo usaremos para denotar *el conjunto de todas las funciones* $f : X \to \mathbb{R}$ *son* ***continuas*** *sobre* X. Con las operaciones algebraicas usuales resulta que $C(X)$ es un poco más que un espacio vectorial sobre $\mathbb{R}$, esto es, si $f, g \in C(X)$ y $\alpha, \beta \in \mathbb{R}$, entonces

(a) $\alpha f + \beta g \in C(X)$,

(b) $fg \in C(X)$, y

(c) $f/g \in C(X)$ siempre que $g \neq 0$.

Una noción más fuerte que la de continuidad fue introducida por primera vez por Heinrich Eduard Heine quien la llamó continuidad uniforme.

Definición 3.1.3. *Una función* $f : (X, d) \to \mathbb{R}$ *es* ***uniformemente continua*** *sobre* X *si, dado* $\varepsilon > 0$, *existe un* $\delta > 0$, *dependiendo sólo de* ε, *con la siguiente propiedad: cualesquiera sean* $x, y \in X$ *satisfaciendo la condición* $d(x, y) < \delta$, *se cumple que* $|f(x) - f(y)| < \varepsilon$.

Es claro que toda función uniformemente continua es continua. Aunque el recíproco no siempre es cierto resulta que, en presencia de compacidad, *ambas nociones son indistinguibles, es decir, cuando* X *es un conjunto compacto cualquier función continua* $f : X \to \mathbb{R}$ *es uniformemente continua*, véase el Teorema 3.1.6. Una clase muy particular de funciones uniformemente continuas y que estudiaremos con más profundidad en el Capítulo 8 es el de función Lipschitz.

Definición 3.1.4. *Una función* $f : (X, d) \to \mathbb{R}$ *se dice que es* M-***Lipschitz****, o simplemente,* ***Lipschitz****, sobre* X, *si existe una constante* $M > 0$ *tal que cualesquiera sean* $x, y \in X$ *se cumple que*

$$|f(x) - f(y)| \le M \cdot d(x, y).$$

Es claro que toda función Lipschitz es uniformemente continua. Más aun, para cada **subconjunto cerrado** F de X, la función distancia $f(x) = \operatorname{dist}(x, F)$ para todo $x \in X$ es **Lipschitz** ya que

$$|f(x) - f(y)| = \big|\operatorname{dist}(x, F) - \operatorname{dist}(y, F)\big| \le d(x, y)$$

para todo $x, y \in X$.

Otra de las caracterizaciones topológicas de continuidad que no le confiere importancia alguna a ningún punto en particular del dominio de la función y que, por lo tanto, es de una gran utilidad es la siguiente.

Teorema 3.1.5. *Para cada función $f : (X, \tau) \to \mathbb{R}$, las siguientes condiciones son equivalentes:*

(1) *f es **continua** sobre X.*

(2) *Para cualquier conjunto abierto V en $\mathbb{R}$, existe un abierto $G \subseteq X$ tal que $f(G) \subseteq V$.*

(3) *Para cualquier conjunto abierto V en $\mathbb{R}$, $f^{-1}(V)$ es abierto en X.*

(4) *Para cualquier conjunto cerrado F en $\mathbb{R}$, $f^{-1}(F)$ es cerrado en X.*

(5) *Para cualquier conjunto $E \subseteq X$, $f(\overline{E}) \subseteq \overline{f(E)}$.*

Prueba. Ejercicio.

3.1.1. Funciones Continuas con Soportes Compactos

Como antes, X representará un espacio métrico o un espacio topológico. El símbolo $C_b(X)$ será usado para designar *las funciones en $C(X)$ que son **acotadas** sobre X*. Este conjunto también es un espacio vectorial sobre $\mathbb{R}$.

Cuando el dominio de una función continua es un conjunto compacto, las propiedades de tal función se convierten en una envidia para el resto de las funciones continuas que no comparten ese dominio. En lo que sigue, veremos algunas de tales propiedades. Comencemos con la primera.

Teorema 3.1.6 (Heine). *Sea (X, d) un espacio métrico y sea K un subconjunto **compacto** de X. Para cualquier función $f : K \to \mathbb{R}$ las siguientes condiciones son equivalentes:*

(1) *f es **continua** sobre K.*

(2) *f es **uniformemente continua** sobre K.*

Prueba. Observe que $(2) \Rightarrow (1)$ independientemente si el dominio de f es, o no, un conjunto compacto. Veamos la implicación $(1) \Rightarrow (2)$. Suponga, para construir una contradicción, que f es continua pero no uniformemente continua. Esto, por supuesto, significa que existe algún $\varepsilon > 0$ con la siguiente propiedad: para cualquiera elección arbitraria de $\delta > 0$, se puede determinar un par de puntos $x_\delta, y_\delta \in X$ con $d(x_\delta, y_\delta) < \delta$ pero tal que $|f(x_\delta) - f(y_\delta)| \geq \varepsilon$. Si para cada $n \in \mathbb{N}$, seleccionamos $\delta = 1/n$, entonces lo anterior nos permite la construcción de un par de sucesiones $(x_n)_{n=1}^\infty$ y $(y_n)_{n=1}^\infty$ en K tales que, para todo $n \in \mathbb{N}$,

$$d(x_n, y_n) < 1/n \qquad \text{pero} \qquad |f(x_n) - f(y_n)| \geq \varepsilon$$

Como K es compacto, podemos extraer una subsucesión $(x_{n_k})_{k=1}^\infty$ de $(x_n)_{n=1}^\infty$ que converge a algún punto $z \in K$. Puesto que

$$d(z, y_{n_k}) \ \leq \ d(z, x_{n_k}) + d(x_{n_k}, y_{n_k}) \ < \ d(z, x_{n_k}) + \frac{1}{n_k}, \qquad \forall k \geq 1$$

se concluye que la subsucesión $(y_{n_k})_{k=1}^\infty$ de $(y_n)_{n=1}^\infty$ también converge a z. Veamos que esto conduce a un imposible. En efecto, por un lado, como f es continua, resulta que

$$\lim_{k\to\infty} \big|f(x_{n_k}) - f(z)\big| \ = \ 0 \qquad \text{y} \qquad \lim_{k\to\infty} \big|f(y_{n_k}) - f(z)\big| \ = \ 0. \tag{1}$$

Mientras que, por el otro lado, para todo $k \geq 1$ se cumple que

$$0 \;<\; \varepsilon \;\leq\; \big|f(x_{n_k}) - f(y_{n_k})\big| \;\leq\; \big|f(x_{n_k}) - f(z)\big| + \big|f(y_{n_k}) - f(z)\big|.$$

Esto último nos indica que uno de los límites

$$\lim_{k\to\infty} \big|f(x_{n_k}) - f(z)\big| \qquad \text{o} \qquad \lim_{k\to\infty} \big|f(y_{n_k}) - f(z)\big|$$

debe ser estrictamente positivo, lográndose de este modo un disparate en relación a (1). Esta contradicción nos convence de que f es uniformemente continua y finaliza la prueba. ■

El siguiente resultado establece que compacidad se preserva por imágenes continuas, es decir:

Teorema 3.1.7. *Sea (X,τ) un espacio topológico y suponga que $K \subseteq X$ es **compacto**. Si $f : X \to \mathbb{R}$ es una **función continua** sobre X, entonces $f(K)$ es **compacto**.*

Prueba. Sea $\mathcal{V} = \{V_\alpha : \alpha \in I\}$ un cubrimiento abierto de $f(K)$. Puesto que f es continua sobre X, la colección $\mathcal{U} = \{f^{-1}(V_\alpha) : \alpha \in I\}$ es un cubrimiento abierto de K el cual, por ser un conjunto compacto, se reduce a un subcubrimiento finito, es decir, existen $\alpha_1, \ldots, \alpha_n$ tal que $K \subseteq \bigcup_{i=1}^n f^{-1}(V_{\alpha_i})$. De esto se sigue que $f(K) \subseteq \bigcup_{i=1}^n V_{\alpha_i}$ y termina la prueba. ■

Puesto que todo conjunto compacto es cerrado y acotado, se sigue del resultado anterior que *si X es compacto, cualquier función continua $f : X \to \mathbb{R}$ es **acotada** sobre X*. Por esta razón, si X es compacto, entonces $C(X) = C_b(X)$. Esto permite definir, para cada $f \in C(X)$, el número

$$\| f \|_\infty \;=\; \sup\{|f(x)| \;:\; x \in X\}.$$

A $\| f \|_\infty$ se le llama la **norma uniforme** o **norma sup** de f. La familia de todos los conjuntos de la forma

$$B(f,r) \;=\; \{g \in C_b(X) \;:\; \| g - f \|_\infty < r\}$$

donde $f \in C_b(X)$ y $r \in \mathbb{R}^+$, constituye una topología sobre $C_b(X)$ llamada la **topología uniforme**. Es un ejercicio sencillo demostrar que $(C_b(X), \|\cdot\|_\infty)$ es un **espacio de Banach**.

Recordemos que un espacio topológico de Hausdorff (X,τ) se dice **localmente compacto** si cada $x \in X$ posee un entorno abierto V_x de x cuya clausura es compacta, es decir, tal que $\overline{V}_x$ es un conjunto compacto.

Definición 3.1.8. *Sea (X,τ) un espacio localmente compacto y sea $f : X \to \mathbb{R}$ una función. El **soporte** de f se define como el conjunto*

$$\operatorname{sop}(f) \;=\; \overline{\{x \in X : f(x) \neq 0\}}.$$

Puesto que $(\overline{A})^c = \operatorname{int}(A^c)$, resulta que

$$\begin{aligned} x \notin \operatorname{sop}(f) \quad &\Leftrightarrow \quad x \in \operatorname{int}\big(\{x \in X \;:\; f(x) = 0\}\big) \\ &\Leftrightarrow \quad \text{existe un entorno abierto } V \text{ de } x \text{ tal que } f(z) = 0 \text{ para todo } z \in V. \end{aligned}$$

De esta observación se deduce que si V_0 es la unión de todos los conjuntos abiertos V_x que contienen a x y para los cuales $f|_{V_x} = 0$, resulta que *V_0 es el conjunto abierto más grande sobre el cual f se anula*. Por esto,

$$\operatorname{sop}(f) \;=\; X \setminus V_0$$

y, por lo tanto, si $x \notin \operatorname{sop}(f)$, entonces $f(x) = 0$. Sin embargo, lo contrario no es, en general, cierto. Por ejemplo, si $f : \mathbb{R} \to \mathbb{R}$ se define como $f(x) = x$, resulta que $\operatorname{sop}(f) = \mathbb{R}$ y $f(0) = 0$.

Definición 3.1.9. *Sea (X, τ) un espacio localmente compacto. Una función continua $f : X \to \mathbb{R}$ se dice que tiene **soporte compacto** sobre X si su soporte, $\operatorname{sop}(f)$, es un conjunto compacto.*

En lo que sigue, el símbolo $C_c(X)$ será usado para denotar *el conjunto de todas las **funciones continuas** $f : X \to \mathbb{R}$ con **soporte compacto**.*

Corolario 3.1.10. *Si $f \in C_c(X)$, entonces existe un **intervalo compacto** $[a, b] \supseteq \operatorname{sop}(f)$ tal que*

$$f(x) \;=\; 0 \quad \text{para todo } x \notin [a, b].$$

Prueba. Sea $f \in C_c(X)$. Si definimos

$$a \;=\; \inf \operatorname{sop}(f) \qquad \text{y} \qquad b \;=\; \sup \operatorname{sop}(f),$$

entonces $[a, b] \supseteq \operatorname{sop}(f)$ y $f(x) = 0$ para todo $x \notin [a, b]$. ■

Con las operaciones usuales de adición y multiplicación por un escalar se comprueba sin dificultad que $C_c(X)$ es un espacio vectorial sobre $\mathbb{R}$. De hecho, $C_c(X)$ es un **retículo vectorial**, en el siguiente sentido: si $f, g \in C_c(X)$, entonces también pertenecen a $C_c(X)$ las funciones

$$f \wedge g \;=\; \min\{f, g\} \qquad \text{y} \qquad f \vee g \;=\; \max\{f, g\}.$$

De esto último se sigue que si $f \in C_c(X)$, entonces $|f| \in C_c(X)$. Para ver esto, recordemos que las partes positiva y negativa de f se definen como

$$f^+ \;=\; \max\{f, 0\} \qquad \text{y} \qquad f^- \;=\; \min\{-f, 0\}$$

y de las relaciones

$$\begin{aligned} |f| \;&=\; f^+ + f^-, & f \;&=\; f^+ - f^-, \\ \max\{f, g\} \;&=\; \frac{1}{2}\big[(f+g) + |f-g|\big], & \min\{f, g\} \;&=\; \frac{1}{2}\big[(f+g) - |f-g|\big] \end{aligned}$$

se sigue que $|f| \in C_c(X)$.

El siguiente resultado garantiza la existencia de funciones continuas con soporte compacto bajo condiciones muy sencillas.

Lema 3.1.11 (Urysohn). *Sea (X, d) un espacio métrico y suponga que existe un conjunto **compacto** K y un conjunto **abierto** G tal que $K \subseteq G \subseteq X$. Entonces existe una función $f \in C_c(X)$ tal que*

(a) $f(x) = 1$ para todo $x \in K$,

(b) $f(x) = 0$ para todo $x \in G^c$, y

(c) $0 \leq f(x) \leq 1$ para todo $x \in \mathbb{R}$.

Prueba. La función f puede ser definida explícitamente del modo siguiente:

$$f(x) \;=\; \frac{\operatorname{dist}(x, G^c)}{\operatorname{dist}(x, K) + \operatorname{dist}(x, G^c)}, \quad x \in X.$$

Puesto que K y G^c son conjuntos cerrados y disjuntos, el denominador de f nunca es cero. Más aun, f es continua sobre X por ser el cociente de dos funciones continuas y claramente ella satisface las condiciones $(a) - (c)$. Por supuesto, $\operatorname{sop}(f) \subseteq G$ y, así, $f \in C_c(X)$. ■

En general, el Lema de Urysohn es válido si X es un *espacio localmente compacto* aunque su demostración es mucho más elaborada.

Corolario 3.1.12. *Sean (X, d) un espacio métrico y F un* ***subconjunto cerrado*** *de X. Para cada $x_0 \in X$ con $x_0 \notin F$, existe $f \in C_c(X)$ tal que $0 \leq f(x) \leq 1$ para todo $x \in X$ con $f(x_0) = 1$.*

Prueba. Tomando $K = \{x_0\}$ y $G = F^c$, resulta que K es compacto, G es abierto y $K \subseteq G$. Se sigue entonces del Lema de Urysohn que existe $f \in C_c(X)$ con las propiedades señaladas. ■

En general, valen las inclusiones siguientes

$$C_c(X) \subseteq C_b(X) \subseteq C(X),$$

y que si X es **compacto**, entonces $C_c(X) = C_b(X) = C(X)$. Nótese que la función $f(x) = e^{-x^2}$ es continua pero no pertenece a $C_c(\mathbb{R})$, aunque es fácil construir una sucesión $(f_n)_{n=1}^{\infty}$ en $C_c(\mathbb{R})$ tal que $f_n \to f$ puntualmente. Un característica importante que posee f es la siguiente: $\lim_{|x|\to\infty} f(x) = 0$, lo cual significa que: dado cualquier $\varepsilon > 0$ existe un $N \in \mathbb{N}$ tal que $|f(x)| < \varepsilon$ siempre que $|x| > N$. Si tomamos $K = [-N, N]$, lo anterior se puede expresar en la forma $|f(x)| < \varepsilon$ para todo $x \in K^c$.

Definición 3.1.13. *Sea (X, τ) un espacio localmente compacto. Una función continua $f : X \to \mathbb{R}$ se* ***anula en el infinito*** *si para cada $\varepsilon > 0$, existe un compacto $K \subseteq X$ tal que*

$$|f(x)| \;<\; \varepsilon \quad \text{para todo } x \in X \setminus K.$$

En lo sucesivo, el símbolo $C_0(\mathbb{R})$ designará *el conjunto de todas las funciones continua $f \in C(X)$ que se* ***anulan en el infinito***. Es fácil establecer que

$$C_c(X) \subseteq C_0(X) \subseteq C_b(X) \subseteq C(X)$$

y que $C_0(X)$ es un **subespacio cerrado** de $(C_b(X), \|\cdot\|_\infty)$. En particular, un espacio de Banach con la norma del supremo. El ejemplo de la función $f(x) = e^{-x^2}$, muestra que $C_0(\mathbb{R})$ no coincide con $C_c(\mathbb{R})$.

Teorema 3.1.14. *La clausura de $C_c(\mathbb{R})$ en la* ***topología uniforme*** *es igual a $C_0(\mathbb{R})$, es decir,*

$$\overline{C_c(\mathbb{R})}^{\|\cdot\|_\infty} \;=\; C_0(\mathbb{R}).$$

Prueba. Es claro que $\overline{C_c(\mathbb{R})}^{\|\cdot\|_\infty} \subseteq C_0(\mathbb{R})$. Para demostrar la otra inclusión, sea $f \in C_0(\mathbb{R})$ y fijemos un $\varepsilon > 0$. Entonces existe un compacto $K \subseteq \mathbb{R}$ tal que $|f(x)| < \varepsilon$ si $x \notin K$. Sea G cualquier conjunto abierto tal que $K \subseteq G$. Por el Lema de Urysohn, existe una función $\varphi \in C_c(\mathbb{R})$ tal que

(a) $\varphi(x) = 1$ para todo $x \in K$,

(b) $\varphi(x) = 0$ para todo $x \in G^c$, y

(c) $0 \leq \varphi(x) \leq 1$ para todo $x \in \mathbb{R}$.

Defina $g = f\varphi$. Entonces $\operatorname{sop}(g) \subseteq \operatorname{sop}(\varphi)$, $g \in C_c(\mathbb{R})$ y se cumple que

$$|f(x) - g(x)| \;=\; |f(x) - f(x)\varphi(x)| \;=\; |f(x)||1 - \varphi(x)|$$

el cual vale 0 si $x \in K$ y es $< \varepsilon$ si $x \notin K$. Por lo tanto, $\| f - g \|_\infty < \varepsilon$ y termina la prueba. ■

Teorema 3.1.15. *Si $f \in C_c(\mathbb{R})$, entonces f es **uniformemente continua** sobre $\mathbb{R}$.*

Prueba. Puesto que $\operatorname{sop}(f)$ es un conjunto compacto, él es cerrado y acotado y, en consecuencia, existen números reales a, b tales que $\operatorname{sop}(f) \subseteq [a, b]$. En particular, por ser f continua sobre $\mathbb{R}$, ella es uniformemente continua sobre $[a, b]$. Por otro lado, como $f = 0$ sobre $\operatorname{sop}(f)^c$ y $[a, b]^c \subseteq \operatorname{sop}(f)^c$, resulta que $f = 0$ sobre $[a, b]^c$. De esto se sigue que f también es uniformemente continua sobre $[a, b]^c$ y, por lo tanto, uniformemente continua sobre $[a, b] \cup [a, b]^c = \mathbb{R}$. ■

3.1.2. Más sobre Funciones Continuas

Sea (X, d) un espacio métrico. Sabemos que toda función continua $f : X \to \mathbb{R}$ transforma sucesiones convergentes en sucesiones convergentes; sin embargo, no ocurre lo mismo para sucesiones de Cauchy, es decir, f no necesariamente transforma sucesiones de Cauchy en sucesiones de Cauchy. Por otro lado, si X es **compacto**, resulta que que *toda función continua $f : X \to \mathbb{R}$ transforma sucesiones de Cauchy en sucesiones de Cauchy*. La razón de fondo de tal hecho radica en que, en este caso, dicha función es uniformemente continua y, por lo tanto, la afirmación anterior se expresa en la forma: *toda función **uniformemente continua** transforma sucesiones de Cauchy en sucesiones de Cauchy*. Otra de las buenas razones de por qué las funciones uniformemente continuas son importantes lo constituye el siguiente resultado:

Teorema 3.1.16 (Extensión Continua). *Sea (X, d) un espacio métrico y suponga que D es un subconjunto no vacío de X. Si $f : D \to \mathbb{R}$ es una función **uniformemente continua**, entonces existe una **única función** $F : \overline{D} \to \mathbb{R}$ tal que*

*(a) F es **uniformemente continua** sobre $\overline{D}$ y*

(b) $F(x) = f(x)$ para todo $x \in D$.

Prueba. Nuestro primer paso es la construcción de F. Para hacer eso, tomemos cualquier $x \in \overline{D}$ y elija una sucesión $(x_n)_{n=1}^{\infty}$ en D tal que $x_n \to x$. Puesto que $(x_n)_{n=1}^{\infty}$ es una sucesión de Cauchy y f es uniformemente continua sobre D, la sucesión $(f(x_n))_{n=1}^{\infty}$ es de Cauchy en $\mathbb{R}$ y, por consiguiente, converge en $\mathbb{R}$. Esto permite definir $F : \overline{D} \to \mathbb{R}$ por

$$F(x) \;=\; \lim_{n\to\infty} f(x_n)$$

para cada $x \in \overline{D}$. Debemos ahora verificar que esta es una buena definición, es decir, no depende de la elección de la sucesión que converge a x. En efecto, sea $(z_n)_{n=1}^{\infty}$ otra sucesión en D, distinta de la sucesión $(x_n)_{n=1}^{\infty}$, que también converge a x. Veamos que

$$\lim_{n\to\infty} f(x_n) \;=\; \lim_{n\to\infty} f(z_n).$$

Fijemos $\varepsilon > 0$ elegido de manera arbitraria. Puesto que f es uniformemente continua sobre D, existe un $\delta > 0$ tal que cualesquiera sean $u, v \in D$ con $d(u, v) < \delta$ se cumple que $|f(u) - f(v)| < \varepsilon$. Ahora bien, como $d(x_n, z_n) \to 0$, resulta que para el δ elegido, existe un $N \in \mathbb{N}$ tal que $d(x_n, z_n) < \delta$ para todo $n \geq N$. Se sigue de lo anterior que

$$|f(x_n) - f(z_n)| \;<\; \varepsilon \quad \text{para todo } n \geq N,$$

lo cual significa que $\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} f(z_n)$. Con esto hemos demostrado la existencia de F.

(i) $F(x) = f(x)$ para todo $x \in D$. En efecto, sea $x \in D$. Tomando $x_n = x$ para todo $n \geq 1$, vemos que
$$F(x) \;=\; \lim_{n\to\infty} f(x_n) \;=\; f(x).$$

(ii) F es uniformemente continua sobre $\overline{D}$. Dado $\varepsilon > 0$, usemos el hecho de que f es uniformemente continua sobre D para elegir un $\delta > 0$ tal que

$$\text{si} \quad u, v \in D \quad \text{con} \quad d(u,v) < \delta \quad \text{entonces} \quad |f(u) - f(v)| < \varepsilon/3. \tag{1}$$

Sean $x, z \in \overline{D}$ tal que $d(x,z) < \delta/3$. Para demostrar que $|F(x) - F(y)| < \varepsilon$, seleccionemos sucesiones $(x_n)_{n=1}^{\infty}$ y $(y_n)_{n=1}^{\infty}$ en D tales que $x_n \to x$ y $z_n \to z$. Escojamos ahora un $N_1 \in \mathbb{N}$ de modo tal que si $n \geq N_1$, entonces

$$d(x_n, x) \;<\; \delta/3 \qquad\qquad \text{y} \qquad\qquad d(z_n, z) \;<\; \delta/3.$$

En particular,

$$d(x_n, z_n) \;\leq\; d(x_n, x) + d(x,z) + d(z_n, z) \;<\; \delta \quad \text{para todo} \;\; n \geq N_1$$

y así, por (1), $|f(x_n) - f(z_n)| < \varepsilon/3$ para todo $n \geq N_1$. De igual modo, como $f(x_n) \to F(x)$ y $f(z_n) \to F(z)$, existe un $N_2 \in \mathbb{N}$ tal que

$$|F(x_n) - F(x)| \;<\; \varepsilon/3 \qquad\qquad \text{y} \qquad\qquad |F(z_n) - F(z)| \;<\; \varepsilon/3.$$

para todo $n \geq N_2$. Sea $N = \max\{N_1, N_2\}$. De lo anterior se sigue que

$$\begin{aligned}
|F(x) - F(z)| \;&\leq\; |F(x) - F(x_N)| + |F(x_N) - F(z_N)| + |F(z_N) - F(z)| \\
&=\; |F(x) - F(x_N)| + |f(x_N) - f(z_N)| + |F(z_N) - F(z)| \\
&<\; \varepsilon/3 + \varepsilon/3 + \varepsilon/3 \;=\; \varepsilon
\end{aligned}$$

lo cual prueba que F es uniformemente continua sobre $\overline{D}$.

(iii) Para verificar que F es única, suponga que G es otra extensión uniformemente continua de f sobre $\overline{D}$ y sea $x \in \overline{D}$. Entonces existe una sucesión $(x_n)_{n=1}^{\infty}$ en D tal que

$$F(x) \;=\; \lim_{n\to\infty} f(x_n) \;=\; \lim_{n\to\infty} G(x_n) \;=\; G(x).$$

La prueba es completa. ■

Una consecuencia inmediata del resultado anterior en combinación con el Teorema 3.1.7 es el siguiente.

Corolario 3.1.17. *Sea $K \subseteq \mathbb{R}^n$ **acotado**. Si $f : K \to \mathbb{R}$ es una función **uniformemente continua** sobre K, entonces $f(K)$ es **acotado**.*

Prueba. Por el Teorema de Extensión Continua, existe una única aplicación uniformemente continua $F : \overline{K} \to \mathbb{R}$ que es una extensión de f. Como K es acotado, también lo es $\overline{K}$ y así, $\overline{K}$ es compacto. Se sigue del Teorema 3.1.7 que $F(\overline{K})$ es compacto y, por lo tanto, acotado. Finalmente, como
$$f(K) \;=\; F(K) \;\subseteq\; F(\overline{K}),$$

resulta que $f(K)$ es acotado. ■

Sean (A, τ_1) y (B, τ_2) espacios topológicos. Una función $f : A \to B$ se dice que es un **homeomorfismo** de A sobre B si ella es biyectiva, continua y su inversa $f^{-1} : B \to A$ también es continua. En este caso se dice que A y B son **homeomorfos** . Si (A, d_1) y (B, d_2) son espacios métricos y $f : A \to B$ es un aplicación biyectiva y continua tal que $d_1(x, y) = d_2(f(x), f(y))$ para todo $x, y \in A$, entonces diremos que A y B son **isométricos** y a la función f se le llama una **isometría sobreyectiva**. Es importante destacar que una función continua y biyectiva no es necesariamente un homeomorfismo. Sin embargo, si su dominio es compacto, se obtiene el siguiente resultado:

Teorema 3.1.18. *Sean K_1 y K_2 espacios topológicos tal que K_1 **es compacto**. Entonces, toda **biyección continua** $f : K_1 \to K_2$ es un **homeomorfismo**.*

Prueba. Sólo tenemos que demostrar que la función inversa $f^{-1} : K_2 \to K_1$ es continua en K_2, es decir, si $F \subseteq K_1$ es cerrado, entonces $\left(f^{-1}\right)^{-1}(F) = f(F)$ es cerrado en K_2. Suponga que F es un subconjunto cerrado en K_1. Como K_1 es compacto, el Teorema 2.2.29 nos garantiza que F es compacto y la continuidad de f nos revela, gracias al Teorema 3.1.7, que $f(F)$ es compacto en K_2. Por una nueva aplicación del Teorema 2.2.29 tenemos que $f(F)$ es cerrado en K_2. ■

Definición 3.1.19. *Sea $f : [a, b] \to \mathbb{R}$ una función. Diremos que f es una **función de Darboux**, o posee la **Propiedad del Valor Intermedio** (PVI) si, para cualesquiera $x, y \in [a, b]$ con $x < y$ y cualquier número c entre $f(x)$ y $f(y)$, existe un $t \in (x, y)$ tal que $f(t) = c$.*

En realidad, uno puede reemplazar el intervalo $[a, b]$ en la definición anterior, por cualquier intervalo $I \subseteq \mathbb{R}$.

Teorema 3.1.20 (Bolzano, Teorema del Valor Intermedio). *Si $f : [a, b] \to \mathbb{R}$ es una función **continua** sobre $[a, b]$, entonces f posee la **PVI**.*

Prueba. Puesto que f es continua en el compacto $[a, b]$ ella es acotada. Sean

$$m = \inf\{f(x) : x \in [a, b]\} \qquad \text{y} \qquad M = \sup\{f(x) : x \in [a, b]\}.$$

Si $m = M$, entonces f es constante y la conclusión es inmediata. Suponga que $m < M$ y sea c tal que $m < c < M$. Las propiedades del ínfimo y del supremo implican la existencia de puntos a_1, b_1 en $[a, b]$ tales que

$$m \leq f(a_1) < c < f(b_1) \leq M.$$

Para simplificar la presentación supondremos que $a_1 < b_1$ (el caso $a_1 > b_1$ se trata de modo similar). Defina

$$E = \{x \in [a, b] : f(x) \leq c\}.$$

Observe que $E \neq \varnothing$ ya que $a_1 \in E$ y, además, está acotado superiormente por b. Sea $t = \sup E$. Para cada $n \in \mathbb{N}$, las propiedades de éste supremo indican la existencia de un elemento $x_n \in E$ tal que $t - 1/n < x_n \leq t$ y, por consiguiente, la sucesión $(x_n)_{n=1}^{\infty}$ converge a t, y entonces, por continuidad, $f(x_n) \to f(t)$. Pero como $x_n \in E$ resulta que $f(x_n) \leq c$ y, así,

$$f(t) = \lim_{n\to\infty} f(x_n) \leq c.$$

Por otra parte, puesto que t es el supremo de E, entonces cualquier $x \in (t, b]$ es cota superior de E y, en consecuencia, $f(x) > c$. De nuevo, por continuidad,

$$f(t) = \lim_{x \to t^+} f(x) \geq c,$$

de donde se concluye que $f(t) = c$. La prueba es completa. ■

En el siglo XIX muchos matemáticos compartían la creencia de que la Propiedad del Valor Intermedio era equivalente a continuidad. Sin embargo, en el año 1875, el matemático francés Jean Gaston Darboux (1842-1917) probó que esa creencia era totalmente falsa, es decir, el recíproco del Teorema de Valor Intermedio es *falso*. Para ver esto, considere la aplicación $f : \mathbb{R} \to \mathbb{R}$ definida por

$$f(x) = \begin{cases} \operatorname{sen}(1/x) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0. \end{cases}$$

Es claro que f es discontinua en $x = 0$. Sin embargo, puesto que el rango de f es el intervalo $[-1, 1]$ se tiene que f posee la **PVI**.

El resultado anterior establece que la **PVI** no es una propiedad exclusiva de las funciones continuas. De hecho, existen funciones que poseen la **PVI** que no son continuas en ningún punto de su dominio (véase el Corolario 6.5.36, página 342).

Teorema 3.1.21 (Darboux). *Si $f : [a, b] \to \mathbb{R}$ es una función **diferenciable** en $[a, b]$, entonces f' posee la **PVI**.*

Prueba. Sean $x, y \in [a, b]$ con $x < y$ y suponga que c es cualquier punto entre $f'(x)$ y $f'(y)$. Sin perder generalidad, podemos asumir que $f'(x) > f'(y)$. Defina la función $G : [x, y] \to \mathbb{R}$ por $G(t) = f(t) - ct$. Puesto que G es diferenciable sobre $[a, b]$, ella es continua sobre dicho intervalo y, en consecuencia, alcanza su máximo en algún punto $t_0 \in [x, y]$. En particular, $G'(t_0) = 0$. Observe ahora que $G'(x) = f'(x) - c > 0$, por lo que $t_0 \neq x$. Similarmente, $G'(y) = f'(y) - c < 0$ y, entonces, $t_0 \neq y$. Esto nos garantiza que $t_0 \in (x, y)$ y, por lo tanto, $0 = G'(t_0) = f'(t_0) - c$. ■

Uno entonces se pregunta, ¿bajo qué condiciones adicionales, si es que existen, la **PVI** implica continuidad? La respuesta viene dada por el siguiente resultado.

Teorema 3.1.22. *Sea $f : [a, b] \to \mathbb{R}$ una **función inyectiva** la cual satisface la* PVI *sobre $[a, b]$. Entonces f es **continua** sobre $[a, b]$.*

Prueba. Nuestra primera tarea es demostrar que f es estrictamente monótona sobre $[a, b]$. Para ver esto, sean $x_1, x_2 \in [a, b]$ con $x_1 < x_2$. Puesto que f es inyectiva, resulta que $f(x_1) \neq f(x_2)$. Suponga que $f(x_1) < f(x_2)$. Si $x_1 < u < x_2$, entonces se debe cumplir que $f(x_1) < f(u) < f(x_2)$. En efecto, de ocurrir la desigualdad $f(u) > f(x_2)$, tendríamos las desigualdades $f(x_1) < f(x_2) < f(u)$ y entonces la **PVI** nos proporcionará un $x \in (x_1, u)$ tal que $f(x) = f(x_2)$ lo que negaría la hipótesis de que f es inyectiva. Similarmente, la desigualdad $f(u) < f(x_1)$ no puede ocurrir.

Para demostrar la continuidad de f asumiremos que ella es estrictamente creciente sobre $[a, b]$ y tomemos $x \in [a, b]$. Si f es discontinua en x, entonces debemos tener que $f(x^-) < f(x)$ o $f(x) < f(x^+)$, donde $f(x^-)$ y $f(x^+)$ son, respectivamente, los límites laterales por la izquierda y por la derecha de f en x. Si, por ejemplo, ocurre que $f(x) < f(x^+)$, entonces la **PVI** no se cumple sobre $[x, x+\delta]$ para cualquier $\delta > 0$. Por esto, debemos tener que $f(x) = f(x^+)$. De modo similar se tiene que $f(x) = f(x^-)$ y termina la prueba. ■

Una fácil aplicación de la Propiedad del Valor Intermedio para funciones continuas es el siguiente el hecho:

Corolario 3.1.23. *Si $f : [a,b] \to \mathbb{R}$ es una función **continua estrictamente creciente** sobre $[a,b]$, entonces $f(J)$ es un intervalo para cualquier intervalo $J \subseteq [a,b]$.*

Prueba. Suponga que $J = (x,y) \subseteq [a,b]$ y sea $x < t < y$. Como f es estrictamente creciente sobre $[a,b]$, se tiene que $f(x) < f(t) < f(y)$ y, por lo tanto,

$$f((x,y)) \subseteq (f(x), f(y)).$$

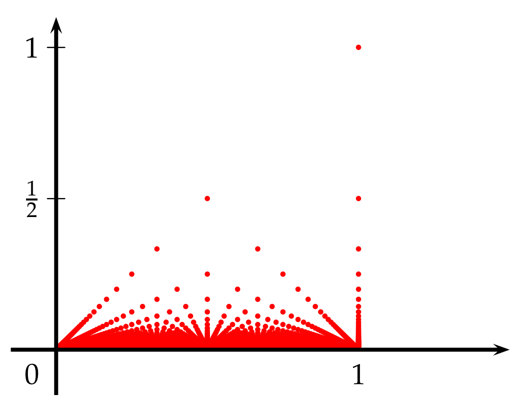
Por otro lado, si $\xi \in (f(x), f(y))$, entonces por ser f continua sobre $[a,b]$, la Propiedad del Valor Intermedio nos garantiza la existencia de un $t \in (x,y)$ tal que $f(t) = \xi$. Esto nos indica que $(f(x), f(y)) \subseteq f((x,y))$ y termina la prueba. ■

La siguiente función es una modificación de la función de Dirichlet. Es un ejemplo interesante de *una función que es continua en los irracionales pero discontinua en los racionales*. Esto, por supuesto, constituye un hecho que sorprende a la imaginación. Pero si ese hecho resulta curioso, no menos curioso es este otro: el Teorema de Categoría de Baire garantiza (véase la consecuencia **TBC**(7) en la página 140) que *es imposible construir una función $f : [0,1] \to \mathbb{R}$ que sea continua en los racionales y discontinua en los irracionales*. Estos, y tantos otros resultados similares, revelan que continuidad es un asunto que hay que analizar con mucha precaución. En lo que sigue designaremos por $\mathbb{Q}^*$ al conjunto de los números racionales en forma irreducible, es decir, $p/q \in \mathbb{Q}^*$ si p y q son primos relativos.

Ejemplo 3.1.1. *Considere la **función de Thomae** $f : \mathbb{R} \to \mathbb{R}$ definida por*

$$f(x) = \begin{cases} \dfrac{1}{q} & \text{si} \quad x = p/q \in \mathbb{Q}^*, \\ 0 & \text{si} \quad x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

*Entonces f es **continua en los irracionales** pero **discontinua en los racionales**.*



Prueba. Veamos, en primer lugar, que f es discontinua sobre $\mathbb{Q}$. Sea $x = p/q \in \mathbb{Q}^*$. Como el conjunto de los números irracionales es denso en $\mathbb{R}$, podemos elegir una sucesión de números irracionales $(x_n)_{n=1}^{\infty}$ tal que $\lim_{n\to\infty} x_n = x$. Por definición, $f(x_n) = 0$ para todo $n \in \mathbb{N}$, mientras que $f(x) = 1/q \neq 0$, es decir, $\lim_{n\to\infty} f(x_n) \neq f(x)$. Esto prueba la discontinuidad de f en $x \in \mathbb{Q}$ y como x es arbitrario, concluimos que f es discontinua sobre $\mathbb{Q}$.

Para probar que f es continua sobre los irracionales, será suficiente, por la periodicidad de f, demostrarla en lo irracionales del intervalo $(0,1)$. Tomemos cualquier $x_0 \in (0,1) \setminus \mathbb{Q}$ y sea $0 < \varepsilon < 1$. Elijamos un $N \in \mathbb{N}$ tal que $\frac{1}{N} < \varepsilon$. Nuestro objetivo es determinar un intervalo abierto con centro en x_0, digamos J, que no contenga ningún racional en forma reducida de la siguiente lista

$$\frac{1}{2}, \quad \frac{1}{3}, \quad \frac{2}{3}, \quad \frac{1}{4}, \quad \frac{3}{4}, \quad \frac{1}{5}, \quad \frac{2}{5}, \quad \frac{3}{5}, \quad \frac{4}{5}, \quad \ldots, \quad \frac{N-1}{N}. \tag{3.1.1}$$

¿Por qué la elección de J es la adecuada? Pues bien, supongamos que hemos obtenido el intervalo J y tomemos cualquier $x \in J$. Notemos ahora que:

- si x es irracional, entonces $f(x) = f(x_0) = 0$ y, en consecuencia,

$$|f(x) - f(x_0)| \;=\; 0 \;<\; \varepsilon.$$

- si x es racional, entonces dicho número no es ninguno de los que aparecen en (3.1.1) y, por consiguiente, su denominador debe ser mayor que N, es decir, x es de la forma $\dfrac{p}{q}$ con $q > N$ y, por lo tanto,

$$|f(x) - f(x_0)| \;=\; |f(x)| \;=\; \frac{1}{q} \;<\; \frac{1}{N} \;<\; \varepsilon.$$

Esto demuestra la continuidad de f en x_0 y la prueba finalizará una vez hallamos construido el intervalo J. El procedimiento para obtener el intervalo abierto J es muy sencillo: en efecto, sea

$$S_N \;=\; \big\{p/q \in (0,1) : \; p, q \text{ son primos relativos con } q \leq N\big\}$$

Es claro que S_N es un conjunto finito y sus elementos son precisamente los puntos que aparecen en la lista (3.1.1). Ahora bien, como S_N es finito, podemos determinar un $\delta > 0$ tal que $J := (x_0 - \delta, x_0 + \delta) \subseteq (0,1)$ y con $S_N \cap (x_0 - \delta, x_0 + \delta) = \varnothing$. Esto termina la prueba. ■

La función de Thomae también recibe los siguientes nombres: *función de las palomitas de maíz, función de las gotas de lluvias, la función reglada y la función Estrellas sobre Babilonia*. Otra función que es continua en los irracionales y discontinua en los racionales es la siguiente.

Ejemplo 3.1.2. *Sea $(r_n)_{n=1}^{\infty}$ una lista de los números racionales en $[0,1]$. Defina la función $f : [0,1] \to \mathbb{R}$ por*

$$f(x) \;=\; \begin{cases} \displaystyle\sum_{n \in D_x} \frac{1}{2^n}, & \text{si } x \in (0,1] \\ 0, & \text{si } x = 0, \end{cases}$$

donde $D_x = \{n \in \mathbb{N} : r_n < x\}$. Entonces f, además de ser estrictamente creciente, es continua en los irracionales y discontinua en los racionales de $[0,1]$.

Prueba. Para demostrar que f es estrictamente creciente, sean $x_1, x_2 \in [0,1]$ con $x_1 < x_2$. Entonces

$$f(x_2) = \sum_{\substack{n \in \mathbb{N}, \\ r_n < x_2}} \frac{1}{2^n} = \sum_{\substack{n \in \mathbb{N}, \\ r_n < x_1}} \frac{1}{2^n} + \sum_{\substack{n \in \mathbb{N}, \\ x_1 \leq r_n < x_2}} \frac{1}{2^n} > \sum_{\substack{n \in \mathbb{N}, \\ r_n < x_1}} \frac{1}{2^n} = f(x_1).$$

Para ver la segunda parte, sea $\varepsilon > 0$. Puesto que la serie $\sum_{n=1}^{\infty} 1/2^n$ converge, existe un $N \in \mathbb{N}$ tal que

$$\sum_{n=N}^{\infty} \frac{1}{2^n} < \varepsilon.$$

Sea y un número irracional en $[0,1]$ y defina

$$\delta = \max\left\{|y - r_j| \,:\, 1 \leq j < N\right\}.$$

Vamos a demostrar que

$$f(x) - f(y) < \varepsilon$$

para cualquier $x \in [0,1]$ que satisfaga $0 < x - y < \delta$. En efecto, observe en primer lugar que, si n es cualquier número natural tal que $y < r_n < x$, entonces $n \geq N$. Para ver esto, suponga por un momento que $n < N$. Como $y < r_n < x$, resulta entonces que $r_n - y < x - y < \delta$ lo cual es imposible por la definición de δ. De lo anterior se deduce que si $x - y < \delta$, entonces

$$f(x) - f(y) = \sum_{\{n \,:\, y < r_n < x\}} \frac{1}{2^n} \leq \sum_{n=N}^{\infty} \frac{1}{2^n} < \varepsilon.$$

Esto prueba la continuidad de f en cualquier irracional $y \in [0,1]$.

Suponga ahora que y es un racional arbitrario en $[0,1]$. Entonces $y = r_m$ para algún $m \in \mathbb{N}$. Admitamos, por un momento, que f es continua en y y escojamos un $\delta > 0$ tal que si $0 < x - y < \delta$, entonces $f(x) - f(y) < \varepsilon$. Puesto que

$$\varepsilon > f(x) - f(y) = \sum_{\{n \,:\, y < r_n < x\}} \frac{1}{2^n} \geq \frac{1}{2^m}$$

entonces se obtiene una contradicción si nuestro ε se elige tan pequeño de modo que $\varepsilon < \frac{1}{2^m}$. De esto se concluye que f es discontinua en cada racional de $[0,1]$ y termina la prueba. ■

3.1.3. Oscilación y Discontinuidad de una Función

Como siempre $V(x,\delta)$ denotará un entorno de x en $[a,b]$, es decir, $V(x,\delta)$ es de la forma

$$V(x,\delta) = (x - \delta, x + \delta) \cap [a,b] \qquad \text{para algún} \quad \delta > 0.$$

Recordemos que una función $f : [a,b] \to \mathbb{R}$ es continua en $x_0 \in [a,b]$ si, dado $\varepsilon > 0$, existe un $\delta > 0$ tal que $|f(x) - f(x_0)| < \varepsilon$ siempre que $x \in V(x_0, \delta)$. Esto nos indica que

$$f(V(x_0,\delta)) \subseteq (f(x_0) - \varepsilon, f(x_0) + \varepsilon).$$

Nuestro primer resultado estable que cualquier función continua que es positiva en un punto posee un entorno alrededor de dicho donde la función sigue siendo positiva.

Teorema 3.1.24. *Sea $f : [a,b] \to \mathbb{R}$ **continua** sobre $[a,b]$ y suponga que $f(x_0) > 0$ para algún $x_0 \in [a,b]$. Entonces existe un $\delta > 0$ tal que $f(x) > 0$ para todo $x \in V(x_0,\delta)$.*

Prueba. Puesto que f es continua en x_0, resulta que tomando $\varepsilon = f(x_0)/2$, podemos encontrar un $\delta > 0$ de modo tal que

$$|f(x) - f(x_0)| < \varepsilon \quad \text{para todo } x \in V(x_0,\delta).$$

En particular, $f(x_0) - \varepsilon < f(x)$ para todo $x \in V(x_0,\delta)$, es decir,

$$f(x) > f(x_0) - f(x_0)/2 = f(x_0)/2 > 0 \quad \text{para todo } x \in V(x_0,\delta).$$

La prueba es completa. ■

Un conjunto $G \subseteq \mathbb{R}$ se dice que es un G_δ si él se puede escribir en la forma $G = \bigcap_{n=1}^{\infty} G_n$, donde cada G_n es un conjunto abierto. Como siempre, el símbolo $\mathrm{PC}(f)$ lo usaremos exclusivamente para representar el conjunto de los puntos de continuidad de f.

Teorema 3.1.25. *Sea $f : [a,b] \to \mathbb{R}$ una función arbitraria. Entonces $\mathrm{PC}(f)$ es un G_δ.*

Prueba. Sea $a \in \mathrm{PC}(f)$. Por definición, para cada $n \geq 1$, existe un entorno $V(a,\delta_n^a)$ de a tal que

$$|f(x) - f(a)| < 1/n \quad \text{siempre que } x \in V(a,\delta_n^a).$$

Para cada $n \geq 1$, sea

$$G_n = \bigcup_{a \in \mathrm{PC}(f)} V(a,\delta_n^a)$$

y defina $G = \bigcap_{n=1}^{\infty} G_n$. Claramente cada G_n es un conjunto abierto conteniendo a $\mathrm{PC}(f)$ y, por lo tanto, G es un G_δ. Nuestra tarea es demostrar que $G = \mathrm{PC}(f)$.

Puesto que $\mathrm{PC}(f) \subseteq G_n$ para todo $n \geq 1$, se sigue que $\mathrm{PC}(f) \subseteq G$. Para demostrar la otra inclusión, sea $x_0 \in G$. Veamos que f es continua en x_0. En efecto, fijemos un $\varepsilon > 0$ arbitrario y seleccionemos un entero positivo k tal que $1/k < \varepsilon/2$. Por definición, $x_0 \in G_k$, lo cual significa que $x_0 \in V(a,\delta_k^a)$ para algún $a \in \mathrm{PC}(f)$. Puesto que $V(a,\delta_k^a)$ es un conjunto abierto, existe un $\delta > 0$ tal que

$$\big(x_0 - \delta,\, x_0 + \delta\big) \subseteq V(a,\delta_k^a).$$

Finalmente, si $x \in \big(x_0 - \delta,\, x_0 + \delta\big)$, entonces

$$|f(x) - f(x_0)| \leq |f(x) - f(a)| + |f(a) - f(x_0)| < 1/k + 1/k < \varepsilon.$$

Esto prueba que f es continua en x_0 y, por lo tanto, $G \subseteq \mathrm{PC}(f)$ y concluye la prueba. ■

Otra forma de demostrar el resultado anterior es a través de la noción de oscilación de una función. Es René Baire quien introduce este concepto en su tesis para "medir" *cuánto salta* una función en una discontinuidad. Recordemos que

*Un punto $x_0 \in [a,b]$ es un **punto de discontinuidad** de f si, y sólo si, existe un $\varepsilon > 0$ con la siguiente propiedad: para cada $\delta > 0$, existe un $x_\delta \in V(x_0,\delta)$ para el cual se cumple que*

$$|f(x_\delta) - f(x_0)| \geq \varepsilon.$$

Lo anterior se puede expresar en la forma:

Una función f es ***discontinua en*** $x_0 \in [a,b]$*, si existe un* $\varepsilon > 0$ *con la siguiente propiedad: para cualquier intervalo abierto I contenido en* $[a,b]$ *con* $x_0 \in I$*, se cumple que*

$$\operatorname{diam}\big(f(I)\big) \;=\; \sup\big\{|f(x)-f(y)| \;:\; x,y \in I\big\} \;\geq\; \varepsilon.$$

La siguiente definición permitirá describir a $\operatorname{Disc}(f)$, *el conjunto de los puntos de discontinuidad de* f, en términos de estos supremos. Recordemos antes, véase el Corolario 2.1.13, que si $f : [a,b] \to \mathbb{R}$ es función acotada, entonces

$$\sup\,\big\{|f(x)-f(y)| \;:\; x,y \in F\big\} \;=\; \sup_{x\in F} f(x) \;-\; \inf_{x\in F} f(x)$$

para cualquier conjunto no vacío $F \subseteq [a,b]$.

Definición 3.1.26. *Sea $f : [a,b] \to \mathbb{R}$ una función acotada sobre $[a,b]$. Para cada subconjunto F de $[a,b]$, definimos la* ***oscilación de f en*** *F como*

$$\begin{aligned}
\operatorname{osc}(f,F) \;&=\; \sup_{x\in F} f(x) \;-\; \inf_{x\in F} f(x)\\
&=\; \sup\,\big\{|f(x)-f(y)| \;:\; x,y \in F\big\}\\
&=\; \operatorname{diam}\big(f(F)\big).
\end{aligned}$$

Por supuesto, si f no es acotada sobre F, pondremos $\operatorname{osc}(f,F) = \infty$. Observe que si $\delta_1 < \delta_2$, entonces para cada $x \in [a,b]$,

$$\operatorname{osc}\big(f, V(x,\delta_1)\big) \;\leq \operatorname{osc}\big(f, V(x,\delta_2)\big),$$

de modo que la función $\omega_x : (0,+\infty) \to \mathbb{R}$ definida por $\omega_x(\delta) = \operatorname{osc}\big(f, V(x,\delta)\big)$ es creciente y no-negativa. Usando lo anterior podemos justificar la siguiente definición:

Definición 3.1.27. *Sea $f : [a,b] \to \mathbb{R}$ una función acotada. Para cada $x \in [a,b]$, la* ***oscilación de f en*** *x se define como*

$$\begin{aligned}
\operatorname{osc}(f,x) \;&=\; \inf_{\delta>0} \operatorname{osc}\big(f, V(x,\delta)\big)\\
&=\; \lim_{\delta\to 0^+} \operatorname{osc}\big(f, V(x,\delta)\big).
\end{aligned}$$

El siguiente resultado caracteriza la continuidad de una función en términos de su oscilación en un punto.

Teorema 3.1.28. *Sea $f : [a,b] \to \mathbb{R}$ una función. Entonces*

(1) $\operatorname{PC}(f) \;=\; \big\{x \in [a,b] \;:\; \operatorname{osc}(f,x) = 0\big\}$.

(2) *Para cada $\varepsilon > 0$, el conjunto $\operatorname{O}_f(\varepsilon) \;=\; \big\{x \in [a,b] \;:\; \operatorname{osc}(f,x) < \varepsilon\big\}$ es abierto en $[a,b]$.*

Prueba. (1) Sea $x \in \operatorname{PC}(f)$. Dado $\varepsilon > 0$ existe, por la continuidad de f en x, un $\delta > 0$ para el cual se cumple que $f(V(x,\delta)) \subseteq \big(f(x)-\varepsilon/2, f(x)+\varepsilon/2\big)$. De aquí se sigue $\operatorname{osc}(f,x) \leq \varepsilon$ y como $\varepsilon > 0$ es arbitrario, concluimos que $\operatorname{osc}(f,x) = 0$. Esto prueba que $\operatorname{PC}(f) \subseteq \{x \in [a,b] \;:\; \operatorname{osc}(f,x) = 0\}$.

Para demostrar la otra inclusión, sea $x \in [a,b]$ tal que $\operatorname{osc}(f,x) = 0$. Entonces, dado $\varepsilon > 0$, existe un $\delta > 0$ tal que $|f(y)-f(z)| < \varepsilon$ para todo $y,z \in V(x,\delta)$. En particular, $|f(x)-f(y)| < \varepsilon$

para todo $y \in V(x,\delta)$. Con esto hemos demostrado que $x \in \mathrm{PC}(f)$ y con ello la prueba de la primera parte del teorema.

(2) Sean $\varepsilon > 0$ y $x \in \mathrm{O}_f(\varepsilon)$. Entonces $\operatorname{osc}(f,x) = \inf_{\delta>0} \operatorname{osc}(f, V(x,\delta)) < \varepsilon$ y, por lo tanto, existe un $\delta_0 > 0$ tal que $\operatorname{osc}(f, V(x,\delta_0)) < \varepsilon$. Sea $\delta = \delta_0/2$ y veamos que

$$V(x,\delta) \subseteq \mathrm{O}_f(\varepsilon).$$

En efecto, sea $y \in V(x,\delta)$. Si $|z-y| < \delta$, entonces

$$|z-x| \leq |z-y| + |y-x| < \delta + \delta = \delta_0$$

y, por consiguiente, se cumple que $V(y,\delta) \subseteq V(x,\delta_0)$. Por esto,

$$|f(z) - f(z')| < \varepsilon \quad \text{para todo} \quad z, z' \in V(y,\delta),$$

es decir, $\operatorname{osc}(f, V(y,\delta)) < \varepsilon$ y así, $\operatorname{osc}(f,y) < \varepsilon$. Como $y \in V(x,\delta)$ es arbitrario, resulta entonces que $V(x,\delta) \subseteq \mathrm{O}_f(\varepsilon)$. Esto termina la prueba. ■

Observemos que el Teorema 3.1.28 nos dice que

$$x \in \mathrm{Disc}(f) \quad \Leftrightarrow \quad \operatorname{osc}(f,x) > 0,$$

es decir,

$$\mathrm{Disc}(f) = \{x \in [a,b] : \operatorname{osc}(f,x) > 0\}.$$

Además, puesto que el conjunto $O_f(\varepsilon)$ es abierto en $[a,b]$ para cada $\varepsilon > 0$, resulta que el conjunto

$$\mathrm{D}_\varepsilon(f) = \{x \in [a,b] : \operatorname{osc}(f,x) \geq \varepsilon\} = [a,b] \setminus \mathrm{O}_f(\varepsilon)$$

es cerrado en $[a,b]$ para cada $\varepsilon > 0$. En particular, $\mathrm{D}_\varepsilon(f)$ es compacto para todo $\varepsilon > 0$. Más aun, $\mathrm{Disc}(f)$ se puede representar en la forma

$$\begin{aligned}
\mathrm{Disc}(f) &= [a,b] \setminus \mathrm{PC}(f) \\
&= \{x \in [a,b] : \operatorname{osc}(f,x) > 0\} \\
&= \bigcup_{n=1}^{\infty} \left\{x \in [a,b] : \operatorname{osc}(f,x) \geq \frac{1}{n}\right\} \\
&= \bigcup_{n=1}^{\infty} \mathrm{D}_{1/n}(f).
\end{aligned}$$

De esto se obtiene que $\mathrm{Disc}(f)$ es un F_σ y, por lo tanto, $\mathrm{PC}(f)$ es un G_δ.

Teorema 3.1.29. *Sea $f : [a,b] \to \mathbb{R}$ una* ***función acotada*** *sobre $[a,b]$. Entonces $\mathrm{Disc}(f) \cap \mathrm{L}_f^-$ es* ***a lo más numerable****, donde $\mathrm{L}_f^- = \{x \in (a,b] : f(x^-) \text{ existe}\}$.*

Prueba. Puesto que $\mathrm{Disc}(f) = \bigcup_{n=1}^{\infty} \mathrm{D}_{1/n}(f)$, resulta que

$$\mathrm{Disc}(f) \cap \mathrm{L}_f^- = \bigcup_{n=1}^{\infty} \left(\mathrm{D}_{1/n}(f) \cap \mathrm{L}_f^-\right)$$

y, en consecuencia, es suficiente demostrar que el conjunto $\mathrm{D}_{1/n}(f)\cap\mathrm{L}_f^-$ es a lo más numerable. Fijemos un $n\in\mathbb{N}$ y sea $x_0\in\mathrm{D}_{1/n}(f)\cap\mathrm{L}_f^-$. Como $x_0\in\mathrm{L}_f^-$, existe un $\delta>0$ tal que

$$|f(x)-f(x_0^-)|\ <\ \frac{1}{2n}\qquad\text{para todo}\quad x\in(x_0-\delta,x_0),$$

de donde se sigue que

$$|f(u)-f(v)|\ <\ \frac{1}{n}\qquad\text{para todo}\quad u,v\in(x_0-\delta,x_0).$$

Esto último nos indica que $\operatorname{osc}(f,(x_0-\delta,x_0))\leq 1/n$ y, en consecuencia,

$$\text{si }\ x\in(x_0-\delta,x_0),\ \text{entonces}\ \operatorname{osc}(f,x)\leq 1/n.$$

Por otro lado, como $x_0\in\mathrm{D}_{1/n}(f)$, se tiene que $\operatorname{osc}(f,x)\geq 1/n$ y por lo tanto, x_0 es el extremo derecho de algún intervalo abierto, digamos J_{x_0}, el cual no contiene puntos de $\mathrm{D}_{1/n}(f)\cap\mathrm{L}_f^-$. Es claro que si x,y son puntos distintos en $\mathrm{D}_{1/n}(f)\cap\mathrm{L}_f^-$, entonces $J_x\cap J_y=\varnothing$. Por esto, la familia de intervalos abiertos $\{J_x : x\in\mathrm{D}_{1/n}(f)\cap\mathrm{L}_f^-\}$ es disjunta, de donde se obtiene que ella es a lo más numerable y, por consiguiente, $\mathrm{D}_{1/n}(f)\cap\mathrm{L}_f^-$ es a lo más numerable. ■

Un argumento enteramente similar nos revela que $\operatorname{Disc}(f)\cap L_f^+$ es a lo más numerable, donde

$$\mathrm{L}_f^+\ =\ \big\{x\in[a,b)\ :\ f(x^+)\ \text{existe}\big\}.$$

De nuevo, sea $f:[a,b]\to\mathbb{R}$ una función arbitraria. Los puntos de $\operatorname{Disc}(f)$ se pueden clasificar en dos grandes categorías: los que son de la primera especie y los que son de la segunda especie.

Definición 3.1.30. *Sea $f:[a,b]\to\mathbb{R}$ una función. Un punto $x\in\operatorname{Disc}(f)$ es llamado un **punto de discontinuidad de la primera especie**, si lo límites*

$$f(x^+)\ =\ \lim_{t\to x+}f(t)\qquad\qquad y\qquad\qquad f(x^-)\ =\ \lim_{t\to x-}f(t)$$

existen.

Por supuesto, $x\in\operatorname{Disc}(f)$ es un punto de discontinuidad de la primera especie significa que

$$f(x^+)\ =\ f(x^-)\ \neq\ f(x)\qquad\qquad \text{o}\qquad\qquad f(x^-)\neq f(x^+).$$

Si ocurre que $f(x^+)=f(x^-)\neq f(x)$, entonces se dice que x es una **discontinuidad removible**. Si por el contrario, $f(x^-)\neq f(x^+)$, entonces diremos que x es una **discontinuidad de salto**. Todas las otras discontinuidades son llamadas de la **segunda especie**. Por ejemplo, las funciones

$$f(x)\ =\ \begin{cases}\dfrac{1}{x^2} & \text{si}\quad x\neq 0\\[2mm] 0 & \text{si}\quad x=0\end{cases}\qquad\qquad \text{y}\qquad\qquad g(x)\ =\ \begin{cases}\operatorname{sen}\left(\dfrac{1}{x}\right) & \text{si}\quad x\neq 0\\[2mm] 0 & \text{si}\quad x=0\end{cases}$$

poseen, ambas, una discontinuidad de la segunda especie en $x=0$.

Definición 3.1.31. *Una función $\varphi : [a,b] \to \mathbb{R}$ se llama una **función en escalera**, o **función escalonada**, si existen números $x_0, x_1, \ldots, x_n$ en $[a,b]$ tales que*

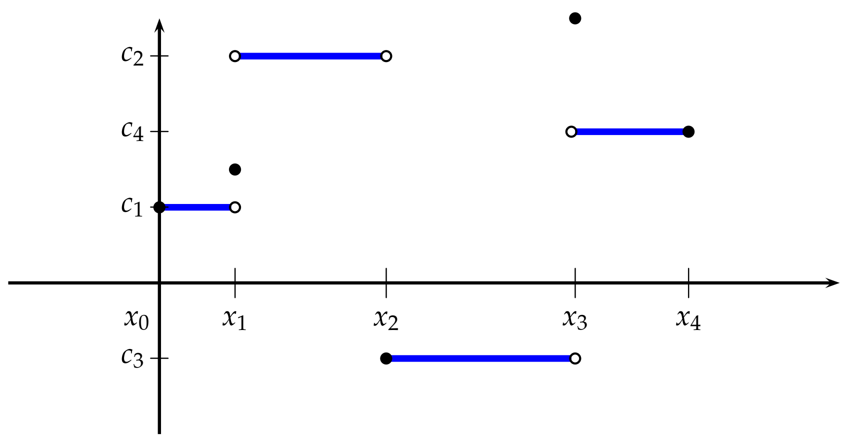
(i) $a = x_0 < x_1 < \cdots < x_n = b$ *y*

(ii) φ *es constante en cada intervalo abierto* (x_{j-1}, x_i), $j = 1, 2, \ldots n$.

Sea $\varphi : [a,b] \to \mathbb{R}$ una función en escalera y suponga que $\varphi(x) = c_j$ para todo $x \in (x_{j-1}, x_j)$ para cada $j \in \{1, \ldots, n\}$. Observe que los valores $\varphi(x_0), \ldots, \varphi(x_n)$ de los puntos extremos de los intervalos $[x_{i-1}, x_{i+1}]$ pueden no coincidir con los números $c_1, \ldots, c_n$. Sin embargo, toda función representada en la forma

$$\varphi(x) \;=\; \sum_{j=1}^{n} c_j \, \chi_{I_j}$$

donde $a = x_0 < x_1 < \cdots < x_n = b$, $I_j = [x_{j-1}, x_i]$, para $j = 1, 2, \ldots n$ y los c_j son números reales, es una función en escalera.



*El conjunto de todas las **funciones escalonadas*** $\varphi : [a,b] \to \mathbb{R}$ será denotado por $\mathrm{Esc}([a,b])$. Es fácil establecer que $\mathrm{Esc}([a,b])$ es un subespacio vectorial de $\mathcal{B}_\infty([a,b])$. Nótese que el número de discontinuidades de cualquier $\varphi \in \mathrm{Esc}([a,b])$ es finito y todas son de salto.

Teorema 3.1.32. *Si $f : [a,b] \to \mathbb{R}$ es una **función con la PVI**, entonces **toda discontinuidad** de f es de la **segunda especie**.*

Prueba. Suponga, para arribar a una contradicción, que f posee una discontinuidad de la primera especie, digamos x_0 y que, por ejemplo, dicha discontinuidad es de salto. En este caso $f(x_0^-) \neq f(x_0^+)$. Para fijar idea, suponga que $f(x_0^-) < f(x_0^+)$ y que c es tal que $c \neq f(x_0)$ y $f(x_0^-) < c < f(x_0^+)$. Ahora bien, puesto que $f(x_0^-) = \lim\limits_{x \to x_0-} f(x)$, tomando $\varepsilon = c - f(x_0^-) > 0$, existe un $\delta_1 > 0$ tal que

$$|f(x) - f(x_0^-)| \;<\; \varepsilon \quad \text{para cada} \quad x \in [a,b] \;\text{ con }\; 0 < x_0 - x < \delta_1. \tag{PVI$_1$}$$

Similarmente, eligiendo $\varepsilon = f(x_0^+) - c > 0$, existe un $\delta_2 > 0$ tal que

$$|f(x) - f(x_0^+)| \;<\; \varepsilon \quad \text{para cada} \quad x \in [a,b] \;\text{ con }\; 0 < x - x_0 < \delta_2. \tag{PVI$_2$}$$

Observe que (PVI_1) implica que

$$f(x) \ < \ c \quad \text{para} \quad x_0 - \delta_1 < x < x_0. \tag{1}$$

Similarmente, (PVI_2) implica que

$$f(x) \ > \ c \quad \text{para} \quad x_0 < x < x_0 + \delta_2. \tag{2}$$

Seleccione dos números cualesquiera en $[a, b]$, digamos x_1 y x_2 tales que

$$x_0 - \delta_1 \ < \ x_1 \ < \ x_0 \ < \ x_2 \ < \ x_0 + \delta_2$$

y observe que de (1) y (2) se obtiene que

$$f(x_1) \ < \ c \ < \ f(x_2)$$

Esto último combinado con el hecho de que $c \neq f(x_0)$, nos indica que no existe $t \in (x_1, x_2)$ para el cual $f(t) = c$ lo que contradice nuestra hipótesis de que f posee la PVI. Suponga ahora que x_0 es una desigualdad removible, es decir,

$$f(x_0^+) \ = \ f(x_0^-) \ \neq \ f(x_0).$$

Sea $L = f(x_0^+) = f(x_0^-)$ y seleccione un $\varepsilon_0 > 0$ tal que $f(x_0) \notin (L - \varepsilon_0, L + \varepsilon_0)$. Sin perder generalidad, asumiremos que $f(x_0) > L + \varepsilon_0$ (el caso $f(x_0) < L - \varepsilon_0$ se trata de modo similar). De la definición de límite se sigue la existencia de un $\delta > 0$ tal que

$$L - \varepsilon_0 \ < \ f(x) \ < \ L + \varepsilon_0 \quad \text{para todo} \quad x \in (x_0 - \delta, x_0 + \delta) \quad x \neq x_0. \tag{PVI$_3$}$$

Finalmente, seleccione cualquier $x_1 \in (x_0 - \delta, x_0 + \delta)$ con $x_1 \neq x_0$ y observe que $L + \varepsilon_0$ está entre $f(x_0)$ y $f(x_1)$. Sin embargo, por (PVI_3), ningún x entre x_0 y x_1 cumple que $f(x) = L + \varepsilon_0$. Esto, por supuesto, contradice la PVI y termina la prueba. ■

Si $f : [a, b] \to \mathbb{R}$ es una función diferenciable en $[a, b]$, entonces el Teorema de Darboux, Teorema 3.1.21, nos dice que f' posee la Propiedad del Valor Intermedio y, por consiguiente, se cumple que:

Corolario 3.1.33. *Si $f : [a, b] \to \mathbb{R}$ es una **función diferenciable** sobre $[a, b]$, entonces **toda discontinuidad** de f' es de la **segunda especie**.*

El siguiente resultado es el ingrediente fundamental para establecer que los puntos de discontinuidad de cualquier función monótona es a lo más numerable.

Teorema 3.1.34. *Si $f : [a, b] \to \mathbb{R}$ es **monótona** sobre $[a, b]$, entonces $f(x^-)$ y $f(x^+)$ existen para cualquier $x \in (a, b)$. En particular, todos los **puntos de discontinuidad** de f son de **salto**.*

Prueba. Sin perder generalidad, asumiremos que f es creciente. Sea $x \in (a, b)$ y considere el conjunto

$$A_x \ = \ \{f(t) \ : \ t < x\}.$$

Como f es creciente, A_x está acotado superiormente por $f(x)$ y, por consiguiente, gracias al Axioma del Supremo, $\sup A_x$ existe y es menor o igual a $f(x)$. Defina entonces

$$\lambda_x \ = \ \sup A_x \ = \ \sup\{f(t) \ : \ t < x\}.$$

Dado cualquier $\varepsilon > 0$, se tiene que $\lambda_x - \varepsilon < \lambda_x$ y se sigue de las propiedades del supremo que existe un $y_0 \in A_x = \{f(t) : t < x\}$ tal que $\lambda_x - \varepsilon < y_0 \le \lambda_x$. Observe que como $y_0 \in A_x$, existe un $t' < x$ tal que $y_0 = f(t')$ y puesto que f es creciente resulta, tomando $\delta = x - t'$, que para todo $z \in [a,b]$ con $x - \delta < z < x$, esto es, $t' < z < x$ se cumple que $y_0 = f(t') \le f(z) \le f(x)$. Ahora bien, como $f(z) \in \{f(t) : t < x\}$ se tiene que $f(z) \le \sup\{f(t) : t < x\} = \lambda_x$, de donde se sigue que $\lambda_x - \varepsilon < y_0 = f(t') \le f(z) \le \lambda_x < \lambda_x + \varepsilon$. De aquí se obtiene que

$$|f(z) - \lambda_x| < \varepsilon \quad \text{siempre que } x - \delta < z < x,$$

lo cual prueba que $\lim_{t\to x^-} f(t) = \lambda_x \le f(x)$. Similarmente, se comprueba que

$$f(x) \ge \lim_{t\to x^+} f(t) = \inf\{f(t) : t > x\},$$

de donde se sigue que los límites $f(x^-)$ y $f(x^+)$ existen y $f(x^-) \le f(x) \le f(x^+)$.

Suponga ahora que $x \in \mathrm{Disc}(f) \cap (a,b)$. Siendo x un punto de discontinuidad de f, resulta que $f(x^-) \le f(x) < f(x^+)$ o bien $f(x^-) < f(x) \le f(x^+)$. En cualquier caso $f(x^+) - f(x^-) > 0$ y termina la prueba. ■

Corolario 3.1.35. *Si $f : [a,b] \to \mathbb{R}$ es **monótona** sobre $[a,b]$, entonces $\mathrm{Disc}(f)$ es **a lo más numerable**.*

Prueba. Suponga que f es creciente. Por el resultado anterior, todos los puntos de $\mathrm{Disc}(f)$ son de salto y, en consecuencia, como estamos asumiendo que f es creciente, resulta que

$$f(x^-) = \lim_{t\to x-} f(t) < \lim_{t\to x+} f(t) = f(x^+)$$

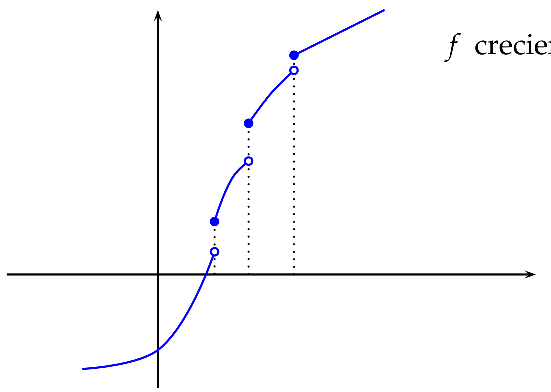
para todo $x \in \mathrm{Disc}(f)$. Seleccione, por cada $x \in \mathrm{Disc}(f)$, un número racional $r_x \in [a,b]$ tal que $\lim_{t\to x-} f(t) < r_x < \lim_{t\to x+} f(t)$. Es claro que si x, y son puntos distintos en $\mathrm{Disc}(f)$, entonces $r_x \neq r_y$. Esto prueba que la aplicación

$$\phi : \mathrm{Disc}(f) \to \mathbb{Q} \qquad \text{definida por} \qquad \phi(x) = r_x$$

es inyectiva y, por lo tanto, como $\mathbb{Q}$ es numerable, resulta que el conjunto $\mathrm{Disc}(f)$ es a lo más numerable. ■

Por supuesto, si $f : \mathbb{R} \to \mathbb{R}$ es **monótona** sobre $\mathbb{R}$, entonces como $\mathbb{R} = \bigcup_{n=1}^{\infty}[-n,n]$ y $f|_{[-n,n]}$ es monótona sobre $[-n,n]$, resulta que $\mathrm{Disc}(f)$ es a lo más numerable.

Definición 3.1.36. *Una función $f : \mathbb{R} \to \mathbb{R}$ se dice que es **continua por la derecha** sobre $\mathbb{R}$ si para cada $x \in \mathbb{R}$, se tiene que $f(x) = f(x^+)$.*



El siguiente resultado muestra que, dada cualquier función creciente, siempre se puede construir una función continua por la derecha.

Corolario 3.1.37. *Si $f : \mathbb{R} \to \mathbb{R}$ es una **función creciente** sobre $\mathbb{R}$, entonces la función $g_+ : \mathbb{R} \to \mathbb{R}$ definida por*

$$g_+(x) \;=\; f(x^+) \quad \text{para todo } x \in \mathbb{R}$$

*es **creciente** y **continua por la derecha** sobre $\mathbb{R}$.*

Prueba. Claramente g_+ es creciente. Para ver que ella también es continua por la derecha, sea $x \in \mathbb{R}$. Entonces

$$g_+(x^+) \;=\; \lim_{t \to x+} g_+(t) \;=\; \lim_{t \to x+} f(t^+) \;=\; f(x^+) \;=\; g_+(x).$$

Fin de la prueba. ■

Denote por $\mathrm{Mon}([a,b])$ *la familia de todas las funciones* $f : [a,b] \to \mathbb{R}$ *que son **monótonas** sobre* $[a,b]$. Observe que por el Teorema 3.1.34, toda función $f \in \mathrm{Mon}([a,b])$ posee límites laterales finitos en cada punto $x \in [a,b]$. Esto permite considerar la siguiente clase de funciones:

Definición 3.1.38. *Una función $f : [a,b] \to \mathbb{R}$ se dice **regulada** sobre $[a,b]$ si ella posee límites laterales finitos en cada punto $x \in [a,b]$, donde convenimos en definir $f(a^-) = f(a)$ y $f(b^+) = f(b)$.*

Denote por $\mathrm{Reg}([a,b])$ *a la familia de todas las funciones* $f : [a,b] \to \mathbb{R}$ *que son **reguladas** sobre* $[a,b]$, y observe que, gracias al Teorema 3.1.34, $\mathrm{Mon}([a,b]) \subsetneq \mathrm{Reg}([a,b])$. Por otro lado, $\mathrm{Mon}([a,b]) \neq \mathrm{Reg}([a,b])$ ya que existen funciones reguladas que no son monótonas. Por ejemplo, cualquier función en escalera pertenece a $\mathrm{Reg}([a,b])$ y, por supuesto, no todas ellas son monótonas. Así mismo, resulta claro que $C([a,b]) \subsetneq \mathrm{Reg}([a,b])$.

3.1.4. Convergencia de Sucesiones de Funciones

Sea X un subconjunto no vacío de $\mathbb{R}$. Para cada $n \in \mathbb{N}$, sea $f_n : X \to \mathbb{R}$ una función. La sucesión de funciones $(f_n)_{n=1}^{\infty}$ se dice que **converge puntualmente** a f en X, donde $f : X \to \mathbb{R}$ si la sucesión numérica $(f_n(x))_{n=1}^{\infty}$ converge a $f(x)$ para cada $x \in X$, en otras palabras, si para

cada $x \in X$ y cada $\varepsilon > 0$, existe un $N \in \mathbb{N}$, que por lo general depende tanto de x así como de ε, tal que
$$\left|f_n(x) - f(x)\right| \;<\; \varepsilon \quad \text{para todo } n \geq N.$$
Observe que esto es equivalente a afirmar lo siguiente: la sucesión $(f_n)_{n=1}^{\infty}$ converge puntualmente (a una cierta función) en X si, para cada $x \in X$ y cada $\varepsilon > 0$, existe un $N = N(x, \varepsilon) \in \mathbb{N}$ tal que
$$\left|f_m(x) - f_n(x)\right| \;<\; \varepsilon \quad \text{para todo } m, n \geq N.$$
Escribiremos $f_n \to f$ *puntualmente* cuando la sucesión $(f_n)_{n=1}^{\infty}$ converja a f puntualmente en X.

Ejemplo 3.1.3. *Para cada $m \in \mathbb{N}$, defina la función $f_m : \mathbb{R} \to \mathbb{R}$ por*
$$f_m(x) \;=\; \lim_{n\to\infty} \big(\cos m!\pi x\big)^{2n}, \quad x \in \mathbb{R}.$$
Entonces $f_m \to \chi_{\mathbb{Q}}$ puntualmente.

Prueba. Sea $x = p/q$ un número racional. Sin perder generalidad, podemos asumir que $q \geq 1$ y $p \in \mathbb{Z}$. Entonces $m!x$ es un entero para cualquier número natural $m \geq q$ y, en consecuencia, $\cos m!\pi x = \pm 1$. Para tales m's resulta que
$$f_m(x) \;=\; \lim_{n\to\infty} \big(\cos m!\pi x\big)^{2n} \;=\; 1$$
y, por lo tanto, $\lim_{m\to\infty} f_m(x) = 1$. Por otro lado, si x es irracional, también lo es $m!x$ para cualquier entero m y, en consecuencia, $m!\pi x$ no es un múltiplo entero de π. De esto se sigue que $|\cos m!\pi x| < 1$ y, en consecuencia,
$$f_m(x) \;=\; \lim_{n\to\infty} \big(\cos m!\pi x\big)^{2n} \;=\; 0.$$
Tomando límite cuando m tiende a infinito, resulta que $\lim_{m\to\infty} f_m(x) = 0$ si x es irracional. En conclusión,
$$\lim_{m\to\infty} f_m(x) \;=\; \chi_{\mathbb{Q}}(x)$$
para cada $x \in \mathbb{R}$, es decir, $f_m \to \chi_{\mathbb{Q}}$ puntualmente. ■

La convergencia puntual de una sucesión de funciones continuas no garantiza que su límite, en caso de existir, sea una función continua. Por ejemplo, si $f_n(x) = x^n$ para todo $x \in [0, 1]$ resulta que la sucesión $(f_n)_{n=1}^{\infty}$ converge puntualmente a la función discontinua f, la cual vale 1 en $x = 1$ y 0 en los puntos restantes. Por consiguiente, se requiere algo más que convergencia puntual para garantizar que la función límite herede la propiedad de continuidad. Esta y otras circunstancias evidencian la necesidad de introducir una noción más fuerte que la convergencia puntual de modo que garantice, por ejemplo, la continuidad de la función límite.

Definición 3.1.39. *Una sucesión de funciones $(f_n)_{n=1}^{\infty}$ definidas sobre un conjunto X y a valores reales se dice que* ***converge uniformemente*** *sobre X a una función $f : X \to \mathbb{R}$ si, dado $\varepsilon > 0$, existe un $N = N(\varepsilon) \in \mathbb{N}$, que depende únicamente de ε, tal que si $n \geq N$ entonces se verifica que*
$$\left|f_n(x) - f(x)\right| \;<\; \varepsilon \quad \textit{para todo} \quad x \in X.$$

Como antes, escribiremos $f_n \to f$ *uniformemente* cuando la sucesión $(f_n)_{n=1}^{\infty}$ converja a f uniformemente. Observe que $(f_n)_{n=1}^{\infty}$ converge uniformemente a una función f sobre X si, y sólo si,

$$\lim_{n\to\infty} \sup\big\{|f_n(x) - f(x)| \,:\, x \in X\big\} \;=\; \lim_{n\to\infty} \|\, f_n - f\,\|_{\infty} \;=\; 0.$$

Resulta claro que la convergencia uniforme implica la convergencia puntual pero no recíprocamente.

Algunos de los enormes beneficios que genera la convergencia uniforme de sucesiones de funciones se muestran en los siguientes cuatro resultados.

Teorema 3.1.40. *Sea $(f_n)_{n=1}^{\infty}$ una sucesión de **funciones continuas** definidas sobre un espacio métrico (X,d). Si $(f_n)_{n=1}^{\infty}$ **converge uniformemente** sobre X a una función $f : X \to \mathbb{R}$, entonces f es **continua** sobre X.*

Prueba. Sea $\varepsilon > 0$. Según la definición de convergencia uniforme, existe un $N \in \mathbb{N}$ tal que si $n \geq N$ se verifica que

$$\big|f_n(x) - f(x)\big| \;<\; \varepsilon/3 \quad \text{para todo} \quad x \in X.$$

Fijemos un $n \geq N$ y sea $x \in X$. Como f_n es continua en x, existe un $\delta > 0$ tal que

$$\big|f_n(x) - f_n(y)\big| \;<\; \varepsilon/3$$

siempre que $d(x,y) < \delta$. De esto resulta que si $d(x,y) < \delta$, entonces

$$\begin{aligned}
\big|f(x) - f(y)\big| \;&=\; \big|f(x) \,-\, f_n(x) \,+\, f_n(x) \,-\, f_n(y) \,+\, f_n(y) \,-\, f(y)\big| \\
&=\; \big|f(x) \,-\, f_n(x)\big| \,+\, \big|f_n(x) \,-\, f_n(y)\big| \,+\, \big|f_n(y) \,-\, f(y)\big| \\
&<\; \varepsilon/3 \,+\, \varepsilon/3 \,+\, \varepsilon/3 \;=\; \varepsilon.
\end{aligned}$$

Esto prueba la continuidad de f en x y como x fue elegido arbitrariamente, concluimos que f es continua sobre X. ■

Un resultado que puede resultar útil en otras situaciones que involucra convergencia uniforme de funciones continuas es el siguiente.

Teorema 3.1.41. *Sea $(f_n)_{n=1}^{\infty}$ una sucesión de **funciones continuas** definidas sobre un espacio métrico (X,d) que **converge uniformemente** sobre X a una función $f : X \to \mathbb{R}$. Si $(x_n)_{n=1}^{\infty}$ es una sucesión en X que converge a un punto $x_0 \in X$, entonces*

$$\lim_{n\to\infty} f_n(x_n) \;=\; f(x_0).$$

Prueba. Sea $\varepsilon > 0$. Queremos demostrar la existencia de un $N \in \mathbb{N}$ tal que

$$|f_n(x_n) - f(x_0)| \;<\; \varepsilon \quad \text{para todo } n \geq N.$$

Por el Teorema 3.1.40 sabemos que f es continua sobre X, de modo que podemos determinar la existencia un $\delta > 0$ para el cual

$$|f(x) - f(x_0)| \;<\; \frac{\varepsilon}{2} \quad \text{siempre que} \quad d(x,x_0) < \delta.$$

Por otro lado, como $x_n \to x_0$, existe, para el δ hallado anteriormente, un $N_1 \in \mathbb{N}$ tal que

$$d(x_n, x_0) < \delta \quad \text{para todo} \quad n \geq N_1.$$

También, puesto que $f_n \to f$ uniformemente, existe un $N_2 \in \mathbb{N}$ tal que

$$|f_n(x) - f(x)| < \varepsilon \quad \text{para todo } n \geq N_2 \text{ y todo } x \in X.$$

Si ahora escojemos $N = \max\{N_1, N_2\}$, resulta que si $n \geq N$, entonces

$$|f_n(x_n) - f(x_0)| \leq |f_n(x_n) - f(x_n)| + |f(x_n) - f(x_0)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

La prueba es completa. ■

Con las notaciones del resultado anterior, si escribimos $x_{mn} = f_n(x_m)$ para todo $m, n \geq 1$, entonces se tiene que

$$\lim_{n\to\infty} \lim_{m\to\infty} x_{mn} = \lim_{m\to\infty} \lim_{n\to\infty} x_{mn} = f(x_0).$$

Por supuesto, existen sucesiones dobles $(x_{mn})_{m,n=1}^{\infty}$ tales que la igualdad anterior no se cumple. Por ejemplo, si tomamos $x_{m,n} = m/(m+n)$, resulta que

$$\lim_{n\to\infty} \lim_{m\to\infty} x_{mn} = 1 \qquad \text{y} \qquad \lim_{m\to\infty} \lim_{n\to\infty} x_{mn} = 0.$$

Tal vez uno de los criterios más importantes respecto a la convergencia uniforme de sucesiones de funciones es el siguiente.

Teorema 3.1.42 (Criterio Uniforme de Cauchy). *Una sucesión $(f_n)_{n=1}^{\infty}$ de funciones a valores reales definidas sobre X. Las siguientes condiciones son equivalentes:*

(1) $(f_n)_{n=1}^{\infty}$ ***converge uniformemente*** *sobre X.*

(2) *Dado $\varepsilon > 0$, existe un $N \in \mathbb{N}$ tal que si $m, n \geq N$ entonces*

$$\big|f_m(x) - f_n(x)\big| < \varepsilon$$

para todo $x \in X$.

Prueba. Suponga que $f_n \to f$ uniformemente sobre X y sea $\varepsilon > 0$. Por definición de convergencia uniforme, existe un $N \in \mathbb{N}$ tal que si $n \geq N$, entonces

$$\big|f_m(x) - f(x)\big| < \varepsilon/2$$

para todo $x \in X$. De aquí se sigue que si $m, n \geq N$ entonces se cumple que

$$\big|f_m(x) - f_n(x)\big| \leq \big|f_m(x) - f(x)\big| + \big|f_n(x) - f(x)\big| < \varepsilon$$

para todo $x \in X$.

Recíprocamente, suponga que la condición (2) de Cauchy es válida. Puesto que $\mathbb{R}$ es completo, para cada $x \in X$, la sucesión $(f_n(x))_{n=1}^{\infty}$ converge a un límite en $\mathbb{R}$ que llamaremos $f(x)$. Por lo tanto, $(f_n)_{n=1}^{\infty}$ converge puntualmente a $f : X \to \mathbb{R}$. Para ver que ella también converge uniformemente, sea $\varepsilon > 0$ y suponga que existe un $N \in \mathbb{N}$ tal que si $m, n \geq N$, la desigualdad

$$\big|f_m(x) - f_n(x)\big| < \varepsilon$$

es válida para todo $x \in X$. Fijando n y haciendo que $m \to \infty$ la desigualdad anterior se convierte en

$$|f(x) - f_n(x)| < \varepsilon$$

que resulta válida para todo $x \in X$ y todo $n \geq N$. ∎

Sea $(f_n)_{n=1}^{\infty}$ una sucesión de funciones definidas sobre X y para cada $n \in \mathbb{N}$, definamos

$$S_n = f_1 + \cdots + f_n.$$

A la sucesión $(S_n)_{n=1}^{\infty}$ la llamaremos la **serie** asociada a $(f_n)_{n=1}^{\infty}$ y denotada en lo sucesivo por $\sum_{n=1}^{\infty} f_n$. La serie $\sum_{n=1}^{\infty} f_n$ **converge puntualmente** a una función f si, para cada x, la sucesión numérica $(S_n(x))_{n=1}^{\infty}$ converge a $f(x)$. En este caso se dice que f es la **suma** de la serie $(f_n)_{n=1}^{\infty}$. Similarmente, la serie $\sum_{n=1}^{\infty} f_n$ **converge uniformemente** sobre X a una función f si, la sucesión $(S_n)_{n=1}^{\infty}$ converge uniformemente sobre X a f.

Es importante advertir que la expresión $\sum_{n=1}^{\infty} f_n$ no indica que la serie converja, es sólo una notación. Se mencionará explícitamente si una serie dada converge o no. Por ejemplo, la serie $\sum_{n=1}^{\infty} x^{n-1}$ converge puntualmente en $(-1, 1)$ y su suma es $f(x) = 1/(1-x)$.

El siguiente test para la convergencia uniforme de una serie dada debido a K. Weierstrass, es muy conveniente.

Teorema 3.1.43 (M-Test de Weierstrass). *Sea $(f_n)_{n=1}^{\infty}$ una sucesión de funciones a valores reales definidas sobre un conjunto X. Suponga que, para cada $n \in \mathbb{N}$, existe una constante no negativa M_n tal que*

$$|f_n(x)| \leq M_n \quad \text{para todo } x \in X.$$

*Si $\sum_{n=1}^{\infty} M_n < \infty$, entonces la serie $\sum_{n=1}^{\infty} f_n$ **converge uniformemente** sobre X.*

Prueba. Sea $\varepsilon > 0$. Como la serie $\sum_{n=1}^{\infty} M_n$ converge, existe un $N \in \mathbb{N}$ tal que si $m > n > N$, entonces $\sum_{j=n+1}^{m} M_j < \varepsilon$. Por lo tanto, para todo $x \in X$ se cumple que

$$|S_m(x) - S_n(x)| \leq \sum_{j=n+1}^{m} |f_j(x)| \leq \sum_{j=n+1}^{m} M_j < \varepsilon$$

siempre que $m > n > N$. Del Criterio Uniforme de Cauchy se sigue que la serie $\sum_{n=1}^{\infty} f_n$ converge uniformemente sobre X. ∎

3.1.5. Una Función Continua Nunca Diferenciable

El objetivo de esta corta sección es presentar un ejemplo sencillo de una función continua nunca diferenciable, es decir, que no posee derivada finita en ningún punto de su dominio. La primera demostración de la existencia de una función continua nunca diferenciable parece provenir del matemático checo Bernard Placidus Tohann Nepomuk Bolzano (1781-1848). Su trabajo matemático pasó casi desapercibido y nunca recibió el reconocimiento que merecía salvo mucho tiempo después de su muerte. Bolzano era contemporáneo de Weierstrass. Además de dar definiciones similares de límite, derivada, continuidad y convergencia, también hizo valiosas contribuciones a la lógica y la teoría de conjuntos (véase, por ejemplo, [15]). Bolzano inventó, alrededor del año 1830, un procedimiento para la construcción de funciones continuas nunca diferenciables que es muy distinta a otras construcciones de funciones con propiedades similares. El proceso de creación de Bolzano es completamente geométrico mientras que las otras construcciones usan series convergentes. Es importante destacar que ese resultado de Bolzano fue dado a conocer en el año 1930, ¡casi 100 años después de haber sido descubierto! En el Volumen 2 de su *Mathematische Werke*, publicado en 1895, aparece por primera vez publicado el artículo de K. Weierstrass donde él demuestra que la función

$$W(x) \;=\; \sum_{n=0}^{\infty} a^n \cos(b^n \pi x)$$

es continua pero nunca diferenciable, siempre que $0 < a < 1$, $ab > 1 + (3\pi/2)$ y b es un entero impar > 1. Cincuenta y cinco años más tarde, en 1916, G. H. Hardy prueba que la función de Weierstrass W sigue siendo continua y nunca diferenciable si además de la condición $0 < a < 1$ se exige que $ab \geq 1$, con $b > 1$, pero sin pedirle que sea un entero impar.

La función que presentamos en esta sección es la de Takagi, dada a conocer a la comunidad matemática por Teiji Takagi (1875-1960) en el año 1903 [126] y una de las más simples. A tal función también se le conoce como el **manjar blanco**, un tipo de pastel francés que tiene una apariencia hinchada que se asemeja de una manera casi caprichosa a la gráfica de la función de Takagi. La tesis de Johan Thim [130] contiene en detalle la construcción de ésta, además, de otras 17 funciones nunca diferenciables.

El siguiente resultado es una de las piezas claves para demostrar que la función de Takagi es nunca diferenciable.

Lema 3.1.44. *Sea $f : [a,b] \to \mathbb{R}$ una **función continua** sobre $[a,b]$ y suponga que $f'(x)$ existe para algún $x \in (a,b)$. Si $(a_n)_{n=1}^{\infty}$ y $(b_n)_{n=1}^{\infty}$ son sucesiones en $[a,b]$ con $a < x_n < x < b_n < b$ y tal que $\lim\limits_{n\to\infty} a_n = x = \lim\limits_{n\to\infty} b_n$, entonces*

$$\lim_{n\to\infty} \frac{f(b_n) - f(a_n)}{b_n - a_n} \;=\; f'(x).$$

Prueba. Puesto que

$$\left|\frac{b_n - x}{b_n - a_n}\right| \;\leq\; \frac{b_n - a_n}{b_n - a_n} \;=\; 1 \qquad \text{y} \qquad \left|\frac{x - a_n}{b_n - a_n}\right| \;\leq\; \frac{b_n - a_n}{b_n - a_n} \;=\; 1,$$

resulta que

$$\begin{aligned}
\left|\frac{f(b_n) - f(a_n)}{b_n - a_n} - f'(x)\right| &= \left|\frac{b_n - x}{b_n - a_n}\left(\frac{f(b_n) - f(x)}{b_n - x} - f'(x)\right) + \frac{x - a_n}{b_n - a_n}\left(\frac{f(a_n) - f(x)}{a_n - x} - f'(x)\right)\right| \\
&\leq \left|\frac{f(b_n) - f(x)}{b_n - x} - f'(x)\right| + \left|\frac{f(a_n) - f(x)}{a_n - x} - f'(x)\right| \\
&\to 0 \quad \text{cuando} \quad n \to \infty.
\end{aligned}$$

Esto finaliza la prueba. ■

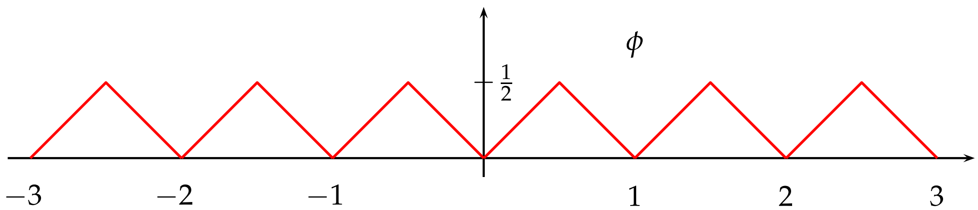
La Función de Takagi

Existen varias formas de introducir la función de Takagi. Una de ellas es la siguiente: considere, en primer lugar, la función $\phi_0 : \mathbb{R} \to \mathbb{R}$ definida por

$$\phi_0(x) \;=\; \begin{cases} x & \text{si } x \in [0, 1/2] \\ 1 - x & \text{si } x \in (1/2, 1] \end{cases}$$

y entonces extienda dicha función, periódicamente, a todo $\mathbb{R}$, esto es, defina

$$\phi(x + m) \;=\; \phi_0(x) \quad \text{para todo} \quad x \in [0, 1] \quad \text{y todo} \quad m \in \mathbb{Z}.$$



Claramente ϕ es continua, no-negativa y $\phi(x) \leq 1/2$ para todo $x \in \mathbb{R}$. Además,

$$\phi(x) \;=\; \phi(1 - x) \quad \text{para cualquier} \quad x \in [0, 1].$$

Nótese que $\phi(m) = 0$ para todo $m \in \mathbb{Z}$. En particular, si x es un racional diádico de orden n, es decir, si $x = m/2^n$ para algún $m \in \mathbb{Z}$, entonces

(a_1^*) $\phi(2^k x) = 0$ siempre que $k \geq n$ ya que, en este caso, $2^k x \in \mathbb{Z}$.

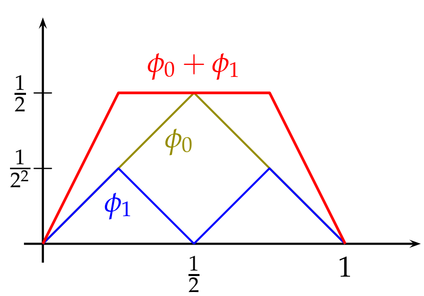
(a_2^*) Más aun, si $u = (m-1)/2^n$ y $v = m/2^n$ son números diádicos consecutivos de orden n tal que los también números diádicos $2^k v$, $2^k u$ pertenecen al intervalo $[0, 1]$ para $0 \leq k < n$, entonces de la linealidad de ϕ se sigue que

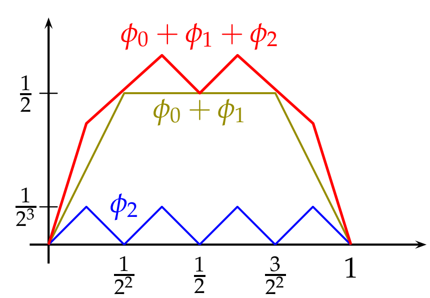
$$\phi(2^k v - 2^k u) \;=\; 2^k \phi(v - u) \;=\; \begin{cases} 2^k \dfrac{1}{2^n} & \text{si } 2^k v,\, 2^k u \in [0, 1/2] \\[2ex] -2^k \dfrac{1}{2^n} & \text{si } 2^k v,\, 2^k u \in (1/2, 1]. \end{cases}$$

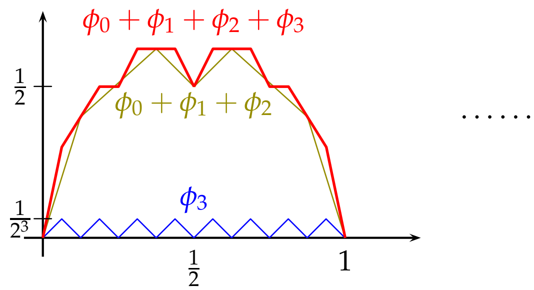
Una vez obtenida la función ϕ, se define la función de Takagi como la serie

$$T(x) \;=\; \sum_{k=0}^{\infty} \frac{1}{2^k} \phi(2^k x), \quad x \in \mathbb{R}.$$

Las gráficas de ϕ_0, $\phi_0 + \phi_1$, $\phi_0 + \phi_1 + \phi_2$, etc. son mostradas en el intervalo $[0, 1]$, donde $\phi_k(x) = 2^{-k}\phi(2^k x)$ para $k = 0, 1, 2, \ldots$







Teorema 3.1.45 (Función de Takagi). *La función de Takagi es* ***continua*** *pero* ***nunca diferenciable*** *sobre* $\mathbb{R}$.

Prueba. Veamos, en primer lugar, que T es continua sobre $\mathbb{R}$. En efecto, para cada $k \in \mathbb{N}_0$, considere la función continua $\phi_k(x) = 2^{-k}\phi(2^k x)$ para todo $x \in \mathbb{R}$. Puesto que $|\phi_k(x)| \leq 1/2^k$ para todo $x \in \mathbb{R}$, el M-test de Weierstrass combinado con el Teorema 3.1.40 nos muestran que T es continua.

Para ver que ϕ no posee derivada finita en ningún punto de $\mathbb{R}$, vamos asumir lo contrario y construir, con la ayuda del Lema 3.1.44, una contradicción. Suponga que $x \in \mathbb{R}$ es un punto arbitrario tal que $T'(x)$ es finito. Sin perder generalidad y, sólo por simplificar los cálculos, supondremos que $x \in (0,1)$. Para cada entero $n \geq 1$, sea $D_n = \{\frac{m}{2^n} : m = 0, 1, \ldots, 2^n\}$. Puesto que $D = \bigcup_{n=1}^{\infty} D_n$ es denso en $[0,1]$, podemos elegir números diádicos sucesivos de orden n, digamos, $u_n = \frac{m-1}{2^n}$ y $v_n = \frac{m}{2^n}$ tales que $u_n < x < v_n$ y $\lim_{n\to\infty} u_n = x = \lim_{n\to\infty} v_n$. Observe que por (a_1^*), si $k \geq n$, entonces $\phi(2^k u_n) = \phi(2^k v_n) = 0$ y, en consecuencia,

$$T(u_n) = \sum_{k=0}^{n-1} \frac{1}{2^k}\,\phi(2^k\,u_n) \qquad \text{y} \qquad T(v_n) = \sum_{k=0}^{n-1} \frac{1}{2^k}\,\phi(2^k\,v_n).$$

De esto resulta que

$$\frac{T(v_n) - T(u_n)}{v_n - u_n} = \sum_{k=0}^{n-1} \frac{1}{2^k}\,\frac{\phi(2^k\,v_n) - \phi(2^k\,u_n)}{v_n - u_n}.$$

Puesto que $v_n - u_n = \frac{1}{2^n}$, se sigue de (a_2^*) que

$$\begin{aligned}\frac{T(v_n) - T(u_n)}{v_n - u_n} &= \sum_{k=0}^{n-1} \frac{1}{2^k}\,\frac{\phi(2^k\,v_n) - \phi(2^k\,u_n)}{v_n - u_n}\\ &= \sum_{k=0}^{n-1} \frac{1}{2^k}\,\frac{\phi(2^k\,(v_n - u_n))}{v_n - u_n}\\ &= \sum_{k=0}^{n-1} \pm 1.\end{aligned}$$

cuyo límite, cuando $n \to \infty$, no converge, lo cual constituye una contradicción ya que el Lema 3.1.44 nos garantiza que

$$T'(x) = \lim_{n\to\infty} \frac{T(v_n) - T(u_n)}{v_n - u_n} \quad \text{existe.}$$

Esta contradicción establece que T no posee derivada finita en ningún punto de $\mathbb{R}$ y termina la prueba. ■

Nota Adicional 3.1.1 El descubrimiento de funciones continuas nunca diferenciables conmocionó a la comunidad matemática de la época que incluso, matemático de la talla de Charles Hermite (1822-1901), en una carta dirigida a Stieltjes fechada el 20 de Mayo de 1893, le decía:

"*Je me détourne avec horreur et effroi de cette plaie lamentable des functions continue qui n'ont pas de dérivé*".

("Me alejo con horror y temor de esta plaga lamentable de las funciones continuas que no poseen derivadas").

Aunque en la actualidad existen variados ejemplos concretos de funciones continuas nunca diferenciables, (véase, por ejemplo, Johan Thim [130]), encontrar una de ellas es casi una proeza y, por supuesto, una curiosidad. Más aun, uno pudiera pensar, lo que es enteramente natural, que este tipo de funciones son excepcionales, que es algo patológico y, de hecho, hasta hace un poco más de cien años esa era la opinión expresada por la mayoría de los matemáticos de la época; pero resulta, y este es lo que fundamentalmente debemos resaltar, que la existencia de tales funciones constituye, desde el punto de vista topológico, la regla y no la excepción. Existe un resultado, llamado el Teorema de Categoría de Baire (véase [23] para conocer una impresionante variedad de aplicaciones de dicho teorema), que permite demostrar la *abundancia* de tales funciones sin exhibir ningún ejemplo en particular.

3.1.6. Funciones Semicontinuas

Recordemos que una función $f : [a, b] \to \mathbb{R}$ es continua en $x_0 \in [a, b]$ si, para cada $\varepsilon > 0$, existe un entorno V de x_0 tal que

$$f(x_0) - \varepsilon \ < \ f(x) \ < \ f(x_0) + \varepsilon \quad \text{para todo} \quad x \in V \cap [a, b].$$

Considerando por separado cada una de las desigualdades anteriores conduce a la siguiente definición.

Definición 3.1.46. *Sea $f : [a, b] \to \mathbb{R}$ una función. Se dice que f es* ***semicontinua superiormente*** *en $x_0 \in [a, b]$ si, para cada $\varepsilon > 0$, existe un entorno V de x_0 tal que*

$$f(x) \ < \ f(x_0) + \varepsilon \quad \textit{para todo} \quad x \in V \cap [a, b].$$

Similarmente, f es ***semicontinua inferiormente*** *en $x_0 \in [a, b]$ si, para cada $\varepsilon > 0$, existe un entorno V de x_0 tal que*

$$f(x_0) - \varepsilon \ < \ f(x) \quad \textit{para todo} \quad x \in V \cap [a, b].$$

La función f se dice que es **semicontinua superiormente** (respectivamente, **semicontinua inferiormente**) sobre (o en) $[a, b]$, si ella es semicontinua superiormente (respectivamente, semicontinua inferiormente) en todos los puntos de $[a, b]$.

Denotemos por $\mathrm{Sci}([a, b])$ *el conjunto de todas las funciones $f : [a, b] \to \mathbb{R}$ que son* ***semicontinuas inferiormente*** *sobre* $[a, b]$. Similarmente, $\mathrm{Scs}([a, b])$ representará *el conjunto de todas las funciones*

$f : [a, b] \to \mathbb{R}$ *que son* ***semicontinuas superiormente*** *sobre* $[a, b]$. Observe que la elección del intervalo $[a, b]$ en la definición de semicontinuidad es intranscendente, por lo que puede ser reemplazado por cualquier intervalo. En particular, cuando el intervalo es $\mathbb{R}$ escribiremos $\mathrm{Scs}(\mathbb{R})$ y $\mathrm{Sci}(\mathbb{R})$.

Definición 3.1.47. *Una función $f : [a, b] \to \mathbb{R}$ se dice que es* ***semicontinua****, en notación $f \in \mathrm{Sc}([a, b])$, si ella es semicontinua superiormente o semicontinua inferiormente.*

Observe que

$$\mathrm{Sc}([a, b]) = \mathrm{Sci}([a, b]) \cup \mathrm{Scs}([a, b]) \qquad \text{y} \qquad C([a, b]) = \mathrm{Sci}([a, b]) \cap \mathrm{Scs}([a, b]).$$

Es fácil establecer que si $f, g \in \mathrm{Sci}([a, b])$, entonces

(a) $f + g \in \mathrm{Sci}([a, b])$.

(b) $\alpha f \in \mathrm{Sci}([a, b])$ para cualquier escalar $\alpha \geq 0$.

(c) $\min\{f, g\} \in \mathrm{Sci}([a, b])$.

Similares consideraciones valen para las funciones en $\mathrm{Scs}([a, b])$. Más aun,

$$f \in \mathrm{Sci}([a, b]) \quad \Leftrightarrow \quad -f \in \mathrm{Scs}([a, b]).$$

Esto último nos permite considerar, en todo lo que sigue, sólo las propiedades de la funciones que son semicontinuas inferiormente. También es claro que si f es semicontinua inferiormente sobre $[a, b]$, entonces ella es semicontinua inferiormente en cualquier subconjunto de $[a, b]$.

Ejemplo 3.1.4. *Las funciones $f, g : [-1, 1] \to \mathbb{R}$ definidas por*

$$f(x) = \begin{cases} -1 & si\ -1 \leq x \leq 0 \\ 1 & si\ \ 0 < x \leq 1 \end{cases} \qquad y \qquad g(x) = \begin{cases} -1 & si\ -1 \leq x < 0 \\ 1 & si\ \ 0 \leq x \leq 1 \end{cases}$$

son, respectivamente, ***semicontinuas inferiormente*** *y* ***superiormente*** *en $x_0 = 0$.*

Teorema 3.1.48. *Sea $f : [a, b] \to \mathbb{R}$ una función. Son equivalentes:*

(1) *f es* ***semicontinua inferiormente*** *sobre $[a, b]$.*

(2) *$G_\xi = \{x \in [a, b] : f(x) > \xi\}$ es* ***abierto*** *en $[a, b]$ para cada $\xi \in \mathbb{R}$.*

Prueba. (1) $\Rightarrow$ (2). Suponga que f es semicontinua inferiormente en $[a, b]$. Fijemos $\xi \in \mathbb{R}$ y sea $x_0 \in G_\xi$. Puesto que $f(x_0) > \xi$, el número $\varepsilon = f(x_0) - \xi > 0$. Usemos ahora el hecho de que f es semicontinua inferiormente en x_0 para hallar un entorno V de x_0 tal que

$$f(x_0) - \varepsilon = \xi < f(x) \qquad \text{para todo} \quad x \in V \cap [a, b].$$

De esto se sigue que el conjunto abierto $V \subseteq G_\xi$ y, por lo tanto, G_ξ es abierto.

(2) $\Rightarrow$ (1). Aceptemos que (2) se cumple y sean $\varepsilon > 0$ y $x_0 \in [a, b]$. Puesto que $G_\xi = f^{-1}((\xi, +\infty))$ es abierto en $[a, b]$ para cualquier $\xi \in \mathbb{R}$, resulta, en particular, que el conjunto $f^{-1}((f(x_0) - \varepsilon, +\infty))$ es abierto y, por supuesto, contiene a x_0. Seleccione un $\delta > 0$ tal que $(x_0 - \delta, x_0 + \delta) \subseteq f^{-1}((f(x_0) - \varepsilon, +\infty))$. Finalmente, si tomamos $V = [a, b] \cap (x_0 - \delta, x_0 + \delta)$ entonces V es un entorno de x_0 en $[a, b]$ y se cumple que

$$f(x_0) - \varepsilon < f(x) \quad \text{para todo} \quad x \in V.$$

Esto termina la prueba. ■

Corolario 3.1.49. *Sea $G \subseteq [a,b]$. Entonces $\chi_G \in \mathrm{Sci}([a,b])$ si, y sólo si, G es **abierto**.*

Prueba. Sea $\xi \in \mathbb{R}$. Si G es abierto, entonces todas las imágenes inversas

$$\chi_G^{-1}((\xi,+\infty)) = \begin{cases} \varnothing & \text{si } \xi \geq 1 \\ G & \text{si } 0 \leq \xi < 1 \\ [a,b] & \text{si } \xi < 0. \end{cases}$$

son abiertas y, por consiguiente, $\chi_G \in \mathrm{Sci}([a,b])$. ■

Aprovechando la caracterización dada en el Teorema 3.1.48 podemos definir la semicontinuidad de funciones $f : X \to \overline{\mathbb{R}}$, donde X es un espacio topológico de Hausdorff del modo siguiente: $f \in \mathrm{Sci}(X)$ si $f^{-1}((\xi,+\infty])$ es abierto en X para cualquier $\xi \in \mathbb{R}$. Análogamente, $f \in \mathrm{Scs}(X)$ si $f^{-1}([-\infty,\xi))$ es abierto en X para cualquier $\xi \in \mathbb{R}$.

En la prueba de la implicación (1) $\Rightarrow$ (2) del Teorema 3.1.48 obtuvimos el siguiente resultado, el cual también es usado como la definición de semicontinuidad inferior en un punto.

Corolario 3.1.50. *Si $f : [a,b] \to \mathbb{R}$ es una función **semicontinua inferiormente** en un punto $x_0 \in [a,b]$, entonces para cada $\xi < f(x_0)$ existe un entorno V de x_0 tal que*

$$f(x) > \xi \qquad \text{para todo} \quad x \in V \cap [a,b].$$

Definición 3.1.51. *Sea $f : [a,b] \to \mathbb{R}$ una función arbitraria. Se define el **límite inferior** de f en $x_0 \in [a,b]$ como*

$$\liminf_{x \to x_0} f(x) = \sup_{\delta > 0} \inf_{x \in V(x_0,\delta)} f(x)$$

donde $V(x_0,\delta) = (x_0 - \delta, x_0 + \delta) \cap [a,b]$.

Otro modo natural de caracterizar a las funciones semicontinuas inferiormente es por medio de su límite inferior.

Teorema 3.1.52. *Sea $f : [a,b] \to \mathbb{R}$ una función y sea $x_0 \in [a,b]$. Son equivalentes:*

*(1) f es **semicontinua inferiormente** en x_0.*

(2) $f(x_0) \leq \liminf_{x \to x_0} f(x)$.

Prueba. Suponga que f es semicontinua inferiormente en el punto x_0. Entonces, dado $\varepsilon > 0$, existe un entorno V de x_0 tal que

$$f(x_0) - \varepsilon < f(x) \quad \text{para todo} \quad x \in V \cap [a,b].$$

Claramente podemos suponer que $V = V(x_0,\delta)$ para algún $\delta > 0$. De lo anterior se sigue que

$$f(x_0) < \inf_{x \in V(x_0,\delta)} f(x) + \varepsilon$$

y como nuestro ε es arbitrario, se concluye que

$$f(x_0) \leq \sup_{\delta > 0} \inf_{x \in V(x_0,\delta)} f(x).$$

Recíprocamente, suponga que $f(x_0) \leq \liminf_{x \to x_0} f(x)$ pero que f no es semicontinua inferiormente en x_0. Esto significa que existe un $\xi < f(x_0)$ tal que, para cualquier $\delta > 0$, existe un $z_\delta \in V(x_0, \delta)$ para el cual $f(z_\delta) < \xi$. Por lo tanto, para cada $\delta > 0$,

$$\inf_{x \in V(x_0,\delta)} f(x) \leq f(z_\delta) < \xi$$

y, en consecuencia,

$$\xi < f(x_0) \leq \liminf_{x \to x_0} f(x) = \sup_{\delta \to 0} \inf_{x \in V(x_0,\delta)} f(x) < \xi,$$

lo cual es absurdo. Esto termina la prueba. ■

Observe que si f es semicontinua inferiormente en x_0 y $(x_n)_{n=1}^{\infty}$ es cualquier sucesión en $[a, b]$ convergiendo a x_0, entonces

$$f(x_0) \leq \liminf_{n \to \infty} f(x_n).$$

De modo similar, el **límite superior** de f en $x_0 \in [a, b]$ viene expresado en la forma

$$\limsup_{x \to x_0} f(x) = \inf_{\delta > 0} \sup_{x \in V(x_0,\delta)} f(x)$$

y procediendo casi de manera idéntica a la demostración del resultado anterior se obtiene que:

Teorema 3.1.53. *Sea $f : [a, b] \to \mathbb{R}$ una función y sea $x_0 \in [a, b]$. Son equivalentes:*

(1) *f es **semicontinua superiormente** en x_0.*

(2) $\limsup_{x \to x_0} f(x) \leq f(x_0)$.

Ejemplo 3.1.5. *Si $f : [a, b] \to \mathbb{R}$ tiene un **mínimo relativo** en $x_0 \in [a, b]$, esto es, existe un entorno V de x_0 tal que $f(x_0) \leq f(x)$ para todo $x \in V$, entonces f es **semicontinua inferiormente** en x_0.*

Ejemplo 3.1.6. *Sea $f : [0, 1] \to \mathbb{R}$ la función de Thomae,* Ejemplo *3.1.1. Sabemos que f es discontinua en cada racional de $[0, 1]$, pero continua en cualquier irracional de $[0, 1]$. Por otro lado, puesto que f tiene un máximo relativo en cada racional $x \in [0, 1]$, se sigue entonces del ejemplo anterior que f **semicontinua superiormente** en $[0, 1]$.*

Otras de las propiedades similares que comparten las funciones semicontinuas con las funciones continuas es la siguiente.

Teorema 3.1.54. *Si $f : [a, b] \to \mathbb{R}$ es una función **semicontinua inferiormente** (respectivamente, **semicontinua superiormente**) sobre $[a, b]$, entonces f **alcanza su mínimo** (respectivamente, **alcanza su máximo**) en algún punto de $[a, b]$.*

Prueba. Suponga que f es semicontinua inferiormente sobre $[a, b]$ y sea

$$m = \text{ínf}\{f(x) : x \in [a, b]\}.$$

Vamos a demostrar que existe un $x_0 \in [a, b]$ tal que $m = f(x_0)$. Para ello, usemos las propiedades del ínfimo y escojamos una sucesión $(x_n)_{n=1}^{\infty}$ en $[a, b]$ tal que $\text{lím}_{n\to\infty} f(x_n) = m$. Por ser $[a, b]$ compacto, se sigue del Teorema 2.1.25 que existe una subsucesión $(x_{n_k})_{k=1}^{\infty}$ de la sucesión $(x_n)_{n=1}^{\infty}$ que converge a algún punto $x_0 \in [a, b]$. Ahora bien, como f es semicontinua inferiormente en x_0, resulta del Teorema 3.1.52 que

$$f(x_0) \leq \liminf_{k\to\infty} f(x_{n_k}) = \lim_{n\to\infty} f(x_n) = m.$$

Esto prueba que $m = f(x_0)$ ya que $m \leq f(x_0)$. ■

En particular, cada función semicontinua inferiormente $f : [a, b] \to \mathbb{R}$ está **acotada por debajo** en $[a, b]$, esto es, existe una constante $m > 0$ tal que

$$m \leq f(x) \quad \text{para todo} \quad x \in [a, b].$$

Similarmente, si $f : [a, b] \to \mathbb{R}$ es semicontinua superiormente, entonces f está **acotada por arriba** en $[a, b]$, lo cual significa que existe una constante $M > 0$ tal que

$$f(x) \leq M \quad \text{para todo} \quad x \in [a, b].$$

3.1.7. Convergencia Puntual en Sc$([a, b])$

El objetivo de esta sección es caracterizar las funciones en Sc$([a, b])$ a través de la convergencia puntual de sucesiones monótonas de funciones continuas. Una de las propiedades que mejor identifica a las funciones semicontinuas son los siguientes dos resultados.

Teorema 3.1.55. *Si $f : [a, b] \to \mathbb{R}$ es una función **semicontinua inferiormente** sobre $[a, b]$, entonces existe una **sucesión creciente** $(f_n)_{n=1}^{\infty}$ en $C([a, b])$ tal que*

$$f(x) = \lim_{n\to\infty} f_n(x) \quad \text{para cada } x \in [a, b].$$

Prueba. Para cada $n \in \mathbb{N}$, defina la función $f_n : [a, b] \to \mathbb{R}$ por

$$f_n(x) = \text{ínf}\{f(t) + n|t - x| : t \in [a, b]\}, \qquad x \in [a, b].$$

Puesto que f es acotada por debajo, la función f_n es finita. Claramente

$$f_1 \leq f_2 \leq \cdots \leq f_n \leq \cdots \leq f. \tag{1}$$

Afirmamos que cada f_n es continua sobre $[a, b]$. Para ver esto, observe que si $x, y \in [a, b]$, entonces

$$\begin{aligned} f_n(x) &= \text{ínf}\{f(t) + n|t - x| : t \in [a, b]\} \\ &\leq \text{ínf}\{f(t) + n|t - y| + n|y - x| : t \in [a, b]\} \\ &= f_n(y) + n|x - y| \end{aligned}$$

De esto último se sigue que

$$|f_n(x) - f_n(y)| \;\leq\; n|x-y|, \tag{2}$$

lo cual muestra que f_n es Lipschitz. En particular, f_n es uniformemente continua sobre $[a,b]$.. Veamos ahora que $f(x) = \lim_{n\to\infty} f_n(x)$ para cada $x \in [a,b]$. En efecto, fijemos $x \in [a,b]$ y observe que, gracias a (1), $\lim_{n\to\infty} f_n(x) \leq f(x)$. Para demostrar la otra desigualdad, tomemos cualquier número ξ y suponga que $f(x) > \xi$. Puesto que f es semicontinua inferiormente en x, el Corolario 3.1.50 nos garantiza la existencia de un $\delta > 0$ tal que

$$f(t) \;>\; \xi \quad \text{para todo } t \in V(x,\delta) = (x-\delta, x+\delta) \cap [a,b].$$

De esto se sigue que

$$\inf\{f(t) \,+\, n|t-x| \;:\; t \in V(x,\delta)\} \;\geq\; \xi.$$

Por otro lado, para cualquier $t \in [a,b] \setminus V(x,\delta)$ se cumple que $n|t-x| \geq n\delta$ y, en consecuencia,

$$\begin{aligned}
&\inf\{f(t) \,+\, n|t-x| \;:\; t \in [a,b] \setminus V(x,\delta)\} \\
&= \inf\{f(t) \;:\; t \in [a,b] \setminus V(x,\delta)\} \,+\, \inf\{n|t-x| \;:\; t \in [a,b] \setminus V(x,\delta)\} \\
&\geq \inf\{f(t) \;:\; t \in [a,b]\} + n\delta \\
&= m + n\delta.
\end{aligned}$$

donde $m = \inf\{f(x) : x \in [a,b]\}$. Si ahora elegimos un $n_0 \in \mathbb{N}$ adecuadamente de modo $m + n_0\delta > \xi$, resultará $f_n(x) \geq \xi$ para todo $n \geq n_0$ y, por consiguiente, $\lim_{n\to\infty} f_n(x) \geq \xi$. Hemos demostrado que

$$\lim_{n\to\infty} f_n(x) \;\geq\; \xi \quad \text{cada vez que} \quad f(x) > \xi.$$

Se sigue del Teorema 2.1.30, página 97, que $\lim_{n\to\infty} f_n(x) \geq f(x)$. Esto completa la prueba. ■

Teorema 3.1.56. *Si $f : [a,b] \to \mathbb{R}$ es una función **semicontinua superiormente**, sobre $[a,b]$, entonces existe una **sucesión decreciente** $(f_n)_{n=1}^{\infty}$ en $C([a,b])$ tal que*

$$f(x) \;=\; \lim_{n\to\infty} f_n(x) \quad \textit{para cada } x \in [a,b].$$

Prueba. Es inmediata del resultado anterior. ■

Los dos teoremas anteriores, Teoremas 3.1.55 y 3.1.56, nos indican que:

Corolario 3.1.57. *Si $f : [a,b] \to \mathbb{R}$ es **semicontinua superiormente** (resp. **semicontinua inferiormente**), entonces $f \in \mathcal{B}_1([a,b])$. En particular, $\mathrm{PC}(f)$ es un G_δ-denso en $[a,b]$ para cualquier $f \in \mathrm{Sc}([a,b])$.*

Lo anterior se puede resumir en la forma:

$$C([a,b]) \;\subseteq\; \mathrm{Sc}([a,b]) \;\subseteq\; \mathcal{B}_1([a,b]).$$

En el siguiente ejemplo se muestra que ninguna de las clases $\mathrm{Sci}([a,b])$ y $\mathrm{Scs}([a,b])$ son cerradas bajo la convergencia puntual, es decir, si $(f_n)_{n=1}^{\infty} \in \mathrm{Sci}([a,b])$ converge puntualmente a una función $f : [a,b] \to \mathbb{R}$, entonces no necesariamente $f \in \mathrm{Sci}([a,b])$. Veamos un ejemplo.

Ejemplo 3.1.7. *Existe una sucesión* $(f_n)_{n=1}^{\infty} \in \mathrm{Sci}([0,1])$ *que* ***converge puntualmente*** *a una función* $f \notin \mathrm{Sci}([0,1])$.

Prueba. Sea $R = \{q_n \in [0,1] \cap \mathbb{Q}\}$ una enumeración de los racionales en $[0,1]$. Para cada $n \in \mathbb{N}$, defina $f_n : [0,1] \to \mathbb{R}$ por $f_n(x) = \chi_{[0,1]\setminus\{q_1,\ldots,q_n\}}(x)$. Puesto que cada f_n alcanza un mínimo relativo en cada racional incluido en $\{q_1, \ldots, q_n\}$, resulta que tal función es semicontinua inferiormente sobre $[0,1]$, pero $\lim_{n\to\infty} f_n = \chi_{[0,1]\setminus R}$ no es semicontinua inferiormente sobre $[0,1]$. ■

A pesar del ejemplo anterior, existen condiciones que garantizan que el límite puntual de una sucesión de funciones semicontinuas (inferiormente o superiormente) preserve dicha propiedad.

Teorema 3.1.58. *Sea* $(f_n)_{n=1}^{\infty}$ *una* ***sucesión creciente*** *en* $\mathrm{Sci}([a,b])$ *y suponga que ella* ***converge puntualmente*** *a una función* $f : [a,b] \to \mathbb{R}$. *Entonces* $f \in \mathrm{Sci}([a,b])$.

Prueba. Fijemos un número real ξ. Puesto que $f_n \to f$ puntualmente, resulta que

$$\{x \in [a,b] : f(x) > \xi\} = \bigcup_{n=1}^{\infty} \{x \in [a,b] : f_n(x) > \xi\}.$$

Como cada conjunto $\{x \in [a,b] : f_n(x) > \xi\}$ es abierto, entonces $\{x \in [a,b] : f(x) > \xi\}$ también es abierto y, por lo tanto, $f \in \mathrm{Sci}([a,b])$ gracias al Teorema 3.1.48. ■

Combinando los Teoremas 3.1.55, 3.1.56 y 3.1.58, se obtiene la siguiente caracterización de las funciones semicontinuas.

Corolario 3.1.59. *Sea* $f : [a,b] \to \mathbb{R}$ *una función.*

(a) $f \in \mathrm{Sci}([a,b])$ *si, y sólo si, existe una* ***sucesión creciente*** $(f_n)_{n=1}^{\infty}$ *en* $C([a,b])$ *que* ***converge puntualmente*** *a* f.

(b) $f \in \mathrm{Scs}([a,b])$ *si, y sólo si, existe una* ***sucesión decreciente*** $(f_n)_{n=1}^{\infty}$ *en* $C([a,b])$ *que* ***converge puntualmente*** *a* f.

3.1.8. Funciones Acotadas en Sc$([a,b])$

Sea $f : [a,b] \to \mathbb{R}$ una **función acotada** sobre $[a,b]$ y considere los conjuntos

$$\mathcal{F}_{\inf}(f) = \{g \in \mathrm{Sci}([a,b]) : g \le f\} \qquad \text{y} \qquad \mathcal{G}^{\sup}(f) = \{h \in \mathrm{Scs}([a,b]) : h \ge f\}.$$

Observe que para cada $x \in [a,b]$, el conjunto

$$F_x = \{g(x) : g \in \mathcal{F}_{\inf}(f)\}$$

es acotado, por lo que la función $f^{\sup} : [a,b] \to \mathbb{R}$ dada por

$$f^{\sup}(x) = \sup\{g(x) : g \in \mathcal{F}_{\inf}(f)\} \quad \text{para cada } x \in [a,b],$$

está bien definida y, además, es semicontinua inferiormente sobre $[a,b]$. Para ver esto, sea $\xi \in \mathbb{R}$ y nótese que

$$\{x \in [a,b] \,:\, f^{\sup}(x) > \xi\} \;=\; \bigcup_{g \in \mathcal{F}_{\inf}(f)} \{x \in [a,b] \,:\, g(x) > \xi\},$$

es un conjunto abierto ya que cada g es semicontinua inferiormente sobre $[a,b]$. Similarmente, para cada $x \in [a,b]$, el conjunto $G_x = \{h(x) \,:\, g \in \mathcal{G}^{\sup}(f)\}$ es acotado y, por lo tanto, la función

$$f_{\inf}(x) \;=\; \text{ínf}\,\{h(x) \,:\, h \in \mathcal{G}^{\sup}(f)\} \quad \text{para cada } x \in [a,b],$$

está bien definida y es semicontinua superiormente sobre $[a,b]$ pues

$$\{x \in [a,b] \,:\, f_{\inf}(x) \geq \xi\} \;=\; \bigcap_{h \in \mathcal{G}^{\sup}(f)} \{x \in [a,b] \,:\, h(x) \geq \xi\},$$

es cerrado para cualquier $\xi \in \mathbb{R}$. Estas consideraciones prueban que:

Teorema 3.1.60. *Sea $f : [a,b] \to \mathbb{R}$ una* ***función acotada*** *sobre $[a,b]$ y considere las familias*

$$\mathcal{F}_{\inf}(f) \;=\; \{g \in \mathrm{Sci}([a,b]) \,:\, g \leq f\} \quad y \quad \mathcal{G}^{\sup}(f) \;=\; \{h \in \mathrm{Scs}([a,b]) \,:\, h \geq f\}.$$

Entonces

$$f^{\sup} \in \mathrm{Sci}([a,b]) \qquad y \qquad f_{\inf} \in \mathrm{Scs}([a,b]) \tag{1a}$$

Además, se cumple que

$$f^{\sup} \;\leq\; f \;\leq\; f_{\inf}. \tag{1b}$$

La información recopilada en el resultado anterior permite obtener la siguiente caracterización de las funciones acotadas semicontinuas.

Teorema 3.1.61. *Sea $f : [a,b] \to \mathbb{R}$ una* ***función acotada*** *sobre $[a,b]$. Son equivalentes:*

(1) *$f \in \mathrm{Sci}([a,b])$ (respectivamente, $f \in \mathrm{Scs}([a,b])$).*

(2) *$f = f^{\sup}$, (respectivamente, $f = f_{\inf}$).*

Prueba. $(2) \Rightarrow (1)$ es inmediata ya que $f^{\sup}$ es semicontinua inferiormente. Suponga entonces que (1) se cumple. Por $(1b)$ del Teorema 3.1.60 sabemos que $f^{\sup} \leq f$, de modo que sólo basta demostrar la otra desigualdad. Suponga, para generar una contradicción, que existe algún $x_0 \in [a,b]$ tal que $f^{\sup}(x_0) < f(x_0)$. Como f es semicontinua inferiormente, el Teorema 3.1.55 nos garantiza la existencia de una sucesión creciente $(f_n)_{n=1}^{\infty}$ de funciones continuas tal que $f_n \leq f$ para todo $n \geq 1$ y $\text{lím}_{n\to\infty} f_n(x) = f(x)$ para cada $x \in [a,b]$. En particular, $\text{lím}_{n\to\infty} f_n(x_0) = f(x_0)$. Tome cualquier $0 < \varepsilon < f(x_0) - f^{\sup}(x_0)$ y elija un $N \in \mathbb{N}$ lo suficientemente grande de modo tal que $0 \leq f(x_0) - f_N(x_0) < \varepsilon$. Puesto que $f_N \in \mathrm{Sci}([a,b])$ y $f_N \leq f$ tenemos que

$$f(x_0) \;<\; f_N(x_0) + \varepsilon \;\leq\; f^{\sup}(x_0) + \varepsilon \;<\; f(x_0).$$

Esta contradicción establece que $f(x) = f^{\sup}(x)$ para todo $x \in [a,b]$ y termina la prueba. ■

3.2. Problemas

(1) Construya un subconjunto en $\mathbb{I} = \mathbb{R} \setminus \mathbb{Q}$ que sea perfecto y nunca-denso.

(2) Sea $f : \mathbb{R} \to \mathbb{R}$ la función definida por

$$f(x) = \begin{cases} x & \text{si } x \text{ es irracional} \\ p \cdot \operatorname{sen}(1/q) & \text{si } x = p/q \text{ es un racional en forma irreducible.} \end{cases}$$

Pruebe que f es continua en los irracionales y en $x = 0$, pero es discontinua en cada número racional distinto de cero.

(3) Considere la función de Dirichlet $D : \mathbb{R} \to \mathbb{R}$ dada por

$$D(x) = \lim_{n\to\infty} \lim_{m\to\infty} \big(\cos(2\pi n! x)\big)^m.$$

Pruebe que $D \notin \mathcal{B}_1(\mathbb{R})$, es decir, no existe ninguna sucesión de funciones continuas, digamos $(f_n)_{n=1}^{\infty}$, tal que $D(x) = \lim_{n\to\infty} f_n(x)$.

(4) Una función continua $p : [a,b] \to \mathbb{R}$ se dice que es **lineal a trozos** si existe una partición $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_n = b\}$ del intervalo $[a,b]$ tal que p es lineal en cada intervalo $[x_{k-1}, x_k]$ $k = 1, \ldots, n$. Pruebe que para cada función continua $f : [a,b] \to \mathbb{R}$ y cada $\varepsilon > 0$, existe una función continua lineal a trozos p tal que $|f(x) - p(x)| < \varepsilon$ para todo $x \in [a,b]$.

(5) ***Teorema de Aproximación de Weierstrass***. Suponga que $f \in C([a,b])$. Pruebe que existe una sucesión $(p_n)_{n=1}^{\infty}$ de polinomios tal que

$$\lim_{n\to\infty} p_n = f \quad \textit{uniformemente sobre} \quad [a,b].$$

(6) Sea $\mathcal{F}$ una familia de funciones definidas sobre un conjunto compacto K y a valores reales.

(i) $\mathcal{F}$ se dice **equi-continua** si, dado $\varepsilon > 0$, existe un $\delta > 0$ tal que para cualquier $f \in \mathcal{F}$ se cumple que

$$|f(x) - f(y)| < \varepsilon$$

para todo $x, y \in K$ con $|x - y| < \delta$.

(ii) $\mathcal{F}$ se dice **puntualmente acotada** si, para cada $x \in K$, existe una constante $M_x > 0$ tal que $|f(x)| \leq M_x$ para toda $f \in \mathcal{F}$.

(iii) $\mathcal{F}$ se dice que es **uniformemente acotada** sobre K si existe una constante $M > 0$ tal que $|f(x)| \leq M$ para toda $f \in \mathcal{F}$ y todo $x \in K$.

(a) Suponga que $(f_n)_{n=1}^{\infty}$ es una sucesión en $C(K)$ convergiendo uniformemente a una función f. Pruebe que $(f_n)_{n=1}^{\infty}$ es equicontinua.

(b) Suponga que $(f_n)_{n=1}^{\infty}$ es una sucesión equi-continua en $C(K)$ puntualmente acotada. Pruebe que $(f_n)_{n=1}^{\infty}$ es uniformemente acotada y, que además, posee una subsucesión que converge uniformemente a una función $f \in C(K)$.

(7) Sea $(a_n)_{n=1}^{\infty}$ una sucesión de números reales con $x_n \neq 0$ para todo $n \geq 1$ y defina

$$p_n \;=\; a_1 a_2 \dots a_n \;=\; \prod_{k=1}^{n} a_k \quad n \geq 1.$$

Recordemos que el **producto infinito** $\prod_{n=1}^{\infty} a_n$ es convergente si la sucesión $(p_n)_{n=1}^{\infty}$ tiene un *límite finito no nulo*. Si la sucesión $(p_n)_{n=1}^{\infty}$ carece de límite finito, o si tiende a cero, se dice que el producto infinito $\prod_{n=1}^{\infty} a_n$ es **divergente**.

Pruebe que una serie $\sum_{n=1}^{\infty} a_n$ converge si, sólo si, el producto infinito $\prod_{n=1}^{\infty} \mathrm{e}^{a_n}$ converge. En consecuencia,

$$\prod_{n=1}^{\infty} \mathrm{e}^{a_n} \;=\; \begin{cases} \mathrm{e}^{\sum_{n=1}^{\infty} a_n} & \text{si } \displaystyle\sum_{n=1}^{\infty} a_n < +\infty \\ 0 & \text{si } \displaystyle\sum_{n=1}^{\infty} a_n = +\infty. \end{cases}$$

(8) Sean $f, g \in \mathrm{Sci}([a,b])$ ambas no-negativas. Pruebe que $f \cdot g \in \mathrm{Sci}([a,b])$.

(9) Pruebe que $f \in \mathrm{Sci}([a,b])$ si, y sólo si, el conjunto $\{(x, f(x)) : f(x) \leq x\}$ es cerrado.

CAPÍTULO 4

Desigualdades de Hölder y Minkowski en $\mathbb{R}^n$

Este capítulo está dedicado a presentar, fundamentalmente, las desigualdades de Hölder y Minkowski en $\mathbb{R}^n$. Para lograr dicho objetivo nos pasearemos brevemente por otras importantísimas desigualdades conocidas como las desigualdades de la Media Aritmética y la Media Geométrica.

4.1. Convexidad

La noción de convexidad juega un papel de primer orden en Matemáticas. Recordemos que si X es un espacio vectorial sobre $\mathbb{R}$, un subconjunto G de X se llama **convexo** si, cualesquiera sean $x, y \in G$ y $\lambda \in [0,1]$, se cumple que

$$\lambda x + (1-\lambda)y \in G.$$

Es claro que cualquier subintervalo de $\mathbb{R}$ es un conjunto convexo. Lo interesante es que en $\mathbb{R}$, *todo conjunto convexo es un intervalo*. Si $x_1, \ldots, x_n$ son elementos de un conjunto G (no necesariamente convexo), entonces cualquier combinación de la forma

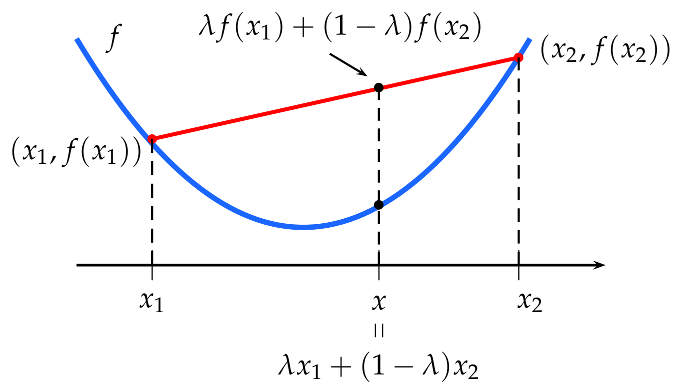
$$\lambda_1 x_1 + \cdots + \lambda_n x_n$$

donde los $\lambda_1, \ldots, \lambda_n$ son no-negativos y satisfacen $\lambda_1 + \cdots + \lambda_n = 1$ es llamada una **combinación convexa** de G. El conjunto de todas las combinaciones convexas de elementos de G constituye un conjunto convexo que llamaremos **la cápsula convexa** de G y denotado por $\operatorname{co}(G)$. De hecho, $\operatorname{co}(G)$ representa el conjunto convexo más pequeño conteniendo a G. Por supuesto, si G es convexo, entonces $G = \operatorname{co}(G)$.

Definición 4.1.1. *Sea I un subintervalo de $\mathbb{R}$. Una función $f : I \to \mathbb{R}$ se dice que es **convexa** en I si*

$$f(\lambda x + (1-\lambda)y) \;\leq\; \lambda f(x) + (1-\lambda)f(y)$$

cualesquiera sean $x, y \in I$ y $\lambda \in [0,1]$.



Geométricamente, la convexidad de f significa que para cada par de puntos $x_1, x_2 \in I$ con $x_1 \neq x_2$, el segmento de línea que une a los puntos $(x_1, f(x_1))$ y $(x_2, f(x_2))$ está por encima de la curva entre x_1 y x_2 (véase la gráfica).

La siguiente desigualdad, válida para cualquier función convexa, es fundamental para la obtención de una versión de la Desigualdad de Jensen para integrales.

Lema 4.1.2. *Si $g : I \to \mathbb{R}$ es una función **convexa** en I, entonces existen $a, b \in \mathbb{R}$ tales que*

$$g(x) \;\geq\; ax + b \quad \text{para todo } x \in I.$$

Prueba. Sean $x, y \in I$ con $x < y$ y sea $x < x_0 < y$. Entonces $x_0 = \lambda x + (1-\lambda)y$, donde $\lambda = \frac{y - x_0}{y - x} \in (0,1)$ y por convexidad,

$$g(x_0) \;\leq\; \lambda g(x) \,+\, (1-\lambda)g(y) \;=\; \frac{y - x_0}{y - x} g(x) \,+\, \left(1 - \frac{y - x_0}{y - x}\right) g(y).$$

Multiplicando el lado izquierdo de la desigualdad anterior por $1 = \frac{(y-x_0)+(x_0-x)}{y-x}$ y reagrupando, resulta que

$$\frac{g(x_0) - g(x)}{x_0 - x} \;\leq\; \frac{g(y) - g(x_0)}{y - x_0},$$

cualquiera sean $x < x_0 < y$ y, por lo tanto, gracias al Teorema 2.1.9, página 87, se tiene que

$$A \;=\; \sup_{x < x_0} \frac{g(x_0) - g(x)}{x_0 - x} \;\leq\; \inf_{y > x_0} \frac{g(y) - g(x_0)}{y - x_0} \;=\; B.$$

Seleccione un a tal que $A \leq a \leq B$ y observe finalmente que

$$g(x) \;\geq\; a(x - x_0) \,+\, g(x_0) \;=\; ax \,+\, (g(x_0) - ax_0).$$

La prueba es completa. ∎

Cualquier recta $L(x) = a(x - x_0) + g(x_0)$ obtenida en el resultado anterior, se llama una **línea tangente** para f en x_0. Si f no es diferenciable en x_0, la pendiente de la recta tangente puede que no esté unívocamente determinada. Sin embrago, cuando f es diferenciable en I, resulta que a se puede reemplazar por $f'(x_0)$ y, en consecuencia, la línea tangente L es única. Más aun, se tiene el siguiente resultado.

Teorema 4.1.3. *Sea I un intervalo abierto y $f : I \to \mathbb{R}$ una función dos veces diferenciable en I. Si $f''(x) \geq 0$ para todo $x \in I$, entonces f es convexa en I.*

Prueba. Sean $x, y \in I$. Para cada $x \leq z \leq y$, existe un escalar $\lambda \in [0,1]$ tal que $z = \lambda x + (1-\lambda)y$. Puesto que $f'' \geq 0$, entonces f' es creciente y, en consecuencia, gracias al Teorema del Valor Medio para derivadas, existen $\xi_1 \in (x,z)$ y $\xi_2 \in (z,y)$ tales que

$$\frac{f(z)-f(x)}{z-x} \;=\; f'(\xi_1) \;\leq\; f'(\xi_2) \;=\; \frac{f(y)-f(z)}{y-z}.$$

Como $1-\lambda = \frac{z-x}{y-x}$, resulta que $z - x = (1-\lambda)(y-x)$ de donde se obtiene que

$$\frac{f(z)-f(x)}{(1-\lambda)(y-x)} \;=\; \frac{f(z)-f(x)}{z-x} \;\leq\; \frac{f(y)-f(z)}{y-z}.$$

Finalmente, si multiplicamos a ambos lados de esta última desigualdad por el número positivo $(1-\lambda)(y-x) > 0$ y teniendo en cuenta, además, que $\frac{1}{\lambda} = \frac{y-x}{y-z}$ resulta entonces que

$$f(\lambda x + (1-\lambda)y) \;=\; f(z) \;\leq\; \lambda f(x) \,+\, (1-\lambda)f(y).$$

La prueba es completa. ∎

El Teorema 4.1.3 permite verificar que la función $f : [0, +\infty) \to \mathbb{R}$ definida por

$$f(x) \;=\; x^p, \quad \text{para} \quad p \geq 1$$

es convexa. Similarmente, la función $g(x) = e^x$ para todo $x \in \mathbb{R}$, es convexa.

Es una tarea fácil de comprobar, usando el Lema 4.1.2, que *toda función convexa cuyo dominio es un intervalo abierto es continua*. Por supuesto, existen funciones convexas definidas sobre un intervalo compacto $I = [a,b]$ que no son continuas. Por ejemplo, la función $f : [0,1] \to \mathbb{R}$ definida por

$$f(x) \;=\; \begin{cases} x^2 & \text{si} \quad x \in (0,1] \\ 1/2 & \text{si} \quad x = 0 \end{cases}$$

es convexa, pero no continua. Se puede probar, además, que: *si f es continua sobre $[a,b]$, entonces f es convexa si, y sólo si,*

$$f\left(\frac{1}{2}x + \frac{1}{2}y\right) \;\leq\; \frac{1}{2}f(x) + \frac{1}{2}f(y).$$

para todo $x, y \in [a,b]$.

4.1.1. Las Desigualdades AM-GM

Algunas desigualdades útiles en Análisis y que ahora presentaremos se derivan de los resultados anteriores. Comencemos.

Teorema 4.1.4 (Desigualdad de Jensen finita). *Sea $f : I \to \mathbb{R}$ una **función convexa** en I y considere cualquier conjunto finito de puntos $x_1, \ldots, x_n \in I$. Entonces*

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \;\leq\; \sum_{i=1}^{n} \lambda_i f(x_i) \tag{1}$$

cualesquiera sean $\lambda_1, \ldots, \lambda_n \in [0,1]$ satisfaciendo $\lambda_1 + \cdots + \lambda_n = 1$.

Prueba. Usemos inducción sobre n. Claramente la conclusión es trivial si $n = 1$. Suponga que la desigualdad (1) es válida para cualquier conjunto de $n-1$ en I sea $n > 1$. Seleccione cualquier conjunto finito de números $\lambda_1, \ldots, \lambda_n \in [0,1]$ satisfaciendo $\lambda_1 + \cdots + \lambda_n = 1$. Defina

$$\beta \; = \; \lambda_1 + \cdots + \lambda_{n-1}.$$

Resulta que $\lambda_n = 1 - \beta$ y como $0 \le \lambda_i/\beta \le 1$ para $i = 1, \ldots, n-1$, entonces $\sum_{i=1}^{n-1} \frac{\lambda_i}{\beta} = 1$, de donde se concluye, usando la hipótesis inductiva, que

$$\begin{aligned} f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \; &\le \; \beta f\left(\sum_{i=1}^{n-1} \frac{\lambda_i}{\beta} x_i\right) + (1-\beta) f(x_n) \\ &\le \; \beta \sum_{i=1}^{n-1} \frac{\lambda_i}{\beta} f(x_i) + (1-\beta) f(x_n) \\ &= \; \sum_{i=1}^{n} \lambda_i f(x_i). \end{aligned}$$

La prueba es completa. ■

Definición 4.1.5. *Si $x_1, \ldots, x_n$ son números positivos, su **media aritmética** y su **media geométrica** se definen, respectivamente, como*

$$\frac{x_1 + \cdots + x_n}{n} \qquad y \qquad \sqrt[n]{x_1 \cdot \ldots \cdot x_n}.$$

Una de las desigualdades más importantes y con un amplio abanico de aplicaciones la constituye la siguiente la desigualdad

$$\sqrt[n]{x_1 \cdot \ldots \cdot x_n} \; \le \; \frac{x_1 + \cdots + x_n}{n}, \qquad (\text{AM-GM})_n$$

llamada la **desigualdad AM-GM**. En [27] se pueden ver 78 demostraciones de dicha desigualdad. De modo más general, si $\lambda_1, \ldots, \lambda_n \in [0,1]$ satisfacen la igualdad $\lambda_1 + \cdots + \lambda_n = 1$, entonces las medias aritmética y geométrica correspondientes a los **pesos** λ_i se definen como

$$\lambda_1 x_1 + \cdots + \lambda_n x_n \qquad \text{y} \qquad x_1^{\lambda_1} \cdot \ldots \cdot x_n^{\lambda_n}.$$

Teorema 4.1.6 (Desigualdad AM-GM). *Sean $x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_n$ números positivos y suponga que $\lambda_1 + \cdots + \lambda_n = 1$. Entonces*

$$x_1^{\lambda_1} \cdot \ldots \cdot x_n^{\lambda_n} \; \le \; \lambda_1 x_1 + \cdots + \lambda_n x_n.$$

Prueba. Sea $f : \mathbb{R} \to \mathbb{R}$ definida por $f(x) = e^x$. Sabemos que f es convexa, de modo que si tomamos $t_i = \ln(x_i)$ para $i = 1, \ldots, n$ tendremos, por la Desigualdad de Jensen, que

$$e^{\lambda_1 t_1 + \cdots + \lambda_n t_n} \; \le \; \lambda_1 e^{t_1} \, + \, \cdots \, + \, \lambda_n e^{t_n}$$

la cual no es otra cosa que la desigualdad requerida ya que $e^{t_i} = x_i$ para cada $i = 1, \ldots, n$. ■

Observe que, tomando $\lambda_i = 1/n$ para $i = 1, \ldots, n$ en el resultado anterior se obtiene la desigualdad $(\text{AM-GM})_n$.

$\|$ ▶ **Existencia de e y su irracionalidad.**

Recordemos que el número e se define, tal vez la más común de todas, como

$$e = \sum_{n=0}^{\infty} \frac{1}{n!}.$$

Nuestro objetivo es ver que él es irracional y usar la Desigualdad AM-GM para expresar a e como

$$e = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n.$$

$\|$ ▶ e es irracional.

Prueba. Observe, en primer lugar, que

$$\begin{aligned}
\frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \frac{1}{(n+3)!} + \cdots &= \frac{1}{(n+1)!}\left(1 + \frac{1}{n+2} + \frac{1}{(n+3)(n+2)} + \cdots\right) \\
&< \frac{1}{(n+1)!}\left(1 + \frac{1}{n+1} + \frac{1}{(n+1)^2} + \cdots\right) \\
&= \frac{1}{(n+1)!} \sum_{k=0}^{\infty} \frac{1}{(n+1)^k} \\
&= \frac{1}{(n+1)!} \cdot \frac{1}{(1 - \frac{1}{n+1})} = \frac{1}{(n+1)!} \cdot \frac{n+1}{n} = \frac{1}{nn!},
\end{aligned}$$

de donde se sigue que

$$1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{n!} < e < 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{n!} + \frac{1}{nn!}. \tag{2}$$

Esta última desigualdad permite estimar a e tanto como se desee con sólo elegir a n lo suficientemente grande: por ejemplo, podemos escribir $e \approx 2{,}718281828459\ldots$ Suponga ahora que e es un racional, digamos $e = p/q$ escrito en forma reducida. Nótese que $q \geq 2$ y que $q! \cdot e$ es un número natural. Si en las desigualdades obtenidas en (2) reemplazamos n por q y las multiplicamos por $q!$, obtenemos

$$q! \cdot \left(1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{q!}\right) < q! \cdot e < q! \cdot \left(1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{q!}\right) + \frac{1}{q}$$

lo cual conduce a la siguiente contradicción: como $q! \cdot e$ y $q!(1 + 1/1! + 1/2! + \cdots + 1/q!)$ son números enteros, resulta que

$$0 < \underbrace{q! \cdot e - q! \cdot \left(1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{q!}\right)}_{\text{un entero}} < \frac{1}{q}$$

Esta contradicción establece que e es un número irracional.

‖ ▶ La Desigualdad AM-GM puede ser utilizada para demostrar que las sucesiones $(a_n)_{n=1}^{\infty}$ y $(b_n)_{n=1}^{\infty}$ definidas por

$$a_n = \left(1+\frac{1}{n}\right)^n \qquad \text{y} \qquad b_n = \left(1+\frac{1}{n}\right)^{n+1}, \quad n = 1, 2, \ldots$$

son monótonas y ambas convergen al mismo número. De hecho, $(a_n)_{n=1}^{\infty}$ es *estrictamente creciente*, mientras que $(b_n)_{n=1}^{\infty}$ es *estrictamente decreciente*. En efecto, si en $(\text{AM-GM})_n$ tomamos:

(a) $x_1 = 1$ y $x_2 = \cdots = x_{n+1} = 1 + \frac{1}{n}$, resulta que

$$\left(1+\frac{1}{n}\right)^{\frac{n}{n+1}} = (x_1 \cdot x_2 \cdot \ldots \cdot x_{n+1})^{\frac{1}{n+1}} < \frac{x_1 + x_2 + \cdots + x_{n+1}}{n+1} = 1 + \frac{1}{n+1},$$

de donde se obtiene, elevando a la potencia $n+1$ a ambos lados de la desigualdad anterior, que

$$a_n < a_{n+1}.$$

(b) Similarmente, si $x_1 = 1$ y $x_2 = \cdots = x_{n+2} = 1 - \frac{1}{n}$, resulta que

$$\left(1-\frac{1}{n+1}\right)^{\frac{n+1}{n+2}} = (x_1 \cdot x_2 \cdot \ldots \cdot x_{n+2})^{\frac{1}{n+2}} < \frac{x_1 + x_2 + \cdots + x_{n+2}}{n+2} = \frac{n+1}{n+2}$$

y, como antes, elevando a la potencia $n+2$ en ambos lados de esta desigualdad y luego tomando recíprocos, se tiene que

$$b_{n+1} < b_n.$$

Puesto que ambas sucesiones son acotadas y satisfacen las desigualdades

$$0 < a_1 < \cdots < a_n < a_{n+1} < \cdots < b_{n+1} < b_n < \cdots < b_1 = 4,$$

se concluye que ellas convergen y lo hacen hacia un único número. Definamos entonces

$$e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^{n+1}$$

De lo anterior se sigue que

$$a_n < a_{n+1} < e < b_{n+1} < b_n \quad \text{para todo} \quad n \geq 1.$$

Existen otras desigualdades más generales que la AM - GM. Ellas son conocidas como las desigualdades de la Media Generalizada.

Teorema 4.1.7 (Desigualdad de la Media Generalizada). *Sean $x_1, \ldots, x_n$ números positivos y sean $\lambda_1, \ldots, \lambda_n \in [0,1]$ tales que $\lambda_1 + \cdots + \lambda_n = 1$. Entonces, para cualquier par $p, q \in \mathbb{R}$ con $0 \leq p < q$, se cumple que*

$$\left(\lambda_1 x_1^p + \cdots + \lambda_n x_n^p\right)^{1/p} \leq \left(\lambda_1 x_1^q + \cdots + \lambda_n x_n^q\right)^{1/q}.$$

Prueba. Si $p = 1$ y $q > 1$, la Desigualdad de la Media Generalizada toma la forma

$$(\lambda_1 x_1 + \cdots + \lambda_n x_n)^q \leq \lambda_1 x_1^q + \cdots + \lambda_n x_n^q$$

la cual sigue inmediatamente de la convexidad de la función $f(x) = x^q$. Suponga ahora que $0 < p < q$. Puesto que $q/p > 1$, de lo anterior vemos que

$$(\lambda_1 x_1^p + \cdots + \lambda_n x_n^p)^{q/p} \leq \lambda_1 (x_1^p)^{q/p} + \cdots + \lambda_n (x_n^p)^{q/p}$$

la cual no es otra cosa que la desigualdad requerida. El caso en que p o q toman valores negativos se deja a cargo del lector. ■

Una de las desigualdades importantes que permite definir una norma completa en los espacios $L_p(\mu)$ para $p \geq 1$, es la siguiente:

Teorema 4.1.8 (Desigualdad de Young). *Sean $x, y \in [0, +\infty)$. Si $p, q \in (1, +\infty)$ son tales que $1/p + 1/q = 1$, entonces*

$$xy \leq \frac{x^p}{p} + \frac{x^q}{q}.$$

Prueba. Esto es consecuencia de la Desigualdad AM - GM ya que

$$xy = (x^p)^{1/p} (y^q)^{1/q} \leq \frac{1}{p} x^p + \frac{1}{q} y^q.$$

■

4.1.2. Las Desigualdades de Hölder y Minkowski

Existen otras dos desigualdades famosas llamadas la Desigualdad de Hölder y la Desigualdad de Minkowski que son las encargadas de garantizarnos que una cierta aplicación definida sobre $\mathbb{R}^n$ constituye, de hecho, una norma sobre $\mathbb{R}^n$.

Recordemos que, para cada $n \in \mathbb{N}$, $\mathbb{R}^n$ denota la colección de todas las n-uplas de números reales $x = (x_1, \ldots, x_n)$. Entonces $\mathbb{R}^n$ es un espacio vectorial sobre $\mathbb{R}$ con las operaciones usuales de sumas de n-uplas y multiplicación de un escalar por una n-upla. La base estándar de $\mathbb{R}^n$ la denotaremos por $\mathcal{B}_n = \{e_1, \ldots, e_n\}$, donde

$$e_1 = (1, 0, \ldots, 0), \quad e_2 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, 0, \ldots, 0, 1).$$

Esto significa que cada $x = (x_1, \ldots, x_n)$ se expresa de modo único en la forma

$$x = x_1 \cdot e_1 + \cdots + x_n \cdot x_n.$$

También, sobre $\mathbb{R}$ se define el *producto interno* usual $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ dado por

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i \cdot y_i \quad \text{para todo} \quad x, y \in \mathbb{R}^n.$$

Este producto interno induce una norma sobre $\mathbb{R}^n$, llamada la *norma euclídea* y definida por

$$\| x \|_2 = \sqrt{\langle x, x \rangle} = (x_1^2 + \cdots + x_n^2)^{1/2}.$$

En general, si para cada número $p \in [1, +\infty)$, definimos la aplicación $\|\cdot\|_p : \mathbb{R}^n \to [0, +\infty)$ por

$$\|(x_1, \ldots, x_n)\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$$

entonces $\|\cdot\|_p$ es una norma sobre $\mathbb{R}^n$, llamada la p-**norma**. Si $p = +\infty$, pondremos

$$\|x\|_\infty = \max_{1 \le i \le n} |x_i|,$$

para cualquier $x \in \mathbb{R}^n$. Es claro que si $p \in \{1, +\infty\}$, entonces $\|\cdot\|_p$ es un norma. De hecho, lo mismo es cierto para cualquier $p \in (1, +\infty)$. Para ver esto, observe que $\|\cdot\|_p$ satisface las propiedades $(1)-(3)$ que definen una norma. Falta verificar la desigualdad triangular la cual es la conclusión del siguiente resultado también conocido como:

Teorema 4.1.9 (Desigualdad de Minkowski). *Sea* $p \in [1, +\infty)$. *Entonces*

$$\left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{1/p} \le \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{1/p}.$$

para cualesquiera $\mathbf{x} = (x_1, \ldots, x_n)$ y $\mathbf{y} = (y_1, \ldots, y_n)$ *en* $\mathbb{R}^n$.

Prueba. Puesto que $p \ge 1$, la función $x \mapsto |x|^p$ es convexa. De esto se sigue que si $\lambda \in [0, 1]$, entonces

$$\begin{aligned}
\|\lambda \cdot x + (1-\lambda) \cdot y\|_p^p &= \sum_{i=1}^{n} |\lambda x_i + (1-\lambda) y_i|^p \\
&\le \sum_{i=1}^{n} \lambda |x_i|^p + (1-\lambda)|y_i|^p \\
&= \lambda \|x\|_p^p + (1-\lambda)\|x\|_p^p
\end{aligned}$$

para todo $x, y \in \mathbb{R}^n$. En particular, si $\|x\|_p = \|y\|_p = 1$, entonces

$$\|\lambda \cdot x + (1-\lambda) \cdot y\|_p \le 1.$$

Para ver el caso general, suponga que x y y son vectores no-nulos en $\mathbb{R}^n$. Entonces $x/\|x\|_p$ y $y/\|y\|_p$ son vectores de norma igual a 1 y así, por la primera parte,

$$\frac{\|x+y\|_p}{\|x\|_p + \|y\|_p} = \left\| \frac{\|x\|_p}{\|x\|_p + \|y\|_p} \frac{x}{\|x\|_p} + \frac{\|y\|_p}{\|x\|_p + \|y\|_p} \frac{y}{\|y\|_p} \right\|_p \le 1$$

Esto termina la prueba. ■

La siguiente desigualdad, comúnmente conocida como la *Desigualdad de Hölder*, o también como *Desigualdad de Cauchy-Schwarz* para el caso cuando $p = q = 2$, se obtiene como consecuencia de la Desigualdad de Young. En muchos textos ella es usada para dar otra demostración de la Desigualdad de Minkowski.

Teorema 4.1.10 (Desigualdad de Hölder). *Sean $x, y \in \mathbb{R}^n$ y suponga que $p, q \in (1, +\infty)$ satisfacen la igualdad $1/p + 1/q = 1$. Entonces*

$$\left|\langle x, y\rangle\right| \;\leq\; \|x\|_p \, \|y\|_q \, .$$

Prueba. Pongamos $x = (x_1, \ldots, x_n)$ y $y = (y_1, \ldots, y_n)$. Por la Desigualdad de Young

$$\begin{aligned}
\left|\langle x, y\rangle\right| \;&\leq\; |x_1 y_1| + \cdots + |x_n y_n| \\
&\leq\; \frac{|x_1|^p + \cdots + |x_n|^p}{p} + \frac{|y_1|^q + \cdots + |y_n|^q}{q} \\
&=\; \frac{\|x\|_p^p}{p} + \frac{\|y\|_q^q}{q}.
\end{aligned}$$

En particular, si $\|x\|_p = \|y\|_q = 1$, entonces $\left|\langle x, y\rangle\right| \leq 1$. Para el caso general, sean x y y vectores no-nulos en $\mathbb{R}^n$. Entonces $x/\|x\|_p$ y $y/\|y\|_q$ son vectores unitarios y así, gracias a la primera parte,

$$\left|\left\langle \frac{x}{\|y\|_p}, \frac{y}{\|y\|_q} \right\rangle\right| \;\leq\; 1.$$

De esto se concluye la prueba. ■

En lo que sigue escribiremos ℓ_p^n para designar a $\mathbb{R}^n$ provisto de la norma $\|\cdot\|_p$, para cualquier $p \in [1, +\infty]$. El siguiente resultado indica que todas las normas $\|\cdot\|_p$ sobre $\mathbb{R}^n$ son equivalente mostrando, además, cómo se comparan tales normas.

Teorema 4.1.11 (Equivalencia de las p-normas sobre $\mathbb{R}^n$). *Todas las p-normas sobre $\mathbb{R}^n$ son* ***equivalentes****, $p \in [1, +\infty]$. Más aun, para todo $x \in \mathbb{R}^n$:*

(1) Si $1 \leq p \leq +\infty$, entonces

$$\|x\|_\infty \;\leq\; \|x\|_p \;\leq\; n^{1/p} \, \|x\|_\infty \, . \tag{1}$$

(2) Si $1 \leq p < q < +\infty$, entonces existe una constante $a > 0$ tal que

$$\|x\|_p \;\leq\; a \cdot \|x\|_q \, . \tag{2}$$

Prueba. Sea $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. Puesto $|x_i| = \sqrt[p]{|x_i|^p} \leq \sqrt[p]{|x_1|^p + \cdots + |x_n|^p} = \|x\|_p$ para cada $i = 1, \ldots, n$, resulta que

$$\|x\|_\infty \;=\; \text{máx}\left\{|x_1|, \ldots, |x_n|\right\} \;\leq\; \|x\|_p \;\leq\; \sqrt[p]{\|x\|_\infty^p + \cdots + \|x\|_\infty^p} \;=\; n^{1/p} \, \|x\|_\infty \, .$$

Esto prueba (1). Para demostrar (2), sea $x \in \mathbb{R}^n$ y suponga que $1 \leq p < q < +\infty$. Defina $r = q/p > 1$ y elija $s > 1$ tal que $1/r + 1/s = 1$. Usando la Desigualdad de Hölder y el hecho de que $q = pr$, se tiene que

$$\begin{aligned}
\|x\|_p^p \;&=\; \sum_{i=1}^n |x_i|^p \cdot 1 \\
&\leq\; \left(\sum_{i=1}^n \left(|x_i|^p\right)^r\right)^{1/r} \left(\sum_{i=1}^n 1^s\right)^{1/s} \\
&=\; \|x\|_q^p \cdot n^{1/s}.
\end{aligned}$$

Puesto que $1/ps = 1/p - 1/q$, resulta que tomando $a = n^{1/ps}$ se obtiene el resultado deseado.■

Una consecuencia importante de estas desigualdades es que cualquier transformación lineal de $\mathbb{R}^n$ en $\mathbb{R}^m$ es continua. Recordemos que toda aplicación $T : \mathbb{R}^n \to \mathbb{R}^m$ que cumple con las propiedades:

$$T(x+y) = T(x) + T(y) \qquad \text{y} \qquad T(\lambda \cdot x) = \lambda \cdot T(x)$$

para todo $x, y \in \mathbb{R}$ y todo $\lambda \in \mathbb{R}$, es llamada una **Transformación Lineal**.

Teorema 4.1.12. *Si $T : (\mathbb{R}^n, \|\cdot\|_p) \to (\mathbb{R}^m, \|\cdot\|_p)$ es una **transformación lineal**, entonces T es **continua**.*

Prueba. Sea $x \in \mathbb{R}^n$ y sea $q \geq 1$ tal que $1/p + 1/q = 1$. Por la linealidad de T tenemos que

$$T(x) = T(x_1 e_1 + \cdots + x_n e_n) = x_1 \cdot T(e_1) + \cdots + x_n \cdot T(e_n)$$

y entonces, por la desigualdad triangular de la norma, resulta que

$$\| T(x) \|_p = \| x_1 \cdot T(e_1) + \cdots + x_n \cdot T(e_n) \|_p \leq \sum_{i=1}^{n} |x_i| \, \| T(e_i) \|_p .$$

Si definimos

$$M = \sqrt[q]{\sum_{i=1}^{n} \| T(e_i) \|_p^q},$$

entonces la Desigualdad de Hölder nos muestra que

$$\| Tx \|_p \leq M \cdot \| x \|_p .$$

Finalmente, si $x, y \in \mathbb{R}$, entonces como $T(x-y) = T(x) - T(y)$ se obtiene que

$$\| T(x) - T(y) \|_p \leq M \cdot \| x - y \|_p$$

y termina la prueba. ■

Si $T : (\mathbb{R}^n, \|\cdot\|) \to (\mathbb{R}^m, \|\cdot\|)$ es una transformación lineal, se sigue del resultado anterior que ella es continua, de hecho T es Lipschitz. Esto permite definir

$$\| T \|_{\text{op}} = \inf \{ c > 0 : \| Tx \| \leq c \| x \| \text{ para todo } x \in \mathbb{R}^n \}.$$

Es fácil establecer que $\|\cdot\|_{\text{op}}$ es una norma sobre el espacio vectorial $\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ *formado por todas las **transformaciones lineales** (continuas) de* $\mathbb{R}^n$ *en* $\mathbb{R}^m$. Más aun, si $T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$, entonces

$$\| T \|_{\text{op}} = \sup \{ \| Tx \| : \| x \| \leq 1 \},$$

de donde resulta que

$$\| T(x) \| \leq \| T \|_{\text{op}} \| x \| \quad \text{para todo} \quad x \in \mathbb{R}^n.$$

En un espacio normado $(X, \|\cdot\|)$, decir que una sucesión $(x_n)_{n=1}^{\infty}$ en **converge** (en la norma) a un vector x, significa que

$$\lim_{n\to\infty} \| x_n - x \| = 0.$$

Observe que, gracias a (1), si la sucesión $(x_n)_{n=1}^{\infty}$ converge a x, entonces

$$\lim_{n\to\infty} \| x_n \| = \| x \| ,$$

pero el recíproco no siempre es cierto. Un subconjunto A de X se dice que **norma-acotado** o, simplemente, **acotado**, si existe una constante $M > 0$ tal que $\| x \| \leq M$ para todo $x \in A$. Si (X, d) es un espacio métrico completo, donde la métrica d es dada por $(2a)$, entonces diremos que $(X, \|\cdot\|)$ es un **espacio de Banach**. El ejemplo clásico de un espacio de Banach lo constituye $(\mathbb{R}^n, \|\cdot\|_p)$, para cada $n \geq 1$, donde

$$\| (x_1, \dots, x_n) \|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

para $p \geq 1$. Dada una sucesión $(x_n)_{n=1}^{\infty}$ en X, diremos que la serie $\sum_{n=1}^{\infty} x_n$ **converge** (en la norma) si la sucesión $(\sum_{k=1}^{n} x_k)_{n=1}^{\infty}$ de sus sumas parciales converge a algún $x \in X$, esto es,

$$\lim_{n\to\infty} \left\| \sum_{k=1}^{n} x_k - x \right\| = 0.$$

Si este es el caso escribiremos, como es usual, $x = \sum_{n=1}^{\infty} x_n$ y diremos que la serie $\sum_{n=1}^{\infty} x_n$ es **convergente** y que x es su **suma**. La serie $\sum_{n=1}^{\infty} x_n$ se dice que es **absolutamente convergente** si

$$\sum_{n=1}^{\infty} \| x_n \| < +\infty.$$

Observe que si $(x_n)_{n=1}^{\infty}$ es cualquier sucesión en X, entonces para cualquier $n \geq 2$,

$$x_n = x_1 + \sum_{k=2}^{n} (x_k - x_{k-1}),$$

lo cual permite concluir que la sucesión $(x_n)_{n=1}^{\infty}$ y la serie $\sum_{k=2}^{\infty} (x_k - x_{k-1})$ poseen un romance en extremo: o ambas convergen o ambas divergen. Este controversial romance interviene de modo directo en la demostración del siguiente resultado fundamental.

Lema 4.1.13. *Sea $(X, \|\cdot\|)$ un espacio normado. Las siguientes condiciones son equivalentes:*

*(1) $(X, \|\cdot\|)$ es un **espacio de Banach**.*

*(2) Toda serie **absolutamente convergente** es **convergente**.*

Prueba. (1) $\Rightarrow$ (2). Suponga que $(X, \|\cdot\|)$ es un espacio de Banach y sea $\sum_{n=1}^{\infty} x_n$ una serie absolutamente convergente. Para demostrar que la serie $\sum_{n=1}^{\infty} x_n$ converge en X, es suficiente demostrar que la sucesión de sus sumas parciales $(s_n)_{n=1}^{\infty}$ es de Cauchy en X, donde $s_n = \sum_{k=1}^{n} x_k$ para cada $n \in \mathbb{N}$. Observe que cualesquiera sean $m, n \in \mathbb{N}$ con $m > n$, se tiene que

$$\| s_m - s_n \| = \left\| \sum_{k=n+1}^{m} x_k \right\| \leq \sum_{k=n+1}^{m} \| x_k \|$$

y como la serie $\sum_{n=1}^{\infty} \| x_n \|$ converge, resulta entonces que $\sum_{k=n+1}^{m} \| x_k \| \to 0$ cuando $m, n \to \infty$. Esto prueba que la sucesión de sumas parciales $(s_n)_{n=1}^{\infty}$ es de Cauchy en X y, en consecuencia, converge ya que $(X, \| \cdot \|)$ es completo.

(2) $\Rightarrow$ (1). Suponga que toda serie absolutamente convergente es convergente y sea $(x_n)_{n=1}^{\infty}$ una sucesión de Cauchy en X. Para demostrar que ella converge en X, es suficiente encontrar una subsucesión de $(x_n)_{n=1}^{\infty}$ que converja. Para lograr tal objetivo, use el hecho de que la sucesión $(x_n)_{n=1}^{\infty}$ es de Cauchy para determinar una subsucesión $(x_{n_k})_{k=1}^{\infty}$ que satisfaga

$$\| x_{k+1} - x_k \| < 2^{-k} \quad \text{para todo} \quad k \geq 1.$$

De esto se sigue que la serie $\sum_{k=1}^{\infty} \| x_{n_{k+1}} - x_{n_k} \|$ es absolutamente convergente ya que

$$\sum_{k=1}^{\infty} \| x_{n_{k+1}} - x_{n_k} \| < \sum_{k=1}^{\infty} 2^{-k} < +\infty.$$

Haciendo uso de nuestra hipótesis, se tiene que la serie $\sum_{k=1}^{\infty} (x_{n_{k+1}} - x_{n_k})$ converge en X y, en consecuencia, como fue observado anteriormente, la subsucesión $(x_{n_k})_{k=1}^{\infty}$ converge pues

$$x_{n_{m+1}} = x_{n_1} + \sum_{k=1}^{m} (x_{n_{k+1}} - x_{n_k}).$$

Esto termina la prueba. ■

4.2. Problemas

(1) Pruebe que una función $f : \mathbb{R} \to \mathbb{R}$ es convexa si, y sólo si,

$$\frac{f(y) - f(x)}{y - x} \leq \frac{f(z) - f(x)}{z - x} \leq \frac{f(z) - f(y)}{z - y}$$

para todo $x, y, z \in \mathbb{R}$ con $x < y < z$. Deduzca de esto que las derivadas laterales f'_- y f'_+ son funciones crecientes.

(2) Sea $f : \mathbb{R} \to \mathbb{R}$ una función convexa y para cada $a, b \in \mathbb{R}$ con $a < b$, considere el número $M = \text{máx}\{-f'_+(a), f'_-(b)\}$. Pruebe que

$$|f(x) - f(y)| \leq M|x - y| \quad \text{para todo} \quad x, y \in [a, b].$$

En particular, f es continua en todo punto de $\mathbb{R}$. Concluya que f es convexa si, y sólo si,

$$f\left(\frac{1}{2}x + \frac{1}{2}y\right) \leq \frac{1}{2}f(x) + \frac{1}{2}f(y).$$

para todo $x, y \in \mathbb{R}$.

(3) Sea $f : \mathbb{R} \to \mathbb{R}$ una función convexa. Pruebe que

$$f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y)$$

para todo $\lambda \in \mathbb{R} \setminus [0, 1]$.

(4) Sea $f : \mathbb{R} \to \mathbb{R}$ una función convexa y sea

$$Z_f \;=\; \{x \in \mathbb{R} \,:\, f \text{ no es diferenciable en } \; x\}.$$

Pruebe que Z_f es a lo más numerable.

(5) Sea $f : \mathbb{R} \to \mathbb{R}$ una función convexa y para cada $x \in \mathbb{R}$ considere el intervalo cerrado

$$\partial f(x) \;=\; [f'_-(x), f'_+(x)].$$

Pruebe que:

(i) f es diferenciable en x si, y sólo si, $\partial f(x)$ consiste de un sólo punto.

(ii) Para cada $x_0 \in \mathbb{R}$ y cada $s \in \partial f(x_0)$, se cumple

$$f(x) \;\geq\; f(x_0) + s \cdot (x - x_0) \quad \text{para todo } x \in \mathbb{R}.$$

(iii) Para cada $x, y \in \mathbb{R}$ con $x < y$, existe $c \in (x, y)$ tal que

$$\frac{f(y) - f(x)}{y - x} \in \partial f(c).$$

(6) Sea $f : (0, +\infty) \to \mathbb{R}$ una función convexa y defina la función $F : (0, +\infty) \to \mathbb{R}$ por

$$F(x) \;=\; \frac{1}{x} \int_0^x f(t)\, dt \quad \text{para todo } \; x > 0.$$

Pruebe que F es convexa.

(7) Sean $x_1, \ldots, x_n, y_1, \ldots, y_n \in (0, +\infty)$. Pruebe que

$$\frac{\sqrt{x_1 y_1} + \cdots + \sqrt{x_n y_n}}{n} \;\leq\; \sqrt{\frac{x_1 + \cdots + x_n}{n}}\, \sqrt{\frac{y_1 + \cdots + y_n}{n}}.$$

(8) Sean $\lambda, \nu \in [0, 1]$ tales que $\lambda + \nu = 1$. Pruebe que

$$\frac{x_1^\lambda y_1^\nu + \cdots + x_n^\lambda y_n^\nu}{n} \;\leq\; \left(\frac{x_1 + \cdots + x_n}{n}\right)^\lambda \left(\frac{y_1 + \cdots + y_n}{n}\right)^\nu.$$

(9) Pruebe que $(C([a, b]), \|\cdot\|_\infty)$ es un espacio de Banach.

(10) Sea $f : [a, b] \to \mathbb{R}$ una función acotada. Pruebe que f es continua en $[a, b]$ si, y sólo si, $Gra(f) = \{(x, f(x)) : x \in [a, b]\}$ es cerrado en $\mathbb{R}^2$.

(11) (a) Sea $(r_n)_{n=1}^\infty$ una enumeración de los racionales en $(0, 1)$ y considere, para cada $n \geq 1$, la función $f_n : [0, 1] \to \mathbb{R}$ definida por

$$f_n(x) \;=\; \begin{cases} 0 & \text{si } \; x \in [0, r_n), \\ 1/2^n & \text{si } \; x \in [r_n, 1]. \end{cases}$$

Sea $f = \sum_{n=1}^\infty f_n$. Pruebe que f es creciente, acotada y discontinua precisamente en los racionales de $(0, 1)$.

(b) Sea $f : [0,1] \to \mathbb{R}$ definida por

$$f(x) \;=\; \sum_{0 \,\le\, \frac{m}{n} \,\le\, x} \frac{1}{n^3}, \quad (m,n) = 1$$

para cada $x \in [0,1]$. Pruebe que f es creciente, continua en los irracionales y discontinua en los racionales.

(12) Sea $W : [0,1] \to \mathbb{R}$ definida por

$$W(x) \;=\; \sum_{n=0}^{\infty} 10^{-n} f(10^n x),$$

donde $f(x) = |x - k|$ siempre que $x \in [k - (1/2), x + (1/2)]$, $k \in \mathbb{Z}$. Pruebe que la función W es continua pero nunca-diferenciable en $[0,1]$.

(13) Sea $G \subseteq \mathbb{R}$. Pruebe que $\chi_G : \mathbb{R} \to \mathbb{R}$ es semicontinua inferiormente (resp. superiormente) si, y sólo si G es abierto (resp. cerrado).

(14) Para cada $x \in [a,b]$, sea $\mathcal{N}_x$ la familia de todos los entornos abiertos de x. Suponga que $f : [a,b] \to \mathbb{R}$ es una función arbitraria y defina

$$\underline{f}(x) \;=\; \sup_{V \in \mathcal{N}_x} \inf_{\zeta \in V} f(\zeta) \qquad \text{y} \qquad \overline{f}(x) \;=\; \inf_{V \in \mathcal{N}_x} \sup_{\zeta \in V} f(\zeta)$$

para cada $x \in [a,b]$. Pruebe que:

(a) Si $(I_n)_{n=1}^{\infty}$ es cualquier sucesión decreciente de intervalos tal que $\ell(I_n) \to 0$ y $x \in int(I_n)$ para cada n, entonces

$$\underline{f}(x) \;=\; \lim_{n\to\infty} \Big(\inf_{\zeta \in I_n} f(\zeta) \Big) \qquad \text{y} \qquad \overline{f}(x) \;=\; \lim_{n\to\infty} \Big(\sup_{\zeta \in I_n} f(\zeta) \Big).$$

(b) $\underline{f} \le f \le \overline{f}$.

(c) $\underline{f}$ es semicontinua inferiormente y $\overline{f}$ es semicontinua superiormente.

(d) f es semicontinua inferiormente (respectivamente, semicontinua superiormente) si, y sólo si, $f = \underline{f}$ (respectivamente, $f = \overline{f}$).

(15) Sea $J \subseteq [0,1]$. Una familia $(V_j)_{j\in J}$ de subconjuntos de $[0,1]$ se dice que es un **cubrimiento creciente** de $[0,1]$ si

$$[0,1] \;=\; \bigcup_{j \in J} V_j \qquad \text{y} \qquad V_i \subseteq V_j \;\; \text{siempre que} \;\; i \le j.$$

Asocie, a cada cubrimiento creciente $(V_j)_{j\in J}$ de $[0,1]$, la función $f : [0,1] \to \mathbb{R}$ definida por

$$f(x) \;=\; \inf\big\{ j \in J \;:\; x \in V_j \big\}.$$

Pruebe que si J es denso en $[0,1]$ y los V_j son cerrados (respectivamente, abiertos), entonces f es semicontinua inferiormente, (respectivamente, superiormente).

(16) Sea $\{A_\alpha : \alpha \in D\}$ una familia arbitraria de conjuntos y sean

$$G = \bigcup_{\alpha \in D} A_\alpha \qquad \text{y} \qquad F = \bigcap_{\alpha \in D} A_\alpha.$$

Pruebe que

$$\chi_G = \sup_{\alpha \in D} \chi_{A_\alpha} \qquad \text{y} \qquad \chi_F = \inf_{\alpha \in D} \chi_{A_\alpha}.$$

(17) Sea $(A_n)_{n=1}^{\infty}$ una sucesión de conjuntos. Pruebe que

$$\chi_{\overline{\lim}_{n\to\infty} A_n} = \overline{\lim_{n\to\infty}} \chi_{A_n} \qquad \text{y} \qquad \chi_{\underline{\lim}_{n\to\infty} A_n} = \underline{\lim_{n\to\infty}} \chi_{A_n}$$

(18) Demuestre que $\sum_{n=0}^{\infty} a_n/n!$ es un número irracional si cada $a_n \in \{0,1\}$ y si $a_n = 1$ para infinitos n's.

CAPÍTULO 5

El Conjunto de Cantor y su Media Hermana

Esta sección esta dedicada fundamental a la construcción del Conjunto Ternario de Cantor, un conjunto que constituye una fuente casi inagotable de ejemplos y contra-ejemplos en muchas ramas de las matemáticas. Su construcción se lleva a cabo por medio de un proceso infinito de eliminación de ciertos intervalos abiertos y lo que queda de tal proceso es el conjunto de Cantor. De inmediato se comienzan a mostrar las casi inimaginables propiedades que dicho conjunto posee. Luego pasamos a construir a sus parientes más cercanos: los así llamados Conjuntos Tipo-Cantor con propiedades muy similares al conjunto de Cantor y se finaliza con la Función de Cantor, una función también con característica muy especiales.

5.1. Representaciones Ternarias y Binarias

Para poder entender en profundidad las propiedades que posee el Conjunto Ternario de Cantor debemos pasearnos antes por conocer las propiedades de las representaciones ternarias de los puntos que habitan en $[0,1]$.

Dada una sucesión de números reales, $(x_n)_{n=1}^{\infty}$, considere la sucesión $(s_n)_{n=1}^{\infty}$ de sus sumas parciales, donde $s_n = \sum_{j=1}^{n} x_j$ para todo $n \in \mathbb{N}$. Si la sucesión $(s_n)_{n=1}^{\infty}$ converge a un punto $s \in \mathbb{R}$, escribiremos

$$s = \sum_{n=1}^{\infty} x_n = \lim_{n\to\infty} s_n$$

y diremos que la serie $\sum_{n=1}^{\infty} x_n$ **converge** a s. Por ejemplo, para cada $x \in \mathbb{R}$ y cada $n \in \mathbb{N}$, sabemos que

$$\sum_{j=0}^{n} x^j = \begin{cases} \dfrac{1-x^{n+1}}{1-x} & \text{si} \quad x \neq 1 \\ n+1 & \text{si} \quad x = 1 \end{cases}$$

y que si $|x| < 1$, entonces $\lim_{n\to\infty} x^n = 0$. De esto se sigue el siguiente resultado conocido como el Teorema de la Serie Geométrica.

Teorema 5.1.1 (Serie Geométrica). *Para cada $x \in \mathbb{R}$, la serie geométrica $\sum_{n=0}^{\infty} x^n$ **converge** en $\mathbb{R}$ si, y sólo si, $|x| < 1$. En este caso,*

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}.$$

Observe que para cualquier $m \geq 0$,

$$\sum_{n=m}^{\infty} x^n = \sum_{n=0}^{\infty} x^{m+n} = x^m \cdot \left(\sum_{n=0}^{\infty} x^n \right) = x^m \left(\frac{1}{1-x} \right),$$

siempre que $|x| < 1$. Por ejemplo, si $x = 1/k$, donde $k \in \{2, 3, \dots, 9\}$, entonces

$$\sum_{n=m+1}^{\infty} \frac{1}{k^n} = \frac{1}{k^{m+1}} \left(\frac{1}{1 - \frac{1}{k}} \right) = \frac{1}{k^m(k-1)}. \tag{1_m^k}$$

En particular, tomando $k = 3$ en la igualdad anterior, resulta que

$$\sum_{n=m+1}^{\infty} \frac{2}{3^n} = \frac{1}{3^m}. \tag{1_m^3}$$

Por consiguiente,

$$\sum_{n=1}^{\infty} \frac{2}{3^n} = 1, \qquad \sum_{n=2}^{\infty} \frac{2}{3^n} = \frac{1}{3}, \qquad \sum_{n=3}^{\infty} \frac{2}{3^n} = \frac{1}{3^2}, \cdots$$

y así,

$$\frac{1}{3} + \sum_{n=2}^{\infty} \frac{2}{3^n} = \frac{2}{3}, \qquad \frac{2}{3} + \sum_{n=3}^{\infty} \frac{2}{3^n} = \frac{7}{3^2}, \cdots$$

5.1.1. Representaciones Ternarias

Suponga que para $n \in \mathbb{N}$, $a_n \in \{0, 1, 2\}$. Entonces, $0 \leq a_n/3 < 1$ y se sigue del Teorema de la Serie Geométrica que la serie $\sum_{n=1}^{\infty} a_n/3^n$ converge a un elemento $x \in \mathbb{R}$. En este caso a x también lo representaremos en la forma

$$x = (\alpha_0, a_1 a_2 a_3 \dots)_3$$

y diremos que $(\alpha_0, a_1 a_2 a_3 \dots)_3$ es una **representación ternaria** de x. Los números 0, 1 y 2 son llamados los **dígitos ternarios**.

Lema 5.1.2. *Sea $x \in \mathbb{R}$ con $0 < x < 1$. Entonces, para cada $n \in \mathbb{N}$, existen **dígitos ternarios** $a_1, \dots, a_n$ tales que*

$$\frac{a_1}{3} + \frac{a_2}{3^2} + \cdots + \frac{a_n}{3^n} \leq x < \frac{a_1}{3} + \frac{a_2}{3^2} + \cdots + \frac{a_n}{3^n} + \frac{1}{3^n}. \tag{1}$$

Prueba. La prueba es por inducción sobre n. Puesto que $0 < x < 1$, entonces $0 < 3x < 3$, de modo que si definimos $a_1 = [3x]$, la parte entera de $3x$, tendremos que a_1 es un dígito ternario tal que $a_1 \leq 3x < a_1 + 1$. De aquí se sigue que

$$\frac{a_1}{3} \leq x < \frac{a_1}{3} + \frac{1}{3},$$

lo cual es (1) para $n = 1$. Suponga que han sido construidos los dígitos ternarios $a_1, \ldots, a_n$ satisfaciendo (1) y defina

$$q_n = \frac{a_1}{3} + \frac{a_2}{3^2} + \cdots + \frac{a_n}{3^n}.$$

Entonces (1) es igual a

$$q_n \leq x < q_n + \frac{1}{3^n}$$

de donde se obtiene que

$$0 \leq 3^{n+1}(x - q_n) < 3.$$

Si ahora definimos $a_{n+1} = [3^{n+1}(x - q_n)]$, tendremos que $a_{n+1} \in \{0, 1, 2\}$ y se cumple que

$$a_{n+1} \leq 3^{n+1}(x - q_n) < a_{n+1} + 1,$$

es decir,

$$q_n + \frac{a_{n+1}}{3^{n+1}} \leq x < q_n + \frac{a_{n+1}}{3^{n+1}} + \frac{1}{3^{n+1}},$$

lo cual es (1) para $k = n + 1$. ■

Puesto que los números 0 y 1 poseen las representaciones ternarias

$$0 = (0{,}000\ldots)_3 \qquad \text{y} \qquad 1 = (0{,}222\ldots)_3,$$

entonces, del Lema 5.1.2, se concluye que:

Corolario 5.1.3. *Cada $x \in [0, 1]$ posee **al menos una representación ternaria**.*

Por otro lado, observe que las fracciones $\frac{1}{3}$, $\frac{2}{3}$ se pueden representar en la forma:

$$\frac{1}{3} = (0{,}100\ldots)_3 = (0{,}0222\ldots)_3, \qquad \frac{2}{3} = (0{,}200\ldots)_3 = (0{,}1222\ldots)_3.$$

En general, cualquier racional de la forma $m/3^n$ admite dos representaciones ternarias, pues

$$\frac{m}{3^n} = \frac{m-1}{3^n} + \frac{1}{3^n} = \frac{m-1}{3^n} + \sum_{k=n+1}^{\infty} \frac{2}{3^k}.$$

Estos ejemplos ponen en evidencian que hay números reales que poseen al menos dos representaciones ternarias. El siguiente resultado establece que la representación ternaria de cualquier $x \in (0, 1)$ no puede ser más de dos.

Teorema 5.1.4. *Sea $x \in \mathbb{R}$ con $0 < x < 1$. Entonces x posee **a lo sumo dos representaciones ternarias**.*

Prueba. Suponga que x posee al menos dos representaciones ternarias, digamos

$$\begin{aligned} x &= (0, a_1a_2a_3\ldots)_3 \\ &= (0, b_1b_2b_3\ldots)_3 \end{aligned}$$

Sea m el primer valor de n para el cual $a_m \neq b_m$, es decir, suponga que

$$a_1 = b_1, \; a_2 = b_2, \; \ldots, \; a_{m-1} = b_{m-1}, \quad \text{pero} \quad a_m \neq b_m$$

Sin perder generalidad podemos asumir que $a_m < b_m$. Como $a_m, b_m \in \{0,1,2\}$, resulta que

$$1 \leq a_m + 1 \leq b_m \leq 2.$$

De aquí se sigue que

$$\begin{aligned} x &= \sum_{n=1}^{m-1} \frac{a_n}{3^n} + \frac{a_m}{3^m} + \sum_{n=m+1}^{\infty} \frac{a_n}{3^n} \\ &\leq \sum_{n=1}^{m-1} \frac{a_n}{3^n} + \frac{a_m}{3^m} + \sum_{n=m+1}^{\infty} \frac{2}{3^n} \\ &= \sum_{n=1}^{m-1} \frac{a_n}{3^n} + \frac{a_m}{3^m} + \frac{1}{3^m} \\ &= \sum_{n=1}^{m-1} \frac{a_n}{3^n} + \frac{a_m+1}{3^m} \\ &\leq \sum_{n=1}^{m-1} \frac{b_n}{3^n} + \frac{b_m}{3^m} \\ &\leq \sum_{n=1}^{m-1} \frac{b_n}{3^n} + \frac{b_m}{3^m} + \sum_{n=m+1}^{\infty} \frac{b_n}{3^n} \\ &= x \end{aligned}$$

de donde se concluye que todas las desigualdades anteriores se convierten en igualdades y, en consecuencia,

$$a_m + 1 = b_m \qquad \text{y} \qquad a_n = 2, \quad b_n = 0 \text{ para todo } n \geq m+1.$$

Estas relaciones muestran que x no puede tener otra representación ternaria. ■

Ejemplo 5.1.1. Recordemos que

$$\begin{aligned} \frac{1}{3} &= \frac{1}{3} + \frac{0}{3^2} + \frac{0}{3^3} + \cdots = (0{,}100\ldots)_3 \\ &= \frac{0}{3} + \frac{2}{3^2} + \frac{2}{3^3} + \cdots = (0{,}022\ldots)_3. \end{aligned}$$

También,

$$\begin{aligned} \frac{2}{3} &= \frac{2}{3} + \frac{0}{3^2} + \frac{0}{3^3} + \cdots = (0{,}200\ldots)_3 \\ &= \frac{1}{3} + \frac{2}{3^2} + \frac{2}{3^3} + \cdots = (0{,}122\ldots)_3 \end{aligned}$$

Lo interesante de las dos representaciones ternarias tanto de $1/3$ así como de $2/3$ es que una de ellas usa el dígito 1, mientras que la otra no lo usa en ningún lugar. En general, todo x con representación ternaria $x = (0, a_1 a_2 \ldots a_{m-1}1000\ldots)_3$ se puede escribir como $x = (0, a_1 a_2 \ldots a_{m-1}0222\ldots)_3$. En efecto, por (1^3_m) se tiene que

$$\begin{aligned}(0, a_1 a_2 \ldots a_{m-1}1000\ldots)_3 &= \sum_{n=1}^{m-1} \frac{a_n}{3^n} + \frac{1}{3^m} \\ &= \sum_{n=1}^{m-1} \frac{a_n}{3^n} + \frac{0}{3^m} + \sum_{n=m+1}^{\infty} \frac{2}{3^n} \\ &= (0, a_1 a_2 \ldots a_{m-1}0222\ldots)_3\end{aligned}$$

Similarmente,

$$\begin{aligned}(0, a_1 a_2 \ldots a_{m-1}1222\ldots)_3 &= \sum_{n=1}^{m-1} \frac{a_n}{3^n} + \frac{1}{3^m} + \sum_{n=m+1}^{\infty} \frac{2}{3^n} \\ &= \sum_{n=1}^{m-1} \frac{a_n}{3^n} + \frac{1}{3^m} + \frac{1}{3^m} \\ &= \sum_{n=1}^{m-1} \frac{a_n}{3^n} + \frac{2}{3^m} \\ &= (0, a_1 a_2 \ldots a_{m-1}2000\ldots)_3\end{aligned}$$

5.1.2. Representaciones Binarias

Similar a la representación ternaria se define la representación binaria de un número real $x \in [0,1]$. En efecto, para cada $n \in \mathbb{N}$, considere $a_n \in \{0,1\}$. Entonces $0 \leq a_n/2^n < 1$ y se sigue del Teorema de la Serie Geométrica que la serie

$$\sum_{n=1}^{\infty} \frac{a_n}{2^n}$$

converge a un único elemento $x \in [0,1]$ al que representaremos en la forma

$$x = (0, a_1 a_2 a_3 \ldots)_2$$

y, como antes, diremos que $(0, a_1 a_2 a_3 \ldots)_2$ es una **representación binaria** de x. Los números 0 y 1 son llamados los **dígitos binarios**.

Es fácil establecer, similar a como se hizo en el caso de la representación ternaria, que todo $x \in [0,1]$ posee al menos una, pero no más de dos, representaciones binarias. Por ejemplo, la única representación binaria de 0 es

$$0 = (0, 000\ldots)_2$$

pero 1 se puede representar de dos formas distintas:

$$1 = (1,000\ldots)_2 = (0,111\ldots)_2$$

Más aun, cualquier $0 < x < 1$ que posea una representación binaria, digamos en la forma, $x = (a_1\, a_2\, \ldots a_n 1000\ldots)_2$ donde los $a_1, \ldots, a_n \in \{0,1\}$, también se puede escribir como

$$x = (0, a_1\, a_2\, \ldots\, a_n\, 0\,111\ldots)_2$$

A tal expresión la llamaremos la **representación binaria infinita** de x.

Corolario 5.1.5. *Si $0 < x < 1$, entonces x posee **una, y sólo una, representación binaria infinita**.*

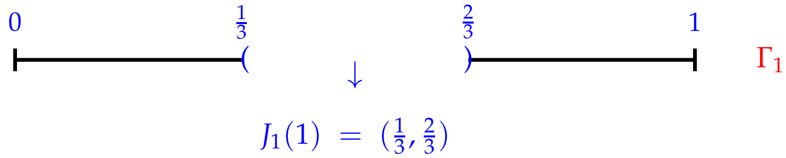
5.1.3. El Conjunto Ternario de Cantor

En el siglo XIX, tres matemáticos descubrieron, cada uno e independientemente de los otros dos, una familia de conjuntos de números reales, ahora conocidos bajo el nombre de **conjuntos tipo-Cantor**, con propiedades muy sorprendentes y que fueron utilizados fundamentalmente para construir algunos contraejemplos problemáticos. El primero de esos conjuntos fue descubierto por el inglés Henry J. S. Smith (1826-1883) en el año 1875. Este trabajo de Smith pasó casi desapercibido para la época por la sencilla razón de que muy poca atención se le prestaba a la investigación matemática proveniente de Inglaterra o de cualquier otra parte de Europa que no fuese Alemania pues el prestigio de las universidades alemanas constituían el centro del mundo matemático. De modo similar, en el año 1881, el brillante físico-matemático italiano Vito Volterra (1860-1940), publicó un resultado similar al de Smith que le sirvió para demostrar la existencia de una derivada acotada que no era Riemann integrable imponiendo, de este modo, una severa limitación al Teorema Fundamental del Cálculo para la integral de Riemann y que originó una profunda revisión de la noción de integral por parte de Lebesgue. Finalmente, en el año 1883 aparece el más conocido y famoso ejemplo de este tipo de conjuntos llamado el **conjunto ternario de Cantor** o simplemente conjunto **SVC**(3) (3 por ternario y **SVC** por **S**mith, **V**olterra y **C**antor).

El conjunto ternario de Cantor, al que denotaremos en lo sucesivo por Γ, es un subconjunto de $[0,1]$ que se construye ejecutando, en infinitos pasos, la eliminación de ciertos subintervalos abiertos en $[0,1]$. Desde su descubrimiento se ha convertido en una suerte de caja de sorpresas: aparte de poseer unas propiedades extraordinarias y, por demás, sorprendentes, lo que le confiere un estatus de privilegio y una fuente casi inagotable de contraejemplos en Analysis, Topología, Teoría de la Medida, etc. dicho conjunto es una herramienta fundamental en la Teoría de los Sistemas Dinámicos, la Teoría de Fractales, etc., El proceso de construcción de dicho conjunto, como ya mencionamos, se llevará a cabo en el intervalo $[0,1]$, aunque debemos advertir que la elección de ese intervalo se hace sólo para simplificar los cálculos por lo que la construcción de un conjunto ternario de Cantor se puede realizar en cualquier intervalo cerrado $[a,b]$ y, en consecuencia, en cualquier intervalo abierto ya que éstos contienen en su interior a un intervalo cerrado y acotado.

‖ ► **Construcción del Conjunto Ternario de Cantor**.

(1) Divida el intervalo $I = [0,1]$ en tres subintervalos de igual longitud y luego elimine el subintervalo abierto que se encuentra ubicado en el centro, es decir, elimine el intervalo $J_1(1) = (\frac{1}{3}, \frac{2}{3})$ y retenga los intervalos cerrados $[0,\ 1/3]$ y $[2/3,\ 1]$.



Definiendo

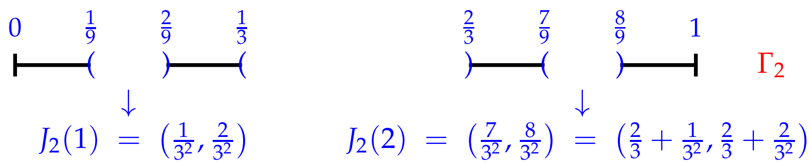
$$\Gamma_1 = [0,1] \setminus J_1 = [0,\, 1/3] \cup [2/3,\, 1],$$

donde $J_1 = J_1(1)$, resulta que Γ_1 es compacto. Más aun, si hacemos

$$F_{11} = [0,\, 1/3] \qquad \text{y} \qquad F_{12} = [2/3,\, 1],$$

se tiene que la longitud de cada uno de estos intervalos cerrados es $1/3$ y, en consecuencia, la longitud de Γ_1 es $2/3$.

(2) En la segunda etapa se subdivide cada uno de los dos intervalos cerrados retenidos anteriormente, F_{11} y F_{12}, en tres partes iguales eliminándose, como antes, los intervalos abiertos centrales $J_2(1) = (\frac{1}{9}, \frac{2}{9})$ y $J_2(2) = (\frac{7}{9}, \frac{8}{9})$, respectivamente, pero conservando los cuatro intervalos restantes.



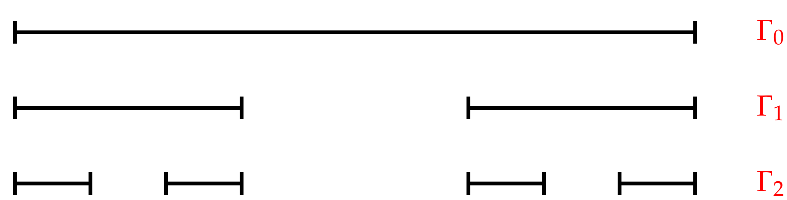
Formemos el conjunto $J_2 = J_1(1) \cup J_2(1) \cup J_2(2)$ y defina

$$\begin{aligned} \Gamma_2 &= [0,1] \setminus \big(J_1(1) \cup J_2(1) \cup J_2(2)\big) \\ &= \bigcup_{j=1}^{2^2} F_{2j}, \end{aligned}$$

donde

$$F_{21} = \left[0,\, \frac{1}{3^2}\right], \;\; F_{22} = \left[\frac{2}{3^2},\, \frac{3}{3^2}\right], \;\; F_{23} = \left[\frac{6}{3^2},\, \frac{7}{3^2}\right], \;\; F_{24} = \left[\frac{8}{3^2},\, 1\right].$$

Como antes, Γ_2 es compacto y su longitud es $4/3^2$ ya que la longitud de cada F_{2j}, $j = 1, \ldots, 2^2$, es $1/3^2$. Más aun, $\Gamma_1 \supseteq \Gamma_2$.



(3) Si se continúa de este modo indefinidamente se obtienen dos sucesiones de conjuntos, $(J_n)_{n=1}^{\infty}$ y $(\Gamma_n)_{n=1}^{\infty}$ con las siguientes propiedades: para cada entero $n \geq 1$,

  (a) J_n es la unión disjunta de $2^n - 1$ intervalos abiertos: $J_n(1), \ldots, J_n(2^n - 1)$.

(*b*) $(\Gamma_n)_{n=1}^{\infty}$ es una sucesión decreciente de conjuntos cerrados, donde cada Γ_n es la unión disjunta de 2^n intervalos cerrados:

$$\Gamma_n = F_{n1} \cup \cdots \cup F_{n2^n} = [0,1] \setminus \big(J_1 \cup J_2 \cup \cdots \cup J_n\big).$$

(*c*) Cada uno de los intervalos que definen a J_{n+1} es el centro de alguno de los intervalos que componen a Γ_n, $n = 1, 2, \ldots$.

(*d*) La longitud de cada una de las componentes tanto de J_n así como de Γ_n es $1/3^n$. Luego, la suma total de todas las longitudes de los intervalos de J_n y Γ_n son, respectivamente:

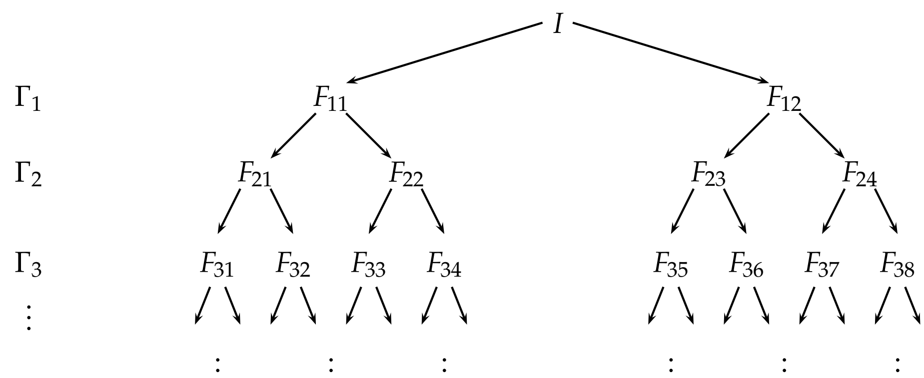
$$\ell(J_n) = 2^{n-1}\left(\frac{1}{3^n}\right) \qquad \text{y} \qquad \ell(\Gamma_n) = 2^n\left(\frac{1}{3^n}\right).$$

Finalmente, sean

$$G = \bigcup_{n=1}^{\infty} J_n \qquad \text{y} \qquad \Gamma = \bigcap_{n=1}^{\infty} \Gamma_n = [0,1] \setminus G.$$

Observe que G es un conjunto abierto no vacío y como $(\Gamma_n)_{n=1}^{\infty}$ es una sucesión decreciente de conjuntos compactos no vacíos, se sigue del Corolario 2.2.35, página 135, que Γ es también no vacío. A este conjunto Γ es al que llamaremos **conjunto ternario de Cantor**, o simplemente, **conjunto de Cantor**.

En el siguiente gráfico se muestra la representación del conjunto ternario de Cantor a través de un árbol binario.



Antes de comenzar a describir algunas de las formidables propiedades que posee el conjunto ternario de Cantor, es imprescindible que podamos entender cómo se expresan las representaciones ternarias de los extremos de cada uno de los intervalos cerrados que definen a Γ_n.

La siguiente lista da una idea precisa de tal proceso. Para cada $n \in \mathbb{N}$, denotemos por $\text{ext}(\Gamma_n)$ los puntos extremos de los intervalos que definen a Γ_n. Así,

$$\text{ext}(\Gamma_1) = \left\{0, \frac{1}{3}, \frac{2}{3}, \frac{3}{3}\right\} \cup \left\{\frac{6}{3^2}, \frac{7}{3^2}, \frac{8}{3^2}, \frac{9}{3^2}\right\} = \bigcup_{k=0}^{3}\left\{\frac{k}{3}\right\}$$

$$\text{ext}(\Gamma_2) = \left\{0, \frac{1}{3^2}, \frac{2}{3^2}, \frac{3}{3^2}\right\} \cup \left\{\frac{6}{3^2}, \frac{7}{3^2}, \frac{8}{3^2}, \frac{9}{3^2}\right\} = \bigcup_{k=0}^{3}\bigcup_{i_1=0}^{1}\left\{\frac{k+3(2i_1)}{3^2}\right\}$$

$$\begin{aligned}
\text{ext}(\Gamma_3) &= \left\{0, \frac{1}{3^3}, \frac{2}{3^3}, \frac{3}{3^3}\right\} \cup \left\{\frac{6}{3^3}, \frac{7}{3^3}, \frac{8}{3^3}, \frac{9}{3^3}\right\} \cup \left\{\frac{18}{3^3}, \frac{19}{3^3}, \frac{20}{3^3}, \frac{21}{3^3}\right\} \cup \left\{\frac{24}{3^3}, \frac{25}{3^3}, \frac{26}{3^3}, \frac{27}{3^3}\right\} \\
&= \bigcup_{k=0}^{3} \bigcup_{i_1=0}^{1} \bigcup_{i_2=0}^{1} \left\{\frac{k+3(2i_1)+3^2(2i_2)}{3^3}\right\} \\
&\ \vdots \\
\text{ext}(\Gamma_n) &= \bigcup_{k=0}^{3} \bigcup_{i_1=0}^{1} \bigcup_{i_2=0}^{1} \cdots \bigcup_{i_{n-1}=0}^{1} \left\{\frac{k+3(2i_1)+3^2(2i_2)+\cdots+3^{n-1}(2i_{n-1})}{3^n}\right\}
\end{aligned}$$

Observe que $\text{ext}\,(\Gamma_1) \subseteq \text{ext}\,(\Gamma_2) \subseteq \cdots \subseteq \text{ext}\,(\Gamma_n) \subseteq \cdots$ Definamos ahora

$$\text{ext}(\Gamma) = \bigcup_{n=1}^{\infty} \text{ext}\,(\Gamma_n) \subseteq \Gamma$$

y nótese que si x es cualquier elemento en $[0,1]$ con representación ternaria finita, digamos $x = \sum_{k=1}^{n} c_k 3^{-k}$, entonces

$$x = \sum_{k=1}^{n} \frac{c_k}{3^k} \in \text{ext}(\Gamma) \quad \text{siempre que} \quad c_1, \ldots, c_n \in \{0,2\}. \tag{5.1.1}$$

Más adelante veremos que $\text{ext}(\Gamma)$ es, de hecho, *denso* en Γ. También observe que los pares centrales en cada uno de los conjuntos que definen a $\text{ext}(\Gamma_n)$, dibujados en azul, constituyen los extremos de los intervalos abiertos $J_n(k)$ eliminados en el n-ésimo paso. Dicha fórmula permite describir tales extremos como

$$\text{ext}(J_n) = \bigcup_{k=1}^{2} \bigcup_{i_1=0}^{1} \bigcup_{i_2=0}^{1} \cdots \bigcup_{i_{n-1}=0}^{1} \left\{\frac{k+3(2i_1)+3^2(2i_2)+\cdots+3^{n-1}(2i_{n-1})}{3^n}\right\},$$

donde $J_n = \bigcup_{k=1}^{2^{n-1}} J_n(k)$. De modo explícito

$$J_1(1) = \left(\frac{1}{3}, \frac{2}{3}\right)$$

$$J_2(1) = \left(\frac{1}{3^2}, \frac{2}{3^2}\right), \qquad J_2(2) = \left(\frac{2}{3}+\frac{1}{3^2}, \frac{2}{3}+\frac{2}{3^2}\right)$$

$$J_3(1) = \left(\frac{1}{3^3}, \frac{2}{3^3}\right), \qquad J_3(2) = \left(\frac{2}{3^2}+\frac{1}{3^3}, \frac{2}{3^2}+\frac{2}{3^3}\right)$$

$$J_3(3) = \left(\frac{2}{3}+\frac{1}{3^3}, \frac{2}{3}+\frac{2}{3^3}\right), \qquad J_3(4) = \left(\frac{2}{3}+\frac{2}{3^2}+\frac{1}{3^3}, \frac{2}{3}+\frac{2}{3^2}+\frac{2}{3^3}\right)$$

$$J_4(1) = \left(\frac{1}{3^4}, \frac{2}{3^4}\right), \qquad J_4(2) = \left(\frac{2}{3^3} + \frac{1}{3^4}, \frac{2}{3^3} + \frac{2}{3^4}\right)$$

$$J_4(3) = \left(\frac{2}{3^2} + \frac{1}{3^4}, \frac{2}{3^2} + \frac{2}{3^4}\right), \qquad J_4(4) = \left(\frac{2}{3^2} + \frac{2}{3^3} + \frac{1}{3^4}, \frac{2}{3^2} + \frac{2}{3^3} + \frac{2}{3^4}\right)$$

$$J_4(5) = \left(\frac{2}{3} + \frac{1}{3^4}, \frac{2}{3} + \frac{2}{3^4}\right), \qquad J_4(6) = \left(\frac{2}{3} + \frac{2}{3^3} + \frac{1}{3^4}, \frac{2}{3} + \frac{2}{3^3} + \frac{2}{3^4}\right)$$

$$\vdots \qquad\qquad \vdots$$

Pongamos

$$\mathrm{ext}\Big(\bigcup_{n=1}^{\infty} J_n\Big) \;=\; \bigcup_{n=1}^{\infty} \mathrm{ext}(J_n) \;\subseteq\; \Gamma.$$

y note que $\mathrm{ext}(\Gamma) = \mathrm{ext}(\bigcup_{n=1}^{\infty} J_n) \cup \{0,1\}$. Lo anterior se puede resumir en la forma siguiente:

($\mathfrak{P}_0$) ‖ *Si para cada* $n \in \mathbb{N}$ *y cada* $1 \leq k \leq 2^{n-1}$ *indicamos con* a_k^n *y* b_k^n *los extremos del intervalo abierto* $J_n(k)$ *borrado en el n-ésimo paso, entonces*

$$a_1^n \;=\; \frac{1}{3^n} \qquad\qquad \text{y} \qquad\qquad b_1^n \;=\; \frac{2}{3^n}$$

y para $k \geq 2$, *existen* $1 \leq n_1 < n_2 < \cdots < n_k < n$ *tales que*

$$a_k^n \;=\; \left(\frac{2}{3^{n_1}} + \frac{2}{3^{n_2}} + \cdots + \frac{2}{3^{n_k}}\right) + \frac{1}{3^n}, \tag{5.1.2}$$

$$b_k^n \;=\; \left(\frac{2}{3^{n_1}} + \frac{2}{3^{n_2}} + \cdots + \frac{2}{3^{n_k}}\right) + \frac{2}{3^n}. \tag{5.1.3}$$

De aquí se concluye que las representaciones ternarias de a_k^n y b_k^n vienen dadas, respectivamente, por

$$\begin{aligned} a_k^n \;&=\; (0, c_1c_2\cdots c_{n-1}100\ldots)_3 \;=\; (0, c_1c_2\cdots c_{n-1}022\ldots)_3, \\ b_k^n \;&=\; (0, c_1c_2\cdots c_{n-1}200\ldots)_3 \end{aligned}$$

donde $c_1, \ldots, c_{n-1} \in \{0,2\}$.

5.1.4. Propiedades del Conjunto Ternario de Cantor

Las siguientes propiedades del conjunto ternario de Cantor, sólo mostramos algunas, son las que le confieren a dicho conjunto los calificativos de extraordinario, sorprendente, especial y, por supuesto, un estatus de privilegio entre los muchos subconjuntos de $\mathbb{R}$ que también poseen propiedades muy especiales.

($\mathcal{K}_1$) Γ ***es compacto***.

Prueba. Observe que cada conjunto Γ_n es compacto por ser cerrado y acotado, o si prefieres, por ser una unión finita de intervalos compactos. Más aun,

$$\Gamma_1 \;\supseteq\; \Gamma_2 \;\supseteq\; \cdots \;\supseteq\; \Gamma_n \;\supseteq\; \cdots$$

Se sigue entonces del Corolario 2.2.35 que $\Gamma = \bigcap_{n=1}^{\infty} \Gamma_n$ es compacto y no vacío. ■

De hecho, todos los puntos extremos de los intervalos que definen a Γ_n para cada $n \in \mathbb{N}$, pertenecen a Γ. Por ejemplo, todas las fracciones

$$0,\ 1,\ \frac{1}{3},\ \frac{2}{3},\ \frac{1}{9},\ \frac{2}{9},\ \frac{7}{9},\ \frac{8}{9},\ \frac{1}{27},\ \frac{2}{27},\ \frac{7}{27},\ \frac{8}{27},\ \frac{19}{27},\ \frac{20}{27},\ \frac{25}{27},\ \frac{26}{27},\ \frac{1}{81},\ \frac{2}{81},\ \cdots$$

están en Γ.

$(\mathcal{K}_2)$ $\mu(\Gamma) = 0$, donde μ es la medida de Lebesgue sobre $\mathbb{R}$. Véase el Corolario 6.3.41, página 283. En particular, Γ *no contiene ningún intervalo*.

Tal vez podamos convencer al lector de que la "longitud" de Γ es cero argumentando del modo siguiente: Recordemos que en el primer paso, en la construcción de Γ, removimos el intervalo abierto $J_1(1)$ de longitud $1/3$. En el segundo paso eliminamos dos intervalos abiertos $J_2(1)$ y $J_2(2)$, cada uno de longitud $1/3^2$, de modo que en estas dos etapas hemos quitado de $[0,1]$ una longitud total de $\frac{1}{3} + 2\frac{1}{3^2}$. En el n-ésimo paso habíamos eliminado 2^{n-1} intervalos $J_n(1), \ldots, J_n(2^{n-1})$ cada uno de ellos de longitud $1/3^n$, por lo que la suma de todas las longitudes eliminadas hasta el paso n es:

$$S_n \;=\; \frac{1}{3} + 2\frac{1}{3^2} + 2^2\frac{1}{3^3} + \cdots + 2^{n-1}\frac{1}{3^n}.$$

Una vez construido Γ, se observa que la totalidad de las longitudes de los intervalos eliminados de $[0,1]$ es la siguiente:

$$\sum_{n=1}^{\infty} \ell(J_n) \;=\; \sum_{n=1}^{\infty} 2^{n-1}\frac{1}{3^n} \;=\; \frac{1}{2}\sum_{n=1}^{\infty}\left(\frac{2}{3}\right)^n \;=\; \frac{1}{2}\cdot\frac{2/3}{1-2/3} \;=\; 1.$$

Pero 1 es la longitud de $[0,1]$, de modo que Γ debe "medir" cero. Aunque esta conclusión es cierta, resulta que cada una de estas afirmaciones necesitan una justificación. Ellas serán dadas más adelante cuando veamos la noción de medida de Lebesgue.

$(\mathcal{K}_3)$ Γ ***es totalmente disconexo***. Puesto que Γ no contiene, gracias al resultado anterior, ningún intervalo y como los únicos subconjuntos conexos de $\mathbb{R}$ son los intervalos y los puntos, Teorema 2.2.7, página 123, resulta que las componentes conexas de Γ son sus puntos. ■

$(\mathcal{K}_4)$ Γ ***es nunca-denso***. Esto es consecuencia inmediata del $(\mathcal{K}_2)$. Más aun, el conjunto $G = [0,1] \setminus \Gamma$ es ***denso*** en $[0,1]$. Para ver esto, observe que como Γ es nunca-denso, entonces $\operatorname{int}(\Gamma) = \varnothing$ y se sigue del Teorema 2.2.39, página 137, que

$$[0,1] \;=\; [0,1] \setminus \operatorname{int}(\Gamma) \;=\; \overline{[0,1] \setminus \Gamma} \;=\; \overline{G}.$$

La siguiente importantísima descripción de los elementos de Γ es la que permite determinar con exactitud si un elemento dado en $[0,1]$ pertenece o no a dicho conjunto.

$(\mathcal{K}_5)$ ***Cada*** $x \in \Gamma$ ***admite una única representación ternaria sin usar el dígito*** 1, es decir,

$$\Gamma \;=\; \left\{ x \in [0,1] \;:\; x = \sum_{n=1}^{\infty} \frac{c_n}{3^n},\ c_n \in \{0,2\} \text{ para todo } n \geq 1 \right\}.$$

Prueba. Pongamos

$$\Gamma^* \;=\; \left\{x \in [0,1] \,:\, x = \sum_{n=1}^{\infty} \frac{c_n}{3^n},\;\; c_n \in \{0,2\} \text{ para todo } n \geq 1\right\}.$$

Sea $x = \sum_{n=1}^{\infty} c_n 3^{-n} \in \Gamma^*$, donde $c_n \in \{0,2\}$ para todo $n \geq 1$. Fijemos un $n \in \mathbb{N}$ y considere la suma parcial $s_n = \sum_{k=1}^{n} \frac{c_k}{3^k}$. Se sigue entonces de (5.1.1) que $s_n \in \Gamma$ para todo $n \geq 1$ y, además,

$$\lim_{n\to\infty} s_n \;=\; \sum_{n=1}^{\infty} \frac{c_n}{3^n} \;=\; x.$$

Como Γ es un conjunto cerrado, resulta que $x \in \Gamma$. Esto prueba que $\Gamma^* \subseteq \Gamma$.

Para demostrar la otra inclusión, sea $x \in \Gamma$ y suponga que $x \notin \Gamma^*$. Ahora bien, como $x \notin \Gamma^*$, al expresar dicho número en su representación ternaria resulta que algunos de sus coeficiente debe ser igual a 1, de modo que no se pierde generalidad en suponer que x tiene la forma

$$x \;=\; \sum_{n=1}^{m-1} \frac{c_n}{3^n} + \frac{1}{3^m} + \sum_{n=m+1}^{\infty} \frac{c_n}{3^n} \tag{1}$$

donde $c_1, \ldots, c_{m-1}, c_{m+1}, c_{m+2}, \ldots \in \{0,2\}$ y m es el primer entero tal que $c_m = 1$. Observe que no todos los c_{m+j} con $j \geq 1$ pueden ser iguales a 0, pues si todos ellos fuesen cero, entonces por (1) y (1^3_m) tendríamos que

$$x \;=\; \sum_{n=1}^{m-1} \frac{c_n}{3^n} + \frac{1}{3^m} \;=\; \sum_{n=1}^{m-1} \frac{c_n}{3^n} + \sum_{n=m+1}^{\infty} \frac{2}{3^n} \in \Gamma^*$$

lo cual es absurdo. Similarmente, si se supone que $c_{m+j} = 2$ para todo $j \geq 1$ se obtiene, como antes, que $x \in \Gamma^*$. Por consiguiente, al menos un $c_{m+j} \neq 0$ y al menos un $c_{m+j} \neq 2$. De esto último se concluye que

$$0 \;<\; \sum_{n=m+1}^{\infty} \frac{c_n}{3^n} \;<\; \sum_{n=m+1}^{\infty} \frac{2}{3^n} \;=\; \frac{1}{3^m},$$

y, por lo tanto, usando (5.1.2) y (5.1.3), vemos que

$$\begin{aligned} a_j^m \;=\; \sum_{n=1}^{m-1} \frac{c_n}{3^n} + \frac{1}{3^m} \;<\; x \;&=\; \sum_{n=1}^{m-1} \frac{c_n}{3^n} + \frac{1}{3^m} + \sum_{n=m+1}^{\infty} \frac{c_n}{3^n} \\ &<\; \sum_{n=1}^{m-1} \frac{c_n}{3^n} + \frac{1}{3^m} + \frac{1}{3^m} \\ &=\; \sum_{n=1}^{m-1} \frac{c_n}{3^n} + \frac{2}{3^m} \;=\; b_j^m \end{aligned}$$

es decir, $x \in J_m(j)$. Pero como $\Gamma_m = [0,1] \setminus (J_m(1) \cup \cdots \cup J_m(2^{m-1}))$, resulta que $x \notin \Gamma_m$ y, en consecuencia, $x \notin \Gamma$. Esta contradicción muestra que $x \in \Gamma^*$ y termina la prueba. ∎

Observe que multiplicar por 3 a un número $x \in \Gamma$ equivale a desplazar el punto en la representación de ternaria de x un lugar hacia la derecha. Por ejemplo, tomando

$$\frac{1}{4} = (0{,}02020202\ldots)_3 \in \Gamma$$

resulta que

$$3 \times \frac{1}{4} = (0{,}20202020\ldots)_3 = \left(\frac{3}{4}\right)_3.$$

Mientras que dividir por 3 equivale a desplazar el punto ternario un lugar hacia la izquierda. Así, por ejemplo,

$$\frac{1}{3} \times \frac{1}{4} = (0{,}00202020\ldots)_3 = \left(\frac{1}{12}\right)_3.$$

Una vez establecida la propiedad $(\mathcal{K}_5)$ se comienzan a descubrir otros atributos importantes e interesantes de Γ. Por ejemplo, una de las propiedades más sorprendente, extraordinaria y, por demás extraña de Γ, es la siguiente:

$(\mathcal{K}_6)$ Γ ***es no-numerable***.

Prueba. Ya hemos visto, Teorema 1.2.17, que

$$\operatorname{card}\big(\{0,1\}^{\mathbb{N}}\big) = \operatorname{card}(\mathscr{P}(\mathbb{N})),$$

y, como $\operatorname{card}(\mathscr{P}(\mathbb{N})) = \mathfrak{c}$, resulta entonces que $\operatorname{card}\big(\{0,1\}^{\mathbb{N}}\big) = \mathfrak{c}$. Para finalizar la prueba, defina

$$\varphi : \{0,1\}^{\mathbb{N}} \to \Gamma \qquad \text{por} \qquad \varphi\big((a_n)_{n=1}^{\infty}\big) = \sum_{n=1}^{\infty} \frac{2a_n}{3^n}$$

Puesto que φ es claramente biyectiva, entonces Γ es no numerable. ■

Otra forma de demostrar lo anterior es a través del Método de la Diagonal de Cantor. Suponga que Γ es numerable y sea $\{x_1, x_2, \ldots\}$ una enumeración de dicho conjunto. Escribamos cada uno de los x_i en su representación ternaria sin usar el dígito 1, esto es:

$$\begin{aligned}
x_1 &= (0, a_{11}a_{12}a_{13}\cdots)_3,\\
x_2 &= (0, a_{21}a_{22}a_{23}\cdots)_3,\\
x_3 &= (0, a_{31}a_{32}a_{33}\cdots)_3,\\
&\ \ \vdots
\end{aligned}$$

donde los $a_{mn} \in \{0,2\}$ para todo $m, n \in \mathbb{N}$. Ahora defina

$$a_n = \begin{cases} 0 & \text{si } a_{nn} = 2 \\ 2 & \text{si } a_{nn} = 0. \end{cases}$$

Resulta que el número $x = (0, a_1a_2a_3\ldots)$ pertenece a Γ, pero no está en la lista anterior. De allí que Γ es, necesariamente, no-numerable. ■

$(\mathcal{K}_7)$ Γ ***es simétrico*** *en el siguiente sentido*:

$$\Gamma \;=\; 1-\Gamma.$$

Prueba. Por $(\mathcal{K}_5)$, cada $x \in \Gamma$ se puede representar en la forma $x = (0, c_1c_2c_3\ldots)_3$, donde $c_n \in \{0,2\}$ para todo $n \geq 1$. Por esto,

$$\begin{aligned} 1-x \;&=\; (0{,}222\ldots)_3 \;-\; (0.c_1c_2c_3\ldots)_3 \\ &=\; \sum_{n=1}^{\infty}\frac{2}{3^n} - \sum_{n=1}^{\infty}\frac{c_n}{3^n} \\ &=\; \sum_{n=1}^{\infty}\frac{2-c_n}{3^n} \\ &=\; \sum_{n=1}^{\infty}\frac{a_n}{3^n} \;\in\; \Gamma \end{aligned}$$

donde $a_n = 2 - c_n \in \{0,2\}$ para todo $n \geq 1$. ■

Observe también que

$$\frac{1}{3^m}\Gamma \subseteq \Gamma$$

se cumple para cualquier $m \geq 1$. En efecto, si $x = \sum_{n=1}^{\infty}\frac{c_n}{3^n} \in \Gamma$, entonces se tiene que

$$\frac{1}{3^m}x \;=\; \frac{1}{3^m}\sum_{n=1}^{\infty}\frac{c_n}{3^n} \;=\; \sum_{n=1}^{\infty}\frac{c_n}{3^{n+m}} \;\in\; \Gamma$$

para cualquier $m \in \mathbb{N}$.

Otros resultados que que se obtienen gracias a la representación de los elementos de Γ dada por $(\mathcal{K}_5)$, consiste en identificar una cantidad infinita de racionales que no son de la forma $\frac{p}{3^n}$ pero que aun pertenecen a Γ. Como Γ es no-numerable, en él se pueden apreciar dos grandes categorías de puntos: los *visibles* y los *ocultos*. Los visibles, como ya hemos visto, son los extremos de los intervalos retenidos en cada paso de su construcción, es decir,

$$0,\ 1,\ \frac{1}{3},\ \frac{2}{3},\ \frac{1}{9},\ \frac{2}{9},\ \frac{7}{9},\ \frac{8}{9},\ \frac{1}{27},\ \frac{2}{27},\ \frac{7}{27},\ \frac{8}{27},\ \frac{19}{27},\ \frac{20}{27},\ \frac{25}{27},\ \frac{26}{27},\ \frac{1}{81},\ \frac{2}{81},\ \ldots$$

Los ocultos, como su nombre lo indica, no están a la vista y, por consiguiente, no son fáciles de detectar. Entre ellos están, por supuesto, los números irracionales de $[0,1]$ que pertenecen a Γ. Identificar tales números irracionales es, como cabe esperar, una tarea harto difícil. Sin embargo, el número de Liouville

$$\sum_{n=1}^{\infty}\frac{2}{3^{n!}} \;=\; (0, 220002000000000000000000200\ldots)_3,$$

su simétrico $1 - \sum_{n=1}^{\infty}\frac{2}{3^{n!}}$ y los que se obtienen multiplicando los dos anteriores por $\frac{1}{3^m}$ para cada $m \in \mathbb{N}$, son de los "pocos" números irracionales conocidos que viven en Γ. Sin embargo, además de los números irracionales, existen ciertas colecciones de fracciones que están ocultas en Γ, entre ellas las que tienen la forma $\frac{1}{3^n+1}$, para todo $n \in \mathbb{N}$. En efecto,

$$
\begin{aligned}
(0,\underbrace{0\ldots0}_{n}\underbrace{2\ldots2}_{n}\underbrace{0\ldots0}_{n}\underbrace{2\ldots2}_{n}0\ldots)_3 &= \\
&= \left(\frac{2}{3^{n+1}} + \frac{2}{3^{n+2}} + \cdots + \frac{2}{3^{2n}}\right) + \left(\frac{2}{3^{3n+1}} + \frac{2}{3^{3n+2}} + \cdots + \frac{2}{3^{4n}}\right) + \cdots \\
&= \left(\frac{2}{3^{n+1}} + \frac{2}{3^{n+2}} + \cdots + \frac{2}{3^{2n}}\right)\left(1 + \frac{1}{3^{2n}} + \frac{1}{3^{4n}} + \frac{1}{3^{6n}} + \cdots\right) \\
&= \frac{3^n-1}{3^{2n}} \cdot \frac{1}{1-\dfrac{1}{3^{2n}}} = \frac{3^n-1}{3^{2n}-1} = \frac{3^n-1}{(3^n+1)(3^n-1)} \\
&= \frac{1}{3^n+1} \in \Gamma.
\end{aligned}
$$

Además, como Γ es simétrico, los siguientes números también forman parte de Γ:

$$1 - \frac{1}{3^n+1} = \frac{3^n}{3^n+1} \quad \text{para todo } n \in \mathbb{N}.$$

y, por supuesto, también todas las fracciones del tipo

$$\frac{1}{3^m}\frac{3^n}{3^n+1} \quad \text{para todo } m, n \in \mathbb{N}.$$

En particular, para $n = 1$ uno obtiene que $\frac{1}{4} = \frac{1}{3+1} \in \Gamma$ y su numerosa familia:

$$\frac{1}{4}, \frac{3}{4}, \frac{1}{3\cdot 4}, \frac{11}{3\cdot 4}, \frac{1}{3^2\cdot 4}, \frac{11}{3^2\cdot 4}, \frac{25}{3^2\cdot 4}, \frac{35}{3^2\cdot 4}, \frac{1}{3^3\cdot 4},$$
$$\frac{11}{3^3\cdot 4}, \frac{25}{3^3\cdot 4}, \frac{35}{3^3\cdot 4}, \frac{73}{3^3\cdot 4}, \frac{83}{3^3\cdot 4}, \frac{97}{3^3\cdot 4}, \frac{107}{3^3\cdot 4}, \cdots$$

todos están en Γ. Pero no sólo los números de la forma $\frac{1}{3^n+1}$ y sus simétricos están en Γ, existen muchos otros racionales ocultos en Γ que no son fáciles de visualizar, como por ejemplo, todas las fracciones del tipo $\frac{2}{3^n-1}$ están en Γ ya que

$$
\begin{aligned}
(0,\underbrace{0\ldots02}_{n}\underbrace{0\ldots02}_{n}0\cdots)_3 &= \frac{2}{3^n} + \frac{2}{3^{2n}} + \frac{2}{3^{3n}} + \cdots \\
&= \frac{2}{3^n}\left[1 + \frac{1}{3^n} + \left(\frac{1}{3^n}\right)^2 + \left(\frac{1}{3^n}\right)^3 + \cdots\right] \\
&= \frac{2}{3^n} \cdot \frac{1}{1-\dfrac{1}{3^n}} = \frac{2}{3^n-1} \in \Gamma,
\end{aligned}
$$

y por simetría, $1 - \frac{2}{3^n-1} \in \Gamma$ para todo $n \in \mathbb{N}$. Por supuesto, los múltiplos $\frac{1}{3^m} \cdot \frac{2}{3^n-1}$ y $\frac{1}{3^m}(1 - \frac{2}{3^n-1})$ están en Γ para todo $m, n \in \mathbb{N}$.

($\mathcal{K}_8$) Γ ***es perfecto***, *es decir, es cerrado y no posee puntos aislados*.

Ya hemos visto que Γ es cerrado. Para demostrar que él no posee puntos aislados es suficiente comprobar que el conjunto

$$\operatorname{ext}(\Gamma) \;=\; \bigcup_{n=1}^{\infty} \operatorname{ext}(\Gamma_n) \;\subseteq\; \Gamma$$

es denso en Γ. En efecto, sea $x = \sum_{n=1}^{\infty} c_n 3^{-n} \in \Gamma$. Dado $\varepsilon > 0$, escojamos un $N \in \mathbb{N}$ tal que $3^{-N} < \varepsilon$. Si ahora definimos

$$s_N \;=\; \sum_{n=1}^{N} \frac{c_n}{3^n},$$

resulta, por (5.1.1), que $s_N \in \operatorname{ext}(\Gamma)$ y así,

$$|x - s_N| \;=\; \sum_{n=N+1}^{\infty} \frac{c_n}{3^n} \;\leq\; \sum_{n=N+1}^{\infty} \frac{2}{3^n} \;=\; \frac{1}{3^N} \;<\; \varepsilon.$$

Esto termina la prueba. ■

Observe que de lo anterior también se obtiene, como caso particular, que el conjunto $G = [0,1] \setminus \Gamma$ es denso en Γ.

($\mathcal{K}_9$) $\Gamma + \Gamma \;=\; [0,2]$ y $\Gamma - \Gamma \;=\; [-1,1]$.

Prueba. (a) Veamos que $\Gamma + \Gamma = [0,2]$. Observe, en primer lugar, que

$$\frac{1}{2}\Gamma \;=\; \left\{ \sum_{n=1}^{\infty} \frac{a_n}{3^n} \;:\; a_n \in \{0,1\} \text{ para todo } \; n \geq 1 \right\}.$$

De esto se concluye que

$$\begin{aligned} \frac{1}{2}\Gamma + \frac{1}{2}\Gamma \;&=\; \left\{ \sum_{n=1}^{\infty} \frac{a_n + b_n}{3^n} \;:\; a_n, b_n \in \{0,1\} \text{ para todo } \; n \geq 1 \right\} \\ &=\; \left\{ \sum_{n=1}^{\infty} \frac{d_n}{3^n} \;:\; d_n \in \{0,1,2\} \text{ para todo } \; n \geq 1 \right\} \\ &=\; [0,1] \end{aligned}$$

y, por lo tanto, $\Gamma + \Gamma = [0,2]$. ■

(b) Veamos ahora que $\Gamma - \Gamma = [-1,1]$.

En efecto, es claro que $\Gamma - \Gamma \subseteq [-1,1]$. Para demostrar la otra inclusión necesitamos definir, para cada $n \in \mathbb{N}$, el conjunto

$$R_n \;=\; \left\{ x \in [0,1] \;:\; x = \sum_{k=1}^{n} \frac{a_k}{3^k}, \;\; a_k \in \{0,1,2\} \right\}.$$

Nuestra primera tarea será demostrar que

$$R_n \subseteq \Gamma - \Gamma.$$

para todo $n \geq 1$. Usaremos, para ello, inducción sobre n. Observe que si $x \in \Gamma$, entonces $x = x - 0 \in \Gamma - \Gamma$, por lo que sólo es necesario analizar los elementos de R_n que no forman parte de Γ. Para $n = 1$, se cumple

$$R_1 = \left\{0, \frac{1}{3}, \frac{2}{3}\right\} \subseteq \Gamma$$

y, por lo tanto, $R_1 \subseteq \Gamma - \Gamma$. Para $n = 2$, nuestro conjunto R_2 contiene nueve elementos, a saber,

$$R_2 = \left\{0, \frac{1}{3}, \frac{2}{3}, \frac{1}{3^2}, \frac{2}{3^2}, \frac{1}{3} + \frac{1}{3^2}, \frac{1}{3} + \frac{2}{3^2}, \frac{2}{3} + \frac{1}{3^2}, \frac{2}{3} + \frac{2}{3^2}\right\}$$

Nótese que, gracias a la definición de $\text{ext}(\Gamma)$, todas las fracciones de R_2 pertenecen a Γ salvo $\frac{1}{3} + \frac{1}{3^2}$ y $\frac{1}{3} + \frac{2}{3^2}$ las cuales se pueden expresar en la forma

$$\frac{1}{3} + \frac{1}{3^2} = (0,11000\ldots)_3 = (0,2000\ldots)_3 - (0,0200\ldots)_3 \in \Gamma - \Gamma$$

$$\frac{1}{3} + \frac{2}{3^2} = (0,12000\ldots)_3 = (0,2200\ldots)_3 - (0,0222\ldots)_3 \in \Gamma - \Gamma$$

Por esto, $R_2 \subseteq \Gamma - \Gamma$.

En general, suponga que para $k = n - 1$ se cumple que $R_{n-1} \subseteq \Gamma - \Gamma$ y sea $x = (0,a_1 \ldots a_n)_3$ cualquier elemento de R_n con $a_n \neq 0$. Por lo observado anteriormente, supondremos que $x \notin \Gamma$. Como

$$x = (0,a_1 \ldots a_{n-1})_3 + \frac{a_n}{3^n}$$

resulta de nuestra hipótesis que existen r_{n-1} y q_{n-1} en Γ tal que

$$(0,a_1 \ldots a_{n-1})_3 = r_{n-1} - q_{n-1}$$

Pongamos

$$r_{n-1} = (0,b_1 \ldots b_{n-1} \ldots)_3 \qquad \text{y} \qquad q_{n-1} = (0,c_1 \ldots c_{n-1} \ldots)_3,$$

donde los dígitos $b_1, c_1, b_2, c_2, \ldots \in \{0, 2\}$. Es importante destacar que, en la representación anterior, r_{n-1} siempre se puede elegir en la forma

$$(0,b_1 \ldots b_{n-1}000\ldots)_3.$$

Por esto, $(0,a_1 \ldots a_{n-1})_3$ se puede escribir como

(1) $(0,a_1 \ldots a_{n-1})_3 = (0,b_1 \ldots b_{n-1}000\ldots)_3 - (0,c_1 \ldots c_{n-1}0000\ldots)_3$ con $c_{n-1} \neq 0$, aunque b_{n-1} puede ser 0, o

(2) $(0,a_1 \ldots a_{n-1})_3 = (0,b_1 \ldots b_{n-1}000\ldots)_3 - (0,c_1 \ldots c_{n-1}0222\ldots)_3$.

Si $a_n = 2$, simplemente hacemos $b_n = 2$ y entonces se tiene que

$$(0, a_1 \ldots a_{n-1}2)_3 \;=\; (0, b_1 \ldots b_{n-1}200\ldots)_3 \;-\; (0, c_1 \ldots c_{n-1}0000\ldots)_3$$

pertenece a $\Gamma - \Gamma$. Por otro lado, si $a_n = 1$, podemos escribir

$$(0, a_1 \ldots a_{n-1}1)_3 \;=\; (0, b_1 \ldots b_{n-1}2000\ldots)_3 \;-\; (0, c_1 \ldots c_{n-1}0222\ldots)_3$$

o

$$(0, a_1 \ldots a_{n-1}1)_3 \;=\; (0, b_1 \ldots b_{n-1}000\ldots)_3 \;-\; (0, c_1 \ldots c_{n-1}c_n000\ldots)_3.$$

Esto completa la prueba del paso inductivo.

Para terminar la demostración, sea $x \in [0,1]$ y suponga que su representación ternaria no es finita, es decir, $x \notin R_n$ para todo $n \geq 1$. Pongamos, $x = (0, a_1 a_2 \ldots)_3$ y observe que

$$x \;=\; \lim_{n\to\infty} p_n,$$

donde $p_n = (0, a_1 a_2 \ldots a_n)_3 \in R_n$ para todo $n \geq 1$. De la discusión anterior sabemos que, por cada $n \in \mathbb{N}$, existen r_n y q_n en Γ tales que $p_n = r_n - q_n$. Ahora bien, puesto que Γ es compacto, se sigue del Teorema de Bolzano-Weierstrass que existe una subsucesión $(r_{n_j})_{j=1}^{\infty}$ de $(r_n)_{n=1}^{\infty}$ convergiendo a un único punto $r \in \Gamma$. La correspondiente subsucesión $(q_{n_j})_{j=1}^{\infty}$ de $(q_n)_{n=1}^{\infty}$ también posee una subsucesión $(q_{n_{j_i}})_{i=1}^{\infty}$ que converge a un punto $q \in \Gamma$. Puesto que toda subsucesión de una sucesión convergente converge al mismo punto, se tiene que:

$$x \;=\; \lim_{i\to\infty} p_{n_{j_i}} \;=\; \lim_{i\to\infty}\left(r_{n_{j_i}} \;-\; q_{n_{j_i}}\right) \;=\; r \;-\; q \;\in\; \Gamma \;-\; \Gamma$$

Finalmente, si $x \in [-1,0]$, entonces $-x \in [0,1]$ y por lo anterior, existen $r, q \in \Gamma$ tal que $x = r - q$. Por esto, $-x = q - r \in \Gamma - \Gamma$ y, en consecuencia,

$$[-1,1] \;\subseteq\; \Gamma - \Gamma.$$

Esto termina la prueba. ■

5.1.5. Conjuntos Tipo-Cantor de Medida Cero

El siguiente es un procedimiento más general que el anterior para producir conjuntos tipo-Cantor de medida cero. Fijemos un $\alpha \in (0,1)$ y sea $\beta = \frac{1-\alpha}{2}$. Nótese que $\beta \in (0, 1/2)$ ya que

$$0 < \alpha < 1 \quad\Rightarrow\quad -1 < -\alpha < 0 \quad\Rightarrow\quad 0 < 1-\alpha < 1 \quad\Rightarrow\quad 0 < \beta < \frac{1}{2}.$$

Considere ahora las funciones $T_0, T_1 : [0,1] \to [0,1]$ definidas por

$$T_0(x) \;=\; \beta x \qquad\qquad \text{y} \qquad\qquad T_1(x) \;=\; \beta x + (1-\beta).$$

Pongamos $K_0 = [0,1]$ y para cada entero $n \geq 1$, defina

$$K_n \;=\; T_0(K_{n-1}) \cup T_1(K_{n-1}).$$

Si definimos

$$\Gamma_\alpha^\beta \;=\; \bigcap_{n=1}^{\infty} K_n,$$

entonces se obtiene un conjunto con características similares al conjunto ternario de Cantor al que llamaremos un **conjunto tipo-Cantor de medida cero**. Observe que cuando $\alpha = 1/3$, entonces $\beta = 1/3$ obteniéndose el conjunto ternario de Cantor, es decir, $\Gamma = \Gamma_{1/3}^{1/3}$.

¿Cómo son, realmente, los conjuntos K_n? Si uno mira cuidadosamente la definición anterior, se observa que

$$T_0([0,1]) \;=\; [0,\,\beta] \qquad \text{y} \qquad T_1([0,1]) \;=\; [1-\beta,\,1],$$

de modo que

$$K_1 \;=\; [0,\beta] \cup [1-\beta, 1].$$

Otra manera de describir el conjunto K_1 consiste en remover del centro de $[0,1]$ el intervalo abierto $(\beta, 1-\beta)$ de longitud $1-2\beta = \alpha$, quedando dos intervalos cerrados cada uno de longitud β. De nuevo, nótese que

$$\begin{aligned} T_0(K_1) \;&=\; \big[0,\,\beta^2\big] \cup \big[\beta(1-\beta),\,\beta\big], \\ T_1(K_1) \;&=\; \big[1-\beta,\,\beta^2+(1-\beta)\big] \cup \big[(1-\beta)+\beta(1-\beta),\,1\big] \end{aligned}$$

por lo que el conjunto K_2 se expresa en la forma:

$$K_2 \;=\; \Big(\big[0,\,\beta^2\big] \cup \big[\beta(1-\beta),\,\beta\big]\Big) \cup \Big(\big[1-\beta,\,\beta^2+(1-\beta)\big] \cup \big[(1-\beta)+\beta(1-\beta),\,1\big]\Big)$$

Observe que K_2 se obtiene de K_1 removiendo del centro de cada uno de los intervalos cerrados de K_1 un intervalo abierto de longitud $\alpha\beta$. Cada uno de los cuatro intervalos cerrados que definen a K_2 tiene longitud β^2, por lo que la longitud de K_2 es $2^2\beta^2 = (1-\alpha)^2$. En general, K_n es la unión disjunta de 2^n intervalos cerrados cada uno de longitud β^n y K_{n+1} se obtiene eliminando del centro de cada intervalo de K_n un intervalo abierto de longitud $\alpha\beta^n$. Por supuesto, la colección $(K_n)_{n=1}^{\infty}$ es una sucesión decreciente de subconjuntos compactos de $[0,1]$ que posee la propiedad de intersección finita y, en consecuencia, por la compacidad de $[0,1]$, resulta que Γ_α^β es no vacío, compacto, nunca-denso y perfecto. ¿Cuál es la longitud o medida de Γ_α^β? De nuestro análisis sabemos que la longitud de cada conjunto K_n es $2^n\beta^n$ y, en consecuencia, la longitud total de Γ_α^β es menor o igual a $2^n\beta^n$. Sin embargo, como $\beta \in (0, 1/2)$, se tiene que $2\beta < 1$ y, por lo tanto, la longitud de K_n tiende a cero cuando n tiende a infinito, lo cual quiere decir que la medida de Γ_α^β es cero (véase el Corolario 6.3.47, página 287). Otra modo de calcular la medida de Γ_α^β es observar que la suma de las longitudes de todos los intervalos que se eliminaron en su construcción es:

$$\begin{aligned} \sum_{n=1}^{\infty} \ell(J_n) \;&=\; \alpha + 2(\alpha\beta) + \cdots + 2^n(\alpha\beta^n) + \cdots \\ &=\; \alpha(1 + 2\beta + \cdots + (2\beta)^n + \cdots) \;=\; \alpha \cdot \frac{1}{1-2\beta} \;=\; 1 \end{aligned}$$

de modo que la medida de Γ_α^β es cero.

Teorema 5.1.6. *Sea* $\operatorname{ext}(\Gamma_\alpha^\beta)$ *el conjunto de los puntos extremos de todos los intervalos abiertos removidos en la construcción de* Γ_α^β. *Entonces*

$$\overline{\operatorname{ext}(\Gamma_\alpha^\beta)} \;=\; \Gamma_\alpha^\beta.$$

Prueba. Sabemos que $\mathrm{ext}(\Gamma_\alpha^\beta) \subseteq \Gamma_\alpha^\beta$ y como Γ_α^β es cerrado, entonces $\overline{\mathrm{ext}(\Gamma_\alpha^\beta)} \subseteq \Gamma_\alpha^\beta$. Para demostrar la otra inclusión, sea $x \in \Gamma_\alpha^\beta$ y seleccionemos un $\varepsilon > 0$ arbitrario. Puesto que Γ_α^β es un conjunto perfecto, x es un punto límite de Γ_α^β, lo cual significa que $(x-\varepsilon, x+\varepsilon) \cap \Gamma_\alpha^\beta \neq \varnothing$ contiene infinitos puntos. Fijemos un $y \in (x-\varepsilon, x+\varepsilon) \cap \Gamma_\alpha^\beta$ y suponga que $x < y$. Observe que gracias a que Γ_α^β no contiene intervalos (por ser nunca-denso), entonces entre el punto x y el punto y existe uno de los intervalos que fue removido en algún paso de la construcción de Γ_α^β. Denotemos a tal intervalo por (a_k, b_k). Puesto que $y \in (x-\delta, x+\delta)$, resulta que $x \in (a_k - \varepsilon, a_k + \varepsilon)$ y $x \in (b_k - \varepsilon, b_k + \varepsilon)$. Esto nos indica que x es un punto límite de a_k y b_k. Como la elección del intervalo se hizo de modo arbitrario, se sigue que $x \in \mathrm{ext}(\Gamma_\alpha^\beta)$. Esto muestra que $\Gamma_\alpha^\beta \subseteq \mathrm{ext}(\Gamma_\alpha^\beta)$ y termina la prueba. ■

Por supuesto, existen otros procedimientos distintos al anterior para generar conjuntos tipo-Cantor de medida cero. Por ejemplo, divida el intervalo $[0,1]$ en 5 partes iguales y remueva los intervalos abiertos que ocupan el 2do y 4to lugar. Repita el paso anterior con los 3 intervalos cerrados que permanecen y elimine el 2do y el 4to intervalos abiertos en cada uno de ellos. Continúe. La intersección de todos los intervalos cerrados que permanecen en cada etapa del proceso es otro conjunto tipo-Cantor que posee las mismas propiedades que el conjunto ternario de Cantor; en particular, su medida también es nula.

Nota Adicional 5.1.1 Aunque parezca extraño, los conjuntos tipo-Cantor Γ_α^β que acabamos de construir son más interesantes y útiles para los números β. La razón de por qué ello es así radica en el hecho de que $1/\beta$ es un factor de escalamiento de Γ_α^β, en otras palabras, $\beta^{-1}(\Gamma_\alpha^\beta \cap [0,\beta]) = \Gamma_\alpha^\beta$. También es importante destacar que cuando α decrece hacia cero, β crece hacia $1/2$ y, en consecuencia, los conjuntos tipo-Cantor Γ_α^β se hacen "más grandes". Esta idea, que no es tan inmediata de verificar, se puede precisar usando la noción de "dimensión fractal". Es importante observar que el hecho de Γ_α^β se hace más grande a medida que β se aproxime a $1/2$ no puede ser probado usando el argumento de cardinalidad ya que tales conjuntos poseen la misma cardinalidad, ni tampoco con un argumento de Teoría de la Medida pues todos ellos poseen medida cero.

Observe que los valores de α y β elegidos en el proceso para generar el conjunto Γ_α^β se pueden tomar de un modo más general, es decir, siempre se pueden elegir cualquier par números α, $\beta \in (0,1)$ tal que $\alpha + \beta < 1$.

Otro hecho curioso que relaciona el conjunto tipo-Cantor Γ_α^β con la Teoría de Números es el siguiente problema geométrico formulado por Roger L. Kraft en [81]: Dado $\beta \in (0, 1/2)$, ¿es posible encontrar un $\lambda \in (0,1)$ de modo que

$$\Gamma_\alpha^\beta \cap \left(\lambda \cdot \Gamma_\alpha^\beta\right) \;=\; \{0\}? \tag{K}$$

Kraft demostró que $\beta_0 = (3-\sqrt{5})/2$ es un valor crítico a la solución de dicho problema. Esto significa que si $\beta < \beta_0$, el problema (K) posee una solución, mientras que, si $\beta \geq \beta_0$, entonces dichos conjuntos son ahora demasiados grandes para proveer una solución. De hecho, Kraft demuestra que, en este caso, la cardinalidad de $\Gamma_\alpha^\beta \cap \left(\lambda \cdot \Gamma_\alpha^\beta\right)$ es infinita. Pero, ¿qué tiene de especial el valor β_0? Bueno, es aquí donde aparece sun conexión con la Teoría de Números. En primer lugar, nótese que

$$\beta_0 \;=\; \frac{3-\sqrt{5}}{2} \;=\; (\phi')^2$$

donde ϕ' es el ***número de oro*** negativo. Más aun, si α_0 es el valor asociado a β_0, entonces

$$\frac{\beta_0}{\alpha_0} = \frac{\beta_0}{1-2\beta_0} = \frac{(3-\sqrt{5})/2}{1-(3-\sqrt{5})} = \frac{1+\sqrt{5}}{2} = \phi.$$

donde ϕ es el ***número de oro*** positivo. Pero, ¿qué es el número de oro? Recordemos que el número de oro es la raíz positiva de la ecuación cuadrática $x^2 - x - 1 = 0$, cuyo valor es $\phi = (1+\sqrt{5})/2$. Ésta ecuación se obtiene, geométricamente, del modo siguiente: Un segmento de línea de longitud, digamos 1, es dividido en dos partes por un punto $x \in (0,1)$

Suponga que el segmento más largo tiene longitud x. Diremos que el segmento $[0,1]$ está divido en la **razón áurea** o **radio de oro** si la relación

$$\frac{x}{1-x} = \frac{1}{x} \tag{1}$$

se cumple, es decir, el subsegmento más largo está relacionado al más pequeño exactamente como el segmento total está relacionado al segmento más largo. Existe, por supuesto, un único número $x \in (0,1)$ que satisface la ecuación (1). Si hacemos la sustitución $\phi = 1/x$ en la igualdad (1) ella nos conduce a la ecuación cuadrática

$$\phi^2 - \phi - 1 = 0.$$

cuyas raíces son $(1+\sqrt{5})/2$ y $(1-\sqrt{5})/2$. La raíz positiva usualmente se denota por el símbolo ϕ y la negativa por ϕ'. Al número $x = \phi$ se le conoce como el ***número de oro***. Este número tiene una importancia fundamental en muchas áreas del conocimiento humano. Por ejemplo en el arte, la arquitectura, la música, la biología, la matemática, etc.

5.1.6. Conjuntos Tipo-Cantor de Medida Positiva

El conjunto ternario de Cantor fue construido removiendo, en cada etapa de su construcción, 2^{n-1} intervalos abiertos, cada uno de ellos de longitud 3^{-n}. En la sección anterior vimos como se construían conjuntos tipo-Cantor de medida cero para cualquier $\alpha \in (0,1)$. Lo que también resulta sorprendente es que se pueden construir conjuntos tipo-Cantor con medida positiva. Por supuesto, el procedimiento para lograrlo va a depender del tamaño de los intervalos abiertos que hay que eliminar en cada paso. La siguiente colección de conjuntos, que denotaremos por $(\Gamma_\alpha)_{\alpha\in(0,1)}$, similares al proceso de construcción del conjunto ternario de Cantor, poseen propiedades idénticas con una única excepción: tales conjuntos poseen *medida positiva*.

($\Vert\blacktriangleright$ Γ_α) Fijemos un $\alpha \in (0,1)$. Tal y como se procedió en la construcción del conjunto ternario de Cantor, removamos del centro de $[0,1]$ un intervalo abierto de longitud $\alpha/2$, digamos,

$$J_1(1) = \left(\frac{1}{2} - \frac{\alpha}{2^2}, \frac{1}{2} + \frac{\alpha}{2^2}\right)$$

y denotemos por $\Gamma_1(\alpha)$ la unión de los dos intervalos cerrados que permanecen, esto es,

$$\Gamma_1(\alpha) \;=\; F_{11} \cup F_{12} \;=\; \left[0,\, \frac{1}{2} - \frac{\alpha}{2^2}\right] \cup \left[\frac{1}{2} + \frac{\alpha}{2^2},\, 1\right].$$

Elimine del centro de F_{11}, así como de F_{12}, los intervalos abiertos

$$J_2(1) \;=\; \left(\frac{1}{2^2} - \frac{3\alpha}{2^4},\, \frac{1}{2^2} - \frac{\alpha}{2^4}\right) \qquad \text{y} \qquad J_2(2) \;=\; \left(\frac{3}{2^2} + \frac{\alpha}{2^4},\, \frac{3}{2^2} + \frac{3\alpha}{2^4}\right)$$

y observe que la longitud de cada uno de ellos es $\alpha/2^3$. Denote por $\Gamma_2(\alpha)$ la unión de los cuatro intervalos cerrados que quedan, vale decir, $\Gamma_2(\alpha) = F_{21} \cup F_{22} \cup F_{23} \cup F_{24}$, donde

$$F_{21} \;=\; \left[0,\, \frac{1}{2^2} - \frac{3\alpha}{2^4}\right] \qquad\qquad F_{22} \;=\; \left[\frac{1}{2^2} - \frac{\alpha}{2^4},\, \frac{1}{2} - \frac{\alpha}{2^2}\right]$$

$$F_{23} \;=\; \left[\frac{1}{2} + \frac{\alpha}{2^2},\, \frac{3}{2^2} + \frac{\alpha}{2^4}\right] \qquad\qquad F_{24} \;=\; \left[\frac{3}{2^2} + \frac{3\alpha}{2^4},\, 1\right]$$

Observe que la longitud de cada uno de estos intervalos cerrados es $\frac{1}{2^2}(1 - \frac{1}{2}\alpha - \frac{1}{2^2}\alpha)$. Repitamos, una vez más, el proceso anterior a cada uno de los intervalos cerrados que conforman a $\Gamma_2(\alpha)$ removiendo de su centro un intervalo abierto de longitud $\alpha/2^5$ y denotemos por $\Gamma_3(\alpha)$ la unión de los 2^3 intervalos cerrados que quedan. Continuando indefinidamente con este procedimiento se obtiene una sucesión decreciente de conjuntos compactos $(\Gamma_n(\alpha))_{n=1}^{\infty}$, donde cada $\Gamma_n(\alpha)$ es la unión disjunta de 2^n intervalos cerrados y acotados. Definamos finalmente

$$\Gamma_\alpha \;=\; \bigcap_{n=1}^{\infty} \Gamma_n(\alpha).$$

Por el Corolario 2.2.35 sabemos que $\Gamma_\alpha \neq \varnothing$. A Γ_α lo llamaremos un conjunto **tipo-Cantor** o conjunto **SVC**(α) de **medida positiva**. En la literatura sobre el tema, a tales conjuntos también se les llaman **conjuntos gordos de Cantor**. La propiedades del conjunto Γ_α (y sus demostraciones) son similares a la del conjunto ternario de Cantor con la excepción de que ellos poseen medida positiva, esto es,

(1) Γ_α *es* ***compacto, nunca-denso, perfecto*** *y* ***totalmente disconexo***.

(2) $\mu(\Gamma_\alpha) = 1 - \alpha > 0$. Para verificar esto, observe que como

$$[0,1] \setminus \Gamma_\alpha \;=\; \bigcup_{n=1}^{\infty} J_n^\alpha,$$

donde cada J_n^α es la unión disjunta de los 2^{n-1} intervalos abiertos borrados en la etapa n-ésima, cada uno de los cuales posee longitud $\alpha/2^{2n-1}$, se tiene entonces que:

$$\begin{aligned}\sum_{n=1}^{\infty} \ell(J_n^\alpha) \;&=\; \frac{\alpha}{2} + 2\left(\frac{\alpha}{2^3}\right) + 2^2\left(\frac{\alpha}{2^5}\right) + 2^3\left(\frac{\alpha}{2^7}\right) + \cdots + 2^{n-1}\left(\frac{\alpha}{2^{2n-1}}\right) + \cdots \\ &=\; \alpha \cdot \left(\sum_{n=1}^{\infty} \frac{1}{2^n}\right) \;=\; \alpha.\end{aligned}$$

De esto se sigue que $\mu(\Gamma_\alpha) = 1 - \alpha > 0$. De nuevo, estas afirmaciones deben ser formalmente justificadas a través de la noción de medida que introduciremos en el siguiente capítulo (véase el Ejemplo 6.4.2, página 306).

En particular, Γ_α *es no-numerable*.

Otros conjuntos tipo-Cantor con medida positiva se pueden construir argumentando del modo siguiente:

($\|\blacktriangleright$ Γ_α^0) Para cada $n \geq 3$, el conjunto $\mathbf{SVC}(n)$ se construye removiendo, en el k-ésimo paso de la iteración, un intervalo abierto de longitud $1/n^k$ del centro de cada uno de los intervalos cerrados obtenidos en el paso anterior. La intersección de todos los intervalos cerrados que permanecen en cada paso es el conjunto $\mathbf{SVC}(n)$. Estos conjuntos son compactos, perfectos, nunca-densos y de medida positiva cuando $n > 3$:

$$\begin{aligned} \mu(\mathbf{SVC}(n)) &= \mu\big([0,1]\big) - \mu\big(\mathbf{SVC}(n)^c\big) \\ &= 1 - \frac{1}{n} - \frac{2}{n^2} - \frac{2^2}{n^3} - \cdots \\ &= 1 - \frac{1}{n}\frac{1}{1-2/n} = \frac{n-3}{n-2}. \end{aligned}$$

($\|\blacktriangleright$ Γ_α^1) Fijemos, como antes, un $\alpha \in (0,1)$. En la construcción de Γ_α removimos, en cada etapa, intervalos abiertos de longitud $\alpha/2^{2n+1}$. Si en su lugar eliminamos intervalos que tengan longitud $\alpha/3^n$, entonces el conjunto Γ_α^1 que se construye posee exactamente las mismas propiedades que las obtenidas en el caso anterior. En particular, $\mu(\Gamma_\alpha^1) = 1 - \alpha$.

($\|\blacktriangleright$ Γ_α^2) Fijemos $\alpha \in (3, +\infty)$ y, como antes, comience eliminando del centro de $[0,1]$ un intervalo abierto de longitud $1/\alpha$. Luego, elimine un intervalo abierto de longitud $1/\alpha^2$ del centro de cada uno de los dos intervalos cerrados que quedaron y así sucesivamente. El conjunto tipo-Cantor Γ_α^2 que se construye intersectando la sucesión decreciente de los conjuntos compactos que se van formando en cada paso es compacto, nunca-denso, perfecto, totalmente disconexo y de medida $(\alpha-3)/(\alpha-2)$. En efecto,

$$\begin{aligned} \sum_{n=1}^{\infty} \ell(J_n^\alpha) &= \frac{1}{\alpha} + 2\frac{1}{\alpha^2} + 2^2\frac{1}{\alpha^3} + \cdots + 2^{n-1}\frac{1}{\alpha^n} + \cdots \\ &= \frac{1}{2}\sum_{n=1}^{\infty}\left(\frac{2}{\alpha}\right)^n = \frac{1}{2}\cdot\frac{2}{\alpha-2} = \frac{1}{\alpha-2}. \end{aligned}$$

Luego, $\mu(\Gamma_\alpha^2) = 1 - \frac{1}{\alpha-2} = \frac{\alpha-3}{\alpha-2}$.

($\|\blacktriangleright$ Γ_α^3) Seleccione un $\alpha \in (0,1)$ y sea $(a_n)_{n=1}^\infty$ una sucesión estrictamente creciente de números reales positivos tal que $\sum_{n=1}^\infty a_n = \alpha$. Sea J_1^1 un intervalo abierto de longitud a_1 removido del centro de $[0,1]$. De los dos intervalos cerrados restantes, remueva del centro de cada uno de ellos intervalos abiertos J_2^1, J_2^2 de longitud total a_2, es decir, $\ell(J_2^1) + \ell(J_2^2) = a_2$. Si se continúa indefinidamente con este proceso, se obtiene una colección numerable $\{J_n^{2^{k-1}} : n \in \mathbb{N}\}$ de intervalos abiertos disjuntos dos a dos tal que

$$\begin{aligned} \sum_{n=1}^{\infty} \ell(J_n) &= \underbrace{\ell(J_1^1)}_{a_1} + \underbrace{\big[\ell(J_2^1) + \ell(J_2^2)\big]}_{a_2} + \underbrace{\big[\ell(J_3^1) + \ell(J_3^2) + \ell(J_3^3) + \ell(J_3^4)\big]}_{a_3} + \cdots \\ &= a_1 + a_2 + a_3 + \cdots = \alpha, \end{aligned}$$

donde $J_n = \bigcup_{k=1}^{n} J_n^{2^{k-1}}$ para cada $n \in \mathbb{N}$. El conjunto $\Gamma_\alpha^3 = [0,1] \setminus \bigcup_{n=1}^{\infty} J_n$ es un conjunto tipo-Cantor cuya medida es $1-\alpha > 0$.

($\|\blacktriangleright$ Γ_α^4) Sea J cualquier subintervalo cerrado y acotado de $\mathbb{R}$ con $\mu(J) > 0$ y sea $\alpha \in [0, \mu(J))$. Seleccione una sucesión $(a_n)_{n=1}^{\infty}$ de números reales positivos tal que

$$\sum_{n=1}^{\infty} 2^{n-1} a_n = \mu(J) - \alpha$$

y, como en el ejemplo anterior, en el primer paso, remueva del centro de J un intervalo abierto de longitud a_1. De los dos intervalos cerrados que permanecen, remueva del centro de cada uno de ellos un intervalo abierto de longitud a_2 y continúe. Sea G la unión (disjunta) de todos los intervalos abiertos que fueron removidos. Entonces

$$\mu(G) = a_1 + 2a_2 + 2^2 a_3 + \cdots + 2^{n-1} a_n + \cdots = \mu(J) - \alpha.$$

El conjunto $\Gamma_\alpha^4 = J \setminus G$ es un conjunto tipo-Cantor de medida α.

Observe que si $J = [0,1]$ y $a_n = 3^{-n}$ para todo $n \in \mathbb{N}$, entonces Γ_α^4 con $\alpha = 0$, es el conjunto ternario de Cantor.

5.1.7. La Función de Cantor

El conjunto ternario de Cantor y sus hermanos, los conjuntos tipo-Cantor, pueden ser usados para construir funciones muy especiales y, por supuesto, tan fascinantes como los mismos conjuntos de Cantor. Tales funciones constituyen un contraejemplo a un número de situaciones diversas en la Teoría de la Medida e Integración

Existen variadas formas de definir la así llamada función de Cantor, función de Cantor-Lebesgue o escalera del Diablo. En esta sección esbozaremos cuatro modos distintos de presentar dicha función. Comencemos:

(1^{a}) **Definición**. Fijemos $n \in \mathbb{N}$ y sean

$$\Gamma_n = F_{n1} \cup \cdots \cup F_{n2^n} \qquad \text{y} \qquad J_n = J_n(1) \cup \cdots \cup J_n(2^n - 1)$$

la unión de todos los intervalos cerrados que permanecen, (respectivamente, todos los intervalos abiertos que fueron eliminados) en la n-ésima etapa en la construcción de Γ. Nótese que $J_1 \subseteq J_2 \subseteq \cdots \subseteq J_n \subseteq \cdots$ y, además,

$$J_{n+1}(2k) = J_n(k) \quad \text{para } k = 1, 2, \ldots, 2^n - 1.$$

Definición 5.1.7. *Para cada $n \in \mathbb{N}$, sea $\varphi_n : [0,1] \to \mathbb{R}$ la **función continua** definida del modo siguiente:*

(a) $\varphi_n(0) = 0$ y $\varphi_n(1) = 1$.

(b) $\varphi_n(x) = k/2^n$ para todo $x \in J_n(k)$, donde $k = 1, 2, \ldots, 2^n - 1$.

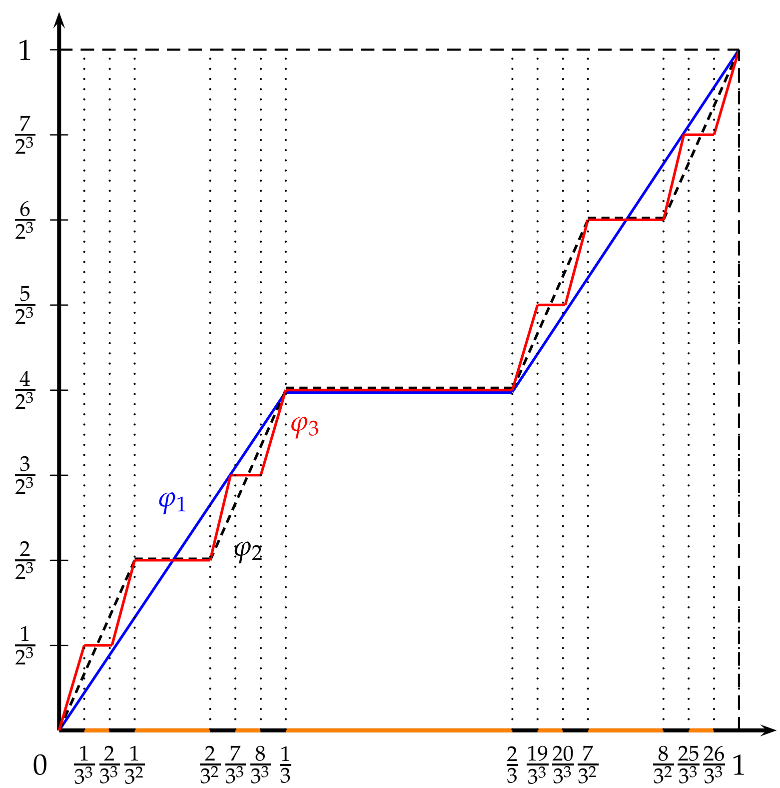
(c) φ_n es lineal sobre cada uno de los intervalos cerrados F_{nk} que conforman a Γ_n.

Observe que φ_n es constante sobre cada uno de los subintervalos abiertos $J_n(k)$ que definen a J_n, tomando el valor $1/2^n$ sobre $J_n(1)$, el valor $2/2^n$ sobre $J_n(2)$, hasta alcanzar el valor $(2^n-1)/2^n$ sobre $J_n(2^n-1)$.

Las funciones φ_n se pueden describir explícitamente del modo siguiente: comenzando con la función identidad, esto es, $\varphi_0(x)=x$ para todo $x\in[0,1]$, entonces

$$\varphi_{n+1}(x) = \begin{cases} \frac{1}{2}\cdot\varphi_n(3x) & \text{si} \quad 0\leq x\leq \frac{1}{3} \\ \frac{1}{2} & \text{si} \quad \frac{1}{3}< x< \frac{2}{3} \\ \frac{1}{2}+\frac{1}{2}\cdot\varphi_n(3x-2) & \text{si} \quad \frac{2}{3}\leq x\leq 1 \end{cases}$$

para cada $n\in\mathbb{N}_0$.



En el gráfico de arriba se muestran las funciones φ_1, φ_2 y φ_3. Es interesante observar que, en el intervalo $[0,1/3]$ la diferencia entre φ_2 y φ_1 no excede a $\frac{1}{2^2}$, $\varphi_1(x)=\varphi_2(x)$ si $x\in(1/3,2/3)$ y en el intervalo $[2/3,1]$ el comportamiento de φ_2 y φ_1 es el mismo que en el intervalo $[0,1/3]$, excepto que sus gráficas son desplazadas hacia arriba por una misma constante, la cual se cancela cuando uno toma la diferencia $\varphi_2-\varphi_1$. Similarmente, la diferencia entre φ_3 y φ_2 sobre el intervalo $[0,1/3^2]$ no excede a $\frac{1}{2^3}$ y, como antes, su comportamiento sobre cada uno de los los intervalos $[2/3^2,1/3]$, $[2/3,7/3^2]$ y $[8/3^2,1]$ son los mismos que en $[0,1/3^2]$ y la igualdad $\varphi_2(x)=\varphi_3(x)$ se cumple para todo $x\in J_3$. Este razonamiento nos lleva a la conclusión de que para entender la diferencia entre φ_n y φ_{n+1} basta con analizarla en cualquier intervalo de

$[0,1]\setminus J_n = \Gamma_n$. En particular, podemos considerar, sin perder generalidad en el razonamiento, el intervalo $[0,1/3^n]$ para saber cómo está acotada la diferencia $\varphi_n - \varphi_{n+1}$. Nótese que sobre dicho intervalo se tiene que $\varphi_n(x) = \frac{3^n}{2^n}x$ para todo $x \in [0,1/3^n]$, mientras que φ_{n+1} viene dada por

$$\varphi_{n+1}(x) = \begin{cases} \dfrac{3^{n+1}}{2^{n+1}}\,x & \text{si } \; x \in [0,\, 1/3^{n+1}] \\ \dfrac{1}{2^{n+1}} & \text{si } \; x \in [1/3^{n+1},\, 2/3^{n+1}] \\ \dfrac{1}{2^{n+1}} + \dfrac{3^{n+1}}{2^{n+1}}\,x & \text{si } \; x \in [2/3^{n+1},\, 3/3^{n+1}] \end{cases}$$

Puesto que la igualdad

$$\varphi_{n+1}(x) = \varphi_n(x)$$

se cumple para todo $x \in J_n$ y cada $n \in \mathbb{N}$, se deduce fácilmente que

$$\begin{aligned} \|\, \varphi_{n+1} - \varphi_n \,\|_\infty &= \sup\Big\{ |\varphi_{n+1}(x) - \varphi_n(x)| \,:\, 0 \leq x \leq 1 \Big\} \\ &= \sup\Big\{ |\varphi_{n+1}(x) - \varphi_n(x)| \,:\, 0 \leq x \leq 1/3^n \Big\} \\ &\leq \frac{1}{2^{n+1}}. \end{aligned}$$

Otro aspecto que hay que observar es que φ_n es inyectiva sobre Γ_n y $\varphi_n(\Gamma_n) = [0,1]$ para todo $n \geq 1$.

Teorema 5.1.8. *La sucesión de funciones $(\varphi_n)_{n=1}^\infty$ **converge uniformemente** sobre $[0,1]$.*

Prueba. Sea $\varepsilon > 0$. Teniendo en cuenta que la serie $\sum_{n=1}^\infty 1/2^n$ converge, se sigue del Criterio de Cauchy para Series que existe un $N \in \mathbb{N}$ tal que si $m > n \geq N$, entonces $\sum_{k=n+1}^{m-1} 1/2^k < \varepsilon$. De esto último se concluye que si $m > n \geq N$, entonces

$$\|\, \varphi_m - \varphi_n \,\|_\infty \;\leq\; \sum_{k=n+1}^{m-1} \|\, \varphi_{k+1} - \varphi_k \,\|_\infty \;\leq\; \sum_{k=n+1}^{m-1} \frac{1}{2^k} \;<\; \varepsilon$$

y así, gracias al Teorema 3.1.42, la sucesión $(\varphi_n)_{n=1}^\infty$ converge uniformemente sobre $[0,1]$. ■

Definición 5.1.9. *La función $\varphi_\Gamma : [0,1] \to [0,1]$ definida por*

$$\varphi_\Gamma(x) \;=\; \lim_{n\to\infty} \varphi_n(x)$$

*para todo $x \in [0,1]$, es llamada la **función de Cantor**, **función de Cantor-Lebesgue** o **escalera del Diablo**.*

(2ª) **Definición**. La siguiente definición de φ_Γ se apoya sobre la integral de Riemann. Para cada $n \geq 1$, considere, como antes, el conjunto $\Gamma_n = \bigcup_{k=1}^{2^n} F_{nk}$ y defina $f_n : [0,1] \to \mathbb{R}$ por

$$f_n(x) \;=\; \left(\frac{3}{2}\right)^n \cdot \chi_{\Gamma_n}(x) \;=\; \begin{cases} \left(\dfrac{3}{2}\right)^n & \text{si } \; x \in \Gamma_n \\ 0 & \text{si } \; x \notin \Gamma_n. \end{cases}$$

La función f_n es claramente discontinua en un conjunto finito de puntos y, por lo tanto, es Riemann integrable gracias al Teorema de Vitali-Lebesgue (véase, (**TVL**$_4$), página 424). Apoyándonos en esta información podemos definir la función $\varphi_n : [0,1] \to \mathbb{R}$ declarando que

$$\varphi_n(x) \ = \ \int_0^x f_n(t)\,dt.$$

El Teorema Fundamental del Cálculo nos asegura que cada función φ_n es Lipschitz y, en particular, continua. Observe también que

$$\varphi_n(0) \ = \ 0 \qquad \text{y} \qquad \varphi_n(1) \ = \ \int_0^1 f_n(t)\,dt \ = \ \left(\frac{3}{2}\right)^n \int_0^1 \chi_{\Gamma_n}(t)\,dt \ = \ \left(\frac{3}{2}\right)^n \cdot \ell(\Gamma_n) \ = \ 1.$$

Tenemos entonces que φ_n es una función continua con $\varphi_n(0) = 0$ y $\varphi_n(1) = 1$ para cada $n \geq 1$. Nótese que φ_n es constante sobre $[0,1] \setminus \Gamma_n$ y, además, lineal con pendiente $(\frac{2}{3})^n$ sobre cada $F_{nk} \subseteq \Gamma_n$. Esto nos indica que cada φ_n es monótona creciente. Sobre cada conjunto $F_{nk} \subseteq \Gamma_n$ se tiene que $f_n(x) = (\frac{3}{2})^n$ para todo $x \in F_{nk}$, mientras que $f_{n+1}(x) = (\frac{3}{2})^{n+1} = (\frac{3}{2})^n f_n(x)$ para x en el primer tercio o último tercio de F_{nk}, e igual a cero en el tercio del centro. Se sigue de esto que

$$\int_{F_{nk}} f_n(t)\,dt \ = \ \int_{F_{nt}} f_{n+1}(t)\,dt \ = \ 2^{-n}. \tag{1}$$

Puesto que $f_n(x) = f_{n+1}(x)$ para todo $x \notin \Gamma_n$, resulta que si α es un punto extremo arbitrario en cualquiera de los intervalos que conforman a Γ_n, entonces

$$\int_0^\alpha f_n(t)\,dt \ = \ \int_0^\alpha f_{n+1}(t)\,dt, \tag{2}$$

de donde se obtiene

$$\varphi_{n+1}(x) \ = \ \varphi_n(x) \quad \text{para todo } x \in \Gamma_n$$

y, en consecuencia,

$$\big|\varphi_n(x) - \varphi_{n+1}(x)\big| \ < \ 2^{-n+1} \quad \text{para todo } x \in [0,1].$$

Lo anterior nos revela que la sucesión $(\varphi_n)_{n=1}^\infty$ es uniformemente de Cauchy y, en consecuencia, converge uniformemente a un función continua φ_Γ: la función de Cantor.

(3ª) **Definición**. Esta es otra forma de definir a φ_Γ usando la representación ternaria de los puntos en $[0,1]$. Sea $x \in [0,1]$ y exprese dicho número en su representación ternaria habitual, digamos

$$x \ = \ \sum_{n=1}^\infty \frac{a_n(x)}{3^n},$$

donde $a_n(x) \in \{0,1,2\}$ para todo $n \geq 1$. Si $x \notin \Gamma$, entonces existe al menos un $n \in \mathbb{N}$ tal que $a_n(x) = 1$. Sea n_x el entero positivo más pequeño para el cual $a_{n_x}(x) = 1$. Si $x \in \Gamma$, entonces todos los $a_n(x)$ son distintos de 1 y, por lo tanto, convenimos en tomar $n_x = +\infty$. Este análisis permite definir la función de Cantor $\varphi_\Gamma : [0,1] \to \mathbb{R}$ por

$$\varphi_\Gamma(x) \ = \ \frac{1}{2^{n_x}} + \frac{1}{2}\sum_{k=1}^{n_x - 1} \frac{a_n(x)}{2^n}.$$

Observe que si $x \in \Gamma$, la igualdad anterior toma la forma

$$\varphi_\Gamma(x) = \frac{1}{2} \sum_{k=1}^{\infty} \frac{a_n(x)}{2^n},$$

donde $a_n(x) \in \{0, 2\}$ para todo $n \geq 1$.

(4ª) **Definición**. Otra forma equivalente de definir la función de Cantor se obtiene de la siguiente manera: para cada $x = \sum_{n=1}^{\infty} \frac{2a_n}{3^n} \in \Gamma$, donde $a_n \in \{0, 1\}$, defina

$$\varphi_\Gamma(x) = \varphi_\Gamma\left(\sum_{n=1}^{\infty} \frac{2a_n}{3^n}\right) = \sum_{n=1}^{\infty} \frac{a_n}{2^n}$$

y extienda φ_Γ a todo $[0, 1]$ poniendo

$$\begin{aligned} \varphi_\Gamma(x) &= \sup\{\varphi_\Gamma(y) : y \in \Gamma,\ y \leq x\} \\ &= \sup\left\{\sum_{n=1}^{\infty} \frac{a_n}{2^n} : \sum_{n=1}^{\infty} \frac{2a_n}{3^n} \leq x,\ a_n \in \{0, 1\}\right\}, \end{aligned}$$

para todo $x \in [0, 1]$.

El siguiente resultado muestra algunas de las características más importantes de la función de Cantor.

Corolario 5.1.10. *La función de Cantor φ_Γ es **continua**, **creciente** y **sobreyectiva**. Además, $\varphi'_\Gamma(x) = 0$ para todo $x \in [0, 1] \setminus \Gamma$.*

Prueba. Sea $(\varphi_n)_{n=1}^{\infty}$ la sucesión de funciones continuas definidas anteriormente. Por el Teorema 5.1.8 sabemos que $\varphi_n \to \varphi_\Gamma$ uniformemente y, entonces, el Teorema 3.1.40, página 177, nos garantiza que φ_Γ es continua sobre $[0, 1]$ y, por supuesto, creciente y sobreyectiva ya que cada φ_n posee esas propiedades. Para demostrar que $\varphi'_\Gamma(x) = 0$ para todo $x \in [0, 1] \setminus \Gamma$, observe que φ_Γ es constante en cada uno de los intervalos abiertos borrados en la construcción de Γ y, por lo tanto, $\varphi'_\Gamma(x) = 0$ para todo $x \in [0, 1] \setminus \Gamma$. ■

5.2. Problemas

(1) Un número $x \in [0, 1]$ es llamado **normal** si éste posee una representación binaria $(0, x_1 x_2 \ldots)_2$ tal que

$$\lim_{n \to \infty} \frac{s_n(x)}{n} = \frac{1}{2},$$

donde $s_n(x) = \operatorname{card}(\{1 \leq i \leq n : x_i = 0\})$. Si E denota el conjuntos de todos los números normales, demuestre que

$$E = \bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} B_{n,k},$$

donde

$$B_{n,k} = \left\{x \in [0, 1] : \left|\frac{s_n(x)}{n} - \frac{1}{2}\right| < \frac{1}{k}\right\}.$$

(2) Sea $\big((a_n, b_n)\big)_{n=1}^{\infty}$ los intervalos componentes del conjunto abierto $[0,1] \setminus \Gamma$. Sean

$$P_m = \left\{\frac{1}{2}(a_i + b_i) : i \in \mathbb{N}\right\} \qquad \text{y} \qquad P_d = \{b_i : i \in \mathbb{N}\}$$

Pruebe que P_m es un G_δ, pero que P_d no lo es.

(3) Pruebe que existe una función monótona $f : \Gamma \to \mathbb{R}$ que no es diferenciable en ningún punto de Γ.

(4) Pruebe que, para cada $n \in \mathbb{N}$, el conjunto

$$F_n = \left\{x \in [0,1] : x = \sum_{k=1}^{n} \frac{c_k}{3^k}, \text{ con } c_k \in \{0,2\} \text{ para todo } k = 1, \ldots, n\right\}$$

consiste de todos los extremos izquierdos de los subintervalos que conforman a Γ_n en la construcción del conjunto ternario de Cantor. Concluya que Γ_n se puede representar en la forma:

$$\Gamma_n = \bigcup_{x \in F_n} \left[x, x + \frac{1}{3^n}\right].$$

(5) **Conjunto n-ario de Cantor**. Sea $n = 2m + 1$, $m = 1, 2, \ldots$ y como antes, sea $\Gamma_0 = [0,1]$. Divida el intervalo $[0,1]$ en n subintervalos de igual longitud y elimine los intervalos abiertos

$$\left(\frac{1}{n}, \frac{2}{n}\right), \quad \left(\frac{3}{n}, \frac{4}{n}\right), \quad \ldots \quad, \left(\frac{n-2}{n}, \frac{n-1}{n}\right).$$

Sea

$$\Gamma_1 = \left[0, \frac{1}{n}\right] \cup \left[\frac{2}{n}, \frac{3}{n}\right] \cup \cdots \cup \left[\frac{n-1}{n}, 1\right]$$

los intervalos que permanecen. Observe que Γ_1 contiene $m+1$ intervalos. A cada uno de los intervalos que forman a Γ_1 divídalos en n subintervalos de igual longitud y remueva de ellos los intervalos abiertos que ocupan el 2do, el 4to,$\ldots$, $2m$-ésimo lugar y continúe. De este modo se obtiene una sucesión de conjuntos cerrados $(\Gamma_k)_{k=1}^{\infty}$ tal que Γ_k contiene, en el k-ésimo paso, $(m+1)^k$ intervalos cada uno de longitud $(1/n)^k$, $k = 0, 1, 2, \ldots$ Pruebe que

$$\Gamma(n) = \bigcap_{k=0}^{\infty} \Gamma_k$$

posee las mismas propiedades que Γ.

(6) Sea $\alpha \in (0,1)$ y defina $\beta = (1-\alpha)/2$. Sea Γ_α^β el conjunto tipo-Cantor construido en la Sección 5.1.5. Pruebe que:

(a) Si $\beta < 1/3$, entonces $\Gamma_\alpha^\beta - \Gamma_\alpha^\beta$ tiene medida cero.

(b) Si $\beta \geq 1/3$, entonces $\Gamma_\alpha^\beta - \Gamma_\alpha^\beta = [-1,1]$.

Advertencia: La prueba no es trivial, véase, por ejemplo, [80].

(7) Pruebe que el grafo de la función de Cantor, como subconjunto del plano, tiene longitud igual a 2.

(8) Pruebe que la función de Cantor φ_Γ es subaditiva, es decir, para todo $x, y \in [0,1]$,

$$\varphi_\Gamma(x+y) \ \leq \ \varphi_\Gamma(x) \ + \ \varphi_\Gamma(y).$$

CAPÍTULO 6

La Medida de Lebesgue en $\mathbb{R}$

6.1. Introducción

En geometría elemental se nos enseña cómo determinar el perímetro y área de un polígono, de un círculo y otras figuras planas más o menos "sencillas". Un poco después, en los cursos de cálculo diferencial e integral se nos muestra cómo se determinan las áreas de figuras más complicadas como aquellas que se encuentran limitadas por una curva continua, digamos $y = f(x)$, un par de ordenadas $x = a$, $x = b$ y el eje X. Esto se hace mediante la integral de Riemann. Sin embargo, dicha integral, lamentablemente, posee ciertas fisuras y limitaciones que no permiten que se pueda integrar una amplia gama de funciones acotadas importantes y, por consiguiente, impide que se pueda determinar el área limitada por funciones de ese tipo. De estos y otros serios problemas con dicha integral surgió la integral de Lebesgue, una nueva integral más amplia y poderosa que la integral de Riemann cuyo desarrollo se sustenta, en primer lugar, sobre la noción de "medida" de un conjunto que desarrollaremos en este capítulo.

La noción matemática de *medida* que se utiliza hoy en día en los textos sobre Teoría de la Medida e Integración fue dada a conocer por Émile Borel en su libro "*Leçons sur la Théorie de Fonctions*" publicado en el año 1898. Con esa noción, Borel pretende representar conceptos tales como: longitud, área, volumen, masa, carga eléctrica, etc. Los objetos a ser medidos son pensados como conjuntos y una medida es una función de conjuntos que es aditiva, es decir, asigna a la unión de cualquier colección finita de conjuntos disjuntos la suma de las medidas de cada uno de ellos. El ingrediente principal que Borel añade a dicha definición y que tendrá un grand impacto, es la noción de "*numerablemente aditiva*". Borel también introduce los conjuntos que él llama medibles pero no estudia sus propiedades. Es Henri Lebesgue quien, posteriormente, analiza de manera rigurosa dichas propiedades logrando obtener una colección especial de "conjuntos medibles" a la que se denominará más tarde como una σ-álgebra. Lo que Lebesgue nos revela es que esta nueva noción introducida por Borel es el marco ideal para desarrollar su obra más preciada: la integral que hoy lleva su nombre. La motivación de Borel para introducir el concepto de medida proviene de su estudio del tamaño del conjunto de puntos sobre el cual ciertas series infinitas convergen, sin embargo, Borel no pudo ver la vinculación de esta idea de medida con la

Teoría de Integración.

6.2. La Medida Exterior de Lebesgue

El objetivo fundamental en esta y las siguientes secciones es la de presentar la construcción clásica de la *medida de Lebesgue* en $\mathbb{R}$, partiendo de la noción de medida exterior de Lebesgue. Es bueno advertir, sin embargo, que casi todos los resultados y procedimientos que desarrollaremos sobre la construcción de la medida de Lebesgue sobre $\mathbb{R}$, se puede extender, casi sin ningún cambio, en el espacio Euclidiano $\mathbb{R}^n$, $n \geq 2$. Los pasos que seguiremos en la construcción de la medida de Lebesgue se desarrollarán de acuerdo al siguiente esquema:

(a) Se comienza con la clase de los intervalos de $\mathbb{R}$ que, como sabemos, se "miden" a través de su *longitud*.

(b) Inmediatamente después se extiende la noción de longitud a *todos los subconjuntos* de $\mathbb{R}$ por medio de una definición que permite asignar a cada subconjunto A de $\mathbb{R}$ un único y bien definido número real extendido no negativo que se denotará por $\mu^*(A)$. Esta función de conjuntos es lo que llamaremos la *medida exterior de Lebesgue*. Lo bueno de esta manera de "medir" los subconjuntos de $\mathbb{R}$ es que ella preserva la longitud de los intervalos de $\mathbb{R}$, es decir, si I es un intervalo de $\mathbb{R}$, entonces $\mu^*(I)$ es exactamente la longitud de I. Además, μ^* es invariante por traslación, lo cual significa que la "medida" de cualquier conjunto y la de sus trasladados son exactamente las mismas. Sin embargo, no todas las propiedades que uno espera se cumplan, son satisfechas por μ^*. Por ejemplo, μ^* no es lo que uno llama una "*medida finitamente aditiva*", ni mucho menos *numerablemente aditiva*, lo cual quiere decir que la media exterior de la unión de una colección finita, o infinita numerable, de subconjuntos de $\mathbb{R}$ disjuntos dos a dos es igual a la suma total de las medidas exteriores de cada uno de esos conjuntos.

(c) La noción de *medida de Lebesgue* finalmente se obtiene cuando se restringe la medida exterior de Lebesgue a una clase muy particular, pero suficientemente amplia, de subconjuntos de $\mathbb{R}$. Dicha clase es lo que comúnmente se llama la *σ-álgebra de Lebesgue*.

En conclusión, la medida de Lebesgue nace como un intento de generalizar la noción de longitud de un intervalo arbitrario a cualquier conjunto de números reales.

Recordemos que si I es cualquier *intervalo acotado* con extremos a y b donde $a < b$, entonces, por definición, su **longitud** viene expresada mediante el número

$$\ell(I) \;=\; b - a.$$

Si $a = b$, el intervalo I se reduce a un punto y, por supuesto, su longitud es cero, es decir, $\ell(\{a\}) = 0$. A tales intervalos se les llama **degenerados**. Por otro lado, todo intervalo que no es degenerado se le denomina **no-degenerado**. Finalmente, si I es un *intervalo no acotado*, su longitud se designa por $\ell(I) = +\infty$. En lo que sigue asumiremos que cuando hablamos, en términos generales, de un intervalo éste puede ser degenerado o no-degenerado. Recordemos que habíamos convenido en definir la "longitud" de una unión finita de intervalos acotados y disjuntos dos a dos como la suma de las longitudes de cada uno de esos intervalos. La siguiente definición es un poco más general.

Definición 6.2.1. *Un subconjunto E de $\mathbb{R}$ se llama **elemental** si E es una unión finita de intervalos no-superpuestos.*

Observe que cada conjunto elemental E se puede representar de muchas formas. Por ejemplo,

$$\begin{aligned} E &= [0,1] \cup [1,2) \cup (2,3] \cup [3,4] \cup \{5\} \\ &= [0,2) \cup (2,3] \cup [3,4] \cup \{5\} \\ &= [0,1] \cup [1,2) \cup (2,4] \cup \{5\} \end{aligned}$$

Sea $\mathfrak{E}$ *la colección de todos los subconjuntos elementales de* $\mathbb{R}$ y consideremos la función de conjuntos

$$\zeta : \mathfrak{E} \to [0, +\infty]$$

definida por

$$\zeta(E) = \ell(I_1) + \cdots + \ell(I_n).$$

para todo $E \in \mathfrak{E}$ representado en la forma $E = I_1 \cup \cdots \cup I_n$. No es difícil comprobar que si $E = J_1 \cup \cdots \cup J_m$ es otra representación de E por medio de intervalos no-superpuestos, entonces $\ell(I_1) + \cdots + \ell(I_n) = \ell(J_1) + \cdots + \ell(J_m)$. Más aun, es fácil establecer que si E_1 y E_2 son elementos disjuntos de $\mathfrak{E}$, entonces

$$\zeta(E_1 \cup E_2) = \zeta(E_1) + \zeta(E_2).$$

Un hecho interesante que hay que destacar es que la familia $\mathfrak{E}$ es un **anillo** de conjuntos, es decir, si $E_1, \ldots, E_n$ es cualquier colección finita de elementos de $\mathfrak{E}$, entonces $\bigcup_{i=1}^n E_i$, así como $\bigcap_{i=1}^n E_i$, ambos permanecen dentro de $\mathfrak{E}$. Similarmente, si $E, F \in \mathfrak{E}$, entonces $E \setminus F \in \mathfrak{E}$. Lo anterior permite deducir la existencia de una *medida* que asigna a cada conjunto elemental un número real positivo extendido bien definido. Se puede probar que dicha medida posee ciertas propiedades interesantes, a pesar de ser muy limitadas. Lo que deseamos indagar es ver si es posible extender dicha noción de *medida* a conjuntos más "complicados" que los elementales, es decir, queremos investigar si es posible construir una función de conjuntos μ que asigne a cada elemento E, en alguna colección $\mathfrak{M}$ más amplia que $\mathfrak{E}$, un número real positivo extendido $\mu(E)$ y que las propiedades que posea dicha función de conjuntos capture perfectamente nuestra noción intuitiva de longitud (y de área y volumen cuando nos ubicamos en los espacios $\mathbb{R}^2$ y $\mathbb{R}^3$ respectivamente). En otras palabras, lo ideal sería obtener una tal μ que verifique las siguientes cuatro propiedades, conocidas como:

El Problema de la Medida de Lebesgue:

(α_1) *$\mu(E)$ esté definida para todo $E \subseteq \mathbb{R}$, es decir, $0 \le \mu(E) \le +\infty$ para todo $E \subseteq \mathbb{R}$;*

(α_2) *$\mu(I) = \ell(I)$ para cualquier intervalo $I \subseteq \mathbb{R}$; es decir, que la* "medida" *de un* ***intervalo*** *sea exactamente la* ***longitud*** *de dicho intervalo;*

(α_3) *μ sea* ***invariante por traslación****, esto es, que la igualdad*

$$\mu(x + E) = \mu(E)$$

se cumpla para todo $E \in \mathfrak{M}$ y todo $x \in \mathbb{R}$, y, finalmente,

(α_4) *μ sea* ***numerablemente aditiva****, o* ***σ-aditiva****, lo cual quiere decir que si $(E_n)_{n=1}^\infty$ es una sucesión disjunta en $\mathscr{P}(\mathbb{R})$, entonces*

$$\mu\left(\bigcup_{n=1}^\infty E_n\right) = \sum_{n=1}^\infty \mu(E_n).$$

El Problema de la Medida fue propuesto por H. Lebesgue en 1904 provocando, desde su formulación, un gran debate: ¿tiene dicho problema una solución positiva? La respuesta es, en general, negativa y la razón de peso de tal aseveración es, como hemos convenido, la aceptación del Axioma de Elección. En efecto, en el transcurso de este capítulo veremos, como lo demostró G. Vitali en 1905, que es imposible, si se acepta el Axioma de Elección, que la medida de Lebesgue, la cual es la medida en la que estamos interesados, retenga las cuatro propiedades $(\alpha_1) - (\alpha_4)$ antes mencionadas. De hecho, lo que dedujo Vitali es que la propiedad (α_1) es insostenible y, en consecuencia, debemos conformarnos con definir a μ sobre una cierta clase $\mathfrak{M} \subsetneqq 2^{\mathbb{R}}$ pero reteniendo las otras tres propiedades. En consecuencia, la aceptación del Axioma de Elección demuestra que El Problema de la Medida posee una solución negativa.

El artifice de la construcción de la Teoría de Integración que expondremos en los siguientes capítulos se debe fundamentalmente a Henry Léon Lebesgue quien nació en Beauvais, Francia, el 28 de Junio del año 1875. En 1894 entra a la École Normale Superieur una de las instituciones de enseñanza de mayor prestigio ubicada en Francia (Paris), graduándose en el año 1897. A partir de ese momento comienza a trabajar sobre su tesis doctoral la cual propone un enfoque completamente revolucionario para resolver los problemas derivados de la integral de Riemann. Durante el período que va de 1899 a 1901 publica los resultados de sus investigaciones. La tesis fue aceptada formalmente en la Sorbona en 1902, y en el lapso de 1902 a 1903 dicta el prestigioso Cours Peccot en el Collège de France el cual va ser publicado bajo el nombre de *Leçons sur l'integration et la recherche des fonctions primitives* (Conferencias sobre la integración y la búsqueda de funciones primitivas). Muere Lebesgue en París el 26 de Julio de 1941.

En todo lo que sigue, denotaremos por $\mathfrak{J}_{\mathbb{R}}$ *la colección de todas las sucesiones de **intervalos abiertos** de* $\mathbb{R}$. Para cada subconjunto A de $\mathbb{R}$ considere la subcolección de $\mathfrak{J}_{\mathbb{R}}$, $\mathfrak{J}_{\mathbb{R}}(A)$, formada por todas las sucesiones $(I_n)_{n=1}^{\infty}$ en $\mathfrak{J}_{\mathbb{R}}$ que satisfacen la condición $A \subseteq \bigcup_{n=1}^{\infty} I_n$, esto es,

$$\mathfrak{J}_{\mathbb{R}}(A) \;=\; \left\{ (I_n)_{n=1}^{\infty} \in \mathfrak{J}_{\mathbb{R}} \;:\; A \;\subseteq\; \bigcup_{n=1}^{\infty} I_n \right\}$$

Por cada $(I_n)_{n=1}^{\infty} \in \mathfrak{J}_{\mathbb{R}}(A)$, consideremos la serie $\sum_{n=1}^{\infty} \ell(I_n)$. Puesto que las longitudes son números reales positivos o $+\infty$, resulta que la serie $\sum_{n=1}^{\infty} \ell(I_n)$ o es convergente, o diverge a $+\infty$. En cualquier caso, $\sum_{n=1}^{\infty} \ell(I_n)$ es un número real extendido no negativo bien definido pues siempre se obtiene el mismo valor independientemente del orden en que se coloquen los términos $\ell(I_n)$. Esto nos permite formular la siguiente definición.

Definición 6.2.2. *Para cada subconjunto A de $\mathbb{R}$ se define la **medida exterior de Lebesgue** de A como*

$$\mu^*(A) \;=\; \inf \left\{ \sum_{n=1}^{\infty} \ell(I_n) \;:\; (I_n)_{n=1}^{\infty} \;\in\; \mathfrak{J}_{\mathbb{R}}(A) \right\}.$$

La definición de medida exterior de Lebesgue que abreviaremos, a partir de este momento, simplemente como ***medida exterior***, permite, de modo inmediato, concluir que:

(1_1) $0 \leq \mu^*(A) \leq +\infty$ *para todo* $A \subseteq \mathbb{R}$. Esto significa que el dominio de la función de conjuntos μ^* es $\mathscr{P}(\mathbb{R})$ y su rango es $[0, +\infty]$, es decir,

$$\mu^* : \mathscr{P}(\mathbb{R}) \to [0, +\infty].$$

(2_2) $\mu^*(A) \leq \sum_{n=1}^{\infty} \ell(I_n)$ *para cualquier sucesión* $(I_n)_{n=1}^{\infty} \in \mathfrak{J}_{\mathbb{R}}(A)$.

(3_3) *Para cada* $\varepsilon > 0$, *siempre se puede elegir, usando las propiedades del ínfimo, una sucesión* $(I_n)_{n=1}^{\infty}$ *en* $\mathfrak{J}_{\mathbb{R}}(A)$ *de modo tal que*

$$\mu^*(A) \leq \sum_{n=1}^{\infty} \ell(I_n) \leq \mu^*(A) + \varepsilon.$$

Observe que podemos tomar $\mu^*(A) = +\infty$ en dichas desigualdades.

(4_4) $\mu^*(A) \leq \mu^*(B)$ *cualesquiera sean los subconjuntos* A *y* B *de* $\mathbb{R}$ *con* $A \subseteq B$. Esto sigue del hecho de que

$$\mathfrak{J}_{\mathbb{R}}(B) \subseteq \mathfrak{J}_{\mathbb{R}}(A).$$

En efecto, si $\left(I_n\right)_{n=1}^{\infty} \in \mathfrak{J}_{\mathbb{R}}(B)$, entonces $B \subseteq \bigcup_{n=1}^{\infty} I_n$ y como $A \subseteq B$, resulta que $A \subseteq \bigcup_{n=1}^{\infty} I_n$, es decir, $\left(I_n\right)_{n=1}^{\infty} \in \mathfrak{J}(A)$. De (2_2) se sigue que, $\mu^*(A) \leq \sum_{n=1}^{\infty} \ell(I_n)$ y como ésta desigualdad es válida para cualquier sucesión $\left(I_n\right)_{n=1}^{\infty} \in \mathfrak{J}_{\mathbb{R}}(B)$, tenemos finalmente que

$$\mu^*(A) \leq \inf\left\{ \sum_{n=1}^{\infty} \ell(I_n) : (I_n)_{n=1}^{\infty} \in \mathfrak{J}_{\mathbb{R}}(B) \right\} = \mu^*(B).$$

A esta desigualdad se le conoce con el nombre de **propiedad monótona** (o de monotonía) de μ^*.

Ya sabemos que μ^* cumple con la condición (α_1), de modo que nuestra tarea inmediata es ver que μ^* satisface las condiciones (α_2) y (α_3) pero no así con la condición (α_4). El siguiente resultado establece que la definición de medida exterior es consistente con la noción de longitud de un intervalo y, por consiguiente, es una extensión de la función de conjuntos ζ antes definida.

Teorema 6.2.3. *Para cada **intervalo*** $I \subseteq \mathbb{R}$, *se cumple que* $\mu^*(I) = \ell(I)$.

Prueba. Suponga, en primer lugar, que I es un intervalo cerrado y acotado, digamos $I = [a, b]$ con $a, b \in \mathbb{R}$ y $a < b$. Puesto que, para cada $\varepsilon > 0$, el intervalo abierto $(a - \varepsilon, b + \varepsilon)$ contiene a I, tenemos, por la definición de μ^*, que $\mu^*(I) \leq \ell((a - \varepsilon, b + \varepsilon)) = b - a + 2\varepsilon$, de donde se sigue que

$$\mu^*(I) \leq b - a.$$

Queda por demostrar que $\mu^*(I) \geq b - a$. Para ver esto último, sea $\varepsilon > 0$ y usemos (3_3) para hallar una sucesión $(I_n)_{n=1}^{\infty}$ en $\mathscr{I}(I)$ tal que

$$\sum_{n=1}^{\infty} \ell(I_n) \leq \mu^*(I) + \varepsilon. \tag{6.2.1}$$

Ahora bien, teniendo en cuenta que el intervalo $I = [a, b]$ es un conjunto compacto y que $[a, b] \subseteq \bigcup_{n=1}^{\infty} I_n$, se sigue entonces del Teorema de Heine-Borel, página 133, que existe un conjunto finito de índices, digamos $n_1, \ldots, n_k$, tal que

$$[a, b] \subseteq \bigcup_{i=1}^{k} I_{n_i}.$$

Una vez establecida la inclusión anterior, podemos suponer, sin afectar el razonamiento que sigue, que $I_{n_i} = (a_i, b_i)$ para $i = 1, \ldots, k$, que ninguno de los I_{n_i} está incluido en ninguno de los intervalos restantes y que $a_1 < a_2 < \cdots < a_k$. Esta última cadena de desigualdades en combinación con $[a,b] \subseteq \bigcup_{i=1}^{k} I_{n_i}$, obliga a concluir que las desigualdades $a_1 < a < b < b_k$, así como $a_i < b_{i-1}$ se cumplen. De esto se sigue que

$$\begin{aligned}
\sum_{n=1}^{\infty} \ell(I_n) \ &\geq \ \sum_{i=1}^{k} \ell(I_{n_i}) \ = \ (b_k - a_k) \ + \ (b_{k-1} - a_{k-1}) \ + \ \cdots \ + \ (b_1 - a_1) \\
&= \ b_k \ - \ (a_k - b_{k-1}) \ - \ (a_{k-1} - b_{k-2}) \ - \ \cdots \ - \ (a_2 - b_1) \ - \ a_1 \\
&> \ b_k \ - \ a_1 \ > \ b \ - \ a.
\end{aligned}$$

De la desigualdad (6.2.1), vemos que

$$\mu^*(I) \ \geq \ \sum_{n=1}^{\infty} \ell(I_n) \ - \ \varepsilon \ > \ b \ - \ a \ - \ \varepsilon,$$

de donde se concluye, por la arbitrariedad del $\varepsilon > 0$, que $\mu^*(I) \geq b - a$.

Demostrada la igualdad $\mu^*(I) = \ell(I)$ para cualquier intervalo compacto $I = [a,b]$, tomemos ahora un intervalo arbitrario I de longitud finita pero no compacto. En este caso, I es necesariamente uno de los siguientes intervalos: $I = (a,b)$, $[a,b)$ o $(a,b]$. Fijemos un número $\varepsilon > 0$ elegido arbitrariamente y considere el intervalo $J = [a + \varepsilon/4, b - \varepsilon/4]$. Observe que J es un intervalo cerrado y acotado tal que $J \subseteq I$ y $\ell(J) > \ell(I) - \varepsilon$. De esto y el hecho de que $I \subseteq \overline{I}$ se sigue que

$$\begin{aligned}
\ell(I) \ - \ \varepsilon \ &< \ \ell(J) \\
&= \ \mu^*(J) && \text{por el primer caso} \\
&\leq \ \mu^*(I) && \text{por } (4_4) \\
&\leq \ \mu^*(\overline{I}) && \text{de nuevo por } (4_4) \\
&= \ \ell(\overline{I}) && \text{por el primer caso} \\
&= \ \ell(I), && \text{por definición}
\end{aligned}$$

es decir, para cada $\varepsilon > 0$, se cumple que

$$\ell(I) \ - \ \varepsilon \ < \ \mu^*(I) \ \leq \ \ell(I),$$

de donde se deduce que $\mu^*(I) = \ell(I)$.

Para finalizar la demostración del teorema, suponga que I es un intervalo de longitud infinita. Dado cualquier entero positivo n, seleccionemos un intervalo cerrado y acotado $J \subseteq I$ con $\ell(J) = n$. Por (4_4) y la demostración de la primera parte se tiene que $\mu^*(I) \geq \mu^*(J) = \ell(J) = n$, es decir, $\mu^*(I) \geq n$ para cualquier $n \in \mathbb{N}$. De aquí se sigue que $\mu^*(I) = +\infty = \ell(I)$ y termina la prueba. ■

En particular, si I es un intervalo degenerado, digamos $I = \{x\}$, entonces

(5_5) $\mu^*(\{x\}) \ = \ \ell(\{x\}) \ = \ 0.$

Más aun, como el conjunto vacío es un subconjunto de cualquier conjunto, resulta que $\varnothing \subseteq \{x\}$ y, por consiguiente, de la monotonía de μ^* se sigue que

(6_6) $\mu^*(\varnothing) \,=\, 0.$

(7_7) $\mu^*(G) \,>\, 0$ *para cualquier conjunto abierto* G *no vacío*. En efecto, como G es un abierto no vacío, él contiene un intervalo abierto, digamos $J \subseteq G$. Se sigue de (4_4) que $\mu^*(G) \geq \mu^*(J) = \ell(J) > 0$.

Veamos ahora que μ^* cumple con (α_3). Nótese que ser *invariante por traslación* en $\mathbb{R}$ significa que la "medida" de un conjunto es independiente de su localización, es decir, podemos desplazarlo a cualquier otro lugar sin afectar su "medida".

Teorema 6.2.4. *μ^* es **invariante por traslación**, es decir, para cualquier subconjunto A de $\mathbb{R}$ y cualquier $x \in \mathbb{R}$, se cumple que*

$$\mu^*(x + A) \;=\; \mu^*(A)$$

Prueba. Sea A un subconjunto de $\mathbb{R}$ y sea $x \in \mathbb{R}$. Si $A = \varnothing$, entonces $x + A = \{x\}$ y, así, por (5_5) y (6_6) tenemos que $\mu^*(x+A) = \mu^*(\{x\}) = 0 = \mu^*(A)$. Supongamos ahora que $A \neq \varnothing$ y sea $x \in \mathbb{R}$. Consideremos primero el caso en que $\mu^*(A) < +\infty$. Dado $\varepsilon > 0$, seleccione, usando (3_3), una sucesión $(I_n)_{n=1}^{\infty}$ de intervalos abiertos tal que

$$A \;\subseteq\; \bigcup_{n=1}^{\infty} I_n \qquad \text{y} \qquad \sum_{n=1}^{\infty} \ell(I_n) \;\leq\; \mu^*(A) \,+\, \varepsilon. \tag{6.2.2}$$

Puesto que claramente

$$x \,+\, A \;\subseteq\; x \,+\, \bigcup_{n=1}^{\infty} I_n \;\subseteq\; \bigcup_{n=1}^{\infty} (x + I_n),$$

y teniendo en cuenta que $\ell(x + I_n) = \ell(I_n)$ para todo $n \in \mathbb{N}$, resulta entonces de (2_2) y (6.2.2) que

$$\begin{aligned} \mu^*(x+A) \;&\leq\; \sum_{n=1}^{\infty} \ell(x + I_n) \;=\; \sum_{n=1}^{\infty} \ell(I_n) \\ &\leq\; \mu^*(A) \,+\, \varepsilon \end{aligned}$$

Siendo $\varepsilon > 0$ arbitrario, concluimos que $\mu^*(x+A) \leq \mu^*(A)$. Para demostrar la otra desigualdad, observe que como $A = -x + (x+A)$, entonces se tiene, gracias a la primera parte, que

$$\mu^*(A) \;=\; \mu^*(-x + (x+A)) \;\leq\; \mu^*(x+A).$$

Por esto $\mu^*(x+A) = \mu^*(A)$. Falta considerar el caso en que $\mu^*(A) = +\infty$, el cual se deja como un fácil ejercicio a cargo del lector. ■

Hasta ahora hemos demostrado que μ^* satisface $(\alpha_1), (\alpha_2)$ y (α_3). La condición (α_4) es una de esas condiciones que no se puede omitir en ninguna teoría que intente "medir" conjuntos de números reales partiendo, por supuesto, de la noción de longitud. Para ver que μ^* no cumple con (α_4) es necesario provocar una partición de $\mathscr{P}(\mathbb{R})$ en dos clases, digamos $\mathfrak{M}$ y $\mathfrak{N}$, de modo que en la clase $\mathfrak{M}$ cualquier colección numerable y disjunta su unión permanezca en ella y la medida exterior sea numerablemente aditiva para tal unión, mientras que en la clase $\mathfrak{N}$ se puede obtener al menos una colección finita o numerable que es disjunta dos a dos, pero que μ^* no es numerablemente aditiva sobre su unión. Esto lo haremos en la sección sobre conjuntos no-medibles. Mientras tanto, demostraremos el siguiente resultado el cual nos muestra que μ^* es *casi numerablemente aditiva*.

Teorema 6.2.5. *μ^* es **numerablemente subaditiva**, es decir, si $(A_n)_{n=1}^{\infty}$ es cualquier colección numerable de subconjuntos de $\mathbb{R}$, se cumple que*

$$\mu^*\Big(\bigcup_{n=1}^{\infty} A_n\Big) \leq \sum_{n=1}^{\infty} \mu^*(A_n). \tag{6.2.3}$$

Prueba. Observe, en primer lugar, que si $\mu^*(A_{n_0}) = +\infty$ para algún $n_0 \in \mathbb{N}$, entonces la monotonía de μ^* nos revela que $\mu^*\big(\bigcup_{n=1}^{\infty} A_n\big) \geq \mu^*(A_{n_0}) = +\infty$, lo cual nos indica que $\mu^*\big(\bigcup_{n=1}^{\infty} A_n\big) = \sum_{n=1}^{\infty} \mu^*(A_n) = +\infty$ y, por consiguiente, (6.2.3) se cumple. Suponga ahora que $\mu^*(A_n) < +\infty$ para todo $n \in \mathbb{N}$. Nótese que:

(1) Si $\sum_{n=1}^{\infty} \mu^*(A_n) = +\infty$, entonces claramente (6.2.3) se cumple ya que por (1_1)

$$\mu^*\Big(\bigcup_{n=1}^{\infty} A_n\Big) \leq +\infty = \sum_{n=1}^{\infty} \mu^*(A_n).$$

(2) Suponga entonces que $\sum_{n=1}^{\infty} \mu^*(A_n) < +\infty$. Fijemos un $\varepsilon > 0$ elegido arbitrariamente y hagamos uso de (3_3) para seleccionar, por cada $n \in \mathbb{N}$, una sucesión $(I_{n,k})_{k=1}^{\infty}$ de intervalos abiertos tal que

$$A_n \subseteq \bigcup_{k=1}^{\infty} I_{n,k} \qquad \text{y} \qquad \sum_{k=1}^{\infty} \ell(I_{n,k}) \leq \mu^*(A_n) + \frac{\varepsilon}{2^n}.$$

Puesto que

$$\bigcup_{n=1}^{\infty} A_n \subseteq \bigcup_{n=1}^{\infty}\bigcup_{k=1}^{\infty} I_{n,k}$$

se sigue entonces de (2_2) que

$$\begin{aligned}\mu^*\Big(\bigcup_{n=1}^{\infty} A_n\Big) &\leq \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} \ell(I_{n,k}) \\ &\leq \sum_{n=1}^{\infty}\Big[\mu^*(A_n) + \frac{\varepsilon}{2^n}\Big] \\ &= \sum_{n=1}^{\infty} \mu^*(A_n) + \varepsilon,\end{aligned}$$

de donde se concluye, por la arbitrariedad de nuestro ε, que

$$\mu^*\Big(\bigcup_{n=1}^{\infty} A_n\Big) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$$

y termina la prueba. ■

Observe que, para cualquier colección finita de subconjuntos de $\mathbb{R}$, digamos, $\{A_1, \ldots, A_n\}$ se cumple que

$$\mu^*\bigg(\bigcup_{k=1}^{n} A_k\bigg) \leq \sum_{k=1}^{n} \mu^*(A_k).$$

En efecto, basta con definir $A_k = \varnothing$ para todo $k \geq n+1$ y aplicar el Teorema 6.2.5. En este caso se dice que μ^* es **finitamente subaditiva**.

Corolario 6.2.6. *Si $A \subseteq \mathbb{R}$ es **numerable**, entonces $\mu^*(A) = 0$.*

Prueba. Suponga que $A = \{a_1, a_2, \ldots\}$. Entonces $A = \bigcup_{n=1}^{\infty}\{a_n\}$ y puesto que $\mu^*(\{a_n\}) = 0$ para todo $n \in \mathbb{N}$, el resultado sigue del Teorema 6.2.5. ■

Por ejemplo, el conjunto de los números racionales $\mathbb{Q} \subseteq \mathbb{R}$, a pesar de ser denso en $\mathbb{R}$, posee medida exterior cero ya que dicho conjunto es numerable. Podemos formularnos la siguiente pregunta: Si un conjunto $A \subseteq \mathbb{R}$ es tal que $\mu^*(A) = 0$, ¿será verdad que A es numerable? Como veremos un poco más adelante, el conjunto ternario de Cantor constituye un contraejemplo a dicha interrogante, pues $\mu^*(\Gamma) = 0$ pero Γ no es numerable. En lo que sigue, el símbolo $\mathfrak{N}_\mu(\mathbb{R})$ denotará *la colección de todos **conjuntos nulos***, es decir,

$$\mathfrak{N}_\mu(\mathbb{R}) = \big\{N \subseteq \mathbb{R} : \mu^*(N) = 0\big\}.$$

Veremos ahora que bajo la Hipótesis del Continuo cualquier conjunto de medida exterior positiva es no-numerable.

Corolario 6.2.7. *Bajo la **Hipótesis del Continuo**, cualquier **conjunto infinito** $A \subseteq \mathbb{R}$ con $\mu^*(A) > 0$ posee la **cardinalidad del continuo**.*

Prueba. Suponga que A es un conjunto infinito con $\mu^*(A) > 0$. Por la Hipótesis del Continuo, la cardinalidad de A es $\aleph_0$, o bien $\mathfrak{c}$. Suponer que la cardinalidad de A es $\aleph_0$ conduce a que A es numerable y, por consiguiente, gracias al Corolario 6.2.6, $\mu^*(A) = 0$ lo que produce una contradicción. Por lo tanto, si $\mu^*(A) > 0$, entonces $\operatorname{card}(A) = \mathfrak{c}$. ■

¿Qué ocurre si no se asume la Hipótesis del Continuo en el Corolario anterior? Es importante advertir que la condición $\mu^*(A) > 0$ no implica que A contenga, en su interior, un intervalo abierto. ¡Existen conjuntos A con $\mu^*(A) > 0$ que son nunca-densos! Sin embargo, la Hipótesis del Continuo no es necesaria en la conclusión del Corolario 6.2.7 como se puede demostrar en el Corolario 6.4.9, página 309.

Una de las nociones de mayor trascendencia y que jugará un papel fundamental en la Teoría de la Medida e Integración es la de conjunto nulo.

Definición 6.2.8. *Un conjunto $N \subseteq \mathbb{R}$ se llama un **conjunto nulo** si $\mu^*(N) = 0$.*

Una de las características excepcionales de los conjuntos nulos es que ellos no tienen ningún efecto cuando se añade o se excluye de un conjunto arbitrario; es decir:

Teorema 6.2.9. *Si $N \subseteq \mathbb{R}$ es un* ***conjunto nulo****, entonces*

$$\mu^*(A \cup N) = \mu^*(A) = \mu^*(A \setminus N)$$

para cualquier conjunto $A \subseteq \mathbb{R}$.

Prueba. Sea $N \subseteq \mathbb{R}$ un conjunto nulo. Se sigue de (4_4) y la subaditividad finita de μ^* que para cualquier conjunto $A \subseteq \mathbb{R}$,

$$\mu^*(A) \leq \mu^*(A \cup N) \leq \mu^*(A) + \mu^*(N) = \mu^*(A), \tag{1}$$

y, por lo tanto, $\mu^*(A \cup N) = \mu^*(A)$. Por otro lado, puesto que $A \cap N$ es un conjunto nulo, resulta que si reemplazamos N por $A \cap N$ y A por $A \setminus N$ en la ecuación (1) se obtiene que $\nu^*(A \setminus N) = \mu^*((A \setminus N) \cup (A \cap N))$; es decir, $\nu^*(A \setminus N) = \mu^*(A)$. ■

Es importante, en este punto, valorar la noción de numerabilidad sub-aditiva demostrada para μ^*; es decir, no se puede sustituir dicha noción por *sub-aditividad no-numerable* pues, en este caso, cualquier subconjunto no-numerable de $\mathbb{R}$ tendría medida exterior nula al ser unión no-numerable de sus puntos. En efecto, si $A = \{x_\alpha : \alpha \in I\}$ es no-numerable y si la sub-aditividad no-numerable fuese cierta, entonces

$$0 \leq \mu^*(A) = \mu^*\left(\bigcup_{\alpha \in I} \{x_\alpha\}\right) \leq \sum_{\alpha \in I} \mu^*(\{x_\alpha\}) = 0,$$

lo cual haría de μ^* una noción sin sentido.

Se sigue de la definición de medida exterior de un conjunto que los conjuntos nulos se pueden expresar de la siguiente forma.

Teorema 6.2.10. *Para un conjunto $N \subseteq \mathbb{R}$, las siguientes condiciones son equivalentes:*

(1) $\mu^*(N) = 0$.

(2) *Dado $\varepsilon > 0$, existe una sucesión $(I_n)_{n=1}^{\infty}$ de intervalos abiertos tal que*

$$N \subseteq \bigcup_{n=1}^{\infty} I_n \qquad y \qquad \sum_{n=1}^{\infty} \ell(I_n) < \varepsilon.$$

Prueba. En efecto, (1) ⇒ (2) es consecuencia inmediata de (3_3), mientras que (2) ⇒ (1) sigue de la definición. ■

El corolario que sigue fue demostrado por primera vez por George Cantor usando el ahora famoso e imprescindible *Método de la Diagonal*. He aquí otro modo de probarlo usando la noción de medida exterior.

Corolario 6.2.11. *El intervalo $[0,1]$ es* ***no-numerable****.*

Prueba. Si aceptamos la Hipótesis del Continuo, el resultado es consecuencia inmediata del Corolario 6.2.7 ya que $\mu^*([0,1]) > 0$. Si no deseamos usar dicha hipótesis, entonces podemos razonar de este otro modo: suponga que $[0,1]$ es numerable. Por el Corolario 6.2.6 sabemos que

$\mu^*([0,1]) = 0$, mientras que por el Teorema 6.2.3, $\mu^*([0,1]) = \ell([0,1]) = 1$, lo que resulta ser una franca e incuestionable contradicción. Por esto $[0,1]$ es no-numerable y finaliza la prueba. ■

Existen varios modos de "*aproximar*" un conjunto arbitrario por conjuntos más simples. El próximo teorema establece que cualquier subconjunto de $\mathbb{R}$ se puede "*aproximar*" por un conjunto abierto que lo contiene y cuya medida exterior difiere muy poco de la del conjunto dado.

Teorema 6.2.12. *Sea $A \subseteq \mathbb{R}$. Dado cualquier $\varepsilon > 0$, existe un **conjunto abierto** $G \subseteq \mathbb{R}$ tal que*

$$A \subseteq G \qquad y \qquad \mu^*(G) \leq \mu^*(A) + \varepsilon. \tag{6.2.4}$$

En particular,

$$\mu^*(A) = \text{ínf}\,\{\mu^*(G) : A \subseteq G,\ G \text{ es abierto}\}. \tag{6.2.5}$$

Prueba. Fijemos un $\varepsilon > 0$ elegido arbitrariamente. Si $\mu^*(A) = +\infty$, entonces basta tomar $G = \mathbb{R}$ para verificar que (6.2.4) se cumple. Supongamos ahora que $\mu^*(A) < +\infty$. Para el $\varepsilon > 0$ dado, usemos (3_3) para escoger una sucesión $(I_n)_{n=1}^\infty$ de intervalos abiertos tal que

$$A \subseteq \bigcup_{n=1}^\infty I_n \qquad \text{y} \qquad \sum_{n=1}^\infty \ell(I_n) \leq \mu^*(A) + \varepsilon.$$

Si ahora definimos $G = \bigcup_{n=1}^\infty I_n$, resulta que G es un conjunto abierto, $A \subseteq G$ y, se sigue de los Teorema 6.2.3 y Teorema 6.2.5 que

$$\begin{aligned} \mu^*(G) &= \mu^*\Big(\bigcup_{n=1}^\infty I_n\Big) \leq \sum_{n=1}^\infty \mu^*(I_n) = \sum_{n=1}^\infty \ell(I_n) \\ &\leq \mu^*(A) + \varepsilon. \end{aligned}$$

Esto prueba (6.2.4). Puesto que $\mu^*(A) \leq \mu^*(G)$, resulta de (6.2.4) y la elección arbitraria del ε, que $\mu^*(A) = \text{ínf}\,\{\mu^*(G) : A \subseteq G,\ G \text{ es abierto}\}$. ■

Recordemos que un subconjunto G de $\mathbb{R}$ se dice que es un G_δ, si existe una sucesión $(G_n)_{n=1}^\infty$ de subconjuntos abiertos no vacíos de $\mathbb{R}$ tal que $G = \bigcap_{n=1}^\infty G_n$. *La colección de todos los conjuntos G_δ incluidos en* $\mathbb{R}$ será denotada por $\mathcal{G}_\delta(\mathbb{R})$.

Corolario 6.2.13. *Sea A un subconjunto no vacío de $\mathbb{R}$. Existe un **conjunto** $G \in \mathcal{G}_\delta(\mathbb{R})$ tal que*

$$A \subseteq G \qquad y \qquad \mu^*(A) = \mu^*(G).$$

Prueba. Para cada $n \in \mathbb{N}$ existe, por el Teorema 6.2.12, un conjunto abierto $G_n \subseteq \mathbb{R}$ tal que

$$A \subseteq G_n \qquad \text{y} \qquad \mu^*(G_n) \leq \mu^*(A) + \frac{1}{n}.$$

Definiendo $G = \bigcap_{n=1}^\infty G_n$, resulta que G es un G_δ, $A \subseteq G$ y entonces, por la monotonía de μ^*, tenemos que $\mu^*(A) \leq \mu^*(G)$. Para obtener la otra desigualdad, notemos que $G \subseteq G_n$ para todo $n \geq 1$ y de nuevo, por la monotonicidad de μ^*, resulta que

$$\mu^*(G) \leq \mu^*(G_n) \leq \mu^*(A) + \frac{1}{n},$$

para todo $n \in \mathbb{N}$. De esto último se sigue que $\mu^*(G) \leq \mu^*(A)$ y así, $\mu^*(A) = \mu^*(G)$. ■

Combinando los dos resultados anteriores se tiene que:

Corolario 6.2.14. *Para cualquier conjunto $A \subseteq \mathbb{R}$,*

$$\mu^*(A) = \inf\{\mu^*(G) : A \subseteq G, G \in \mathcal{G}_\delta(\mathbb{R})\}.$$

6.2.1. Condiciones bajo la cual μ^* es σ-aditiva

Hasta ahora hemos demostrado que nuestra medida exterior μ^* satisface las propiedades (α_1), (α_2) y (α_3), excepto la numerabilidad aditiva. Desafortunadamente μ^* no es numerablemente aditiva hecho que demostraremos un poco más adelante al construir una sucesión de subconjuntos disjuntos dos a dos de $\mathbb{R}$, digamos $(V_n)_{n=1}^{\infty}$, tal que

$$\mu^*\left(\bigcup_{n=1}^{\infty} V_n\right) \neq \sum_{n=1}^{\infty} \mu^*(V_n).$$

Abordar la construcción de dicha sucesión va a demorar un largo pero agradable lapso de tiempo, por lo que, en compensación, mostraremos en lo que sigue que, bajo ciertas condiciones, uno puede garantizar, tanto en el caso finito así como en el caso infinito, la aditividad de μ^*.

Recordemos que si A y B son subconjuntos de $\mathbb{R}$, la **distancia** entre ellos viene dada por

$$\operatorname{dist}(A, B) = \inf\{|a - b| : a \in A, b \in B\}.$$

Diremos que A y B están ***estrictamente separados*** si $\operatorname{dist}(A, B) > 0$. Por ejemplo, del Teorema 2.2.32, página 134, sabemos que cualquier par de conjuntos compactos y disjuntos están estrictamente separados. Observe que si $\operatorname{dist}(A, B) > 0$, entonces necesariamente A y B son disjuntos. Además, si I es un intervalo abierto intersectando a ambos conjuntos, resulta claro que $\ell(I) > \operatorname{dist}(A, B)$. En este caso siempre es posible subdividir a I en piezas más pequeñas, digamos $\{I_1, \ldots, I_k\}$, de modo que la longitud de cada una de ellos sea menor que $\operatorname{dist}(A, B)$. La finalidad de hacer eso es poder garantizar que los intervalos I_i que intersecten, digamos al conjunto A, no intersecten a B y viceversa.

Teorema 6.2.15 (Carathéodory). *Si $\{A_1, \ldots, A_n\}$ es una **familia finita** de subconjuntos de $\mathbb{R}$ tal que $\operatorname{dist}(A_i, A_j) > 0$ para todo par de índices i, j con $i \neq j$, entonces*

$$\mu^*\left(\bigcup_{i=1}^{n} A_i\right) = \mu^*(A_1) + \cdots + \mu^*(A_n).$$

Prueba. Un fácil argumento inductivo permite reducir la prueba sólo a dos conjuntos. Suponga entonces que A y B son dos conjuntos tal que $\operatorname{dist}(A, B) > 0$. Por la subaditivad finita de μ^*, tenemos que

$$\mu^*(A \cup B) \leq \mu^*(A) + \mu^*(B).$$

Para demostrar la otra desigualdad, fijemos un $\varepsilon > 0$ elegido de modo arbitrario. Usemos (3_3) para seleccionar una sucesión $(I_n)_{n=1}^{\infty}$ de intervalos abiertos tal que

$$A \cup B \subseteq \bigcup_{n=1}^{\infty} I_n \qquad \text{y} \qquad \sum_{n=1}^{\infty} \ell(I_n) \leq \mu^*(A \cup B) + \varepsilon.$$

Dividiendo cada intervalo I_n en subintervalos más pequeños si fuese necesario, podemos suponer y, así lo haremos, que la longitud de cada I_n es menor que $\operatorname{dist}(A, B)$, es decir,

$$\ell(I_n) \ < \ \operatorname{dist}(A, B) \quad n \geq 1.$$

La finalidad de hacer esto, como ya lo mencionamos anteriormente, es poder garantizar que cada intervalo I_n intersecte o bien al conjunto A, o bien al conjunto B, pero no a ambos. Consideremos ahora los conjuntos de índices

$$F_1 \ = \ \{n \in \mathbb{N} \ : \ I_n \cap A \neq \varnothing\} \qquad \text{y} \qquad F_2 \ = \ \{n \in \mathbb{N} \ : \ I_n \cap B \neq \varnothing\}.$$

De acuerdo a nuestra suposición, tenemos que $F_1 \cap F_2 = \varnothing$, aunque pueden haber intervalos que no intersecten a ninguno de los dos conjuntos. Se sigue de esto que la sucesión $(I_n)_{n \in F_1}$ cubre a A, mientras que $(I_n)_{n \in F_2}$ cubre a B, es decir,

$$A \ \subseteq \ \bigcup_{n \in F_1} I_n \qquad \text{y} \qquad B \ \subseteq \ \bigcup_{n \in F_2} I_n.$$

Finalmente,

$$\mu^*(A) \ + \ \mu^*(B) \ \leq \ \sum_{n \in F_1} \ell(I_n) \ + \ \sum_{n \in F_2} \ell(I_n) \ \leq \ \sum_{n=1}^{\infty} \ell(I_n) \ \leq \ \mu^*(A \cup B) \ + \ \varepsilon.$$

Puesto que $\varepsilon > 0$ es arbitrario, se concluye que $\mu^*(A) + \mu^*(B) \leq \mu^*(A \cup B)$. Fin de la prueba. ■

Corolario 6.2.16. *Si $(K_i)_{i=1}^n$ es una sucesión finita de* ***subconjuntos compactos disjuntos dos a dos****, entonces*

$$\mu^*\left(\bigcup_{i=1}^n K_i\right) \ = \ \sum_{i=1}^n \mu^*(K_i).$$

Prueba. Observe, en primer lugar, que $\operatorname{dist}(K_i, K_j) > 0$ para todo $i \neq j$. En efecto, la definición de $\operatorname{dist}(K_i, K_j)$ nos garantiza que, para cada $k \in \mathbb{N}$, existen $x_k \in K_i$ y $y_k \in K_j$ tales que

$$|x_k - y_k| \ < \ \operatorname{dist}(K_i, K_j) \ + \ \frac{1}{k}$$

Usando ahora el hecho de que los conjuntos K_i son compactos, podemos seleccionar subsucesiones $(x_{k_l})_{l=1}^{\infty}$ y $(y_{k_l})_{l=1}^{\infty}$ de $(x_k)_{k=1}^{\infty}$ y $(y_k)_{k=1}^{\infty}$ respectivamente, tales que

$$\lim_{l \to \infty} x_{k_l} \ = \ x \in K_i \qquad \text{y} \qquad \lim_{l \to \infty} y_{k_l} \ = \ y \in K_j.$$

Puesto que $K_i \cap K_j = \varnothing$, se sigue que $x \neq y$ y, en consecuencia,

$$0 \ < \ |x - y| \ = \ \lim_{l \to \infty} |x_{k_l} - y_{k_l}| \ < \ \operatorname{dist}(K_i, K_j) \ + \ \lim_{l \to \infty} \frac{1}{k_l} \ = \ \operatorname{dist}(K_i, K_j).$$

El resultado sigue ahora del Teorema 6.2.15. ■

El siguiente resultado dice que para ciertos subconjuntos de $\mathbb{R}$ uno puede obtener igualdad en el Teorema 6.2.5.

Teorema 6.2.17. *Si $(I_n)_{n\in J}$ es una colección a lo más numerable de* ***intervalos compactos no-superpuestos*** *en $\mathbb{R}$, entonces*

$$\mu^*\Big(\bigcup_{n\in J} I_n\Big) = \sum_{n\in J}\mu^*(I_n).$$

Prueba. Fijemos $\varepsilon > 0$ y suponga, en primer lugar, que J es finito, digamos, $J = \{1,2,\dots,n\}$. Por la sub-aditividad de μ^*, tenemos que

$$\mu^*\Big(\bigcup_{i=1}^{n} I_i\Big) \leq \sum_{i=1}^{n}\mu^*(I_i) = \sum_{i=1}^{n}\ell(I_i). \tag{1}$$

Observe que si I_i es de la forma $I_i = [a_i, b_i]$ con $a_i < b_i$, entonces siempre podemos seleccionar un intervalo compacto, digamos K_i, de modo que $K_i \subseteq I_i$ y $\ell(K_i) > \ell(I_i) - \varepsilon/n$. En efecto, basta tomar $K_i = [a_i + \varepsilon/4n,\, b_i - \varepsilon/4n]$, para asegurarnos que

$$\ell(K_i) = b_i - a_i - \frac{\varepsilon}{2n} > \ell(I_i) - \frac{\varepsilon}{n}.$$

Observe que la colección de intervalos compactos $\{K_1,\dots,K_n\}$ son disjuntos dos a dos y, puesto que $\bigcup_{i=1}^n I_i \supseteq \bigcup_{i=1}^n K_i$, aplicando una vez más la sub-aditividad de μ^* y el Corolario 6.2.16, vemos que

$$\mu^*\Big(\bigcup_{i=1}^{n} I_i\Big) \geq \mu^*\Big(\bigcup_{i=1}^{n} K_i\Big) = \sum_{i=1}^{n}\mu^*(K_i) = \sum_{i=1}^{n}\ell(K_i) > \sum_{i=1}^{n}\ell(I_i) - \varepsilon.$$

Como ε es arbitrario, concluimos que

$$\mu^*\Big(\bigcup_{i=1}^{n} I_i\Big) \geq \sum_{i=1}^{n}\ell(I_i).$$

Esta última desigualdad combinada con (1) nos muestra que

$$\mu^*\Big(\bigcup_{i=1}^{n} I_i\Big) = \sum_{i=1}^{n}\ell(I_i).$$

Esto termina la prueba del caso finito. Para finalizar la demostración, suponga ahora que $J = \mathbb{N}$. Nótese que, por la sub-aditividad de μ^* y la primera parte, tenemos que la desigualdad

$$\mu^*\Big(\bigcup_{i=1}^{\infty} I_i\Big) \geq \mu^*\Big(\bigcup_{i=1}^{n} I_i\Big) = \sum_{i=1}^{n}\ell(I_i)$$

se cumple para todo $n \geq 1$. Por esto,

$$\sum_{i=1}^{\infty}\ell(I_i) = \lim_{n\to\infty}\sum_{i=1}^{n}\ell(I_i) \leq \mu^*\Big(\bigcup_{i=1}^{\infty} I_i\Big).$$

La desigualdad opuesta sigue del Teorema 6.2.5. ∎

Un poco más adelante, una vez definido lo que es un conjunto medible, daremos otra condición balo la cual una sucesión disjunta de conjuntos $(A_n)_{n=1}^{\infty}$ satisface

$$\mu^*\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \sum_{n=1}^{\infty}\mu^*(A_n).$$

6.2.2. Conjuntos de Contenido Cero

Existe una noción menos general que la de conjunto de medida exterior cero llamada conjunto de contenido cero en el sentido de Jordan que es de mucha utilidad para obtener una caracterización de las funciones que son Riemann integrables. Una motivación para su estudio es la noción de conjunto nulo la cual, como sabemos, se puede establecer en los siguientes términos: para cualquier subconjunto A de $\mathbb{R}$ las siguientes condiciones son equivalentes:

(1) $\mu^*(A) = 0$.

(2) *dado* $\varepsilon > 0$, *existe una sucesión* $(I_n)_{n=1}^{\infty}$ *de intervalos abiertos tal que*

$$A \subseteq \bigcup_{n=1}^{\infty} I_n \qquad \text{y} \qquad \sum_{n=1}^{\infty} \ell(I_n) < \varepsilon.$$

Si la colección de intervalos abiertos $\{I_n : n \in \mathbb{N}\}$ que cubre al conjunto A en (2) en dicha equivalencia se puede elegir siempre finita, entonces el conjunto A recibe un nombre muy especial.

Definición 6.2.18. *Un conjunto $A \subseteq \mathbb{R}$ se dice que tiene* **contenido cero** *o es de* ***medida cero en el sentido de Jordan*** *si, para cada $\varepsilon > 0$, existe una colección finita $\{I_1, \ldots, I_k\}$ de intervalos abiertos tal que*

$$A \subseteq \bigcup_{n=1}^{k} I_n \qquad y \qquad \sum_{n=1}^{k} \ell(I_k) < \varepsilon.$$

Si A tiene contenido cero, escribiremos $c_{\mu^*}(A) = 0$. Es claro que si $c_{\mu^*}(A) = 0$, entonces $\mu^*(A) = 0$. El recíproco se cumple si A es compacto.

Lema 6.2.19. *Sea A un subconjunto de $\mathbb{R}$. Si A es* ***compacto*** *y $\mu^*(A) = 0$, entonces $c_{\mu^*}(A) = 0$.*

Prueba. Suponga que A es un compacto tal que $\mu^*(A) = 0$. Sea $\varepsilon > 0$. Puesto que $\mu^*(A) = 0$, existe una sucesión $(I_n)_{n=1}^{\infty}$ de intervalos abiertos tal que

$$A \subseteq \bigcup_{n=1}^{\infty} I_n \qquad \text{y} \qquad \sum_{n=1}^{\infty} \ell(I_n) < \varepsilon.$$

Pero como A es compacto, el cubrimiento $\mathcal{V} = \{I_n : n \in \mathbb{N}\}$ se reduce a un cubrimiento finito, esto es, existe una colección finita $\{I_1, \ldots, I_k\}$ en $\mathcal{V}$ tal que

$$A \subseteq \bigcup_{n=1}^{k} I_n \qquad \text{y} \qquad \sum_{n=1}^{k} \ell(I_n) \leq \sum_{n=1}^{\infty} \ell(I_n) < \varepsilon.$$

Esto prueba que $c_{\mu^*}(A) = 0$. ■

El siguiente resultado establece que en la definición de conjuntos de contenido cero uno puede elegir intervalos cerrados en lugar de intervalos abiertos.

Lema 6.2.20. *Un conjunto A es de* ***contenido cero*** *si, y sólo si, para cada $\varepsilon > 0$, existe una colección finita $\{J_1, \ldots, J_k\}$ de* ***intervalos cerrados*** *tal que*

$$A \subseteq \bigcup_{n=1}^{k} J_n \qquad y \qquad \sum_{n=1}^{k} \ell(J_k) < \varepsilon, .$$

Prueba. ($\Rightarrow$) Suponga que A es de contenido cero y sea $\varepsilon > 0$. Entonces existe una colección finita $\{I_1, \ldots, I_k\}$ de intervalos abiertos tal que

$$A \subseteq \bigcup_{n=1}^{k} I_n \qquad \text{y} \qquad \sum_{n=1}^{k} \ell(I_k) < \varepsilon,$$

Haciendo $J_i = \overline{I_i}$ para cada índice i y teniendo en cuenta que $A \subseteq \bigcup_{n=1}^{k} J_n$ y que $\ell(I_i) = \ell(J_i)$ se termina la prueba de esta implicación.

($\Leftarrow$) Sea $\varepsilon > 0$ y suponga que existe una colección finita $\{J_1, \ldots, J_k\}$ de intervalos cerrados tal que

$$A \subseteq \bigcup_{n=1}^{k} J_n \qquad \text{y} \qquad \sum_{n=1}^{k} \ell(J_k) < \varepsilon/2.$$

Pongamos $J_i = [u_i, v_i]$, $i = 1, 2, \ldots, k$ y considere los intervalos abiertos

$$I_i = \big(u_i - \varepsilon/4k,\, v_i + \varepsilon/4k\big), \quad i = 1, 2, \ldots, k.$$

Entonces $A \subseteq \bigcup_{n=1}^{k} I_n$ y

$$\sum_{i=1}^{k} \ell(I_i) = \sum_{i=1}^{k} \left((v_i - u_i) + \frac{\varepsilon}{2k} \right) < \frac{\varepsilon}{2} + k \cdot \frac{\varepsilon}{2k} = \varepsilon.$$

La aprueba es completa. ■

El lema anterior nos indica que A *es de contenido cero si, y sólo si,* $\overline{A}$*, la clausura de* A*, es de contenido cero.* Esto último, por supuesto, no se cumple para conjuntos de medida exterior cero. En efecto, el conjunto $\mathbb{Q}$ es de medida exterior cero, pero $\overline{\mathbb{Q}} = \mathbb{R}$ no lo es. Resulta claro que cualquier conjunto finito es de contenido cero, aunque existen conjuntos infinitos numerables que pueden o no tener contenido cero. Por ejemplo, del resultado anterior se sigue que $\mathbb{Q}$ es un conjunto numerable que no tiene contenido cero. Sin embargo, el conjunto $A = \{1/n : n \geq 1\}$ es tal que $c_{\mu^*}(A) = 0$. Para ver esto último, sea $\varepsilon > 0$ y escoja un $n \in \mathbb{N}$ tal que $\frac{1}{n} < \frac{\varepsilon}{4} < \frac{1}{n-1}$. Si ahora definimos

$$J_1 = \left(\frac{-\varepsilon}{4}, \frac{\varepsilon}{4} \right) \qquad \text{y} \qquad J_k = \left(\frac{1}{k-1} - \frac{\varepsilon}{4n}, \frac{1}{k-1} + \frac{\varepsilon}{4n} \right)$$

para $k = 2, 3, \ldots, n$, entonces $A \subseteq \bigcup_{k=1}^{n} J_k$ y

$$\sum_{k=1}^{n} \ell(J_k) = \frac{\varepsilon}{2} + (n-1)\frac{\varepsilon}{2n} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Por esto, A es de contenido cero. Por otro lado, el conjunto ternario de Cantor Γ es un conjunto no-numerable que es de contenido cero. En efecto, recordemos que $\Gamma \subseteq \Gamma_n$ para todo $n \geq 1$, donde Γ_n consiste de la unión de los 2^n intervalos cerrados y acotados que no fueron eliminados en el n-ésimo paso en la construcción de Γ y cuya longitud es $(2/3)^n$. Si fijamos un $\varepsilon > 0$ arbitrario y elegimos un $n \in \mathbb{N}$ de modo que $(2/3)^n < \varepsilon$, vemos que $c_{\mu^*}(\Gamma) = 0$.

Teorema 6.2.21. *Si $a < b$, entonces el intervalo $[a, b] \subseteq \mathbb{R}$ **no tiene contenido cero**. De hecho, si $\{I_1, \ldots, I_n\}$ es un cubrimiento finito de $[a, b]$ por intervalos cerrados, entonces*

$$\sum_{i=1}^{n} \ell(I_i) \ \geq \ b - a.$$

Prueba. La demostración es por inducción sobre el número de intervalos que cubren a $[a, b]$. El caso $n = 1$ es inmediato. Suponga que el teorema es cierto para recubrimientos con n intervalos y sea $\{I_1, \ldots, I_{n+1}\}$ un cubrimiento de $[a, b]$ por intervalos cerrados. Se puede suponer (cambiando los subíndices si es necesario) que $a \in I_1$. Pongamos $I_1 = [\alpha, \beta]$. Entonces $\alpha \leq a \leq \beta$. Si $\alpha \geq b$, entonces $\ell(I_1) \geq b - a$ y, por consiguiente, $\sum_{i=1}^{n} \ell(I_i) \geq b - a$. Si $\beta < b$, entonces $[\beta, b]$ es recubierto por $\{I_2, \ldots, I_{n+1}\}$. Por lo tanto, $\ell(I_2) + \cdots + \ell(I_{n+1}) \geq b - \beta$ y, en consecuencia,

$$\sum_{i=1}^{n} \ell(I_i) \ \geq \ (\beta - a) + (b - \beta) \ = \ b - a.$$

La prueba es completa. ■

Corolario 6.2.22. *Sea $(A_n)_{n=1}^{\infty}$ una sucesión de subconjuntos **compactos** de $\mathbb{R}$. Las siguientes condiciones son equivalentes:*

(1) *A_n tiene **contenido cero** para todo $n \geq 1$.*

(2) *$\bigcup_{n=1}^{\infty} A_n$ tiene **medida cero**.*

Prueba. (1) $\Rightarrow$ (2). Suponga que $c_{\mu^*}(A_n) = 0$ para todo $n \geq 1$. Entonces $\mu^*(A_n) = 0$ para todo $n \in \mathbb{N}$ y se sigue del Teorema 6.2.5 que $\mu^*(\bigcup_{n=1}^{\infty} A_n) = 0$. Esto prueba que $\bigcup_{n=1}^{\infty} A_n$ tiene medida cero.

(2) $\Rightarrow$ (1). Suponga que (2) se cumple. Entonces, por ser μ^* monótona, se tiene que para cada $k \geq 1$, $\mu^*(A_k) \leq \mu^*(\bigcup_{n=1}^{\infty} A_n) = 0$, de donde se sigue que $c_{\mu^*}(A_k) = 0$ para todo $k \geq 1$ gracias al Lema 6.2.19. ■

6.3. La Medida de Lebesgue

El objetivo de esta sección es construir una clase muy especial de subconjuntos de $\mathbb{R}$, llamada la σ-álgebra de Lebesgue, sobre la cual la medida exterior de Lebesgue es numerablemente aditiva y derivar algunos resultados fundamentales con dicha "medida".

6.3.1. La σ-álgebra de Lebesgue

La imposibilidad de la medida exterior μ^* de ser numerablemente aditiva, hecho que probaremos al final de este capítulo, deja hasta ahora el trabajo incompleto en el sentido de obtener una "*medida*" que generalize la noción de longitud, mida cada uno de los subconjuntos de $\mathbb{R}$, sea invariante por traslación y numerablemente aditiva. Un modo inteligente de salir de este impasse es considerar, en lugar de la clase $\mathscr{P}(\mathbb{R})$, una sub-clase mucho más pequeña, como por ejemplo: *la sub-clase de todos aquellos subconjuntos de $\mathbb{R}$ para los cuales μ^* sea numerablemente aditiva*. Denotemos, por el momento, a dicha sub-clase por $\mathfrak{M}$. Aunque la idea es, en principio, muy buena, persisten, sin embargo, algunos problemas que no son de solución inmediata. Por ejemplo, ¿cómo identificar a los elementos de esa sub-clase?, es decir, ¿cómo se determina si un determinado conjunto está o no en la clase $\mathfrak{M}$? ¿Qué tipos de operaciones de conjuntos son permitidas dentro de $\mathfrak{M}$? ¿Qué tan grande es $\mathfrak{M}$?, esto es, ¿cuál es su cardinalidad? ¿Cuál es su utilidad?, etc. Intentaremos, en lo que sigue, abordar una definición adecuada de ciertos conjuntos muy especiales que permitan que μ^* sea numerablemente aditiva para tales conjuntos y luego darle respuestas a las interrogantes antes formuladas.

¿Pero, por dónde debemos comenzar? Recordemos que el Teorema 6.2.12 nos garantiza que, dado cualquier subconjunto A de $\mathbb{R}$ y cualquier $\varepsilon > 0$, podemos determinar la existencia de un subconjunto abierto $G \supseteq A$ tal que

$$\mu^*(G) \ < \ \mu^*(A) \ + \ \varepsilon. \tag{$\star$}$$

Lo que tenemos en mente es investigar qué tipos de conjuntos se pueden *aproximar* por conjuntos abiertos, es decir, ¿es posible elegir a G de modo que $\mu^*(G \setminus A) < \varepsilon$? Suponga, por un momento, que $\mu^*(A) < +\infty$ y observemos que de la desigualdad $(\star)$ se tiene que

$$\mu^*(G) - \mu^*(A) \ < \ \varepsilon.$$

Sin embrago, a pesar de que $G = A \cup \big(G \setminus A\big)$ es una unión disjunta, la sub-aditividad finita de μ^* lo único que puede asegurar es que

$$\mu^*(G) \ \leq \ \mu^*(A) \ + \ \mu^*(G \setminus A),$$

y, por consiguiente,

$$\mu^*(G) \ - \ \mu^*(A) \ \leq \ \mu^*(G \setminus A).$$

No existe, en consecuencia, argumento alguno que conduzca a la igualdad

$$\mu^*(G \setminus A) \ = \ \mu^*(G) \ - \ \mu^*(A)$$

y, en consecuencia, poder asegurar, usando $(\star)$, que

$$\mu^*(G \setminus A) \ < \ \varepsilon. \tag{$\star\star$}$$

¿Por qué esto es importante? Pues bien, lo que veremos en lo inmediato es que la colección $\mathfrak{M}$ de todos aquellos conjuntos A para los cuales $(\star\star)$ se cumple para algún abierto $G \supseteq A$ es una muy buena elección en el sentido de que dicha colección es una σ-álgebra, tiene cardinalidad $2^{\mathfrak{c}}$ y μ^* es numerablemente aditiva sobre $\mathfrak{M}$. Esto pequeño análisis permite justificar la siguiente definición.

Definición 6.3.1. *Un conjunto $E \subseteq \mathbb{R}$ se dice que es* ***medible según Lebesgue****, o simplemente* ***medible*** *si, para cada $\varepsilon > 0$, existe un conjunto abierto $G \supseteq E$ tal que*

$$\mu^*(G \setminus E) \ \leq \ \varepsilon.$$

Con frecuencia escribiremos "conjunto medible" en lugar de "conjunto medible según Lebesgue". Denotaremos por $\mathfrak{M}_\mu(\mathbb{R})$ *a la familia de todos los subconjuntos de* $\mathbb{R}$ *que son* ***medibles según Lebesgue***, esto es,

$$\mathfrak{M}_\mu(\mathbb{R}) \ = \ \big\{E \subseteq \mathbb{R} \ : \ E \text{ es medible según Lebesgue}\big\}.$$

Observe que

$Z \notin \mathfrak{M}_\mu(\mathbb{R})$ significa que existe un $\varepsilon > 0$ con la siguiente propiedad: para cada conjunto abierto $G \supseteq Z$ se cumple que $\mu^*(G \setminus Z) > \varepsilon$.

Suponga, además, que $\mu^*(Z) < +\infty$ y seleccione una sucesión $(I_n)_{n=1}^{\infty}$ de intervalos abiertos tal que

$$Z \subseteq \bigcup_{n=1}^{\infty} I_n \qquad \text{y} \qquad \sum_{n=1}^{\infty} \ell(I_n) < \mu^*(Z) + \varepsilon.$$

Pongamos $G = \bigcup_{n=1}^{\infty} I_n$. Resulta que G es un conjunto abierto conteniendo a Z y, por lo anterior, $\mu^*(G) < \mu^*(Z) + \varepsilon$, es decir, $\mu^*(G) - \mu^*(Z) < \varepsilon$. Por esto,

$$\mu^*(G \setminus Z) > \varepsilon > \mu^*(G) - \mu^*(Z).$$

Esta observación permite señalar que *sólamente los conjuntos medibles* pueden garantizar la validez de la desigualdad $\mu^*(G \setminus E) \leq \varepsilon$.

¿Cuáles conjuntos son medibles? Para verificar que la colección $\mathfrak{M}_\mu(\mathbb{R})$ es una buena elección debemos comprobar que ella es no vacía y contiene a casi todos conjuntos que son de utilidad en el Análisis. Lo primero que debemos destacar es que $\mathfrak{M}_\mu(\mathbb{R})$ es, de hecho, una colección muy grande. En efecto, si $\mathcal{O}_\mathbb{R}$ denota *la colección de todos los **subconjuntos abiertos** de* $\mathbb{R}$, entonces

(M_1) $\mathcal{O}_\mathbb{R} \subseteq \mathfrak{M}_\mu(\mathbb{R})$.

En efecto, si A es un *subconjunto abierto* de $\mathbb{R}$ y si $\varepsilon > 0$, entonces tomando $G = A$, resulta que $G \supseteq A$ y, por consiguiente, $\mu^*(G \setminus G) = 0 < \varepsilon$. En particular,

$$\mathbb{R},\ \varnothing \in \mathfrak{M}_\mu(\mathbb{R})$$

con lo cual queda establecido que $\mathfrak{M}_\mu(\mathbb{R})$ satisface (σ_1).

(M_2) $\operatorname{card}\big(\mathfrak{M}_\mu(\mathbb{R})\big) = 2^{\mathfrak{c}}$.

El conjunto ternario de Cantor Γ, como se demuestra en el Corolario 6.3.41, página 283, es un conjunto nulo y, en consecuencia, $\mathscr{P}(\Gamma) \subseteq \mathfrak{M}_\mu(\mathbb{R})$ (véase el Teorema 6.3.2). Puesto que $\operatorname{card}(\Gamma) = \mathfrak{c}$, resulta que

$$\operatorname{card}\big(\mathfrak{M}_\mu(\mathbb{R})\big) \geq \operatorname{card}(\mathscr{P}(\Gamma)) = 2^{\mathfrak{c}}.$$

Por otro lado, como $\mathfrak{M}_\mu(\mathbb{R}) \subseteq \mathscr{P}(\mathbb{R})$ y la cardinalidad de $\mathscr{P}(\mathbb{R})$ es $2^{\mathfrak{c}}$, se tiene que $\operatorname{card}\big(\mathfrak{M}_\mu(\mathbb{R})\big) \leq 2^{\mathfrak{c}}$ y, así

$$\operatorname{card}\big(\mathfrak{M}_\mu(\mathbb{R})\big) = 2^{\mathfrak{c}}.$$

Si bien es cierto que hay tantos conjuntos medibles como subconjuntos de $\mathbb{R}$, existen, sin embargo, elementos de $\mathscr{P}(\mathbb{R})$ que no pertenecen a $\mathfrak{M}_\mu(\mathbb{R})$.

(M_3) $\mathfrak{M}_\mu(\mathbb{R}) \subsetneqq \mathscr{P}(\mathbb{R})$.

Sólo resta encontrar un subconjunto de $\mathbb{R}$ que no sea medible. Esto hecho, sin embargo, tendrá que esperar hasta que lleguemos a la Sección 6.5.1, página 310. Allí se demuestra que el conjunto

$$\neg\,\mathfrak{M}_\mu(\mathbb{R}) = \mathscr{P}(\mathbb{R}) \setminus \mathfrak{M}_\mu(\mathbb{R})$$

es no vacío. Cualquier elemento de $\neg\,\mathfrak{M}_\mu(\mathbb{R})$ será llamado un ***conjunto no-medible***.

El siguiente resultado, el cual es fundamental en la Teoría de la Medida, establece que *todo conjunto nulo es medible*.

Teorema 6.3.2. *Si* $E \subseteq \mathbb{R}$ *y* $\mu^*(E) = 0$, *entonces* E *es* ***medible*** *según Lebesgue.*

Prueba. Suponga que $\mu^*(E) = 0$ y sea $\varepsilon > 0$. Por el Teorema 6.2.12 existe un conjunto abierto $G \supseteq E$ tal que

$$\mu^*(G) \leq \mu^*(E) + \varepsilon = \varepsilon.$$

Puesto que $G \setminus E \subseteq G$, resulta, por la monotonía de μ^*, que

$$\mu^*(G \setminus E) \leq \mu^*(G) \leq \varepsilon.$$

Esto prueba que $E \in \mathfrak{M}_\mu(\mathbb{R})$. ■

Del Teorema 6.3.2 y el Corolario 6.2.6 se concluye que:

(a) *Cualquier* ***subconjunto a lo más numerable*** *es* ***medible***. En particular, los conjuntos $\mathbb{N}$, $\mathbb{Z}$ y $\mathbb{Q}$ son medibles.

(b) Γ, *el* ***conjunto ternario de Cantor***, *aunque es no-numerable, también es* ***medible*** *pues* $\mu^*(\Gamma) = 0$. (Véase el Corolario 6.3.41, página 283).

(c) *En general, cualquier* ***subconjunto*** *de un* ***conjunto nulo*** *es* ***medible***, esto es,

$$\mu^*(A) = 0 \quad \text{y} \quad E \subseteq A \quad \Rightarrow \quad E \in \mathfrak{M}_\mu(\mathbb{R}),$$

lo cual es equivalente a expresarlo de este otro modo:

$$\mu^*(A) = 0 \quad \Rightarrow \quad \mathscr{P}(A) \subseteq \mathfrak{M}_\mu(\mathbb{R}).$$

Esta propiedad de la medida exterior de Lebesgue recibe el nombre de **completitud** y se dice entonces que μ^* es una ***medida completa*** cuando ella se restringe a la clase $\mathfrak{M}_\mu(\mathbb{R})$. Es importante advertir que: *no es cierto que cualquier subconjunto de un conjunto medible sea medible.*

Nuestra siguiente misión es escudriñar cuáles otras propiedades (importantes) posee la familia $\mathfrak{M}_\mu(\mathbb{R})$. Los siguientes resultados van a mostrar que $\mathfrak{M}_\mu(\mathbb{R})$ es, en realidad, una σ**-álgebra**, es decir, una clase de subconjuntos de $\mathbb{R}$ que es cerrada bajo complementos y uniones numerables, además de verificar otras propiedades muy especiales. En general, cualquier σ-álgebra sirve para definir la colección de los conjuntos que queremos "medir".

Definición 6.3.3. *Una familia no vacía* $\mathcal{M}$ *de subconjuntos de un conjunto no vacío* X *se dice que es una* σ***-álgebra*** *de* X *si:*

(σ_1) $X \in \mathcal{M}$,

(σ_2) $X \setminus E \in \mathcal{M}$ *siempre que* $E \in \mathcal{M}$, *y*

(σ_1) $\bigcup_{n=1}^{\infty} E_n \in \mathcal{M}$ *para cualquier colección numerable* $\{E_n : n = 1, 2, \ldots\} \subseteq \mathcal{M}$.

Si las condiciones (σ_1) y (σ_2) en la definición anterior se mantienen pero (σ_3) se sustituye por la siguiente:

(σ_4) $\bigcup_{i=1}^{n} E_i \in \mathcal{M}$ para cualquier colección finita $\{E_i : i = 1, 2, \ldots, n\} \subseteq \mathcal{M}$,

entonces diremos que $\mathcal{M}$ es una **álgebra de conjuntos**.

Observe que cualquier σ-álgebra es una álgebra pero no recíprocamente. Veamos ahora que la colección $\mathfrak{M}_\mu(\mathbb{R})$ es cerrada bajo uniones numerables, es decir, satisface (σ_3).

Teorema 6.3.4. *Si $(E_n)_{n=1}^{\infty}$ es una sucesión en $\mathfrak{M}_\mu(\mathbb{R})$, entonces $\bigcup_{n=1}^{\infty} E_n \in \mathfrak{M}_\mu(\mathbb{R})$.*

Prueba. Fijemos $\varepsilon > 0$. Para cada $n \in \mathbb{N}$ podemos seleccionar, usando el hecho de que E_n es medible, un conjunto abierto $G_n \supseteq E_n$ de modo tal que

$$\mu^*(G_n \setminus E_n) \;\leq\; \frac{\varepsilon}{2^n}.$$

Tomando $G = \bigcup_{n=1}^{\infty} G_n$, resulta que G es abierto, $G \supseteq \bigcup_{n=1}^{\infty} E_n$, y

$$G \setminus \bigcup_{n=1}^{\infty} E_n \;=\; \left(\bigcup_{n=1}^{\infty} G_n\right) \setminus \left(\bigcup_{n=1}^{\infty} E_n\right) \;\subseteq\; \bigcup_{n=1}^{\infty} (G_n \setminus E_n).$$

Por esto y la sub-aditividad numerable de μ^* tenemos que

$$\mu^*\left(G \setminus \bigcup_{n=1}^{\infty} E_n\right) \;\leq\; \sum_{n=1}^{\infty} \mu^*(G_n \setminus E_n) \;\leq\; \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} \;=\; \varepsilon.$$

Esto termina la prueba de que $\bigcup_{n=1}^{\infty} E_n$ es medible. ■

Para verificar que $\mathfrak{M}_\mu(\mathbb{R})$ es una σ-álgebra sólo nos resta demostrar (σ_2). Esto requiere un poquito de trabajo. Comencemos.

Lema 6.3.5. *Cualquier conjunto **compacto** $K \subseteq \mathbb{R}$ es **medible según Lebesgue**.*

Prueba. En primer lugar, nótese que como K es compacto, podemos encontrar un intervalo cerrado y acotado I de modo que $K \subseteq I$ y así, $\mu^*(K) \leq \ell(I) < +\infty$.

Fijemos $\varepsilon > 0$. Por el Teorema 6.2.12 existe un conjunto abierto, necesariamente acotado, $G \supseteq K$ tal que

$$\mu^*(G) \;\leq\; \mu^*(K) \;+\; \varepsilon.$$

Para demostrar que $\mu^*(G \setminus K) \leq \varepsilon$ vamos a proceder del modo siguiente. Siendo $G \setminus K$ un conjunto abierto, se sigue del Lema 2.2.10, página 124, que dicho conjunto se puede escribir en la forma

$$G \setminus K \;=\; \bigcup_{n=1}^{\infty} I_n,$$

donde $(I_n)_{n=1}^{\infty}$ es una sucesión de intervalos compactos no-superpuestos. Para cualquier $k \in \mathbb{N}$, tenemos que $\bigcup_{n=1}^{k} I_k \subseteq G \setminus K \subseteq G$ y, por lo tanto, $K \cup \bigcup_{n=1}^{k} I_n \subseteq G$. Más aun, los conjuntos K

y $\bigcup_{n=1}^{k} I_k$ son compactos y disjuntos. El Corolario 6.2.16 combinado con el Teorema 6.2.17 nos garantizan que

$$\mu^*(K) + \sum_{n=1}^{k} \ell(I_n) = \mu^*(K) + \mu^*\Big(\bigcup_{n=1}^{k} I_n\Big) = \mu^*\Big(K \cup \bigcup_{n=1}^{k} I_n\Big) \leq \mu^*(G).$$

Puesto que $\mu^*(K) < +\infty$, la anterior desigualdad implica que

$$\sum_{n=1}^{k} \ell(I_n) \leq \mu^*(G) - \mu^*(K) \leq \varepsilon,$$

y como $k \geq 1$ es arbitrario, concluimos de la desigualdad anterior que la serie $\sum_{n=1}^{\infty} \ell(I_n)$ converge y que $\sum_{n=1}^{\infty} \ell(I_n) \leq \varepsilon$. De nuevo, por el Teorema 6.2.17, se sigue que

$$\mu^*(G \setminus K) = \sum_{n=1}^{\infty} \ell(I_n) \leq \varepsilon.$$

Esto nos dice que K es medible y finaliza la prueba. ■

Corolario 6.3.6. *Todo subconjunto **cerrado** de $\mathbb{R}$ es **medible según Lebesgue**.*

Prueba. Suponga que $F \subseteq \mathbb{R}$ es cerrado. Como $\mathbb{R} = \bigcup_{n=1}^{\infty}[-n, n]$, entonces a F lo podemos reescribir en la forma

$$F = F \cap \mathbb{R} = \bigcup_{n=1}^{\infty} \big(F \cap [-n, n]\big).$$

Ya que cada $F \cap [-n, n]$ es un subconjunto compacto, el Lema 6.3.5 nos revela que él es medible y, así, por el Teorema 6.3.4, se concluye que F es medible. ■

Estamos ahora en posesión de los argumentos para demostrar que el complemento de cualquier conjunto medible es medible.

Teorema 6.3.7. *Si $E \in \mathfrak{M}_\mu(\mathbb{R})$, entonces $E^c = \mathbb{R} \setminus E \in \mathfrak{M}_\mu(\mathbb{R})$.*

Prueba. Suponga que $E \in \mathfrak{M}_\mu(\mathbb{R})$ y para cada $n \in \mathbb{N}$, escojamos un conjunto abierto $G_n \supseteq E$ tal que $\mu^*(G_n \setminus E) < 1/n$. Defina $F_n = G_n^c$. Entonces F_n es cerrado y, en consecuencia, medible gracias al Corolario 6.3.6. Sea

$$F = \bigcup_{n=1}^{\infty} F_n = \bigcup_{n=1}^{\infty} G_n^c.$$

El Teorema 6.3.2 garantiza que F es medible y, además, $F \subseteq E^c$. Sea $Z = E^c \setminus F$. Fijando cualquier $k \in \mathbb{N}$ tenemos que

$$Z = E^c \setminus \bigcup_{n=1}^{\infty} G_n^c \subseteq E^c \setminus G_k^c = G_k \setminus E,$$

de donde se tiene que,

$$\mu^*(Z) \leq \mu^*\big((G_k \setminus E)\big) < \frac{1}{k},$$

para todo $k \in \mathbb{N}$. De aquí se sigue que $\mu^*(Z) = 0$, y entonces, por el Teorema 6.3.2, Z es medible. Por esto, $E^c = F \cup Z$ también es medible y termina la prueba. ■

Con el resultado anterior hemos finalizado la prueba de que:

Corolario 6.3.8. *$\mathfrak{M}_\mu(\mathbb{R})$ es una σ-**álgebra**.*

A $\mathfrak{M}_\mu(\mathbb{R})$ la llamaremos la **σ-álgebra de Lebesgue**. Cuando no exista ninguna ocasión de producir una confusión, a los elementos de $\mathfrak{M}_\mu(\mathbb{R})$ los llamaremos simplemente **conjuntos medibles**. Varias consecuencias se derivan inmediatamente a partir de la información de que $\mathfrak{M}_\mu(\mathbb{R})$ es una σ-álgebra. Por ejemplo,

(β_1) $\bigcap_{n=1}^{\infty} E_n \in \mathfrak{M}_\mu(\mathbb{R})$ para cualquier sucesión $(E_n)_{n=1}^{\infty}$ incluida en $\mathfrak{M}_\mu(\mathbb{R})$.

En efecto, como $E_n \in \mathfrak{M}_\mu(\mathbb{R})$ para todo $n \in \mathbb{N}$, entonces $E_n^c \in \mathfrak{M}_\mu(\mathbb{R})$ y, por lo tanto, $\bigcup_{n=1}^{\infty} E_n^c \in \mathfrak{M}_\mu(\mathbb{R})$. De las Leyes de Morgan se sigue que

$$\bigcap_{n=1}^{\infty} E_n = \left(\bigcup_{n=1}^{\infty} E_n^c\right)^c \in \mathfrak{M}_\mu(\mathbb{R}).$$

En particular,

$$\bigcup_{n=1}^{m} E_n \in \mathfrak{M}_\mu(\mathbb{R}) \qquad \text{y} \qquad \bigcap_{n=1}^{m} E_n \in \mathfrak{M}_\mu(\mathbb{R})$$

cualesquiera sean $E_1, \ldots, E_m$ en $\mathfrak{M}_\mu(\mathbb{R})$.

(β_2) Si $\mathcal{G}_\delta(\mathbb{R})$ y $\mathcal{F}_\sigma(\mathbb{R})$ denotan a las familias de todos los conjuntos G_δ y F_σ de $\mathbb{R}$, respectivamente, entonces

$$\mathcal{G}_\delta(\mathbb{R}) \subseteq \mathfrak{M}_\mu(\mathbb{R}) \qquad \text{y} \qquad \mathcal{F}_\sigma(\mathbb{R}) \subseteq \mathfrak{M}_\mu(\mathbb{R}).$$

En general, cada una de las familias

$$\mathcal{G}_{\delta\sigma}(\mathbb{R}) = (\mathcal{G}_\delta)_\sigma(\mathbb{R}), \quad \mathcal{F}_{\sigma\delta}(\mathbb{R}) = (\mathcal{F}_\sigma)_\delta(\mathbb{R}), \quad \mathcal{G}_{\delta\sigma\delta}(\mathbb{R}), \quad \mathcal{F}_{\sigma\delta\sigma}(\mathbb{R}), \ldots$$

están en $\mathfrak{M}_\mu(\mathbb{R})$, donde cada elemento de $\mathcal{G}_{\delta\sigma}(\mathbb{R})$ es una unión de elementos de $\mathcal{G}_\delta(\mathbb{R})$, similarmente, todo elemento de $\mathcal{F}_{\sigma\delta}(\mathbb{R})$ es una intersección de elementos de $\mathcal{F}_\sigma(\mathbb{R})$, etcétera.

(β_3) Sea $(E_n)_{n=1}^{\infty}$ una sucesión en $\mathfrak{M}_\mu(\mathbb{R})$. Entonces existe una **sucesión disjunta** $(F_n)_{n=1}^{\infty}$ en $\mathfrak{M}_\mu(\mathbb{R})$ tal que $F_n \subseteq E_n$ para todo $n \geq 1$ y

$$\bigcup_{n=1}^{\infty} E_n = \bigcup_{n=1}^{\infty} F_n.$$

En efecto, pongamos $F_1 = E_1$ y para $n \geq 2$, defina

$$F_n = E_n \setminus \bigcup_{j=1}^{n-1} E_j.$$

Es claro que la sucesión $(F_n)_{n=1}^{\infty}$ satisface las propiedades requeridas.

(β_4) $(E_n)_{n=1}^{\infty}$ es una sucesión de **conjuntos nulos**, entonces $\bigcup_{n=1}^{\infty} E_n$ es un **conjunto nulo**. Además,

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n)$$

sin exigir que ellos sean disjuntos. Una consecuencia inmediata de (β_4) es que $\mathbb{R}$ *no se puede descomponer como una unión infinita numerable de conjuntos numerables*.

En lo que sigue, si $X \in \mathfrak{M}_\mu(\mathbb{R})$, escribiremos

$$\mathfrak{M}_\mu(X) = \mathfrak{M}_\mu(\mathbb{R}) \cap X = \{B \cap X : B \in \mathfrak{M}_\mu(\mathbb{R})\} = \{E \in \mathfrak{M}_\mu(\mathbb{R}) : E \subseteq X\}.$$

Es fácil ver que $\mathfrak{M}_\mu(X)$ es una σ-álgebra de subconjuntos de X a la que llamaremos la **σ-álgebra de Lebesgue inducida** por X o la **σ-álgebra de Lebesgue de** X.

Ya hemos visto que todo subconjunto de $\mathbb{R}$ se puede aproximar por conjuntos abiertos. Una consecuencia importante del Teorema 6.3.7 es el siguiente resultado, el cual establece que cada conjunto medible se puede aproximar por un algún conjunto cerrado.

Corolario 6.3.9. *Sea $E \in \mathfrak{M}_\mu(\mathbb{R})$. Para cada $\varepsilon > 0$, existe un **subconjunto cerrado** $F \subseteq E$ tal que $\mu^*(E \setminus F) < \varepsilon$.*

Prueba. Por el Teorema 6.3.7, $E^c \in \mathfrak{M}_\mu(\mathbb{R})$ y, por lo tanto, de la definición de conjunto medible, existe un conjunto abierto $G \supseteq E^c$ tal que $\mu^*(G \setminus E^c) < \varepsilon$. Tomando $F = G^c$, resulta que F es cerrado, $F \subseteq E$ y, además, como $E \setminus F \subseteq G \setminus E^c$, se tiene entonces que $\mu^*(E \setminus F) < \varepsilon$. ■

El siguiente resultado combinado con el corolario anterior permitirá, en el caso de conjuntos medibles de medida finita, aproximar tales conjuntos por conjuntos compactos.

Lema 6.3.10. *Sea $A \subseteq \mathbb{R}$ con $\mu^*(A) < +\infty$. Entonces, para cada $\varepsilon > 0$, existe un **conjunto acotado** $A_0 \subseteq A$ tal que $\mu^*(A \setminus A_0) < \varepsilon$. Además, si A es **medible**, entonces A_0 también se puede elegir medible.*

Prueba. Sea $\varepsilon > 0$. Por el Teorema 6.2.12 existe un conjunto abierto G conteniendo a A tal que $\mu^*(G) \leq \mu^*(A) + 1 < +\infty$. Un llamado al Teorema 2.2.4 nos permite escribir a G en la forma $G = \bigcup_{n=1}^{\infty} I_n$, donde los I_n son intervalos abiertos y disjuntos dos a dos. El Teorema 6.2.17 nos dice que $\sum_{n=1}^{\infty} \ell(I_n) = \mu^*(G) < +\infty$ y, por lo tanto, podemos elegir un $N \in \mathbb{N}$ lo suficientemente grande de modo que $\sum_{n=N+1}^{\infty} \ell(I_n) < \varepsilon$. Sea ahora

$$A_0 = A \cap \left(\bigcup_{n=1}^{N} I_n\right).$$

Tenemos que $A_0 \subset A$ y como todos los intervalos I_n son de longitud finita, resulta que A_0 es acotado. Nótese que si A es medible, también lo es A_0. Finalmente, como $A \subseteq \bigcup_{n=1}^{\infty} I_n$, se sigue que $A \setminus A_0 \subseteq \bigcup_{n=N+1}^{\infty} I_n$ y, por lo tanto,

$$\mu^*(A \setminus A_0) \leq \mu^*\left(\bigcup_{n=N+1}^{\infty} I_n\right) \leq \sum_{n=N+1}^{\infty} \ell(I_n) < \varepsilon.$$

Esto termina la prueba. ■

6.3.2. La σ-álgebra de Borel

Existe otra σ-álgebra, conocida como la σ-álgebra de Borel, que juega un papel de primer orden en la Teoría de la Medida sobre conjuntos con una estructura topológica. Observe que si X es un conjunto arbitrario, entonces $\mathscr{P}(X)$ y $\{\varnothing, X\}$ son σ-álgebras. Cualquier otra σ-álgebra $\mathcal{A}$ de X verifica que

$$\{\varnothing, X\} \subseteq \mathcal{A} \subseteq \mathscr{P}(X).$$

Un procedimiento de carácter general que permite construir σ-álgebras partiendo de una clase arbitraria de subconjuntos de un conjunto dado es el siguiente.

Teorema 6.3.11 (σ-álgebra generada). *Sea X un conjunto no vacío y sea $\mathcal{C}$ una colección arbitraria no vacía de subconjuntos de X. Entonces* ***existe una única σ-álgebra*** *$\sigma(\mathcal{C})$ sobre X que verifica lo siguiente:*

(a) $\mathcal{C} \subseteq \sigma(\mathcal{C})$, y

(b) Si $\mathcal{B}$ es otra σ-álgebra sobre X conteniendo a $\mathcal{C}$, entonces $\sigma(\mathcal{C}) \subseteq \mathcal{B}$.

Prueba. Considere, en primer lugar, la colección

$$\mathfrak{F} = \{\mathcal{B} : \mathcal{B} \text{ es una } \sigma\text{-álgebra sobre } X \text{ con } \mathcal{C} \subseteq \mathcal{B}\}.$$

Observe que $\mathfrak{F} \neq \varnothing$ ya que $\mathscr{P}(X) \in \mathfrak{F}$. Defina

$$\sigma(\mathcal{C}) = \bigcap_{\mathcal{B} \in \mathfrak{F}} \mathcal{B}.$$

Es un ejercicio sencillo establecer que $\sigma(\mathcal{C})$ es una σ-álgebra conteniendo a $\mathcal{C}$. Más aun, por construcción, ella satisface la propiedad (b) y, por supuesto, es única. ■

A la σ-álgebra $\sigma(\mathcal{C})$ del teorema anterior se le llama la **σ-álgebra generada** por $\mathcal{C}$ y es la σ-álgebra *más pequeña conteniendo a* $\mathcal{C}$. Observe que:

(a) Si $\mathcal{D} \subseteq \mathcal{C}$, entonces $\sigma(\mathcal{D}) \subseteq \sigma(\mathcal{C})$.

(b) Si $\mathcal{A}$ es una σ-álgebra, entonces $\sigma(\mathcal{A}) = \mathcal{A}$. En particular,

$$\sigma(\sigma(\mathcal{C})) = \sigma(\mathcal{C}).$$

La definición de σ-álgebra generada es, desde el punto de visto práctico, casi inoperante: ella no ofrece ningún indicio de cómo se reconoce que un conjunto determinado pertenece o no a dicha σ-álgebra. En las próximas líneas intentaremos esbozar un procedimiento que permite conocer cómo son sus elementos. El programa para alcanzar tal objetivo comienza fijando un conjunto no vacío X y después se considera $\mathcal{C}$, una colección no vacía de subconjuntos de X. Nuestro segundo paso es definir $\mathcal{C}_0 = \mathcal{C} \cup \{\varnothing, X\}$ (en el supuesto de que $\varnothing$ y X no están en la colección) y luego considerar la colección

$$\mathcal{C}_1 = \{A \subseteq X : A \in \mathcal{C} \text{ o } A^c \in \mathcal{C}\}. \tag{1}$$

Por supuesto, ésta colección puede no ser cerrada bajo uniones numerables, de modo que es apropiado considerar la familia

$$\mathcal{B}_2 = \left\{ \bigcup_{j=1}^{\infty} A_j : (A_j)_{j=1}^{\infty} \in \mathcal{C}_1 \right\}.$$

Así, podemos repetir el primer argumento y definir

$$\mathcal{C}_2 = \{A \subseteq X : A \in \mathcal{B}_2 \text{ o } A^c \in \mathcal{B}_2\}.$$

Observemos de nuevo de que no hay garantía de que la colección $\mathcal{C}_2$ sea cerrada bajo uniones numerables y, como antes, cabe considerar la familia $\mathcal{B}_3$ formada por todas las uniones numerables de conjuntos pertenecientes a $\mathcal{C}_2$. Si se continúa inductivamente con este proceso se obtienen, por cada $n > 1$, el par de familias

$$\mathcal{B}_n = \left\{\bigcup_{j=1}^{\infty} A_j : (A_j)_{j=1}^{\infty} \in \mathcal{C}_{n-1}\right\} \qquad \text{y} \qquad \mathcal{C}_n = \{A \subseteq X : A \in \mathcal{B}_n \text{ o } A^c \in \mathcal{B}_n\}.$$

Nótese que $\mathcal{C}_1 \subseteq \mathcal{C}_2 \subseteq \cdots$ y que ninguna de esas colecciones de conjuntos es igual a $\sigma(\mathcal{C})$, de modo que es conveniente considerar la unión $\bigcup_{n=1}^{\infty} \mathcal{C}_n$. Sin embargo, esta unión no es, en general, igual a $\sigma(\mathcal{C})$. ¿Cuántas veces hay que realizar este proceso para alcanzar a la σ-álgebra $\sigma(\mathcal{C})$? Para que éste ocurra necesitamos hacerlo ω_1-veces, lo que implica utilizar inducción transfinita. Fijemos entonces un ordinal α con $\alpha < \omega_1$. Si α es un *ordinal límite*, defina

$$\mathcal{C}_\alpha = \bigcup_{\beta<\alpha} \mathcal{C}_\beta.$$

Si α *no es un ordinal límite*, entonces él posee un predecesor inmediato, digamos $\alpha = \beta^+$, y entonces se define

$$\mathcal{B}_\alpha = \left\{\bigcup_{j=1}^{\infty} A_j : (A_j)_{j=1}^{\infty} \in \mathcal{C}_\beta\right\}. \tag{2}$$

Finalmente, sea

$$\mathcal{C}_\alpha = \{A \subseteq X : A \in \mathcal{B}_\alpha \text{ o } A^c \in \mathcal{B}_\alpha\}. \tag{3}$$

Observe que: para cada $\alpha < \omega_1$, $\mathcal{B}_\alpha \subseteq \mathcal{C}_\alpha$. Además, se cumple que

Lema 6.3.12. *Si $\beta < \alpha < \omega_1$, entonces*

$$\mathcal{C}_\beta \subseteq \mathcal{C}_\alpha.$$

Prueba. Considere el conjunto

$$\Lambda = \left\{\alpha < \omega_1 : \bigcup_{\beta<\alpha} \mathcal{C}_\beta \subseteq \mathcal{C}_\alpha\right\}.$$

Claramente $\alpha = 1 \in \Lambda$. Sea $\alpha \in \omega_1$ y veamos que la condición $\operatorname{Seg}(\alpha) \subseteq \Lambda$ implica que $\alpha \in \Lambda$. En efecto, si α es un ordinal límite, entonces se sigue de la definición de $\mathcal{C}_\alpha$ que $\alpha \in \Lambda$. Asuma ahora que α posee un predecesor inmediato, digamos β. Por la definición de $\mathcal{C}_\alpha$, tenemos que $\mathcal{C}_\beta \subseteq \mathcal{C}_\alpha$. Sea $\gamma \in \operatorname{Seg}(\alpha) = \operatorname{Seg}(\beta^+) = \operatorname{Seg}(\beta) \cup \{\beta\}$. Si $\gamma = \beta$, entonces como fue observado anteriormente, $\mathcal{C}_\gamma \subseteq \mathcal{C}_\alpha$. Suponga que $\gamma \in \operatorname{Seg}(\beta)$. Puesto que $\beta \in \operatorname{Seg}(\alpha) \subseteq \Lambda$, se tiene que

$$\mathcal{C}_\gamma \subseteq \bigcup_{\zeta<\beta} \mathcal{C}_\zeta \subseteq \mathcal{C}_\beta \subseteq \mathcal{C}_\alpha.$$

Con esto se ha demostrado que $\mathcal{C}_\gamma \subseteq \mathcal{C}_\alpha$ para todo $\gamma \in \operatorname{Seg}(\alpha)$, lo cual implica que $\bigcup_{\gamma<\alpha} \mathcal{C}_\gamma \subseteq \mathcal{C}_\alpha$ y, por lo tanto, $\alpha \in \Lambda$. Un llamado al Principio de Inducción Transfinita nos revela entonces que $\Lambda = \omega_1$ y termina la prueba. ∎

Teorema 6.3.13 (σ-álgebra generada). *Sea X un conjunto no vacío y sea $\mathcal{C}$ una colección no vacía de subconjuntos de X. Entonces*

$$\sigma(\mathcal{C}) = \bigcup_{\alpha<\omega_1} \mathcal{C}_\alpha,$$

donde la colección $(\mathcal{C}_\alpha)_{\alpha<\omega_1}$ se define como en (2) y (3).

Prueba. Pongamos $\mathcal{C}_{\omega_1} = \bigcup_{\alpha<\omega_1} \mathcal{C}_\alpha$ y veamos que $\sigma(\mathcal{C}) = \mathcal{C}_{\omega_1}$. En primer lugar vamos a demostrar que $\mathcal{C}_{\omega_1}$ es una σ-álgebra conteniendo a $\mathcal{C}$. Para ello, defina

$$\Lambda = \{\alpha < \omega_1 : \mathcal{C}_\alpha \text{ es cerrada bajo complementos}\}.$$

Si logramos que demostrar que $\Lambda = \omega_1$, entonces concluiremos que $\mathcal{C}_{\omega_1}$ es cerrada bajo complementos. Sea $\alpha < \omega_1$ y suponga que $\Lambda_\alpha = \mathrm{Seg}(\alpha) \subseteq \Lambda$. Observe que si $\alpha = 1$ o si α tiene un predecesor inmediato, entonces se sigue de la definición de $\mathcal{C}_\alpha$ que $\alpha \in \Lambda$. Suponga entonces que α es un ordinal límite y sea $E \in \mathcal{C}_\alpha = \bigcup_{\beta<\alpha} \mathcal{C}_\beta$. Entonces $E \in \mathcal{C}_\beta$ para algún $\beta \in \Lambda_\alpha$. Puesto que $\Lambda_\alpha \subseteq \Lambda$, se concluye $\mathcal{C}_\beta$ es cerrada bajo complementos. Por esto, $E^c \in \mathcal{C}_\beta \subseteq \mathcal{C}_\alpha$, lo cual nos dice que $\mathcal{C}_\alpha$ es cerrada bajo complementos y, en consecuencia, $\alpha \in \Lambda$. En conclusión, hemos demostrado que para cualquier ordinal $\alpha < \omega_1$, la condición $\Lambda_\alpha \subseteq \Lambda$ implica que $\alpha \in \Lambda$. El Principio de Inducción Transfinita nos revela entonces que $\Lambda = \omega_1$.

Para terminar de verificar que $\mathcal{C}_{\omega_1}$ es una σ-álgebra sólo nos queda probar que es cerrada bajo uniones numerables. Sea $(E_j)_{j=1}^\infty \subseteq \mathcal{C}_{\omega_1}$. Entonces existe una colección numerable $\{\alpha_j : j \in \mathbb{N}\}$ tal que $E_j \in \mathcal{C}_{\alpha_j}$. Por el Teorema 1.3.32, página 63, sabemos que $\beta = \sup\{\alpha_j : j \in \mathbb{N}\}$ existe. Sea $\alpha_0 = \beta^+$. Puesto que $\alpha_j \leq \beta$ para todo $j \in \mathbb{N}$, resulta del Lema 6.3.12 que $\mathcal{C}_{\alpha_j} \subseteq \mathcal{C}_\beta$ de donde se obtiene que $(E_j)_{j=1}^\infty \subseteq \mathcal{C}_\beta$. Finalmente, por definición, $\bigcup_{j=1}^\infty E_j \subseteq \mathcal{B}_{\alpha_0} \subseteq \mathcal{C}_{\alpha_0} \subseteq \bigcup_{\alpha<\omega_1} \mathcal{C}_\alpha = \mathcal{C}_{\omega_1}$.

Claramente, $\mathcal{C} \subseteq \mathcal{C}_{\omega_1}$, de modo que $\mathcal{C}_{\omega_1}$ es una σ-álgebra conteniendo a $\mathcal{C}$ y en consecuencia,

$$\sigma(\mathcal{C}) \subseteq \mathcal{C}_{\omega_1}.$$

Para demostrar la otra inclusión, sea

$$\Lambda = \{\alpha < \omega_1 : \mathcal{C}_\alpha \subseteq \sigma(\mathcal{C})\}.$$

Por la definición de $\mathcal{C}_\alpha$ se tiene que: si $\Lambda_\alpha = \mathrm{Seg}(\alpha) \subseteq \Lambda$, entonces $\alpha \in \Lambda$. Se sigue del Principio de Inducción Transfinita que $\Lambda = \omega_1$ y de esto se sigue que $\mathcal{C}_\alpha \subseteq \sigma(\mathcal{C})$ para todo $\alpha < \omega_1$ y, por lo tanto,

$$\mathcal{C}_{\omega_1} = \bigcup_{\alpha<\omega_1} \mathcal{C}_\alpha \subseteq \sigma(\mathcal{C}).$$

La prueba es completa. ■

Observe que al ser $\mathcal{C}_{\omega_1}$ una σ-álgebra, no se generan nuevos conjuntos en $\mathcal{C}_{\omega_1}$ tomando complementos, así como tampoco formando uniones numerables de sus elementos, de modo tal que el proceso realmente finaliza en $\aleph_1$-pasos.

Otra manera de caracterizar a $\sigma(\mathcal{C})$ es mirar las colecciones numerables que habitan en $\mathcal{C}$ y considerar las σ-álgebras que ellas generan. Se obtiene entonces el siguiente resultado.

Teorema 6.3.14. *Sea X un conjunto no vacío y sea $\mathcal{C}$ una colección no vacía de subconjuntos de X. Entonces*

$$\sigma(\mathcal{C}) = \bigcup_{\substack{\mathcal{F}\subseteq\mathcal{C} \\ \mathcal{F} \text{ numerable}}} \sigma(\mathcal{F}).$$

Prueba. Pongamos
$$\mathcal{A} = \bigcup_{\substack{\mathcal{F}\subseteq\mathcal{C} \\ \mathcal{F}\text{ numerable}}} \sigma(\mathcal{F})$$
y veamos que $\mathcal{A} = \sigma(\mathcal{C})$. En primer lugar, observe que para cualquier colección numerable $\mathcal{F} \subseteq \mathcal{C}$, se tiene que $\sigma(\mathcal{F}) \subseteq \sigma(\mathcal{C})$, de donde se sigue
$$\mathcal{A} = \bigcup_{\substack{\mathcal{F}\subseteq\mathcal{C} \\ \mathcal{F}\text{ numerable}}} \sigma(\mathcal{F}) \subseteq \sigma(\mathcal{C}).$$
Para verificar la otra inclusión es suficiente demostrar que $\mathcal{A}$ es una σ-álgebra conteniendo a $\mathcal{C}$. Sea $E \in \mathcal{A}$. Entonces $E \in \sigma(\mathcal{F})$ para alguna colección numerable $\mathcal{F} \subseteq \mathcal{C}$. Puesto que $\sigma(\mathcal{F})$ es una σ-álgebra, resulta que $E^c \in \sigma(\mathcal{F}) \subseteq \mathcal{A}$. Esto muestra que $\mathcal{A}$ es cerrada bajo complementos. Suponga ahora que $(E_n)_{n=1}^{\infty}$ es una sucesión en $\mathcal{A}$. Entonces, para cada $n \in \mathbb{N}$, existe un conjunto numerable $\mathcal{F}_n \subseteq \mathcal{C}$ tal que $E_n \in \sigma(\mathcal{F}_n)$. Si hacemos $\mathcal{F} = \bigcup_{n=1}^{\infty} \mathcal{F}_n$, se tiene que $\mathcal{F}$ es numerable y
$$\bigcup_{n=1}^{\infty} E_n \in \bigcup_{n=1}^{\infty} \sigma(\mathcal{F}_n) \subseteq \mathcal{A}.$$
Con esto se concluye que $\mathcal{A}$ es cerrada bajo uniones numerables y, en consecuencia, es una σ-álgebra. Falta comprobar que $\mathcal{A}$ contiene a $\mathcal{C}$. Para ver esto, sea $E \in \mathcal{C}$. Puesto que $\mathcal{F} = \{E\}$ es una colección numerable (posee un único elemento), resulta que $E \in \sigma(\mathcal{F}) \subseteq \mathcal{A}$, es decir, $\mathcal{C} \subseteq \mathcal{A}$. Por esto, $\sigma(\mathcal{C}) \subseteq \mathcal{A}$ y termina la prueba. ■

Definición 6.3.15. *Sea (X, τ) un espacio topológico y denote por $\mathcal{O}_X$ la familia de todos los subconjuntos abiertos de X. La σ-álgebra generada por $\mathcal{O}_X$ es llamada la **σ-álgebra de Borel** y denotada por $\mathfrak{Bo}(X)$.*

Los elementos de la σ-álgebra $\mathfrak{Bo}(X)$ serán llamados ***conjuntos de Borel*** o, simplemente, ***borelianos*** de X. Recordemos que si $Y \subseteq X$, entonces un conjunto $V \subseteq Y$ es abierto en Y si, y sólo si, $V = G \cap Y$ para algún conjunto abierto $G \subseteq X$. Fijemos un conjunto $Y \in \mathfrak{Bo}(X)$ y considere la familia $\mathcal{O}_Y = \{G \cap Y : G \in \mathcal{O}_X\}$, entonces
$$\mathfrak{Bo}(Y) = \sigma(\mathcal{O}_Y) = \mathfrak{Bo}(X) \cap Y = \big\{E \cap Y : E \in \mathfrak{Bo}(X)\big\}.$$
A los elementos de $\mathfrak{Bo}(Y)$ los llamaremos los **conjuntos de Borel de** Y.

Por definición, $\mathfrak{Bo}(\mathbb{R}) = \sigma(\mathcal{O}_{\mathbb{R}})$, pero no se necesitan a todos los conjuntos abiertos para generar a $\mathfrak{Bo}(\mathbb{R})$. De hecho, existen algunas colecciones "más simples" que $\mathcal{O}_{\mathbb{R}}$ que generan a $\mathfrak{Bo}(\mathbb{R})$.

Ejemplo 6.3.1. *Considere las siguientes colecciones:*
$$\begin{aligned}
\mathcal{A}_1 &= \{I \subseteq \mathbb{R} : I \text{ es un intervalo abierto}\}\\
\mathcal{A}_2 &= \{J \subseteq \mathbb{R} : J \text{ es un intervalo cerrado}\}\\
\mathcal{A}_3 &= \{I \subseteq \mathbb{R} : I \text{ es un intervalo de la forma } (a,b],\ \ a,b \in \mathbb{R},\ \ a \leq b\}\\
\mathcal{A}_4 &= \{F \subseteq \mathbb{R} : F \text{ es un conjunto cerrado}\}\\
\mathcal{A}_5 &= \{K \subseteq \mathbb{R} : K \text{ es un conjunto compacto}\}\\
\mathcal{A}_6 &= \{I \subseteq \mathbb{R} : I \text{ es un intervalo de la forma } (-\infty,a],\ \ a \in \mathbb{R}\}\\
\mathcal{A}_7 &= \{I \subseteq \mathbb{R} : I \text{ es un intervalo de la forma } (a,+\infty),\ \ a \in \mathbb{R}\}
\end{aligned}$$

Entonces se cumple que:

$$\mathfrak{Bo}(\mathbb{R}) \;=\; \sigma(\mathcal{A}_j), \quad para\ j = 1, 2, \ldots, 7. \tag{1}$$

Prueba. Puesto que $\mathcal{A}_1 \subseteq \mathcal{O}_{\mathbb{R}}$, resulta que

$$\sigma(\mathcal{A}_1) \subseteq \mathfrak{Bo}(\mathbb{R}).$$

Por otro lado, como todo conjunto abierto es una unión numerable y disjunta de conjuntos abiertos, se tiene que $\mathcal{O}_{\mathbb{R}} \subseteq \sigma(\mathcal{A}_1)$ y, por consiguiente,

$$\mathfrak{Bo}(\mathbb{R}) \subseteq \sigma(\mathcal{A}_1).$$

Esto prueba que $\mathfrak{Bo}(\mathbb{R}) = \sigma(\mathcal{A}_1)$. Para ver, por ejemplo, que $\mathfrak{Bo}(\mathbb{R}) = \sigma(\mathcal{A}_2)$, observe que para cualquier intervalo cerrado $[a, b]$,

$$[a, b] \;=\; \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, \, b + \frac{1}{n}\right) \in \sigma(\mathcal{O}_{\mathbb{R}})$$

y así, $\mathcal{A}_2 \subseteq \mathfrak{Bo}(\mathbb{R})$. De esto se sigue que $\sigma(\mathcal{A}_2) \subseteq \mathfrak{Bo}(\mathbb{R})$. Recíprocamente, como todo intervalo abierto (a, b) se puede escribir en la forma

$$(a, b) \;=\; \bigcup_{n=1}^{\infty} \left[a + \frac{1}{n}, \, b - \frac{1}{n}\right],$$

resulta que $\mathcal{A}_1 \subseteq \sigma(\mathcal{A}_2)$. Por esto y la primera parte se concluye que $\mathfrak{Bo}(\mathbb{R}) = \sigma(\mathcal{A}_1) = \sigma(\mathcal{A}_2)$. Similarmente, como

$$(a, b) \;=\; \bigcup_{n=1}^{\infty} \left(a, \, b - \frac{1}{n}\right] \qquad \text{y} \qquad (a, b] \;=\; \bigcap_{n=1}^{\infty} \left(a, \, b + \frac{1}{n}\right)$$

se deduce que $\mathfrak{Bo}(\mathbb{R}) = \sigma(\mathcal{A}_3)$. El resto de las afirmaciones en (1) se dejan a cargo del lector. ■

Recordemos que si (X, d) es un ***espacio métrico separable***, entonces existe una colección numerable $\mathcal{V} = \{V_1, V_2, \ldots\}$ de subconjuntos abiertos de X con la siguiente propiedad: para cualquier abierto no vacío $G \subseteq X$, existe una subcolección de $\mathcal{V}$, digamos $(V_{n_k})_{k=1}^{\infty}$, tal que $G = \bigcup_{k=1}^{\infty} V_{n_k}$. A la colección $\mathcal{V}$ se le llama una **base numerable** para X.

Definición 6.3.16. *Sea $(X, \mathcal{A})$ un espacio medible. Si $\mathcal{A} = \sigma(\mathcal{C})$ para alguna subcolección numerable $\mathcal{C}$ de $\mathcal{A}$, entonces diremos que $\mathcal{A}$, es **numerablemente generada**.*

Lema 6.3.17. *Si (X, d) es un **espacio métrico separable**, entonces $\mathfrak{Bo}(X)$ es **numerablemente generada**.*

Prueba. Sea $\mathcal{V}$ una base numerable para X. Puesto que $\mathcal{V} \subseteq \mathcal{O}_X \subseteq \mathfrak{Bo}(X)$, resulta que $\sigma(\mathcal{V}) \subseteq \mathfrak{Bo}(X)$. Por otro lado, como cualquier $G \in \mathcal{O}_X$ puede ser escrito como una unión numerable de elementos de $\mathcal{V}$, se tiene que $\mathcal{O} \subseteq \sigma(\mathcal{V})$ y ya que $\sigma(\mathcal{V})$ es una σ-álgebra conteniendo a $\mathcal{O}_X$, resulta que $\mathfrak{Bo}(X) \subseteq \sigma(\mathcal{V})$. Esto termina la prueba. ■

6.3.3. La Cardinalidad de la σ-álgebra de Borel

Como $\mathcal{O}_{\mathbb{R}} \subseteq \mathfrak{M}_\mu(\mathbb{R})$ y $\mathfrak{M}_\mu(\mathbb{R})$ es una σ-álgebra, resulta de la definición de $\mathfrak{Bo}(\mathbb{R})$ que

$$\mathfrak{Bo}(\mathbb{R}) \subseteq \mathfrak{M}_\mu(\mathbb{R})$$

de donde se tiene que $\operatorname{card}(\mathfrak{Bo}(\mathbb{R})) \leq \operatorname{card}(\mathfrak{M}_\mu(\mathbb{R})) = 2^{\mathfrak{c}}$.

Veremos a continuación que la σ-álgebra de Borel, $\mathfrak{Bo}(\mathbb{R})$, en comparación con la σ-álgebra de los conjuntos medibles según Lebesgue es, en el sentido de cardinalidad, muchísimo más *pequeña*. De hecho, del próximo resultado se deducirá que la gran mayoría de los conjuntos medibles según Lebesgue nunca son borelianos lo que evidencia lo extraño que puede ser un conjunto medible Lebesgue que no sea boreliano.

Teorema 6.3.18. $\operatorname{card}\big(\mathfrak{Bo}(\mathbb{R})\big) = \mathfrak{c}$.

Prueba. Sea $(\mathcal{C}_\alpha)_{\alpha<\omega_1}$ la colección definida anteriormente tal que $\mathfrak{Bo}(\mathbb{R}) = \bigcup_{\alpha<\omega_1} \mathcal{C}_\alpha$, donde $\mathcal{C}_0 = \mathcal{O}_{\mathbb{R}}$. Usaremos inducción transfinita para demostrar que $\operatorname{card}(\mathcal{C}_\alpha) = \mathfrak{c}$ para todo $\alpha < \omega_1$. En efecto, por (6) del Teorema 1.3.47, página 74, sabemos que $\mathfrak{c} = \operatorname{card}(\mathcal{O}_{\mathbb{R}}) = \operatorname{card}(\mathcal{C}_0)$. Sea $\alpha < \omega_1$ y suponga que hemos demostrado que $\operatorname{card}(\mathcal{C}_\beta) = \mathfrak{c}$ para todo $\beta < \alpha$. Si α es un ordinal límite, entonces $\mathcal{C}_\alpha = \bigcup_{\beta<\alpha} \mathcal{C}_\beta$ y como el conjunto $\operatorname{Seg}(\alpha) = \{\beta : \beta < \alpha\}$ es infinito numerable, se tiene que

$$\operatorname{card}(\mathcal{C}_\alpha) = \operatorname{car}\Big(\bigcup_{\beta<\alpha} \mathcal{C}_\beta\Big) = \aleph_0 \cdot \mathfrak{c} = \mathfrak{c}.$$

Suponga ahora que α posee un predecesor inmediato β. Entonces, por definición,

$$\mathcal{B}_\alpha = \Big\{ \bigcup_{j=1}^{\infty} A_j : (A_j)_{j=1}^{\infty} \in \mathcal{C}_\beta \Big\}$$

y, en consecuencia, $\operatorname{card}(\mathcal{B}_\alpha) = \operatorname{card}(\mathcal{C}_\beta)^{\aleph_0} = \mathfrak{c}^{\aleph_0} = \mathfrak{c}$. Puesto que

$$\mathcal{C}_\alpha = \big\{ A \subseteq X : A \in \mathcal{B}_\alpha \text{ o } A^c \in \mathcal{B}_\alpha \big\} = \mathcal{B}_\alpha \cup \mathcal{B}_\alpha^c,$$

resulta que $\operatorname{card}(\mathcal{C}_\alpha) = \operatorname{card}(\mathcal{B}_\alpha) + \operatorname{card}(\mathcal{B}_\alpha) = \mathfrak{c} + \mathfrak{c} = \mathfrak{c}$. Esto termina la prueba de que $\operatorname{card}(\mathcal{C}_\alpha) = \mathfrak{c}$ para todo $\alpha < \omega_1$. Finalmente,

$$\operatorname{card}\big(\mathfrak{Bo}(\mathbb{R})\big) = \operatorname{card}\Big(\bigcup_{\alpha<\omega_1} \mathcal{C}_\alpha\Big) = \mathfrak{c} \cdot \mathfrak{c} = \mathfrak{c}.$$

La prueba es completa. ■

Corolario 6.3.19 (Lebesgue). $\mathfrak{Bo}(\mathbb{R}) \subsetneqq \mathfrak{M}_\mu(\mathbb{R})$.

Prueba. Puesto que

$$\operatorname{card}\big(\mathfrak{Bo}(\mathbb{R})\big) = \mathfrak{c} < 2^{\mathfrak{c}} = \operatorname{card}(\mathfrak{M}_\mu(\mathbb{R}))$$

el resultado sigue. ■

Del resultado anterior se deduce que existen conjuntos medibles que no son borelianos. ¿Dónde habitan tales conjuntos? La respuesta es sencilla: en los conjuntos nulos.

Corolario 6.3.20. *Existen* ***conjuntos nulos*** *que* ***no son borelianos****, esto es,* $\mathfrak{N}_\mu(\mathbb{R}) \nsubseteq \mathfrak{Bo}(\mathbb{R})$. *Más aun, si* X *es cualquier* ***subconjunto no-numerable*** *de* $\mathbb{R}$*, entonces existe un conjunto* $A \subseteq X$ *tal que* $A \notin \mathfrak{Bo}(\mathbb{R})$.

Prueba. Puesto que $\Gamma \in \mathfrak{N}_\mu(\mathbb{R})$, entonces,

$$\mathscr{P}(\Gamma) \subseteq \mathfrak{N}_\mu(\mathbb{R}) \subseteq \mathfrak{M}_\mu(\mathbb{R}).$$

Por otro lado, como $\text{card}(\mathscr{P}(\Gamma)) = \text{card}(\mathfrak{M}_\mu(\mathbb{R})) = 2^{\mathfrak{c}}$, resulta, por una nueva aplicación del Teorema de Cantor, que:

$$2^{\mathfrak{c}} = \text{card}(\mathfrak{N}_\mu(\mathbb{R})) > \mathfrak{c}.$$

Esto nos revela que $\mathfrak{N}_\mu(\mathbb{R}) \nsubseteq \mathfrak{Bo}(\mathbb{R})$. La prueba de la segunda parte es simple. En efecto, por el Teorema 6.3.18 y el Teorema de Cantor se tiene que:

$$\mathfrak{c} = \text{card}(\mathfrak{Bo}(\mathbb{R})) = \text{card}(\mathbb{R}) = \text{card}(X) < \text{card}(\mathscr{P}(X)) = 2^{\mathfrak{c}},$$

es decir, existe al menos un conjunto $A \subseteq X$ tal que $A \notin \mathfrak{Bo}(\mathbb{R})$. La prueba es completa. ■

Observe que el argumento de cardinalidad usado para demostrar, en el corolario anterior, la existencia de un conjunto no boreliano en cualquier conjunto no-numerable no es aplicable para obtener un conjunto no-medible según Lebesgue. De hecho, dicha prueba no afirma que A sea medible según Lebesgue. Más adelante veremos, véase el Ejercicio 6.7.1 (d), página 354, que existe otro modo de probar la existencia de conjuntos medibles y, en particular, de conjuntos nulos que no son borelianos sin usar la noción de cardinalidad.

El siguiente resultado constituye otro modo de pensar a la σ-álgebra de Borel. Como siempre, las colecciones $\mathcal{O}_\mathbb{R}$ y $\mathcal{F}_\mathbb{R}$ son, respectivamente, los conjuntos abiertos y cerrados de $\mathbb{R}$.

Lema 6.3.21. *Sea* $\mathfrak{S}(\mathbb{R})$ *la colección más pequeña de subconjuntos de* $\mathbb{R}$ *que cumple:*

(a) $\mathcal{O}_\mathbb{R}, \mathcal{F}_\mathbb{R} \subseteq \mathfrak{S}(\mathbb{R})$,

(b) $\mathfrak{S}(\mathbb{R})$ ***es cerrada bajo la formación de intersecciones numerables****, y*

(c) $\mathfrak{S}(\mathbb{R})$ ***es cerrada bajo la formación de uniones numerables disjuntas****.*

Entonces $\mathfrak{Bo}(\mathbb{R}) = \mathfrak{S}(\mathbb{R})$.

Prueba. Sea $\mathfrak{S}_0 = \{A : A \in \mathfrak{S}(\mathbb{R}) \text{ y } A^c \in \mathfrak{S}(\mathbb{R})\}$. Es claro que $\mathfrak{S}_0 \subseteq \mathfrak{S}(\mathbb{R}) \subseteq \mathfrak{Bo}(\mathbb{R})$. Si logramos demostrar que $\mathfrak{S}_0$ es una σ-álgebra conteniendo a los subconjuntos abiertos de $\mathbb{R}$, tendríamos que $\mathfrak{S}_0 = \mathfrak{S}(\mathbb{R}) = \mathfrak{Bo}(\mathbb{R})$ y terminaría la prueba.

Por definición, $\mathfrak{S}_0$ contiene a los conjuntos abiertos y es cerrada bajo complementos. Suponga ahora que $(A_n)_{n=1}^\infty$ es una sucesión de conjuntos en $\mathfrak{S}_0$. Si definimos $F_1 = A_1$ y para $n \geq 2$,

$$F_n = A_n \setminus \bigcup_{j=1}^{n-1} A_j = A_n \cap \left(\bigcap_{j=1}^{n} A_j^c\right)$$

resulta que cada $F_n \in \mathfrak{S}(\mathbb{R})$ y como $\bigcup_{n=1}^\infty A_n = \bigcup_{n=1}^\infty F_n$ se tiene que $\bigcup_{n=1}^\infty A_n \in \mathfrak{S}(\mathbb{R})$ y, por supuesto, también $(\bigcup_{n=1}^\infty A_n)^c = \bigcap_{n=1}^\infty A_n^c \in \mathfrak{S}(\mathbb{R})$ ya que cada $A_n^c \in \mathfrak{S}(\mathbb{R})$. De esto se sigue que $\bigcup_{n=1}^\infty A_n \in \mathfrak{S}_0$. Con esto quedad demostrado que $\mathfrak{S}_0$ es una σ-álgebra conteniendo a los subconjuntos abiertos de $\mathbb{R}$ y finaliza la prueba. ■

A pesar de ser $\mathfrak{Bo}(\mathbb{R})$ un conjunto muy pequeño, en relación a su cardinalidad, con respecto a $\mathfrak{M}_\mu(\mathbb{R})$, curiosamente $\mathfrak{M}_\mu(\mathbb{R})$ es sólo un "poquito más grande" que $\mathfrak{Bo}(\mathbb{R})$ en el siguiente sentido: como veremos en el Teorema 6.3.55, página 293, se tendrá que

$$\mathfrak{M}_\mu(\mathbb{R}) \;=\; \mathfrak{Bo}(\mathbb{R}) \cup \mathfrak{N}_\mu(\mathbb{R}).$$

Este hecho nos muestra que ***todo conjunto medible se diferencia de un conjunto boreliano sólo por un conjunto de medida cero***. Otra manera de decir esto es que $\mathfrak{Bo}(\mathbb{R})$ es la **completación** de $\mathfrak{M}_\mu(\mathbb{R})$. Por ésta y muchas otras razones, los conjuntos borelianos son extremadamente importantes.

Nota Adicional 6.3.1 El resultado anterior, Corolario 6.3.19, nos muestra que *existen conjuntos medibles según Lebesgue que no son medibles según Borel* sin mostrar ni un sólo ejemplo concreto del tal fenómeno y, además, que la totalidad de los conjuntos de Borel es infinitamente más pequeño que la de todos los conjuntos medibles según Lebesgue, en otras palabras, la "*gran mayoría*" de los conjuntos medibles según Lebesgue **nunca** son borelianos.

Aunque a Lebesgue le parecía dudoso que alguien pudiera jamás nombrar un conjunto medible según Lebesgue que no fuese de Borel, el matemático ruso Nicolas Lusin (1883-1950) fue el primero en construir explícitamente un conjunto medible Lebesgue que no era de Borel. El ejemplo es el siguiente: Considere el conjunto S compuesto de todas las fracciones continuas $[a_0; a_1, a_2, \ldots]$ que cumplen con la siguiente propiedad: existe una subsucesión $(n_j)_{j=1}^{\infty}$ tal que a_{n_j} es un divisor de $a_{n_{j+1}}$ para todo $j \in \mathbb{N}$. Lusin demostró que S es medible según Lebesgue (de hecho, un conjunto analítico) pero no es boreliano. En la próxima sección veremos una versión más general de este hecho. Por otro lado, en el espacio de Banach con la norma del supremo, $C([a,b])$, formado por todas las funciones $f : [a,b] \to \mathbb{R}$ que son continuas, el conjunto $D \subseteq C([a,b])$ de las funciones que son diferenciables en todo punto de $[a,b]$ constituye un conjunto que no es de Borel. Un poco más adelante, en la sección sobre Ejemplos y Contraejemplos usando la Función de Cantor, veremos otro modo de mostrar conjuntos medibles según Lebesgue que no son borelianos.

La siguiente observación puede ser de interés al lector: El Axioma de Elección (en su versión numerable) es el responsable de que exista un continuum de conjuntos borelianos. Sin embargo, rechazando dicho axioma se puede construir un modelo **ZF** + $\neg$**AC** en el cual pueden ocurrir cosas como estas: (véase, [101], p. 103):

(a) $\operatorname{card}(\mathfrak{Bo}(\mathbb{R})) > \mathfrak{c}$.

(b) Cualquier conjunto de números reales sería un conjunto de Borel y la medida de Lebesgue dejaría de ser numerablemente aditiva.

(c) La jerarquía de los conjuntos de Borel se colapsaría en los primeros cuatro niveles.

En estas circunstancias, y desde el punto de vista del analista, la colección $\operatorname{card}(\mathfrak{Bo}(\mathbb{R}))$ sería poco útil.

La jerarquía boreliana

Siguiendo la convención establecida en la Teoría Descriptiva de Conjuntos, definamos las colecciones, comenzando con $\alpha = 1$,

$$\mathbf{\Sigma}_1^0 \;=\; \big\{G \subseteq \mathbb{R} \,:\, G \text{ es abierto}\big\} \;=\; \mathcal{O}_{\mathbb{R}}$$

$$\mathbf{\Pi}_1^0 \;=\; \big\{\mathbb{R} \setminus G \,:\, G \in \mathbf{\Sigma}_1^0\big\} \;=\; \mathcal{F}_{\mathbb{R}}$$

y, para cada $1 < \alpha < \omega_1$, defina

$$\mathbf{\Sigma}^0_\alpha = \Big(\bigcup_{\beta<\alpha} \mathbf{\Pi}^0_\beta\Big)_\sigma, \qquad \mathbf{\Pi}^0_\alpha = \Big(\bigcup_{\beta<\alpha} \mathbf{\Sigma}^0_\beta\Big)_\delta \qquad \text{y} \qquad \mathbf{\Delta}^0_\alpha = \mathbf{\Sigma}^0_\alpha \cap \mathbf{\Pi}^0_\alpha.$$

Así, un conjunto $A \in \mathbf{\Sigma}^0_\alpha$ para $\alpha > 1$, si existe una sucesión de conjuntos $\big(A_n\big)_{n=1}^\infty$ tal que $A_n \in \mathbf{\Pi}^0_{\beta_n}$ para algún $\beta_n < \alpha$ y $A = \bigcup_{n=1}^\infty A_n$. Similarmente, un conjunto $A \in \mathbf{\Pi}^0_\alpha$ para $\alpha > 1$, si existe una sucesión de conjuntos $\big(A_n\big)_{n=1}^\infty$ tal que $A_n \in \mathbf{\Sigma}^0_{\beta_n}$ para algún $\beta_n < \alpha$ y $A = \bigcap_{n=1}^\infty A_n$. Se verifica entonces que

$$\mathfrak{Bo}(\mathbb{R}) = \bigcup_{\alpha<\omega_1} \mathbf{\Sigma}^0_\alpha = \bigcup_{\alpha<\omega_1} \mathbf{\Pi}^0_\alpha = \bigcup_{\alpha<\omega_1} \mathbf{\Delta}^0_\alpha$$

es decir, los conjuntos de Borel se construyen en una jerarquía de longitud ω_1 conocida como la **jerarquía de Borel**. Observe que

$$\mathbf{\Sigma}^0_2 = \mathcal{F}_\sigma, \quad \mathbf{\Pi}^0_2 = \mathcal{G}_\delta, \quad \mathbf{\Sigma}^0_3 = \mathcal{G}_{\delta\sigma}, \quad \mathbf{\Pi}^0_3 = \mathcal{F}_{\sigma\delta}, \ \ldots$$

En este sentido, cada conjunto de Borel posee una plantilla de construcción, conocido como **código de Borel**, que detalla exactamente cómo ellos han sido construidos a partir de los conjuntos abiertos.

$$\begin{array}{ccccccccccccc}
\mathbf{\Sigma}^0_1 & \subseteq & \mathbf{\Sigma}^0_2 & \subseteq & \mathbf{\Sigma}^0_3 & \subseteq & \mathbf{\Sigma}^0_4 & \subseteq & \cdots & \subseteq & \mathbf{\Sigma}^0_\alpha & \cdots \\
 & \rotatebox{-45}{$\subseteq$}\,\rotatebox{45}{$\subseteq$} & & \rotatebox{-45}{$\subseteq$}\,\rotatebox{45}{$\subseteq$} & & \rotatebox{-45}{$\subseteq$}\,\rotatebox{45}{$\subseteq$} & & \rotatebox{-45}{$\subseteq$}\,\rotatebox{45}{$\subseteq$} & \cdots & \rotatebox{-45}{$\subseteq$}\,\rotatebox{45}{$\subseteq$} & & \cdots \\
\mathbf{\Pi}^0_1 & \subseteq & \mathbf{\Pi}^0_2 & \subseteq & \mathbf{\Pi}^0_3 & \subseteq & \mathbf{\Pi}^0_4 & \subseteq & \cdots & \subseteq & \mathbf{\Pi}^0_\alpha & \cdots
\end{array}$$

El hecho de que la formación de todas esas colecciones debe ser llevada a cabo hasta el primer ordinal no-numerable para obtener a $\mathfrak{Bo}(\mathbb{R})$ se puede evidenciar con el siguiente argumento: suponga, por ejemplo, que elegimos una sucesión $\big(A_n\big)_{n=1}^\infty$ de modo que

$$A_1 \in \mathbf{\Sigma}^0_2 \setminus \mathbf{\Sigma}^0_1, \quad A_2 \in \mathbf{\Sigma}^0_3 \setminus \mathbf{\Sigma}^0_2, \quad A_3 \in \mathbf{\Sigma}^0_4 \setminus \mathbf{\Sigma}^0_3, \ \ldots$$

El conjunto $\bigcup_{n=1}^\infty A_n$, ¿a cuál de las familias $(\mathbf{\Sigma}^0_k)_{k=1}^\infty$ pertenece? La respuesta es que no hay garantía de que esté en alguna de ellas. La única certeza que tenemos es que pertenece a alguna de las familias obtenidas en el paso $\omega + 1$. Si se continúa con este proceso de selección de sucesiones se puede apreciar lo que afirmamos anteriormente.

6.3.4. Conjuntos Analíticos

En esta sección veremos otro modo de demostrar la existencia de un conjunto medible que no es de Borel. Esto requiere introducir una familia muy especial de conjuntos medibles según Lebesgue, denominados conjuntos analíticos, que contiene a los conjuntos de Borel propiamente.

Este método permite la existencia de conjuntos medibles que no son borelianos sin apelar al argumento de cardinalidad.

La motivación para introducir la noción de los conjuntos analíticos es el siguiente resultado el cual es una pieza clave para determinar si un conjunto dado es o no un conjunto de Borel partiendo de una función dada.

Lema 6.3.22 (Imagen e imagen inversa de borelianos). *Sea $f : \mathbb{R} \to \mathbb{R}$ una función.*

(1) *Si f es **estrictamente creciente**, entonces $f^{-1}(B) \in \mathfrak{Bo}(\mathbb{R})$, para cualquier $B \in \mathfrak{Bo}(\mathbb{R})$.*

(2) *Si f es **continua e inyectiva**, entonces $f(B) \in \mathfrak{Bo}(\mathbb{R})$, para cualquier $B \in \mathfrak{Bo}(\mathbb{R})$.*

Prueba. (1) Considere la siguiente colección de subconjuntos de $\mathbb{R}$,

$$\mathcal{B} \;=\; \big\{B \subseteq \mathbb{R} \,:\, f^{-1}(B) \in \mathfrak{Bo}(\mathbb{R})\big\}.$$

Es un ejercicio sencillo establecer que $\mathcal{B}$ es una σ-álgebra. En efecto, es claro que $\varnothing \in \mathcal{B}$. Si $B \in \mathcal{B}$, entonces $\mathbb{R} \setminus B$ también está en $\mathcal{B}$ ya que

$$f^{-1}\big(\mathbb{R} \setminus B\big) \;=\; f^{-1}(\mathbb{R}) \setminus f^{-1}(B) \;=\; \mathbb{R} \setminus f^{-1}(B) \;\in\; \mathfrak{Bo}(\mathbb{R}).$$

Sea $(B_n)_{n=1}^{\infty}$ una sucesión de conjuntos en $\mathcal{B}$. Entonces

$$f^{-1}\bigg(\bigcup_{n=1}^{\infty} B_n\bigg) \;=\; \bigcup_{n=1}^{\infty} f^{-1}(B_n) \;\in\; \mathfrak{Bo}(\mathbb{R}).$$

Esto prueba que $\bigcup_{n=1}^{\infty} B_n \in \mathcal{B}$ y así, $\mathcal{B}$ es una σ-álgebra.

Como f es estrictamente creciente, ella es, en particular, sobreyectiva y, así, f es biyectiva. Para cada $a \in \mathbb{R}$ seleccione un x tal que $a = f(x)$ y observe que $f^{-1}\big((a, +\infty)\big) = (x, +\infty) \in \mathcal{B}$. Esto prueba que $(a, +\infty) \in \mathcal{B}$ y, similarmente, $(-\infty, b) \in \mathcal{B}$ para cada $b \in \mathbb{R}$. De aquí se sigue que $(a, b) \in \mathcal{B}$ para cualesquiera $a < b$ ya que

$$f^{-1}\big((a, b)\big) \;=\; f^{-1}(a, +\infty) \cap f^{-1}(-\infty, b) \;\in\; \mathfrak{Bo}(\mathbb{R}).$$

De lo anterior se concluye que $\mathcal{O} \subseteq \mathcal{B}$ y, por lo tanto, $\mathfrak{Bo}(\mathbb{R}) \subseteq \mathcal{B}$.

(2) Como f es inyectiva, ella es estrictamente monótona y, en consecuencia, biyectiva. Considere

$$\mathcal{C} \;=\; \big\{B \subseteq \mathbb{R} \,:\, f(B) \in \mathfrak{Bo}(\mathbb{R})\big\}.$$

De la biyectividad de f se sigue que $\mathcal{C}$ es una σ-álgebra. Veamos que $\mathcal{C}$ contiene a $\mathcal{O}$. En efecto, puesto que f es continua y biyectiva, ella envía intervalos abiertos en intervalos abiertos. Sea $G \in \mathcal{O}$ y exprese dicho conjunto como una unión disjunta de intervalos abiertos, digamos, $G = \bigcup_{n=1}^{\infty} J_n$. De la biyectividad de f se sigue que $f(G) = \bigcup_{n=1}^{\infty} f(J_n) \in \mathfrak{Bo}(\mathbb{R})$ y, por consiguiente, $\mathcal{O} \subseteq \mathcal{C}$. De aquí se deduce que $\mathfrak{Bo}(\mathbb{R}) \subseteq \mathcal{C}$. ■

Del resultado anterior se concluye que si f es un homeomorfismo, entonces dicha función preserva conjuntos de Borel, así como imágenes inversa, esto es:

Corolario 6.3.23. *Si $f : \mathbb{R} \to \mathbb{R}$ es un **homeomorfismo**, entonces f **preserva conjuntos borelianos** así como sus **imágenes inversas**, es decir, para todo $B \subseteq \mathbb{R}$*

$$f^{-1}(B) \;\in\; \mathfrak{Bo}(\mathbb{R}) \quad\Leftrightarrow\quad B \;\in\; \mathfrak{Bo}(\mathbb{R}) \quad\Leftrightarrow\quad f(B) \;\in\; \mathfrak{Bo}(\mathbb{R}).$$

Es importante señalar que la validez de la segunda equivalencia del corolario anterior no se extiende a los conjuntos medibles. Esto significa, véase el ejemplo dado en el Ejercicio 6.7.1 (c), página 354, que existe un homeomorfismo $f:\mathbb{R}\to\mathbb{R}$ y existe al menos un conjunto medible $E\in\mathfrak{M}_\mu(\mathbb{R})$ tal que $f(E)\notin\mathfrak{M}_\mu(\mathbb{R})$.

Definición 6.3.24. *Una función $f:\mathbb{R}\to\mathbb{R}$ se dice que es* ***medible según Borel*** *si $f^{-1}(B)\in\mathfrak{Bo}(\mathbb{R})$ para cualquier $B\in\mathfrak{Bo}(\mathbb{R})$.*

Se sigue del Ejercicio 6.3.1 (7) que una función $f:\mathbb{R}\to\mathbb{R}$ es medible según Borel si $f^{-1}((a,+\infty))\in\mathfrak{Bo}(\mathbb{R})$ para cualquier $a\in\mathbb{R}$. Observe que, gracias al Lema 6.3.22, toda función estrictamente creciente $f:\mathbb{R}\to\mathbb{R}$ es medible según Borel. Por supuesto, si $f:\mathbb{R}\to\mathbb{R}$ es continua, entonces f es medible según Borel. En particular, ***todo homeomorfismo*** $f:\mathbb{R}\to\mathbb{R}$ ***es medible según Borel***.

Fijemos ahora una función medible Borel $f:\mathbb{R}\to\mathbb{R}$ y sea $B\in\mathfrak{Bo}(\mathbb{R})$. ¿Es $f(B)\in\mathfrak{Bo}(\mathbb{R})$? Por supuesto, si f es continua e inyectiva, el Lema 6.3.22 provee una respuesta positiva; sin embargo, en el caso general, no hay garantía de que $f(B)\in\mathfrak{Bo}(\mathbb{R})$. En esta parte estamos interesados en investigar la clase de todos los conjuntos $A\subseteq\mathbb{R}$ que son imágenes de un algún conjunto de Borel bajo alguna aplicación medible Borel $f:\mathbb{R}\to\mathbb{R}$.

Sean X,Y conjuntos arbitrarios. Recordemos que la proyección sobre X, $\mathrm{proy}_1:X\times Y\to X$, viene dada por $\mathrm{proy}_1(x,y)=x$ para todo $(x,y)\in X\times Y$.

Definición 6.3.25. *Un conjunto $A\subseteq\mathbb{R}$ se dice que es un* ***conjunto analítico*** *si satisface una de las siguientes condiciones equivalentes:*

(a) *Existe una* ***función medible Borel*** *$f:\mathbb{R}\to\mathbb{R}$ y un conjunto $B\in\mathfrak{Bo}(\mathbb{R})$ tal que $f(B)=A$.*

(b) *Existe una* ***función continua*** *$f:\mathbb{R}\to\mathbb{R}$ y un conjunto $B\in\mathfrak{Bo}(\mathbb{R})$ tal que $f(B)=A$.*

(c) *Existe un* ***subconjunto cerrado*** *$B\subseteq\mathbb{R}\times\mathcal{N}$ cuya proyección sobre $\mathbb{R}$ es A, esto es, $\mathrm{proy}_1(B)=A$.*

(d) *Existe una* ***función continua*** *$f:\mathcal{N}\to\mathbb{R}$ tal que $f(\mathcal{N})=A$.*

En general, si se reemplaza el conjunto $\mathbb{R}$ por cualquier espacio Polaco X, entonces la noción de conjunto analítico en X permanece inalterable. Sin embargo, en estas notas, sólo nos interesa su estudio restringido al espacio Polaco $\mathbb{R}$. La condición (d) de la definición anterior es, tal vez, la más practica y, por supuesto, la más utilizada. El siguiente resultado establece que $\mathbb{R}$ es analítico.

Lema 6.3.26. *Existe una* ***función continua*** *$f:\mathcal{N}\to\mathbb{R}$ tal que $f(\mathcal{N})=\mathbb{R}$.*

Prueba. Sea $D=\{q_n:n\in\mathbb{N}\}$ una enumeración de todos los números racionales. Puesto que cada $x\in\mathbb{R}$ es el límite de alguna subsucesión del conjunto D, uno puede intentar definir la aplicación $g:\mathcal{N}\to\mathbb{R}$ por

$$g\big((n_k)_{k=1}^{\infty}\big)\ =\ \lim_{k\to\infty} q_{n_k}.$$

El problema con esta definición es que el límite del lado derecho puede que no exista. Más aun, en el supuesto que dicho límite exista, no hay garantía de que la aplicación g sea continua. Para remediar esta situación, debemos proceder con más cuidado. Para cada $\alpha=(n_k)_{k=1}^{\infty}\in\mathcal{N}$ defina $x_1^\alpha=q_{n_1}$ y para todo $k\geq 1$, sea

$$x_{k+1}^\alpha\ =\ \begin{cases} q_{n_{k+1}} & \text{si}\ \ |x_k^\alpha-q_{n_{k+1}}|<2^{-k},\\ x_k^\alpha & \text{en otro caso.}\end{cases}$$

La sucesión así obtenida $(x_k^\alpha)_{k=1}^\infty$ es claramente de Cauchy en $\mathbb{R}$ y, en consecuencia, converge a un único $x^\alpha \in \mathbb{R}$. Definamos ahora $f : \mathcal{N} \to \mathbb{R}$ por

$$f(\alpha) = x^\alpha$$

para cada $\alpha = (n_k)_{k=1}^\infty \in \mathcal{N}$. Veamos que f es continua. En efecto, sean $\alpha = (n_k)_{k=1}^\infty$ y $\beta = (m_k)_{k=1}^\infty$ dos elementos distintos de $\mathcal{N}$ y suponga que $N = \text{ínf}\{k \in \mathbb{N} : n_k \neq m_k\}$. Entonces $d(\alpha, \beta) = 2^{-N}$ y, por consiguiente, los N primeros términos de las sucesiones correspondientes, $(x_k^\alpha)_{k=1}^\infty$ y $(x_k^\beta)_{k=1}^\infty$ coinciden. De esta información se tiene que todos los demás términos de ambas sucesiones están a una distancia menor que $1/2^N$ de x_N^α y x_N^β respectivamente y, por lo tanto, la continuidad de f sigue de esto. Finalmente, puesto que D es denso en $\mathbb{R}$, se tiene que f es sobreyectiva y termina la prueba. ■

Observe que la función $f : \mathcal{N} \to \mathbb{R}$ construida en el lema anterior es también inyectiva, aunque f no es un homeomorfismo. Una razón para que eso no ocurra es que $\mathcal{N}$ es totalmente disconexo, es decir, $\mathcal{N}$ contiene una base cuyos elementos son abiertos y cerrados al mismo tiempo y, por supuesto, $\mathbb{R}$ no posee una tal base.

Lema 6.3.27. *Los siguientes conjuntos son **homeomorfos** a $\mathcal{N}$:*

(a) $\mathbb{N} \times \mathcal{N}$.

(b) $\mathcal{N}^{\mathbb{N}}$.

Prueba. (a) La aplicación $\varphi : \mathcal{N} \to \mathbb{N} \times \mathcal{N}$ definida por

$$\varphi((n_1, n_2, n_3, \ldots)) = (n_1, (n_2, n_3, \ldots)) \tag{2}$$

es claramente un homeomorfismo. Para demostrar la parte (b), considere la función $\phi : \mathcal{N} \to \mathcal{N}^{\mathbb{N}}$ dada por

$$\phi(n_1, n_2, \ldots) = ((n_1, n_2, \ldots), (n_2, n_3, \ldots), (n_3, n_4, \ldots), \ldots).$$

No es difícil establecer que ϕ es un homeomorfismo. ■

Como es costumbre, denotaremos por $\mathbf{\Sigma}_1^1$ *a la familia de todos los subconjuntos analíticos de* $\mathbb{R}$. Una consecuencia inmediata de nuestra definición es que los conjuntos analíticos se preservan bajo funciones continuas.

Teorema 6.3.28. *Si $f : \mathbb{R} \to \mathbb{R}$ es una **función continua** y $A \in \mathbf{\Sigma}_1^1$, entonces $f(A) \in \mathbf{\Sigma}_1^1$.*

Prueba. Como A es analítico, existe una función continua $g : \mathcal{N} \to \mathbb{R}$ tal que $g(\mathcal{N}) = A$. La función $h : \mathcal{N} \to \mathbb{R}$ dada por $h(x) = f(g(x))$ es continua y verifica que $h(\mathcal{N}) = f(A)$, es decir, $f(A) \in \mathbf{\Sigma}_1^1$. ■

Se sigue de la definición de conjunto analítico que *cualquier intervalo abierto (cerrado) es un conjunto analítico*. Nos proponemos demostrar que $\mathbf{\Sigma}_1^1$ es una clase muy amplia para tenerla en consideración. De hecho, vamos a demostrar que

$$\mathfrak{Bo}(\mathbb{R}) \subseteq \mathbf{\Sigma}_1^1 \subseteq \mathfrak{M}_\mu(\mathbb{R}).$$

Lema 6.3.29. *Si $(A_n)_{n=1}^\infty$ es una sucesión de **conjuntos analíticos**, entonces $\bigcup_{n=1}^\infty A_n$ y $\bigcap_{n=1}^\infty A_n$ son **analíticos**.*

Prueba. Sea $n \in \mathbb{N}$. Como A_n es analítico, escojamos una función continua $f_n : \mathcal{N} \to \mathbb{R}$ tal que $f_n(\mathcal{N}) = A_n$. Defina $g : \mathbb{N} \times \mathcal{N} \to \mathbb{R}$ por

$$g((n,(n_1,n_2,\ldots))) \;=\; f_n(n_1,n_2,\ldots)$$

Es claro que g es continua y, por lo tanto, $f = g \circ \varphi$ también lo es, donde φ es la función dada por (2). De aquí se sigue que

$$f(\mathcal{N}) \;=\; g\big(\varphi(\mathcal{N})\big) \;=\; g(\mathbb{N} \times \mathcal{N}) \;=\; \bigcup_{n=1}^{\infty} g(\{n\} \times \mathcal{N}) \;=\; \bigcup_{n=1}^{\infty} f_n(\mathcal{N}) \;=\; \bigcup_{n=1}^{\infty} A_n.$$

Esto prueba que $\bigcup_{n=1}^{\infty} A_n \in \Sigma_1^1$. Para demostrar que $\bigcap_{n=1}^{\infty} A_n \in \Sigma_1^1$, primero considere el conjunto

$$\Delta \;=\; \Big\{x = (x_1,x_2,\ldots) \in \mathcal{N}^{\mathbb{N}} \,:\, f_m(x_m) = f_n(x_n) \text{ para todo } m,n \in \mathbb{N}\Big\}$$

y observe que Δ es cerrado en $\mathcal{N}^{\mathbb{N}}$. Finalmente, definiendo $g : \Delta \to \mathbb{R}$ mediante la igualdad $g(x) = f_1(x_1)$ resulta que g es continua y $g(\Delta) = \bigcap_{n=1}^{\infty} A_n$. Extienda g, haciendo uso del Teorema de Extensión de Tietze (Teorema 7.2.19, página 391), a una función continua f definida sobre $\mathcal{N}^{\mathbb{N}}$ y luego aplique el Lema 6.3.27 (b) para finalizar la prueba. ■

Estamos ahora en posesión de los argumentos para demostrar que:

Teorema 6.3.30. $\mathfrak{Bo}(\mathbb{R}) \subseteq \mathbf{\Sigma}_1^1 \subseteq \mathfrak{M}_\mu(\mathbb{R})$.

Prueba. Puesto que todo conjunto abierto no vacío es unión numerable de intervalos abiertos y todo intervalo abierto es analítico, se sigue del resultado anterior que $\mathcal{O} \subseteq \mathbf{\Sigma}_1^1$. Además, como $\mathbf{\Sigma}_1^1$ satisface las condiciones (a), (b) y (c) del Lema 6.3.21 se tiene que $\mathfrak{Bo}(\mathbb{R}) \subseteq \mathbf{\Sigma}_1^1$.

Para demostrar que todo conjunto analítico es medible, suponga que A es un conjunto analítico y escoja una función continua $f : \mathcal{N} \to \mathbb{R}$ tal que $f(\mathcal{N}) = A$. Sin perder generalidad, asumiremos que $\mu^*(A) < +\infty$. Defina, para cada $k \in \mathbb{N}$, el conjunto

$$\mathcal{L}(k) \;=\; \Big\{(m_n)_{n=1}^{\infty} \in \mathcal{N} \,:\, m_1 \leq k\Big\}.$$

Observe que $\mathcal{L}(1) \subseteq \mathcal{L}(2) \subseteq \cdots$ y $\mathcal{N} = \bigcup_{k=1}^{\infty} \mathcal{L}(k)$, de donde resulta que

$$A \;=\; f(\mathcal{N}) \;=\; f\bigg(\bigcup_{k=1}^{\infty} \mathcal{L}(k)\bigg) \;=\; \bigcup_{k=1}^{\infty} f(\mathcal{L}(k)).$$

Puesto que la sucesión $\big(f(\mathcal{L}(k))\big)_{k=1}^{\infty}$ es creciente, se sigue del Corolario 6.3.46, página 286, que

$$\mu^*(A) \;=\; \lim_{k\to\infty} \mu^*\big(f(\mathcal{L}(k))\big).$$

Fijemos un $\varepsilon > 0$ arbitrario y seleccione un entero $n_1 > 1$ lo suficientemente grande de modo que

$$\mu^*\big(f(\mathcal{L}(n_1))\big) \;>\; \mu^*(A) - \varepsilon.$$

Como $\mathcal{L}(n_1) = \bigcup_{k=1}^{\infty} \mathcal{L}(n_1,k)$, un argumento similar al anterior produce un entero positivo n_2 tal que

$$\mu^*\big(f(\mathcal{L}(n_1,n_2))\big) \;>\; \mu^*(A) - \varepsilon.$$

Continuando de este modo se obtiene una sucesión $(n_i)_{i=1}^{\infty} \in \mathcal{N}$ tal que

$$\mu^*\big(f(\mathcal{L}(n_1, n_2, \ldots, n_k))\big) \;>\; \mu^*(A) - \varepsilon \tag{3}$$

para todo $k \in \mathbb{N}$. Nótese que

$$\mathcal{L}(n_1, n_2, \ldots, n_k) \;=\; \Big\{(m_n)_{n=1}^{\infty} \in \mathcal{N} \,:\, m_i \leq n_i \;\text{ para }\; i = 1, \ldots, k\Big\}$$

para cada $k \in \mathbb{N}$. Si ahora definimos

$$L \;=\; \bigcap_{k=1}^{\infty} \mathcal{L}(n_1, n_2, \ldots, n_k) \;=\; \big\{(m_n)_{n=1}^{\infty} \in \mathcal{N} \,:\, m_i \leq n_i \;\text{ para }\; i = 1, 2, \ldots\big\},$$

resulta que L es compacto y así, por continuidad, $K = f(L)$ es un subconjunto compacto de A. Nuestra segunda estrategia es verificar que $\mu(K) > \mu^*(A) - \varepsilon$. En efecto, no es difícil establecer que

$$K \;=\; \bigcap_{k=1}^{\infty} \overline{f\big(\mathcal{L}(n_1, n_2, \ldots, n_k)\big)}$$

y entonces, como $f\big(\mathcal{L}(n_1, n_2, \ldots, n_k)\big) \subseteq \overline{f\big(\mathcal{L}(n_1, n_2, \ldots, n_k)\big)}$, se tiene que

$$\mu\big(\,\overline{f(\mathcal{L}(n_1, n_2, \ldots, n_k))}\,\big) \;\geq\; \mu^*\big(f(\mathcal{L}(n_1, n_2, \ldots, n_k))\big) \;>\; \mu^*(A) - \varepsilon.$$

Más aun, puesto que la sucesión $\big(\,\overline{f(\mathcal{L}(n_1, n_2, \ldots, n_k))}\,\big)_{n=1}^{\infty}$ es decreciente, el Teorema 6.3.47, página 287, nos garantiza que

$$\mu(K) \;=\; \lim_{k\to\infty} \mu\big(\,\overline{f(\mathcal{L}(n_1, n_2, \ldots, n_k))}\,\big) \;>\; \mu^*(A) - \varepsilon.$$

Como K es medible, se sigue del Criterio de Carathéodory, Teorema 6.3.58, página 295, que para todo conjunto $X \subseteq \mathbb{R}$

$$\mu^*(X) \;\geq\; \mu^*(K) + \mu^*(X \setminus K) \;\geq\; \mu^*(A) - \varepsilon + \mu^*(X \setminus A)$$

Puesto que $\varepsilon > 0$ es arbitrario, se tiene que $\mu^*(X) \geq \mu^*(A) + \mu^*(X \setminus A)$ para cualquier conjunto $X \subseteq \mathbb{R}$. Esto prueba, por una nueva aplicación de Criterio de Carathéodory, que A es medible y termina la prueba. ■

Es importante observar que, en general, el complemento de un conjunto analítico no es necesariamente un conjunto analítico. De hecho, el complemento de un conjunto analítico A es un conjunto analítico si, y sólo si, A es un boreliano.

Acabamos de ver que $\mathfrak{Bo}(\mathbb{R}) \subseteq \mathbf{\Sigma}_1^1$ y entonces es lógico indagar si $\mathbf{\Sigma}_1^1 \subseteq \mathfrak{Bo}(\mathbb{R})$. Por supuesto, si la respuesta es afirmativa, el estudio y análisis de los conjuntos analíticos no tendrá ningún valor, salvo el de encontrar propiedades adicionales de los conjuntos de Borel. De modo que nuestra siguiente tarea es la de demostrar la existencia de conjuntos analíticos que no son borelianos. Su construcción se basa sobre una noción de conjunto llamado "conjunto universal", el cual cual consiste en hallar un conjunto que sirva como modelo para determinar a todos los conjuntos de una cierta clase dada.

Sean X, Y conjuntos no vacíos y sea $S \subseteq X \times Y$. Si $(x_0, y_0) \in X \times Y$, la x_0-**sección** y la y_0-**sección** de S se definen, respectivamente, como

$$S_{x_0} \;=\; \big\{y \in Y \,:\, (x_0, y) \in S\big\} \qquad\qquad \text{y} \qquad\qquad S^{y_0} \;=\; \big\{x \in X \,:\, (x, y_0) \in S\big\}.$$

Definición 6.3.31. *Sea X un conjunto no vacío y sea $\mathcal{C}$ una familia de subconjuntos de X. Un subconjunto S de $X \times \mathcal{N}$ es **universal para la clase** $\mathcal{C}$ si*

$$\mathcal{C} \;=\; \big\{S^y \,:\, y \in \mathcal{N}\big\}.$$

Lo que queremos es demostrar la existencia de un conjunto cerrado en $\mathbb{R} \times \mathcal{N}$ que es universal para la clase $\boldsymbol{\Sigma}_1^1$ de los conjuntos analíticos. Esto requiere un resultado previo. Antes, recordemos que $\mathcal{O}_{\mathbb{R}}$ representa a la familia de todos los subconjuntos abiertos de $\mathbb{R}$, mientras que $\mathcal{F}$ constituye la familia de todos los subconjuntos cerrados de $\mathbb{R}$.

Lema 6.3.32. *Existe un **conjunto abierto** U (respectivamente, un **conjunto cerrado** F) en $\mathbb{R} \times \mathcal{N}$ que es **universal** para $\mathcal{O}_{\mathbb{R}}$ (respectivamente, para $\mathcal{F}$).*

Prueba. Sea $\mathcal{B} = (V_n)_{n=1}^{\infty}$ una base numerable para $\mathbb{R}$. Por ejemplo, podemos tomar como $(V_n)_{n=1}^{\infty}$ la colección numerable de todos los intervalos abiertos cuyos extremos son números racionales. En primer lugar, para cada $k, n \in \mathbb{N}$, considere el conjunto $U(k,n) \subseteq \mathbb{R} \times \mathcal{N}$ definido por

$$\begin{aligned} U(k,n) \;&=\; V_n \times \Big\{(m_j)_{j=1}^{\infty} \in \mathcal{N} \,:\, m_k = n\Big\} \\ &=\; V_n \times \mathbb{N}_1 \times \cdots \times \mathbb{N}_{k-1} \times \{n\} \times \mathbb{N}_{k+1} \times \cdots \end{aligned}$$

donde $\mathbb{N}_j = \mathbb{N}$ para todo $j \in \mathbb{N}$. Puesto que cada conjunto $\mathbb{N}_1 \times \cdots \times N_{k-1} \times \{n\} \times N_{k+1} \times \cdots$ es abierto en $\mathcal{N}$, resulta que $U(k,n)$ es abierto en $\mathbb{R} \times \mathcal{N}$ para todo $k, n \in \mathbb{N}$. Definamos ahora $U \subseteq \mathbb{R} \times \mathcal{N}$ como

$$\begin{aligned} U \;&=\; \bigcup_{(k,n) \in \mathbb{N} \times \mathbb{N}} U(k,n) \\ &=\; \Big\{\big(x,\, (n_k)_{k=1}^{\infty}\big) \in \mathbb{R} \times \mathcal{N} \,:\, x \in V_{n_k} \text{ para algún } k \in \mathbb{N}\Big\}. \end{aligned}$$

Entonces U es abierto en $\mathbb{R} \times \mathcal{N}$. Observe que si $y = (n_k)_{k=1}^{\infty} \in \mathcal{N}$, entonces de la definición de U se tiene que la y-sección de U viene dada por:

$$U^y \;=\; \Big\{x \in \mathbb{R} \,:\, \big(x, (n_k)_{k=1}^{\infty}\big) \in U\Big\} \;=\; \bigcup_{k=1}^{\infty} V_{n_k}.$$

Falta verificar que

$$\mathcal{O}_{\mathbb{R}} \;=\; \big\{U^y \,:\, y \in \mathcal{N}\big\}.$$

Para ver esto último, sea G un conjunto abierto arbitrario en $\mathbb{R}$. Puesto $(V_n)_{n=1}^{\infty}$ es una base para X, se tiene que $G = \bigcup_{k=1}^{\infty} V_{n_k}$ para alguna subsucesión $(V_{n_k})_{k=1}^{\infty}$ de $\mathcal{B}$. Luego, nuestro abierto será de la forma U^y, donde $y = (n_k)_{k=1}^{\infty} \in \mathcal{N}$. Esto prueba que $\mathcal{O}_{\mathbb{R}} \subseteq \big\{U^y \,:\, y \in \mathcal{N}\big\}$. La otra inclusión es consecuencia inmediata del hecho de que cada V_{n_k} es abierto y, por lo tanto, $U^y \in \mathcal{O}_{\mathbb{R}}$ para cada $y = (n_k)_{k=1}^{\infty} \in \mathcal{N}$. Esto termina la prueba de que U es universal para $\mathcal{O}_{\mathbb{R}}$.

Sea $U \subseteq \mathbb{R} \times \mathcal{N}$ el conjunto universal obtenido anteriormente. Entonces $F = (\mathbb{R} \times \mathcal{N}) \setminus U$ es cerrado en $\mathbb{R} \times \mathcal{N}$ y como

$$F^y \;=\; \mathbb{R} \setminus U^y \;=\; \bigcap_{k=1}^{\infty} V_{n_k}^c$$

para todo $y = (n_k)_{k=1}^{\infty} \in \mathcal{N}$, resulta que F es universal para la colección $\mathcal{F}$ de todos los subconjuntos cerrados de $\mathbb{R}$ y termina la prueba. ■

Observe que ninguna propiedad adicional de $\mathbb{R}$ fue usada en la demostración del resultado anterior, salvo el hecho de que $\mathbb{R}$ posee una base numerable, de modo que la conclusión de dicho teorema es válido si se reemplaza a $\mathbb{R}$ por cualquier espacio topológico que sea 2º numerable. En particular, puesto que $\mathbb{R} \times \mathcal{N}$ es 2º numerable, podemos encontrar un conjunto cerrado $F \subseteq (\mathbb{R} \times \mathcal{N}) \times \mathcal{N}$ que es universal para la clase de todos los subconjuntos cerrados de $\mathbb{R} \times \mathcal{N}$. Esta observación es clave para obtener el siguiente resultado el cual es fundamental para exhibir conjuntos analíticos que no son borelianos.

Teorema 6.3.33. *Existe un* ***conjunto cerrado*** A *en* $\mathbb{R} \times \mathcal{N}$ *que es* ***universal*** *para la familia* $\mathbf{\Sigma}_1^1$.

Prueba. Sea $\mathcal{F}$ la colección formada por todos los subconjuntos cerrados de $\mathbb{R} \times \mathcal{N}$. Por el Lema 6.3.32 existe un conjunto cerrado $F \subseteq (\mathbb{R} \times \mathcal{N}) \times \mathcal{N}$ que es universal para la clase $\mathcal{F}$. Por supuesto, esto significa que

$$\mathcal{F} = \Big\{ F^{y} \subseteq \mathbb{R} \times \mathcal{N} : y \in \mathcal{N} \Big\}. \tag{4}$$

Ahora, considere la función continua $f : (\mathbb{R} \times \mathcal{N}) \times \mathcal{N} \to \mathbb{R} \times \mathcal{N}$ dada por

$$f\big(x, (m_k)_{k=1}^{\infty}, (n_k)_{k=1}^{\infty}\big) = \big(x, (n_k)_{k=1}^{\infty}\big)$$

y sea $A = f(F)$. Resulta que el conjunto A es el candidato que buscamos. En efecto:

(i) A es cerrado. Sigue directamente de la definición de f.

(ii) $\mathbf{\Sigma}_1^1 = \{A^y : y \in \mathcal{N}\}$.

Para ver esto último, sea $C \in \mathbf{\Sigma}_1^1$. Por la condición (b) de la Definición 6.3.25, existe un conjunto cerrado $H \subseteq \mathbb{R} \times \mathcal{N}$ tal que $\operatorname{proy}_1(H) = C$, donde $\operatorname{proy}_1 : \mathbb{R} \times \mathcal{N} \to \mathbb{R}$ es la proyección sobre $\mathbb{R}$ dada por $\operatorname{proy}_1(x, y) = x$ para todo $x \in \mathbb{R}$ y todo $y \in \mathcal{N}$. Se sigue de (4) que $H = F^{y_0}$ para algún $y_0 \in \mathcal{N}$, de donde se obtiene que $C = \operatorname{proy}_1(F^{y_0})$.

Afirmamos que: $A^y = \operatorname{proy}_1\big(F^{y}\big)$ para cualquier $y \in \mathcal{N}$.

Fijemos $y = (n_k)_{k=1}^{\infty} \in \mathcal{N}$. Si $x \in A^y = \{x \in \mathbb{R} : (x, y) \in A\}$, entonces $\big(x, (n_k)_{k=1}^{\infty}\big) \in A$, lo cual significa que existe un $(m_k)_{k=1}^{\infty} \in \mathcal{N}$ tal que $\big(x, (m_k)_{k=1}^{\infty}, (n_k)_{k=1}^{\infty}\big) \in F$ y $f\big(x, (m_k)_{k=1}^{\infty}, (n_k)_{k=1}^{\infty}\big) = \big(x, (n_k)_{k=1}^{\infty}\big)$. Por otro lado, como $\big(x, (m_k)_{k=1}^{\infty}, (n_k)_{k=1}^{\infty}\big) \in F$, resulta que $\big(x, (m_k)_{k=1}^{\infty}\big) \in F^{y}$ y, por lo tanto, $x \in \operatorname{proy}_1(F^{y})$. Esto prueba que

$$A^y \subseteq \operatorname{proy}_1\big(F^{y}\big).$$

Recíprocamente, si $x \in \operatorname{proy}_1(F^{y})$, entonces existe $(m_k)_{k=1}^{\infty} \in \mathcal{N}$ tal que $\big(x, (m_k)_{k=1}^{\infty}, y\big) \in F$ y, en consecuencia, $\big(x, (m_k)_{k=1}^{\infty}\big) \in A$. Esto nos muestra que $x \in A^y$ y, así,

$$\operatorname{proy}_1\big(F^{y}\big) \subseteq A^y.$$

Por esto, $A^y = \operatorname{proy}_1\big(F^{y}\big)$ y termina la prueba de nuestra afirmación. Usando esto y el hecho de que $C = \operatorname{proy}_1(F^{y_0})$ se tiene que

$$\mathbf{\Sigma}_1^1 \subseteq \big\{A^y : y \in \mathcal{N}\big\}.$$

Para verificar la otra inclusión, sea $A^y \in \{A^y : y \in \mathcal{N}\}$. Como F^y es un subconjunto cerrado de $\mathbb{R} \times \mathcal{N}$, se sigue de la condición (c) de la Definición 6.3.25 que $\text{proy}_1(F^y)$ es analítico, es decir, $A^y = \text{proy}_1(F^y) \in \mathbf{\Sigma}_1^1$. Con esto queda establecido que $\mathbf{\Sigma}_1^1 = \{A^y : y \in \mathcal{N}\}$ y termina la prueba de nuestro teorema. ■

Puesto que $\mathcal{N}$ es un espacio métrico, completo y separable, podemos aplicar el Teorema 6.3.33 a éste espacio en lugar de $\mathbb{R}$ para obtener un conjunto en $\mathcal{N} \times \mathcal{N}$ que es universal para la familia de los subconjuntos analíticos de $\mathcal{N}$. Esta observación permite simplificar la demostración del siguiente resultado.

Teorema 6.3.34. *Existe un **conjunto analítico** que **no es medible** según Borel.*

Prueba. La demostración consistirá en construir un conjunto analítico A en $\mathbb{R}$ cuyo complemento no es analítico. La existencia de un tal conjunto A conduce, por supuesto, a concluir que A no puede ser un conjunto de Borel ya que el complemento de todo conjunto de Borel es de Borel. Por la observación del parágrafo anterior, seleccione un subconjunto cerrado F de $\mathcal{N} \times \mathcal{N}$ que es universal para la colección de todos los subconjuntos analíticos de $\mathcal{N}$ y sea D la diagonal en $\mathcal{N} \times \mathcal{N}$, esto es,

$$D = \{(y,y) : y \in \mathcal{N}\}.$$

Definamos ahora

$$A = \{y \in \mathcal{N} : (y,y) \in F \cap D\}.$$

Puesto que $F \cap D$ es cerrado en $\mathcal{N} \times \mathcal{N}$ y $\text{proy}_1(F \cap D) = A$ resulta, por la Definición 6.3.25, que A es analítico. Veamos que su complemento A^c no es analítico. Suponga, para generar una contradicción, que A^c es analítico. Puesto que F es universal para la familia de los conjuntos analíticos de $\mathcal{N}$, existe un $y_0 \in \mathcal{N}$ tal que la y_0-sección $F^{y_0} = A^c$ y es esta igualdad la que produce la contradicción. En efecto, nos preguntamos, ¿el elemento y_0, pertenece o no al conjunto A^c?. Aquí está la respuesta:

(a) Si $y_0 \in A^c$, entonces, por definición, $(y_0, y_0) \notin F$ y, por consiguiente, $y_0 \notin F^{y_0} = A^c$, es decir, $y_0 \notin A^c$, lo cual es absurdo.

(b) Si $y_0 \notin A^c$, entonces $(y_0, y_0) \in F$ y, en consecuencia, $y_0 \in F^{y_0} = A^c$, lo cual conduce, de nuevo, a una contradicción.

De esto se concluye que A^c no puede ser analítico y, en particular, A no puede ser un conjunto de Borel. ■

El hecho de que $\text{card}(\mathcal{N}) = 2^{\aleph_0} = \text{card}(\mathfrak{Bo}(\mathbb{R}))$, nos indica que ningún argumento usando la cardinalidad de ambos conjuntos puede ser utilizado para hacer visible la diferencia entre $\mathfrak{Bo}(\mathbb{R})$ y $\mathbf{\Sigma}_1^1$. Aunque la demostración de que $\mathfrak{Bo}(\mathbb{R})) \subsetneq \mathbf{\Sigma}_1^1$ es un tanto complicada, observe que en ella no se usó el Axioma de Elección. Es claro que un conjunto analítico que no sea boreliano debe ser bastante complicado y, por supuesto, muy extraño.

6.3.5. La Medida de Lebesgue

Una de las nociones más importantes en el desarrollo de la Integral de Lebesgue es la de Medida de Lebesgue. Antes, introduciremos una definición general.

Definición 6.3.35. *Sean* X *un conjunto y* $\mathcal{M}$ *una* σ*-álgebra de subconjuntos de* X. *Una* ***medida*** *sobre* $\mathcal{M}$ *es una función de conjuntos* $m : \mathcal{M} \to [0, +\infty]$ *que verifica las siguientes propiedades:*

(a) $m(\varnothing) = 0$, *y*

(b) m *es* ***numerablemente aditiva****, es decir, si* $(A_n)_{n=1}^{\infty}$ *es una sucesión disjunta en* $\mathcal{M}$, *entonces*

$$m\bigg(\bigcup_{n=1}^{\infty} A_n\bigg) \;=\; \sum_{n=1}^{\infty} m(A_n).$$

El hecho de que $\mathfrak{M}_\mu(\mathbb{R})$ es una σ-álgebra en combinación con el Teorema 6.3.37, permiten justificar la siguiente definición.

Definición 6.3.36 (Medida de Lebesgue). *La restricción de* μ^* *a* $\mathfrak{M}_\mu(\mathbb{R})$, *denotada por* μ, *se llama la* ***medida de Lebesgue*** *en* $\mathbb{R}$, *esto es,*

$$\mu \;=\; \mu^*\big|_{\mathfrak{M}_\mu(\mathbb{R})} : \mathfrak{M}_\mu(\mathbb{R}) \to [0, +\infty]$$

donde $\mu(E) = \mu^*(E)$ *para todo* $E \in \mathfrak{M}_\mu(\mathbb{R})$.

Similarmente, *la restricción de* μ *a la* σ*-álgebra de Borel* $\mathfrak{Bo}(\mathbb{R})$ *se llama la* ***medida de Borel*** *sobre* $\mathbb{R}$ *y será denotada, en lo sucesivo, por* μ_0.

Por definición, la medida de Lebesgue μ *hereda todas las propiedades* de la medida exterior μ^*. En particular,

(a) μ es **completa**, es decir, *si* $E \subseteq \mathbb{R}$ *con* $\mu(E) = 0$, *entonces* E *y todos sus subconjuntos son medibles según Lebesgue*. A diferencia de la medida de Lebesgue, la medida de Borel no es completa (véase el Ejercicio 6.7.1 (e), página 354).

(b) $\mu(E) \le \mu(F)$ si $E \subseteq F$ con $E, F \in \mathfrak{M}_\mu(\mathbb{R})$.

(c) $\mu(\{x\}) = 0$ para todo $x \in \mathbb{R}$.

Teorema 6.3.37 (μ es numerablemente aditiva). *Si* $(E_n)_{n=1}^{\infty}$ *es una* ***sucesión disjunta*** *de conjuntos en* $\mathfrak{M}_\mu(\mathbb{R})$, *entonces*

$$\mu\bigg(\bigcup_{n=1}^{\infty} E_n\bigg) \;=\; \sum_{n=1}^{\infty} \mu(E_n).$$

Prueba. Puesto que $\mathfrak{M}_\mu(\mathbb{R})$ es una σ-álgebra, tenemos que $\bigcup_{n=1}^{\infty} E_n \in \mathfrak{M}_\mu(\mathbb{R})$. Suponga, en primer lugar, que cada E_n es acotado. Por la subaditividad de μ, tenemos que

$$\mu\bigg(\bigcup_{n=1}^{\infty} E_n\bigg) \;\le\; \sum_{n=1}^{\infty} \mu(E_n), \tag{6.3.1}$$

de modo que nuestro objetivo es demostrar la otra desigualdad. Fijemos $\varepsilon > 0$. Puesto que $E_n \in \mathfrak{M}_\mu(\mathbb{R})$, se sigue del Corolario 6.3.9 que existe un conjunto cerrado $F_n \subseteq E_n$ tal que

$$\mu(E_n \setminus F_n) \;<\; \frac{\varepsilon}{2^n}.$$

Como cada conjunto F_n es cerrado y está incluido en el conjunto acotado E_n, resulta que los F_n son compactos y disjuntos dos a dos. Aquí es donde entra en escena el Corolario 6.2.16 que, junto a la monotonicidad de μ, nos dice que, para cualquier $k \in \mathbb{N}$,

$$\sum_{n=1}^{k} \mu(F_n) \;=\; \mu\Big(\bigcup_{n=1}^{k} F_n\Big) \;\le\; \mu\Big(\bigcup_{n=1}^{k} E_n\Big) \;\le\; \mu\Big(\bigcup_{n=1}^{\infty} E_n\Big).$$

Puesto que esta última desigualdad es válida para todo $k \in \mathbb{N}$, resulta que la sucesión de sumas parciales $\big(\sum_{n=1}^{k} \mu(F_n)\big)_{k=1}^{\infty}$, por ser creciente y acotada, converge gracias al Teorema 2.1.23, esto es,

$$\sum_{n=1}^{\infty} \mu(F_n) \;=\; \lim_{k\to\infty} \sum_{n=1}^{k} \mu(F_n) \;\le\; \mu\Big(\bigcup_{n=1}^{\infty} E_n\Big). \tag{6.3.2}$$

Observe que si $\mu\big(\bigcup_{n=1}^{\infty} E_n\big) = +\infty$, entonces de la desigualdad (6.3.1) se sigue que $\sum_{n=1}^{\infty} \mu(E_n) = +\infty$ y no hay nada más que demostrar. Suponga entonces que $\mu\big(\bigcup_{n=1}^{\infty} E_n\big) < +\infty$. Por (6.3.2) tenemos que $\sum_{n=1}^{\infty} \mu(F_n) < +\infty$ y, así,

$$\begin{aligned}
\sum_{n=1}^{\infty} \mu(E_n) \;&=\; \sum_{n=1}^{\infty} \mu\big(F_n \cup (E_n \setminus F_n)\big) \;\le\; \sum_{n=1}^{\infty} \Big(\mu(F_n) \;+\; \mu(E_n \setminus F_n)\Big)\\
&\le\; \sum_{n=1}^{\infty} \Big(\mu(F_n) \;+\; \frac{\varepsilon}{2^n}\Big) \;=\; \sum_{n=1}^{\infty} \mu(F_n) \;+\; \varepsilon\\
&\le\; \mu\Big(\bigcup_{n=1}^{\infty} E_n\Big) \;+\; \varepsilon.
\end{aligned}$$

Siendo ε arbitrario, se concluye que $\sum_{n=1}^{\infty} \mu(E_n) \le \mu\Big(\bigcup_{n=1}^{\infty} E_n\Big)$ y, por lo tanto,

$$\sum_{n=1}^{\infty} \mu(E_n) \;=\; \mu\Big(\bigcup_{n=1}^{\infty} E_n\Big). \tag{1}$$

Esto completa la prueba para el caso en que cada E_n es acotado. Para el caso general defina, para cada par de enteros $n, k \ge 1$, el conjunto

$$E_{n,k} \;=\; \big\{x \in E_n \;:\; k-1 \le |x| < k\big\}.$$

Entonces $\{E_{n,k} : n, k \in \mathbb{N}\}$ es una colección numerable de conjuntos medibles, acotados y disjuntos dos a dos. Además, para cada $n \in \mathbb{N}$,

$$E_n \;=\; \bigcup_{k=1}^{\infty} E_{n,k}.$$

Aplicando dos veces la igualdad (1), válida para conjuntos acotados, se obtiene que

$$\mu\Big(\bigcup_{n=1}^{\infty} E_n\Big) \;=\; \mu\Big(\bigcup_{n=1}^{\infty}\bigcup_{k=1}^{\infty} E_{n,k}\Big) \;=\; \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} \mu(E_{n,k}) \;=\; \sum_{n=1}^{\infty} \mu(E_n).$$

La prueba es completa. ■

La expresión μ es **σ-aditiva** también es frecuentemente utilizada como sinónimo de μ es numerablemente aditiva. Observe que si $\{E_1, \ldots, E_n\}$ es cualquier colección finita de conjuntos medibles disjuntos dos a dos, entonces si en el resultado anterior hacemos $E_{n+1} = E_{n+2} = \cdots = \varnothing$, resulta que

$$\mu\Big(\bigcup_{k=1}^{n} E_k\Big) \;=\; \sum_{k=1}^{n} \mu(E_k).$$

En este caso se dice que μ es ***finitamente aditiva*** .

Es importante destacar que la medida de Lebesgue es **σ-finita**, es decir, existe una sucesión $(E_n)_{n=1}^{\infty}$ de conjuntos medibles tal que

$$\mathbb{R} \;=\; \bigcup_{n=1}^{\infty} E_n \qquad \text{y} \qquad \mu(E_n) \;<\; +\infty, \quad n = 1, 2 \ldots$$

Por ejemplo, podemos escribir a $\mathbb{R}$ como $\mathbb{R} = \bigcup_{n=1}^{\infty}[-n, n]$ y observar que $\mu([-n, n]) < +\infty$ para todo $n \in \mathbb{N}$. Nótese que siempre se puede y, en ocasiones es conveniente hacerlo de esa forma, elegir la sucesión $(E_n)_{n=1}^{\infty}$ de conjuntos medibles en la definición anterior, o bien disjunta (véase (β_3) en la página 261), o bien creciente.

El Teorema 6.3.37 permite calcular la medida de *cualquier* conjunto abierto por medio del siguiente argumento. Sabemos que todo subconjunto abierto no vacío G de $\mathbb{R}$ se puede representar en la forma $G = \bigcup_{n=1}^{\infty} I_n$, donde $(I_n)_{n=1}^{\infty}$ es una sucesión de intervalos abiertos disjuntos dos a dos. Se sigue entonces del Teorema 6.3.37 que

$$\mu(G) \;=\; \sum_{n=1}^{\infty} \mu(I_n) \;=\; \sum_{n=1}^{\infty} \ell(I_n).$$

Corolario 6.3.38. *Sea $A \subseteq \mathbb{R}$ con $\mu^*(A) < +\infty$. Entonces existe un conjunto medible G de tipo $\mathcal{G}_\delta$ con la siguiente propiedad: si E es medible y $A \subseteq E \subseteq G$, entonces $\mu(G \setminus E) = 0$.*

Prueba. Por el Corolario 6.2.13 existe un G_δ medible G, tal que

$$A \subseteq G \qquad \text{y} \qquad \mu(G) = \mu^*(A) < +\infty.$$

Si E es medible y $A \subseteq E \subseteq G$, entonces $\mu^*(A) \leq \mu^*(E) \leq \mu^*(G)$, de donde se sigue que $\mu^*(A) = \mu(E) = \mu(G)$. Por esto y puesto que μ es finitamente aditiva se tiene que

$$\mu(E) \;=\; \mu(G) \;=\; \mu(G \cap E) + \mu(G \setminus E) \;=\; \mu(E) + \mu(G \setminus E).$$

Finalmente, el hecho de que $\mu(E) < +\infty$, conduce a que $\mu(G \setminus E) = 0$. ■

Al comienzo de la Sección 6.3 cuando introducimos la noción de conjunto medible, vimos que si $Z \notin \mathfrak{M}_\mu(\mathbb{R})$ con $\mu^*(Z) < +\infty$, entonces para cualquier conjunto medible G conteniendo a Z se cumplía que

$$\mu^*(G \setminus Z) \;>\; \mu(G) - \mu^*(Z).$$

Para conjuntos medibles, sin embargo, este hecho nunca ocurre como lo muestra el siguiente resultado.

Corolario 6.3.39. *Sea $E \in \mathfrak{M}_\mu(\mathbb{R})$. Si $F \subseteq E$ es **medible** con $\mu(F) < +\infty$, entonces*

$$\mu(E \setminus F) \;=\; \mu(E) \;-\; \mu(F).$$

Prueba. Nótese que como $E = F \cup (E \setminus E)$ es una unión disjunta de conjuntos medibles, de la aditividad finita de μ se sigue que $\mu(E) = \mu(F) + \mu(E \setminus F)$ y entonces $\mu(E \setminus F) = \mu(E) - \mu(F)$, gracias a que $\mu(F)$ es finita. ■

Observe que si F es un conjunto cerrado y acotado, entonces existe un intervalo compacto, digamos $J = [a, b]$, tal que $F \subseteq J$. Si $G = J \setminus F$, resulta que G es abierto y por lo anterior y el Corolario 6.3.39, se tiene que

$$\mu(F) \;=\; \ell(J) - \mu(G) \;=\; (b-a) - \sum_{n=1}^{\infty} \ell(I_n),$$

donde $\big(I_n\big)_{n=1}^{\infty}$ es una sucesión disjunta de intervalos abierto cuya unión es G.

El siguiente corolario es una de la buenas razones de por qué la medida de Lebesgue es exquisita: cualquier conjunto medible de medida finita se puede aproximar por un compacto.

Corolario 6.3.40 (Aproximación por compactos). *Sea $E \in \mathfrak{M}_\mu(\mathbb{R})$ con $\mu(E) < +\infty$. Entonces, para cada $\varepsilon > 0$, existe un **conjunto compacto** K tal que*

$$K \subseteq E \qquad\qquad y \qquad\qquad \mu(E \setminus K) \;<\; \varepsilon.$$

En particular,

$$\mu(E) \;=\; \sup\big\{\mu(K) \;:\; K \subseteq E,\; K \text{ es compacto}\big\}.$$

Prueba. Sea $\varepsilon > 0$. Por el Lema 6.3.10 existe un subconjunto medible y acotado $E_0 \subseteq E$ tal que $\mu(E \setminus E_0) < \varepsilon/2$. Más aun, por el Corolario 6.3.9, existe un conjunto cerrado $K \subseteq E_0$ y, por consiguiente, acotado, tal que $\mu(E_0 \setminus K) < \varepsilon/2$. Resulta que K es compacto y $K \subseteq E$. En particular, $\mu(K) \le \mu(E) < +\infty$. Por otro lado, como $E \setminus K = \big(E \setminus E_0\big) \cup \big(E_0 \setminus K\big)$ es la unión disjunta de dos conjuntos medibles, obtenemos, por la aditividad finita de μ, que

$$\mu(E \setminus K) \;=\; \mu(E \setminus E_0) \;+\; \mu(E_0 \setminus K) \;<\; \frac{\varepsilon}{2} \;+\; \frac{\varepsilon}{2} \;=\; \varepsilon$$

Finalmente, puesto que $\mu(K) < +\infty$, se obtiene del corolario anterior que $\mu(E) - \mu(K) = \mu(E \setminus K) < \varepsilon$ y, en consecuencia,

$$\mu(E) \;<\; \mu(K) \;+\; \varepsilon.$$

Esto prueba que $\mu(E) = \sup\big\{\mu(K) \;:\; K \subseteq E,\; K \text{ compacto}\big\}$. ■

Hemos visto, Corolario 6.2.6, que si A es cualquier subconjunto numerable de $\mathbb{R}$, entonces $\mu^*(A) = 0$ y habíamos anunciado que el recíproco de ese corolario era falso. En efecto, el conjunto ternario de Cantor Γ es un contraejemplo.

Corolario 6.3.41. *Sea Γ el **conjunto ternario de Cantor**. Entonces Γ es no-numerable con $\mu(\Gamma) = 0$.*

Prueba. Ya hemos demostrado que Γ es no-numerable. Para probar que Γ es de medida cero, recordemos que, para cada $n \in \mathbb{N}$, $J_n = \bigcup_{k=1}^{2^{n-1}} J_n(k)$, donde $J_n(1), \ldots, J_n(2^{n-1})$ son los intervalos abiertos disjuntos que fueron eliminados en la n-ésima etapa en la construcción de Γ, los cuales tienen, cada uno, longitud $1/3^n$. Luego,

$$\mu(J_n) \;=\; \ell\big(J_n(1)\big) + \cdots + \ell\big(J_n(2^{n-1})\big) \;=\; 2^{n-1}\frac{1}{3^n}.$$

Por esto, la longitud de todos los intervalos borrados es, gracias al Teorema 6.3.37,

$$\mu\Big(\bigcup_{n=1}^{\infty} J_n\Big) \;=\; \sum_{n=1}^{\infty}\mu(J_n) \;=\; \sum_{n=1}^{\infty}\frac{2^{n-1}}{3^n} \;=\; \frac{1}{3}\sum_{n=0}^{\infty}\Big(\frac{2}{3}\Big)^n \;=\; 1$$

y así, por el Corolario 6.3.39,

$$\mu(\Gamma) \;=\; \mu\Big([0,1] \setminus \bigcup_{n=1}^{\infty} J_n\Big) \;=\; \mu([0,1]) \,-\, \mu\Big(\bigcup_{n=1}^{\infty} J_n\Big) \;=\; 0.$$

La prueba es completa. ■

¿Cuál es la medida de la unión de dos conjuntos medibles que no son disjuntos? La respuesta viene dada por el siguiente resultado.

Corolario 6.3.42. *Sean $E, F \in \mathfrak{M}_\mu(\mathbb{R})$. Entonces*

$$\mu(E \cup F) \,+\, \mu(E \cap F) \;=\; \mu(E) \,+\, \mu(F).$$

Prueba. Si $\mu(E)$ o $\mu(F)$ es $+\infty$, entonces la igualdad es inmediata. Suponga entonces que tanto $\mu(E)$, así como $\mu(F)$, son finitos. Entonces $\mu(E \cap F) < +\infty$ y se sigue del Corolario 6.3.39 que $\mu\big(F \setminus (E \cap F)\big) = \mu(F) - \mu(E \cap F)$. Por esto,

$$\begin{aligned} \mu(E \cup F) \;&=\; \mu\Big(E \cup \big(F \setminus (E \cap F)\big)\Big) \\ &=\; \mu(E) \,+\, \mu\big(F \setminus (E \cap F)\big) \\ &=\; \mu(E) \,+\, \mu(E) \,-\, \mu(E \cap F) \end{aligned}$$

y termina la prueba. ■

El Teorema 6.3.37 sirve, como se muestra a continuación, para obtener otro criterio para que la medida exterior de la unión de una sucesión disjunta de subconjuntos de $\mathbb{R}$ sea igual a la suma de las medidas exteriores de los conjuntos de la sucesión.

Corolario 6.3.43. *Sea $(A_n)_{n=1}^{\infty}$ una sucesión arbitraria de subconjuntos de $\mathbb{R}$. Si existe una* ***sucesión disjunta*** *$(E_n)_{n=1}^{\infty}$ en $\mathfrak{M}_\mu(\mathbb{R})$ con $A_n \subseteq E_n$ para todo $n \in \mathbb{N}$, entonces*

$$\mu^*\Big(\bigcup_{n=1}^{\infty} A_n\Big) \;=\; \sum_{n=1}^{\infty}\mu^*(A_n).$$

Prueba. Seleccionemos, haciendo uso del Corolario 6.2.13, de un conjunto G_δ, digamos G, tal que

$$G \supseteq \bigcup_{n=1}^{\infty} A_n \qquad \text{y} \qquad \mu(G) = \mu^*\Big(\bigcup_{n=1}^{\infty} A_n\Big).$$

Puesto que $G \in \mathfrak{M}_\mu(\mathbb{R})$, resulta que para cada $n \geq 1$, el conjunto $E'_n = G \cap E_n$ es medible y $A_n \subseteq E'_n$. Siendo $(E'_n)_{n=1}^{\infty}$ una sucesión disjunta de conjuntos medibles según Lebesgue con $\bigcup_{n=1}^{\infty} E'_n \subseteq G$, se tiene, por la monotonía de μ^* y el Teorema 6.3.37, que

$$\mu^*\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \mu(G) \geq \mu\Big(\bigcup_{n=1}^{\infty} E'_n\Big) = \sum_{n=1}^{\infty} \mu(E'_n) \geq \sum_{n=1}^{\infty} \mu^*(A_n).$$

La desigualdad contraria $\mu^*\big(\bigcup_{n=1}^{\infty} A_n\big) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$ sigue del Teorema 6.2.5 y termina la prueba. ■

Uno de los rasgos sobresalientes que debe poseer cualquier medida que extienda la longitud de los intervalos de $\mathbb{R}$ es que ella sea invariante por traslación, es decir, que la medida de cualquier conjunto y su trasladado sean iguales.

Teorema 6.3.44 (μ es invariante por traslación). *Sea E un subconjunto de $\mathbb{R}$. La siguientes condiciones son equivalentes:*

(1) $E \in \mathfrak{M}_\mu(\mathbb{R})$.

(2) $x + E \in \mathfrak{M}_\mu(\mathbb{R})$ para todo $x \in \mathbb{R}$.

Además, se cumple que

$$\mu(E) = \mu(x + E) \quad \textit{para todo} \quad x \in \mathbb{R}.$$

Prueba. Suponga que $E \in \mathfrak{M}_\mu(\mathbb{R})$ y sea $\varepsilon > 0$. Entonces existe un conjunto abierto $G \supseteq E$ tal que $\mu^*(G \setminus E) \leq \varepsilon$. Fijemos $x \in \mathbb{R}$. Claramente $x + G$ es un conjunto abierto, $x + G \supseteq x + E$ y como

$$(x + G) \setminus (x + E) \subseteq x + (G \setminus E)$$

resulta, por la propiedad de ser μ^* invariante por traslación, que

$$\mu\big((x + G) \setminus (x + E)\big) \leq \mu\big(x + (G \setminus E)\big) = \mu^*\big(x + (G \setminus E)\big) = \mu^*(G \setminus E) \leq \varepsilon.$$

Esto prueba que $x + E \in \mathfrak{M}_\mu(\mathbb{R})$. Suponga ahora que $x + E \in \mathfrak{M}_\mu(\mathbb{R})$ para todo $x \in \mathbb{R}$. Fijemos un $x \in \mathbb{R}$. Entonces, de la primera parte, se tiene que $E = -x + (x + E)$ es medible. La igualdad $\mu(E) = \mu(x + E)$ sigue del Teorema 6.2.4 y termina la prueba. ■

Recordemos, Teorema 2.1.44, que si $\big(E_n\big)_{n=1}^{\infty}$ es una sucesión monótona creciente, entonces

$$\lim_{n \to \infty} E_n = \bigcup_{n=1}^{\infty} E_n.$$

Los siguientes dos teoremas establecen que la medida de Lebesgue es *continua* en el sentido de que si una sucesión monótona de conjuntos medibles converge, entonces la sucesión de sus medidas también converge. Esos dos hechos jugarán un papel importante en nuestro desarrollo.

Teorema 6.3.45 (μ es continua). *Sea $(E_n)_{n=1}^{\infty}$ una sucesión **monótona creciente** en $\mathfrak{M}_\mu(\mathbb{R})$. Entonces*

$$\mu\big(\lim_{n\to\infty} E_n\big) \;=\; \mu\Big(\bigcup_{n=1}^{\infty} E_n\Big) \;=\; \lim_{n\to\infty} \mu(E_n).$$

Prueba. Si para algún $N \in \mathbb{N}$ ocurre que $\mu(E_N) = +\infty$, entonces

$$\mu\Big(\bigcup_{n=1}^{\infty} E_n\Big) \;\geq\; \mu(E_N) \;=\; +\infty$$

y puesto que $E_1 \subseteq E_2 \subseteq \cdots \subseteq E_N \subseteq \cdots$ tendremos que $\mu(E_n) \geq \mu(E_N) = +\infty$ para todo $n \geq N$ y, en consecuencia, $\mu\big(\bigcup_{n=1}^{\infty} E_n\big) = \lim_{n\to\infty} \mu(E_n) = +\infty$.

Suponga ahora que $\mu(E_n) < +\infty$ para todo $n \in \mathbb{N}$. Como $E_1 \subseteq E_2 \subseteq \cdots \subseteq E_n \subseteq \cdots$, podemos escribir

$$\bigcup_{n=1}^{\infty} E_n \;=\; E_1 \cup (E_2 \setminus E_1) \cup \cdots \cup (E_{n+1} \setminus E_n) \cup \cdots$$

lo cual es una unión numerable y disjunta de conjuntos medibles. Usando el hecho de que μ es numerablemente aditiva y el Corolario 6.3.39, tenemos que

$$\begin{aligned}
\mu\Big(\bigcup_{n=1}^{\infty} E_n\Big) &= \mu(E_1) \;+\; \sum_{n=1}^{\infty} \mu(E_{n+1} \setminus E_n) \\
&= \mu(E_1) \;+\; \lim_{n\to\infty} \sum_{k=1}^{n} \Big(\mu(E_{k+1}) - \mu(E_k)\Big) \\
&= \lim_{n\to\infty} \mu(E_{n+1}).
\end{aligned}$$

Esto finaliza la prueba. ■

En general, el resultado anterior se puede usar para demostrar que μ^* también es continua.

Corolario 6.3.46 (μ^* es continua). *Sea $(E_n)_{n=1}^{\infty}$ una sucesión **monótona creciente** de subconjuntos de $\mathbb{R}$ no necesariamente medibles. Entonces*

$$\mu^*\Big(\bigcup_{n=1}^{\infty} E_n\Big) \;=\; \lim_{n\to\infty} \mu^*(E_n).$$

Prueba. Claramente $\lim_{n\to\infty} \mu^*(E_n) \leq \mu^*\big(\bigcup_{n=1}^{\infty} E_n\big)$. Para demostrar la otra desigualdad usemos el Corolario 6.3.38 para seleccionar, por cada $n \in \mathbb{N}$, un conjunto G_δ, digamos $G_n \supseteq E_n$ tal que $\mu(G_n) = \mu^*(E_n)$. Observe que si reemplazamos G_n por $\bigcap_{m=n}^{\infty} G_m$, podemos asumir que $(G_n)_{n=1}^{\infty}$ es una sucesión monótona creciente de conjuntos medibles y, así, gracias al Teorema 6.3.45,

$$\begin{aligned}
\lim_{n\to\infty} \mu^*(E_n) &= \lim_{n\to\infty} \mu(G_n) \;=\; \mu\Big(\bigcup_{n=1}^{\infty} G_n\Big) \\
&= \mu^*\Big(\bigcup_{n=1}^{\infty} G_n\Big) \;\geq\; \mu^*\Big(\bigcup_{n=1}^{\infty} E_n\Big).
\end{aligned}$$

Esto termina la prueba. ■

Teorema 6.3.47 (μ es continua). *Sea $(E_n)_{n=1}^{\infty}$ una sucesión monótona **decreciente** en $\mathfrak{M}_\mu(\mathbb{R})$. Si para algún $n_0 \in \mathbb{N}$, $\mu(E_{n_0}) < +\infty$, entonces*

$$\mu\big(\lim_{n\to\infty} E_n\big) \;=\; \mu\bigg(\bigcap_{n=1}^{\infty} E_n\bigg) \;=\; \lim_{n\to\infty} \mu(E_n).$$

Prueba. Sin perder generalidad, asumiremos que $\mu(E_1) < +\infty$. Observe, en primer lugar, que $\lim_{n\to\infty} \mu(E_n)$ existe y es finito. En efecto, como $E_1 \supseteq E_2 \supseteq \cdots \supseteq E_n \supseteq \cdots$ y $\mu(E_1) < +\infty$, entonces

$$+\infty \;>\; \mu(E_1) \;\geq\; \mu(E_2) \;\geq\; \cdots \;\geq\; \mu(E_n) \;\geq\; \cdots \;\geq\; 0$$

y, por lo tanto, por ser la sucesión $(\mu(E_n))_{n=1}^{\infty}$ decreciente y acotada, ella converge. Por esto, $\lim_{n\to\infty} \mu(E_n) \leq \mu(E_1)$. Sea $H = \bigcap_{n=1}^{\infty} E_n$ y pongamos $F_n = E_n \setminus E_{n+1}$ para $n = 1, 2, \ldots$ Resulta que los conjuntos F_n son medibles y disjuntos dos a dos. Más aun, como $E_1 \setminus H = \bigcup_{n=1}^{\infty} F_n$ se tiene, por el Teorema 6.3.47 y el Corolario 6.3.39, que

$$\begin{aligned}
\mu(E_1) - \mu(H) \;&=\; \mu(E_1 \setminus H) \;=\; \mu\bigg(\bigcup_{n=1}^{\infty} F_n\bigg)\\
&=\; \sum_{n=1}^{\infty} \mu(F_n) \;=\; \lim_{n\to\infty} \sum_{k=1}^{n} \Big(\mu(E_{k+1}) - \mu(E_k)\Big)\\
&=\; \mu(E_1) \;-\; \lim_{n\to\infty} \mu(E_n).
\end{aligned}$$

Finalmente, como $\mu(E_1) < +\infty$ se concluye que $\mu\big(\bigcap_{n=1}^{\infty} E_n\big) = \lim_{n\to\infty} \mu(E_n)$. ■

Nota Adicional 6.3.2 Es importante destacar que la hipótesis $\mu(E_{n_0}) < +\infty$ en el Teorema 6.3.47 no se puede suprimir ya el resultado pudiera no ser cierto. Por ejemplo, si para cada $n \in \mathbb{N}$, definimos $E_n = [n, +\infty)$, entonces $(E_n)_{n=1}^{\infty}$ es una sucesión decreciente de subconjuntos medibles para la cual $\mu(E_n) = +\infty$ para todo $n \in \mathbb{N}$. En particular, $\lim_{n\to\infty} \mu(E_n) = +\infty$. Por otro lado, $\bigcap_{n=1}^{\infty} E_n = \varnothing$ y, en consecuencia, $\mu(\bigcap_{n=1}^{\infty} E_n) = 0$.

6.3.6. El Lema de Borel-Cantelli

En esta sección mostraremos cómo las nociones de límites superior e inferior de sucesiones de conjuntos medibles permiten demostrar algunos resultados no triviales en la Teoría de la Medida. Recordemos que dada una sucesión $(E_n)_{n=1}^{\infty}$ de subconjuntos de $\mathbb{R}$, los conjuntos

$$\varliminf_{n\to\infty} E_n \;=\; \bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} E_k \qquad \text{y} \qquad \varlimsup_{n\to\infty} E_n \;=\; \bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} E_k$$

son llamados los límites inferior y superior, respectivamente, de la sucesión $(E_n)_{n=1}^{\infty}$.

Lema 6.3.48. *Sea $(E_n)_{n=1}^{\infty}$ una sucesión de **conjuntos medibles**. Entonces*

(a) $\mu\Big(\varliminf_{n\to\infty} E_n\Big) \leq \varliminf_{n\to\infty} \mu(E_n)$.

(b) $\mu\Big(\varlimsup_{n\to\infty} E_n\Big) \geq \varlimsup_{n\to\infty} \mu(E_n)$ *siempre que* $\mu\big(\bigcup_{n=1}^{\infty} E_n\big) < +\infty$.

Prueba. (a) Para cada $n \geq 1$, sea $G_n = \bigcap_{k=n}^{\infty} E_k$. Puesto que $(G_n)_{n=1}^{\infty}$ es una sucesión monótona creciente, se sigue del Teorema 6.3.45 que

$$\mu\Big(\varliminf_{n\to\infty} E_n\Big) \;=\; \mu\Big(\bigcup_{n=1}^{\infty} G_n\Big) \;=\; \lim_{n\to\infty} \mu(G_n).$$

Fijemos un $n \in \mathbb{N}$ y observe que $G_n \subseteq E_{n+k}$ y $\mu(G_n) \leq \mu(E_{n+k})$ para todo $k \in \mathbb{N}$. De esto se sigue que

$$\mu(G_n) \;\leq\; \varliminf_{k\to\infty} \mu(E_{n+k}) \;=\; \varliminf_{k\to\infty} \mu(E_k)$$

para cada $n \in \mathbb{N}$ y, por consiguiente,

$$\mu\Big(\varliminf_{n\to\infty} E_n\Big) \;\leq\; \varliminf_{n\to\infty} \mu(E_n).$$

(b) La prueba es similar a la del caso anterior usando, esta vez, el Teorema 6.3.47. ■

Nota Adicional 6.3.3 De nuevo, la condición de que $\mu(\bigcup_{n=1}^{\infty} E_n) < +\infty$, en el lema anterior, no se puede omitir. En efecto, para la sucesión de conjuntos $(E_n)_{n=1}^{\infty}$, donde $E_n = [n, n+1)$, se cumple que $\mu(\bigcup_{n=1}^{\infty} E_n) = +\infty$, $\varlimsup_{n\to\infty} \mu(E_n) = 1$, pero $\mu\big(\varlimsup_{n\to\infty} E_n\big) = 0$.

Estamos listo para presentar el celebre Lema de Borel-Cantelli. Sus aplicaciones son de tal importancia en la Teoría de Probabilidades que el Dr. Tapas Kumar Chandra ha escrito un libro, **The Borel-Cantelli Lemma**, [33], dedicado exclusivamente a mostrar algunas de las aplicaciones de dicho lema.

Lema 6.3.49 (Borel-Cantelli). *Sea $(E_n)_{n=1}^{\infty} \subseteq \mathfrak{M}_\mu(\mathbb{R})$. Si $\sum_{n=1}^{\infty} \mu(E_n) < +\infty$, entonces*

$$\mu\Big(\varlimsup_{n\to\infty} E_n\Big) \;=\; 0.$$

Prueba. Para cada $n \in \mathbb{N}$, defina $F_n = \bigcup_{k=n}^{\infty} E_k$. Resulta entonces que $(F_n)_{n=1}^{\infty}$ es una sucesión decreciente de conjuntos medibles con $\mu(F_1) < +\infty$. Se sigue ahora del Teorema 6.3.47 que

$$\mu\Big(\bigcap_{n=1}^{\infty} F_n\Big) \;=\; \lim_{n\to\infty} \mu(F_n) \;=\; 0,$$

y como $\displaystyle\varlimsup_{n\to\infty} E_n = \bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} E_n = \bigcap_{n=1}^{\infty} F_n$, entonces $\mu\big(\varlimsup_{n\to\infty} E_n\big) = 0$. ■

Al resultado anterior lo llamaremos el *primer Lema de Borel-Cantelli*. Observe que la conclusión de este resultado también se puede enunciar del modo siguiente: *la colección de los puntos $x \in \mathbb{R}$ que pertenecen a infinitos E_n tiene medida cero*. El primer Lema de Borel-Cantelli es un caso especial de un famoso teorema llamado el Teorema de Beppo Levi o Teorema de la Convergencia Monótona que veremos más adelante cuando estudiemos la convergencia de sucesiones de funciones integrables según Lebesgue. El *segundo Lema de Borel-Cantelli* es, en cierto sentido, el recíproco del primero bajo una hipótesis adicional: la independencia de los conjuntos. Él es consecuencia de la Ley Fuerte de los Grandes Números de Kolmogorov para una sucesión de variables aleatorias independientes.

Definición 6.3.50. *Una colección finita* $\{E_n : n = 1, \ldots, N\}$ *de conjuntos medibles se dice que es* ***independiente*** *si*

$$\mu(E_{n_1} \cap \cdots \cap E_{n_k}) \;=\; \mu(E_{n_1}) \cdot \ldots \cdot \mu(E_{n_k})$$

para todo $k \leq N$, $n_1 < \cdots < n_k \leq N$. *Una colección infinita numerable* $(E_n)_{n=1}^{\infty}$ *de conjuntos medibles se llama* ***independiente*** *si cada subcolección finita de ella es independiente.*

Lema 6.3.51 (Borel-Cantelli). *Sea* $(E_n)_{n=1}^{\infty}$ *una sucesión de* ***conjuntos independientes*** *incluidos en* $[0,1]$. *Si* $\sum_{n=1}^{\infty} \mu(E_n) = +\infty$, *entonces*

$$\mu\Big(\overline{\lim_{n\to\infty}}\, E_n\Big) \;=\; 1$$

Prueba. Gracias a las Leyes de Morgan, tenemos que

$$\big(\overline{\lim_{n\to\infty}}\, E_n\big)^c \;=\; \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} E_k^c,$$

de modo que para demostrar que $\mu\big(\overline{\lim_{n\to\infty}}\, E_n\big) = 1$ será suficiente probar que

$$\mu\Big(\big(\overline{\lim_{n\to\infty}}\, E_n\big)^c\Big) \;=\; \mu\Big(\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} E_k^c\Big) \;=\; 0,$$

o lo que es lo mismo, ver que $\mu\big(\bigcap_{k=n}^{\infty} E_k^c\big) = 0$ para todo $n \geq 1$. Fijemos un tal $n \geq 1$. Ahora bien, para cada $m \geq 1$, los conjuntos $E_n, E_{n+1}, \ldots, E_{n+m}$ son independientes y como $1 - x \leq e^{-x}$, resulta entonces que

$$\mu\Big(\bigcap_{k=n}^{\infty} E_k^c\Big) \;\leq\; \mu\Big(\bigcap_{k=n}^{n+m} E_k^c\Big) \;=\; \prod_{k=n}^{n+m} \big(1 - \mu(E_k)\big) \;\leq\; \prod_{k=n}^{n+m} e^{-\mu(E_k)}$$

Por otro lado, puesto que $\sum_{n=1}^{\infty} \mu(E_n) = +\infty$, entonces $\sum_{k=n}^{\infty} \mu(E_k) = +\infty$, de modo que $\sum_{k=n}^{n+m} \mu(E_k) \to +\infty$ cuando $m \to \infty$. Finalmente, como $\lim_{x\to\infty} e^{-x} = 0$, se sigue que el producto $\prod_{k=n}^{n+m} e^{-\mu(E_k)} \to 0$ cuando $m \to \infty$ (véase Ejercicios 3.2 (3), página 191). ■

Uno puede obtener la misma conclusión del segundo Lema de Borel-Cantelli sin ser la sucesión independiente, pero imponiendo otra condición.

Ejemplo 6.3.2. *Sea* $(E_n)_{n=1}^{\infty}$ *una sucesión de conjuntos medibles en* $[0,1]$ *con la siguiente propiedad: para cada* $A \in \mathfrak{M}_\mu(\mathbb{R})$,

$$\mu(A) \;>\; 0 \quad \Rightarrow \quad \sum_{n=1}^{\infty} \mu(A \cap E_n) \;=\; +\infty.$$

Entonces $\mu\big(\overline{\lim_{n\to\infty}}\, E_n\big) = 1$.

Prueba. Es suficiente demostrar, gracias al Teorema 6.3.47, que

$$\mu\Big(\bigcup_{k=m}^{\infty} E_k\Big) \;=\; 1 \quad \text{para todo} \quad m \geq 1$$

Suponga que esto es falso. Entonces existe un entero $m > 1$ tal que

$$\mu\Big(\bigcup_{k=m}^{\infty} E_k\Big) \ < \ 1.$$

Pongamos $A = \bigcap_{k=m}^{\infty} E_k^c$. Entonces $\mu(A) > 0$ pero

$$\sum_{n=1}^{\infty} \mu(A \cap E_n) \ = \ \sum_{n=1}^{m-1} \mu(A \cap E_n) \ < \ +\infty,$$

lo que contradice nuestra hipótesis. ■

La siguiente desigualdad, demostrada por Ivan D. Arandelović [6], permite obtener algunos resultados clásicos importantes de manera sencilla.

Teorema 6.3.52 (Arandelović). *Sea $E \in \mathfrak{M}_\mu(\mathbb{R})$ con $\mu(E) > 0$ y sea $(x_n)_{n=1}^{\infty}$ una sucesión **acotada** de números reales. Entonces*

$$\mu(E) \ \leq \ \mu\Big(\overline{\lim_{n\to\infty}}\,(x_n + E)\Big).$$

Prueba. Sea K un conjunto compacto contenido en E. Puesto que $x_n + K \subseteq x_n + E$ para todo $n \in \mathbb{N}$, se sigue de la relación $\overline{\lim_{n\to\infty}}\,(x_n + K) \subseteq \overline{\lim_{n\to\infty}}\,(x_n + E)$ que

$$\mu\Big(\overline{\lim_{n\to\infty}}\,(x_n + K)\Big) \ \leq \ \mu\Big(\overline{\lim_{n\to\infty}}\,(x_n + E)\Big).$$

Puesto que μ es invariante por traslación, resulta que $\mu(K) = \mu(x_n + K)$ para todo $n \in \mathbb{N}$ y, por consiguiente,

$$\mu(K) \ = \ \lim_{n\to\infty} \mu(x_n + K) \ = \ \overline{\lim_{n\to\infty}}\, \mu(x_n + K).$$

Ahora bien, como la sucesión $(x_n)_{n=1}^{\infty}$ es acotada y K es compacto, resulta que el conjunto $\bigcup_{n=1}^{\infty}(x_n + K)$ es acotado y, por lo tanto,

$$\mu\Big(\bigcup_{n=1}^{\infty}(x_n + K)\Big) \ < \ +\infty.$$

Invocando el Lema 6.3.48 (b), vemos que

$$\mu(K) \ = \ \overline{\lim_{n\to\infty}}\, \mu(x_n + K) \ \leq \ \mu\Big(\overline{\lim_{n\to\infty}}\,(x_n + K)\Big) \ \leq \ \mu\Big(\overline{\lim_{n\to\infty}}\,(x_n + E)\Big).$$

Sabemos, Corolario 6.3.40, que $\mu(E) = \sup\{\mu(K) : K \subseteq E, K \text{ compacto}\}$ y como el compacto $K \subseteq E$ fue elegido de modo arbitrario, se sigue que

$$0 \ < \ \mu(E) \ \leq \ \mu\Big(\overline{\lim_{n\to\infty}}\,(x_n + E)\Big).$$

La prueba es completa. ■

La conclusión más importante que hay que destacar del resultado anterior es que el conjunto $\limsup(x_n + E) \neq \varnothing$, de hecho, *es no-numerable*. Las siguientes dos aplicaciones del teorema anterior son bien conocidas y con variadas demostraciones.

Corolario 6.3.53 (H. Steinhaus). *Sea $E \in \mathfrak{M}_\mu(\mathbb{R})$ con $\mu(E) > 0$. Entonces existe una sucesión de puntos distintos $(x_n)_{n=1}^{\infty}$ en E tal que $|x_n - x_m|$ es un* ***número racional*** *para cualesquiera $m, n \in \mathbb{N}$ con $m \neq n$.*

Prueba. Sea $(r_n)_{n=1}^{\infty}$ una sucesión arbitraria pero acotada de números racionales tal que $r_n \neq r_m$ para todo $m, n \in \mathbb{N}$ con $m \neq n$. Por el Teorema de Arandelović se tiene que

$$0 \;<\; \mu(E) \;\leq\; \mu\Big(\overline{\lim_{n\to\infty}}\,(r_n + E)\Big),$$

de donde se sigue que el conjunto $\limsup(r_n + E)$ es no vacío. Fijemos un elemento $x \in \limsup(r_n + E)$. Por el Teorema 2.1.41 sabemos que existe una cantidad infinita de términos de la sucesión $(r_n + E)_{n=1}^{\infty}$ que contienen a x, es decir, existe una sucesión de enteros positivos $(n_j)_{j=1}^{\infty}$ tal que $x \in r_{n_j} + E$ para todo $j = 1, 2, \ldots$. Por cada $j \in \mathbb{N}$, seleccionemos un punto x_j en E de modo tal que

$$x \;=\; r_{n_j} \;+\; x_j$$

Entonces, para cualesquiera $i \neq j$, se tiene que $|x_i - x_j| \;=\; |r_{n_i} - r_{n_j}| \;\neq\; 0$ y, en consecuencia, $x_i \neq x_j$ para $i \neq j$. De aquí se concluye que

$$|x_i - x_j| \;=\; |r_{n_i} - r_{n_j}| \;\in \mathbb{Q}$$

y termina la prueba. ■

Nota Adicional 6.3.4 El argumento clásico para demostrar el corolario anterior es como sigue. Suponga, en primer lugar, que el conjunto medible E de medida positiva es acotado, digamos $E \subseteq [a, b]$. El siguiente razonamiento permite encontrar un par de elementos distintos $x, y \in E$ tal que $|x - y|$ es un racional. Sea $\{r_1, r_2, \ldots\}$ una enumeración, sin repetición, de los números racionales en $[a, b]$ y para cada $n \in \mathbb{N}$, sea $E_n = r_n + E$. Como cada E_n es medible y μ es invariante por traslación, tenemos que $\mu(E_n) = \mu(E)$ para todo $n \in \mathbb{N}$. Afirmamos que existen $m_1 \neq n_1$ en $\mathbb{N}$ tal que $E_{m_1} \cap E_{n_1} \neq \varnothing$. En efecto, si ocurre que $E_m \cap E_n = \varnothing$ para todo $m \neq n$, entonces

$$\mu\Big(\bigcup_{n=1}^{\infty} E_n\Big) \;=\; \sum_{n=1}^{\infty} \mu(E_n) \;=\; \sum_{n=1}^{\infty} \mu(E) \;=\; +\infty,$$

y como $\bigcup_{n=1}^{\infty} E_n \subseteq [a, 2b]$, resulta que $\mu(\bigcup_{n=1}^{\infty} E_n) \leq 2b - a < +\infty$, lo que origina una contradicción. Esto establece nuestra afirmación. Sea $z \in E_{m_1} \cap E_{n_1}$. Entonces existen $x, y \in E$ tales que $z = r_{m_1} + x = r_{n_1} + y$ y, en consecuencia,

$$|x - y| \;=\; |r_{m_1} - r_{n_1}| \;\in\; \mathbb{Q}.$$

Para demostrar el caso general, sea $I_n = [-n, n]$ para cada $n \in \mathbb{N}$ y definamos

$$E_1 \;=\; E \cap I_1 \qquad\qquad \text{y} \qquad\qquad E_n \;=\; E \cap (I_n \setminus I_{n-1}) \quad \text{para } n = 2, 3, \ldots$$

Como E_n es medible y $0 < \mu(E_n) < +\infty$, se sigue del primer caso que existen $x_n \neq y_n$ en E_n tal que $|x_n - y_n|$ es racional. Más aun, puesto que $E = \bigcup_{n=1}^{\infty} E_n$ es una unión disjunta, los elementos en la sucesión $(x_n)_{n=1}^{\infty} \subseteq E$ son distintos dos a dos y se cumple que $|x_n - y_n|$ es un racional para cada $n \in \mathbb{N}$.

Aunque un conjunto *acotado* E con *medida positiva* puede tener interior vacío (por ejemplo, los irracionales en $[0,1]$, o cualquier conjunto tipo-Cantor de medida positiva posee tal propiedad), lo que resulta sorprendente, sin embargo, es que el conjunto $E-E$ siempre tiene interior no-vacío. Este es el contenido del siguiente resultado el cual posee, como veremos en el transcurso de estas notas, demostraciones muy variadas.

Teorema 6.3.54 (Steinhaus). *Sea $E \in \mathfrak{M}_\mu(\mathbb{R})$ con $\mu(E) > 0$. Entonces $E-E$* ***contiene un entorno abierto*** *G del cero. En particular,*

$$E \cap (g+E) \neq \varnothing$$

para todo $g \in G$.

Prueba. Puesto que $E = \bigcup_{n=1}^{\infty}(E \cap [-n,n])$ no se pierde generalidad en asumir que E es acotado. Usemos el Corolario 6.3.40 para escribir $\mu(E) = \sup\{\mu(K) : K \subseteq E,\ K \text{ compacto}\} > 0$. Seleccione ahora un compacto $K \subseteq E$ con $\mu(K) > 0$. Puesto que $K-K \subseteq E-E$, será suficiente demostrar que $K-K$ posee un entorno abierto del cero. Suponga, para generar una contradicción, que $K-K$ no contiene ningún entorno abierto del cero. Puesto que $K-K$ es un conjunto compacto y $0 \in K-K$, entonces existe una sucesión $(x_n)_{n=1}^{\infty}$ en $\mathbb{R}$ tal que

$$\lim_{n\to\infty} x_n \;=\; 0 \qquad\qquad \text{y} \qquad\qquad \{x_n : n \in \mathbb{N}\} \cap (K-K) \;=\; \varnothing. \tag{1}$$

Por el Teorema 6.3.52,

$$0 \;<\; \mu(K) \;\leq\; \mu\Big(\varlimsup_{n\to\infty}(K-x_n)\Big)$$

y, en consecuencia, el conjunto $\limsup(K-x_n)$ es no vacío. Seleccionemos un punto de dicho conjunto, digamos, $x \in \limsup(K-x_n)$. Esto significa que $x + x_n \in K$ para infinitos valores de n, lo cual es equivalente a afirmar que existe una sucesión estrictamente creciente $(n_j)_{j=1}^{\infty}$ de enteros positivos tal que $x + x_{n_j} \in K$ para todo $j \in \mathbb{N}$. Como $\lim_{j\to\infty} x_{n_j} = 0$ y K es cerrado, se concluye que $x \in K$. Escojamos ahora, por cada $j \in \mathbb{N}$, un punto $a_j \in K$ tal que $x + x_{n_j} = a_j$. De lo anterior se obtiene que

$$x_{n_j} \;=\; a_j \,-\, x \;\in\; K-K$$

lo cual es imposible por (1). Esta contradicción nos asegura que $K-K$ posee un entorno abierto del cero. Para demostrar la última parte, sea G un entorno abierto del 0 incluido en $E-E$ y sea $g \in G$. Entonces existen $v_1, v_2 \in E$ tal que $g = v_1 - v_2$. De aquí se sigue que

$$v_1 \;=\; g \,+\, v_2 \;\in\; g \,+\, E,$$

lo cual prueba que $v_1 \in E \cap (g+E)$ y termina la prueba. ■

Nota Adicional 6.3.5 Es interesante observar que la conclusión del resultado de Steinhaus no es una propiedad exclusiva de los conjuntos medibles con medida positiva. En efecto, recordemos que el conjunto ternario de Cantor Γ satisface $\Gamma - \Gamma = [-1,1]$ y, por lo tanto, contiene un entorno abierto del 0.

6.3.7. Criterios de Medibilidad

Sea E un subconjunto no vacío de $\mathbb{R}$. ¿Cómo se determina si E es o no un conjunto medible según Lebesgue?. El objetivo de esta sección es presentar algunos de los criterios más conocidos y usados que permiten identificar a los subconjuntos medibles de $\mathbb{R}$.

La definición de conjunto medible que hemos dado al comienzo de esta sección usa explícitamente la noción topológica de conjuntos abiertos. La siguiente caracterización de conjuntos medibles, aunque sigue estando relacionada con la topología de $\mathbb{R}$, en el sentido de ella se formula en término de conjuntos abiertos, cerrados, G_δ y F_σ, tiene un mérito tremendo: describe *explícitamente* cómo son los conjuntos medibles.

Teorema 6.3.55. *Sea $E \subseteq \mathbb{R}$. Las siguientes condiciones son equivalentes:*

(a) $E \in \mathfrak{M}_\mu(\mathbb{R})$.

(b) Para cada $\varepsilon > 0$, existe un conjunto cerrado $F \subseteq E$ tal que $\mu^*(E \setminus F) < \varepsilon$.

(c) $E = H \cup Z$, donde H es un F_σ y $\mu^*(Z) = 0$.

(d) $E = H \setminus Z$, donde H es un G_δ y $\mu^*(Z) = 0$.

Prueba. $(a) \Rightarrow (b)$. Es el Corolario 6.3.9.

$(b) \Rightarrow (c)$. Suponga que (b) se cumple y seleccione, por cada $n \in \mathbb{N}$, un conjunto cerrado $F_n \subseteq E$ tal que $\mu^*(E \setminus F_n) < 1/n$. Sean $H = \bigcup_{n=1}^{\infty} F_n$ y $Z = E \setminus H$. Entonces $E = H \cup Z$, H es un conjunto F_σ y como

$$\mu^*(E \setminus H) \;=\; \mu^*\Big(E \cap \Big(\bigcup_{n=1}^{\infty} F_n\Big)^c\Big) \;=\; \mu^*\Big(\bigcap_{n=1}^{\infty} (E \setminus F_n)\Big) \;\leq\; \mu^*(E \setminus F_n) \;<\; \frac{1}{n}$$

para todo $n \in \mathbb{N}$, resulta que $\mu^*(Z) = \mu^*(E \setminus H) = 0$.

$(c) \Rightarrow (a)$. Suponga que (c) se cumple. Como H y Z son medibles, entonces $E = H \cup Z$ también lo es.

$(a) \Rightarrow (d)$. Suponga que $E \in \mathfrak{M}_\mu(\mathbb{R})$. Para cada $n \in \mathbb{N}$, escojamos un abierto $G_n \supseteq E$ tal que $\mu^*(G_n \setminus E) < 1/n$. Si $H = \bigcap_{n=1}^{\infty} G_n$, entonces H es un conjunto G_δ conteniendo a E, por lo que si ahora definimos $Z = H \setminus E$, resultará que $Z \subseteq G_n \setminus E$ para todo $n \in \mathbb{N}$ y, por consiguiente, $\mu^*(Z) \leq \mu^*(G_n \setminus E) < 1/n$. Ya que esta última desigualdad es válida para todo $n \in \mathbb{N}$, concluimos que $\mu^*(Z) = 0$.

$(d) \Rightarrow (a)$. Es inmediata ya que $H \setminus Z$ es medible. ■

Las partes (d) y (c) del resultado anterior es interesante por lo siguiente: *para obtener, a partir de los conjuntos abiertos, la familia de los conjuntos medibles según Lebesgue* tan sólo se requiere llevar a cabo el siguiente procedimiento:

Primero: *Se comienza con la familia $\mathcal{O}_{\mathbb{R}}$ constituida por todos los conjuntos abiertos*.

Segundo: *Se forma ahora la familia $\mathcal{G}_\delta(\mathbb{R})$ de todas las intersecciones numerables de elementos de $\mathcal{O}_{\mathbb{R}}$, es decir, los conjuntos de tipo G_δ*, y finalmente,

Tercero: *Se construye la familia $\mathfrak{N}_\mu^\sigma(\mathbb{R})$ de todos los conjuntos nulos N que están incluidos en conjuntos de tipo G_δ*.

Cuarto: *El conjunto de las diferencias*

$$\mathcal{G}_\delta(\mathbb{R}) \setminus \mathfrak{N}^\sigma_\mu(\mathbb{R}) \;=\; \big\{ H \setminus Z \,:\; H \in \mathcal{G}_\delta(\mathbb{R}),\; Z \subseteq E \in \mathcal{G}_\delta(\mathbb{R}),\; \mu(E) = 0 \big\}$$

constituye, según (d), la familia de los conjuntos medibles según Lebesgue. También, si uno comienza con los conjuntos cerrados, construye la colección $\mathcal{F}_\sigma(\mathbb{R})$ de los conjuntos F_σ's y luego formamos la unión $\mathcal{F}_\sigma(\mathbb{R}) \cup \mathfrak{N}_\mu(\mathbb{R})$, se obtiene:

$$\mathfrak{M}_\mu(\mathbb{R}) \;=\; \big\{ B \cup Z \,:\; B \in \mathcal{F}_\sigma(\mathbb{R}),\; Z \subseteq E \in \mathcal{F}_\sigma(\mathbb{R}),\; \mu(E) = 0 \big\}.$$

Puesto que $\mathcal{F}_\sigma(\mathbb{R}) \subseteq \mathfrak{Bo}(\mathbb{R}) \subseteq \mathfrak{M}_\mu(\mathbb{R})$, lo anterior también se puede expresar en la forma:

$$\mathfrak{M}_\mu(\mathbb{R}) \;=\; \mathfrak{Bo}(\mathbb{R}) \,\cup\, \mathfrak{N}_\mu(\mathbb{R}). \tag{MBN}$$

Esta última igualdad revela un hecho crucial: ya habíamos descubierto que $\mathfrak{N}_\mu(\mathbb{R})$ no está incluido en $\mathfrak{Bo}(\mathbb{R})$, de modo que *si queremos localizar conjuntos medibles según Lebesgue que no sean borelianos tenemos que buscarlos precisamente en* $\mathfrak{N}_\mu(\mathbb{R})$.

Otra manera de reconocer a los conjuntos medibles, similar al resultado anterior, es a través del siguiente teorema.

Teorema 6.3.56. *Sea $E \subseteq \mathbb{R}$. Son equivalentes:*

(1) *$E \in \mathfrak{M}_\mu(\mathbb{R})$.*

(2) *Para cada $\varepsilon > 0$, existen un **conjunto abierto** G y un **conjunto cerrado** F tales que*

$$F \subseteq E \subseteq G \qquad y \qquad \mu(G \setminus F) \;<\; \varepsilon.$$

(3) *Existe un conjunto $G \in \mathcal{G}_\delta(\mathbb{R})$ y un conjunto $F \in \mathcal{F}_\sigma(\mathbb{R})$ tales que*

$$F \subseteq E \subseteq G \qquad y \qquad \mu(G \setminus F) \;=\; 0.$$

Prueba. (1) $\Rightarrow$ (2). Por definición de conjunto medible, existe un conjunto abierto $G \supseteq E$ tal que $\mu(G \setminus E) < \varepsilon/2$, y por el Corolario 6.3.9 existe un conjunto cerrado $F \subseteq E$ tal que $\mu(E \setminus F) < \varepsilon/2$. Finalmente, como $G \setminus F = (G \setminus E) \cup (E \setminus F)$, entonces

$$\mu(G \setminus F) \;=\; \mu(G \setminus E) \,+\, \mu(E \setminus F) \;<\; \varepsilon.$$

(2) $\Rightarrow$ (3). Suponga que (2) se cumple. Para cada $n \in \mathbb{N}$, seleccione, usando nuestra suposición, un conjunto abierto G_n y un conjunto cerrado F_n tales que

$$F_n \subseteq E \subseteq G_n \qquad \text{y} \qquad \mu(G_n \setminus F_n) < 1/n.$$

Tomando

$$F = \bigcup_{n=1}^{\infty} F_n \qquad \text{y} \qquad G = \bigcap_{n=1}^{\infty} G_n,$$

vemos que G es un G_δ, F es un F_σ, $F \subseteq E \subseteq G$, y como $G \setminus F \subseteq G_n \setminus F_n$ para todo $n \in \mathbb{N}$, se concluye que $\mu(G \setminus F) = 0$.

(3) $\Rightarrow$ (1). Suponga que los conjuntos F y G satisfacen las condiciones impuestas en (3). Observe que como $E \setminus F \subseteq G \setminus F$ y $\mu(G \setminus F) = 0$, entonces $E \setminus F \in \mathfrak{M}_\mu(\mathbb{R})$ y, en consecuencia, $E = F \cup (E \setminus F) \in \mathfrak{M}_\mu(\mathbb{R})$. ■

Nótese que las dos caracterizaciones anteriores sólo usan propiedades topológicas de $\mathbb{R}$ y, por consiguiente, son susceptibles de ser generalizadas a espacios topológicos arbitrarios; sin embargo, es una caracterización sin utilidad para subconjuntos en espacios desprovistos de estructura topológica, por lo que si queremos formular una noción de medibilidad en un contexto abstracto, es decir, en cualquier conjunto donde no intervenga noción alguna de topología, debemos buscar otras formas de representarlos. La siguiente manera de caracterizar a los conjuntos medibles, atribuida a Constantin Carathéodory (1873–1950), es interesante ya que no requiere ningún concepto topológico en su formulación.

Definición 6.3.57. *Un conjunto $E \subseteq \mathbb{R}$ se dice que satisface el* ***Criterio de Carathéodory*** *si, para cualquier subconjunto A de $\mathbb{R}$, se cumple que*

$$\mu^*(A) \;=\; \mu^*(A \cap E) \;+\; \mu^*(A \cap E^c) \tag{6.3.3}$$

Nótese que por ser $A = (A \cap E) \cup (A \cap E^c)$ una unión disjunta, la subaditividad de μ^* siempre garantiza que

$$\mu^*(A) \;\leq\; \mu^*(A \cap E) + \mu^*(A \cap E^c),$$

de modo que, para verificar que un conjunto E satisface el Criterio de Carathéodory sólo es suficiente demostrar que

$$\mu^*(A) \;\geq\; \mu^*(A \cap E) \;+\; \mu^*(A \cap E^c)$$

para todo $A \subseteq \mathbb{R}$. La siguiente caracterización de medibilidad es utilizado con frecuencia como la definición habitual de conjunto medible.

Teorema 6.3.58 (Carathéodory). *Sea $E \subseteq \mathbb{R}$. Las siguientes condiciones son equivalentes:*

(1) $E \in \mathfrak{M}_\mu(\mathbb{R})$.

(2) E *satisface el* ***Criterio de Carathéodory***.

Prueba. (1) $\Rightarrow$ (2). Suponga que E es medible y sea A cualquier subconjunto de $\mathbb{R}$. Puesto que $A = (A \cap E) \cup (A \cap E^c)$, donde los conjuntos $A \cap E$ y $A \cap E^c$ están contenidos, respectivamente, en los conjuntos medibles y disjuntos E y E^c, resulta entonces del Corolario 6.3.43 que

$$\mu^*(A) \;=\; \mu^*(A \cap E) \;+\; \mu^*(A \cap E^c).$$

(2) $\Rightarrow$ (1). Suponga que el conjunto E satisface el Criterio de Carathéodory. Para cada $n \in \mathbb{N}$, sea

$$E_n \;=\; \big\{x \in E \;:\; |x| \leq n\big\} \;=\; E \cap [-n, n].$$

Entonces cada conjunto E_n es acotado y $E = \bigcup_{n=1}^{\infty} E_n$. Por cada entero $n \in \mathbb{N}$, usando de nuevo el Corolario 6.2.13, escojamos un conjunto G_δ, digamos G_n, tal que

$$G_n \supseteq E_n \qquad \text{y} \qquad \mu(G_n) \;= \mu^*(E_n).$$

Observe que como $E_n \subseteq E$ y $G_n \supseteq E_n$, entonces $E_n = (E_n \cap E) \subseteq (G_n \cap E)$. Si ahora reemplazamos A por G_n en (6.3.3) se tiene que

$$\mu^*(E_n) \;=\; \mu(G_n) \;=\; \mu^*(G_n \cap E) \;+\; \mu^*(G_n \cap E^c) \;\geq\; \mu^*(E_n) \;+\; \mu^*(G_n \cap E^c).$$

Puesto que que cada E_n es acotado, resulta que $\mu^*(E_n) < +\infty$ y se sigue la desigualdad anterior que $\mu^*(G_n \cap E^c) = 0$. Sea $Z_n = G_n \cap E^c$, $n = 1, 2, \ldots$ Del Teorema 6.3.2 sabemos que Z_n es medible y, por lo tanto,

$$G = \bigcup_{n=1}^{\infty} G_n \qquad \text{y} \qquad Z = \bigcup_{n=1}^{\infty} Z_n$$

son medibles. Finalmente, como

$$G \setminus E = G \cap E^c = \left(\bigcup_{n=1}^{\infty} G_n\right) \cap E^c = \bigcup_{n=1}^{\infty} (G_n \cap E^c) = \bigcup_{n=1}^{\infty} Z_n = Z,$$

resulta que $E = G \setminus (G \setminus E) = G \setminus Z \in \mathfrak{M}_\mu(\mathbb{R})$. ■

Nota Adicional 6.3.6 (a) Es importante destacar que si en el Criterio de Carathéodory la condición "*para todo* $A \subseteq \mathbb{R}$" se reemplaza por "*para todo* $A \in \mathfrak{M}_\mu(\mathbb{R})$", entonces $E \in \mathfrak{M}_\mu(\mathbb{R})$. De hecho, algo más fuerte se puede afirmar. Del Corolario 6.2.13 y los argumentos dados en la demostración del resultado anterior se tiene que:

(b) *si el Criterio de Carathéodory se cumple* ***para todo conjunto*** $A \in \mathcal{G}_\delta$, *entonces* $E \in \mathfrak{M}_\mu(\mathbb{R})$.

Corolario 6.3.59. *Sea* $A \subseteq \mathbb{R}$ *con* $\mu^*(A) < +\infty$. *Si existe un* ***conjunto medible*** $E \subseteq A$ *tal que* $\mu(E) = \mu^*(A)$, *entonces* A *es* ***medible***.

Prueba. Siendo E medible, se sigue del Teorema de Carathéodory que,

$$\begin{aligned} \mu^*(A) &= \mu^*(A \cap E) + \mu^*(A \cap E^c) \\ &= \mu^*(E) + \mu^*(A \cap E^c) \\ &= \mu(E) + \mu^*(A \cap E^c). \end{aligned}$$

Por otro lado, como $\mu^*(A) < +\infty$ y $E \subseteq A$, entonces $\mu^*(E) < +\infty$ y, en consecuencia,

$$0 = \mu^*(A) - \mu(E) = \mu^*(A \cap E^c).$$

Esto nos muestra que $A \cap E^c$ es medible y, por consiguiente, $A = E \cup (A \cap E^c)$ también es medible. ■

Teorema 6.3.60. *Sea* $E \in \mathfrak{M}_\mu(\mathbb{R})$ *con* $\mu(E) < +\infty$. *Para cada conjunto* $A \subseteq E$, *las siguientes condiciones son equivalentes:*

(1) $A \in \mathfrak{M}_\mu(\mathbb{R})$.

(2) $\mu(E) = \mu^*(A) + \mu^*(E \setminus A)$.

Prueba. (1) $\Rightarrow$ (2). Si $A \in \mathfrak{M}_\mu(\mathbb{R})$, entonces

$$\mu(E) = \mu(E \cap A) + \mu(E \cap A^c) = \mu(A) + \mu(E \setminus A).$$

(2) $\Rightarrow$ (1). Suponga que (2) se cumple. Por el Corolario 6.3.38 existe un conjunto G_1 de tipo $\mathcal{G}_\delta$ tal que $A \subseteq G_1$ y $\mu(G_1) = \mu^*(A)$. Sea $G = G_1 \cap E$. Entonces $G \in \mathfrak{M}_\mu(\mathbb{R})$, $A \subseteq G$, $E \setminus G \subseteq E \setminus A$ y $\mu(G) = \mu^*(A)$. Nuestra hipótesis y el hecho de que G es medible implican que

$$\mu^*(A) + \mu^*(E \setminus A) = \mu(E) = \mu(E \cap G) + \mu(E \setminus G) = \mu(G) + \mu(E \setminus G).$$

De esto se sigue que $\mu^*(E \setminus A) = \mu(E \setminus G)$ y, entonces, por el Corolario 6.3.59, el conjunto $E \setminus A$ es medible. Finalmente, $A = E \setminus (E \setminus A) \in \mathfrak{M}_\mu(\mathbb{R})$. ■

Del resultado anterior se tiene que si $E \subseteq [0,1]$ y

$$\mu^*(E) + \mu^*([0,1] \setminus E) = 1,$$

entonces E es medible.

Ejemplo 6.3.3. *Ya hemos visto, Teorema 6.2.17, que si A y B son subconjuntos de $\mathbb{R}$ que cumplen la desigualdad $\operatorname{dist}(A,B) > 0$, entonces $\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$. Si se omite la condición $\operatorname{dist}(A,B) > 0$, pero se impone esta otra: existe un* ***conjunto medible*** *E tal que*

$$A \subseteq E \qquad y \qquad B \cap E = \varnothing$$

entonces la conclusión es la misma, es decir, $\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$.

Prueba. Como E es medible, el Teorema de Carathéodory nos dice

$$\begin{aligned} \mu^*(A \cup B) &= \mu^*\big((A \cup B) \cap E\big) + \mu^*\big((A \cup B) \cap E^c\big) \\ &= \mu^*(A \cap E) + \mu^*(B \cap E^c) \\ &= \mu^*(A) + \mu^*(B). \end{aligned}$$

Podemos ahora dar otra condición bajo la cual la medida exterior es numerablemente aditiva. Aquí supondremos que $\mathfrak{M}_\mu(\mathbb{R}) \subsetneqq \mathscr{P}(\mathbb{R})$, es decir, que existen conjuntos no-medibles.

Teorema 6.3.61. *Sea $(E_n)_{n=1}^\infty$ una* ***sucesión disjunta*** *en $\mathfrak{M}_\mu(\mathbb{R})$ con $\mu(E_n) > 0$ para todo $n \in \mathbb{N}$. Si $(A_n)_{n=1}^\infty$ es una sucesión de* ***conjuntos no-medibles*** *tal que $A_n \subseteq E_n$ para todo $n \geq 1$, entonces*

$$\bigcup_{n=1}^\infty A_n \notin \mathfrak{M}_\mu(\mathbb{R}) \qquad y \qquad \mu^*\Big(\bigcup_{n=1}^\infty A_n\Big) = \sum_{n=1}^\infty \mu^*(A_n).$$

Prueba. Claramente $\bigcup_{n=1}^\infty A_n \notin \mathfrak{M}_\mu(\mathbb{R})$. En efecto, suponer que $\bigcup_{n=1}^\infty A_n \in \mathfrak{M}_\mu(\mathbb{R})$ conduce a que $(\bigcup_{n=1}^\infty A_n) \cap E_m = A_m \in \mathfrak{M}_\mu(\mathbb{R})$ para cualquier $m \in \mathbb{N}$, lo cual es imposible. Veamos ahora que

$$\mu^*\Big(\bigcup_{n=1}^\infty A_n\Big) = \sum_{n=1}^\infty \mu^*(A_n).$$

Puesto que μ^* es numerablemente subaditiva, sólo necesitamos demostrar que

$$\mu^*\Big(\bigcup_{n=1}^\infty A_n\Big) \geq \sum_{n=1}^m \mu^*(A_n)$$

para cualquier $m \in \mathbb{N}$. Veamos esto último. Como μ^* es monótona, resulta que

$$\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \geq \mu^*\left(\bigcup_{n=1}^{m} A_n\right)$$

para todo $m \geq 1$. Demostremos ahora, por inducción, que

$$\mu^*\left(\bigcup_{n=1}^{m} A_n\right) = \sum_{n=1}^{m} \mu^*(A_n).$$

La igualdad es trivial para $m = 1$. Suponga que ella se cumple para $m = k$. Puesto que $A_{k+1} \subseteq E_{k+1}$ y $\bigcup_{n=1}^{k} A_n \subseteq E_{k+1}^c$, entonces la medibilidad de E_{k+1} implica, aplicando el Criterio de Carathéodory y la hipótesis inductiva, que

$$\begin{aligned}\mu^*\left(\bigcup_{n=1}^{k+1} A_n\right) &= \mu^*\left(\bigcup_{n=1}^{k+1} A_n \cap E_{k+1}\right) + \mu^*\left(\bigcup_{n=1}^{k+1} A_n \cap E_{k+1}^c\right) \\ &= \mu^*(A_{k+1}) + \mu^*\left(\bigcup_{n=1}^{k} A_n\right) \\ &= \mu^*(A_{k+1}) + \sum_{n=1}^{k} \mu^*(A_n) = \sum_{n=1}^{k+1} \mu^*(A_n)\end{aligned}$$

La prueba es completa. ■

Finalizamos esta sección con una nueva e interesante caracterización de medibilidad para conjuntos con medida exterior finita. J. E. Littlewood simplificó la conexión de la Teoría de La Medida con el Análisis Real Clásico mostrando tres principios, conocidos ahora como los Tres Principios de Littlewood, que los vinculan (es necesario señalar que Littlewood no tuvo ninguna contribución en la prueba de esos principios).

Teorema 6.3.62 (Primer Principio de Littlewood). *Sea $E \subseteq \mathbb{R}$ con $\mu^*(E) < +\infty$. Las siguientes condiciones son equivalentes:*

(1) $E \in \mathfrak{M}_\mu(\mathbb{R})$.

(2) *Dado $\varepsilon > 0$, existe una* ***familia finita de intervalos abiertos*** $\{I_1, \ldots, I_n\}$, ***disjuntos dos a dos****, tal que*

$$\mu^*\left(E \,\triangle\, \bigcup_{i=1}^{n} I_i\right) < \varepsilon.$$

Prueba. (1) $\Rightarrow$ (2). Suponga que $E \in \mathfrak{M}_\mu(\mathbb{R})$ y sea $\varepsilon > 0$. Como $\mu^*(E) < +\infty$, el Corolario 6.3.40 nos garantiza la existencia de un conjunto compacto $K \subseteq E$ tal que $\mu(E \setminus K) < \varepsilon/2$. Siendo K un conjunto medible, existe un conjunto abierto $G \supseteq K$ tal que $\mu(G \setminus K) < \varepsilon/2$. Escribamos a G como $G = \bigcup_{n=1}^{\infty} I_n$, donde $(I_n)_{n=1}^{\infty}$ es una sucesión de intervalos abiertos disjuntos dos a dos. Puesto que K es compacto y $K \subseteq G = \bigcup_{n=1}^{\infty} I_n$, existe una subcolección finita $\{I_{n_1}, \ldots, I_{n_k}\}$ de $(I_n)_{n=1}^{\infty}$ tal que $K \subseteq \bigcup_{i=1}^{k} I_{n_i}$. Por esto,

$$\begin{aligned}
\mu^*\Big(E \triangle \bigcup_{i=1}^{k} I_{n_i}\Big) &= \mu^*\Big(\Big(E \setminus \bigcup_{i=1}^{k} I_{n_i}\Big) \cup \Big(\bigcup_{i=1}^{k} I_{n_i} \setminus E\Big)\Big) \\
&= \mu\Big(E \setminus \bigcup_{i=1}^{k} I_{n_i}\Big) + \mu\Big(\bigcup_{i=1}^{k} I_{n_i} \setminus E\Big) \\
&\leq \mu(E \setminus K) + \mu(G \setminus E) \\
&< \varepsilon.
\end{aligned}$$

(2) $\Rightarrow$ (1). Suponga que (2) se cumple y sea $\varepsilon > 0$. Lo que queremos es demostrar la existencia de un conjunto abierto $G \supseteq E$ tal que $\mu^*(G \setminus E) < \varepsilon$. Sea $\{I_1, \ldots, I_n\}$ una colección finita de intervalos abiertos bajo la cual la desigualdad $\mu^*\big(E \triangle \bigcup_{i=1}^{n} I_i\big) < \varepsilon/3$ se cumple. Pongamos $O_1 = \bigcup_{i=1}^{n} I_i$. Observe que como $E \setminus O_1$ y $O_1 \setminus E$ son subconjuntos de $E \triangle \bigcup_{i=1}^{n} I_i$, entonces

$$\mu^*\big(E \setminus O_1\big) < \frac{\varepsilon}{3} \qquad \text{y} \qquad \mu^*\big(O_1 \setminus E\big) < \frac{\varepsilon}{3}.$$

Invocando el Teorema 6.2.12 podemos garantizar la existencia de un conjunto abierto O_2 de modo tal que

$$O_2 \supseteq E \setminus O_1 \qquad \text{y} \qquad \mu^*(O_2) \leq \mu^*\big(E \setminus O_1\big) + \frac{\varepsilon}{3}.$$

Esta última desigualdad combinada con la desigualdad $\mu^*\big(E \setminus O_1\big) < \varepsilon/3$ produce

$$\mu^*(O_2) < \frac{2\varepsilon}{3}.$$

Definamos ahora $G = O_1 \cup O_2$. Como $E = \big(E \setminus O_1\big) \cup \big(E \cap O_1\big)$ y G es abierto, resulta que $E \subseteq G$ y

$$\begin{aligned}
\mu^*(G \setminus E) &= \mu^*\big((O_1 \cup O_2) \setminus E\big) \\
&\leq \mu^*\Big(O_1 \setminus E\Big) + \mu^*\big(O_2 \setminus E\big) \\
&< \frac{\varepsilon}{3} + \mu^*(O_2) \\
&< \varepsilon.
\end{aligned}$$

La prueba es completa. ■

Los otros dos principios de Littlewood, conocidos como los Teorema de Lusin y Severini-Egoroff, serán dados a conocer en la sección sobre funciones medibles desarrolladas un poco más adelante.

6.3.8. Medida Interior

El concepto de *medida interior* es de mucha utilidad para entender de modo más cabal el fenómeno de la no existencia de subconjuntos medibles que viven en conjuntos que poseen medida exterior positiva. Recordemos, Corolario 6.3.40, que si E es un conjunto medible de **medida finita**, entonces

$$\mu(E) = \sup\{\mu(K) : K \subseteq E, K \text{ compacto}\}.$$

En general, para conjuntos arbitrarios, se tiene la siguiente definición.

Definición 6.3.63. *Sea A un subconjunto de $\mathbb{R}$. La **medida interior de Lebesgue** de A se define por*

$$\mu_*(A) = \sup\{\mu(K) : K \subseteq A, K \text{ compacto}\}.$$

Como antes, escribiremos *medida interior* en lugar de medida interior de Lebesgue. Algunas de las propiedades que posee la medida interior se dan a continuación. Observe la similitud de algunas de ellas con la de la medida exterior.

Lema 6.3.64. *Sea A un subconjunto de $\mathbb{R}$. Entonces las siguientes condiciones se cumplen:*

(i_1) $\mu_*(A) \leq \mu^*(A)$.

(i_2) *Si* $B \subseteq A$, *entonces* $\mu_*(B) \leq \mu_*(A)$.

(i_3) $\mu_*(x + A) = \mu_*(A)$ *para cualquier* $x \in \mathbb{R}$.

(i_4) *Si* $(A_n)_{n=1}^{\infty}$ *es una **sucesión disjunta** de subconjuntos de* $\mathbb{R}$, *entonces*

$$\mu_*\Big(\bigcup_{n=1}^{\infty} A_n\Big) \geq \sum_{n=1}^{\infty} \mu_*(A_n).$$

Prueba. (i_1). Sean K y G subconjuntos arbitrarios de $\mathbb{R}$ de modo tal que K sea compacto, G abierto y se cumpla que $K \subseteq A \subseteq G$. Puesto que $\mu(K) \leq \mu(G)$, resulta que si mantenemos a G fijo y variamos a K tendremos que

$$\mu_*(A) = \sup\{\mu(K) : K \subseteq A, K \text{ compacto}\} \leq \mu(G).$$

Por otro lado, por el Teorema 6.2.12 sabemos que

$$\mu^*(A) = \inf\{\mu(G) : A \subseteq G, G \text{ abierto}\}$$

y, por lo tanto, $\mu_*(A) \leq \mu^*(A)$.

(i_2) sigue directamente de la definición, mientras que (i_3) es consecuencia de la definición y del hecho de que μ es invariante por traslación.

(i_4). Escojamos, para cada $n \in \mathbb{N}$, un conjunto compacto $K_n \subseteq A_n$. La sucesión $(K_n)_{n=1}^{\infty}$ es medible y disjunta. Sea $N \in \mathbb{N}$ arbitrario. Puesto que $\bigcup_{n=1}^{N} K_n$ es un conjunto compacto incluido en $\bigcup_{n=1}^{\infty} A_n$ entonces por la definición de μ_* y la aditividad de μ tenemos que

$$\mu_*\Big(\bigcup_{n=1}^{\infty} A_n\Big) \geq \mu\Big(\bigcup_{n=1}^{N} K_n\Big) = \sum_{n=1}^{N} \mu(K_n).$$

Con N fijo y tomando el supremo sobre todos los compactos K_n que habitan en A_n, resulta que

$$\mu_*\Big(\bigcup_{n=1}^{\infty} A_n\Big) \;\geq\; \sum_{n=1}^{N} \mu_*(A_n).$$

Si ahora hacemos que N tienda a $+\infty$, obtendremos el resultado deseado. ■

Una de las grandes diferencias entre la medida exterior y la medida interior de Lebesgue es la siguiente: *si E es cualquier conjunto tal que $\mu^*(E) = 0$, entonces E siempre es medible. Sin embrago, es posible obtener un conjunto no medible V tal que $\mu_*(V) = 0$.* Por ejemplo, el conjunto de Vitali, que construiremos en la próxima sección, es un tal conjunto.

El siguiente resultado, que no es ninguna sorpresa, es una nueva caracterización de conjuntos medibles con medida finita.

Teorema 6.3.65 (La Igualdad $\mu^* = \mu_*$). *Sea E un subconjunto de $\mathbb{R}$. Si E es **medible**, entonces*

$$\mu_*(E) \;=\; \mu^*(E). \tag{6.3.4}$$

Recíprocamente, si $\mu_*(E) = \mu^*(E) < +\infty$, entonces $E \in \mathfrak{M}_\mu(\mathbb{R})$.

Prueba. Suponga que E es medible. Si $\mu^*(E) < +\infty$, entonces la igualdad (6.3.4) sigue del Corolario 6.3.40. Si $\mu^*(E) = +\infty$, entonces con $\varepsilon = 1$, podemos invocar el Corolario 6.3.9 para obtener un subconjunto cerrado $F \subseteq E$ tal que $\mu^*(E \setminus F) < 1$. Por esto,

$$\mu^*(E) \;=\; \mu^*\big(F \cup (E \setminus F)\big) \;\leq\; \mu^*(F) \;+\; \mu^*(E \setminus F) \;<\; \mu^*(F) \;+\; 1$$

y, por consiguiente, $\mu^*(F) = +\infty$. Para cada $n \in \mathbb{N}$, sea $K_n = F \cap [-n, n]$. Observe que como $(K_n)_{n=1}^{\infty}$ es una sucesión monótona creciente de compactos con $\bigcup_{n=1}^{\infty} K_n = F$, el Teorema 6.3.45 nos garantiza entonces que

$$\lim_{n\to\infty} \mu(K_n) \;=\; \mu(F) \;=\; +\infty.$$

De esto se deduce que $\mu_*(E) \geq \sup_{n\geq 1} \mu(K_n) = +\infty = \mu^*(E)$ y termina la prueba de la igualdad (6.3.4).

Para demostrar el recíproco, suponga que $\mu_*(E) = \mu^*(E) < +\infty$ y fijemos un $\varepsilon > 0$ elegido arbitrariamente. Seleccionemos un conjunto compacto $K \subseteq E$ de modo tal que

$$\mu(K) \;\geq\; \mu_*(E) \;-\; \frac{\varepsilon}{2} \;\geq\; \mu^*(E) \;-\; \frac{\varepsilon}{2}.$$

Invoquemos ahora el Teorema 6.2.12 para hallar un conjunto abierto $G \supseteq E$ que satisfaga

$$\mu(G) \;\leq\; \mu^*(E) \;+\; \frac{\varepsilon}{2}.$$

Finalmente, de las dos desigualdades anteriores podemos concluir que

$$\begin{aligned} \mu^*(G \setminus E) \;\leq\; \mu(G \setminus K) \;&=\; \mu(G) \;-\; \mu(K) \\ &=\; \Big(\mu(G) - \mu^*(E)\Big) \;+\; \Big(\mu^*(E) - \mu(K)\Big) \;\leq\; \varepsilon, \end{aligned}$$

lo cual prueba que E es medible. ■

El Teorema 6.2.12 combinado con el Teorema 6.3.65, conducen a una de las representaciones más importantes de la medida de Lebesgue, comúnmente conocida con el nombre de ***regularidad*** de la medida de Lebesgue.

Corolario 6.3.66 (Regularidad de μ). *Si $E \in \mathfrak{M}_\mu(\mathbb{R})$, entonces*

$$\begin{aligned} \mu(E) &= \inf\{\mu(G) : E \subseteq G,\ G \text{ abierto}\} \\ &= \sup\{\mu(K) : K \subseteq E,\ K \text{ compacto}\}. \end{aligned}$$

Los siguientes dos resultados muestran una estrecha e interesante relación entre la medida exterior y la medida interior. Debemos resaltar que tales resultados son fundamentales para conocer algunas de las extrañas propiedades que poseen los conjuntos no-medibles que estudiaremos un poco más abajo.

Teorema 6.3.67. *Sea $E \in \mathfrak{M}_\mu(\mathbb{R})$. Entonces para cualquier conjunto $A \subseteq E$, se verifica que*

$$\mu(E) = \mu_*(A) + \mu^*(E \setminus A).$$

Prueba. Por cada $\varepsilon > 0$ seleccionemos, usando la definición de $\mu_*(A)$, un subconjunto compacto $K \subseteq A$ tal que $\mu(K) > \mu_*(A) - \varepsilon$. De esto se sigue que

$$\mu(E) = \mu(K) + \mu(E \setminus K) > \mu_*(A) - \varepsilon + \mu^*(E \setminus A)$$

y como nuestro ε es arbitrario, se concluye que

$$\mu(E) \geq \mu_*(A) + \mu^*(E \setminus A).$$

Para demostrar la otra desigualdad, trabajemos ahora con el conjunto $E \setminus A$ e invoquemos el Corolario 6.2.13 para hallar un conjunto G_δ, digamos G, tal que $E \setminus A \subseteq G \subseteq E$ y se cumpla que $\mu(G) = \mu^*(E \setminus A)$. Puesto que $E \setminus G$ es medible y $E \setminus G \subseteq A$, se sigue del Teorema 6.3.65 que $\mu(E \setminus G) = \mu^*(E \setminus G) = \mu_*(E \setminus G) \leq \mu_*(A)$ y, en consecuencia,

$$\mu(E) = \mu(G) + \mu(E \setminus G) = \mu^*(E \setminus A) + \mu(E \setminus G) \leq \mu^*(E \setminus A) + \mu_*(A).$$

Combinando las desigualdades anteriores obtenemos que

$$\mu(E) = \mu_*(A) + \mu^*(E \setminus A).$$

Esto termina la prueba. ∎

Teorema 6.3.68. *Sea $A \subseteq \mathbb{R}$ un conjunto arbitrario. Las siguientes condiciones son equivalentes:*

(1) $\mu_*(A^c) = 0$.

(2) $A \cap E \neq \varnothing$ *para cualquier conjunto* $E \in \mathfrak{M}_\mu(\mathbb{R})$ *con* $\mu(E) > 0$.

(3) $\mu(E) = \mu^*(A \cap E)$ *para cualquier conjunto* $E \in \mathfrak{M}_\mu(\mathbb{R})$.

(4) $\mu(I) = \mu^*(A \cap I)$ *para cualquier intervalo abierto* $I \subseteq \mathbb{R}$.

Prueba. (1) $\Rightarrow$ (2) Sea $A \subseteq \mathbb{R}$ satisfaciendo (1) y sea $E \in \mathfrak{M}_\mu(\mathbb{R})$ con $\mu(E) > 0$. Si $A \cap E = \varnothing$, entonces $E \subseteq A^c$ y se sigue del Lema 6.3.64 y el Teorema 6.3.65 que

$$\mu_*(A^c) \geq \mu_*(E) = \mu(E) > 0$$

lo cual contradice a (1).

(2) $\Rightarrow$ (3) Suponga que (2) se cumple y sea $E \in \mathfrak{M}_\mu(\mathbb{R})$. Entonces

$$\mu^*(A \cap E) \;\leq\; \mu^*(E) \;=\; \mu(E).$$

Veamos que la desigualdad estricta $\mu^*(A \cap E) < \mu(E)$ no puede ocurrir. Suponga que existe algún conjunto medible $E \subseteq \mathbb{R}$ tal que

$$\mu^*(A \cap E) \;<\; \mu(E).$$

Por el Corolario 6.2.13, existe un conjunto medible G tal que $A \cap E \subseteq G$ y $\mu(G) = \mu^*(A \cap E)$. De allí que $\mu(G) = \mu^*(A \cap E) < \mu(E)$ y, por consiguiente,

$$\mu(G \cap E) \;\leq\; \mu(G) \;<\; \mu(E). \qquad (\alpha)$$

Ahora bien, puesto que $\mu(E) < +\infty$, resulta de (α) que $\mu(G \cap E) < +\infty$ y, por lo tanto, usando una vez más (α), se obtiene que

$$\mu(E \setminus (G \cap E)) \;=\; \mu(E) - \mu(G \cap E) \;>\; 0.$$

Por otro lado, como $A \cap E \subseteq G \cap E$, entonces $E \setminus (G \cap E) \subseteq E \setminus (A \cap E) \subseteq A^c$ lo cual muestra la existencia de un conjunto medible $E_0 = E \setminus (G \cap E)$ de medida positiva disjunto de A. Esto, por supuesto, contradice a (2) y, en consecuencia, la suposición de que $\mu^*(A \cap E) < \mu(E)$ es falsa.

(3) $\Rightarrow$ (4) Es trivial ya que todo intervalo abierto es un conjunto medible.

(4) $\Rightarrow$ (1) Suponga que $\mu_*(A^c) > 0$. Se sigue de la definición de $\mu_*(A^c)$ que podemos elegir un conjunto compacto $K \subseteq A^c$ tal que $\mu(K) > 0$. Para cada $n \in \mathbb{N}$, pongamos $J_n = (-n, n)$. Puesto que

$$K \;=\; K \cap \mathbb{R} \;=\; K \cap \left(\bigcup_{n=1}^{\infty} J_n \right) \;=\; \bigcup_{n=1}^{\infty} (K \cap J_n),$$

y la sucesión $\big(K \cap J_n\big)_{n=1}^{\infty}$ es creciente, la continuidad de μ nos garantiza entonces que

$$\mu(K) \;=\; \lim_{n\to\infty} \mu\big(K \cap J_n\big),$$

y, por lo tanto, $\mu\big(K \cap J_n\big) > 0$ para algún n suficientemente grande. Pongamos $J = J_n$ para tal n. Observe ahora que como $K \subseteq A^c$, entonces $A \cap J \cap K = \varnothing$ y, por consiguiente,

$$A \cap J \;=\; (A \cap J \cap K) \cup (A \cap J \cap K^c) \;=\; A \cap J \cap K^c \;\subseteq\; J \cap K^c.$$

Se sigue de esto que

$$\begin{aligned} \mu(J) \;&=\; \mu^*(A \cap J) \;\leq\; \mu^*(J \cap K^c) \;=\; \mu(J \cap K^c) \\ &=\; \mu(J \setminus (J \cap K)) \;=\; \mu(J) - \mu(J \cap K) \;<\; \mu(J) \end{aligned}$$

lo que constituye una incuestionable inconsistencia. La prueba es completa. ■

Recordemos que un subconjunto D de $\mathbb{R}$ es denso en $\mathbb{R}$ si $D \cap V \neq \varnothing$ para todo conjunto abierto no vacío $V \subseteq \mathbb{R}$. Puesto que todo conjunto abierto en $\mathbb{R}$ tiene medida positiva, se sigue de (2) del resultado anterior que *si $\mu_*(A^c) = 0$, entonces A es denso en $\mathbb{R}$*. También, nótese que los intervalos abiertos en la condición (4) del resultado anterior pueden ser reemplazados por intervalos cerrados. Obsérvese que cuando $A = \mathbb{R}$ todas esas condiciones se satisfacen.

6.4. Conjuntos Medibles con Propiedades Especiales

En esta parte nos proponemos mostrar la existencia de algunos conjuntos medibles que poseen propiedades que desbordan la imaginación.

El siguiente argumento constituye otra manera de demostrar que Γ, el conjunto ternario de Cantor, es nunca-denso.

Teorema 6.4.1. *Sea E un **conjunto cerrado no vacío** de $\mathbb{R}$ y suponga que $\mu(E) = 0$. Entonces E es **nunca-denso**. En particular, Γ es nunca-denso en $[0, 1]$.*

Prueba. Vamos a demostrar que cada intervalo abierto $I \subseteq \mathbb{R}$ contiene un subintervalo abierto J tal que $J \cap E = \varnothing$. En efecto, sea I un intervalo abierto arbitrario no vacío. Puesto $\mu(E) = 0$, el intervalo I no puede estar contenido en E y, en consecuencia, $I \cap (\mathbb{R} \setminus E) \neq \varnothing$. Ahora bien, como $\mathbb{R} \setminus E$ es abierto, el conjunto abierto no vacío $I \cap (\mathbb{R} \setminus E)$ contiene un intervalo J el cual es disjunto de E y termina la prueba. ■

Otra forma, tal vez más sencilla, de demostrar el resultado anterior es observar que si $\operatorname{int}(E) \neq \varnothing$ y J es un intervalo abierto incluido en $\operatorname{int}(E)$, entonces $0 = \mu(E) \geq \mu(J) > 0$. Esta contradicción nos indica que E es nunca-denso.

El hecho de todo conjunto nunca-denso no contiene ningún intervalo abierto en su interior puede hacer suponer que su medida es siempre cero. Sin embargo, tal presunción es falsa: *todo conjunto tipo-Cantor de medida positiva es cerrado y nunca-denso*. Lo que es realmente sorprendente es la existencia de un conjunto cerrado, nunca-denso, de medida positiva que, además, está contenido en el complemento de un G_δ-denso, es decir, un conjunto de primera categoría que también es no-numerable.

Teorema 6.4.2 (Conjunto nunca-denso de medida positiva). *(a) Para cada $\varepsilon > 0$, existe un **conjunto abierto y denso** $G_\varepsilon \subseteq \mathbb{R}$ tal que $\mu(G_\varepsilon) \leq \varepsilon$.*

*(b) Existe un **conjunto medible** $E \subseteq \mathbb{R}$ con las siguientes propiedades:*

*(b_1) E es **cerrado**, **nunca-denso** con $\mu(E) > 0$ y*

*(b_2) existe un conjunto G_δ-**denso**, digamos $G \subseteq \mathbb{R}$, con $\mu(G) = 0$ tal que $E \subseteq \mathbb{R} \setminus G$.*

*En particular, E es de **primera categoría**.*

Prueba. (a) Sea $\varepsilon > 0$ y suponga que $(q_n)_{n=1}^\infty$ es una lista de los números racionales. Considere el conjunto

$$G_\varepsilon = \bigcup_{n=1}^{\infty} \left(q_n - \frac{\varepsilon}{2^{n+1}}, q_n + \frac{\varepsilon}{2^{n+1}} \right)$$

Observe que cada G_ε es un conjunto abierto y, además, denso en $\mathbb{R}$ ya que contiene a $\mathbb{Q}$. Más aun,

$$\mu(G_\varepsilon) \leq \sum_{n=1}^{\infty} \mu\left(\left(q_n - \frac{\varepsilon}{2^{n+1}}, q_n + \frac{\varepsilon}{2^{n+1}} \right) \right) = \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon.$$

Esto último nos indica que el conjunto $G_\varepsilon \subsetneqq \mathbb{R}$.

(b) Seleccione, usando la parte (a), el conjunto $E = \mathbb{R} \setminus G_{1/2}$. Resulta entonces que E es cerrado y como no contiene a ningún número racional dicho conjunto es nunca-denso. Más aun,

$$\mu(E) = \mu(\mathbb{R} \setminus G_{1/2}) \geq \mu([0, 1] \setminus G_{1/2}) = \mu([0, 1]) - \mu(G_{1/2}) \geq \frac{1}{2}.$$

¿Puede el lector exhibir algún número irracional en E? Para demostrar la segunda parte, sea $G = \bigcap_{n=1}^{\infty} G_{1/n}$. Por el Teorema de Categoría de Baire, G es un G_δ-denso y, por consiguiente, no-numerable. En particular, $\mathbb{R} \setminus G$ es de primera categoría. Además, puesto que $\mu(G) \leq \mu(G_{1/n}) \leq 1/n$ para todo $n \geq 1$, se tiene que $\mu(G) = 0$ y concluye la prueba. ■

Teorema 6.4.3. *El conjunto de los números reales se puede expresar en la forma $\mathbb{R} = A \cup G$, donde A es de **primera categoría** con $\mu(A) > 0$, G es un G_δ-**denso** con $\mu(G) = 0$ y $A \cap G = \varnothing$.*

Prueba. Sea $(q_n)_{n=1}^{\infty}$ una lista de los números racionales en $\mathbb{R}$ y como en el resultado anterior, defina

$$G_\varepsilon = \bigcup_{n=1}^{\infty} \left(q_n - \frac{\varepsilon}{2^{n+1}}, q_n + \frac{\varepsilon}{2^{n+1}} \right)$$

para cada $0 < \varepsilon < 1$. Entonces $G = \bigcap_{n=1}^{\infty} G_{1/n}$ es, gracias al Teorema de Categoría de Baire, un G_δ-denso en $\mathbb{R}$ con $\mu(G) = 0$. De allí que, si definimos $A = \mathbb{R} \setminus G$ tendremos que A es de primera categoría, $\mathbb{R} = A \cup G$ y $\mu(A) > 0$. ■

Nota Adicional 6.4.7 Lo que nos muestra el Teorema 6.4.3 es que $\mathbb{R}$ se puede descomponer en los dos conjuntos A y N que son, desde dos puntos de vista diferentes, **pequeños** y **grandes** a la vez. En efecto, la noción de conjunto de primera categoría es, desde el punto de vista topológico, un objeto "pequeño", por lo que A es *pequeño* desde la perspectiva topológica pero *grande* desde la perspectiva de medida. Lo mismo ocurre con el conjunto N: desde la óptica de la Teoría de la Medida N es *pequeño* pero *inmensamente grande* topológicamente.

Corolario 6.4.4. *Si $\mathbb{L}$ es un **conjunto de Lusin**, entonces $\mathbb{L}$ es de **segunda categoría** y de **medida cero**.*

Prueba. Usando el resultado anterior, escriba a $\mathbb{R}$ como $\mathbb{R} = A \cup N$, donde A es de primera categoría y N es un conjunto nulo. Si $\mathbb{L}$ es un conjunto de Lusin, entonces él se puede escribir como la unión disjunta $\mathbb{L} = (\mathbb{L} \cap A) \cup (\mathbb{L} \cap N)$. Ahora bien, como A es primera categoría, se tiene que $\mathrm{card}(\mathbb{L} \cap A) \leq \aleph_0$, es decir, $\mathbb{L} \cap A$ es a lo más numerable y, por lo tanto, $\mu(\mathbb{L} \cap A) = 0$. Por otro lado, como $\mathbb{L} \cap N$ es un subconjunto de un conjunto nulo, resulta que $\mu(\mathbb{L} \cap N) = 0$. Por esto,

$$\mu(\mathbb{L}) = \mu(\mathbb{L} \cap A) + \mu(\mathbb{L} \cap N) = 0.$$

La prueba es completa. ■

Podemos afinar un poco más el resultado anterior si permitimos la siguiente definición.

Definición 6.4.5. *Un conjunto $A \subseteq \mathbb{R}$ se dice que es **fuertemente de medida cero** si para cada sucesión $(\varepsilon_n)_{n=1}^{\infty}$ de números reales positivos, se puede encontrar una sucesión de intervalos abiertos $(I_n)_{n=1}^{\infty}$ tal que*

$$A \subseteq \bigcup_{n=1}^{\infty} I_n \qquad y \qquad \mu(I_n) < \varepsilon_n \quad \text{para todo} \quad n \geq 1.$$

Como la definición de conjunto fuertemente de medida cero no impone restricciones sobre los números positivos ε_n, considerar el caso en que todo ellos son iguales conduce a la conclusión: *todo conjunto **fuertemente de medida** cero es de **medida cero***. Sin embargo, el recíproco de ésta conclusión no es, en general cierto: *existen conjuntos nulos que no son fuertemente de medida cero*.

Teorema 6.4.6. *Si $\mathbb{L}$ es un **conjunto de Lusin**, entonces $\mathbb{L}$ es **fuertemente de medida cero**.*

Prueba. Sea $(\varepsilon_n)_{n=1}^{\infty}$ una sucesión de números reales positivos. Deseamos encontrar una sucesión de intervalos abiertos $(I_n)_{n=1}^{\infty}$ tal que

$$\mathbb{L} \subseteq \bigcup_{n=1}^{\infty} I_n \qquad \text{y} \qquad \mu(I_n) < \varepsilon_n \quad \text{para todo} \quad n \geq 1.$$

Sea $\{q_n : n = 1, 2, \ldots\}$ una enumeración de los racionales en $\mathbb{R}$ y construya, tal como se hizo en la demostración del Teorema 6.4.2, una sucesión $(I_n)_{n=1}^{\infty}$ de intervalos abiertos de modo que $q_n \in I_n$, $\mu(I_n) < \varepsilon_{2n}$ y $\mathbb{R} \nsubseteq \bigcup_{n=1}^{\infty} I_n$. Con esta prescripción resulta que el conjunto $A = \mathbb{R} \setminus \bigcup_{n=1}^{\infty} I_n$ es un conjunto cerrado y nunca-denso ya que no contiene a ningún número racional. Por consiguiente, A es primera categoría y como $\mathbb{L}$ es un conjunto de Lusin se tiene que $\mathbb{L} \cap A$ es a lo más numerable. Sea $\{x_n : n \in N\}$ una enumeración (finita o infinita numerable) de los elemento de $\mathbb{L} \cap A$ y para cada $n \in N$, seleccione un intervalo abierto J_n tal que $x_n \in J_n$ y $\mu(J_n) < \varepsilon_{2n+1}$ para todo $n \in N$. De todo lo anterior se sigue que $\mathbb{L} \subseteq \big(\bigcup_{n=1}^{\infty} I_n\big) \cup \big(\bigcup_{n \in N} J_n\big)$ lo cual muestra que $\mathbb{L}$ es fuertemente de medida cero. ■

Sabemos que todo conjunto medible de medida positiva es no-numerable y que existen conjuntos no-numerables de medida cero. Consideremos este otro hecho: si denotamos por $\mathbb{I}_{[0,1]}$ el conjunto de los números irracionales en $[0,1]$, entonces $\mathbb{I}_{[0,1]}$ es no numerable, denso en $[0,1]$ y, además, se cumple que $\mu(\mathbb{I}_{[0,1]}) = 1$. En base a este ejemplo podemos formularnos la siguiente pregunta: *¿es cualquier conjunto medible, denso y no numerable de medida positiva*? El siguiente es un contraejemplo a nuestra pregunta.

Ejemplo 6.4.1. *Si definimos $E = \Gamma \cup \mathbb{Q}_{[0,1]}$, donde $\mathbb{Q}_{[0,1]}$ es el conjunto de los racionales en $[0,1]$ y Γ es el conjunto ternario de Cantor, resulta que E es un **conjunto medible**, **no numerable**, **denso** en $[0,1]$ y, sin embargo, $\mu(E) = 0$.*

El conjunto ternario de Cantor es un ejemplo de un *conjunto compacto, perfecto y nunca-denso de medida nula*. Podemos preguntarnos: *¿es cierto que todo conjunto compacto, perfecto y nunca-denso de $\mathbb{R}$ posee medida cero*? La respuesta, como lo muestra el siguiente resultado, es que cualquier conjunto tipo-Cantor de medida positiva es un contraejemplo a dicha presunción.

Ejemplo 6.4.2. *Existe un conjunto **compacto**, **perfecto** y **nunca-denso** $P \subseteq \mathbb{R}$ con $\mu(P) > 0$.*

Prueba. Fijemos $\alpha \in (0,1)$ y sea Γ_α el conjunto tipo-Cantor de medida positiva construido en la Sección 5.1.6, página 229. Resulta que Γ_α es compacto, perfecto y nunca-denso. Además, como

$$[0,1] \setminus \Gamma_\alpha = \bigcup_{n=1}^{\infty} J_n^\alpha$$

donde cada J_n^α es la unión disjunta de los 2^{n-1} intervalos abiertos borrados en la etapa n-ésima, cada uno de los cuales tiene longitud $\alpha/2^{2n-1}$, se tiene que:

$$\begin{aligned}
\mu\big([0,1] \setminus \Gamma_\alpha\big) &= \sum_{n=1}^{\infty} \mu(J_n^\alpha) \\
&= \frac{\alpha}{2} + 2\left(\frac{\alpha}{2^3}\right) + 4\left(\frac{\alpha}{2^5}\right) + 8\left(\frac{\alpha}{2^7}\right) + \cdots + 2^{n-1}\left(\frac{\alpha}{2^{2n-1}}\right) + \cdots \\
&= \alpha \cdot \left(\sum_{n=1}^{\infty} \frac{1}{2^n}\right) = \alpha.
\end{aligned}$$

Por el Corolario 6.3.39 se sigue entonces que $\mu(\Gamma_\alpha) = 1 - \alpha > 0$. ■

Otra manera de obtener un conjunto perfecto como el del ejemplo anterior es considerar el conjunto $\mathbb{I}$ de los números irracionales en $[0,1]$. Puesto que $\mathbb{I}$ es medible con $\mu(\mathbb{I}) = 1$, el Teorema 6.3.56, página 294, permite seleccionar un conjunto cerrado $F \subseteq \mathbb{I}$ tal que $\mu(\mathbb{I} \setminus F) < 1/2$. De esto se sigue que $\mu(F) = \mu(\mathbb{I}) - \mu(\mathbb{I} \setminus F) > 1/2$. En particular, F es no-numerable. Por el Teorema de Cantor-Bendixson, página 130, existen un conjunto perfecto P y conjunto a lo más numerable N tal que $F = P \cup N$. Como $\mu(F) = \mu(P) + \mu(N) = \mu(P)$, se tiene entonces que $\mu(P) > 0$. Más aun, P es compacto y nunca-denso ya que no contiene ningún número racional en $[0,1]$.

Teorema 6.4.7. *Cualquier subconjunto G_δ **no-numerable** de $\mathbb{R}$ contiene un **subconjunto no-numerable, cerrado, nunca-denso** y de **medida cero**.*

Prueba. Sea G subconjunto G_δ y no-numerable de $\mathbb{R}$. Seleccione una sucesión $(G_n)_{n=1}^{\infty}$ de conjuntos abiertos tal que $G = \bigcap_{n=1}^{\infty} G_n$. Denote por F el conjunto de todos los puntos de condensación de G y observe que, gracias al Teorema 2.2.21, página 130, $F \neq \varnothing$. Además, cualquier punto de F es un punto límite de F. Sean J_0 y J_1 intervalos cerrados y disjuntos tales que, para cada $i \in \{0,1\}$,

(a_1) $J_i \subseteq G_1$,

(b_1) $\mu(J_i) \leq 1/3$ y

(c_1) $\operatorname{int}(J_i) \cap F \neq \varnothing$.

Observe que $G_2 \cap J_0$ y $G_2 \cap J_1$ son ambos no vacíos pues cada uno de ellos contienen puntos de condensación de G. Seleccione intervalos cerrados y disjuntos J_{00}, J_{01}, J_{10} y J_{11} tales que, para cada $i, j \in \{0,1\}$,

(a_2) $J_{ij} \subseteq G_2 \cap J_i$,

(b_2) $\mu(J_{ij}) \leq 1/3^2$, y

(c_2) $\operatorname{int}(J_{ij}) \cap F \neq \varnothing$.

Continuando de este modo indefinidamente, se obtiene una familia de intervalos cerrados

$$\mathcal{J} = \Big\{ J_{i_1 i_2 \ldots i_n} : n \in \mathbb{N},\, i_j \in \{0,1\} \Big\}$$

tales que cada uno de ellos tiene un punto en común con F y, además, verifican, para cada $n \in \mathbb{N}$, lo siguiente:

(a_3) $J_{i_1 i_2 \ldots i_{n+1}} \subseteq G_n \cap J_{i_1 i_2 \ldots i_n}$ y

(b_3) $\mu(J_{i_1 i_2 \ldots i_n}) \leq 1/3^n$.

Defina, para cada $n \in \mathbb{N}$, el conjunto

$$F_n = \bigcup_{i_1=0}^{1} \bigcup_{i_2=0}^{1} \cdots \bigcup_{i_n=0}^{1} J_{i_1 i_2 \ldots i_n}$$

y sea $F = \bigcap_{n=1}^{\infty} F_n$. Veamos que F tiene las propiedades deseadas. En efecto, observe en primer lugar que F es cerrado ya que cada F_n lo es. Además, como $\mu(F_n) \leq 2^n/3^n$ para todo $n \geq 1$ y la sucesión $(F_n)_{n=1}^{\infty}$ es decreciente, resulta que

$$\mu(F) \;=\; \lim_{n\to\infty} \mu(F_n) \;=\; 0.$$

Se sigue del Teorema 6.4.1 que F es nunca-denso. Finalmente, para ver que F es no numerable, observe que a cada $x \in F$, le corresponde una única sucesión $(i_j)_{j=1}^{\infty}$ de 0's y 1's tal que $x \in J_{i_1 i_2 \ldots i_n}$ para todo $n \geq 1$. Recíprocamente, sea $(i_j)_{j=1}^{\infty}$ una sucesión con $i_j \in \{0,1\}$ y observe que como la sucesión $(J_{i_1 i_2 \ldots i_n})_{n=1}^{\infty}$ es decreciente y sus diámetros tienden a cero, el Teorema de Encaje de Cantor nos garantiza que $\bigcap_{n=1}^{\infty} J_{i_1 i_2 \ldots i_n} = \{x\}$. Este argumento revela que existe una identificación entre los puntos de F y las sucesiones de 0's y 1's, es decir, $2^{\mathbb{N}}$. Puesto que $\operatorname{card}(2^{\mathbb{N}}) = \mathfrak{c}$, resulta que $\operatorname{card}(F) = \mathfrak{c}$ y termina la prueba. ■

La parte (a) siguiente resultado establece que los subconjuntos medibles de un conjunto medible de medida positiva son super abundantes: hay tantos como la cardinalidad del continuo, lo cual nos hace recordar la Propiedad del Valor Intermedio.

Teorema 6.4.8. *Sea E un **conjunto medible** y **acotado** con $\mu(E) > 0$. Para cada $q \in (0, \mu(E))$, se tiene que:*

*(a) Existe un **conjunto medible** $F \subseteq E$ tal que $\mu(F) = q$.*

*(b) Existe un **conjunto perfecto** $P \subseteq E$ con $q < \mu(P) < \mu(E)$.*

*(c) Existen $x, y \in E$ tales que $|x - y|$ es **irracional**.*

Prueba. (a) Defina la función $\varphi : [0, +\infty) \to [0, +\infty)$ por

$$\varphi(x) \;=\; \mu\big(E \cap [-x,\, x]\big) \quad \text{para todo } x \in [0, +\infty).$$

Se verifica fácilmente que:

(i) $\varphi(0) = 0$, y

(ii) φ es creciente y continua.

Puesto que E es acotado, existe un $x_0 \in (0, +\infty)$ tal que $E \subseteq [-x_0, x_0]$ y, en consecuencia,

$$\varphi(x_0) \;=\; \mu\big(E \cap [-x_0,\, x_0]\big) \;=\; \mu(E).$$

Puesto que $q \in (0, \mu(E))$, el Teorema del Valor Intermedio para funciones continuas garantiza la existencia de un $x_q \in (0,\, x_0)$ tal que $\varphi(x_q) = q$. Si ahora definimos $F = E \cap [-x_q,\, x_q]$, vemos que F es un subconjunto medible de E verificando

$$\mu(F) \;=\; \mu\big(E \cap [-x_q,\, x_q]\big) \;=\; \varphi(x_q) \;=\; q.$$

Esto finaliza la prueba de (a)

Para demostrar (b) seleccione, haciendo uso de la parte (a), un conjunto medible $B \subseteq E$ tal que $\mu(B) = (q + \mu(E))/2$. Por otro lado, como B es medible, se sigue del Teorema 6.3.55 que existe un conjunto cerrado $F \subseteq B$ tal que $\mu(B \setminus F) < \mu(B) - q$ y, por lo tanto,

$$\mu(B) - \mu(F) \;=\; \mu(B \setminus F) \;<\; \mu(B) \,-\, q.$$

De aquí se obtiene que $\mu(F) > q$. Finalmente, gracias al Teorema de Cantor-Bendixson, página 130, existen un conjunto perfecto P y conjunto a lo más numerable N tal que $F = P \cup N$ y como $\mu(F) = \mu(P) + \mu(N) = \mu(P)$, resulta que $0 < \mu(P) < \mu(E)$ lo que finaliza la prueba de (b).

Para ver (c), sea P el conjunto perfecto obtenido en (b) y fijemos un $x \in E$. Por el Teorema de Categoría de Baire sabemos que P posee la cardinalidad del continuo y, en consecuencia, es no-numerable. Por lo tanto, el conjunto $A_x = \{|x-y| : y \in P\}$ es no-numerable, de donde se sigue que $|x-y|$ es irracional para algún $y \in P$. ■

Suponga, por ejemplo, que $E = [0,1]$. El Teorema 6.4.8 parte (a) nos muestra que el **rango** de μ es el intervalo $[0,1]$, es decir,

$$\mu\big(\mathfrak{M}_\mu([0,1])\big) = \big\{\mu(E) : E \in \mathfrak{M}_\mu([0,1])\big\} = [0,1].$$

Este resultado es la versión unidimensional de un famoso teorema conocido con el nombre de **Teorema de Convexidad de Lyapunov**.

Ya habíamos visto que, en presencia de la Hipótesis del Continuo, cualquier subconjunto E de $\mathbb{R}$ con medida exterior positiva posee la cardinalidad del continuo y nos habíamos formulado la pregunta de si la conclusión era la misma en ausencia de esa hipótesis. La respuesta es que no se requiere la Hipótesis del Continuo para obtener dicha conclusión.

Corolario 6.4.9. *Sea* $E \in \mathfrak{M}_\mu(\mathbb{R})$ *con* $\mu(E) > 0$. *Entonces* $\operatorname{card}(E) = \mathfrak{c}$.

Prueba. Por la parte (b) del resultado anterior sabemos que E contiene un conjunto perfecto P de medida positiva y como todo conjunto perfecto posee la cardinalidad del continuo (consecuencia del Teorema de Categoría de Baire), resulta que $\operatorname{card}(E) = \mathfrak{c}$.

Otra manera de verificar esto es invocar el Teorema de Steinhaus para obtener un intervalo abierto, digamos $J = (-\delta, \delta)$, tal que $J \subseteq E - E$. De allí que:

$$\mathfrak{c} = \operatorname{card}(J) \leq \operatorname{card}(E-E) = \operatorname{card}(E)\cdot\operatorname{card}(E) \leq \operatorname{card}(\mathbb{R})\cdot\operatorname{card}(\mathbb{R}) = \mathfrak{c}\cdot\mathfrak{c} = \mathfrak{c}.$$

Es nos dice que $\operatorname{card}(E)\cdot\operatorname{card}(E) = \mathfrak{c}$, de donde se deduce que $\operatorname{card}(E) = \mathfrak{c}$. La prueba es completa. ■

6.5. Conjuntos no-medibles

La medida exterior de Lebesgue μ^* restringida a la σ-álgebra $\mathfrak{M}_\mu(\mathbb{R})$ es la que hemos denominado *la medida de Lebesgue sobre* $\mathbb{R}$. Lebesgue no sabía si su medida resolvía **El Problema de la Medida** formulado por él que enunciamos a comienzo de este capítulo, es decir, él no sabía si cualquier conjunto acotado de $\mathbb{R}$ era medible según Lebesgue. En 1905, G. Vitali demostró, usando el Axioma de Elección, que **El Problema de la Medida** de Lebesgue no tenía solución al construir un subconjunto acotado V de $\mathbb{R}$ que no pertenecía a la clase $\mathfrak{M}_\mu(\mathbb{R})$.

Esta sección trata sobre varias construcciones diferentes de *conjuntos no-medibles* según Lebesgue. Todas esas ellas se sustentan sobre la base del Axioma de Elección o, en su defecto, de alguna de sus formas equivalentes tales como el Lema de Zorn, o el Principio del Buen-Orden, y otras sobre la Hipótesis del Continuo. Una vez demostrada la existencia de conjuntos no-medibles, se abrió una especie de Caja de Pandora de tales "monstruos". De hecho, en *todo conjunto con medida* *exterior positiva siempre habita un tal monstruo* lo que permite concluir que la *cantidad* de conjuntos no-medibles es tan numerosa como la de los conjuntos medibles, de hecho, existen $2^{\mathfrak{c}}$ de tales monstruos. No obstante, Sierpiński afirma que la *diversidad de los conjuntos no-medibles es **infinitamente más grande** que la de los conjuntos medibles*. Intentar obtener extensiones de μ a todo $\mathscr{P}(\mathbb{R})$ ha generado algunos de los problemas más interesantes en la Teoría Moderna de Conjuntos: la existencia de los *grandes cardinales*.

6.5.1. Conjunto de Vitali

Aunque el Axioma de Elección es el responsable de muchos resultados hermosos, elegantes y poderosos, él es igualmente responsable de la construcción de una variedad de hechos de apariencia contradictorios y de unas cuantas monstruosidades. Una de tales monstruosidades es el irrenunciable conjunto no-medible de Vitali. La primera persona en exhibir un conjunto realmente extraño y que destruía las aspiraciones de algunos matemáticos de la época de construir una "*medida*" que midiera a todos y cada uno de los subconjuntos de $\mathbb{R}$ y que preservase, además, la noción de longitud de un intervalo, fue la del matemático italiano Giuseppe Vitali (1875-1932). En efecto, en el año 1905, Vitali publicó un artículo titulado "*Sul problema della misura dei gruppi di punti di una retta*" [134] donde, haciendo uso del poderoso, pero polémico *Axioma de Elección*, construía un conjunto que *no era medible* en el sentido de Lebesgue. Algunos años después, y sin conocimiento previo del trabajo de Vitali, tanto Edward Burr Van Vleck (1863-1943) [136], así como Paul Lévy (1886-1971) [94] construían cada uno de ellos, pero separadamente, conjuntos con la misma propiedad de ser no-medibles según Lebesgue. Esos ejemplos dieron origen a la construcción de innumerables conjuntos no-medibles con propiedades monstruosamente excepcionales. Sobre tales "monstruosidades" solo mencionaremos algunas de ellas al final de esta sección. ¿Por qué decimos que el conjunto de Vitali es realmente extraño? Pues bien, supongamos que V es el conjunto *no-medible* construido por Vitali. Como se demostrará un poco más adelante, resulta que dicho conjunto satisface las dos propiedades siguientes:

(aa) $\mu^*(V) > 0$, y

(bb) $\mu([0,1]) = \mu^*([0,1] \setminus V)$.

Como ya hemos visto, la primera condición es obvia pues sabemos, Teorema 6.3.2, que si $\mu^*(V) = 0$, entonces V es un conjunto medible según Lebesgue, lo cual conduce a un contrasentido. En particular, V es no-numerable. La segunda condición es, por supuesto, la que produce lo que hemos llamado un "*conjunto extraño*". En primer lugar, V es un conjunto que posee "*masa*" pues su medida exterior es positiva; sin embargo, la condición (bb) nos revela un hecho realmente paradójico y, por lo tanto, sorprendente: el conjunto V viola la *ley de la conservación de la masa*, pues al sustraer de $[0,1]$ un conjunto que contiene "masa", en este caso $\mu^*(V) > 0$, el conjunto resultante $[0,1] \setminus V$ sigue teniendo "*la misma masa*" que $[0,1]$.

Giuseppe Vitali se graduó en la Scuola Normale Superiore de Pisa en 1899. Trabajó dos años con Ulisse Dini antes de entrar a la política como representante del Partido Socialista en el Ayuntamiento de Génova. Con la llegada al poder de los fascistas en 1922 y la disolución del Partido Socialista, regresa a las Matemáticas. En 1926 sufrió un derrame cerebral dejando la mitad de su cuerpo paralizado, lo cual no le impidió seguir haciendo importantes contribuciones al Análisis hasta su muerte ocurrida en 1932 de un ataque al corazón.

El siguiente resultado, producto de la mente brillante de G. Vitali, establece la existencia de un conjunto no-medible en el intervalo $[0,1]$ fabricado usando el imprescindible Axioma de Elección.

Teorema 6.5.1 (Vitali). *En el intervalo $[0,1]$ existe un subconjunto no vacío V que **no es medible** según Lebesgue.*

Prueba. Nuestro primer paso es construir una partición muy especial de $[0,1]$ a partir de la siguiente relación de equivalencia: dados dos números cualesquiera $x, y \in [0,1]$, escribiremos

$$x \sim y \qquad \text{si, y sólo si,} \qquad y - x \in \mathbb{Q}.$$

Es fácil establecer que $\sim$ es una relación de equivalencia sobre $[0,1]$. Consideremos ahora la familia $\mathcal{A} = \{[x] : x \in [0,1]\}$ formada por todas las clases de equivalencias determinadas por la relación $\sim$, esto es,

$$[x] = \{y \in [0,1] : y \sim x\} = \{y \in [0,1] : x - y \in \mathbb{Q}\}$$

para cada $x \in [0,1]$. Por ejemplo, escribiendo $\mathbb{Q}_{[0,1]} = \mathbb{Q} \cap [0,1]$, tenemos

$$\begin{aligned}
[0] &= \{y \in [0,1] : 0 - y \in \mathbb{Q}\} &&= \mathbb{Q}_{[0,1]} \\
[1/2] &= \{y \in [0,1] : 1/2 - y \in \mathbb{Q}\} &&= \frac{1}{2} + \mathbb{Q}_{[0,1]} \\
[\sqrt{2}/2] &= \{y \in [0,1] : 1/\sqrt{2} - y \in \mathbb{Q}\} &&= \frac{1}{\sqrt{2}} + \mathbb{Q}_{[0,1]} \\
[\pi/4] &= \{y \in [0,1] : \pi/4 - y \in \mathbb{Q}\} &&= \frac{\pi}{4} + \mathbb{Q}_{[0,1]},
\end{aligned}$$

y así sucesivamente. Observe que

(a) Cada clase de equivalencia $[x]$ es numerable.

(b) Si x es racional, entonces cada miembro de $[x]$ es un número racional.

(c) Si x es irracional, entonces cada miembro de $[x]$ es un número irracional.

(d) Puesto que $x \in [x]$ para cada $x \in [0,1]$, resulta que

$$[0,1] = \bigcup_{x \in [0,1]} [x],$$

(e) Para cualquier par de elementos distintos $x, y \in [0,1]$, las clases $[x]$ y $[y]$ en $\mathcal{A}$ o *son iguales*, o *son disjuntas*. En efecto, sean $x, y \in [0,1]$ con $x \neq y$. Entonces debe ocurrir que: $x \sim y$, o bien $x \not\sim y$.

(e_1) Si $x \sim y$, resulta entonces que $x - y \in \mathbb{Q}$. Veamos que esto conduce a la igualdad $[x] = [y]$. En efecto, sea $a \in [x]$ y seleccione un $r_a \in \mathbb{Q}$ tal que $x - a = r_a$. Como $x - y \in \mathbb{Q}$, existe $r \in \mathbb{Q}$ tal que $x = y + r$, de donde se sigue que

$$a = x - r_a = (y + r) - r_a = y + (r - r_a) \in A_y.$$

Esto es, $[x] \subseteq [y]$. De modo enteramente similar se prueba que $[y] \subseteq [x]$.

(e_2) Si $x \not\sim y$, entonces $x - y \notin \mathbb{Q}$. La conclusión, en este caso, es que $[x] \cap [y] = \varnothing$. Suponga, por un momento, que $[x] \cap [y] \neq \varnothing$ y seleccione un $a \in [x] \cap [y]$. Sean r_x y r_y números racionales tales que

$$x - a = r_x \qquad \text{y} \qquad y - a = r_y.$$

Entonces $x - y = r_y - r_x \in \mathbb{Q}$, lo que contradice el hecho de que $x - y \notin \mathbb{Q}$. Por esto, $[x] \cap [y] = \varnothing$.

Observe que la colección $\mathcal{A} = \{[x] : x \in [0,1]\}$ es no-numerable y constituye una partición del intervalo $[0,1]$ gracias a las condiciones (d) y (e). En particular, $\operatorname{card}(\mathcal{A}) = \mathfrak{c}$.

Usemos ahora el Axioma de Elección para construir un nuevo conjunto, llamémoslo V, eligiendo de cada clase de equivalencia $[x]$ uno, y sólo un punto. Cualquier conjunto V construido de este modo se le denomina un ***conjunto de Vitali***. Observe que,

$$v \in V \text{ si, y sólo si, existe un único } x \in [0,1] \text{ tal que } V \cap [x] = \{v\}.$$

Más aun, V es un subconjunto *no-numerable* de $[0,1]$. Para finalizar la demostración del teorema sólo resta probar que el conjunto V no puede ser medible según Lebesgue.

(f) ***V no es medible según Lebesgue***.

Primera Prueba. Sea $D = \{r_n : n = 0, 1, \ldots\}$ una enumeración de los racionales en $[-1,1]$ sin duplicaciones, con $r_0 = 0$. Para cada $n \in \mathbb{N}_0$, definamos

$$V_n = r_n + V$$

La sucesión $(V_n)_{n=0}^{\infty}$ posee las siguientes dos propiedades:

(f_1) $(V_n)_{n=0}^{\infty}$ *es una colección disjunta*.

Afirmamos que $V_n \cap V_m = \varnothing$ si $r_n \neq r_m$. Suponga, para generar una contradicción, que $V_n \cap V_m \neq \varnothing$ y sea $x \in V_n \cap V_m$. Entonces existen $x', x'' \in V$ tales que

$$x = r_n + x' = r_m + x''.$$

De esto se sigue que $x' - x'' = r_m - r_n \in \mathbb{Q}$ y así, gracias a (e_1), tenemos que $[x'] = [x'']$, es decir, x' y x'' pertenecen a una misma clase de equivalencia y, entonces, como V fue construido eligiendo un único punto en cada clase de equivalencia, resulta que $x' = x''$. De esto se deduce que $r_n = r_m$, lo cual es imposible. Esta contradicción establece que $V_n \cap V_m = \varnothing$.

(f_2) $[0,1] \subseteq \bigcup_{n=0}^{\infty} V_n \subseteq [-1,2]$.

Para demostrar estas inclusiones, sea $a \in [0,1] = \bigcup_{x \in [0,1]} [x]$. Entonces existe un único $x_0 \in [0,1]$ tal que $a \in [x_0]$. Sea v el único elemento de $[x_0]$ que pertenece a V. Entonces, $a \sim v$; es decir, $a - v = r_{n_0}$ para algún racional r_{n_0}. Esto nos dice que $a \in V_{n_0} \subseteq \bigcup_{n=0}^{\infty} V_n$. La otra inclusión es inmediata.

Para verificar que V no es medible según Lebesgue, vamos a usar el hecho de μ es numerablemente aditiva e invariante por traslación. Suponga, por un momento, que *V es medible según Lebesgue*. La invariancia por traslación de μ, Teorema 6.3.44, nos revela que cada V_n es medible y que $\mu(V_n) = \mu(V)$ para todo $n \in \mathbb{N}_0$. Existen sólo dos posibilidades para V: o bien $\mu(V) = 0$, o bien, $\mu(V) > 0$:

(g_1) Suponga que $\mu(V) = 0$. Como μ es numerablemente aditiva, con la primera inclusión en (f_2) se origina la siguiente contradicción:

$$1 \;=\; \mu([0,1]) \;\leq\; \mu\Big(\bigcup_{n=0}^{\infty} V_n\Big) \;=\; \sum_{n=0}^{\infty} \mu(V_n) \;=\; \sum_{n=0}^{\infty} \mu(V) \;=\; 0.$$

(g_2) Si ahora suponemos que $\mu(V) > 0$, entonces la segunda inclusión en (f_2) nos conduce a esta otra contradicción:

$$3 \;\geq\; \mu\Big(\bigcup_{n=0}^{\infty} V_n\Big) \;=\; \sum_{n=0}^{\infty} \mu(V_n) \;=\; \sum_{n=0}^{\infty} \mu(V) \;=\; +\infty.$$

Con las contradicciones obtenidas en (g_1) y (g_2) queda demostrado que V no puede ser medible según Lebesgue y termina la prueba. ■

Segunda Prueba. Suponga que V es medible y, como antes, sea $D = \{r_n \,:\, n = 0, 1, \ldots\}$ una enumeración de los racionales en $[-1, 1]$ sin duplicaciones con $r_0 = 0$. Puesto que

$$[0,1] \;=\; \bigcup_{x\,\in\,[0,1]} [x] \;\subseteq\; \bigcup_{n=0}^{\infty} (r_n + V),$$

entonces la subaditividad de μ^* más el hecho de μ^* también es invariante por traslación, revelan que $1 = \mu^*([0,1]) \leq \sum_{n=0}^{\infty} \mu^*(V)$ y, por lo tanto, $\mu^*(V) > 0$. Se sigue del Teorema de Steinhaus, Teorema 6.3.54, que $V - V$ contiene un intervalo abierto $J = (-\delta, \delta)$ para algún $\delta > 0$ tal que $(x + V) \cap V \neq \varnothing$ para todo $x \in (-\delta, \delta)$. Sin embargo, si x es un racional distinto de cero, entonces se sigue de (e_2) que $(x + V) \cap V = \varnothing$. Esta contradicción establece el resultado. ■

Una vez establecido que V es un conjunto no-medible, del Teorema 6.3.44 resulta que *cada uno de los conjuntos* V_n *definidos anteriormente también es no-medible*. Más aun, $[0,1] \setminus V$ *es no-medible* lo que conduce a otra sorpresa: $[0,1]$ *se puede expresar como la unión de dos conjuntos disjuntos no-medibles*: $[0,1] = V \cup ([0,1] \setminus V)$. Otras consecuencias interesantes que se obtienen del conjunto V de Vitali son las siguientes.

(h_1) $\mu^*(V) > 0$ pues, suponer que $\mu^*(V) = 0$, conlleva a que V sería medible lo que resulta ser imposible.

(h_2) $\mu_*(V) = 0$.

Prueba. Sea $K \subseteq V$ con K compacto. Veamos que $\mu(K) = 0$. En efecto, como $K \subseteq V$, resulta que $r_n + K \subseteq r_n + V = V_n$ para todo $n \in \mathbb{N}_0$ y, en consecuencia, por ser $(V_n)_{n=0}^{\infty}$ una sucesión disjunta también lo es la sucesión de conjuntos medibles $(r_n + K)_{n=0}^{\infty}$. Por otro lado, teniendo en cuenta que $D = \{r_n \,:\, n = 0, 1, \ldots\}$ y K son conjuntos medibles y acotados, entonces $D + K$ es medible y acotado y, por lo tanto, $\mu(D + K) < +\infty$. Usando ahora el hecho de que μ es numerablemente aditiva e invariante por traslación, tenemos que

$$\begin{aligned} +\infty \;>\; \mu(D+K) \;&=\; \mu\Big(\bigcup_{n=0}^{\infty} (r_n + K)\Big) \\ &=\; \sum_{n=0}^{\infty} \mu\big(r_n + K\big) \;=\; \sum_{n=0}^{\infty} \mu(K). \end{aligned}$$

De esto último se concluye que $\mu(K) = 0$ y como K era un subconjunto compacto arbitrario incluido en V, resulta que

$$\mu_*(V) \;=\; \sup\big\{\mu(K) \,:\, K \subseteq V,\; K \text{ compacto}\big\} \;=\; 0.$$

(h_3) *Si E y F son conjuntos medibles de $[0,1]$ tales que $E \subseteq V \subseteq F$, entonces $\mu(E) = 0$ y $\mu(F) = 1$.*

Prueba. Suponga que $E \subseteq V$ es medible. Por el Lema 6.3.64 y el Teorema 6.3.65 se tiene que

$$0 \;\leq\; \mu(E) \;=\; \mu^*(E) \;=\; \mu_*(E) \;\leq\; \mu_*(V) \;=\; 0.$$

Para demostrar que $\mu(F) = 1$, observe que $F^c \subseteq V^c \subseteq E^c$. Como V^c es no-medible, se sigue de la primera parte que $\mu(F^c) = 0$ y, por consiguiente, $1 = \mu([0,1]) = \mu(F) + \mu(F^c) = \mu(F)$.

(h_4) $\mu([0,1]) = \mu^*([0,1] \setminus V)$.

Prueba. Por el Teorema 6.3.67 sabemos que

$$\mu([0,1]) = \mu_*(V) + \mu^*([0,1] \setminus V)$$

y como $\mu_*(V) = 0$ el resultado sigue.

(h_5) μ^* ***no es numerablemente aditiva****, ni aun* ***finitamente aditiva****.*

Prueba. La sucesión de conjuntos no medibles $(V_n)_{n=0}^{\infty}$ satisface

$$\mu^*\Big(\bigcup_{n=0}^{\infty} V_n\Big) \;<\; \sum_{n=0}^{\infty} \mu^*(V_n).$$

En efecto, puesto que $\mu^*(V_n) = \mu^*(V) > 0$, si μ^* fuese numerablemente aditiva obtendríamos la siguiente contradicción:

$$3 \;\geq\; \mu^*\Big(\bigcup_{n=0}^{\infty} V_n\Big) \;=\; \sum_{n=0}^{\infty} \mu^*(V_n) \;=\; \sum_{n=0}^{\infty} \mu^*(V) \;=\; +\infty.$$

Por otro lado, como $[0,1] = V \cup \big([0,1] \setminus V\big)$ y $\mu^*(V) > 0$, resulta de (h_3) que

$$1 \;=\; \mu([0,1]) \;=\; \mu^*([0,1] \setminus V) \;<\; \mu^*([0,1] \setminus V) \;+\; \mu^*(V).$$

lo cual prueba que μ^* tampoco es finitamente aditiva. ■

(h_6) $\operatorname{int}(V) = \varnothing$. En efecto, si $\operatorname{int}(V) \neq \varnothing$, entonces V contendría un intervalo abierto, digamos I, y se seguiría de (h_3) que $\mu(I) = 0$ lo que es imposible. En particular, ***V no es de primera categoría***. Es consecuencia inmediata del Teorema de Categoría de Baire ya que la sucesión $(r_n + V)_{n=1}^{\infty}$ cubre a $[0,1]$ y cada uno de ellos tiene interior vacío.

La condición (h_3) es realmente curiosa: *cualquier subconjunto medible incluido en V posee medida cero*. Como veremos un poco más abajo, el conjunto de Bernstein y muchos otros conjuntos no-medibles, poseen esa condición. En general, el siguiente resultado nos da información de cómo se obtienen conjuntos no-medibles.

Teorema 6.5.2. *Sea $M \subseteq [0,1]$ tal que:*

*(a) $\mu(N) = 0$ para cualquier **conjunto medible** $N \subseteq M$ y*

*(b) $\mu(E) = 1$ para cualquier **conjunto medible** $E \supseteq M$.*

*Entonces M **no es medible** según Lebesgue.*

Prueba. Suponga, para generar una contradicción, que M es medible. Tomando $N = M$ y $E = M$, se tiene que $N \subseteq M \subseteq E$ y entonces, por (a), $\mu(M) = \mu(N) = 0$, mientras que (b) nos dice que $\mu(M) = \mu(E) = 1$ lo que es, evidentemente, una contradicción. Por esto, M no es medible según Lebesgue. ■

Al comienzo de este capítulo habíamos formulado el Problema de la Medida de Lebesgue y adelantamos la opinión de que **la medida de Lebesgue** no podía asignar a cada subconjunto de $\mathbb{R}$ un número real extendido. Pues bien, con la prueba de la existencia de conjuntos no-medibles establecida anteriormente, el cual es posible bajo el imperio del **Axioma de Elección**, tenemos ahora la certeza de que

$$\mathfrak{M}_\mu(\mathbb{R}) \subsetneqq \mathscr{P}(\mathbb{R}),$$

de donde resulta que la condición (α_1) en *el Problema de la Medida de Lebesgue no se cumple para la medida de Lebesgue* y, en consecuencia, **dicho problema no admite solución**.

En lo sucesivo escribiremos

$$\neg\,\mathfrak{M}_\mu(\mathbb{R}) \;=\; \mathscr{P}(\mathbb{R}) \setminus \mathfrak{M}_\mu(\mathbb{R})$$

para designar a *la familia de todos los subconjuntos de $\mathbb{R}$ que son **no-medibles según Lebesgue***.

Ya hemos visto que los conjuntos no-medibles no pueden vivir dentro de conjuntos cuya medida exterior es nula; en particular, ***todo conjunto no-medible es no-numerable***. Este hecho demuestra que:

> *La noción de medida exterior de Lebesgue más el Axioma de Elección sirven, además del Método de la Diagonal de Cantor y los números ordinales, como instrumentos para obtener conjuntos no-numerables.*

¿Podemos encontrar conjuntos no-medibles en cualquier conjunto cuya medida exterior es positiva? El siguiente resultado, dado a conocer por Hans Rademacher (1892-1969) y que también descansa sobre el Axioma de Elección, establece que la respuesta es siempre afirmativa.

Teorema 6.5.3 (Rademacher). *Sea $A \subseteq \mathbb{R}$ con $\mu^*(A) > 0$. Entonces existe un conjunto no vacío $V \subseteq A$ tal que $V \notin \mathfrak{M}_\mu(\mathbb{R})$.*

Prueba. Sin perder generalidad en el razonamiento, asumiremos que $A \subseteq [0,1]$. En efecto, puesto que $\mathbb{R} = \bigcup_{n\in\mathbb{Z}}[n, n+1]$, resulta que $A = A \cap \mathbb{R} = \bigcup_{n\in\mathbb{Z}} A \cap [n, n+1]$ y, entonces

$$0 \;<\; \mu^*(A) \;\leq\; \sum_{n\in\mathbb{Z}} \mu^*\big(A \cap [n, n+1]\big).$$

De aquí se sigue la existencia de al menos un $n_0 \in \mathbb{Z}$ tal que

$$\mu^*\big(A \cap [n_0, n_0+1]\big) \;>\; 0.$$

Sea $B_0 = A \cap [n_0, n_0 + 1]$ y observe que el conjunto $A_0 = -n_0 + B \subseteq [0,1]$ satisface $\mu^*(A_0) > 0$. Por consiguiente, si logramos demostrar que A_0 contiene un conjunto no-medible, entonces B también contendría un conjunto no-medible y, en consecuencia, A tendría la misma propiedad.

Supongamos entonces que $A \subseteq [0,1]$ y sea V el conjunto no-medible de Vitali obtenido anteriormente. Como antes, sea $(r_n)_{n=1}^{\infty}$ una enumeración de los racionales, sin repetición, en $[-1,1]$ y sea $V_n = r_n + V$ para todo $n \in \mathbb{N}$.

Suponga ahora, para fabricar una contradicción, que *cualquier subconjunto de A es medible*. Observe que, para cada $n \in \mathbb{N}$, el conjunto $V_n \cap A$, siendo un subconjunto de A es, por nuestra suposición, medible y, por supuesto, también está incluido en V_n. Afirmamos que:

$$\mu\big(V_n \cap A\big) \;=\; 0 \qquad \text{para todo} \quad n \in \mathbb{N}. \tag{1}$$

En efecto, si para algún $n \in \mathbb{N}$, ocurre que $\mu\big(V_n \cap A\big) > 0$, entonces usando el hecho de que μ_* es monótona y que $\mu_*(V_n) = \mu_*(V) = 0$ tendremos, por el Teorema 6.3.65, página 301, que

$$0 \;<\; \mu\big(V_n \cap A\big) \;=\; \mu_*\big(V_n \cap A\big) \;\leq\; \mu_*\big(V_n\big) \;=\; 0$$

lo cual es imposible. Esto prueba nuestra afirmación. Finalmente, como $[0,1] \subseteq \bigcup_{n=1}^{\infty} V_n$, entonces $A \subseteq \bigcup_{n=1}^{\infty}(V_n \cap A)$ de donde se sigue, usando (1), que

$$0 \;<\; \mu^*\big(A\big) \;\leq\; \sum_{n=1}^{\infty} \mu^*\big(V_n \cap A\big) \;=\; 0,$$

lo que origina, una vez más, una contradicción. Por consiguiente, la suposición de que cualquier subconjunto de A es medible conduce a una contradicción y, en consecuencia, nuestro conjunto A debe contener algún conjunto no-medible. ■

Varias consecuencia se pueden derivar del resultado anterior. Comencemos con la primera:

Corolario 6.5.4. *Para cada $\varepsilon > 0$, existe un **conjunto no-medible** V tal que*

$$0 \;<\; \mu^*(V) \;<\; \varepsilon.$$

Prueba. Como $\mu([0, \varepsilon/2]) = \varepsilon/2 > 0$, el Teorema de Rademacher termina la prueba. ■

El siguiente corolario confirma, una vez más, que μ^* no es finitamente aditiva.

Corolario 6.5.5. *Sea $E \in \mathfrak{M}_\mu(\mathbb{R})$ con $\mu(E) > 0$. Entonces para cualquier **conjunto no-medible** $V \subseteq E$ se tiene que*

$$\mu^*(E) \;<\; \mu^*(V) \,+\, \mu^*(E \setminus V).$$

Prueba. Sin perder generalidad, asumiremos que E es acotado. Sea V un subconjunto no-medible incluido en E (el cual existe por el Teorema de Rademacher). Si ocurriese que

$$\mu^*(E) \;=\; \mu^*(V) \,+\, \mu^*(E \setminus V),$$

entonces el Teorema 6.3.60, página 296, nos diría que V es medible. Esta contradicción combinado con la subaditividad de μ^* finaliza la prueba. ■

Ya vimos en (h_4) que el complemento V^c del conjunto no-medible de Vitali, el cual también es no-medible, satisface la igualdad $\mu^*(V^c) = \mu([0,1])$. En el siguiente resultado se prueba que esa es la regla y no la excepción.

Teorema 6.5.6. *Sea $E \in \mathfrak{M}_\mu(\mathbb{R})$ con $\mu(E) > 0$. Entonces existe un **conjunto no-medible** $A \subseteq E$ tal que*

$$\mu^*(A) \;=\; \mu(E).$$

Prueba. En primer lugar, vamos a demostrar la existencia de un conjunto no-medible $A \subseteq E$, tal que

$$\mu^*(A) \geq \frac{1}{2}\mu(E). \tag{1}$$

En efecto, seleccione, usando el Teorema de Rademacher, un conjunto no-medible $V_0 \subseteq E$.

(a) Si $0 < \mu(E) < +\infty$, entonces por el Corolario 6.5.5 se tiene que

$$\mu(E) \;<\; \mu^*(V_0) \,+\, \mu^*(E \setminus V_0),$$

de donde se sigue que

$$\mu^*(V_0) \;>\; \frac{1}{2}\mu(E) \qquad\qquad \text{o} \qquad\qquad \mu^*(E \setminus V_0) \;>\; \frac{1}{2}\mu(E). \tag{2}$$

También, como $E \setminus V_0$ es no-medible, resulta que tomando $A = V_0$ o $A = E \setminus V_0$, según se cumpla (2), el resultado (1) sigue.

(b) Si $\mu(E) = +\infty$, entonces, usando el hecho de que μ es σ-finita, elija una sucesión disjunta $(E_n)_{n=1}^{\infty}$ en $\mathfrak{M}_\mu(\mathbb{R})$ tal que $E = \bigcup_{n=1}^{\infty} E_n$ y $\mu(E_n) < +\infty$ para todo $n \geq 1$ (por supuesto, podemos suponer que $0 < \mu(E_n) < +\infty$ para todo $n \geq 1$). Por cada $n \in \mathbb{N}$ seleccione, invocando la parte (a), un conjunto no-medible $A_n \subseteq E_n$ tal que $\mu^*(A_n) > \frac{1}{2}\mu(E_n)$. Sea $A = \bigcup_{n=1}^{\infty} A_n$. Por el Teorema 6.3.61, página 297, se tiene que

$$A \notin \mathfrak{M}_\mu(\mathbb{R}) \qquad\qquad \text{y} \qquad\qquad \mu^*(A) \;=\; \sum_{n=1}^{\infty} \mu^*(A_n) \;\geq\; \frac{1}{2}\sum_{n=1}^{\infty} \mu^*(A_n) \;=\; \frac{1}{2}\mu(E).$$

Esto finaliza la demostración de la desigualdad (1).

Pongamos ahora $r_0 = \mu(E)$ y $B_0 = \varnothing$. Usemos la primera parte para elegir un conjunto no-medible $A_1 \subseteq E$ tal que $\mu^*(A_1) \geq \frac{1}{2}\mu(E) = \frac{r_0}{2}$. También, por el Corolario 6.2.13, página 249, existe $B_1 \in \mathfrak{M}_\mu(\mathbb{R})$ tal que

$$A_1 \subseteq B_1 \subseteq E \setminus B_0 \qquad\qquad \text{y} \qquad\qquad \mu(B_1) \;=\; \mu^*(A_1).$$

Sea $r_1 = \mu(E \setminus B_1)$. Claramente, $0 \leq r_1 \leq r_0/2$. Si $r_1 = 0$ nos detenemos ya que, en este caso, $\mu^*(A_1) = \mu(E)$. Suponga entonces que $r_1 > 0$. Con un argumento similar al anterior, podemos seleccionar dos conjuntos A_2 y B_2 tales que

$$A_2 \notin \mathfrak{M}_\mu(\mathbb{R}), \qquad B_2 \in \mathfrak{M}_\mu(\mathbb{R}), \qquad A_2 \subseteq B_2 \subseteq E \setminus (B_0 \cup B_1) \qquad \text{y} \qquad \mu(B_2) \;=\; \mu^*(A_2).$$

Defina $r_2 = \mu(E \setminus (B_0 \cup B_1))$. Entonces $0 \leq r_2 \leq r_1/2 \leq r_0/2^2$. Suponga que $(A_i)_{i=1}^n$, $(B_i)_{i=0}^n$ y $(r_i)_{i=0}^n$ han sido construidos de modo tal que, para cada $i \in \{1, 2, \ldots, n\}$,

$$A_i \notin \mathfrak{M}_\mu(\mathbb{R}), \qquad B_i \in \mathfrak{M}_\mu(\mathbb{R}), \qquad A_i \subseteq B_i \subseteq E \setminus \left(\bigcup_{i=0}^{n-1} B_i\right), \qquad \mu(B_i) \;=\; \mu^*(A_i)$$

y

$$r_i \;=\; \mu\Big(E \setminus \bigcup_{i=0}^{n-1} B_i\Big) \;\leq\; r_0/2^i.$$

Claramente la familia de conjuntos medibles $\{B_0, B_1, \ldots, B_n\}$ es disjunta y entonces, por el Teorema 6.3.61 se tiene que $\bigcup_{i=1}^{n} A_i \notin \mathfrak{M}_\mu(\mathbb{R})$ y

$$\begin{aligned} \mu^*\Big(\bigcup_{i=1}^{n} A_i\Big) \;&=\; \sum_{i=1}^{n} \mu^*(A_i) \;=\; \sum_{i=1}^{n} \mu(B_i) \\ &=\; \mu\Big(\bigcup_{i=1}^{n} B_i\Big) \;=\; \mu(E) - r_n. \end{aligned}$$

Si $r_n = 0$ paramos la construcción ya que el conjunto $A = \bigcup_{i=1}^{n} A_i$ satisface, gracias a la igualdad anterior, la conclusión del teorema. Si $r_n > 0$, entonces procediendo como antes seleccione conjuntos $A_{n+1} \notin \mathfrak{M}_\mu(\mathbb{R})$ y $B_{n+1} \in \mathfrak{M}_\mu(\mathbb{R})$ tales que

$$A_{n+1} \subseteq B_{n+1} \subseteq E \setminus \Big(\bigcup_{i=0}^{n} B_i\Big), \qquad \mu(B_{n+1}) \;=\; \mu^*(A_{n+1}).$$

Defina $r_{n+1} = \mu(E \setminus \bigcup_{i=0}^{n} B_i)$. Si el proceso no termina, se obtienen las tres sucesiones $(A_n)_{n=1}^{\infty}$, $(B_n)_{n=1}^{\infty}$ y $(r_n)_{n=1}^{\infty}$. Sea $A = \bigcup_{n=1}^{\infty} A_n$. Por el Teorema 6.3.61 se tiene que $A \notin \mathfrak{M}_\mu(\mathbb{R})$ y

$$\mu^*(A) \;=\; \sum_{n=1}^{\infty} \mu^*(A_n) \;=\; \sum_{n=1}^{\infty} \mu(B_n) \;=\; \mu\Big(\bigcup_{n=1}^{\infty} B_n\Big).$$

De esto se sigue que, para todo $n \geq 1$,

$$\mu(E) \;\geq\; \mu^*(A) \;\geq\; \mu\Big(\bigcup_{i=1}^{n} B_i\Big) \;=\; \mu(E) \;-\; r_n \;\geq\; \mu(E)(1 - 1/2^n)$$

y, en consecuencia, tomando límite cuando $n \to \infty$, se obtiene que $\mu^*(A) = \mu(E)$. ■

Otra consecuencia del Teorema de Rademacher es que *en cualquier conjunto tipo-Cantor de medida positiva habita un conjunto no-medible* y, por consiguiente:

Corolario 6.5.7. *Existen conjuntos **cerrados, nunca-densos** de **medida positiva** conteniendo **conjuntos no-medibles**.*

6.5.2. Conjunto no-medible de un Grupo Aditivo

En el construcción del conjunto de Vitali usamos el hecho de que $\mathbb{Q}$ es un subgrupo aditivo de $\mathbb{R}$. En esta parte vamos a construir un conjunto no-medible $\mathbf{A}$ en un cierto subgrupo aditivo de $\mathbb{R}$ tal que cualquier conjunto medible E incluido en $\mathbf{A}$ o en $\mathbf{A}^c$ posee medida cero. Fijemos un número irracional ξ y considere los conjuntos

$$\begin{aligned} G_1 &= \{r + n\xi : r \in \mathbb{Q},\ n \in \mathbb{Z}\} = \mathbb{Q} + \xi\mathbb{Z} \\ G_2 &= \{r + 2n\xi : r \in \mathbb{Q},\ n \in \mathbb{Z}\} \\ G_3 &= \{r + (2n+1)\xi : r \in \mathbb{Q},\ n \in \mathbb{Z}\} \end{aligned}$$

Se comprueba fácilmente que G_1 y G_2 son subgrupos aditivos de $\mathbb{R}$,

$$G_2 \cap G_3 = \varnothing, \qquad G_3 = G_2 + \xi \qquad \text{y} \qquad G_1 = G_2 \cup G_3.$$

Defina la relación $\sim$ sobre $\mathbb{R}$ del modo siguiente: para todo $x, y \in \mathbb{R}$,

$$x \sim y \qquad \Leftrightarrow \qquad x - y \in G_1.$$

Se comprueba fácilmente que $\sim$ es una relación de equivalencia sobre $\mathbb{R}$. Usemos ahora el Axioma de Elección para formar un subconjunto A_0 de $\mathbb{R}$ que contenga exactamente un punto de cada una de las clases de equivalencias determinadas por $\sim$. Sea

$$\mathbf{A} = A_0 + G_2,$$

es decir, $\mathbf{A}$ consiste de todos los puntos que tienen la forma $a + g_2$ para algún $a \in A_0$ y algún $g_2 \in G_2$. Veamos que $\mathbf{A}$ posee las propiedades requeridas.

(a) *Cualquier subconjunto medible E de $\mathbf{A}$ es de medida cero*. Para ver esto, suponga que existe un conjunto medible $E \subseteq \mathbf{A}$ con $\mu(E) > 0$. Se sigue del Teorema de Steinhaus que existe un intervalo abierto, digamos $J = (-\varepsilon, \varepsilon)$, tal que $J \subseteq E - E \subseteq \mathbf{A} - \mathbf{A}$. Puesto que G_3 es denso en $\mathbb{R}$, resulta que $G_3 \cap J \neq \varnothing$; en particular, $G_3 \cap (\mathbf{A} - \mathbf{A}) \neq \varnothing$. Sin embargo, esto último es imposible que ocurra. En efecto, cada elemento de $\mathbf{A} - \mathbf{A}$ es de la forma $a_1 - a_2 + g_2$ donde $a_1, a_2 \in A_0$ y $g_2 \in G_2$ y, en consecuencia, no puede pertenecer a G_3 (la relación $a_1 - a_2 + g_2 = g_3$ implicaría que $a_1 = a_2$ y $g_2 = g_3$ lo cual es imposible ya que $G_2 \cap G_3 = \varnothing$). Esta contradicción establece que $\mu(E) = 0$.

(b) *Cualquier subconjunto medible E de $\mathbf{A}^c$ es de medida cero*. Para verificar esta afirmación, observe que $\mathbf{A}^c = A_0 + G_3$, de allí que $\mathbf{A}^c = \mathbf{A} + \xi$. De esto último se sigue que cualquier subconjunto medible de $\mathbf{A}^c$ es de la forma $E + \xi$ para algún subconjunto medible E de $\mathbf{A}$. Puesto que todos los subconjuntos medibles de $\mathbf{A}$ son de medida cero y μ es invariante por traslación, resulta que los subconjuntos medibles de $\mathbf{A}^c$ también son de medida cero.

(c) ***$\mathbf{A}$ no es medible según Lebesgue***. En efecto, si $\mathbf{A}$ fuese medible, entonces también lo sería $\mathbf{A}^c$ y $\mathbb{R} = \mathbf{A} \cup \mathbf{A}^c$. Sea E cualquier subconjunto medible de $\mathbb{R}$ con $\mu(E) > 0$. Entonces $E \cap \mathbf{A}$ es un conjunto medible incluido en $\mathbf{A}$ y, por lo tanto, su medida es cero. Similarmente, $\mu(E \cap \mathbf{A}^c) = 0$. Estos hechos producen la siguiente contradicción:

$$0 < \mu(E) = \mu(E \cap \mathbf{A}) + \mu(E \cap \mathbf{A}^c) = 0.$$

6.5.3. Conjunto Saturado no-medible

En esta sección veremos cómo se construye, con un argumento muy similar al de la sección anterior, un caso muy particular de un conjunto no-medible pero esta vez con propiedades relacionadas con la medida interior. En los ejercicios, al final de este capítulo, se proponen otras formas de obtener tales conjuntos. En este sentido, también es apropiado echarle una miradita al libro de M. Kuczma [82].

Definición 6.5.8. *Un conjunto $A \subseteq \mathbb{R}$ se dice que es **saturado no-medible** si*

$$\mu_*(A) \;=\; \mu_*(A^c) \;=\; 0.$$

Nótese que *si existe algún conjunto saturado no-medible $A \subseteq \mathbb{R}$, entonces automáticamente él es no-medible según Lebesgue*, lo cual justifica su nombre. En efecto, suponga A es medible. Entonces, A^c también es medible y como $\mathbb{R} = A \cup A^c$, se sigue de la aditividad de μ y del Teorema 6.3.65 que

$$+\infty \;=\; \mu(\mathbb{R}) \;=\; \mu(A) + \mu(A^c) \;=\; \mu_*(A) + \mu_*(A^c) \;=\; 0$$

lo que, por supuesto, es imposible.

Una consecuencia del Teorema 6.3.65 y del Teorema 6.3.68 es que la existencia de un conjunto saturado no-medible da origen a una cantidad enorme de conjuntos no-medibles.

Corolario 6.5.9. *Si $A \subseteq \mathbb{R}$ es un **conjunto saturado no-medible**, entonces*

$$\mu_*(A \cap E) \;=\; 0 \qquad\qquad y \qquad\qquad \mu^*(A \cap E) \;=\; \mu(E)$$

*para cualquier conjunto $E \in \mathfrak{M}_\mu(\mathbb{R})$. En particular, $A \cap E$ es **no-medible** para cualquier conjunto $E \in \mathfrak{M}_\mu(\mathbb{R})$ con $\mu(E) > 0$.*

Prueba. Suponga que A es un conjunto saturado no-medible y sea $E \in \mathfrak{M}_\mu(\mathbb{R})$. Puesto que $A \cap E \subseteq A$, entonces $\mu_*(A \cap E) \leq \mu_*(A) = 0$ y, por lo tanto, $\mu_*(A \cap E) = 0$. Por otro lado, como $\mu_*(A^c) = 0$, entonces el Teorema 6.3.68 nos asegura que $\mu^*(A \cap E) = \mu(E)$.

Finalmente, sea $E \in \mathfrak{M}_\mu(\mathbb{R})$ con $\mu(E) > 0$. Si $A \cap E$ fuese medible, entonces el Teorema 6.3.65 y la primera parte no conduciría a la siguiente contradicción:

$$0 \;=\; \mu_*(A \cap E) \;=\; \mu^*(A \cap E) \;=\; \mu(E) \;>\; 0.$$

Esto termina la prueba. ■

Teorema 6.5.10. *En $\mathbb{R}$ existen **conjuntos saturados no-medibles**.*

Prueba. Un conjunto no-medible tipo Vitali se puede obtener argumentando como la en la sección 6.5.2. Fijemos un número irracional ξ y considere el subgrupo aditivo de $\mathbb{R}$,

$$D_\xi \;=\; \{m + n\xi \,:\, m, n \in \mathbb{Z}\} \;=\; \mathbb{Z} + \xi\mathbb{Z}.$$

Por el Teorema de Kronecker, Teorema 2.2.37, página 135, sabemos que D_ξ es denso y, además, numerable en $\mathbb{R}$. Para cada $x, y \in \mathbb{R}$ escribamos:

$$x \sim y \quad \text{si, y sólo si,} \quad x - y \in D_\xi.$$

Esta relación es, por supuesto, una relación de equivalencia, lo cual implica que la colección $([x])_{x\in\mathbb{R}}$ de todas las clases de equivalencias determinadas por $\sim$, constituye una partición de $\mathbb{R}$. Observe que $[x] = x + D_\xi$ para cada $x \in \mathbb{R}$. Usemos de nuevo el Axioma de Elección para construir un conjunto $\mathbb{V}$ el cual contiene exactamente un punto de cada clase $[x]$, $x \in \mathbb{R}$. Nótese que si $v_1, v_2 \in \mathbb{V}$ y ellos son distintos, resulta que tales números están en clases de equivalencias diferentes, por lo que $v_1 - v_2 \notin D_\xi$. En otras palabras, si $v_1, v_2 \in \mathbb{V}$, entonces

$$v_1 - v_2 \in \mathbb{V} \quad \text{significa que} \quad v_1 = v_2. \tag{$*$}$$

De esto se concluye que

$$(\mathbb{V} - \mathbb{V}) \cap D_\xi = \{0\}. \tag{1}$$

Para demostrar que $\mathbb{V}$ no es medible según Lebesgue, supondremos lo contrario para producir una contradicción. Suponga entonces que $\mathbb{V}$ es medible. En este caso se tiene que $\mu(\mathbb{V}) = 0$ o $\mu(\mathbb{V}) > 0$. Veamos que ambas situaciones conducen a una contradicción.

(a) Suponga que $\mu(\mathbb{V}) = 0$. Nuestra primera tarea es verificar que

$$\mathbb{R} = \bigcup_{x\in D_\xi} (x + \mathbb{V}). \tag{2}$$

En efecto, sea $z \in \mathbb{R}$. Puesto que $\mathbb{R} = \bigcup_{x\in\mathbb{R}}[x]$ y la colección $\{[x] : x \in \mathbb{R}\}$ es disjunta, entonces existe un único $x \in \mathbb{R}$ tal que $z \in [x]$ y, por lo tanto, $z - y \in D_\xi$ para cualquier $y \in [x]$. Sea v el único elemento de $\mathbb{V}$ que pertenece a $[x]$. Entonces $u = z - v \in D_\xi$, de donde se sigue que $z = u + v \in u + \mathbb{V}$ y así, (2) se cumple. Ahora bien, como D_ξ es numerable y μ es numerablemente aditiva e invariante por traslación, resulta que

$$+\infty = \mu(\mathbb{R}) = \sum_{x\in D_\xi} \mu(x + D_\xi) = \sum_{x\in D_\xi} \mu(D_\xi) = 0.$$

Esta contradicción impide que $\mu(\mathbb{V}) = 0$.

(b) Suponga que $\mu(\mathbb{V}) > 0$. En este caso podemos hacer uso de la Teorema de Steinhaus, Teorema 6.3.54, para obtener un intervalo abierto, digamos $J = (-\varepsilon, \varepsilon) \subseteq \mathbb{V} - \mathbb{V}$. Aquí viene la contradicción: como D_ξ es denso en $\mathbb{R}$, resulta que $J \cap D_\xi \neq \varnothing$ y, de hecho, tal intersección contiene infinitos puntos; sin embargo, por (1) sabemos que $(\mathbb{V} - \mathbb{V}) \cap D_\xi = \{0\}$ lo cual niega lo anterior y, por lo tanto, tampoco puede ocurrir que $\mu(\mathbb{V}) > 0$. La conclusión es simple: $\mathbb{V}$ no puede ser medible.

Nos apoyaremos en el conjunto no-medible $\mathbb{V}$ que acabamos de obtener para construir nuestro conjunto saturado no-medible. Consideremos el conjunto

$$H = \{2m + n\xi : m, n \in \mathbb{Z}\}.$$

Sabemos que H es denso en $\mathbb{R}$ (véase la Observación 2.2.2 después del Teorema de Kronecker) y, por consiguiente, $H_1 = H + 1$ también lo es. Afirmamos que el conjunto

$$\mathbb{S} = \mathbb{V} + H$$

es saturado no-medible. Para verificar esta afirmación, vamos a demostrar, en primer lugar, que $(\mathbb{S} - \mathbb{S}) \cap H_1 = \varnothing$. En efecto, sea $x \in \mathbb{S} - \mathbb{S}$ y suponga que $x \in H_1$. Puesto que $x \in \mathbb{S} - \mathbb{S}$, existen $v_1, v_2 \in \mathbb{V}$ y $h_1, h_2 \in H$ tales que

$$x = (v_1 + h_1) - (v_2 + h_2) = (v_1 - v_2) + (h_1 - h_2).$$

Además, como $H, H_1 \subseteq D_\xi$ y D_ξ es un subgrupo, se tiene que $h_1 - h_2 \in D_\xi$. Como hemos supuesto que $x \in H_1$, resulta que $x - (h_1 - h_2) \in D_\xi$ y, en consecuencia, $v_1 - v_2 = x - (h_1 - h_2) \in D_\xi$. Se sigue de $(*)$ que $v_1 = v_2$ y así, $x = h_1 - h_2$. Esto, por supuesto, origina la siguiente contradicción: $x \in H \cap H_1 = H \cap (H+1) = \varnothing$ ya que $x = h_1 - h_2 \in H$ y $x \in H_1 = H + 1$. Esta incongruencia muestra que

$$(\mathbb{S} - \mathbb{S}) \cap H_1 \;=\; \varnothing. \tag{3}$$

Observe que como H_1 es denso en $\mathbb{R}$, la igualdad (3) nos indica que $\operatorname{int}(\mathbb{S} - \mathbb{S}) = \varnothing$ y, por consiguiente, gracias a la Teorema de Steinhaus, se tiene que $\mu^*(\mathbb{S}) = 0$. Finalmente, como $\mu_*(\mathbb{S}) \leq \mu^*(\mathbb{S})$ se concluye que $\mu_*(\mathbb{S}) = 0$.

Veamos ahora que $\mu_*(\mathbb{S}^c) = 0$. En efecto, como $\mu^*(\mathbb{V}) > 0$, podemos seleccionar un conjunto medible E tal que $\mathbb{V} \subseteq E$ y $\mu(E) = \mu^*(\mathbb{V})$. Usemos el Teorema de Densidad de Lebesgue, Teorema 9.1.33, página 480, para hallar un conjunto medible $N \subseteq E$ con $\mu(N) = 0$ tal que

$$\lim_{\delta \to 0} \frac{\mu\big(E \cap (x - \delta, x + \delta)\big)}{\mu\big((x - \delta, x + \delta)\big)} \;=\; 1 \tag{4}$$

para todo $x \in E \setminus N$. El hecho de que $\mu^*(\mathbb{V}) > 0$ impide que $\mathbb{V} \subseteq N$ de modo que $\mathbb{V} \cap (E \setminus N) \neq \varnothing$. Fijemos un punto $x_0 \in \mathbb{V}$ verificando (4) y sea $0 < c < 1$. Entonces existe un $\delta > 0$ tal que

$$\mu^*\big(\mathbb{V} \cap (x - \delta/2, x + \delta/2)\big) \;>\; c \cdot \mu\big((x - \delta/2, x + \delta/2)\big) \;=\; c \cdot \delta.$$

De esto se sigue que si J es cualquier intervalo con centro en cualquier punto de $x_0 + D_\xi$ y de longitud menor que δ, entonces

$$\mu^*(\mathbb{S} \cap J) \;>\; c \cdot \mu(J).$$

Tomemos ahora cualquier intervalo abierto $I \subseteq \mathbb{R}$. Puesto que $x_0 + D_\xi$ es denso en $\mathbb{R}$, resulta que $I \cap (x_0 + D_\xi) \neq \varnothing$ y, por consiguiente, la familia $\mathcal{V}$ de todos los intervalos $J_x \subseteq I$ cuyos centros son puntos de $x_0 + D_\xi$ y poseen longitud menor que δ, constituyen un cubrimiento de Vitali. Un llamado al Teorema del Cubrimiento de Vitali, Teorema 9.1.28, página 473, nos garantiza la existencia de una sucesión disjunta de conjuntos en $\mathcal{V}$, digamos $(J_n)_{n=1}^{\infty}$, tal que

$$\mu\Big(I \setminus \bigcup_{i=1}^{\infty} J_i\Big) \;=\; 0.$$

Lo anterior permite obtener

$$\begin{aligned} \mu(I) \;\geq\; \mu^*(\mathbb{S} \cap I) \;&=\; \mu^*\Big(\mathbb{S} \cap \bigcup_{i=1}^{\infty} J_i\Big) \;+\; \mu^*\Big(\mathbb{S} \cap \Big(I \setminus \bigcup_{i=1}^{\infty} J_i\Big)\Big) \\ &=\; \mu^*\Big(\mathbb{S} \cap \bigcup_{i=1}^{n} J_i\Big) \;=\; \sum_{i=1}^{\infty} \mu^*(\mathbb{S} \cap J_i) \;>\; c \sum_{i=1}^{\infty} \mu(J_i) \;=\; c\,\mu(I). \end{aligned}$$

Haciendo que $c \to 1$ se obtiene que

$$\mu(I) \;=\; \mu^*(\mathbb{S} \cap I)$$

para cualquier intervalo abierto $I \subseteq \mathbb{R}$. Un llamado al Teorema 6.3.68 nos revela que $\mu_*(\mathbb{S}^c) = 0$ y termina la prueba. ■

6.5.4. Conjunto de Bernstein

En esta sección mostraremos la existencia de *un conjunto no-medible* $\mathbf{B}$ *tal que cualquier conjunto medible* E *incluido en* $\mathbf{B}$ *o en* $\mathbf{B}^c$*, posee medida cero.*

Recordemos, Teorema 1.3.49, página 78, que $\mathbf{B} \subseteq \mathbb{R}$ es un ***conjunto de Bernstein*** *si él, así como su complemento, intersectan a cualquier subconjunto cerrado, no-numerable de* $\mathbb{R}$. Nuestro objetivo, en esta sección, es demostrar que cualquier conjunto de Bernstein no es medible según Lebesgue.

En lo inmediato, también es necesario recordar que: *si* $\mathbf{B}$ *es un conjunto de Bernstein, entonces*

(B_1) $\mathbb{R} \setminus \mathbf{B}$ *también lo es,* y

(B_2) *si* $F \subseteq \mathbf{B}$ *es cerrado, entonces* F *es numerable.*

Teorema 6.5.11 (Bernstein). *Si* E *es un* ***subconjunto medible*** *de* $\mathbf{B}$*, entonces* $\mu(E) = 0$*. En particular,* $\mathbf{B} \notin \mathfrak{M}_\mu(\mathbb{R})$*.*

Prueba. Sea E un subconjunto medible de $\mathbf{B}$ y fijemos un conjunto compacto arbitrario $K \subseteq E$. Puesto que K es cerrado, sabemos por (B_2) que K es numerable y, en consecuencia, $\mu(K) = 0$. Se sigue del Corolario 6.3.66, página 302, que

$$\mu(E) \;=\; \sup\big\{\mu(K) : K \subseteq E,\ K \text{ compacto}\big\} \;=\; 0.$$

De modo enteramente similar se prueba que cualquier conjunto medible según Lebesgue incluido en $\mathbb{R} \setminus \mathbf{B}$ tiene medida cero. Finalmente, si $\mathbf{B}$ fuese medible, entonces también lo sería $\mathbb{R} \setminus \mathbf{B}$ y ambos conjuntos tendrían medida cero. Esto, por supuesto, generaría la siguiente contradicción:

$$+\infty \;=\; \mu(\mathbb{R}) \;=\; \mu(\mathbf{B}) \;+\; \mu(\mathbb{R} \setminus \mathbf{B}) \;=\; 0.$$

En conclusión, $\mathbf{B}$ no puede ser medible según Lebesgue. ■

Cuando estudiamos el conjunto no-medible según Vitali, vimos que todo conjunto medible con medida positiva contiene un conjunto no-medible según Lebesgue. El resultado anterior también permite obtener una respuesta inmediata a tal afirmación.

Corolario 6.5.12. *Sea* E *un subconjunto* ***medible*** *de* $\mathbb{R}$ *con* $\mu(E) > 0$*. Entonces* $E \cap \mathbf{B}$ *y* $E \cap \mathbf{B}^c$ *son* ***no-medibles*** *según Lebesgue.*

Prueba. Si ambos conjuntos fuesen medibles, entonces por el Teorema 6.5.11 tendríamos que $\mu(E \cap \mathbf{B}) = 0 = \mu(E \cap \mathbf{B}^c)$, de donde se seguiría que $\mu(E) = \mu(E \cap \mathbf{B}) + \mu(E \cap \mathbf{B}^c) = 0$. Suponga ahora que uno de ellos es medible, por ejemplo, que $E \cap \mathbf{B}$ es medible. Entonces $E \setminus (E \cap \mathbf{B}) = E \cap \mathbf{B}^c$ sería medible lo cual es imposible por la primera parte. ■

Por ejemplo, los conjuntos tipo-Cantor Γ_α que poseen medida positiva son conjuntos cerrados, no-numerables y, en consecuencia, intersectan a $\mathbf{B}$. Por lo tanto, $\Gamma_\alpha \cap \mathbf{B}$ es no-medible según Lebesgue. En el siguiente resultado se muestra la existencia de conjuntos no-medibles que también son nunca-densos.

Corolario 6.5.13. *Existe un conjunto* ***no-medible*** *y, por consiguiente,* ***no-numerable*** *que es* ***nunca-denso*** *con la propiedad adicional de que todo conjunto cerrado incluido en él es* ***numerable****.*

Prueba. Sea $\mathbf{B}$ un conjunto de Bernstein y sea Γ_α un conjunto tipo-Cantor con $\mu(\Gamma_\alpha) = \alpha > 0$. Entonces $W = \mathbf{B} \cap \Gamma_\alpha$ es, por el corolario anterior, un conjunto no-medible, nunca-denso y, por supuesto, cualquier subconjunto cerrado de él es numerable. ■

Si denotamos por $\mathcal{V}_\mu(\mathbb{R})$, (respectivamente, $\mathcal{B}_\mu(\mathbb{R})$) *la subfamilia de* $\neg\mathfrak{M}_\mu(\mathbb{R})$ *conteniendo a los* ***conjuntos de Vitali*** (respectivamente, *los* ***conjuntos de Bernstein***) se tiene, véase [77], Theorem 3, p. 22, que

$$\mathcal{V}_\mu(\mathbb{R}) \cap \mathcal{B}_\mu(\mathbb{R}) \neq \varnothing.$$

6.5.5. Conjunto de Sierpiński

Si en lugar de considerar *conjuntos de primera categoría* en la construcción del conjunto Lusin, lo reemplazamos por *conjuntos de medida de Lebesgue cero* se obtiene, aceptando la Hipótesis del Continuo, un nuevo tipo de conjunto al que se le denominará conjunto de Sierpiński. Como antes, $\mathfrak{N}_\mu(\mathbb{R})$ denotará la colección de todos los conjuntos nulos.

Teorema 6.5.14 (Sierpiński). *Bajo la* ***Hipótesis del Continuo****, existe un conjunto* $\mathfrak{S} \subseteq \mathbb{R}$*, llamado* ***conjunto de Sierpiński****, tal que:*

(a) $\mathfrak{S}$ *es* ***no-numerable****, y*

(b) $\operatorname{card}(\mathfrak{S} \cap N) \leq \aleph_0$ *para cualquier conjunto* $N \in \mathfrak{N}_\mu(\mathbb{R})$.

Prueba. Denote por $\mathfrak{N}_\delta(\mathbb{R})$ *la colección de todos aquellos elementos de* $\mathfrak{N}_\mu(\mathbb{R})$ *que son de tipo* G_δ. Puesto que $\{x\}$ es un G_δ para cada $x \in \mathbb{R}$, la Hipótesis del Continuo nos revela que

$$\aleph_1 = 2^{\aleph_0} = \operatorname{card}(\{\{x\} : x \in \mathbb{R}\}) \leq \operatorname{card}(\mathfrak{N}_\delta(\mathbb{R})) \leq \operatorname{card}(\mathcal{G}_\delta) = 2^{\aleph_0},$$

y así, $\mathfrak{N}_\delta(\mathbb{R})$ puede ser escrito en términos de una sucesión transfinita, es decir, en la forma $\mathfrak{N}_\delta(\mathbb{R}) = \{N_\alpha : \alpha < \omega_1\}$. Sea $x_1 \in \mathbb{R}$ tal que $x_1 \notin N_1 \in \mathfrak{N}_\delta(\mathbb{R})$. Fijemos un $\xi < \omega_1$ y supongamos que la familia $(x_\alpha)_{\alpha<\xi}$ ha sido construida. Consideremos el conjunto

$$X_\xi = \Big(\bigcup_{\alpha<\xi} N_\alpha\Big) \cup \{x_\alpha : \alpha < \xi\}$$

y observemos que como ξ es numerable y $\mathfrak{M}_\mu(\mathbb{R})$ es una σ-álgebra, X_ξ es un conjunto nulo, por lo que $\mathbb{R} \setminus X_\xi$ es no vacío. Si ahora elegimos un $x_\xi \in \mathbb{R} \setminus X_\xi$, entonces el Principio de Inducción Transfinita da por finalizada la construcción de $(x_\xi)_{\xi<\omega_1}$. Definamos ahora

$$\mathfrak{S} = \{x_\xi : \xi < \omega_1\}.$$

Afirmamos que $\mathfrak{S}$ posee las propiedades requeridas. En efecto:

(*a*) $\mathfrak{S}$ es no numerable. Esto sigue del hecho de que la aplicación $f : \omega_1 \to \mathfrak{S}$ que asigna a cada $\alpha \in \omega_1$ el elemento $f(\alpha) = x_\alpha$, es biyectiva.

(*b*) Sea N un conjunto nulo. Por el Corolario 6.2.13, página 249, existe un conjunto G_δ, llamémoslo N_0, tal que $N \subseteq N_0$ y $\mu(N) = \mu(N_0)$, es decir, $N_0 \in \mathfrak{N}_\delta(\mathbb{R})$. Puesto que $\mathfrak{N}_\delta(\mathbb{R})$ consiste de la familia de todos los subconjuntos nulos que son G_δ, existe un $\alpha < \omega_1$ tal que $N_0 = N_\alpha \in \mathfrak{N}_\delta(\mathbb{R})$. Por construcción, sabemos que

$$\mathfrak{S} \cap N \subseteq \mathfrak{S} \cap N_\alpha \subseteq \{x_\beta : \beta < \alpha\}$$

y como éste último conjunto es numerable, tenemos que $\mathfrak{S} \cap N$ es a lo más numerable.

Esto termina la prueba. ■

El hecho de que $\mathfrak{S}$ no contiene subconjuntos medibles no-numerables de medida cero conduce a que dicho conjunto es no-medible.

Teorema 6.5.15. *$\mathfrak{S}$ **no es medible** según Lebesgue.*

Prueba. Es suficiente demostrar que $\mathfrak{S}$ **no contiene subconjuntos medibles no-numerables**. En efecto, suponga, para generar una contradicción, que E es un subconjunto medible no-numerable incluido en $\mathfrak{S}$. Puesto que $\mathfrak{S} \cap E = E$ es no-numerable, entonces la definición de $\mathfrak{S}$ nos indica que $\mu(E) > 0$. Ahora bien, como μ es regular, Corolario 6.3.66,

$$\mu(E) = \sup\{\mu(K) : K \subseteq E,\ K \text{ compacto}\}$$

y, por lo tanto, existe un conjunto compacto $K \subseteq E$ tal que $\mu(K) > 0$. Por supuesto, como $K \subseteq \mathfrak{S}$ y $\mu(K) > 0$, dicho conjunto es no-numerable y, además, es un G_δ pues él es cerrado. Gracias al Teorema 6.4.7 tenemos que K contiene un subconjunto no-numerable de medida cero, lo cual contradice la definición de $\mathfrak{S}$. Esto termina la prueba de nuestra afirmación. Finalmente, si $\mathfrak{S}$ fuese medible según Lebesgue, entonces la inclusión $\mathfrak{S} \subseteq \mathfrak{S}$ nos indica la existencia un subconjunto medible no-numerable incluido en $\mathfrak{S}$ lo cual es imposible por la primera parte. ■

Nota Adicional 6.5.8 Observe que *$\mathfrak{S}$ es de primera categoría*. En efecto, escriba a $\mathbb{R}$ como $\mathbb{R} = A \cup N$ donde A es de primera categoría y N es un conjunto de medida cero. Entonces $\mathfrak{S} = \mathfrak{S} \cap \mathbb{R} = (\mathfrak{S} \cap A) \cup (\mathfrak{S} \cap N)$. Puesto que $\mathfrak{S}$ es un conjunto de Sierpiński y $\mu(N) = 0$, resulta que $\mathfrak{S} \cap N$ es a lo más numerable, de modo que todos los elementos de N, salvo una cantidad a lo más numerable, están contenidos en A el cual es de primera categoría y, por lo tanto, $\mathfrak{S}$ es de primera categoría.

Es claro que *bajo la **Hipótesis del Continuo**, cualquier conjunto N de cardinalidad menor que $\mathfrak{c}$ es numerable y, en consecuencia, de medida cero* y entonces, similar al caso de la existencia de conjuntos de Lusin, se tiene que:

Teorema 6.5.16. *La **Hipótesis del Continuo** es equivalente a la existencia de un conjunto de Sierpiński y a que cualquier conjunto de cardinalidad menor que $\mathfrak{c}$ es de medida cero.*

Otra manera de obtener conjuntos no-medibles según Lebesgue sin usar la Teoría **ZFC** + **CH** es la siguiente:

Teorema 6.5.17 (Sierpiński). *Sea X un subconjunto no-numerable de $\mathbb{R}$. Las siguientes dos condiciones son equivalentes dentro de la Teoría* **ZF** + **DC**.

(1) *X es un **conjunto de Sierpiński**.*

(2) *Cada subconjunto **no-numerable** de X es **no-medible** según Lebesgue.*

En general, si aceptamos la Hipótesis del Continuo en el sistema **ZF** + **DC**, entonces existen conjuntos no-medibles según Lebesgue. La demostración de éstos hechos se pueden ver, por ejemplo, en [75].

Como antes, denote por $\mathfrak{N}_\mu(\mathbb{R})$ *la familia de todos los conjuntos nulos* y por $\mathcal{A}_{\rm pc}(\mathbb{R})$ *todos los subconjuntos de $\mathbb{R}$ que son de primera categoría*. Una relación fenomenal y, por supuesto, fundamental entre los conjuntos de primera categoría y los conjuntos nulos fue hallado por P. Erdös y W. Sierpiński cuando demostraron que:

Teorema 6.5.18 (Erdös-Sierpiński). *Asumiendo la **Hipótesis del Continuo**, existe una **función biyectiva** $f:\mathbb{R}\to\mathbb{R}$ con la siguiente propiedad: para todo $A\subseteq\mathbb{R}$,*

(a) $f(A)\in\mathfrak{N}_\mu(\mathbb{R})$ si, y sólo si $A\in\mathcal{A}_{\rm pc}(\mathbb{R})$.

(b) $f(A)\in\mathcal{A}_{\rm pc}(\mathbb{R})$ si, y sólo si $A\in\mathfrak{N}_\mu(\mathbb{R})$.

6.5.6. Ultrafiltros no-medibles

Usando la noción de ultrafiltro sobre un conjunto X, introducida por H. Cartan, es posible construir otros tipos de conjuntos no-medibles en $\mathbb{R}$. La existencia de tales objetos, los ultrafiltros, se sustentan sobre el Lema de Zorn el cual, como sabemos, es equivalente al Axioma de Elección. La noción de filtro, al igual que la noción de red, son generalizaciones del concepto de sucesión. Como sabemos, las sucesiones son herramientas fundamentales para caracterizar, por ejemplo, en espacios métricos, compacidad, funciones continuas y muchas otras propiedades. Cuando dicho espacio no satisface el primer axioma de numerabilidad, hay que apelar a la noción de redes, o en su defecto, la de filtro para caracterizar tales objetos.

Definición 6.5.19 (Cartan). *Sea X un conjunto no vacío. Un **filtro** sobre X es una colección no vacía $\mathcal{F}$ de subconjuntos de X que cumple con las siguientes propiedades:*

(a) $\varnothing\notin\mathcal{F}$.

(b) Si $A, B\in\mathcal{F}$, entonces $A\cap B\in\mathcal{F}$.

(c) Si $A\in\mathcal{F}$ y $A\subseteq B\subseteq X$, entonces $B\in\mathcal{F}$.

Observe que si $\mathcal{F}$ es un filtro sobre X, entonces la condición (c) garantiza que $X\in\mathcal{F}$. Además, si $A\subseteq X$, entonces A y $X\setminus A$ no pueden, simultáneamente, pertenecer a $\mathcal{F}$, pues ello forzaría a que $\varnothing = A\cap(X\setminus A)$ pertenezca a $\mathcal{F}$. Más aun, de (b) se sigue que si $A_1,\dots,A_n$ están en $\mathcal{F}$, entonces $\bigcap_{i=1}^n A_i\in\mathcal{F}$ para cualquier $n\geq 2$.

Tres ejemplos importantes de filtros son los siguientes:

(1) Sea X un conjunto no vacío y fijemos un punto $x\in X$. Si

$$\mathcal{F}_x = \{A\subseteq X : x\in A\},$$

entonces $\mathcal{F}_x$ es un filtro sobre X, llamado el **filtro principal** sobre X en x. En general, si $A\subseteq X$ es un subconjunto no vacío, entonces

$$\mathcal{F}_A = \{F\subseteq X : A\subseteq F\},$$

es un filtro sobre X.

(2) Sea X un conjunto infinito y sea

$$\mathcal{F}_{\sf co}(X) = \{A\subseteq X : X\setminus A \text{ es finito}\}.$$

Entonces $\mathcal{F}_{\sf co}(X)$ es un filtro sobre X, al que llamaremos ***filtro co-finito***. Cuando $X=\mathbb{N}$, al filtro $\mathcal{F}_{\sf co}(\mathbb{N})$ se le llama **filtro de Fréchet**. Observe que *ningún subconjunto finito de $\mathbb{N}$ puede pertenecer a $\mathcal{F}_{\sf co}(\mathbb{N})$.*

(3) Nótese que, gracias a la propiedad (b) de la definición anterior, todo filtro $\mathcal{F}$ sobre X posee la **Propiedad de Intersección Finita**, es decir, si $A_1, \ldots, A_n$ son elementos arbitrarios de $\mathcal{F}$, entonces $\bigcap_{i=1}^{n} A_i \neq \varnothing$. Además, si $\mathcal{G} \subseteq \mathscr{P}(X)$ posee la Propiedad de Intersección Finita, entonces la colección

$$\mathcal{F}_{\mathcal{G}} = \Big\{ A \subseteq X : \bigcap_{i=1}^{n} A_i \subseteq A \text{ para alguna colección finita } \{A_1, \ldots, A_n\} \subseteq \mathcal{G} \Big\}$$

es un filtro conteniendo a $\mathcal{G}$. A $\mathcal{F}_{\mathcal{G}}$ se le llama el ***filtro generado*** por $\mathcal{G}$.

Definición 6.5.20. *Sea $\mathcal{F}$ un filtro sobre X. Diremos que $\mathcal{F}$ es un **ultrafiltro** si $\mathcal{F}$ no está contenido propiamente en ningún otro filtro sobre X; en otras palabras, si para cualquier filtro $\mathcal{G}$ sobre X se tiene que*

$$\mathcal{F} \subseteq \mathcal{G} \Rightarrow \mathcal{G} = \mathcal{F}.$$

Los ultrafiltros son, por lo general, objetos muy extraños salvo los que son de la forma $\mathcal{F}_x$. Aunque existen ultrafiltros distintos de los anteriores, ellos son imposible de describir explícitamente. Además, si $\operatorname{card}(X) \geq \aleph_0$, entonces cada ultrafiltro sobre X posee la cardinalidad del continuo.

Nótese que: *Cualquier **filtro principal** es un **ultrafiltro***. En general, si X es un conjunto finito, cualquier ultrafiltro sobre X es principal. Si X es infinito, entonces existen ultrafiltros sobre X que no son principales. Ultrafiltros que no son principales se llaman **libres** o **no-principales**.

Teorema 6.5.21. *Sea X un conjunto no vacío. $\mathcal{F}$ es un **ultrafiltro** sobre X si, y sólo si, para cualquier conjunto $A \subseteq X$, se cumple que $A \in \mathcal{F}$ o bien $X \setminus A \in \mathcal{F}$.*

Prueba. . Suponga que $\mathcal{F}$ es un ultrafiltro y sea $A \subseteq X$. Si ocurre que $A \notin \mathcal{F}$ y $X \setminus A \notin \mathcal{F}$, entonces la colección $\mathcal{G} = \{B \subseteq X : A \cup B \in \mathcal{F}\}$ es un filtro conteniendo propiamente a $\mathcal{F}$ ya que $A \in \mathcal{G} \setminus \mathcal{F}$ lo que, por supuesto, contradice el hecho de que $\mathcal{F}$ es un ultrafiltro.

Para demostrar el recíproco, suponga que para cualquier conjunto $A \subseteq X$, ocurre que $A \in \mathcal{F}$ o bien $X \setminus A \in \mathcal{F}$, pero que $\mathcal{F}$ no es un ultrafiltro. Entonces existe un filtro $\mathcal{G}$ conteniendo propiamente $\mathcal{F}$, esto es, $\mathcal{F} \subsetneq \mathcal{G}$. Sea $A_0 \in \mathcal{G} \setminus \mathcal{F}$. Puesto que $\mathcal{G}$ es un filtro, no puede ocurrir que $X \setminus A_0 \in \mathcal{G}$ y, en consecuencia, como $\mathcal{F} \subseteq \mathcal{G}$, también es imposible que $X \setminus A_0 \in \mathcal{F}$. Hemos entonces demostrado que si $\mathcal{F}$ no es un ultrafiltro, existe al menos un conjunto $A_0 \subseteq X$ tal que ni A_0 ni $X \setminus A_0$ pertenecen a $\mathcal{F}$, lo que constituye una contradicción a nuestra hipótesis. Por esto, $\mathcal{F}$ es un ultrafiltro. ■

Otra forma de pensar a los ultrafiltros es por medio de medidas finitamente aditivas que toman sólo dos valores definidas sobre la σ-álgebra $\mathscr{P}(X)$. Sea X un conjunto no vacío y sea $\lambda : \mathscr{P}(X) \to \{0,1\}$ una *medida finitamente aditiva* tal que $\lambda(X) = 1$. Considere la colección

$$\mathbb{U} = \{U \subseteq X : \lambda(U) = 1\}.$$

Veamos que $\mathbb{U}$ es un ultrafiltro.

(a) Claramente $X \in \mathbb{U}$ pero $\varnothing \notin \mathbb{U}$.

(b) También es claro que si $U \in \mathbb{U}$ y $U \subseteq V \subseteq X$, entonces $V \in \mathbb{U}$.

(c) Sean $U, V \in \mathbb{U}$ y suponga que $U \cap V \notin \mathbb{U}$. Entonces

$$1 \,=\, \lambda(U) \,=\, \lambda(U \setminus V) + \lambda(U \cap V) \,=\, \lambda(U \setminus V)$$

y, similarmente, $\lambda(V \setminus U) = 1$. De esto se sigue que

$$\lambda(U \cup V) \,=\, \lambda(U \setminus V) + \lambda(U \cap V) + \lambda(V \setminus U) \,=\, 2$$

lo cual es absurdo ya que $U \cup V \in \mathbb{U}$ por la parte (b). Todo lo anterior confirma que $\mathbb{U}$ es un filtro. Finalmente, para cualquier conjunto $U \subseteq X$ se cumple que $U \in \mathbb{U}$ o $X \setminus U \in \mathbb{U}$. En efecto, si, por ejemplo, $U \notin \mathbb{U}$, entonces de la relación

$$1 \,= \lambda(X) \,=\, \lambda(U) + \lambda(X \setminus U) \,=\, \lambda(X \setminus U)$$

se sigue que $X \setminus U \in \mathbb{U}$ y, en consecuencia, $\mathbb{U}$ es un ultrafiltro. Además, $\mathbb{U}$ es no-principal.

Recíprocamente, si $\mathbb{U}$ es un ultrafiltro no-principal sobre X, entonces considerando la función de conjuntos $\lambda : \mathscr{P}(X) \to \{0,1\}$ definida por

$$\lambda(A) \,=\, \begin{cases} 1 & \text{si } A \in \mathbb{U} \\ 0 & \text{si } A \notin \mathbb{U} \end{cases}$$

se tiene que λ es una medida finitamente aditiva tomando sólo dos valores y satisfaciendo $\lambda(X) = 1$. En efecto, claramente $\lambda(X) = 1$ ya que $X \in \mathbb{U}$. Sean $A_1, A_2 \in X$ con $A_1 \cap A_2 = \varnothing$. Nótese que si $A_1 \cup A_2 \in \mathbb{U}$, entonces $A_1 \in \mathbb{U}$ o $A_2 \in \mathbb{U}$ pero ambos no pueden estar en $\mathbb{U}$ ya que $A_1 \cap A_2 = \varnothing$. En consecuencia,

$$1 \,=\, \lambda(A_1 \cup A_2) \,=\, \lambda(A_1) + \lambda(A_2).$$

Por otro lado, si $A_1 \cup A_2 \notin \mathbb{U}$, entonces $A_1 \notin \mathbb{U}$ y $A_2 \notin \mathbb{U}$ y, así,

$$0 \,=\, \lambda(A_1 \cup A_2) \,=\, \lambda(A_1) + \lambda(A_2).$$

Lo que acabamos de demostrar se puede enunciar del modo siguiente:

Teorema 6.5.22. *Existe una **biyección** entre el conjunto de **todos los ultrafiltros no-principales** sobre X y el conjunto de **todas las medidas finitamente aditivas** $\lambda : \mathscr{P}(X) \to \{0,1\}$.*

La siguiente caracterización es una simple, pero elegante, forma de interpretar a los ultrafiltros libres.

Lema 6.5.23 (Ultrafiltros libres). *Sea $\mathbb{U}$ un **ultrafiltro** sobre X. Las siguientes afirmaciones son equivalentes:*

(1) *$\mathbb{U}$ es **libre**.*

(2) $\bigcap_{A \in \mathbb{U}} A \,=\, \varnothing$.

(3) *Para cualquier $A \in \mathbb{U}$, $\operatorname{card}(A) \geq \aleph_0$.*

Prueba. (1) $\Rightarrow$ (2). Suponga que $\bigcap_{A\in\mathbb{U}} A \neq \varnothing$ y sea $x \in \bigcap_{A\in\mathbb{U}} A$. Entonces $x \in A$ para todo $A \in \mathbb{U}$, por lo que $\mathbb{U} \subseteq \mathcal{F}_x$. Puesto que $\mathbb{U}$ es un filtro maximal, resulta que $\mathbb{U} = \mathcal{F}_x$, lo cual es imposible ya que $\mathbb{U}$ es libre.

(2) $\Rightarrow$ (3). Sea $A \in \mathbb{U}$ y suponga que $\mathrm{card}(A) = n \in \mathbb{N}$. Usando el Teorema 6.5.21 se obtiene la existencia de un $a \in A$ tal que $\{a\} \in \mathbb{U}$. Se sigue de esto que $\mathbb{U} = \mathcal{F}_a$ lo cual es imposible.

(3) $\Rightarrow$ (1). Es inmediata. ■

La existencia de ultrafiltros libres fue probada por primera vez por A. Tarski y su demostración se sustenta sobre el Axioma de Elección (equivalente al Lema de Zorn). Es importante destacar que no se puede probar la existencia de ultrafiltros libres sin el Axioma de Elección.

Teorema 6.5.24 (Teorema del Ultrafiltro). *Sean X un conjunto no vacío y $\mathcal{F}_0$ un **filtro** sobre X. Entonces existe un **ultrafiltro** $\mathbb{U}$ sobre X tal que $\mathcal{F}_0 \subseteq \mathbb{U}$.*

Prueba. Considere la familia

$$\mathfrak{F} = \{\mathcal{F} \subseteq \mathscr{P}(X) : \mathcal{F} \text{ es un filtro sobre } X \text{ con } \mathcal{F}_0 \subseteq \mathcal{F}\}.$$

Ordene a $\mathfrak{F}$ parcialmente con la relación de inclusión $\subseteq$. Sea $\mathcal{C}$ una cadena en $\mathfrak{F}$. Tomando $\mathcal{G} = \bigcup_{A\in\mathcal{C}} \mathcal{A}$, entonces es fácil ver que $\mathcal{G}$ es un filtro sobre X y, por lo tanto, $\mathcal{G} \in \mathfrak{F}$. Además, $\mathcal{G}$ es una cota superior para $\mathcal{C}$. Se sigue del Lema de Zorn que $\mathfrak{F}$ posee un elemento maximal, digamos $\mathbb{U}$. Claramente $\mathbb{U}$ es un ultrafiltro conteniendo a $\mathcal{F}_0$. ■

‖ ▶ Ultrafiltros Libres sobre $\mathbb{N}$.

Hagamos $X = \mathbb{N}$. Un resultado de B. Pospíšil estable que existen tantos ultrafiltros sobre $\mathbb{N}$ como subconjuntos de $\mathbb{N}$, es decir, si

$$\beta\mathbb{N} = \{\mathcal{U} \subseteq \mathscr{P}(\mathbb{N}) : \mathcal{U} \text{ es un ultrafiltro sobre } \mathbb{N}\},$$

entonces $\mathrm{card}(\beta\mathbb{N}) = 2^{\mathfrak{c}}$. Considere ahora el filtro de Fréchet $\mathcal{F}_{\mathrm{co}}(\mathbb{N})$. Éste filtro no es un ultrafiltro ya que, por ejemplo, $\mathbb{N} = \mathrm{P} \cup \mathrm{I}$ (los pares y los impares) y tanto P, así como I, no están en $\mathcal{F}_{\mathrm{co}}(\mathbb{N})$. Sin embargo, el Teorema 6.5.24 nos garantiza la existencia de un ultrafiltro $\mathbb{U}_{\mathrm{co}} \in \beta\mathbb{N}$ tal que $\mathcal{F}_{\mathrm{co}}(\mathbb{N}) \subseteq \mathbb{U}_{\mathrm{co}}$. Observe que, en este caso, y gracias al Lema 6.5.23 (3), se tiene que $\mathbb{U}_{\mathrm{co}}$ *es **libre***. En efecto, suponga que existe un conjunto $A \in \mathbb{U}_{\mathrm{co}}$ tal que $\mathrm{card}(A) < \infty$. Puesto que $\mathbb{U}_{\mathrm{co}}$ es un ultrafiltro sobre $\mathbb{N}$, resulta del Teorema 6.5.21 que $\mathbb{N} \setminus A \notin \mathbb{U}_{\mathrm{co}}$. Sin embrago, como $\mathbb{N} \setminus (\mathbb{N} \setminus A) = A$ y $\mathrm{card}(A) < \infty$, se tiene que $\mathbb{N} \setminus A \in \mathcal{F}_{\mathrm{co}}(\mathbb{N}) \subseteq \mathbb{U}_{\mathrm{co}}$. Esta contradicción establece que $\mathrm{card}(A) \geq \aleph_0$ y, entonces, por el Lema 6.5.23 (3), $\mathbb{U}_{\mathrm{co}}$ es un ultrafiltro libre.

Recíprocamente, cualquier ultrafiltro libre $\mathbb{U}$ sobre $\mathbb{N}$ contiene a $\mathcal{F}_{\mathrm{co}}(\mathbb{N})$. En efecto, suponga que algún conjunto $A \in \mathcal{F}_{\mathrm{co}}(\mathbb{N})$ no está en $\mathbb{U}$. Entonces, A es infinito por pertenecer a $\mathcal{F}_{\mathrm{co}}(\mathbb{N})$, pero a su vez es finito por estar fuera de $\mathbb{U}$. Esta contradicción establece que $\mathcal{F}_{\mathrm{co}}(\mathbb{N}) \subseteq \mathbb{U}$.

Lo anterior se puede resumir en el siguiente:

Teorema 6.5.25. *Sea $\mathbb{U} \in \beta\mathbb{N}$. Son equivalentes:*

(1) $\mathbb{U}$ *es **libre**.*

(2) $\mathcal{F}_{\mathrm{co}}(\mathbb{N}) \subseteq \mathbb{U}$.

Recordemos que un racional diádico en $(0,1)$ es cualquier número de la forma $m/2^n$ con $m \in \{1,\ldots,2^n-1\}$ y que cada uno de tales números posee exactamente dos representaciones binarias: las que terminan en cero y las que todos sus términos son, a partir de un cierto lugar, 1. Por ejemplo,

$$\begin{aligned}\frac{1}{2} &= (0{,}100\ldots)_2 = \sum_{n=1}^{\infty}\frac{a_n}{2^n}, \quad \text{donde } a_1 = 1 \text{ y } a_n = 0 \text{ para todo } n \geq 2 \\ &= (0{,}011\ldots)_2 = \sum_{n=1}^{\infty}\frac{a_n}{2^n}, \quad \text{donde, } a_1 = 0 \text{ y } a_n = 1 \text{ para todo } n \geq 2.\end{aligned}$$

Por otro lado, si $x \in (0,1)$ no es un racional diádico, entonces él posee una única representación binaria infinita. En lo que sigue, siempre asumiremos que cualquier racional diádico en $(0,1)$ se representa en su forma binaria con infinitos 1, de modo que cualquier $x \in (0,1)$ admite una, y sólo una, representación binaria infinita.

Fijemos un *ultrafiltro libre* $\mathbb{U} \in \beta\mathbb{N}$. Por el Lema 6.5.23 sabemos que cada conjunto $A \in \mathbb{U}$ es infinito. Asociemos, a cada conjunto $A \in \mathbb{U}$, el número

$$x_A = \sum_{n\in A}\frac{1}{2^n} \in [0,1]$$

y considere, finalmente, el conjunto

$$\Lambda_{\mathbb{U}} = \left\{\sum_{n\in A}\frac{1}{2^n} : A \in \mathbb{U}\right\}.$$

Nótese que la inclusión $\mathcal{F}_{\mathsf{co}}(\mathbb{N}) \subseteq \mathbb{U}$ garantiza que cualquier racional diádico expresado en su forma binaria con infinitos 1 pertenece a $\Lambda_{\mathbb{U}}$. En efecto, si $x \in (0,1)$ es diádico, entonces el conjunto

$$A_x = \{n \in \mathbb{N} : a_n = 1\}$$

es, por nuestra suposición, infinito y, por lo tanto, $\mathbb{N} \setminus A_x$ es finito, lo cual prueba que $A_x \in \mathcal{F}_{\mathsf{co}}(\mathbb{N})$. Existen dos aspectos importantes que se deben considerar respecto al conjunto $\Lambda_{\mathbb{U}}$.

(α_1) $x \in \Lambda_{\mathbb{U}} \quad \Leftrightarrow \quad 1 - x \in [0,1] \setminus \Lambda_{\mathbb{U}}$. En efecto, puesto que $\sum_{n=1}^{\infty}\frac{1}{2^n} = 1$, entonces

$$x \in \Lambda_{\mathbb{U}} \quad \Leftrightarrow \quad 1 - x = 1 - \sum_{n\in A}\frac{1}{2^n} = \sum_{n\in\mathbb{N}\setminus A}\frac{1}{2^n} \notin \Lambda_{\mathbb{U}}$$

ya que $\mathbb{U}$ es un ultrafiltro ($A \in \mathbb{U} \Leftrightarrow \mathbb{N} \setminus A \notin \mathbb{U}$).

(α_2) $\Lambda_{\mathbb{U}}$ *es un* ***conjunto absorbente***; en otras palabras, si $x = (x_n)_{n=1}^{\infty} \in \Lambda_{\mathbb{U}}$ y si $y = (y_n)_{n=1}^{\infty}$ es tal que el conjunto $\{n \in \mathbb{N} : x_n \neq y_n\}$ es finito, entonces $y \in \Lambda_{\mathbb{U}}$. Para demostrar esta afirmación, basta verificar que

$$A \in \mathbb{U} \qquad \Leftrightarrow \qquad A' \in \mathbb{U}$$

para cualesquiera conjuntos A, A' que difieran por un conjunto finito. En efecto, como $\mathbb{U}$ es un filtro, entonces la condición (c) de su definición nos muestra que:

$$A \in \mathbb{U} \quad \Rightarrow \quad A \cup F \in \mathbb{U} \quad \text{para cualquier conjunto finito } F \subseteq \mathbb{N}$$

También,

$$A \in \mathbb{U} \quad \Rightarrow \quad A \setminus F \in \mathbb{U} \quad \text{para cualquier conjunto finito } F \subseteq \mathbb{N}$$

ya que $A \setminus F = A \cap (\mathbb{N} \setminus F) \in \mathbb{U}$ pues $\mathbb{N} \setminus F \in \mathcal{F}_{\mathsf{co}}(\mathbb{N}) \subseteq \mathbb{U}$. Por esto, $\Lambda_{\mathbb{U}}$ es absorbente.

Teorema 6.5.26 (Sierpiński). *El conjunto $\Lambda_{\mathbb{U}}$ **no es medible según Lebesgue**.*

Prueba. Vamos a suponer, para generar una contradicción, que $\Lambda_{\mathbb{U}}$ es medible. Nótese que por (α_1) se tiene que $[0,1] \setminus \Lambda_{\mathbb{U}}$ es la reflexión de $\Lambda_{\mathbb{U}}$ en $x = 1/2$, lo cual significa que ambos conjuntos tienen la misma medida, de donde se tiene que $1 = \mu(\Lambda_{\mathbb{U}}) + \mu([0,1] \setminus \Lambda_{\mathbb{U}}) = 2\mu(\Lambda_{\mathbb{U}})$; es decir,

$$\mu(\Lambda_{\mathbb{U}}) = \frac{1}{2}. \tag{1}$$

Fijemos una partición arbitraria de $[0,1]$ formada por 2^m subintervalos, digamos $I_1, \ldots, I_{2^m}$. Asumiremos que $\mu(I_k) = 2^{-m}$ para todo $k \in \{1, 2, \ldots, 2^m\}$. Afirmamos que

$$\mu(\Lambda_{\mathbb{U}} \cap I_1) = \cdots = \mu(\Lambda_{\mathbb{U}} \cap I_{2^m}). \tag{2}$$

De hecho, lo que vamos a demostrar es que cada $\Lambda_{\mathbb{U}} \cap I_k$ es un trasladado de $\Lambda_{\mathbb{U}} \cap I_1$. De modo más preciso, veamos que

$$x \in \Lambda_{\mathbb{U}} \quad \Leftrightarrow \quad x \pm 2^{-m} \in \Lambda_{\mathbb{U}}.$$

Esto, sin embargo, es consecuencia de (α_2). Se sigue entonces de (1) y (2) que para cada $k \in \{1, 2, \ldots, 2^m\}$,

$$\begin{aligned}
\frac{1}{2} &= \mu(\Lambda_{\mathbb{U}}) = \mu\big(\Lambda_{\mathbb{U}} \cap [0,1]\big) \\
&= \mu(\Lambda_{\mathbb{U}} \cap I_1) + \cdots + \mu(\Lambda_{\mathbb{U}} \cap I_{2^m}) \\
&= 2^m \mu(\Lambda_{\mathbb{U}} \cap I_k),
\end{aligned}$$

es decir,

$$\mu(\Lambda_{\mathbb{U}} \cap I_k) = \frac{1}{2}\frac{1}{2^m} = \frac{1}{2}\mu(I_1) \quad \text{para } k = 1, \ldots, 2^m. \tag{3}$$

Nótese que (3) dice que sólamente la mitad de la medida de los intervalos I_k pertenecen a $\Lambda_{\mathbb{U}}$.

Sea $\varepsilon > 0$ y usemos el Primer Principio de Littlewood, Teorema 6.3.62, página 298, para seleccionar una colección finita $\{J_1, \ldots, J_n\}$ de intervalos abiertos y disjuntos tal que

$$\mu\Big(\Lambda_{\mathbb{U}} \triangle \bigcup_{i=1}^{n} J_i\Big) < \frac{\varepsilon}{2}.$$

Escoja ahora un m lo suficientemente grande de modo que la unión $\bigcup_{i=1}^{n} J_i$ pueda ser aproximada a menos de $\varepsilon/2$ por una subcolección $\{I_{n_1}, \ldots, I_{n_k}\}$ de la partición $\{I_1, \ldots, I_{2^m}\}$ de $[0,1]$; es decir,

$$\mu\Big(\bigcup_{i=1}^{n} J_i \setminus \bigcup_{j=1}^{k} I_{i_j}\Big) < \frac{\varepsilon}{2}.$$

Observe que

$$\mu\Big(\Lambda_{\mathbb{U}} \setminus \bigcup_{j=1}^{k} I_{i_j}\Big) \leq \mu\Big(\Lambda_{\mathbb{U}} \setminus \bigcup_{i=1}^{n} J_i\Big) + \mu\Big(\bigcup_{i=1}^{n} J_i \setminus \bigcup_{j=1}^{k} I_{i_j}\Big) < \varepsilon$$

y entonces

$$\begin{aligned}\mu(\Lambda_{\mathbb{U}}) &= \mu\Big(\Lambda_{\mathbb{U}} \setminus \bigcup_{j=1}^{k} I_{i_j}\Big) + \mu\Big(\Lambda_{\mathbb{U}} \cap \bigcup_{j=1}^{k} I_{i_j}\Big)\\ &< \varepsilon + \sum_{j=1}^{k} \mu(\Lambda_{\mathbb{U}} \cap I_{i_j}) = \varepsilon + \frac{k}{2}\mu(I_1)\\ &= \varepsilon + \frac{k}{2^{m+1}}.\end{aligned}$$

Si tomamos $\varepsilon < \frac{1}{2}(1 - \frac{k}{2^m})$ se obtiene que $\mu(\Lambda_{\mathbb{U}}) < \frac{1}{2}$. Esta contradicción establece que $\Lambda_{\mathbb{U}}$ no puede ser medible Lebesgue y termina la prueba. ■

6.5.7. Conjunto de Lévy

Casi al mismo tiempo en que Vitali construyó su ejemplo de un conjunto no-medible, Paul Lévy (1886-1971) [94] obtuvo un ejemplo similar. La construcción del conjunto no-medible de Lévy se fabrica sobre una base de Hamel en $\mathbb{R}_{\mathbb{Q}}$, donde $\mathbb{R}_{\mathbb{Q}}$ es el espacio vectorial $\mathbb{R}$ sobre $\mathbb{Q}$. Estas bases de Hamel han sido objeto de un estudio profundo derivándose de ellas algunos resultados sorprendentes como se muestra un poco más abajo. Quien por primera vez consideró a $\mathbb{R}$ como un espacio vectorial sobre $\mathbb{Q}$ fue Georg Karl Wilhelm Hamel (1877-1954) quien construyó una tal base usando métodos transfinitos. A dicha base se le llamará posteriormente *base de Hamel*. El objetivo de Hamel era construir una solución no trivial (o de modo equivalente, discontinua) de la ecuación funcional de Cauchy

$$f(x+y) = f(x) + f(y) \qquad \text{para todo} \quad x, y \in \mathbb{R}.$$

Hamel pudo demostrar que existían funciones $f : \mathbb{R} \to \mathbb{R}$ que satisfacían dicha ecuación pero que eran discontinuas en todo punto de $\mathbb{R}$. Una prueba de este hecho será dada en el Corolario 6.5.31, página 339.

En lo que sigue, a $\mathbb{R}$, como espacio vectorial sobre $\mathbb{Q}$, lo denotaremos por $\mathbb{R}_{\mathbb{Q}}$. Por el Teorema 1.3.7 sabemos que todo espacio vectorial sobre un cuerpo no trivial posee una base de Hamel, en particular, $\mathbb{R}_{\mathbb{Q}}$ posee una base de Hamel que denotaremos por $\mathcal{H} = \{x_\alpha : \alpha \in D\}$. Recordemos que $\mathcal{H}$ queda determinada por las dos propiedades siguientes:

(a) $\mathcal{H}$ es linealmente independiente, y

(b) $\mathrm{Lin}(\mathcal{H}, \mathbb{Q}) = \mathbb{R}$, donde

$$\mathrm{Lin}(\mathcal{H}, \mathbb{Q}) = \big\{q_1x_1 + \cdots + q_kx_k : \ q_1, \ldots, q_k \in \mathbb{Q}, \ x_1, \ldots, x_k \in \mathcal{H}, \ k \in \mathbb{N}\big\}.$$

Un hecho simple, pero que siempre hay que tener presente (véase el Corolario 1.2.11, página 16), es el siguiente:

‖ ► *Si A es un conjunto **infinito numerable**, también lo es la familia*

$$\mathcal{P}_{\mathrm{fin}}(A) = \big\{F \subseteq A : F \text{ es finito}\big\}.$$

Teorema 6.5.27 (Hamel). *Si $\mathcal{H}$ es una **base de Hamel** en $\mathbb{R}_{\mathbb{Q}}$, entonces*

$$\operatorname{card}(\mathcal{H}) \;=\; \mathfrak{c}.$$

Prueba. Aceptando la Hipótesis del Continuo. Suponga, por un momento, que $\mathcal{H}$ es numerable y escribamos

$$\mathcal{H} \;=\; \{x_1, x_2, \ldots\}.$$

Por el resultado anterior, la familia $\mathcal{P}_{\text{fin}}(\mathcal{H})$ es numerable. Para cada $F \in \mathcal{P}_{\text{fin}}(\mathcal{H})$, digamos $F = \{x_{n_1}, \ldots, x_{n_k}\}$, sea $\operatorname{Lin}(F, \mathbb{Q})$ el conjunto formado por todas las combinaciones lineales racionales de elementos de F, es decir,

$$\operatorname{Lin}(F, \mathbb{Q}) \;=\; \big\{q_1 x_{n_1} + \cdots + q_k x_{n_k} : \; (q_1, \ldots, q_k) \in \mathbb{Q}^k\big\}.$$

Puesto que la aplicación $\varphi : \mathbb{Q}^k \times F \to \operatorname{Lin}(F, \mathbb{Q})$ dada por

$$\varphi\big(q_1, \ldots, q_k, x_{n_1}, \ldots, r_{n_k}\big) \;=\; q_1 x_{n_1} + \cdots + q_k x_{n_k}$$

es claramente sobreyectiva, se sigue entonces de la numerabilidad de $\mathbb{Q}^k \times F$ que $\operatorname{Lin}(F, \mathbb{Q})$ es también numerable y, por consiguiente, el conjunto

$$\bigcup_{F \in \mathcal{P}_{\text{fin}}(\mathcal{H})} \operatorname{Lin}(F, \mathbb{Q})$$

es numerable. Observe finalmente que si $x \in \mathbb{R}$, entonces $x \in \operatorname{Lin}(F, \mathbb{Q})$ para algún $F \in \mathcal{P}_{\text{fin}}(\mathcal{H})$ y, en consecuencia,

$$\mathbb{R} \;=\; \bigcup_{F \in \mathcal{P}_{\text{fin}}(\mathcal{H})} \operatorname{Lin}(F, \mathbb{Q}).$$

Esto último nos revela que $\mathbb{R}$ es numerable lo que constituye una flagrante contradicción. Por esto, $\mathcal{H}$ es no-numerable y, entonces, por una aplicación de la Hipótesis del Continuo se concluye que $\operatorname{card}(\mathcal{H}) = \mathfrak{c}$.

Evitando la Hipótesis del Continuo. Suponga que $\mathcal{H}$ es no-numerable y sea $\mathfrak{m} = \operatorname{card}(\mathcal{H})$. Para cada $n \in \mathbb{N}$, considere el conjunto

$$\mathcal{H}_n \;=\; \big\{F \subseteq \mathcal{H} : \operatorname{card}(F) = n\big\}.$$

Observe que $\operatorname{card}(\mathcal{H}_n) \leq \mathfrak{m}^n = \mathfrak{m}$ y, por lo tanto,

$$\operatorname{card}\left(\bigcup_{n=1}^{\infty} \mathcal{H}_n\right) \;\leq\; \aleph_0 \cdot \mathfrak{m} \;=\; \mathfrak{m}.$$

Por otro lado, para cualquier $F \in \bigcup_{n=1}^{\infty} \mathcal{H}_n$ se tiene que $\operatorname{card}(\operatorname{Lin}(F, \mathbb{Q})) = \aleph_0$ y como

$$\mathbb{R} \;=\; \operatorname{Lin}(\mathcal{H}, \mathbb{Q}) \;=\; \bigcup_{F \in \bigcup_{n=1}^{\infty} \mathcal{H}_n} \operatorname{Lin}(F, \mathbb{Q})$$

resulta que $\mathfrak{c} = \operatorname{card}(\mathbb{R}) \leq \mathfrak{m} \cdot \aleph_0 = \mathfrak{m} \leq \mathfrak{c}$. Esto nos revela que $\mathfrak{m} = \mathfrak{c}$ y, así, $\operatorname{card}(\mathcal{H}) = \mathfrak{c}$. Fin de la prueba. ■

Fijemos una base de Hamel $\mathcal{H}$ en $\mathbb{R}_{\mathbb{Q}}$. Algunas propiedades interesantes que posee $\mathcal{H}$ son las siguientes:

$(\mathfrak{H})_1$ $a\mathcal{H}$ es una base de Hamel para cualquier $a \in \mathbb{R} \setminus \{0\}$.

$(\mathfrak{H})_2$ $\mathcal{H} \cap (a + \mathcal{H}) = \varnothing$ para todo $a \in \mathcal{H}$.

En efecto, suponga que $\mathcal{H} \cap (a + \mathcal{H}) \neq \varnothing$ para algún $a \in \mathcal{H}$ y sea $y \in \mathcal{H} \cap (a + \mathcal{H})$. Seleccionemos un $x \in \mathcal{H}$ de modo que $y = a + x$. También, como $y \in \mathcal{H}$, resulta que el conjunto $\{a, x, y\} \subseteq \mathcal{H}$ es linealmente dependiente, lo cual es imposible. Por esto, $(\mathfrak{H})_2$ se cumple.

$(\mathfrak{H})_3$ $a + \mathcal{H} \subseteq \mathbb{R} \setminus \mathcal{H}$ para todo $a \in \mathcal{H}$.

$(\mathfrak{H})_4$ No existe ninguna base de Hamel $\mathcal{H}$ tal que $ab \in \mathcal{H}$ para todo $a, b \in \mathcal{H}$.

Por el Teorema de Hamel sabemos que $\mathrm{card}(\mathcal{H}) = \mathfrak{c}$, por lo que podemos escribir $\mathcal{H} = \{x_\alpha : \alpha \in [0,1]\}$. Observe que, por definición, para cada $x \in \mathbb{R}$, existe $F \in \mathcal{P}_{\mathrm{fin}}(\mathcal{H})$, digamos $F = \{x_{\alpha_1}, \ldots, x_{\alpha_n}\}$, y coeficientes racionales únicos $q_1(x), \ldots, q_n(x)$ tales que

$$x = q_1(x)x_{\alpha_1} + q_2(x)x_{\alpha_2} + \cdots + q_n(x)x_{\alpha_n}.$$

Teorema 6.5.28 (Lévy). *Para cada* ***base de Hamel*** *$\mathcal{H}$ en $\mathbb{R}_{\mathbb{Q}}$, existe un subconjunto $V_{\mathcal{H}}$ de $\mathbb{R}$ que* ***no es medible*** *según Lebesgue.*

Prueba. Escribamos $\mathcal{H} = \{x_\alpha : \alpha \in [0,1]\}$ y fijemos un índice α_0 de $[0,1]$. Considere ahora el subespacio vectorial de $\mathbb{R}_{\mathbb{Q}}$ generado por el conjunto $\mathcal{H} \setminus \{x_{\alpha_0}\}$. Denotemos a este subespacio por $V_{\mathcal{H}}$. Observe que, por definición, $x_{\alpha_0} \notin V_{\mathcal{H}}$. Nuestro objetivo es demostrar que $V_{\mathcal{H}}$ no puede ser medible según Lebesgue. Suponga lo contrario, es decir, que $V_{\mathcal{H}} \in \mathfrak{M}_\mu(\mathbb{R})$ y sea $(q_n)_{n=0}^\infty$ una enumeración sin duplicaciones de $\mathbb{Q}$ con $q_0 = 0$. Si para cada $n \in \mathbb{N}_0$ definimos $V_n = q_n x_{\alpha_0} + V_{\mathcal{H}}$, resulta entonces que:

(a) V_n es medible y $\mu(V_n) = \mu(V_{\mathcal{H}})$ para todo $n \geq 0$.

(b) $V_m \cap V_n = \varnothing$ para todo $m, n \in \mathbb{N}$ con $m \neq n$.

En efecto, suponga por un momento que $V_m \cap V_n \neq \varnothing$ y sea $x \in V_m \cap V_n$. Entonces existen x_{α_m} y x_{α_n} en $V_{\mathcal{H}}$ tales que

$$x = q_m x_{\alpha_0} + x_{\alpha_m} = q_n x_{\alpha_0} + x_{\alpha_n}.$$

De esto se sigue que $(q_m - q_n)x_{\alpha_0} = x_{\alpha_n} - x_{\alpha_m} \in V_{\mathcal{H}}$, pues $V_{\mathcal{H}}$ es un espacio vectorial. Pero como $q_m \neq q_n$, resulta que $x_{\alpha_0} = (q_m - q_n)^{-1}(x_{\alpha_n} - x_{\alpha_m}) \in V_{\mathcal{H}}$, lo cual es imposible.

(c) $\mathbb{R} = \displaystyle\bigcup_{n=0}^\infty V_n$.

(d) $V_{\mathcal{H}} - V_{\mathcal{H}} = V_{\mathcal{H}}$, pues $V_{\mathcal{H}}$ es un espacio vectorial.

Observe que si $\mu(V_{\mathcal{H}}) = 0$, entonces de (a), (c) y la numerabilidad de μ se tiene que

$$+\infty = \mu(\mathbb{R}) = \sum_{n=1}^\infty \mu(V_n) = \sum_{n=1}^\infty \mu(V_{\mathcal{H}}) = 0,$$

lo que obviamente es un disparate. Suponga ahora que $\mu(V_{\mathcal{H}}) > 0$ y nótese que, gracias a (b),

$$V_{\mathcal{H}} \cap (q_n x_{\alpha_0} + V_{\mathcal{H}}) = \varnothing \tag{$**$}$$

para cualquier $q_n \neq 0$. Por otro lado, como $\mu(V_{\mathcal{H}}) > 0$, el Teorema de Steinhaus, Teorema 6.3.54, nos garantiza la existencia de un entorno abierto G_0 del 0 tal que $V_{\mathcal{H}} \cap (g + V_{\mathcal{H}}) \neq \varnothing$ para todo $g \in G_0$. Sin embargo, como $\mathbb{Q}$ es denso, podemos elegir un $q_n \in \mathbb{Q}$ lo suficientemente pequeño de modo que $q_n x_{\alpha_0} \in G_0$ y, por consiguiente,

$$V_{\mathcal{H}} \cap (q_n x_{\alpha_0} + V_{\mathcal{H}}) \neq \varnothing$$

lo que constituye una violación a $(**)$. Esta contradicción establece que la condición $\mu(V_{\mathcal{H}}) > 0$ tampoco puede ocurrir y termina la prueba. ■

Nota Adicional 6.5.9 Es importante destacar que en la prueba del Teorema de Lévy se obtuvo que $\mathbb{R}$ *se puede representar como una unión numerable y disjunta de subconjuntos no-medibles*. Otro aspecto interesante es que, si bien es cierto que la existencia de una base de Hamel conduce a la creación de un conjunto no-medible según Lebesgue surge, de modo natural, la pregunta de si existen bases de Hamel que sean no-medibles. La respuesta, como era de esperarse, es que existen bases de Hamel que son medibles y otras que no lo son. Algunas otras propiedades relacionadas con bases de Hamel son las siguientes:

(1) ***Cualquier subconjunto medible de $V_{\mathcal{H}}$ posee medida 0.***

Prueba. Sea $E \subseteq V_{\mathcal{H}}$ medible y suponga que $\mu(E) > 0$. El Teorema de Steinhaus muestra que $E - E$ contiene un intervalo abierto J con 0 en su interior. Por lo tanto,

$$J \subseteq E - E \subseteq V_{\mathcal{H}} - V_{\mathcal{H}} = V_{\mathcal{H}},$$

de donde se concluye, por ser $V_{\mathcal{H}}$ un espacio vectorial, que $V_{\mathcal{H}} = \mathbb{R}$. Esta contradicción establece que $\mu(E) = 0$. ■

(2) ***Si $\mathcal{H}$ es una base de Hamel medible según Lebesgue, entonces $\mu(\mathcal{H}) = 0$.***

Prueba. Suponga que $\mu(\mathcal{H}) > 0$ y sea $x_0 \in \mathcal{H}$. Entonces $\mathcal{H}' = \mathcal{H} \setminus \{x_0\}$ es medible con $\mu(\mathcal{H}') > 0$. Invoquemos el Teorema de Steinhaus para obtener un intervalo abierto $J = (-\varepsilon, \varepsilon)$ tal que $J \subseteq \mathcal{H}' - \mathcal{H}'$. Por otro lado, como $x_0 \in \mathbb{R}$, el Principio de Arquímedes nos garantiza la existencia de un número racional $q \neq 0$ de modo tal que $qx_0 \in J$ y, en consecuencia, $qx_0 \in \mathcal{H} - \mathcal{H}$. De allí que existen $u, v \in \mathcal{H}$ tales que $qx_0 = u - v$; en otras palabras, $qx_0 - (u - v) = 0$. Esto nos revela que $\mathcal{H}$ no es linealmente independiente y termina la prueba. ■

(3) ***Existe una base de Hamel que es medible según Lebesgue.***

Prueba. Sea Γ el conjunto ternario de Cantor. Sabemos que Γ es medible de medida cero, por lo que todo subconjunto de él es medible. Más aun, como

$$\Gamma + \Gamma = [0, 2], \tag{2a}$$

resulta que Γ genera a $\mathbb{R}_{\mathbb{Q}}$, es decir, $\mathrm{Lin}(\Gamma, \mathbb{Q}) = \mathbb{R}$. Para ver esto, sea $x \in \mathbb{R}$ y escoja $q \in \mathbb{Q}$ tal que $qx \in [0, 2]$. Use ahora $(2a)$ para determinar un par de elementos $u, v \in \Gamma$ tal que $u + v = qx$. De aquí se sigue que $x \in \mathrm{Lin}(\Gamma, \mathbb{Q})$ y, en consecuencia, $\mathrm{Lin}(\Gamma, \mathbb{Q}) = \mathbb{R}$. Un llamado al Corolario 1.3.9, página 52, nos revela que existe una base de Hamel $\mathcal{H} \subseteq \Gamma$. Por supuesto, $\mathcal{H}$ es medible. ■

(4) ***Bajo la* Hipótesis del Continuo*, existe una base de Hamel que es un conjunto de Bernstein. Por consiguiente, existen bases de Hamel que no son medibles según Lebesgue*.**

Prueba. Sea $\mathcal{F}_{\mathfrak{c}}$ la familia de todos los subconjuntos cerrados no numerables de $\mathbb{R}$. Sabemos que $\operatorname{card}(\mathcal{F}_{\mathfrak{c}}) = \mathfrak{c}$. La Hipótesis del Continuo permite expresar a $\mathcal{F}_{\mathfrak{c}}$ en la forma $\mathcal{F}_{\mathfrak{c}} = \{F_\alpha : \alpha < \omega_1\}$. Nuestro objetivo es usar el Método de Inducción Transfinita para obtener una familia $\mathbb{Q}$-linealmente independiente, digamos $H = \{x_\alpha : \alpha < \omega_1\}$ tal que $x_\alpha \in F_\alpha$, para cada $\alpha < \omega_1$. Si G es un subconjunto no vacío de $\mathbb{R}$, denote por $\operatorname{Lin}(G, \mathbb{Q})$ al conjunto de todas las combinaciones $\mathbb{Q}$-lineales de elemento de G, esto es, $x \in \operatorname{Lin}(G, \mathbb{Q})$ si, y sólo si

$$x = q_1 x_{\alpha_1} + \cdots + q_n x_{\alpha_n},$$

donde $x_{\alpha_1}, \ldots, x_{\alpha_n} \in G$ y $q_1, \ldots, q_n \in \mathbb{Q}$. Sea $x_0 \in F_0$ con $x_0 \neq 0$. Suponga que, para un ordinal $\beta < \omega_1$, la familia $\mathbb{Q}$-linealmente independiente $G_\beta = \{x_\alpha : \alpha < \beta\}$ ha sido definida. Considere el conjunto

$$H_\beta = \operatorname{Lin}\big(G_\beta, \mathbb{Q}\big).$$

Puesto que $\aleph_0 = \operatorname{card}(H_\beta) < \mathfrak{c}$ y F_β es no-numerable, resulta que $F_\beta \setminus H_\beta \neq \varnothing$. Seleccione un punto $x_\beta \in F_\beta \setminus H_\beta$. De este modo queda construida la familia requerida

$$H = \big\{x_\alpha : \alpha < \omega_1\big\}.$$

Ahora, usando el Corolario 1.3.8, página 52, encuentre una base de Hamel $\mathcal{H}$ tal que $H \subseteq \mathcal{H}$. Afirmamos que $\mathcal{H}$ es un conjunto de Bernstein. En efecto, sea F un conjunto cerrado no-numerable. Entonces $F = F_\alpha$ para algún $\alpha < \omega_1$ y, así, por construcción, $x_\alpha \in F_\alpha \cap H \subseteq F \cap \mathcal{H}$. De modo similar, se verifica que $F \cap (\mathbb{R} \setminus \mathcal{H}) \neq \varnothing$ y termina la prueba. ■

(5) ***Ninguna base de Hamel es un conjunto analítico. En particular, ninguna base de Hamel es medible según Borel*.**

Prueba. Sea $\mathcal{H}$ una base de Hamel y suponga que es un conjunto analítico. Fijemos $x_{\alpha_0} \in \mathcal{H}$ y considere $\mathcal{B} = \mathcal{H} \setminus \{x_{\alpha_0}\}$. Para cada $k \in \mathbb{N}$, sea $\mathcal{B}^{(k)}$ el conjunto de todos los números reales que pueden ser representados en la forma

$$q_1 x_{\alpha_1} + \cdots + q_k x_{\alpha_k}$$

donde $x_{\alpha_1}, \ldots, x_{\alpha_k} \in \mathcal{B}$ y $q_1, \ldots, q_k \in \mathbb{Q} \setminus \{0\}$ con $q_i \neq q_j$ para $i \neq j$. Finalmente, sea $\mathcal{B}_q = q \cdot \mathcal{B}$ para cada $q \in \mathbb{Q} \setminus \{0\}$. Observe ahora que:

(a) $\mathcal{B}$ y $\mathcal{B}_q$ son claramente conjuntos analíticos.

(b) $\mathcal{B}^{(1)} = \bigcup_{q \in \mathbb{Q} \setminus \{0\}} \mathcal{B}_q$ es un conjunto analítico por el Lema 6.3.29, página 274.

(c) $\mathcal{B}^{(k)} = \mathcal{B}^{(1)} \oplus \mathcal{B}^{(1)} \oplus \cdots \oplus \mathcal{B}^{(1)}$ $(k - \text{veces})$ es un conjunto analítico. En efecto, en primer lugar, observe que si H y K son conjuntos analíticos, entonces $H + K$, por ser la proyección del conjunto analítico $H \times K$ sobre la recta $y = x$, es también analítico. Luego, por inducción, si $H_1, \ldots, H_n$ son conjuntos analíticos, entonces $H_1 + \cdots + H_n$ es analítico.

De esto se sigue que $\bigcup_{k=1}^{\infty} \mathcal{B}^{(k)}$ es un conjunto analítico, en particular, medible según Lebesgue gracias al Teorema 6.3.30. Sin embargo, esto no es posible ya que

$$\bigcup_{k=1}^{\infty} \mathcal{B}^{(k)} = V_{\mathcal{H}},$$

donde $V_{\mathcal{H}}$ es el subespacio vectorial generado por $\mathcal{B}$, el cual, por el Teorema de Levy, no es medible según Lebesgue. Esta contradicción establece que $\mathcal{H}$ no puede ser un conjunto analítico. ■

Combinado (3) y (5) se obtiene:

(6) ***Existe un conjunto medible según Lebesgue que no es medible según Borel.***

(7) ***Existencia de funciones de Cauchy no-continuas.***

Recordemos que una función $f : \mathbb{R} \to \mathbb{R}$ es llamada una **función de Cauchy** si ella es aditiva, esto es, si

$$f(x+y) = f(x) + f(y)$$

para todo $x, y \in \mathbb{R}$. Es fácil comprobar que toda función de Cauchy f es $\mathbb{Q}$-**lineal**, es decir,

$$f(qx) = qf(x)$$

para todo $x \in \mathbb{R}$ y todo $q \in \mathbb{Q}$. Veamos esto. En primer lugar, observe que

$$f(0) = 0$$

ya que $f(0) = f(0+0) = 2f(0)$. De aquí se deduce que

$$f(-x) = -f(x).$$

Más aun, para cada $x \in \mathbb{R}$ y cada $n \in \mathbb{N}$, se sigue por inducción que

$$f(2x) = 2f(x), \quad f(3x) = 3f(x), \quad \cdots, \quad f(nx) = nf(x).$$

Si n es un entero negativo, entonces $-n$ es un entero positivo y, así,

$$f(nx) = f(-(-nx)) = -f(-nx) = -(-n)f(x) = nf(x).$$

Sea ahora $q = m/n$ un número racional. Entonces, $mx = n(qx)$ y, por lo tanto,

$$mf(x) = f(mx) = f(n(qx)) = nf(qx),$$

de donde se tiene que

$$f(qx) = qf(x).$$

Si tomamos $a = f(1)$, resulta de lo anterior que

$$f(q) = a \cdot q \quad \text{para todo } q \in \mathbb{Q}. \tag{FC$_1$}$$

En presencia de continuidad las funciones $\mathbb{Q}$-lineales se convierten automáticamente en $\mathbb{R}$-lineales (= lineales).

Teorema 6.5.29. *Si $f : \mathbb{R} \to \mathbb{R}$ es una* ***función de Cauchy continua****, entonces f es de la forma*

$$f(x) = a \cdot x \quad \text{para todo } x \in \mathbb{R},$$

donde $a = f(1)$.

Prueba. Sea $x \in \mathbb{R}$. Como $\mathbb{Q}$ es denso en $\mathbb{R}$, seleccione una sucesión $(q_n)_{n=1}^{\infty}$ en $\mathbb{Q}$ tal que $q_n \to x$. Ahora por (FC_1) y la continuidad de f resulta que

$$f(x) = \lim_{n\to\infty} f(q_n) = \lim_{n\to\infty} a \cdot q_n = a \cdot x$$

La prueba es completa. ∎

Uno puede preguntarse: *¿existen funciones de Cauchy que no son continuas*? Aunque nunca se ha construido una tal función, el Axioma de Elección garantiza la existencia de tales monstruos.

Lema 6.5.30 (Existencia de funciones aditivas). *Sea $\mathcal{H}$ una* ***base de Hamel*** *en $\mathbb{R}_{\mathbb{Q}}$. Para cada función arbitraria $g : \mathcal{H} \to \mathbb{R}$, existe una* ***función de Cauchy*** *$f : \mathbb{R} \to \mathbb{R}$ tal que $f(x) = g(x)$ para todo $x \in \mathcal{H}$.*

Prueba. Pongamos $\mathcal{H} = \{x_\alpha : \alpha \in [0,1]\}$. Para cada $x \in \mathbb{R}$, existen $x_{\alpha_1}, \ldots, x_{\alpha_n}$ en $\mathcal{H}$ y $q_1, \ldots, q_n$ en $\mathbb{Q}$ tales que $x = q_1 x_{\alpha_1} + \cdots + q_n x_{\alpha_n}$. Defina $f : \mathbb{R} \to \mathbb{R}$ por

$$f(x) = q_1 g(x_{\alpha_1}) + \cdots + q_n g(x_{\alpha_n}).$$

Observe que f está bien definida ya que, al ser $\mathcal{H}$ una base de Hamel, la elección de los números $x_{\alpha_1}, \ldots, x_{\alpha_n}, q_1, \ldots, q_n$ son únicos. También es claro que $f(x) = g(x)$ para cualquier $x \in \mathcal{H}$. Resta por demostrar que f es aditiva. Sean $x, y \in \mathbb{R}$. Entonces

$$x = q_1 x_{\alpha_1} + \cdots + q_n x_{\alpha_n} \qquad \text{y} \qquad y = r_1 x_{\beta_1} + \cdots + r_m x_{\beta_m}$$

donde $x_{\alpha_1}, \ldots x_{\alpha_n}, x_{\beta_1}, \ldots, x_{\beta_m}$ son miembros de la base $\mathcal{H}$ y los números $q_1, \ldots, q_n$ y $r_1, \ldots, r_m$ son racionales. Observe que los dos conjuntos $\{x_{\alpha_1}, \ldots, x_{\alpha_n}\}$ y $\{x_{\beta_1}, \ldots, x_{\beta_m}\}$ pueden tener miembros en común. Sea $\{z_1, \ldots, z_k\}$ la unión de esos dos conjuntos. Entonces $k \leq m + n$, y se tiene que

$$x = a_1 z_1 + \cdots + a_k z_k \qquad \text{y} \qquad y = b_1 z_1 + \cdots + b_k z_k,$$

donde los números $a_1, b_1, \ldots, a_k, b_k$ son racionales, algunos de los cuales pueden ser ceros. Ahora,

$$x + y = (a_1 + b_1) z_1 + \cdots + (a_k + b_k) z_k$$

y

$$\begin{aligned} f(x+y) &= f\big((a_1+b_1)z_1 + \cdots + (a_k+b_k)z_k\big) \\ &= (a_1+b_1)g(z_1) + \cdots + (a_k+b_k)g(z_k) \\ &= \big[a_1 g(z_1) + \cdots + a_k g(z_k)\big] + \big[\, b_1 g(z_1) + \cdots + b_k g(z_k)\big] \\ &= f(x) + f(y). \end{aligned}$$

Esto termina la prueba. ∎

Corolario 6.5.31 (Hamel). *Existen **funciones de Cauchy** $f : \mathbb{R} \to \mathbb{R}$ **que no son continuas**.*

Prueba. Sea $\mathcal{H} = \{x_\alpha : \alpha \in [0,1]\}$ una base de Hamel en $\mathbb{R}_{\mathbb{Q}}$ y fije cualquier elemento $h \in \mathcal{H}$. Defina $g_h : \mathcal{H} \to \mathbb{R}$ del modo siguiente:

$$g_h(x) = \begin{cases} 0 & \text{si } x \in \mathcal{H} \setminus \{h\} \\ 1 & \text{si } x = h. \end{cases}$$

Por el Lema 6.5.30, existe una función de Cauchy $f : \mathbb{R} \to \mathbb{R}$ tal que $f(x) = g_h(x)$ para todo $x \in \mathcal{H}$. Afirmamos que f no puede ser continua. En efecto, suponga por un momento que f es continua. Entonces existe una constante a tal que $f(x) = ax$ para todo $x \in \mathbb{R}$. Veamos que esto conduce a una contradicción. Seleccione un $x \in \mathcal{H}$ arbitrario con $x \neq h$ y observe que como $hf(x) = f(hx) = f(xh) = xf(h)$, entonces se tiene que

$$0 = \frac{f(x)}{x} = \frac{f(h)}{h} = \frac{1}{h}.$$

Esta contradicción establece que f no puede ser continua y finaliza la prueba. ■

Observe que la función f del corolario anterior también se puede elegir *inyectiva*. Para ver esto, sea $\phi : [0,1] \to [0,1]$ una aplicación inyectiva arbitraria distinta de la identidad. Si ahora imponemos la condición de que $f(x_\alpha) = x_{\phi(\alpha)}$ para todo $x_\alpha \in \mathcal{H}$ resulta que f es inyectiva.

(8) ***Si** $f : \mathbb{R} \to \mathbb{R}$ **es una función de Cauchy no-continua, entonces***

$$\mathrm{Graf}(f) = \{(x, f(x)) \in \mathbb{R}^2 : x \in \mathbb{R}\}$$

***es denso en** $\mathbb{R}^2$.*

Prueba. Sea $(x,y) \in \mathbb{R}^2$ y sea G un conjunto abierto en $\mathbb{R}^2$ conteniendo a (x,y). Veamos que $G \cap \mathrm{Graf}(f) \neq \varnothing$. Puesto que f no es lineal, existen $a \neq 0$ y $b \neq 0$ tales que

$$\frac{f(a)}{a} \qquad \text{y} \qquad \frac{f(b)}{b}$$

son diferentes. Por consiguiente, los vectores

$$u = (a, f(a)) \qquad \text{y} \qquad v = (b, f(b))$$

son linealmente independientes y, por lo tanto, forman una base de $\mathbb{R}^2$. De esto se sigue que existen números reales p, q tales que

$$(x,y) = pu + qv.$$

Ahora bien, como $\mathbb{Q}^2$ es denso en $\mathbb{R}^2$, entonces $G \cap \mathbb{Q}^2 \neq \varnothing$ y, en consecuencia, existen números racionales r_p y r_q tales que

$$r_p u + r_q v \in G.$$

Observe que

$$\begin{aligned} r_p u + r_q v &= \big(r_p a + r_q b,\ r_p f(a) + r_q f(b)\big) \\ &= \big(r_p a + r_q b,\ f(r_p a + r_q b)\big) \in G \cap \mathrm{Graf}(f) \end{aligned}$$

lo cual prueba que $\mathrm{Graf}(f)$ es denso en $\mathbb{R}^2$. ■

(9) ***Las funciones de Cauchy no-continuas son más abundantes que las que son continuas.***

Teorema 6.5.32. *En* **ZFC** *existen:*

(a) $2^{\aleph_0}$ *funciones de Cauchy continuas.*

(b) $2^{2^{\aleph_0}}$ *funciones de Cauchy que no son continuas.*

Prueba. (a) Para cada $a \in \mathbb{R}$, la función $f_a : \mathbb{R} \to \mathbb{R}$ definida por $f_a(x) = a \cdot x$ para todo $x \in \mathbb{R}$ es una función de Cauchy continua. Más aun, gracias a (6), toda función de Cauchy $f : \mathbb{R} \to \mathbb{R}$ que es continua es de la forma $f = f_a$, donde $a = f(1)$. Esto prueba la parte (a).

(b) Por Teorema 6.5.27, $\mathbb{R}_{\mathbb{Q}}$ posee una base de Hamel $\mathcal{H}$ con cardinalidad $2^{\aleph_0}$. Puesto que cualquier aplicación $g : \mathcal{H} \to \mathbb{R}$ se puede extender unívocamente a una función aditiva $f : \mathbb{R} \to \mathbb{R}$, Lema 6.5.30, existen precisamente

$$\operatorname{card}\left(\mathbb{R}^{\mathcal{H}}\right) = \operatorname{card}(\mathbb{R})^{\operatorname{card}(\mathcal{H})} = \left(2^{\aleph_0}\right)^{2^{\aleph_0}} = 2^{\aleph_0 \cdot 2^{\aleph_0}} = 2^{2^{\aleph_0}}$$

funciones de Cauchy y ya que sólo existen, por (a), $2^{\aleph_0}$ funciones de Cauchy continuas, resulta entonces que existen $2^{2^{\aleph_0}}$ de tales funciones que no son continuas. ■

Una vez que el Axioma de Elección entra en acción, se pueden construir ciertos conjuntos y funciones que, por lo general, poseen algunas patologías extrañas. Otros ejemplos raros, sin embargo, no requieren el uso de tal herramienta. Por ejemplo, Henri Lebesgue, quien no era creyente del Axioma de Elección, construyó una función con una patología muy rara: *siempre es sobreyectiva*, es decir:

Definición 6.5.33. *Sea $f : \mathbb{R} \to \mathbb{R}$ una función. Se dice que f es **siempre-sobreyectiva** si ocurre que $f((a,b)) = \mathbb{R}$ para cualquier intervalo abierto $(a,b) \subseteq \mathbb{R}$ con $a < b$.*

El siguiente resultado nos dice cómo se caracterizan las funciones que son siempre-sobreyectivas.

Teorema 6.5.34 (Lebesgue). *Sea $f : \mathbb{R} \to \mathbb{R}$ una función arbitraria. Las siguientes condiciones son equivalentes:*

*(1) f es **siempre-sobreyectiva**.*

*(2) $f^{-1}(t)$ es **denso en** $\mathbb{R}$ para cualquier $t \in \mathbb{R}$.*

Prueba. Suponga que f es siempre-sobreyectiva pero que $f^{-1}(t)$ no es denso en $\mathbb{R}$ para algún $t \in \mathbb{R}$. Esto, por supuesto, significa que existe algún intervalo abierto (a,b) en $\mathbb{R}$ tal que $f^{-1}(t) \cap (a,b) = \varnothing$. Selecciones un $x \in f^{-1}(t)$ de modo tal que $x \notin (a,b)$ y observe ahora que $t = f(x) \notin f((a,b)) = \mathbb{R}$. Esta contradicción prueba que $f^{-1}(t)$ es denso en $\mathbb{R}$.

Recíprocamente, suponga que (2) se cumple y sea $t \in \mathbb{R}$. Fijemos un intervalo abierto (a,b) en $\mathbb{R}$. Como $f^{-1}(t)$ es denso en $\mathbb{R}$, resulta que $f^{-1}(t) \cap (a,b) \neq \varnothing$. Si se selecciona cualquier $x \in f^{-1}(t) \cap (a,b)$, tendremos que $f(x) = t$. ■

Teorema 6.5.35 (Lebesgue). *Existen funciones $f : \mathbb{R} \to \mathbb{R}$ que son **siempre-sobreyectivas**.*

Prueba. Denote por $\mathcal{A} = \{A_x : x \in \mathbb{R}\}$ la colección de todas las clases de equivalencias determinadas por la relación de equivalencia:

$$x \sim y \qquad \Leftrightarrow \qquad x - y \in \mathbb{Q}.$$

Tal como se demostró en la prueba de la existencia de un conjunto no-medible al estilo Vitali, resulta que $\mathcal{A}$ es una colección no-numerable, donde cada par de sus miembros o son iguales o son disjuntos. En particular,

$$t \in A_x \qquad \Leftrightarrow \qquad A_t = A_x.$$

De lo anterior se deduce que $\operatorname{card}(\mathcal{A}) = \mathfrak{c}$. Teniendo en cuenta que también $\operatorname{card}(\mathbb{R}) = \mathfrak{c}$, podemos determinar una biyección $g : \mathcal{A} \to \mathbb{R}$. Defina ahora la función $f : \mathbb{R} \to \mathbb{R}$ por

$$f(x) = g(A_x) \quad \text{para todo } x \in \mathbb{R}.$$

Veamos que f es *siempre-sobreyectiva*. Para ver esto, fijemos $a, b \in \mathbb{R}$ con $a < b$. Como $\mathbb{Q}$ es denso en $\mathbb{R}$, resulta que cada traslado $x + \mathbb{Q} = A_x$ también es denso en $\mathbb{R}$ y, en consecuencia $A_x \cap (a, b) \neq \varnothing$ para todo $x \in \mathbb{R}$. Sea $y \in f((a, b))$ y seleccione un $t \in (a, b)$ tal que $y = f(t)$. Puesto que

$$t \in (a, b) = \bigcup_{x \in \mathbb{R}} (a, b) \cap A_x$$

resulta que existe un único $x \in \mathbb{R}$ tal que $t \in (a, b) \cap A_x$. En consecuencia, $A_t = A_x$ y por lo tanto, $f(t) = g(A_t) = g(A_x)$. De esto se sigue que

$$f((a, b)) = \{f(t) : t \in (a, b)\} = \{g(A_x) : x \in \mathbb{R}\} = g(\mathcal{A}) = \mathbb{R}.$$

La prueba es completa. ∎

Otra forma de demostrar la existencia de funciones siempre-sobreyectivas es como sigue:

Segunda Prueba del Teorema 6.5.35. Sea $\mathcal{I}_{\mathbb{Q}}$ la colección de todos los intervalos abiertos no vacíos con extremos racionales. Puesto que $\mathcal{I}_{\mathbb{Q}}$ es numerable podemos considerar una enumeración de sus elementos, digamos:

$$\mathcal{I}_{\mathbb{Q}} = \{(r_1, s_1), (r_2, s_2), \ldots\}.$$

Nuestro próximo objetivo es construir una sucesión $(\Gamma_n)_{n=1}^{\infty}$ de conjuntos ternarios de Cantor con las siguientes propiedades:

(a) $\Gamma_n \subseteq (r_n, s_n)$ para todo $n \in \mathbb{N}$, y

(b) $\Gamma_m \cap \Gamma_n = \varnothing$ para todo $m \neq n$.

En efecto, sea Γ_1 cualquier conjunto ternario de Cantor incluido en (r_1, s_1) (simplemente construya Γ_1 en cualquier subintervalo cerrado $I \subseteq (r_1, s_1)$). Suponga ahora que hemos construidos conjuntos ternarios de Cantor $\Gamma_1, \ldots, \Gamma_n$ satisfaciendo las condiciones (a) y (b). Para construir Γ_{n+1} se procederá del modo siguiente: como cada Γ_i para $i = 1, \ldots, n$ es de medida cero, también los es su unión $\Gamma_1 \cup \cdots \cup \Gamma_n$. Esto nos indica que

$$(r_{n+1}, s_{n+1}) \nsubseteq \Gamma_1 \cup \cdots \cup \Gamma_n$$

ya que $\mu((r_{n+1}, s_{n+1})) > 0$. Considere el conjunto

$$V = (r_{n+1}, s_{n+1}) \cap \big(\Gamma_1 \cup \cdots \cup \Gamma_n\big)^c.$$

Observe que V es abierto y no vacío. Fijemos ahora cualquier $x \in V$ y escoja un $\varepsilon > 0$ de modo que el intervalo $(x-\varepsilon, x+\varepsilon) \subseteq V$. Finalmente, construya un conjunto ternario de Cantor $\Gamma_{n+1} \subseteq (x-\varepsilon, x+\varepsilon)$. Esto termina la construcción de la sucesión $(\Gamma_n)_{n=1}^{\infty}$.
Ahora bien, como $\operatorname{card}(\Gamma_n) = \mathfrak{c}$ para todo $n \in \mathbb{N}$, podemos seleccionar una función biyectiva $f_n : \Gamma_n \to \mathbb{R}$ y entonces definir $f : \mathbb{R} \to \mathbb{R}$ por

$$f(x) = \begin{cases} f_n(x) & \text{si } x \in \Gamma_n \text{ para algún } n \in \mathbb{N}, \\ 0 & \text{si } x \notin \displaystyle\bigcup_{n=1}^{\infty} \Gamma_n. \end{cases}$$

Para verificar que f es siempre-sobreyectiva, sea I un intervalo abierto no vacío y escoja un intervalo $(r_n, s_n) \in \mathscr{J}_{\mathbb{Q}}$ tal que $(r_n, s_n) \subseteq I$. Por construcción $\Gamma_n \subseteq (r_n, s_n)$ y, en consecuencia,

$$\mathbb{R} = f_n(\Gamma_n) = f(\Gamma_n) \subseteq f(I) \subseteq \mathbb{R}.$$

De aquí se concluye que $f(I) = \mathbb{R}$ y termina la prueba. ■

Observe que el gráfico de f es *siempre denso* en $\mathbb{R}^2$. Un hecho que es realmente sorprendente es el siguiente resultado.

Corolario 6.5.36. *Cualquier **función siempre-sobreyectiva** $f : \mathbb{R} \to \mathbb{R}$ posee la **Propiedad del Valor Intermedio** pero es **discontinua** en todo punto de $\mathbb{R}$.*

Prueba. Sea $f : \mathbb{R} \to \mathbb{R}$ una función siempre-sobreyectiva. Afirmamos que f *no es continua en ningún punto de* $\mathbb{R}$. En efecto, sea $x_0 \in \mathbb{R}$ y suponga que f es continua en x_0. Tomando $V = (f(x_0) - 1, f(x_0) + 1)$ existe, por la continuidad de f en x_0, un intervalo abierto I_{x_0} conteniendo a x_0 tal que $f(I_{x_0}) \subseteq V$. Pero como f es siempre-sobreyectiva, resulta que $f(I_{x_0}) = \mathbb{R} \subseteq (f(x_0) - 1, f(x_0) + 1)$. Esta contradicción establece que f es discontinua en todo punto de $\mathbb{R}$. También es claro que f posee la Propiedad del Valor Intermedio. ■

Invocando el Teorema 3.1.22, página 164, vemos que la función f del resultado anterior nunca puede ser inyectiva.

En [61], páginas 31-33 y en el Ejercicio 6.9 (2), página 364, se muestran explícitamente la construcción de otras dos funciones que son siempre-sobreyectivas. Más aun, *se puede demostrar la existencia de una función* $f : \mathbb{R} \to \mathbb{R}$ *tal que* $f(P) = \mathbb{R}$ *para todo conjunto perfecto* $P \subseteq \mathbb{R}$ (véase, por ejemplo, [59], Example 3.12, p. 124).

El siguiente resultado muestra que las funciones de Cauchy que no son continuas se parecen mucho a las funciones que son siempre-sobreyectivas.

Teorema 6.5.37. *Si $f : \mathbb{R} \to \mathbb{R}$ es una **función de Cauchy no continua**, entonces $f((a,b))$ es **denso** en $\mathbb{R}$ para cualquier intervalo abierto $(a,b) \subseteq \mathbb{R}$.*

Prueba. Sea (a, b) un intervalo abierto. Para demostrar que $f((a, b))$ es denso en $\mathbb{R}$, tomemos cualquier intervalo abierto $J \subseteq \mathbb{R}$ y veamos que $f((a, b)) \cap J \neq \varnothing$. En efecto, puesto que $\mathrm{Gra}(f)$ es denso en $\mathbb{R}^2$ resulta que

$$\mathrm{Gra}(f) \cap \big[(a, b) \times J\big] \neq \varnothing.$$

Seleccione un $(x, f(x)) \in \mathrm{Gra}(f) \cap \big[(a, b) \times J\big]$. Entonces $x \in (a, b)$ y $f(x) \in J$, lo cual muestra que $f(x) \in f((a, b)) \cap J$. Por esto $f((a, b))$ es denso en $\mathbb{R}$ y termina la prueba. ■

Teorema 6.5.38. *Sea $f : \mathbb{R} \to \mathbb{R}$ una* ***función de Cauchy****. Las siguientes condiciones son equivalentes:*

(1) *f es* ***siempre-sobreyectiva****.*

(2) *f es* ***sobreyectiva*** *pero* ***no inyectiva****.*

Prueba. $(1) \Rightarrow (2)$. Suponga que (1) se cumple. Como $\mathbb{R}$ es un intervalo abierto, resulta de la definición de función siempre-sobreyectiva que $f(\mathbb{R}) = \mathbb{R}$ lo cual prueba que f es sobreyectiva. Por otro lado, como $f^{-1}(t)$ es denso en $\mathbb{R}$ para cada $t \in \mathbb{R}$, Teorema 6.5.34, se tiene que cualquier par de elementos distintos $x, y \in f^{-1}(t)$ satisfacen $f(x) = f(y)$. Esto prueba que f no es inyectiva.
Otra manera de probar que f no es inyectiva es recordar que el conjunto de los puntos de discontinuidad de cualquier función monótona es a lo más numerable, Corolario 3.1.35, página 174, suponer que f es inyectiva y entonces aplicar el Corolario 6.5.36 para arribar a una contradicción.
$(2) \Rightarrow (1)$. Suponga que (2) se cumple. Para demostrar que f es siempre-sobreyectiva es suficiente, gracias al Teorema 6.5.34, verificar que $f^{-1}(t)$ es denso en $\mathbb{R}$ para todo $t \in \mathbb{R}$. Esto lo haremos en dos pasos: el primero es comprobar que $f^{-1}(0)$ es denso en $\mathbb{R}$. En efecto, puesto que f no es inyectiva resulta que

$$f^{-1}(0) \;=\; \big\{x \in \mathbb{R} : \; f(x) = 0\big\}$$

contiene elementos distintos de cero. Sea (a, b) un intervalo abierto y sea $x \in f^{-1}(0)$, $x \neq 0$. Seleccione un número racional q tal que $qx \in (a, b)$ y observe que como f es es una función de Cauchy, entonces $f(qx) = qf(x) = 0$, es decir, $qx \in f^{-1}(0) \cap (a, b)$. Esto prueba que $f^{-1}(0)$ es denso en $\mathbb{R}$.
Veamos ahora que para cualquier t en $\mathbb{R} \setminus \{0\}$, el conjunto $f^{-1}(t)$ es denso en $\mathbb{R}$. En efecto, fijemos $t \in \mathbb{R} \setminus \{0\}$ y usemos el hecho de que f es sobreyectiva para determinar un $x_0 \in \mathbb{R}$ tal que $f(x_0) = t$. Puesto que $x_0 + f^{-1}(0) = f^{-1}(t)$ resulta de la primera parte que $f^{-1}(t)$ es denso en $\mathbb{R}$ y termina la prueba. ■

No queremos finalizar esta sección sin antes informar sobre esta otra dirección para la obtención de conjuntos no-medibles ligeramente diferente a los ofrecidos anteriormente. El Teorema de Hahn-Banach, una de las Joyas de la Corona del Análisis Funcional, se demuestra usando el Lema de Zorn el cual, como sabemos, es equivalente al Axioma de Elección; sin embargo dicho teorema es estrictamente más débil que el Axioma de Elección: uno no puede probar la validez del Axioma de Elección usando **ZF** + **HB**, donde **HB** es el Axioma de que el Teorema de Hahn-Banach se cumple para todo espacio vectorial V. Lo que resulta interesante, asumiendo **ZF** + **HB** como un sistema de axiomas de la Teoría de Conjuntos, es que es posible construir conjuntos no-medibles según Lebesgue. En efecto, Matt Foreman y Friedrich Wehrung demostraron, en un artículo publicado en la revista **Studia Math**. 138(1991), p. 13-19, que *el Teorema de Hahn-Banach implica la existencia de conjuntos no-medibles*.

6.6. Notas Breves sobre El Problema de la Medida

El Problema de la Medida de Lebesgue, después de la solución negativa ofrecida por Vitali, se convirtió en otros problemas al que hemos llamado el Problema de la Medida. En su versión abstracta resulta ser profundo y complicado. En efecto, entre 1929 y 1930, S. Banach lo concibe como un problema de la Teoría de Conjuntos en lugar de la visión geométrica que de el se tenía. Esta nueva visión del problema permitió la creación de una nueva rama de la Teoría de Conjuntos conocida como la Teoría de los Grandes Cardinales: sofisticada, elegante, pero difícil. En esas aguas aun navega dicho problema. La modificación del Problema de la Medida de Lebesgue derivó algunas variantes. En esta sección nos ocuparemos de revisar muy brevemente algunas de ellas. En primer lugar, veremos que la medida de Lebesgue no puede extenderse a todo $\mathscr{P}(\mathbb{R})$, es decir, no existe ninguna función de conjuntos $m : \mathscr{P}(\mathbb{R}) \to \mathbb{R}$ que sea numerablemente aditiva, invariante por traslación y, además, que satisfaga $m(E) = \mu(E)$ para todo $E \in \mathfrak{M}_\mu(\mathbb{R})$. Es a partir de este resultado, y por consiguiente, del Axioma de Elección, de donde comienzan a complicarse las cosas: es la punta del iceberg del Problema de la Medida. En efecto, en la búsqueda por obtener posibles extensiones de la medida de Lebesgue a σ-álgebras más grandes que $\mathfrak{M}_\mu(\mathbb{R})$ hay que pasearse, por supuesto, por varias consideraciones del problema: por ejemplo, si se insiste en obtener una extensión definida sore todo $\mathscr{P}(\mathbb{R})$, entonces dicha extensión no puede, por el resultado ya anunciado, ser, al mismo tiempo, invariante por traslación y numerablemente aditiva, de modo que se presenta un nuevo problema con varias aristas. Necesariamente hay que abandonar una de las dos condiciones anteriores, es decir, o se abandona la aditividad numerable o, en su defecto, la invarianza por traslación. Cada una de estas consideraciones ha generado una investigación sorprendente y, a veces, difícil. Por otro lado, si no se requiere una extensión a todo $\mathscr{P}(\mathbb{R})$, ¿cuál es la mayor σ-álgebra sobre la cual se puede extender μ preservando la aditividad numerable y(o) la invarianza por traslación? Todos estos problemas presentan soluciones que, en ciertos casos, son parciales. Todos ellos son, como lo considera Banach, esencialmente problemas de la Teoría de Conjuntos en lugar de problemas geométricos. De hechos, los Grandes Cardinales de la Teoría Moderna de Conjuntos cuyo análisis quedan fuera del alcance de los objetivos de estas notas, son sus aliados imprescindibles. Sin embargo, ciertas informaciones son dadas al lector interesado.

6.6.1. El Problema de la Medida de Lebesgue y El Axioma de Elección

Sea $\mathcal{A}$ una σ-álgebra de subconjuntos de $\mathbb{R}$ y $m : \mathcal{A} \to [0, +\infty]$ una función de conjuntos. Diremos que m es una **extensión** de μ si cumple con las siguientes condiciones:

(a) $\mathfrak{M}_\mu(\mathbb{R}) \subseteq \mathcal{A}$.

(b) $m(E) = \mu(E)$ para todo $E \in \mathfrak{M}_\mu(\mathbb{R})$.

Si m es una extensión de μ que es, además, numerablemente aditiva e invariante por traslación, entonces diremos que m es una **extensión continua** de μ. Si, más aun, $\mathcal{A} = \mathscr{P}(\mathbb{R})$, entonces a

m la llamaremos una **extensión universal** de μ.

Ahora analizaremos brevemente dos puntos de vistas del problema de la extensión universal de la medida de Lebesgue ambos relacionados con el Axioma de Elección.

(1º). Si se acepta el Axioma de Elección, entonces la demostración de Vitali de la existencia de conjuntos no-medibles no sólo muestra que $\mathfrak{M}_\mu(\mathbb{R}) \neq \mathscr{P}(\mathbb{R})$, sino, además, que *no puede existir ninguna extensión universal* $m : \mathscr{P}(\mathbb{R}) \to [0, +\infty]$ *de la medida de Lebesgue*. En efecto, el lector atento habrá observado que la construcción del conjunto V de Vitali no depende sobre la definición específica de la medida de Lebesgue sino sobre las tres últimas propiedades enunciadas en el Problema de la Medida, es decir,

(α_2) $\mu(I) = \ell(I)$ *para cualquier intervalo* $I \subseteq \mathbb{R}$,

(α_3) μ *es invariante por traslación* y

(α_4) μ *es numerablemente aditiva*,

de modo que si se asume la existencia de una extensión universal $m : \mathscr{P}(\mathbb{R}) \to [0, +\infty]$ de μ, entonces se puede construir, de modo idéntico al conjunto de Vitali, un conjunto $V \subseteq \mathbb{R}$ tal que $V \notin \mathscr{P}(\mathbb{R})$. Este disparate confirma que la aceptación del Axioma de Elección nos conduce al siguiente hecho:

Corolario 6.6.1. *No existe ninguna* ***extensión universal*** $m : \mathscr{P}(\mathbb{R}) \to [0, +\infty]$ *de* μ.

Este resultado nos conduce a la siguiente interrogante:

> *¿Qué tan lejos podemos extender la medida de Lebesgue y qué propiedades puede preservar, si existe, una tal extensión?*

Es problema fue examinado en profundidad bajo suposiciones diferentes:

(a) ¿Podemos extender a μ a una medida continua $m : \mathcal{A} \to [0, +\infty]$ tal que $\mathfrak{M}_\mu(\mathbb{R}) \subsetneqq \mathcal{A} \subsetneqq \mathscr{P}(\mathbb{R})$? Una vieja técnica de J. Łoś y E. Marczewski [96] garantiza que siempre se puede extender la medida de Lebesgue a una σ-álgebra $\mathcal{A}$ que capture cualquier colección finita de subconjuntos no-medibles, es decir, si $\mathcal{A}$ es la σ-álgebra generada por $\mathfrak{M}_\mu(\mathbb{R}) \cup \mathcal{C}$, donde C una colección finita, pero arbitraria, de subconjuntos no-medibles según Lebesgue. Sin embargo, si la colección $\mathcal{C}$ es infinita numerable entonces el problema no se puede decidir en **ZFC**. De hecho, T. Carlson demostró en [30] que es consistente con **ZFC** admitir que para cualquier colección numerable $\mathcal{C}$ de subconjuntos no-medibles, existe una extensión $m : \mathcal{A} \to [0, +\infty]$ de μ con $\mathcal{A} = \sigma(\mathfrak{M}_\mu(\mathbb{R}) \cup \mathcal{C})$.

(b) Si se insiste en obtener una extensión m a todo $\mathscr{P}(\mathbb{R})$, podemos, en principio, exigir que m sea numerablemente aditiva o invariante por traslación, pero no ambas cosas. Si la exigencia es que no sea numerablemente aditiva, entonces el problema se bifurca en dos ramas: que m que sea finitamente aditiva o que ella sea no-numerablemente aditiva. ¿Cuáles son las consecuencias de asumir cada una de esas condiciones? Estos dos aspectos serán tratados brevemente un poco más abajo.

(2º). Si no se acepta el Axioma de Elección, la prueba de la existencia de un conjunto no-medible al estilo Vitali no trabaja. Podemos, en consecuencia, formularnos esta otra pregunta: *¿es indispensable el uso de dicho axioma para garantizar la existencia de conjuntos no medibles según Lebesgue*? La investigación de este problema permitió un avance espectacular de la Teoría Moderna de Conjuntos pues se tuvo que postular la existencia de ciertos cardinales, los así llamados cardinales inaccesibles, lográndose demostrar que: *si existe un cardinal inaccesible, entonces se puede probar que la respuesta es sí*. En efecto, en 1970, Robert Solovay [122] demostró que:

Teorema 6.6.2 (Solovay). *Si el modelo* **ZFC + existe un cardinal inaccesible** *es consistente, entonces también es consistente el modelo* **ZF + DC + cualquier conjunto de números reales es medible según Lebesgue**.

Solovay sugirió, en su artículo, que tal vez el uso de un cardinal inaccesible pudiera ser no necesario. Sin embargo, unos años más tarde, en 1984, S. Shelah [120] probó que tal condición si es necesaria; en otras palabras, si **ZF + DC + todos los subconjuntos de $\mathbb{R}$ son medibles según Lebesgue** es consistente, entonces también lo es **ZFC + existe un cardinal inaccesible**. Como consecuencia de esos resultados se tiene que: *no se puede probar que* $\mathfrak{M}_\mu(\mathbb{R}) \neq \mathscr{P}(\mathbb{R})$ *sin la ayuda del Axioma de Elección*.

6.6.2. El Problema de la Medida de Lebesgue y la Hipótesis del Continuo

¿Cómo entra la Hipótesis del Continuo en la escena del Problema de la Medida? Recordemos que la existencia de conjuntos no-medibles según Lebesgue (por la aceptación del Axioma de Elección) nos indica que la medida de Lebesgue no puede extenderse a todo $\mathscr{P}(\mathbb{R})$, esto significa que no existe ninguna función de conjuntos m definida sobre $\mathscr{P}(\mathbb{R})$ que sea numerablemente aditiva, invariante por traslación y tal que $m(E) = \mu(E)$ para todo conjunto $E \in \mathfrak{M}_\mu(\mathbb{R})$. En vista de éste resultado, podemos intentar investigar, sin prescindir del Axioma de Elección, un problema más general, abandonando una de las dos condiciones que impiden tal extensión: la numerabilidad aditiva o la invarianza por traslación. Nos preguntamos entonces: ¿Qué ocurre si se omite, por ejemplo, el requerimiento de que la medida $m : \mathscr{P}(\mathbb{R}) \to [0, +\infty]$ sea invariante por traslación pero se mantiene la numerabilidad aditiva? Banach propone sustituir la condición geométrica de la invarianza por traslación por esta esta otra condición: $m(\{x\}) = 0$ para todo $x \in \mathbb{R}$, ya que no sabríamos cómo demostrar que los conjuntos finitos y los numerables tienen medida cero, una condición que debería cumplir cualquier medida que intente extender la medida de Lebesgue. En [7], S. Banach propone esta nueva generalización del problema:

El Problema de la Medida Generalizada. *¿Existe una función de conjuntos $m : \mathscr{P}(\mathbb{R}) \to [0, +\infty)$* ***numerablemente aditiva*** *tal que $m(\{x\}) = 0$ para todo $x \in \mathbb{R}$?*.

En [8] Banach y Kuratowski demostraron que la bajo la Hipótesis del Continuo la única solución al Problema de la Medida Generalizada es la medida nula. Para la prueba necesitaremos algunos hechos.

Definición 6.6.3. *Sea X un conjunto no vacío. Una* ***medida continua*** *sobre X es una función de conjuntos $m : \mathscr{P}(\mathbb{R}) \to [0, +\infty)$ tal que:*

(a) $m(\{x\}) = 0$ para cualquier $x \in X$, y

(b) m es numerablemente aditiva.

Observe que si m es una medida continua sobre X, entonces ella es finita, es decir, $m(X) < +\infty$ y, además, finitamente aditiva. Más aun, las propiedades (a) y (b) implican que si A es cualquier conjunto numerable incluido en X, entonces $m(A) = 0$. Además,

(c) $m(A) \leq m(B)$ siempre que $A \subseteq B \subseteq X$.

De esto último se observa que si $A \subseteq B \subseteq X$ y $m(B) = 0$, entonces $m(A) = 0$. Por supuesto, ninguna medida de conteo sobre $\mathscr{P}(\mathbb{R})$ es continua.

Si existe una medida continua sobre X con $m(X) > 0$, entonces las condiciones (a) y (b) nos indican que X es *no-numerable* y, en consecuencia, *la existencia de medidas continuas sobre X*

dependen únicamente de la cardinalidad de X. Nótese también que, si X y Y son conjuntos con la misma cardinalidad y si existe una medida continua m sobre X, entonces existe una medida continua m' sobre Y y viceversa. En efecto, el hecho de X y Y poseen la misma cardinalidad, significa que existe una biyección $f : X \to Y$. Defina $m' : \mathscr{P}(Y) \to [0,1]$ por $m'(A) = m(f^{-1}(A))$ para todo $A \in \mathscr{P}(Y)$. En particular, como $\mathbb{R}$ y ω_1 poseen la misma cardinalidad, resulta de la observación anterior que la existencia de una medida continua sobre $\mathbb{R}$ da origen a una medida continua sobre ω_1 y recíprocamente.

En el siguiente resultado se establece que el ideal de los conjuntos nulos es estable bajo uniones numerables.

Lema 6.6.4. *Si* $m : \mathscr{P}(X) \to [0, +\infty)$ *es una* ***medida continua*** *sobre* X *y si definimos*

$$\mathfrak{I} \;=\; \{A \subseteq X \,:\, m(A) = 0\},$$

entonces $\mathfrak{I}$ *es* ***cerrada bajo uniones numerables****, es decir, si* $(A_n)_{n=1}^{\infty}$ *es una sucesión en* $\mathfrak{I}$*, entonces* $\bigcup_{n=1}^{\infty} A_n \in \mathfrak{I}$.

Prueba. Sea $(A_n)_{n=1}^{\infty}$ una sucesión en $\mathfrak{I}$. Definiendo la sucesión $(B_n)_{n=1}^{\infty}$ en X por

$$B_1 \;=\; A_1 \qquad \text{y} \qquad B_n \;=\; A_n \setminus \left(\bigcup_{j=1}^{n-1} A_j\right) \quad \text{para todo entero } n \geq 2,$$

resulta que ella es disjunta y $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n$. Puesto que $B_n \subseteq A_n$ para todo $n \geq 1$, se tiene que $m(B_n) = 0$ para todo $n \geq 1$ y, entonces, por la propiedad (c) se concluye que

$$m\left(\bigcup_{n=1}^{\infty} A_n\right) \;=\; m\left(\bigcup_{n=1}^{\infty} B_n\right) \;=\; \sum_{n=1}^{\infty} m(B_n) \;=\; 0$$

y termina la prueba. ■

Lema 6.6.5. *Si existe una* ***medida continua*** m *sobre un conjunto* X*, entonces cualquier subcolección disjunta*

$$\mathcal{L} \;\subseteq\; \{m(A) \,:\, A \notin \mathfrak{I}\}$$

es ***a lo más numerable****. En otras palabras, no existe ninguna colección no-numerable y disjunta de conjuntos en* X *donde todos sus elementos poseen medida estrictamente positiva.*

Prueba. Puesto que $m(X) < +\infty$, la familia $\{m(A) : m(A) \in \mathcal{L}\}$ es sumable (véase la Definición 2.1.52, página 114). Se sigue del Corolario 2.1.56 que $\mathcal{L}$ es a lo más numerable. ■

He aquí la conección del Problema de la Medida Generalizada con la Hipótesis del Continuo. Este resultado fue demostrado en el año 1929 por los matemáticos Stefan Banach (1892-1945) y Kazimierz Kuratowski (1896-1980) [8].

Teorema 6.6.6 (Banach-Kuratowski). *Si existe una* ***medida continua*** *sobre* $\mathbb{R}$*, entonces la* Hipótesis del Continuo *es* ***falsa****.*

Prueba. Suponga que existe una medida continua m sobre $\mathbb{R}$. Por la observación anterior, podemos admitir que m está definida sobre $\mathscr{P}(\omega_1)$. Para generar una contradicción, vamos a consentir que la Hipótesis del Continuo es verdadera, es decir, $2^{\aleph_0} = \aleph_1$. Considere la siguiente familia de subconjuntos de ω_1:

$$\mathcal{I} \;=\; \big\{X \subseteq \omega_1 \,:\, m(X) = 0\big\}.$$

Esta familia satisface las siguientes propiedades:

(a_1) $\{\alpha\} \in \mathcal{I}$ *para cualquier* $\alpha \in \omega_1$. (Sigue de la definición de m).

(a_2) $\bigcup_{n=1}^{\infty} X_n \in \mathcal{I}$ *para cualquier sucesión* $\big(X_n\big)_{n=1}^{\infty}$ *con* $X_n \in \mathcal{I}$ *para todo* $n \in \mathbb{N}$. (Por el Lema 6.6.4).

(a_3) $\mathcal{I}^c = \mathscr{P}(\omega_1) \setminus \mathcal{I} = \{A \subseteq X \,:\, m(A) > 0\}$ *es a lo más numerable.* (Es consecuencia del Lema 6.6.5). Por consiguiente, no existe ninguna colección no-numerable y disjunta de subconjuntos de ω_1 de medida positiva.

Ahora construiremos una matriz infinita $(A_{\alpha n})_{\omega_1\times\mathbb{N}_0}$ donde cada $A_{\alpha n} \subseteq \omega_1$ del modo siguiente: para cada $\xi < \omega_1$, existe una función $f_\xi : \omega_0 \to \omega_1$ tal que $\xi \subseteq \operatorname{Rang}(f_\xi)$. Seleccionemos una tal función f_ξ por cada $\xi < \omega_1$ y defina

$$A_{\alpha n} \;=\; \big\{\xi < \omega_1 \,:\, f_\xi(n) = \alpha\big\} \quad \text{para todo } \alpha < \omega_1,\ n < \omega_0.$$

La matriz infinita $(A_{\alpha n})_{\omega_1\times\mathbb{N}_0}$ goza de las siguientes propiedades:

(b_1) $A_{\alpha n} \cap A_{\beta n} = \varnothing$ para todo $\alpha \neq \beta$ en ω_1 y cualquier $n \in \mathbb{N}_0$.

(b_2) $\omega_1 \setminus \bigcup_{n=0}^{\infty} A_{\alpha n}$ es a lo más numerable para cualquier $\alpha < \omega_1$.

La veracidad de (b_1) sigue inmediatamente. En efecto, si $\xi \in A_{\alpha n} \cap A_{\beta\, n}$, entonces $f_\xi(n) = \alpha$ y $f_\xi(n) = \beta$, de donde se tiene que $\alpha = \beta$. Para ver que (b_2) también se cumple, observe que si $\xi \notin A_{\alpha\, n}$ para todo n, entonces $\alpha \notin \operatorname{Rang}(f_\xi)$ y, por lo tanto, $\xi < \alpha$. Esto nos dice que (b_2) es verdadero.

Fijemos un $\alpha < \omega_1$. Afirmamos que existe algún $n_\alpha \in \mathbb{N}_0$ tal que $A_{\alpha n_\alpha} \notin \mathcal{I}$. En efecto, suponga, para generar una contradicción, que $A_{\alpha n} \in \mathcal{I}$ para todo $n \in \mathbb{N}_0$. Como

$$\omega_1 \;=\; \bigg(\omega_1 \setminus \bigcup_{n=0}^{\infty} A_{\alpha n}\bigg) \cup \bigcup_{n=0}^{\infty} A_{\alpha n},$$

resulta de (b_1), (b_2), (a_1) y (a_2) que

$$\begin{aligned} 1 \;=\; m(\omega_1) \;&=\; m\bigg(\omega_1 \setminus \bigcup_{n=0}^{\infty} A_{\alpha n}\bigg) \;+\; m\bigg(\bigcup_{n=0}^{\infty} A_{\alpha n}\bigg) \\ &=\; m\bigg(\bigcup_{n=0}^{\infty} A_{\alpha n}\bigg) \;=\; \sum_{n=0}^{\infty} m(A_{\alpha n}) \\ &=\; 0. \end{aligned}$$

Esta contradicción confirma nuestra afirmación. Ahora bien, como ω_0 es numerable y ω_1 no lo es, resulta del Principio del Palomar $\mathfrak{c}$-Infinito, página 16, que debe existe algún $k \in \mathbb{N}_0$ de modo que el conjunto $\{\alpha \in \omega_1 : n_\alpha = k\}$ es no-numerable. Sea entonces

$$\mathcal{L} \;=\; \big\{A_{\alpha k} \,:\, n_\alpha = k\big\}.$$

Esta familia es, por lo que acabamos de ver, no-numerable. También, gracias a (b_1) ella es disjunta y se cumple, por nuestra afirmación, que $A_{\alpha k} \notin \mathcal{I}$ para cada $A_{\alpha k} \in \mathcal{L}$. Esto, por supuesto, contradice la propiedad (a_3). La conclusión ahora es inequívoca: suponer que la Hipótesis del Continuo es verdadera conduce a una contradicción. Fin de la prueba. ■

En la misma dirección del resultado probado por Banach y Kuratowski, S. Ulam obtuvo la siguiente generalización:

Teorema 6.6.7 (Ulam). *Sea X un conjunto con $\operatorname{card}(X) = \aleph_1$. Si $m : \mathscr{P}(X) \to [0, +\infty)$ es numerablemente aditiva y $m(\{x\}) = 0$ para todo $x \in X$, entonces m es idénticamente cero.*

Prueba. Como $\operatorname{card}(\omega_1) = \aleph_1$, el Principio del Buen-Orden y el Teorema 1.3.35, página 66, nos garantizan que X puede ser bien-ordenado de tal modo que, para cada $y \in X$, el conjunto $\operatorname{Seg}(y) = \{x \in X : x \prec y\}$ es a lo más numerable. Seleccione, por cada $y \in X$, una aplicación inyectiva $T(\cdot, y) : \operatorname{Seg}(y) \to \mathbb{N}$ y observe que, para cada par $(x, y) \in X \times X$ con $x \prec y$, $T(x, y)$ es un número natural. Consideremos ahora, para cada $x \in X$ y cada $n \in \mathbb{N}$, el conjunto

$$A_x^n = \{y \in X : x \prec y,\ T(x, y) = n\}.$$

Afirmamos que, para cada $n \in \mathbb{N}$, la familia $(A_x^n)_{x \in X}$ es disjunta. En efecto, suponga que $x, z \in X$ con $x \neq z$ y que $A_x^n \cap A_z^n \neq \varnothing$. Sea $y \in A_x^n \cap A_z^n$. Esto significa que

$$x \prec y, \quad z \prec y \qquad \text{y} \qquad T(x, y) = n = T(y, z). \tag{1}$$

Como $\preceq$ es un orden total, se tiene que $x \prec z$ o $z \prec x$. Si se asume que $x \prec z$, tendremos, por la inyectividad de $T(\cdot, y)$, que $T(x, y) \neq T(z, y)$ lo cual, gracias a (1), es imposible. El mismo razonamiento se aplica si $z \prec x$. Esto prueba nuestra afirmación. Ahora bien, como m es finita y numerablemente aditiva, resulta que la familia $(m(A_x^n))_{x \in X}$ es sumable para cada $n \in \mathbb{N}$, de donde se obtiene, invocando al Corolario 2.1.56, página 116, que el conjunto

$$X_n = \{x \in X : m(A_x^n) > 0\}$$

es a lo más numerable. Por esto, el conjunto $X_* = \bigcup_{n=1}^{\infty} X_n$ es numerable y como X es no-numerable, resulta que $X \neq X_*$, esto es, existe al menos un $x \in X$ tal que $x \notin X_*$, lo cual significa que $m(A_x^n) = 0$ para todo $n \in \mathbb{N}$. Definiendo $A = \bigcup_{n=1}^{\infty} A_x^n$, vemos que $m(A) = 0$. Para finalizar la prueba sólo nos resta demostrar que $X \setminus A$ está contenido en $\operatorname{Seg}(x)$. En efecto, como $\operatorname{Seg}(x)$ es a lo más numerable, entonces $X \setminus A$ también sería a lo más numerable y, en consecuencia, teniendo en cuenta que $m(\{x\}) = 0$ para todo $x \in X$, tendríamos que $m(X \setminus A) = 0$ y concluiría la prueba. Veamos entonces que $X \setminus A \subseteq \operatorname{Seg}(x)$. Suponga, por contradicción, que algún $y \in X \setminus A$ se encuentra fuera de $\operatorname{Seg}(x)$. Entonces $y \succ x$ y, por lo tanto, $y \in A_x^n$ donde $n = T(x, y)$, es decir, $y \in A$ lo cual es absurdo. La prueba es completa. ■

En particular, si la **Hipótesis del Continuo** es verdadera, entonces $\operatorname{card}([0,1]) = \aleph_1$ y se obtiene el resultado de Banach-Kuratowski.

Corolario 6.6.8 (Banach-Kuratowski). *Bajo la **Hipótesis del Continuo**, si $m : \mathscr{P}([0,1]) \to [0,1]$ es es numerablemente aditiva y $m(\{x\}) = 0$ para todo $x \in X$, entonces $m([0,1]) = 0$.*

6.6.3. El Problema de la Medida de Lebesgue y la Aditividad Finita

El Problema de la Medida de Lebesgue se puede formular, casi sin ningún cambio, en $\mathbb{R}^n$ para cualquier n. En efecto, la única modificación que consideraremos es la ampliación de la noción de invarianza por traslación que se traduce, en este caso, a la invarianza por isometrías, es decir, una medida $m : \mathscr{P}(\mathbb{R}) \to [0, +\infty]$ es **invariante por isometrías** si $m(A) = m(\varphi(A))$ para cualquier isometría $\varphi : \mathbb{R}^n \to \mathbb{R}^n$. Recordemos que el grupo de la isometrías en $\mathbb{R}^n$ está compuesto por todas las transformaciones geométricas formadas por traslaciones, rotaciones y reflexiones que no alteran las distancias de un conjunto. Hemos visto que la numerabilidad aditiva, así como la invarianza por traslación de la medida de Lebesgue μ son los responsables, con la colaboración del Axioma de Elección, de la no existencia de una extensión universal de μ. También, en la sección anterior descubrimos que si se abandona el requerimiento de la invarianza por traslación, entonces la única *extensión continua* que se obtiene sobre $\mathbb{R}$ es la nula siempre que la Hipótesis del Continuo esté presente. Ahora veremos qué ocurre si se omite la numerabilidad aditiva pero se mantiene la invarianza por isometrías. Estas consideraciones nos conducen a analizar las siguientes dos posibilidades: suponer, o bien la aditividad finita, o la κ-aditividad o aditividad no-numerable. Veamos el primer caso, es decir, intentaremos encontrar, en lugar de una medida numerablemente aditiva, una "*medida finitamente aditiva*" definida sobre $\mathscr{P}(\mathbb{R})$ que extienda a μ, con el requerimiento adicional de ser invariante por isometrías, en otras palabras, *¿existe una* ***extensión universal*** *finitamente aditiva e invariante por isometrías de la medida de Lebesgue en* $\mathbb{R}^n$, es decir, *¿se puede determinar la existencia una función de conjuntos* m *que cumpla:*

(β_1) $0 \leq m(A) \leq +\infty$ *para todo* $A \subseteq \mathbb{R}^n$.

(β_2) $m(Q) = \mu(Q)$ *para cualquier paralelepípedo* $Q \subseteq \mathbb{R}^n$.

(β_3) ma *es invariante por isometrías*.

(β_4) ma *es finitamente aditiva*?

Este problema es conocido como **El Problema de la Medida de Hausdorff** y posee dos respuestas que son realmente sorprendentes pues depende de la dimensión n del espacio. En efecto, en 1914, Felix Hausdorff (1869-1942), un astrónomo Judío-Alemán a quien no le incomodaba usar el Axioma de Elección, estudió en gran profundidad dicho problema demostrando que si $n \geq 3$, entonces no existía ninguna extensión universal finitamente aditiva e invariante por isometrías de la medida de Lebesgue en $\mathbb{R}^n$. De hecho, Hausdorff, con la complicidad del Axioma de Elección, demuestra que existe una descomposición de la superficie de una esfera de radio 1 (en $\mathbb{R}^3$) en cuatro conjuntos A, B, C y D con las siguientes propiedades: D es numerable, $A \equiv B \equiv C$ y $A \equiv (B \cup C)$ ($\equiv$ significa que los conjuntos son congruentes entre sí, es decir, existe una biyección isométrica entre ellos). Por consiguiente, si existiese una función de conjuntos λ que satisfaga las condiciones $(\alpha_1), (\alpha_2), (\alpha_3)$ y $(\alpha_4)'$ se tendría que la medida de D sería λ-nula y que

$$m(A) = m(B) = m(C) = \frac{1}{3} \qquad \text{y} \qquad m(A) = m(B \cup C) = \frac{1}{2}$$

lo que conduciría a una contradicción. Esta descomposición será conocida posteriormente, gracias a Borel, como la **Paradoja de Hausdorff**. Sin embargo, a Hausdorff le fue imposible resolver dicho problema para $n = 1$ y $n = 2$.

La investigación de Hausdorff fue continuada por el joven matemático polaco Stefan Banach quien también era, al igual que la mayoría de los matemáticos polacos, partidario de usar el Axioma de Elección sin limitaciones. Banach demostró, en 1923, con la ayuda del Principio del

Buen-Orden, equivalente al Axioma de Elección, que **El Problema de la Medida de Hausdorff** posee solución positiva en los espacios $\mathbb{R}$ y $\mathbb{R}^2$. Esto significa que en $\mathbb{R}$ y $\mathbb{R}^2$ la medida de Lebesgue no es la única medida que es finitamente aditiva, normalizada e invariante por isometrías. Este hermoso resultado de Banach parece ser el único mejoramiento razonable de la medida de Lebesgue. Sin embargo, como afirma K. Cielsielski en [34]: "si queremos mantener la Teoría de la Medida de Lebesgue uniforme para todas las dimensiones, no podemos aceptar la solución de Banach sobre $\mathbb{R}^2$ como la mejor solución al Problema de la Medida"

A causa de esta disparidad del Problema de la Medida de Hausdorff entre los casos $n \in \{1, 2\}$ y $n \geq 3$, Banach se propone investigar en profundidad la prueba de Hausdorff para el último caso. Al mismo tiempo, Alfred Tarski, un alumno de Sierpiński, investiga un problema similar. Banach y Tarski escribieron juntos un artículo [9] demostrando, entre otras cosas que:

> *Es imposible particionar un polígono $P \subseteq \mathbb{R}^2$ en un número finito de piezas y luego reensamblar dichas piezas para obtener otro polígono conteniendo estrictamente a P. Sin embargo, contrario a lo que ocurre en $\mathbb{R}^2$, ellos prueban que si S_1 y S_2 son dos esferas de $\mathbb{R}^3$, entonces es posible descomponer a S_1 en un números finito de piezas y luego reensamblar dichas piezas por medio de traslaciones y rotaciones para obtener exactamente a S_2.*

Esa descomposición, que posteriormente será llamada la **Paradoja de Banach-Tarski**, sólo es posible usando el Axioma de Elección. Es importante destacar que aunque el resultado anterior fue nombrada una paradoja, dicho resultado es, en realidad un TEOREMA de la Teoría de la Medida. A modo de broma algunos matemáticos sugieren que la aplicación de la Paradoja de Banach-Tarski podría resolver todos los problemas monetarios de cualquier país: basta con aplicar dicho resultado a cada moneda para convertirla en dos del mismo valor.

El resultado inesperado de la duplicación de un conjunto en la Paradoja de Banach-Tarski sucede porque la medida es finitamente aditiva. ¿Qué ocurre si removemos esta condición?, es decir, ¿si eliminamos la aditividad finita garantizamos que paradojas como la Paradoja de Banach-Tarski no ocurrirán?

Los resultados de Hausdorff, Banach y Tarski atrajeron la atención de John von Neumann (1903-1957) quien, en 1929, arrojó luz a la Paradoja de Banach-Tarski y, al mismo tiempo, propuso una nueva generalización del Problema de la Medida de Hausdorff. En efecto, von Neumann demuestra que la dicotomía en el Problema de la Medida de Hausdorff no se debía a la naturaleza de la recta, el plano o el espacio Euclidiano, es decir, no dependía de la dimensión del espacio sino del grupo de isometrías utilizado. Para ello, él hizo lo siguiente: tomó un conjunto arbitrario X, M un subconjunto cualquiera de X y $\mathfrak{G}$ cualquier grupo de aplicaciones biyectivas de X sobre sí mismo. Entonces definió lo siguiente:

*m es una $(X, M, \mathfrak{G})$-**medida** si cumple con las siguientes condiciones:*

(a) $m(M) = 1$.

(b) $m(E) = m(\varphi(E))$ *para cualquier* $\varphi \in \mathfrak{G}$ y $E \subseteq X$.

(c) $m(E \cup F) = m(E) + m(F)$ si E, F *son disjuntos*.

Por supuesto, este nuevo enfoque alteró el Problema de la Medida de Hausdorff en dos aspectos fundamentales: en primer lugar, se reemplazó el espacio $\mathbb{R}^n$ por un conjunto arbitrario y, en segundo lugar, se consideró un grupo arbitrario en lugar del grupo de las isometrías de $\mathbb{R}^n$. von Neumann entonces demuestra una condición suficiente para la existencia de una $(X, M, \mathfrak{G})$-medida y prueba que esta condición se satisface para el grupo de isometrías en el caso $n = 1, 2$, pero no para $n \geq 3$. Similarmente, él demuestra que si $\mathfrak{G}$ es un grupo libre con dos generadores y sin puntos fijos, como es el caso del grupo de isometrías de $\mathbb{R}^n$ con $n \geq 3$, entonces una tal $(X, M, \mathfrak{G})$-medida no existe.

En el mismo año 1929, pero un poco después del trabajo de von Neumann, Banach generalizó el Problema de la Medida de Hausdorff en una dirección totalmente diferente. Su nuevo punto de vista colocó el problema en el marco de la Teoría de Conjuntos en lugar del punto de vista geométrico o, aun, de la Teoría de Grupos de von Neumann.

La aditividad no-numerable o κ-aditividad entra en el juego del Problema de la Medida del modo siguiente: Asumiendo la Hipótesis del Continuo Generalizada, Banach pudo demostrar, en 1930, que si existe una medida continua no-trivial sobre X, entonces existe un **cardinal inaccesible** $\kappa \leq \operatorname{card}(X)$. Para lograr esto, Banach se ve en la necesidad de promover dos cambios al resultado anteriormente demostrado por él y Kuratowski: el primer cambio consistió en reemplazar la Hipótesis del Continuo por la Hipótesis del Continuo Generalizada y, el segundo, la condición de aditividad numerable la reemplaza por la de κ-aditividad o adtividad no-numerable.

Definición 6.6.9. *Sean X un conjunto no-vacío X y κ un cardinal no-numerable. Una función de conjuntos $m : \mathscr{P}(X) \to [0, +\infty)$ se llama κ-**aditiva** si para cualquier cardinal $\lambda < \kappa$ y cualquier colección disjunta de subconjuntos de X, $(A_\alpha)_{\alpha<\lambda}$, se cumple que*

$$m\Big(\bigcup_{\alpha<\lambda} A_\alpha\Big) = \sum_{\alpha<\lambda} m(A_\alpha).$$

Observe que como $m(X) < +\infty$, entonces la familia $\{m(A_\alpha) : \alpha < \lambda\}$ es sumable y, en consecuencia, la colección $\{\alpha : m(A_\alpha) > 0\}$ es a lo más numerable. Nótese que la $\aleph_1$-aditividad es de hecho la numerabilidad aditiva.

Con base en esas hipótesis y algunas nuevas definiciones, Banach demuestra que:

Teorema 6.6.10. *Sea κ un cardinal no-numerable. Si sobre un conjunto X se puede definir una medida $m : \mathscr{P}(X) \to [0, +\infty)$ κ-**aditiva**, entonces κ es un **cardinal inaccesible**.*

Como corolario a este resultado, Banach demuestra que si κ es *el primer cardinal inaccesible*, $\kappa > \aleph_0$ y el número cardinal de un conjunto X es igual a $\lambda < \kappa$, entonces, si existe una medida sobre X, ella sólo puede ser finitamente aditiva. De este modo concluye que el *continuo es demasiado pequeño* para aceptar una medida numerablemente aditiva. Algunas referencias para ampliar estas notas son, por ejemplo, [100, 101, 137].

6.6.4. El Problema de la Medida de Lebesgue y el Axioma de Determinación

Sin el Axioma de Elección existe otra vía de resolver el Problema de la Medida. En efecto, el Axioma de Determinancia es quien ofrece esa solución. Esta otra dirección de la investigación del Problema de la Medida de Lebesgue consiste en añadir al sistema **ZF** un nuevo axioma conocido como el Axioma de Determinación, **AD**, *el cual estipula que ciertos juegos infinitos con información perfecta jugados entre dos personas siempre son determinados; es decir, uno de los jugadores posee una estrategia ganadora*. En 1962, J. Mycielski y H. Steinhauss introdujeron ese axioma. Aunque los autores no afirman que este nuevo axioma sea “intuitivamente verdadero”, ellos consideran que dicho axioma es muy interesante pero hacen notar que éste posee un problema de consistencia, es decir, la cuestión de si **ZF** + **AD** es consistente si, y sólo si, es consistente **ZF**, es aun un problema por resolver. La investigación en profundidad de las consecuencias del Axioma de

Determinación es, tal vez, el desarrollo más distintivo y fascinante de la Teoría Moderna de Conjuntos, incluyendo las profundas y sorprendentes conexiones con los grandes cardinales. Una de las magníficas consecuencias del Axioma de Determinación es que el Problema de la Medida de Lebesgue posee una solución positiva. Esto hace que el Axioma de Determinación sea altamente deseable.

‖ ► **El Juego G**(X, A).

Sean X un conjunto no-vacío y A un subconjunto de $X^{\mathbb{N}}$. El juego infinito $\mathrm{G}(X, A)$ se lleva a cabo del modo siguiente: dos jugadores α y β eligen, alternativamente, elementos $x_0, x_1, x_2, \ldots$ en el conjunto X. El jugador α es quien inicia la partida escogiendo sucesivamente la subsucesión $x_0, x_2, x_4, \ldots$, mientras que el jugado β elige la .subsucesión $x_1, x_3, x_5, \ldots$ Al jugador α se le declara **ganador** del juego si la sucesión $(x_n)_{n=1}^{\infty}$ pertenece al conjunto A. En caso contrario, el ganador es el jugador β. Este juego es lo que se llama un **juego con información perfecta** ya que cada jugador conoce cuál ha sido la elección de su oponente en cualquier paso anterior en el juego.

El juego $\mathrm{G}(X, A)$ *se llama* ***determinado*** *si uno de los jugadores posee una* ***estrategia ganadora***.

Una *estrategia* para uno de los dos jugadores en el juego $\mathrm{G}(X, A)$ es una función que indica al jugador el elemento x_n que tiene que elegir en el n-ésimo paso, tomando como referencia la sucesión $x_0, x_1, \ldots, x_{n-1}$ de elementos elegidos hasta ese momento. Una *estrategia es ganadora* si jugando de acuerdo con la estrategia se gana el juego, sin importar cuál es la elección de su oponente en cada paso.

‖ ► **El Axioma de Determinación**: **G**$(\mathbb{N}, A)$ **es determinado para cualquier** $A \subseteq \mathbb{N}^{\mathbb{N}}$.

Lo primero que debemos advertir es que el Axioma de Determinación *es inconsistente con el* Axioma de Elección *en el sentido de que en* **ZFC** *existe un juego que no es determinado*. De hecho, si $\mathbf{B}$ es un conjunto de Bernstein, entonces $\mathrm{G}(\mathbb{R}, \mathbf{B})$ no es determinado (véase, por ejemplo, [41], p. 408). Mejor aun, los resultados que se muestran a continuación así lo confirma.

Mencionaremos ahora varios hechos que son sorprendentes e increíbles cuando se acepta el Axioma de Determinación (véase, por ejemplo, [74], pag. 627-630 y [71], pag. 150-156):

Bajo el **Axioma de Determinación** se cumple que:

(1) *Cualquier subconjunto de* $\mathbb{R}$ *es medible Lebesgue, es decir,* $\mathscr{P}(\mathbb{R}) = \mathfrak{M}_{\mu}(\mathbb{R})$.

(2) *La medida de Lebesgue es numerablemente aditiva sobre* $\mathscr{P}(\mathbb{R})$.

(3) *De* (1) *y* (2) *se concluye que* El Problema de la Medida de Lebesgue *en* $\mathbb{R}$ *tiene solución positiva*.

(4) *No existen descomposiciones paradójicas como la de Banach-Tarski*.

(4) *No existen ultrafiltros libres sobre* $\mathbb{N}$.

(5) *Todas las soluciones de la ecuación de Cauchy* $f(x+y) = f(x) + f(y)$ *son continuas*.

(6) *El espacio vectorial* $\mathbb{R}_{\mathbb{Q}}$ *no posee base de Hamel de dimensión infinita*.

(7) $\aleph_1$ *es un cardinal medible. En particular, inaccesible*.

El último comentario que queremos ofrecer es que en el desarrollo axiomático de la Teoría de la Probabilidad iniciada por A. Kolmogorov en 1933, se estableció que la *numerabilidad aditiva de la medida* era de importancia fundamental para tal desarrollo de modo que, en ese contexto, se requiere de la numerabilidad aditiva de la medida aun a expensa de que no se pueda "medir" a todos los subconjuntos de $\mathbb{R}$. En general, la numerabilidad aditiva es, matemáticamente, una condición muy poderosa; sin embargo, la condición finitamente aditiva es un concepto mucho más manejable y menos restrictivo.

6.7. Ejemplos y Contraejemplos Usando la Función de Cantor

Esta corta sección es relevante pues nos muestra algunos extraordinarios e increíbles contraejemplos usando la función de Cantor.

Ejercicio 6.7.1. *Sea φ_Γ la **función de Cantor** sobre $[0,1]$ y defina $\Phi_\Gamma : [0,1] \to [0,2]$ por*

$$\Phi_\Gamma(x) \;=\; x + \varphi_\Gamma(x)$$

para todo $x \in [0,1]$. Entonces:

(a) *Φ_Γ es un **homeomorfismo**.*

Prueba. Puesto que $\Phi_\Gamma(0) = 0$ y $\Phi_\Gamma(1) = 2$ se sigue de la Propiedad del Valor Intermedio que para cualquier $y \in (0,2)$, existe un $x \in (0,1)$ tal que $y = \Phi_\Gamma(x)$. Esto prueba que $\Phi_\Gamma([0,1]) = [0,2]$. Por otro lado, como φ_Γ es creciente y la función identidad $\mathrm{Id}(x) = x$ es estrictamente creciente, resulta que Φ_Γ también es estrictamente creciente. En particular, Φ_Γ es una función continua y biyectiva. Se sigue del Teorema 3.1.18 que Φ_Γ es un homeomorfismo. ■

(b) *Φ_Γ **no preserva la medida de conjuntos medibles**, es decir, la igualdad*

$$\mu\big(\Phi_\Gamma(E)\big) \;=\; \mu(E)$$

no se cumple para todo $E \in \mathfrak{M}_\mu(\mathbb{R})$.

Prueba. Nuestra tarea es demostrar que $\mu\big(\Phi_\Gamma(\Gamma)\big) = 1$. Para verificar esta afirmación, sea $\big\{J_n(k) : n \in \mathbb{N},\ k = 1, \ldots, 2^n - 1\big\}$ el conjunto de todos los intervalos abiertos borrados en la construcción de Γ y suponga que $(J_n)_{n=1}^\infty$ es una lista de tales intervalos. Sabemos que dicha colección es disjunta. Sea $x_n = \varphi_\Gamma(J_n)$ y note que

$$\Phi_\Gamma(J_n) \;=\; (\mathrm{Id} + \varphi_\Gamma)\big(J_n\big) \;=\; J_n + x_n.$$

Además, la sucesión $(J_n + x_n)_{n=1}^\infty$ es disjunta. Sea $G = \bigcup_{n=1}^\infty J_n$. Entonces

$$\Phi_\Gamma(G) \;=\; (\mathrm{Id} + \varphi_\Gamma)\bigg(\bigcup_{n=1}^\infty J_n\bigg) \;=\; \bigcup_{n=1}^\infty (J_n + x_n) \;\in\; \mathfrak{M}_\mu(\mathbb{R})$$

y así, como μ es invariante por traslación y numerablemente aditiva,

$$\mu\big(\Phi_\Gamma(G)\big) \;=\; \sum_{n=1}^\infty \mu\big(J_n + x_n\big) \;=\; \sum_{n=1}^\infty \mu\big(J_n\big) \;=\; \mu(G) \;=\; 1.$$

Por otro lado, teniendo en cuenta que

$$\Phi_\Gamma(\Gamma) \;=\; \Phi_\Gamma([0,1] \setminus G) \;=\; [0,2] \setminus \Phi_\Gamma(G) \;\in\; \mathfrak{M}_\mu(\mathbb{R})$$

resulta entonces que

$$\begin{aligned} \mu\big(\Phi_\Gamma(\Gamma)\big) &= \mu\big([0,2]\big) - \mu\big(\Phi_\Gamma(G)\big) = 2 - 1 = 1 \\ &\neq 0 = \mu(\Gamma) \end{aligned}$$

Esto termina la prueba. ■

En particular, Φ_Γ *es un ejemplo de un* ***homeomorfismo que transforma un conjunto de medida cero en un conjunto de medida positiva****.*

(c) Φ_Γ ***no preserva conjuntos medibles según Lebesgue****, es decir, la implicación*

$$E \in \mathfrak{M}_\mu(\mathbb{R}) \quad \Rightarrow \quad \Phi_\Gamma(E) \in \mathfrak{M}_\mu(\mathbb{R})$$

no siempre se cumple.

Prueba. Como $\Phi_\Gamma(\Gamma) > 0$, el Teorema 6.5.3 nos garantiza la existencia de un subconjunto no-medible $A^* \subseteq \Phi_\Gamma(\Gamma)$. Por ser Φ_Γ biyectiva,

$$\Phi_\Gamma^{-1}(A^*) \subseteq \Phi_\Gamma^{-1}\big(\Phi_\Gamma(\Gamma)\big) = \Gamma$$

y puesto que $\mu(\Gamma) = 0$, resulta que todos sus subconjuntos son medibles, en particular, $\Phi_\Gamma^{-1}(A^*)$ es **medible**. Finalmente, usando de nuevo el hecho de que Φ_Γ es biyectiva, tenemos que $\Phi_\Gamma(E) = A^*$, donde hemos puesto $E = \Phi_\Gamma^{-1}(A^*)$. ■

En particular, Φ_Γ *es un ejemplo de un* ***homeomorfismo que transforma un conjunto medible Lebesgue en un conjunto no-medible****.*

(d) ***Existe conjunto medible Lebesgue que no es medible según Borel****. De modo más preciso: si A^* es un* ***subconjunto no-medible*** *en $\Phi_\Gamma(\Gamma)$, entonces*

$$\Phi_\Gamma^{-1}(A^*) \in \mathfrak{M}_\mu(\mathbb{R}) \setminus \mathfrak{Bo}(\mathbb{R}).$$

Prueba. Recordemos que $\Phi_\Gamma^{-1}(A^*) \subseteq \Gamma$ y, en consecuencia, medible según Lebesgue. Por otro lado, como todo homeomorfismo transforma conjuntos de Borel en conjuntos de Borel, Lema 6.3.22, página 272, resulta que si $\Phi_\Gamma^{-1}(A^*)$ fuese de Borel, tendríamos que $A^* = \Phi_\Gamma(\Phi_\Gamma^{-1}(A^*))$ sería un conjunto de Borel, en particular, medible según Lebesgue lo cual es imposible por nuestra hipótesis. Por esto, el conjunto $E = \Phi_\Gamma^{-1}(A^*)$ no es un conjunto de Borel. ■

En particular, Γ ***contiene un subconjunto que no es medible según Borel****.* Véase también el Ejemplo (4) de la página 337. Esta observación no debería ser una sorpresa. En efecto, del hecho de que

$$\mathfrak{M}_\mu(\mathbb{R}) = \mathfrak{Bo}(\mathbb{R}) \cup \mathfrak{N}_\mu(\mathbb{R}),$$

resulta que para localizar conjuntos medibles según Lebesgue que no son borelianos, sólo hay que buscarlos en conjuntos de medida cero.

(e) *μ_0, la **medida de Borel**, no es completa.*

Prueba. De (c) y (d) sabemos que existe un conjunto no-medible según Lebesgue $A^* \subseteq \Phi_\Gamma(\Gamma)$ tal que

(a) $\Phi_\Gamma^{-1}(A^*) \subseteq \Gamma$, y

(b) $\Phi_\Gamma^{-1}(A^*) \in \mathfrak{M}_\mu(\mathbb{R}) \setminus \mathfrak{Bo}(\mathbb{R})$.

Puesto que $\mu_0(\Gamma) = 0$ y $\Phi_\Gamma^{-1}(A^*)$ no es un boreliano, resulta que $\mu_0(\Phi_\Gamma^{-1}(A^*))$ no está definido y, en consecuencia, no puede satisfacer la condición $\mu_0(\Phi_\Gamma^{-1}(A^*)) = 0$. ■

Sea $f : [a, b] \to \mathbb{R}$ una **función continua** y denote por G_f *el conjunto de todos los **puntos de constancia** de f*, es decir, *$x \in G_f$ si, y sólo si, f es constante en un entorno abierto de x*. Por ejemplo, si $f = \varphi_\Gamma$, se tiene entonces que $G_f = [0, 1] \setminus \Gamma$. Observe que G_f es siempre un **conjunto abierto** en la topología relativa de $[a, b]$.

Teorema 6.7.1. *Sea $f : [a, b] \to \mathbb{R}$ una **función continua** y **monótona**. Las siguientes condiciones son equivalentes:*

(1) $\mu([a, b] \setminus G_f) = 0$.

(2) $f^{-1}(A) \in \mathfrak{M}_\mu([a, b])$ *para todo conjunto $A \subseteq \mathbb{R}$.*

Prueba. (1) $\Rightarrow$ (2). Siendo G_f un conjunto abierto (en la topología relativa de $[a, b]$), existe una sucesión disjunta $(I_n)_{n=1}^\infty$ de intervalos abiertos tal que $G_f = \bigcup_{n=1}^\infty I_n$ y f es constante en cada intervalo I_n, $n = 1, 2, \ldots$ Sea E el conjunto de los puntos extremos de todos los intervalos I_n. Considere ahora las restricciones de f a los conjuntos $G_f \cup E$ y $[a, b] \setminus (G_f \cup E)$ y llamémoslas, respectivamente, f_0 y f_1; es decir,

$$f_0 \;=\; f|_{G_f \cup E} \qquad\qquad \text{y} \qquad\qquad f_1 \;=\; f|_{[a,b] \setminus (G_f \cup E)}.$$

Nótese que

$$f \;=\; f_0 + f_1 \qquad\qquad \text{y} \qquad\qquad G_f \cup E = \bigcup_{n=1}^\infty \overline{I}_n.$$

Si A es un subconjunto arbitrario de $\mathbb{R}$, resulta que $f_0^{-1}(A)$ es un F_σ incluido en $[a, b]$ y

$$f_1^{-1}(A) \;\subseteq\; [a, b] \setminus G_f.$$

Puesto que $f^{-1}(A) = f_0^{-1}(A) + f_1^{-1}(A)$, entonces la igualdad $\mu([a, b] \setminus G_f) = 0$ implica que $f_1^{-1}(A)$ es medible y, en consecuencia, $f^{-1}(A)$ es medible según Lebesgue.

(2) $\Rightarrow$ (1). El hecho de f es monótona nos indica que f_1 es inyectiva y

$$f_0(G_f \cup E) \;\cap\; f_1([a,\, b] \setminus (G_f \cup E)) \;=\; \varnothing. \tag{$*$}$$

Suponga que $\mu([a, b] \setminus G_f) > 0$. En particular, como E es numerable, se tiene que

$$\mu([a, b] \setminus (G_f \cup E)) > 0.$$

Una llamado al Teorema 6.5.3, página 315, nos garantiza la existencia de un conjunto no-medible $V \subseteq [a,b] \setminus (G_f \cup E)$. Puesto que f_1 es inyectiva, la igualdad $(*)$ nos revela que $f^{-1}(f(V)) = V$. Esto, por supuesto, contradice nuestra hipótesis (2). ■

Puesto que $[0,1] \setminus G_{\varphi_\Gamma} = \Gamma$, se tiene que:

(f) $\varphi_\Gamma^{-1}(A)$ ***es medible según Lebesgue en*** $[0,1]$ ***para cualquier conjunto*** $A \subseteq \mathbb{R}$.

Recordemos, Lema 6.3.22, página 272, que toda función continua e inyectiva transforma conjuntos de Borel en conjuntos de Borel. La función de Cantor, aunque es continua pero no inyectiva, posee la misma propiedad:

(g) $\varphi_\Gamma(B) \in \mathfrak{Bo}(\mathbb{R})$ ***para todo conjunto de Borel*** $B \subseteq [0,1]$.

Prueba. Sea B un conjunto de Borel incluido en $[0,1]$. Entonces

$$\varphi_\Gamma(B) \;=\; \varphi_\Gamma\big(B \cap G_{\varphi_\Gamma}\big) \;\cup\; \varphi_\Gamma\big(B \cap E\big) \;\cup\; \varphi_\Gamma\big(B \setminus (G_{\varphi_\Gamma} \cup E)\big),$$

donde E es el conjunto de los puntos extremos de las componente de G_{φ_Γ}. Los conjuntos $\varphi_\Gamma\big(B \cap G_{\varphi_\Gamma}\big)$ y $\varphi_\Gamma\big(B \cap E\big)$ son a lo más numerables. Por otro lado, como la restricción de φ_Γ al conjunto $[0,1] \setminus (G_{\varphi_\Gamma} \cup E)$ es un homeomorfismo, resulta que $\varphi_\Gamma\big(B \setminus (G_{\varphi_\Gamma} \cup E)\big)$ es un conjunto de Borel y, en consecuencia, $\varphi_\Gamma(B)$ es un conjunto de Borel. ■

Para cada función $f : [a,b] \to \mathbb{R}$ defina

$$\mathrm{Ra}(f) \;=\; \big\{y \in \mathbb{R} \,:\, \mathrm{card}\big(f^{-1}(\{y\}) \;=\; +\infty\big)\big\}.$$

Una función $f : [a,b] \to \mathbb{R}$ se dice que satisface la **condición de Banach** si

$$\mu\big(\mathrm{Ra}(f)\big) \;=\; 0. \tag{1}$$

(h) φ_Γ *satisface la* ***condición de Banach***.

Prueba. Para cada $y \in \mathrm{Ra}(\varphi_\Gamma)$, sean $x'_y = \inf \varphi_\Gamma^{-1}(\{y\})$ y $x''_y = \sup \varphi_\Gamma^{-1}(\{y\})$. Como φ_Γ es creciente y constante sobre cada intervalo abierto de $G = [0,1] \setminus \Gamma$, resulta que $\varnothing \neq (x'_y, x''_y) \subseteq \varphi_\Gamma^{-1}(\{y\})$, de donde se obtiene que

$$(x'_{y_1}, x''_{y_1}) \cap (x'_{y_2}, x''_{y_2}) \;=\; \varnothing \quad \text{siempre que} \quad y_1 \neq y_2.$$

Puesto que cualquier colección disjunta de intervalos abiertos es a lo más numerable, se tiene que $\mathrm{Ra}(\varphi_\Gamma)$ es a lo más numerable. De esto se sigue que (1) se cumple. ■

(i) φ_Γ *es* ***uniformemente continua****, pero no es* ***Lipschitz*** *(*Véase el Corolario 9.1.10, página 460*).*

(j) φ_Γ *no es* ***absolutamente continua*** *(*Véase el Ejemplo 9.1.8, página 512*).*

6.8. Ejercicios Resueltos

Esta sección está dedicada, como su nombre lo indica, a desarrollar ciertos resultados complementarios que pueden ser de algún interés al lector.

Ejercicio 6.8.1. *Sea E un* ***subconjunto cerrado*** *de $[0,1]$ con $\mu(E)=1$. Pruebe que $E=[0,1]$.*

Prueba. Suponga que $E \neq [0,1]$ y sea $V=[0,1]\setminus E$. Entonces V es un abierto no vacío y, por consiguiente, $\mu(V)>0$. Puesto que $[0,1]=E\cup([0,1]\setminus E)=E\cup V$, resulta que $1=\mu([0,1])=\mu(E)+\mu(V)>1$. Esta contradicción establece que $E=[0,1]$. ■

Ejercicio 6.8.2. *Sea $D\subseteq[0,1]$* ***medible*** *con $\mu(D)=1$. Demuestre que D es denso en $[0,1]$. En particular, si D es cerrado con $\mu(D)=1$, entonces $D=[0,1]$.*

Prueba. Suponga que D no es denso en $[0,1]$. Esto significa que existe un conjunto abierto no vacío $G\subseteq[0,1]$ tal que $D\cap G=\varnothing$. Se sigue entonces de la aditividad de μ que

$$1 \ \geq\ \mu(D\cup G)\ =\ \mu(D)\ +\ \mu(G)\ =\ 1\ +\ \mu(G)\ >\ 1$$

pues $\mu(G)>0$. Esta contradicción establece que D es denso en $[0,1]$. ■

Ejercicio 6.8.3. *Sea $E\in\mathfrak{M}_\mu(\mathbb{R})$ tal que $\mu(E)=0$. Pruebe que $\operatorname{int}(E)=\varnothing$.*

Prueba. Claramente $\mu(\operatorname{int}(E))=0$. Si suponemos que $\operatorname{int}(E)\neq\varnothing$, entonces existe un intervalo abierto $J\subseteq\operatorname{int}(E)$ y por lo tanto $0=\mu(\operatorname{int}(E))\geq\mu(J)>0$. Esta contradicción establece que $\operatorname{int}(E)=\varnothing$. ■

Ejercicio 6.8.4. *Sea $E\subseteq[0,1]$ tal que $\mu(\operatorname{int}(E))=\mu(\overline{E})$. Pruebe que $E\in\mathfrak{M}_\mu(\mathbb{R})$.*

Prueba. Puesto $\partial E=\overline{E}\setminus\operatorname{int}(E)$, entonces $\mu(\partial E)=\mu(\overline{E})-\mu(\operatorname{int}(E))=0$. Por otro lado, como $E\setminus\operatorname{int}(E)\subseteq$, resulta que $E\setminus\operatorname{int}(E)\subseteq\partial E$ es medible y, así, $E=\operatorname{int}(E)\cup(E\setminus\operatorname{int}(E))$ es medible ya que $\operatorname{int}(E)$ es un conjunto abierto. ■

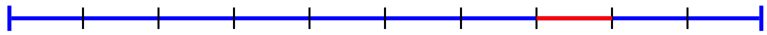
Ejercicio 6.8.5. *Sea $E\in\mathfrak{M}_\mu(\mathbb{R})$ tal que $\mu(E)>0$. Pruebe que existen $x,y\in E$ tal que $|x-y|$ es irracional.*

Prueba. Si para todo $x,y\in E$ ocurre que $|x-y|$ es racional, entonces $E\subseteq x+\mathbb{Q}$ para cualquier $x\in E$, de donde se sigue que E es numerable y, por lo tanto, $\mu(E)=0$.

Ejercicio 6.8.6. *Sea E el conjunto de todos los puntos en $[0,1]$ cuya representación decimal no usa el dígito 7. Pruebe que $\mu(E)=0$.*

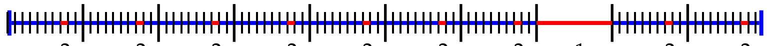
Prueba. El conjunto E se puede construir, geométricamente, del siguiente modo: comience dividiendo el intervalo $[0,1]$ en 10 subintervalos cerrados no-superpuestos de igual longitud, digamos $J_1=\{I_0^1, I_1^1,\ldots,I_9^1\}$ y elimine de éste conjunto el octavo intervalo semi-cerrado $I_7^1=[7/10,8/10)$. Observe que

$$I_7^1\cap E\ =\ \varnothing \qquad\qquad \text{y} \qquad\qquad \mu(I_7^1)\ =\ \frac{1}{10}.$$



En el siguiente paso, subdivida cada uno de los intervalos en $J_1 \setminus \{I_7^1\}$ en 10 partes iguales y denotemos a estos 90 intervalos por J_2, es decir,

$$J_2 = \{I_{00}^2, \ldots, I_{09}^2, I_{10}^2, \ldots, I_{19}^2, \ldots, I_{90}^2, \ldots, I_{99}^2\}$$



Como antes, elimine de J_2 los 9 intervalos semi-cerrados I_{k7}^2, $k \in \{0, 1, \ldots, 9\}$, $k \neq 7$, y observe de nuevo que

$$\left(\bigcup_{k=0,\, k\neq 7}^{9} I_{k7}^2 \right) \cap E = \varnothing \qquad \text{y} \qquad \mu\left(\bigcup_{k=0,\, k\neq 7}^{9} I_{k7}^2 \right) = 9 \cdot \frac{1}{10^2}$$

En el n-ésimo paso, subdivida cada uno de los 9^{n-1} intervalos restantes en 10 subintervalos de igual longitud y remueva el octavo intervalo semi-cerrado de cada uno de ellos. La unión de todos estos intervalos borrados posee longitud $9^{n-1}/10^n$. La intersección de todos los intervalos que permanecen en cada subdivisión es el conjunto E. Luego,

$$\begin{aligned} \mu(E) &= 1 - \sum_{n=1}^{\infty} \frac{9^{n-1}}{10^n} = 1 - \frac{1}{9} \sum_{n=1}^{\infty} \left(\frac{9}{10} \right)^n \\ &= 1 - \frac{1}{9} \cdot \frac{9/10}{1 - 9/10} = 1 - \left(\frac{1}{9} \cdot 9 \right) = 0. \end{aligned}$$

■

Ejercicio 6.8.7. *Sea E el conjunto de todos los puntos en $[0,1]$ cuya representación binaria posee un cero en todas las posiciones pares. Pruebe que $\mu(E) = 0$.*

Prueba. Para cada $n \in \mathbb{N}$, sea

$$E_n = \{ x \in [0,1] : x = (0, a_1\, 0\, a_3\, 0\, a_5\, 0\, a_7 \ldots 0\, a_{2n+1}\, a_{2n+2} \ldots)_2 \}$$

y observe que

$$\begin{aligned} E_1 &= \left[0, \frac{1}{4}\right] \cup \left[\frac{2}{4}, \frac{3}{4}\right] \\ E_2 &= \left[0, \frac{1}{4^2}\right] \cup \left[\frac{2}{4^2}, \frac{3}{4^2}\right] \cup \left[\frac{8}{4^2}, \frac{9}{4^2}\right] \cup \left[\frac{10}{4^2}, \frac{11}{4^2}\right] \\ &\vdots \\ E_n &= \bigcup_{k=1}^{2^n} J_k, \quad \text{donde } \mu(J_k) = \frac{1}{4^n} \text{ para } k = 1, \ldots, 2^n. \end{aligned}$$

Por lo tanto, para todo $n \geq 1$ se cumple que

$$\mu(E_n) \;=\; \sum_{k=1}^{2^n} \mu(J_k) \;=\; \frac{2^n}{4^n} \;=\; \frac{1}{2^n}.$$

Además, $E_1 \supseteq E_2 \supseteq E_3 \supseteq \cdots$ y $E = \bigcap_{n=1}^{\infty} E_n$. Puesto que $\mu(E_1) < \infty$, resulta, por una aplicación del Teorema 6.3.47, que

$$\mu(E) \;=\; \lim_{n\to\infty} \mu(E_n) \;=\; \lim_{n\to\infty} \frac{1}{2^n} \;=\; 0.$$

Ejercicio 6.8.8. *Sea $E \in \mathfrak{M}_\mu(\mathbb{R})$ con $\mu(E) > 0$. Diremos que $x \in \mathbb{R}$ es un* ***punto de Lebesgue*** *de E si $\mu(E \cap J_x) > 0$ para todo intervalo abierto J_x conteniendo a x. Sea*

$$E_+ \;=\; \big\{x \in \mathbb{R} \,:\, x \text{ es un punto de Lebesgue de } E\big\}.$$

Pruebe que:

(1) *E_+ es perfecto.*

(2) $\mu(E \setminus E_+) = 0$.

Prueba. (1) En primer lugar vamos a demostrar que E_+ es cerrado. En efecto, sea x un punto de acumulación de E_+. Para comprobar que $x \in E_+$ debemos demostrar que $\mu(E \cap J_x) > 0$ para cualquier intervalo abierto J_x conteniendo a x. Sea entonces J_x un intervalo abierto conteniendo a x. Como x un punto de acumulación de E_+ resulta, por definición, que J_x contiene puntos de E_+ distintos de x. Sea $x_0 \in J_x \cap E_+$ tal que $x_0 \neq x$. Puesto que $x_0 \in E_+$ y J_x es un intervalo abierto conteniendo a x_0 se tiene que $\mu(E \cap J_x) > 0$, lo cual prueba que $x \in E_+$. Se sigue entonces del Corolario 2.2.13, que E_+ es cerrado.

Veamos ahora que $E'_+ = E_+$. Como E_+ es cerrado, el Corolario 2.2.13, nos dice que $E'_+ \subseteq E_+$. Sea ahora $x \in E_+$. Necesitamos demostrar que $x \in E'_+$, es decir, que cualquier entorno de x contiene puntos de E_+ distintos de x. Seleccione entonces un intervalo abierto arbitrario J_x conteniendo a x y veamos que existe al menos un $z_0 \in (E \cap J_x) \setminus \{x\}$ tal que $z_0 \in E_+$. En efecto, suponga, para construir una contradicción, que cada uno de los puntos $z \in (E \cap J_x) \setminus \{x\}$ no pertenece a E_+. Observe que $z \notin E_+$, significa que existe un intervalo abierto I_z conteniendo a z tal que $\mu(E \cap I_z) = 0$. Sea

$$G \;=\; \bigcup_{z\in(E\cap J_x)\setminus\{x\}} I_z.$$

Entonces G es un conjunto abierto el cual se puede escribir como una unión numerable y disjunta de intervalos de la colección $\{I_z : z \in (E \cap J_x) \setminus \{x\}\}$, véase el Teorema 2.2.4, página 122. Escriba entonces a G como

$$G \;=\; \bigcup_{n=1}^{\infty} I_{z_n}$$

y observe que $x \notin G$ y $\mu(E \cap I_{z_n}) = 0$ para todo $n \geq 1$. Más aun,

$$\mu(E \cap G) \;=\; \sum_{n=1}^{\infty} \mu(E \cap I_{z_n}) \;=\; 0.$$

Ahora bien, como $x \in E_+$ resulta, por definición, que $\mu(E \cap J_x) > 0$ y como además $E \cap J_x = E \cap (G \cup \{x\})$, se tiene entonces que

$$0 < \mu(E \cap J_x) = \mu\big(E \cap (G \cup \{x\})\big) = \mu(E \cap G) + \mu(\{x\}) = 0.$$

Esta contradicción establece que existe al menos un $z_0 \in (E \cap J_x) \setminus \{x\}$ tal que $z_0 \in E_+$ y, en consecuencia, $x \in E'_+$.

(2) Observe que, para cada $x \in E \setminus E_+$ existe un intervalo abierto J_x conteniendo a x tal que $\mu(E \cap J_x) = 0$. Sea

$$G = \bigcup_{x \in E \setminus E_+} J_x.$$

Entonces G es abierto el cual, como sabemos, se puede escribir como una unión numerable y disjunta de intervalos de la colección $\{J_x : x \in E \setminus E_+\}$, digamos $G = \bigcup_{n=1}^{\infty} J_{x_n}$. De aquí se sigue que

$$\mu(E \cap G) = \sum_{n=1}^{\infty} \mu(E \cap J_{x_n}) = 0,$$

y como $E \setminus E_+ \subseteq E \cap G$, resulta entonces que $\mu(E \setminus E_+) = 0$. ■

Si tomamos $E = \mathbb{Q}$, resulta que E es medible y tanto él, así como su complemento, son ambos densos en $\mathbb{R}$; sin embargo, $\mu(E \cap J) = 0$ y $\mu(E^c \cap J) > 0$ para cualquier intervalo abierto $J \subseteq \mathbb{R}$. El siguiente ejemplo muestra la existencia de un conjunto medible E de medida positiva tal que él y su complemento son ambos densos en $[a, b]$ y $\mu(E \cap J) > 0$ y $\mu(E^c \cap J) > 0$ para cualquier intervalo abierto $J \subseteq \mathbb{R}$.

Ejercicio 6.8.9. *Existe un conjunto $E \in \mathfrak{M}_\mu([a, b])$ con la siguiente propiedad: para cada **intervalo abierto** $J \subseteq [a, b]$ se cumple que*

$$\mu(E \cap J) > 0 \qquad \text{y} \qquad \mu(E^c \cap J) > 0$$

Prueba. Sea $(I_n)_{n=1}^{\infty}$ una enumeración de todos los intervalos abiertos con extremos racionales incluidos en $[a, b]$. Elija, haciendo uso del Teorema 6.4.2, página 304, un conjunto nunca-denso $D_1 \subseteq I_1$ con $\mu(D_1) > 0$. Puesto que D_1 es nunca-denso, existe un intervalo abierto $(a_1, b_1) \subseteq I_1$ tal que $(a_1, b_1) \cap D_1 = \varnothing$. Por una nueva aplicación del Teorema 6.4.2, seleccione un conjunto nunca-denso $D_2 \subseteq (a_1, b_1)$ de medida positiva. Tenemos entonces dos conjuntos nunca-densos D_1 y D_2, ambos con medida positiva, tales que

$$D_1 \cap D_2 = \varnothing \qquad \text{y} \qquad D_1 \cup D_2 \subseteq I_1.$$

Como $D_1 \cup D_2$ es nunca-denso y el intervalo I_2 es un abierto no vacío, podemos seleccionar un intervalo $(a_2, b_2) \subseteq I_2$ disjunto de $D_1 \cup D_2$. Escoja un conjunto nunca-denso $D_3 \subseteq (a_2, b_2)$ de medida positiva y use el hecho de que D_3 es nunca-denso para elegir un intervalo abierto $(a_3, b_3) \subseteq (a_2, b_2)$ disjunto de D_3. De nuevo, dentro del intervalo (a_3, b_3), construya un conjunto nunca-denso D_4 de medida positiva. De esta forma, hemos construido otros dos conjuntos nunca-densos D_3 y D_4, cada uno con medida positiva, tales que

$$D_3 \cap D_4 = \varnothing \qquad \text{y} \qquad D_3 \cup D_4 \subseteq I_2.$$

Continuando indefinidamente con este procedimiento, se logra generar una extraña sucesión de conjuntos $(D_{2n-1} \cup D_{2n})_{n=1}^{\infty}$, cada uno de los cuales es nunca-denso y, además, satisfacen las siguientes relaciones:

$$D_{2n-1} \cap D_{2n} = \varnothing, \qquad D_{2n-1} \cup D_{2n} \subseteq I_n \qquad \text{y} \qquad \mu(D_n) > 0$$

para todo $n \geq 1$. Sea $E = \bigcup_{n=1}^{\infty} D_{2n}$. Veamos que E tiene las propiedades deseadas. En efecto, sea $J = (\alpha, \beta)$ un intervalo abierto arbitrario incluido en $[a, b]$ y seleccione un $n \in \mathbb{N}$ de modo tal que $I_n \subseteq J$. Por construcción, $D_{2n} \subseteq I_n$ y así, $D_{2n} \subseteq E \cap J$. Por esto

$$\mu(E \cap J) \geq \mu(D_n) > 0.$$

Similarmente, $D_{2n-1} \subseteq E^c \cap J$ y, en consecuencia,

$$\mu(E^c \cap J) \geq \mu(D_{2n-1}) > 0.$$

La prueba es completa. ■

Ejercicio 6.8.10. *Sea V un **conjunto no-medible**. Pruebe que existe un $\varepsilon > 0$ tal que cualesquiera sean los conjuntos medibles E y F satisfaciendo $V \subseteq E$ y $V^c \subseteq F$, entonces se cumple que*

$$\mu(E \cap F) \geq \varepsilon.$$

Prueba. Suponga que la conclusión es falsa. Esto significa que para cada entero $n \geq 1$, existen conjuntos medibles E_n y F_n tales que

$$V \subseteq E_n, \qquad V^c \subseteq F_n \qquad \text{y} \qquad \mu(E_n \cap F_n) < 1/n.$$

Observe que los conjuntos $E = \bigcap_{n=1}^{\infty} E_n$ y $F = \bigcap_{n=1}^{\infty} F_n$ son medibles y satisfacen:

$$V \subseteq E, \qquad V^c \subseteq F \qquad \text{y} \qquad \mu(E \cap F) = 0.$$

Ahora bien, puesto que E es medible, se sigue del Criterio de Carathéodory, Teorema 6.3.58, que cualquiera sea el conjunto $A \subseteq \mathbb{R}$ la igualdad

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c)$$

siempre es válida. Fijemos un conjunto $A \subseteq \mathbb{R}$. Como $E^c \subseteq V^c \subseteq F$, resulta que

$$F = F \cap (E \cup E^c) = (F \cap E) \cup E^c$$

y, en consecuencia,

$$A \cap F \subseteq (A \cap E \cap F) \cup (A \cap E^c).$$

De esto último y la igualdad $\mu^*(E \cap F) = \mu(E \cap F) = 0$, se concluye que

$$\begin{aligned} \mu^*(A \cap F) &\leq \mu^*(A \cap E \cap F) + \mu^*(A \cap E^c) \\ &\leq \mu^*(E \cap F) + \mu^*(A \cap E^c) \\ &= \mu^*(A \cap E^c) \end{aligned}$$

y, por consiguiente,

$$\mu^*(A \cap F) + \mu^*(A \cap E) \leq \mu^*(A \cap E^c) + \mu^*(A \cap E) = \mu^*(A) \tag{1}$$

Finalmente, como

$$\begin{aligned} V \subseteq E \quad &\Rightarrow \quad \mu^*(A \cap V) \leq \mu^*(A \cap E) \quad \text{y} \\ V^c \subseteq F \quad &\Rightarrow \quad \mu^*(A \cap V^c) \leq \mu^*(A \cap F), \end{aligned}$$

entonces de (1) se obtiene que

$$\mu^*(A \cap V) + \mu^*(A \cap V^c) \leq \mu^*(A).$$

Puesto que el conjunto A fue seleccionado de modo arbitrario, se concluye, por una nueva aplicación del Criterio de Carathéodory, que nuestro conjunto V es medible. Esta contradicción finaliza la prueba. ■

Ejercicio 6.8.11. *La función característica del conjunto de Cantor $\chi_\Gamma : [0,1] \to \mathbb{R}$ es discontinua sobre Γ y continua sobre $[0,1] \setminus \Gamma$. Lo mismo es cierto para χ_{Γ_α}, donde Γ_α es cualquier conjunto tipo-Cantor en $[0,1]$ de medida cero.*

Prueba. Sea $x \in \Gamma$. Como Γ no contiene intervalos, resulta que para cada $n \in \mathbb{N}$, el intervalo $(x - 1/n, x + 1/n) \nsubseteq \Gamma$ y, en consecuencia, podemos elegir un $x_n \in (x - 1/n, x + 1/n)$ de modo tal que $x_n \notin \Gamma$. Claramente $x_n \to x$, pero

$$0 = \chi_\Gamma(x_n) \not\longrightarrow \chi_\Gamma(x) = 1.$$

Esto prueba que χ_Γ es discontinua en x y como dicho punto era arbitrario, se concluye que χ_Γ es discontinua sobre Γ.

Sea ahora $x \in [0,1] \setminus \Gamma$ y sea $\varepsilon > 0$. Recuerde que como $[0,1] \setminus \Gamma = \bigcup_{n=1}^\infty J_n$ donde $J_n = J_n(1) \cup \cdots \cup J_n(2^n - 1)$ es la unión de los $2^n - 1$ intervalos abiertos borrados en la n-ésima etapa en la construcción de Γ, entonces existe un único $n \in \mathbb{N}$ tal que $x \in J_n(k)$ para algún $k \in \{1, \ldots, 2^n - 1\}$. Si tomamos $\delta < 1/3^n$ resulta que cualquier y que satisfaga $0 < |x - y| < \delta$ se tiene que $y \in J_n(k)$ y, en consecuencia,

$$|\chi_\Gamma(x) - \chi_\Gamma(y)| = |0 - 0| < \varepsilon.$$

Esto demuestra que χ_Γ es continua en x y termina la aprueba. ■

Ejercicio 6.8.12. *Existe una función $F : [0,1] \to \mathbb{R}$ de clase C^1 con una cantidad no-numerable de puntos críticos.*

Prueba. Deseamos construir una función cuyos ceros sean exactamente los puntos de Γ. Comencemos. Puesto que Γ^c es un subconjunto abierto de $\mathbb{R}$, existe una sucesión disjunta $(I_n)_{n=1}^\infty$ de intervalos abiertos tal que $\Gamma^c = \bigcup_{n=1}^\infty I_n$. Para cada $n \in \mathbb{N}$, sea K_n un conjunto cerrado incluido en I_n y usemos el Lema de Urysohn para determinar una función $f_n : \mathbb{R} \to [0,1]$ de clase C^∞ tal que

$$f_n = 1 \quad \text{sobre } K_n \qquad \text{y} \qquad f_n = 0 \quad \text{fuera de } I_n.$$

En particular, para cada $n \geq 1$, $f_n(x) = 0$ para todo $x \in \Gamma$. Defina $f : \mathbb{R} \to [0,1]$ por

$$f(x) = \sum_{n=1}^{\infty} \frac{f_n(x)}{2^n}.$$

Veamos que f es continua sobre $\mathbb{R}$. En efecto, dado $\varepsilon > 0$, escojamos un $N \in \mathbb{N}$ de modo que

$$\sum_{n=m+1}^{\infty} \frac{1}{2^n} < \frac{\varepsilon}{2}$$

para todo $m \geq N$. Usemos ahora la continuidad uniforme de cada f_n para determinar un $\delta_n > 0$ tal que

$$|f_n(x) - f_n(y)| < \varepsilon/2 \quad \text{siempre que } |x - y| < \delta_n.$$

Si escojemos $\delta > 0$ de modo que $\delta < \min\{\delta_1, \ldots, \delta_m\}$, resultará que si $|x - y| < \delta$, entonces

$$\begin{aligned} |f(x) - f(y)| &= \left| \sum_{n=1}^{\infty} \frac{f_n(x) - f_n(y)}{2^n} \right| \leq \sum_{n=1}^{\infty} \frac{|f_n(x) - f_n(y)|}{2^n} \\ &\leq \sum_{n=1}^{m} \frac{|f_n(x) - f_n(y)|}{2^n} + \sum_{n=m+1}^{\infty} \frac{1}{2^n} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Esto completa la prueba de que f es continua. Finalmente, defina $F : [0,1] \to \mathbb{R}$ por

$$F(x) = \int_0^x f(t)\, dt.$$

Como $f > 0$, se tiene que F es creciente y positiva. Por el Teorema Fundamental del Cálculo, $F'(x) = f(x)$ y, así, los puntos críticos de F son los ceros de f el cual incluye al conjunto de Cantor. Falta verificar que si $x, y \in \Gamma$ con $x \neq y$, entonces $F(x) \neq F(y)$. En efecto, sean $x, y \in \Gamma$ con $x < y$. Como Γ es totalmente disconexo, existe un $z \notin \Gamma$ tal que $x < z < y$. Por esto, existe un $n_0 \in \mathbb{N}$ tal que $f_{n_0}(z) \neq 0$. La continuidad de f_{n_0} nos garantiza la existencia de un intervalo alrededor de z, $J_z \subseteq [x, y] \subseteq [0,1]$ tal que $f_{n_0}(w) > 0$ para todo $w \in J_z$. De aquí se sigue que

$$\int_{J_z} \frac{f_{n_0}(t)}{2^{n_0}}\, dt > 0$$

y, por lo tanto,

$$F(x) = \int_0^x f(t)\, dt < \int_0^x f(t)\, dt + \int_{J_z} \frac{f_{n_0}(t)}{2^{n_0}}\, dt \leq \int_0^x f(t)\, dt + \int_x^y f(t)\, dt = F(y).$$

La prueba es completa. ■

6.9. Problemas

(1) Pruebe que si $\mu^*(A) = 0$, entonces $\mu^*(A \cup B) = \mu^*(B)$ para cualquier conjunto $B \subseteq \mathbb{R}$.

(2) Sean A y B subconjuntos de $\mathbb{R}$. Pruebe que si $\mu^*(A \triangle B) = 0$, entonces $\mu^*(A) = \mu^*(B)$.

(3) Sea E un subconjunto de $\mathbb{R}$ y suponga que $\lambda \neq 0$. Pruebe que $\mu^*(\lambda E) = |\lambda|\mu^*(E)$.

(4) Sea $(I_k)_{k=1}^n$ una colección finita de intervalos abiertos cuya unión cubre a los racionales de $[0,1]$. Demuestre que $\sum_{k=1}^n \ell(I_k) \geq 1$.

(5) Sea $X \in \mathfrak{M}_\mu(\mathbb{R})$. Demuestre que $\mathfrak{M}(X,\mathbb{R}) = \{E \in \mathfrak{M}_\mu(\mathbb{R}) : E \subseteq X\}$ es una σ-álgebra sobre X.

(6) De un ejemplo de una familia de conjuntos $\mathcal{A}$ tal que $\sigma(\mathcal{A}) = \mathfrak{M}_\mu(\mathbb{R})$.

(7) Sea $(\mathcal{A}_\alpha)_{\alpha\in D}$ es una familia arbitraria de σ-álgebras sobre un conjunto X. Pruebe que $\bigcap_{\alpha\in D}\mathcal{A}_\alpha$ es también una σ-álgebra sobre X.

(8) Para cada $\alpha < \omega_1$ defina

$$\Delta_\alpha^0 \ = \ \mathbf{\Sigma}_\alpha^0 \cap \mathbf{\Pi}_\alpha^0.$$

Pruebe que:

(a) $\mathbf{\Sigma}_\beta^0 \cap \mathbf{\Pi}_\beta^0 \subseteq \Delta_\alpha^0$ para cualquier $1 \leq \beta < \alpha$.

(b) $\mathbf{\Sigma}_\alpha^0 = \left\{ \bigcup_{n=1}^\infty V_n : V_n \in \bigcup_{\beta<\alpha} \mathbf{\Pi}_\beta^0 \right\}$ para cualquier $\alpha \geq 2$.

(c) Si $f : \mathbb{R} \to \mathbb{R}$ es una **función continua**, entonces

$$A \in \mathbf{\Sigma}_\alpha^0 \ \Rightarrow \ f^{-1}(A) \in \mathbf{\Sigma}_\alpha^0 \qquad \text{y} \qquad B \in \mathbf{\Pi}_\alpha^0 \ \Rightarrow \ f^{-1}(B) \in \mathbf{\Pi}_\alpha^0.$$

(d) $\mathfrak{Bo}(\mathbb{R}) \ = \ \bigcup_{\alpha<\omega_1} \mathbf{\Sigma}_\alpha^0 \ = \ \bigcup_{\alpha<\omega_1} \mathbf{\Pi}_\alpha^0.$

(e) Para cada $\alpha < \omega_1$, $\mathrm{card}(\mathbf{\Sigma}_\alpha^0) = \mathrm{card}(\mathbf{\Pi}_\alpha^0) = \mathfrak{c}$. Concluya que $\mathrm{card}(\mathfrak{Bo}(\mathbb{R})) = \mathfrak{c}$.

(9) Sea $A \subseteq \mathbb{R}$ y fijemos un $\delta > 0$. Una colección numerable $(V_n)_{n=1}^\infty$ de subconjuntos de $\mathbb{R}$ se llama un δ-**cubrimiento** de A si

(a) $\mathrm{diam}(V_n) < \delta$ para todo $n \geq 1$, y

(b) $A \subseteq \bigcup_{n=1}^\infty V_n$.

Defina

$$H_\delta(A) \ = \ \inf\left\{ \sum_{n=1}^\infty \mathrm{diam}(V_n) \ : \ (V_n)_{n=1}^\infty \text{ es un } \delta\text{-cubrimiento de } A \right\}.$$

Demuestre que:

(1) Si $\delta_1 < \delta_2$, entonces $H_{\delta_1}(A) \leq H_{\delta_2}(A)$ para todo $A \subseteq \mathbb{R}$.

(2) Para todo $A \subseteq \mathbb{R}$,

$$\mu^*(A) \ = \ \lim_{\delta\to 0} H_\delta(A).$$

(10) El siguiente ejercicio indica que en la definición de $\mu^*(A)$ se puede considerar, en lugar de colecciones numerables de intervalos de abiertos, cualquier colección numerable de intervalos de longitud finita. Sea $A \subseteq \mathbb{R}$. Pruebe que

$$\mu^*(A) \;=\; \text{ínf}\left\{ \sum_{n=1}^{\infty} \ell(I_n) \;:\; (I_n)_{n=1}^{\infty} \in \mathscr{B}_i \right\}$$

donde $\mathscr{B}_i$ es cada una de las siguientes colecciones:

$$\mathscr{B}_1 \;=\; \left\{ (I_n)_{n=1}^{\infty} \;:\; A \subseteq \bigcup_{n=1}^{\infty} I_n,\; I_n = [a_n, b_n] \text{ para todo } n \in \mathbb{N} \right\}$$

$$\mathscr{B}_2 \;=\; \left\{ (I_n)_{n=1}^{\infty} \;:\; A \subseteq \bigcup_{n=1}^{\infty} I_n,\; I_n = (a_n, b_n] \text{ para todo } n \in \mathbb{N} \right\}$$

$$\mathscr{B}_3 \;=\; \left\{ (I_n)_{n=1}^{\infty} \;:\; A \subseteq \bigcup_{n=1}^{\infty} I_n,\; I_n = [a_n, b_n) \text{ para todo } n \in \mathbb{N} \right\}$$

$$\mathscr{B}_4 \;=\; \left\{ (I_n)_{n=1}^{\infty} \;:\; A \subseteq \bigcup_{n=1}^{\infty} I_n,\; I_n \text{ es de longitud finita para todo } n \in \mathbb{N} \right\}$$

(11) Sea $E \in \mathfrak{M}_\mu(\mathbb{R})$ y acotado. Si para algún conjunto infinito numerable y acotado Λ de números reales se cumple que $(a+E)_{a\in\Lambda}$ es disjunta, entonces demuestre que $\mu(E) = 0$. La conclusión es falsa si el conjunto Λ se toma finito. Por ejemplo, si $E = [0,1)$ y $\Lambda = \{0,1\}$, entonces E y $1+E$ son disjuntos, pero $\mu(E) > 0$. Similarmente, si Λ no es acotado, entonces tomando $E = [0,1)$ y $\Lambda = \mathbb{Z}$, resulta que la sucesión $(a+E)_{a\in\Lambda}$ es disjunta, pero $\mu(E) > 0$. ¿Qué ocurre si Λ es no-numerable?

(12) Sea $E \subseteq [0,1]$ un conjunto medible con la siguiente propiedad: existe un número $\delta > 0$ para el cual $\mu(E \cap J) > \delta \cdot \ell(J)$ para cada intervalo abierto $J \subseteq [0,1]$. Demuestre que $\mu(E) = 1$.

(13) Sea $E \subseteq \mathbb{R}$ un conjunto con la siguiente propiedad: existe un número $0 < \delta < 1$ tal que, para cada intervalo abierto (a,b) el conjunto $E \cap (a,b)$ puede ser cubierto por una colección numerable de intervalos $(I_n)_{n=1}^{\infty}$ tal que

$$\sum_{n=1}^{\infty} \ell(I_n) \;<\delta \cdot (b-a).$$

Demuestre que $\mu(E) = 0$.

(14) Sea $E \in \mathfrak{M}_\mu(\mathbb{R})$ con $\mu(E) > 0$. Demuestre que $\mu(E + \mathbb{Q}) = +\infty$, donde

$$E + \mathbb{Q} \;=\; \{x + q \;:\; x \in E,\; q \in \mathbb{Q}\}.$$

(15) Sea $E \in \mathfrak{M}_\mu(\mathbb{R})$ tal que $\mu(E) = 0$. ¿Es cierto que $\mu(E \cdot E) = 0$? Justifique su respuesta.

(16) Sea $E \in \mathfrak{M}_\mu(\mathbb{R})$. Suponga que $\mu(E \triangle (x+E)) = 0$ para todo x en algún subconjunto denso de $\mathbb{R}$. Demuestre que $\mu(E) = 0$ o $\mu(\mathbb{R} \setminus E) = 0$.

(17) Demuestre que para cualquier medible $E \subseteq \mathbb{R}$ con $\mu(E) = 1$, existe un conjunto medible $B \subseteq E$ tal que $\mu(B) = 1/2$.

(18) Sea $E \in \mathfrak{M}_\mu(\mathbb{R})$ y suponga que

$$\mu\big(E \cap (a,b)\big) \ \leq \ \frac{b-a}{2}$$

para todo $a, b \in \mathbb{R}$ con $a < b$. Pruebe que $\mu(E) = 0$.

(19) Construya un subconjunto de $\mathbb{R}$ que sea denso, de primera categoría y tenga cardinalidad $\mathfrak{c}$.

(20) Pruebe que:

(a) Si $A \subseteq \mathbb{R}$ es analítico, entonces $f^{-1}(A)$ es analítico.

(b) $A \subseteq \mathbb{R}$ es analítico si, y sólo si, existe un conjunto cerrado $F \subseteq \mathbb{R} \times \mathcal{N}$ cuya proyección $\mathrm{proy}_1(F) = A$.

(c) $A \subseteq \mathbb{R}$ es analítico si, y sólo si, existe una función continua y sobreyectiva $\Phi : \mathbb{R} \setminus \mathbb{Q} \to A$.

(d) Si $A_1, \ldots, A_n$ son subconjuntos analíticos de $\mathbb{R}$, entonces $A_1 \times \cdots \times A_n$ es analítico en $\mathbb{R}^n$.

(21) Sea $(q_n)_{n=1}^\infty$ una enumeración de los racionales y defina el conjunto

$$G \ = \ \bigcup_{n=1}^\infty \left(q_n - \frac{1}{n^2}, \, q_n + \frac{1}{n^2} \right).$$

Demuestre que, para cualquier conjunto cerrado $F \subseteq \mathbb{R}$, se cumple que $\mu(G \triangle F) > 0$.

(22) Sea $A \subseteq \mathbb{R}$ un conjunto **no-medible** con $\mu^*(A) < +\infty$. Pruebe que existe un $\delta > 0$ tal que $\mu(E) \leq \mu^*(A) - \delta$ para cualquier conjunto medible $E \subseteq A$.

(23) Pruebe que existe un conjunto no-medible $A \subseteq \mathbb{R}$ tal que

$$\mu_*(A \cap E) \ = \ 0 \qquad \text{y} \qquad \mu^*(A \cap E) \ = \ \mu(E)$$

para todo conjunto medible E con $\mu(E) > 0$.

(24) Sea E un conjunto medible no-numerable con $\mu(E) > 0$. Demuestre que E contiene un subconjunto no-numerable F con $\mu(F) = 0$.

(25) Pruebe que existe una familia no-numerable $\{N_t : t \in \mathbb{R}\}$ se subconjuntos de $\mathbb{R}$ con las siguientes propiedades: (a) cada N_t es numerable y (b) $N_s \cap N_t$ es finito para todo $s \neq t$.

(26) **Definición 6.9.1 (Borel,1919).** *Un conjunto $M \subseteq \mathbb{R}$ se dice que es **fuertemente de medida cero** si para cada sucesión $(\varepsilon_n)_{n=1}^\infty$ de números reales positivos, existe una sucesión $(I_n)_{n=1}^\infty$ de intervalos abiertos tal que*

$$M \ \subseteq \ \bigcup_{n=1}^\infty I_n \qquad \text{y} \qquad \mu(I_n) \ < \ \varepsilon_n \quad \text{para todo } n \geq 1.$$

Pruebe que:

(a) Cualquier conjunto numerable es fuertemente de medida cero.

(b) Si M es fuertemente de medida cero, entonces $\mu(M) = 0$.

(c) Pruebe que Γ, el conjunto ternario de Cantor, no es fuertemente de medida cero.

(d) Ningún conjunto conteniendo a un conjunto perfecto es fuertemente de medida cero.

(e) Si $f : [0,1] \to \mathbb{R}$ es una **función continua** y $M \subseteq [0,1]$ es fuertemente de medida cero, entonces $f(M)$ es fuertemente de medida cero. Deduzca de esto que existen conjuntos nulos que no son fuertemente de medida cero.

(f) Bajo la Hipótesis del Continuo, existen conjuntos no-numerables que son fuertemente de medida cero.

Conjetura de Borel. **Ningún conjunto no-numerable de $\mathbb{R}$ es fuertemente de medida cero**.

(27) (a) Sean $E, F \in \mathfrak{M}_\mu(\mathbb{R})$ con $\mu(E) > 0$ y $\mu(F) > 0$. Demuestre que existen conjunto nulos $M \subseteq E$ y $N \subseteq F$ tal que $M + N$ es no-medible.

(b) Sea $E \in \mathfrak{M}_\mu(\mathbb{R})$. Pruebe que existe $N \subseteq E$ tal que $\mu(N) = 0$ y $N + N \notin \mathfrak{M}_\mu(\mathbb{R})$.

(c) Pruebe que existe un conjunto $N \subseteq \Gamma$ tal que $N + N$ es un conjunto de Bernstein en $[0,2]$.

(28) Sea $\mathcal{H} \subseteq \mathbb{R}_{\mathbb{Q}}$ una base de Hamel. Demuestre que:

(a) $\mu_*(\mathcal{H}) = 0$.

(b) Si $E \subseteq \mathcal{H}$ es medible, entonces $\mu(E) = 0$.

(c) Si $D \subseteq \mathcal{H}$ es un conjunto a lo más numerable, entonces $\operatorname{Lin}(\mathcal{H} \setminus D, \mathbb{Q})$ es un conjunto saturado no-medible.

(29) Sea $X \subseteq \mathbb{R}$. Una base de Hamel $\mathcal{H} \subseteq X$ es llamada una ***base de Burstin relativa a*** X si $\mathcal{H} \cap B \neq \varnothing$ para cualquier conjunto de Borel no-numerable $B \subseteq X$. Pruebe que:

(a) Si $\mu(\mathbb{R} \setminus X) = 0$, entonces cualquier base de Burstin relativa a X es un conjunto saturado no-medible.

(b) Si $\operatorname{Lin}(X, \mathbb{Q}) = \mathbb{R}$, entonces existe base de Burstin relativa a X.

(c) Cualquier base de Burstin (relativa a $\mathbb{R}$) es un conjunto saturado no-medible.

(30) Pruebe que, bajo la Hipótesis de Continuo, existe una base de $\mathcal{H} \subseteq \mathbb{R}_{\mathbb{Q}}$ que es un conjunto de Lusin.

(31) Sea $f : \mathbb{R} \to \mathbb{R}$ una función de Cauchy. Demuestre que las siguientes condiciones son equivalentes:

(a) f es continua en un punto x_0.

(b) f es monótona creciente.

(c) $f(x) \geq 0$ para todo $x \geq 0$.

(d) f está acotada por arriba sobre cualquier intervalo finito.

(e) f está acotada por abajo sobre cualquier intervalo finito.

(f) f es diferenciable.

(g) $f(x) = ax$ para algún $a \in \mathbb{R}$.

(32) Sea $f : \mathbb{R} \to \mathbb{R}$ una función arbitraria no idénticamente nula. Cualquier conjunto de la forma $L = f^{-1}(\{y\})$ donde $y \in \mathbb{R}$, es llamado un **conjunto de nivel** para f. Suponga que f es una **función de Cauchy**, $f \neq 0$. Pruebe que:

(a) Si L_1 y L_2 son dos conjuntos de nivel de f, entonces existe un $a \in \mathbb{R}$ tal que $L_1 = a + L_2$.

(b) Si L es un conjunto de nivel de f, entonces $\operatorname{card}(L) = 1$ o $\operatorname{card}(L) \geq \aleph_0$.

(c) $\mu_*(L) = 0$ para cualquier conjunto de nivel L de f.

(d) Si L es un conjunto de nivel de f, entonces $\mu(L) = 0$ o L es un conjunto saturado no-medible.

(33) Para cada cardinal $\mathfrak{m}$ tal que $\aleph_0 \leq \mathfrak{m} \leq \mathfrak{c}$, existe una función de Cauchy $f : \mathbb{R} \to \mathbb{R}$ tal que $\operatorname{card}(f(\mathbb{R})) = \mathfrak{m}$ y todos los conjuntos de nivel de f son saturados no-medibles.

(34) Bajo la Hipótesis del Continuo, existe una función de Cauchy $f : \mathbb{R} \to \mathbb{R}$ tal que $f(\mathbb{R}) = \mathbb{R}$ y todos los conjuntos de nivel de f tienen la cardinalidad del continuo y medida cero.

(35) Defina la función $f : \mathbb{R} \to \mathbb{R}$ por

$$f(x) = \begin{cases} \lim\limits_{n\to\infty} \tan(n!\pi\, x) & \textit{si } \lim\limits_{n\to\infty} \tan(n!\pi\, x) \textit{ existe.} \\ 0 & \textit{si } \lim\limits_{n\to\infty} \tan(n!\pi\, x) \textit{ no existe.} \end{cases}$$

Pruebe que f posee las siguientes propiedades:

(a) $f(x+q) = f(x)$ para todo $x \in \mathbb{R}$ y todo $q \in \mathbb{Q}$.

(b) f es sobreyectiva.

(c) f es siempre-sobreyectiva.

(36) Pruebe que cualquier función acotada y creciente $F : [0,1] \to \mathbb{R}$ que satisfaga:

$$F\Big(\frac{x}{3}\Big) = \frac{F(x)}{3} \qquad \text{y} \qquad F(1-x) = 1 - F(x)$$

para todo $x \in [0,1]$ es la función de Cantor.

CAPÍTULO 7

Funciones Medibles

En esta capítulo estudiaremos un tipo muy general de funciones a las que llamaremos funciones medibles. Tales funciones son las que permiten desarrollar una teoría de integración mucho más amplia y satisfactoria que la Teoría de Integración de Riemann. Para lograr tal objetivo se requiere que los dominios de tales funciones sean subconjuntos medibles de $\mathbb{R}$ con rango en el conjunto de los números reales extendidos. Estas exigencias conducen a que casi todas las operaciones que se pueden hacer con ellas se preserven, salvo contadas excepciones. Por ejemplo, todas las operaciones algebraicas entre funciones medibles, así como tomar límites puntuales, tomar supremos e ínfimos, etc., producen funciones medibles, mientras que la composición de tales funciones no garantiza, en general, que ella sea medible. El programa sobre las funciones medibles incluye el estudio de tres resultados fundamentales que jugarán un papel de primer orden en el estudio sobre la integral de Lebesgue: el primero establece que *toda función medible* se puede aproximar por una *función simple*, el segundo dice que *la convergencia puntual* de funciones medibles *es casi convergencia uniforme*, y, finalmente, el tercero afirma que *toda función medible* es *casi continua*.

7.1. Definición

Recordemos que una función $f : \mathbb{R} \to \mathbb{R}$ es continua si, y sólo si, $f^{-1}(G)$ es un conjunto abierto para cualquier conjunto abierto G de $\mathbb{R}$. Una pequeña variación de este concepto conduce a la definición de función medible, de hecho, una función f será medible si, y sólo, si $f^{-1}(G)$ es un conjunto medible para cualquier conjunto abierto G de $\mathbb{R}$.

Definición 7.1.1. *Sea $E \subseteq \mathbb{R}$. Una función $f : E \to \overline{\mathbb{R}}$ se dice que es* ***medible Lebesgue****, o simplemente* ***medible****, sobre E si $E \in \mathfrak{M}_\mu(\mathbb{R})$ y para cada $a \in \mathbb{R}$, el conjunto*

$$f^{-1}\big((a, +\infty]\big) = \big\{x \in E : f(x) > a\big\}$$

es medible.

Tal como acabamos de definirla, el dominio de una función medible siempre será un conjunto medible. Más aun, asumiremos, a partir de este momento, que *el dominio de una función, sea ésta medible o no, será siempre un conjunto medible* salvo que explícitamente declaremos lo contrario.

En ocasiones escribiremos el símboolo E^a_f, o con más frecuencia $[f > a]$, para denotar el conjunto $\{x \in E : f(x) > a\}$ para cada $a \in \mathbb{R}$. En particular, si $a \in \{-\infty, +\infty\}$

$$E^{\infty}_f = \{x \in E : |f(x)| = \infty\}.$$

En general, si X es un conjunto arbitrario y $\mathcal{A}$ es una σ-álgebra de subconjuntos de X, entonces diremos que $f : X \to \overline{\mathbb{R}}$ es $\mathcal{A}$-**medible** si $\{x \in X : f(x) > a\} \in \mathcal{A}$ para cualquier $a \in \mathbb{R}$. En consecuencia, decir que $f : E \to \overline{\mathbb{R}}$ es una función medible según Lebesgue, significa que ella es $\mathfrak{M}_\mu(E)$-medible, donde $\mathfrak{M}_\mu(E)$ es *la colección de todos los subconjuntos medibles según Lebesgue incluidos en* E.

En lo que sigue, E será un conjunto medible fijo y el símbolo $\mathcal{F}_\mu(E)$ denotará a *la familia de todas las* ***funciones medibles*** $f : E \to \overline{\mathbb{R}}$.

Ejemplo 7.1.1. *Si $f : E \to \mathbb{R}$ es **continua**, entonces $f \in \mathcal{F}_\mu(E)$.*

Prueba. Como f es continua, para cualquier $a \in \mathbb{R}$, se cumple que $f^{-1}\big((a, +\infty)\big)$ es un conjunto abierto en $\mathbb{R}$ y, por lo tanto, medible. De esto se sigue

$$\{x \in E : f(x) > a\} = f^{-1}\big((a, +\infty)\big) \cap E \in \mathfrak{M}_\mu(\mathbb{R})$$

y termina la prueba. ■

En particular, cualquier función $f \in C_c(\mathbb{R})$ es medible. Si para cada $\alpha \in \mathbb{R}$ y cada $f \in C_c(\mathbb{R})$, consideramos el conjunto

$$K_{f,\alpha} = \{x \in \mathbb{R} : f(x) \geq \alpha\},$$

resulta que $K_{f,\alpha}$ es un conjunto G_δ-compacto. La σ-álgebra $\mathfrak{Ba}(\mathbb{R})$ generada por la colección

$$\mathcal{C} = \{K_{f,\alpha} : f \in C_c(\mathbb{R}),\ \alpha \in \mathbb{R}\}$$

es llamada la **σ-álgebra de Baire**. Observe que $\mathfrak{Ba}(\mathbb{R})$ es la σ-álgebra más pequeña tal que cada función $f \in C_c(\mathbb{R})$ es $\mathfrak{Ba}(\mathbb{R})$-medible. Puesto que cualquier conjunto G_δ-compacto es un boreliano, se tiene que $\mathfrak{Ba}(\mathbb{R}) \subseteq \mathfrak{Bo}(\mathbb{R})$. Es fácil demostrar que, en realidad, $\mathfrak{Ba}(\mathbb{R}) = \mathfrak{Bo}(\mathbb{R})$.

Ejemplo 7.1.2. *Sea $A \subseteq E$. Entonces $\chi_A \in \mathcal{F}_\mu(E)$ si, y sólo si, $A \in \mathfrak{M}_\mu(\mathbb{R})$.*

Prueba. En efecto, para cada $a \in \mathbb{R}$,

$$\Big\{x \in E : \chi_A(x) > a\Big\} = \begin{cases} E & \text{si} \quad a < 0 \\ A & \text{si} \quad 0 \leq a < 1 \\ \varnothing & \text{si} \quad a \geq 1. \end{cases}$$

es medible si, y sólo si, $A \in \mathfrak{M}_\mu(\mathbb{R})$. ■

De este resultado se tiene que *la existencia de conjuntos no-medibles generan funciones no-medibles*.

Teorema 7.1.2. *Sean* $E \in \mathfrak{M}_\mu(\mathbb{R})$ *y* $f : E \to \overline{\mathbb{R}}$. *Las siguientes condiciones son equivalentes:*

(1) $f \in \mathcal{F}_\mu(E)$.

(2) $\{x \in E \,:\, f(x) \geq a\} \in \mathfrak{M}_\mu(\mathbb{R})$ *para cada* $a \in \mathbb{R}$.

(3) $\{x \in E \,:\, f(x) < a\} \in \mathfrak{M}_\mu(\mathbb{R})$ *para cada* $a \in \mathbb{R}$.

(4) $\{x \in E \,:\, f(x) \leq a\} \in \mathfrak{M}_\mu(\mathbb{R})$ *para cada* $a \in \mathbb{R}$.

Prueba. Para ver que $(1) \Rightarrow (2)$ simplemente observe que

$$\{x \in E \,:\, f(x) \geq a\} \;=\; \bigcap_{n=1}^{\infty} \{x \in E \,:\, f(x) > a - 1/n\} \;\in\; \mathfrak{M}_\mu(\mathbb{R}),$$

mientras que $(2) \Rightarrow (3)$ sigue de la relación

$$\{x \in E \,:\, f(x) < a\} \;=\; E \setminus \{x \in E \,:\, f(x) \geq a\} \;\in\; \mathfrak{M}_\mu(\mathbb{R}).$$

$(3) \Rightarrow (4)$ se demuestra de modo similar a la implicación $(1) \Rightarrow (2)$. Finalmente, que $(4) \Rightarrow (1)$, se obtiene directamente tomando complemento, esto es,

$$\{x \in E \,:\, f(x) > a\} \;=\; E \setminus \{x \in E \,:\, f(x) \leq a\} \;\in\; \mathfrak{M}_\mu(\mathbb{R}).$$

La prueba es completa. ■

El siguiente resultado constituye una útil caracterización de las funciones medibles la cual depende fundamentalmente del Lema 6.3.22.

Teorema 7.1.3. *Sean* $E \in \mathfrak{M}_\mu(\mathbb{R})$ *y* $f : E \to \mathbb{R}$ *una función. Son equivalentes:*

(1) $f \in \mathcal{F}_\mu(E)$.

(2) $f^{-1}(B) \in \mathfrak{M}_\mu(E)$ *para cada* ***conjunto de Borel*** $B \subseteq \mathbb{R}$.

Prueba. $(1) \Rightarrow (2)$. Suponga que $f \in \mathcal{F}_\mu(E)$ y defina

$$\mathcal{B}_f \;=\; \{B \subseteq \mathbb{R} \,:\, f^{-1}(B) \in \mathfrak{M}_\mu(E)\}.$$

Tal como se demostró en el Lema 6.3.22, página 272, $\mathcal{B}_f$ es una σ-álgebra conteniendo a $\mathfrak{Bo}(\mathbb{R})$. En particular, $f^{-1}(B) \in \mathfrak{M}_\mu(E)$ para cualquier conjunto de Borel B. La implicación $(2) \Rightarrow (1)$ es inmediata. ■

Es de interés saber cómo son los conjuntos de nivel de una función medible, es decir, los conjuntos de la forma $\{x \in E : f(x) = a\}$ para cada $a \in \mathbb{R}$. Por ejemplo, si f es medible, todos sus conjuntos de nivel son medible como lo muestra el siguiente corolario, aunque el recíproco no es, en general, cierto.

Corolario 7.1.4. *Si* $f \in \mathcal{F}_\mu(E)$, *entonces para cada* $a \in \overline{\mathbb{R}}$,

$$\{x \in E \,:\, f(x) = a\} \;\in\; \mathfrak{M}_\mu(\mathbb{R}).$$

Prueba. Si $a \in \mathbb{R}$, entonces

$$\{x \in E \,:\, f(x) = a\} \;=\; \{x \in E \,:\, f(x) \geq a\} \cap \{x \in E \,:\, f(x) \leq a\}$$

es medible, mientras que si $a = +\infty$, entonces

$$\{x \in E \,:\, f(x) = +\infty\} \;=\; \bigcap_{n=1}^{\infty} \{x \in E \,:\, f(x) \geq n\} \;\in\; \mathfrak{M}_\mu(\mathbb{R}).$$

El caso $a = -\infty$ se demuestra de forma similar. ■

El siguiente resultado es un contraejemplo al recíproco del corolario anterior.

Ejemplo 7.1.3. *Existe una función **no-medible** $f : \mathbb{R} \to \mathbb{R}$ tal que*

$$\{x \in \mathbb{R} \,:\, f(x) = a\} \in \mathfrak{M}_\mu(\mathbb{R})$$

para todo $a \in \mathbb{R}$.

Prueba. Usemos el Teorema de Rademacher, Teorema 6.5.3, para seleccionar un subconjunto no-medible V incluido en $(0, +\infty)$ y defina $f : \mathbb{R} \to \mathbb{R}$ por

$$f(x) \;=\; \begin{cases} \;\;x^2 & \text{si} \quad x \in V \\ -x^2 & \text{si} \quad x \in \mathbb{R} \setminus V. \end{cases}$$

Sea $a \in \mathbb{R}$. Entonces

$$\{x \in \mathbb{R} \,:\, f(x) = a\} \;=\; \begin{cases} \{0\} \in \mathfrak{M}_\mu(\mathbb{R}) & \text{si} \quad a = 0 \\ \{\sqrt{a},\; -\sqrt{a}\,\} \in \mathfrak{M}_\mu(\mathbb{R}) & \text{si} \quad a > 0 \\ \{\sqrt{-a},\; -\sqrt{-a}\,\} \in \mathfrak{M}_\mu(\mathbb{R}) & \text{si} \quad a < 0. \end{cases}$$

Sin embargo, tomando $a = 0 \notin V$ resulta que

$$\{x \in \mathbb{R} \,:\, f(x) > 0\} \;=\; V \;\notin\; \mathfrak{M}_\mu(\mathbb{R}).$$

Esto nos revela que f no es medible. ■

En lugar de considerar todos los números reales para comprobar si una función dada es medible, es suficiente verificar la condición de medibilidad en un subconjunto denso (sea o no numerable) de $\mathbb{R}$ tal como se prueba en el siguiente resultado.

Teorema 7.1.5. *Sean $E \in \mathfrak{M}_\mu(\mathbb{R})$, D un subconjunto **denso** de $\mathbb{R}$ y $f : E \to \overline{\mathbb{R}}$ una función. Son equivalentes:*

(1) $f \in \mathcal{F}_\mu(E)$.

(2) $\{x \in E : f(x) > r\} \in \mathfrak{M}_\mu(\mathbb{R})$ *para todo $r \in D$.*

Prueba. Es claro que $(1) \Rightarrow (2)$. Para demostrar la otra implicación, suponga que (2) se cumple y sea $a \in \mathbb{R}$. Como D es denso en $\mathbb{R}$, existe una sucesión estrictamente creciente $(r_n)_{n=1}^{\infty}$ en D tal que $r_n \to a$. Sabemos, por hipótesis, que $\{x \in E : f(x) > r_n\} \in \mathfrak{M}_\mu(\mathbb{R})$ para todo $n \geq 1$ y, en consecuencia,

$$\{x \in E : f(x) \geq a\} \;=\; \bigcap_{n=1}^{\infty}\{x \in E : f(x) > r_n\} \;\in\; \mathfrak{M}_\mu(\mathbb{R}).$$

Se sigue del Teorema 7.1.2 que f es medible. ■

Nota Adicional 7.1.1 En el Teorema 7.1.2 podemos imponer la condición de que $a \in \overline{\mathbb{R}}$. En efecto, suponga que $f \in \mathcal{F}_\mu(E)$. Si $a = -\infty$, entonces

$$\big\{x \in E \;:\; f(x) \leq -\infty\big\} \;=\; \big\{x \in E \;:\; f(x) = -\infty\big\}$$

es medible gracias al Corolario 7.1.4. Por otro lado,

$$\big\{x \in E \;:\; f(x) \geq -\infty\big\} \;=\; E$$

el cual también es medible. Similarmente, los conjuntos $\big\{x \in E : f(x) \leq \infty\big\} = E$ y $\big\{x \in E : f(x) \geq +\infty\big\} = \big\{x \in E : f(x) = +\infty\big\}$ son medibles. De modo que todas las equivalencias del Teorema 7.1.2 son validas para cualquier $a \in \overline{\mathbb{R}}$; esto es,

Corolario 7.1.6. *Sea $f : E \to \overline{\mathbb{R}}$ una función. Las siguientes condiciones son equivalentes:*

(1) $f \in \mathcal{F}_\mu(E)$.

(2) $\big\{x \in E \;:\; f(x) \geq a\big\} \in \mathfrak{M}_\mu(\mathbb{R})$ para cada $a \in \overline{\mathbb{R}}$.

(3) $\big\{x \in E \;:\; f(x) < a\big\} \in \mathfrak{M}_\mu(\mathbb{R})$ para cada $a \in \overline{\mathbb{R}}$.

(4) $\big\{x \in E \;:\; f(x) \leq a\big\} \in \mathfrak{M}_\mu(\mathbb{R})$ para cada $a \in \overline{\mathbb{R}}$.

En la Teoría de Integración de Lebesgue, como veremos posteriormente, los conjuntos nulos siempre pueden ser ignorados. Por esta razón conviene considerar una propiedad que los identifique. De modo preciso:

Definición 7.1.7. *Una propiedad* **P**, *relativa a los elementos de un conjunto $E \in \mathfrak{M}_\mu(\mathbb{R})$, se dirá que se cumple μ-**casi-siempre**, o simplemente **casi-siempre** sobre E, si el conjunto de puntos donde ella no se verifica es de medida cero, es decir,*

$$\mu\big(\big\{x \in E : \mathbf{P}(x) \text{ es falso}\big\}\big) \;=\; 0.$$

En ocasiones escribiremos "$\mathbf{P}(x)$ se cumple para **casi-todo** x" en lugar de decir que dicha propiedad se cumple μ-casi-siempre. Con frecuencia, abreviaremos la expresión μ-casi-siempre por μ-c.s.

Existen varias situaciones que son de interés en estas notas referente a la noción casi-siempre. Por ejemplo, si $f : E \to \overline{\mathbb{R}}$ es medible y

$$\mu\big(\{x \in E \;:\; f(x) = \pm\infty\}\big) \;=\; 0,$$

entonces diremos que f es **finita casi-siempre** sobre E, o que $f(x)$ es **finita** para **casi-todo** $x \in E$. De igual modo, convenimos en decir que f está **acotada casi-siempre** sobre E, si existe una constante $M > 0$ y un conjunto medible $A \subseteq E$ con $\mu(A) = 0$ tal que

$$|f(x)| \leq M \quad \text{para todo} \quad x \notin A.$$

Sean $f, g \in \mathfrak{F}_\mu(E)$. La expresión $f = g$ **casi-siempre** sobre E, significa que

$$\mu\big(\{x \in E : f(x) \neq g(x)\}\big) = 0.$$

En particular, diremos que f es **nula casi-siempre** sobre E si $f = 0$ casi-siempre sobre E. Similarmente, la función f se dice que es **continua casi-siempre** sobre E, si $\mu(\mathrm{Disc}(f)) = 0$. Suponga que $(f_n)_{n=1}^\infty$ es una sucesión en $\mathfrak{F}_\mu(E)$. Se dirá que $f_n \to f$ **puntualmente casi-siempre**, o simplemente, que **converge casi-siempre** sobre E, si

$$\mu\big(\{x \in E : \lim_{n\to\infty} f_n(x) \neq f(x)\}\big) = 0.$$

La noción de ser **diferenciable casi-siempre** se define de manera similar.

Teorema 7.1.8. *Sea* $\mathbf{P}$ *una propiedad relativa a los elementos de un conjunto* $E \in \mathfrak{M}_\mu(\mathbb{R})$. *Las siguientes afirmaciones son equivalentes:*

(1) $\mathbf{P}(x)$ *se cumple* ***casi-siempre*** *sobre* E.

(2) *Para cada* $\varepsilon > 0$, *existe un* ***conjunto abierto*** G *con* $\mu(G) < \varepsilon$ *tal que* $\mathbf{P}(x)$ *se cumple para todo* $x \in E \setminus G$.

Prueba. (1) $\Rightarrow$ (2). Suponga que (1) se cumple y sea $\varepsilon > 0$. Puesto que $\mu(N) = 0$, donde $N = \big\{x \in E : \mathbf{P}(x) \text{ es falso}\big\}$, resulta que N es medible y, por lo tanto, existe un conjunto abierto G tal que $N \subseteq G$ y $\mu(G \setminus N) < \varepsilon$. Por supuesto, $\mu(G) < \varepsilon$ ya que $G = N \cup (G \setminus N)$. Finalmente, como $\mathbf{P}(x)$ se cumple para todo $x \in E \setminus N$ y puesto que $N \subseteq G$, entonces con más razón $\mathbf{P}(x)$ se cumple si $x \in E \setminus G$. Esto prueba (2).

(2) $\Rightarrow$ (1). Suponga ahora que (2) se cumple. Por cada $n \in \mathbb{N}$, escoja un conjunto abierto G_n con $\mu(G_n) < 1/n$ tal que la propiedad de $\mathbf{P}(x)$ se cumple para todo $x \in E \setminus G_n$. Sea

$$N = \bigcap_{n=1}^{\infty} G_n.$$

Claramente $\mu(N) = 0$ y, por supuesto, $\mathbf{P}(x)$ es falsa si $x \in N$. Esto termina la prueba. ■

Teorema 7.1.9. *Sea* $f \in \mathfrak{F}_\mu(E)$ *y suponga que* $g : E \to \overline{\mathbb{R}}$ *es una función tal que* $f = g$ ***casi-siempre*** *sobre* E. *Entonces* $g \in \mathfrak{F}_\mu(E)$.

Prueba. Sea $A = \{x \in E : f(x) = g(x)\}$. Entonces $A \in \mathfrak{M}_\mu(\mathbb{R})$, $\mu(E \setminus A) = 0$ y todos los subconjuntos de $E \setminus A$ están en $\mathfrak{M}_\mu(\mathbb{R})$. Tomemos ahora cualquier $r \in \mathbb{R}$ y observe que como f es medible, $\{x \in E : f(x) > a\} \cap A \in \mathfrak{M}_\mu(\mathbb{R})$. Igualmente, $\{x \in E : g(x) > a\} \cap (E \setminus A) \in \mathfrak{M}_\mu(\mathbb{R})$ ya que dicho conjunto está incluido en $E \setminus A$. Por esto, el conjunto

$$\begin{aligned}
\{x \in E : g(x) > a\} &= \Big(\{x \in E : g(x) > a\} \cap A\Big) \cup \Big(\{x \in E : g(x) > a\} \cap (E \setminus A)\Big) \\
&= \Big(\{x \in E : f(x) > a\} \cap A\Big) \cup \Big(\{x \in E : g(x) > a\} \cap (E \setminus A)\Big)
\end{aligned}$$

es medible y termina la prueba. ■

La importancia del resultado anterior radica en que el conjunto $\mathcal{F}_\mu(E)$ se puede particionar en clases de equivalencias del modo siguiente: sean $f, g \in \mathcal{F}_\mu(E)$. Si definimos

$$f \sim g \qquad \text{si, y sólo si,} \qquad \mu\big(\{x \in E : f(x) \neq g(x)\}\big) \ = \ 0$$

entonces $\sim$ es una relación de equivalencia sobre $\mathcal{F}_\mu(E)$ y, por lo tanto, el conjunto

$$\mathcal{F}_\mu(E)/\sim \ = \ \big\{[f] \ : \ f \in \mathcal{F}_\mu(E)\big\}$$

particiona a $\mathcal{F}_\mu(E)$ en clases de equivalencias, donde hemos puesto

$$[f] \ = \ \big\{g \in \mathcal{F}_\mu(E) \ : \ f \sim g\big\}$$

para cada $f \in \mathcal{F}_\mu(E)$. En ciertas ocasiones escribiremos $\mathcal{L}^0(E,\mu)$ para denotar el conjunto $\mathcal{F}_\mu(E)/\sim$.

7.2. Propiedades Básicas

En esta sección se demuestran las propiedades básicas de las funciones medibles; es otras palabras, se comprueba que $\mathcal{F}_\mu(E)$ es, en realidad, un espacio vectorial sobre $\mathbb{R}$. De hecho, $\mathcal{F}_\mu(E)$ es una *álgebra* de funciones. Más aun, límites puntuales de sucesiones de funciones medibles permanecen medibles, etc.

Teorema 7.2.1 (Propiedades Algebraicas). *Sean $f, g \in \mathcal{F}_\mu(E)$ y sea $\lambda \in \mathbb{R}$. Entonces*

(1) $f + g \in \mathcal{F}_\mu(E)$,

(2) $\lambda f \in \mathcal{F}_\mu(E)$, y

(3) $f \cdot g$, $|f| \in \mathcal{F}_\mu(E)$.

Prueba. (1) Sea $a \in \mathbb{R}$. Si $f(x) + g(x) > a$, entonces $a - g(x) < f(x)$ y usemos el hecho de que $\mathbb{Q}$ es denso en $\mathbb{R}$ para elegir un $r \in \mathbb{Q}$ tal que $a - g(x) < r < f(x)$. De esto se sigue que

$$\big\{x \in E \ : \ f(x) + g(x) > a\big\} = \bigcup_{r \in \mathbb{Q}} \Big(\big\{x \in E \ : \ f(x) > r\big\} \cap \big\{x \in E \ : \ g(x) > a - r\big\}\Big)$$

es medible y, en consecuencia, $f + g \in \mathcal{F}_\mu(E)$.

(2) Es inmediata.

(3) En primer lugar vamos a probar que $f^2 \in \mathcal{F}_\mu(E)$. En efecto, sea $a \in \mathbb{R}$. Si $a < 0$, entonces $\big\{x \in E \ : \ f^2(x) > a\big\} = E \in \mathfrak{M}_\mu(E)$. Por otro lado, si $a \geq 0$, entonces

$$\big\{x \in E \ : \ f^2(x) > a\big\} = \big\{x \in E \ : \ f(x) > \sqrt{a}\big\} \cup \big\{x \in E \ : \ f(x) < -\sqrt{a}\big\} \in \mathfrak{M}_\mu(\mathbb{R}).$$

Esto prueba que $f^2 \in \mathcal{F}_\mu(E)$ y, por consiguiente, por (1) y (2),

$$f \cdot g \ = \ \frac{1}{4}\Big[(f + g)^2 \ - \ (f - g)^2\Big] \ \in \ \mathcal{F}_\mu(E).$$

Finalmente, como el conjunto

$$\{x \in E : |f|(x) < a\} = \begin{cases} \varnothing & \text{si} \quad a \leq 0 \\ \{x \in E : f(x) < a\} \cap \{x \in E : f(x) > -a\} & \text{si} \quad a > 0, \end{cases}$$

es medible, se concluye que $|f| \in \mathcal{F}_\mu(E)$ y termina la prueba. ■

Del resultado anterior se sigue inmediatamente que si $f, g \in \mathcal{F}_\mu(E)$, entonces

(a) $\{x \in E : f(x) < g(x)\} = \{x \in E : (g - f)(x) > 0\} \in \mathfrak{M}_\mu(E)$, y

(b) $\{x \in E : f(x) = g(x)\} = \{x \in E : (g - f)(x) = 0\} \in \mathfrak{M}_\mu(E)$.

Sabemos que la necesidad de considerar límites de sucesiones o series de funciones es fundamental en Matemáticas. Resulta, en consecuencia, importante determinar si el límite de una sucesión de funciones medibles permanece siendo medible. Un escenario general para indagar sobre una respuesta comienza con el siguiente resultado.

Teorema 7.2.2. *Sea* $(f_n)_{n=1}^\infty \subseteq \mathcal{F}_\mu(E)$. *Entonces:*

(a) $\inf f_n,\ \sup f_n \in \mathcal{F}_\mu(E)$.

(b) $\liminf f_n,\ \limsup f_n \in \mathcal{F}_\mu(E)$.

Prueba. Sea $a \in \mathbb{R}$ y defina

$$g(x) = \sup\{f_n(x) : n = 1, 2, \ldots\} \qquad \text{y} \qquad h(x) = \inf\{f_n(x) : n = 1, 2, \ldots\}$$

para todo $x \in E$. Veamos que, para cada $\alpha \in \mathbb{R}$, se cumple que

$$g^{-1}((\alpha, +\infty]) = \bigcup_{n=1}^\infty f_n^{-1}((\alpha, +\infty]) \qquad \text{y} \qquad h^{-1}((\alpha, +\infty]) = \bigcap_{n=1}^\infty f_n^{-1}((\alpha, +\infty]).$$

En efecto, sea $x \in g^{-1}((\alpha, +\infty])$. Entonces $g(x) > \alpha$ y se sigue de las propiedades del supremo que existe por lo menos un $m \in \mathbb{N}$ tal que $\alpha < f_m(x) \leq g(x)$. De esto se sigue que $x \in f_m^{-1}((\alpha, +\infty]) \subseteq \bigcup_{n=1}^\infty f_n^{-1}((\alpha, +\infty])$ y, así,

$$g^{-1}((\alpha, +\infty]) \subseteq \bigcup_{n=1}^\infty f_n^{-1}((\alpha, +\infty]).$$

Recíprocamente, si $x \in \bigcup_{n=1}^\infty f_n^{-1}((\alpha, +\infty])$, entonces $x \in f_m^{-1}((\alpha, +\infty])$ para algún $m \in \mathbb{N}$ y, en consecuencia, $g(x) \geq f_m(x) > \alpha$, es decir, $x \in g^{-1}((\alpha, +\infty])$. Esto prueba que

$$\bigcup_{n=1}^\infty f_n^{-1}((\alpha, +\infty]) \subseteq g^{-1}((\alpha, +\infty]).$$

Por lo tanto,

$$g^{-1}((\alpha, +\infty]) = \bigcup_{n=1}^\infty f_n^{-1}((\alpha, +\infty]) \in \mathfrak{M}_\mu(E).$$

y, en consecuencia, g es medible. Un argumento enteramente similar nos revela que h es medible y termina la prueba de (a). Para demostrar (b), observe que como

$$\liminf f_n = \sup_{n\geq 1} \inf_{k\geq n} f_n \qquad \text{y} \qquad \limsup f_n = \inf_{k\geq n} \sup_{k\geq n} f_n,$$

entonces $\liminf f_n$ y $\limsup f_n$ son, por la primera parte, medibles. La prueba es completa. ■

Nótese que por (a) del resultado anterior se tiene que si $f_1, \ldots, f_n \in \mathcal{F}_\mu(E)$, entonces las funciones

$$f = \min\{f_1, \ldots, f_n\} \qquad \text{y} \qquad g = \max\{f_1, \ldots, f_n\}$$

ambas pertenecen a $\mathcal{F}_\mu(E)$. Suponga ahora que $(f_n)_{n=1}^{\infty}$ es una sucesión en $\mathcal{F}_\mu(E)$ que *converge puntualmente* a una función f. Puesto que para cada $x \in E$,

$$\liminf_{n\to\infty} f_n(x) = \limsup_{n\to\infty} f_n(x) = f(x),$$

resulta del Teorema 7.2.2 que f es medible. Pero, ¿qué ocurre si la sucesión $(f_n)_{n=1}^{\infty}$ converge casi-siempre a f? ¿Podemos afirmar que $f \in \mathcal{F}_\mu(E)$? La respuesta es, por supuesto, afirmativa. En efecto, observe que

$$\begin{aligned} D_f &= \{x \in E : \lim_{n\to\infty} f_n(x) = f(x) \text{ existe}\} \\ &= \{x \in E : \liminf_{n\to\infty} f_n(x) = f(x) = \limsup_{n\to\infty} f_n(x)\} \end{aligned}$$

es, por el teorema anterior, un subconjunto medible de E, mientras que

$$E \setminus D_f = \{x \in E : \lim_{n\to\infty} f_n(x) \text{ no existe}\}$$

posee, por definición, medida cero. Esto prueba que f es medible. Más aun, en el escenario de la Teoría de Integración de Lebesgue, puesto que los conjuntos medibles de medida cero no afectan la integración, resulta que podemos redefinir $f(x)$ como queramos para los $x \in E \setminus D_f$. Por esta razón, y partir de ahora:

> *Si $(f_n)_{n=1}^{\infty}$ es una sucesión en $\mathcal{F}_\mu(E)$ que* ***converge casi-siempre*** *a una función f, convenimos en redefinir $f(x) = 0$ para todo $x \in E \setminus D_f$, salvo que explícitamente declaremos lo contrario.*

Con esta observación se concluye que:

Corolario 7.2.3. *Si $(f_n)_{n=1}^{\infty}$ es una sucesión en $\mathcal{F}_\mu(E)$ que* ***converge casi-siempre*** *sobre E a una función $f : E \to \overline{\mathbb{R}}$, entonces $f \in \mathcal{F}_\mu(E)$.*

Recordemos que $\mathcal{B}_1([a,b])$ representa *el espacio vectorial de todas las funciones $f : [a,b] \to \mathbb{R}$ que son de la primera clase de Baire*, es decir, $f \in \mathcal{B}_1([a,b])$ si, y sólo si, existe una sucesión $(f_n)_{n=1}^{\infty}$ de funciones continuas tal que $f(x) = \lim_{n\to\infty} f_n(x)$ para cada $x \in [a,b]$. Se sigue entonces del resultado anterior que

$$\mathcal{B}_1([a,b]) \subseteq \mathcal{F}_\mu([a,b]).$$

Teorema 7.2.4. *Sea $f : E \to \mathbb{R}$ una función arbitraria y suponga que $(E_n)_{n=1}^{\infty}$ es una sucesión de conjuntos medibles tal que $E = \bigcup_{n=1}^{\infty} E_n$. Si $f|_{E_n}$ es medible para cada $n \geq 1$, entonces $f \in \mathcal{F}_\mu(E)$.*

Prueba. Es fácil verificar que la igualdad

$$\{x \in E : f(x) > a\} = \bigcup_{n=1}^{\infty} \{x \in E_n : f(x) > a\} = \bigcup_{n=1}^{\infty} \{x \in E_n : f|_{E_n}(x) > a\}$$

se cumple que para cada $a \in \mathbb{R}$. Esto prueba que f es medible. ■

Ya hemos visto que toda función continua $f : E \to \mathbb{R}$ es medible. Pero, ¿qué ocurre si f es continua casi-siempre? La respuesta, como era de esperarse, es que también en este caso f es medible.

Teorema 7.2.5. *Si $f : E \to \overline{\mathbb{R}}$ es **continua casi-siempre** sobre $E \in \mathfrak{M}_\mu(\mathbb{R})$, entonces $f \in \mathcal{F}_\mu(E)$.*

Prueba. Puesto que $f : E \to \mathbb{R}$ es continua casi-siempre en E, el conjunto $\mathrm{Disc}(f)$, de sus puntos de discontinuidades, posee medida cero y, por consiguiente, como μ es una medida completa todos los subconjuntos de $\mathrm{Disc}(f)$ son medibles. Sea $a \in \mathbb{R}$. Como

$$\{x \in E : f(x) > a\} = \{x \in E \setminus \mathrm{Disc}(f) : f(x) > a\} \cup \{x \in \mathrm{Disc}(f) : f(x) > a\},$$

resulta que $\{x \in \mathrm{Disc}(f) : f(x) > a\} \in \mathfrak{M}_\mu(E)$ ya que él es un subconjunto de $\mathrm{Disc}(f)$. Por lo tanto, para demostrar que $f \in \mathcal{F}_\mu(E)$ será suficiente comprobar que $E_a = \{x \in E \setminus \mathrm{Disc}(f) : f(x) > a\}$ es medible. Ahora bien, como f es continua en cada uno de los puntos de $E \setminus \mathrm{Disc}(f)$, resulta entonces que para cada $x \in E \setminus \mathrm{Disc}(f)$, existe un $\delta_x > 0$ tal que

$$f(t) > a \qquad \text{siempre que} \qquad t \in \big(E \setminus \mathrm{Disc}(f)\big) \cap (x - \delta_x, x + \delta_x).$$

Sea

$$G = \bigcup_{x \in E_a} (x - \delta_x, x + \delta_x).$$

y observe que como G es un conjunto abierto, entonces él es medible y, en consecuencia,

$$E_a = G \cap \big(E \setminus \mathrm{Disc}(f)\big) \in \mathfrak{M}_\mu(E).$$

La prueba es completa. ■

Ejemplo 7.2.1. *Del* Ejemplo 3.1.1 *sabemos que la función de Thomae $f : \mathbb{R} \to \mathbb{R}$ definida por*

$$f(x) = \begin{cases} \dfrac{1}{q} & si \quad x = p/q \in \mathbb{Q}^* \\ 0 & si \quad x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

*es continua en los irracionales pero discontinua en los racionales y, en consecuencia, f es **continua casi-siempre** ya que $\mu(\mathbb{Q}) = 0$. Por lo tanto, por el resultado anterior, f es **medible**.*

Combinado el Corolario 7.2.3 con el Teorema 7.2.5 se obtiene el siguiente resultado.

Teorema 7.2.6. *Si $f : [a, b] \to \mathbb{R}$ es **diferenciable casi-siempre** sobre $[a, b]$, entonces $f' \in \mathcal{F}_\mu([a, b])$.*

Prueba. Puesto que f es diferenciable casi-siempre ella es continua casi-siempre y entonces, del Teorema 7.2.5, se sigue que $f \in \mathcal{F}_\mu([a,b])$. Extienda f al intervalo $[a,b+1)$ definiendo $f(x) = f(b)$ para todo $x \in (b,b+1)$. Elija ahora una sucesión $(\delta_n)_{n=1}^\infty$ en $(0,1)$ que converja a 0 y defina, para cada $n \in \mathbb{N}$, la función $f_n : [a,b] \to \mathbb{R}$ por

$$f_n(x) \;=\; \frac{f(x+\delta_n) \;-\; f(x)}{\delta_n}.$$

Resulta que $(f_n)_{n=1}^\infty$ es una sucesión de funciones medibles que converge puntualmente a f' casi-siempre y, así, gracias al Corolario 7.2.3 vemos que $f' \in \mathcal{F}_\mu([a,b])$. ■

Una subclase importante de funciones en $\mathcal{F}_\mu(E)$ son las así llamadas funciones medibles según Borel.

Definición 7.2.7. *Una función $f : E \to \mathbb{R}$ se dice que es **medible Borel** sobre E si*

(a) E es un conjunto de Borel y

(b) para cada $a \in \mathbb{R}$, el conjunto $\{x \in E : f(x) > a\}$ es un conjunto de Borel.

Claramente cualquier función medible Borel es medible según Lebesgue. Similarmente, si $E \in \mathfrak{Bo}(\mathbb{R})$, entonces toda ***función continua*** $f : E \to \mathbb{R}$ es medible Borel. Observe que si f es medible Borel si, y sólo si, $f^{-1}(G)$ es un conjunto de Borel para cualquier conjunto abierto $G \subseteq \mathbb{R}$.

En el teorema sobre las propiedades algebraicas de $\mathcal{F}_\mu(E)$ no se incluyó la de la composición de funciones. La razón de este hecho es que, en general, la composición de dos funciones medibles no es necesariamente una función medible. Concretamente, la composición $g \circ f$ siendo f continua y g medible no es necesariamente medible. Veamos esto.

Ejemplo 7.2.2. *Existen funciones medibles Lebesgue cuya **composición no es medible Lebesgue**.*

Prueba. Sea Φ_Γ la función del Ejercicio 6.7.1, página 354, esto es,

$$\Phi_\Gamma(x) \;=\; x + \varphi_\Gamma(x)$$

para todo $x \in [0,1]$, donde φ_Γ es la función de Cantor. De dicho Ejercicio sabemos que $E = \Phi_\Gamma^{-1}(A)$ es un conjunto medible según Lebesgue pero que no es medible Borel, donde A es un subconjunto no medible incluido en $\Phi_\Gamma(\Gamma)$. Sea $f = \Phi_\Gamma^{-1}$ y considere la función $g : [0,1] \to \mathbb{R}$ definida por $g(x) = \chi_E(x)$. Observe que f y g son medibles ya que f es continua y E es medible. Sin embargo, $g \circ f \notin \mathcal{F}_\mu([0,2])$. En efecto,

$$\begin{aligned}\{x \in [0,2] : (g\circ f)(x) > 1/2\} \;&=\; \{x \in [0,2] : f(x) \in E\}\\ &=\; f^{-1}(E)\\ &=\; \Phi_\Gamma(E) \;=\; A \;\notin\; \mathfrak{M}_\mu(\mathbb{R}).\end{aligned}$$

Sin embargo, si se intercambian los papeles en el ejemplo anterior, es decir, si f es medible y g es continua, entonces $g \circ f$ siempre es medible. De hecho, vale el siguiente resultado, el cual muestra la importancia de las funciones medibles según Borel.

Teorema 7.2.8. *Sea $f : E \to \mathbb{R}$ una **función medible** y suponga que $g : \mathbb{R} \to \mathbb{R}$ es **medible Borel**. Entonces $g \circ f \in \mathcal{F}_\mu(E)$.*

Prueba. Para cada $a \in \mathbb{R}$,

$$\{x \in E \,:\, g(f(x)) > a\} \;=\; \big(g \circ f\big)^{-1}\big((a, +\infty)\big) \;=\; f^{-1}\big(g^{-1}\big((a, +\infty)\big)\big).$$

Observe ahora que, gracias a que g es medible Borel, el conjunto $g^{-1}\big((a, +\infty)\big)$ es un conjunto medible Borel. Se sigue del Teorema 7.1.3 que $f^{-1}\big(g^{-1}\big((a, +\infty)\big)\big)$ es medible Lebesgue. ■

7.2.1. Aproximación de Funciones Medibles

El objetivo de esta sección es obtener uno de los resultados fundamentales en la Teoría de Integración: aproximar cualquier función medible por medio de funciones simples. En la Teoría de la Integral de Riemann las funciones en escaleras, definidas sobre un intervalo cerrado y acotado $[a, b]$, son las piezas claves para tal desarrollo. Esas funciones asumen, por definición, sólo un número finito de valores siendo constante en el interior de cada uno de los intervalos cerrados asociados a una partición finita de $[a, b]$. Para la Integral de Lebesgue las funciones en escaleras son reemplazadas por funciones simples, una noción un poco más general que las funciones en escalera pues sustituyen intervalos por conjuntos medibles, pero que siguen asumiendo sólo un número finito de valores.

Fijemos un conjunto medible E. Recordemos que una **partición medible** de E es una colección finita $\{E_1, \ldots, E_n\}$ de subconjuntos medibles no-vacíos incluidos en E que son disjuntos dos a dos y tales que $E = E_1 \cup \cdots \cup E_n$.

Definición 7.2.9. *Una función $\varphi \in \mathcal{F}_\mu(E)$ se llama **función simple** si se puede escribir en la forma*

$$\varphi \;=\; \sum_{i=1}^{n} a_i \cdot \chi_{E_i},$$

donde $\{E_1, \ldots, E_n\} \subseteq \mathfrak{M}_\mu(\mathbb{R})$ es una partición medible de E y los a_i, $i = 1, \ldots, n$ son números reales distintos dos a dos.

Observe que cualquier función simple es acotada ya que ella sólo asume un número finito de valores. En general, existen muchas formas de representar a una función simple y, por supuesto, no se requiere, en principio, que los conjuntos medibles E_i formen una partición medible de E. Sin embargo, si los E_i no son disjuntos, entonces es fácil ver que si $a_1, \ldots, a_m$ son los valores distintos que φ asume, resulta que los conjuntos

$$F_i \;=\; \{x \in E : \varphi(x) = a_i\} \qquad i = 1, \ldots, m$$

son medibles, disjuntos dos a dos y se cumple que

$$\varphi \;=\; \sum_{i=1}^{n} a_i \cdot \chi_{F_i}.$$

Por esta razón a la representación de cualquier función simple como la expresada en la definición anterior se acostumbra denominarla la **representación canónica** o **forma canónica** de φ.

La ventaja de expresar cada función simple como en la definición anterior, es que dicha representación es única. En lo que sigue denotaremos por $\mathcal{S}_\mu(E)$ *el conjunto de todas la funciones simples* $\varphi : E \to \mathbb{R}$. Es un ejercicio sencillo establecer que $\mathcal{S}_\mu(E)$ es un subespacio vectorial de $\mathcal{F}_\mu(E)$. En efecto, sean $\varphi, \psi \in \mathcal{S}_\mu(E)$ escritas en sus formas canónicas, esto es,

$$\varphi \;=\; \sum_{i=1}^{n} a_i \cdot \chi_{E_i} \qquad \text{y} \qquad \psi \;=\; \sum_{j=1}^{m} b_j \cdot \chi_{F_j}$$

donde $\{E_1, \ldots, E_n\}$ y $\{F_1, \ldots, F_m\}$ son particiones medibles de E y los correspondientes números a_i son distintos entre sí, así como también los b_j. De esto se sigue que los conjuntos $E_i \cap F_j \in \mathfrak{M}_\mu(\mathbb{R})$ son disjuntos y

$$(\varphi + \psi)(x) \;=\; a_i + b_j \qquad \text{para todo } x \in E_i \cap F_j.$$

Por esto,

$$\varphi + \psi \;=\; \sum_{i=1}^{n} \sum_{j=1}^{m} (a_i + b_j) \cdot \chi_{E_i \cap F_j} \;\in\; \mathcal{S}_\mu(E).$$

El hecho de que $a \cdot \varphi \in \mathcal{S}_\mu(E)$ para cualquier $a \in \mathbb{R}$ se comprueba más fácilmente. Observe, además, que

$$\varphi \cdot \psi \;=\; \sum_{i=1}^{n} \sum_{j=1}^{m} a_i b_j \cdot \chi_{E_i \cap F_j} \;\in\; \mathcal{S}_\mu(E).$$

Suponga que $\varphi = \sum_{i=1}^{n} a_i \cdot \chi_{E_i}$ es una función simple. Si los conjuntos E_i que aparecen en la definición de φ son *intervalos de longitud finita*, entonces nuestra función φ coincide con la noción de **función en escalera** estudiada con anterioridad. Usaremos, en lo sucesivo, el símbolo $\mathrm{Esc}(E)$ para denotar el conjunto de las funciones en escaleras $f : E \to \mathbb{R}$.

Como ya habíamos informado, el propósito de esta sección es probar uno de los teoremas fundamentales sobre funciones medibles: *las funciones simples aproximan cualquier función medible*. Éste resultado es crucial y sirve para muchos propósitos. Por ejemplo, es de utilidad para demostrar el Teorema de Lusin, también se utiliza para probar numerosos resultados importantes en la Teoría de Integración tales como el Teorema de la Convergencia Monótona y, en consecuencia, extender la noción de integral de funciones medibles acotadas a funciones medibles no acotadas, etc.

Teorema 7.2.10 (Densidad de las Funciones Simples). *Sea $f : E \to \overline{\mathbb{R}}$ una función medible **no-negativa**. Entonces existe una sucesión de funciones $(\varphi_n)_{n=1}^{\infty}$ en $\mathcal{S}_\mu(E)$ tal que*

(a) $0 \leq \varphi_1 \leq \cdots \leq \varphi_n \leq \cdots \leq f$ *y*

(b) $\lim_{n \to \infty} \varphi_n(x) = f(x)$ *para cada* $x \in E$.

*Más aun, si f es **acotada**, la convergencia en (b) es **uniforme**.*

Prueba. Para cada $n \in \mathbb{N}$, construya la partición $\mathcal{P}_n = \{I_{n,k} : k = 1, 2, \ldots, n2^n\}$ en el intervalo $[0, n)$ formada por $n2^n$ subintervalos cada uno de longitud $1/2^n$, es decir,

$$I_{n,k} \;=\; \left[\frac{k-1}{2^n}, \frac{k}{2^n}\right), \qquad \text{para } k = 1, 2, \ldots, n2^n.$$

Sea $I_n = [n, +\infty)$ y defina los conjuntos E_n y $E_{n,k}$ como sigue:

$$E_n = f^{-1}(I_n) = \{x \in E : f(x) \geq n\},$$

$$E_{n,k} = f^{-1}(I_{n,k}) = \left\{x \in E : \frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n}\right\}, \qquad \text{para } k = 1, 2, \ldots, n2^n.$$

Como f es medible, los conjuntos E_n y $E_{n,k}$ que acabamos de definir son medibles y, además, disjuntos dos a dos. Considere ahora, para cada $n \geq 1$, la función simple $\varphi_n : E \to \mathbb{R}$ definida por

$$\varphi_n = \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \chi_{E_{n,k}} + n\chi_{E_n}.$$

Veamos que la sucesión $(\varphi_n)_{n=1}^{\infty}$ satisface las condiciones (a) y (b).

(a) Para establecer que $(\varphi_n)_{n=1}^{\infty}$ es una sucesión creciente y no-negativa, observe primeramente que $\varphi_n \geq 0$ para todo $n \geq 1$. Para cada $n \in \mathbb{N}$ y cada $k = 1, 2, \ldots, n2^n$ se tiene que

$$I_{n,k} = I_{n+1,2k-1} \cup I_{n+1,2k} = \left[\frac{k-1}{2^n}, \frac{k-1}{2^n} + \frac{1}{2^{n+1}}\right) \cup \left[\frac{k-1}{2^n} + \frac{1}{2^{n+1}}, \frac{k}{2^n}\right),$$

y, por lo tanto,

$$\begin{aligned} E_{n,k} &= f^{-1}\left(\left[\frac{k-1}{2^n}, \frac{k-1}{2^n} + \frac{1}{2^{n+1}}\right)\right) \cup f^{-1}\left(\left[\frac{k-1}{2^n} + \frac{1}{2^{n+1}}, \frac{k}{2^n}\right)\right) \\ &= E_{n+1,2k-1} \cup E_{n+1,2k} \end{aligned}$$

Fijemos $n \geq 1$ y sea $x \in E$. Entonces existe un único k tal que $x \in E_{n,k}$. Si x pertenece a $E_{n+1,2k-1}$, resultará que

$$\frac{k-1}{2^n} \leq f(x) < \frac{k-1}{2^n} + \frac{1}{2^{n+1}} \qquad \text{y} \qquad \varphi_n(x) = \frac{k-1}{2^n} = \varphi_{n+1}(x),$$

mientras que si $x \in E_{n+1,2k}$, tendremos que $\dfrac{k-1}{2^n} + \dfrac{1}{2^{n+1}} \leq f(x) < \dfrac{k}{2^n}$ y, así,

$$\varphi_n(x) = \frac{k-1}{2^n} < \frac{k-1}{2^n} + \frac{1}{2^{n+1}} = \varphi_{n+1}(x)$$

En cualquier caso, $\varphi_n \leq \varphi_{n+1}$ para cada $n \in \mathbb{N}$.

(b) Para demostrar que $\varphi_n \to f$ puntualmente, considere, en primer lugar, el conjunto

$$E_f^{\infty} = \bigcap_{n=1}^{\infty} E_n = \{x \in E : f(x) = +\infty\}.$$

Si $x \in E_f^{\infty}$, entonces $x \in E_n$ para todo $n \in \mathbb{N}$ y, en consecuencia, $\varphi_n(x) = n$, de donde se sigue que

$$f(x) = \lim_{n\to\infty} \varphi_n(x) = +\infty \qquad \text{para cada} \quad x \in E_f^{\infty}.$$

Suponga ahora que $x \notin E_f^\infty$, esto es, $x \in E \setminus E_\infty = \bigcup_{n=1}^\infty (E \setminus E_n)$. Puesto que $E_1 \supseteq E_2 \supseteq \cdots \supseteq E_n \supseteq \cdots$, resulta que la sucesión $(E \setminus E_n)_{n=1}^\infty$ es creciente, de donde se deduce la existencia un $N \in \mathbb{N}$ tal que $x \in E \setminus E_n$ para todo $n \geq N$, es decir, $f(x) < n$ para todo $n \geq N$. Por esto,

$$x \in \{x \in E : 0 \leq f(x) < n\} = \bigcup_{k=1}^{n2^n} E_{n,k}$$

para cada $n \geq N$. Fijemos un $n \geq N$ y observe que como los conjuntos $E_{n,k}$ son disjuntos, existe un único $k \in \{1, 2, \ldots, n2^n\}$ tal que $x \in E_{n,k}$ y, por lo tanto,

$$\frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n} \qquad \text{y} \qquad \varphi_n(x) = \frac{k-1}{2^n}.$$

De lo anterior se concluye que

$$0 \leq f(x) - \varphi_n(x) < \frac{1}{2^n}$$

para todo $n \geq N$, lo cual es equivalente a afirmar que $\varphi_n(x) \to f(x)$. Puesto que $\mu(E_f^\infty) = 0$, concluimos que $\varphi_n \to f$ casi-siempre. Finalmente, si f es acotada, entonces de lo demostrado anteriormente resulta claro que $\varphi_n \to f$ uniformemente. La prueba es completa. ■

Observe que no hay nada de especial en usar particiones diádicas en la prueba del resultado anterior. Sólo se hizo en beneficio de la elegancia y simplicidad de la misma.

Sea $f : E \to \overline{\mathbb{R}}$ una función medible. Recordemos que las partes positivas y negativas de f son las funciones f^+ y f^- definidas por

$$f^+(x) = \text{máx}\{f(x), 0\} \qquad \text{y} \qquad f^-(x) = \text{mín}\{f(x), 0\}$$

para todo $x \in E$. Resulta que f^+ y f^- son medibles, no-negativas y se cumple que

$$f = f^+ - f^- \qquad \text{y} \qquad |f| = f^+ + f^-.$$

Usemos el Teorema 7.2.10 para seleccionar sucesiones crecientes, simples y no-negativas, digamos $(\varphi_n)_{n=1}^\infty$ y $(\phi_n)_{n=1}^\infty$ tales que

$$\lim_{n\to\infty} \varphi_n = f^+ \qquad \text{y} \qquad \lim_{n\to\infty} \phi_n = f^-.$$

Resulta que si definimos $\psi_n = \varphi_n - \phi_n$ para cada $n \geq 1$, se tiene el siguiente resultado.

Corolario 7.2.11 (Densidad de $\mathcal{S}_\mu(E)$). *Sea $f : E \to \mathbb{R}$ una **función medible**. Entonces existe una sucesión $(\psi_n)_{n=1}^\infty$ de funciones simples tal que*

(a) $|\psi_n| \leq |f|$ para todo $n \geq 1$ y

(b) $\lim_{n\to\infty} \psi_n(x) = f(x)$ para cada $x \in E$.

*Más aun, si f es **acotada**, la convergencia en (b) es **uniforme**.*

Sea $f : E \to \mathbb{R}$ una función medible. La **función signo de** f que denotaremos por $\text{sign}(f) : E \to \mathbb{R}$ se define como

$$\text{sign}(f)(x) = \begin{cases} 1 & \text{si } f(x) > 0 \\ 0 & \text{si } f(x) = 0 \\ -1 & \text{si } f(x) < 0 \end{cases}$$

Observe que $\text{sign}(f)$ es una función simple que cumple $|f| = f \cdot \text{sign}(f)$.

7.2.2. Los Teoremas de Severini-Egoroff y de Lusin

El objetivo de esta sección es presentar, fundamentalmente, dos resultados fantásticos. En términos generales, dichos resultados establecen que: (1°) la convergencia puntual de sucesiones de funciones medibles es *casi* convergencia uniforme (Teorema de Severini-Egoroff, también conocido como el **Tercer Principio de Littlewood**) y (2°) toda función medible es *casi* una función continua (Teorema de Lusin, al que también se le denomina el **Segundo Principio de Littlewood**). La palabra *casi* tiene, en este contexto, un significado muy especial: significa que existe, en cada caso, un conjunto medible que difiere muy poco del dominio de las funciones, donde la convergencia es uniforme, respectivamente, la restricción de la función a dicho conjunto es continua.

El siguiente resultado, también conocido bajo el nombre de Teorema de Egoroff, fue primeramente demostrado por Carlo Severini (1872-1951) en 1910. Debido al hecho, tal vez, de estar escrito en italiano permaneció casi desconocido fuera de Italia. Un año después, Dmitri Egoroff (1869-1931) publicó dicho resultado obtenido de manera independiente, el cual comienza a ser ampliamente conocido bajo ese nombre.

Teorema 7.2.12 (Severini-Egoroff). *Sea $E \in \mathfrak{M}_\mu(\mathbb{R})$ con $\mu(E) < +\infty$ y suponga que $(f_n)_{n=1}^\infty$ es una sucesión de funciones medibles a valores reales definidas sobre E. Si $(f_n)_{n=1}^\infty$* ***converge casi-siempre*** *a una función $f : E \to \mathbb{R}$, entonces para cada $\varepsilon > 0$, existe un conjunto medible $F_\varepsilon \subseteq E$ tal que*

$$\mu(E \setminus F_\varepsilon) < \varepsilon \qquad \text{y} \qquad f_n \to f \ \textbf{uniformemente} \text{ sobre } F_\varepsilon.$$

Prueba. Por el Corolario 7.2.3 sabemos que la función $f \in \mathcal{F}_\mu(E)$. Sea

$$B = \Big\{ x \in E : \lim_{n\to\infty} f_n(x) = f(x) \in \mathbb{R} \Big\}$$

y para cada par de enteros positivos n, k considere el conjunto

$$B_{n,k} = \bigcap_{m=n}^{\infty} \Big\{ x \in B : \big| f_m(x) - f(x) \big| < \frac{1}{k} \Big\}.$$

Entonces

(1) $B_{n,k} \in \mathfrak{M}_\mu(\mathbb{R})$ para todo $n, k \in \mathbb{N}$,

(2) $B = \bigcup_{k=1}^\infty B_{n,k}$.

En efecto, es claro que $\bigcup_{k=1}^\infty B_{n,k} \subseteq B$. Para demostrar la otra inclusión, sea $x \in B$. Entonces $\lim_{n\to\infty} f_n(x) = f(x)$ y, por consiguiente, para cada $k \in \mathbb{N}$ existe un $n_0 \in \mathbb{N}$ tal que

$$\big| f_m(x) - f(x) \big| < \frac{1}{k} \qquad \text{para todo} \quad n \geq n_0.$$

Esto prueba que $x \in B_{n_0,k} \subseteq \bigcup_{k=1}^\infty B_{n,k}$.

(3) $\lim\limits_{n\to\infty} \mu(B_{n,k}) = \mu(B)$.

Para ver esto último, observe que, para cada $k \in \mathbb{N}$, la sucesión $(B_{n,k})_{n=1}^\infty$ es creciente y, en consecuencia, por el Teorema 6.3.45

$$\lim_{n\to\infty} \mu(B_{n,k}) = \mu(B).$$

Fijemos un $\varepsilon > 0$ elegido arbitrariamente. Puesto que $\mu(E) < +\infty$, resulta que $\mu(B) < +\infty$ y de la última parte se deduce que, para cada $k \in \mathbb{N}$, se puede seleccionar un $n_k \in \mathbb{N}$ tal que

$$\mu(B \setminus B_{n,k}) \ = \ \mu(B) \ - \ \mu(B_{n,k}) \ < \ \frac{\varepsilon}{2^k}, \quad \text{para todo } \ n \geq n_k.$$

Definamos

$$F_\varepsilon \ = \ \bigcap_{k=1}^{\infty} B_{n_k,k}.$$

Entonces $F_\varepsilon \in \mathfrak{M}_\mu(\mathbb{R})$ y como

$$B \setminus F_\varepsilon \ = \ \bigcap_{k=1}^{\infty} (B \setminus B_{n_k,k}),$$

se tiene que

$$\mu(B \setminus F_\varepsilon) \ \leq \ \sum_{k=1}^{\infty} \mu\big(B \setminus B_{n_k,k}\big) \ < \ \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} \ = \ \varepsilon.$$

Falta demostrar que $f_n \to f$ uniformemente sobre F_ε. Para este fin, escojamos un $k_0 \in \mathbb{N}$ de modo tal que $1/k_0 < \varepsilon$. Ahora bien, para todo $n \geq n_{k_0}$ y cualquier $x \in F_\varepsilon = \bigcap_{k=1}^{\infty} B_{n_k,k}$, resulta que $x \in B_{n_{k_0},k_0}$ y, en consecuencia,

$$\big|f_n(x) \ - \ f(x)\big| \ < \ \frac{1}{k_0} \ < \ \varepsilon.$$

Esto es, $\sup_{x \in F_\varepsilon} \big|f_n(x) \ - \ f(x)\big| < \varepsilon$ para todo $n \geq n_{k_0}$ y termina la prueba. ■

Observe que podemos reemplazar el conjunto medible F_ε en el Teorema de Severini-Egoroff por un conjunto abierto. En efecto, puesto que $F_\varepsilon \subseteq E$, la definición de medida exterior nos permite escoger un conjunto abierto $O \supseteq F_\varepsilon$ tal que

$$\mu(O) \ < \ \mu(F_\varepsilon) + \varepsilon.$$

Por esto, $\mu(E \setminus O) \leq \mu(E \setminus F_\varepsilon) < \varepsilon$.

El Teorema de Severini-Egoroff permite justificar la siguiente definición.

Definición 7.2.13. *Sea $(f_n)_{n=1}^{\infty}$ una sucesión de funciones medibles definidas sobre un conjunto medible E. Diremos que $(f_n)_{n=1}^{\infty}$ **converge casi-uniformemente** a una función medible f si, para cada $\varepsilon > 0$, existe un conjunto medible $F_\varepsilon \subseteq E$ para el cual*

$$\mu(E \setminus F_\varepsilon) \ < \ \varepsilon \qquad \text{y} \qquad f_n \to f \ \text{ uniformemente sobre } \ F_\varepsilon.$$

Observe que el recíproco del Teorema de Severini-Egoroff es válido, es decir, si E es medible con $\mu(E) < +\infty$ y $f_n \to f$ casi-uniformemente sobre E, entonces $f_n \to f$ casi-siempre. En efecto, para cada $n \in \mathbb{N}$, seleccione un conjunto medible $F_n \subseteq E$ tal que

$$\mu(E \setminus F_n) \ < \ 1/n \qquad \text{y} \qquad f_n \to f \quad \text{uniformemente sobre } F_n.$$

Sin perder generalidad, podemos asumir que la sucesión $\big(F_n\big)_{n=1}^{\infty}$ es creciente. Sea $F = \bigcup_{n=1}^{\infty} F_n$. Entonces, como $\mu(E) < +\infty$, se sigue del Teorema 6.3.47 que $\mu(E \setminus F) = \lim_{n\to\infty} \mu(E \setminus F_n) = 0$ y, en consecuencia, para cualquier $x \in F$ se tiene que $f_n(x) \to f(x)$. Esto termina la prueba.

Corolario 7.2.14. *Sea $E \in \mathfrak{M}_\mu(\mathbb{R})$ con $\mu(E) < +\infty$. Suponga que $(f_n)_{n=1}^\infty$ es una sucesión en $\mathcal{F}_\mu(E)$ y que $f \in \mathcal{F}_\mu(E)$. Son equivalentes:*

(1) $f_n \to f$ ***casi-siempre*** *sobre* E.

(2) $f_n \to f$ ***casi-uniformemente*** *sobre* E.

Nota Adicional 7.2.2 (1) Es importante destacar que el conjunto medible E en el Teorema de Severini-Egoroff debe tener **medida finita** pues, en caso contrario, la conclusión pudiera ser falsa. En efecto, tome $E = \mathbb{R}$ y defina, para cada $n \in \mathbb{N}$, la función $f_n : \mathbb{R} \to \mathbb{R}$ por

$$f_n(x) \;=\; \chi_{[n,n+1)}(x) \qquad \text{para todo } x \in \mathbb{R}.$$

Si consideramos la función $f : \mathbb{R} \to \mathbb{R}$ dada por $f(x) = 0$ para todo $x \in \mathbb{R}$, vemos que $f_n \to f$ puntualmente, pero la convergencia no es uniforme. Para ver esto último, sea $\varepsilon = 1$ y escojamos un conjunto $E \in \mathfrak{M}_\mu(\mathbb{R})$ tal que $\mu(E) < 1$. En este caso resulta que

$$(\mathbb{R} \setminus E) \cap [n, n+1) \;\neq\; \varnothing \qquad \text{para todo } n \geq 1.$$

Afirmamos que $(f_n)_{n=1}^\infty$ no converge uniformemente sobre $\mathbb{R} \setminus E$. En efecto, seleccione, por cada $n \geq 1$, un punto $x_n \in (\mathbb{R} \setminus E) \cap [n, n+1)$ y observe que

$$\big|f_n(x_n) \;-\; f(x_n)\big| \;=\; \big|f_n(x_n)\big| \;=\; 1.$$

(2) Otra observación que hay que enfatizar en el Teorema de Severini-Egoroff es que la condición $\mu(E \setminus F_\varepsilon) < \varepsilon$ no puede ser reemplazada por $\mu(E \setminus F_\varepsilon) = 0$. En efecto, defina la sucesión de funciones medibles $(f_n)_{n=1}^\infty$ por

$$f_n(x) \;=\; \begin{cases} 0 & \text{si } \; x = 0 \\ n & \text{si } \; 0 < x \leq 1/n \\ \dfrac{1}{n} & \text{si } \; 1/n < x \leq 1 \end{cases}$$

Se comprueba sin dificultad que $f_n \to 0$ puntualmente sobre $[0, 1]$. Si ahora se considera cualquier conjunto medible $F \subseteq [0, 1]$ con $\mu(F) = 1$, tendremos que $\max\{f_n(x) : x \in F\} = n$ y, por supuesto, la convergencia no puede ser uniforme sobre F.

(3) Otro aspecto que puede ser de algún interés en el Teorema de Severini-Egoroff es que el conjunto F_ε se puede elegir cerrado. En efecto, una vez obtenido el conjunto E_ε para el cual $f_n \to f$ uniformemente sobre E_ε y satisfaciendo $\mu(E \setminus E_\varepsilon) < \varepsilon/2$, escoja, haciendo uso del Teorema 6.3.56, página 294, un conjunto cerrado $F_\varepsilon \subseteq E_\varepsilon$ tal que $\mu(E_\varepsilon \setminus F_\varepsilon) < \varepsilon/2$. Es claro que $\mu(E \setminus E_\varepsilon) < \varepsilon$ y, además, $f_n \to f$ uniformemente sobre F_ε.

Garantizar convergencia uniforme de sucesiones de funciones que convergen puntualmente es una tarea difícil, sin embargo, el siguiente criterio es uno de los más conocido permitiendo convergencia uniforme sin utilizar ninguna herramienta sobre la Teoría de la Medida.

Teorema 7.2.15 (Dini). *Sea $K \subseteq \mathbb{R}$ **compacto**. Sea $f \in C(K)$ y suponga que $(f_n)_{n=1}^{\infty}$ es una sucesión en $C(K)$ tal que*

*(a) $f_n \to f$ **puntualmente** sobre K y*

*(b) $(f_n)_{n=1}^{\infty}$ es **decreciente**.*

*Entonces $f_n \to f$ **uniformemente** sobre K.*

Prueba. Para cada $n \geq 1$, defina $g_n = f_n - f$. Puesto que $(f_n)_{n=1}^{\infty}$ es decreciente, se tiene que $(g_n)_{n=1}^{\infty}$ es una sucesión decreciente de funciones continuas no-negativas que converge a 0 puntualmente. Sea ahora $\varepsilon > 0$ y considere, para cada $n \geq 1$, el conjunto

$$G_n = g_n^{-1}((-\infty, \varepsilon)).$$

Puesto que g_n es continua, G_n es un conjunto abierto, pero además, teniendo en cuenta que $g_n \geq g_{n+1}$, resulta que $G_n \subseteq G_{n+1}$. Veamos ahora que $K = \bigcup_{n=1}^{\infty} G_n$. Para ver esto, sea $x \in K$. Como $\lim_{n\to\infty} g_n(x) = 0$, existe un $N \in \mathbb{N}$ tal que $g_N(x) < \varepsilon$ y, en consecuencia, $x \in G_N \subseteq \bigcup_{n=1}^{\infty} G_n$. Esto prueba que $K = \bigcup_{n=1}^{\infty} G_n$ y, así, por compacidad, existen $G_{n_1}, \ldots, G_{n_k}$ tal que $K = \bigcup_{j=1}^{k} G_{n_j}$. Sin perder generalidad, podemos suponer que $n_1 < n_2 < \cdots < n_k$ y, entonces, como $G_{n_1} \subseteq \cdots \subseteq G_{n_k}$, tenemos que $K = G_{n_k}$. Esto último significa que $g_{n_k}(x) < \varepsilon$ para todo $x \in K$, y puesto que la sucesión $(g_n)_{n=1}^{\infty}$ es decreciente, resulta que $g_n(x) < \varepsilon$ para todo $n \geq n_k$ y todo $x \in K$. Con esto queda demostrado que $\lim_{n\to\infty} g_n = 0$ uniformemente, o sea, que $f_n \to f$ uniformemente. La prueba es completa. ■

Observe que el resultado de Dini sigue siendo válido si cambiamos $C(K)$ por $C_c(\mathbb{R})$. En efecto, si $(f_n)_{n=1}^{\infty}$ es una sucesión no-negativa y decreciente en $C_c(\mathbb{R})$, entonces tomando $K = K_1$, donde K_1 es el soporte de f_1, resulta que la sucesión $(f_n)_{n=1}^{\infty} \in C(K)$ converge a 0 uniformemente sobre K.

Ya hemos visto que toda función continua $f : E \to \mathbb{R}$ es medible. El Teorema de Lusin establece que el recíproco es "*casi*" cierto en el siguiente sentido: toda función medible es *casi-continua*. En otras palabras, para cada $\varepsilon > 0$, existe un conjunto cerrado $F \subseteq E$ tal que $\mu(E \setminus F) < \varepsilon$ y $f|_F$ es continua sobre F. Un ingrediente fundamental en la demostración del resultado de Lusin es el siguiente lema.

Lema 7.2.16. *Sea $\varphi \in \mathcal{S}_\mu(E)$. Entonces, dado $\varepsilon > 0$, existe un **conjunto cerrado** $F \subseteq E$ tal que*

$$\mu(E \setminus F) < \varepsilon \qquad \text{y} \qquad \varphi|_F : F \to \mathbb{R} \quad \text{es **continua** sobre } F.$$

Prueba. Escribamos a φ en su forma canónica, esto es,

$$\varphi = \sum_{i=1}^{n} a_i \cdot \chi_{E_i}$$

donde $\{E_1, \ldots, E_n\}$ una partición medible y finita de E y los a_i son números reales distintos dos a dos. Por cada $i \in \{1, 2, \ldots, n\}$ seleccione, haciendo uso del Teorema 6.3.56, un conjunto cerrado $F_i \subseteq E_i$ tal que $\mu(E_i \setminus F_i) < \varepsilon/n$. Defina $F = F_1 \cup \cdots \cup F_n$. Entonces F es cerrado, $F \subseteq E$ y $\mu(E \setminus F) < \varepsilon$.

Veamos ahora que $\varphi|_F$ es continua sobre F. Sea $x \in F$ y suponga que $(x_n)_{n=1}^{\infty}$ es una sucesión en E que converge a x. Como $x \in F$, existe un $k \in \{1, \ldots, n\}$ tal que $x \in F_k$. Por otro lado, puesto que $\{F_1, \ldots, F_n\}$ es una familia disjunta, se sigue que ningún punto perteneciente a F_k puede ser un punto límite de ningún otro conjunto F_j con $k \neq j$, es decir, esto último lo que nos indica es que si $x \in F_k$ y $x_n \to x$, entonces debe existir un $N \in \mathbb{N}$ de modo tal que $x_n \in F_k$ para todo $n \geq N$ y, por lo tanto, $\varphi(x_n) = \varphi(x) = a_k$ para todo $n \geq N$. Esto prueba que $\varphi(x_n) \to \varphi(x)$, y así, φ es continua en x. La prueba es completa. ■

Teorema 7.2.17 (Lusin). *Sea $E \in \mathfrak{M}_\mu(\mathbb{R})$ con $\mu(E) < +\infty$ y sea $f \in \mathcal{F}_\mu(E)$. Dado $\varepsilon > 0$, existe un **conjunto cerrado** $F \subseteq E$ tal que*

$$\mu(E \setminus F) \ < \ \varepsilon \qquad \text{y} \qquad f|_F : F \to \mathbb{R} \quad \textit{es } \textbf{\textit{continua}} \textit{ sobre } F.$$

Prueba. Seleccionemos, aplicando el Corolario 7.2.11, una sucesión de funciones simples $(\varphi_n)_{n=1}^\infty$ tal que

$$f(x) \ = \ \lim_{n\to\infty} \varphi_n(x) \quad \text{para cada } x \in E.$$

Sea $\varepsilon > 0$. Como $\mu(E) < +\infty$, podemos invocar el Teorema de Severini-Egoroff para obtener un conjunto medible $K \subseteq E$ tal que

$$\mu(E \setminus K) \ < \varepsilon/2 \qquad \text{y} \qquad \varphi_n \to f \ \text{ uniformemente sobre } K.$$

Usemos ahora el Lema 7.2.16 para elegir, por cada $n \in \mathbb{N}$, un conjunto cerrado $F_n \subseteq K$ tal que

$$\mu(K \setminus F_n) \ < \varepsilon/2^{n+1} \qquad \text{y} \qquad \varphi|_{F_n} : F_n \to \mathbb{R} \ \text{ es continua sobre } F_n.$$

Definamos $F = \bigcap_{n=1}^\infty F_n$. Resulta que F es cerrado, $F \subseteq K$ y

$$\mu(K \setminus F) \ \leq \ \mu\Big(\bigcup_{n=1}^\infty K \setminus F_n\Big) \ \leq \ \sum_{n=1}^\infty \mu(K \setminus F_n) \ < \ \frac{\varepsilon}{2}.$$

De esto se sigue que $\mu(E \setminus F) \leq \mu(E \setminus K) + \mu(K \setminus F) < \varepsilon$. Más aun, como $\varphi_n|_{F_n}$ es continua sobre F_n, entonces con más razón $\varphi_n|_F$ es continua sobre F. Por otro lado, puesto que $\varphi_n \to f$ uniformemente sobre K, resulta que $\varphi_n|_F \to f|_F$ uniformemente sobre F y entonces, por una aplicación del Teorema 3.1.40, página 177, vemos que $f|_F$ es continua sobre F. ■

Nota Adicional 7.2.3 Es importante destacar que la conclusión del Teorema de Lusin no es que f es continua sobre F, sino que la **restricción de f sobre el conjunto** F es continua sobre dicho conjunto. Esto, por supuesto, es una diferencia fundamental. Por ejemplo, sea $f = \chi_{\mathbb{Q}}$ la función de Dirichlet definida sobre todo $\mathbb{R}$. Fijemos $\varepsilon > 0$ y escojamos, usando el hecho de que $\mu(\mathbb{Q}) = 0$, un conjunto abierto G tal que

$$\mathbb{Q} \subseteq G \qquad \text{y} \qquad \mu(G) \ < \ \varepsilon.$$

Si ahora definimos $F = \mathbb{R} \setminus G$, resulta que $\mu(\mathbb{R} \setminus F) = \mu(G) < \varepsilon$, y puesto que $f|_F \equiv 0$, se tiene que $f|_F$ es continua. Sin embargo, $f = \chi_{\mathbb{Q}}$, considerada como una función sobre $\mathbb{R}$, no es continua sobre F.

Similar a la justificación de la noción de convergencia casi-uniforme, el Teorema 7.2.17 permite apoyar la siguiente definición.

Definición 7.2.18. *Sea $f \in \mathcal{F}_\mu(E)$. Se dice que f es* ***casi-continua*** *sobre E, si dado $\varepsilon > 0$, existe un conjunto medible $F \subseteq E$ tal que*

$$\mu(E \setminus F) < \varepsilon \qquad \text{y} \qquad f|_F \quad \text{es continua sobre } F.$$

La distinción entre una función que es casi-continua y una que es continua casi-siempre es ahora clara: $f \in \mathcal{F}_\mu(E)$ es *casi-continua* sobre E significa que, dado $\varepsilon > 0$, existe un conjunto medible $F \subseteq E$ tal que

$$\mu(E \setminus F) < \varepsilon \qquad \text{y} \qquad f|_F \quad \text{es continua sobre } F.$$

Mientras que f es *continua casi-siempre* sobre E significa que existe un conjunto medible $F \subseteq E$ tal que

$$\mu(E \setminus F) = 0 \qquad \text{y} \qquad f \quad \text{es continua sobre } F.$$

Por consiguiente, toda función que es continua casi-siempre es casi-continua, pero no recíprocamente.

Un resultado importante que permite que ciertas funciones definidas sobre un subconjunto de $\mathbb{R}$ puedan ser extendidas, bajo ciertas condiciones, a una función definida sobre todo $\mathbb{R}$ es el siguiente.

Teorema 7.2.19 (Extensión de Tietze). *Sea $F \subseteq \mathbb{R}$* ***cerrado*** *y suponga que $f : F \to \mathbb{R}$ es una* ***función continua*** *sobre F. Entonces* ***existe una función continua*** *$g : \mathbb{R} \to \mathbb{R}$ tal que $g(x) = f(x)$ para todo $x \in F$.*

Prueba. Podemos suponer que el conjunto cerrado F no es un intervalo acotado ni superiormente ni inferiormente, pues en caso contrario una extensión trivial de f sería la función $g : \mathbb{R} \to \mathbb{R}$ definida por

$$g(x) = \begin{cases} f(x) & \text{si } x \in F \\ f(\sup F) & \text{si } x > \sup F \\ f(\inf F) & \text{si } x < \inf F \end{cases}$$

Sea $((a_n, b_n))_{n=1}^{\infty}$ una sucesión de intervalos abiertos y disjuntos cuya unión es $\mathbb{R} \setminus F$. Defina ahora la función $g : \mathbb{R} \to \mathbb{R}$ por

$$g(x) = \begin{cases} f(x) & \text{si } x \in F \\ \dfrac{f(b_n) - f(a_n)}{b_n - a_n}(x - a_n) + f(a_n) & \text{si } x \in (a_n, b_n) \subseteq \mathbb{R} \setminus F. \end{cases}$$

Es fácil verificar que g cumple con los requerimientos exigidos. ■

Es importante resaltar que es necesario que el conjunto F en el resultado anterior sea cerrado. En efecto, la función continua $f : (0,1) \to \mathbb{R}$ definida por $f(x) = 1/x$ no admite ninguna extensión continua a todo $\mathbb{R}$. Más aun, si la función f en el Teorema de Extensión de Tietze se elige acotada, entonces la extensión continua g se puede elegir poseyendo la misma cota que f.

Teorema 7.2.20 (Lusin). *Sea $f : \mathbb{R} \to \mathbb{R}$ una función arbitraria. Son equivalentes:*

(1) $f \in \mathcal{F}_\mu(\mathbb{R})$.

(2) *Para cada $\varepsilon > 0$, existe un **conjunto cerrado** F y una **función continua** $g : \mathbb{R} \to \mathbb{R}$ tal que*

$$\mu(\mathbb{R} \setminus F) < \varepsilon \qquad y \qquad f = g \quad \text{sobre} \quad F.$$

Prueba. (1) $\Rightarrow$ (2). Sea $(I_n)_{n=1}^\infty$ una enumeración de los intervalos en la colección $\{[n, n+1) : n \in \mathbb{Z}\}$. Dado $\varepsilon > 0$, por el Teorema 7.2.17 existe, por cada $n \in \mathbb{N}$, un conjunto cerrado $F_n \subseteq I_n$ tal que $\mu(I_n \setminus F_n) < \varepsilon/2^n$ y $f|_{F_n}$ es continua sobre F_n. Defina $F = \bigcup_{n=1}^\infty F_n$. Puesto que los F_n son cerrados y disjuntos dos a dos, F es cerrado,

$$\mu(\mathbb{R} \setminus F) \leq \sum_{n=1}^\infty \mu(I_n \setminus F_n) < \sum_{n=1}^\infty \frac{\varepsilon}{2^n} = \varepsilon,$$

y la función $f|_F$ es continua sobre F. Estamos ahora en condiciones de invocar el Teorema de Extensión de Tietze, el cual nos garantiza la existencia de una función continua $g : \mathbb{R} \to \mathbb{R}$ tal que $g = f$ sobre F.

(2) $\Rightarrow$ (1). Suponga que (2) se cumple. Entonces, para cada $n \in \mathbb{N}$, existe un conjunto cerrado F_n y una función continua $g_n : \mathbb{R} \to \mathbb{R}$ tal que $\mu(\mathbb{R} \setminus F_n) < 1/n$ y $g_n = f$ sobre F_n. Puesto que $f|_{F_n}$ es continua sobre F_n se tiene que ella es medible para cada $n \geq 1$. Considere el conjunto medible $F_0 = \mathbb{R} \setminus \bigcup_{n=1}^\infty F_n$, y observe que para todo $n \geq 1$

$$\mu(F_0) = \mu\bigg(\bigcap_{k=1}^\infty \mathbb{R} \setminus F_k\bigg) \leq \mu(\mathbb{R} \setminus F_n) < \frac{1}{n}.$$

Esto nos dice que $\mu(F_0) = 0$, de donde resulta que $f|_{F_0}$ también es medible y, entonces, por el Teorema 7.2.4, f es medible. ■

Observe que el conjunto $\mathbb{R} \setminus F$, en el resultado anterior, es exactamente

$$\mathbb{R} \setminus F = \{x \in \mathbb{R} : f(x) \neq g(x)\}.$$

El principio establecido en el Teorema de Luzin puede ser formulado de modo que recuerde muy de cerca el Teorema de Aproximación de Weierstrass.

Corolario 7.2.21 (Fréchet). *Si $f : [a, b] \to \mathbb{R}$ es una función medible, entonces existe una sucesión $(f_n)_{n=1}^\infty$ en $C([a, b])$ que **converge puntualmente** a f **casi-siempre** sobre $[a, b]$.*

Prueba. Por el Teorema de Lusin, para cada $n \in \mathbb{N}$, existe un conjunto cerrado $F_n \subseteq [a, b]$ y una función continua $f_n : [a, b] \to \mathbb{R}$ tal que

$$\mu([a, b] \setminus F_n) < \frac{1}{n} \qquad \text{y} \qquad f_n = f \quad \text{sobre} \quad F_n.$$

Observe que si $x \notin F_n$, entonces $|f_n(x) - f(x)| > 0$ de donde se sigue que para cada $k \in \mathbb{N}$,

$$\left\{x \in [a, b] : |f_n(x) - f(x)| \geq \frac{1}{k}\right\} \subseteq [a, b] \setminus F_n$$

y, por lo tanto,

$$0 \leq \lim_{n\to\infty} \mu\left(\left\{x \in [a,b] : |f_n(x) - f(x)| \geq \frac{1}{k}\right\}\right) \leq \lim_{n\to\infty} \mu([a,b] \setminus F_n) = 0$$

para cada $k \geq 1$. De esto último se concluye la existencia de una sucesión estrictamente creciente $(n_k)_{k=1}^{\infty}$ de enteros positivos tal que los conjuntos medibles

$$E_k = \left\{x \in [a,b] : |f_{n_k}(x) - f(x)| \geq \frac{1}{k}\right\}$$

satisfacen $\mu(E_k) < 1/2^k$ para todo $k \geq 1$. A partir de este punto, la prueba es idéntica a la del Teorema de Riesz pero no hace daño volver a repetirla. Defina

$$E = \limsup_{n\to\infty} E_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k.$$

Puesto que la sucesión $(\bigcup_{k=n}^{\infty} E_k)_{n=1}^{\infty}$ es decreciente, se sigue del Teorema 6.3.45 que

$$\begin{aligned} 0 \leq \mu(E) &= \lim_{n\to\infty} \mu\left(\bigcup_{k=n}^{\infty} E_k\right) \\ &\leq \lim_{n\to\infty} \sum_{k=n}^{\infty} \mu(E_k) \\ &< \lim_{n\to\infty} \frac{1}{2^{n-1}} = 0, \end{aligned}$$

es decir, $\mu(E) = 0$. Afirmamos que la sucesión $(f_{n_k})_{k=1}^{\infty}$ converge puntualmente a f sobre $[a,b] \setminus E$. En efecto, sea $x \in [a,b] \setminus E$ y sea $\varepsilon > 0$. Puesto que $x \notin E$, se sigue del Teorema 2.1.41, página 105, que existe un $N_1 \in \mathbb{N}$ tal que $x \notin E_k$ para todo $k \geq N_1$. Escojamos ahora un $N \geq N_1$ que satisfaga $1/N < \varepsilon$. Entonces, para todo $k \geq N$ se tiene que $x \notin E_k$ y

$$\left|f_{n_k}(x) - f(x)\right| < \frac{1}{k} < \varepsilon.$$

Esto completa la prueba. ■

Es importante tener presente la siguiente observación referente al resultado anterior. En la conclusión de dicho resultado hay que destacar que no podemos suprimir convergencia puntual casi-siempre sólo por convergencia puntual, pues en este caso tendríamos que las clases $\mathcal{B}_1([a,b])$ y $\mathcal{F}_\mu([a,b])$ serían idénticas lo cual es falso. ¡Existen funciones medibles que no son de la primera clase de Baire!. Sin embargo vale el siguiente resultado.

Corolario 7.2.22 (Fréchet). *Si $f : [a,b] \to \mathbb{R}$ es una función medible, entonces existe un conjunto F_σ, digamos F, tal que $\mu([a,b] \setminus F) = 0$ y $f|_F$ es de **la primera clase de Baire** sobre F.*

Prueba. Use el Teorema de Lusin para seleccionar, por cada $n \in \mathbb{N}$, un conjunto cerrado $F_n \subseteq [a,b]$ y una función continua $f_n : [a,b] \to \mathbb{R}$ tal que

$$\mu([a,b] \setminus F_n) < \frac{1}{n} \qquad \text{y} \qquad f_n = f \quad \text{sobre } F_n.$$

Defina $H_n = \bigcup_{j=1}^{n} F_j$, y observe que cada conjunto H_n es cerrado y f restringida a H_n es continua. Haciendo $F = \bigcup_{n=1}^{\infty} H_n$, resulta que F es un F_σ y $\mu([a,b] \setminus F) = 0$. Invocando el Teorema de Extensión de Tietze existe, por cada $n \in \mathbb{N}$, una función continua $f_n : \mathbb{R} \to \mathbb{R}$ tal que $f_n(x) = f(x)$ para todo $x \in H_n$. Es claro que $f_n \to f$ puntualmente sobre F y, por consiguiente, $f|_F$ es de la primera clase de Baire. Esto termina la prueba. ■

7.2.3. Convergencia en Medida

La siguiente noción de convergencia, dada por F. Riesz en 1909, será importante para establecer comparaciones con otros modos de convergencia. Una buena justificación para su existencia la daremos cuando hayamos definido el espacio $L_1(\mu)$ en el Capítulo 8 y se considere el Teorema de la Convergencia Acotada. Como siempre, en toda esta sección, asumiremos que E es un subconjunto medible de $\mathbb{R}$.

Definición 7.2.23. *Una sucesión $(f_n)_{n=1}^{\infty} \subseteq \mathcal{F}_\mu(E)$ se dice que* ***converge en medida*** *a una función $f \in \mathcal{F}_\mu(E)$ si, para cada $\varepsilon > 0$,*

$$\lim_{n\to\infty} \mu\big(\{x \in E : |f_n(x) - f(x)| \geq \varepsilon\}\big) = 0,$$

En este caso escribiremos $f_n \xrightarrow{\mu} f$.

El primer aspecto que hay que destacar, en relación a la convergencia en medida, es que la función límite es finita casi-siempre, esto es:

Teorema 7.2.24. *Sea $(f_n)_{n=1}^{\infty} \subseteq \mathcal{F}_\mu(E)$. Si $f_n \to f \in \mathcal{F}_\mu(E)$* ***en medida****, entonces f es* ***finita casi-siempre****.*

Prueba. Sea $\varepsilon > 0$ y para cada $n \geq 1$ pongamos $E_n = \bigcap_{k=n}^{\infty} \{x \in E : |f_k(x) - f(x)| \geq \varepsilon\}$. Observe que como $f_n \to f$ en medida y

$$E_n \subseteq \{x \in E : |f_n(x) - f(x)| \geq \varepsilon\} \quad \text{para todo } n \geq 1,$$

resulta que $\lim_{n\to\infty} \mu(E_n) = 0$ y, entonces, puesto que la sucesión $(E_n)_{n=1}^{\infty}$ es creciente, se sigue de la continuidad de la medida, Teorema 6.3.45, página 286, que

$$\mu\Big(\bigcup_{n=1}^{\infty} E_n\Big) = \lim_{n\to\infty} \mu(E_n) = 0.$$

Finalmente, para cualquier $x \in E \setminus \bigcup_{k=n}^{\infty} E_n$, se tiene que $|f_n(x) - f(x)| < \varepsilon$ para todo $n \geq 1$, de donde se deduce que $f(x)$ es finito y termina la prueba. ■

Cuando $\mu(E) = 1$, los probabilistas gustan llamar a la convergencia en medida **convergencia en probabilidad**. Los Teoremas de Riesz y Lebesgue dados más abajo, muestran claramente la diferencia entre las nociones de convergencia puntual y convergencia en medida.

Nuestro primer resultado el cual compara la convergencia simple con la convergencia en medida se debe a Lebesgue.

Teorema 7.2.25 (Lebesgue). *Sea $E \in \mathfrak{M}_\mu(\mathbb{R})$ con $\mu(E) < +\infty$. Si $f : E \to \mathbb{R}$ es una función medible y $(f_n)_{n=1}^\infty$ es una sucesión en $\mathcal{F}_\mu(E)$ convergiendo a f* ***puntualmente casi-siempre****, entonces $f_n \to f$* ***en medida****.*

Prueba. Fijemos un $\varepsilon > 0$ elegido de modo arbitrario y para cada $n \in \mathbb{N}$ considere el conjunto medible $E_n = \{x \in E : |f_n(x) - f(x)| \geq \varepsilon\}$. Sean

$$F_k = \bigcup_{n=k}^{\infty} E_n \qquad \text{y} \qquad F = \bigcap_{k=1}^{\infty} F_k.$$

Puesto que $f_n \to f$ casi-siempre, resulta que $\mu(F) = 0$. Más aun, teniendo en cuenta que $\mu(E) < +\infty$, se sigue entonces del Lema 6.3.48, página 287, que

$$0 = \mu(F) = \mu\big(\overline{\lim_{k\to\infty}}\, \mu(F_k)\big) \geq \overline{\lim_{k\to\infty}}\, \mu(F_k) \geq \overline{\lim_{n\to\infty}}\, \mu(E_n).$$

Esto muestra que $\lim_{n\to\infty} \mu(E_n) = 0$ y termina la prueba. ■

De nuevo, es importante advertir que el requerimiento $\mu(E) < +\infty$ en el Teorema 7.2.25 no se puede suprimir.

Ejemplo 7.2.3. *Sea $E = \mathbb{R}$ y considere, para cada $n \geq 1$, la función $f_n : \mathbb{R} \to \mathbb{R}$ definida por $f_n = \chi_{[n,n+1]}$ y sea $f = 0$. Entonces $f_n \to 0$* ***puntualmente*** *sobre $\mathbb{R}$, pero*

$$\mu\big(\{x \in \mathbb{R} : |f_n(x) - f(x)| \geq 1\}\big) = \mu([n, n+1]) = 1 \not\to 0.$$

Esto es, $f_n \not\to f$ ***en medida****.*

En el siguiente ejemplo se prueba que el recíproco del Teorema de Lebesgue es falso, es decir, convergencia en medida no implica, en general, convergencia puntual.

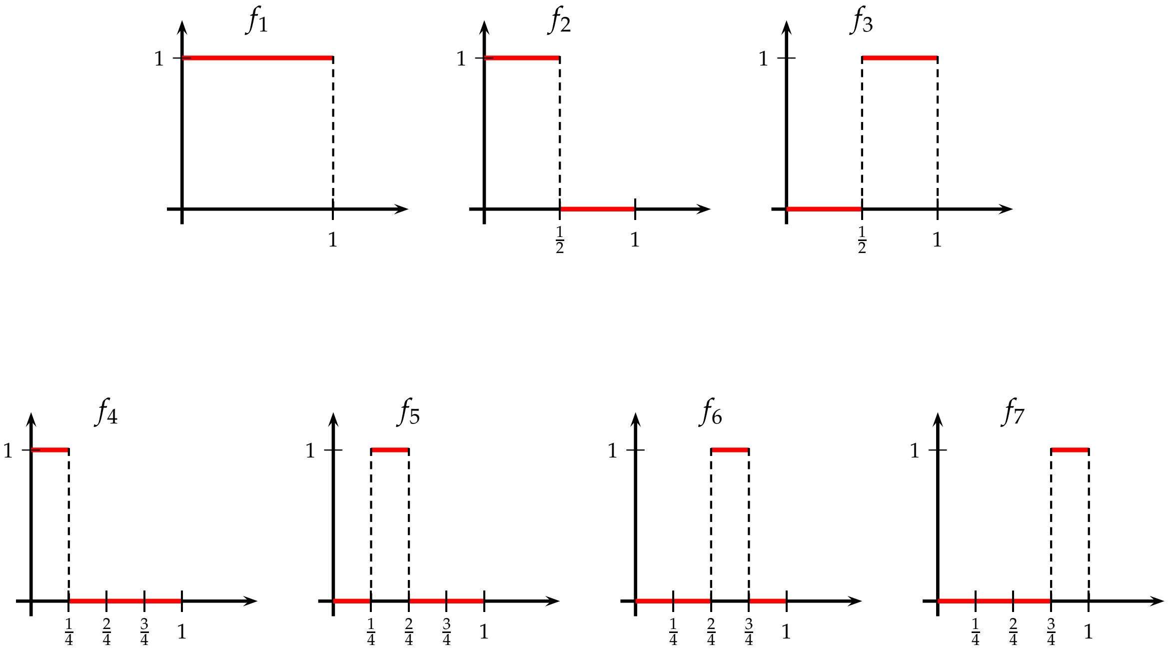
Ejemplo 7.2.4. *Existe una sucesión de funciones medibles definidas sobre $[0,1]$ que* ***converge en medida*** *pero* ***no converge*** *en ningún punto del intervalo.*

Prueba. Defina, para cada $n \in \mathbb{N}$, la función $f_n : [0,1] \to \mathbb{R}$ por

$$f_n = \chi_{\left[\frac{j}{2^k}, \frac{j+1}{2^k}\right]}, \quad \text{donde } n = j + 2^k, \;\; j = 0, 1, \ldots, 2^k - 1.$$

Esto es,

$$\begin{aligned}
f_1 &= \chi_{[0,1]} \\
f_2 &= \chi_{[0,1/2]}, \quad f_3 = \chi_{[1/2,1]} \\
f_4 &= \chi_{[0,1/2^2]}, \quad f_5 = \chi_{[1/2^2,2/2^2]}, \quad f_6 = \chi_{[2/2^2,3/2^2]}, \quad f_7 = \chi_{[3/2^2,1]} \\
&\vdots \\
f_{2^n} &= \chi_{[0,1/2^n]}, \quad f_{2^n+1} = \chi_{[1/2^n,2/2^n]}, \quad \ldots, \quad f_{2^n+(2^n-1)} = \chi_{[(2^n-1)/2^n,1]} \\
&\vdots
\end{aligned}$$



Es claro que, para cada $\varepsilon > 0$, se cumple que

$$\{x \in [0,1] : |f_n(x)| \geq \varepsilon\} = [j/2^n, (j+1)/2^n]$$

para todo $n \in \mathbb{N}$ y algún $0 \leq j < 2^n$. Por lo tanto,

$$\mu(\{x \in [0,1] : |f_n(x)| \geq \varepsilon\}) = 1/2^n.$$

De aquí se sigue que $f_n \xrightarrow{\mu} 0$. Por otro lado, si $x \in [0,1]$, la sucesión $(f_n(x))_{n=1}^{\infty}$ contiene infinitos 0's y un número infinito de 1's. En consecuencia, la sucesión $(f_n)_{n=1}^{\infty}$ nunca converge puntualmente sobre $[0,1]$. Observe, sin embargo, que la subsucesión $(f_{2^n})_{n=1}^{\infty}$ converge puntualmente a 0. ■

Aunque el ejemplo anterior no arroja ninguna duda, existe, sin embargo, un resultado donde el requerimiento de la existencia de una subsucesión $(f_{n_k})_{k=1}^{\infty}$ hace verdadero que convergencia en medida implique convergencia casi-siempre para alguna subsucesión.

Teorema 7.2.26 (Riesz). *Sea $(f_n)_{n=1}^{\infty} \subseteq \mathcal{F}_\mu(E)$ tal que $f_n \to f \in \mathcal{F}_\mu(E)$ en **medida**. Entonces existe una subsucesión $(f_{n_k})_{k=1}^{\infty}$ de $(f_n)_{n=1}^{\infty}$ tal que $f_{n_k} \to f$ **puntualmente casi-siempre** sobre E.*

Prueba. La construcción de la subsucesión la haremos usando inducción. Escoja $n_1 \in \mathbb{N}$ tal que

$$\mu\Big(\{x \in E : |f_{n_1}(x) - f(x)| \geq 1\}\Big) < \frac{1}{2}$$

y suponga que hemos construido los números $n_1 < n_2 < \cdots < n_k$. Seleccione ahora un n_{k+1} de modo tal que $n_{k+1} > n_k$ y

$$\mu\bigg(\Big\{x \in E : |f_{n_{k+1}}(x) - f(x)| \geq \frac{1}{k+1}\Big\}\bigg) < \frac{1}{2^{k+1}}.$$

Esto completa la construcción de la subsucesión $(f_{n_k})_{k=1}^{\infty}$. Para cada $k, n \geq 1$, sean

$$E_k = \left\{ x \in E : \left| f_{n_k}(x) - f(x) \right| \geq \frac{1}{k} \right\} \qquad \text{y} \qquad A_n = \bigcup_{k=n}^{\infty} E_k$$

y observe que

$$\mu(A_n) \leq \sum_{k=n}^{\infty} \mu(E_k) < \sum_{k=n}^{\infty} \frac{1}{2^k} = \frac{1}{2^{n-1}}$$

Definimos ahora

$$A = \limsup_{n\to\infty} E_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k = \bigcap_{n=1}^{\infty} A_n.$$

Puesto que la sucesión $(A_n)_{n=1}^{\infty}$ es decreciente y $\mu(A_1) < \sum_{k=1}^{\infty} 1/2^k < +\infty$, se sigue del Teorema 6.3.45 que

$$0 \leq \mu(A) = \lim_{n\to\infty} \mu(A_n) < \lim_{n\to\infty} \frac{1}{2^{n-1}} = 0,$$

es decir, $\mu(A) = 0$. Afirmamos que la sucesión $(f_{n_k})_{k=1}^{\infty}$ converge puntualmente a f sobre $E \setminus A$. En efecto, sea $x \in E \setminus A$ y sea $\varepsilon > 0$. Puesto que $x \notin A$, se sigue del Teorema 2.1.41, página 105, que existe un $N_1 \in \mathbb{N}$ tal que $x \notin A_k$ para todo $k \geq N_1$. Escojamos ahora un $N \geq N_1$ que satisfaga $1/N < \varepsilon$. Entonces, para todo $k \geq N$, se tiene que $x \notin A_k$ y

$$\left| f_{n_k}(x) - f(x) \right| < \frac{1}{k} < \varepsilon.$$

Esto completa la prueba. ∎

Una pequeña modificación en las hipótesis del Teorema de Riesz conduce al siguiente resultado.

Teorema 7.2.27. *Sean $f \in \mathcal{F}_\mu(E)$ y $(f_n)_{n=1}^{\infty} \subseteq \mathcal{F}_\mu(E)$. Suponga que*

*(a) f_n es una **función creciente** para cada $n \geq 1$ y*

(b) $f_n \xrightarrow{\mu} f$.

Entonces $\lim_{n\to\infty} f_n(x) = f(x)$ para cada $x \in \mathrm{PC}(f)$.

Prueba. Sea $x_0 \in \mathrm{PC}(f)$ y suponga que $f_n(x_0) \not\to f(x_0)$. Esto significa que existen un $\varepsilon > 0$ y una subsucesión $(f_{n_k})_{k=1}^{\infty}$ de $(f_n)_{n=1}^{\infty}$ tal que $|f_{n_k}(x_0) - f(x_0)| \geq \varepsilon$ para todo $k \geq 1$. Fijemos un tal k y observe que

$$f_{n_k}(x_0) - f(x_0) \geq \varepsilon \qquad \text{o} \qquad f_{n_k}(x_0) - f(x_0) \leq -\varepsilon. \tag{1}$$

Por otro lado, como f es continua en x_0, existe un $\delta > 0$ tal que

$$f(x_0) - \frac{\varepsilon}{2} < f(x) < f(x_0) + \frac{\varepsilon}{2}$$

siempre que $x \in V(x_0, \delta) = [a, b] \cap (x_0 - \delta, x_0 + \delta)$. Usando ahora la hipótesis (a) para la función f_{n_k} resulta que $f_{n_k}(x) \geq f_{n_k}(x_0)$ para $x > x_0$ y así, si la primera desigualdad en (1) se cumple, entonces

$$f_{n_k}(x) - f(x) > f_{n_k}(x_0) - f(x_0) - \frac{\varepsilon}{2} > \frac{\varepsilon}{2}$$

para todo $x \in (x_0, x_0 + \delta)$. Por esto,

$$\mu\big(\{x \in [a,b] : |f_{n_k}(x) - f(x)| > \varepsilon/2\}\big) \;\geq\; \delta,$$

lo cual es una contradicción. Esto termina la prueba. ■

Ya hemos visto que convergencia casi-uniforme implica convergencia casi-siempre. El siguiente resultado establece que ella también implica convergencia en medida.

Teorema 7.2.28. *Sean $f \in \mathcal{F}_\mu(E)$ y $(f_n)_{n=1}^\infty \subseteq \mathcal{F}_\mu(E)$. Si $f_n \to f$ **casi-uniformemente** sobre E, entonces $(f_n)_{n=1}^\infty$ **converge en medida** a f.*

Prueba. Para demostrar que $f_n \to f$ en medida, sea $\varepsilon > 0$ y seleccione un conjunto medible $F \subseteq E$ tal que

$$\mu(F) \;<\; \varepsilon \qquad\qquad \text{y} \qquad\qquad f_n \to f \text{ uniformemente sobre } E \setminus F.$$

Esto último nos permite elegir un $N \in \mathbb{N}$ tal que

$$|f_n(x) - f(x)| \;<\; \varepsilon \quad \text{para todo } \; x \in E \setminus F \text{ y todo } \; n \geq N.$$

Finalmente, como $E = (E \setminus F) \cup F$, resulta que si $n \geq N$, entonces

$$\mu\big(\{x \in E \setminus F : |f_n(x) - f(x)| \geq \varepsilon\}\big) \;=\; \mu(\varnothing) \;=\; 0$$

y, en consecuencia,

$$\begin{aligned}
\mu\big(\{x \in E : |f_n(x) - f(x)| \geq \varepsilon\}\big) \;&=\; \mu\big(\{x \in E \setminus F : |f_n(x) - f(x)| \geq \varepsilon\}\big) \\
&\quad + \; \mu\big(\{x \in F : |f_n(x) - f(x)| \geq \varepsilon\}\big) \\
&\leq \; \mu(F) \;<\; \varepsilon.
\end{aligned}$$

Esto prueba que $f_n \to f$ en medida y termina la prueba. ■

7.3. Ejercicios Resueltos

El siguiente resultado establece que toda función medible y finita c.s. es casi-acotada en el siguiente sentido.

Ejercicio 7.3.1. ***Definición**. Una función medible $\varphi : E \to \mathbb{R}$ se dice que es $\aleph_0$-**simple** si ella asume sólo una cantidad infinita numerable de valores.*

*Pruebe que si $f : E \to \mathbb{R}$ es una función medible, entonces existe una sucesión $(\varphi_n)_{n=1}^\infty$ de funciones $\aleph_0$-simples tal que $\varphi_n \to f$ **uniformemente** sobre E.*

Prueba. Para cada $n \in \mathbb{Z}$, y cada $m \in \mathbb{N}$, considere el intervalo

$$I_{n,m} \;=\; [(n-1)/m,\, n/m).$$

Observe que, fijando $m \in \mathbb{N}$, la colección $(I_{n,m})_{n\in\mathbb{Z}}$ constituye una partición de $\mathbb{R}$ y, en consecuencia, la colección $(E_{n,m})_{n\in\mathbb{Z}}$ donde $E_{n,m} = f^{-1}(I_{n,m})$, produce una partición de E para cada $m \in \mathbb{N}$. Defina ahora la función $\aleph_0$-simple $\varphi_m : E \to \mathbb{R}$ por

$$\varphi_m(x) \;=\; \sum_{n\in\mathbb{Z}} \frac{n-1}{m}\cdot \chi_{E_{n,m}}.$$

Fijemos $x \in E = \bigcup_{n=1}^{\infty} E_{n,m}$. Entonces existe un $n \in \mathbb{Z}$ tal que $x \in E_{n,m}$ y, por consiguiente,

$$\varphi_m(x) \;=\; \frac{n-1}{m} \;\leq\; f(x) \;<\; \frac{n}{m}.$$

De aquí se sigue que $f(x) - \varphi_m(x) < 1/m$ para cada $x \in E$ y cada $m \geq 1$. Sea $\varepsilon > 0$ y seleccione un $n \in \mathbb{N}$ de modo que $n > 1/\varepsilon$. Entonces

$$f - \varphi_n \;=\; |f - \varphi_n| \;<\; 1/n \;<\; \varepsilon \quad \text{sobre } E,$$

lo cual nos revela que $\varphi_n \to f$ uniformemente sobre E. ■

Ejercicio 7.3.2. *Sea $f : \mathbb{R} \to \overline{\mathbb{R}}$ una **función medible** y **finita casi-siempre**. Entonces, dado $\varepsilon > 0$, existe un conjunto medible $F \subseteq \mathbb{R}$ tal que*

$$\mu(\mathbb{R}\setminus F) \;<\; \varepsilon \qquad\qquad y \qquad\qquad f \text{ es \textbf{acotada} sobre } F.$$

Prueba. Para cada entero $n \geq 1$, sean

$$E_n \;=\; \big\{x \in \mathbb{R} : |f(x)| > n\big\} \qquad\qquad \text{y} \qquad\qquad E_f^{\infty} = \big\{x \in \mathbb{R} : |f(x)| = \infty\big\}.$$

Por hipótesis, sabemos que $\mu(E_f^{\infty}) = 0$. Más aun, como

$$E_1 \supseteq E_2 \supseteq E_3 \supseteq \cdots \qquad\qquad \text{y} \qquad\qquad E_f^{\infty} \;=\; \bigcap_{n=1}^{\infty} E_n$$

el Teorema 6.3.45 nos asegura que $\lim_{n\to\infty} \mu(E_n) = \mu(E_f^{\infty}) = 0$. De aquí que, dado $\varepsilon > 0$, existe un $N \in \mathbb{N}$ tal que $\mu(E_N) < \varepsilon$. Defina

$$F \;=\; \mathbb{R}\setminus E_N \;=\; =\; \big\{x \in E : |f(x)| \leq N\big\}.$$

Entonces $\mu(\mathbb{R}\setminus F) = \mu\big(\{x \in E : |f(x)| > N\}\big) < \varepsilon$ y se cumple que $|f(x)| \leq N$ para todo $x \in F$, es decir, f es acotada sobre F. ■

Ya hemos visto que si $f : \mathbb{R} \to \mathbb{R}$ una función de Cauchy continua, entonces f es de la forma $f(x) = a \cdot x$ para todo $x \in \mathbb{R}$ y algún $a \in \mathbb{R}$. El siguiente resultado establece que la ausencia de continuidad de una función de Cauchy indica que dicha función no puede ser medible

Ejercicio 7.3.3 (Banach). *Sea $f : \mathbb{R} \to \mathbb{R}$ una **función de Cauchy**. Las siguientes condiciones son equivalentes:*

(1) *f es **continua**.*

(2) *f es **medible**.*

Prueba. (1) $\Rightarrow$ (2) ya fue demostrado.

(2) $\Rightarrow$ (1). Daremos tres demostraciones de esta implicación.

(**Primera Demostración**). Suponga que f es medible. Se sigue del Ejercicio 7.3.2 que existe un conjunto medible E con $\mu(E) > 0$ tal que f es acotada sobre E. Sea $M > 0$ tal que $|f(x)| \leq M$ para todo $x \in E$. Entonces, para todo $x, y \in E$ se tiene que

$$|f(x-y)| \ = \ |f(x) - f(y)| \ \leq \ |f(x)| + |f(y)| \ \leq \ 2M.$$

Por otro lado, por el Teorema de Steinhaus, Teorema 6.3.54, sabemos que $E - E$ contiene un intervalo, digamos $(-\alpha, \alpha) \subseteq E - E$ para algún $\alpha > 0$. De aquí se sigue que si $|x| < \alpha/n$, entonces se cumple que $|f(x)| \leq 2M/n$. Finalmente, sea $x \in \mathbb{R}$ y seleccione un $r \in \mathbb{Q}$ tal que $|x - r| < \alpha/n$ para algún $n \in \mathbb{N}$. De (FC_1), página 337, sabemos que $f(r) = rf(1)$ para todo $r \in \mathbb{Q}$, de donde se tiene que

$$\begin{aligned}
|f(x) - xf(1)| \ &= \ |f(x) - f(r) + f(r) - xf(1)| \\
&= \ |f(x) - f(r) + (r - x)f(1)| \\
&\leq \ |f(x) - f(r)| + |x - r||f(1)| \\
&\leq \ \frac{2M}{n} + \frac{\alpha}{n}|f(1)|,
\end{aligned}$$

de modo que si n se elige lo suficientemente grande se tendrá que $f(x) = xf(1)$ para todo $x \in \mathbb{R}$. Esto prueba que f es continua en x.

(**Segunda Demostración**). Como antes, suponga que f es medible pero que f no es continua sobre $\mathbb{R}$. En particular, f no es la función la constante y, por lo tanto, existen $a, b \in \mathbb{R} \setminus \{0\}$, $a \neq b$, tal que $f(a) \neq f(b)$. Sin perder generalidad en el razonamiento, podemos suponer, y así lo haremos, que

$$f(a) \ = \ 1 \qquad\qquad \text{y} \qquad\qquad f(b) \ = \ 0.$$

Para cada $n \in \mathbb{Z}$, seleccione un $q_n \in \mathbb{Q}$ tal que

$$|na \ - \ q_n b| \ < \ \frac{1}{2}$$

(esto es posible pues el conjunto $\mathbb{Q}_b = \{qb : q \in \mathbb{Q}\}$ es denso en $\mathbb{R}$) y sea $A_n = f^{-1}([n, n+1))$. Defina ahora

$$B_0 \ = \ A_0 \cap \left[-\frac{1}{2}, \frac{3}{2} \right]$$

y para $n \neq 0$, sea

$$B_n \ = \ B_0 + (na - q_n b) \ = \ \big\{ x + (na - q_n b) \ : \ x \in B_0 \big\}.$$

Nuestra primera tarea es demostrar que $A_n \cap [0,1] \subseteq B_n$. Para ver esto, observe en primer lugar que por ser f $\mathbb{Q}$-lineal, se cumple que

$$f(na - q_n b) \ = \ f(na) - f(q_n b) = nf(a) - q_n f(b) \ = \ n.$$

Sea $x \in A_n \cap [0,1]$. Como $x \in A_n$, se tiene que $n \leq f(x) < n+1$ y, por lo tanto,

$$0 \ \leq \ f(x) - n \ = \ f(x - (na - q_n b)) \ < \ 1.$$

Esto nos dice que $x-(na-q_nb)\in A_0$. Por otro lado, puesto que $x\in[0,1]$ y $-1/2<na-q_nb<1/2$, resulta que $-1/2 < x-(na-q_nb) < 3/2$ y, en consecuencia, $y = x-(na-q_nb)\in A_0\cap[-1/2,3/2]$, esto es, $y\in B_0$ y, por lo tanto, $x=y+(na-q_nb)\in B_n$. De aquí se concluye que

$$A_n\cap[0,1]\ \subseteq\ B_n\ \subseteq\ [-1,2]$$

y, por consiguiente,

$$[0,1]\ =\ \Big(\bigcup_{n\in\mathbb{Z}} A_n\Big)\cap[0,1]\ =\ \bigcup_{n\in\mathbb{Z}}(A_n\cap[0,1])\ \subseteq\ \bigcup_{n\in\mathbb{Z}} B_n\subseteq\ [-1,2]. \tag{$*$}$$

Puesto que f es medible, resulta que A_0 y, por lo tanto, B_n son medibles y $\mu(B_0)=\mu(B_n)$ para todo $n\in\mathbb{Z}$. De $(*)$ se sigue que

$$1\ =\ \mu([0,1])\ \leq\ \sum_{n\in\mathbb{Z}}\mu(B_n)\ =\ \sum_{n\in\mathbb{Z}}\mu(B_0)\ \leq\ \mu([-1,2])\ =\ 3.$$

Esto, por supuesto, es lo que genera la contradicción, pues al ser B_0 medible su medida, o es 0, o es positiva. En el primer caso se obtiene que $1\leq 0$, mientras que el segundo caso se consigue que $+\infty\leq 3$. Esto prueba que $B_0\notin\mathfrak{M}_\mu(\mathbb{R})$ y, por lo tanto, f no puede ser medible. Esta contradicción establece que f es continua en $\mathbb{R}$.

(**Tercera Demostración**). Suponga, de nuevo, que f es medible, pero que no es continua sobre $\mathbb{R}$. Entonces existe al menos un $x_0\in\mathbb{R}$ para el cual f es discontinua en dicho punto. Esto significa que existe un $\varepsilon>0$ y una sucesión convergente $(x_n)_{n=1}^\infty$ tal que

$$\lim_{n\to\infty} x_n\ =\ x_0 \qquad \text{pero} \qquad |f(x_n)-f(x_0)|\ >\ \varepsilon\quad \text{para todo }\ n\geq 1. \tag{A}$$

Por el Teorema de Lusin, Teorema 7.2.20, existe un conjunto cerrado $F\subseteq\mathbb{R}$ tal que

$$\mu(F)\ >\ 0 \qquad\qquad \text{y} \qquad\qquad f|_F\quad \text{es continua sobre } F.$$

La convergencia de la sucesión $(x_n)_{n=1}^\infty$ garantiza que ella es acotada y, en consecuencia, por el Teorema de Arandelović, Teorema 6.3.52, página 290, tenemos que

$$\mu\Big(\limsup_{n\to\infty}(-x_n+F)\Big)\ >\ \mu(F)\ >\ 0.$$

En particular, $\limsup_{n\to\infty}(-x_n+F)\neq\varnothing$. De esto se deduce la existencia de un $z\in\mathbb{R}$ y una subsucesión $(x_{n_k})_{k=1}^\infty$ de $(x_n)_{n=1}^\infty$ tal que $\{z+x_{n_k} : k=1,2\ldots\}\subseteq F$. Ahora bien, puesto que $x_{n_k}\to x_0$, se tiene que $z+x_{n_k}\to z+x_0\in F$ ya que F es cerrado y entonces, por la continuidad de la restricción $f|_F$ sobre F, resulta que

$$\lim_{k\to\infty}\big[f(z+x_{n_k})\ -\ f(z+x_0)\big]\ =\ 0.$$

Usando ahora el hecho de que f es aditiva, se concluye que

$$\lim_{k\to\infty}\big[f(z+x_{n_k})\ -\ f(z+x_0)\big]\ =\ \lim_{k\to\infty}\big[f(x_{n_k})\ -\ f(x_0)\big]\ =\ 0$$

lo cual está en franca contradicción con (A). Esto termina la prueba. ■

Ejercicio 7.3.4. *Sea* $f : \mathbb{R} \to \mathbb{R}$ *una función. Son equivalentes:*

(1) f *es* ***continua casi-siempre***.

(2) *Para cada* ***conjunto abierto*** *no-vacío* $V \subseteq \mathbb{R}$,

$$f^{-1}(V) \;=\; G \cap N,$$

donde G *es un abierto de* $\mathbb{R}$, *y* $N \in \mathfrak{M}_\mu(\mathbb{R})$ *con* $\mu(N) = 0$.

Prueba. (1) $\Rightarrow$ (2). Suponga que f es continua casi-siempre y sea V un subconjunto abierto no-vacío de $\mathbb{R}$. Puesto que $\mathbb{R} = \mathrm{PC}(f) \cup \mathrm{Disc}(f)$, entonces

$$f^{-1}(V) \;=\; \big(f^{-1}(V) \cap \mathrm{PC}(f)\big) \cup \big(f^{-1}(V) \cap \mathrm{Disc}(f)\big),$$

y, además, como $\mu(\mathrm{Disc}(f)) = 0$, resulta que $\mu(f^{-1}(V) \cap \mathrm{Disc}(f)) = 0$. Tome cualquier punto $x \in f^{-1}(V) \cap \mathrm{PC}(f)$ y observe que como f es continua en x, existe un entorno G_x de x tal que $f(G_x) \subseteq V$. Teniendo en cuenta que $G_x \subseteq f^{-1}(V) \cap \mathrm{PC}(f)$, si ahora definimos

$$G \;=\; \bigcup_{x \in f^{-1}(V) \cap \mathrm{PC}(f)} G_x \qquad \text{y} \qquad N \;=\; f^{-1}(V) \cap \mathrm{Disc}(f)$$

resulta que G es abierto y

$$f^{-1}(V) \;=\; \big(f^{-1}(V) \cap \mathrm{PC}(f)\big) \cup \big(f^{-1}(V) \cap \mathrm{Disc}(f)\big) \;\subseteq\; G \cup N \;\subseteq f^{-1}(V).$$

(2) $\Rightarrow$ (1). Suponga que (2) se cumple y sea $x \in \mathrm{Disc}(f)$. Esto significa que existe un conjunto abierto V con $f(x) \in V$ tal que, para cualquier entorno abierto G de x, $f(G) \nsubseteq V$. Ahora bien, como $f(x) \in V$ y V es abierto, existe un intervalo abierto $I_{s,r} = (s - r, s + r) \subseteq V$ con $s, r \in \mathbb{Q}$ tal que $f(x) \in I_{s,r}$. Sea $\mathcal{G} = \{G : G \text{ es un entorno abierto de } x \text{ con } f(G) \nsubseteq I_{s,r}\}$. Por hipótesis, para tales intervalos $I_{s,r}$ se cumple que

$$f^{-1}(I_{s,r}) \;=\; G_{s,r} \cup N_{s,r}$$

donde $G_{s,r} \in \mathcal{G}$ y $\mu(N_{s,r}) = 0$. Puesto que $x \in f^{-1}(I_{s,r})$ pero $x \notin G_{s,r}$, se sigue que la igualdad anterior que $x \in N_{s,r}$. Esto prueba que $\mathrm{Disc}(f) \subseteq \bigcup\{N_{s,r} : s, r \in \mathbb{Q}\}$ y como $\mathbb{Q}$ es numerable, se concluye que

$$\mu(\mathrm{Disc}(f)) \;\le\; \sum_{s,r \in \mathbb{Q}} \mu(N_{s,r}) \;=\; 0.$$

Esto demuestra que f es continua casi-siempre y termina la prueba. ■

Ejercicio 7.3.5. *Existen funciones* $f, g : [0,1] \to \mathbb{R}$ *que son* ***continuas casi-siempre***, *pero* $g \circ f$ *no lo es.*

Prueba. Sea f la función de Thomae y considere la función $g : [0,1] \to \mathbb{R}$ definida por

$$g(x) \;=\; \begin{cases} 1 & \text{si} \quad x = 1/n, \quad n = 1, 2, \ldots \\ 0 & \text{en otro caso.} \end{cases}$$

Entonces $g \circ f = \chi_{\mathbb{Q}}$ y esta función no es continua casi-siempre. ■

El siguiente resultado establece que la relación de equivalencia $\sim$ introducida en $\mathcal{F}_\mu(E)$ no es de importancia en el espacio de las funciones continuas, es decir, *ningún par de funciones continuas distintas pueden ser iguales casi-siempre* ya que la medida del conjunto de los puntos donde ellas no coinciden es estrictamente positivo.

Ejercicio 7.3.6. *Sean* $f, g \in C([a,b])$ *y suponga que* $f = g$ ***casi-siempre***. *Demuestre que* $f = g$ *sobre* $[a,b]$.

Prueba. Sea $N = \{x \in [a,b] : f(x) \neq g(x)\}$. Entonces $\mu(N) = 0$. Queremos demostrar que $N = \varnothing$. En efecto, suponga que $N \neq \varnothing$. Puesto que $f - g$ es continua, se tiene que

$$[a,b] \setminus N \;=\; \{x \in [a,b] : f(x) = g(x)\} \;=\; (f-g)^{-1}(\{0\})$$

es un conjunto cerrado y, por consiguiente, N es un conjunto abierto no vacío, por lo que su medida $\mu(N) > 0$. Esta contradicción establece que $N = \varnothing$, y así, $f = g$ sobre $[a,b]$. ■

Observe que la elección del intervalo $[a,b]$ no intervino para nada en la conclusión del ejemplo anterior, por lo que dicho intervalo puede ser sustituido por cualquier conjunto medible $\Omega \subseteq \mathbb{R}$.

7.4. Problemas

(1) Sea X un subconjunto localmente compacto de $\mathbb{R}$. Demuestre que $\mathfrak{Ba}(X) = \mathfrak{Bo}(X)$.

(2) Sea $E \in \mathfrak{M}_\mu(\mathbb{R})$. Demuestre que

$$\mu^*(E \cup A) + \mu^*(E \cap A) \;=\; \mu^*(E) + \mu^*(A)$$

para cualquier conjunto $A \subseteq \mathbb{R}$. (Use el Criterio de Carathéodory dos veces.)

(3) Sean $A, B \in \mathscr{P}(\mathbb{R})$. Demuestre que

(*a*) $\chi_{A\cap B} = \chi_A \cdot \chi_B$.
(*b*) $\chi_{A\cup B} = \chi_A + \chi_B - \chi_{A\cap B}$.
(*c*) $\chi_{A^c} = 1 - \chi_A$.
(*d*) $\chi_{A\triangle B} = |\chi_A - \chi_B|$.
(*e*) $\chi_{A\setminus B} = \chi_A(1 - \chi_B)$.

(4) Sea $(A_n)_{n=1}^{\infty}$ una sucesión de conjuntos. Demuestre que

$$\chi_{\underline{\lim} A_n} \;=\; \underline{\lim}\, \chi_{A_n}.$$

(5) Dada una función $f : E \to \mathbb{R}$, de ejemplos que demuestren que la condición *"f es continua casi-siempre en E"* no implica ni es implicada por la condición *"existe una función continua $g : E \to \mathbb{R}$ tal que $f = g$ casi-siempre"*.

(6) Suponga que $f : \mathbb{R} \to \mathbb{R}$ es medible. Pruebe que:

(*a*) Si $g : \mathbb{R} \to \mathbb{R}$ es monótona, entonces $g \circ f$ es medible.

(*b*) Si $g : \mathbb{R} \to \mathbb{R}$ es sobreyectiva y existe alguna constante $c > 0$ tal que

$$|g(x) - g(y)| \;\geq\; c|x - y| \quad \text{para todo} \quad x, y \in \mathbb{R},$$

entonces $f \circ g$ es medible.

(7) Sea $E \in \mathfrak{M}_\mu(\mathbb{R})$ con $\mu(E) < +\infty$ y suponga que $(f_n)_{n=1}^{\infty} \subseteq \mathcal{F}_\mu(E)$ es una sucesión **puntualmente acotada**, es decir, para cada $x \in E$, existe una constante $M_x > 0$ tal que $|f_n(x)| \leq M_x$ para todo $n \geq 1$. Demuestre que, para cada $\varepsilon > 0$, existe un conjunto cerrado $F_\varepsilon \subseteq E$ y un número $M > 0$ tal que

$$\mu(E \setminus F_\varepsilon) < \varepsilon \qquad \text{y} \qquad |f_n(x)| \leq M \quad \text{para todo } n \geq 1 \text{ y todo } x \in F_\varepsilon.$$

(8) Sea $(f_n)_{n=1}^{\infty}$ una sucesión de funciones medibles definidas sobre un conjunto medible y suponga que ella converge a una función f casi-siempre. Pruebe que existe una sucesión de subconjuntos medibles $\big(E_n\big)_{n=1}^{\infty}$ incluidos en E tal que $\mu(E \setminus \bigcup_{j=1}^{\infty} E_j) = 0$ y la sucesión $(f_n)_{n=1}^{\infty}$ converge uniformemente a f sobre cualquier E_j.

(9) Sea $(f_n)_{n=1}^{\infty}$ una sucesión de funciones medibles. Demuestre que si la sucesión $(f_n)_{n=1}^{\infty}$ converge en medida a f y también a g, entonces $f = g$ casi-siempre.

(10) Sea $(f_n)_{n=1}^{\infty}$ una sucesión en $\mathcal{F}_\mu(\mathbb{R})$ y sea $(\varepsilon_n)_{n=1}^{\infty}$ una sucesión de números positivos que converge a 0. Demuestre que si

$$\sum_{n=1}^{\infty} \mu\Big(\big\{x \in \mathbb{R} : |f_n(x)| \geq \varepsilon_n\big\}\Big) < +\infty,$$

entonces $f_n \to 0$ casi-siempre.

(11) Sea $(f_n)_{n=1}^{\infty}$ una sucesión en $\mathcal{F}_\mu(\mathbb{R})$ y sea $a > 0$. Demuestre que si

$$\sum_{n=1}^{\infty} \mu\Big(\big\{x \in \mathbb{R} : |f_n(x)| > a\big\}\Big) < +\infty,$$

entonces $\limsup\limits_{n\to\infty} f_n \leq a$ casi-siempre.

(12) Sea $E \in \mathfrak{M}_\mu(\mathbb{R})$ con $\mu(E) < +\infty$, y sea $(f_n)_{n=1}^{\infty}$ una sucesión en $\mathcal{F}_\mu(E)$. Muestre que $f_n \to f$ en medida si, sólo si, cada **subsucesión** $(f_{n_k})_{k=1}^{\infty}$ de $(f_n)_{n=1}^{\infty}$ posee una subsucesión que converge puntualmente $\mu - c.s.$ sobre E.

(13) Sean $f, g, h \in \mathcal{F}_\mu(E)$. Demuestre que, para cada $\varepsilon > 0$,

$$\mu\big([\,|f-g| > \varepsilon]\big) \leq \mu\big([\,|f-h| > \varepsilon/2\,]\big) + \mu\big([\,|h-g| > \varepsilon/2\,]\big)$$

donde $[\,|\varphi| > c\,] = \{x \in E : |\varphi(x)| > c\}$.

CAPÍTULO 8

La Integral de Riemann

8.1. Introducción

Tanto Isaac Newton (1642-1727) así como Gottfried Wilhelm Leibniz (1646-1716) comparten, juntos, el mérito de ser los creadores del Cálculo Diferencial e Integral. Dicho cálculo es, por supuesto, la puerta de entrada a la matemática avanzada. La integral, tal como fue ideada por Newton, se define como una anti-derivada; en otras palabras, integrar una función f al estilo Newton, consistía en encontrar otra función, digamos F, cuya derivada sea precisamente f. Esto es lo que se conoce como el Problema de las Primitivas. En lenguaje moderno, esto se expresa como

$$\int_a^b f\,dx \;=\; F(b) - F(a), \tag{1}$$

donde $F' = f$. Dicha expresión se conoce generalmente como la fórmula de Newton-Leibniz. En este sentido, una función $f : [a,b] \to \mathbb{R}$ se dice que es Newton integrable si existe una función diferenciable F tal que $F' = f$ y la integral de Newton de f se define como en (1). Es muy probable que el éxito de la fórmula de Newton-Leibniz propiciara el criterio de "la integral" como una primitiva o antiderivada. Aunque la fórmula se aplica al problema general de encontrar áreas debajo de una curva, los matemáticos de la época, al parecer, no sintieron la necesidad de definir con mayor rigor "la integral" en términos de límites de sumas asociadas a particiones. Tal vez el gran éxito del método de Newton-Leibniz, los distrajo. Augustin-Louis Cauchy (1789-1857) fue el primero en interpretar, en la década de 1820, *la integral como límite de sumas de áreas*. Más tarde, el alemán Georg Friedrich Bernhard Riemann (1826-1866), los franceses Jean Gaston Darboux (1842-1917), Henri Leon Lebesgue (1875-1941) y muchos otros brillantes matemáticos se hicieron cargo del asunto al liderar el camino en la construcción de un cálculo más profundo y riguroso. Ciertamente Riemann y, posteriormente, Darboux y Lebesgue jugaron un papel de primer orden en dicha construcción, éste último no sólo por su teorema de 1902 donde establece las condiciones necesarias y suficientes para la integrabilidad, en el sentido de Riemann, de las funciones acotadas, sino también por inventar su propia integral que permitió una generalización significativa de la integral de Riemann.

Como fue expresado anteriormente, Cauchy fue la primera persona en desarrollar de manera rigurosa la definición de integral de una función acotada $f : [a, b] \to \mathbb{R}$ como el límite de las sumas de las áreas determinadas por particiones finitas del intervalo $[a, b]$, demostrando que si f es continua, entonces ella se podía integrar de acuerdo con esa definición. Después que Cauchy presentó sus ideas a principios de la década de 1820, hubo alguna discusión sobre lo mal que una función puede comportarse y seguir siendo integrable, pero, ¿qué "tan mal" debía comportarse una función acotada para dejar de ser integrable? La respuesta la tenía Peter Gustav Lejeune-Dirichlet (1805-1859) quien propuso una función acotada que no podía manejar la integral de Cauchy. La función en cuestión era, en el lenguaje moderno, la archiconocida función característica de los racionales en $[0, 1]$ que, como sabemos, se define igual cero en los irracionales y uno en los números racionales. Por supuesto, la integral de Riemann tampoco fue capaz de integrar tal función. Georg Riemann dio la definición de su integral en 1854 como parte de su Habilitationschrift, una especie de segunda disertación que debían escribir, con carácter obligatorio, todos los aspirantes que deseaban enseñar en las universidades alemanas. Curiosamente, la integral que propuso Riemann no era el punto principal de su disertación, sino que desempeñaba un papel de apoyo de su estudio sobre "*La representación de cualquier función por medio de funciones trigonométricas*". El matemático francés Gaston Darboux elaboró una definición de la integral que era intuitivamente más practica y amigable que la de Riemann pues no tenía que lidiar con los puntos elegidos en los intervalos de cada partición. Más tarde se descubriría que la integral de Darboux era, en realidad, equivalente a la formulación original de la integral de Riemann, por lo que él nunca tuvo "su propia integral". Con la creación de la integral de Riemann comienzan, al poco tiempo, a aparecer algunos problemas. Uno de ellos consistía en determinar si existían, a propósito de la fórmula de Newton-Leibniz, derivadas acotadas que no eran integrables en el sentido de Riemann. Es aquí cuando entra en escena Vito Volterra (1860-1940) creando una admirable, pero extraña, función que tenía una derivada acotada en todos los puntos del intervalo $[0, 1]$ y, sin embargo, no podía ser integrada en el sentido de Riemann; es decir, la fórmula de Newton-Leibniz no se cumple para dicha función. Otro problema consistió en lo poco adecuado que era su proceso de convergencia: el límite puntual de una sucesión de funciones Riemann integrables no siempre es una función Riemann integrable. Estos y algunos otros hechos estimularon la búsqueda de un concepto más poderoso de integral que comienza con los esfuerzos de Camille Jordan, Emile Borel, René Baire y muchos otros, culminando con la definición de Henri Lebesgue basada en la idea ingeniosa de cambiar las particiones del dominio de una función por particiones en el rango. Además de desarrollar su propia integral, Lebesgue resolvió con ella casi todas las limitaciones que presentaba la integral de Riemann, demostrando, en 1902, un fantástico y formidable resultado sobre cómo se caracterizan las funciones que son Riemann integrales. Sin embrago, dicha integral no resolvía satisfactoriamente el problema de las primitivas por lo que, en un nuevo intento por restablecer la fórmula de Newton-Leibniz, se sembraron varias nuevas semillas en el jardín de las integrales que resolvían satisfactoriamente el problema de las primitivas. Tales integrables serán conocidas bajo los nombres de: integral de Denjoy, Perron y de Henstock-Kurzweil. Esta última integral florece con mucho más brillo ya que resuelve satisfactoriamente muchos de problemas que las integrales de Riemann y, aun, la integral Lebesgue no podían resolver. En particular, revela el camino que conduce al paraíso de la fórmula de Newton-Leibniz. Esto, por supuesto, es una pequeñísima parte de la historia que contaremos en este y subsiguientes capítulos. Más aun, como la integral Henstock-Kurzweil también posee sus propias limitaciones, la búsqueda de una integral que se aproxime cada vez a la "perfección", aun continúa.

8.2. La Integral de Newton

En el campo matemático, como ya hemos mencionado, se ha ido sembrando poco a poco un amplio jardín de integrales. Cada una de ellas tiene, por supuesto, una motivación o razón de ser para su existencia. De hecho, muchas de ellas admiten varios puntos de vistas distintos para su presentación o construcción. En el caso que nos atañe en esta sección, la integral de Riemann, haremos una presentación un poco informal e incompleta. Sólo nos interesa resaltar algunas de las notables deficiencias que habitan en su interior con el único propósito de resaltar las bondades de las otras dos integrales que desarrollaremos en los siguientes capítulos y que corrigen de modo significativo, y casi por completo, las imperfecciones de la integral de Riemann. Sin embargo, a pesar de sus pocos defectos, es a la integral de Riemann a la que aun se le sigue rindiendo honores en todos los cursos de Cálculo, aunque, por supuesto, nunca ha dejado de tener acérrimos detractores de eminentes matemáticos. Por ejemplo, Jean Diedudonné en [49], p. 146, dice de ella:

> *"... Se puede sospechar que, de no haber sido por su prestigioso nombre, hace ya tiempo que se habría dejado de considerar, pues (con toda la reverencia que se debe al genio de Riemann) es muy claro para todo matemático que esta teoría tiene actualmente la importancia de un ejercicio medianamente interesante en la teoría general de la medida y de la integración. Sólo un sentido conservador a ultranza de la tradición académica la ha dejado congelada como capítulo notable de los programas, mucho después de transcurrido el momento histórico en que tuvo verdadero significado"*

Un poco más conciliador son los matemáticos R. Bartle, R. Henstock, J. Kurzweil, E. Schechter, S. Scwabik y R. Výborný quienes lanzan a la luz pública, vía Internet, una carta firmada por ellos donde sugieren, a los que ya han escrito o están por escribir libros de Cálculo, lo siguiente:

> *"Ha sido sólo un accidente de la historia que sea la integral de Riemann la que se usa en todos los libros de Cálculo. La integral de Henstock-Kurzweil fue descubierta un poco más tarde pero es en todo, o casi todos los aspectos, una mejor integral. Así que permítanos sugerirles que en las siguientes ediciones de sus libros de Cálculo, presenten ustedes ambas, la integral de Riemann y la de Henstock-Kurzweil..."*

Esta otra opinión la manifiesta William Dunham en su libro [54], p. 208, al afirmar:

> *"... existe una cierta ironía en el hecho de que la persona que finalmente entendió cabalmente la integral de Riemann fue la misma que muy pronto la haría obsoleta: Henri Lebesgue".*

Igor Kluvánek en su artículo *What is wrong with the teaching of calculus*? [78] afirma lo siguiente:

> *"Es bastante obvio que si alguien ha leído y entendido el artículo de Riemann en el que presentó la integral que hoy lleva su nombre, o por lo menos notó el objetivo para el cual Riemann utilizó esta definición, entonces buscaría, sin duda alguna, incluir la integral de Lebesgue en la enseñanza elemental del cálculo diferencial e integral. Sólo entonces el respeto al legado de B. Riemann puede ser realmente pagado".*

La integral de Riemann es una de las herederas del concepto original de integral desarrollado independientemente por Isaac Newton y Gottfried Leibniz. Ellos descubrieron que el problema de determinar áreas (y otros problemas físicos) puede ser expresado como un problema de primitivas. De modo específico, y usando el lenguaje moderno, se requería resolver el problema con valores iniciales

$$\frac{dy}{dx} = f(x), \qquad y(x_0) = x_0 \tag{2}$$

donde f es una función arbitraria y (x_0, y_0) es un punto dado. Encontrar cualquier primitiva o antiderivada de f resolvía satisfactoriamente ese problema. La integral de Newton, dada en la siguiente definición, es una solución formal de ese problema

Recordemos que una función $F : [a,b] \to \mathbb{R}$ es **diferenciable en** $[a,b]$ significa que el límite

$$F'(x_0) \;=\; \lim_{x \to x_0} \frac{F(x) - F(x_0)}{(x - x_0)}$$

existe para cada $x_0 \in (a,b)$ y, en los puntos extremos a y b, los límites

$$F'(a) \;=\; \lim_{x \to a^+} \frac{F(x) - F(a)}{(x - a)} \qquad \text{y} \qquad F'(b) \;=\; \lim_{x \to b^-} \frac{F(x) - F(b)}{(x - b)}$$

existen. Una función $f : [a,b] \to \mathbb{R}$ se dice que es una **derivada** si existe una función diferenciable en $[a,b]$, digamos $F : [a,b] \to \mathbb{R}$, tal que $F'(x) = f(x)$ para todo $x \in [a,b]$.

Definición 8.2.1 (Integral de Newton). *Una función $f : [a,b] \to \mathbb{R}$ se dice que es **Newton integrable** sobre $[a,b]$ si f es una derivada, en otras palabras, si existe una función diferenciable $F : [a,b] \to \mathbb{R}$ tal que $F'(x) = f(x)$ para todo $x \in [a,b]$.*

Denote por $\mathcal{N}([a,b])$ *el conjunto de todas las funciones $f : [a,b] \to \mathbb{R}$ que son **Newton integrables** sobre* $[a,b]$. Si f es Newton integrable sobre $[a,b]$ con $F' = f$, entonces la **integral de Newton** de f se define como:

$$(\mathcal{N})\int_a^b f\,dx \;=\; F(b) - F(a). \tag{3}$$

La expresión dada en (3) se llama la **fórmula de Newton-Leibniz** y el problema de recobrar una función de su derivada se le conoce con el nombre de: **el Problema de las Primitivas**. La función derivable F satisfaciendo $F' = f$ se llama una **primitiva** o **antiderivada** de f. Este es el concepto original de integración de Newton y Leibniz concebido como la operación inversa de la diferenciación, es decir, la anti-diferenciación. Es importante destacar que una vez obtenida una primitiva F de f, toda función de la forma $F + c$, donde c es cualquier constante, también es una primitiva de f. Por lo tanto, la definición anterior requiere chequear que $F(b) - F(a)$ no depende sobre cuál primitiva se elige. En efecto, si G es otra primitiva de f, entonces el Teorema del Valor Medio nos garantiza la existencia de una constante c tal que $G(x) = F(x) + c$ para todo $x \in [a,b]$, de donde se concluye que $G(b) - G(a) = F(b) - F(a)$.

Es un ejercicio sencillo verificar que $\mathcal{N}([a,b])$ es un espacio vectorial sobre $\mathbb{R}$. Además, si $f, g \in \mathcal{N}([a,b])$, entonces se cumple que:

(a) $(\mathcal{N})\displaystyle\int_a^b f\,dx \le (\mathcal{N})\int_a^b g\,dx$ si $f \le g$,

(b) $(\mathcal{N})\displaystyle\int_a^b f\,dx = (\mathcal{N})\int_a^c f\,dx + (\mathcal{N})\int_c^b f\,dx$ para cualquier $c \in [a,b]$,

(c) $\left|(\mathcal{N})\displaystyle\int_a^b f\,dx\right| \le (\mathcal{N})\int_a^b |f|\,dx$ siempre que $|f|$ sea Newton integrable sobre $[a,b]$.

Sin embargo, si $f \in \mathcal{N}([a,b])$, entonces no siempre es verdad que $|f| \in \mathcal{N}([a,b])$ como se comprueba fácilmente tomando la función derivada:

$$f(x) \;=\; [x\,\cos(\pi/x)]' \quad \text{para todo } x \in [0,1].$$

En efecto, es claro que $F(x) = x\cos(\pi/x)$ es una primitiva para f tal que $F(0) = 0$ y $F(1) = -1$. Por esto,

$$(\mathcal{N})\int_0^1 f\,dx \;=\; F(1) - F(0) \;=\; -1.$$

Por otro lado, si suponemos que $|f| \in \mathcal{N}([a,b])$, entonces, tomando $a_k = 2/(2k+1)$ y $b_k = 1/k$ para cada $k \in \mathbb{N}$, resulta que $F(a_k) = 0$, $F(b_k) = (-1)^k/k$ y como

$$0 \;<\; a_k \;<\; b_k \;<\; a_{k-1} \;<\; b_{k-1} \;<\; \cdots \;<\; a_1 \;<\; b_1 \;<\; 1,$$

entonces

$$(\mathcal{N})\int_{a_k}^{b_k} |f|\,dx \;\geq\; \left|(\mathcal{N})\int_{a_k}^{b_k} f\,dx\right| \;=\; \left|F(b_k) \;-\; F(a_k)\right| \;=\; \frac{1}{k}.$$

Por lo tanto, para cualquier $n \in \mathbb{N}$ se tiene que

$$\sum_{k=1}^{n} \;\leq\; \sum_{k=1}^{n}(\mathcal{N})\int_{a_k}^{b_k} |f|\,dx \;\leq\; (\mathcal{N})\int_0^1 |f|\,dx.$$

Puesto que la serie $\sum_{k=1}^{\infty}$ diverge, se concluye que $|f| \notin \mathcal{N}([a,b])$.

Por supuesto, la integral de Newton es una integral puramente *descriptiva*: ella no ofrece ningún método para construir una integral. Por consiguiente, la fórmula de Newton-Leibniz es una *solución formal* del problema formulado en (2). Observe que, desde el punto de vista de Newton, la integral se concibe como el proceso inverso de la diferenciación. Sin embargo, éste punto de vista va a cambiar con el tiempo.

8.3. Construcción de la Integral de Riemann

Cauchy fue el primero en definir, para una **función continua** $f : [a,b] \to \mathbb{R}$, **la integral de** f como

$$(\mathcal{C})\int_a^b f\,dx \;=\; \lim_{\|\mathcal{P}\|\to 0} \sum_{k=1}^{n} f(x_k)\cdot(x_k - x_{k-1}),$$

donde $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ es cualquier subconjunto finito de $[a,b]$ satisfaciendo $a = x_0 < x_1 < \cdots < x_n = b$ y $\|\mathcal{P}\| = \max\{x_k - x_{k-1} : k = 1, 2, \ldots, n\}$. Con esa definición Cauchy logra demostrar que si $F(x) = (\mathcal{C})\int_a^x f\,dt$ para cada $x \in [a,b]$, entonces F es diferenciable sobre $[a,b]$ y $F' = f$: *este es el primer enunciado riguroso del Teorema Fundamental del Cálculo.*

La integral de Riemann, formulada por Bernhard Riemann en la década de 1850, permite integrar una clase de funciones mucho más amplia que la de las funciones continuas. Su definición, aunque es muy similar a la de Cauchy posee, sin embargo, una *sutil diferencia*: en lugar de elegir el extremo derecho x_k de cada intervalo $[x_{k-1}, x_k]$ y formar la suma $\sum_{k=1}^{n} f(x_k)\cdot(x_k - x_{k-1})$, Riemann *escoge un punto arbitrario* t_k en cada intervalo $[x_{k-1}, x_k]$ y considera, como Cauchy, la suma $\sum_{k=1}^{n} f(t_k)\cdot(x_k - x_{k-1})$. Por supuesto, esa pequeña diferencia, que en apariencia parece insignificante, da origen a una integral que es mucho más amplia y poderosa que la de Cauchy permitiéndole, desde su creación, ser la preferida de todos. Además de sencilla y elegante, la integral de Riemann posee, como es bien conocido, un amplio abanico de aplicaciones. Todos los libros de Cálculo Diferencial e Integral moderno la utilizan como su herramienta fundamental para el cálculo de longitudes, áreas, volúmenes, centros de masas, etc. Sin embargo, con el correr de los tiempos, del edificio construido por Riemann comienzan a brotar ciertas grietas que la debilitan y limitan enormemente. Una de sus primeras limitaciones consiste en que sólo las **funciones acotadas**, definidas sobre un intervalo **cerrado** y **acotado**, son susceptibles de poder ser integradas en el sentido de Riemann quedando por fuera, por supuesto, todas las funciones que no son acotadas y también las acotadas o no-acotadas definidas, por ejemplo, en cualquier intervalo de longitud infinita. Otra limitante, tal vez la más importante, es su mal comportamiento con respecto a la convergencia puntual de sucesiones: esto significa que el límite puntual de una sucesión de funciones integrables en el sentido de Riemann no es necesariamente integrable según Riemann. Finalmente, el Teorema Fundamental de Cálculo sólo es aplicable a funciones primitivas cuya derivada sea integrable según Riemann. Todas esas limitaciones conducen a una conclusión incuestionable: dicha integral es inapropiada, en comparación con otras integrales que surgirán a partir de ella, para ser considerada como la integral preferida por los matemáticos. A pesar de estos sinsabores, compartimos la apreciación de Florencio del Castillo cuando afirma en [47], p. 152, lo siguiente:

> *"La integral de Riemann es de alcance limitado y es insuficiente para una gran parte de los problemas de la Matemática y de la técnica, pero las diversas formulaciones satisfactorias de otras integrales son generalizaciones de dicho concepto. Creo que esta razón sería suficiente para motivar el estudio de la mencionada integral de Riemann como iniciación a la teoría de otras integrales. Sólo así se apreciaría el valor y la razón de posteriores generalizaciones".*

Sea $[a, b]$ un intervalo cerrado y acotado de $\mathbb{R}$. Una **partición** de $[a, b]$ es una colección finita de puntos $\{x_0, x_1, \ldots, x_n\}$ de $[a, b]$ tal que

$$a = x_0 < x_1 < x_2 < \cdots < x_n = b.$$

Denotaremos a una tal partición por $\mathcal{P}$. Los puntos de $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ dividen al intervalo $[a, b]$ en n subintervalos cerrados no-superpuestos

$$I_1 = [x_0, x_1], \qquad I_2 = [x_1, x_2], \qquad \cdots, \qquad I_n = [x_{n-1}, x_n]$$

a los que llamaremos **intervalos asociados a** $\mathcal{P}$. En lo que sigue, el término partición será usada indistintamente para referirnos a los puntos de $\mathcal{P}$, o a los intervalos asociados a $\mathcal{P}$. A cada partición $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ de $[a, b]$, le asignaremos un número, al que llamaremos la **norma** de $\mathcal{P}$, y que se define como

$$\|\mathcal{P}\| = \text{máx}\big\{(x_1 - x_0), (x_2 - x_1), \ldots, (x_n - x_{n-1})\big\}.$$

Un **conjunto de etiquetas** para los intervalos asociados a la partición $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ es cualquier colección finita de puntos de $[a, b]$, digamos $e = \{t_1, t_2, \ldots, t_n\}$, tal que $t_i \in [x_{i-1}, x_i]$ para $i = 1, \ldots, n$. El conjunto de pares ordenados

$$\mathcal{P}_e = \big\{\big(t_i, [x_{i-1}, x_i]\big) : i = 1, \ldots, n\big\},$$

donde $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ es una partición de $[a, b]$ y $e = \{t_1, t_2, \ldots, t_n\}$ es un conjunto de etiquetas de los intervalos asociados a $\mathcal{P}$, lo llamaremos una **partición etiquetada** de $[a, b]$. Dado un **número positivo** δ, una partición etiquetada $\mathcal{P}_e = \big\{\big(t_i, [x_{i-1}, x_i]\big) : i = 1, \ldots, n\big\}$ se dice que es δ-**fina**, o **subordinada** a δ, si para cada $i \in \{1, \ldots, n\}$,

$$[x_{i-1}, x_i] \subseteq (t_i - \delta, t_i + \delta).$$

Observe que *si* $\mathcal{P}_e$ *es una* ***partición etiquetada*** *de* $[a,b]$, *entonces* $\| \mathcal{P}_e \| \leq 2\delta$ *si, y sólo si,* $\mathcal{P}_e$ *está subordinada a* δ.

Sea $\mathcal{B}_\infty([a,b])$ *el espacio vectorial de todas las funciones* $f : [a,b] \to \mathbb{R}$ *que son* ***acotadas*** *sobre* $[a,b]$. Si $f \in \mathcal{B}_\infty([a,b])$ y $\mathcal{P}_e = \big(t_i,\, [x_{i-1},\, x_i]\big)_{i=1}^n$ es una partición etiquetada de $[a,b]$, entonces al número

$$S(f, \mathcal{P}_e) \; = \; \sum_{i=1}^n f(t_i)(x_i - x_{i-1}) \tag{3}$$

se le llama una **suma de Riemann de** f **asociada a** $\mathcal{P}_e$. Observe que si $f \geq 0$, entonces las sumas de Riemann de f se pueden interpretar como áreas que, a medida que la partición se hace más fina, se aproximan al área que se encuentra debajo de la curva $S = \{(x, f(x)) : x \in [a,b]\}$ limitada por el eje de las x y las rectas $x = a$ y $x = b$. La usual definición de la integral de Riemann, formulada originalmente por B. Riemann en 1854, puede ser expresada en la forma:

Definición 8.3.1. *Sea* $f : [a,b] \to \mathbb{R}$ *una función acotada sobre* $[a,b]$. *Se dice que* f *es* ***Riemann integrable*** *sobre* $[a,b]$, *si existe un número real* A *con la siguiente propiedad: para cada* $\varepsilon > 0$, *existe una constante* $\delta_\varepsilon > 0$ *tal que*

$$\big|S(f, \mathcal{P}_e) \, - \, A\big| \; < \; \varepsilon$$

para cualquier partición etiquetada $\mathcal{P}_e = \big(t_i,\, [x_{i-1},\, x_i]\big)_{i=1}^n$ *de* $[a,b]$ *subordinada a* δ_ε.

El número A de la definición anterior, si existe, es único y se le llama la **integral de Riemann** de f y se denotará, en lo sucesivo, por

$$A \; = \; (\mathrm{R})\int_a^b f\,dx.$$

En lo que sigue, el símbolo $\mathcal{R}([a,b])$ será usado para denotar *el conjunto de todas las funciones* $f \in \mathcal{B}_\infty([a,b])$ *que son* ***Riemann integrables*** *sobre* $[a,b]$.

Es importante destacar que, en la definición de la integral de f, se exigió que ella fuese acotada; sin embargo, tal exigencia no es necesaria pues el acotamiento de f sigue si ella es Riemann integrable. Observe también que la existencia del número positivo δ_ε en la definición de la integral de Riemann es, en definitiva, quien ejerce el control de las particiones: *toda partición etiquetada* $\mathcal{P}$ *que tenga una* ***norma*** *menor que* δ *participa del banquete, es decir, permite que la diferencia* $|S(f, \mathcal{P}) - A|$ *sea siempre menor que* ε.

8.3.1. La Integral de Darboux - Su Construcción

En ocasiones resulta más práctico considerar las así llamadas *sumas de Darboux* para evaluar si una función dada es o no Riemann integrable. Tales sumas no requieren el uso explícito de etiquetas.

Como antes, fijemos una función acotada $f : [a,b] \to \mathbb{R}$ y sea $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ una partición de $[a,b]$. Si $\{I_1, I_2, \ldots, I_n\}$ son los intervalos asociados a $\mathcal{P}$, definimos los números

$$m_i \; = \; \inf_{x \in I_i} f(x) \qquad \text{y} \qquad M_i \; = \; \sup_{x \in I_i} f(x)$$

para $i = 1, 2, \ldots, n$. Darboux, en lugar de evaluar a f en un conjunto de etiquetas y luego multiplicarlo por $(x_i - x_{i-1})$, lo que él hace es multiplicar directamente M_i, así como m_i, por

$(x_i - x_{i-1})$ para obtener una "suma superior" y una "suma inferior" similar a las sumas de Riemann. La idea es que a medida que se refina la partición $\mathcal{P}$, se obtiene una familia de sumas superiores que decrecen y una familia de sumas inferiores que crecen. Si tales familias se "encuentran" en el límite, entonces se dice que f es integrable en el sentido de Darboux.

Definición 8.3.2. *Sea $f : [a, b] \to \mathbb{R}$ una función acotada. Las **sumas inferior y superior de Darboux** de f asociadas a una partición $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ de $[a, b]$ son definidas, respectivamente, por los números*

$$\mathsf{L}(f, \mathcal{P}) = \sum_{i=1}^{n} m_i \cdot (x_i - x_{i-1}) \qquad y \qquad \mathsf{U}(f, \mathcal{P}) = \sum_{i=1}^{n} M_i \cdot (x_i - x_{i-1}).$$

Teniendo en cuenta que $\sum_{i=1}^{n}(x_i - x_{i-1}) = b - a$ para cualquier partición $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ de $[a, b]$, resulta que si $f : [a, b] \to \mathbb{R}$ es una función acotada tal que $m \leq f(x) \leq M$ para todo $x \in [a, b]$, entonces

$$m \cdot (b - a) \leq \mathsf{L}(f, \mathcal{P}) \leq \mathsf{U}(f, \mathcal{P}) \leq M \cdot (b - a).$$

Denotemos por $\mathbb{P}[a, b]$ *el conjunto de todas las particiones de* $[a, b]$. Si $\mathcal{P}$ y $\mathcal{Q}$ son dos particiones de $[a, b]$, diremos que $\mathcal{Q}$ es **más fina** que $\mathcal{P}$ si $\mathcal{P} \subseteq \mathcal{Q}$. Observe que $\mathcal{Q}$ es más fina que $\mathcal{P}$ implica que cada uno de los subintervalos asociados a $\mathcal{P}$ es, o uno de los intervalos asociados de $\mathcal{Q}$, o una unión finita de ellos.

Lema 8.3.3. *Sea $f : [a, b] \to \mathbb{R}$ una **función acotada** y sean $\mathcal{P}, \mathcal{Q} \in \mathbb{P}[a, b]$. Si $\mathcal{Q}$ es **más fina** que $\mathcal{P}$, entonces*

$$\mathsf{L}(f, \mathcal{P}) \leq \mathsf{L}(f, \mathcal{Q}) \qquad y \qquad \mathsf{U}(f, \mathcal{P}) \geq \mathsf{U}(f, \mathcal{Q}).$$

Prueba. Suponga que $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ y que $\mathcal{Q}$ contiene, por el momento, sólo un punto más que $\mathcal{P}$, esto es, $\mathcal{Q} = \{x_0,\, x_1,\, \ldots,\, x_{i-1},\, x^*,\, x_i,\, \ldots,\, x_n\}$. Observe que

$$I_i = [x_{i-1}, x_i] = [x_{i-1}, x^*] \cup [x^*, x_i]$$

y los restante intervalos asociados de $\mathcal{P}$ y $\mathcal{Q}$ coinciden. Pongamos

$$M_i' = \sup\{f(x) : x \in [x_{i-1}, x^*]\} \quad \text{y} \quad M_i'' = \sup\{f(x) : x \in [x^*, x_i]\}.$$

Entonces

$$M_i' \leq M_i \qquad \text{y} \qquad M_i'' \leq M_i$$

y puesto que $x_i - x_{i-1} = (x_i - x^*) + (x^* - x_{i-1})$, se obtiene que

$$M_i(x_i - x_{i-1}) = M_i(x_i - x^*) + M_i(x^* - x_{i-1}) \geq M_i'(x_i - x^*) + M_i''(x^* - x_{i-1}).$$

Por esto,

$$\mathsf{U}(f, \mathcal{P}) - \mathsf{U}(f, \mathcal{Q}) = M_i(x_i - x_{i-1}) - \big(M_i'(x_i - x^*) + M_i''(x^* - x_{i-1})\big) \geq 0.$$

Suponga ahora que $\mathcal{Q}$ contiene k puntos más que $\mathcal{P}$. Entonces existe un conjunto finito de particiones $\mathcal{Q}_1, \ldots, \mathcal{Q}_{k-1}$ de $[a, b]$ tal que

$$\mathcal{P} \subseteq \mathcal{Q}_1 \subseteq \cdots \subseteq \mathcal{Q}_{k-1} \subseteq \mathcal{Q}$$

donde cada partición se obtiene de la anterior añadiéndole exactamente un punto. Entonces

$$\mathsf{U}(f, \mathcal{P}) \geq \mathsf{U}(f, \mathcal{Q}_1) \geq \cdots \geq \mathsf{U}(f, \mathcal{Q}_{k-1}) \geq \mathsf{U}(f, \mathcal{Q}).$$

El caso $\mathsf{L}(f, \mathcal{P}) \leq \mathsf{L}(f, \mathcal{Q})$ se prueba de modo enteramente similar. ∎

Corolario 8.3.4. *Sea $f : [a,b] \to \mathbb{R}$ una **función acotada**. Entonces se cumple que*

$$\mathsf{L}(f,\mathcal{P}) \ \leq \ \mathsf{U}(f,\mathcal{Q}),$$

cualesquiera sean $\mathcal{P},\mathcal{Q} \in \mathbb{P}[a,b]$.

Prueba. Hagamos $\mathcal{R} = \mathcal{P} \cup \mathcal{Q}$ y observe que como $\mathcal{R}$ es más fina que $\mathcal{P}$ y $\mathcal{Q}$, el Teorema 8.3.3 nos revela que

$$\mathsf{L}(f,\mathcal{P}) \ \leq \ \mathsf{L}(f,\mathcal{R}) \ \leq \ \mathsf{U}(f,\mathcal{R}) \ \leq \ \mathsf{U}(f,\mathcal{Q}).$$

Esto termina la prueba. ■

Consideremos ahora los siguientes subconjuntos de $\mathbb{R}$:

$$\mathcal{A} \ = \ \Big\{\mathsf{L}(f,\mathcal{P}) \ : \ \mathcal{P} \in \mathbb{P}[a,b]\Big\} \qquad \text{y} \qquad \mathcal{B} \ = \ \Big\{\mathsf{U}(f,\mathcal{Q}) \ : \ \mathcal{Q} \in \mathbb{P}[a,b]\Big\}$$

Puesto que $\mathcal{A}$ y $\mathcal{B}$ son conjuntos acotados se sigue del Teorema 8.3.3 que

$$a \ \leq \ b \quad \text{para todo} \quad a \in \mathcal{A} \ \text{y todo} \quad b \in \mathcal{B}$$

y, por lo tanto, gracias a la primera parte del Teorema 2.1.9, página 87, se concluye que

$$\sup\Big\{\mathsf{L}(f,\mathcal{P}) \ : \ \mathcal{P} \in \mathbb{P}[a,b]\Big\} \ \leq \ \text{ínf}\Big\{\mathsf{U}(f,\mathcal{Q}) \ : \ \mathcal{Q} \in \mathbb{P}[a,b]\Big\}. \tag{4}$$

Definición 8.3.5. *Sea $f : [a,b] \to \mathbb{R}$ una función acotada. La **integral inferior** e **integral superior** de Darboux se definen, respectivamente, como los números*

$$\underline{(\mathrm{D})}\int_a^b f(x)\,dx \ = \ \sup\Big\{\mathsf{L}(f,\mathcal{P}) \ : \ \mathcal{P} \in \mathbb{P}[a,b]\Big\}$$

y

$$\overline{(\mathrm{D})}\int_a^b f(x)\,dx \ = \ \text{ínf}\Big\{\mathsf{U}(f,\mathcal{Q}) \ : \ \mathcal{Q} \in \mathbb{P}[a,b]\Big\}.$$

Se sigue de (4) que

$$\underline{(\mathrm{D})}\int_a^b f(x)\,dx \ \leq \ \overline{(\mathrm{D})}\int_a^b f(x)\,dx.$$

Definición 8.3.6. *Una función acotada $f : [a,b] \to \mathbb{R}$ se dice que es **Darboux integrable** sobre $[a,b]$ si*

$$\underline{(\mathrm{D})}\int_a^b f(x)dx \ = \ \overline{(\mathrm{D})}\int_a^b f(x)dx$$

*Si éste es el caso, entonces la **integral de Darboux de f sobre** $[a,b]$ se define como ese valor común al que denotaremos por*

$$(\mathrm{D})\int_a^b f(x)\,dx.$$

Esta definición de integral fue establecida por Gaston Darboux en 1875. Lo que resulta interesante es que dicha integral es equivalente a la integral de Riemann. El símbolo $\mathcal{D}\mathrm{ar}([a,b])$ lo usaremos para denotar *el conjunto de las funciones en* $\mathcal{B}_\infty([a,b])$ *que son Darboux integrables sobre* $[a,b]$.

Ejemplo 8.3.1. *Si* $f:[a,b]\to\mathbb{R}$ *es una* ***función constante****, digamos* $f(x)=k$ *para todo* $x\in[a,b]$*, entonces* $f\in\mathcal{D}\mathrm{ar}([a,b])$ *y*

$$(\mathrm{D})\int_a^b f\,dx \;=\; k(b-a).$$

Prueba. Sea $\mathcal{P}=\{x_0,x_1,\ldots,x_n\}$ cualquier partición de $[a,b]$. Para cada $i=1,2,\ldots,n$ se cumple que

$$m_i \;=\; \text{ínf}\{f(x):x\in[x_{i-1},x_i]\} \;=\; \sup\{f(x):x\in[x_{i-1},x_i]\} \;=\; M_i \;=\; k.$$

Luego,

$$\mathsf{L}(f,\mathcal{P}) \;=\; \mathsf{U}(f,\mathcal{P}) \;=\; k(b-a),$$

de donde se sigue que

$$(\mathrm{D})\int_a^b f\,dx \;=\; k(b-a).$$

Es importante destacar que no cualquier función acotada es Darboux integrable. En efecto, una de las primeras funciones patológicas que eludía el criterio de integrabilidad de Cauchy y, por supuesto, también el de Riemann fue la función característica de los racionales $f=\chi_{\mathbb{Q}}:[0,1]\to\mathbb{R}$, propuesta por primera vez por G. Dirichlet.

Ejemplo 8.3.2. *La función de Dirichlet* $f=\chi_{\mathbb{Q}}:[0,1]\to\mathbb{R}$ *es* ***acotada*** *pero* ***no es Darboux integrable****.*

Prueba. Si $\mathcal{P}=\{x_0,x_1,\ldots,x_n\}$ es cualquier partición de $[0,1]$, entonces

$$m_k \;=\; \text{ínf}\,\big\{f(x):x\in[x_{k-1},x_k]\big\} \;=\; 0, \qquad\qquad M_k \;=\; \sup\,\big\{f(x):x\in[x_{k-1},x_k]\big\} \;=\; 1,$$

y así,

$$\mathsf{L}(f,\mathcal{P}) \;=\; 0 \qquad\text{y}\qquad \mathsf{U}(f,\mathcal{P}) \;=\; 1.$$

Por esto,

$$\underline{(\mathrm{D})}\int_a^b f(x)dx \;=\; 0 \qquad\text{y}\qquad \overline{(\mathrm{D})}\int_a^b f(x)dx \;=\; 1,$$

de donde se concluye que $f\in\mathcal{B}_\infty([a,b])\setminus\mathcal{D}\mathrm{ar}([a,b])$. ■

Recordemos que si $f:[a,b]\to\mathbb{R}$ es una función acotada, entonces la oscilación de f sobre cualquier conjunto $F\subseteq[a,b]$ se define como

$$\operatorname{osc}(f,F) \;=\; \sup_{x\in F} f(x) \;-\; \inf_{x\in F} f(x).$$

En particular, si $\mathcal{P}=\{x_0,x_1,\ldots,x_n\}$ es una partición de $[a,b]$ y si $\{I_1,I_2,\ldots,I_n\}$ son los intervalos asociados a $\mathcal{P}$, entonces

$$\operatorname{osc}(f,I_i) \;=\; M_i \;-\; m_i, \quad i=1,2,\ldots,n,$$

para $i = 1, 2, \ldots, n$ y, por lo tanto,

$$\begin{aligned} \mathsf{U}(f, \mathcal{P}) - \mathsf{L}(f, \mathcal{P}) &= \sum_{i=1}^{n} (M_i - m_i) \cdot (x_i - x_{i-1}) \\ &= \sum_{i=1}^{n} \operatorname{osc}(f, I_i) \cdot (x_i - x_{i-1}) \end{aligned} \tag{8.3.1}$$

El siguiente criterio de integrabilidad se debe a Riemann y vincula la noción de integral con continuidad. Constituye una forma simple de caracterizar a las funciones acotadas que son Darboux integrables.

Teorema 8.3.7 (Riemann-Darboux). *Sea $f : [a, b] \to \mathbb{R}$ una **función acotada**. Las siguientes son equivalentes:*

(1) $f \in \mathcal{D}\mathrm{ar}([a, b])$.

(2) *Para cada $\varepsilon > 0$, existen particiones $\mathcal{P}$ y $\mathcal{Q}$ de $[a, b]$ tal que $\mathsf{U}(f, \mathcal{P}) - \mathsf{L}(f, \mathcal{Q}) < \varepsilon$.*

(3) *Para cada $\varepsilon > 0$, existe una partición $\mathcal{P}$ de $[a, b]$ tal que $\mathsf{U}(f, \mathcal{P}) - \mathsf{L}(f, \mathcal{P}) < \varepsilon$.*

(4) *Para cada $\varepsilon > 0$, existe una partición $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ de $[a, b]$ tal que*

$$\sum_{i=1}^{n} \operatorname{osc}(f, I_i) \cdot (x_i - x_{i-1}) < \varepsilon.$$

Prueba. Las implicaciones $(1) \Leftrightarrow (2)$ siguen del Teorema 2.1.9, página 87. Sea $\varepsilon > 0$ y suponga que (2) se cumple. Definamos $\mathcal{R} = \mathcal{P} \cup \mathcal{Q}$. Entonces, por el Lema 8.3.3, se tiene que

$$\mathsf{L}(f, \mathcal{P}) \leq \mathsf{L}(f, \mathcal{R}) \leq \mathsf{U}(f, \mathcal{R}) \leq \mathsf{U}(f, \mathcal{Q}).$$

y, por consiguiente, $\mathsf{U}(f, \mathcal{R}) - \mathsf{L}(f, \mathcal{R}) < \varepsilon$. Esto prueba (3). Claramente (3) implica (2). Finalmente $(3) \Leftrightarrow (4)$ sigue de (8.3.1). ■

Observe que si $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ es cualquier partición de $[a, b]$ y f es Darboux integrable, entonces

$$\mathsf{L}(f, \mathcal{R}) \leq (\mathrm{D})\int_a^b f\, dx \leq \mathsf{U}(f, \mathcal{R}). \tag{8.3.2}$$

Más aun, si $\{t_1, t_2, \ldots, t_n\}$ es un conjunto arbitrario de etiquetas de $\mathcal{P}$, entonces

$$m_i \leq f(t_i) \leq M_i$$

para cada $i = 1, 2, \ldots, n$, por lo que

$$\mathsf{L}(f, \mathcal{P}) \leq \sum_{i=1}^{n} f(t_i) \cdot (x_i - x_{i-1}) \leq \mathsf{U}(f, \mathcal{P}). \tag{8.3.3}$$

8.3.2. Equivalencia de las Integrales de Riemann y Darboux

El siguiente resultado establece que la integral de Darboux y la integral de Riemann son exactamente la misma cosa.

Teorema 8.3.8 (Igualdad Riemann-Darboux). *Sea $f : [a,b] \to \mathbb{R}$ una **función acotada**. Las siguientes son equivalentes:*

(1) $f \in \mathcal{D}\mathrm{ar}([a,b])$.

(2) $f \in \mathcal{R}([a,b])$.

En este caso,

$$\text{(R)}\int_a^b f\,dx \;=\; \text{(D)}\int_a^b f\,dx.$$

Prueba. Suponga que $f \in \mathcal{D}\mathrm{ar}([a,b])$ y sea $A = \text{(D)}\int_a^b f\,dx$. Dado $\varepsilon > 0$ existe, gracias al Teorema 8.3.7, una partición $\mathcal{P}_\varepsilon$ de $[a,b]$ tal que

$$\mathsf{U}(f,\mathcal{P}_\varepsilon) - \mathsf{L}(f,\mathcal{P}_\varepsilon) \;<\; \varepsilon.$$

Defina $\delta = \|\mathcal{P}_\varepsilon\|$ y sea $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ cualquier partición más fina que $\mathcal{P}_\varepsilon$. Resulta que $\|\mathcal{P}\| \leq \delta$ y si $\{t_1, t_2, \ldots, t_n\}$ es un conjunto arbitrario de etiquetas de $\mathcal{P}$, tendremos que

$$\mathsf{L}(f,\mathcal{P}) \;\leq\; \sum_{i=1}^{n} f(t_i)\cdot(x_i - x_{i-1}) \;\leq\; \mathsf{U}(f,\mathcal{P})$$

y, por supuesto, también se cumple que

$$\mathsf{L}(f,\mathcal{P}) \;\leq\; \text{(D)}\int_a^b f\,dx \;\leq\; \mathsf{U}(f,\mathcal{P}).$$

De estas dos desigualdades se sigue que

$$\begin{aligned}\left|\sum_{i=1}^{n} f(t_i)(x_i - x_{i-1}) \;-\; \text{(D)}\int_a^b f\,dx\right| \;&\leq\; \mathsf{U}(f,\mathcal{P}) - \mathsf{L}(f,\mathcal{P})\\ &<\; \mathsf{U}(f,\mathcal{P}_\varepsilon) - \mathsf{L}(f,\mathcal{P}_\varepsilon)\\ &<\; \varepsilon.\end{aligned}$$

Esto prueba que f es Riemann integrable y así, por la unicidad de la integral de Riemann,

$$\text{(R)}\int_a^b f\,dx \;=\; \text{(D)}\int_a^b f\,dx.$$

Para demostrar la otra implicación, suponga que (2) se cumple y sea $\varepsilon > 0$. Fijemos una partición $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ de $[a,b]$ con $\|\mathcal{P}\| \leq \delta$ tal que la desigualdad

$$\left|\sum_{i=1}^{n} f(s_i)\cdot(x_i - x_{i-1}) \;-\; \text{(R)}\int_a^b f\,dx\right| \;<\; \varepsilon/2 \tag{$*$}$$

es válida para cualquier conjunto arbitrario de etiquetas $\{s_1, s_2, \ldots, s_n\}$ de $\mathcal{P}$. Como $m_i = \inf\{f(x) : x \in [x_{i-1}, x_i]\}$, resulta que para cada $i \in \{1, 2, \ldots, n\}$, existe, por las propiedades del ínfimo, un $t_i \in [x_{i-1}, x_i]$ tal que

$$f(t_i) \; < \; m_i \; + \; \frac{\varepsilon}{2(b-a)}.$$

Por esto, para el conjunto de etiquetas $\{t_1, t_2, \ldots, t_n\}$, se cumple que

$$\begin{aligned}\sum_{i=1}^{n} f(t_i)\cdot(x_i - x_{i-1}) \; &< \; \sum_{i=1}^{n}\Big(m_i\cdot(x_i - x_{i-1}) \; + \; \frac{\varepsilon}{2(b-a)}(x_i - x_{i-1})\Big) \\ &= \; \mathsf{L}(f, \mathcal{P}) \; + \; \frac{\varepsilon}{2}\end{aligned}$$

y así, usando $(*)$, vemos que

$$\underline{\text{(D)}}\int_a^b f\,dx \; \geq \; \mathsf{L}(f, \mathcal{P}) \; > \; \sum_{i=1}^{n} f(t_i)\cdot(x_i - x_{i-1}) \; - \; \frac{\varepsilon}{2} \; > \; \text{(R)}\int_a^b f\,dx - \varepsilon.$$

Como ε es arbitrario, concluimos que

$$\underline{\text{(D)}}\int_a^b f\,dx \; \geq \; \text{(R)}\int_a^b f\,dx.$$

De modo similar, pero ahora trabajando con $M_i = \sup\{f(x) : x \in [x_{i-1}, x_i]\}$, se prueba que

$$\underline{\text{(D)}}\int_a^b f\,dx \; \leq \; \text{(R)}\int_a^b f\,dx.$$

Finalmente, teniendo en cuenta que $\underline{\text{(D)}}\int_a^b f\,dx \leq \overline{\text{(D)}}\int_a^b f\,dx$, resulta que

$$\text{(R)}\int_a^b f\,dx \; \leq \; \underline{\text{(D)}}\int_a^b f\,dx \; \leq \; \overline{\text{(D)}}\int_a^b f\,dx \; \leq \; \text{(R)}\int_a^b f\,dx$$

y finaliza la prueba. ■

Del resultado anterior se sigue que los conjuntos $\mathcal{D}\mathrm{ar}([a,b])$ y $\mathcal{R}([a,b])$ coinciden y, además, se cumple que

$$\text{(R)}\int_a^b f\,dx \; = \; \text{(D)}\int_a^b f\,dx$$

para cualquier $f \in \mathcal{R}([a,b]) = \mathcal{D}\mathrm{ar}([a,b])$. El número (R) $\int_a^b f\,dx$ será denotado, en lo sucesivo, por $\int_a^b f\,dx$.

Recordemos que una función acotada $\varphi : [a,b] \to \mathbb{R}$ es una **función en escalera** si existe una partición $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ de $[a,b]$ tal que φ es constante sobre cada uno de los intervalos abiertos (x_{i-1}, x_i), $i = 1, \ldots, n$. Si $c_i = \varphi(t_i)$, para cualquier $t_i \in (x_{i-1}, x_i)$, definimos la integral de Riemann de φ como

$$\int_a^b \varphi\,dx \; = \; \sum_{i=1}^{n} c_i(x_i - x_{i-1}).$$

Este análisis permite reformular la integral de Riemann del modo siguiente: si $f : [a,b] \to \mathbb{R}$ es Riemann integrable, entonces

$$\begin{aligned}\int_a^b f\,dx &= \sup\left\{ \int_a^b \varphi\,dx \ :\ \varphi \in \mathrm{Esc}([a,b]),\ \ \varphi \leq f \right\} \\ &= \text{ínf}\left\{ \int_a^b \psi\,dx \ :\ \psi \in \mathrm{Esc}([a,b]),\ \ \psi \geq f \right\}.\end{aligned}$$

Si en algún momento existe, además de la integral de Riemann otra integral diferente, escribiremos (R) $\int_a^b f\,dx$ para diferenciarla de la otra.

¿Qué tipo de funciones habitan en $\mathcal{R}([a,b])$? El siguiente resultado establece que contiene, en principio, a todas las funciones continuas definidas sobre $[a,b]$.

Teorema 8.3.9. $C([a,b]) \subseteq \mathcal{R}([a,b])$.

Prueba. Sea $f \in C([a,b])$. Puesto que f es continua y $[a,b]$ es compacto, resulta que f es uniformemente continua gracias al Teorema 3.1.6. Por lo tanto, dado $\varepsilon > 0$, existe un $\delta > 0$ tal que cualesquiera sean $x,y \in [a,b]$ con $|x-y| < \delta$ se cumple que

$$|f(x) - f(y)| \ < \ \varepsilon/(b-a).$$

Sea $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ cualquier partición de $[a,b]$ con $\|\,\mathcal{P}\,\| < \delta$. Entonces, para cada $i = 1, 2, \ldots, n$,

$$\mathrm{osc}(f, I_i) \ = \ \sup\left\{|f(x) - f(y)| : x,y \in [x_{i-1}, x_i]\right\} \ < \ \frac{\varepsilon}{(b-a)},$$

de donde se sigue que

$$\sum_{i=1}^{n} \mathrm{osc}(f, I_i) \cdot (x_i - x_{i-1}) \ < \ \frac{\varepsilon}{(b-a)} \sum_{i=1}^{n} (x_i - x_{i-1}) \ = \ \varepsilon.$$

Un llamado al Teorema 8.3.7 nos revela que $f \in \mathcal{R}([a,b])$. ■

Nota Adicional 8.3.1 Nótese que si $f \in C_c(\mathbb{R})$, entonces, por el Corolario 3.1.10, página 159, existe un intervalo compacto $[a,b] \supseteq \mathrm{sop}(f)$ tal que $f(x) = 0$ para todo $x \notin [a,b]$ y, por lo tanto, $f|_{[a,b]} \in C([a,b])$. Este hecho nos indica, gracias al resultado anterior, que siempre podemos considerar a f como una función Riemann integrable y definir

$$\int f\,dt \ := \ \int_a^b f\,dt.$$

8.3.3. El Teorema de Vitali-Lebesgue

El objetivo fundamental de esta sección es presentar el formidable e imprescindible resultado demostrado en el año 1907 por Giuseppe Vitali y Henri Lebesgue, independientemente uno del otro, el cual limita el proceso de integración de Riemann a una clase de funciones que están suficientemente próximas a las funciones continuas. Ya hemos visto que toda función continua

$f : [a,b] \to \mathbb{R}$ es Riemann integrable, y es fácil ver que si f es discontinua en un punto y, en general, en un conjunto finito de puntos, entonces f también es Riemann integrable. De modo que es enteramente natural preguntarse: *¿Cuán discontinua debe ser una función acotada* $f : [a,b] \to \mathbb{R}$ *para que ella sea Riemann integrable*? Vitali y Lebesgue encontraron un modo elegante de caracterizar a las funciones integrables según Riemann en función de la medida de Lebesgue. De modo específico, ellos demostraron que *las funciones Riemann integrables que no son continuas en todo su dominio, son precisamente las funciones acotadas que son* ***continuas casi-siempre***, es decir, si escribimos

$$\mathcal{R}([a,b]) \;=\; \big(\mathcal{R}([a,b]) \setminus C([a,b])\big) \cup C([a,b])$$

entonces $\mathcal{R}([a,b]) \setminus C([a,b])$ consiste de todas las funciones acotadas que son continuas casi-siempre. Este hecho permite entender por qué $\mathcal{R}([a,b])$ no puede contener funciones que son *muy discontinuas* como, por ejemplo, la función característica de los racionales en $[a,b]$.

La siguiente demostración del Teorema de Vitali-Lebesgue utiliza un resultado que es muy poco conocido pero que posee un abanico muy amplio de aplicaciones, nos referimos al Lema de Cousin.

Definición 8.3.10. *Sea* $\delta : [a,b] \to \mathbb{R}$ *una función* ***estrictamente positiva****. Una partición etiquetada* $\mathcal{P}_e = (t_i, [x_{i-1}, x_i])_{i=1}^n$ *de* $[a,b]$ *se llama* δ*-**fina**, o* ***subordinada*** *a* δ*, si para cada índice* $i \in \{1, 2, \ldots, n\}$*, se cumple que*

$$t_i \,\in\, [x_{i-1},\, x_i] \,\subseteq\, \big(t_i - \delta(t_i),\, t_i + \delta(t_i)\big).$$

A la función estrictamente positiva δ de la definición anterior la llamaremos **función gauge** o **calibrador**. Si ahora consideramos la función $\gamma : [a,b] \to \mathcal{I}([a,b])$ definida por

$$\gamma(t) \;=\; (t - \delta(t),\, t + \delta(t)),$$

donde $\mathcal{I}([a,b])$ es la colección de todos los subintervalos abiertos de $\mathbb{R}$, resulta que cada etiqueta t_i controla al intervalo $[x_{i-1}, x_i]$ a través de la función γ. Por esta razón, a la función γ se le suele llamar la **función que controla a** $\mathcal{P}_e$ asociada a δ. Observe que no existe, o no se exige, ninguna condición de continuidad sobre la función δ.

El siguiente resultado, el cual es la puerta de entrada a la Integral de Henstock-Kurzweil, establece que cualquier función $\delta : [a,b] \to \mathbb{R}$ estrictamente positiva siempre tiene asociada una partición etiquetada δ-fina.

Lema 8.3.11 (Cousin). *Si* $\delta : [a,b] \to \mathbb{R}$ *es un* ***calibrador****, entonces existe una* ***partición etiquetada*** *en* $[a,b]$ *que es* δ*-**fina**.*

Prueba. Suponga, para generar una contradicción, que no existe ninguna partición etiquetada de $[a,b]$ que sea δ-fina. Pongamos $I_0 = [a,b]$ y sea $c_0 = (a+b)/2$ el punto medio de I_0. Afirmamos que al menos uno de los dos intervalos $[a, c_0]$, $[c_0, b]$ no posee particiones etiquetadas δ-finas. En efecto, si ambos intervalos tuviesen una partición etiquetada δ-fina, entonces se seguiría de la Observación 1 que $[a,b]$ tendría una partición etiquetada δ-fina, lo cual negaría nuestra suposición. Llamemos $I_1 = [a_1, b_1]$ el intervalo que no posee ninguna partición etiquetada δ-fina (si ambos intervalos no poseen particiones etiquetadas δ-finass, elija uno cualquiera de ellos y nómbrelo I_1.) Sea $c_1 = (a_1 + b_1)/2$ el punto medio de I_1. Razonando como antes, al menos uno de los intervalos $[a_1, c_1]$, $[c_1, b_1]$ no posee ninguna partición etiquetada δ-fina. Elija el que no posee particiones etiquetadas δ-finas y nómbrelo $I_2 = [a_2, b_2]$. Sea c_3 el punto medio de I_2.

Continuando indefinidamente con este proceso, se obtiene una sucesión $(I_n)_{n=1}^{\infty}$ de intervalos compactos tales que:

$$I_0 \supseteq I_1 \supseteq I_2 \supseteq \cdots \supseteq I_n \supseteq I_{n+1} \supseteq \cdots \qquad \text{y} \qquad \ell(I_n) \to 0.$$

Se sigue del Teorema de Encaje de Cantor que existe un único $x_0 \in I$ tal que $\bigcap_{n=1}^{\infty} I_n = \{x_0\}$. Por otro lado, como $\delta(x_0) > 0$, podemos invocar el Principio de Arquímedes para obtener un $n \in \mathbb{N}$ tal que

$$\ell(I_n) = \frac{b-a}{2^n} < \delta(x_0).$$

Esta desigualdad nos dice que $I_n \subseteq (x_0 - \delta(x_0), x_0 + \delta(x_0))$ y, en consecuencia, el par (x_0, I_n) es, trivialmente, una partición etiquetada δ-fina de I_n. Esto, por supuesto, es contrario a la construcción de I_n, de modo que lo que habíamos supuesto es falso. La prueba es completa. ■

Otra manera de demostrar el resultado procede como sigue. Consideremos el conjunto

$$E = \{x \in [a,b] : \text{existe una partición etiquetada } \delta\text{-fina de } [a,x]\}.$$

El conjunto E es no vacío. En efecto, como $\delta(a) > 0$, escoja un x tal que $a < x < a + \delta(a)$. Es claro que $\{(x, [a, x])\}$ constituye una partición etiquetada δ-fina de $[a, x]$ y, por lo tanto, $x \in E$. Ahora bien, como E es no vacío y acotado superiormente por b, el Axioma del Supremo nos garantiza que $\sup E$ existe. Pongamos $x^* = \sup E$ y observe que $x^* \in [a, b]$. Afirmamos que $x^* \in E$. En efecto, como $x^* = \sup E$, entonces debe ocurrir que o bien $x^* \in E$, o existe un $u \in E$ tal que $u < x^*$. De darse el primer caso paramos. Si es el último caso el que ocurre, entonces puesto que $u \in E$, existe, por definición, una partición etiquetada δ-fina $\mathcal{P}_e$ de $[a, u]$. Agreguemos ahora el elemento $(x^*, [u, x^*])$ a dicha partición, es decir, sea $\mathcal{Q}_e = \mathcal{P}_e \cup (x^*, [u, x^*])$. Resulta entonces que $\mathcal{Q}_e$ es una partición etiquetada δ-fina de $[a, x^*]$ y, por consiguiente, $x^* \in E$.

La prueba finalizará una vez que logremos demostrar que $x^* = b$. Suponga, para construir una contradicción, que $x^* < b$. Elija ahora un punto arbitrario $v \in [a, b]$ de modo tal que $x^* < v < x^* + \delta(x^*)$. Puesto que $x^* \in E$, existe una partición etiquetada δ-fina de $[a, x^*]$, a la que llamaremos $\mathcal{P}_e^1$ y considere la partición $\mathcal{R}_e = \mathcal{P}_e^1 \cup (x^*, [x^*, v])$. Tenemos entonces que $\mathcal{R}_e$ es una partición etiquetada δ-fina de $[a, v]$ y, por consiguiente, $v \in E$. Teniendo en cuenta que x^* es el supremo de E, resulta que $v \le x^*$, lo que, por supuesto, contradice la elección v. Por esto. $x^* = b$ y termina la prueba. ■

Estamos ahora en posesión de los argumentos para demostrar el formidable resultado de Vitali-Lebesgue.

Teorema 8.3.12 (Vitali-Lebesgue). *Sea $f : [a, b] \to \mathbb{R}$ una función **acotada**. Las siguientes condiciones son equivalentes:*

(1) $f \in \mathcal{R}([a, b])$.

*(2) f es **continua casi-siempre**.*

Prueba. (1) $\Rightarrow$ (2) Suponga que $f \in \mathcal{R}([a, b])$ y fijemos un $\delta > 0$. En primer lugar, vamos a demostrar que el conjunto $\mathrm{D}_\delta = \{x \in [a, b] : \mathrm{osc}(f, x) \ge \delta\}$ tiene contenido cero, es decir, $\mathrm{c}_{\mu^*}(\mathrm{D}_\delta) = 0$. Sea $\varepsilon > 0$ elegido arbitrariamente y usemos el hecho de que $f \in \mathcal{R}([a, b])$ para hallar, según el Teorema 8.3.7, una partición $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ de $[a, b]$ tal que

$$\sum_{i=1}^{n} \mathrm{osc}(f, I_i) \cdot (x_i - x_{i-1}) < \varepsilon\, \delta,$$

donde $I_i = [x_{i-1}, x_i]$, $i = 1, \ldots, n$. Sea

$$J \;=\; \big\{k \in \{1, \ldots, n\} \;:\; (x_{k-1}, x_k) \cap \mathrm{D}_\delta \neq \varnothing\big\}$$

y observe que si $x \in (x_{k-1}, x_k) \cap \mathrm{D}_\delta$, entonces $\delta \leq \operatorname{osc}(f, x) \leq \operatorname{osc}(f, I_k)$, de donde se sigue que

$$\delta \sum_{k\in J}(x_k - x_{k-1}) \;\leq\; \sum_{k\in J} \operatorname{osc}(f, I_k)\cdot(x_k - x_{k-1}) \;\leq\; \sum_{i=1}^{n} \operatorname{osc}(f, I_i)\cdot(x_i - x_{i-1}) \;<\; \varepsilon\,\delta,$$

lo cual prueba que $\sum_{k\in J} \ell(I_k) = \sum_{k\in J}(x_k - x_{k-1}) < \varepsilon$. Además, como $\mathrm{D}_\delta \subseteq \bigcup_{k\in J} I_k$, resulta del Lema 6.2.20, página 253, que $\mathrm{c}_{\mu^*}(\mathrm{D}_\delta) = 0$. En particular, $\mu(\mathrm{D}_\delta) = 0$. Finalmente, como $\mathrm{Disc}(f) = \bigcup_{n=1}^{\infty} \mathrm{D}_{1/n}$, el cual se obtuvo como consecuencia del Teorema 3.1.28, resulta entonces que

$$0 \;\leq\; \mu\big(\mathrm{Disc}(f)\big) \;\leq\; \sum_{n=1}^{\infty} \mu(\mathrm{D}_{1/n}) \;=\; 0.$$

y así, f es continua casi-siempre.

(2) $\Rightarrow$ (1) Suponga que f es continua casi-siempre y sea $\varepsilon > 0$. Para demostrar que $f \in \mathcal{R}([a, b])$, vamos a probar que existe una partición $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ de $[a, b]$ para la cual se cumple que

$$\sum_{i=1}^{n} \operatorname{osc}(f, I_i)\cdot(x_i - x_{i-1}) \;<\; \varepsilon$$

e invocar después el Teorema 8.3.7. Esto lo haremos construyendo, en primer lugar, una función $\delta : [a, b] \to \mathbb{R}$ estrictamente positiva para luego aplicar el Lema de Cousin.

Si $f = 0$, no hay nada que probar. Suponga entonces que $f \neq 0$ y observe que $M = \sup\big\{|f(x)| : x \in [a, b]\big\} > 0$. Más aun, para cualquier subintervalo $J \subseteq [a, b]$ se cumple que

$$\operatorname{osc}(f, J) \;\leq\; 2M.$$

Ahora bien, como f es continua $\mu -$c.s., resulta que $\mu\big(\mathrm{Disc}(f)\big) = 0$ y, en consecuencia, existe una colección numerable $\{J_n : n \in \mathbb{N}\}$ de intervalos abiertos y disjuntos dos a dos tal que

$$\mathrm{Disc}(f) \;\subseteq\; \bigcup_{n=1}^{\infty} J_n \qquad \text{y} \qquad \sum_{n=1}^{\infty} \ell(J_n) \;<\; \frac{\varepsilon}{M}$$

Podemos ahora definir una función estrictamente positiva $\delta : [a, b] \to (0, +\infty)$ por medio del siguiente argumento:

(i) Si $z \notin \mathrm{Disc}(f)$, entonces la continuidad de f en z nos permite elegir un número $\delta(z) > 0$ tal que $|f(x) - f(z)| < \varepsilon/4(b - a)$ para cada $x \in [a, b]$ que satisfaga $|x - z| < \delta(z)$.

(ii) Si $z \in \mathrm{Disc}(f) \subseteq \bigcup_{n=1}^{\infty} J_n$, entonces existe un único intervalo J_n tal que $z \in J_n$ y como J_n es abierto, existe un $\delta(z) > 0$ tal que $\big(z - \delta(z), z - \delta(z)\big) \subseteq J_n$.

Esto termina la demostración de la existencia de la función $\delta : [a, b] \to \mathbb{R}$ estrictamente positiva. Usemos ahora el Lema de Cousin para obtener una partición etiquetada δ-fina de $[a, b]$, digamos $\mathcal{P}_e = \big(t_i, [x_{i-1}, x_i]\big)_{i=1}^{n}$. Esto significa que para cada índice $i \in \Lambda = \{1, 2, \ldots, n\}$ se cumple que

$$t_i \;\in\; [x_{i-1}, x_i] \;\subseteq\; \big(t_i - \delta(t_i), t_i + \delta(t_i)\big).$$

Defina ahora

$$\Lambda_1 = \{ i \in \Lambda : t_i \notin \mathrm{Disc}(f) \} \qquad \text{y} \qquad \Lambda_2 = \{ i \in \Lambda : t_i \in \mathrm{Disc}(f) \}.$$

Puesto que los intervalos son no-superpuestos se tiene que

$$\begin{aligned}
\sum_{i=1}^{n} \mathrm{osc}(f, I_i)\cdot(x_i - x_{i-1}) &= \sum_{i\in\Lambda_1} \mathrm{osc}(f, I_i)\cdot(x_i - x_{i-1}) + \sum_{i\in\Lambda_2} \mathrm{osc}(f, I_i)\cdot(x_i - x_{i-1}) \\
&< \sum_{i\in\Lambda_1} \frac{\varepsilon}{2(b-a)}\cdot(x_i - x_{i-1}) + \sum_{i\in\Lambda_2} 2M\cdot(x_i - x_{i-1}) \\
&\leq \frac{\varepsilon}{2(b-a)} \sum_{i\in\Lambda_1}(x_i - x_{i-1}) + 2M \sum_{i\in\Lambda_2} \ell(J_i) \\
&\leq \frac{\varepsilon}{2(b-a)} \sum_{i=1}^{n}(x_i - x_{i-1}) + 2M \sum_{i=1}^{\infty} \ell(J_i) \\
&< \frac{\varepsilon}{2(b-a)}(b-a) + 2M\frac{\varepsilon}{4M} = \varepsilon.
\end{aligned}$$

Un llamado al Teorema 8.3.7 nos dice que $f \in \mathcal{R}([a,b])$ y finaliza la prueba. ■

Nota Adicional 8.3.2 El Teorema de Vitali-Lebesgue es uno de los resultados fundamentales de la teoría de integración de Riemann pues revela exactamente cómo se caracterizan a las funciones acotadas que son Riemann integrables. Es un resultado profundo que relaciona a la integral de Riemann con un aspecto muy especial de la teoría de la medida: los conjuntos de medida cero. Existe, en la literatura, muchas otras formas interesantes y diferentes de demostrar dicho resultado. Por ejemplo, véase, T. Apostol [5], M. Botsko [19], J. LaVita [87], L. Levine [92], R. Gordon [67], etc.

¿Cuál es la importancia, además de servir como un importante ingrediente en la prueba del Teorema de Vitali-Lebesgue, del Lema de Cousin? Pues bien, el Axioma de Supremo, como sabemos, es equivalente a cada una de las siguientes declaraciones:

(a) *Cualquier sucesión de Cauchy converge* ($\mathbb{R}$ es completo).

(b) *Cualquier sucesión monótona acotada converge.*

(c) *Cualquier sucesión acotada contiene una subsucesión convergente* (Teorema de Bolzano-Weierstrass).

(d) *La intersección de cualquier sucesión encajada de intervalos cerrados y acotados es no vacía* (Teorema de Encaje de Cantor).

Por otro lado, en la demostración del Lema de Cousin vimos que el Axioma de Supremo implicaba dicho lema y que éste último implica, por ejemplo, el Teorema de Bolzano-Weierstrass. En consecuencia, *el Lema de Cousin también es equivalente a cada una de las declaraciones anteriores*. Pero además, el Lema de Cousin es una herramienta interesante y poderosa que es utilizada para demostrar muchos resultados del Análisis, como por ejemplo:

(1) *Si $f \in C([a,b])$, entonces f posee la Propiedad del Valor Intermedio.*

(2) *Si* $f \in C([a,b])$, *entonces* f *es acotada.*

(3) *Si* $f \in C([a,b])$, *entonces* f *es uniformemente continua.*

(4) *Si* $f \in C([a,b])$, *entonces* f *es Riemann integrable.*

Y un largo etcétera. La demostración de estos y, muchos otros resultados de Análisis, se pueden ver, por ejemplo, en [67].

8.3.4. Consecuencias del Teorema de Vitali-Lebesgue

La importancia del Teorema de Vitali-Lebesgue quedará evidenciada en las siguientes aplicaciones y muchas otras, algunas de las cuales aparecerán en el transcurso de estas notas. Aunque la demostración del Teorema de Vitali-Lebesgue no es elemental, aunque tampoco es difícil, algunas de sus aplicaciones son realmente maravillosas y muy simples.

(**TVL**$_0$) Si $f \in \mathcal{R}([a,b])$, entonces $\mathrm{PC}(f)$ es un G_δ-**denso** en $[a,b]$.

Prueba. Sea $f \in \mathcal{R}([a,b])$. Por el Teorema de Vitali-Lebesgue $\mu(\mathrm{Disc}(f)) = 0$ y, por lo tanto,

$$\mu(\mathrm{PC}(f)) \;=\; \mu([a,b]) - \mu(\mathrm{Disc}(f)) \;=\; b - a.$$

Suponga, por un momento, que $\mathrm{PC}(f))$ no es denso en $[a,b]$. Esto significa que en $[a,b]$ habita un cierto conjunto abierto no vacío V tal que $V \cap \mathrm{PC}(f) = \varnothing$. Por supuesto, esto último nos dice que $V \subseteq \mathrm{Disc}(f)$ y, en consecuencia, $\mu(V) = 0$ lo cual es imposible pues todo conjunto abierto no vacío posee medida estrictamente positiva. Se sigue entonces que $\mathrm{PC}(f)$ es denso en $[a,b]$ y, además, un G_δ, gracias al Teorema 3.1.25, página 168. ■

Los conjuntos G_δ-densos son, desde el punto de vista de su cardinalidad, *mucho más numerosos* que los que son simplemente densos: por ejemplo, $\mathbb{Q}$ es un conjunto denso que nunca puede aspirar a ser un G_δ-denso (una consecuencia del Teorema de Categoría de Baire) y, por supuesto, $\mathrm{card}(\mathbb{Q}) = \aleph_0$, mientras que cualquier conjunto que sea G_δ-denso posee la cardinalidad del continuo; por ejemplo, $\mathbb{I} = \mathbb{R} \setminus \mathbb{Q}$ es un G_δ-denso cuya cardinalidad es $\mathfrak{c}$. Por otro lado, a pesar de que toda función Riemann integrable posee "muchísimos puntos de continuidad", de hecho, un conjunto G_δ-denso de tales puntos gracias al resultado anterior, resulta que todas aquellas funciones acotadas que no lo son deben contener una "cantidad horrenda" de discontinuidades, en otras palabras, *funciones acotadas que no son Riemann integrables* deben ser *extremadamente discontinuas* pues, según el Teorema de Vitale-Lebesgue el conjunto de puntos de discontinuidad de cualquier

$$f \in \mathcal{B}_\infty([a,b]) \setminus \mathcal{R}([a,b])$$

es de medida positiva.

(**TVL**$_1$) Si $f, g \in \mathcal{R}([a,b])$ y $c \in \mathbb{R}$, entonces $f+g$, $f \cdot g$, $c \cdot f \in \mathcal{R}([a,b])$.

Prueba. Suponga que $f, g \in \mathcal{R}([a,b])$. Entonces $\mu(\mathrm{Disc}(f)) = 0 = \mu(\mathrm{Disc}(g))$ y como

$$\mathrm{PC}(f) \cap \mathrm{PC}(g) \;\subseteq\; \mathrm{PC}(f+g)$$

se sigue que $\mathrm{Disc}(f+g) \subseteq \mathrm{Disc}(f) \cup \mathrm{Disc}(g)$ y así, $\mu(\mathrm{Disc}(f+g)) = 0$. Esto prueba que $f+g \in \mathcal{R}([a,b])$. Similarmente se demuestra que $f \cdot g \in \mathcal{R}([a,b])$. Nótese que si g es una función constante, digamos $g(x) = c$ para todo $x \in [a,b]$, entonces $g \in \mathcal{R}([a,b])$ y, por lo anterior, se tiene que $c \cdot f \in \mathcal{R}([a,b])$. ■

($\mathbf{TVL_2}$) Si $f \in \mathcal{R}([a,b])$, entonces $|f| \in \mathcal{R}([a,b])$.

Prueba. Esto sigue del hecho de que si f es continua en $x \in [a,b]$, entonces $|f|$ también es continua en x y, por lo tanto,

$$\mathrm{Disc}(|f|) \subseteq \mathrm{Disc}(f).$$

Pero como $0 \leq \mu(\mathrm{Disc}(|f|)) \leq \mu(\mathrm{Disc}(f)) = 0$, el Teorema de Vitali-Lebesgue nos dice que $|f|$ es Riemann integrable. ■

El recíproco del resultado anterior es falso, es decir, si $|f| \in \mathcal{R}([a,b])$, entonces no es cierto, en general, que $f \in \mathcal{R}([a,b])$. En efecto, basta considerar la función $f : [0,1] \to \mathbb{R}$ dada por

$$f(x) = \begin{cases} 1 & \text{si} \quad x \in \mathbb{Q}_{[0,1]} \\ -1 & \text{si} \quad x \notin [0,1] \setminus \mathbb{Q}_{[0,1]}. \end{cases}$$

para verificar nuestra afirmación. Este ejemplo muestra que, a pesar de ser ***la integral de Riemann una integral absoluta***, en el sentido de que: $f \in \mathcal{R}([a,b]) \Rightarrow |f| \in \mathcal{R}([a,b])$, la implicación recíproca no es válida.

($\mathbf{TVL_3}$) Toda función **monótona acotada** $f : [a,b] \to \mathbb{R}$ es Riemann integrable, es decir,

$$\mathrm{Mon}([a,b]) \subseteq \mathcal{R}([a,b]).$$

Prueba. Sea $f \in \mathrm{Mon}([a,b])$. Por el Corolario 3.1.35 sabemos que $\mathrm{Disc}(f)$ es a lo más numerable y, por consiguiente, $\mu(\mathrm{Disc}(f)) = 0$. El Teorema de Vitali-Lebesgue nos garantiza entonces que $f \in \mathcal{R}([a,b])$. ■

($\mathbf{TVL_4}$) En general, cualquier función $f : [a,b] \to \mathbb{R}$ que es **acotada** y **continua excepto sobre un subconjunto a lo más numerable** de $[a,b]$, es Riemann integrable.

Sin embargo, existen funciones Riemann integrables que son discontinuas sobre un conjunto no-numerable. Por ejemplo, si $\chi_\Gamma : [0,1] \to \mathbb{R}$ es la función característica del conjunto ternario de Cantor Γ, entonces $\chi_\Gamma \in \mathcal{R}([0,1])$. En efecto, por el Ejercicio 6.8.11, página 363, sabemos que $\mathrm{Disc}(\chi_\Gamma) = \Gamma$ y como $\mu(\Gamma) = 0$, resulta del Teorema de Vitali-Lebesgue que $\chi_\Gamma \in \mathcal{R}([0,1])$. Por otro lado, si χ_{Γ_α} es la función característica de Γ_α, donde Γ_α es un conjunto tipo-Cantor con $\mu(\Gamma_\alpha) > 0$, entonces $\chi_{\Gamma_\alpha} \notin \mathcal{R}([0,1])$ pues $\mathrm{Disc}(\chi_{\Gamma_\alpha}) = \Gamma_\alpha$. Más aun, $\chi_{\mathbb{Q}\cap[0,1]} \notin \mathcal{R}([0,1])$ pues $\mathrm{Disc}(\chi_{\mathbb{Q}\cap[0,1]}) = [0,1]$. Los dos últimos ejemplos evidencian, una vez más, que

$$\mathcal{R}([a,b]) \subsetneqq \mathcal{B}_\infty([a,b]).$$

($\mathbf{TVL_5}$) Si φ es una **función en escalera**, entonces $\varphi \in \mathcal{R}([a,b])$, esto es,

$$\mathrm{Esc}([a,b]) \subseteq \mathcal{R}([a,b]).$$

Prueba. Puesto que el número de discontinuidades de cualquier función en escalera φ es finito, entonces $\varphi \in \mathcal{R}([a,b])$ gracias al Teorema de Vitali-Lebesgue. Otra manera de ver esto es recordar que $\mathrm{Esc}([a,b]) \subseteq \mathcal{D}ar([a,b])$ y como $\mathcal{D}ar([a,b]) = \mathcal{R}([a,b])$ el resultado sigue. ■

Observe que si $G \subseteq [a,b]$ es cualquier conjunto abierto no vacío, entonces $\chi_G \in \mathcal{R}([a,b])$. En efecto, escoja una colección numerable $(J_n)_{n=1}^\infty$ de intervalos abiertos disjuntos dos a dos tal que $G = \bigcup_{n=1}^\infty J_n$ y tomemos ahora un $x \in G$. Entonces existe un único $n \geq 1$ tal que $x \in J_n$ y, en consecuencia, $\chi_G(x) = \chi_{J_n}(x)$. Por consiguiente, $\mathrm{Disc}(\chi_G)$ es numerable y así, por el Teorema de Vitali-Lebesgue, $\chi_G \in \mathcal{R}([a,b])$.

(**TVL**$_6$) Si $(f_n)_{n=1}^{\infty}$ es una sucesión en $\mathcal{R}([a,b])$ que **converge uniformemente** a una **función acotada** $f : [a,b] \to \mathbb{R}$, entonces $f \in \mathcal{R}([a,b])$.

Prueba. Por el Teorema de Vitali-Lebesgue, $\mu(\mathrm{Disc}(f_n)) = 0$ para todo $n \in \mathbb{N}$. Afirmamos que $\mathrm{Disc}(f) \subseteq \bigcup_{n=1}^{\infty} \mathrm{Disc}(f_n)$. En efecto, si $x \notin \bigcup_{n=1}^{\infty} \mathrm{Disc}(f_n)$, entonces f_n es continua en x para todo $n \in \mathbb{N}$ y puesto que $f_n \to f$ uniformemente, resulta que f es continua en x, es decir, $x \notin \mathrm{Disc}(f)$. Finalmente, como

$$0 \ \leq \ \mu(\mathrm{Disc}(f)) \leq \sum_{n=1}^{\infty} \mu(\mathrm{Disc}(f_n)) \ = \ 0,$$

una nueva aplicación del Teorema de Vitali-Lebesgue nos revela que $f \in \mathcal{R}([a,b])$. Esto termina la prueba. ■

(**TVL**$_7$) La **función de Thomae** $f : [0,1] \to \mathbb{R}$,

$$f(x) \ = \ \begin{cases} \dfrac{1}{q} & \text{si} \quad x = \dfrac{p}{q} \in \mathbb{Q}_{\mathrm{irre}} \cap [0,1] \\ 0 & \text{si} \quad x \in [0,1] \setminus \mathbb{Q}_{[0,1]}. \end{cases}$$

es Riemann integrable.

Prueba. Por el Ejemplo 3.1.1, página 165, sabemos que f es discontinua sólo en $\mathbb{Q}_{[0,1]}$ y como tal conjunto es de medida nula, el resultado sigue del Teorema de Vitali-Lebesgue. ■

(**TVL**$_8$) Si $f \in \mathcal{R}([a,b])$ y $g : f([a,b]) \to \mathbb{R}$ es **continua**, entonces $g \circ f \in \mathcal{R}([a,b])$.

Prueba. Como $f \in \mathcal{R}([a,b])$, el Teorema de Vitali-Lebesgue nos muestra que f es continua $\mu -$ c.s. y ya que g es continua sobre $f([a,b])$, resulta que $g \circ f$ es continua en todo punto $x \in \mathrm{PC}(f)$. Por lo tanto, $\mathrm{Disc}(g \circ f) \subseteq \mathrm{Disc}(f)$ y como $\mu(\mathrm{Disc}(f)) = 0$, el resultado sigue por una nueva aplicación del Teorema de Vitali-Lebesgue. ■

Es importante destacar que *la composición de dos funciones Riemann integrables no es necesariamente Riemann integral*. Por ejemplo, si $f : [0,1] \to [0,1]$ es la función de Thomae y si $g : [0,1] \to \mathbb{R}$ se define como

$$g(x) \ = \ \begin{cases} 1 & \text{si} \quad x \in (0,1] \\ 0 & \text{si} \quad x = 0, \end{cases}$$

entonces $f, g \in \mathcal{R}([0,1])$ y, sin embargo,

$$(g \circ f)(x) \ = \ \begin{cases} 1 & \text{si} \quad x \in [0,1] \setminus \mathbb{Q}[0,1] \\ 0 & \text{si} \quad x \in \mathbb{Q}[0,1], \end{cases}$$

no es Riemann integrable.

Otra caracterización de las funciones Riemann integrables y que puede resultar de utilidad es la siguiente, demostrada por Leo M. Levine [92] en 1977.

(**TVL**$_9$) Sea $f : [a,b] \to \mathbb{R}$ una **función acotada** sobre $[a,b]$. Son equivalentes

(1) $f \in \mathcal{R}([a,b])$.

(2) $\mu([a,b] \setminus L_f) = 0$, donde $L_f = \{x \in [a,b] : f(x^-) \text{ existe}\}$.

Prueba. Suponga que $f \in \mathcal{R}([a,b])$. Por el Teorema de Vitali-Lebesgue se tiene que $\mu(\text{Disc}(f)) = 0$ y puesto que

$$\begin{aligned}\text{Disc}(f) &= \big(\text{Disc}(f) \cap L_f\big) \cup \big(\text{Disc}(f) \cap ([a,b] \setminus L_f)\big) \\ &= \big(\text{Disc}(f) \cap L_f\big) \cup \big([a,b] \setminus L_f\big),\end{aligned}$$

resulta que $\mu([a,b] \setminus L_f) = 0$.

Recíprocamente, suponga que $\mu([a,b] \setminus L_f) = 0$. Del Teorema 3.1.29, página 170, sabemos que $\text{Disc}(f) \cap L_f$ es a lo más numerable y, en consecuencia, posee medida cero. El resultado sigue de la igualdad anterior. ■

($\mathbf{TVL}_{10}$) Si $f \in \mathcal{R}([a,b])$, entonces f es **medible según Lebesgue**.

Prueba. Suponga que $f \in \mathcal{R}([a,b])$. Por el Teorema de Vitali-Lebesgue f es continua $\mu -$ c.s. y se sigue del Teorema 7.2.5 que ella es medible. ■

8.3.5. Propiedades Básicas de la Integral de Riemann

Nuestra primer objetivo a la vista es demostrar que la integral de Riemann es una aplicación lineal sobre $\mathcal{R}([a,b])$.

Teorema 8.3.13. *Sean $f, g \in \mathcal{R}([a,b])$ y $c \in \mathbb{R}$. Entonces $f+g$, $cf \in \mathcal{R}([a,b])$ y se cumple que*

(1) $\displaystyle\int_a^b (f+g)\,dx = \int_a^b f\,dx + \int_a^b g\,dx,$

(2) $\displaystyle\int_a^b (cf)\,dx = c\int_a^b f\,dx.$

Prueba. (1) Por ($\mathbf{TVL}_1$) sabemos que $f+g \in \mathcal{R}([a,b])$. Sean ahora $\mathcal{P}$ y $\mathcal{Q}$ particiones arbitrarias de $[a,b]$. Entonces

$$\begin{aligned}\int_a^b (f+g)\,dx &\leq \mathsf{U}(f+g, \mathcal{P}\cup\mathcal{Q}) \\ &\leq \mathsf{U}(f, \mathcal{P}\cup\mathcal{Q}) + \mathsf{U}(g, \mathcal{P}\cup\mathcal{Q}) \\ &\leq \mathsf{U}(f, \mathcal{P}) + \mathsf{U}(g, \mathcal{Q})\end{aligned}$$

Fijando la partición $\mathcal{Q}$, vemos que

$$\begin{aligned}\int_a^b (f+g)\,dx &\leq \sup\big\{\mathsf{U}(f,\mathcal{P}) \,:\, \mathcal{P} \in \mathbb{P}[a,b]\big\} + \mathsf{U}(g,\mathcal{Q}) \\ &= \int_a^b f\,dx + \mathsf{U}(g,\mathcal{Q})\end{aligned}$$

y como $\mathcal{Q}$ es arbitraria, concluimos que

$$\int_a^b (f+g)\,dx \;\leq\; \int_a^b f\,dx \,+\, \sup\left\{\mathsf{U}(g,\mathcal{Q}) \,:\, \mathcal{Q}\in\mathbb{P}[a,b]\right\}$$
$$= \int_a^b f\,dx \,+\, \int_a^b g\,dx.$$

Si ahora usamos las sumas inferiores de Darboux se comprueba igualmente que

$$\int_a^b (f+g)\,dx \;\geq\; \int_a^b f\,dx \,+\, \int_a^b g\,dx.$$

y con ello termina la prueba de (1). La demostración de (2) se deja a cargo del lector. ■

Teorema 8.3.14. *Sea $f\in\mathcal{R}([a,b])$ con $f\geq 0$. Si*

$$\int_a^b f\,dx \;=\; 0,$$

*entonces $f=0$ **casi-siempre**. En particular, si f es **continua**, entonces $f=0$.*

Prueba. En primer lugar, vamos a demostrar que el conjunto

$$E_n \;=\; \left\{x\in[a,b] \,:\, f(x)\geq 1/n\right\}$$

posee medida cero para todo $n\in\mathbb{N}$. En efecto, fijemos un $n\in\mathbb{N}$ y sea $\varepsilon>0$. Como $f\in\mathcal{R}([a,b])$ y $\int_a^b f=0$, se sigue de la definición de integral que existe una partición $\mathcal{P}=\{x_0,x_1,\ldots,x_n\}$ de $[a,b]$ tal que $\mathsf{U}(f,\mathcal{P})<\varepsilon/n$. Sea $J=\{i : E_n\cap[x_{i-1},x_i]\neq\varnothing\}$ y observe que si $i\in J$, entonces $M_i\geq 1/n$. Por esto,

$$\frac{\varepsilon}{n} \;>\; \mathsf{U}(f,\mathcal{P}) \;\geq\; \sum_{i\in J} M_i(x_i-x_{i-1}) \;\geq\; \frac{1}{n}\sum_{i\in J}(x_i-x_{i-1}),$$

lo cual implica que $\sum_{i\in J}(x_i-x_{i-1})=\sum_{i\in J}\ell(I_i)<\varepsilon$. Esto prueba que E_n es de contenido cero, en particular, de medida cero. Finalmente, como

$$E \;=\; \left\{x\in[a,b] \,:\, f(x)\neq 0\right\} \;=\; \bigcup_{n=1}^{\infty} E_n$$

resulta, por la subaditividad de μ, que $\mu(E)=0$. ■

Teorema 8.3.15. *Para cada $f\in\mathcal{R}([a,b])$ con $f\geq 0$, se cumple que $\int_a^b f\,dx\geq 0$. En particular, si $f,g\in\mathcal{R}([a,b])$ con $f\leq g$, entonces*

$$\int_a^b f\,dx \;\leq\; \int_a^b g\,dx.$$

Prueba. Sea $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ cualquier partición de $[a, b]$ y observe que como $f \geq 0$, entonces $m_i = \sup\{f(x) : x \in [x_{i-1}, x_i]\} \geq 0$ para cada $i = 1, 2, \ldots, n$, de donde se obtiene que

$$\int_a^b f\, dx \;\geq\; \mathsf{L}(f, \mathcal{P}) \;\geq\; 0.$$

Si $f \leq g$, entonces $g - f \geq 0$ y el resultado sigue de lo anterior y el Teorema 8.3.13. ■

Del resultado anterior se sigue que

Corolario 8.3.16. *Si $f \in \mathcal{R}([a, b])$, entonces $|f| \in \mathcal{R}([a, b])$ y se cumple que*

$$\left|\int_a^b f\, dx\right| \;\leq\; \int_a^b |f|\, dx.$$

En particular, si $m \leq f(x) \leq M$ para todo $x \in [a, b]$, entonces

$$m \cdot (b - a) \;\leq\; \int_a^b f\, dx \;\leq\; M \cdot (b - a).$$

Prueba. Sea $f \in \mathcal{R}([a, b])$. Por (**TVL**$_2$) sabemos que $|f| \in \mathcal{R}([a, b])$ y puesto que

$$-|f| \;\leq\; f \;\leq\; |f|,$$

se sigue del resultado anterior que

$$-\int_a^b |f|\, dx \;\leq\; \int_a^b f\, dx \;\leq\; \int_a^b |f|\, dx,$$

lo cual significa que

$$\left|\int_a^b f\, dx\right| \;\leq\; \int_a^b |f|\, dx.$$

La segunda parte es consecuencia de la primera y del hecho de que $\int_a^b 1\, dx = b - a$. ■

Teorema 8.3.17 (Teorema del Valor Medio para Integrales). *Sea $f \in C([a, b])$. Entonces existe un número $c \in [a, b]$ tal que*

$$\int_a^b f\, dx \;=\; f(c) \cdot (b - a).$$

Prueba. Puesto que f es continua sobre el compacto $[a, b]$, ella es acotada y asume su máximo y su mínimo sobre tal intervalo, es decir, existen $x_0, x_1 \in [a, b]$ tales que

$$f(x_0) \;=\; \text{ínf}\{f(x) \;:\; x \in [a, b]\} \qquad \text{y} \qquad f(x_1) \;=\; \sup\{f(x) \;:\; x \in [a, b]\}.$$

Ahora bien, como $f(x_0) \leq f(x) \leq f(x_1)$ para cualquier $x \in [a, b]$, se sigue del Corolario 8.3.16 que

$$f(x_0) \cdot (b - a) \;\leq\; \int_a^b f\, dx \;\leq\; f(x_1) \cdot (b - a),$$

en otras palabras,

$$f(x_0) \;\leq\; \frac{\displaystyle\int_a^b f\,dx}{b-a} \;\leq\; f(x_1).$$

Puesto que f es continua, ella satisface la Propiedad del Valor Intermedio y, por lo tanto, existe $c \in [a,b]$ tal que

$$f(c) \;=\; \frac{\displaystyle\int_a^b f\,dx}{b-a}.$$

Esto termina la prueba. ■

Resulta de mucha utilidad y, además, conveniente definir la integral $\int_a^b f\,dx$ aun si $a \not< b$. De allí que, suponga que $b < a$ y que $f \in \mathcal{R}([b,a])$. Convenimos en definir

$$\int_a^b f\,dx \;=\; -\int_b^a f\,dx \qquad \text{y} \qquad \int_a^a f\,dx \;=\; 0.$$

Además, si $x \in [a,b]$ y f es Riemann integrable sobre cada uno de los intervalos $[a,x]$ y $[x,b]$, entonces $f \in \mathcal{R}([a,b])$ y

$$\int_a^b f\,dx \;=\; \int_a^x f\,dx + \int_x^b f\,dx.$$

La prueba se deja a cargo del lector.

8.3.6. El Teorema Fundamental del Cálculo

Desde sus inicios, la derivada y la integral fueron creadas como procesos independientes una de la otra. La noción de derivada estuvo motivada por el problema de encontrar líneas tangentes a una curva, mientras que la definición de integral se desarrolla con el deseo de determinar áreas debajo de una curva. El Teorema Fundamental del Cálculo para la integral de Riemann revela una notable y extraordinaria relación entre ambos procesos. Fundamentalmente, dicho resultado es una declaración acerca de la relación inversa entre la diferenciación y la integración sustentada, por supuesto, sobre la base de una definición particular de integral. Por consiguiente, un Teorema Fundamental del Cálculo debe establecer: (1) los criterios para la existencia de la integral, (2) una declaración sobre la relación inversa entre la función que es integrada y la integral, y (3) un método para evaluar la integral. La integral de Riemann, aunque proporciona tales criterios, se limita sólo a funciones integrables que poseen primitivas. Una histórica conección entre la derivada y la integral es contada, de modo agradable, por Fyodor A. Medvedev en el Capítulo 4 de [98].

La demostración del Teorema Fundamental del Cálculo requiere tan sólo del teorema, probablemente más utilizado en Análisis: el Teorema del Valor Medio para Derivadas el cual establece que:

Teorema 8.3.18 (Teorema del Valor Medio para Derivadas). *Sea $f \in C([a,b])$ y suponga que f es **diferenciable** en $[a,b]$. Entonces existe un $t \in [a,b]$ tal que*

$$f(b) - f(a) \;=\; f'(t)\cdot(b-a).$$

Prueba. Considere la aplicación $F : [a,b] \to \mathbb{R}$ definida por

$$F(x) \;=\; f(x) \,-\, f(a) \,-\, \frac{f(b)-f(a)}{b-a}(x-a), \quad \text{para todo } x \in [a,b].$$

Observe que por ser F continua en $[a,b]$ ella alcanza su máximo y (o) su mínimo en algún punto $t \in [a,b]$. Por otro lado, la diferenciabilidad de F en t garantiza que $F'(t) = 0$, de donde se obtiene el resultado deseado. ■

Aunque ya hemos definido lo que es una primitiva o anti-derivada de una función, no hace daño volver a recordarla.

Definición 8.3.19. *Sea $f : [a,b] \to \mathbb{R}$ una función. Una función $F : [a,b] \to \mathbb{R}$ se dice que es una **primitiva**, o **anti-derivada** de f, si F es diferenciable en $[a,b]$ y $F' = f$.*

En este caso, también se dice que f **posee una primitiva**, o que f es una **derivada**. En lo que sigue escribiremos

$$\mathcal{P}\mathrm{rimi}([a,b]) \;=\; \big\{ f : [a,b] \to \mathbb{R} \;:\; f \text{ posee una primitiva} \big\}.$$

Observe que si $f \in \mathcal{P}\mathrm{rimi}([a,b])$ y F es una primitiva de f, entonces $F_c(x) = F(x) + c$ satisface $F_c' = f$ donde c es una constante arbitraria. Por consiguiente, toda $f \in \mathcal{P}\mathrm{rimi}([a,b])$ posee infinitas primitivas. Sin embargo, si $f \in \mathcal{R}([a,b])$, entonces la función

$$F(x) = \int_a^x f\,dt \quad \text{para todo } x \in [a,b],$$

es la única primitiva de f satisfaciendo $F(a) = 0$.

Por otro lado, es importante destacar los siguientes dos hechos:

(1) *Una función $f : [a,b] \to \mathbb{R}$ puede poseer una primitiva F y no ser de la forma $F(x) = \int_a^x f\,dt$.* En efecto, la función $f : [0,1] \to \mathbb{R}$ definida por

$$f(x) \;=\; \begin{cases} 2x\,\mathrm{sen}(1/x^2) - \dfrac{2}{x}\cos(1/x^2) & \text{si } \; 0 < x \leq 1, \\ 0 & \text{si } \; x = 0 \end{cases}$$

posee una primitiva $F : [0,1] \to \mathbb{R}$ dada por

$$F(x) \;=\; \begin{cases} x^2\,\mathrm{sen}(1/x^2) & \text{si } \; 0 < x \leq 1, \\ 0 & \text{si } \; x = 0. \end{cases}$$

Sin embargo, F no es de la forma $F(x) = \int_a^x f\,dt$ ya que $F' = f$ no es Riemann integrable pues no es acotada.

(2) *No cualquier función Riemann integrable $f : [a,b] \to \mathbb{R}$ posee una primitiva.* Para ver esto, observe que toda función $f : [a,b] \to \mathbb{R}$ que posea una discontinuidad de salto en algún punto $c \in (a,b)$ es garantía suficiente para impedir que ella tenga una primitiva. ¿Por qué esto es así? La respuesta es simple: recordemos que una discontinuidad de salto en un punto $c \in (a,b)$ significa que los límites $f(c^-) = \lim_{x \to c^-} f(x)$ y $f(c^+) = \lim_{x \to c^+} f(x)$ existen pero son distintos. Por lo tanto, si una función f con una discontinuidad de salto posee una primitiva F sobre $[a,b]$,

entonces se generaría un contrasentido ya que, gracias al Corolario 3.1.33, página 173, sabemos que todas las discontinuidades de $F' = f$ son de la segunda especie.

Por supuesto, uno puede construir funciones que son Riemann integrables que no poseen primitivas y sin discontinuidades de salto. Por ejemplo, considere las funciones $g, f : [0,1] \to \mathbb{R}$ definidas por

$$g(x) = \begin{cases} \operatorname{sen}(1/x) & \text{si } 0 < x \leq 1, \\ 1 & \text{si } x = 0, \end{cases} \qquad f(x) = \begin{cases} g(x) & \text{si } 0 < x \leq 1, \\ 0 & \text{si } x = 0. \end{cases}$$

Es claro que $f, g \in \mathcal{R}([a,b])$. Por otro lado, la función $h : [0,1] \to \mathbb{R}$ definida por

$$h(x) = \begin{cases} x^2 \cos(1/x) & \text{si } 0 < x \leq 1, \\ 0 & \text{si } x = 0 \end{cases}$$

es derivable con $h' = \varphi + f$, donde φ es continua en $[0,1]$. Luego,

$$\Phi(x) = \int_0^x \varphi(t)\, dt \quad \text{para todo } 0 \leq x \leq 1,$$

es una primitiva de φ y, en consecuencia, $h - \Phi$ es una primitiva de f. Veamos ahora que g no posee primitiva. En efecto, suponga, por un momento, que g posee una primitiva. Resulta entonces que $g - f$ también posee una primitiva, digamos, F. Esto, por supuesto, contradice el Teorema del Valor Intermedio para derivadas, Teorema 3.1.21, página 164, que establece que que si una función derivada toma dos valores, entonces toma todos los valores intermedios, lo cual es imposible ya que $F' = g - f$ toma únicamente los valores 0 y 1.

Estos ejemplos nos indican que la Teoría de Integración de Riemann no resuelve satisfactoriamente el problema de la búsqueda de primitivas. Sin embargo, cuando una función Riemann integrable posee una primitiva, el Teorema Fundamental del Cálculo es la herramienta ideal que permite evaluar dichas integrales de un modo realmente simple.

Teorema 8.3.20 (Primer Teorema Fundamental del Cálculo). *Sea $f \in \mathcal{R}([a,b])$. Si $F : [a,b] \to \mathbb{R}$ es una* ***primitiva*** *de f, entonces*

$$\int_a^b f\, dx = F(b) - F(a).$$

Prueba. Sea $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ cualquier partición de $[a,b]$. Puesto que F es diferenciable en $[a,b]$, el Teorema de Valor Medio para Derivadas nos garantiza, para cada $i \in \{1, 2, \ldots, n\}$, la existencia de un $t_i \in (x_{i-1}, x_i)$ tal que

$$F(x_i) - F(x_{i-1}) = F'(t_i) \cdot (x_i - x_{i-1}) = f(t_i) \cdot (x_i - x_{i-1}).$$

Por esto,

$$\begin{aligned} F(b) - F(a) &= \big[F(x_1) - F(x_0)\big] + \big[F(x_2) - F(x_1)\big] + \cdots + \big[F(x_n) - F(x_{n-1})\big] \\ &= \sum_{i=1}^{n} f(t_i)(x_i - x_{i-1}). \end{aligned}$$

Por otro lado, teniendo en cuenta que (véase (8.3.3), página 415)

$$\mathsf{L}(f,\mathcal{P}) \;\leq\; \sum_{i=1}^{n} f(t_i)(x_i - x_{i-1}) \;\leq\; \mathsf{U}(f,\mathcal{P}),$$

resulta de lo anterior que

$$\mathsf{L}(f,\mathcal{P}) \;\leq\; F(b) - F(a) \;\leq\; \mathsf{U}(f,\mathcal{P}).$$

Puesto que esta última desigualdad es válida para cualquier partición $\mathcal{P}$ de $[a,b]$, se obtiene, usando ahora el hecho de que f es Riemann integrable, que

$$\int_a^b f\,dx \;=\; F(b) \;-\; F(a).$$

La prueba es completa. ■

Nota Adicional 8.3.3 El Primer Teorema Fundamental del Cálculo establece que cualquier función Riemann integrable que posea una primitiva es Newton integrable, es decir,

$$f \in \mathcal{R}([a,b]) \cap \mathcal{P}\mathrm{rimi}([a,b]) \quad \Rightarrow \quad f \in \mathcal{N}([a,b]).$$

Sin embargo, como ya hemos visto, existen funciones Riemann integrables que no poseen primitivas. Esto no dice que, en general,

$$\mathcal{R}([a,b]) \not\subseteq \mathcal{N}([a,b]).$$

Por otro lado, la función de Volterra $F : [a,b] \to \mathbb{R}$, véase el Teorema 8.3.29, página 444, es un ejemplo de una función la cual posee una derivada $F'(x) = f(x)$ para todo $x \in [a,b]$ tal que f es acotada pero no es Riemann integrable, en otras palabras,

$$\mathcal{N}([a,b]) \not\subseteq \mathcal{R}([a,b]).$$

Tal vez el lector se sienta más cómodo en recordar la forma más corriente de formular el Primer Teorema Fundamental del Cálculo:

Primer Teorema Fundamental del Cálculo: *Si* $F : [a,b] \to \mathbb{R}$ *es diferenciable sobre* $[a,b]$ *y* $F' \in \mathcal{R}([a,b])$, *entonces*

$$\int_a^b F'\,dx \;=\; F(b) \;-\; F(a).$$

El Segundo Teorema Fundamental del Cálculo para la integral de Riemann trata sobre la diferenciabilidad de la función $F : [a,b] \to \mathbb{R}$ definida por

$$F(x) \;=\; \int_a^x f(t)\,dt \quad \text{para todo } x \in [a,b],$$

donde $f \in \mathcal{R}([a,b])$. Recordemos que una función función $f : [a,b] \to \mathbb{R}$ se llama Lipschitz si existe una constante $M > 0$ tal que

$$|f(x) - f(y)| \;\leq\; M|x - y|$$

para todo $x, y \in [a,b]$. Es claro que toda función Lipschitz es uniformemente continua.

Teorema 8.3.21 (Segundo Teorema Fundamental del Cálculo). *Sea $f \in \mathcal{R}([a,b])$ y defina la función $F : [a,b] \to \mathbb{R}$ por*

$$F(x) \;=\; \int_a^x f(t)\,dt \quad \text{para todo } x \in [a,b].$$

*Entonces F es **Lipschitz** sobre $[a,b]$. Más aun, si f es **continua** en $x_0 \in [a,b]$, entonces F es **diferenciable** en x_0 y $F'(x_0) = f(x_0)$.*

Prueba. Para demostrar que F es Lipschitz, observe que si $f = 0$ sobre $[a,b]$, entonces también $F = 0$ sobre $[a,b]$ y finaliza la prueba. Si éste no es el caso, sea

$$M \;=\; \sup\{|f(x)| \,:\, x \in [a,b]\}.$$

Observe que $M > 0$. Sean $x_1, x_2 \in [a,b]$ elegidos arbitrariamente. Entonces,

$$\begin{aligned} |F(x_1) - F(x_2)| &= \left| \int_a^{x_1} f\,dt - \int_a^{x_2} f\,dt \right| \\ &= \left| \int_{x_1}^{x_2} f\,dt \right| \\ &\leq \int_{x_1}^{x_2} |f|\,dt \;\leq\; M|x_1 - x_2|, \end{aligned}$$

de donde se obtiene que F es Lipschitz. En particular, F es continua sobre $[a,b]$.

Para demostrar la segunda parte, suponga que f es continua en $x_0 \in (a,b)$; el lector podrá suplir los detalles si $x_0 = a$ o $x_0 = b$. Por definición,

$$F'(x_0) \;=\; \lim_{h\to 0} \frac{F(x_0+h) - F(x_0)}{h}.$$

Suponga primero que $h > 0$. Entonces

$$F(x_0+h) - F(x_0) \;=\; \int_{x_0}^{x_0+h} f\,dt.$$

Definamos

$$m_h = \inf\{f(x) \,:\, x \in [x_0, x_0+h]\} \qquad \text{y} \qquad M_h = \sup\{f(x) \,:\, x \in [x_0, x_0+h]\}.$$

Del Corolario 8.3.16 se sigue que

$$m_h \cdot h \;\leq\; \int_{x_0}^{x_0+h} f\,dt \;\leq\; M_h \cdot h$$

y, por lo tanto,

$$m_h \;\leq\; \frac{F(x_0+h) - F(x_0)}{h} \;\leq\; M_h.$$

Es fácil ver que las desigualdades anteriores también se cumplen si $h < 0$. Puesto que f es continua en x_0, resulta que

$$\lim_{h\to 0} m_h \;=\; f(x_0) \;=\; \lim_{h\to 0} M_h,$$

lo cual demuestra que

$$F'(x_0) \;=\; \lim_{h\to 0} \frac{F(x_0+h)-F(x_0)}{h} \;=\; f(x_0)$$

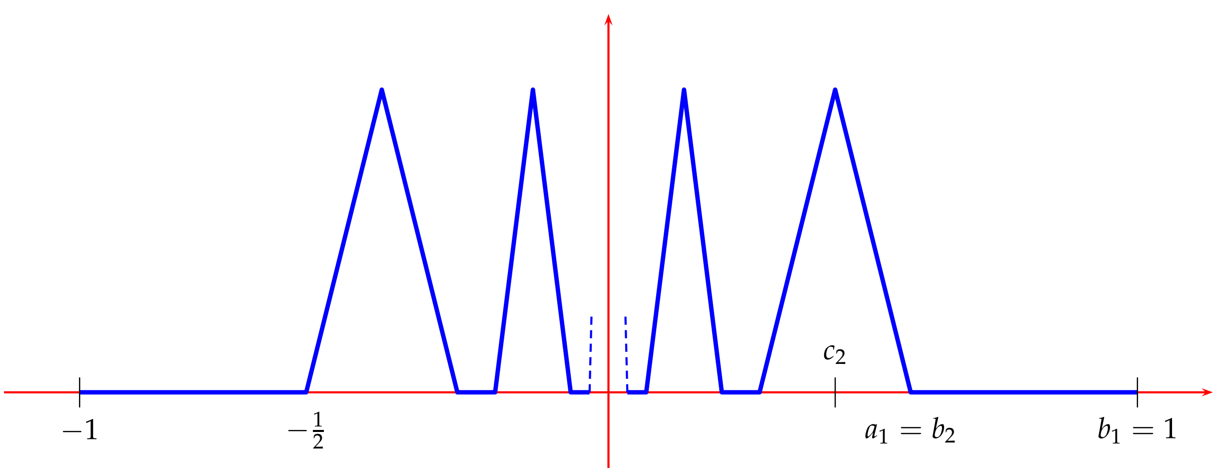
y termina la prueba. ■

Es importante destacar que en el Segundo Teorema Fundamental del Cálculo, la continuidad de f en $x_0 \in [a,b]$ es crucial. Para ver por qué esto es así, vamos a construir una función f que no es continua en $x_0 = 0$ tal que $F(x) = \int_a^x f(t)\,dt$ no es diferenciable en dicho punto. Para cada $n \in \mathbb{N}$, sean

$$a_n \;=\; 2^{-n}, \quad b_n \;=\; 2^{-(n-1)} \quad \text{y} \quad c_n \;=\; \frac{1}{2}(a_n+b_n).$$

Considere ahora la función $f : [-1,1] \to \mathbb{R}$ definida por

$$f(x) \;=\; \begin{cases} \dfrac{x-a_{2n}}{c_{2n}-a_{2n}} & \text{si } a_{2n} \le n \le c_{2n} \\[2mm] \dfrac{b_{2n}-x}{b_{2n}-c_{2n}} & \text{si } c_{2n} \le n \le b_{2n} \\[2mm] 0 & \text{si } a_{2n-1} < x < b_{2n-1} \quad \text{y} \quad x=0 \\[2mm] f(-x) & \text{si } -1 \le x < 0. \end{cases}$$



Observe que la función f es continua en todo punto de $[-1,1]$ excepto en $x = 0$. Por consiguiente, por (TVL_4), $f \in \mathcal{R}([-1,1])$. Sin embargo, la función $F : [-1,1] \to \mathbb{R}$ dada por

$$F(x) \;=\; \int_{-1}^{x} f(t)\,dt$$

no es diferenciable en $x = 0$. En efecto, si $h = b_{2n}$, entonces

$$\frac{F(h)-F(0)}{h} \;=\; \frac{1}{h}\int_0^h f(t)\,dt \;=\; \frac{1}{h}\sum_{n=i}^{\infty} 2^{-2n-1} \;=\; \frac{3^{-1}2^{-(2^i-1)}}{2^{-(2^i-1)}} \;=\; \frac{1}{3},$$

mientras que si $h = b_{2n+1}$, entonces

$$\frac{F(h)-F(0)}{h} = \frac{1}{h}\int_0^h f(t)\,dt = \frac{1}{h}\sum_{n=i}^{\infty} 2^{-2n+1} = \frac{3^{-1}2^{-(2^i+1)}}{2^{-2^i}} = \frac{1}{6},$$

De lo anterior se sigue que

$$\lim_{h\to 0}\frac{F(h)-F(0)}{h}$$

no existe y, en consecuencia, F no es diferenciable en $x=0$.

Una consecuencia inmediata del Teorema 8.3.21, es el siguiente resultado el cual establece que *toda función continua* $f:[a,b]\to\mathbb{R}$ *posee una primitiva.*

Corolario 8.3.22. *Sea* $f\in C([a,b])$ *y defina* $F:[a,b]\to\mathbb{R}$ *por*

$$F(x) = \int_a^x f\,dt \quad \textit{para todo } x\in[a,b].$$

Entonces F *es* ***diferenciable*** *en* $[a,b]$ *y* $F'(x)=f(x)$ *para todo* $x\in[a,b]$*; en otras palabras,*

$$C([a,b]) \subseteq \mathcal{N}([a,b]).$$

¿Qué ocurre si, en el resultado anterior, no se exige que f sea continua sobre $[a,b]$, pero se requiere que ella siga siendo Riemann integrable? En este caso se obtiene lo siguiente.

Corolario 8.3.23. *Sea* $f\in\mathcal{R}([a,b])$ *y defina* $F:[a,b]\to\mathbb{R}$ *por*

$$F(x) = \int_a^x f\,dt \quad \textit{para todo } x\in[a,b].$$

Entonces $F'=f$ ***casi-siempre*** *sobre* $[a,b]$.

Prueba. Puesto que $f\in\mathcal{R}([a,b])$, el Teorema de Vitali-Lebesgue nos dice que f es continua casi-siempre, es decir, $\mu(\mathrm{Disc}(f))=0$. Por el Segundo Teorema Fundamental del Cálculo, Teorema 8.3.21, tenemos que $F'(x)=f(x)$ para todo $x\in[a,b]\setminus\mathrm{Disc}(f)$. Esto termina la prueba. ■

Ejemplo 8.3.3. *Ssea* $f:[0,1]\to\mathbb{R}$ *definida por*

$$f(x) = \begin{cases} 1 & \textit{si } x\in\{1,1/2,1/3,\ldots\} \\ 0 & \textit{si } x\in[0,1]\setminus\{1,1/2,1/3,\ldots\}. \end{cases}$$

Resulta que $f\in\mathcal{R}([0,1])$ *y*

$$F(x) = \int_0^x f\,dx = 0 \quad \textit{para todo } x\in[0,1].$$

Esto muestra que F *es* ***diferenciable*** *en* $[0,1]$ *pero* $F'\neq f$, *aunque* $F'=f$ *casi-siempre en* $[0,1]$ *gracias al Corolario 8.3.23.*

El resultado anterior sugiere considerar un tipo de "primitivas" cuya derivada no coincida con f en todo punto de su dominio.

Definición 8.3.24. *Una función $f : [a, b] \to \mathbb{R}$ posee una* ***primitiva - c.s.*** *si existe una función continua $F : [a, b] \to \mathbb{R}$ y un conjunto medible $E \subseteq [a, b]$ con $\mu(E) = 0$ tal que $F'(t)$ existe y es igual a $f(t)$ para todo $t \in [a, b] \setminus E$.*

Observe que si $t \in E$, entonces $F'(t)$ no existe, o si existe, no es igual a $f(t)$. Si el conjunto E en la definición anterior es *a lo más numerable*, entonces diremos que f posee una $\aleph_0$-**primitiva**

8.3.7. Limitaciones y Deficiencias de la Integral de Riemann

El interés de esta sección es presentar algunas de las limitaciones y deficiencias encontradas en la integral de Riemann que producen ciertas fisuras en su edificación y que, por lo tanto, la objetan los matemáticos, desde el punto de vista teórico, como su integral preferida. Comencemos:

(1) La primera seria limitación de la integral de Riemann es que sólo las **funciones acotadas** definidas sobre un intervalo cerrado y acotado son susceptibles de ser integradas, por lo que *todas las funciones que no son acotadas y muchas otras que son acotadas* no entran en el juego.

(2) La segunda limitación es que la integral de Riemann está confinada a **intervalos cerrados y acotados**, es decir, dicha integral no admite una definición en intervalos no acotados y, menos aun, en subconjuntos arbitrarios de $\mathbb{R}$. Sin embargo, existe una manera de **extenderla a intervalos infinitos** a través del paso al límite: éste proceso es conocido como **la integral impropia de Riemann**. Por ejemplo, si $f : \mathbb{R} \to \mathbb{R}$ es una función tal $f \in \mathcal{R}([-x, x])$ para cada $x > 0$, podemos definir la **integral de Cauchy-Riemann** como

$$(\mathrm{CR}) \int_{-\infty}^{+\infty} f(t)\, dt \;:=\; \lim_{x\to\infty} \int_{-x}^{x} f(t)\, dt \tag{1}$$

siempre que dicho límite exista. Sin embargo, esta extensión de la integral de Riemann no funciona adecuadamente en el siguiente sentido: una propiedad básica que debe poseer toda *buena integral* es que ella sea invariante por traslación, es decir, que la integral no cambie si f se traslada hacia la derecha o hacia la izquierda, vale decir, que la igualdad

$$(\mathrm{CR}) \int_{-\infty}^{+\infty} f(t)\, dt \;=\; (\mathrm{CR}) \int_{-\infty}^{+\infty} f(t - t_0)\, dt \tag{2}$$

se cumpla para cualquier $t_0 \in \mathbb{R}$. Sin embargo, esto propiedad no es, en general, válida para la definición dada en (1). En efecto, si definimos

$$f(t) \;=\; \begin{cases} 1 & \text{si } x > 0 \\ 0 & \text{si } x = 0 \\ -1 & \text{si } x < 0, \end{cases}$$

resulta que

$$\int_{-x}^{x} f(t)\, dt \;=\; \int_{-x}^{0} f(t)\, dt + \int_{0}^{x} f(t)\, dt \;=\; -x + x \;=\; 0$$

para todo $x \in \mathbb{R}$. Si ahora trasladamos hacia la derecha a f por una unidad, tendremos que

$$\int_{-x}^{x} f(t-1)\,dt \;=\; \int_{-x}^{1} f(t-1)\,dt + \int_{1}^{x} f(t-1)\,dt \;=\; -(x+1)+(x-1) \;=\; -2$$

para todo $x > 1$ lo que, por supuesto, es inaceptable para la validez de la ecuación (2).

Para intervalos acotados la integral de Riemann sufre del mismo defecto que el caso anterior, es decir, puede ocurrir que una función $f : [a,b] \to \mathbb{R}$ tenga la particularidad de ser Riemann integrable en todo intervalo $[c,b]$ con $a < c < b$ y, sin embargo, f no sea integrable sobre $[a,b]$. Por ejemplo, la función

$$f(x) \;=\; \begin{cases} 1/x & \text{si } x \in (0,1] \\ 0 & \text{si } x = 0 \end{cases}$$

no es Riemann integrable sobre $[0,1]$ ya que ella no es acotada, pero satisface $f \in \mathcal{R}([c,1])$ para todo $0 < c < 1$.

En general, suponga que f es Riemann integrable en todo intervalo $[c,b]$ con $a < c < b$. Nótese que esto garantiza que f es acotada sobre $[c,b]$ para todo $c \in (a,b)$ pero no hay garantía que f sea acotada sobre $[a,b]$. Si f no es acotada sobre $[a,b]$ se define la **integral impropia de Riemann**, o **integral de Cauchy-Riemann**, sobre $[a,b]$ como

$$(\mathrm{CR})\int_a^b f\,dt \;=\; \lim_{c\to a^+}\int_c^b f\,dt$$

siempre que dicho límite exista. Una definición similar se obtiene si f es Riemann integrable en todo intervalo $[a,c]$ con $a < c < b$ pero no es acotada sobre $[a,b]$. En general, esta definición se puede dar para cualquier $c \in (a,b)$ siempre que f no sea acotada sobre $[a,b]$.

(3) Una de las deficiencias más notables de la integral de Riemann trata sobre la **convergencia puntual de sucesiones de funciones Riemann integrables**. ¿Qué significa esto? Pues bien, suponga que $f : [a,b] \to \mathbb{R}$ es una función acotada y que $(f_n)_{n=1}^{\infty}$ es una sucesión en $\mathcal{R}([a,b])$ tal que $\lim_{n\to\infty} f_n(x) = f(x)$ para cada $x \in [a,b]$.

(a) ¿Es f Riemann integrable?

(b) Si f es Riemann integrable, ¿será cierto que

$$\lim_{n\to\infty}\int_a^b f_n\,dx \;=\; \int_a^b f\,dx\,?$$

La respuesta es, en general, negativa en ambos casos. Para la primera, sea $(q_n)_{n=1}^{\infty}$ una lista (sin duplicaciones) de los números racionales en $[0,1]$ y para cada $n \in \mathbb{N}$, considere la función acotada $f_n : [0,1] \to \mathbb{R}$ definida por

$$f_n(x) \;=\; \begin{cases} 1 & \text{si } x \in \{q_1, q_2, \ldots, q_n\} \\ 0 & \text{si } x \in [0,1] \setminus \{q_1, q_2, \ldots, q_n\}. \end{cases}$$

Puesto que cada f_n posee sólo un número finito de discontinuidades, ella es, por el Teorema de Vitali-Lebesgue, Riemann integrable. Sin embargo, para cada $x \in [0,1]$,

$$\lim_{n\to\infty} f_n(x) \;=\; \chi_{\mathbb{Q}\cap[0,1]}(x)$$

y ya sabemos que la función límite $\chi_{\mathbb{Q}\cap[0,1]} \notin \mathcal{R}([0,1])$.

Para responder negativamente a la segunda interrogante, considere las funciones acotadas $f, f_n : [0,1] \to \mathbb{R}$ definidas por $f = 0$, y para $n \geq 1$,

$$f_n(x) \;=\; \begin{cases} n & \text{si } 0 < x < 1/n \\ 0 & \text{si } x \in [0,1] \setminus (0,1/n). \end{cases}$$

Observe que tanto f, así como cada f_n, son Riemann integrables, $f_n \to f$ puntualmente y sin embargo,

$$1 \;=\; \lim_{n\to\infty} \int_a^b f_n\,dx \;\neq\; \int_a^b f\,dx \;=\; 0.$$

Estos dos ejemplos indican que un teorema de convergencia para la integral de Riemann que cumpla (b) tiene que requerir una condición más fuerte que la convergencia puntual, o una o varias condiciones adicionales además de la convergencia puntual. Un primer resultado positivo en esta dirección, el cual es una generalización de un teorema de Cauchy para funciones continuas, es el siguiente:

Teorema 8.3.25 (Convergencia Uniforme). *Si $(f_n)_{n=1}^{\infty}$ es una sucesión de funciones en $\mathcal{R}([a,b])$ que* ***converge uniformemente*** *a una función $f : [a,b] \to \mathbb{R}$, entonces $f \in \mathcal{R}([a,b])$ y se cumple que*

$$\lim_{n\to\infty} \int_a^b f_n\,dx \;=\; \int_a^b f\,dx. \tag{CU}$$

Prueba. Por el Teorema de Vitali-Lebesgue sabemos que $f \in \mathcal{R}([a,b])$, véase (**TVL**$_6$). Para demostrar (CU), sea $\varepsilon > 0$. Como $f_n \to f$ uniformemente, existe un $N \in \mathbb{N}$ tal que si $n \geq N$ entonces

$$|f_n(x) - f(x)| \;<\; \frac{\varepsilon}{b-a} \quad \text{para todo } x \in [a,b].$$

Se sigue del Corolario 8.3.16 que, para todo $n \geq N$,

$$\left| \int_a^b f_n\,dx - \int_a^b f\,dx \right| \;\leq\; \int_a^b |f_n - f|\,dx \;<\; \frac{\varepsilon}{b-a}(b-a) \;=\; \varepsilon.$$

Esto prueba (CU). ■

El siguiente resultado también garantiza una respuesta positiva a la interrogante planteada en (b). Aunque en éste se prescinde de la convergencia uniforme, se requiere, sin embargo, además de la convergencia puntual, que la sucesión de funciones sea monótona creciente y que la función limite sea Riemann integrable. Una manera elegante de demostrar este resultado se logra usando el Lema de Cousin.

Teorema 8.3.26 (Convergencia Monótona). *Sea $(f_n)_{n=1}^{\infty}$ una sucesión en $\mathcal{R}([a,b])$ tal que:*

(a) $f_n \to f$ ***puntualmente*** *sobre $[a,b]$ y*

(b) $(f_n)_{n=1}^{\infty}$ es ***monótona creciente****.*

Si $f \in \mathcal{R}([a,b])$, entonces

$$\lim_{n\to\infty} \int_a^b f_n\,dx \;=\; \int_a^b f\,dx.$$

Prueba. Para cada $n \in \mathbb{N}$, sea $g_n = f - f_n$. Entonces $(g_n)_{n=1}^{\infty}$ es una sucesión de funciones no-negativas, decreciente y Riemann integrables que, además, converge puntualmente a 0. Puesto que las funciones Riemann integrables son acotadas, uno puede escoger una constante $M > 0$ tal que $0 \leq g_1 \leq M$. En particular,

$$0 \leq g_n \leq M \quad \text{para todo } n \geq 1,$$

esto debido a que la sucesión $(g_n)_{n=1}^{\infty}$ es decreciente. Como $g_n \in \mathcal{R}([a,b])$ para cada $n \in \mathbb{N}$, el Teorema de Vitali-Lebesgue nos dice que existe un conjunto $E_n \subseteq [a,b]$ de medida cero tal que g_n es continua en cada punto de $[a,b] \setminus E_n$. Sea

$$E = \bigcup_{n=1}^{\infty} E_n$$

y observe que $\mu(E) = 0$, gracias a la subaditividad numerable de μ. Esto último significa que, para un $\varepsilon > 0$ dado, existe una colección numerable $\big(J_n\big)_{n=1}^{\infty}$ de intervalos abiertos y disjuntos dos a dos tales que

$$E \subseteq \bigcup_{n=1}^{\infty} J_n \qquad \text{y} \qquad \sum_{n=1}^{\infty} \ell(J_n) < \frac{\varepsilon}{4M}.$$

Por supuesto, cada punto $z \in E$ pertenece a un único intervalo J_n y como J_n es abierto, se puede elegir un número $\delta(z) > 0$ tal que

$$\big(z - \delta(z),\, z + \delta(z)\big) \subseteq J_n. \tag{1a}$$

Por otro lado, como $g_n \to 0$ puntualmente, uno puede seleccionar, para cada $z \in [a,b] \setminus E$, un entero positivo $N(z,\varepsilon)$ tal que

$$g_n < \frac{\varepsilon}{4(b-a)}$$

para todo $n \geq N(z,\varepsilon)$. Miremos ahora a la función $g_{N(z,\varepsilon)}$. Ésta función es continua sobre el conjunto $[a,b] \setminus E_{N(z,\varepsilon)}$ y ya que $[a,b] \setminus E \subseteq [a,b] \setminus E_{N(z,\varepsilon)}$, resulta que $g_{N(z,\varepsilon)}$ es continua en cada punto $z \in [a,b] \setminus E$. Por lo tanto, para cada $z \in [a,b] \setminus E$, podemos determinar un $\delta(z) > 0$ tal que

$$g_{N(z,\varepsilon)}(x) < \frac{\varepsilon}{4(b-a)}$$

para todo $x \in \big(z - \delta(z),\, z + \delta(z)\big) \cap [a,b]$. Debido a que la sucesión $(g_n)_{n=1}^{\infty}$ es decreciente, la desigualdad anterior sigue siendo valida para todo $x \in \big(z - \delta(z),\, z + \delta(z)\big) \cap [a,b]$ y cualquier entero $n \geq N(z,\varepsilon)$, es decir,

$$g_n(x) < \frac{\varepsilon}{4(b-a)}. \tag{2a}$$

El análisis anterior nos muestra que la función $z \to \delta(z)$ es estrictamente positiva para todo $z \in [a,b]$. De acuerdo al Lema de Cousin, página 419, uno puede elegir una partición etiquetada δ-fina $\mathcal{P}_e = (s_i, [x_{i-1}, x_i])_{i=1}^{n}$ de $[a,b]$, lo cual quiere decir que

$$s_i \in [x_{i-1}, x_i] \subseteq \big(s_i - \delta(s_i),\, s_i + \delta(s_i)\big)$$

para $i = 1, 2, \ldots, n$. Fijemos ahora un conjunto arbitrario de etiquetas $\{t_1, \ldots, t_n\}$ en los intervalos asociados a la partición $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ y sean

$$\Lambda_1 = \{i : t_i \in E\} \qquad \text{y} \qquad \Lambda_2 = \{i : t_i \in [a,b] \setminus E\}.$$

Definamos
$$N_\varepsilon \;=\; \max\{N(t_i,\varepsilon)\;:\; i=1,2,\ldots,n\},$$
y observe que:

(i) Si $t_i\in E$, entonces por $(1a)$ existe un único intervalo J_{n_i} tal que
$$t_i \;\in\; [x_{i-1},x_i] \;\subseteq\; \big(t_i-\delta(t_i),\, t_i+\delta(t_i)\big) \;\subseteq\; J_{n_i},$$
de donde se obtiene que
$$\sum_{i\in\Lambda_1}(x_i-x_{i-1}) \;\leq\; \sum_{i\in\Lambda_1}\ell(J_{n_i}) \;\leq\; \sum_{n=1}^{\infty}\ell(J_n) \;<\; \frac{\varepsilon}{4M}.$$

(ii) Si $t_i\in[a,b]\setminus E$ y $n\geq N_\varepsilon$, resulta de $(2a)$ que $g_n(t_i)<\varepsilon/4(b-a)$.

Estas dos observaciones nos indican que si $n\geq N_\varepsilon$, entonces
$$\begin{aligned}\sum_{i=1}^{n} g_n(t_i)(x_i-x_{i-1}) \;&=\; \sum_{i\in\Lambda_1} g_n(t_i)(x_i-x_{i-1}) \;+\; \sum_{i\in\Lambda_2} g_n(t_i)(x_i-x_{i-1})\\ &<\; M\cdot\frac{\varepsilon}{4M} \;+\; \frac{\varepsilon}{4(b-a)}\cdot(b-a) \;=\; \varepsilon,\end{aligned}$$
y, por consiguiente,
$$\begin{aligned}\int_a^b g_n\,dx \;&=\; \inf\big\{\mathsf{L}(g_n,\mathcal{Q})\;:\;\mathcal{Q}\in\mathbb{P}[a,b]\big\}\\ &\leq\; \mathsf{L}(g_n,\mathcal{P}) \;\leq\; \sum_{i=1}^{n} g_n(t_i)(x_i-x_{i-1}) \;<\; \varepsilon\end{aligned}$$
para cualquier $n\geq N_\varepsilon$. Esto último es, por supuesto, equivalente a que
$$\lim_{n\to\infty}\int_a^b g_n\,dx \;=\; 0$$
y finaliza la prueba. ■

Es importante destacar que la integrabilidad del límite f en el Teorema de la Convergencia Monótona no puede ser omitida de la hipótesis pues, como se sabe, las condiciones (a) y (b) no son suficientes para garantizar que $f\in\mathcal{R}([a,b])$. Por tal motivo, el Teorema de la Convergencia Monótona no proporciona un criterio de integrabilidad.

En el siguiente resultado, establecido originalmente por Cesare Arzelà (1885), también garantiza la igualdad (CU), aunque tampoco constituye un criterio de integrabilidad á la Riemann pues, similar al resultado anterior, hay que exigir que la función límite sea Riemann integrable y que la sucesión sea uniformemente acotada.

Definición 8.3.27. *Una familia $\mathcal{F}$ de funciones a valores reales definidas sobre un conjunto X se dice que es **uniformemente acotada** sobre X, si existe una constante $M>0$ tal que*
$$|f(x)| \;\leq\; M \quad \text{para toda } f\in\mathcal{F} \text{ y todo } x\in X.$$

Nótese que si $(f_n)_{n=1}^{\infty}$ es una sucesión de funciones en $\mathcal{B}_\infty([a,b])$, entonces ella es **uniformemente acotada** si existe una constante $M > 0$ tal que

$$\sup_{n\in\mathbb{N}} \| f_n \|_\infty \;\leq\; M,$$

donde, para cada $n \geq 1$, $\| f_n \|_\infty \;=\; \sup_{x\in[a,b]} |f_n(x)|$.

Teorema 8.3.28 (Convergencia Acotada). *Sea $(f_n)_{n=1}^{\infty}$ una sucesión en $\mathcal{R}([a,b])$ tal que*

*(a) $f_n \to f$ **puntualmente** sobre $[a,b]$ y*

*(b) $(f_n)_{n=1}^{\infty}$ es **uniformemente acotada**.*

Si $f \in \mathcal{R}([a,b])$, entonces

$$\lim_{n\to\infty} \int_a^b f_n\,dx \;=\; \int_a^b f\,dx.$$

Prueba. Suponga que (a) y (b) se cumplen. En este caso se tiene que f es acotada pero no hay garantía de que ella sea Riemann integrable. Suponga entonces que $f \in \mathcal{R}([a,b])$. Observe que la sucesión $(|f_n - f|)_{n=1}^{\infty}$ es una sucesión de funciones Riemann integrables, no-negativas y uniformemente acotada que, además, converge puntualmente a la función cero. Si el teorema puede ser demostrado para este caso especial, entonces la desigualdad

$$\left| \int_a^b f_n\,dx - \int_a^b f\,dx \right| \;\leq\; \int_a^b |f_n - f|\,dx$$

indica que la conclusión del teorema también se cumple. Por consiguiente, sólo es necesario probar este caso especial del Teorema de la Convergencia Acotada. Suponga entonces que $f = 0$, $f_n \geq 0$ y que $\sup_{n\in\mathbb{N}} \| f_n \|_\infty \leq M$ para algún $M > 0$. Aceptemos por un momento que la conclusión es falsa, es decir, que

$$\lim_{n\to\infty} \int_a^b f_n\,dx \;>\; 0.$$

Sin perder generalidad, podemos asumir que $\int_a^b f_n\,dx > 0$ para todo $n \geq 1$. Seleccionemos un $\varepsilon > 0$ arbitrario, pero suficientemente pequeño, de modo que

$$\int_a^b f_n\,dx \;>\; 4\varepsilon(b-a) \quad \text{para todo} \quad n \geq 1.$$

Fijemos un $n \in \mathbb{N}$ y escojamos, por ser $f_n \in \mathcal{R}([a,b])$, una partición de $[a,b]$, digamos $\mathcal{P} = \{x_0, x_1, \ldots, x_k\}$, tal que

$$\int_a^b f_n\,dx - \sum_{i=1}^{k} m_i(f_n)(x_i - x_{i-1}) \;\leq\; \left| \int_a^b f_n\,dx - \sum_{i=1}^{k} m_i(f_n)(x_i - x_{i-1}) \right| \;<\; 2\varepsilon(b-a)$$

donde $m_i(f_n) = \inf\{f_n(x) : x \in [x_{i-1}, x_i]\}$ para $i = 1, 2, \ldots, k$. En particular,

$$\sum_{i=1}^{k} m_i(f_n)(x_i - x_{i-1}) \;>\; 2\varepsilon(b-a).$$

Sean $J_0 = \{i : m_i(f_n) < \varepsilon\}$ y $J_1 = \{i : m_i(f_n) \geq \varepsilon\}$ y considere el conjunto $V_n = \bigcup_{i \in J_1} I_i$. Observe que

$$f_n(x) \geq \varepsilon \quad \text{para todo} \quad x \in V_n.$$

Más aun,

$$\begin{aligned}
2\varepsilon(b-a) &< \sum_{i=1}^{k} m_i(f_n)(x_i - x_{i-1}) \\
&= \sum_{i \in J_0} m_i(f_n)(x_i - x_{i-1}) + \sum_{i \in J_1} m_i(f_n)(x_i - x_{i-1}) \\
&< \sum_{i \in J_0} \varepsilon \cdot (x_i - x_{i-1}) + \sum_{i \in J_1} M \cdot (x_i - x_{i-1}) \\
&\leq \varepsilon \cdot (b-a) + M \cdot \ell(V_n),
\end{aligned}$$

de donde se sigue que $\ell(V_n) > \delta$, donde $\delta = \varepsilon(b-a)/M$. Puesto que la elección de n se hizo de modo arbitrario, este proceso genera una sucesión $(V_n)_{n=1}^{\infty}$ de conjuntos cerrados tal que $\ell(V_n) > \delta$ para todo $n \geq 1$. Pongamos

$$G = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} V_k = \limsup_{n \to \infty} V_n.$$

Por el Ejercicio 8.5 (6), página 451, $G \neq \varnothing$. Esto significa que para cada $z \in G$, existen infinitos n's para los cuales $z \in V_n$. Para tales n's se tiene que $f_n(z) \geq \varepsilon$ lo que conduce a una contradicción ya que $(f_n(z))_n$ converge a cero. Esta contradicción indica que $\lim_{n\to\infty} \int_a^b f_n\, dx = 0$ y termina la prueba. ■

Nota Adicional 8.3.4 Dos comentarios son necesarios referente al Teorema de la Convergencia Acotada para la integral de Riemann. El primero es que a dicho teorema se le presta muy poca atención en los cursos básicos de Análisis por el simple hecho de que la integral de Riemann no es la integral preferida por los matemáticos en tales cursos. Ella ha sido reemplazada, casi en su totalidad, por la integral de Lebesgue la cual es mucho más amplia y donde el Teorema de la Convergencia Acotada (para la integral de Lebesgue) es muy fácil de probar. Hay que destacar, por oro lado, que en la integral de Lebesgue no se exige que la función límite en el Teorema de la Convergencia Acotada sea Lebesgue integrable. El segundo comentario, y tal vez el más importante, es que dicho teorema es difícil de demostrar (hasta este momento) usando sólo las herramientas de la integral de Riemann. Recordemos que la demostración del Teorema de la Convergencia Acotada para la integral de Riemann hace uso del Ejercicio 8.5 (6) y es en este punto donde la demostración de ese resultado comienza por hacerse difícil y tediosa. Para más información recomendamos [64, 93, 97]. La demostración dada por Jonathan W. Lewin [93] puede pensarse como más simple que las de sus predecesores.

(4) Otra deficiencia de la integral de Riemann tiene que ver con la **propiedad casi-siempre**, es decir, si $f, g : [a,b] \to \mathbb{R}$ son funciones acotadas con $f \in \mathcal{R}([a,b])$ y $f = g \quad \mu -$ c.s., entonces no necesariamente se cumple que $g \in \mathcal{R}([a,b])$. En efecto, si se considera la función $f = 0$

sobre $[0,1]$, entonces $f \in \mathcal{R}([0,1])$. Sin embargo, la función característica de los racionales en $[0,1]$, $g = \chi_{\mathbb{Q}\cap[0,1]}$ no es Riemann integrable, aunque ella satisface $f = g \ \ \mu -$ c.s. ya que $\{x \in [0,1] : f(x) \neq g(x)\} = \mathbb{Q}_{[0,1]}$ tiene medida cero.

Sin embargo, si $f, g \in \mathcal{R}([a,b])$ y $f = g \ \mu -$ c.s., entonces

$$\int_a^b f\,dx = \int_a^b g\,dx.$$

Véase el Ejercicio 8.4.4 en la página 448.

(5) La última deficiencia que abordaremos en estas notas tiene que ver con el Primer Teorema Fundamental del Cálculo, el cual establece que para cada $x \in [a,b]$, la igualdad

$$f(x) - f(a) = \int_a^x f'(t)\,dt$$

es válida siempre que f sea diferenciable sobre $[a,b]$ y $f' \in \mathcal{R}([a,b])$.

¿Por qué es necesario que f', en el Primer Teorema Fundamental del Cálculo, sea Riemann integrable? ¿Es posible encontrar una función acotada $f : [a,b] \to \mathbb{R}$ que sea diferenciable con f' sea acotada pero no Riemann integrable? La respuesta, desafortunadamente, (¿o afortunadamente?) es que una tal función existe. El primero en construir una función con esas características fue Vito Volterra en 1881 basado en la siguiente observación de Ulisse Dini:

Dini, 1878. *Si una función no constante $f : [a,b] \to \mathbb{R}$ posee una derivada acotada sobre $[a,b]$, y f' se anula sobre un subconjunto denso de $[a,b]$, entonces f' no puede ser Riemann integrable.*

Aunque a Dini le fue imposible construir un ejemplo, el genio admirable de Volterra utilizó esa idea para construir, en primer lugar, un conjunto tipo-Cantor de medida positiva, luego una función cuya derivada era acotada pero discontinua precisamente sobre dicho conjunto, y finalmente aplicar el Teorema de Vitali-Lebesgue para concluir que tal función no es Riemann integrable. Puesto que estamos interesados en una primitiva, sus discontinuidades no pueden ser de la primera especie. En efecto, gracias al Teorema de Darboux, Teorema 3.1.20, sabemos que toda derivada satisface la Propiedad del Valor Intermedio y, por lo tanto, sus discontinuidades son de la segunda especie, Corolario 3.1.33, página 173. Una función que sirve como modelo en la construcción del ejemplo de Volterra es la derivada de la función $f : [0,1] \to \mathbb{R}$ definida por

$$f(x) = \begin{cases} x^2 \operatorname{sen}\left(\dfrac{1}{x}\right) & \text{si} \quad 0 < x \leq 1 \\ 0 & \text{si} \quad x = 0. \end{cases}$$

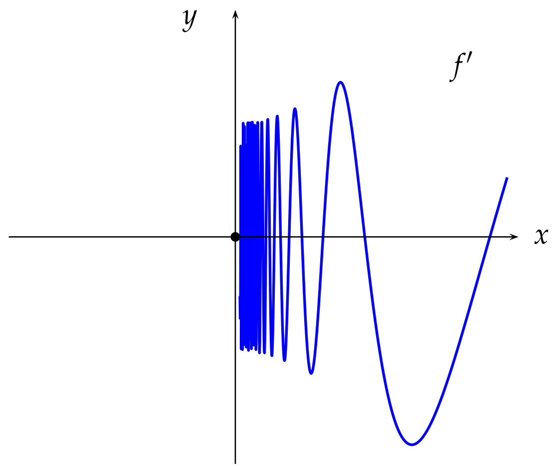
Observe que si $x \in (0,1]$, entonces $f'(x) = 2x\operatorname{sen}(1/x) - \cos(1/x)$ y cuando $x = 0$, se tiene que $f'(0) = \lim_{t\to 0^+} t^2\,\frac{\operatorname{sen}(1/t)}{t} = 0$. Por esto,

$$f'(x) = \begin{cases} 2x\operatorname{sen}\left(\dfrac{1}{x}\right) - \cos\left(\dfrac{1}{x}\right) & \text{si} \quad 0 < x \leq 1 \\ 0 & \text{si} \quad x = 0. \end{cases}$$

la cual es claramente acotada pues

$$|f'(x)| \leq 2|x|\left|\operatorname{sen}\left(\frac{1}{x}\right)\right| + \left|\cos\left(\frac{1}{x}\right)\right| \leq 3$$

para todo $x \in [0,1]$ y, por supuesto, f' no tiene límite en el punto $x = 0$, es decir, 0 es una discontinuidad de la segunda especie para f'. Por lo tanto, su oscilación



en dicho punto es positiva. La idea entonces es hacer una copia de f, modificarla un poco y luego pegarla en cada uno de los intervalos abiertos que fueron eliminados en la construcción de un conjunto tipo-Cantor de medida positiva, de modo tal que el comportamiento de la nueva función en cada uno de los extremos de los intervalos removidos sea exactamente igual que el comportamiento de f en $x = 0$. Vayamos a los detalles.

Teorema 8.3.29 (Volterra). *Existe una función $F : [0,1] \to \mathbb{R}$ la cual es* ***diferenciable*** *en todo punto de $[0,1]$ con F'* ***acotada*** *sobre $[0,1]$, pero F'* ***no es Riemann integrable*** *sobre $[0,1]$.*

Prueba. Sea Γ_α el conjunto tipo-Cantor con $\mu(\Gamma_\alpha) > 0$ construido en la Sección 5.1.6, página 229 y sea $((a_n, b_n))_{n=1}^{\infty}$ la lista, disjunta, de todos los intervalos abiertos borrados en su construcción. Para cada $n \in \mathbb{N}$, considere la función $G_n : [a_n, b_n] \to \mathbb{R}$ definida por $G_n(x) = f(x - a_n)$, para todo $x \in [a_n, b_n]$, esto es

$$G_n(x) \;=\; \begin{cases} (x-a_n)^2 \operatorname{sen}\left(\dfrac{1}{x-a_n}\right) & \text{si} \quad a_n < x \leq b_n \\ 0 & \text{si} \quad x = a_n. \end{cases}$$

Como ya fue establecido,

$$G_n'(x) \;=\; \begin{cases} 2(x-a_n) \operatorname{sen}\left(\dfrac{1}{x-a_n}\right) - \cos\left(\dfrac{1}{x-a_n}\right) & \text{si} \quad a_n < x \leq b_n \\ 0 & \text{si} \quad x = a_n. \end{cases}$$

De hecho, como el grafo de G_n oscila indefinidamente cuando x se aproxima a a_n, existen infinitos extremos relativos. Sea

$$E \;=\; \{x \in (a_n, b_n) \,:\, x \;\text{ es un extremo relativo de } G_n\}.$$

Entonces $G_n'(x) = 0$ para cada $x \in E$. Observe que entre dos extremos relativos consecutivos siempre existe un x tal que $|G'(x)| = 1$. Este hecho nos permite concluir que G' no es continua en $x = a_n$. En efecto, la sucesión $(t_j)_{j=1}^{\infty}$ definida por $t_j = a_n + 1/j\pi$ para $j \geq 1$ satisface

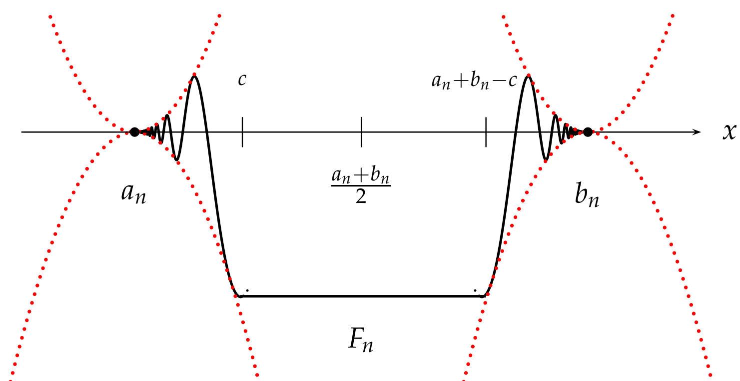
$$\lim_{j\to\infty} t_j \;=\; a_n \qquad \text{pero} \qquad \lim_{j\to\infty} |G_n'(t_j)| \;=\; 1 \;\neq\; 0 \;=\; |G_n'(a_n)|.$$

Sea c el mayor número en $(a_n, (a_n + b_n)/2)$ para el cual $G_n'(c) = 0$ y escoja un d en $((a_n + b_n)/2, b_n)$ tal que $c - a_n = b_n - d$. En este caso $d = a_n + b_n - c$ y se cumple que

$$(c - a_n)^2 \operatorname{sen}\left(\frac{1}{c - a_n}\right) = -(b_n - d)^2 \operatorname{sen}\left(\frac{1}{d - b_n}\right).$$

Defina ahora $F_n : [a_n, b_n] \to \mathbb{R}$ por

$$F_n(x) = \begin{cases} (x - a_n)^2 \operatorname{sen}\left(\dfrac{1}{x - a_n}\right) & \text{si} \quad a_n < x \leq c \\ (c - a_n)^2 \operatorname{sen}\left(\dfrac{1}{c - a_n}\right) & \text{si} \quad c < x < d \\ -(x - b_n)^2 \operatorname{sen}\left(\dfrac{1}{x - b_n}\right) & \text{si} \quad d \leq x < b_n \end{cases}$$



Nótese que F_n exhibe el mismo comportamiento oscilatorio que G_n tanto en a_n como en b_n. Además, F_n es constante en el intervalo (c, d), de modo que F_n' existe en cualquier punto de $[a_n, b_n]$. Como ya fue establecido, resulta que F_n' es acotada sobre $[a_n, b_n]$, pero no es continua en a_n ni tampoco en b_n. Más aun,

$$\begin{aligned} |F_n(x)| &\leq |x - a_n|^2 \leq |x - b_n|^2 && \text{si } a_n \leq x \leq c, \\ |F_n(x)| &\leq |c - a_n|^2 \leq |x - a_n|^2 && \text{si } c < x < d, \\ |F_n(x)| &\leq |d - b_n|^2 \leq |x - b_n|^2 && \text{si } c < x < d, \\ |F_n(x)| &\leq |x - b_n|^2 \leq |x - a_n|^2 && \text{si } d \leq x \leq b_n. \end{aligned}$$

De donde se sigue que $|F_n(x)|$ está acotada por ambas $|x - a_n|^2$ y $|x - b_n|^2$ para todo $x \in [a_n, b_n]$.

Para completar la construcción de nuestra función, defina $F : [0, 1] \to \mathbb{R}$ por

$$F(x) = \begin{cases} F_n(x) & \text{si} \quad x \in (a_n, b_n) \subseteq \bigcup_{j=1}^{\infty} (a_j, b_j) \\ 0 & \text{si} \quad x \in [0, 1] \setminus \bigcup_{n=1}^{\infty} (a_n, b_n) = \Gamma_\alpha. \end{cases}$$

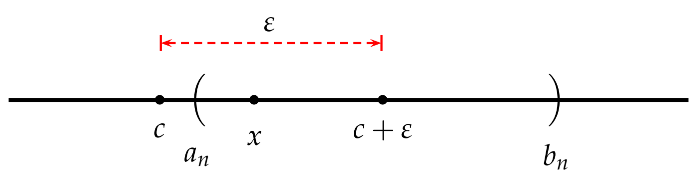
(a) F es **diferenciable** en cada punto de $[0,1]$.

Fijemos cualquier $c \in [0,1]$. Si $c \in \bigcup_{n=1}^{\infty}(a_n,\, b_n)$, entonces por lo visto anteriormente tenemos que $F'(c)$ existe. Suponga ahora que $c \in \Gamma_\alpha$ y probemos que

$$F'_+(c) \;=\; \lim_{x\to c^+} \frac{F(x)-F(c)}{x-c} \;=\; 0.$$

Sea $\varepsilon > 0$ y sea $x \in (c, c+\varepsilon)$. Si $x \in \Gamma_\alpha$, entonces $F(x) = F(c) = 0$ y, por lo tanto, $F'_+(c) = 0$. Suponga entonces que $x \notin \Gamma_\alpha$. En este caso, $x \in \bigcup_{n=1}^{\infty}(a_n,\, b_n)$ y, en consecuencia, existe un único $n \geq 1$ tal que $x \in (a_n, b_n)$. Por esto,

$$\left|\frac{F(x)-F(c)}{x-c}\right| \;\leq\; \frac{|F_n(x)|}{|x-a_n|} \;\leq\; \frac{|x-a_n|^2}{|x-a_n|} \;=\; |x-a_n| \;<\; \varepsilon.$$



Esto muestra que $F'_+(c) = 0$. Con un argumento similar se prueba que $F'_-(c) = 0$ y, por consiguiente, $F'(c)$ existe para todo $c \in [0,1]$.

(b) F' es **acotada** sobre $[0,1]$.

Esto sigue del hecho de que $|F'(x)| = |F'_n(x)| \leq 3$ para todo $x \in [0,1]$.

(c) $\mathrm{Disc}(F') = \Gamma_\alpha$.

Sea $c \in \Gamma_\alpha$. Del Teorema 5.1.6, página 227, existe una subsucesión $(a_{n_k})_{k=1}^{\infty}$ de $(a_n)_{n=1}^{\infty}$ convergiendo a c. Para cada $k \in \mathbb{N}$, existe un entero $q_k > k$ tal que

$$|F'(x_k)| \;=\; |F'_{n_k}(x_k)| \;=\; 1 \qquad \text{donde } \; x_k \;=\; a_{n_k} + \frac{1}{q_k \pi}.$$

La sucesión $(x_k)_{k=1}^{\infty}$ converge c, pero la sucesión $(F'(x_k))_{k=1}^{\infty}$ no converge al punto $F'(c) = 0$. Esto prueba que F' no es continua en $c \in \Gamma_\alpha$ y como c era arbitrario, resulta que

$$\mathrm{Disc}(F') \;=\; \Gamma_\alpha.$$

Puesto que $\mu(\Gamma_\alpha) > 0$, se sigue del Teorema de Vitali-Lebesgue que la función F' no es Riemann integrable. ■

Acabamos de ver que existe una función continua $f : [a,b] \to \mathbb{R}$ tal que f' es acotada pero discontinua sobre un conjunto medible de medida positiva y, en consecuencia, no es Riemann integrable. Este ejemplo nos muestra que la integral de Riemann no recupera derivadas, es decir, la fórmula de Newton-Leibniz no se satisface. Por otro lado, por el Corolario 2.2.47 sabemos que $f' \in \mathcal{B}_1([a,b])$ y, por lo tanto, $\mathrm{PC}(f')$ es un G_δ-denso en $[a,b]$. Esto muestra que

$$\mathcal{R}([a,b]) \;\not\subseteq\; \mathcal{B}_1([a,b]).$$

Por supuesto, también es cierto que $\mathcal{B}_1([a,b]) \not\subseteq \mathcal{R}([a,b])$. Más adelante veremos que ambas clases están contenidas en el conjunto de las funciones que son Lebesgue integrables.

8.4. Ejercicios Resueltos

Ejercicio 8.4.1. *Sea* $f \in \mathcal{R}([a,b])$. *Pruebe que, para cada* $\varepsilon > 0$, *existe una función en escalera* $\varphi : [a,b] \to \mathbb{R}$ *tal que*

$$\left| \int_a^b (f - \varphi)\, dx \right| < \varepsilon.$$

Prueba. Sea $\varepsilon > 0$. Por el Teorema de Riemann-Darboux, Teorema 8.3.8, sabemos que existe una partición $\mathcal{P} = \{x_0, x_1, \dots, x_n\}$ tal que

$$\left| \sum_{i=1}^{n} f(t_i)(x_i - x_{i-1}) - \int_a^b f\, dx \right| < \varepsilon, \tag{1}$$

para cualquier conjunto arbitrario de etiquetas $\{t_1, \dots, t_n\}$ de $\mathcal{P}$. Fijemos un tal conjunto de etiquetas y defina, para cada $i \in \{1, \dots, n\}$,

$$\varphi_i(x) = \begin{cases} f(t_i) & \text{si } x \in I_i, \\ 0 & \text{si } x \in [a,b] \setminus I_i \end{cases}$$

donde

$$I_1 = [x_0, x_1), \quad I_2 = [x_1, x_2), \quad \cdots, \quad I_n = [x_{n-1}, x_n].$$

Observe que como $\varphi_i = f(t_i)\chi_{I_i}$, resulta que $\varphi = \varphi_1 + \cdots + \varphi_n$ es una función en escalera y, en consecuencia,

$$\int_a^b \varphi\, dx = \int_{x_0}^{x_1} \varphi_1\, dx + \cdots + \int_{x_{n-1}}^{x_n} \varphi_n\, dx = \sum_{i=1}^{n} f(t_i)(x_i - x_{i-1}).$$

De (1) sigue que

$$\left| \int_a^b f\, dx - \int_a^b \varphi\, dx \right| = \left| \int_a^b f\, dx - \sum_{i=1}^{n} f(t_i)(x_i - x_{i-1}) \right| < \varepsilon.$$

La prueba es completa. ■

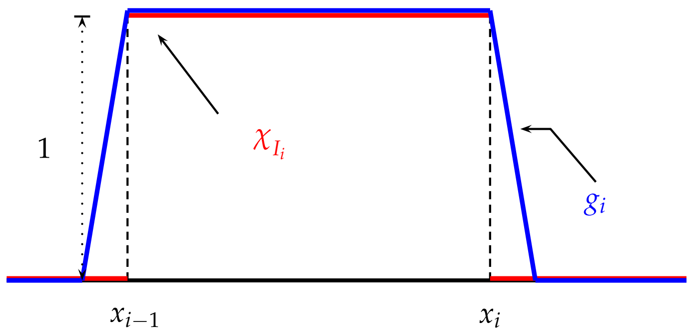
Ejercicio 8.4.2. *Sea* $\varphi : [a,b] \to \mathbb{R}$ *una **función en escalera**. Pruebe que, para cada* $\varepsilon > 0$, *existe una **función continua*** $g : [a,b] \to \mathbb{R}$ *tal que*

$$\left| \int_a^b \varphi\, dx - \int_a^b g\, dx \right| < \varepsilon.$$

Prueba. Sea $\varphi = \sum_{i=1}^{n} c_i \chi_{I_i}$ una función en escalera y sea $\varepsilon > 0$. Construya, para cada entero $i \in \{1, \dots, n\}$, una función continua $g_i : [a,b] \to \mathbb{R}$, como se muestra en la figura (dibujada en azul), tal que

$$\int_a^b |\chi_{I_i} - g_i|\, dx < \frac{\varepsilon}{n(1 + |c_i|)}.$$

(En este caso, la integral representa el área de dos triángulos).



Entonces $g = c_1 g_1 + \cdots + c_n g_n$ es una función continua sobre $[a, b]$ satisfaciendo

$$\left| \int_a^b \varphi - \int_a^b g\, dx \right| \leq \sum_{i=1}^{n} \int_a^b |c_i||\chi_{I_i} - g_i|\, dx < \sum_{i=1}^{n} \frac{|c_i|}{n(1+|c_i|)} \cdot \varepsilon < \varepsilon.$$

Esto termina la prueba. ■

Ejercicio 8.4.3. *Sea* $f \in \mathcal{R}([a, b])$*. Para cada* $\varepsilon > 0$*, existe* $g \in C([a, b])$ *tal que*

$$\left| \int_a^b f\, dx - \int_a^b g\, dx \right| < \varepsilon.$$

Prueba. Sigue de los dos ejercicios anteriores. ■

Ejercicio 8.4.4. *Sean* $f, g \in \mathcal{R}([a, b])$*. Demuestre que las siguientes condiciones son equivalentes:*

(1) $\displaystyle\int_a^x f\, dt = \int_a^x g\, dt$ *para todo* $x \in [a, b]$.

(2) $\displaystyle\int_c^d f\, dt = \int_c^d g\, dt$ *para todo* $c, d \in [a, b]$.

(3) $f(x) = g(x)$ *para todo* $x \in \mathrm{PC}(f) \cap \mathrm{PC}(g)$.

(4) $\{x \in [a, b] : f(x) = g(x)\}$ *es* ***denso*** *en* $[a, b]$.

(5) $\displaystyle\int_a^b |f(t) - g(t)|\, dt = 0$.

Prueba. (1) $\Rightarrow$ (2). Sean $c, d \in [a, b]$ y suponga que $c < d$. Entonces

$$\begin{aligned} \int_a^c f\, dt + \int_c^d f\, dt &= \int_a^d f\, dt = \int_a^d g\, dt \\ &= \int_a^c g\, dt + \int_c^d g\, dt \\ &= \int_a^c f\, dt + \int_c^d g\, dt \end{aligned}$$

de donde se sigue (2).

(1) $\Rightarrow$ (3). Sean F y G definidas por

$$F(x) = \int_a^x f\,dt \qquad \text{y} \qquad G(x) = \int_a^x f\,dt$$

para todo $x \in [a,b]$ y sea $x_0 \in \mathrm{PC}(f) \cap \mathrm{PC}(g)$. Por el Segundo Teorema Fundamental del Cálculo, tenemos que F y G son diferenciables en x_0 con $F'(x_0) = f(x_0)$ y $G'(x_0) = g(x_0)$. Pero por hipótesis, $F(x) = G(x)$ para todo $x \in [a,b]$ de donde se sigue que $F'(x_0) = G'(x_0)$, y así, $f(x_0) = g(x_0)$.

(3) $\Rightarrow$ (4). Como $f, g \in \mathcal{R}([a,b])$, entonces (**TL$_0$**) nos dice que $\mathrm{PC}(f)$ y $\mathrm{PC}(g)$ son ambos G_δ-densos y, así, por el Teorema de Categoría de Baire $\mathrm{PC}(f) \cap \mathrm{PC}(g)$ es denso en $[a,b]$. Finalmente, por (3), se tiene que

$$\mathrm{PC}(f) \cap \mathrm{PC}(g) = \{x \in [a,b] : f(x) = g(x)\}$$

es denso en $[a,b]$.

(4) $\Rightarrow$ (5). Suponga que $D = \{x \in [a,b] : f(x) = g(x)\}$ es denso en $[a,b]$ y pongamos $h = |f - g|$. Sea $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ una partición arbitraria de $[a,b]$. Fijemos $i \in \{1, 2, \ldots, n\}$. Puesto que D es denso en $[a,b]$ resulta que $D \cap (x_{i-1}, x_i) \neq \varnothing$ y, por lo tanto, $h(x) = 0$ para todo $x \in D \cap (x_{i-1}, x_i)$. Esto último combinado con el hecho de que $h \geq 0$, conduce a que

$$m_i = \inf\{h(x) : x \in [x_{i-1}, x_i]\} = 0$$

y, por lo tanto, $\mathsf{L}(h, \mathcal{P}) = 0$. Finalmente, como $h \in \mathcal{R}([a,b])$ se tiene que

$$\int_a^b |f - g|\,dx = \int_a^b h\,dx = \sup\{\mathsf{L}(h, \mathcal{P}) : \mathcal{P} \in \mathbb{P}[a,b]\} = 0.$$

(5) $\Rightarrow$ (1). Suponga que (5) se cumple y sea $x \in [a,b]$. Como $\int_a^b |f - g|\,dx = 0$, entonces $\int_a^x |f - g|\,dx = 0$, y se sigue del Corolario 8.3.16 que $\int_a^x (f - g)\,dx = 0$. Esto termina la prueba.■

Ejercicio 8.4.5. *Sea $f \in \mathcal{R}([a,b])$ tal que $f \geq 0$ sobre $[a,b]$. Pruebe que si $\int_a^b f\,dx = 0$, entonces*

$$D = \{x \in [a,b] : f(x) = 0\}$$

*es **denso** en $[a,b]$.*

Prueba. Si consideremos la función idénticamente nula $g : [a,b] \to \mathbb{R}$, entonces el resultado sigue del Ejercicio 8.4.4 (5) $\Rightarrow$ (4). ■

Ejercicio 8.4.6. *Sea $f \in C([a,b])$ y suponga que $\int_a^b x^n f(x)\,dx = 0$ para todo entero $n \geq 0$. Demuestre que $f = 0$.*

Prueba. Sea $p_n(x) = a_0 + a_1 x + \cdots + a_n x^n$ cualquier polinomio de grado $n \geq 0$ y observe que

$$\int_a^b p_n(x) f(x)\,dx = \sum_{k=0}^{n} a_k \int_a^b x^k f(x)\,dx = 0.$$

Seleccionemos, usando el Teorema de Aproximación de Weierstrass, página 191, una sucesión de polinomios $(p_n)_{n=1}^{\infty}$ tal que $p_n \to f$ uniformemente sobre $[a,b]$. Puesto f es acotada, resulta que $fp_n \to f^2$ uniformemente y, en consecuencia, por el Teorema 8.3.25 se tiene que

$$\int_a^b f^2(x)\,dx \;=\; \lim_{n\to\infty}\int_a^b p_n(x)f(x)\,dx \;=\; 0.$$

Del Teorema 8.3.14 se sigue que $f^2 = 0$ y así, $f = 0$. ■

Ejercicio 8.4.7. *Sea* $f : [a,b] \to \mathbb{R}$ *una función **continua** sobre* $[a,b]$ *. Entonces se cumple que*

$$\lim_{n\to\infty} n\int_a^{a+1/n} f(t)\,dt \;=\; f(a)$$

Prueba. Defina la función $F : [a,b] \to \mathbb{R}$ por

$$F(x) \;=\; \int_a^x f(t)\,dt.$$

Por el Segundo Teorema Fundamental del Cálculo se tiene que $F' = f$ sobre $[a,b]$. En particular, $f(a) = F'(a)$. Más aun, como $F(a) = 0$, se tiene que

$$\begin{aligned} f(a) \;=\; F'(a) \;&=\; \lim_{n\to\infty}\frac{F(a+1/n)-F(a)}{1/n} \\ &=\; \lim_{n\to\infty}\frac{F(a+1/n)}{1/n} \\ &=\; \lim_{n\to\infty} n\int_a^{a+1/n} f(t)\,dt. \end{aligned}$$

Esto termina la prueba. ■

Ejercicio 8.4.8. *Sea* $f : [0,1] \to \mathbb{R}$ *una función **continua** sobre* $[0,1]$ *y suponga que*

$$\int_0^x f\,dt \;=\; \int_x^1 f\,dt, \quad \textit{para todo } x \in [0,1]. \qquad (\star)$$

Pruebe que $f = 0$.

Prueba. Defina $F : [0,1] \to \mathbb{R}$ por

$$F(x) \;=\; \int_0^x f\,dt, \quad \text{para todo } x \in [0,1].$$

Puesto que f es continua, el Segundo Teorema Fundamental del Cálculo nos asegura que F es diferenciable y $F' = f$. Tomando $x = 1$ en la igualdad $(\star)$ se tiene que

$$F(1) \;=\; \int_0^1 f\,dt \;=\; \int_1^1 f\,dt \;=\; 0$$

y, en consecuencia, para cada $x \in [0,1]$,

$$F(x) \;=\; \int_0^x f\,dt \;=\; \int_x^1 f\,dt \;=\; \int_0^1 f\,dt - \int_0^x f\,dt \;=\; -F(x).$$

De aquí se sigue que $F(x) = 0$ y, en consecuencia, $f = 0$. ■

8.5. Problemas

(1) Sea $f \in \mathcal{R}([a,b])$ y suponga que f es continua en algún punto $c \in [a,b]$. Pruebe que si $f(c) > 0$, entonces $\int_a^b f\,dx > 0$.

(2) Sea $f \in \mathcal{R}([a,b])$ y suponga que $f > 0$. Pruebe que, para cada $y \in \left[0, \int_a^b f\,dx\right]$ existe un $x \in [a,b]$ tal que

$$y = \int_a^x f(t)\,dt.$$

(3) Sea $f \in C([0,1])$ y suponga que

$$|f(x)| \le \int_0^x f(t)\,dt \quad \text{para todo} \quad x \in [0,1].$$

Demuestre que $f = 0$.

(4) Sea $f \in C([a,b])$ y suponga que $f(x) > 0$ para todo $x \in [a,b]$. Pruebe que

$$\left(\int_a^b f(x)\,dx\right)\left(\int_a^b \frac{1}{f(x)}\,dx\right) \ge (b-a)^2.$$

(5) Sea $f : [a,b] \to \mathbb{R}$ una función acotada sobre $[a,b]$. Suponga que existe una sucesión $(\mathcal{P}_n)_{n=1}^\infty$ de particiones de $[a,b]$ tal que

$$\lim_{n\to\infty} \Big(\mathsf{U}(f,\mathcal{P}_n) - \mathsf{L}(f,\mathcal{P}_n)\Big) = 0.$$

Pruebe que $f \in \mathcal{R}([a,b])$ y que

$$\int_a^b f\,dx = \lim_{n\to\infty} \mathsf{U}(f,\mathcal{P}_n) = \lim_{n\to\infty} \mathsf{L}(f,\mathcal{P}_n).$$

(6) Para cada $n \in \mathbb{N}$, suponga que V_n es una unión finita de subintervalos no-superpuestos de $[a,b]$, digamos $J_{n1}, \ldots, J_{nk_n}$, tal que

$$\ell(V_n) = \ell(J_{n1}) + \cdots + \ell(J_{nk_n}).$$

Pruebe que si para algún $\delta > 0$, $\ell(V_n) > \delta$ para todo $n \ge 1$, entonces existe un punto $z \in [a,b]$ tal que $z \in V_n$ para infinitos n's.

(7) Pruebe que si $f, g \in \mathcal{R}([a,b])$ y $f = g$ $\mu -$c.s., entonces

$$\int_a^b f\,dx = \int_a^b g\,dx.$$

CAPÍTULO 9

Diferenciación y un Teorema de Lebesgue

En los cursos de Cálculo uno aprende que si una función f es derivable en un punto x, entonces f es continua en dicho punto. También se aprende de la existencia de funciones continuas que no son derivables en ciertos puntos. Sin embargo, en el siglo XIX, muchos matemáticos de la época tenían la firme convicción de que toda *función continua* tenían derivadas en un número significativo de puntos. Lo que comienza a ser sorpresivo y puesto en duda por algunos matemáticos de la época es la existencia de funciones continuas que no son derivables en ninguno punto de su dominio tal como lo demostró Weierstrass y entonces uno se pregunta: ¿qué funciones poseen la particularidad de que ellas son diferenciables casi-siempre? La respuesta, sin dilación, es: *las funciones de "variación de acotada"*, o de modo más general, *las funciones monótonas*. La demostración de este hecho será dado un poco más abajo y se le conoce con el nombre de Teorema de Diferenciabilidad de Lebesgue, un resultado considerado de máxima importancia pues, hace más de un siglo, la existencia de funciones extrañas y con perfiles polémicos estaba minando el entusiasmo que hasta ese momento disfrutaba el Análisis Real. Es el Teorema de Diferenciabilidad de Lebesgue quien logra restablecer la confianza en el Análisis.

9.1. Funciones Absolutamente Continuas y de Variación Acotada

En esta sección estudiaremos las funciones absolutamente continuas y de variación acotada definidas sobre un intervalo $[a, b]$. El resultado fundamental que demostraremos en relación con las funciones de variación acotada es que cualquiera de ella se puede expresar como diferencia de dos funciones monótonas crecientes. Posteriormente se introducen los cubrimientos de Vitali que constituyen una herramienta clave para demostrar el Teorema de Diferenciación de Lebesgue. Las funciones absolutamente continuas deben su importancia debido la hecho de que ellas son las que garantizan la validez del Teorema Fundamental del Cálculo para la integral de Lebesgue.

9.1.1. Funciones Lipschitz y la condición $(\mathbf{N})$ de Lusin

Las funciones Lipschitz son una subclase muy particular de las funciones continuas que tienen propiedades realmente sorprendentes: por ejemplo, ellas poseen el peculiar encanto de preservar conjuntos de medida cero, es decir, transforman conjuntos de medida cero en conjuntos de medida cero. Algunas de las propiedades importantes que tales funciones poseen y que nos son de utilidad serán presentadas fundamentalmente en el transcurso de esta sección.

Definición 9.1.1. *Una función $f : [a,b] \to \mathbb{R}$ es llamada **M-Lipschitz** sobre $[a,b]$, si existe una constante $M > 0$ tal que*

$$|f(x) - f(y)| \ \leq \ M\,|x - y| \tag{1}$$

*para todo $x, y \in [a,b]$. La más pequeña de las constantes M que satisfacen (1) se llama la **constante de Lipschitz** de f.*

Por supuesto, no hay nada de espacial en tener al intervalo $[a,b]$ como dominio de f. En general, si (X_1, d) y (X_2, ϱ) son espacios métricos, entonces una función $f : X_1 \to X_2$ se llama M-Lipschitz si

$$\varrho(f(x), f(y)) \ \leq \ M\,d(x,y)$$

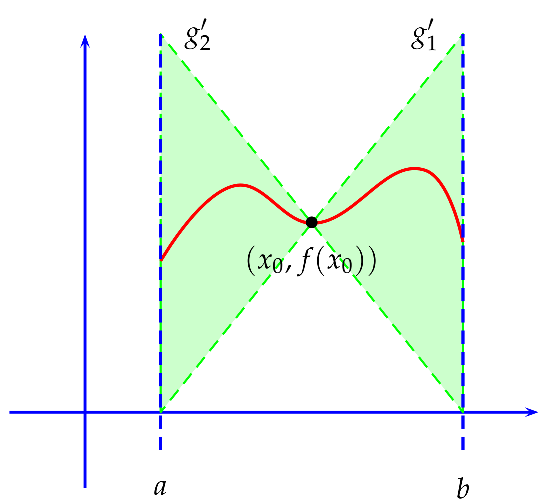
para todo $x, y \in X_1$ y alguna constante $M > 0$. En ocasiones, escribiremos "f es Lipschitz", en lugar de "f es M-Lipschitz". También es importante observar que existen funciones Lipschitz no triviales. Por ejemplo, para cada $x_0 \in [a,b]$, la función $f(x) = |x - x_0|$ para todo $x \in [a,b]$ es 1-Lipschitz. Las funciones Lipschitz tienen una interpretación geométrica muy simple: sea $f : [a,b] \to \mathbb{R}$ una función M-Lipschitz y considere las rectas

$$g_1(x) \ = \ Mx \qquad \text{y} \qquad g_2(x) \ = \ -Mx.$$

Si $(x, f(x_0))$ es un punto cualquiera del gráfico de f, entonces la traslación de las dos rectas anteriores pasando por el punto $(x_0, f(x_0))$:

$$\ell_1(x) \ = \ M(x - x_0) + f(x_0) \qquad \text{y} \qquad \ell_2(x) \ = \ -M(x - x_0) + f(x_0)$$

forman un doble cono (tal como se muestra en verde en la figura adjunta) de modo tal que el gráfico de f siempre permanece completamente dentro de dichos conos.



Denote por $\mathrm{Lip}([a,b])$ *el conjunto de todas las funciones* $f : [a,b] \to \mathbb{R}$ *que son* **M-*Lipschitz*** *sobre* $[a,b]$. Observe que toda función Lipschitz $f : [a,b] \to \mathbb{R}$ es *uniformemente continua*. En particular, como f es continua y $[a,b]$ es compacto, resulta que f es acotada y se cumple que

$$\mathrm{Lip}([a,b]) \subseteq C([a,b]) \subseteq \mathcal{R}([a,b]) \subseteq \mathcal{B}_\infty([a,b]) \cap \mathcal{F}_\mu([a,b]).$$

De esta relación se sigue que *cualquier función* $f : [a,b] \to \mathbb{R}$ *que no es acotada no puede ser Lipschitz*. Es un ejercicio sencillo demostrar que $\mathrm{Lip}([a,b])$ es un espacio vectorial sobre $\mathbb{R}$. Más aun, si $f, g \in \mathrm{Lip}([a,b])$, entonces $|f|, f \cdot g \in \mathrm{Lip}([a,b])$.

Por supuesto, existen funciones uniformemente continuas que no son Lipschitz. Por ejemplo, la función $f(x) = \sqrt[3]{x}$ sobre $[-1,1]$ es uniformemente continua pero no es Lipschitz. Para ver esto último, para cada $n \in \mathbb{N}$, tome $x_n = -1/n^3$ y $y_n = 1/n^3$. Entonces

$$\frac{f(x_n) - f(y_n)}{x_n - y_n} = \frac{-1/n - 1/n}{-1/n^3 - 1/n^3} = \frac{-2/n}{-2/n^3} = n^2.$$

Como no existe M para el cual $n^2 < M$ para todo n, resulta que f no puede ser Lipschitz.

Un aspecto importante en la Teoría de la Medida es investigar qué funciones tienen la propiedad de preservar conjuntos de medida cero. Tales funciones reciben un nombre muy especial.

Definición 9.1.2. *Sea* X *un conjunto medible. Una función* $f : X \to \mathbb{R}$ *se dice que satisface la* ***condición* (N) *de Lusin***, *si* $\mu^*\big(f(E)\big) = 0$ *para cualquier conjunto medible* $E \subseteq X$ *con* $\mu(E) = 0$.

Para funciones continuas, el siguiente resultado caracteriza a las funciones que satisfacen la condición (N) de Lusin.

Teorema 9.1.3 (Rademacher). *Sea* $f : [a,b] \to \mathbb{R}$ *una* ***función continua*** *sobre* $[a,b]$. *Las siguientes condiciones son equivalentes:*

(1) f *satisface* ***la condición* (N) *de Lusin***.

(2) $f(E) \in \mathfrak{M}_\mu(\mathbb{R})$ *para cualquier* ***conjunto medible*** $E \subseteq [a,b]$.

Prueba. (1) $\Rightarrow$ (2). Suponga que (1) se cumple y sea $E \subseteq [a,b]$ medible. Por el Teorema 6.3.55, página 293, existe un conjunto de tipo $\mathcal{F}_\sigma$, digamos $F = \bigcup_{n=1}^\infty F_n$, donde cada F_n es un conjunto cerrado, tal que

$$F \subseteq E \qquad \text{y} \qquad \mu(E \setminus F) = 0.$$

Como cada $F_n \subseteq [a,b]$, resulta que él es compacto y así, por continuidad, $f(F_n)$ también es compacto. Por esto, $f(F_n) \in \mathfrak{M}_\mu(\mathbb{R})$ y, por lo tanto,

$$f(F) = f\bigg(\bigcup_{n=1}^\infty F_n\bigg) = \bigcup_{n=1}^\infty f(F_n) \in \mathfrak{M}_\mu(\mathbb{R}).$$

Puesto que $\mu(E \setminus F) = 0$ se tiene, por nuestra hipótesis, que $\mu^*\big(f(E \setminus F)\big) = 0$ y, en consecuencia, $f(E \setminus F) \in \mathfrak{M}_\mu(\mathbb{R})$. Finalmente, como $E = (E \setminus F) \cup F$, entonces $f(E) = f(E \setminus F) \cup f(F) \in \mathfrak{M}_\mu(\mathbb{R})$.

(2) $\Rightarrow$ (1). Suponga que (2) se cumple y sea E un subconjunto medible de $[a,b]$ con $\mu(E) = 0$. Para construir una contradicción, suponga que $\mu\big(f(E)\big) > 0$. Seleccionemos, invocando el

Teorema 6.5.3, un conjunto no-medible $V \subseteq f(E)$. Como $\mu(E) = 0$ y $E \cap f^{-1}(V) \subseteq E$, resulta que $E \cap f^{-1}(V) \in \mathfrak{M}_\mu(\mathbb{R})$ y, por consiguiente, usando nuestra hipótesis, vemos que

$$V = f\big(E \cap f^{-1}(V)\big) \in \mathfrak{M}_\mu(\mathbb{R}).$$

Esta contradicción establece que $\mu\big(f(E)\big) = 0$ y termina la prueba. ■

Ya hemos visto que una función continua no necesita transformar conjuntos medibles en conjuntos medibles. En efecto, en la Sección 6.7, página 354, Ejercicio 6.7.1 (c), demostramos que la función de Cantor era una ejemplo de tal función. Otra manera de ver esto es usando el resultado de Rademacher.

Corolario 9.1.4. *Existe una función* ***uniformemente continua*** $f : [0,1] \to \mathbb{R}$ *y al menos un conjunto medible* $E \subseteq [0,1]$ *tal que* $f(E)$ ***no es medible según Lebesgue****.*

Prueba. Sea $f = \varphi_\Gamma$ la función de Cantor. Si fuese cierto que φ_Γ transforma conjuntos medibles en conjuntos medibles, entonces el Teorema de Rademacher nos diría que $\varphi_\Gamma(E)$ satisface la condición (N) de Lusin. En particular, como $\mu(\Gamma) = 0$, tendríamos que $\mu(\varphi_\Gamma(\Gamma)) = 0$ lo cual es absurdo ya que $\varphi_\Gamma(\Gamma) = [0,1]$. Fin de la prueba. ■

En la parte (2) del Teorema de Rademacher podemos restringirnos a subconjuntos compactos siempre que la función f satisfaga cierta propiedad adicional.

Teorema 9.1.5 (Foran). *Sea* $f : [a,b] \to \mathbb{R}$ *una* ***función continua*** *y* ***estrictamente creciente****. Son equivalentes:*

(1) f *satisface* ***la condición* (N) *de Lusin****.*

(2) $\mu(f(K)) = 0$ *para cada* ***conjunto compacto*** $K \subseteq [a,b]$ *con* $\mu(K) = 0$.

Prueba. (1) $\Rightarrow$ (2) es inmediata. Para demostrar la otra implicación, suponga que (2) se cumple y sea E un subconjunto medible de $[a,b]$ con $\mu(E) = 0$. Observe que como f *es continua y estrictamente creciente, entonces ella es un homeomorfismo sobre su rango. Por esto, todo conjunto compacto* $K' \subseteq f(E)$ *es de la forma* $K' = f(K)$ *para algún conjunto compacto* $K \subseteq E$. En efecto, como f es continua, $K = f^{-1}(K')$ es compacto, y así, usando el hecho de que $f : [a,b] \to [f(a), f(b)]$ es un homeomorfismo, tenemos que $K' = f(K)$. Se sigue de la regularidad de la medida de Lebesgue, Corolario 6.3.66, página 302 y de nuestra hipótesis, que

$$\begin{aligned} \mu(f(E)) &= \sup\{\mu(K') : K' \subseteq f(E),\ K' \text{ compacto}\} \\ &= \sup\{\mu(f(K)) : K \subseteq E,\ K \text{ compacto}\} \\ &= 0. \end{aligned}$$

Esto termina la prueba. ■

Observe que, en realidad, uno puede sustituir conjunto compacto por conjunto cerrado en la parte (2) del Teorema de Foran. El siguiente resultado establece, entre otras cosas, que toda función Lipschitz satisface la condición (N) de Lusin.

Teorema 9.1.6. *Sea $f : [a,b] \to \mathbb{R}$ una función arbitraria.*

*(a) Si f es **diferenciable** sobre $[a,b]$ y f' es **acotada** sobre $[a,b]$, entonces f es **Lipschitz**. En particular, si f es **continuamente diferenciable** sobre $[a,b]$, entonces f es **Lipschitz**.*

*(b) Si f es **Lipschitz** con constante de Lipschitz M, entonces*

$$\mu^*(f([\alpha,\beta])) \leq M(\beta - \alpha)$$

para cualquier intervalo $[\alpha,\beta] \subseteq [a,b]$.

*(c) Si f es **Lipschitz**, entonces f satisface **la condición $(\mathbf{N})$ de Lusin**.*

Prueba. (a) Como f' es acotada, existe una constante $M > 0$ tal que $|f'(x)| \leq M$ para todo $x \in [a,b]$. Sean ahora $x, y \in [a,b]$ con $x \neq y$. Sin perder generalidad, asumiremos que $x < y$. Por el Teorema del Valor Medio, existe un $\xi \in (x,y)$ tal que

$$f'(\xi) \;=\; \frac{f(y) - f(x)}{y - x}.$$

Por esto,

$$\left|\frac{f(y) - f(x)}{y - x}\right| \;=\; |f'(\xi)| \;\leq\; M$$

y, así, $|f(x) - f(y)| \;\leq\; M\,|x - y|$. Como la elección de $x, y \in [a,b]$ fue arbitraria, se concluye que f es Lipschitz.

Suponga ahora que f es continuamente diferenciable en $[a,b]$. Entonces f' es continua en $[a,b]$ y, en consecuencia, acotada sobre $[a,b]$. El hecho de que f es Lipschitz sigue de la primera parte.

(b) Puesto que f es continua en el compacto $[a,b]$, ella alcanza su máximo y su mínimo. Sean entonces α y β en $[a,b]$ tal que $f(\alpha) \leq f(x) \leq f(\beta)$ para todo $x \in [a,b]$. Para simplificar la presentación, asumiremos que $\alpha < \beta$. Entonces $f([\alpha,\beta]) \subseteq [f(\alpha), f(\beta)]$ y se tiene que

$$\mu^*\big(f([\alpha,\beta])\big) \;\leq\; f(\beta) - f(\alpha) \;\leq\; M(\beta - \alpha).$$

(c) Suponga que f es Lipschitz con constante de Lipschitz M y sea $\varepsilon > 0$. Si $E \subseteq [a,b]$ con $\mu(E) = 0$, entonces podemos elegir un conjunto abierto $G \supseteq E$ tal que $\mu(G) < \varepsilon/M$. Escribamos a G como una unión numerable y disjunta de intervalos abiertos, digamos $G = \bigcup_{n=1}^{\infty}(a_n, b_n)$. Entonces, $\mu(G) = \sum_{n=1}^{\infty}(b_n - a_n)$ y, como además,

$$f(E) \;\subseteq\; f(G) \;=\; f\Big(\bigcup_{n=1}^{\infty}(a_n, b_n)\Big) \;=\; \bigcup_{n=1}^{\infty} f\big((a_n, b_n)\big) \;\subseteq\; \bigcup_{n=1}^{\infty} f\big([a_n, b_n]\big).$$

resulta entonces de la subaditividad numerable de μ^* y de (b) que

$$\begin{aligned}
\mu^*(f(E)) \;&\leq\; \sum_{n=1}^{\infty} \mu^*\big(f([a_n, b_n)]\big) \\
&\leq\; \sum_{n=1}^{\infty} M \cdot (b_n - a_n) \\
&\leq\; K \cdot \mu(G) \;<\; \varepsilon.
\end{aligned}$$

Como la elección de $\varepsilon > 0$ se hizo de modo arbitrario, se tiene que $\mu^*(f(E)) = 0$. En particular, $f(E) \in \mathfrak{M}_\mu(\mathbb{R})$ y $\mu(f(E)) = 0$. ■

Es interesante destacar que si bien es cierto que toda función Lipschitz (la cual es continua) satisface la condición (N) de Lusin, existen funciones continuas, de hecho *homeomorfismos*, que no satisfacen tal propiedad (véase, por ejemplo, Ejercicio 6.7.1 (b), página 354). Por otro lado, las funciones diferenciables que son acotadas son indistinguibles de las funciones Lipschitz como se muestra a continuación.

Teorema 9.1.7. *Sea $f : [a, b] \to \mathbb{R}$ una* ***función diferenciable*** *sobre $[a, b]$. Las siguientes condiciones son equivalentes:*

(1) $f \in \mathrm{Lip}([a, b])$.

(2) f' *es* ***acotada*** *sobre $[a, b]$.*

Prueba. (1) $\Rightarrow$ (2). Suponga que (1) se cumple. Puesto que $f \in \mathrm{Lip}([a, b])$, existe una constante $M > 0$ para la cual se cumple que $|f(x) - f(y)| \leq M|x - y|$ cualesquiera sean $x, y \in [a, b]$. Por otro lado, por ser f diferenciable en $[a, b]$ tenemos que, para cada $y \in [a, b]$,

$$|f'(y)| = \lim_{\substack{x \to y \\ x \neq y}} \frac{|f(x) - f(y)|}{|x - y|} \leq \lim_{\substack{x \to y \\ x \neq y}} \frac{M \cdot |x - y|}{|x - y|} = M$$

y así, f' es acotada sobre $[a, b]$.

(2) $\Rightarrow$ (1). Sigue Teorema 9.1.6 (a). ■

Del Teorema del Valor Medio para Derivadas se deduce que si $F : [a, b] \to \mathbb{R}$ es una función diferenciable sobre $[a, b]$ y $F'(x) = 0$ para todo $x \in [a, b]$, entonces F es constante sobre $[a, b]$. Sin embargo, existen funciones que no son constantes pero cuya derivada es cero casi-siempre. Una tal ejemplo es la función de Cantor. Esto permite justificar la siguiente definición.

Definición 9.1.8. *Una función continua $f : [a, b] \to \mathbb{R}$ con la propiedad de que $f' = 0$ casi-siempre sobre $[a, b]$ es llamada una* ***función singular*** *sobre $[a, b]$.*

El siguiente resultado asegura que si f es una función Lipschitz y, además, singular, entonces ella es constante. Este hecho será utilizado para dar una demostración de un Teorema Fundamental del Cálculo más general que el Primer Teorema Fundamental del Cálculo, aunque también se puede usar para dar otra demostración del Teorema de Diferenciabilidad de Lebesgue.

Teorema 9.1.9. *Si $f : [a, b] \to \mathbb{R}$ una función* ***Lipschitz*** *y* ***singular*** *sobre $[a, b]$, entonces f es constante sobre $[a, b]$.*

Prueba. Sea $\varepsilon > 0$. Teniendo en cuenta que f es Lipschitz, existe una constante $M > 0$ tal que

$$|f(x) - f(y)| \leq M|x - y| \quad \text{para todo } x, y \in [a, b].$$

Sea $E = \{x \in [a, b] : f'(x) = 0\}$. Puesto que $\mu([a, b] \setminus E) = 0$, podemos seleccionar una sucesión $(I_n)_{n=1}^{\infty}$ de intervalos abiertos tal que

$$[a, b] \setminus E \subseteq \bigcup_{n=1}^{\infty} I_n \qquad \text{y} \qquad \sum_{n=1}^{\infty} \ell(I_n) < \frac{\varepsilon}{2M}. \tag{1*}$$

Considere ahora las colecciones $\mathcal{F}_1$ y $\mathcal{F}_2$ de intervalos cerrados $[c,d] \subseteq [a,b]$ que satisfacen las siguientes condiciones:

(α) $[c,d] \in \mathcal{F}_1$ si, y sólo si, $|f(d) - f(c)| < \frac{\varepsilon}{2(b-a)}(d-c)$, y

(β) $[c,d] \in \mathcal{F}_2$ si, y sólo si, $[c,d] \subseteq I_n$, para algún n.

Sea

$$S = \Big\{ x \in (a,b] \, : \, \text{existe una partición } \mathcal{P}_x = \{x_0, x_1, \ldots, x_n = x\} \text{ de } [a,x] \\ \text{tal que } [x_{k-1}, x_k] \in \mathcal{F}_1 \cup \mathcal{F}_2, \;\; k = 1, 2, \ldots, n \Big\}$$

Afirmamos que $S \neq \varnothing$. En efecto,

(i) Si $a \in E$, entonces

$$0 = f'(a) = \lim_{x \to a^+} \frac{f(x) - f(a)}{x - a}$$

y, por lo tanto, existe un $\rho > 0$ tal que

$$\left| \frac{f(x) - f(a)}{x - a} \right| < \frac{\varepsilon}{2(b-a)} \qquad \text{para todo} \quad x \in (a, a+\rho).$$

Si ahora fijamos un $x \in (a, a+\rho)$, vemos que

$$|f(x) - f(a)| < \frac{\varepsilon}{2(b-a)}(x-a)$$

lo cual significa que $[a,x] \in \mathcal{F}_1$ y, así, $x \in S$.

(ii) Si $a \in [a,b] \setminus E$, entonces de (1^*) se sigue que existe un $n \in \mathbb{N}$ tal que $x \in I_n$ y como I_n es un intervalo abierto, existe un $x > a$ tal que $[a,x] \subseteq I_n$ y, de nuevo, $x \in S$. Nuestro siguiente paso es demostrar que

$$|f(x) - f(a)| \leq \varepsilon \tag{2^*}$$

para todo $x \in S$. Para ver esto, fijemos un $x \in S$ y sea $\mathcal{P}_x = \{x_0, x_1, \ldots, x_n = x\}$ una partición de $[a,x]$ tal que $[x_{k-1}, x_k] \in \mathcal{F}_1 \cup \mathcal{F}_2$, para $k = 1, 2, \ldots, n$ y sean

$$J_1 = \{k \, : \, [x_{k-1}, x_k] \in \mathcal{F}_1\} \qquad \text{y} \qquad J_2 = \{k \, : \, [x_{k-1}, x_k] \in \mathcal{F}_2\}.$$

Puesto que

$$f(x) - f(a) = \sum_{k=1}^{n} f(x_k) - f(x_{k-1}),$$

resulta que

$$\begin{aligned}
|f(x)-f(a)| &\leq \sum_{k=1}^{n} |f(x_k)-f(x_{k-1})| \\
&= \sum_{k\in J_1} |f(x_k)-f(x_{k-1})| + \sum_{k\in J_2} |f(x_k)-f(x_{k-1})| \\
&< \sum_{k\in J_1} \frac{\varepsilon}{2(b-a)}(x_k-x_{k-1}) + \sum_{k\in J_2} M(x_k-x_{k-1}) \\
&\leq \frac{\varepsilon}{2(b-a)}(b-a) + M\sum_{n=1}^{\infty} \ell(I_n) \\
&< \frac{\varepsilon}{2} + M\frac{\varepsilon}{2M} = \varepsilon.
\end{aligned}$$

Sea $c = \sup S$. Entonces $a < c \leq b$ y así, gracias a la continuidad de f, se puede concluir que $|f(c)-f(a)| \leq \varepsilon$. En efecto, como $c = \sup S$, existe una sucesión $(x_n)_{n=1}^{\infty}$ en S tal que $x_n \to c$. Ahora, por (2^*) y la continuidad de f, tenemos que

$$|f(c)-f(a)| = \lim_{n\to\infty} |f(x_n)-f(a)| \leq \varepsilon.$$

Veamos que $c = b$. Para construir una contradicción, suponga que $c < b$.

(a) Si $c \in E$, entonces $f'(c) = 0$ y, como en (i), seleccionemos un número $\rho > 0$ tal que $[c-\rho, c+\rho] \subseteq (a,b)$ y

$$|f(x)-f(c)| < \frac{\varepsilon}{2(b-a)}|x-c| \qquad \text{si} \quad x \in [c-\rho, c+\rho]. \tag{3^*}$$

Puesto que $c-\rho$ no es cota superior de S, podemos escoger un $x \in S$ tal que $c-\rho < x \leq c$. Pero como $x \in S$, existe una partición $\mathcal{P}_x = \{x_0, x_1, \ldots, x_n\}$ de $[a,x]$ tal que $[x_{k-1}, x_k] \in \mathcal{F}_1 \cup \mathcal{F}_2$, para $k = 1,2,\ldots,n$. Claramente $\mathcal{P} = \{x_0, x_1, \ldots, x_n, x_{n+1}\}$ es una partición de $[a, c+\rho]$ donde $x_{n+1} = c+\rho$. Además, por (3^*) sabemos que $[x, c+\rho] \in \mathcal{F}_1$ por lo que $[x_{k-1}, x_k] \in \mathcal{F}_1 \cup \mathcal{F}_2$, para $k = 1,2,\ldots,n+1$. Esto prueba que $c+\rho \in S$, lo cual es imposible.

(b) Si $c \in [a,b] \setminus E$, entonces existe un $n \geq 1$ tal que $c \in I_n$ y como I_n es abierto, existe un $\rho > 0$ tal que $[c-\rho, c+\rho] \subseteq I_n$. De nuevo, existe un $x \in S$ tal que $c-\rho < x \leq c$. Un argumento similar al caso anterior conduce a que $c+\rho \in S$, lo que de nuevo genera una contradicción. Por esto $c = b$ y

$$|f(b)-f(a)| \leq \varepsilon.$$

Puesto que ε es un número positivo arbitrario, se concluye que $f(a) = f(b)$ y claramente $f(x) = f(a)$ para todo $x \in [a,b]$ lo cual finaliza la prueba. ■

Del resultado anterior se sigue que:

Corolario 9.1.10. *Existe una función $f : [0,1] \to \mathbb{R}$ **uniformemente continua** que **no es Lipschitz** sobre $[0,1]$.*

Prueba. La función de Cantor es no-constante, uniformemente continua, pero no es Lipschitz ya que, gracias al Corolario 5.1.10, página 236, ella es singular. ∎

Ya hemos visto, Teorema 8.3.21 y Corolario 8.3.23, que si $f \in \mathcal{R}([a,b])$, entonces su integral indefinida $F(x) = \int_a^x f\,dt$ para cada $x \in [a,b]$, es una función Lipschitz que satisface $F' = f$ $\mu -$ c.s. El siguiente resultado incluye el Primer Teorema Fundamental del Cálculo pues toda función que tenga una derivada acotada es, por el Teorema 9.1.6 (a), Lipschitz.

Teorema 9.1.11 (Teorema Fundamental del Cálculo). *Sea $f \in \mathcal{R}([a,b])$. Si $F : [a,b] \to \mathbb{R}$ es una función **Lipschitz** tal que $F' = f$ casi-siempre sobre $[a,b]$, entonces*

$$\int_a^b f\,dx \;=\; F(b) \;-\; F(a).$$

Prueba. Sea $G : [a,b] \to \mathbb{R}$ definida por

$$G(x) \;=\; \int_a^x f\,dt \quad \text{para cada } x \in [a,b].$$

Por el Teorema 8.3.21 y el Corolario 8.3.23 sabemos que G es Lipschitz y $G' = f$ $\mu -$ c.s. Esto y nuestra hipótesis nos muestran que

$$(G-F)'(x) \;=\; G'(x) - F'(x) \;=\; f(x) - f(x) \;=\; 0 \quad \mu - \text{c.s.}$$

Además, como $F' = f$ es acotada, resulta del Teorema 9.1.6 que F también es Lipschitz, por lo que $G - F$ es Lipschitz. Un llamado al Teorema 9.1.9 nos garantiza la existencia de una constante C tal que $G(x) = F(x) + C$ para todo $x \in [a,b]$. Observe que cuando $x = a$ se obtiene que $0 = F(a) + C$, es decir, $C = -F(a)$ y para $x = b$,

$$\int_a^b f\,dt \;=\; G(b) \;=\; F(b) - F(a).$$

Esto termina la prueba. ∎

9.1.2. Funciones de Variación Acotada

En un libro escrito en 1878, U. Dini demuestra que si f es una función que posee derivadas de Dini (véase la Definición 9.1.37) acotadas, entonces f se puede escribir como diferencia de dos funciones monótona crecientes. Tres años después, Camille Jordan encontró una caracterización que es equivalente a la de las funciones que pueden ser representadas como diferencias de dos funciones monótona crecientes y a las que él llamó **funciones de variación acotada**. Las funciones de variación acotada definidas sobre un intervalo cerrado y acotado $[a,b]$, constituyen un subconjunto muy especial de las funciones acotadas que son Riemann integrables. Su estudio es fundamental en el desarrollo del Teorema Fundamental del Cálculo para la integral de Lebesgue, ya que toda función de variación acotada es diferenciable casi-siempre, así como el cálculo de longitudes de arco y su relación con la integral de Riemann-Stieltjes.

Definición 9.1.12. *Sean $f : [a,b] \to \mathbb{R}$ una función y $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ una partición de $[a,b]$. La **variación** de f **sobre** $\mathcal{P}$ se define como*

$$V(f,\mathcal{P}) \;=\; \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|$$

*Diremos que f es de **variación acotada** sobre $[a,b]$ si*

$$V(f,[a,b]) \;=\; \sup\big\{V(f,\mathcal{P}) \;:\; \mathcal{P} \in \mathbb{P}[a,b]\big\} \;<\; +\infty.$$

Si $f : [a,b] \to \mathbb{R}$ no es de variación acotada en $[a,b]$ escribiremos $V(f,[a,b]) = +\infty$. En lo que sigue, el símbolo $\mathrm{BV}([a,b])$ será usado para denotar *el conjunto de todas las funciones $f : [a,b] \to \mathbb{R}$ que son de **variación acotada** sobre $[a,b]$.*

Lo primero que debemos destacar es que

Teorema 9.1.13. *Toda **función** de **variación acotada** es **acotada**, esto es,*

$$\mathrm{BV}([a,b]) \;\subseteq\; \mathcal{B}_\infty([a,b]).$$

Prueba. Sea $f \in \mathrm{BV}([a,b])$ y veamos que existe una constante $M > 0$ tal que $|f(x)| \leq M$ para todo $x \in [a,b]$. Fijemos entonces un $x \in [a,b]$ elegido de modo arbitrario y considere la siguiente partición $\mathcal{P} = \{a, x, b\}$ de $[a,b]$. Puesto f es de variación acotada, tenemos que

$$|f(x) - f(a)| \;\leq\; |f(x) - f(a)| + |f(b) - f(x)| \;\leq\; V(f,[a,b]).$$

De esto se sigue que

$$|f(x)| \;\leq\; |f(a)| + |f(x) - f(a)| \;\leq\; |f(a)| + V(f,[a,b]),$$

de modo que si tomamos $M = |f(a)| + V(f,[a,b])$, el cual es un número que no depende de x, tendremos que f es acotada. ■

De la prueba del resultado anterior se obtiene, en particular, que

$$\|f\|_\infty \;\leq\; |f(a)| + V(f,[a,b]).$$

Es interesante también observar que

$$\mathrm{BV}([a,b]) \;=\; \bigcup_{n=1}^{\infty} \mathrm{BV}([a,b],n),$$

donde $\mathrm{BV}([a,b],n) = \big\{f \in \mathrm{BV}([a,b]) : V(f,[a,b]) \leq n\big\}$ para cada $n \in \mathbb{N}$. El siguiente es un ejemplo de una función acotada que no es de variación acotada.

Ejemplo 9.1.1. *La función de Dirichlet $\chi_{\mathbb{Q}} : [0,1] \to \mathbb{R}$ es **acotada** sobre $[0,1]$, pero no es de **variación acotada** sobre $[0,1]$.*

Prueba. Suponga, por un momento, que $\chi_{\mathbb{Q}}$ es de variación acotada y seleccione un entero positivo $n > V(\chi_{\mathbb{Q}}, [0,1])$. Para este n, considere la partición $\mathcal{P} = \{x_0, x_1, \ldots, x_{n+2}\}$ de $[0,1]$, donde cada x_j con índice impar es un número irracional y los de índices pares son números racionales. De esto se sigue que

$$\begin{aligned} n \;>\; V(\chi_{\mathbb{Q}}, [0,1]) \;&\geq\; \sum_{j=1}^{n+2} |\chi_{\mathbb{Q}}(x_j) - \chi_{\mathbb{Q}}(x_{j-1})| \\ &\geq\; \sum_{j=2}^{n+1} |\chi_{\mathbb{Q}}(x_j) - \chi_{\mathbb{Q}}(x_{j-1})| \\ &=\; |1-0| \,+\, |0-1| \,+\, \cdots \,+\, |1-0| \\ &=\; n \end{aligned}$$

lo cual es imposible. Por esto, $\chi_{\mathbb{Q}}$ no es de variación acotada. ■

Es un ejercicio sencillo establecer que $\mathrm{BV}([a,b])$ es un *espacio vectorial sobre* $\mathbb{R}$, esto es, si $f, g \in \mathrm{BV}([a,b])$ entonces

(a) $f + g \in \mathrm{BV}([a,b])$, y

(b) $\alpha f \in \mathrm{BV}([a,b])$ para cualquier $\alpha \in \mathbb{R}$.

Más aun

(c) $fg \in \mathrm{BV}([a,b])$.

Ejemplo 9.1.2. *Si $f : [a,b] \to \mathbb{R}$ es **constante** sobre $[a,b]$, entonces $f \in \mathrm{BV}([a,b])$ y $V(f, [a,b]) = 0$.*

Prueba. Esto sigue del hecho de que f es acotada.

Ejemplo 9.1.3. *Si $f : [a,b] \to \mathbb{R}$ es una **función monótona** sobre $[a,b]$, entonces $f \in \mathrm{BV}([a,b])$ y*

$$V(f, [a,b]) \;=\; |f(b) - f(a)|.$$

Prueba. Suponga que f es creciente y sea $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ una partición de $[a,b]$. Entonces

$$V(f, \mathcal{P}) \;=\; \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| \;=\; \sum_{i=1}^{n} \big(f(x_i) - f(x_{i-1})\big) \;=\; f(b) - f(a)$$

y, en consecuencia, $V(f, [a,b]) \;=\; f(b) - f(a)$. Por supuesto, si f es monótona decreciente, entonces $V(f, [a,b]) = f(a) - f(b)$ y termina la prueba. ■

Corolario 9.1.14. *Si $f : [a,b] \to \mathbb{R}$ es **Lipschitz**, entonces $f \in \mathrm{BV}([a,b])$.*

Prueba. Sea $M > 0$ la constante de Lipschitz para f. Si $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ es una partición de $[a,b]$, entonces

$$V(f, \mathcal{P}) \;=\; \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| \;\leq\; \sum_{i=1}^{n} M(x_{i-1} - x_i) \;=\; M(b-a)$$

y, en consecuencia, $V(f, [a,b]) \leq M(b-a) < +\infty$. ■

Corolario 9.1.15. *Si $f \in \mathcal{R}([a,b])$ y F es su* ***integral indefinida****, esto es, $F(x) = \int_a^x f\,dt$ para cada $x \in [a,b]$, entonces $F \in \mathrm{BV}([a,b])$.*

Prueba. Por el Segundo Teorema Fundamental del Cálculo, Teorema 8.3.21, sabemos que F es Lipschitz y el resultado sigue del corolario anterior. ■

La continuidad de una función $f : [a,b] \to \mathbb{R}$ no garantiza, en general, que ella sea de variación acotada. En el siguiente resultado se establecen ciertas condiciones bajo la cual una función continua es de variación acotada.

Corolario 9.1.16. *Sea $f : [a,b] \to \mathbb{R}$ una* ***función continua*** *sobre $[a,b]$. Si f es* ***diferenciable*** *y f' es* ***acotada*** *sobre $[a,b]$, entonces $f \in \mathrm{BV}([a,b])$.*

Prueba. Esto es consecuencia inmediata del Teorema 9.1.6 y el Corolario 9.1.14. ■

Observe que, gracias al resultado de Volterra, una función diferenciable f puede tener derivada acotada y su derivada f' no ser Riemann integrable. Sin embargo, *si f' es Riemann integral*, entonces ella es acotada y por lo tanto, por el corolario anterior, de variación acotada. Más aun, en este caso, se tiene la siguiente conclusión.

Teorema 9.1.17. *Sea $f : [a,b] \to \mathbb{R}$ una función. Si f es* ***diferenciable*** *y $f' \in \mathcal{R}([a,b])$, entonces $f \in \mathrm{BV}([a,b])$ y*

$$V(f,[a,b]) = \int_a^b |f'(x)|\,dx. \tag{$*$}$$

Prueba. El hecho de que f sea de variación acotada sigue del corolario anterior. Para demostrar $(*)$, sea $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ una partición de $[a,b]$. Como f' es Riemann integral, el Teorema Fundamental del Cálculo nos garantiza que

$$\int_{x_{i-1}}^{x_i} f'\,dx = f(x_i) - f(x_{i-1})$$

para $i = 1, 2, \ldots, n$ y así,

$$\begin{aligned}\sum_{i=1}^n |f(x_i) - f(x_{i-1})| &= \sum_{i=1}^n \left| \int_{x_{i-1}}^{x_i} f'\,dx \right| \\ &\leq \sum_{i=1}^n \int_{x_{i-1}}^{x_i} |f'|\,dx \\ &= \int_a^b |f'|\,dx.\end{aligned}$$

De esto se deduce que

$$V(f,[a,b]) \leq \int_a^b |f'(x)|\,dx.$$

Para demostrar la otra desigualdad, sea $\varepsilon > 0$. Puesto que f' es Riemann integral, también lo es $|f'|$ y, en consecuencia, podemos determinar la existencia de un $\delta > 0$ tal que

$$\left| \int_a^b |f'|\,dx - \sum_{i=1}^n |f'(t_i)|(x_i - x_{i-1}) \right| < \varepsilon \tag{1}$$

para cualquier partición etiquetada $\mathcal{P}_e = (t_i, [x_{i-1}, x_i])_{i=1}^n$ de $[a, b]$ para la cual $\|\mathcal{P}\| \leq \delta$. Apliquemos el Teorema del Valor Medio a f para obtener, para cada $i = 1, \ldots, n$, un $\xi_i \in (x_{i-1}, x_i)$ tal que

$$f(x_i) - f(x_{i-1}) \;=\; f'(\xi_i)(x_i - x_{i-1}).$$

Considerando las etiquetas $(\xi_i)_{i=1}^n$ de $\mathcal{P}$, resulta de (1) y lo anterior que

$$\int_a^b |f'|\, dx \,-\, \varepsilon \;<\; \sum_{i=1}^n |f'(\xi_i)|(x_i - x_{i-1}) \;=\; \sum_{i=1}^n \big|f(x_i) - f(x_{i-1})\big| \;\leq\; V(f, [a, b]).$$

Como ε es arbitrario, concluimos que

$$\int_a^b |f'|\, dx \;\leq\; V(f, [a, b]).$$

La prueba es completa. ■

Ya hemos visto que si $f \in \mathcal{R}([a, b])$, entonces su integral indefinida F es una función de variación acotada. El siguiente resultado indica una relaciona más estrecha entre la variación total de F y la integral de $|f|$.

Lema 9.1.18. *Si $f \in \mathcal{R}([a, b])$, entonces su integral indefinida $F(x) = \int_a^x f(t)\, dt$ para todo $x \in [a, b]$, posee la siguiente propiedad: para cualquier partición $\mathcal{P} = \{x_0, x_1, \ldots x_n\}$ de $[a, b]$, se cumple que*

$$\sum_{i=1}^n |F(x_i) - F(x_{i-1})| \;\leq\; \int_a^b |f|\, dx.$$

En particular,

$$V(F, [a, b]) \;\leq\; \int_a^b |f|\, dx.$$

Prueba. Sean $F, G : [a, b] \to \mathbb{R}$ las integrales indefinidas de f y $|f|$ respectivamente. Puesto que G es Lipschitz y $G' = |f|$ $\mu -$c.s., se sigue del Teorema Fundamental del Cálculo, Teorema 9.1.11, que

$$\int_{x_{i-1}}^{x_i} |f|\, dx \;=\; G(x_i) - G(x_{i-1})$$

para cada $i = 1, 2, \ldots, n$. Finalmente, de las propiedades de la integral, se tiene que

$$\begin{aligned}
\sum_{i=1}^n |F(x_i) - F(x_{i-1})| \;&=\; \sum_{i=1}^n \left| \int_a^{x_i} f\, dx - \int_a^{x_{i-1}} f\, dx \right| \;=\; \sum_{i=1}^n \left| \int_{x_i}^{x_{i-1}} f\, dx \right| \\
&\leq\; \sum_{i=1}^n \int_{x_i}^{x_{i-1}} |f|\, dx \;=\; \sum_{i=1}^n \big(G(x_i) - G(x_{i-1})\big) \\
&=\; G(b) - G(a) \;=\; \int_a^b |f|\, dx.
\end{aligned}$$

Esto termina la prueba. ■

Lema 9.1.19. *Sea $f : [a,b] \to \mathbb{R}$ una función. Si $f \in \mathrm{BV}([a,b])$, entonces $f \in \mathrm{BV}([c,d])$ para cualquier subintervalo cerrado $[c,d] \subseteq [a,b]$ y*

$$V(f,[a,b]) \;=\; V(f,[a,c]) + V(f,[c,b])$$

para cada $c \in (a,b)$. En particular, si f es de variación acotada sobre $[a,c]$ y también sobre $[c,b]$, entonces $f \in \mathrm{BV}([a,b])$.

Prueba. Suponga que $f \in BV([a,b])$ y sea $\mathcal{Q} = \{y_0, y_1, \dots, y_n\}$ una partición de $[c,d] \subseteq [a,b]$. Entonces $\mathcal{P} = \{x_0, x_1, \dots, x_{n+1}\}$, donde

$$a \;=\; x_0,\; x_1 = y_0,\; x_2 \;=\; y_1,\; \cdots,\; x_{n-1} \;=\; y_n,\; x_{n+1} \;=\; b$$

es una partición de $[a,b]$ tal que

$$\begin{aligned} +\infty \;>\; V(f,[a,b]) \;&\geq\; \sum_{j=1}^{n} |f(x_j) - f(x_{j-1})| \\ &\geq\; \sum_{j=2}^{n-1} |f(x_j) - f(x_{j-1})| \\ &=\; \sum_{i=1}^{n} |f(y_i) - f(y_{i-1})|. \end{aligned}$$

Se sigue que $V(f,[c,d]) \leq V(f,[a,b]) < +\infty$ y, en consecuencia, $f \in \mathrm{BV}([c,d])$. Fijemos $c \in (a,b)$ y sea $\mathcal{P} = \{x_0, x_1, \dots, x_n\}$ una partición de $[a,b]$. Para simplificar la presentación, supondremos que $x_i = c$ para algún $i \in \{1, \dots, n-1\}$. Entonces

$$\mathcal{P}_1 \;=\; \{x_0, x_1, \dots, x_i\} \qquad \text{y} \qquad \mathcal{P}_2 \;=\; \{x_i, x_{i+1}, \dots, x_n\}$$

son particiones de $[a,c]$ y $[c,b]$ respectivamente, tales que

$$\begin{aligned} \sum_{k=1}^{n} |f(x_k) - f(x_{k-1})| \;&=\; \sum_{k=1}^{i} |f(x_k) - f(x_{k-1})| + \sum_{k=i+1}^{n} |f(x_k) - f(x_{k-1})| \\ &\leq\; V(f,[a,c]) + V(f,[c,b]) \end{aligned}$$

De aquí se sigue que

$$V(f,[a,b]) \;\leq\; V(f,[a,c]) + V(f,[c,b]).$$

Para demostrar la otra dirección de la desigualdad anterior, sea $\varepsilon > 0$. Escojamos ahora particiones

$$\mathcal{P}_1 \;=\; \{x_0, x_1, \dots, x_i\} \qquad \text{y} \qquad \mathcal{P}_2 \;=\; \{x_i, x_{i+1}, \dots, x_n\}$$

de $[a,c]$ y $[c,b]$ respectivamente, tales que

$$\sum_{k=1}^{i} |f(x_k) - f(x_{k-1})| \;>\; V(f,[a,c]) - \frac{\varepsilon}{2} \qquad \text{y} \qquad \sum_{k=i+1}^{n} |f(x_k) - f(x_{k-1})| \;>\; V(f,[c,b]) - \frac{\varepsilon}{2}.$$

Entonces $\mathcal{P} = \mathcal{P}_1 \cup \mathcal{P}_2 = \{x_0, x_1, \ldots, x_n\}$ es una partición de $[a,b]$ tal que

$$\begin{aligned} V(f,[a,b]) &\geq \sum_{k=1}^{i} |f(x_k) - f(x_{k-1})| + \sum_{k=i+1}^{n} |f(x_k) - f(x_{k-1})| \\ &> V(f,[a,c]) + V(f,[c,b]) - \varepsilon. \end{aligned}$$

Como $\varepsilon > 0$ es arbitrario, se concluye que

$$V(f,[a,b]) \geq V(f,[a,c]) + V(f,[c,b]).$$

La última parte (b) es consecuencia de la primera. ■

Lema 9.1.20. *Sea $f \in \mathrm{BV}([a,b])$. La función $V_f : [a,b] \to \mathbb{R}$ definida por*

$$V_f(x) = V(f,[a,x]) \qquad \text{para todo} \quad x \in [a,b]$$

es ***monótona creciente****. En particular, si f no es constante en ningún subintervalo de $[a,b]$, entonces V_f es estrictamente creciente.*

Prueba. Sean $x_1, x_2 \in [a,b]$ con $x_1 < x_2$. Puesto que f es de variación acotada, el Lema 9.1.19 nos garantiza que

$$V(f,[a,x_2]) = V(f,[a,x_1]) + V(f,[x_1,x_2]),$$

de donde se obtiene que

$$V_f(x_2) - V_f(x_1) = V(f,[a,x_2]) - V(f,[a,x_1]) = V(f,[x_1,x_2]) \geq 0.$$

La última parte sigue del Ejemplo 9.1.2. La prueba es completa. ■

Observe que, por definición y el resultado anterior, se obtiene la siguiente desigualdad:

Corolario 9.1.21. *Sea $f \in \mathrm{BV}([a,b])$. Si $x_1, x_2 \in [a,b]$ con $x_1 < x_2$, entonces se cumple que*

$$|f(x_2) - f(x_1)| \leq V(f,[x_1,x_2]) = V_f(x_2) - V_f(x_1).$$

Una de las representaciones más intrigantes de las funciones de variación acotada se debe a Marie Ennemond Camille Jordan (1838-1922). Ella expresa, de modo agradable, de qué están hechas las funciones de variación acotada.

Teorema 9.1.22 (Jordan). *Sea $f : [a,b] \to \mathbb{R}$ una función. Son equivalentes:*

(1) $f \in \mathrm{BV}([a,b])$.

(2) *Existen* ***funciones crecientes*** $g_1, g_2 : [a,b] \to \mathbb{R}$ *tales que* $f = g_1 - g_2$.

Prueba. (1) $\Rightarrow$ (2). Sea $g_1 = V_f$, donde V_f es la función monótona creciente del lema anterior. Si ahora definimos $g_2 : [a,b] \to \mathbb{R}$ por

$$g_2(x) = g_1(x) - f(x) \qquad \text{para todo} \quad x \in [a,b]$$

vemos que $f = g_1 - g_2$. Nos falta demostrar que g_2 también es monótona creciente. Sean $x_1, x_2 \in [a, b]$ con $x_1 < x_2$. Entonces, por el Lema 9.1.19

$$\begin{aligned} g_1(x_2) - g_1(x_1) &= V(f, [x_1, x_2]) \\ &\geq |f(x_2) - f(x_1)| \\ &\geq f(x_2) - f(x_1) \end{aligned}$$

y, por consiguiente, $g_1(x_2) - f(x_2) \geq g_1(x_1) - f(x_1)$, es decir, $g_2(x_2) \geq g_2(x_1)$.

(2) $\Rightarrow$ (1). Sea $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ una partición arbitraria de $[a, b]$. Entonces

$$\begin{aligned} \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| &= \sum_{i=1}^{n} \left|\big(g_1(x_i) - g_2(x_i)\big) - \big(g_1(x_{i-1}) - g_2(x_{i-1})\big)\right| \\ &\leq \sum_{i=1}^{n} |g_1(x_i) - g_1(x_{i-1})| + \sum_{i=1}^{n} |g_2(x_i) - g_2(x_{i-1})| \\ &\leq g_1(b) - g_1(a) + g_2(b) - g_2(a), \end{aligned}$$

de donde se sigue que $V(f, [a, b]) < +\infty$. Esto termina la prueba. ■

Observe que si $f \in \mathrm{BV}([a, b])$, entonces las funciones g_1 y g_2 en el teorema anterior se pueden elegir **estrictamente crecientes**. En efecto, sean g_1, g_2 las funciones monótonas crecientes obtenidas en el Teorema 9.1.22 tales que $f = g_1 - g_2$. Tomando

$$f_1(x) = x + g_1(x) \qquad \text{y} \qquad f_2(x) = x + g_2(x)$$

para todo $x \in [a, b]$, vemos que f_1 y f_2 son estrictamente crecientes y se cumple, además, que $f = f_1 - f_2$.

Una de las primeras consecuencias del Teorema de Jordan en combinación con el Corolario 3.1.35, página 174 y el Teorema de Vitali-Lebesgue es el siguiente corolario.

Corolario 9.1.23. *Si $f \in \mathrm{BV}([a, b])$, entonces $\mathrm{Disc}(f)$ es **a lo más numerable**. En particular,*

$$\mathrm{BV}([a, b]) \subseteq \mathcal{R}([a, b]).$$

Prueba. Sea $f \in \mathrm{BV}([a, b])$. Por el Teorema de Jordan, $f = f_1 - f_2$, donde f_1, f_2 son funciones crecientes. Por el Corolario 3.1.35, el conjunto de discontinuidad de cada f_i, $i = 1, 2$, es a lo más numerable y, en consecuencia, $\mathrm{Disc}(f)$ es a lo más numerable. El hecho de que $f \in \mathcal{R}([a, b])$ sigue del Teorema de Vitali-Lebesgue. ■

Recuerde que $\mathrm{Mon}([a, b])$ es el conjunto de las funciones $f : [a, b] \to \mathbb{R}$ que son monótonas. Del Teorema 3.1.34 sabemos que $\mathrm{Mon}([a, b]) \subseteq \mathrm{Reg}([a, b])$, donde $\mathrm{Reg}([a, b])$ es el conjunto de todas las funciones reguladas, y del Teorema de Jordan se sigue que $\mathrm{BV}([a, b]) = [\mathrm{Mon}([a, b])]$, donde $[A]$ denota el espacio lineal generado por A, es decir, el espacio lineal más pequeño que contiene a A. Puesto que $\mathrm{Reg}([a, b])$ es un espacio vectorial, resulta que

$$\mathrm{BV}([a, b]) \subseteq \mathrm{Reg}([a, b]).$$

Teorema 9.1.24. *Sea* $f \in \mathrm{BV}([a,b])$ *y defina* $V_f : [a,b] \to \mathbb{R}$ *por* $V_f(x) = V(f,[a,x])$ *para todo* $x \in [a,b]$*. Son equivalentes:*

(1) f *es* ***continua sobre*** $[a,b]$.

(2) V_f *es* ***continua sobre*** $[a,b]$.

Prueba. (1) $\Rightarrow$ (2). Suponga que f es continua sobre $[a,b]$ y sea $c \in [a,b]$. Queremos demostrar que

$$\lim_{x\to c^+} V_f(x) \;=\; V_f(c) \;=\; \lim_{x\to c^-} V_f(x).$$

Demostraremos en primer lugar que $\lim_{x\to c^+} V_f(x) = V_f(c)$. Sea $\varepsilon > 0$ y suponga entonces que $a \leq c < b$. Por la continuidad de f en c, existe un $\delta > 0$ tal que

$$|f(x) - f(c)| \;<\; \varepsilon \quad \text{para todo} \quad x \in (c, c+\delta).$$

Por la definición de $V(f,[c,b])$, existe una partición $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ tal que

$$V(f,[c,b]) \;-\; \frac{\varepsilon}{2} \;<\; \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|.$$

Si $x_1 \in (c, c+\delta)$ nos detenemos. En caso contrario, escojamos un $x \in (c, c+\delta)$ y formemos la nueva partición $\mathcal{P}_x = \{y_0, y_1, \ldots, y_{n+1}\}$, donde

$$y_0 \;=\; x_0, \quad y_1 \;=\; x, \quad y_2 \;=\; x_1, \quad \cdots, \quad y_{n+1} \;=\; x_n.$$

Observe que $V(f,[x,b]) \geq \sum_{i=2}^{n+1} |f(y_i) - f(y_{i-1})|$. Por esto,

$$\begin{aligned}
V_f(x) \;-\; V_f(c) \;&=\; V(f,[c,x]) \;=\; V(f,[c,b]) \;-\; V(f,[x,b]) \\
&< \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| \;+\; \frac{\varepsilon}{2} \;-\; \sum_{i=2}^{n+1} |f(y_i) - f(y_{i-1})| \\
&\leq \sum_{i=1}^{n+1} |f(y_i) - f(y_{i-1})| \;+\; \frac{\varepsilon}{2} \;-\; \sum_{i=2}^{n+1} |f(y_i) - f(y_{i-1})| \\
&= |f(x) - f(c)| \;+\; \frac{\varepsilon}{2} \\
&< \varepsilon.
\end{aligned}$$

De aquí se deduce que $\lim_{x\to c^+} V_f(x) = V_f(c)$. Un argumento enteramente similar conduce a que $\lim_{x\to c^-} V_f(x) = V_f(c)$ y, por lo tanto, V_f es continua en c. Como c es arbitrario, tenemos que V_f es continua sobre $[a,b]$.

(2) $\Rightarrow$ (1). Suponga que V_f es continua sobre $[a,b]$ y sea $c \in [a,b]$. Dado $\varepsilon > 0$, existe un $\delta > 0$ tal que

$$|V_f(x) - V_f(c)| \;<\; \varepsilon \quad \text{para todo} \quad x \in (c-\delta, c+\delta) \cap [a,b].$$

Nótese que si $c < x$, entonces

$$|f(x) - f(c)| \;\leq\; V(f,[c,x]) \;=\; V_f(x) - V_f(c) \;<\; \varepsilon.$$

Similarmente, si $x < c$, entonces

$$|f(x) - f(c)| \leq V(f, [x, c]) = V_f(c) - V_f(x) < \varepsilon.$$

Esto nos dice que f es continua en c y termina la prueba. ■

Combinando el resultado anterior con el Teorema de Jordan se obtiene el siguiente corolario.

Corolario 9.1.25. *Sea $f \in \mathrm{BV}([a, b])$. Si f es* ***continua*** *sobre $[a, b]$, entonces existen dos funciones continuas y estrictamente crecientes f_1, f_2 tales que $f = f_1 - f_2$.*

Dada una función $f : [a, b] \to \mathbb{R}$ y una partición $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$, pongamos

$$\mathrm{O}(f, \mathcal{P}) = \sum_{i=1}^{n} (M_i - m_i),$$

donde $M_i = \sup\{f(x) : x \in [x_{i-1}, x_i]\}$ y $m_i = \inf\{f(x) : x \in [x_{i-1}, x_i]\}$. Finalizamos esta sección con un resultado que establece que toda función continua de variación acotada se puede representar como un límite de las sumas $V(f, \mathcal{P})$ y $\mathrm{O}(f, \mathcal{P})$ cuando $\|\mathcal{P}\| \to 0$.

Teorema 9.1.26. *Sea $f : [a, b] \to \mathbb{R}$ una función* ***continua*** *de* ***variación acotada*** *sobre $[a, b]$. Entonces*

$$\lim_{\|\mathcal{P}\| \to 0} V(f, \mathcal{P}) = V(f, [a, b]) = \lim_{\|\mathcal{P}\| \to 0} \mathrm{O}(f, \mathcal{P}).$$

Prueba. Vamos a demostrar, en primer lugar, que $\lim_{\|\mathcal{P}\| \to 0} V(f, \mathcal{P}) = V(f, [a, b])$. Sea $\varepsilon > 0$ y seleccione, usando el hecho de que f es de variación acotada, una partición $\mathcal{P}_0 = \{y_0, y_1, \ldots, y_{m+1}\}$ de $[a, b]$ tal que

$$V(f, [a, b]) - \frac{\varepsilon}{2} < V(f, \mathcal{P}_0). \tag{1}$$

Ahora bien, teniendo en cuenta que f es uniformemente continua en $[a, b]$, escoja un $\delta_1 > 0$ de modo que tal que

$$|f(x) - f(y)| < \frac{\varepsilon}{4m} \quad \text{siempre que} \quad |x - y| < \delta_1. \tag{2}$$

Sea $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ cualquier partición de $[a, b]$ con $\|\mathcal{P}\| < \delta$, donde $\delta < \delta_1$ se elige lo suficientemente pequeño de modo que cada intervalo $[x_{k-1}, x_k]$ contenga a lo sumo un punto $y_j \in \mathcal{P}_0$. Considere la partición $\mathcal{P}' = \mathcal{P} \cup \mathcal{P}_0$ y observe que como $\mathcal{P}'$ es un refinamiento de $\mathcal{P}_0$, resulta que

$$V(f, \mathcal{P}') \geq V(f, \mathcal{P}_0).$$

De inmediato, vamos a comparar las sumas $V(f, \mathcal{P})$ y $V(f, \mathcal{P}')$. Observe que la suma $V(f, \mathcal{P})$ consiste de los n términos $|f(x_i) - f(x_{i-1})|$ y que ellos coinciden con los términos de $V(f, \mathcal{P}')$, con la excepción de aquellos que corresponden a los intervalos $[x_{k-1}, x_k]$ que contienen un punto y_j con $1 \leq j \leq m$. Por consiguiente, el término $|f(x_{k-1}) - f(x_k)|$ es sustituido por $|f(y_j) - f(x_{k-1})| + |f(x_k) - f(y_j)|$, de donde se deduce que $V(f, \mathcal{P}') - V(f, \mathcal{P})$ consiste de m términos de la forma

$$|f(y_j) - f(x_{k-1})| + |f(x_k) - f(y_j)| - |f(x_k) - f(x_{k-1})|.$$

y como $|x_k - y_j| < \delta_1$ y $|y_j - x_{k-1}| < \delta_1$, se sigue entonces de (2) que

$$\begin{aligned} 0 &\leq |f(y_j) - f(x_{k-1})| + |f(x_k) - f(y_j)| - |f(x_k) - f(x_{k-1})| \\ &\leq |f(y_j) - f(x_{k-1})| + |f(x_k) - f(y_j)| \\ &< \frac{\varepsilon}{4m} + \frac{\varepsilon}{4m} = \frac{\varepsilon}{2m}. \end{aligned}$$

Por esto,

$$V(f, \mathcal{P}') - V(f, \mathcal{P}) < m \cdot \frac{\varepsilon}{2m} = \frac{\varepsilon}{2}$$

y así, usando (1), vemos que

$$V(f, [a,b]) - \frac{\varepsilon}{2} < V(f, \mathcal{P}') < V(f, \mathcal{P}) + \frac{\varepsilon}{2},$$

es decir,

$$V(f, [a,b]) - V(f, \mathcal{P}) < \varepsilon \quad \text{para cada} \quad \mathcal{P} \in \mathbb{P}[a,b] \ \text{ con } \|\mathcal{P}\| < \delta.$$

Esto último significa, por supuesto, que $\lim_{\|\mathcal{P}\| \to 0} V(f, \mathcal{P}) = V(f, [a,b])$.

Para demostrar que $V(f, [a,b]) = \lim_{\|\mathcal{P}\| \to 0} \mathrm{O}(f, \mathcal{P})$ sea, de nuevo, $\varepsilon > 0$ y escojamos un $\delta > 0$, usando la primera parte, tal que

$$V(f, [a,b]) - \varepsilon < V(f, \mathcal{P}) \tag{3}$$

para cada partición $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ de $[a,b]$ con $\|\mathcal{P}\| < \delta$. Fijemos una tal partición $\mathcal{P}$ y observe que como $V(f, \mathcal{P}) \leq \mathrm{O}(f, \mathcal{P})$, entonces, por (3), se tiene que

$$V(f, [a,b]) - \varepsilon < \mathrm{O}(f, \mathcal{P}). \tag{4}$$

Por otro lado, tome cualquier partición $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ de $[a,b]$ y sean t_i y s_i puntos en $[x_{i-1}, x_i]$ donde f alcanza su máximo y su mínimo respectivamente, esto es, $M_i = f(t_i)$ y $m_i = f(s_i)$. Añada los puntos $t_1, s_1, \ldots, t_n, s_n$ a la partición $\mathcal{P}$ y denote a esta nueva partición por $\mathcal{P}'$. Puesto que la suma $\mathrm{O}(f, \mathcal{P}')$ contiene a los términos $|f(t_i) - f(s_i)| = M_i - m_i$ resulta que

$$V(f, [a,b]) \geq V(f, \mathcal{P}') = \sum_{i=1}^{n} (M_i - m_i) + K,$$

donde K son los otro términos positivos que definen a $V(f, \mathcal{P}')$. De aquí se sigue que

$$V(f, [a,b]) \geq \mathrm{O}(f, \mathcal{P}) \quad \text{para cualquier } \mathcal{P} \in \mathbb{P}[a,b].$$

Este hecho combinado con (4) produce la desigualdad

$$V(f, [a,b]) - \varepsilon < \mathrm{O}(f, \mathcal{P}) \leq V(f, [a,b])$$

la cual es válida para toda partición $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ de $[a,b]$ con $\|\mathcal{P}\| < \delta$. Esto demuestra que $V(f, [a,b]) = \lim_{\|\mathcal{P}\| \to 0} \mathrm{O}(f, \mathcal{P})$ y termina la prueba. ■

9.1.3. Cubrimientos de Vitali

La noción de compacidad establece que cada *cubrimiento abierto* de un conjunto compacto se puede reducir a un subcubrimiento finito. En esta corta sección se considerará un tipo muy especial de cubrimientos. Se trata de estimar $\mu^*(E)$, donde E es un conjunto arbitrario que puede ser cubierto por una cierta colección de intervalos "*pequeños*" y el objetivo es obtener una subcolección (finita o infinita), pero *disjunta*, con la propiedad de que la unión de los miembros de la subcolección *cubra casi por completo* al conjunto dado. El Teorema del Cubrimiento de Vitali, descubierto por G. Vitali en el año 1907 cuando intentaba extender el Teorema Fundamental del Cálculo a $\mathbb{R}^2$, permite que una tal subcolección se pueda obtener. Muy pronto dicho teorema se convertiría en una pieza clave en la prueba de un resultado fundamental sobre la diferenciabilidad de las funciones de variación acotadas conocido como el Teorema de Diferenciación de Lebesgue.

Definición 9.1.27. *Sean $E \subseteq \mathbb{R}$ y $\mathcal{V}$ una colección de subintervalos cerrados no-degenerados de $\mathbb{R}$. Diremos que $\mathcal{V}$ es un* ***cubrimiento de E en el sentido de Vitali****, o que es un* ***cubrimiento de Vitali*** *de E si, para cada $\varepsilon > 0$, existe un intervalo $J \in \mathcal{V}$ tal que*

$$x \in J \qquad y \qquad \ell(J) < \varepsilon.$$

Es importante destacar los siguientes hechos respecto a los cubrimientos de Vitali. El primero es un ejemplo sencillo de un cubrimiento de Vitali, pero los otros ejemplos serán utilizados en varias pruebas posteriores

(1) *Sea $E = [0,1]$ y sea $\mathbb{Q}_{[0,1]}$ los números racionales en $[0,1]$. Entonces la colección*

$$\mathcal{V} = \left\{ \left[q - \frac{1}{n},\, q + \frac{1}{n}\right] : q \in \mathbb{Q}_{[0,1]},\ n \in \mathbb{N} \right\}$$

*constituye un c**ubrimiento de E en el sentido de Vitali**.*

Para comprobar nuestra afirmación seleccione, de modo arbitrario, un $x \in E$ y sea $\varepsilon > 0$. El Principio de Arquímedes nos permite elegir un $n \in \mathbb{N}$ de modo que $1/n < \varepsilon$ y ahora use el hecho de que $\mathbb{Q}_{[0,1]}$ es denso en $[0,1]$ para hallar un $q \in \mathbb{Q}_{[0,1]}$ tal que $|x - q| < 1/2n$. Esto último implica que

$$x \in I_{n,q} = \left[q - \frac{1}{2n},\, q + \frac{1}{2n}\right] \in \mathcal{V} \qquad \text{y} \qquad \ell(I_{n,q}) = \frac{1}{n} < \varepsilon.$$

(2) *Sea $\mathcal{V}$ un* ***cubrimiento de Vitali*** *de $E \subseteq \mathbb{R}$. Si G es un* ***conjunto abierto*** *conteniendo puntos de E, entonces G contiene* ***intervalos*** *pertenecientes a $\mathcal{V}$. En particular,*

$$\mathcal{V}_G = \{J \in \mathcal{V} : J \subseteq G\}$$

constituye un ***cubrimiento de Vitali*** *de $E \cap G$.*

En efecto, sea $x \in E \cap G$. Puesto que $x \in G$ y G es abierto, existe un $\varepsilon > 0$ tal que $(x - \varepsilon, x + \varepsilon) \subseteq G$. Por otro lado, como $x \in E$ y $\mathcal{V}$ es un cubrimiento de Vitali, para el ε hallado, existe un intervalo $J \in \mathcal{V}$ con $x \in J$ y $\ell(J) < \varepsilon$. De esto se concluye que los intervalos de $\mathcal{V}$ que contienen a x y son de longitud menor o igual que ε están totalmente contenidos en G. Esto prueba que $\mathcal{V}_G$ es un cubrimiento de Vitali de $E \cap G$.

De lo anterior se sigue que:

(3) *Si G es un **conjunto abierto** conteniendo a E, entonces $\mathcal{V}_G$ es un **cubrimiento de** E **en el sentido de Vitali.***

(4) *Sea E un conjunto arbitrario y suponga que para cada $x \in E$ se puede escoger una **sucesión estrictamente decreciente** $(y_n^x)_{n=1}^{\infty}$ **convergiendo** a x. Entonces la colección*

$$\mathcal{V} \;=\; \big\{[x, y_n^x] \;:\; x \in E,\; n \in \mathbb{N}\big\}$$

*es un **cubrimiento de** E **en el sentido de Vitali**.*

Para demostrar esto, sean $x \in E$ y $\varepsilon > 0$. Puesto que $y_n^x \to x$, existe un $n \geq 1$ tal que $|x - y_n^x| < \varepsilon$. Claramente

$$x \in [x, y_n^x] \in \mathcal{V} \qquad \text{y} \qquad \ell([x, y_n^x]) \;=\; y_n^x - x \;<\; \varepsilon.$$

Hechas estas consideraciones pasemos ahora a demostrar el siguiente resultado.

Teorema 9.1.28 (Cubrimiento de Vitali). *Sea $E \subseteq \mathbb{R}$ con $\mu^*(E) < +\infty$ y sea $\mathcal{V}$ un cubrimiento de E en el sentido de Vitali. Entonces, para cada $\varepsilon > 0$, existe una **colección finita** $\{I_1, \ldots, I_n\}$ de **intervalos disjuntos** en $\mathcal{V}$ tal que*

$$\mu^*\Big(E \setminus \bigcup_{j=1}^{n} I_j\Big) \;<\; \varepsilon.$$

Además, existe una sucesión disjunta $(I_n)_{n=1}^{\infty}$ en $\mathcal{V}$ tal que $\mu^*\big(E \setminus \bigcup_{j=1}^{\infty} I_j\big) = 0$.

Prueba. Sea $\varepsilon > 0$. Una aplicación del Teorema 6.2.12 permite que podamos seleccionar un conjunto abierto G tal que

$$E \;\subseteq\; G \qquad \text{y} \qquad \mu(G) \;<\; \mu^*(E) + \varepsilon.$$

Como $\mu^*(E) < +\infty$, resulta entonces que $\mu(G) < +\infty$. De las observaciones hechas anteriormente sabemos que la familia

$$\mathcal{V}_0 \;=\; \Big\{I \in \mathcal{V} \;:\; I \subseteq G\Big\}$$

sigue siendo un cubrimiento de E en el sentido de Vitali, por lo que no se pierde generalidad en el razonamiento asumir que $I \subseteq G$ para cada $I \in \mathcal{V}$.

Sea I_1 cualquier intervalo de $\mathcal{V}$. Si $E \subseteq I_1$, entonces $\mu^*(E \setminus I_1) = 0 < \varepsilon$ y termina la prueba. Suponga entonces que $E \nsubseteq I_1$ y sea $x \in E \setminus I_1$. Como $\mathcal{V}$ es un cubrimiento de Vitale de E, existe un intervalo $J \in \mathcal{V}$ conteniendo a x que no intersecta a I_1. Esto prueba que la familia $\mathcal{F}_1 = \{I \in \mathcal{V} : I \cap I_1 = \varnothing\}$ es no vacía. Considere

$$k_1 \;=\; \sup\big\{\ell(I) \;:\; I \in \mathcal{F}_1\big\}.$$

Como ninguno de los intervalos en $\mathcal{V}$ son degenerados, $k_1 > 0$. Más aun, teniendo en cuenta que $I \subseteq G$ para todo $I \in \mathcal{F}_1$, resulta que $\ell(I) \leq \mu(G) < +\infty$, de donde se concluye que $k_1 < +\infty$. Usemos las propiedades del supremo para seleccionar un intervalo $I_2 \in \mathcal{F}_1$ de modo tal que

$$\frac{1}{2}k_1 \;<\; \ell(I_2).$$

Observe que $I_1 \cap I_2 = \varnothing$. Si $\mu^*(E \setminus (I_1 \cup I_2)) = 0$ nos detenemos pues el teorema ha quedado demostrado. En caso contrario, defina $\mathcal{F}_2 = \{I \in \mathcal{V} : I \cap (I_1 \cup I_2) = \varnothing\}$. Como antes, se prueba que $\mathcal{F}_2 \neq \varnothing$ y, en consecuencia,

$$k_2 \;=\; \sup\big\{\ell(I) \,:\, I \in \mathcal{F}_2\big\}.$$

satisface la relación $0 < k_2 < +\infty$. Una vez más, seleccione un intervalo $I_3 \in \mathcal{F}_2$ tal que

$$\frac{1}{2}k_2 \;<\ell(I_3).$$

Si $\mu^*(E \setminus (I_1 \cup I_2 \cup I_3)) = 0$ paramos. Si no, defina como antes la colección $\mathcal{F}_3 = \{I \in \mathcal{V} : I \cap (I_1 \cup I_2 \cup I_3) = \varnothing\}$ y sea

$$k_3 \;=\; \sup\big\{\ell(I) \,:\, I \in \mathcal{F}_3\big\},$$

y se continua inductivamente con este proceso. Existen dos posibilidades:

(a) o bien el proceso anterior finaliza al cabo de n pasos, es decir, existen n intervalos disjuntos dos a dos $I_1, \ldots, I_n$ en $\mathcal{V}$ tal que

$$\mu^*(E \setminus (I_1 \cup \cdots \cup I_n)) \;=\; 0 \;<\; \varepsilon$$

y con ello queda establecido el teorema,

(b) o bien el proceso continúa indefinidamente, esto es,

$$\mu^*(E \setminus (I_1 \cup \cdots \cup I_n)) \;>\; 0 \quad \text{para todo } n \geq 1.$$

De ocurrir (b) se generan tres sucesiones, $(\mathcal{F}_n)_{n=1}^{\infty}$, $\big(I_n\big)_{n=1}^{\infty}$ y $(k_n)_{n=1}^{\infty}$ satisfaciendo las siguientes propiedades:

(1) $\mathcal{F}_1 \supseteq \mathcal{F}_2 \supseteq \cdots \supseteq \mathcal{F}_n \supseteq \cdots$.

(2) $\big(I_n\big)_{n=1}^{\infty}$ es una *sucesión disjunta* en $\mathcal{V}$ que cumple con

$$I \,\cap\, \left(\bigcup_{k=1}^{n} I_k\right) \;=\; \varnothing \quad \text{para todo } I \in \mathcal{F}_n, \qquad n = 1, 2, \ldots.$$

(3) $(k_n)_{n=1}^{\infty}$ satisface las relaciones

$$k_n \;=\; \sup\big\{\ell(I) \,:\, I \in \mathcal{F}_n\big\} \qquad \text{y} \qquad \frac{1}{2}k_n \;<\; \ell(I_{n+1}), \qquad n = 1, 2, \ldots$$

Puesto que la sucesión $\big(I_n\big)_{n=1}^{\infty}$ es disjunta,

$$\sum_{n=1}^{\infty} \ell(I_n) \;=\; \sum_{n=1}^{\infty} \mu(I_n) \;=\; \mu\!\left(\bigcup_{n=1}^{\infty} I_n\right) \;\leq\; \mu(G) \;<\; +\infty.$$

De allí que $\lim_{n\to\infty} \ell(I_n) = 0$. En particular, $\lim_{n\to\infty} k_n = 0$ gracias a (3). Usemos ahora la convergencia de la serie $\sum_{n=1}^{\infty} \ell(I_n)$ para hallar un entero $N \geq 1$ de modo tal que

$$\sum_{n=N+1}^{\infty} \ell(I_n) \;<\; \frac{\varepsilon}{5}.$$

Una vez que la desigualdad anterior ha sido establecida escojamos, por cada $n > N$, un intervalo cerrado J_n con el mismo centro que I_n y 5 veces su longitud. De lo anterior se sigue que

$$\mu\Big(\bigcup_{n=N+1}^{\infty} J_n\Big) \;\leq\; \sum_{n=N+1}^{\infty} \mu(J_n) \;=\; \sum_{n=N+1}^{\infty} 5\,\ell(I_n) \;=\; 5\sum_{n=N+1}^{\infty} \ell(I_n) \;<\; \varepsilon.$$

Si logramos demostrar que

$$E \setminus \bigcup_{n=1}^{N} I_n \subseteq \bigcup_{n=N+1}^{\infty} J_n, \tag{$\circledast$}$$

tendríamos entonces que

$$\mu^*\Big(E \setminus \bigcup_{n=1}^{N} I_n\Big) \;\leq\; \mu\Big(\bigcup_{n=N+1}^{\infty} J_n\Big) \;<\; \varepsilon \tag{$\circledast_1$}$$

y finalizaría la demostración de la primera parte del teorema. Para ver que $(\circledast)$ se cumple, elija cualquier $x \in E \setminus \bigcup_{n=1}^{N} I_n$. Puesto que $\bigcup_{n=1}^{N} I_n$ es un conjunto cerrado y $x \notin \bigcup_{n=1}^{N} I_n$, podemos usar de nuevo el hecho de que $\mathcal{V}$ es un cubrimiento de Vitale de E para seleccionar un intervalo $I_x \in \mathcal{V}$ conteniendo a x y que no intersecte a ninguno de los intervalos $I_1, \ldots, I_N$. Esto, por supuesto, significa que $I_x \in \mathcal{F}_N$, es decir,

$$I_x \,\cap \Big(\bigcup_{n=1}^{N} I_n\Big) \;=\; \varnothing. \tag{$\circledast_2$}$$

Afirmamos que $I_x \cap I_n \neq \varnothing$ *para algún* $n > N$.

En efecto, suponga que $I_x \cap I_n = \varnothing$ para todo $n > N$. Combinando esto último con $(\circledast)_2$ se tiene que $I_x \in \mathcal{F}_n$ para todo $n \in \mathbb{N}$, de donde se sigue que $0 \leq \ell(I_x) \leq k_n$ para cualquier $n \in \mathbb{N}$ y así,

$$0 \;\leq\; \ell(I_x) \;\leq\; \lim_{n\to\infty} k_n \;=\; 0.$$

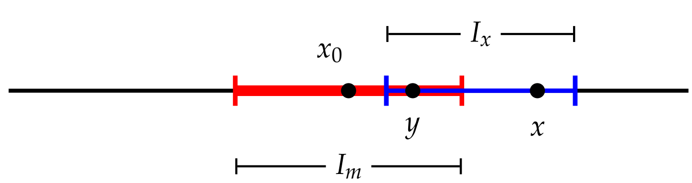
Este hecho nos revela que $I_x \in \mathcal{V}$ es un intervalo degenerado, lo cual es imposible por definición. Nuestra afirmación ha sido demostrada. Sea m el entero positivo más pequeño para el cual

$$I_x \cap I_m \;\neq\; \varnothing.$$

La elección de m y nuestra afirmación nos muestra que $m > N$ y también que $I_x \in \mathcal{F}_{m-1}$, de donde resulta, gracias a (3), que

$$\ell(I_x) \;\leq\; k_{m-1} \;<\; 2\ell(I_m).$$

Sea x_0 el centro del intervalo I_m y sea $y \in I_x \cap I_m$. Entonces



$$
\begin{aligned}
|x - x_0| &\leq |x - y| + |y - x_0| \\
&\leq \ell(I_x) + \frac{1}{2}\ell(I_m) \\
&< 2\ell(I_m) + \frac{1}{2}\ell(I_m) \\
&= \frac{5}{2}\ell(I_m) = \frac{1}{2}\ell(J_m).
\end{aligned}
$$

Se sigue de esto que $x \in J_m$ para algún $m > N$ y, por lo tanto, $x \in \bigcup_{n=N+1}^{\infty} J_n$. Hemos demostrado que

$$E \setminus \bigcup_{n=1}^{N} I_n \subseteq \bigcup_{n=N+1}^{\infty} J_n$$

y entonces $(\circledast)_1$ se cumple. Con esto queda establecida la prueba de la primera parte.

Para demostrar la última parte, observe que de $(\circledast)_1$ se tiene que

$$
\begin{aligned}
\mu^*\left(E \setminus \bigcup_{n=1}^{\infty} I_n\right) &\leq \mu^*\left(E \setminus \bigcup_{n=1}^{N} I_n\right) \\
&\leq \mu\left(\bigcup_{n=N+1}^{\infty} J_n\right) \\
&< \varepsilon
\end{aligned}
$$

y como esta desigualdad es válida para cualquier $\varepsilon > 0$ se concluye que

$$\mu^*\left(E \setminus \bigcup_{n=1}^{\infty} I_n\right) = 0.$$

Esto completa la prueba. ■

Nota Adicional 9.1.1 Del Teorema del Cubrimiento de Vitali sabemos que si $\mathcal{V}$ es un cubrimiento de Vitali para E con $\mu^*(E) < +\infty$, entonces dado $\varepsilon > 0$, existe una colección finita $\{I_1, \ldots, I_n\}$ de intervalos disjuntos en $\mathcal{V}$ tal que

$$\mu^*\left(E \setminus \bigcup_{j=1}^{n} I_j\right) < \varepsilon. \tag{1}$$

En particular,

$$\mu^*\left(E \setminus \bigcup_{j=1}^{\infty} J_j\right) \leq \mu^*\left(E \setminus \bigcup_{j=1}^{n} I_j\right) < \varepsilon. \tag{2}$$

para cualquier colección numerable y disjunta $(J_n)_{n=1}^{\infty} \subseteq \mathcal{V}$ tal que $J_j = I_j$ para $j = 1, \ldots, n$. Similarmente, como $E = \big(E \cap \bigcup_{j=1}^{n} I_j\big) \cup \big(E \setminus \bigcup_{j=1}^{n} I_j\big)$, entonces de (1) se obtiene que

$$\begin{aligned} \mu^*(E) &\leq \mu^*\Big(E \cap \bigcup_{j=1}^{n} I_j\Big) + \mu^*\Big(E \setminus \bigcup_{j=1}^{n} I_j\Big) \\ &\leq \mu^*\Big(\bigcup_{j=1}^{n} I_j\Big) + \mu^*\Big(E \setminus \bigcup_{j=1}^{n} I_j\Big) \\ &< \sum_{j=1}^{n} \ell(I_j) + \varepsilon. \end{aligned} \tag{3}$$

El siguiente resultado es la versión "*no acotada*" del Teorema del Cubrimiento de Vitali.

Teorema 9.1.29 (Cubrimiento de Vitali). *Sea $E \subseteq \mathbb{R}$ con $\mu^*(E) = +\infty$ y sea $\mathcal{V}$ un cubrimiento de E en el sentido de Vitali. Entonces existe una* ***colección infinita*** *$\{I_1, I_2, \ldots\}$ de* ***intervalos disjuntos*** *en $\mathcal{V}$ tal que*

$$\mu^*\Big(E \setminus \bigcup_{j=1}^{\infty} I_j\Big) = 0.$$

Prueba. Para cada $n \geq 1$, considere el conjunto $E_n = E \cap (-n, n)$. Puesto que $\mu^*(E_n) < +\infty$, podemos usar el Teorema 9.1.28 para hallar, por cada $k \geq 1$, una colección finita $\{I_1, \ldots, I_{n_k}\}$ de intervalos disjuntos en $\mathcal{V}$ tal que

$$\mu^*\Big(E_k \setminus \bigcup_{j=1}^{n_k} I_j\Big) < \frac{1}{k}.$$

Observe que, para cada $k \in \mathbb{N}$, aquellos miembros de $\mathcal{V}$ que están contenidos en el conjunto abierto $G_k = (-k-1, k+1) \setminus \bigcup_{j=1}^{n_k} I_j$ constituyen un cubrimiento de Vitali del conjunto $E_{k+1} \setminus \bigcup_{j=1}^{n_k} I_j$, y otra vez, por una nueva aplicación del Teorema 9.1.28, existen miembros disjuntos $I_{n_k+1}, \ldots, I_{n_{k+1}}$ de $\mathcal{V}$ que no intersectan a ningún miembro de $\{I_1, \ldots, I_{n_k}\}$ y que satisfacen la desigualdad

$$\mu^*\bigg(\Big(E_{k+1} \setminus \bigcup_{j=1}^{n_k} I_j\Big) \setminus \bigcup_{j=n_k+1}^{n_{k+1}} I_j\bigg) < \frac{1}{k+1}$$

o lo que es lo mismo,

$$\mu^*\Big(E_{k+1} \setminus \bigcup_{j=1}^{n_{k+1}} I_j\Big) < \frac{1}{k+1}. \tag{1}$$

Continuando indefinidamente de este modo se obtiene una sucesión estrictamente creciente $(n_k)_{k=1}^{\infty}$ de enteros positivos y una sucesión $(I_n)_{n=1}^{\infty}$ de intervalos disjuntos en $\mathcal{V}$ tal que

$$\mu^*\Big(E_k \setminus \bigcup_{j=1}^{n_k} I_j\Big) < \frac{1}{k}$$

para todo $k \geq 1$ y, por lo tanto,

$$\mu^*\Big(E_k \setminus \bigcup_{j=1}^{\infty} I_j\Big) \;\leq\; \mu^*\Big(E_k \setminus \bigcup_{j=1}^{n_k} I_j\Big) \;<\; \frac{1}{k}$$

para todo $k \geq 1$. Puesto que $E_k \setminus \bigcup_{j=1}^{\infty} I_j \subseteq E_{k+1} \setminus \bigcup_{j=1}^{\infty} I_j$ resulta que

$$\mu^*\Big(E_k \setminus \bigcup_{j=1}^{\infty} I_j\Big) \;<\; \frac{1}{n} \quad \text{para todo} \quad n > k.$$

De esto se deduce que $\mu^*\big(E_k \setminus \bigcup_{j=1}^{\infty} I_j\big) = 0$ y así, $E_k \setminus \bigcup_{j=1}^{\infty} I_j$ es medible para todo $k \geq 1$. Se sigue de la continuidad de la medida, Teorema 6.3.45, que

$$\mu\Big(E \setminus \bigcup_{j=1}^{\infty} I_j\Big) \;=\; \mu\Big(\bigcup_{k=1}^{\infty}\Big(E_k \setminus \bigcup_{j=1}^{\infty} I_j\Big)\Big) \;=\; \lim_{k\to\infty} \mu\Big(E_k \setminus \bigcup_{j=1}^{\infty} I_j\Big) \;=\; 0.$$

La prueba es completa. ■

Ya hemos visto que *la unión de cualquier* ***colección numerable*** *de conjuntos medibles es medible*. En particular, la unión de cualquier colección numerable de intervalos (de cualquier tipo) es un conjunto medible. Pero, ¿qué se puede decir si la colección es una familia *no-numerable* de conjuntos medibles? Una consecuencia curiosa e interesante del Teorema del Cubrimiento de Vitali es que uniones arbitrarias de intervalos sigue siendo medible.

Corolario 9.1.30. *Sea $\mathcal{I}$* ***cualquier colección de subintervalos*** *de $\mathbb{R}$. Entonces $\bigcup_{J\in\mathcal{I}} J$ es* ***medible****.*

Prueba. Sea $E = \bigcup_{J\in\mathcal{I}} J$ y considere la colección $\mathcal{V}$ formada por todos aquellos intervalos cerrados I para los cuales $I \subseteq J$ para algún $J \in \mathcal{I}$. Entonces $\mathcal{V}$ es un cubrimiento de Vitali de E y se sigue del Teorema del Cubrimiento de Vitali que existe una colección numerable y disjunta $(I_n)_{n=1}^{\infty}$ de $\mathcal{V}$ con la propiedad de que

$$\mu^*\Big(E \setminus \bigcup_{n=1}^{\infty} I_n\Big) \;=\; 0.$$

En particular, $E \setminus \bigcup_{n=1}^{\infty} I_n$ es medible. Finalmente, como $I_n \subseteq E$ para todo $n \geq 1$, tenemos que

$$E \;=\; \Big(E \setminus \bigcup_{n=1}^{\infty} I_n\Big) \cup \bigcup_{n=1}^{\infty} I_n$$

es medible. Fin de la prueba. ■

9.1.4. El Teorema de Densidad de Lebesgue

El objetivo de esta sección es presentar otro resultado de H. Lebesgue demostrado en 1904, conocido como el Teorema de Densidad de Lebesgue, el cual establece que si E es un conjunto medible arbitrario, entonces casi todos sus puntos son puntos de densidad. Ese resultado es una pieza fundamental para la construcción de una nueva topología sobre $\mathbb{R}$, llamada la *topología densidad*, que resulta ser más fina que la topología ordinaria o estándar y que permite identificar a los conjuntos de primera categoría (respecto a la topología densidad) con los conjuntos de medida cero. Dicha topología permite, además, una nueva noción de continuidad, llamada "*aproximadamente continua*", con la siguiente propiedad: si $f : [a, b] \to \mathbb{R}$ es acotada y aproximadamente continua, entonces existe una función diferenciable F tal que $F' = f$. Estos hechos han generado un desarrollo espectacular creándose una nueva área de investigación que fusiona ciertos aspectos de la Teoría de la Medida e Integración con la noción topológica de categoría (según Baire), (véase, por ejemplo, [66], Chapter 14, y las referencias allí citadas).

El siguiente teorema afirma que si un subconjunto E de un intervalo I es "igualmente distribuido" en todo el intervalo, entonces o bien dicho conjunto es un conjunto nulo o es un conjunto de medida absoluta, es decir, el complemento de un conjunto nulo. Por ejemplo, no es posible obtener un conjunto medible $E \subseteq [0, 1]$, para el cual $\mu(E \cap [a, b]) = \mu([a, b])/2$ para todo $0 < a < b < 1$. Siempre habrá intervalos pequeños con una "concentración alta" de puntos de E y otro con una baja concentración. Dicho de otra manera:

Teorema 9.1.31. *Sea $E \in \mathfrak{M}_\mu(\mathbb{R})$ con $0 < \mu(E) < +\infty$. Para cada $\varepsilon > 0$, existe un* ***intervalo abierto*** *$J_\varepsilon \subseteq \mathbb{R}$ tal que*

$$\mu(E \cap J_\varepsilon) \;>\; (1-\varepsilon) \cdot \ell(J_\varepsilon).$$

Prueba. Sea $\mathscr{I}$ la familia de todos los intervalos abiertos y defina

$$\alpha \;=\; \sup\left\{ \frac{\mu(E \cap J)}{\ell(J)} \;:\; J \in \mathscr{I} \right\}.$$

Veamos que $\alpha = 1$. En efecto, es claro que $\alpha \leq 1$. Para demostrar la otra desigualdad, sea $(J_n)_{n=1}^{\infty}$ una sucesión en $\mathscr{I}$ tal que $E \subseteq \bigcup_{n=1}^{\infty} J_n$. Puesto que $E = \bigcup_{n=1}^{\infty}(E \cap J_n)$, entonces

$$\mu(E) \;\leq\; \sum_{n=1}^{\infty} \mu(E \cap J_n) \;=\; \sum_{n=1}^{\infty} \frac{\mu(E \cap J_n)}{\ell(J_n)} \ell(J_n) \;\leq\; \alpha \sum_{n=1}^{\infty} \ell(J_n)$$

y tomando el ínfimo sobre todas las sucesiones $(J_n)_{n=1}^{\infty}$ en $\mathscr{I}$ que cubren a E, se obtiene (observe que E es medible) que

$$\mu(E) \;\leq\; \alpha \cdot \mu(E).$$

Como $0 < \mu(E) < +\infty$, concluimos que $\alpha \geq 1$. Así, $\alpha = 1$. Finalmente, dado $\varepsilon > 0$, escojamos un $J_\varepsilon \in \mathscr{I}$ tal que

$$1 - \varepsilon \;<\; \frac{\mu(E \cap J_\varepsilon)}{\ell(J_\varepsilon)}.$$

Esto termina la prueba. ■

Hemos demostrado que, para cada $\varepsilon \in (0, 1)$, no importa que tan próximo al cero él esté, siempre se puede hallar un *intervalo abierto* en el que la "densidad relativa" de E es al menos $1 - \varepsilon$. A menudo es útil tener todos estos intervalos centrados en un punto determinado llamado punto de densidad.

Definición 9.1.32. *Sea $E \subseteq \mathbb{R}$ un conjunto medible. Un punto $x \in \mathbb{R}$ es llamado un* ***punto de densidad de Lebesgue*** *de E si*

$$\lim_{\varepsilon \to 0^+} \left\{ \frac{\mu(E \cap J)}{\ell(J)} \;:\; x \in J,\; \ell(J) < \varepsilon \right\} \;=\; 1.$$

El límite en la definición anterior será abreviado como:

$$\delta(x,E) \;=\; \operatorname*{lím}_{J\to x} \frac{\mu(E\cap J)}{\ell(J)}.$$

Algunos autores se refieren a $\delta(x,E)$ como la **densidad balanceada** de E en x siempre que x sea el centro de cada intervalo J con $\ell(J) < \varepsilon$.

El siguiente es un resultado más fuerte que el Teorema 9.1.31 conocido con el nombre de Teorema de Densidad de Lebesgue.

Teorema 9.1.33 (Densidad de Lebesgue). *Sea $E \in \mathfrak{M}_\mu(\mathbb{R})$. Entonces existe un **conjunto nulo** $D \subseteq E$ tal que todo punto $x \in E\setminus D$ es un **punto de densidad de Lebesgue** de E.*

Prueba. Para cada $n \geq 1$, defina

$$D_n \;=\; \left\{x \in E\cap(-n,n) \,:\, \operatorname*{lím}_{J\to x} \frac{\mu(E\cap J)}{\ell(J)} < 1-\frac{1}{n}\right\}$$

y sea $D = \bigcup_{n=1}^{\infty} D_n$. Si logramos demostrar que $\mu^*(D_n) = 0$ para todo $n \geq 1$, entonces $\mu(D) = 0$ y la prueba habrá terminado. Suponga, para construir una contradicción, que $\mu(D_N) > 0$ para algún $N \in \mathbb{N}$. Considere ahora la colección $\mathcal{V}$ de todos los intervalos cerrados J que satisfacen la siguiente condición:

$$\textit{existe un punto } \;\; x \in J \;\; \textit{ tal que } \;\; x \in D_N \;\; \textit{ y } \;\; \frac{\mu(E\cap J)}{\ell(J)} < 1-\frac{1}{N}.$$

Claramente $\mathcal{V}$ es un cubrimiento de D_N en el sentido de Vitali. Seleccione, con $\varepsilon = \mu^*(D_N)/N$, un conjunto abierto G tal que

$$D_N \subseteq G \qquad\qquad \text{y} \qquad\qquad \mu(G) \;<\; (1+1/N)\cdot\mu^*(D_N).$$

Se sigue que $\mathcal{V}_G = \{J \in \mathcal{V} \,:\, J \subseteq G\}$ también es un cubrimiento de D_N en el sentido de Vitali y, en consecuencia, gracias al Teorema del Cubrimiento de Vitali podemos elegir una sucesión disjunta $(J_n)_{n=1}^{\infty}$ de intervalos en $\mathcal{V}_G$ tal que

$$\mu^*\Big(D_N \setminus \bigcup_{n=1}^{\infty} J_n\Big) \;=\; 0.$$

De esto se deduce que

$$\begin{aligned}\mu^*(D_N) \;&\leq\; \sum_{n=1}^{\infty} \mu^*(D_N\cap J_n) \;\leq\; \sum_{n=1}^{\infty} (1-1/N)\cdot\ell(J_n)\\ &=\; (1-1/N)\cdot\mu(G) \;<\; (1-1/N)(1+1/N)\cdot\mu^*(D_N),\end{aligned}$$

es decir,

$$\mu^*(D_N) \;<\; (1-1/N^2)\mu^*(D_N).$$

Puesto que estamos asumiendo que $\mu^*(D_N) > 0$, se logra la siguiente contradicción: $1 < 1-1/N^2$. Esta contradicción establece que $\mu^*(D_N) = 0$ y entonces $\mu(D) = 0$. Nótese que

$$D \;=\; \left\{x \in E \,:\, \operatorname*{lím}_{J\to x} \frac{\mu(E\cap J)}{\ell(J)} < 1\right\}.$$

La prueba es completa. ■

Si E posee interior no-vacío, entonces todos sus puntos interiores son obviamente puntos de densidad de E, aunque pueden existir conjuntos con interior vacío con muchos puntos de densidad. Ejemplos de tales conjuntos son los números irracionales o cualquier conjunto tipo-Cantor.

Observe que si E es un conjunto medible, entonces para cada intervalo no-degenerado J,

$$\mu(E \cap J) + \mu(E^c \cap J) = \ell(J)$$

y, por consiguiente,

$$\frac{\mu(E \cap J)}{\ell(J)} + \frac{\mu(E^c \cap J)}{\ell(J)} = 1. \tag{1}$$

Se sigue del Teorema de Densidad de Lebesgue que existe un conjunto nulo $D' \subsetneq E^c$ tal que todos los puntos de $E^c \setminus D'$ son puntos de densidad de Lebesgue, es decir, si $z \in E^c \setminus D'$,

$$\lim_{J \to z} \frac{\mu(E^c \cap J)}{\ell(J)} = 1$$

y, en consecuencia, de (1) se sigue que

$$\lim_{J \to z} \frac{\mu(E \cap J)}{\ell(J)} = 0. \tag{2}$$

Los puntos $z \in \mathbb{R}$ que satisfacen (2) se llaman **puntos de dispersión de Lebesgue** de E. Tenemos así que:

Corolario 9.1.34. *Sea $E \subseteq \mathbb{R}$ un **conjunto medible**.*

*(a) Un punto $x \in \mathbb{R}$ es un **punto de densidad de Lebesgue** de E si, y sólo, si, x es un **punto de dispersión** de E^c.*

*(b) **Casi todos** los puntos de E son **puntos de densidad de Lebesgue** de E y **casi todos** los puntos de E^c son puntos de **dispersión de Lebesgue** de E*

Existen, por supuesto, otras modos distintos de demostrar el Teorema de Densidad de Lebesgue que pueden ser de interés para el lector. Por ejemplo, Claude-Alain Faure [58] usa sólo propiedades sencillas de la medida exterior así como el Lema del Sol Naciente, Lema 9.1.46, página 497. Por su parte, Luděk Zajiček [143] sólo utiliza algunas propiedades simples de las funciones medibles.

En lo que sigue, si E es un conjunto medible, escribiremos

$$E^d = \{x \in \mathbb{R} : x \text{ es punto de densidad de Lebesgue de } E\}.$$

El Teorema de Densidad de Lebesgue establece entonces que

$$\mu(E \triangle E^d) = 0 \quad \text{para cualquier conjunto } E \in \mathfrak{M}_\mu(\mathbb{R}).$$

Un conjunto medible $E \subseteq \mathbb{R}$ satisfaciendo la condición de que *cualquier punto* de E es un punto de densidad de Lebesgue de E, esto es, $E \subseteq E^d$, es llamado un **conjunto d-abierto**. Por ejemplo, cualquier conjunto abierto es d-abierto. Similarmente, conjuntos los cuales son complementos de conjuntos de medida cero son también d-abiertos.

Teorema 9.1.35 (Topología densidad). *La colección τ_D de todos los subconjuntos de $\mathbb{R}$ que son d-abiertos forman una topología sobre $\mathbb{R}$.*

Prueba. (i) Es claro que $\varnothing$ y $\mathbb{R}$ pertenecen a τ_D.

(ii) Sea $(E_\alpha)_{\alpha\in I}$ una familia arbitraria en τ_D y sea $E = \bigcup_{\alpha\in I} E_\alpha$. Para demostrar que $E \in \tau_D$, debemos verificar que él es medible y que cualquier punto de E es un punto de densidad de Lebesgue de E. Veamos primero esto último. Sea $x \in E$. Entonces existe un $\alpha_0 \in I$ tal que $x \in E_{\alpha_0}$. Puesto que x es un punto de densidad de Lebesgue de E_{α_0} se tiene que

$$1 \;=\; \lim_{J\to x} \frac{\mu(E_{\alpha_0}\cap J)}{\ell(J)} \;\leq\; \lim_{J\to x} \frac{\mu(E\cap J)}{\ell(J)} \;\leq\; 1,$$

de modo que x es también un punto de densidad de Lebesgue de E. Para demostrar que E es medible, observe que a cada $x \in E$, le corresponde un $\alpha \in I$ tal que $x \in E_\alpha$ y, en consecuencia, existe un $\varepsilon_{x,\alpha} > 0$ con la siguiente propiedad: si J es cualquier intervalo cerrado tal que $x \in J$ y $\ell(J) < \varepsilon_{x,\alpha}$, entonces

$$\frac{\mu(E_\alpha\cap J)}{\ell(J)} \;>\; \frac{1}{2}.$$

Para cada uno de esos intervalos J, seleccione un conjunto cerrado $F_{\alpha,J}$ tal que

$$F_{\alpha,J} \subseteq E_\alpha \cap J \qquad \text{y} \qquad \mu(F_{\alpha,J}) \;>\; \frac{\ell(J)}{2}.$$

Sea $\mathcal{V}$ la colección de tales $F_{\alpha,J}$. Resulta que $\mathcal{V}$ es un cubrimiento de E en el sentido de Vitali y, por lo tanto, existe, por el Teorema del Cubrimiento de Vitali, una colección numerable y disjunta $(F_n)_{n=1}^\infty$ en $\mathcal{V}$ tal que

$$\mu^*\left(E \setminus \bigcup_{n=1}^\infty F_n\right) \;=\; 0.$$

Esto prueba que $E \setminus \bigcup_{n=1}^\infty F_n$ es medible. Finalmente, como cada F_n es medible y $F_n \subseteq E$, se tiene que

$$E \;=\; \left(E \setminus \bigcup_{n=1}^\infty F_n\right) \cup \bigcup_{n=1}^\infty F_n$$

es medible.

(iii) Sean ahora $E_1, \ldots, E_n$ conjuntos en τ_D y veamos que $E_1 \cap \cdots \cap E_n \in \tau_D$. Claramente $E_1 \cap \cdots \cap E_n$ es medible. Sea $x \in E_1 \cap \cdots \cap E_n$ y sea $\varepsilon > 0$. Como $x \in E_i$ para cada $i = 1, \ldots, n$, escoja un $\delta_i > 0$ tal que si $x \in J$ y $\ell(J) < \delta_i$, entonces

$$\frac{\mu(E_i\cap J)}{\ell(J)} \;>\; 1 - \frac{\varepsilon}{n}.$$

Definiendo $\delta = \min\{\delta_1, \ldots, \delta_n\}$, resulta que si J es cualquier intervalo tal que $x \in J$ y $\ell(J) < \delta$, entonces

$$\mu(E_i^c \cap J) \;\leq\; \frac{\varepsilon}{n}\ell(J) \quad \text{y así,} \quad \mu\big((E_1^c \cup \cdots \cup E_n^c) \cap J\big) \;\leq\; \varepsilon \cdot \ell(J).$$

De esto se sigue que

$$\frac{\mu\big((E_1 \cap \cdots \cap E_n) \cap J\big)}{\ell(J)} \;>\; 1 - \varepsilon.$$

Puesto que $\varepsilon > 0$ fue elegido de modo arbitrario, se concluye que x es un punto de densidad de Lebesgue de $E_1 \cap \cdots \cap E_n$. Esto termina la prueba. ■

A la topología τ_D del teorema anterior se le llama la **topología densidad** de $\mathbb{R}$. Si $f : \mathbb{R} \to \mathbb{R}$ es una función, entonces existen cuatro modos, usando la topología ordinaria o estándar de $\mathbb{R}$, τ_O y la topología densidad sobre el dominio y el rango de f, en las que f puede ser continua. Por ejemplo, la colección

$$C_{DO} = \{f : (\mathbb{R}, \tau_D) \to (\mathbb{R}, \tau_O) : f \text{ es continua}\}$$

representa a las funciones que llamaremos aproximadamente continuas. Estas funciones fueron introducidas por A. Denjoy en 1915. Él demostró que *una función $f : \mathbb{R} \to \mathbb{R}$ es medible si, y sólo si, f es aproximadamente continua en casi todos los puntos de* $\mathbb{R}$.

Recordemos que una función $f : [a, b] \to \mathbb{R}$ es continua en un punto x_0 si, y sólo si, x_0 es un punto interior del conjunto $\{x \in [a, b] : |f(x) - f(x_0)| < \varepsilon\} = f^{-1}\big(V(f(x_0), \varepsilon)\big)$ para cada $\varepsilon > 0$, donde $V(f(x_0), \varepsilon) = (f(x_0) - \varepsilon, f(x_0) + \varepsilon)$. Otro modo, que puede resultar más conveniente, de describir a las funciones aproximadamente continuas con dominio un intervalo compacto y que es muy similar a la noción anterior es la siguiente.

Definición 9.1.36. *Sea $f : [a, b] \to \mathbb{R}$ una función y sea $x_0 \in [a, b]$. Se dice que f es **aproximadamente continua** en x_0, si x_0 es un punto de densidad de Lebesgue del conjunto $f^{-1}\big(V(f(x_0), \varepsilon)\big)$ para cada $\varepsilon > 0$.*

Observe que, gracias al Corolario 9.1.34, decir que f es aproximadamente continua en x_0, significa que x_0 es un punto de dispersión de Lebesgue del conjunto $\{x \in [a, b] : |f(x) - f(x_0)| \geq \varepsilon\}$ para cada $\varepsilon > 0$, es decir,

$$\lim_{J \to x} \frac{\mu\big(\{x \in [a, b] : |f(x) - f(x_0)| \geq \varepsilon\} \cap J\big)}{\ell(J)} = 0,$$

para cada $\varepsilon > 0$.

9.1.5. El Teorema de Diferenciabilidad de Lebesgue

El objetivo de esta sección es demostrar otro extraordinario resultado de H. Lebesgue el cual establece *la diferenciabilidad casi-siempre de cualquier función monótona*. Frigyes Riesz y Béla Sz.-Nagy consideran a dicho teorema como uno de los más notables e interesantes en el Análisis Real. Lebesgue demostró su resultado en el año 1904 para una función continua y monótona, [88], sin embargo, es preciso decirlo, su demostración requiere de una definición más elaborada de diferenciabilidad que la definición habitual de derivada, la cual fue formulada por U. Dini en el año de 1877. La idea original de Dini era poder extender el Teorema Fundamental del Cálculo (para la integral de Riemann) a funciones que no poseían derivadas (véase el Teorema 3.1.45, página 182). Ya para esa momento se conocían ejemplos de funciones continuas definidas sobre un intervalo compacto (y, por consiguiente, Riemann integrables) que no eran diferenciables en ninguno punto de su dominio, tal como la función de Weierstrass $W(x) = \sum_{n=1}^{\infty} a^n \cos(b^n \pi x)$ la cual es continua pero nunca diferenciable, siempre que $0 < a < 1$, $ab > 1 + (3\pi/2)$ y b es un entero impar > 1. Muy pronto, otras patologías de éste tipo comenzaron a surgir en el

Análisis provocando cierta incertidumbre y desconfianza a muchos miembros de la comunidad matemática. Por supuesto, si f es una tal función nunca-diferenciable, entonces el límite

$$\lim_{x\to a} \frac{f(x)-f(c)}{x-c}$$

no tiene sentido para ningún c en el dominio de f. Weierstrass creía, una creencia compartida por otros, que era sólo cuestión de tiempo que alguien encontrara una función continua y monótona que fuese nunca-diferenciable. Por supuesto, Lebesgue demostró que Weierstrass estaba equivocado. Dini había puesto en duda la existencia de una tal función, aunque no la descartaba. En un libro que tendría una gran influencia en la comunidad matemática, Dini [45] demuestra de manera rigurosa un resultado que da condiciones bajo la cual continuidad implica diferenciabilidad: *si f es continua sobre $[a,b]$ y si $f(x)+Ax$ es monótona a trozos para todos los valores de A excepto una cantidad finita, entonces f es diferenciable sobre un conjunto denso de $[a,b]$*. Fue Lebesgue quien termina con la polémica al demostrar que *toda función continua y monótona es diferenciable casi-siempre*. De este modo Lebesgue logra *restaurar la confianza en la armonía del Análisis Matemático*. En 1910, Georg Faber demostró que continuidad no era necesaria: *cualquier función monótona es diferenciable casi-siempre*. Un año después, William H. Young y su esposa Grace C. Young publicaron una demostración independiente de la de Faber donde, de nuevo, muestran que continuidad no es imprescindible. Pero, ¿cómo se aplica el Teorema Fundamental del Cálculo a funciones que son continuas y nunca-diferenciables como la de Weierstrass? Dini, para atacar ese problema, introduce cuatro derivadas por medio de los límites superior e inferior y logra demostrar que si $\mathrm{D}f$ es una cualquiera de sus "derivadas", entonces se cumple

$$\int_a^b \mathrm{D}f(x)\,dx \;=\; f(b)-f(a).$$

Definición 9.1.37. *Sea $f:[a,b]\to\mathbb{R}$ una función. Para cada $x\in[a,b)$ definamos*

$$\mathrm{D}^+f(x) \;=\; \overline{\lim_{y\to x+}}\, \frac{f(y)-f(x)}{y-x} \;=\; \inf_{\delta>0}\sup\left\{\frac{f(y)-f(x)}{y-x} \;:\; x<y<x+\delta\right\}$$

$$\mathrm{D}_+f(x) \;=\; \underline{\lim_{y\to x+}}\, \frac{f(y)-f(x)}{y-x} \;=\; \sup_{\delta>0}\inf\left\{\frac{f(y)-f(x)}{y-x} \;:\; x<y<x+\delta\right\}.$$

*A los números $\mathrm{D}^+f(x)$ y $\mathrm{D}_+f(x)$ se les llaman, respectivamente, las **derivadas superior** e **inferior** de f a la **derecha** de x. Similarmente, las **derivadas superior** e **inferior** de f a la **izquierda** de $x\in(a,b]$ se definen, respectivamente, como*

$$\mathrm{D}^-f(x) \;=\; \overline{\lim_{y\to x-}}\, \frac{f(y)-f(x)}{y-x} \;=\; \inf_{\delta>0}\sup\left\{\frac{f(y)-f(x)}{y-x} \;:\; x-\delta<y<x\right\}$$

$$\mathrm{D}_-f(x) \;=\; \underline{\lim_{y\to x-}}\, \frac{f(y)-f(x)}{y-x} \;=\; \sup_{\delta>0}\inf\left\{\frac{f(y)-f(x)}{y-x} \;:\; x-\delta<y<x\right\}.$$

Observe que las cuatro derivadas

$$\mathrm{D}_+f,\ \mathrm{D}^+f,\ \mathrm{D}_-f,\ \mathrm{D}^-f \,:\, [a,b] \;\to\; \overline{\mathbb{R}},$$

a las que llamaremos las **derivadas de Dini**, *siempre existen* en $\overline{\mathbb{R}}$. Se sigue de las propiedades de los límites superior e inferior que

$$\mathrm{D}_+f(x) \;\leq\; \mathrm{D}^+f(x) \qquad \text{y} \qquad \mathrm{D}_-f(x) \;\leq\; \mathrm{D}^-f(x) \tag{D1}$$

para todo $x \in [a,b]$. Si ocurre que $\mathrm{D}_+f(x) = \mathrm{D}^+f(x)$ para algún $x \in [a,b]$, entonces diremos que f es **diferenciable** o **derivable a la derecha** de x. Similarmente, si $\mathrm{D}_-f(x) = \mathrm{D}^-f(x)$ para algún $x \in [a,b]$, diremos que f es **diferenciable** o **derivable a la izquierda** de x. Observe que f es **diferenciable** o **derivable** en $x \in (a,b)$ con derivada finita si, y sólo si,

$$\mathrm{D}_+f(x) \;=\; \mathrm{D}^+f(x) \;=\; \mathrm{D}_-f(x) \;=\; \mathrm{D}^-f(x) \;\neq\; \pm\infty \tag{D2}$$

Por supuesto, f es diferenciable en los puntos extremos $x = a$ o $x = b$ significa que

$$\mathrm{D}_+f(a) \;=\; \mathrm{D}^+f(a) \qquad \text{o} \qquad \mathrm{D}_-f(b) \;=\; \mathrm{D}^-f(b)$$

y ambos valores son finitos.

Finalmente, las **derivadas superior** e **inferior** de f en $x \in [a,b]$ se definen como:

$$\overline{\mathrm{D}}f(x) \;=\; \lim_{\delta\to 0^+} \sup\left\{\frac{f(y)-f(x)}{y-x} \,:\, 0 < |y-x| < \delta\right\} \;=\; \max\left\{\mathrm{D}^+f(x),\, \mathrm{D}^-f(x)\right\},$$

$$\underline{\mathrm{D}}f(x) \;=\; \lim_{\delta\to 0^+} \inf\left\{\frac{f(y)-f(x)}{y-x} \,:\, 0 < |y-x| < \delta\right\} \;=\; \max\left\{\mathrm{D}_+f(x),\, \mathrm{D}_-f(x)\right\}.$$

Teorema 9.1.38 (Dini). *Sea $f : [a,b] \to \mathbb{R}$ una función **Riemann integrable** y sea*

$$F(x) \;=\; \int_a^x f(t)\,dt \quad \textit{para todo} \;\; x \in [a,b].$$

*Si $\mathrm{D}F$ es una de las cuatro derivadas de Dini de F, entonces $\mathrm{D}F$ es **acotada** y **Riemann integrable** sobre $[a,b]$. Además, se cumple que*

$$\int_a^b \mathrm{D}F(t)\,dt \;=\; F(b) - F(a).$$

Prueba. La prueba la haremos sólo para D^+F. Para cada $x \in [a,b]$, defina

$$l(x) \;=\; \varliminf_{t\to x} f(t) \qquad \text{y} \qquad L(x) \;=\; \varlimsup_{t\to x} f(t).$$

Puesto que f es acotada (por ser Riemann integrable), entonces l y L también lo son. Sea $c \in (a,b)$ y observe que para $x > c$,

$$\Big(\inf_{t\in[c,x]} f(t)\Big)(x-c) \;\leq\; \int_c^x f(t)\,dt \;\leq\; \Big(\sup_{t\in[c,x]} f(t)\Big)(x-c),$$

de donde se sigue que

$$\mathrm{D}^+F(c) \;=\; \overline{\lim_{x\to c+}}\, \frac{1}{x-c}\int_c^x f(t)\,dt \;\leq\; \overline{\lim_{x\to c+}}\Big(\sup_{t\in[c,x]} f(t)\Big) \;\leq\; L(c),$$

$$\mathrm{D}^+F(c) \;=\; \overline{\lim_{x\to c+}}\, \frac{1}{x-c}\int_c^x f(t)\,dt \;\geq\; \overline{\lim_{x\to c+}}\Big(\inf_{t\in[c,x]} f(t)\Big) \;\geq\; l(c).$$

Estas desigualdades nos indican que D^+F también es acotada. Puesto que las desigualdades

$$\overline{\lim_{x\to c}}\, \mathrm{D}^+F(x) \;\leq\; \overline{\lim_{x\to c}}\ \overline{\lim_{y\to x+}}\Big(\sup_{t\in[x,y]} f(t)\Big) \;=\; L(c) \qquad \text{y}$$

$$\underline{\lim_{x\to c}}\, \mathrm{D}^+F(x) \;\geq\; \underline{\lim_{x\to c}}\ \underline{\lim_{y\to x+}}\Big(\inf_{t\in[x,y]} f(t)\Big) \;=\; l(c),$$

también se cumplen, resulta que la oscilación de D^+F en cada $x\in[a,b]$ es menor o igual a la oscilación de f y, en consecuencia, por el Teorema de Riemann-Darboux, Teorema 8.3.7, D^+F es Riemann integrable. Más aun, si la oscilación de f en c es cero, es decir, si f es continua en c, entonces $\mathrm{D}^+F(c) = f(c)$.

Por otro lado, puesto que $f\in\mathcal{R}([a,b])$, resulta que el conjunto de los puntos de continuidad de f, $\mathrm{PC}(f)$, es un G_δ-denso en $[a,b]$ (véase la aplicación (TVL_0) del Teorema de Vitali-Lebesgue) de modo que no importa cuán fina sea la partición que se elija en $[a,b]$, cualquier subintervalo no-degenerado de la misma contiene puntos de $\mathrm{PC}(f)$ y, por lo tanto, en tales puntos f y D^+F coinciden. Esto nos revela que para cualquier partición de $[a,b]$, podemos hallar sumas de Riemann para f y D^+F que son iguales y, por consiguiente, sus integrales son iguales (véase el Teorema 15.3 de [113] si requiere ver más detalles). ■

Las demostraciones que siguen dependen, en gran medida, del Teorema del Cubrimiento de Vitali visto en la sección anterior.

Lema 9.1.39. *Si $f:[a,b]\to\mathbb{R}$ es una función **creciente**, entonces las cuatro derivadas de Dini D_+f, D^+f, D_-f, D^-f son **medibles** y **finitas casi-siempre** sobre $[a,b]$.*

Prueba. Puesto que f es creciente, todas las derivadas de Dini son no-negativas. En primer lugar, vamos a demostrar que D^+f es medible. Los argumentos para las restantes derivadas son similares. Para cada $n\in\mathbb{N}$, defina $g_n:[a,b]\to\overline{\mathbb{R}}$ por

$$u_n(x) \;=\; \sup_{0<h<1/n} \frac{f(x+h)-f(x)}{h}.$$

Puesto que

$$\mathrm{D}^+f(x) \;=\; \lim_{n\to\infty} u_n(x)$$

es suficiente demostrar que cada u_n es medible. Observe que como f es creciente, resulta que f es medible y, en consecuencia, cada cociente $(f(x+h)-f(x))/h$ con $h\in(0,1/n)$ fijo, es medible. Sin embargo, aun no podemos afirmar que cada u_n sea medible. Lo que vamos a hacer es considerar una colección numerable de tales cocientes, vale decir, sea Q_n los racionales en $(0,1/n]$ y defina $v_n:[a,b]\to\overline{\mathbb{R}}$ por

$$v_n(x) \;=\; \sup_{h\in Q_n} \frac{f(x+h)-f(x)}{h}.$$

Por supuesto, cada v_n es medible y por construcción, $v_n \leq u_n$ sobre $[a,b]$. Para establecer la otra desigualdad, fijemos un $x \in [a,b]$ y sea $\varepsilon > 0$. Seleccione, usando la definición de u_n, un $\tau \in (0, 1/n]$ de modo que

$$\frac{f(x+\tau)-f(x)}{\tau} \;>\; u_n(x)-\varepsilon.$$

Para el τ hallado, existe un $h \in Q_n$ con $h \geq \tau$ tal que

$$\frac{1}{h} \;<\; \frac{1}{h} \;+\; \frac{\varepsilon}{|f(x+\tau)|+|f(x)|+1}.$$

Como f es creciente, resulta que $f(x+\tau) \leq f(x+h)$ y, por lo tanto,

$$\frac{f(x+h)-f(x)}{h} \;+\; \varepsilon \;\geq\; \frac{f(x+\tau)-f(x)}{\tau} \;>\; u_n(x) \;-\; \varepsilon.$$

De esto último se sigue que $v_n(x) \geq u_n(x) - 2\varepsilon$. Siendo $\varepsilon > 0$ arbitrarios, se concluye que $u_n(x) = v_n(x)$ para todo $x \in [a,b]$ y, en consecuencia, cada u_n es medible. Por esto, $\mathrm{D}^+ f(x) = \lim_{n\to\infty} u_n(x)$ es medible.

Veamos ahora que $\mathrm{D}^+ f(x) < +\infty$ para casi-todo $x \in [a,b]$. En efecto, sea

$$A \;=\; \big\{x \in (a,b) \;:\; D^+ f(x) = +\infty\big\}$$

y suponga, para construir una contradicción, que $\mu^*(A) = \alpha > 0$. Seleccione un número $M > 0$ de modo que

$$0 \;\leq\; f(b)-f(a) \;<\; \frac{\alpha}{2}M.$$

Sea $x \in A$. Puesto que $\mathrm{D}^+ f(x) = +\infty$, para cada $n \in \mathbb{N}$, existe un $\delta_n > 0$ tal que

$$\sup\left\{\frac{f(y)-f(x)}{y-x} \;:\; x < y < x+\delta_n\right\} \;\geq\; M$$

y ahora seleccione un y_n^x tal que

$$x \;<\; y_n^x \;<\; x+\delta_n \qquad\qquad \text{y} \qquad\qquad \frac{f(y_n^x)-f(x)}{y_n^x - x} \;>\; M \tag{1}$$

Resulta que la colección $\mathcal{V} = \big\{[x, y_n^x] \;:\; x \in A,\; n \in \mathbb{N}\big\}$ es un cubrimiento de Vitali de A y, entonces, por el Teorema del Cubrimiento de Vitali, existe una colección finita y disjunta $\{[x_i, y_i] : i = 1, \ldots, n\}$ de intervalos en $\mathcal{V}$ tal que (tomando $\varepsilon = \alpha/2$)

$$\mu^*\left(A \setminus \bigcup_{i=1}^{n}[x_i, y_i]\right) \;<\; \frac{\alpha}{2}.$$

En particular,

$$\alpha \;=\; \mu^*(A) \;<\; \sum_{i=1}^{n}(y_i - x_i) \;+\; \frac{\alpha}{2}.$$

y así,

$$\sum_{i=1}^{n}(y_i - x_i) \;>\; \frac{\alpha}{2}.$$

Por (1) y esto último se llega a

$$\sum_{i=1}^{n} f(y_i) - f(x_i) \;\geq\; \sum_{i=1}^{n} M \cdot (y_i - x_i) \;>\; \frac{\alpha}{2} \cdot M \;>\; f(b) - f(a)$$

lo cual resulta ser imposible ya que f es creciente. Esta contradicción establece que $\mu^*(A) = 0$ y, por lo tanto, $\mathrm{D}^+ f(x) < +\infty$ para casi-todo $x \in [a,b]$. ■

El siguiente resultado, de importancia fundamental en la teoría de la medida e integración de Lebesgue, se le conoce por el nombre de Teorema de Diferenciabilidad de Lebesgue, pero en honor a la verdad debería llamársele el **Teorema de Diferenciabilidad de Lebesgue-Faber**.

Teorema 9.1.40 (Diferenciabilidad de Lebesgue). *Si $f : [a,b] \to \mathbb{R}$ es una función **monótona** sobre $[a,b]$, entonces f' **existe casi-siempre** sobre $[a,b]$.*

Prueba. Sin perder generalidad, supondremos que f es creciente. El lema anterior nos dice que las cuatro derivadas de Dini son finitas casi-siempre y, además, no-negativas, de modo que si consideramos el conjunto

$$Z \;=\; \big\{ x \in [a,b] \;:\; \mathrm{D}_+ f(x) = \mathrm{D}^+ f(x) = \mathrm{D}_- f(x) = \mathrm{D}^- f(x) = +\infty \big\}.$$

tendremos que $\mu(Z) = 0$. Más aun, puesto que las desigualdades

$$\mathrm{D}_+ f(x) \leq \mathrm{D}^+ f(x) \qquad \text{y} \qquad \mathrm{D}_- f(x) \leq \mathrm{D}^- f(x)$$

se cumplen para todo $x \in [a,b] \setminus Z$, resulta que para demostrar que f es diferenciable en x, sólo debemos verificar, en vista de las dos desigualdades anteriores, que

$$\mathrm{D}^- f(x) \;\leq\; \mathrm{D}_+ f(x) \;\leq\; \mathrm{D}^+ f(x) \;\leq\; \mathrm{D}_- f(x) \quad \text{para todo} \quad x \in [a,b] \setminus Z.$$

Sean entonces

$$\begin{aligned} A \;&=\; \big\{ x \in [a,b] \setminus Z \;:\; \mathrm{D}^- f(x) \;<\; \mathrm{D}_+ f(x) \big\}, \\ B \;&=\; \big\{ x \in [a,b] \setminus Z \;:\; \mathrm{D}_+ f(x) \;<\; \mathrm{D}^+ f(x) \big\}, \qquad \text{y} \\ C \;&=\; \big\{ x \in [a,b] \setminus Z \;:\; \mathrm{D}^+ f(x) \;<\; \mathrm{D}_- f(x) \big\} \end{aligned}$$

y veamos que cada uno de esos conjuntos tiene medida cero. Puesto que la prueba es similar en los tres casos, sólo demostraremos una de ellas. Veamos, por ejemplo, que $\mu(B) = 0$. Suponga, para producir una contradicción, que $\mu(B) > 0$. Observe que

$$B \;=\; \bigcup_{p,q \in \mathbb{Q}} \big\{ x \in [a,b] \setminus Z \;:\; \mathrm{D}_+ f(x) < p < q < \mathrm{D}^+ f(x) \big\}.$$

y escribamos $B_{p,q} = \big\{ x \in [a,b] \setminus Z \;:\; \mathrm{D}_+ f(x) < p < q < \mathrm{D}^+ f(x) \big\}$ para cada par de racionales $p < q$. Puesto que estamos asumiendo que $\mu(B) > 0$, entonces la subaditividad numerable de μ nos dice que

$$0 \;<\; \mu(B) \;\leq\; \sum_{p,q \in \mathbb{Q}} \mu(B_{p,q}),$$

y, en consecuencia, $\mu(B_{p,q}) > 0$ para algún par de racionales $p < q$. Fijemos un par de tales racionales $p < q$ y pongamos $\beta = \mu(B_{p,q})$. Nuestra estrategia consistirá en demostrar que

$$q(\beta - 2\varepsilon) \ < \ p(\beta + \varepsilon) \qquad \text{para cada } \varepsilon > 0,$$

lo que conducirá a que $q < p$ originándose, de este modo, la contradicción buscada. Veamos los detalles. Sea $\varepsilon > 0$ y seleccionemos un conjunto abierto $G \subseteq (a,b)$ tal que

$$B_{p,q} \ \subseteq \ G \qquad \text{y} \qquad \mu(G) \ < \ \mu(B_{p,q}) + \varepsilon \ = \ \beta + \varepsilon. \tag{1}$$

Sea $x \in B_{p,q}$. Puesto que $D_+ f(x) < p$ y G es abierto, podemos elegir, para cada $n \in \mathbb{N}$, un número real y_n^x tal que

$$x \ < \ y_n^x \ < \ x + \frac{1}{n}, \qquad [x, y_n^x] \subseteq G \qquad \text{y} \qquad \frac{f(y_n^x) - f(x)}{y_n^x - x} \ < \ p. \tag{2}$$

La colección $\mathcal{V}_p = \big\{[x, y_n^x] : x \in B_{p,q},\ n \in \mathbb{N}\big\}$ es claramente un cubrimiento de Vitali de $B_{p,q}$. Se sigue del Teorema del Cubrimiento de Vitali que existe una colección finita y disjunta de intervalos cerrados $\{[x_i, y_i] : i = 1, \ldots, m\}$ en $\mathcal{V}_p$ tal que

$$\mu^*\Big(B_{p,q} \setminus \bigcup_{i=1}^{m} [x_i, y_i]\Big) \ < \ \varepsilon. \tag{3}$$

En particular, las desigualdad dadas en (1) y (2) conducen a esta otra desigualdad:

$$\begin{aligned} \sum_{i=1}^{m} f(y_i) - f(x_i) \ &< \ \sum_{i=1}^{m} p(y_i - x_i) \\ &\le \ p \cdot \mu(G) \\ &< \ p \cdot (\beta + \varepsilon). \end{aligned}$$

Sea

$$C \ = \ B_{p,q} \cap \bigcup_{i=1}^{m} [x_i, y_i]$$

y observe que

$$B_{p,q} \ = \ \Big(B_{p,q} \setminus \bigcup_{i=1}^{m} [x_i, y_i]\Big) \cup C$$

Se sigue de (3) que

$$\mu^*(C) \ > \ \beta \ - \ \varepsilon.$$

Veamos ahora cómo se construye un cubrimiento de Vitali de C a partir de los intervalos $[x_1, y_1], \ldots, [x_m, y_m]$. Sea $u \in C$. Puesto que $u \in B_{p,q} \cap [x_i, y_i]$ para algún $i = 1, \ldots, m$, resulta que $D^+ f(u) > q$ por lo que existe, para cada $n \in \mathbb{N}$, un z_n^x tal que

$$u \ < \ z_n^x \ < \ u + \frac{1}{n}, \qquad [u, z_n^u] \subseteq (x_i, y_i) \qquad \text{y} \qquad \frac{f(z_n^u) - f(u)}{z_n^u - u} \ > \ p. \tag{4}$$

La colección $\mathcal{V}_q = \{[u, z_n^u] : u \in C,\ n \in \mathbb{N}\}$ es un cubrimiento de Vitali de C y de nuevo, por el Teorema del Cubrimiento de Vitali, existe una colección finita y disjunta de intervalos cerrados $\{[u_j, v_j] : j = 1, \ldots, n\}$ en $\mathcal{V}_q$ tal que

$$\mu^*\left(C \setminus \bigcup_{j=1}^{n}[u_j, v_j]\right) \ < \ \varepsilon.$$

En particular,

$$\sum_{j=1}^{n}(v_j - u_j) \ > \ \mu^*(C) \ - \ \varepsilon.$$

Usando lo anterior, en combinación con (4), se logra la desigualdad

$$\begin{aligned}\sum_{j=1}^{n} f(v_j) - f(u_j) \ &> \ \sum_{j=1}^{n} q(v_j - u_j)\\ &> \ q \cdot (\mu^*(C) - \varepsilon)\\ &> \ q \cdot (\beta - 2\varepsilon).\end{aligned}$$

Para cada $i = 1, \ldots, m$, considere todos los intervalos $[u_j, v_j]$ en $\mathcal{V}_q$ que están contenidos en (x_i, y_i), es decir, sea

$$J_i \ = \ \Big\{ j \in \mathbb{N} : \ [u_j, v_j] \in \mathcal{V}_q, \ [u_j, v_j] \subseteq (x_i, y_i) \Big\}$$

y observe que, como f es creciente,

$$\sum_{j \in J_i} f(v_j) - f(u_j) \ \leq \ f(y_i) - f(x_i).$$

Esta observación combinada con las anteriores desigualdades conducen a

$$\begin{aligned}q \cdot (\beta - 2\varepsilon) \ &< \ \sum_{j=1}^{n} f(v_j) - f(u_j)\\ &= \ \sum_{i=1}^{m} \sum_{j \in J_i} f(v_j) - f(u_j)\\ &\leq \ \sum_{i=1}^{m} f(y_i) - f(x_i)\\ &< \ p \cdot (\beta + \varepsilon).\end{aligned}$$

Con esto hemos demostrado que

$$q(\beta - 2\varepsilon) \ < \ p(\beta + \varepsilon) \qquad \text{para cada } \varepsilon > 0$$

de donde se deduce que $p > q$. Esta contradicción termina la prueba. ∎

Nota Adicional 9.1.2 Del Teorema de Diferenciabilidad de Lebesgue sabemos que si $f : [a,b] \to \mathbb{R}$ es una función creciente, entonces $f'(x)$ es finito casi-siempre y, además, que $f'(x) \geq 0$ para todo $x \in [a,b]$; es decir, $f'(x) \in [0, +\infty]$ para todo $x \in [a,b]$. En lo que sigue, convenimos en redefinir f' de modo que $f'(x)$ siempre sea finito para cualquier $x \in [a,b]$. Para hacer esto, considere

$$E = \{x \in [a,b] \ : \ f'(x) = +\infty\}.$$

Puesto que $\mu(E) = 0$, resulta que si redefinimos f' exigiendo que $f'(x) = 0$ para todo $x \in E$, tendremos que $f'(x)$ *es finito para todo* $x \in [a,b]$.

Corolario 9.1.41 (Lebesgue). *Si $f \in \mathrm{BV}([a,b])$, entonces $f'(x)$ **existe casi-siempre** sobre $[a,b]$.*

Prueba. Por el Teorema de Jordan $f = f_1 - f_2$, donde f_1 y f_2 son funciones crecientes. El resultado sigue del Teorema de Diferenciabilidad de Lebesgue. ■

Del Corolario 9.1.14 sabemos que $\mathrm{Lip}([a,b]) \subseteq \mathrm{BV}([a,b])$, de donde resulta que

Corolario 9.1.42 (Rademacher). *Si $f : [a,b] \to \mathbb{R}$ es una función **Lipschitz**, entonces $f'(x)$ **existe casi-siempre** sobre $[a,b]$.*

Un hecho bien conocido establece que si $f : [a,b] \to \mathbb{R}$ es una función diferenciable en $[a,b]$ y $f'(x) > 0$ para todo $x \in [a,b]$, entonces f es creciente. El siguiente resultado es una generalización de esa declaración.

Teorema 9.1.43. *Sea $f : [a,b] \to \mathbb{R}$ una función **continua** sobre $[a,b]$. Si $\mathrm{D}^+ f(x) > 0$ para todo $x \in [a,b)$, entonces f es **estrictamente creciente** sobre $[a,b]$.*

Prueba. Es suficiente demostrar que f es creciente. En efecto, una vez demostrado lo anterior, suponga que f no es estrictamente creciente. Esto, por supuesto, significa que existe un subintervalo J donde f es constante y, por consiguiente, en cada punto interior x de J tendríamos que $\mathrm{D}^+ f(x) = 0$ lo que contradice el hecho de que $\mathrm{D}^+ f(x) > 0$.

Para demostrar que f es creciente, vamos a suponer lo contrario y generar una contradicción. Pues bien, como estamos admitiendo que f no es creciente, entonces existen números c y d tales que

$$a \leq c < d \leq b \qquad \text{y} \qquad f(c) > f(d).$$

Sea y cualquier punto en $(f(d), f(c))$. Puesto que f es continua, ella satisface la Propiedad del Valor Intermedio y, en consecuencia, existe un $z \in (c,d)$ tal que $f(z) = y$. Esto último nos indica que el conjunto

$$\{x \in [a,b] \ : \ f(x) = y\} \cap [c,d]$$

es no vacío. Sea

$$x_0 = \sup\{x \in [c,d] \ : \ f(x) = y\}.$$

Claramente, $x_0 \leq d$. Afirmamos que $x_0 < d$. En efecto, si fuese $x_0 = d$, entonces las propiedades del supremo nos permite construir una sucesión $(x_n)_{n=1}^{\infty}$ en $[c,d]$ tal que $x_0 = \lim_{n\to\infty} x_n$ y $f(x_n) = y$ para todo $n \geq 1$. Ahora bien, por la continuidad de f se tiene que $f(d) = f(x_0) =$

$\lim_{n\to\infty} f(x_n) = y$ y esto último es lo que genera una contradicción con el hecho de que $y \in (f(d), f(c))$. Una vez demostrado que $x_0 < d$, resulta que

$$f(x) \ < \ y \quad \text{para todo} \quad x \in (x_0, d]$$

Más aun, como f es continua, el conjunto $\{x \in [a,b] \,:\, f(x) = y\} = f^{-1}(\{y\})$ es cerrado y, por lo tanto, $y = f(x_0)$. De esto se obtiene que

$$\frac{f(x) - f(x_0)}{x - x_0} \ < \ 0 \quad \text{para todo} \quad x \in (x_0, d],$$

y, por consiguiente,

$$\mathrm{D}^+ f(x_0) \ = \ \limsup_{\delta\to 0^+} \left\{ \frac{f(x) - f(x_0)}{x - x_0} \ : \ x_0 < x < x_0 + \delta \right\} \ \leq \ 0$$

contradiciendo la hipótesis de que $\mathrm{D}^+ f(x) > 0$ para todo $x \in [a,b)$. Esto termina la prueba. ■

El siguiente es un resultado que se le atribuye a Guido Fubini pero que no tiene nada que ver con su famoso teorema comúnmente llamado "El Teorema de Fubini" que trata sobre el producto de medidas e integrales multiples.

Teorema 9.1.44 (Diferenciabilidad de Fubini). *Sea $(f_n)_{n=1}^{\infty}$ una sucesión de funciones a valores reales definidas sobre $[a,b]$ tal que*

(a) $(f_n)_{n=1}^{\infty}$ es ***monótona****, y*

(b) $s(x) = \sum_{n=1}^{\infty} f_n(x)$ ***existe*** *y es* ***finita*** *para cada $x \in [a,b]$.*

Entonces

$$s'(x) \ = \ \sum_{n=1}^{\infty} f_n'(x),$$

existe *y es* ***finita*** *para casi-todo $x \in [a,b]$.*

Prueba. Asumiremos que la sucesión $(f_n)_{n=1}^{\infty}$ es creciente. Más aun, considerando la función $f_n - f_n(a)$, podemos suponer (y así lo haremos) que $f_n \geq 0$ para todo $n \geq 1$. Con estas premisas tenemos que $s = \sum_{n=1}^{\infty} f_n$ es una función creciente y no-negativa y, en consecuencia, gracias al Teorema de Diferenciabilidad de Lebesgue tenemos que $s'(x)$ existe y es finita para casi-todo $x \in [a,b]$.

Considere las sumas parciales $s_n = f_1 + \cdots + f_n$ y el resto $r_n = s - s_n$ para cada $n \geq 1$. Como antes, cada una de estas funciones posee una derivada finita c.s. y, por consiguiente, existe un conjunto medible $A \subseteq [a,b]$ tal que $\mu(A) = 0$ y para todo $n \geq 1$, las derivadas

$$s_n'(x) \ = \ f_1'(x) \ + \ \cdots \ + \ f_n'(x) \qquad\qquad \text{y} \qquad\qquad s'(x)$$

existen y son finitas para cada $x \in E = [a,b] \setminus A$. Fijemos $x \in (a,b)$ y elijamos cualquier $h > 0$ para el cual $x + h \in (a,b)$. Se sigue de la igualdad

$$\frac{s(x+h) - s(x)}{h} \ = \ \frac{s_n(x+h) - s_n(x)}{h} \ + \ \frac{r_n(x+h) - r_n(x)}{h}$$

que

$$\frac{s_n(x+h)-s_n(x)}{h} \leq \frac{s(x+h)-s(x)}{h}$$

y esta última desigualdad implica que $s_n'(x) \leq s'(x)$ para todo $x \in E$. Además, como también se cumple que $s_n'(x) \leq s_{n+1}'(x)$, tenemos entonces que

$$s_n'(x) \leq s_{n+1}'(x) \leq s'(x)$$

para $x \in E$ y todo $n \geq 1$. De aquí se sigue que

$$\sum_{m=1}^{\infty} f_m'(x) = \lim_{n\to\infty} s_n'(x) \leq s'(x)$$

para cada $x \in E$. Resta demostrar que $\lim_{n\to\infty} s_n'(x) = s'(x)$ casi-siempre. Puesto que la sucesión $(s_n'(x))_{n=1}^{\infty}$ es creciente para cada $x \in E$, es suficiente probar que $(s_n')_{n=1}^{\infty}$ admite una subsucesión que converge c.s. a s'. Para este fin, sea $(n_k)_{k=1}^{\infty}$ una sucesión estrictamente creciente de enteros positivos tal que

$$\sum_{k=1}^{\infty} \big(s(b) - s_{n_k}(b)\big) < +\infty.$$

Para cada n_k y para cualquier $x \in (a,b)$ se tiene que

$$0 \leq s(x) - s_{n_k}(x) \leq s(b) - s_{n_k}(b)$$

lo cual nos muestra que el lado izquierdo de esta desigualdad está acotado por los términos de una serie convergente y, en consecuencia, la serie $\sum_{k=1}^{\infty} \big(s(x) - s_{n_k}(x)\big)$ converge. Más aun, como los términos de esta serie son funciones monótonas, ella tiene derivada finita c.s., de modo que el argumento utilizado anteriormente revela que $\sum_{n=1}^{\infty} f_n'(x)$ converge c.s. y también que $\sum_{k=1}^{\infty} \big(s'(x) - s_{n_k}'(x)\big)$ converge c.s. De esto último vemos que $\lim_{n\to\infty} s_{n_k}'(x) = s'(x)$ c.s. y termina la prueba. ■

Hemos visto que la función de Cantor es un ejemplo clásico de una función monótona creciente que es singular. Con la ayuda de dicha función y el Teorema de Diferenciación de Fubini se puede demostrar la existencia de funciones singulares estrictamente crecientes.

Ejemplo 9.1.4. *Existe una función* $\widehat{\varphi} : \mathbb{R} \to \mathbb{R}$ *que es* ***estrictamente creciente*** *y* ***singular****.*

Prueba. Extienda la función de Cantor a todo $\mathbb{R}$ definiendo

$$\varphi_\Gamma(x) = \begin{cases} 0 & \text{si } \; x < 0 \\ \varphi_\Gamma(x) & \text{si } \; x \in [0,1] \\ 1 & \text{si } \; x > 1. \end{cases}$$

Sea $((a_n,b_n))_{n=1}^{\infty}$ una enumeración de los intervalos abiertos con extremos racionales contenidos en $[0,1]$ y defina $\widehat{\varphi}$ por

$$\widehat{\varphi}(x) = \sum_{n=1}^{\infty} \frac{1}{2^n}\, \varphi_\Gamma\left(\frac{x-a_n}{b_n-a_n}\right), \quad \text{para todo } \; x \in [0,1].$$

Gracias al M-Test de Weierstrass, $\widehat{\varphi}$ es continua. Para ver que ella es estrictamente creciente, sean $x_1, x_2 \in [0,1]$ con $x_1 < x_2$ y escoja números racionales a_n, b_n tales que $x_1 < a_n < b_n < x_2$. Puesto que $x_1 - a_n < 0$, resulta que

$$0 \ = \ \varphi_\Gamma\left(\frac{x_1 - a_n}{b_n - a_n}\right) \ < \ \varphi_\Gamma\left(\frac{x_2 - a_n}{b_n - a_n}\right)$$

de donde se sigue que $\widehat{\varphi}(x_1) < \widehat{\varphi}(x_2)$. Finalmente, por el Teorema de Diferenciación de Fubini,

$$\widehat{\varphi}'(x) \ = \ \sum_{n=1}^{\infty} \frac{-a_n}{2^n(b_n - a_n)}\, \varphi'_\Gamma\left(\frac{x - a_n}{b_n - a_n}\right) \ = \ 0$$

casi-siempre. ■

Si $f : [a,b] \to \mathbb{R}$ es una función creciente, el Teorema de Diferenciabilidad de Lebesgue nos dice que $E = \{x \in [a,b] \ : \ f'(x) = +\infty\}$ es un conjunto nulo. El siguiente resultado es un recíproco a dicha conclusión.

Ejemplo 9.1.5. *Sea E un subconjunto medible de $[0,1]$ con $\mu(E) = 0$. Entonces existe una* ***función continua*** *y* ***creciente*** *$f : [0,1] \to \mathbb{R}$ tal que f* ***no es derivable*** *en ningún punto de E.*

Prueba. Puesto que $\mu(E) = 0$ seleccione, por cada $n \in \mathbb{N}$, una colección numerable de intervalos abiertos $(I_k^n)_{k=1}^{\infty}$ tal que

$$E \ \subseteq \ \bigcup_{k=1}^{\infty} I_k^n \qquad \text{y} \qquad \sum_{k=1}^{\infty} \ell(I_k^n) \ < \ \frac{1}{2^n}. \tag{1}$$

Sea $G_n = \{I_1^n, I_2^n, \ldots\}$ y considere

$$G \ = \ \bigcup_{n=1}^{\infty} G_n \ = \ \big\{I_n^k \ : \ k,\, n \in \mathbb{N}\big\}.$$

Para simplificar la presentación, denote por $\{I_1, I_2, \ldots\}$ la lista de los intervalos de G. Se verifica fácilmente que $E \subseteq \bigcup_{n=1}^{\infty} I_n$, $\mu(\bigcup_{n=1}^{\infty} I_n) < 1$ y además, de (1), se tiene que:

Sea $x \in E$. Para cada entorno V_x de x y cada $n \in \mathbb{N}$, existen enteros positivos $i_1, \ldots, i_n$ con $i_1 < \cdots < i_n$ tal que

$$x \ \in \ I_{i_k} \ \subseteq \ V_x \quad \text{para cada } k \in \{1, \ldots, n\} \tag{2}$$

Para cada $n \in \mathbb{N}$, pongamos $I_n = (a_n, b_n)$ y defina la función $f_n : [0,1] \to \mathbb{R}$ por

$$f_n(x) \ = \ \begin{cases} 0 & \text{si} \quad x < a_n, \\ x - a_n & \text{si} \quad a_n \leq x \leq b_n, \\ b_n - a_n & \text{si} \quad x > b_n. \end{cases}$$

Claramente f_n es continua, creciente y se cumple que

$$0 \ \leq \ f_n(x)\; b_n - a_n \ = \ \ell(I_n) \tag{3}$$

para todo $x \in [a, b]$. Finalmente, definamos $f : [a, b] \to \mathbb{R}$ por

$$f(x) = \sum_{n=1}^{\infty} f_n(x) \qquad \text{para todo} \quad x \in [a, b].$$

Veamos que f posee las propiedades deseadas.

(1) f *es continua*. En efecto, puesto que $\sum_{n=1}^{\infty} \ell(I_n) < 1$, se sigue (3) y el M-Test de Weierstrass que la serie $\sum_{n=1}^{\infty} f_n$ converge uniformemente a f y, así, por el Teorema 3.1.40, f es continua.

(2) f *es creciente*. Esto es consecuencia del hecho de cada f_n lo es.

(3) f *no es diferenciable en ningún punto de* E. Sea $x \in E$ y sea V_x un entorno de x. Fijemos un $n \in \mathbb{N}$. Se sigue de (2) que existen $i_1, \ldots, i_n$ con $i_1 < \cdots < i_n$ tal que

$$x \in I_{i_k} \subseteq V_x \quad \text{para cada } k \in \{1, \ldots, n\}.$$

Tome cualquier $y \in I_{i_1} \cap \cdots \cap I_{i_n}$ con $y \neq x$. Entonces

$$\frac{f(y) - f(x)}{y - x} \geq \sum_{k=1}^{n} \frac{f_{i_k}(y) - f_{i_k}(x)}{y - x} = \sum_{k=1}^{n} 1 = n.$$

Como la elección de n fue hecha de manera arbitraria, se concluye que f no es derivable en x y termina la prueba. ■

En el siguiente ejemplo, el cual es una aplicación del Teorema de Diferenciabilidad de Lebesgue, supondremos que todo número racional p/q está escrito en forma irreducible, es decir, p y q son primos relativos entre sí. Como antes, denotaremos a los números irracionales por $\mathbb{I}$. Un simple hecho acerca de aproximación Diofántica establece que:

Existe una cantidad no-numerable de números reales x que no admiten aproximaciones racionales $x \approx \frac{p}{q}$ satisfaciendo $\left| x - \frac{p}{q} \right| < \frac{1}{4q^3}$.

Ejemplo 9.1.6. *Sea $(a_n)_{n=1}^{\infty}$ una sucesión en $\mathbb{R}$ con $a_n > 0$ para todo $n \in \mathbb{N}$ y suponga que la serie $\sum_{n=1}^{\infty} \frac{1}{a_n} < +\infty$. Si*

$$E = \left\{ x \in [0, 1] \cap \mathbb{I} : \text{ existe } p \text{ t.q. } \left| x - \frac{p}{q} \right| < \frac{1}{q \cdot a_q} \;\; \textit{para infinitos} \;\; q > 1 \right\},$$

entonces $\mu(E) = 0$.

Prueba. Defina la función $f : [0, 1] \to \mathbb{R}$ por

$$f(x) = \begin{cases} \dfrac{1}{q \cdot a_q} & \text{si } \; x = \dfrac{p}{q}, \\ 0 & \text{si } \; x \in [0, 1] \cap \mathbb{I}. \end{cases}$$

Veamos, en primer lugar, que $f \in \mathrm{BV}([0, 1])$. Observe que para cada entero $q > 1$, existen a lo sumo q racionales en forma reducida p/q pertenecientes al intervalo $[0, 1]$. Fijemos un $q \in \mathbb{N}$

con $q > 1$ y sea $\mathcal{P} = \{x_0, x_1, \ldots, x_q\}$ cualquier partición de $[0,1]$. Si $x_i = p/q$ para algún $i = 1, \ldots, q$ tendremos que

$$\sum_{i=1}^{q} |f(x_i)| \le q \cdot \frac{1}{q \cdot a_q} = \frac{1}{a_q}.$$

En cualquier caso,

$$\sum_{i=1}^{q} |f(x_i) - f(x_{i-1})| \le 2 \cdot \sum_{i=1}^{q} |f(x_i)| \le 2 \cdot \frac{1}{a_q} < 2 \cdot \sum_{n=1}^{\infty} \frac{1}{a_n} < \infty.$$

Esto prueba que $f \in \mathrm{BV}([0,1])$. Por el Teorema de Diferenciabilidad de Lebesgue, $f'(x)$ existe casi-siempre, pero como $\mu([0,1] \cap \mathbb{I}) = 1$, lo anterior significa que $f'(x)$ existe para casi-todos los números irracionales. Sea

$$F = \{x \in [0,1] \cap \mathbb{I} : f'(x) \text{ existe.}\}.$$

Tenemos entonces que $\mu([0,1] \setminus F) = 0$. Veamos que $E \subseteq [0,1] \setminus F$, o lo que es lo mismo, vamos a demostrar que si $x \notin [0,1] \setminus F$, entonces $x \notin E$. Suponga que $x \notin [0,1] \setminus F$. Entonces $x \in F$ y como $[0,1] \cap \mathbb{I}$ es denso en $[0,1]$, existe una sucesión $(x_n)_{n=1}^{\infty}$ en $[0,1] \cap \mathbb{I}$ convergiendo a x con $x_n \neq x$ para todo $n \in \mathbb{N}$. Puesto que $f'(x)$ existe, se tiene, en particular, que

$$f'(x) = \lim_{x_n \to x} \frac{f(x_n) - f(x)}{x_n - x} = 0.$$

De aquí se sigue que, dado $\varepsilon = 1$, existe un $\delta > 0$ tal que

$$\left| \frac{f(y) - f(x)}{y - x} \right| < 1 \qquad \text{para todo } y \in (x - \delta, x + \delta) \cap [0,1]. \tag{1}$$

Observe que los y en $(x - \delta, x + \delta) \cap [0,1]$ que son relevantes son los racionales. Por esto, si $y = p/q \in (x - \delta, x + \delta) \cap [0,1]$, entonces (1) se transforma en

$$\frac{|f(p/q)|}{\left| \frac{p}{q} - x \right|} = \frac{1}{q \cdot a_q \left| \frac{p}{q} - x \right|} < 1,$$

es decir,

$$\frac{1}{q \cdot a_q} < \left| \frac{p}{q} - x \right| \qquad \text{siempre que} \quad \left| \frac{p}{q} - x \right| < \delta. \tag{2}$$

Suponga, para producir una contradicción, que $x \in E$. Esto significa que existen infinitos $q > 1$ tales que

$$\left| \frac{p}{q} - x \right| < \frac{1}{q \cdot a_q} \qquad (3).$$

Para tales q se obtiene, usando (2), que

$$\left| \frac{p}{q} - x \right| \ge \delta$$

y, entonces, por (3)

$$\frac{1}{q \cdot a_q} > \left|\frac{p}{q} - x\right| \geq \delta > 0$$

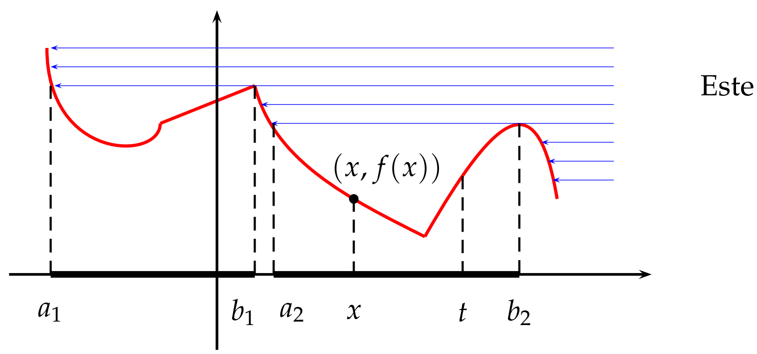
para infinitos $q > 1$. Esto último es lo que genera la contradicción pues en este caso se tiene que

$$0 = \lim_{q\to\infty} \frac{1}{q \cdot a_q} \geq \delta > 0.$$

Hemos demostrado que $E \subseteq [0,1] \setminus F$ y, por lo tanto, $\mu(E) = 0$. La prueba es completa. ■

Nota Adicional 9.1.3 Existen otras maneras diferentes de demostrar el Teorema de Diferenciabilidad de Lebesgue sin usar el Teorema del Cubrimiento de Vitali. Por ejemplo, en 1932 F. Riesz [112] utilizó un resultado ya demostrado por él, llamado el Lema del Sol Naciente, para dar una nueva prueba del Teorema de Diferenciabilidad de Lebesgue.

Definición 9.1.45. *Sea $f : [a,b] \to \mathbb{R}$ una función continua y sea $x \in [a,b]$. Diremos que x es **punto de sombra** de f si existe un $t \in [a,b]$ con $t > x$ tal que $f(t) > f(x)$.*



La razón de ser de ésta definición se indica en el gráfico anterior: las rectas paralelas son los rayos del sol que salen por el Este (estamos mirando hacia al Norte). Observe que x es un punto de sombra de f si $(x, f(x))$ está en la sombra proyectada por los rayos solares.

Lema 9.1.46 (Lema del Sol Naciente). *Sea $f : [a,b] \to \mathbb{R}$ una función **continua** sobre $[a,b]$. El conjunto G de todos los puntos de sombra de f:*

$$G = \{x \in (a,b) : f(t) > f(x) \text{ para algún } t > x\}$$

*es **abierto**. Más aun, si (a_n, b_n) es una componente de G, entonces $f(a_n) \leq f(b_n)$.*

Prueba. Sea $x \in (a,b)$ un punto de sombra de f. Entonces existe $t \in (a,b)$ con $t > x$, tal que $f(t) > f(x)$. Usemos la continuidad de f para hallar un pequeño entorno V_t de t no conteniendo a x tal que $f(z) > f(x)$ para cada $z \in V_t$. Resulta que $(z_0, t) \subseteq G$, donde $z_0 = \inf V_t$ y, por consiguiente, G es un conjunto abierto. Por el Teorema 2.2.4 sabemos que G se puede escribir en la forma

$$G = \bigcup_{n=1}^{\infty} (a_n, b_n)$$

donde los intervalos $((a_n, b_n))_{n=1}^{\infty}$ son disjuntos dos a dos. Queda por demostrar que $f(a_n) \leq f(b_n)$ para cualquier $n \geq 1$. Para ver esto, fijemos un $n \geq 1$ y para cada $x \in (a_n, b_n)$, defina

$$A_x = \{y \in [x, b_n] : f(y) \geq f(x)\}.$$

Puesto que $x \in A_x$, resulta que $A_x \neq \varnothing$ y, además, está acotado por b_n. Se sigue del Axioma del Supremo que A_x posee un supremo, digamos, $t = \sup A_x$. Es claro que $t \in A_x$ y, por lo tanto, $f(x) \leq f(t)$. Afirmamos que $t = b_n$. En efecto, suponga, para obtener una contradicción, que $t < b_n$. Observe, en primer lugar, que $f(b_n) < f(x)$. Además, como $t < b_n$, resulta que $t \in (a_n, b_n)$ es un punto de sombra de f y, en consecuencia, existe un $z > t$ tal que $f(z) > f(t)$. De esto se sigue que

$$f(z) > f(t) \geq f(x) > f(b_n). \tag{1}$$

Puesto que t es el supremo de los $y \in [x, b_n]$ para los cuales $f(y) \geq f(x)$, entonces z debe ser mayor que b_n, lo que combinado con (1), nos muestra que b_n debe ser un punto de sombra de f, lo cual es imposible. Esta contradicción nos revela que $t = b_n = \sup A_x$ y, por lo tanto, $f(b_n) \geq f(x)$. Finalmente, seleccione una sucesión decreciente $(x_k)_{k=1}^{\infty}$ en (a_n, b_n) que converge a a_n. Por lo anterior se tiene que $f(x_k) \leq f(b_n)$ para todo $k \geq 1$ y, entonces, la continuidad de f nos asegura que $f(a_n) = \lim_{k\to\infty} f(x_k) \leq f(b_n)$ y termina la prueba. ■

Pare ver otros métodos de probar el Teorema de Diferenciabilidad de Lebesgue, véase, por ejemplo, John W. Hagood [69], Giorgio Letta [91], Michael W. Botsko [19], etc.

9.1.6. Funciones Absolutamente Continuas

La existencia de una derivada acotada que no es Riemann integrable conduce a reconocer que el problema de recuperar cualquier función diferenciable de su derivada no se puede resolver por medio del Teorema Fundamental del Cálculo para la integral de Riemann y, de hecho, tampoco lo resuelve de modo general la integral de Lebesgue, es decir, la idílica fórmula de Newton-Leibniz no es, en general, válida ni para la integral de Riemann ni aun para la de Lebesgue. Con el tiempo, los matemáticos se dieron cuenta que había que conformarse, en general, con una desigualdad en lugar de una igualdad si se lograba debilitar un poco las condiciones del problema: se trata, en consecuencia, de recuperar funciones que tengan derivadas sólo casi-siempre y, además, que cumplan ciertas hipótesis adicionales. Para ello, Lebesgue considera una clase muy especial y, por demás extraña, de funciones continuas llamadas funciones *absolutamente continuas* que tienen la particularidad de satisfacer el Teorema Fundamental del Cálculo para la integral de Lebesgue. Tales funciones fueron descubiertas por Axel Harnack en 1884, pero quien las nombra de esa manera fue G. Vitali en 1905. El objetivo se esta sección es, entonces, estudiar las propiedades de las funciones absolutamente continuas.

Definición 9.1.47. *Una función $f : [a, b] \to \mathbb{R}$ es **absolutamente continua** sobre $[a, b]$ si para cada $\varepsilon > 0$, se puede determinar un $\delta > 0$ con la siguiente propiedad: para cualquier colección finita de intervalos no-superpuestos $\{[a_i, b_i] : i = 1, \ldots, n\}$ incluidos en $[a, b]$ que satisfaga $\sum_{i=1}^{n}(b_i - a_i) < \delta$, se cumple que*

$$\sum_{i=1}^{n} |f(b_i) - f(a_i)| < \varepsilon.$$

Si hacemos $E=\bigcup_{i=1}^{n}[a_i,b_i]$, entonces podemos reemplazar la expresión $\sum_{i=1}^{n}(b_i-a_i)<\delta$, en la definición anterior, por $\mu(E)<\delta$. Observe también que podemos tomar, en dicha definición, en lugar de una colección finita de intervalos cerrados no-superpuestos, una colección infinita de intervalos abiertos y disjuntos dos a dos (véase el Ejercicio 9.3 $3(c)$, página 523). Similarmente, podemos reemplazar el intervalo $[a,b]$ por $\mathbb{R}$.

Denote por $\mathrm{AC}([a,b])$ *el conjunto de todas las funciones* $f:[a,b]\to\mathbb{R}$ *que son* ***absolutamente continuas*** *sobre* $[a,b]$. Es claro que

$$\mathrm{AC}([a,b])\subseteq C([a,b]).$$

Es un ejercicio sencillo verificar que:

(1) $\mathrm{AC}([a,b])$ es un **espacio vectorial** sobre $\mathbb{R}$.

(2) $|f|\in\mathrm{AC}([a,b])$, siempre que $f\in\mathrm{AC}([a,b])$.

Más aun,

(3) $f\cdot g\in\mathrm{AC}([a,b])$, siempre que $f,g\in\mathrm{AC}([a,b])$.

Para ver esto último, observe que

$$\begin{aligned}|f(x)g(x)-f(y)g(y)| &= |f(x)\big(g(x)-g(y)\big)+g(y)\big(f(x)-f(y)\big)|\\ &\le \|f\|_\infty|g(x)-g(y)|+\|g\|_\infty|f(x)-f(y)|\end{aligned}$$

de donde se sigue que fg es absolutamente continua sobre $[a,b]$. En efecto, sea $\varepsilon>0$ y usemos el hecho de que $f,g\in\mathrm{AC}([a,b])$ para seleccionar un $\delta>0$ tal que si $\{[a_i,b_i]:i=1,\ldots,n\}$ es una colección arbitraria de intervalos no-superpuestos en $[a,b]$ satisfaciendo $\sum_{i=1}^{n}(b_i-a_i)<\delta$, entonces se cumple que

$$\sum_{i=1}^{n}|f(b_i)-f(a_i)|<\frac{\varepsilon}{2(\|g\|_\infty+1)}\quad\text{y}\quad\sum_{i=1}^{n}|g(b_i)-g(a_i)|<\frac{\varepsilon}{2(\|f\|_\infty+1)}.$$

Finalmente,

$$\sum_{i=1}^{n}|f(b_i)g(b_i)-f(a_i)g(a_i)|\le\|f\|_\infty\sum_{i=1}^{n}|g(b_i)-g(a_i)|+\|g\|_\infty\sum_{i=1}^{n}|f(b_i)-f(a_i)|<\varepsilon.$$

Teorema 9.1.48. *Si* $f:[a,b]\to\mathbb{R}$ *una función* ***Lipschitz***, *entonces* $f\in\mathrm{AC}([a,b])$.

Prueba. Suponga que f es Lipschitz con constante Lipschitz M. Para demostrar que f es absolutamente continua en $[a,b]$, sea $\varepsilon>0$ y seleccione un $\delta>0$ de modo que $M\delta<\varepsilon$. Si $\{[a_i,b_i]:i=1,\ldots,n\}$ es cualquier colección finita de intervalos no-superpuestos en $[a,b]$ satisfaciendo $\sum_{i=1}^{n}(b_i-a_i)<\delta$, entonces

$$\sum_{i=1}^{n}|f(b_i)-f(a_i)|\le\sum_{i=1}^{n}M(b_i-a_i)=M\sum_{i=1}^{n}(b_i-a_i)<M\delta<\varepsilon.$$

Esto prueba que $f\in\mathrm{AC}([a,b])$. ■

Ejemplo 9.1.7. (a) *Se sigue del Segundo Teorema Fundamental del Cálculo que para cada* $f \in \mathcal{R}([a,b])$*, su integral indefinida*

$$F(x) \;=\; \int_a^x f\,dx, \quad x \in [a,b]$$

es Lipschitz y entonces, absolutamente continua, gracias al resultado anterior.

(b) *La función* $f : [0,1] \to \mathbb{R}$ *definida por*

$$f(x) \;=\; \begin{cases} x^2 \operatorname{sen}(1/x) & si \quad x \in (0,1] \\ 0 & si \quad x = 0 \end{cases}$$

posee, como ya hemos visto, derivada acotada en $[0,1]$ *y, en consecuencia, gracias al Teorema 9.1.6, página 457, tenemos que* f *es* ***Lipschitz****. En particular,* $f \in \mathrm{AC}([0,1])$*, por el Teorema 9.1.48.*

(c) *En general, para cada* $\alpha, \beta \in (0,+\infty)$*, la función* $f_{\alpha\beta} : [0,1] \to \mathbb{R}$ *definida por*

$$f_{\alpha\beta}(x) \;=\; \begin{cases} x^\alpha\, sen(1/x^\beta) & si \quad x \in (0,1] \\ 0 & si \quad x = 0 \end{cases}$$

es ***absolutamente continua*** *sobre* $[0,1]$ *si* $\alpha > \beta$*, pero no lo es si* $\alpha \leq \beta$*.*

Prueba. Si $\alpha > \beta$, entonces

$$f_{\alpha\beta}(x) \;=\; \int_0^x \left(\alpha\, t^{\alpha-1} \operatorname{sen}\!\Big(\frac{1}{t^\beta}\Big) \;-\; \beta\, t^{\alpha-\beta-1} \cos\Big(\frac{1}{t^\beta}\Big)\right) dt$$

y el integrando es integrable sobre $[0,1]$. Se sigue de la parte (a) que f es absolutamente continua. Para demostrar que $f_{\alpha\beta}$ no es absolutamente continua cuando $\alpha \leq \beta$, es suficiente verificar que ella no es de variación acotada sobre $[0,1]$. Para ver esto último, considere la partición $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ en $[0,1]$ tal que

$$0 \;<\; \frac{1}{(n\pi + \frac{\pi}{2})^{1/\beta}} \;<\; \frac{1}{((n-1)\pi + \frac{\pi}{2})^{1/\beta}} \;<\; \cdots \;<\; \frac{1}{(\pi + \frac{\pi}{2})^{1/\beta}} \;<\; 1.$$

La motivación para tal elección es que

$$\operatorname{sen}(1/x_m^\beta) \;=\; \begin{cases} \;\;1 & \text{si} \quad m \quad \text{es par} \\ -1 & \text{si} \quad m \quad \text{es impar} \end{cases}$$

y, en consecuencia,

$$f_{\alpha\beta}(x_m) \;=\; \begin{cases} \;\;x_m^\alpha & \text{si} \quad m \quad \text{es par} \\ -x_m^\alpha & \text{si} \quad m \quad \text{es impar.} \end{cases}$$

Por esto,

$$\begin{aligned}\sum_{m=1}^{n}|f_{\alpha\beta}(x_m)-f_{\alpha\beta}(x_{m-1})| &= \sum_{m=1}^{n}\left|(-1)^{\alpha}\left(x_m^{\alpha}+x_{m-1}^{\alpha}\right)\right| \\ &= \sum_{m=1}^{n}\left(x_m^{\alpha}+x_{m-1}^{\alpha}\right) \\ &= 2\sum_{m=1}^{n-1}x_m^{\alpha} + x_n + x_0 \\ &\geq \sum_{m=1}^{n-1}x_m^{\alpha} = \sum_{m=1}^{n-1}\left(m\pi+\frac{\pi}{2}\right)^{-\alpha/\beta}.\end{aligned}$$

De aquí se sigue que

$$\lim_{n\to\infty}\sum_{m=1}^{n-1}\left(m\pi+\frac{\pi}{2}\right)^{-\alpha/\beta} = \infty \quad\Leftrightarrow\quad \alpha\leq\beta.$$

Esto prueba que $V(f_{\alpha\beta},[0,1])=+\infty$ si, y sólo si, $\alpha\leq\beta$. ■

Otra manera de verificar directamente que $f_{\alpha\beta}$ es de variación acotada cuando $\alpha>\beta$ es a través de la obtención de una apropiada estimación de su derivada. En efecto, observe que como

$$f'_{\alpha\beta}(x) = \alpha x^{\alpha-1}\mathrm{sen}(1/x^{\beta}) - \beta x^{\alpha-\beta-1}\cos(1/x^{\beta}),$$

entonces

$$|f'_{\alpha\beta}(x)| \leq \alpha x^{\alpha-1} + \beta x^{\alpha-\beta-1},$$

de modo que si definimos $\delta=\min\{\alpha,\alpha-\beta\}$, resulta que $\delta>0$ y entonces, para cualquier partición $\mathcal{P}=\{x_0,x_1,\ldots,x_n\}$ de $[0,1]$, el Teorema Fundamental del Cálculo nos muestra que

$$\begin{aligned}\sum_{i=1}^{n}|f_{\alpha\beta}(x_i)-f_{\alpha\beta}(x_{i-1})| &\leq \sum_{i=1}^{n}\int_{x_{i-1}}^{x_i}|f'_{\alpha\beta}(t)|\,dt \\ &\leq (|\alpha|+|\beta|)\int_0^1 t^{\delta-1}\,dt \\ &= \frac{|\alpha|+|\beta|}{\delta} < +\infty.\end{aligned}$$

Esto prueba que $f_{\alpha\beta}$ es de variación acotada cuando $\alpha>\beta$.

Teorema 9.1.49. *Si $f:[a,b]\to\mathbb{R}$ es una función **absolutamente continua** sobre $[a,b]$, entonces $f\in\mathrm{BV}([a,b])$. En particular, $f'(x)$ **existe** para casi-todo $x\in[a,b]$.*

Prueba. Puesto que f es absolutamente continua, existe un $\delta>0$, correspondiente a $\varepsilon=1$, tal que

$$\sum_{i=1}^{n}|f(b_i)-f(a_i)| < 1 \tag{1}$$

para cualquier colección finita de intervalos no-superpuestos $\{[a_i, b_i] : i = 1, \ldots, n\}$ en $[a, b]$ que satisfaga $\sum_{i=1}^{n}(b_i - a_i) < \delta$. Seleccione el entero más pequeño, digamos N, de modo que $N > (b-a)/\delta$, y defina

$$x_k \;=\; a + k \cdot \frac{b-a}{N} \qquad \text{para } k = 0, 1, \ldots, N.$$

Puesto que $\mathcal{P} = \{x_0, x_1, \ldots, x_N\}$ es una partición de $[a, b]$ para la cual se cumple que $x_k - x_{k-1} = (b-a)/N < \delta$ para todo $k = 1, \ldots, N$, resulta entonces de (1) que $|f(x_k) - f(x_{k-1})| < 1$ y, por lo tanto,

$$V(f, \mathcal{P}) \;=\; \sum_{k=1}^{N} |f(x_k) - f(x_{k-1})| \;<\; N.$$

De esto se sigue que $f \in \mathrm{BV}([a, b])$. La última parte es consecuencia del Corolario 9.1.41. ■

Los últimos dos teoremas nos indican que

$$\mathrm{Lip}([a, b]) \;\subseteq \mathrm{AC}([a, b]) \;\subseteq\; \mathrm{BV}([a, b]).$$

Sin embargo, es fácil establecer que ambas inclusiones son propias. Para la primera, la función $f(x) = \sqrt{x}$ para todo $x \in [0, 1]$ es un ejemplo de una función absolutamente continua que no es Lipschitz, mientras que la función de Cantor es un ejemplo de una función de variación acotada (de hecho, uniformemente continua) que no es absolutamente continua (véase el Ejemplo 9.1.8, página 512). Las funciones Lipschitz tienen, con respecto a las funciones absolutamente continuas y a las de variación acotada, la siguiente virtud: *su derivada siempre es acotada*. Veamos por qué esto es así. Sea $f \in \mathrm{Lip}([a, b])$ con constante de Lipschitz M. Como f es de variación acotada, resulta que $f'(x)$ existe para casi-todo $x \in [a, b]$. Sin embargo, si x es cualquier punto en $[a, b]$, resulta que

$$|f'(x)| \;=\; \lim_{y \to x} \frac{|f(x) - f(y)|}{|x - y|} \;\leq\; M,$$

es decir, el conjunto $E_\infty = \{x \in [a, b] : |f'(x)| = +\infty\} = \varnothing$ y, por lo tanto, f' está acotada sobre $[a, b]$.

El Teorema de Jordan nos dice que cada función $f \in \mathrm{BV}([a, b])$ se puede representar en la forma $f = f_1 - f_2$ donde f_1 y f_2 son funciones no-decrecientes. En particular, toda función f absolutamente continua también se representa, por el Teorema 9.1.49, como la diferencia de dos funciones no-decrecientes. El siguiente resulta dice algo más preciso.

Teorema 9.1.50. *Si $f : [a, b] \to \mathbb{R}$ es una* ***absolutamente continua*** *sobre $[a, b]$, entonces f se puede expresar como la diferencia de dos* ***funciones crecientes absolutamente continuas****.*

Prueba. Por el Lema 9.1.20, página 467, sabemos que la función $V_f : [a, b] \to \mathbb{R}$ definida por

$$V_f(x) \;=\; V(f, [a, x]) \qquad \text{para todo} \quad x \in [a, b]$$

es monótona creciente. Sea $\varepsilon > 0$ y usemos el hecho de que f es absolutamente continua para hallar un $\delta > 0$ de modo que

$$\sum_{i=1}^{n} |f(b_i) - f(a_i)| \;<\; \varepsilon$$

para cualquier familia finita y casi-disjunta, digamos $\mathcal{P} = \{[a_1, b_1], \ldots, [a_n, b_n]\}$, que satisfaga la desigualdad $\sum_{i=1}^{n}(b_i - a_i) < \delta$. Sea $E = \bigcup_{i=1}^{n}[a_i, b_i]$. Si logramos demostrar que

$$\sum_{i=1}^{n} |V_f(b_i) - V_f(a_i)| \ < \ \varepsilon$$

entonces tendremos que V_f es absolutamente continua. Para ver esto, escoja, por cada $i \in \{1, \ldots, n\}$, una partición finita $\mathcal{P}_i = \{x_{i1}, \ldots, x_{ik_i}\}$ del intervalo $[a_i, b_i]$ y observe que si E_0 es la unión de todos los subintervalos de cada partición $\mathcal{P}_i$, entonces $\mu(E_0) = \mu(E) < \delta$. Puesto que f es absolutamente continua, resulta que

$$\sum_{i=1}^{n} V(f, [a_i, b_i]) \ < \ \varepsilon$$

Sin embargo, como $V_f(b_i) - V_f(a_i) = V(f, [a_i, b_i])$ y las particiones $\mathcal{P}_i$ fueron escogidas de modo arbitrarias, se sigue que $\sum_{i=1}^{n} |V_f(b_i) - V_f(a_i)| < \varepsilon$ y, así, V_f es absolutamente continua. Finalmente, como $\mathrm{AC}([a, b])$ es un espacio vectorial, vemos que las funciones $f_1 = V_f$ y $f_2 = f - f_1$ son absolutamente continuas y $f = f_1 - f_2$. ■

Cuando estudiamos las funciones Lipschitz vimos que todas ellas satisfacían la condición (N) de Lusin, Teorema 9.1.6. El siguiente resultado muestra que las funciones absolutamente continuas se comportan de igual manera.

Teorema 9.1.51. *Si $f \in \mathrm{AC}([a, b])$, entonces f satisface la **condición (N) de Lusin**. En particular, f transforma **conjuntos medibles** en **conjuntos medibles**; es decir,*

$$E \in \mathfrak{M}_\mu([a, b]) \quad \Rightarrow \quad f(E) \in \mathfrak{M}_\mu([a, b]).$$

Prueba. Sea $E \subseteq [a, b]$ un conjunto medible con $\mu(E) = 0$ y sea $\varepsilon > 0$. Como f es absolutamente continua, existe un $\delta > 0$ para el cual

$$\sum_{n=1}^{m} |f(b_n) - f(a_n)| \ < \ \varepsilon \tag{1}$$

para cualquier colección finita $\{[a_n, b_n] : n = 1, \ldots, m\}$ de intervalos no-superpuestos en $[a, b]$ que satisfaga $\sum_{n=1}^{m}(b_n - a_n) < \delta$. Ahora bien, como $\mu(E) = 0$, existe un conjunto abierto G tal que

$$E \subseteq G \subseteq [a, b] \qquad \text{y} \qquad \mu(G) \ < \ \delta.$$

Siendo G abierto, existe una sucesión disjunta de intervalos abiertos $\big((\alpha_n, \beta_n)\big)_{n=1}^{\infty}$ incluidos en $[a, b]$ tal que $G = \bigcup_{n=1}^{\infty}(\alpha_n, \beta_n)$. Puesto que f es continua en cada compacto $[\alpha_n, \beta_n]$, seleccione puntos $x_n, y_n \in [\alpha_n, \beta_n]$ en los cuales f alcanza su mínimo y su máximo. Si perder generalidad, asumiremos que $x_n < y_n$ para todo $n \geq 1$. Esto nos dice que f transforma cada uno de los intervalos $[\alpha_n, \beta_n]$ en un intervalo con extremos $f(x_n)$ y $f(y_n)$. Observe que

$$\sum_{n=1}^{m}(y_n - x_n) \ \leq \ \sum_{n=1}^{m}(\beta_n - \alpha_n) \ \leq \ \sum_{n=1}^{\infty}(\beta_n - \alpha_n) \ = \ \mu(G) \ < \ \delta$$

y entonces, por (1),

$$\sum_{n=1}^{m} |f(y_n) - f(x_n)| \ < \ \varepsilon.$$

Tomando límite cuando $m \to \infty$, obtenemos que $\sum_{n=1}^{\infty} |f(y_n) - f(x_n)| \leq \varepsilon$. Por otro lado, como

$$f(E) \subseteq f(G) \ = \ \bigcup_{n=1}^{\infty} f\big((\alpha_n, \beta_n)\big) \subseteq \bigcup_{n=1}^{\infty} f\big([\alpha_n, \beta_n]\big)$$

se obtiene, finalmente, que

$$0 \ \leq \ \mu(f(E)) \ \leq \ \sum_{n=1}^{\infty} \mu\big(f\big([\alpha_n, \beta_n]\big)\big) \ = \ \sum_{n=1}^{\infty} |f(y_n) - f(x_n)| \ \leq \ \varepsilon.$$

Siendo $\varepsilon > 0$ arbitrario, concluimos que $\mu(f(E)) = 0$. Puesto que f es continua, la última parte es consecuencia del Teorema de Rademacher, página 455. ■

Combinando el Teorema 9.1.49 con el Teorema 9.1.51 se tiene que

Corolario 9.1.52. *Si $f \in \mathrm{AC}([a,b])$, entonces f es **continua**, de **variación acotada** y satisface la **condición (N) de Lusin**.*

Lo que resulta un poco sorprendente es que el recíproco del resultado anterior se cumple, es decir, las tres condiciones: *f es continua, de variación acotada y satisface la condición* (N) *de Lusin*, garantizan que f es absolutamente continua. Este resultado fue demostrado por S. Banach, e independientemente, por M. A. Zarecki. Tendremos ocasión de dar dos demostraciones de este resultado: una al final de esta sección y otra en el próximo capítulo.

Corolario 9.1.53. *Sea $f : [a,b] \to \mathbb{R}$ una **función absolutamente continua** sobre $[a,b]$. Entonces f transforma **conjuntos cerrados** en **conjuntos cerrados**. En particular, f transforma conjuntos G_δ en conjuntos G_δ.*

Prueba. Sea $F \subseteq [a,b]$ un conjunto cerrado. Puesto que $[a,b]$ es compacto resulta que F es compacto. Por otro lado, como f es continua, ella transforma conjuntos compactos en conjuntos compactos y, en consecuencia, $f(F)$ es compacto, en particular, cerrado. Suponga ahora que G es un conjunto G_δ incluido en $[a,b]$ y sea $(A_n)_{n=1}^{\infty}$ una sucesión de conjuntos cerrados tal que $G = \bigcup_{n=1}^{\infty} A_n$. Puesto que

$$f\Big(\bigcup_{n=1}^{\infty} A_n\Big) \ = \ \bigcup_{n=1}^{\infty} f(A_n)$$

se concluye que f transforma conjuntos G_δ en conjuntos G_δ. ■

Corolario 9.1.54. *Se $f : [a,b] \to \mathbb{R}$ una función **absolutamente continua** sobre $[a,b]$. Si*

$$E_\infty \ = \ \big\{x \in [a,b] \ : \ |f'(x)| = \infty\big\},$$

entonces $\mu(f(E_\infty)) = 0$.

Prueba. Por el Teorema de Diferenciabilidad de Lebesgue, f' existe casi-siempre de modo que $\mu(E_\infty) = 0$. Como f es absolutamente continua, ella satisface, por el teorema anterior, la condición (N) de Lusin y, en consecuencia, $\mu(f(E_\infty)) = 0$. ∎

El siguiente resultado es otra forma elegante de caracterizar la absoluta continuidad de una función f en presencia de una condición extra.

Teorema 9.1.55. *Sea $f : [a,b] \to \mathbb{R}$ una **función continua** y **creciente** sobre $[a,b]$. Las siguientes condiciones son equivalentes:*

(1) *f es **absolutamente continua** sobre $[a,b]$.*

(2) *Para cada $\varepsilon > 0$, existe un $\delta > 0$ tal que para cada conjunto medible $E \subseteq [a,b]$, se cumple que*

$$\mu(E) < \delta \quad \Rightarrow \quad \mu^*(f(E)) < \varepsilon.$$

Prueba. (1) $\Rightarrow$ (2). Suponga que f es absolutamente continua sobre $[a,b]$ y sea $\varepsilon > 0$. Usemos el Ejercicio 9.3 $3(c)$ para hallar un $\delta > 0$ tal que

$$\sum_{n=1}^{\infty} |f(b_n) - f(a_n)| < \varepsilon$$

siempre que $\{(a_n, b_n) : n \in \mathbb{N}\}$ es una colección infinita numerable de intervalos abiertos y disjuntos incluidos en $[a,b]$ satisfaciendo $\sum_{n=1}^{\infty}(b_n - a_n) < \delta$. Sea ahora $E \subseteq [a,b]$ un conjunto medible con $\mu(E) < \delta/2$. Escoja una colección numerable y disjunta $\{I_n = (a_n, b_n) : n \in \mathbb{N}\}$ de intervalos abiertos tal que

$$E \subseteq \bigcup_{n=1}^{\infty} I_n \qquad \text{y} \qquad \mu\left(\bigcup_{n=1}^{\infty} I_n\right) < \delta/2 + \mu(E) < \delta.$$

Se sigue entonces que

$$\sum_{n=1}^{\infty} |f(b_n) - f(a_n)| < \varepsilon.$$

Puesto que f es continua y creciente, resulta que cada conjunto $f(I_n)$ es abierto, $f(E) \subseteq \bigcup_{n=1}^{\infty} f(I_n)$ y $\mu(I_n) = b_n - a_n$. Se tiene entonces que

$$\mu^*(f(E)) \leq \sum_{n=1}^{\infty} |f(b_n) - f(a_n)| < \varepsilon.$$

Para ver que (2) $\Rightarrow$ (1), suponga que (2) se cumple y escoja un $\delta > 0$ correspondiente al ε dado en la condición (2). Sea $\{a_n, b_n) : n = 1, \ldots, k\}$ una colección finita de intervalos abiertos y disjuntos tal que $E = \bigcup_{n=1}^{k}(a_n, b_n)$ y $\mu(E) < \delta$. Puesto que f es creciente, se tiene que la colección $\{f((a_n, b_n)) : n = 1, \ldots, k\}$ es disjunta y, entonces, la continuidad de f nos asegura que $\mu(f((a_n, b_n))) = f(b_n) - f(a_n)$. Por esto,

$$\sum_{n=1}^{k} |f(b_n) - f(a_n)| = \mu\left(\bigcup_{n=1}^{k}(a_n, b_n)\right) < \varepsilon.$$

La prueba es completa. ∎

Otro hecho que merece nuestra atención es que la composición de funciones absolutamente continuas no es necesariamente absolutamente continua. En efecto, las funciones f y g definidas por

$$f(x) = \begin{cases} x^2\,|\mathrm{sen}(1/x)| & \text{si} \quad x \in (0,1] \\ 0 & \text{si} \quad x = 0 \end{cases}$$

y $g(x) = \sqrt{x}$ son absolutamente continuas, pero $g \circ f$ no lo es. Sin embargo, vale el siguiente resultado.

Teorema 9.1.56. *Sea $g \in \mathrm{AC}([c,d])$ y suponga que g es **estrictamente monótona** sobre $[a,b]$. Si $f \in \mathrm{AC}([a,b])$ donde $g([c,d]) = [a,b]$, entonces $f \circ g \in \mathrm{AC}([c,d])$.*

Prueba. Suponga que g es monótona estrictamente creciente y sea $\varepsilon > 0$. Como f es absolutamente continua, existe un $\delta > 0$ tal que

$$\sum_{i=1}^{n} |f(b_i) - f(a_i)| \; < \; \varepsilon \tag{1}$$

para cualquier colección finita $\{[a_i, b_i] : i = 1, \ldots, n\}$ en $[a,b]$ de intervalos no-superpuestos que satisfaga $\sum_{i=1}^{n}(b_i - a_i) < \delta$. Ahora bien, como g también es absolutamente continua, existe un δ_1 tal que, para cualquier colección finita de intervalos no-superpuestos $\{[c_i, d_i] : i = 1, \ldots, m\}$ en $[c,d]$ que satisfaga $\sum_{i=1}^{m}(d_i - c_i) < \delta_1$, entonces se cumple que

$$\sum_{i=1}^{m} (g(d_i) - g(c_i)) \; < \; \delta \qquad (2).$$

Puesto que g es continua y estrictamente creciente resulta que $\{[g(c_i), g(d_i)] : i = 1, \ldots, m\}$ es una colección finita de intervalos no-superpuestos en $[a,b]$ satisfaciendo (2) de donde se obtiene, gracias a (1), que

$$\sum_{i=1}^{n} |f(g(d_i)) - f(g(c_i))| \; < \; \varepsilon$$

y termina la prueba. ∎

Si uno considera la colección $\mathrm{C(N)}([a,b])$ *de las funciones* $f : [a,b] \to \mathbb{R}$ *que satisfacen la condición* (N) *de Lusin*, entonces $\mathrm{AC}([a,b]) \subseteq \mathrm{C(N)}([a,b])$. Observe que $\mathrm{C(N)}([a,b])$ contiene funciones que no son continuas sobre $[a,b]$. Por ejemplo, toda función simple f definida sobre $[a,b]$ pertenece a $\mathrm{C(N)}([a,b])$: En efecto, sea E un subconjunto medible de $[a,b]$ con $\mu(E) = 0$. Como f es simple, su rango es un conjunto finito y, por lo tanto, de medida cero. De esto se sigue que $\mu(f(E)) = 0$. Sin embargo, $\mathrm{C(N)}([a,b])$ no es un espacio vectorial sobre $\mathbb{R}$. Este hecho se comprueba al final de esta sección donde se muestra un ejemplo de dos funciones en $\mathrm{C(N)}([a,b])$ cuya suma no satisface la condición (N) de Lusin, aunque la composición de dos funciones que satisfacen la condición (N) de Lusin también lo es. Veremos en lo inmediato que toda función que pertenece a $\mathrm{C(N)}([a,b])$ y es singular es necesariamente constante. Para demostrar ese resultado, recordemos que la derivada de una función puede ser usada para estimar el rango de la función en términos de su dominio de definición. En efecto, en el Teorema 9.1.6, página 457, vimos que si $f : [a,b] \to \mathbb{R}$ es una función continua y diferenciable en $[a,b]$ tal que $|f'| \leq M$ para alguna constante $M > 0$, entonces $\mu(f(I)) \leq M\mu(I)$ para cualquier intervalo $I \subseteq [a,b]$. El próximo resultado provee una estimación similar sobre cualquier conjunto $E \subseteq [a,b]$ en donde $f'(x)$ exista y sea acotada. De modo preciso.

Lema 9.1.57. *Sea $f : [a,b] \to \mathbb{R}$ una función y sea $E \subseteq [a,b]$. Suponga que f es **diferenciable** en cada punto de E y que $M = \sup_{x \in E} |f'(x)| < \infty$. Entonces*

$$\mu^*(f(E)) \;\leq\; M \cdot \mu^*(E).$$

Prueba. Sea $\varepsilon > 0$. Para cada $x \in E$, se tiene que

$$\lim_{y \to x} \left| \frac{f(y) - f(x)}{y - x} \right| \;=\; |f'(x)| \leq\; M.$$

Escojamos un $\delta > 0$ tal que

$$\big|f(y) - f(x)\big| \;\leq\; M|x - y| \quad \text{para todo } y \in (x - \delta, x + \delta) \cap [a,b]. \tag{1}$$

Defina, para cada $n \in \mathbb{N}$, el conjunto

$$E_n \;=\; \left\{ x \in E \,:\, \big|f(y) - f(x)\big| \leq M|x - y| \;\text{ si }\; y \in \left(x - \frac{1}{n}, x + \frac{1}{n}\right) \cap [a,b] \right\}.$$

Es fácil ver que $E_1 \subseteq E_2 \subseteq \cdots$ y, además, $E = \bigcup_{n=1}^{\infty} E_n$. En efecto, sea $x \in E$. Eligiendo un $n_0 \in \mathbb{N}$ lo suficientemente grande de modo que $1/n_0 < \delta$, entonces (1) nos garantiza que $x \in E_{n_0}$ y, en consecuencia, $E \subseteq \bigcup_{n=1}^{\infty} E_n$. La otra inclusión es inmediata. Se sigue del Corolario 6.3.46 que

$$\mu^*(E) \;=\; \lim_{n \to \infty} \mu^*(E_n). \tag{2}$$

Por otro lado, como la sucesión $(f(E_n))_{n=1}^{\infty}$ también es monótona creciente y

$$f(E) \;=\; f\left(\bigcup_{n=1}^{\infty} E_n \right) \;=\; \bigcup_{n=1}^{\infty} f(E_n),$$

entonces una nueva aplicación del Corolario 6.3.46 nos revela que

$$\mu^*(f(E)) \;=\; \lim_{n \to \infty} \mu^*(f(E_n)). \tag{3}$$

Usando ahora la definición de medida exterior podemos elegir, por cada $n \in \mathbb{N}$, una sucesión $(I_{n,k})_{k=1}^{\infty}$ de intervalos abiertos tales que:

$$E_n \;\subseteq\; \bigcup_{k=1}^{\infty} I_{n,k} \qquad \text{y} \qquad \sum_{k=1}^{\infty} \ell(I_{n,k}) \;<\; \mu^*(E_n) + \varepsilon. \tag{4}$$

Dividiendo cada uno de los intervalos $I_{n,k}$ en un número finito de subintervalos si fuese necesario, podemos asumir que $\ell(I_{n,k}) < 1/n$ para todo $k \geq 1$. De (4) se sigue que

$$E_n \;=\; E_n \cap \left(\bigcup_{k=1}^{\infty} I_{n,k} \right) \;=\; \bigcup_{k=1}^{\infty} (E_n \cap I_{n,k})$$

y, por lo tanto,

$$f(E_n) \;=\; f\Big(\bigcup_{k=1}^{\infty}(E_n\cap I_{n,k})\Big) \;=\; \bigcup_{k=1}^{\infty} f(E_n\cap I_{n,k}). \tag{5}$$

Fijemos un $n\in\mathbb{N}$ y tome dos puntos cualesquiera $x_1, x_2 \in E_n\cap I_{n,k}$. Entonces

$$|f(x_1)-f(x_2)| \;\leq\; M|x_1-x_2| \;\leq\; M\cdot\ell(I_{n,k}),$$

lo cual prueba que el $\operatorname{diam}(f(E_n\cap I_{n,k}))\leq M\cdot\ell(I_{n,k})$ y, por consiguiente,

$$\mu^*(f(E_n\cap I_{n,k})) \;\leq\; M\cdot\ell(I_{n,k}).$$

Entonces, por (5) y (4) se tiene que

$$\begin{aligned}
\mu^*(f(E_n)) \;&=\; \mu^*\Big(\bigcup_{k=1}^{\infty} f(E_n\cap I_{n,k})\Big)\\
&\leq\; \sum_{k=1}^{\infty}\mu^*\big(f(E_n\cap I_{n,k})\big)\\
&\leq\; \sum_{k=1}^{\infty} M\cdot\ell(I_{n,k})\\
&\leq\; M\cdot\big(\mu^*(E_n)+\varepsilon\big).
\end{aligned}$$

Esto último, en combinación con (3) y (2), producen

$$\begin{aligned}
\mu^*(f(E)) \;&=\; \lim_{n\to\infty}\mu^*(f(E_n))\\
&\leq\; M\cdot\big(\lim_{n\to\infty}\mu^*(E_n)+\varepsilon\big)\\
&\leq\; M\cdot\big(\mu^*(E)+\varepsilon\big).
\end{aligned}$$

Puesto que ε es arbitrario, se concluye que $\mu^*(f(E))\leq M\cdot\mu^*(E)$ y termina la prueba. ∎

Varias consecuencias interesantes se pueden deducir del Lema 9.1.57. Por ejemplo:

Corolario 9.1.58. *Sea $f:[a,b]\to\mathbb{R}$ una función arbitraria y suponga que f es* ***diferenciable*** *sobre un conjunto arbitrario $E\subseteq[a,b]$.*

(a) Si $Z=\{x\in E:\ f'(x)=0\}$, entonces $\mu(f(Z))=0$.

(b) Si $\mu(f(E))=0$, entonces $f'(x)=0$ para casi-todo $x\in E$.

Prueba. (a). Sea $n\in\mathbb{N}$. Puesto que $|f'(x)|\leq 1/n$ para todo $x\in Z$, se sigue del Lema 9.1.57 que $\mu(f(Z))\leq\frac{1}{n}\mu(Z)$, de donde se concluye, tomando límite cuando $n\to\infty$, que $\mu(f(Z))=0$.

(b). Sea $N=\{x\in E:|f'(x)|>0\}$. Vamos a demostrar que $\mu(N)=0$. Para ver esto, considere, para cada $k\in\mathbb{N}$, el conjunto

$$N_k=\Big\{x\in N:\ |f(x)-f(y)|\geq\frac{|x-y|}{k}\ \text{para todo}\ \ y\in(x-1/k,x+1/k)\cap[a,b]\Big\}.$$

Observe que como

$$N = \bigcup_{k=1}^{\infty} N_k,$$

entonces es suficiente demostrar que $\mu(N_k) = 0$ para todo $k \geq 1$. Fijemos un $k \geq 1$ y sea $F = J \cap N_k$, donde J es un intervalo de longitud menor que $1/k$. Nuestra tarea es ver que $\mu(F) = 0$. En efecto, como $\mu(f(E)) = 0$ y $F \subseteq E$, entonces para cada $\varepsilon > 0$, podemos encontrar una sucesión $(J_n)_{n=1}^{\infty}$ de intervalos abiertos tal que

$$f(F) \subseteq \bigcup_{n=1}^{\infty} J_n \qquad \text{y} \qquad \sum_{k=1}^{\infty} \ell(J_n) < \frac{\varepsilon}{k}.$$

Sea $E_n = f^{-1}(J_n) \cap F$ para $n = 1, 2, \ldots$ y nótese que como $f(E_n) \subseteq J_n$, entonces

$$\sup_{x,y \in E_n} |f(x) - f(y)| \leq \ell(J_n),$$

y así,

$$\sum_{n=1}^{\infty} k \sup_{x,y \in E_n} |f(x) - f(y)| \leq k \sum_{k=1}^{\infty} \ell(J_n) < \varepsilon.$$

Finalmente, puesto que $F \subseteq \bigcup_{k=1}^{\infty} E_n$ y $E_n \subseteq J \cap N_k$, resulta que

$$\begin{aligned}
\mu^*(F) &\leq \sum_{n=1}^{\infty} \mu^*(E_n) \leq \sum_{n=1}^{\infty} \operatorname{diam}(E_n) \\
&= \sum_{n=1}^{\infty} \sup_{x,y \in E_n} |x - y| \leq \sum_{n=1}^{\infty} k \sup_{x,y \in E_n} |f(x) - f(y)| \\
&< \varepsilon.
\end{aligned}$$

Siendo $\varepsilon > 0$ arbitrario, concluimos que $\mu(F) = 0$ y termina la prueba. ■

Observe que en la parte (a) del corolario anterior no podemos reemplazar la frase "f es **siempre diferenciable** sobre $[a, b]$" por "f es **diferenciable casi-siempre** sobre $[a, b]$". En efecto, tomando $f = \varphi_\Gamma$, la función de Cantor, resulta que por ser f una función singular se tiene que $Z = [0, 1] \setminus \Gamma$, mientras que $\mu(f(Z)) = \mu(\varphi_\Gamma(Z)) = 1$.

El corolario anterior también puede ser usado para obtener una simple e interesante caracterización de las funciones singulares.

Corolario 9.1.59. *Sea $f : [a, b] \to \mathbb{R}$ una **función no constante** sobre $[a, b]$ tal f' **existe casi-siempre** sobre $[a, b]$. Las siguientes condiciones son equivalentes:*

(1) *f es **singular**.*

(2) *Existe un **conjunto medible** $Z \subseteq [a, b]$ tal que*

$$\mu([a, b] \setminus Z) = 0 \qquad y \qquad \mu(f(Z)) = 0.$$

Prueba. (1) $\Rightarrow$ (2). Suponga que f es singular y sea

$$Z \;=\; \{x \in [a,b] \,:\, f'(x) = 0\}.$$

Entonces Z es medible, $\mu([a,b] \setminus Z) = 0$ y se sigue del Corolario 9.1.58 (a) que $\mu(f(Z)) = 0$.

(2) $\Rightarrow$ (1). Suponga que (2) se cumple. De nuevo, por la parte (b) del Corolario 9.1.58 se tiene que $f'(x) = 0$ para casi todo $x \in Z$. Puesto que $\mu([a,b] \setminus Z) = 0$ se concluye que $f'(x) = 0$ para casi todo $x \in [a,b]$. La prueba es completa. ■

Gracias al Teorema 9.1.7 sabemos que cualquier función diferenciable $f : [a,b] \to \mathbb{R}$ con derivada acotada es Lipschitz y, en consecuencia, satisface la condición (N) de Lusin (Teorema 9.1.6 (c)). En el siguiente corolario se obtiene la misma conclusión en ausencia de acotamiento.

Corolario 9.1.60. *Sea $f : [a,b] \to \mathbb{R}$ una función arbitraria. Si f es **diferenciable** sobre $[a,b]$, entonces ella satisface **la condición (N) de Lusin**.*

Prueba. Sea E un subconjunto incluido en $[a,b]$ tal que $\mu(E) = 0$ y, para cada $n \geq 1$, considere el conjunto

$$E_n \;=\; \big\{x \in E \,:\, |f'(x)| \leq n\big\}.$$

Como $E_n \subseteq E$, se tiene que $\mu(E_n) = 0$ para todo $n \geq 1$ y, entonces, por el Lema 9.1.57, se concluye que $\mu^*(f(E_n)) \leq n \cdot \mu(E_n) = 0$. Finalmente, como μ es numerablemente subaditiva y $f(E) = \bigcup_{n=1}^{\infty} f(E_n)$, resulta que $\mu(f(E)) = 0$ y termina la prueba. ■

Corolario 9.1.61. *Sea $f : [a,b] \to \mathbb{R}$ una función **continua** satisfaciendo la **condición (N) de Lusin**. Si $f'(x) = 0$ para casi-todo $x \in [a,b]$, entonces f es **constante** sobre $[a,b]$.*

Prueba. Sea $Z = \{x \in [a,b] : f'(x) = 0\}$ y sea $D = [a,b] \setminus Z$. Entonces $\mu(D) = 0$ y así, por hipótesis, $\mu(f(D)) = 0$. Del Corolario 9.1.58 se sigue que $\mu(f(Z)) = 0$ y, en consecuencia, $\mu(f([a,b])) = \mu(f(D)) + \mu(f(Z)) = 0$. Por otro lado, como f es continua, resulta que $f([a,b])$ es un intervalo y como dicho intervalo posee medida cero, la única opción que le queda es que él se reduzca a un punto. Por esto, f es constante y finaliza la prueba. ■

Como toda función absolutamente continua es continua y satisface, por el Teorema 9.1.51, la condición (N) de Lusin, se sigue entonces del corolario anterior que:

Corolario 9.1.62. *Si $f \in \mathrm{AC}([a,b])$ y $f'(x) = 0$ para casi-todo $x \in [a,b]$, entonces f es **constante** sobre $[a,b]$. En particular, si $f, g \in \mathrm{AC}([a,b])$ y $f'(x) = g'(x)$ para casi-todo $x \in [a,b]$, entonces $f - g$ es constante sobre $[a,b]$.*

Puesto que toda función Lipschitz es absolutamente continua, resulta que el Teorema 9.1.9, página 458, es consecuencia inmediata del corolario anterior. Otra manera de demostrar el Corolario 9.1.62 usando el Teorema del Cubrimiento de Vitali es como sigue:

Otra prueba del Corolario 9.1.62. Vamos a demostrar que $f(c) = f(a)$ para cualquier $c \in (a,b]$. Fijemos entonces un $c \in (a,b]$ y considere el conjunto

$$E_c \;=\; \big\{x \in [a,c] \,:\, f'(x) = 0\big\}.$$

Puesto que, por hipótesis, f es singular, resulta que $\mu([a,b] \setminus E_b) = 0$. En particular, se tiene que $\mu([a,c] \setminus E_c) = 0$ y, en consecuencia, $\mu(E_c) = c - a$ ya que $[a,c] = ([a,c] \setminus E_c) \cup E_c$. Sea $\varepsilon > 0$ y usemos el hecho de que f es absolutamente continua para hallar un $\delta > 0$ tal que

$$\sum_{i=1}^{n} \left| f(d_i) - f(c_i) \right| \;<\; \varepsilon/2$$

para cualquier colección finita de intervalos no-superpuestos $\{[c_i, d_i] : 1 \leq i \leq n\}$ en $[a,b]$ satisfaciendo la desigualdad $\sum_{i=1}^{n}(d_i - c_i) < \delta$. Observe que si $x \in E_c$, entonces existe un y_x tal que $[x, y_x] \subseteq (a,c)$ y para todo $x < y < y_x$,

$$\frac{f(y) - f(x)}{y - x} \;<\; \frac{\varepsilon}{2(c-a)}. \tag{I$_1$}$$

Este hecho nos permite considerar la siguiente familia $\mathcal{V}$ de subintervalos cerrados de (a,c):

$$\mathcal{V} \;=\; \bigcup_{x \in E_c} \{[x,y] \;:\; x < y < y_x \quad \text{y} \quad (\mathrm{I}_1) \quad \text{se cumple}\}.$$

Resulta que $\mathcal{V}$ es un cubrimiento de Vitali para E_c y, por lo tanto, por el Teorema del Cubrimiento de Vitali, existe una colección finita de subconjuntos no-superpuestos $\{[x_i, y_i] : 1 \leq i \leq n\}$ en $\mathcal{V}$ tal que

$$\mu\Big(E_c \setminus \bigcup_{i=1}^{n} [x_i, y_i]\Big) \;<\; \delta.$$

En particular, véase el Comentario Adicional 9.1.3 (3), se tiene que

$$c - a \;=\; \mu(E_c) \;<\; \sum_{i=1}^{n} (y_i - x_i) \;+\; \delta. \tag{I$_2$}$$

Para simplificar la presentación, vamos a asumir que los puntos extremos de los intervalos $[x_1, y_1], \ldots, [x_n, y_n]$ están dispuestos en orden creciente (esto, por supuesto, no cambia la desigualdad anterior), es decir, suponga que

$$a \;<\; x_1 \;<\; y_1 \;<\; x_2 \;<\; y_2 \;<\; \cdots \;<\; y_{k-1} \;<\; x_n \;<\; y_n \;<\; c.$$

Si definimos $a = y_0$ y $x_{n+1} = c$, tendremos que $\mathcal{P} = \{y_0, x_1, y_1, \ldots, x_n, y_n, x_{n+1}\}$ es una partición de $[a,c]$ y, por lo tanto,

$$\sum_{i=0}^{n} (x_{i+1} - y_i) \;+\; \sum_{i=1}^{n} (y_i - x_i) \;=\; c - a,$$

lo que combinado con (I_2) produce la desigualdad

$$\sum_{i=0}^{n} (x_{i+1} - y_i) \;=\; (c-a) \;-\; \sum_{i=1}^{n} (y_i - x_i) \;<\; \delta.$$

Invocando ahora el hecho de que f es absolutamente continua, resulta que

$$\sum_{i=0}^{n} |f(x_{i+1}) - f(y_i)| \;<\; \frac{\varepsilon}{2}.$$

Por otro lado, puesto que los $[x_i, y_i] \in \mathcal{V}$ para $i = 1, \dots, n,$ se sigue de (I_1) que

$$\frac{f(y_i) - f(x_i)}{y_i - x_i} < \frac{\varepsilon}{2(c-a)}$$

y, en consecuencia,

$$\sum_{i=1}^{n} |f(y_i) - f(x_i)| < \frac{\varepsilon}{2(c-a)} \sum_{i=1}^{n} (y_i - x_i) \le \frac{\varepsilon}{2}.$$

Finalmente,

$$\begin{aligned} |f(c) - f(a)| &\le \sum_{i=0}^{n} |f(x_{i+1}) - f(y_i)| + \sum_{i=1}^{n} |f(y_i) - f(x_i)| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Como $\varepsilon > 0$ es arbitrario, tenemos que $f(c) = f(a)$. La prueba es completa. ■

El siguiente ejemplo, el cual es consecuencia del corolario anterior, es emblemático en la Teoría de Integración de Lebesgue. Es un ejemplo de una función uniformemente continua, singular, pero que no es absolutamente continua.

Ejemplo 9.1.8. *La* ***función de Cantor*** $\varphi_\Gamma : [0,1] \to [0,1]$ ***no es absolutamente continua sobre*** $[a,b]$.

Prueba. En efecto, por el Corolario 5.1.10, página 236, sabemos que $\varphi'_\Gamma = 0$ $\mu -$c.s. Puesto que φ_Γ no es constante, ella no puede, gracias al Corolario 9.1.62, ser absolutamente continua. ■

9.1.7. El Teorema de Banach-Zarecki

El en Corolario 9.1.52 hemos demostrado que cualquier función absolutamente continua es de variación acotada y posee la condición (N) de Lusin. Resulta que esas dos propiedades caracterizan a la clase de las funciones absolutamente continuas. Este resultado es conocido con el nombre de Teorema de Banach-Zarecki (o Banach-Zaretski), el cual cuenta con una amplia literatura y, por supuesto, también con una amplia variedad de pruebas, casi todas distintas. Aquí seguimos la demostración dada por Vasile Ene [56] que no utiliza la noción de integral dada por Lebesgue. Más adelante, cuando hallamos introducido la integral de Lebesgue, daremos otra demostración de este resultado el cual se obtendrá como consecuencia del Teorema de Diferenciabilidad de Lebesgue y otros resultados adicionales.

El Teorema de Banach-Zarecki es un resultado clásico del Análisis Real con muchas aplicaciones, principalmente en el Análisis Funcional, así como también en ciertos fenómenos que ocurren en Física e Ingeniería. Éste teorema fue probado por primera vez por el matemático polaco Stefan Banach e independientemente por el ruso M. A. Zarecki. Por supuesto, tal resultado admite una variedad de extensiones tanto a funciones con dominio en cubos de $\mathbb{R}^n$, así como variaciones en las hipótesis.

Fijemos una función $f : [a,b] \to \mathbb{R}$ y definamos

$$\mathrm{Ra}(f) = \big\{y \in f([a,b]) : \mathrm{card}\big(f^{-1}(\{y\})\big) > 1\big\}.$$

Observe que $y \in \mathrm{Ra}(f)$ significa que la ecuación $y = f(x)$ posee más de una solución y por lo tanto f no es inyectiva. Por supuesto, la inyectividad de f implica que $\mathrm{Ra}(f) = \varnothing$.

Lema 9.1.63. *Sea $f : [a, b] \to \mathbb{R}$ una función. Si $(E_\alpha)_{\alpha\in D}$ es una familia arbitraria de subconjuntos de $[a, b]$, entonces*

$$\Big(\bigcap_{\alpha\in D} f(E_\alpha)\Big) \setminus \mathrm{Ra}(f) \ \subseteq\ f\Big(\bigcap_{\alpha\in D} E_\alpha\Big) \ \subseteq\ \bigcap_{\alpha\in D} f(E_\alpha).$$

Prueba. Sea $y \in \big(\bigcap_{\alpha\in D} f(E_\alpha)\big) \setminus \mathrm{Ra}(f)$. Puesto que $y \in \bigcap_{\alpha\in D} f(E_\alpha)$, se tiene que $y \in f(E_\alpha)$ para cada $\alpha \in D$ y, por lo tanto existe, por cada $\alpha \in D$, un $x_\alpha \in E_\alpha$ tal que $y = f(x_\alpha)$. Pero como $y \notin \mathrm{Ra}(f)$, existe un único punto $x \in [a, b]$ tal que $y = f(x)$, de donde se sigue que $x_\alpha = x$ para todo $\alpha \in D$. Esto prueba que $x \in E_\alpha$ para todo $\alpha \in D$ y, así, $x \in \bigcap_{\alpha\in D} E_\alpha$. Por esto, $y = f(x) \in f\big(\bigcap_{\alpha\in D} E_\alpha\big)$. La última inclusión es bien conocida. ■

Lema 9.1.64. *Sea $f : [a, b] \to \mathbb{R}$ una función **creciente** sobre $[a, b]$. Entonces*

*(a) $\mathrm{Ra}(f)$ es **a lo más numerable**. En particular, $\mathrm{Ra}(f) \in \mathfrak{Bo}(\mathbb{R})$.*

(b) Si $G \subseteq [a, b]$ es un G_δ, entonces $f(G) \in \mathfrak{Bo}(\mathbb{R})$. En particular,

$$\mu(f(G)) \ =\ \mu\Big(\bigcap_{n=1}^{\infty} f(G_n)\Big),$$

donde $G = \bigcap_{n=1}^{\infty} G_n$ con G_n un conjunto abierto para todo $n \in \mathbb{N}$,

Prueba. (a) Para cada $y \in \mathrm{Ra}(f)$, sean $x'_y = \inf f^{-1}(\{y\})$ y $x''_y = \sup f^{-1}(\{y\})$. Como f es creciente, resulta que $\varnothing \neq (x'_y, x''_y) \subseteq f^{-1}(\{y\})$, de donde se obtiene que

$$(x'_{y_1}, x''_{y_1}) \cap (x'_{y_2}, x''_{y_2}) \ =\ \varnothing \quad \text{siempre que} \quad y_1 \neq y_2.$$

Puesto que cualquier colección disjunta de intervalos abiertos es a lo más numerable, se sigue que $\mathrm{Ra}(f)$ es a lo más numerable.

(b) Sea $G = \bigcap_{n=1}^{\infty} G_n$, donde cada G_n es un conjunto abierto. Por el Lema 9.1.63 tenemos que

$$\Big(\bigcap_{n=1}^{\infty} f(G_n)\Big) \setminus \mathrm{Ra}(f) \ \subseteq\ f\Big(\bigcap_{n=1}^{\infty} G_n\Big) \ \subseteq\ \bigcap_{n=1}^{\infty} f(G_n). \tag{1}$$

Por otro lado, como f es creciente, cada $f(G_n)$ es una unión numerable de intervalos (algunos de los cuales puede ser degenerado) y, por consiguiente, un conjunto boreliano. Se sigue de (1) que $f\big(\bigcap_{n=1}^{\infty} G_n\big)$ es también un conjunto de Borel. La última parte es consecuencia de (a) y de (1). ■

Lema 9.1.65. *Sea $f : [a, b] \to \mathbb{R}$ una función **continua**, **creciente** y satisfaciendo la **condición $(\mathbf{N})$ de Lusin**. Entonces $f \in \mathrm{AC}([a, b])$.*

Prueba. Suponga que f no es absolutamente continua. Entonces existe un $\varepsilon > 0$ con la siguiente propiedad: para cada $\delta > 0$, existe una colección finita y disjunta de intervalos abiertos $\big\{(a_1, b_1), \ldots, (a_k, b_k)\big\}$ en $[a, b]$ tal que

$$\sum_{i=1}^{k}(b_i - a_i) \ <\ \delta \qquad \text{pero} \qquad \sum_{i=1}^{k}|f(b_i) - f(a_i)| \ \geq\ \varepsilon.$$

Para cada $n \in \mathbb{N}$, seleccione, usando lo anterior, una colección finita y disjunta de intervalos abiertos $\{((a_1^n, b_1^n), \ldots, (a_{k_n}^n, b_{k_n}^n)\}$ en $[a,b]$ tal que

$$\sum_{i=1}^{k_n}(b_i^n - a_i^n) < \frac{1}{2^n} \qquad \text{pero} \qquad \sum_{i=1}^{k_n}|f(b_i^n) - f(a_i^n)| \geq \varepsilon.$$

Pongamos

$$E_n = \bigcup_{i=1}^{k_n}(a_i^n, b_i^n) \qquad \text{y} \qquad E = \bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty}E_n$$

y observe que si $F_k = \bigcup_{n=k}^{\infty}E_n$, entonces la convergencia de la serie $\sum_{n=1}^{\infty}1/2^n$ nos revela que

$$\lim_{k\to\infty}\mu(F_k) \leq \lim_{k\to\infty}\sum_{n=k}^{\infty}\mu(E_n) < \lim_{k\to\infty}\sum_{n=k}^{\infty}1/2^n = \lim_{k\to\infty}\frac{1}{2^{k+1}} = 0.$$

Así, gracias al Teorema 6.3.47, página 287, se tiene que

$$\mu(E) = \lim_{k\to\infty}\mu\Big(\bigcup_{n=k}^{\infty}E_n\Big) = \lim_{k\to\infty}\mu(F_k) = 0.$$

Puesto que f satisface la condición (N) de Lusin, resulta que $\mu(f(E)) = 0$. Por otro lado, como $E_k \subseteq F_k$ tenemos que

$$\mu(f(F_k)) \geq \mu(f(E_k)) = \sum_{i=1}^{n_k}\big(f(b_i^k) - f(a_i^k)\big) \geq \varepsilon$$

y entonces, por la desigualdad anterior, el Lema 9.1.64 y el Teorema 6.3.47 se genera la siguiente contradicción

$$0 = \mu(f(E)) = \mu\Big(\bigcap_{k=1}^{\infty}f(F_k)\Big) = \lim_{k\to\infty}\mu(f(E_k)) \geq \varepsilon > 0.$$

Este absurdo establece que $f \in AC([a,b])$ y termina la prueba. ■

Lema 9.1.66. *Sea $f : [a,b] \to \mathbb{R}$ una función* ***continua*** *sobre $[a,b]$ y suponga que $P \subseteq [a,b]$ es un conjunto perfecto con $\{a,b\} \subseteq P$. Si $((a_n,b_n))_{n=1}^{\infty}$ son los intervalos contiguos a P, entonces*

$$|f(b) - f(a)| \leq \mu\big(f([a,b])\big) \leq \mu(P) + \sum_{n=1}^{\infty}\operatorname{osc}(f,[a_n,b_n]).$$

Prueba. Puesto que $[a,b] = P \cup \bigcup_{n=1}^{\infty}(a_n,b_n)$, entonces

$$f([a,b]) = f(P) \cup \bigcup_{n=1}^{\infty}f((a_n,b_n))$$

y como f es continua, resulta que $f([a,b])$ es un intervalo compacto conteniendo tanto a $f(a)$ así como a $f(b)$. Por esto,

$$\begin{aligned}|f(b) - f(a)| \leq \mu\big(f([a,b])\big) &\leq \mu(P) + \sum_{n=1}^{\infty}\mu\big(f((a_n,b_n))\big)\\ &\leq \mu(P) + \sum_{n=1}^{\infty}\operatorname{osc}(f,[a_n,b_n])\end{aligned}$$

Esto termina la prueba. ■

Teorema 9.1.67. *Sea $f : [a,b] \to \mathbb{R}$ una función* ***continua*** *y de* ***variación acotada*** *sobre $[a,b]$. Son equivalentes:*

(1) *f satisface* ***la condición (N) de Lusin****.*

(2) *V_f satisface* ***la condición (N) de Lusin****.*

Prueba. (1) $\Rightarrow$ (2). Por el Lema 9.1.20 y el Teorema 9.1.24 sabemos que V_f es continua y monótona creciente. Además, uno puede suponer, y así lo haremos, que V_f es estrictamente creciente. Suponga entonces que (1) se cumple pero que V_f no satisface la condición (N) de Lusin. Esto significa, en vista del Teorema de Foran, página 456, que existe un conjunto compacto $K \subseteq [a,b]$ con $\mu(E) = 0$ tal que $\alpha = \mu(V_f(E)) > 0$. Sean $c = \operatorname{ínf} K$ y $d = \sup K$. Claramente, $c, d \in K$. Más aun, como K es cerrado, el Teorema de Cantor-Bendixson, página 130, nos dice que él se puede escribir en la forma $K = P \cup Z$, donde P es perfecto y Z es a lo más numerable. Por esto, y puesto que $\mu(K) = \mu(P)$, asumiremos, sin perder generalidad en el razonamiento, que K es perfecto. Sean $\{(a_n, b_n) : n = 1, 2, \dots\}$ los intervalos abiertos contiguos a K y para cada entero $n > 1$ considere los n intervalos cerrados

$$[c_1, d_1] \cup [c_2, d_2] \cup \cdots \cup [c_n, d_n] \;=\; [c,d] \setminus \bigcup_{i=1}^{n-1} (a_i, b_i). \tag{0a}$$

Sea $d_n = \operatorname{máx}\{(d_1 - c_1), \dots, (d_n - c_n)\}$ y observe que como $\mu(K) = 0$, entonces $\operatorname{lím}_{n\to\infty} d_n = 0$. Considere cualquier partición $\mathcal{P} = \{c = x_0, x_1, \dots, x_n = d\}$ de $[c,d]$ con las siguientes propiedades: $x_i - x_{i-1} < d_n$ para $i = 1, \dots, n$ y ninguno de los puntos x_j está en algún intervalo (c_i, d_i). Sean

$$U_n \;=\; \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|$$

y

$$V_n \;=\; \sum_{i=1}^{n} |f(d_i) - f(c_i)| \;+\; \sum_{i=1}^{n-1} V(f, [a_i, b_i]). \tag{1a}$$

Entonces $U_n \leq V_n \leq V(f, [c,d])$. En virtud del Teorema 9.1.26 sabemos que

$$\operatorname*{lím}_{n\to\infty} U_n \;=\; V(f,[c,d]) \qquad \text{y} \qquad V(f,[c,d]) \;=\; \operatorname*{lím}_{n\to\infty} V_n.$$

De la última igualdad escoja un entero $N \geq 1$ de modo tal que

$$V_n \;>\; V(f,[c,d]) - \frac{\alpha}{2} \quad \text{para todo} \quad n \geq N \tag{2a}$$

Por el Lema 9.1.19, tenemos que

$$\begin{aligned} V(f,[c,d]) \;&=\; \sum_{i=1}^{n} V(f,[c_i,d_i]) \;+\; \sum_{i=1}^{n-1} V(f,[a_i,b_i]) \\ &=\; \sum_{i=1}^{n} \big(V_f(d_i) - V_f(c_i)\big) \;+\; \sum_{i=1}^{n-1} V(f,[a_i,b_i]). \end{aligned}$$

De $(1a)$, $(2a)$ y la última igualdad se tiene que, para cada $n \geq N$,

$$\sum_{i=1}^{n} |f(d_i) - f(c_i)| \;>\; \sum_{i=1}^{n} \big(V_f(d_i) - V_f(c_i)\big) \;-\; \frac{\alpha}{2}. \tag{3a}$$

Puesto que

$$\sum_{i=1}^{n} \big(V_f(d_i) - V_f(c_i)\big) \;>\; \mu(V_f(K)) \;=\; \alpha,$$

resulta de $(3a)$ que

$$\sum_{i=1}^{n} |f(d_i) - f(c_i)| \;>\; \frac{\alpha}{2}. \tag{4a}$$

Para cada $i = 1, \ldots, n$, considere $J_i = \{j \in \mathbb{N} : [a_j, b_j] \subseteq [c_i, d_i]\}$ y observe que, gracias a $(0a)$, cada índice $j \in J_i$ debe satisfacer que $j \geq n$. Ahora bien, teniendo en cuenta que $\mu(K) = 0$, el Lema 9.1.66 nos revela que

$$|f(d_i) - f(c_i)| \;\leq\; \sum_{j \,\in\, J_i} \operatorname{osc}(f, [a_j, b_j]),$$

para cada $i = 1, \ldots, n$, de donde se obtiene que

$$\sum_{i=1}^{n} |f(d_i) - f(c_i)| \;\leq\; \sum_{j=n}^{\infty} \operatorname{osc}(f, [a_j, b_j]). \tag{5a}$$

Combinando $(4a)$ y $(5a)$ produce, para todo $n \geq N$, la desigualdad

$$\frac{\alpha}{2} \;\leq\; \sum_{j=n}^{\infty} \operatorname{osc}(f, [a_j, b_j]). \tag{6a}$$

Finalmente, como la serie $\sum_{j=1}^{\infty} \operatorname{osc}(f, [a_j, b_j]) \leq V(f, [c, d])$, resulta que

$$\lim_{n\to\infty} \sum_{j=n}^{\infty} \operatorname{osc}(f, [a_j, b_j]) \;=\; 0 \tag{7a}$$

de modo que la combinación de $(6a)$ y $(7a)$ generan la siguiente contradicción:

$$0 \;<\; \frac{\alpha}{2} \;\leq\; 0.$$

Esto termina la prueba de la implicación $(1) \Rightarrow (2)$.

$(2) \Rightarrow (1)$. Suponga que (2) se cumple. Puesto que V_f es continua, creciente y satisface la condición (N) de Lusin, el Lema 9.1.65 nos asegura que $V_f \in \mathrm{AC}([a, b])$. Dado $\varepsilon > 0$ elija, usando el hecho de que $V_f \in \mathrm{AC}([a, b])$, un $\delta > 0$ tal que

$$\sum_{i=1}^{n} \big(V_f(x_i) - V_f(x_{i-1})\big) \;<\; \varepsilon \tag{8a}$$

para cualquier colección finita $\{[x_{i-1}, x_i] : i = 1, \ldots, n\}$ de subintervalos no-rampante de $[a, b]$ que satisfaga la desigualdad $\sum_{i=1}^{n}(x_i - x_{i-1}) < \delta$. Teniendo en cuenta que

$$|f(d) - f(c)| \ \leq \ V(f, [c, d]) \ = \ V_f(d) - V_f(c)$$

se cumple para cualquier intervalo $[c, d] \subseteq [a, b]$, se sigue de esto y $(8a)$ que

$$\sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| \ \leq \ \sum_{i=1}^{n} \big(V_f(x_i) - V_f(x_{i-1})\big) \ < \ \varepsilon$$

para cualquier colección finita $\{[x_{i-1}, x_i] : i = 1, \ldots, n\}$ de subintervalos no-rampante de $[a, b]$ satisfaciendo $\sum_{i=1}^{n}(x_i - x_{i-1}) < \delta$. Esto prueba que $f \in \mathrm{AC}([a, b])$. El hecho de que f satisface la condición (N) de Lusin sigue ahora del Teorema 9.1.51. ■

Los preparativos para la demostración del Teorema de Banach-Zarecki han sido totalmente completados.

Teorema 9.1.68 (Banach-Zarecki). *Sea $f : [a, b] \to \mathbb{R}$ una función. Las condiciones siguientes son equivalentes.*

(1) $f \in \mathrm{AC}([a, b])$.

(2) f es ***continua****, de* ***variación acotada*** *y satisface* ***la condición*** $(\mathbf{N})$ ***de Lusin***

Prueba. Ya hemos visto, Corolario 9.1.52, que (1) ⇒ (2). Para ver que (2) ⇒ (1) observe que la implicación (1) ⇒ (2) del Teorema 9.1.67 nos dice que V_f satisface la condición (N) de Lusin. Finalmente, la demostración de que $f \in \mathrm{AC}([a, b])$ está incluida en la implicación (2) ⇒ (1) del Teorema 9.1.67. ■

Una consecuencia inmediata del Teorema de Banach-Zarecki es el siguiente corolario.

Corolario 9.1.69. *Sea $f : [a, b] \to \mathbb{R}$ una función de* ***variación acotada*** *sobre $[a, b]$. Si f es* ***diferenciable*** *sobre $[a, b]$, entonces $f \in \mathrm{AC}([a, b])$.*

Prueba. Como f es diferenciable sobre $[a, b]$, ella es continua y también, por hipótesis, de variación sobre $[a, b]$. Además, por el Corolario 9.1.60 sabemos que f satisface la condición (N) de Lusin y, entonces, $f \in \mathrm{AC}([a, b])$ gracias al Teorema de Banach-Zarecki. ■

9.2. Ejercicios Resueltos

Ejercicio 9.2.1. *En este ejercicio se demuestra de nuevo que si E es medible, entonces $E - E$ contiene un* ***intervalo abierto****. Sea $E \subseteq \mathbb{R}$ medible. Pruebe que existe un $\delta > 0$ tal que*

$$E \cap (t + E) \ \neq \ \varnothing \quad \text{para todo} \quad t \in (-\delta, \delta). \tag{1}$$

En particular, $(-\delta, \delta) \subseteq E - E$.

Prueba. Fijemos $t \in \mathbb{R}$. Nuestro primer objetivo es demostrar que:

$(*)$ Si $F \subseteq (a,b)$ y $F \cap (t+F) = \varnothing$, entonces

$$2\mu(F) \ \leq \ (b-a) + |t|.$$

Veamos esto. Puesto que $\mu(F) = \mu(t+F)$, se tiene que si $F \cap (t+F) = \varnothing$, resulta entonces que $\mu(F \cup (t+F)) = 2\mu(F)$ y, en consecuencia, cualquier intervalo que contenga tanto a F, así como a $t+F$, tendrá longitud $\geq 2\mu(F)$. Observe ahora que si $t > 0$, entonces $(a, b+t)$ contiene a F y a su trasladado $t+F$. Similarmente, si $t < 0$, entonces $(a+t, b)$ contiene a ambos. En cualquier caso, la longitud de dicho intervalo es $(b-a) + |t|$.

Para demostrar (1) hagamos uso del Teorema de Densidad de Lebesgue, Teorema 9.1.33, para hallar un conjunto de medida cero $D \subseteq E$ tal que todos los puntos de $E \setminus D$ son puntos de densidad de Lebesgue, esto es, para todo $x \in E \setminus D$, se cumple que

$$\lim_{J \to x} \frac{\mu(E \cap J)}{\ell(J)} \ = \ 1.$$

Fijemos $x \in E \setminus D$ y con $\varepsilon = 1/2$, seleccione un intervalo abierto conteniendo a x, digamos $J = (a,b)$, tal que

$$\left| \frac{\mu(E \cap (a,b))}{\ell((a,b))} \ - \ 1 \right| \ < \ \frac{1}{2}.$$

En particular,

$$\frac{\mu(E \cap (a,b))}{\ell((a,b))} \ > \ \frac{1}{2}. \tag{2}$$

Pongamos $F = E \cap (a,b)$. De la desigualdad (2) se tiene que $\mu(F) > (b-a)/2$. Defina $\delta = 2\mu(F) - (b-a)$. Afirmamos que si $t \in (-\delta, \delta)$, entonces $E \cap (t+E) \neq \varnothing$. En efecto, suponga que $t \in (-\delta, \delta)$. Si ocurre que $|t| < 2\mu(F) - (b-a)$, entonces se sigue de $(*)$ que $F \cap (t+F) \neq \varnothing$ y, por lo tanto, $E \cap (t+E) \neq \varnothing$ que es lo que se quería demostrar.

Para demostrar que $(-\delta, \delta) \subseteq E - E$, sea $t \in (-\delta, \delta)$. Como $E \cap (t+E) \neq \varnothing$, entonces

$$\begin{aligned} x \in E \cap (t+E) \quad &\Leftrightarrow \quad x \in E \quad \text{y} \quad x \in t+E \\ &\Leftrightarrow \quad x \in E \quad \text{y} \quad x = t+e \quad \text{para algún} \quad e \in E \\ &\Leftrightarrow \quad x \in E \quad \text{y} \quad t = x-e \quad \text{para algún} \quad e \in E \\ &\Rightarrow \quad t = x - e \in E - E. \end{aligned}$$

Esto prueba que $(-\delta, \delta) \subseteq E - E$. ■

Ejercicio 9.2.2. *La función* $f : [0,1] \to \mathbb{R}$ *definida por*

$$f(x) \ = \ \begin{cases} x \operatorname{sen}\left(\dfrac{1}{x}\right) & si \quad x \in (0,1] \\ 0 & si \quad x = 0 \end{cases}$$

es una ***función continua*** *que no es de* ***variación acotada****.*

Prueba. Considere, para cada $n \in \mathbb{N}$, la partición $\mathcal{P}_n = \{0, x_1, x_2, \ldots, x_{n-1}, x_n, 1\}$, donde

$$x_k \ = \ \frac{1}{\pi(k+1/2)} \qquad \text{para} \ \ k = 0, 1, 2, \ldots, n.$$

Puesto que

$$|f(x_k) - f(x_{k-1})| \;=\; \frac{1}{\pi(k+1/2)} + \frac{1}{\pi(k-1/2)} \;>\; \frac{2}{k\pi}$$

para cada $k = 0, 1, 2, \ldots, n,$ resulta que

$$\begin{aligned} V(f, \mathcal{P}_n) &= |f(x_n) - f(x_0)| + |f(x_{n-1}) - f(x_n)| + \cdots + |f(1) - f(x_0)| \\ &\geq \sum_{k=1}^{n} |f(x_{k-1}) - f(x_k)| \\ &> \frac{2}{\pi}\sum_{k=1}^{n}\frac{1}{k} \end{aligned}$$

y como la serie $\sum_{k=1}^{\infty} 1/k$ diverge, se tiene que f no puede ser de variación acotada, aunque dicha función es continua en $[0, 1]$. ■

Ejercicio 9.2.3. *Sea $f : [a, b] \to \mathbb{R}$ una función que satisface la* PVI*. Demuestre que si $f \in \mathrm{BV}([a, b])$, entonces f es* ***continua*** *sobre $[a, b]$.*

Prueba. Suponga, para obtener una contradicción, que f es discontinua en algún punto de $[a, b]$. Puesto que f posee la PVI, el Teorema 3.1.33 nos asegura que las discontinuidades de f son de la segunda especie. Por otro lado, como f es de variación acotada, el Teorema de Jordan nos revela que ella es la diferencia de dos funciones crecientes y como toda función creciente posee, según el Teorema 3.1.34, sólo discontinuidades de la primera especie se sigue que las discontinuidades de f son de la primera especie. Esta contradicción establece que f no puede poseer puntos de discontinuidad o, lo que es lo mismo, f es continua sobre $[a, b]$. ■

Ejercicio 9.2.4. *La función $f : [0, 1] \to \mathbb{R}$ definida por*

$$f(x) \;=\; \begin{cases} x^{3/2}\cos\left(\dfrac{\pi}{x}\right) & si \quad 0 < x \leq 1 \\ 0 & si \quad x = 0. \end{cases}$$

es de ***variación acotada*** *sobre $[0, 1]$*

Prueba. Veamos que f es diferenciable en $[0, 1]$. En efecto, para cada $x \in (0, 1]$, la función $f(x)$ tiene derivada

$$f'(x) \;=\; \frac{3}{2}\,x^{1/2}\cos\left(\frac{\pi}{x}\right) + \frac{\pi}{2}\,\mathrm{sen}\left(\frac{\pi}{x}\right)$$

y en $x = 0$,

$$f'(0) \;=\; \lim_{x\to 0}\frac{f(x) - f(0)}{x - 0} \;=\; \lim_{x\to 0} x^{1/2}\cos\left(\frac{\pi}{x}\right) \;=\; 0.$$

Además, para cualquier $x \in [0, 1]$,

$$|f'(x)| \;\leq\; \left|\frac{3}{2}\,x^{1/2}\cos\left(\frac{\pi}{x}\right)\right| + \left|\frac{\pi}{2}\,\mathrm{sen}\left(\frac{\pi}{x}\right)\right| \;\leq\; \frac{3}{2} + \frac{\pi}{2}.$$

Esto prueba que f' es acotada y se sigue del Corolario 9.1.16 que $f \in \mathrm{BV}([0, 1])$.

Ejercicio 9.2.5. *Sea* $f : [a,b] \to \mathbb{R}$ *una función. Son equivalentes*

(1) $f \in \operatorname{Reg}([a,b])$.

(2) *Existe una sucesión* $(s_n)_{n=1}^{\infty}$ *de funciones en escaleras tal que* $s_n \to f$ ***uniformemente*** *sobre* $[a,b]$.

Prueba. (1) $\Rightarrow$ (2). Suponga que $f \in \operatorname{Reg}([a,b])$ y fijemos un $x \in [a,b]$. Puesto que $f \in \operatorname{Reg}([a,b])$, resulta que los límites

$$f(x^-) = \lim_{t \to x-} f(t) \qquad \text{y} \qquad f(x^+) = \lim_{t \to x+} f(t)$$

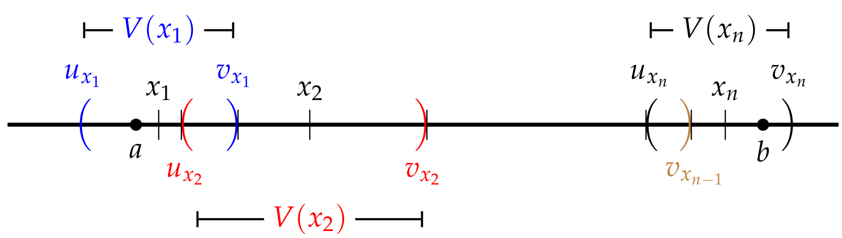
existen y, en consecuencia, para cada $n \in \mathbb{N}$, se pueden determinar intervalos abiertos no vacíos $(u_x, x) \cap [a,b]$ y $(x, v_x) \cap [a,b]$ tales que

$$|f(s) - f(t)| < 1/n$$

siempre que $s,t \in (u_x, x) \cap [a,b]$ o $s,t \in (x, v_x) \cap [a,b]$. Si ahora definimos $V(x) = (u_x, v_x)$, resulta que $V(x)$ es un intervalo abierto conteniendo a x y, en consecuencia, la familia $\mathcal{V} = \{V(x) : x \in (a,b)\}$ cubre a $[a,b]$. Por compacidad, podemos escoger un subcubrimiento finito de $\mathcal{V}$, digamos $V(x_1), \ldots, V(x_n)$. Para simplificar la presentación, asumiremos que los x_i han sido elegidos de modo que:

(i) $x_1 < x_2 < \cdots < x_n$ y

(ii) ningún $V(x_i)$ está contenido en la unión de los restantes.



Una vez aceptado lo anterior, considere la partición $\mathcal{P} = \{y_0, y_1, \ldots, y_m\}$, donde

$$a = y_0,\ y_1 = x_1,\ y_2 = u_{x_2},\ y_3 = v_{x_1}\ \ldots,\ y_{m-2} = v_{x_{n-1}},\ y_{m-1} = x_n,\ y_m = b.$$

Observe que, por construcción, si $s,t \in (y_{i-1}, y_i)$, entonces $|f(s) - f(t)| < 1/n$. Defina ahora $s_n : [a,b] \to \mathbb{R}$ por

$$s_n(x) = \begin{cases} f(x) & \text{si} \quad x \in \{y_0, y_1, \ldots, y_m\}, \\ f((y_{i-1} + y_i)/2) & \text{si} \quad y_{i-1} < x < y_i, \quad i = 1, 2, \ldots, m \end{cases}$$

De esta definición resulta claro que $|f(x) - s_n(x)| < 1/n$ para todo $x \in [a,b]$ y todo $n \geq 1$. Esto prueba que $s_n \to f$ uniformemente sobre $[a,b]$.

Para demostrar que (2) implica (1), suponga que existe una sucesión $(s_n)_{n=1}^{\infty}$ de funciones en escaleras tal que $s_n \to f$ uniformemente sobre $[a,b]$. Esto significa que, dado $\varepsilon > 0$ existe un s_{n_0} tal que

$$|f(z) - s_{n_0}(z)| < \frac{\varepsilon}{3} \tag{1a}$$

para todo $z \in [a,b]$. Sea $x \in [a,b)$ y pongamos $s = s_{n_0}$. Como $s \in \text{Reg}([a,b])$, entonces el límite $s(x^+) = \text{lím}_{z \to x+} s(z)$ existe y, por consiguiente, podemos encontrar un $y \in [a,b]$ con $x < y \leq b$ tal que

$$|s(z) - s(z')| < \frac{\varepsilon}{3} \quad \text{para todo} \quad z, z' \in (x,y).$$

De esto último y $(1a)$ se sigue que para todo $z, z' \in (x,y)$,

$$\begin{aligned} |f(z) - f(z')| &\leq |f(z) - s(z)| + |s(z) - s(z')| + |s(z') - f(z')| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \end{aligned}$$

Por el Criterio de Cauchy se tiene que $f(x^+)$ existe. Hemos demostrado que $f(x^+)$ existe para cualquier $x \in [a,b)$. De la misma manera se prueba que $f(x^-)$ existe para cualquier $x \in (a,b]$. ■

Ejercicio 9.2.6. *Si* $f \in \text{Reg}([a,b])$, *entonces* $\text{Disc}(f)$ *es* ***a lo más numerable****. En particular,*

$$\text{Reg}([a,b]) \subseteq \mathcal{R}([a,b])$$

Prueba. Sea $f \in \text{Reg}([a,b])$. Por el resultado anterior, existe una sucesión $(s_n)_{n=1}^{\infty}$ de funciones en escaleras tal que $s_n \to f$ uniformemente. Puesto que cada s_n es discontinua en un número finito de puntos, resulta que $\bigcup_{n=1}^{\infty} \text{Disc}(s_n)$ es a lo más numerable. Más aun, como f es continua precisamente en los puntos donde cada s_n es continua, resulta que $\bigcap_{n=1}^{\infty} \text{PC}(s_n) \subseteq \text{PC}(f)$, de donde se sigue

$$\text{Disc}(f) \subseteq \bigcup_{n=1}^{\infty} \text{Disc}(s_n).$$

Para demostrar la última parte, observe que como cada $s_n \in \mathcal{R}([a,b])$ para todo $n \in \mathbb{N}$ y la convergencia es uniforme, resulta del Teorema 8.3.25 que $f \in \mathcal{R}([a,b])$.

Otra manera más directa de ver que $\text{Reg}([a,b]) \subseteq \mathcal{R}([a,b])$ consiste en observar que si $f \in \text{Reg}([a,b])$, entonces $\mu(\text{Disc}(f)) = 0$ y luego invocar el Teorema de Vitali-Lebesgue para concluir que $f \in \mathcal{R}([a,b])$. ■

El siguiente ejemplo muestra que $\text{Reg}([a,b]) \neq \mathcal{R}([a,b])$. En efecto, si consideramos la función $f : [0,1] \to \mathbb{R}$ definida por

$$f(x) = \begin{cases} 1 & \text{si} \quad x = 1/n, \quad n \in \mathbb{N}, \\ 0 & \text{si} \quad x \in [0,1] \setminus \{1, 1/2, 1/3, \ldots\} \end{cases}$$

entonces se verifica, por el Teorema de Vitali-Lebesgue, que $f \in \mathcal{R}([a,b])$ pues $\mu(\text{Disc}(f)) = 0$ ya que $\text{Disc}(f) = \{1, 1/2, 1/3, \ldots\}$. Sin embargo, $f \notin \text{Reg}([a,b])$. Para ver esto último, suponga que $f \in \text{Reg}([a,b])$ y sea $0 < \varepsilon < 1/2$. Escojamos una función en escalera $\varphi = \sum_{i=1}^{n} \chi_{I_i}$ donde $I_i = [x_{i-1}, x_i]$ $i = 1, \ldots, n$ tal que

$$|f(x) - \varphi(x)| < \varepsilon \quad \text{para todo } x \in [a,b]$$

Sea n_0 el entero más pequeño tal que $n_0 x_1 > 1$, el cual existe gracias al Principio de Arquímedes, y pongamos $a_1 = 1/n_0$. Seleccione, finalmente, cualquier número $a_2 \in (a_1, x_1)$. Se sigue de la definición de f que $f(a_1) = 0$, $f(a_2) = 1$, y se cumple que

$$|c_1| = |c_1 - 0| = |\varphi(a_1) - f(a_1)| < \varepsilon \quad \text{y} \quad |1 - c_1| = |f(a_2) - \varphi(a_2)| < \varepsilon. \tag{1}$$

Ahora bien, como $|c_1| + |1 - c_1| \geq |c_1 + 1 - c_1| = 1$, entonces uno de los números $|c_1|$ o $|1 - c_1|$ debe ser mayor que $1/2$ y, en consecuencia, por (1), debemos tener que $\varepsilon \geq 1/2$. Esta contradicción establece que f no puede ser una función regulada.

Se sigue del ejemplo anterior que si $f \in \mathrm{Reg}([a, b])$, entonces f es Riemann integrable y, por lo tanto, $\int_a^b f\,dx$ existe. El siguiente es un argumento que permite justificar la existencia de una integral para cualquier $f \in \mathrm{Reg}([a, b])$, sin pasar por $\mathcal{R}([a, b])$ y que puede ser usada para construir la integral de Riemann.

Ejercicio 9.2.7. *Si $f \in \mathrm{Reg}([a, b])$, entonces existe una sucesión $(s_n)_{n=1}^{\infty}$ en $\mathrm{Esc}([a, b])$ que converge uniformemente sobre $[a, b]$ a f tal que*

$$\int_a^b f\,dx \;=\; \lim_{n\to\infty} \int_a^b s_n\,dx.$$

Prueba. Dotemos al espacio vectorial $\mathrm{Esc}([a, b])$ con la métrica del supremo, esto es, si $\varphi, \phi \in \mathrm{Esc}([a, b])$, entonces $d_\infty(\varphi, \phi) = \sup_{x\in[a,b]} |\varphi(x) - \phi(x)|$. Resulta del Ejercicio 9.2.5 que

$$\overline{\mathrm{Esc}([a, b])}^{\,d_\infty} \;=\; \mathrm{Reg}([a, b])$$

Sea $s \in \mathrm{Esc}([a, b])$. Por definición, existe una partición $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ de $[a, b]$ tal que $s = \sum_{i=1}^{n} c_i \chi_{I_i}$, donde $I_i = (x_{i-1}, x_i)$ para $i = 1, 2, \ldots, n$. Pongamos $\|s\|_\infty = \sup\{|c_i| : i = 1, \ldots, n\}$ y defina la aplicación $\Phi : \mathrm{Esc}([a, b]) \to \mathbb{R}$ por

$$\Phi(s) \;=\; \sum_{i=1}^{n} c_i (x_i - x_{i-1}).$$

Observe que como $|\Phi(s)| \;\leq\; \|s\|_\infty (b - a)$, resulta que Φ es una aplicación uniformemente continua y, además, lineal. Se sigue del Teorema 3.1.16 que Φ admite una única extensión (lineal) uniformemente continua

$$\widehat{\Phi} : \mathrm{Reg}([a, b]) \to \mathbb{R}$$

tal que $\widehat{\Phi}(s) = \Phi(s)$ para toda $s \in \mathrm{Esc}([a, b])$. Esto permite definir la **integral** de cualquier $f \in \mathrm{Reg}([a, b])$ como

$$\int_a^b f\,dx \;:=\; \widehat{\Phi}(f).$$

Resulta de lo anterior que

$$\int_a^b f\,dx \;=\; \lim_{n\to\infty} \int_a^b s_n\,dx \tag{1}$$

donde $(s_n)_{n=1}^{\infty}$ es cualquier sucesión en $\mathrm{Esc}([a, b])$ que converge uniformemente a f. Observe que, por construcción, si $(s'_n)_{n=1}^{\infty}$ es cualquier otra sucesión en $\mathrm{Esc}([a, b])$ que converge uniformemente a f, entonces

$$\lim_{n\to\infty} \int_a^b s_n\,dx \;=\; \lim_{n\to\infty} \int_a^b s'_n\,dx.$$

Ésta integral es trivialmente lineal, y satisface casi todas las propiedades interesantes que posee la integral de Riemann. Por ejemplo:

(a) Toda $f \in \mathrm{Reg}([a, b])$ posee una $\aleph_0$-primitiva Para ver esto, sólo es suficiente, gracias al Ejercicio 9.2.5, probar el resultado para funciones en escaleras. Suponga entonces que f es de la forma $f = \sum_{i=1}^{n} c_i \chi_{I_i}$ y defina $F : [a,b] \to \mathbb{R}$ de modo que en cada intervalo $I_k = (x_{k-1}, x_k)$ $(1 \le k \le n)$ sea

$$F(x) \;=\; c_k(x - x_{k-1}) \;+\; \sum_{i=1}^{k-1} c_i(x_i - x_{i-1}).$$

Entonces se verifica que F es una $\aleph_0$-primitiva de f.

(b) Si $f \in \operatorname{Reg}([a,b])$ y si F es una $\aleph_0$-primitiva de f, entonces

$$\int_a^b f\,dx \;=\; F(b) - F(a).$$

9.3. Problemas

(1) Sean $f, g : \mathbb{R} \to \mathbb{R}$. Pruebe que $\mathrm{D}^+(f+g)(x) \le \mathrm{D}^+ f(x) + \mathrm{D}^+ g(x)$. De un ejemplo para ilustrar que la desigualdad anterior puede ser estricta.

(2) Pruebe que ninguna función continua $f : \mathbb{R} \to \mathbb{R}$ puede tener $\mathrm{D}^+ f(x) = \infty$ para todo $x \in \mathbb{R}$. De un ejemplo de una función $f : \mathbb{R} \to \mathbb{R}$ tal que $\mathrm{D}^+ f(x) = \infty$ para todo $x \in \mathbb{R}$.

(3) Demuestre que el conjunto

$$\{x \in \mathbb{R} : \mathrm{D}^+ f(x) < \mathrm{D}_- f(x)\}$$

no puede ser no-numerable.

(4) Sea $f : \mathbb{R} \to \mathbb{R}$ una función y sea $\lambda \in [-\infty, +\infty]$. λ es llamado un **número derivado** para f en $x_0 \in \mathbb{R}$ si existe una sucesión $(x_n)_{n=1}^{\infty}$ con $\lim_{n\to\infty} x_n = x_0$ tal que

$$\lambda \;=\; \lim_{n\to\infty} \frac{f(x_n) - f(x_0)}{x_n - x_0}.$$

Demuestre que f tiene una derivada en x_0 si, y sólo si, todos sus números derivados coinciden.

(5) Sea $\mathcal{F}$ una familia de funciones M-Lipschitz sobre un espacio métrico (X, d). Pruebe que la función

$$F(x) \;=\; \inf\{f(x) \;:\; f \in \mathcal{F}\}, \quad x \in X,$$

es M-Lipschitz sobre X siempre que $F(x)$ sea finita en algún punto $x \in X$. Deduzca de esto último que toda función M-Lipschitz $f : A \to \mathbb{R}$, donde $A \subseteq X$, se puede extender a una función M-Lipschitz $F : X \to \mathbb{R}$.

(6) Una función $f : [a,b] \to \mathbb{R}$ se llama α-**Hölder** si existen constantes $\alpha > 0$ y $M > 0$ tales que, para todo $x, y \in [a,b]$,

$$|f(x) - f(y)| \;\le\; M|x-y|^{\alpha}.$$

Pruebe que:

(i) Si $\alpha > 1$, entonces f' existe y es cero en $[a,b]$. En particular, f es una función constante.

(ii) Si $0 < \alpha < 1$, existe una función α-Hölder que no es diferenciable en ningún punto de $[a,b]$.

(iii) Si $\alpha = 1$, entonces f es diferenciable casi-siempre.

(7) Sea $f \in \mathrm{BV}([a,b])$. Demuestre que si $f \geq c$ sobre $[a,b]$ para alguna constante $c > 0$, entonces $\frac{1}{f} \in \mathrm{BV}([a,b])$.

(8) Demuestre que si $f, g \in \mathrm{BV}([a,b])$, entonces $fg \in \mathrm{BV}([a,b])$ y

$$V(fg,[a,b]) \;\leq\; V(f,[a,b])\,\|g\|_\infty \;+\; V(g,[a,b])\,\|f\|_\infty.$$

(9) Sean $f_n, f \in \mathrm{BV}([a,b])$ para todo $n \in \mathbb{N}$. Si $f(x) = \lim_{n\to\infty} f_n(x)$ para cada $x \in [a,b]$, entonces pruebe que

$$V(f,[a,b]) \;\leq\; \liminf_{n\to\infty} V(f_n,[a,b]).$$

(10) Demuestre que si $f \in \mathrm{BV}([a,b]) \cap C([a,b])$, entonces

$$\lim_{n\to\infty} \sum_{i=1}^{n} \left| f\left(\frac{i}{n}\right) \;-\; f\left(\frac{i-1}{n}\right)\right|^2 \;=\; 0.$$

(11) Pruebe que si $f \in \mathrm{Lip}([a,b])$ con constante de Lipschitz M, entonces

$$\mu^*(f(E)) \;\leq\; M\mu^*(E)$$

para cualquier conjunto $E \subseteq [a,b]$.

(12) Sea $f : [a,b] \to \mathbb{R}$ una **función continua** y de **variación acotada** sobre $[a,b]$. Pruebe que

$$V(f,[a,b]) \;=\; \lim_{h\to 0+} \frac{1}{h} \int_a^{b-h} |f(x+h) - f(x)|\, dx.$$

(13) Sea $f : [a,b] \to \mathbb{R}$ una **función diferenciable** y de **variación acotada**. Pruebe que si f' es **continua** en $x_0 \in [a,b]$, entonces V_f es **diferenciable** en x_0.

(14) La función $f : [0,2] \to \mathbb{R}$ definida por

$$f(x) \;=\; \begin{cases} x \operatorname{sen}\left(\dfrac{\pi}{x}\right) & \text{si} \quad 0 < x \leq 2 \\ 0 & \text{si} \quad x = 0 \end{cases}$$

es Riemann integrable pero no es de variación acotada.

(15) Sea $f, g : [a,b] \to \mathbb{R}$ funciones y suponga que existen constantes $M, c > 0$ tales que

$$|f(x) - f(y)| \;\leq\; M|g(x) - g(y)|^c \qquad \text{para todo} \quad x, y \in [a,b].$$

Demuestre que $f \in \mathrm{BV}([a,b])$ siempre que $g \in BV([a,b])$.

(16) Sea $(f_n)_{n=1}^\infty$ una sucesión de funciones en $\mathrm{BV}([a,b])$. Suponga que existe una constante $M > 0$ tal que $V(f_n,[a,b]) \leq M$ para todo $n \in \mathbb{N}$. Pruebe que si $f_n(x) \to f(x)$ para cada $x \in [a,b]$, entonces $f \in \mathrm{BV}([a,b])$ y $V(f,[a,b]) \leq M$.

(17) Sea $f : [0,1] \to \mathbb{R}$ definida por

$$f(x) \;=\; \sum_{n=1}^{\infty} a_n x^n,$$

donde $\sum_{n=1}^\infty |a_n| < +\infty$. Demuestre que $f \in \mathrm{BV}([a,b])$ y $V(f,[0,1]) \leq \sum_{n=1}^\infty |a_n|$.

(18) Sea Γ el conjunto ternario de Cantor. Pruebe que si $f : \Gamma \to \mathbb{R}$ es una función continua y $\mu(f(\Gamma)) = 0$, entonces f se puede extender a una función continua sobre $[0,1]$ satisfaciendo la condición (N) de Lusin. (Ayuda: haga f diferenciable fuera de Γ).

(19) Exprese cada $x \in \Gamma$ en la forma

$$x = \sum_{n=1}^{\infty} \frac{2c_n(x)}{3^n},$$

donde $c_n(x) \in \{0,1\}$ y construya dos funciones continuas $f_1, f_2 : [0,1] \to \mathbb{R}$ satisfaciendo la condición (N) de Lusin tal que

$$f_1(x) = \sum_{n=1}^{\infty} \frac{c_{2n}(x)}{3^n} \qquad \text{y} \qquad f_2(x) = \sum_{n=1}^{\infty} \frac{c_{2n+1}(x)}{3^n}$$

para todo $x \in \Gamma$. Pruebe, además, que $f_1 + f_2$ no satisface la condición (N) de Lusin.

(20) Sea $f : [a,b] \to \mathbb{R}$ una función. Pruebe que las siguientes afirmaciones son equivalentes:

(a) f es **absolutamente continua** sobre $[a,b]$.

(b) Para cada $\varepsilon > 0$, existe un $\delta > 0$ tal que $\sum_{i=1}^{n} |f(b_i) - f(a_i)| < \varepsilon$ siempre que $\{(a_i, b_i) : i = 1, \ldots, n\}$ es una colección finita y disjunta de subintervalos abiertos incluidos en $[a,b]$ satisfaciendo $\sum_{i=1}^{n}(b_i - a_i) < \delta$.

(c) Para cada $\varepsilon > 0$, existe un $\delta > 0$ tal que $\sum_{i=1}^{\infty} |f(b_i) - f(a_i)| < \varepsilon$ siempre que $\{(a_i, b_i) : i \in \mathbb{N}\}$ es una colección infinita y disjunta de intervalos abiertos incluidos en $[a,b]$ satisfaciendo $\sum_{i=1}^{\infty}(b_i - a_i) < \delta$.

(d) Para cada $\varepsilon > 0$, existe un $\delta > 0$ tal que $\sum_{i=1}^{\infty} \operatorname{osc}(f, [a_i, b_i]) < \varepsilon$ siempre que $\{(a_i, b_i) : i \in \mathbb{N}\}$ es una colección infinita y disjunta de subintervalos abiertos incluidos en $[a,b]$ satisfaciendo $\sum_{i=1}^{\infty}(b_i - a_i) < \delta$.

Nota. Demuestre que si reemplazamos el intervalo $[a,b]$ por $\mathbb{R}$, las equivalencias dadas anteriormente siguen siendo validas para cualquier función $f : \mathbb{R} \to \mathbb{R}$ que sea absolutamente continua.

(21) Sea $f : [0,1] \to \mathbb{R}$ definida por

$$f(x) = \begin{cases} x^2 \left|\operatorname{sen}\left(\dfrac{1}{x}\right)\right| & \text{si } \ x \in (0,1], \\ 0 & \text{si } \ x = 0. \end{cases}$$

y considere la función $g : [0,1] \to \mathbb{R}$ dada por $g(x) = \sqrt{x}$. Pruebe que f y g son absolutamente continuas y que la compuesta $f \circ g$ también lo es, pero que $g \circ f$ no es absolutamente continua.

(22) Pruebe que si f y g son absolutamente continuas y, además, g es monótona, entonces $f \circ g$ es absolutamente continua.

(23) Sean f y g **funciones absolutamente continuas**. Pruebe que $f \circ g$ es **absolutamente continua** si, y sólo si, ella es de **variación acotada**.

CAPÍTULO 10

La Integral de Lebesgue

El ejemplo de Volterra de una función diferenciable $f : [0,1] \to \mathbb{R}$, con f' acotada pero no Riemann integral, le permitió a Henri Léon Lebesgue (1875-1941) pensar en profundidad la noción de integral que hasta ese momento se había desarrollado. Ese ejemplo le condujo a observar que diferenciación e integración no pueden considerarse como procesos inversos en el marco de la teoría de las funciones Riemann integrables, de allí que emprende la tarea de definir una nueva integral que corrija las deficiencias presentadas por la integral de Riemann. De modo incuestionable, su integral fue capaz de integrar cualquier función medible acotada, en particular, cualquier derivada acotada restituyendo, para las derivadas acotadas, el paraíso propuesto por la fórmula de Newton-Leibniz. Pero, además, logra probar que su integral es de mayor alcance que la de Riemann al demostrar sus potentes resultados referentes a la convergencia puntual de funciones integrables según Lebesgue. Su idea simple, pero brillante, fue considerar *particiones en el rango* de una función acotada f, en lugar de las particiones en el dominio tal como lo habían hecho Cauchy y Riemann. Es decir, Lebesgue abandona el enfoque de sus predecesores al demostrar que haciendo particiones en el dominio de un intervalo $[a,b]$ cada vez más fina no da buenos resultados para funciones que poseen demasiadas discontinuidades. Tomar particiones en el rango y considerar imágenes inversas de tales conjuntos le obligó a considerar conjuntos más complicados que los intervalos en el dominio de la función dada lo que le condujo, a su vez, a asignar a cada uno de tales conjuntos, una "*longitud*" o "medida". Esta circunstancia lo lleva a desarrollar la Teoría de la Medida y, por ende, el de las Funciones Medibles. Ese enfoque, definitivamente, marcó la diferencia. Esta historia que ya ha sido contada innumerable veces y que, una vez más, contaremos en estas notas forman parte de esos momentos de genialidad que algunas mentes brillantes, como la de Henri Leon Lebesgue, desarrolló por más de una década para legarnos una de las nociones más potente, fascinante y hermosa de las matemáticas: *la integral de Lebesgue*.

10.1. Construcción de la Integral de Lebesgue

La integral de Lebesgue posee varios enfoques para su construcción. H. Lebesgue, siguiendo las ideas de E. Borel, pensó el problema de integración (*le problème d'intégration*) del modo siguiente: asignar, a cada función acotada $f : [a, b] \to \mathbb{R}$ un único número que satisfaga ciertas condiciones. Una vez descritas las propiedades que él quiere que su "integral" posea, intenta deducir la integral de esas propiedades. El nombró a tal proceso una *definición descriptiva*, en contraste con las *definiciones constructivas*, tal como la desarrollada para obtener la integral de Riemann. Sin embargo, en estas notas, seguiremos el enfoque constructivo, es decir, nos ocuparemos de definir, en primer lugar, la integral de Lebesgue para funciones medibles y acotadas $f : X \to \mathbb{R}$, donde X es cualquier conjunto medible de medida finita similar a la de Riemann pero tomando particiones en el rango en lugar de hacerlo en el dominio de la función como lo hizo Riemann. Una vez mostradas las poderosas propiedades que dicha integral posee, se extiende dicha noción de integral de Lebesgue a cualquier función medible, sea o no acotada, y definida sobre un conjunto medible arbitrario.

Primera Suposición. **Salvo que se indique lo contrario, $X \subseteq \mathbb{R}$ siempre denotará un subconjunto medible con $\mu(X) < +\infty$ y $f : X \to \mathbb{R}$ una función medible y acotada**.

Sea $f : X \to \mathbb{R}$ una función medible y acotada con $\mu(X) < +\infty$ y sean m, M números reales tales que $m \leq f(x) \leq M$ para todo $x \in X$. Siempre podemos suponer, y así lo haremos, que los números m y M están definidos por:

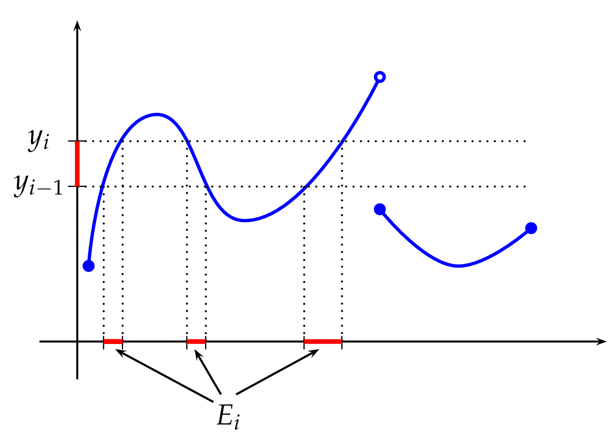
$$m = \inf\{f(x) : x \in X\} \qquad \text{y} \qquad M = \sup\{f(x) : x \in X\}.$$

Considere una partición arbitraria $\mathcal{P} = \{y_0, y_1, \ldots, y_n\}$ del intervalo $[m, M]$, donde

$$m = y_0 < y_1 < \cdots < y_n = M$$

y defina los conjuntos

$$E_i = X \cap f^{-1}\big([y_{i-1}, y_i)\big), \quad i = 1, \ldots, n.$$



Estas consideraciones *siempre las asumiremos* en esta sección. Puesto que f es una función medible, se sigue que los conjuntos $E_1, E_2, \ldots, E_n$ son medibles y disjuntos dos a dos. Más aun,

(α) $X = \bigcup_{i=1}^{n} E_i$,

(β) $\mu(X) = \sum_{i=1}^{n} \mu(E_i)$ y

(γ) $y_{i-1} \leq f(x) < y_i$ para todo $x \in E_i$.

Con esta información se introducen, como en el caso de la integral de Riemann, las **sumas inferior** y **superior de Lebesgue** para f, respectivamente, como

$$\mathrm{L}_\mu(f,\mathcal{P}) = \sum_{i=1}^{n} y_{i-1} \cdot \mu(E_i) \qquad \text{y} \qquad \mathrm{U}_\mu(f,\mathcal{P}) = \sum_{i=1}^{n} y_i \cdot \mu(E_i),$$

Observe que

$$m \cdot \mu(X) \leq \mathrm{L}_\mu(f,\mathcal{P}) \leq \mathrm{U}_\mu(f,\mathcal{P}) \leq M \cdot \mu(X) \tag{L1}$$

y si definimos $\|\mathcal{P}\| = \text{máx}\{(y_i - y_{i-1}) : i = 1,\dots,n\}$, entonces se cumple que

$$0 \leq \mathrm{U}_\mu(f,\mathcal{P}) - \mathrm{L}_\mu(f,\mathcal{P}) \leq \|\mathcal{P}\| \cdot \mu(X). \tag{L2}$$

Algunas observaciones sobre las sumas de Lebesgue, análogas a las sumas de Riemann, son requeridas para poder definir la integral de Lebesgue.

Teorema 10.1.1. *Si $\mathcal{P},\mathcal{Q}$ son **particiones** de $[m,M]$ con $\mathcal{P} \subseteq \mathcal{Q}$, entonces*

$$\mathrm{L}_\mu(f,\mathcal{P}) \leq \mathrm{L}_\mu(f,\mathcal{Q}) \qquad y \qquad \mathrm{U}_\mu(f,\mathcal{P}) \geq \mathrm{U}_\mu(f,\mathcal{Q}).$$

Prueba. Suponga, en principio, que $\mathcal{Q}$ contiene exactamente un punto adicional, digamos y, que no está en $\mathcal{P}$ y que $y \in (y_{i-1}, y_i)$. Entonces $\mathrm{U}_\mu(f,\mathcal{P}) - \mathrm{U}_\mu(f,\mathcal{Q})$ consiste de tres términos. Específicamente, si

$$A = X \cap f^{-1}([y_{i-1}, y)) \qquad \text{y} \qquad B = X \cap f^{-1}([y, y_i)),$$

entonces $E_i = A \cup B$ y

$$\begin{aligned} \mathrm{U}_\mu(f,\mathcal{P}) - \mathrm{U}_\mu(f,\mathcal{Q}) &= y_i \cdot \mu(E_i) - y \cdot \mu(A) - y_i \cdot \mu(B) \\ &\geq y_i \cdot \mu(E_i) - y_i \cdot \mu(A) - y_i \cdot \mu(B) \\ &= 0. \end{aligned}$$

Esto prueba que $\mathrm{U}_\mu(f,\mathcal{P}) \geq \mathrm{U}_\mu(f,\mathcal{Q})$ para este caso particular. Suponga ahora que $\mathcal{Q}$ contiene k puntos más que $\mathcal{P}$. Entonces existe un conjunto finito de particiones $\mathcal{Q}_1,\dots,\mathcal{Q}_{k-1}$ de $[a,b]$ tal que

$$\mathcal{P} \subseteq \mathcal{Q}_1 \subseteq \cdots \subseteq \mathcal{Q}_{k-1} \subseteq \mathcal{Q}$$

donde cada partición se obtiene de la anterior añadiéndole exactamente un punto. Entonces

$$\mathrm{U}_\mu(f,\mathcal{P}) \geq \mathrm{U}_\mu(f,\mathcal{Q}_1) \geq \cdots \geq \mathrm{U}_\mu(f,\mathcal{Q}_{k-1}) \geq \mathrm{U}_\mu(f,\mathcal{Q}).$$

De modo similar, se demuestra que $\mathrm{L}_\mu(f,\mathcal{P}) \leq \mathrm{L}_\mu(f,\mathcal{Q})$ y termina la prueba. ■

Nótese que si $\mathcal{P}$ y $\mathcal{Q}$ son particiones arbitrarias del intervalo $[m, M]$, entonces, tomando $\mathcal{R} = \mathcal{P} \cup \mathcal{Q}$, resulta del Teorema 10.1.1 que

$$\mathrm{L}_\mu(f, \mathcal{P}) \leq \mathrm{L}_\mu(f, \mathcal{R}) \leq \mathrm{U}_\mu(f, \mathcal{R}) \leq \mathrm{U}_\mu(f, \mathcal{Q}).$$

De esto y el Teorema 2.1.9, página 87, se deduce que

$$\sup_{\mathcal{P} \in \mathbb{P}[m,M]} \mathrm{L}_\mu(f, \mathcal{P}) \leq \inf_{\mathcal{Q} \in \mathbb{P}[m,M]} \mathrm{U}_\mu(f, \mathcal{Q}).$$

Definición 10.1.2. *Sea $f : X \to \mathbb{R}$ una función medible y acotada sobre X. Las **integrales inferior** y **superior de Lebesgue** de f se definen, respectivamente, como*

$$\mathcal{L}_\mu(f) = \sup_{\mathcal{P} \in \mathbb{P}[m,M]} \mathrm{L}_\mu(f, \mathcal{P}) \qquad y \qquad \mathcal{U}_\mu(f) = \inf_{\mathcal{P} \in \mathbb{P}[m,M]} \mathrm{U}_\mu(f, \mathcal{P})$$

Fijemos una partición $\mathcal{P}$ de $[m, M]$ elegida de manera arbitraria. Observe entonces que, por (L1),

$$m \cdot \mu(X) \leq \mathrm{L}_\mu(f, \mathcal{P}) \leq \mathcal{L}_\mu(f) \leq \mathcal{U}_\mu(f) \leq \mathrm{U}_\mu(f, \mathcal{P}) \leq M \cdot \mu(X) \tag{L3}$$

y como por (L2),

$$0 \leq \mathrm{U}_\mu(f, \mathcal{P}) - \mathrm{L}_\mu(f, \mathcal{P}) \leq \|\mathcal{P}\| \cdot \mu(X),$$

resulta que

$$0 \leq \mathcal{U}_\mu(f) - \mathcal{L}_\mu(f) \leq \|\mathcal{P}\| \cdot \mu(X). \tag{L4}$$

Suponga ahora que $\mu(X) > 0$. Entonces, dado cualquier $\varepsilon > 0$, si elegimos nuestra partición $\mathcal{P}$ de modo tal que $\|\mathcal{P}\| < \varepsilon/\mu(X)$ tendremos, de la desigualdad anterior, que

$$0 \leq \mathcal{U}_\mu(f) - \mathcal{L}_\mu(f) < \varepsilon.$$

Puesto que la elección del $\varepsilon > 0$ se hizo de forma arbitraria, se concluye que

$$\mathcal{L}_\mu(f) = \mathcal{U}_\mu(f).$$

Observe que, gracias a (L 4), ésta última igualdad se cumple trivialmente si ocurre que $\mu(X) = 0$. Este análisis justifica la siguiente definición.

Definición 10.1.3. *Sea X un conjunto medible con $\mu(X) < +\infty$ y sea $f : X \to \mathbb{R}$ una función medible y acotada sobre X. Diremos que que f es **Lebesgue integrable** sobre X si $\mathcal{L}_\mu(f) = \mathcal{U}_\mu(f)$. En este caso, la **integral de Lebesgue** de f sobre X se define como*

$$(\mathrm{L})\int_X f\, d\mu = \mathcal{L}_\mu(f) = \mathcal{U}_\mu(f). \tag{IL1}$$

Cuando no exista ningún posibilidad de interpretar erróneamente la notación, usaremos el símbolo $\int_X f\, d\mu$ en lugar de $(\mathrm{L})\int_X f\, d\mu$, mientras que el uso de esta última estará reservada sólo cuando exista, además de la integral de Lebesgue, otra integral en el mismo contexto. En particular, si X es el intervalo cerrado $[a, b]$ usaremos uno de los símbolos

$$\int_{[a,b]} f\, d\mu \qquad \text{o} \qquad \int_a^b f\, d\mu$$

para denotar a la integral de Lebesgue.

Observe que si $\mu(X) < +\infty$ y $f : X \to \mathbb{R}$ es una función **medible** y **acotada** sobre X, entonces $|f|$ también es medible y acotada sobre X y, en consecuencia, de nuestra definición de la integral de Lebesgue tenemos que:

Teorema 10.1.4. *Sean $\mu(X) < +\infty$ y $f : X \to \mathbb{R}$ una función **medible** y **acotada** sobre $[a, b]$. Las siguientes condiciones son equivalentes:*

(1) *f es **Lebesgue integrable** sobre X.*

(2) *$|f|$ es **Lebesgue integrable** sobre X.*

Ya establecida la definición de integral de Lebesgue, tal vez sea propicia la ocasión para comparar, tal como lo hizo Lebesgue en 1927, la diferencia entre las integrales de Riemann y la suya: imaginemos que en una caja tenemos muchísimas monedas de distintas denominaciones y queremos contar cuánto dinero hay ella: para Lebesgue, integrar al estilo Riemann consistía en dividir la cantidad de monedas de la caja en n grupos, contar cuánto dinero hay en cada grupo y luego sumar cada una de esas cantidades. Por otro lado, integrar siguiendo el estilo Lebesgue consiste, en primer lugar, agrupar las monedas según sus denominaciones para formar n grupos y luego contar cuántas monedas de cada denominación hay en cada uno de los grupos: digamos que en el grupo 1 existen k_1 monedas de 1 bolívar, en el segundo grupo hay k_2 monedas de 2 bolívares, hasta llegar al n-ésimo grupo donde hay k_n monedas de n bolívares. Después sumamos $1 \cdot k_1 + 2 \cdot k_2 + \cdots + n \cdot k_n$. *Sin duda, la manera de integrar al estilo Lebesgue parece más eficiente*!

Nota Adicional 10.1.1 Ya vimos que cualquier función Riemann integrable es medible y como también es acotada, ella es *integrable en el sentido de Lebesgue*. Más aun, como veremos un poco más adelante, su integral coincide con la integral de Lebesgue. Estos hechos permiten adelantar la opinión de que la integral de Lebesgue es mucho más amplia que la integral de Riemann en el sentido de que toda función Riemann integrable es Lebesgue integrable, pero no cualquier función Lebesgue integrable es Riemann integrable como se muestra en el siguiente ejemplo:

Ejemplo 1. *La función $f : [0, 1] \to \mathbb{R}$ definida por $f(x) = \chi_{\mathbb{Q}}(x)$ es **Lebesgue integrable** pero no **Riemann integrable** sobre $[0, 1]$.*

Prueba. Puesto que el conjunto de puntos donde f es discontinua es $[0, 1]$, el Teorema de Vitali-Lebesgue nos dice que f no puede ser Riemann integrable, mientras que f, por ser medible y acotada, es Lebesgue integrable. ■

Definición 10.1.5. *Si $E \subseteq X$ es cualquier conjunto medible y $f : X \to \mathbb{R}$ es una función medible y acotada sobre X, definimos*

$$\int_E f \, d\mu \;=\; \int_X f \cdot \chi_E \, d\mu.$$

En lo que sigue, el símbolo $\mathcal{L}_{\text{fin}}(X, \mu)$ será usado para designar *el conjunto de todas las funciones medibles $f : X \to \mathbb{R}$ que son Lebesgue integrables sobre X, donde X es un conjunto medible de medida finita*. De lo expresado anteriormente se tiene que

$$\mathcal{R}([a, b]) \;\subseteq\; \mathcal{L}_{\text{fin}}([a, b], \mu)$$

10.1.1. Propiedades de la Integral de Lebesgue

En esta sección presentaremos algunas de las propiedades clásicas de la integral de Lebesgue.

Teorema 10.1.6. *Si $f \in \mathcal{L}_{\text{fin}}(X,\mu)$ con $m \leq f(x) \leq M$ para todo $x \in X$, entonces*

$$m \cdot \mu(X) \;\leq\; \int_X f\,d\mu \;\leq\; M \cdot \mu(X). \tag{L5}$$

*En particular, si $g : X \to \mathbb{R}$ es **acotada**, entonces $f \cdot g \in \mathcal{L}_{\text{fin}}(X,\mu)$.*

Prueba. Es consecuencia inmediata de $(L\,3)$. ■

Del resultado anterior se tiene que:

$$\mu(X) \;=\; 0 \quad \Rightarrow \quad \int_X f\,d\mu \;=\; 0. \tag{L6}$$

Este hecho, el cual es de una importancia fundamental en la teoría de integración, establece que la evaluación de una función Lebesgue integrable sobre cualquier conjunto de medida nula es cero y, por lo tanto, no afecta el proceso de integración, es decir, los conjuntos de medida cero se pueden considerar como despreciables en la Teoría de Integración de Lebesgue.

Otra consecuencia inmediata del Teorema 10.1.6 es la siguiente: si $f : X \to \mathbb{R}$ es una función constante, digamos $f(x) = k$, entonces $m = M = k$ y se sigue de $(L5)$ que

$$\int_X f\,d\mu \;=\; k \cdot \mu(X).$$

En particular, si $E \subseteq X$ es medible y $f(x) = 1$ para todo $x \in X$, entonces

$$\int_E d\mu \;=\; \int_X \chi_E\,d\mu \;=\; \mu(E).$$

Corolario 10.1.7. *Si $f \in \mathcal{L}_{\text{fin}}(X,\mu)$ con $f \geq 0$ sobre X, entonces*

$$\int_X f\,d\mu \;\geq\; 0. \tag{L7}$$

Prueba. Puesto que $m = \inf\{f(x) : x \in X\} \geq 0$, por una nueva aplicación de $(L5)$ se obtiene la desigualdad $(L7)$. ■

Teorema 10.1.8 (σ-aditividad de la integral). *Sea $f \in \mathcal{L}_{\text{fin}}(X,\mu)$. Si $(E_n)_{n=1}^{\infty}$ es una **sucesión disjunta** de subconjuntos medibles de X tal que $X = \bigcup_{k=1}^{\infty} E_k$, entonces*

$$\int_{\bigcup\limits_{k=1}^{\infty} E_k} f\,d\mu \;=\; \sum_{k=1}^{\infty} \int_{E_k} f\,d\mu.$$

Prueba. Suponga, en primer lugar, que $X = E_1 \cup E_2$ y sea $\mathcal{P} = \{y_0, y_1, \ldots, y_n\}$ una partición del rango $[m, M]$ de f. Para cada $j \in \{1, \ldots, n\}$, sea

$$A_j \;=\; X \cap f^{-1}\big([y_{j-1}, y_j)\big) \;=\; (E_1 \cup E_2) \cap f^{-1}\big([y_{j-1}, y_j)\big).$$

Resulta que los conjuntos

$$A_{1j} = E_1 \cap f^{-1}\big([y_{j-1}, y_j)\big) \qquad \text{y} \qquad A_{2j} = E_2 \cap f^{-1}\big([y_{j-1}, y_j)\big)$$

son medibles, disjuntos dos a dos y $A_j = A_{1j} \cup A_{2j}$ para $j = 1, \ldots, n$. Por esto,

$$\mathrm{U}_\mu(f, \mathcal{P}) \;=\; \sum_{j=1}^{n} y_j \cdot \mu(A_j) \;=\; \sum_{j=1}^{n} y_j \cdot \mu(A_{1j}) + \sum_{j=1}^{n} y_j \cdot \mu(A_{2j})$$

de donde se deduce que

$$\int_X f\, d\mu \;=\; \int_{E_1} f\, d\mu \;+\; \int_{E_2} f\, d\mu.$$

Un fácil argumento inductivo nos convence que la igualdad anterior es válida para cualquier colección finita $\{E_1, \ldots, E_n\}$ de conjuntos medibles, disjuntos dos a dos tal que $X = E_1 \cup \cdots \cup E_n$, esto es,

$$\int_{\bigcup_{j=1}^{n} E_j} f\, d\mu \;=\; \int_{E_1} f\, d\mu \;+\; \cdots \;+\; \int_{E_n} f\, d\mu. \tag{1}$$

Para demostrar el caso general, sea $(E_n)_{n=1}^{\infty}$ una sucesión disjunta de subconjuntos medibles de X tal que $X = \bigcup_{k=1}^{\infty} E_k$. Como

$$\mu(X) \;=\; \sum_{k=1}^{\infty} \mu(E_k) \;<\; +\infty,$$

resulta que

$$\lim_{n\to\infty} \sum_{k=n+1}^{\infty} \mu(E_k) \;=\; 0. \tag{2}$$

de modo que si definimos $R_n \;=\; \bigcup_{k=n+1}^{\infty} E_k$, entonces usando (1) se obtiene la igualdad

$$\int_{\bigcup_{k=1}^{\infty} E_k} f\, d\mu \;=\; \sum_{k=1}^{n} \int_{E_k} f\, d\mu \;+\; \int_{R_n} f\, d\mu. \tag{3}$$

Ahora bien, en virtud de $(\mathrm{L5})$, se tiene que

$$m \cdot \sum_{k=n+1}^{\infty} \mu(E_k) \;=\; m \cdot \mu(R_n) \;\leq\; \int_{R_n} f\, d\mu \;\leq\; M \cdot \mu(R_n) \;=\; M \cdot \sum_{k=n+1}^{\infty} \mu(E_k)$$

y así, gracias a (2), podemos derivar la igualdad

$$\lim_{n\to\infty} \int_{R_n} f\, d\mu \;=\; 0.$$

Tomando límite cuando $n \to \infty$ en (3) y teniendo en cuenta el resultado anterior, se concluye que

$$\int_{\bigcup\limits_{k=1}^{\infty} E_k} f\,d\mu \;=\; \sum_{k=1}^{\infty} \int_{E_k} f\,d\mu.$$

Esto termina la prueba. ∎

Usando los últimos dos resultados se concluye que la función de conjuntos ν_f asociada a f es una medida sobre X, en otras palabras:

Corolario 10.1.9. *Si $f : X \to \mathbb{R}$ es una función **medible no-negativa** y **Lebesgue integrable** sobre X, entonces la función de conjuntos $\nu_f : \mathfrak{M}_\mu(X) \to \mathbb{R}$ definida por*

$$\nu_f(E) \;=\; \int_E f\,d\mu, \quad E \in \mathfrak{M}_\mu(X).$$

*es una **medida** sobre X.*

Observe que la medida ν_f verifica la relación: $\mu(E) = 0$ implica que $\nu_f(E) = 0$. En particular, se sigue del la "continuidad de la medida" que:

Corolario 10.1.10. *Sea $f \in \mathcal{L}_{\text{fin}}(X,\mu)$ con $f \geq 0$. Si $(E_n)_{n=1}^{\infty}$ es una **sucesión creciente** de subconjuntos medibles de X, entonces*

$$\int_{\bigcup\limits_{n=1}^{\infty} E_n} f\,d\mu \;=\; \lim_{n\to\infty} \int_{E_n} f\,d\mu,$$

en otras palabras, $\nu_f(\bigcup_{n=1}^{\infty} E_n) = \lim\limits_{n\to\infty} \nu_f(E_n)$.

Prueba. Pongamos $A_1 = E_1$ y para cada $n \geq 2$ defina $A_n = E_n \setminus E_{n-1}$. Resulta que $(A_n)_{n=1}^{\infty}$ es una sucesión disjunta de conjuntos medibles satisfaciendo las siguientes propiedades:

(a) $\displaystyle\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} E_n$

(b) $\displaystyle\bigcup_{k=1}^{n} A_k = E_n$ para cada $n \geq 1$.

Se sigue de la σ-aditividad de la integral y las propiedades anteriores que

$$\begin{aligned}
\int_{\bigcup\limits_{n=1}^{\infty} E_n} f\,d\mu \;&=\; \int_{\bigcup\limits_{n=1}^{\infty} A_n} f\,d\mu \;=\; \sum_{n=1}^{\infty} \int_{A_n} f\,d\mu \\
&=\; \lim_{n\to\infty} \sum_{k=1}^{n} \int_{A_k} f\,d\mu \;=\; \lim_{n\to\infty} \int_{\bigcup\limits_{k=1}^{n} A_k} f\,d\mu \\
&=\; \lim_{n\to\infty} \int_{E_n} f\,d\mu.
\end{aligned}$$

La prueba es completa. ∎

Teorema 10.1.11. *$\mathcal{L}_{\text{fin}}(X,\mu)$ es un **espacio vectorial** sobre $\mathbb{R}$. Aún más, si $f, g \in \mathcal{L}_{\text{fin}}(X,\mu)$, entonces*

(a) $\displaystyle\int_X (f+g)\,d\mu \;=\; \int_X f\,d\mu + \int_X g\,d\mu.$

(b) $\displaystyle\int_X (af)\,d\mu \;=\; a\int_X f\,d\mu,$ *para cualquier* $a\in\mathbb{R}$.

Prueba. (a) Sean $\mathcal{P}=\{y_0,y_1,\ldots,y_n\}$ y $\mathcal{Q}=\{z_0,z_1,\ldots,z_m\}$ particiones del rango de f y g respectivamente y considere los conjuntos

$$A_k \;=\; X\cap f^{-1}\big([y_{i-1},y_i)\big), \qquad\qquad B_i \;=\; X\cap f^{-1}\big([z_{j-1},z_j)\big)$$

y

$$C_{ik} \;=\; B_i\cap A_k, \quad i=1,\ldots,m, \quad k=1,\ldots,n.$$

Cada una de las familias $\mathcal{A}=\{A_i : i=1,\ldots,n\}$ y $\mathcal{B}=\{B_j : j=1,\ldots,m\}$ son medibles y disjuntas, por lo que $\mathcal{C}=\{C_{ik} : i,k\}$ también es medible, disjunta y, además,

$$X \;=\; \bigcup_{i=1}^{m}\bigcup_{k=1}^{n} C_{ik}.$$

Se sigue del Teorema 10.1.8 que

$$\int_X (f+g)\,d\mu \;=\; \sum_{i=1}^{m}\sum_{k=1}^{n}\int_{C_{ik}}(f+g)\,d\mu$$

Fijemos i,k y observe que para cada $x\in C_{ik}$ se tiene que

$$y_{k-1}\,+\,z_{i-1} \;\le\; f(x)\,+\,g(x) \;<\; y_k\,+\,z_i$$

y entonces, por (L5), resulta que

$$(y_{k-1}\,+\,z_{i-1})\mu(C_{ik}) \;\le\; \int_{C_{ik}}(f+g)\,d\mu \;\le\; (y_k\,+\,z_i)\mu(C_{ik})$$

Combinando todas estas desigualdades, se obtiene

$$\sum_{i=1}^{m}\sum_{k=1}^{n}(y_{k-1}\,+\,z_{i-1})\cdot\mu(C_{ik}) \;\le\; \int_X (f+g)\,d\mu \;\le\; \sum_{i=1}^{m}\sum_{k=1}^{n}(y_k\,+\,z_i)\cdot\mu(C_{ik}). \tag{T1}$$

Evaluemos ahora cada una de las sumas que aparecen en la desigualdad anterior. Por ejemplo,

$$\sum_{i=1}^{m}\sum_{k=1}^{n} y_k\cdot\mu(C_{ik}) \;=\; \sum_{k=1}^{n} y_k\cdot\left(\sum_{i=1}^{m}\mu(C_{ik})\right)$$

y como

$$\begin{aligned}\sum_{i=1}^{m}\mu(C_{ik}) \;&=\; \mu\Big(\bigcup_{i=1}^{m} C_{ik}\Big) \;=\; \mu\Big(\bigcup_{i=1}^{m} B_i\cap A_k\Big) \;=\; \mu\Big(A_k\cap\bigcup_{i=1}^{m} B_i\Big) \;=\; \mu(A_k\cap X)\\ &=\; \mu(A_k)\end{aligned}$$

resulta que

$$\sum_{i=1}^{m}\sum_{k=1}^{n} y_k \cdot \mu(C_{ik}) \;=\; \sum_{k=1}^{n} y_k \cdot \left(\sum_{i=1}^{m} \mu(C_{ik})\right) \;=\; \sum_{k=1}^{n} y_k \cdot \mu(A_k) \;=\; \mathrm{U}_\mu(f, \mathcal{P}).$$

De modo enteramente similar se evalúan las otras sumas que aparecen en (T1), estableciéndose la siguiente desigualdad

$$\mathrm{L}_\mu(f, \mathcal{P}) + \mathrm{L}_\mu(g, \mathcal{Q}) \;\leq\; \int_X (f+g)\, d\mu \;\leq\; \mathrm{U}_\mu(f, \mathcal{P}) + \mathrm{U}_\mu(g, \mathcal{Q}).$$

Tomando ínfimo y supremo sobre las particiones de los rangos de f y g respectivamente, y teniendo en cuenta que f, g y $f+g$ son Lebesgue integrables, se obtiene el resultado deseado.

(b) Si $a = 0$, no hay nada que probar. Admitamos entonces que $a \neq 0$ y, para fijar idea, suponga que $a > 0$. Sea $\mathcal{P} = \{y_0, y_1, \ldots, y_n\}$ una partición del rango $[m, M]$ de f y, como antes, considere los conjuntos medibles y disjuntos

$$A_i \;=\; X \cap f^{-1}\big([y_{i-1}, y_i)\big), \quad i = 1, \ldots, n.$$

De nuevo, por el Teorema 10.1.8,

$$\int_X (af)\, d\mu \;=\; \sum_{i=1}^{n} \int_{A_i} (af)\, d\mu.$$

Observe que, para cada $x \in A_i$ se verifica la desigualdad

$$ay_{i-1} \;\leq\; af(x) \;<\; ay_i$$

de modo que, por una nueva aplicación de (L5), se obtiene que

$$a \cdot y_{i-1} \cdot \mu(A_i) \;\leq\; \sum_{i=1}^{n} \int_{A_i} (af)\, d\mu \;<\; a \cdot y_i \cdot \mu(A_i), \quad i = 1, \ldots, n.$$

Sumandos todas éstas desigualdades nos conduce a

$$a \cdot \mathrm{L}(f, \mathcal{P}) \;\leq\; \int_X (af)\, d\mu \;\leq\; a \cdot \mathrm{U}(f, \mathcal{P}),$$

y ya que f es Lebesgue integrable se concluye que

$$\int_X (af)\, d\mu \;=\; a \int_X f\, d\mu.$$

Finalmente, si $a < 0$, entonces como $-a > 0$ se sigue de la primera parte y de (1) que

$$0 \;=\; \int_X \big[(af) \,+\, (-a)f)\big]\, d\mu \;=\; \int_X (af)\, d\mu \,+\, (-a)\int_X f\, d\mu$$

y termina la prueba. ■

Aprovechando el hecho de $\mathcal{L}_{\text{fin}}(X,\mu)$ es un espacio vectorial, podemos definir la integral de cualquier función simple. En efecto, si $\varphi : X \to \mathbb{R}$ es una función simple expresada en su forma canónica, digamos, $\varphi = \sum_{j=1}^{n} a_j \cdot \chi_{E_j}$, definimos

$$\int_X \varphi \, d\mu \;=\; \sum_{j=1}^{n} a_j \cdot \mu(E_j).$$

Un ejercicio sencillo permite verificar que si φ se representa de cualquier otra forma distinta a la canónica, entonces su integral no cambia.

Corolario 10.1.12. *Sean $f, g \in \mathcal{L}_{\text{fin}}(X,\mu)$. Si $f \leq g$ $\mu -$ c.s., entonces $\int_X f \, d\mu \leq \int_X g \, d\mu$. En particular, se cumple la desigualdad:*

$$\left| \int_X f \, d\mu \right| \;\leq\; \int_X |f| \, d\mu.$$

Prueba. Puesto que la integral no se afecta sobre conjuntos de medida cero, asumiremos que $f \leq g$ sobre X. Como $g - f \geq 0$, se sigue de (L7) y el Teorema 10.1.11 que

$$0 \;\leq\; \int_X (g-f) \, d\mu \;=\; \int_X g \, d\mu \;-\; \int_X f \, d\mu.$$

Para demostrar la última desigualdad, observe que como $|f| \in \mathcal{L}_{\text{fin}}(X,\mu)$ y $-|f| \leq f \leq |f|$, entonces se sigue, por la primera parte, que

$$\left| \int_X f \, d\mu \right| \;\leq\; \int_X |f| \, d\mu.$$

La prueba es completa. ■

Nota Adicional 10.1.2 Haciendo uso de los resultados de esta sección, podemos extender la integral de Lebesgue a cualquier función $f : X \to \overline{\mathbb{R}}$ que sea **medible** y **acotada casi-siempre**.

Teorema 10.1.13. *Sea $f : X \to \overline{\mathbb{R}}$ una función **medible** y **acotada casi-siempre** sobre X. Entonces f es **Lebesgue integrable** sobre X.*

Prueba. Como f es acotada casi-siempre, existe un número $M > 0$ y un conjunto medible $E \subseteq X$ tal que

$$\mu(E) \;=\; 0 \qquad\qquad \text{y} \qquad\qquad |f(x)| \;\leq\; M \quad \text{para todo } x \in X \setminus E,$$

Por lo tanto, usando la σ-aditividad de la integral, el Corolario 10.1.12 y (L6), se tiene que

$$\left| \int_X f \, d\mu \right| \;\leq\; \int_{X \setminus E} |f| \, d\mu \;+\; \int_E |f| \, d\mu \;=\; \int_{X \setminus E} |f| \, d\mu \;\leq\; M \cdot \mu(X),$$

lo cual prueba que f integrable según Lebesgue. El símbolo $\mathcal{L}_{\text{fin}}(X,\mu)$ lo seguiremos usando para denotar *el espacio vectorial de todas las funciones medibles $f : X \to \overline{\mathbb{R}}$ que son **acotadas casi-siempre** sobre X.*

Una desigualdad importante que establece una cierta relación entre la integral de Lebesgue y la medida de ciertos conjuntos que probablemente son muy grandes, pero que limita el tamaño de tales conjuntos es el siguiente.

Corolario 10.1.14 (Desigualdad de Chebyshev). *Si $f \in \mathcal{L}_{\text{fin}}(X,\mu)$ con $f \geq 0$ sobre X, entonces*

$$\mu(\{x \in X : f(x) \geq c\}) \;\leq\; \frac{1}{c} \cdot \int_X f\, d\mu.$$

para todo $c > 0$.

Prueba. Para cada $c > 0$, considere el conjunto $E_c = \{x \in X : f(x) \geq c\}$. Si definimos $g = c \cdot \chi_{E_c}$, resulta que g es una función medible, acotada y, por lo tanto, Lebesgue integrable. Puesto que $0 \leq g \leq f$, se sigue del corolario anterior que

$$\int_X f\, d\mu \;\geq\; \int_X c \cdot \chi_{E_c}\, d\mu \;=\; c \cdot \mu(E_c).$$

Esto termina la prueba. ■

Corolario 10.1.15. *Sea $f \in \mathcal{L}_{\text{fin}}(X,\mu)$ con $f \geq 0$ sobre X. Las siguientes condiciones son equivalentes:*

(1) $f = 0$ $\mu-$c.s. *sobre* X.

(2) $\displaystyle\int_X f\, d\mu \;=\; 0.$

Prueba. Suponga que $\int_X f\, d\mu = 0$. Usando la Desigualdad de Chebyshev vemos que para cada número natural n, $\mu(\{x \in X : f(x) > 1/n\}) = 0$ y como

$$\{x \in X : f(x) > 0\} \;=\; \bigcup_{n=1}^{\infty} \{x \in X : f(x) > 1/n\}$$

la numerabilidad aditiva de la medida nos revela que $\mu(\{x \in X : f(x) > 0\}) = 0$. Esto prueba que $f = 0$ $\mu-$c.s.

Suponga ahora que $f = 0$ $\mu-$c.s. Si definimos $A = \{x \in X : f(x) > 0\}$, resulta que $\mu(A) = 0$ y $f = 0$ sobre $X \setminus A$. Por (L6) se tiene que

$$\int_X f\, d\mu \;=\; \int_{X\setminus A} f\, d\mu + \int_A f\, d\mu \;=\; \int_{X\setminus A} 0\, d\mu + 0 \;=\; 0.$$

La prueba es completa. ■

Observe que la hipótesis $f \geq 0$ en el corolario anterior no se puede omitir. En efecto, la función $f : [-1,1] \to \mathbb{R}$ definida por

$$f(x) \;=\; \begin{cases} 1 & \text{si} \quad x \in [0,1] \\ -1 & \text{si} \quad x \in [-1,0) \end{cases}$$

satisface

$$\int_{[-1,1]} f\, d\mu \;=\; \int_{[-1,0)} f\, d\mu + \int_{[0,1]} f\, d\mu \;=\; -1+1 \;=\; 0$$

y, sin embargo, $f \neq 0$ $\mu-$c.s.

Uno de los primeros éxitos incuestionables de la integral de Lebesgue fue el siguiente resultado.

Teorema 10.1.16 (Lebesgue). *Si $f \in \mathcal{R}([a,b])$, entonces f es* ***Lebesgue integrable*** *y*

$$(\mathrm{R})\int_a^b f\,dx \;=\; (\mathrm{L})\int_a^b f\,d\mu.$$

Prueba. Suponga que $f \in \mathcal{R}([a,b])$. Por la consecuencia (**TVL**$_{10}$) del Teorema de Vitali-Lebesgue, página 426, sabemos que f es medible y, además, acotada. Se sigue entonces que f es Lebesgue integrable. Para demostrar la igualdad de las integrales, tomemos una partición arbitraria $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ de $[a,b]$ y sea

$$M_i \;=\; \sup\{f(x) : x \in [x_{i-1}, x_i]\} \quad \text{para } i = 1, \ldots, n.$$

Como f es acotada, seleccione un $M > 0$ tal que $|f(x)| \leq M$ para todo $x \in [a,b]$ y elija cualquier partición $\mathcal{Q} = \{y_0, y_1, \ldots, y_m\}$ de $[-M, M]$ de modo tal que cada uno de los números $M_1, \ldots, M_n$ pertenezcan a dicha partición. Considere los conjuntos medibles y disjuntos

$$E_j \;=\; f^{-1}\big([y_{j-1}, y_j)\big) \quad \text{para } j = 1, \ldots, m$$

y observe que

$$\begin{aligned}
\mathsf{U}_\mu(f, \mathcal{Q}) \;&=\; \sum_{j=1}^{m} y_j \cdot \mu(E_j) \\
&=\; \sum_{j=1}^{m} \sum_{i=1}^{n} y_j \cdot \mu(E_j \cap [x_{i-1},\, x_i]) \\
&\leq\; \sum_{j=1}^{m} \sum_{i=1}^{n} M_i \cdot \mu(E_j \cap [x_{i-1},\, x_i]) \\
&\leq\; \sum_{i=1}^{n} M_i \cdot \mu([x_{i-1},\, x_i]) \\
&=\; \mathsf{U}(f, \mathcal{P})
\end{aligned}$$

Similarmente, si la partición $\mathcal{Q}$ contiene a cada uno de los números

$$m_i \;=\; \inf\{f(x) : x \in [x_{i-1}, x_i]\} \quad \text{para } i = 1, \ldots, n.$$

se tendrá entonces que $\mathsf{L}_\mu(f, \mathcal{Q}) \geq \mathsf{L}(f, \mathcal{P})$. Esto prueba que

$$\mathsf{L}(f, \mathcal{P}) \;\leq\; \mathsf{L}_\mu(f, \mathcal{Q}) \;\leq\; \mathsf{U}_\mu(f, \mathcal{Q}) \;\leq\; \mathsf{U}(f, \mathcal{P})$$

y como f es Riemann integrable se concluye que ambas integrales coinciden, esto es,

$$(\mathrm{R})\int_a^b f\,dx \;=\; (\mathrm{L})\int_a^b f\,d\mu$$

y termina la prueba. ■

10.1.2. Los Poderosos Teoremas de Convergencias

Como en la sección anterior supondremos que X es un **conjunto medible de medida finita**. Considere una función $f \in \mathcal{L}_{\text{fin}}(X, \mu)$ y suponga que existe una sucesión $(f_n)_{n=1}^{\infty}$ en $\mathcal{L}_{\text{fin}}(X, \mu)$ tal que

$$\lim_{n\to\infty} f_n = f \quad \mu - \text{c.s.}$$

¿Será cierto que la igualdad

$$\lim_{n\to\infty} \int_X f_n \, d\mu = \int_X \lim_{n\to\infty} f_n \, d\mu = \int_X f \, d\mu \tag{ILI}$$

se cumple? La validez de la igualdad (ILI) es lo que llamaremos **el intercambio entre el límite y la integral**, o **el paso del límite bajo el signo integral**. Por ejemplo, si consideramos la sucesión de funciones $(f_n)_{n=1}^{\infty}$, donde cada $f_n : [0,1] \to \mathbb{R}$ viene definida por

$$f_n(x) = \begin{cases} 0 & \text{si} \quad x \geq 2/n, \\ n & \text{si} \quad x = 1/n, \\ 0 & \text{si} \quad x = 0, \\ \text{lineal} & \text{sobre los intervalos } [0, 1/n] \text{ y } [1/n, 2/n], \end{cases}$$

tendremos que $\int_{[0,1]} f_n \, d\mu = 1$ para cada $n \geq 2$. Si ahora definimos $f = 0$ sobre $[0,1]$, resulta que

$$\lim_{n\to\infty} f_n = f \quad \text{pero} \quad \lim_{n\to\infty} \int_{[0,1]} f_n \, d\mu \neq \int_{[0,1]} f \, d\mu.$$

Este ejemplo muestra que la convergencia puntual por sí sola no es suficiente para garantizar el paso del límite bajo el signo integral. ¿Qué condiciones adicionales hay que imponerle a la convergencia puntual para que se cumpla la igualdad (ILI)? La respuesta viene dada por cuatro resultados fundamentales que demostraremos en esta sección y que demuestran, entre otros hechos, por qué la integral de Lebesgue es más potente que la de Riemann. Comenzaremos por el más sencillo, el cual es similar al Teorema de la Convergencia Uniforme para funciones Riemann integrables.

Teorema 10.1.17 (Convergencia Uniforme). *Sea $(f_n)_{n=1}^{\infty}$ una sucesión en $\mathcal{L}_{\text{fin}}(X, \mu)$ tal que $f_n \to f$* ***uniformemente*** *sobre X. Entonces $f \in \mathcal{L}_{\text{fin}}(X, \mu)$ y se cumple que*

$$\lim_{n\to\infty} \int_X f_n \, d\mu = \int_X f \, d\mu.$$

Prueba. Como la convergencia es uniforme y cada f_n es acotada, tenemos que f es acotada y, además, medible gracias al Corolario 7.2.3, página 379. Por esto, f es Lebesgue integrable. Sea $\varepsilon > 0$. Usemos de nuevo la convergencia uniforme para seleccionar un $N \in \mathbb{N}$ tal que

$$|f_n(x) - f(x)| < \varepsilon/\mu(X) \quad \text{para todo} \quad n \geq N \text{ y todo } x \in X.$$

Se sigue entonces de la linealidad y monotonicidad de la integral y del Corolario 10.1.12 que, para todo $n \geq N$,

$$\left| \int_X f_n \, d\mu - \int_X f \, d\mu \right| = \left| \int_X (f_n - f) \, d\mu \right| \leq \int_X |f_n - f| \, d\mu < \frac{\varepsilon}{\mu(X)} \cdot \mu(X) = \varepsilon.$$

Esto último nos muestra que $\lim_{n\to\infty}\int_X f_n\, d\mu = \int_X f\, d\mu$ y termina la prueba. ■

El Teorema de la Convergencia Acotada, **TCA**, para funciones Riemann integrables **exige**, para su validez, que la función límite sea Riemann integrable. Su equivalente, para funciones Lebesgue integrables, no requiere de tal exigencia sobre la función límite: su integrabilidad se obtiene como consecuencia directa de las hipótesis. El siguiente resultado fue establecido por primera vez por H. Lebesgue en su tesis de 1902. Aunque su demostración sigue fácilmente del Teorema de la Convergencia Dominada dado un poco más abajo, uno puede usar el Teorema de Severini-Egoroff como un peldaño para obtener una prueba del mismo.

Teorema 10.1.18 (Convergencia Acotada). *Sea $(f_n)_{n=1}^{\infty}$ una sucesión en $\mathcal{L}_{\mathrm{fin}}(X,\mu)$ **uniformemente acotada** sobre X y suponga que existe una función $f : X \to \mathbb{R}$ tal que*

$$\lim_{n\to\infty} f_n \;=\; f \;\; \textit{casi-siempre}.$$

Entonces $f \in \mathcal{L}_{\mathrm{fin}}(X,\mu)$ y

$$\lim_{n\to\infty}\int_X f_n\, d\mu \;=\; \int_X f\, d\mu.$$

Prueba. Sea $M > 0$ una constante tal que $\sup_{n\in\mathbb{N}} \| f_n \|_\infty \leq M$. Puesto que $f_n \to f$ casi-siempre, se sigue del Corolario 7.2.3 que f es medible, y como claramente $|f| \leq M$, resulta que f es Lebesgue integrable. Asuma que $\mu(X) > 0$, en caso contrario el resultado es trivial gracias a (L6). Dado $\varepsilon > 0$, el Teorema de Severini-Egoroff, página 386, nos garantiza la existencia de un conjunto medible $F_\varepsilon \subseteq X$ con $\mu(F_\varepsilon) > 0$ tal que

$$\mu(X \setminus F_\varepsilon) \;<\; \frac{\varepsilon}{4M} \qquad\qquad \text{y} \qquad\qquad f_n \to f \quad \text{uniformemente sobre } F_\varepsilon.$$

De esto último podemos derivar la existencia de un $N \in \mathbb{N}$ tal que, para todo $n \geq N$ se cumple que

$$|f_n(x) - f(x)| \;<\; \frac{\varepsilon}{2\mu(F_\varepsilon)} \quad \text{cualquiera sea } x \in F_\varepsilon.$$

La σ-aditividad de la integral de Lebesgue y otros resultados ya establecidos en esta sección nos indican que si $n \geq N$, entonces

$$\begin{aligned}
\left| \int_X (f_n - f)\, d\mu \right| \;&=\; \left| \int_{F_\varepsilon} (f_n - f)\, d\mu \;+\; \int_{X\setminus F_\varepsilon} (f_n - f)\, d\mu \right| \\
&\leq\; \int_{F_\varepsilon} |f_n - f|\, d\mu \;+\; \int_{X\setminus F_\varepsilon} |f_n - f|\, d\mu \\
&<\; \frac{\varepsilon}{2\mu(F_\varepsilon)}\mu(F_\varepsilon) \;+\; 2M\cdot\mu(X\setminus F_\varepsilon) \\
&<\; \varepsilon
\end{aligned}$$

Esto termina la prueba. ■

Es importante destacar que el **TCA** posee dos grandes inconvenientes:

(1$^{\text{o}}$) No se puede extender a conjuntos medibles de **medida infinita**.

(2^{o}) No se aplica a **funciones no-acotadas**.

El siguiente lema, demostrado por Pierre Joseph Louis Fatou (1878-1929) en su tesis doctoral de 1906, establece una útil desigualdad, en lugar de una igualdad, para cualquier sucesión de funciones no-negativas, medibles y acotadas que no requiere, explícitamente, la convergencia de la sucesión dada.

Lema 10.1.19 (Fatou). *Sea $(f_n)_{n=1}^{\infty}$ una sucesión de **funciones no-negativas** en $\mathcal{L}_{\text{fin}}(X,\mu)$. Entonces*

$$\int_X \liminf_{n\to\infty} f_n\, d\mu \;\leq\; \liminf_{n\to\infty} \int_X f_n\, d\mu.$$

Prueba. Observe, en primer lugar, que la función $\liminf_{n\to\infty} f_n$ es medible y no-negativa, de modo que $\int_X \liminf_{n\to\infty} f_n\, d\mu$ tiene sentido sólo cuando $\liminf_{n\to\infty} f_n$ esté acotada. Suponga, entonces, que $\liminf_{n\in\mathbb{N}} f_n \leq M$ para alguna constante $M > 0$. Para cada $k \geq 1$, sea

$$g_k \;=\; \inf\{f_k, f_{k+1}, \ldots\}.$$

Resulta que $(g_k)_{k=1}^{\infty}$ es una sucesión creciente de funciones medibles no-negativas tal que

$$\lim_{k\to\infty} g_k \;=\; \liminf_{k\to\infty} f_k.$$

Más aun, como $g_k \leq f_n$ para todo $n \geq k$, se tiene que

$$\int_X g_k\, d\mu \;\leq\; \int_X f_n\, d\mu \quad \text{para todo} \quad n \geq k$$

y, en consecuencia,

$$\int_X g_k\, d\mu \;\leq\; \liminf_{n\to\infty} \int_X f_n\, d\mu.$$

Finalmente, como $\sup_{k\geq 1} g_k \leq M$, se sigue del Teorema de la Convergencia Acotada que

$$\int_X \liminf_{n\to\infty} f_n\, d\mu \;=\; \int_X \lim_{k\to\infty} g_k\, d\mu \;=\; \lim_{k\to\infty} \int_X g_k\, d\mu \;\leq\; \liminf_{n\to\infty} \int_X f_n\, d\mu.$$

La prueba es completa. ■

Es importante observar que si la sucesión de funciones no-negativas $(f_n)_{n=1}^{\infty}$, en el resultado anterior, se supone **uniformemente acotada**, entonces automáticamente la función $\liminf_{n\to\infty} f_n$ resulta ser acotada (y medible) y, en consecuencia, Lebesgue integrable. Más aun, si se mantiene la hipótesis de que la sucesión $(f_n)_{n=1}^{\infty}$ sea uniformemente acotada pero no se exige que ellas sean no-negativas, entonces la desigualdad de Fatou aun es válida.

El mismo año de 1906, Beppo Levi (1875-1961) publicó un artículo [11] en donde demuestra uno de los resultados de mayor utilidad en la Teoría de Integración: el Teorema de la Convergencia Monótona, **TCM** .

Teorema 10.1.20 (Convergencia Monótona). *Sea $f : X \to \mathbb{R}$ una función **acotada**. Si $(f_n)_{n=1}^{\infty}$ es una sucesión de **funciones no-negativas** en $\mathcal{L}_{\text{fin}}(X,\mu)$ tal que*

*(a) $(f_n)_{n=1}^{\infty}$ es **monótona creciente** y*

(b) $\lim_{n\to\infty} f_n \;=\; f$ casi-siempre.

Entonces $f \in \mathcal{L}_{\text{fin}}(X,\mu)$ y

$$\lim_{n\to\infty} \int_X f_n\, d\mu \;=\; \int_X f\, d\mu.$$

Prueba. Observe que la condición (b) implica que la función f es medible y, además, acotada por hipótesis. Por esto, $f \in \mathcal{L}_{\text{fin}}(X, \mu)$ y como $f_n \leq f$ para todo $n \geq 1$, resulta que

$$0 \;\leq\; \int_X f_n \, d\mu \;\leq\; \int_X f \, d\mu \quad \text{para todo } n \geq 1.$$

De esto se sigue que la sucesión de números reales $\left(\int_X f_n \, d\mu \right)_{n=1}^{\infty}$ es creciente y acotada y, por consiguiente, converge. Además, $\lim_{n\to\infty} \int_X f_n \, d\mu \leq \int_X f \, d\mu$. Finalmente, aplicando el Lema de Fatou se obtiene que

$$\begin{aligned}
\lim_{n\to\infty} \int_X f_n \, d\mu \;\leq\; \int_X f \, d\mu \;&=\; \int_X \lim_{n\to\infty} f_n \, d\mu \\
&=\; \int_X \liminf_{n\to\infty} f_n \, d\mu \;\leq\; \liminf_{n\to\infty} \int_X f_n \, d\mu \\
&=\; \lim_{n\to\infty} \int_X f_n \, d\mu.
\end{aligned}$$

Esto muestra que

$$\lim_{n\to\infty} \int_X f_n \, d\mu \;=\; \int_X f \, d\mu$$

y termina la prueba. ■

Nota Adicional 10.1.3 La hipótesis de que la sucesión $(f_n)_{n=1}^{\infty}$ sea **creciente** en el Teorema de la Convergencia Monótona es fundamental. Por ejemplo, si $X = [0,1]$ y $f_n = n\chi_{(0,1/n)}$ para cada $n \geq 1$, resulta que la sucesión $(f_n)_{n=1}^{\infty}$ no es creciente, satisface $\lim_{n\to\infty} f_n = 0$, pero

$$\lim_{n\to\infty} \int_{[0,1]} f_n \;=\; 1.$$

Similarmente, el requerimiento de que f esté acotada no se puede omitir. En efecto, si definimos la sucesión de funciones $(f_n)_{n=1}^{\infty}$ por

$$f_n(x) \;=\; \begin{cases} x^{-1} & \text{si} \quad 1/n \leq x \leq 1 \\ 0 & \text{si} \quad 0 \leq x < 1/n \end{cases} \qquad \text{y} \qquad f(x) \;=\; \begin{cases} x^{-1} & \text{si} \quad 0 < x \leq 1 \\ 0 & \text{si} \quad x = 0 \end{cases}$$

resulta que $f_n \to f$ puntualmente en $[0,1]$, cada f_n es Lebesgue integrable, pero f no es Lebesgue integrable. Observe

$$\lim_{n\to\infty} \int_{[0,1]} f_n \, d\mu \;=\; \lim_{n\to\infty} \ln(n).$$

Tal vez pueda ser de utilidad saber que: *el Teorema de la Convergencia Monótona y el Lema de Fatou son proposiciones equivalentes*. En efecto, ya vimos que el Lema de Fatou fue usado para derivar el Teorema de la Convergencia Monótona. Por otro lado, en la demostración del Lema de Fatou se utilizó el Teorema de la Convergencia Acotada que puede, obviamente, ser sustituido por el Teorema de la Convergencia Monótona y, en consecuencia, ambos resultados son equivalentes.

El Teorema de la Convergencia Monótona representa un instrumento adecuado para integrar series de funciones medibles no-negativas. De hecho, el siguiente resultado también es equivalente a dicho teorema.

Corolario 10.1.21 (Beppo Levi). *Sea $(f_n)_{n=1}^{\infty}$ una **sucesión no-negativa** en $\mathcal{L}_{\text{fin}}(X,\mu)$ y suponga que la serie $\sum_{n=1}^{\infty} f_n \leq M$ para alguna constante $M > 0$. Entonces*

$$\sum_{n=1}^{\infty} \int_X f_n \, d\mu \; < \; +\infty \qquad \text{y} \qquad \int_X \left(\sum_{n=1}^{\infty} f_n \right) d\mu \; = \; \sum_{n=1}^{\infty} \int_X f_n \, d\mu.$$

Prueba. Para cada $n \geq 1$, sea $g_n = f_1 + \cdots + f_n$. Entonces $(g_n)_{n=1}^{\infty}$ es una sucesión no-negativa y creciente de funciones Lebesgue integrables tal que $g_n \to \sum_{n=1}^{\infty} f_n$ puntualmente. De aquí se sigue que $\sum_{n=1}^{\infty} f_n$ es Lebesgue integrable y así, por el Teorema de la Convergencia Monótona y la aditividad de la integral se tiene que

$$\int_X \left(\sum_{n=1}^{\infty} f_n \right) d\mu \; = \; \int_X \lim_{n\to\infty} g_n \, d\mu \; = \; \lim_{n\to\infty} \int_X g_n \, d\mu \; = \; \sum_{n=1}^{\infty} \int_X f_n \, d\mu.$$

La prueba es completa. ■

El Teorema 10.1.8 ahora es una fácil consecuencia del Teorema de la Convergencia Monótona. En efecto, en este caso sólo es suficiente considerar la sucesión de funciones medibles no-negativas $(f_n)_{n=1}^{\infty}$, donde $f_n = f \cdot \chi_{E_1 \cup \cdots \cup E_n}$ y aplicar el Teorema de la Convergencia Monótona a dicha sucesión.

Aunque el Teorema de la Convergencia Monótona no es válido, en general, para sucesiones de funciones que no son monótonas crecientes, el siguiente resultado es un formidable sustituto para dicho teorema pues él garantiza el intercambio entre el límite y la integral con un mínimo de requerimientos sin imponer que la sucesión sea monótona creciente. El Teorema de la Convergencia Dominada, **TCD**, apareció en la tesis de Lebesgue de 1902 publicada en *Annali di Matematica* como un artículo titulado «Intégrale, longueur, aire». Este teorema es, probablemente, el más importante y poderoso resultado que usaremos sobre el intercambio entre el límite y la integral.

Teorema 10.1.22 (Convergencia Dominada). *Sea $f : X \to \overline{\mathbb{R}}$ una **función medible** y sea $(f_n)_{n=1}^{\infty}$ una sucesión de funciones medibles definidas sobre X tal que:*

(a) $\lim_{n\to\infty} f_n \; = \; f$ *casi-siempre sobre* X *y*

(b) existe una función $g \in \mathcal{L}_{\text{fin}}(X,\mu)$ *tal que* $|f_n| \leq g$ *para todo* $n \geq 1$.

Entonces $\{f, f_1, f_2, \ldots\} \subseteq \mathcal{L}_{\text{fin}}(X,\mu)$ *y*

$$\lim_{n\to\infty} \int_X f_n \, d\mu \; = \; \int_X f \, d\mu.$$

Prueba. Sin perder generalidad, asumiremos que $f_n \to f$ puntualmente. Por consiguiente, f es medible. Como g es Lebesgue integrable y $|f_n| \leq g$ resulta que f_n es Lebesgue integrable para todo $n \geq 1$. Más aun, puesto que $|f| = \lim_{n\to\infty} |f_n| \leq g$ entonces f también es Lebesgue integrable. Ahora bien, como $\lim_{n\to\infty} \int_X f_n \, d\mu$ existe si, y sólo si, el límite inferior y el límite superior de la sucesión $(\int_X f_n \, d\mu)_{n=1}^{\infty}$ son iguales, resulta que todo lo que tenemos que demostrar es que

$$\liminf_{n\to\infty} \int_X f_n \, d\mu \; = \; \limsup_{n\to\infty} \int_X f_n \, d\mu.$$

Para verificar esto último, observe que las sucesiones $(f_n+g)_{n=1}^{\infty}$ y $(g-f_n)_{n=1}^{\infty}$ son no-negativas y medibles y, en consecuencia, podemos aplicar el Lema de Fatou para obtener

$$\int_X f d\mu + \int_X g\, d\mu \;=\; \int_X (f+g)\, d\mu \;\leq\; \liminf_{n\to\infty} \int_X (f_n+g)\, d\mu \;=\; \int_X g\, d\mu + \liminf_{n\to\infty} \int_X f_n\, d\mu,$$

$$\int_X g d\mu - \int_X f\, d\mu \;=\; \int_X (g-f)\, d\mu \;\leq\; \liminf_{n\to\infty} \int_X (g-f_n)\, d\mu \;=\; \int_X g\, d\mu - \limsup_{n\to\infty} \int_X f_n\, d\mu$$

Cancelando el número $\int_X g\, d\mu$ en ambas desigualdades se obtiene que

$$\limsup_{n\to\infty} \int_X f_n\, d\mu \;\leq\; \int_X f d\mu \;\leq\; \liminf_{n\to\infty} \int_X f_n\, d\mu$$

y el teorema queda demostrado. ■

Observe que como $\mu(X) < +\infty$, cualquier función constante definida sobre X es Lebesgue integrable y, por consiguiente, el Teorema de la Convergencia Acotada es consecuencia inmediata del Teorema de la Convergencia Dominada si tomamos $g(x) = M$ para todo $x \in X$, donde $M \in \mathbb{R}$.

10.1.3. El Teorema Fundamental del Cálculo de Lebesgue

Además de los problemas de convergencia que se originaron cuando físicos y matemáticos investigaban el problema de las series trigonométricas, Lebesgue estaba interesado, especialmente, en resolver otra de las limitaciones de la integral de Riemann: el Teorema Fundamental del Cálculo. Recordemos que el ejemplo construido por Volterra de una derivada acotada que no es Riemann integrable impedía que el Teorema Fundamental del Cálculo para la integral de Riemann se pudiese aplicar en estos casos. Con su integral, Lebesgue resuelve, utilizando entre otros resultados su nuevo e interesante Teorema de la Convergencia Acotada, satisfactoriamente dicho problema para las *funciones medibles con derivadas acotadas*. El siguiente resultado es una versión del Teorema Fundamental del Cálculo para la integral de Lebesgue. Un poco más adelante daremos un resultado más general, Teorema 10.2.41, página 615, que establece con toda precisión, cuáles funciones satisfacen la igualdad dada en el siguiente teorema.

Teorema 10.1.23 (Fundamental del Cálculo). *Sea $f : [a,b] \to \mathbb{R}$ una función **diferenciable** en cada punto de $[a,b]$. Si f' es **acotada** sobre $[a,b]$, entonces f' es **Lebesgue integrable** sobre $[a,b]$ y se cumple que*

$$\int_a^x f'\, d\mu \;=\; f(x) - f(a)$$

para cada $x \in [a,b]$.

Prueba. Por el Teorema 7.2.6, página 380, sabemos que f' es medible y, además, acotada por hipótesis. Se sigue entonces que f' es Lebesgue integrable sobre $[a,b]$. Para demostrar la última parte del teorema, sea M una constante positiva tal que $|f'(x)| \leq M$ para todo $x \in [a,b]$ y extienda f al intervalo $[a,b+1]$ permitiendo que $f(x) = f'(b)(x-b) + f(b)$ para $x \in (b,b+1]$. Observe que la función extendida f es continua y diferenciable sobre $[a,b+1]$. Con esta información podemos definir, para cada $n \in \mathbb{N}$, la función $f_n : [a,b] \to \mathbb{R}$ dada por

$$f_n(x) \;=\; n\big(f(x+1/n) - f(x)\big).$$

Por el Teorema del Valor Medio, para cada $n \in \mathbb{N}$ y cada $x \in [a, b]$, existe un $z_n^x \in (x, x + 1/n)$ tal que

$$f_n(x) = \frac{f(x + 1/n) - f(x)}{1/n} = f'(z_n^x).$$

Se sigue que $|f_n(x)| \leq M$ para todo $x \in [a, b]$ y cualquier $n \in \mathbb{N}$, es decir, la sucesión $(f_n)_{n=1}^{\infty}$ está uniformemente acotada. Además, puesto que $f_n \to f'$, el Teorema de la Convergencia Acotada nos garantiza que

$$\int_a^b f' \, d\mu = \lim_{n\to\infty} \int_a^b f_n \, d\mu.$$

Más aun, como f es continua sobre $[a, b+1]$, ella es Riemann integrable y, en consecuencia, el Segundo Teorema Fundamental del Cálculo para la integral de Riemann implica que (véase el Ejercicio 8.4.7, página 450),

$$\lim_{n\to\infty} n \int_a^{a+1/n} f \, dt = f(a) \qquad \text{y} \qquad \lim_{n\to\infty} n \int_b^{b+1/n} f \, dt = f(b).$$

Usando el Teorema 10.1.16 y un simple cambio de variable (recuerde que f es Riemann integrable) se obtiene que

$$\begin{aligned}
\int_a^b f' \, d\mu &= \lim_{n\to\infty} \int_a^b f_n \, d\mu \\
&= \lim_{n\to\infty} \left(n \int_a^b f(t + 1/n) \, dt - n \int_a^b f(t) \, dt \right) \\
&= \lim_{n\to\infty} \left(n \int_{a+1/n}^{b+1/n} f(t) \, dt - n \int_a^b f(t) \, dt \right) \\
&= \lim_{n\to\infty} n \int_b^{b+1/n} f(t) \, dt - \lim_{n\to\infty} n \int_a^{a+1/n} f(t) \, dt \\
&= f(b) - f(a).
\end{aligned}$$

Un argumento enteramente similar nos convence de la validez de la igualdad anterior para cualquier $x \in (a, b)$. Esto termina la prueba. ■

Aunque el resultado anterior resuelve satisfactoriamente el problema de la primitiva para derivadas acotadas, resulta que cuando la derivada no es acotada y oscila muy rápidamente alrededor de un punto, el Teorema Fundamental del Cálculo no se aplica en este contexto.

10.1.4. La Integral de Lebesgue via Funciones Simples

En la sección anterior vimos cómo se construyó la integral de Lebesgue para una función medible y acotada $f : X \to \mathbb{R}$, donde X era un conjunto medible de medida finita y probamos algunas de las formidables propiedades que dicha integral posee. En esta sección le daremos el toque final a tal construcción: veremos cómo se extiende el concepto de integral de Lebesgue a cualquier función $f : X \to \overline{\mathbb{R}}$, no necesariamente acotada, donde X es un conjunto medible arbitrario, es decir, cuya medida puede ser finita o infinita.

Como medida de precaución, debemos alertar al lector que en esta sección utilizaremos el mismo símbolo $\int_X f\,d\mu$ para denotar las "*distintas extensiones*" de la integral de Lebesgue desarrollada en la sección anterior. Esperamos que esto no le cause confusión.

(a) **Segunda Suposición**. **Como antes, supondremos que $X \in \mathfrak{M}_\mu(\mathbb{R})$ con $\mu(X) < +\infty$ y que $f : X \to \overline{\mathbb{R}}$ es una función medible no-negativa arbitraria pero no necesariamente acotada**.

Aprovechando la definición de integral de cualquier función medible y acotada, podemos extender dicha noción a toda función $f : X \to \overline{\mathbb{R}}$ **medible**, **no-negativa**, y **no necesariamente acotada** como sigue: considere el conjunto

$$\mathcal{A}_f \;=\; \left\{ \int_X \phi\,d\mu \;:\; 0 \leq \phi \leq f,\;\; \phi \text{ es medible y acotada} \right\}$$

el cual, obviamente, es no vacío y está incluido en $[0, +\infty)$. Por lo tanto, puede suceder que $\sup \mathcal{A}_f = +\infty$. Si, por el contrario, ocurrir que $\sup \mathcal{A}_f < +\infty$, entonces diremos que f es **Lebesgue integrable** sobre X y definimos la integral de f sobre X como

$$\int_X f\,d\mu \;=\; \sup\left\{ \int_X \phi\,d\mu \;:\; 0 \leq \phi \leq f,\;\; \phi \text{ es medible y acotada} \right\}. \tag{IL2}$$

Observe que:

> "Para cualquier función medible no-negativa $f : X \to \overline{\mathbb{R}}$, $\int_X f\,d\mu$ **siempre existe** como un elemento de $\overline{\mathbb{R}}$, y que sólo las que satisfacen la condición $\int_X f\,d\mu < +\infty$ son llamadas Lebesgue integrables."

Es claro que si f es no-negativa y acotada, entonces la integral que acabamos de definir coincide con la integral (IL1) de la sección anterior ya que, en este caso, $\int_X f\,d\mu \in \mathcal{A}_f$ y, en consecuencia, $\int_X f\,d\mu = \sup \mathcal{A}_f$. Sin embargo, si f no es acotada esta nueva integral es una **extensión** de la anterior siempre que f sea no-negativa.

Aunque la definición que acabamos de ofrecer de la integral de una función medible no-negativa es inobjetable, el conjunto $\mathcal{A}_f$ puede resultar ser demasiado grande para intentar evaluar dicha integral. En la practica, el siguiente subconjunto de $\mathcal{A}_f$ es más conveniente y "*apropiado*" para trabajar con él:

$$\mathcal{S}_f \;=\; \left\{ \int_X \varphi\,d\mu \;:\; 0 \leq \varphi \leq f,\; \varphi \text{ es simple} \right\},$$

el cual es, definitivamente, "más pequeño" que $\mathcal{A}_f$ y, aun, posee la propiedad de generar la misma integral dada en (IL2), es decir,

Teorema 10.1.24. *Si $f : X \to \overline{\mathbb{R}}$ es una función **medible no-negativa** con $\mu(X) < +\infty$, entonces*

$$\int_X f\,d\mu \;=\; \sup \mathcal{S}_f. \tag{IL3}$$

Prueba. Puesto que $\mathcal{S}_f \subseteq \mathcal{A}_f$, resulta que $\sup \mathcal{S}_f \leq \sup \mathcal{A}_f$. Para demostrar la otra desigualdad, sea $\varepsilon > 0$ y seleccione, usando (IL2), una función medible y acotada g con $0 \leq g \leq f$ tal que

$$\sup \mathcal{A}_f \;<\; \int_X g\,d\mu \;+\; \frac{\varepsilon}{2}. \tag{1}$$

Siendo g una función medible no-negativa, el Teorema de la Densidad de las Funciones Simples, Teorema 7.2.10, página 383, nos garantiza la existencia de una función simple no-negativa $\varphi \leq g$ tal que $g < \varphi + \varepsilon/2\mu(X)$. De esto se sigue que

$$\int_X g\,d\mu \;<\; \int_X \varphi\,d\mu + \frac{\varepsilon}{2} \;\leq\; \sup \mathcal{S}_f + \frac{\varepsilon}{2}$$

y entonces, usando (1), se obtiene que $\sup \mathcal{A}_f < \sup \mathcal{S}_f + \varepsilon$. Puesto que ε es arbitrario, se concluye que $\sup \mathcal{A}_f \leq \sup \mathcal{S}_f$ y termina la prueba. ■

Si bien es cierto que el conjunto $\mathcal{S}_f$ es más pequeño que $\mathcal{A}_f$ persiste, aun, un problema: dicho conjunto sigue siendo todavía muy grande ya que su cardinalidad es $\mathfrak{c}$, por lo que puede resultar difícil evaluar $\int_X f\,d\mu$ directamente del resultado anterior. Un modo de sortear con éxito esta dificultad es a través del Teorema de la Convergencia Monótona para funciones medibles no-negativas. Es importante destacar que, a diferencia de lo que ocurre con el Teorema de la Convergencia Monótona para funciones no-negativas en $\mathcal{L}_{\text{fin}}(X,\mu)$, esta nueva formulación del teorema no requiere que las funciones involucradas sean necesariamente Lebesgue integrables, ni que la función límite sea acotada: *¡ el intercambio entre el límite y la integral es válido aun si ellas no son Lebesgue integrables !*

Teorema 10.1.25 (Convergencia Monótona). *Sea $(f_n)_{n=1}^{\infty}$ una **sucesión creciente** de funciones **medibles no-negativas** definidas sobre X y sea $f : X \to \overline{\mathbb{R}}$ una función tal que $\lim\limits_{n\to\infty} f_n = f$ casi-siempre sobre X. Entonces*

$$\int_X f\,d\mu \;=\; \lim_{n\to\infty} \int_X f_n\,d\mu.$$

Prueba. Claramente f es no-negativa y medible. Es suficiente, para demostrar el resultado, suponer que $f_n(x) \to f(x)$ para cada $x \in X$. Como $f_n \leq f$ para todo $n \geq 1$ y la sucesión $(f_n)_{n=1}^{\infty}$ es creciente, resulta que $(\int_X f_n\,d\mu)_{n=1}^{\infty}$ es una sucesión creciente y acotada por $\int_X f\,d\mu$ en $\overline{\mathbb{R}}$ y, en consecuencia,

$$\lim_{n\to\infty} \int_X f_n\,d\mu \;\leq\; \int_X f\,d\mu. \tag{1}$$

Fijemos una función $\varphi \in \mathcal{S}_f^{+}(X,\mu)$, esto es, φ es simple con $0 \leq \varphi \leq f$ y sea $\lambda \in (0,1)$. Considere, por cada $n \in \mathbb{N}$, el conjunto

$$E_n \;=\; \big\{x \in X : f_n(x) \geq \lambda \cdot \varphi(x)\big\}.$$

Claramente cada conjunto E_n es medible y como $f_1 \leq f_2 \leq \cdots$ se tiene que $E_1 \subseteq E_2 \subseteq \cdots$ Afirmamos que $X = \bigcup_{n=1}^{\infty} E_n$. En efecto, sea $x \in X$ y observe que $f(x) \geq \varphi(x) \geq \lambda \cdot \varphi(x)$. Veamos que existe un $m \geq 1$ tal que $f_m(x) \geq \lambda \cdot \varphi(x)$. Si lo anterior fuese falso tendríamos que $f_m(x) < \lambda \cdot \varphi(x)$ para todo $m \geq 1$ de donde resultaría que $f(x) = \lim_{m\to\infty} f_m(x) < \lambda \cdot \varphi(x)$. Esta contradicción establece que $f_m(x) \geq \lambda \cdot \varphi(x)$ para algún $m \geq 1$ y, en consecuencia, $x \in E_m$. Por esto, $X \subseteq \bigcup_{n=1}^{\infty} E_n$ y como la inclusión recíproca es trivialmente válida, tenemos que $X = \bigcup_{n=1}^{\infty} E_n$. De las propiedades de la integral se sigue que, para cualquier $n \geq 1$,

$$\int_X f_n\,d\mu \;\geq\; \int_{E_n} f_n\,d\mu \;\geq\; \lambda \int_{E_n} \varphi\,d\mu.$$

Puesto que φ es acotada y $\mu(X) < \infty$, se concluye del Corolario 10.1.10 que

$$\lim_{n\to\infty} \int_X f_n\,d\mu \;\geq\; \lambda \lim_{n\to\infty} \int_{E_n} \varphi\,d\mu \;=\; \lambda \int_X \varphi\,d\mu.$$

Ahora bien, como la elección de φ se hizo de forma arbitraria resulta, tomando el supremo sobre todas las funciones simples no-negativas que están por debajo de f en la desigualdad anterior, que

$$\lim_{n\to\infty}\int_X f_n\,d\mu \;\geq\; \lambda\cdot\sup\left\{\int_X \phi\,d\mu \;:\; 0\leq\phi\leq f,\ \phi \text{ es simple}\right\} \;=\; \lambda\int_X f\,d\mu$$

gracias al Lema **??**. Finalmente, puesto que esta última desigualdad es válida para cualquier $0<\lambda<1$, entonces, tomando límite cuando $\lambda\to 1$ se concluye que

$$\lim_{n\to\infty}\int_X f_n\,d\mu \;\geq\; \int_X f\,d\mu.$$

Combinado (1) con esta última desigualdad se obtiene el resultado deseado y con ello termina la prueba. ■

Observe que, si por cada $n\in\mathbb{N}$ definimos la **función truncada** $f_n^T : X\to\mathbb{R}$ por

$$f_n^T \;=\; \min\{f,\, n\}$$

resulta que cada una de ellas es no-negativa, medible, acotada y $f_n^T \leq f$ para todo $n\geq 1$. Más aún, como la sucesión $(f_n^T)_{n=1}^\infty$ es creciente y converge puntualmente a f, el Teorema de la Convergencia Monótona, Teorema 10.1.25, nos garantiza que

$$\int_X f\,d\mu \;=\; \lim_{n\to\infty}\int_X f_n^T\,d\mu.$$

Esta es otra manera muy usada de extender la integral de Lebesgue a funciones medible no-negativas a partir de la integral de funciones medibles acotadas. Aquí está otra: Si f es una función medible no-negativa, seleccione, usando el Teorema 7.2.10, página 383, una sucesión creciente de funciones simples no-negativas $(\varphi_n)_{n=1}^\infty$ tal que $\varphi_n\leq f$ para todo $n\geq 1$ y

$$\lim_{n\to\infty}\varphi_n(x) \;=\; f(x), \quad \text{para cada} \quad x\in X.$$

El Teorema de la Convergencia Monótona, Teorema 10.1.25, nos garantiza entonces que

$$\int_X f\,d\mu \;=\; \lim_{n\to\infty}\int_X \varphi_n\,d\mu. \tag{LFS}$$

Puesto que la igualdad (LFS) no depende sobre cuál sucesión creciente de funciones simples no-negativas se elige convergiendo a f, podemos concluir que:

> *Para evaluar $\int_X f\,d\mu$, cuando f es no-negativa, sólo es suficiente computar $\lim_{n\to\infty}\int_X \varphi_n\,d\mu$, donde $(\varphi_n)_{n=1}^\infty$ es cualquier **sucesión creciente de funciones simples no-negativas** convergiendo puntualmente a f.*

Esto, por supuesto, nos permite dar la siguiente definición:

Definición 10.1.26 (Integral vía Funciones Simples). *Sea $f : X\to\overline{\mathbb{R}}$ una función medible no-negativa con $\mu(X)<+\infty$. La **integral** de f sobre X se define como*

$$\int_X f\,d\mu \;=\; \lim_{n\to\infty}\int_X \varphi_n\,d\mu, \tag{IL3*}$$

donde $(\varphi_n)_{n=1}^\infty$ es cualquier sucesión creciente de funciones simples no-negativas convergiendo puntualmente a f.

Como antes, si $E \subseteq X$ es medible y $f \geq 0$, definimos

$$\int_E f\, d\mu \;=\; \int_X f \cdot \chi_E \, d\mu.$$

En lo que sigue, la siguiente notación será usada:

$$\mathcal{S}_f^+(X,\mu) \;=\; \Big\{\varphi \in \mathcal{S}_\mu(X) \,:\, 0 \leq \varphi \leq f\Big\},$$

donde $\mathcal{S}_\mu(X)$ *es el conjunto de todas las funciones simples definidas sobre* X.

‖ ► Propiedades de la Integral de Lebesgue Extendida

Denotemos por $\mathcal{L}_1^+(X,\mu)$ *el conjunto de todas las funciones* $f : X \to \overline{\mathbb{R}}$ *medibles y no-negativas que son Lebesgue integrables sobre* X. Observe que si $f, g \in \mathcal{L}_1^+(X,\mu)$, entonces $f+g$, $af \in \mathcal{L}_1^+(X,\mu)$ para cualquier $a \geq 0$. Además, se cumple que:

(i) $\displaystyle\int_X (f+g)\, d\mu = \int_X f\, d\mu + \int_X g\, d\mu.$

En particular, si F y G son subconjuntos medibles y disjuntos de X, entonces

$$\int_{F\cup G} f\, d\mu \;=\; \int_F f\, d\mu + \int_G f\, d\mu.$$

(ii) $\displaystyle\int_X (af)\, d\mu = a\int_X f\, d\mu$ si $a \geq 0$

(iii) $\displaystyle\int_X f\, d\mu \leq \int_X g\, d\mu$ siempre que $f \leq g$.

(iv) $\displaystyle\int_A f\, d\mu \leq \int_B f\, d\mu$ para cualquiera conjuntos medibles A, B con $A \subseteq B \subseteq X$.

Prueba. Veamos cómo (i) sigue fácilmente usando el Teorema de la Convergencia Monótona. En efecto, por el Teorema de la Densidad de las Funciones Simples, Teorema 7.2.10, página 383, seleccione sucesiones crecientes de funciones simples no-negativas $(\varphi_n)_{n=1}^\infty$ y $(\psi_n)_{n=1}^\infty$ tales que

$$\lim_{n\to\infty} \varphi_n \;=\; f \qquad\qquad \text{y} \qquad\qquad \lim_{n\to\infty} \psi_n \;=\; g$$

puntualmente. Se sigue del Teorema de la Convergencia Monótona que

$$\int_X f\, d\mu \;=\; \lim_{n\to\infty}\int_X \varphi_n\, d\mu \qquad\qquad \text{y} \qquad\qquad \int_X g\, d\mu \;=\; \lim_{n\to\infty}\int_X \psi_n\, d\mu.$$

Más aun, puesto que $(\varphi_n+\psi_n)_{n=1}^\infty$ es una sucesión creciente de funciones simples no-negativas convergiendo puntualmente a $\varphi+\psi$ resulta, por una nueva aplicación del Teorema de la

Convergencia Monótona, que

$$\begin{aligned}\int_X (f+g)\,d\mu &= \lim_{n\to\infty}\int_X (\varphi_n + \psi_n)\,d\mu \\ &= \lim_{n\to\infty}\int_X \varphi_n\,d\mu + \lim_{n\to\infty}\int_X \psi_n\,d\mu \\ &= \int_X f\,d\mu + \int_X g\,d\mu.\end{aligned}$$

Para ver la última parte, considere $f_1 = f\cdot\chi_F$ y $f_2 = f\cdot\chi_G$. Se sigue de la primera parte que

$$\begin{aligned}\int_{F\cup G} f\,d\mu &= \int_X f\cdot\chi_{F\cup G}\,d\mu = \int_X f\cdot(\chi_F+\chi_G)\,d\mu \\ &= \int_X (f_1+f_2)\,d\mu = \int_X f_1\,d\mu + \int_X f_2\,d\mu \\ &= \int_F f\,d\mu + \int_G f\,d\mu.\end{aligned}$$

Los detalles de las demostraciones de las partes (ii), (iii) y (iv) son fáciles y se dejan a cargo del lector. ■

Es importante observar que la Desigualdad de Chebyshev se cumple para cualquier una función medible no-negativa sea ésta acotada o no.

Teorema 10.1.27 (Desigualdad de Chebyshev). *Si $f : X \to \overline{\mathbb{R}}$ es una función **medible no-negativa** sobre X, entonces*

$$\mu(\{x\in X : f(x)\geq c\}) \leq \frac{1}{c}\cdot\int_X f\,d\mu.$$

*para todo $c>0$. En particular, si f es **Lesbesgue integrable** sobre X, entonces*

$$\lim_{c\to\infty}\mu(\{x\in X : f(x)\geq c\}) = 0.$$

Prueba. Si $\int_X f\,d\mu = +\infty$, entonces la desigualdad se cumple trivialmente. Si ahora suponemos que $\int_X f\,d\mu < +\infty$, resulta de la desigualdad $c\cdot\chi_{X_c} \leq f\cdot\chi_{X_c}$ que

$$\int_X f\,d\mu \geq \int_X c\cdot\chi_{X_c}\,d\mu = c\cdot\mu(X_c)$$

donde hemos puesto $X_c = \{x\in X : f(x)\geq c\}$. Fin de la prueba. ■

El siguiente resultado es una fácil consecuencia de la Desigualdad de Chebyshev.

Corolario 10.1.28. *Sea $f\in\mathcal{L}_1^+(X,\mu)$. Entonces*

*(a) f es **finita casi-siempre** sobre X,*

*(b) $\int_X f\,d\mu = 0$ si, sólo si, $f=0$ **casi-siempre** sobre X.*

Prueba. (a) Debemos demostrar que $\mu(E_f^\infty) = 0$, donde

$$E_f^\infty \;=\; \{x \in X : f(x) = +\infty\} \;=\; \bigcap_{n=1}^{\infty} \{x \in X : f(x) \geq n\}.$$

Puesto que $\int_X f\, d\mu < +\infty$, la Desigualdad de Chebyshev nos garantiza que, para todo $n \geq 1$,

$$\mu(\{x \in X : f(x) \geq n\}) \;\leq\; \frac{1}{n} \cdot \int_X f\, d\mu \;<\; +\infty.$$

Ahora bien, como $\mu(E_1) < +\infty$ y la sucesión $(\{x \in X : f(x) \geq n\})_{n=1}^\infty$ es decreciente, se sigue del Teorema 6.3.47, página 287, que

$$0 \;\leq\; \mu(E_f^\infty) \;=\; \lim_{n\to\infty} \mu(\{x \in X : f(x) \geq n\}) \;\leq\; \lim_{n\to\infty} \frac{1}{n} \cdot \int_X f\, d\mu \;=\; 0.$$

Esto nos indica que $\mu(E_f^\infty) = 0$ y, por lo tanto, f es finita casi-siempre. La demostración de (b) es similar a la del Corolario 10.1.15 y se deja a cargo del lector. ■

Corolario 10.1.29. *Sea $f : X \to \overline{\mathbb{R}}$ una función* ***medible no-negativa****. Si $\mu(X) = 0$, entonces*

$$\int_X f\, d\mu \;=\; 0.$$

Prueba. Puesto que $\mu(X) = 0$, entonces $\int_X \varphi\, d\mu = 0$ para cualquier función simple no-negativa $\varphi \leq f$ y, en consecuencia,

$$\int_X f\, d\mu \;=\; \sup\left\{ \int_X \varphi\, d\mu \;:\; \varphi \in \mathcal{S}_f^+(X,\mu) \right\} \;=\; 0. \tag{L8}$$

La prueba es completa. ■

Uno puede usar el corolario anterior para justificar la existencia de una función Lebesgue integrable que no es acotada en ningún entorno de los puntos de su dominio de definición.

Ejemplo 10.1.1. *Sea $\mathbb{Q}_{[0,1]} = \{q_n : n = 1, 2, \ldots\}$ los números racionales en $[0,1]$ y defina la función $f : [0,1] \to \mathbb{R}$*

$$f(x) \;=\; \begin{cases} n & si \quad x = q_n, \\ 0 & si \quad x \in [0,1] \setminus \{q_n : n = 1, 2, \ldots\}. \end{cases}$$

Entonces f es Lebesgue integrable, pero no es acotada en ningún entorno de puntos de $[0,1]$.

Prueba. En efecto, por la densidad de $\mathbb{Q}_{[0,1]}$ en $[0,1]$, resulta claro que f no es acotada en ningún entorno de cada punto de $[0,1]$. Además, como f es no-negativa, resulta, en virtud del corolario anterior, que $\int_{\mathbb{Q}_{[0,1]}} f\, d\mu = 0$ ya que $\mu(\mathbb{Q}_{[0,1]}) = 0$. Finalmente,

$$\int_{[0,1]} f\, d\mu \;=\; \int_{\mathbb{Q}_{[0,1]}} f\, d\mu + \int_{[0,1]\setminus\mathbb{Q}_{[0,1]}} f\, d\mu \;=\; 0.$$

Esto prueba que f es Lebesgue integrable. ■

El Teorema de Beppo Levi para funciones no-negativas en $\mathcal{L}_{\text{fin}}(X,\mu)$ requiere la convergencia de la serie. Para el caso general se puede omitir tal requerimiento.

Corolario 10.1.30 (Beppo Levi). *Si $(f_n)_{n=1}^{\infty}$ es una sucesión de funciones **medibles no-negativas** definidas sobre X, entonces*

$$\int_X \left(\sum_{n=1}^{\infty} f_n\right) d\mu \;=\; \sum_{n=1}^{\infty} \int_X f_n\, d\mu. \tag{1}$$

*En particular, si $\sum_{n=1}^{\infty} \int_X f_n\, d\mu$ es **finita**, entonces la serie $\sum_{n=1}^{\infty} f_n$ es **finita casi-siempre**.*

Prueba. La primera parte es idéntica a la del Corolario 10.1.21 y, en consecuencia, se omite. Par demostrar la parte final, suponga que $\sum_{n=1}^{\infty} \int_X f_n\, d\mu$ es finito. De la igualdad (1) se sigue que $\int_X \left(\sum_{n=1}^{\infty} f_n\right) d\mu$ también es finita, es decir, $\sum_{n=1}^{\infty} f_n$ es Lebesgue integrable. Se sigue del Corolario 10.1.28 que $\sum_{n=1}^{\infty} f_n$ es finita $\mu -$ c.s. y termina la prueba. ■

Una consecuencia interesante del Teorema de Beppo Levi es el siguiente corolario.

Corolario 10.1.31. *Si $f : X \to \overline{\mathbb{R}}$ es una función **medible no-negativa**, entonces la función de conjuntos $\nu_f : \mathfrak{M}_\mu(X) \to [0, +\infty]$ definida por*

$$\nu_f(E) \;=\; \int_E f\, d\mu, \quad E \in \mathfrak{M}_\mu(X)$$

*es una **medida** sobre X.*

Prueba. Para cada $n \in \mathbb{N}$, considere la función $f_n = f \cdot \chi_{E_n}$. Se sigue del Teorema de Beppo Levi que

$$\begin{aligned}
\sum_{n=1}^{\infty} \nu_f(E_n) \;&=\; \sum_{n=1}^{\infty} \int_{E_n} f\, d\mu \\
&=\; \sum_{n=1}^{\infty} \int_X f_n\, d\mu \;=\; \int_X \left(\sum_{n=1}^{\infty} f_n\right) d\mu \\
&=\; \int_X \left(f \cdot \sum_{n=1}^{\infty} \chi_{E_n}\right) d\mu \;=\; \int_X f \cdot \chi_{\bigcup_{n=1}^{\infty} E_n}\, d\mu \\
&=\; \nu_f\left(\bigcup_{n=1}^{\infty} E_n\right).
\end{aligned}$$

La prueba es completa. ■

En particular, si $f \in \mathcal{L}_1^+(X, \mu)$ y $(E_n)_{n=1}^{\infty}$ es una **sucesión monótona** de subconjuntos medibles de X, entonces

$$\nu_f\left(\bigcap_{n=1}^{\infty} E_n\right) \;=\; \lim_{n\to\infty} \nu_f(E_n)$$

si la sucesión es **decreciente** (observe que $\nu(E_1) < +\infty$); mientras que

$$\nu_f\left(\bigcup_{n=1}^{\infty} E_n\right) \;=\; \lim_{n\to\infty} \nu_f(E_n)$$

si la sucesión es **creciente**.

Como en el caso del Teorema de la Convergencia Monótona, el Lema de Fatou se expresa de la siguiente forma.

Lema 10.1.32 (Lema de Fatou). *Sea $(f_n)_{n=1}^{\infty}$ una sucesión de funciones* ***medibles no-negativas*** *definidas sobre X. Entonces*

$$\int_X \liminf_{n\to\infty} f_n \, d\mu \ \leq \ \liminf_{n\to\infty} \int_X f_n \, d\mu.$$

Prueba. Para cada $k \geq 1$, sea

$$g_k \ = \ \inf\{f_k, f_{k+1}, \ldots\}.$$

Resulta que $(g_k)_{k=1}^{\infty}$ es una sucesión creciente de funciones medibles no-negativas para la cual se cumple que $\lim\limits_{k\to\infty} g_k = \liminf\limits_{k\to\infty} f_k$. Se sigue entonces del Teorema de la Convergencia Monótona que

$$\int_X \liminf_{n\to\infty} f_n \, d\mu \ = \ \int_X \lim_{k\to\infty} g_k \, d\mu \ = \ \lim_{k\to\infty} \int_X g_k \, d\mu. \tag{1}$$

Más aun, como $g_k \leq f_n$ para todo $n \geq k$, se tiene que

$$\int_X g_k \, d\mu \ \leq \ \int_X f_n \, d\mu \quad \text{para todo} \quad n \geq k$$

y, en consecuencia,

$$\int_X g_k \, d\mu \ \leq \ \liminf_{n\to\infty} \int_X f_n \, d\mu \quad \text{para todo} \quad k \geq 1.$$

Esto último, combinado con (1), termina la prueba. ∎

Observe que la desigualdad en el Lema de Fatou puede ser estricta. En efecto, considerando $f_n = -\chi_{[n-1,n]}$ para $n = 1, 2, \ldots$, resulta que $\liminf_{n\to\infty} f_n = 0$, mientras que $\int_{\mathbb{R}} f_n \, d\mu = 1$ para todo $n \geq 1$. Por otro lado, la existencia de los límites $\lim_{n\to\infty} f_n$ y $\lim_{n\to\infty} \int_X f_n \, d\mu$ no son imprescindibles en el Lema de Fatou. Por supuesto, si ambos límites existen, entonces

$$\int_X \lim_{n\to\infty} f_n \, d\mu \ \leq \ \lim_{n\to\infty} \int_X f_n \, d\mu.$$

Corolario 10.1.33. *Sea $(f_n)_{n=1}^{\infty}$ una sucesión de funciones* ***medibles no-negativas*** *definidas sobre X y suponga que*

(a) $\lim\limits_{n\to\infty} f_n = f$ *puntualmente sobre X y*

(b) $f_n \leq f$ *para todo $n \geq 1$.*

Entonces

$$\int_X f \, d\mu \ = \ \lim_{n\to\infty} \int_X f_n \, d\mu.$$

Prueba. Puesto que $f_n \leq f$ para todo $n \geq 1$, se tiene que $\int_X f_n\,d\mu \leq \int_X f\,d\mu$ para todo $n \geq 1$ y, en consecuencia,

$$\limsup_{n\to\infty} \int_X f_n\,d\mu \;\leq\; \int_X f\,d\mu.$$

Por otro lado, gracias al Lema de Fatou sabemos que

$$\int_X f\,d\mu \;=\; \int_X \liminf_{n\to\infty} f_n\,d\mu \;\leq\; \liminf_{n\to\infty} \int_X f_n\,d\mu,$$

lo que combinado con la desigualdad anterior produce

$$\limsup_{n\to\infty} \int_X f_n\,d\mu \;\leq\; \int_X f\,d\mu \;\leq\; \liminf_{n\to\infty} \int_X f_n\,d\mu.$$

Esto termina la prueba. ■

Teorema 10.1.34 (Convergencia Dominada). *Sea $(f_n)_{n=1}^{\infty}$ una sucesión de funciones en $\mathcal{L}_1^+(X,\mu)$ convergiendo casi-siempre a una función $f : X \to \overline{\mathbb{R}}$. Si existe una función $g \in \mathcal{L}_1^+(X,\mu)$ tal que $f_n \leq g$ para todo $n \geq 1$. Entonces $f \in \mathcal{L}_1^+(X,\mu)$ y*

$$\int_X f\,d\mu \;=\; \lim_{n\to\infty} \int_X f_n\,d\mu.$$

Prueba. Puesto que $0 \leq f_n \leq g$ para todo $n \geq 1$ y $f_n \to f$ casi siempre, resulta que $f \geq 0$ casi-siempre. Se sigue del Lema de Fatou que

$$\int_X f\,d\mu \;=\; \int_X \liminf_{n\to\infty} f_n\,d\mu \;\leq\; \liminf_{n\to\infty} \int_X f_n\,d\mu \;\leq\; \int_X g\,d\mu$$

Esto prueba que $f \in \mathcal{L}_1^+(X,\mu)$ y

$$\int_X f\,d\mu \;\leq\; \liminf_{n\to\infty} \int_X f_n\,d\mu. \tag{1}$$

Por otro lado, como $g - f_n \geq 0$ para todo $n \geq 1$, una nueva aplicación del Lema de Fatou nos indica que

$$\int_X \liminf_{n\to\infty} (g - f_n)\,d\mu \;\leq\; \liminf_{n\to\infty} \int_X (g - f_n)\,d\mu. \tag{1}$$

Pero

$$\liminf_{n\to\infty} \int_X (g - f_n)\,d\mu \;=\; \int_X g\,d\mu \;-\; \int_X \limsup_{n\to\infty} f_n\,d\mu$$

y,

$$\liminf_{n\to\infty} \int_X (g - f_n)\,d\mu \;\geq\; \int_X g\,d\mu \;-\; \limsup_{n\to\infty} \int_X f_n\,d\mu.$$

De donde se deduce que

$$\limsup_{n\to\infty} \int_X f_n\,d\mu \;\leq\; \int_X \limsup_{n\to\infty} f_n\,d\mu. \tag{2}$$

Combinando las desigualdades (1) y (2) se tiene que

$$\begin{aligned}\limsup_{n\to\infty}\int_X f_n\,d\mu \;&\leq\; \int_X \limsup_{n\to\infty} f_n\,d\mu \\ &=\; \int_X \liminf_{n\to\infty} f_n\,d\mu \;\leq\; \liminf_{n\to\infty}\int_X f_n\,d\mu\end{aligned}$$

y así,

$$\int_X f\,d\mu \;=\; \lim_{n\to\infty}\int_X f_n\,d\mu.$$

La prueba es completa. ■

Nota Adicional 10.1.4 De (L8) sabemos que si f es medible y no-negativa, entonces se cumple que $\int_A f\,d\mu = 0$ para cualquier conjunto $A \subseteq X$ con $\mu(A) = 0$, de modo que los valores de f en los puntos de un conjunto de medida nula no afectan el valor de la integral sobre el conjunto X y, en consecuencia, se puede prescindir de tales conjuntos en el proceso de integración. Más aun, del Corolario 10.1.28 sabemos que si f es una función medible no-negativa la cual es Lebesgue integrable, entonces f es finita casi-siempre, es decir, $0 \leq f(x) < +\infty$ para todo $x \in E_0 = X \setminus E_f^{\infty}$. Como $\mu(E_f^{\infty}) = 0$, resulta de lo anterior y las propiedades de la integral que

$$\int_X f\,d\mu \;=\; \int_{E_0} f\,d\mu + \int_{E_f^{\infty}} f\,d\mu \;=\; \int_{E_0} f\,d\mu.$$

Esa igualdad nos revela que *la integral de f sobre X es exactamente igual a la integral de f sobre el conjunto donde ella es finita*. Esta información es muy valiosa y, en consecuencia, a partir de este momento, y sólo para efectos de integración, **asumiremos que todas las funciones medibles que son Lebesgue integrables son a valores reales**. Igual consideración haremos para la convergencia casi-siempre, es decir, **supondremos, cuando exista convergencia casi-siempre de funciones integrables, que la convergencia es simplemente convergencia puntual**.

(b) **Tercera Suposición**. **Suponga que $X \in \mathfrak{M}_\mu(\mathbb{R})$ con $\mu(X) < +\infty$ y $f : X \to \overline{\mathbb{R}}$ es una función medible arbitraria**.

Considere las partes positiva y negativa de f, esto es,

$$f^+ \;=\; \max\{f, 0\} \qquad\qquad \text{y} \qquad\qquad f^- \;=\; -\min\{f, 0\}.$$

Entonces f^+ y f^- son funciones medibles no-negativas y así, por el análisis anterior, sus respectivas integrales, $\int_X f^+\,d\mu$ y $\int_X f^-\,d\mu$ están bien definidas como elementos de $\overline{\mathbb{R}}$. Si *al menos una de esas integrales es finita*, y sólo en este caso, se define **la integral de** f sobre X como

$$\int_X f\,d\mu \;=\; \int_X f^+\,d\mu \;-\; \int_X f^-\,d\mu. \tag{IL4}$$

Definición 10.1.35. *Sea $f : X \to \overline{\mathbb{R}}$ una función medible arbitraria. Diremos que f es* ***Lebesgue integrable*** *sobre X si las integrales*

$$\int_X f^+ \, d\mu \qquad y \qquad \int_X f^- \, d\mu$$

son ambas finitas.

Observe que:

$$\begin{aligned} f \text{ es Lebesgue integrable} \quad &\Leftrightarrow \quad \int_X f^+ \, d\mu \ \text{ y } \ \int_X f^- \, d\mu \ \text{ son finitas} \\ &\Leftrightarrow \quad \int_X f^+ \, d\mu - \int_X f^- \, d\mu \ < \ \infty \\ &\Leftrightarrow \quad \int_X f^+ \, d\mu + \int_X f^- \, d\mu \ < \ \infty \end{aligned}$$

Por consiguiente, si f es Lebesgue integrable, definimos la **integral** de $|f|$ como

$$\int_X |f| \, d\mu \ = \ \int_X f^+ \, d\mu \ + \ \int_X f^- \, d\mu. \tag{IL5}$$

Algunos autores usan el término "*sumable*" en lugar de Lebesgue integrable. Si $\mu(X) < +\infty$, el símbolo $\mathcal{L}_1(X, \mu)$ *será usado para representar el conjunto de todas las funciones medibles $f : X \to \overline{\mathbb{R}}$ tal que*

$$\int_X |f| \, d\mu \ < \ +\infty.$$

Como antes, si $E \subseteq X$ es medible y $f \in \mathcal{L}_1(X, \mu)$, definimos

$$\int_E f \, d\mu \ = \ \int_X f \cdot \chi_E \, d\mu.$$

Una primera consecuencia de la definición anterior es que si f es Lebesgue integrable sobre X, entonces

$$\left| \int_X f \, d\mu \right| \ \leq \ \int_X |f| \, d\mu. \tag{L9}$$

En efecto,

$$\left| \int_X f \, d\mu \right| \ = \ \left| \int_X f^+ \, d\mu \ - \ \int_X f^- \, d\mu \right| \ \leq \ \int_X f^+ \, d\mu \ + \ \int_X f^- \, d\mu \ = \ \int_X |f| \, d\mu.$$

De esto último se deduce que si $f \in \mathcal{L}_1(X, \mu)$, entonces

$$\mu(X) = 0 \qquad \Rightarrow \qquad \int_X f \, d\mu = 0. \tag{L10}$$

En efecto, como $|f| \geq 0$, se sigue del Corolario 10.1.29 que $\int_X |f| \, d\mu = 0$ y la conclusión sigue de la desigualdad anterior.

Nota Adicional 10.1.5 Para poder verificar que $\mathcal{L}_1(X,\mu)$ es un espacio vectorial, debemos estar seguros que podemos definir, sin ambigüedad, $f+g$ cualesquiera sean las funciones $f,g \in \mathcal{L}_1(X,\mu)$. Este hecho, en principio, presenta un problema con aquellos valores $x \in X$ para los cuales, por ejemplo, $f(x) = +\infty$ y $g(x) = -\infty$. ¿Cómo se resuelve esta situación? Pues bien, como f^+ y f^- son funciones Lebesgue integrables no-negativas, el Corolario 10.1.28 nos garantiza que ellas son finitas casi-siempre y, por consiguiente, el conjunto $A_f = \{x \in X : f^+(x) = f^-(x) = +\infty\}$ posee medida nula. Similarmente, el conjunto $B_g = \{x \in X : g^+(x) = g^-(x) = +\infty\}$ tiene medida cero. Si definimos $E_0 = A_f \cap B_g$, resulta que E_0 también posee medida cero. Observe que si $x \in E_0$, entonces

$$f^+(x) - g^-(x) \;=\; +\infty - \infty \qquad\qquad \text{y} \qquad\qquad g^+(x) - f^-(x) \;=\; +\infty - \infty$$

y como $f+g = (f^+ - f^-) + (g^+ - g^-) = (f^+ - g^-) + (g^+ - f^-)$, se tiene que

$$\int_{E_0} (f+g)\, d\mu \;=\; 0.$$

Por esta razón, siempre podemos redefinir $f+g$ sobre E_0 como se nos antoje, de modo que su integral sobre E_0 siempre será nula y, en consecuencia, *convenimos en redefinir* $(f+g)(x) = 0$ *para todo* $x \in E_0$ *y suponer, en consecuencia, que* $f+g$ *está bien definida sobre* X *cualesquiera sean* $f,g \in \mathcal{L}_1(X,\mu)$.

Una vez aceptada la observación anterior, se prueba que $\mathcal{L}_1(X,\mu)$ es un espacio vectorial sobre $\mathbb{R}$. Además, se cumple que,

Teorema 10.1.36. *Si* $f,g \in \mathcal{L}_1(X,\mu)$, *entonces*

(*i*) $\displaystyle\int_X (f+g)\, d\mu \;=\; \int_X f\, d\mu \;+\; \int_X g\, d\mu.$

(*ii*) $\displaystyle\int_X (af)\, d\mu \;=\; a\int_X f\, d\mu$ *para cualquier* $a \in \mathbb{R}$.

Prueba. (i) Hagamos $h = f+g$ y observe que como $h^+ - h^- = f^+ - f^- + g^+ - g^-$ entonces

$$h^+ \,+\, f^- \,+\, g^- \;=\; f^+ \,+\, g^+ \,+\, h^-.$$

Puesto que las integrales de cada una de estas funciones no-negativas es finita, integrando término a término, se obtiene

$$\int_X h^+\, d\mu \,+\, \int_X f^-\, d\mu \,+\, \int_X g^-\, d\mu \;=\; \int_X f^+\, d\mu \,+\, \int_X g^+\, d\mu \,+\, \int_X h^-\, d\mu$$

lo cual es equivalente a escribir

$$\int_X (f+g)\, d\mu \;=\; \int_X h\, d\mu \;=\; \int_X (h^+ - h^-)\, d\mu \;=\; \int_X f\, d\mu \,+\, \int_X g\, d\mu.$$

La prueba de (ii) se deja a cargo del lector. ■

Por supuesto, si $(f_n)_{n=1}^{\infty}$ una sucesión en $\mathcal{L}_1(X,\mu)$ que converge uniformemente sobre X a una función f entonces f es Lebesgue integrable sobre X y para cualquier $E \in \mathfrak{M}_\mu(X)$,

$$\lim_{n\to\infty} \int_E |f_n|\, d\mu \;=\; \int_E |f|\, d\mu. \tag{L11}$$

El Teorema de la Convergencia Monótona para funciones crecientes no necesariamente no-negativas, en general, no se cumple. Sin embargo, añadiendo alguna hipótesis adicional se obtiene un resultado similar. Los siguientes dos resultados dan fe de ello.

Teorema 10.1.37. *Sea $(f_n)_{n=1}^{\infty}$ una sucesión **monótona creciente** en $\mathcal{L}_1(X,\mu)$. Si*

(a) $\lim\limits_{n\to\infty} f_n = f$ puntualmente sobre X y

*(b) $f_n \geq g$ para todo $n \geq 1$, donde g es una **función Lebesgue integrable** sobre X,*

entonces

$$\int_X f\, d\mu \;=\; \lim_{n\to\infty} \int_X f_n\, d\mu.$$

Prueba. Considere la sucesión de funciones medibles no-negativas $(f_n - g)_{n=1}^{\infty}$ y aplique el Teorema de la Convergencia Monótona para obtener

$$\begin{aligned}
\lim_{n\to\infty} \int_X f_n\, d\mu - \int_X g\, d\mu \;&=\; \lim_{n\to\infty} \int_X (f_n - g)\, d\mu \\
&=\; \int_X \lim_{n\to\infty} (f_n - g)\, d\mu \\
&=\; \int_X f\, d\mu - \int_X g\, d\mu.
\end{aligned}$$

Puesto que g es Lebesgue integrable, podemos cancelar $\int_X g\, d\mu$ a ambos lados de la igualdad anterior y termina la prueba. ■

Una de las bondades del Teorema de la Convergencia Monótona consiste en chequear si la integral de una función límite es o no finita.

Teorema 10.1.38 (Convergencia Monótona). *Sea $(f_n)_{n=1}^{\infty}$ una sucesión **monótona** en $\mathcal{L}_1(X,\mu)$. Suponga que*

(a) $\lim\limits_{n\to\infty} f_n = f$ puntualmente sobre X y

(b) $\lim\limits_{n\to\infty} \int_X f_n\, d\mu < +\infty$.

Entonces $f \in \mathcal{L}_1(X,\mu)$ y

$$\int_X f\, d\mu \;=\; \lim_{n\to\infty} \int_X f_n\, d\mu.$$

Prueba. Suponga que la sucesión $(f_n)_{n=1}^{\infty}$ es decreciente (el caso creciente se argumenta de modo similar). Considerando ahora la sucesión $(f_1 - f_n)_{n=1}^{\infty}$, resulta que ella es creciente, no-negativa y converge puntualmente a $f_1 - f$. Se sigue entonces del Teorema de la Convergencia Monótona que

$$\int_X (f_1 - f)\,d\mu \;=\; \lim_{n\to\infty} \int_X (f_1 - f_n)\,d\mu \;=\; \int_X f_1\,d\mu \;-\; \lim_{n\to\infty} \int_X f_n\,d\mu.$$

Puesto que, por hipótesis, el lado derecho de ésta igualdad es finito, tenemos que $f_1 - f$ es Lebesgue integrable sobre X y, en consecuencia, $f = f_1 - (f_1 - f)$ también es Lebesgue integrable. Aplicando la propiedad de aditividad de la integral y usando la igualdad anterior se tiene que

$$\int_X f\,d\mu \;=\; \int_X f_1\,d\mu \;-\; \int_X (f_1 - f)\,d\mu \;=\; \lim_{n\to\infty} \int_X f_n\,d\mu.$$

Fin de la prueba. ■

Sabemos que convergencia puntual c.s. implica convergencia en medida (Teorema 7.2.25, página 395) pero no recíprocamente. Un hecho interesante que permite validar la convergencia en medida como un concepto natural, es que uno puede debilitar la hipótesis del Teorema de la Convergencia Acotada para obtener la misma conclusión si, en lugar de exigir convergencia casi-siempre, se pide convergencia en medida.

Teorema 10.1.39 (Convergencia Acotada). *Sea $(f_n)_{n=1}^{\infty}$ una sucesión de funciones en el espacio vectorial $\mathcal{L}_1(X,\mu)$ tal que*

(a) $f_n \to f$ en ***medida*** *y*

(b) $(f_n)_{n=1}^{\infty}$ es ***uniformemente acotada****.*

Entonces $f \in \mathcal{L}_1(X,\mu)$ y

$$\lim_{n\to\infty} \int_X f_n\,d\mu \;=\; \int_X f\,d\mu.$$

Prueba. Para simplificar la prueba supondremos que $\mu(X) = 1$. El hecho de que la sucesión $(f_n)_{n=1}^{\infty}$ es uniformemente acotada, garantiza la existencia de una constante $M > 0$ tal que $|f_n(x)| \leq M$ para todo $n \geq 1$ y todo $x \in X$. Nuestro primer objetivo será demostrar que $|f| \leq M$ casi-siempre. Para ver esto, pongamos $[|g| > c] = \{x \in X : |g(x)| > c\}$ y observe que como $|f| \leq |f_n| + |f - f_n| \leq M + |f - f_n|$ para todo $n \in \mathbb{N}$, resulta que si $|f_n(x) - f(x)| \leq 1/k$, entonces $|f(x)| \leq M + 1/k$, de donde se sigue

$$\big[\,|f| > M + 1/k\,\big] \;\subseteq\; \big[\,|f_n - f| > 1/k\,\big]$$

para cada $k \in \mathbb{N}$. Por esto, si $k \in \mathbb{N}$, entonces

$$\mu\big([\,|f| > M + 1/k\,]\big) \;\leq\; \mu\big([\,|f_n - f| > 1/k\,]\big) \quad \text{para todo} \quad n \geq 1$$

y usando el hecho de que $f_n \xrightarrow{\mu} f$, se tiene que

$$0 \;\leq\; \mu\big([\,|f| > M + 1/k\,]\big) \;\leq\; \lim_{n\to\infty} \mu\big([\,|f_n - f| > 1/k\,]\big) \;=\; 0.$$

Por esto,

$$\mu\big([\,|f| > M\,]\big) \;=\; \mu\bigg(\bigcup_{k=1}^{\infty} [\,|f| > M + 1/k\,]\bigg) \;=\; 0,$$

con lo cual queda demostrado que $|f| \leq M$ c.s. y, por lo tanto, f es Lebesgue integrable.

Sea $\varepsilon > 0$ y use de nuevo la hipótesis de que $f_n \to f$ en medida para seleccionar un $N \in \mathbb{N}$ tal que

$$\mu\big([\,|f_n - f| > \varepsilon/3]\big) \;<\; \varepsilon/3M \quad \text{para todo} \quad n \geq N.$$

Finalmente, si $n \geq N$, entonces

$$\begin{aligned}
\left|\int_X (f_n - f)\, d\mu\right| \;&\leq\; \int_X |f_n - f|\, d\mu \\
&= \int_{[\,|f_n - f| > \varepsilon/3]} |f_n - f|\, d\mu \;+\; \int_{[\,|f_n - f| \leq \varepsilon/3]} |f_n - f|\, d\mu \\
&< 2M \cdot \frac{\varepsilon}{3M} \;+\; \varepsilon/3 \;=\; \varepsilon.
\end{aligned}$$

La prueba es completa. ■

El Teorema de la Convergencia Dominada para funciones en $\mathcal{L}_1(X, \mu)$ se formula de modo enteramente similar a la del Teorema 10.1.22.

Teorema 10.1.40 (Convergencia Dominada). *Sea $(f_n)_{n=1}^{\infty}$ una sucesión de funciones medibles definidas sobre X y suponga que $g \in \mathcal{L}_1(X, \mu)$. Si se cumple que*

(a) $\lim\limits_{n\to\infty} f_n \;=\; f$ puntualmente sobre X y

(b) $|f_n| \leq g$ para todo $n \geq 1$,

entonces $f, f_n \in \mathcal{L}_1(X, \mu)$ y

$$\int_X f\, d\mu \;=\; \lim_{n\to\infty} \int_X f_n\, d\mu. \tag{1}$$

En particular,

$$\lim_{n\to\infty} \int_X |f_n - f|\, d\mu \;=\; 0. \tag{2}$$

Prueba. La prueba de (1) se omite por ser idéntica a la del Teorema 10.1.22. Para demostrar la segunda igualdad, observe que si consideramos la sucesión $(f_n - f)_{n=1}^{\infty}$ resulta que

$$\lim_{n\to\infty} |f_n - f| \;=\; 0 \qquad \text{y} \qquad |f_n - f| \;\leq\; 2g, \quad n \geq 1$$

y, en consecuencia, por la primera parte se obtiene (2). ■

Nota Adicional 10.1.6 Es importante destacar que la condición "$|f_n| \leq g$ para todo $n \geq 1$, donde $g \in \mathcal{L}_1(X, \mu)$" en el Teorema de la Convergencia Dominada es equivalente a requerir que la función $\sup_n |f_n|$ sea **Lebesgue integrable** sobre X.

Corolario 10.1.41. *Bajo las hipótesis del teorema precedente, se tiene que*

$$\int_X f \cdot \varphi\, d\mu \;=\; \lim_{n\to\infty} \int_X f_n \cdot \varphi\, d\mu$$

*para cualquier **función medible** y **acotada** $\varphi : X \to \mathbb{R}$.*

Prueba. Sea φ una función medible y acotada. Si M es una constante tal que $|\varphi| \leq M$. entonces $M \cdot g \in \mathcal{L}_1(X, \mu)$ y se cumple que

(a) $\lim\limits_{n\to\infty} f_n \cdot \varphi = f \cdot \varphi$ y

(b) $|f_n \cdot \varphi| \leq M \cdot g$ para todo $n \geq 1$.

Se sigue del Teorema de la Convergencia Dominada que

$$\int_X f \cdot \varphi \, d\mu \;=\; \lim_{n\to\infty} \int_X f_n \cdot \varphi \, d\mu$$

y termina la prueba. ■

El Teorema de Beppo Levi para funciones Lebesgue integrable requiere, para su validez, una condición adicional.

Teorema 10.1.42 (Beppo Levi). *Sea $(f_n)_{n=1}^\infty$ una sucesión en $\mathcal{L}_1(X, \mu)$. Si $\sum_{n=1}^\infty \int_X |f_n| \, d\mu$* ***converge****, entonces $\sum_{n=1}^\infty f_n$ converge $\mu - \text{c.s.}$ a una función Lebesgue integrable sobre X. Más aun,*

$$\left| \int_X \left(\sum_{n=1}^\infty f_n \right) d\mu \right| \;\leq\; \int_X \left| \sum_{n=1}^\infty f_n \right| d\mu \;\leq\; \sum_{n=1}^\infty \int_X |f_n| \, d\mu$$

y

$$\int_X \left(\sum_{n=1}^\infty f_n \right) d\mu \;=\; \sum_{n=1}^\infty \int_X f_n \, d\mu.$$

Prueba. La función $g = \sum_{n=1}^\infty |f_n|$ es medible y no-negativa. Del Teorema de la Convergencia Monótona o, en su defecto, del Corolario 10.1.21 sabemos que

$$\int_X g \, d\mu \;=\; \sum_{n=1}^\infty \int_X |f_n| \, d\mu \;<\; \infty.$$

Esto nos dice que g es Lebesgue integrable sobre X y, en consecuencia, finita μ-c.s., la cual podemos asumir que es finita por una observación hecha anteriormente. Más aun,

$$\left| \sum_{n=1}^\infty f_n \right| \;\leq\; \sum_{n=1}^\infty |f_n| \;<\; \infty$$

Esta desigualdad nos revela que la serie $\sum_{n=1}^\infty f_n$ converge absolutamente a una función f que satisface $|f| \leq g$. Por supuesto,

$$\left| \int_X f \, d\mu \right| \;\leq\; \int_X |f| \, d\mu \;\leq\; \int_X g \, d\mu$$

la cual no es otra cosa que la prueba de la primera parte.

Para demostrar la segunda afirmación, haremos uso del Teorema de la Convergencia Dominada. Para cada $n \geq 1$, considere la sucesión de funciones Lebesgue integrables $(g_n)_{n=1}^\infty$, donde

$g_n = \sum_{k=1}^{n} f_k$. Puesto que $g_n \to \sum_{k=1}^{\infty} f_k$ y $|g_n| \leq g$ para todo $n \geq 1$, el Teorema de la Convergencia Dominada nos garantiza que

$$\begin{aligned}\int_X \Big(\sum_{n=1}^{\infty} f_n\Big)\, d\mu &= \int_X \Big(\lim_{n\to\infty} \sum_{k=1}^{n} f_k\Big)\, d\mu \\ &= \lim_{n\to\infty} \sum_{k=1}^{n} \int_X f_k\, d\mu \\ &= \sum_{n=1}^{\infty} \int_X f_n\, d\mu.\end{aligned}$$

La prueba es completa. ■

Hasta ahora hemos visto cómo, poco a poco, se ha ido extendiendo la noción de integral de Lebesgue comenzando con la integral de una función medible y acotada hasta llegar a la integral de cualquier función medible, pero siempre en el marco de que el dominio de tales funciones fuese un conjunto medible arbitrario pero de medida finita. Resta ver cómo se extiende ese proceso de integración sobre cualquier conjunto medible de medida infinita. Tal vez algún lector encuentre dicho proceso un poco tedioso, sin embargo, y en beneficio de nuestra presentación, creemos que dicho enfoque es de gran utilidad al estudiante ya que demuestra los resultados más importantes de cada extensión con un mínimo esfuerzo. Por otro lado, bueno es decirlo, existen otros desarrollos o presentaciones de la integral de Lebesgue que parten de otro enfoque. Por ejemplo, uno de los más usados comienza directamente de la noción de integral para funciones simples no-negativas definidas sobre un conjunto medible de medida arbitraria, luego se extiende el proceso a funciones medibles no-negativas, hasta llegar a la definición de integral de cualquier función medible. Este método tiene sus ventajas: va más directo a la construcción de la integral de Lebesgue y no requiere "repetir", en cada paso, los resultados fundamentales probados con el método anterior. Lo que sigue es una breve incursión de ese enfoque, aunque sustentado en la construcción anterior.

En el Teorema 10.1.24 vimos que la integral de una función medible, no-negativa $f : X \to \overline{\mathbb{R}}$, donde $\mu(X) < +\infty$, se podía expresar en la forma

$$\int_X f\, d\mu = \sup\Big\{ \int_X \varphi\, d\mu \ :\ 0 \leq \varphi \leq f,\ \varphi \text{ es simple}\Big\}.$$

Lo que resulta interesante es que esa definición de integral también se puede utilizar aun si $\mu(X) = +\infty$. En efecto, sea f una función medible no-negativa. El Teorema 7.2.10 siempre garantiza la existencia de una función simple no negativa φ tal que $\varphi \leq f$, de forma tal que podemos considerar el supremo de las integrales de tales funciones simples del modo siguiente:

(c) **Cuarta Suposición**. **Suponga ahora que $X \in \mathfrak{M}_\mu(\mathbb{R})$ con $0 \leq \mu(X) \leq +\infty$ y $f : X \to \overline{\mathbb{R}}$ es una función medible no-negativa arbitraria**.

Definimos

$$\int_X f\, d\mu = \sup\Big\{ \int_X \varphi\, d\mu \ :\ 0 \leq \varphi \leq f,\ \varphi \text{ es simple}\Big\}. \tag{IL6}$$

Observe que si φ es una función simple no-negativa con representación canónica

$$\varphi = \sum_{j=1}^{n} a_j \cdot \chi_{E_j},$$

su integral viene dada por la expresión

$$\int_X \varphi \, d\mu = \sum_{j=1}^{n} a_j \cdot \mu(E_j).$$

Esta definición, sin embargo, requiere de un pequeño análisis si $\mu(X) = +\infty$. En efecto, si $\mu(X) = +\infty$ y $\{E_1, \ldots, E_n\}$ es una partición medible finita de X, entonces al menos uno de los conjuntos, digamos E_{j_0}, debe tener medida infinita. Si ocurre que $a_{j_0} = 0$, entonces $0 \cdot \infty = 0$ como fue establecido en $\overline{\mathbb{R}}$, por lo que tales conjuntos se pueden omitir en la representación canónica de φ ya que ellos no afectan el proceso de integración. Por otro lado, si $a_{j_0} > 0$, entonces $a_{j_0} \cdot \mu(E_{j_0}) = +\infty$ y, por lo tanto, $\int_X \varphi \, d\mu = +\infty$. De ocurrir esto último tendríamos que $\int_X f \, d\mu = +\infty$, de modo que la igualdad (IL6) se cumple trivialmente. La conclusión final es clara: *para definir la integral de f usando* (IL6), *sólo debemos considerar funciones simples no-negativas cuyas integrales sean finitas*. Sin embargo, para que φ tenga integral finita ella, necesariamente, debe ser de la forma

$$\varphi = \sum_{j=0}^{n} a_j \cdot \chi_{E_j},$$

donde $a_0 = 0$, $\mu(E_0) = +\infty$ y $\mu(E_j) < +\infty$ para $j = 1, \ldots, n$. Observe que, en este caso,

$$\mu\big(\{x \in X : \varphi(x) \neq 0\}\big) < +\infty.$$

Este acuerdo lo mantendremos de aquí en adelante.

Como antes, diremos que una función medible y no-negativa $f : X \to \overline{\mathbb{R}}$ es **Lebesgue integrable** sobre X si $\int_X f \, d\mu$, definida por (IL6), es finita y entonces escribiremos $\mathcal{L}_1^+(X, \mu)$ *para designar, indistintamente si $\mu(X)$ es o infinito, el conjunto de todas las funciones medibles, no-negativas que son Lebesgue integrables sobre X.*

(d) **Quinta y Última Suposición**. **Finalmente, suponga que $X \in \mathfrak{M}_\mu(\mathbb{R})$ con $0 \leq \mu(X) \leq +\infty$ y $f : X \to \overline{\mathbb{R}}$ es una función medible arbitraria**.

En este caso, definimos **la integral de** f sobre X como

$$\int_X f \, d\mu = \int_X f^+ \, d\mu - \int_X f^- \, d\mu$$

siempre que una (o ambas) de las funciones f^+ o f^- pertenezcan a $\mathcal{L}_1^+(X, \mu)$, es decir, que una de las integrales $\int_X f^+ \, d\mu$ o $\int_X f^- \, d\mu$ sea finita. Como antes, diremos que f es **Lebesgue integrable** sobre X si $\int_X f^+ \, d\mu$ y $\int_X f^- \, d\mu$ son ambas finitas. En este caso $|f|$ también es integrable sobre X y entonces se define

$$\int_X |f| \, d\mu = \int_X f^+ \, d\mu + \int_X f^- \, d\mu < +\infty.$$

En particular, si f es Lebesgue integrable, entonces

$$\left|\int_X f\,d\mu\right| \;\le\; \int_X |f|\,d\mu.$$

Tal y como se hizo anteriormente, convenimos en seguir usando el símbolo $\mathcal{L}_1(X,\mu)$ para denotar *el conjunto de todas las funciones medibles* $f : X \to \overline{\mathbb{R}}$ *que son Lebesgue integrables sobre* X, *donde la medida del conjunto medible* X *puede ser finita o infinita*. Resulta que $\mathcal{L}_1(X,\mu)$ es un espacio vectorial sobre $\mathbb{R}$ y se cumple que si $f,g \in \mathcal{L}_1(X,\mu)$, entonces

(i) $\displaystyle\int_X (f+g)\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu$, y

(ii) $\displaystyle\int_X (a\cdot f)\,d\mu = a\int_X f\,d\mu$ para todo $a \in \mathbb{R}$.

Por supuesto, todos los resultados sobre convergencia: el Teorema de la Convergencia Acotada, el de la Convergencia Monótona y el de la Convergencia Dominada son válidos en $\mathcal{L}_1(X,\mu)$. Igualmente se cumple el Lema de Fatou. Los detalles se dejan a cargo del lector.

Nota Adicional 10.1.7 Recordemos que en $\overline{\mathbb{R}}$ definimos $0\cdot\infty = 0$. El por qué de este convenio tiene que ver con la propiedad (ii) del resultado anterior. En efecto, para que exista coherencia con dicha propiedad, debemos aceptar que $0\cdot\infty = 0$ ya que, si tomamos $a = 0$ y $f = \chi_{\mathbb{R}}$, entonces $0 = 0\cdot\chi_{\mathbb{R}}$ y, por lo tanto,

$$0 \;=\; \int_{\mathbb{R}} 0\cdot\chi_{\mathbb{R}}\,d\mu \;=\; 0\cdot\int_{\mathbb{R}}\chi_{\mathbb{R}}\,d\mu \;=\; 0\cdot\mu(\mathbb{R}) \;=\; 0\cdot\infty.$$

Algunas de las hipótesis de ciertos resultados ya vistos no se pueden debilitar y, menos aun, suprimir. Por ejemplo:

(a) En el **Teorema de la Convergencia Uniforme**, el requerimiento de que $\mu(X)$ sea finito no se puede omitir. En efecto, sea $X = \mathbb{R}$ y para cada $n \in \mathbb{N}$, considere la función

$$f_n \;=\; \frac{1}{n}\chi_{[0,n)}$$

Pongamos $f = 0$. Resulta que $f_n \to f$ uniformemente, pero

$$\int_X f_n\,d\mu \;=\; 1.$$

(b) En el **Teorema de la Convergencia Monótona**, la sucesión monótona creciente $(f_n)_{n=1}^{\infty}$ no siempre puede ser reemplazada por una sucesión monótona decreciente. Para ver esto, considere la sucesión decreciente $(f_n)_{n=1}^{\infty}$, donde $f_n = \chi_{[n,+\infty)}$. Entonces $f_n \to 0$ pero

$$\lim_{n\to\infty}\int_{\mathbb{R}} f_n\,d\mu$$

no existe. Sin embargo, *si* f_1 *es* ***Lebesgue integrable*** *sobre* X *y* $(f_n)_{n=1}^{\infty}$ *es* ***monótona decreciente*** *con* $f_n \to f$, *entonces se cumple que*

$$\lim_{n\to\infty}\int_X f_n\,d\mu \;=\; \int_X f\,d\mu.$$

Los casos fundamentalmente de interés referentes al espacio $\mathcal{L}_1(X,\mu)$ ocurren cuando X es un intervalo de la forma $[a,b]$, $[0,+\infty)$ o $\mathbb{R}$. Con frecuencia, si $X=\mathbb{R}$, escribiremos $\mathcal{L}_1(\mu)$ en lugar $\mathcal{L}_1(\mathbb{R},\mu)$. También, si $X=[a,b]$, pondremos $\mathcal{L}_1([a,b])$ en lugar de $\mathcal{L}_1([a,b],\mu)$ y, en algunos casos, la notación

$$\int_a^b f\,d\mu \qquad \text{en lugar de} \qquad \int_{[a,b]} f\,d\mu$$

será usada. Más aun, en muchas ocasiones trabajaremos con el espacio $\mathcal{L}_1([0,1])$ en lugar de $\mathcal{L}_1([a,b])$, aunque debe quedar claro que los resultados válidos en el primer espacio también lo son para el segundo espacio. La única finalidad de elegir el intervalo $[0,1]$ es porque ello simplifica considerablemente los cálculos y no por otra razón. Puesto que

$$\int_X f\,d\mu \;=\; \int_{\mathbb{R}} f\cdot\chi_X\,d\mu$$

para cualquier conjunto medible $X\subseteq\mathbb{R}$, no hay ningún peligro real en trabajar exclusivamente en el espacio $\mathcal{L}_1(\mu)$ como el caso típico (y esto es precisamente lo que vamos a hacer salvo que se indique otra cosa). Si $f\in\mathcal{L}_1(\mu)$, entonces, en ocasiones, sustituiremos el símbolo

$$\int_{\mathbb{R}} f\,d\mu \qquad \text{por} \qquad \int_{-\infty}^{+\infty} f\,d\mu.$$

De la Teoría de Integración de Riemann sabemos que si $f:[0,+\infty)\to\mathbb{R}$ es una función tal que $f\in\mathcal{R}([0,n])$ para todo $n\geq 1$, entonces la **integral impropia** de f se define como

$$\int_0^{\infty} f\,dx \;=\; \lim_{n\to\infty}\int_0^n f\,dx.$$

siempre que dicho límite exista en $\mathbb{R}$, es decir, sea un número finito. Sin embargo, en el caso de la integración de Lebesgue, la validez de la igualdad anterior es un teorema, es decir:

Teorema 10.1.43. *Sea $f\in\mathcal{L}_1(\mu)$. Entonces*

(a) $\displaystyle\int_{-\infty}^{+\infty} f\,d\mu = \lim_{n\to+\infty}\int_{-\infty}^{n} f\,d\mu.$

Además,

(b) $\displaystyle\lim_{n\to+\infty}\int_{n}^{+\infty} f\,d\mu = 0.$

(c) $\displaystyle\int_{-\infty}^{+\infty} f\,d\mu = \lim_{n\to+\infty}\int_{-n}^{n} f\,d\mu.$

Prueba. Para obtener (a) considere la sucesión $(f_n)_{n=1}^{\infty}$, donde

$$f_n \;=\; f\cdot\chi_{(-\infty,n]} \quad n=1,2,\ldots$$

Note que cada $f_n \in \mathcal{L}_1(\mu)$, $f_n \to f$ puntualmente y $|f_n| \leq |f|$ para todo $n \geq 1$. El Teorema de la Convergencia Dominada nos garantiza que (a) se cumple. Puesto que

$$\int_n^{\infty} f\, d\mu \;=\; \int_{-\infty}^{+\infty} f\, d\mu \;-\; \int_{-\infty}^{n} f\, d\mu$$

resulta que (b) sigue de (a). Para la parte (c) considere $f_n = f \cdot \chi_{[-n,n]}$ y aplique, de nuevo, el Teorema de la Convergencia Dominada. ■

Finalizamos esta sección con el siguiente resultado conocido como la Fórmula de Cavalieri y una elegante aplicación.

Teorema 10.1.44 (Fórmula de Cavalieri). *Sea $f : X \to \overline{\mathbb{R}}$ una **función medible no-negativa**. Entonces*

$$\int_X f\, d\mu \;=\; \int_0^{+\infty} \mu\big(\{x \in X \,:\, f(x) > t\}\big)\, d\mu(t).$$

Prueba. Suponga, en primer lugar, que f es una función simple, digamos $f = \sum_{j=1}^{N} a_j \cdot \chi_{E_j}$, donde los a_j son distintos y cada $E_j = \{x \in X : f(x) = a_j\}$. Para cada $t > 0$, sea $\alpha_j(t) = \chi_{[0,\,a_j]}(t)$. Entonces

$$\mu\big(\{x \in X \,:\, f(x) > t\}\big) \;=\; \sum_{j=1}^{N} \alpha_j(t) \cdot \mu(E_j),$$

de donde se sigue que

$$\begin{aligned}
\int_0^{+\infty} \mu\big(\{x \in X \,:\, f(x) > t\}\big)\, d\mu(t) \;&=\; \sum_{j=1}^{N} \bigg(\int_0^{+\infty} \alpha_j(t)\, dt \bigg) \mu(E_j) \\
&=\; \sum_{j=1}^{N} a_i \cdot \mu(E_j) \;=\; \int_X f\, d\mu.
\end{aligned}$$

Esto prueba que la fórmula de Cavalieri es válida para funciones simples. Suponga ahora que f es medible no-negativa y seleccione una sucesión creciente $(\varphi_n)_{n=1}^{\infty}$ de funciones simples tal que $\varphi_n \to f$ puntualmente. Puesto que la sucesión de conjuntos medibles $(\{x \in X \,:\, \varphi_n(x) > t\})_{n=1}^{\infty}$ es creciente para cada $t > 0$, se tiene que

$$\lim_{n\to\infty} \mu\big(\{x \in X \,:\, \varphi_n(x) > t\}\big) \;=\; \mu\big(\{x \in X \,:\, f(x) > t\}\big)$$

y, en consecuencia, gracias al Teorema de la Convergencia Dominada y a la primera parte se concluye que

$$\begin{aligned}
\int_X f\, d\mu \;&=\; \lim_{n\to\infty} \int_X \varphi_n\, d\mu \\
&=\; \lim_{n\to\infty} \int_0^{+\infty} \mu\big(\{x \in X \,:\, \varphi_n(x) > t\}\big)\, d\mu(t) \\
&=\; \int_0^{+\infty} \mu\big(\{x \in X \,:\, f(x) > t\}\big)\, d\mu(t).
\end{aligned}$$

La prueba es completa. ■

Corolario 10.1.45. *Sea $f : X \to \mathbb{R}$ una* ***función medible Lebesgue*** *y suponga que existe una constante $c > 0$ tal que*

$$\mu\big(\{x \in X : |f(x)| > t\}\big) \leq \frac{c}{t^2}$$

para todo $t > 0$. Entonces, para cualquier conjunto medible $E \subseteq X$ se cumple que

$$\int_E |f|\, d\mu \leq C_1\sqrt{\mu(E)},$$

donde $C_1 = 2\sqrt{c}$.

Prueba. Por la Fórmula de Cavalieri se tiene que

$$\begin{aligned}
\int_E |f|\, d\mu &= \int_0^{+\infty} \mu\big(\{x \in X : |f(x)| > t\} \cap E\big)\, d\mu \\
&= \int_0^{\sqrt{\frac{c}{\mu(E)}}} \mu\big(\{x \in X : |f(x)| > t\} \cap E\big)\, d\mu + \int_{\sqrt{\frac{c}{\mu(E)}}}^{+\infty} \mu\big(\{x \in X : |f(x)| > t\} \cap E\big)\, d\mu \\
&\leq \mu(E)\sqrt{\frac{c}{\mu(E)}} + \int_{\sqrt{\frac{c}{\mu(E)}}}^{+\infty} \frac{c}{t^2}\, d\mu = \sqrt{c}\sqrt{\mu(E)} + \left[\frac{-c}{t}\right]_{\sqrt{\frac{c}{\mu(E)}}}^{+\infty} \\
&= \sqrt{c}\sqrt{\mu(E)} + \frac{c}{\sqrt{\frac{c}{\mu(E)}}} = C_1\sqrt{\mu(E)}.
\end{aligned}$$

10.1.5. Integrales Dependiendo de un Parámetro

En muchas aplicaciones aparecen, de modo natural, integrales que dependen continuamente de un parámetro. Por ejemplo, uno puede encontrarse en una situación como la siguiente:

$$F(t) = \int_X f(t,x)\, dx, \qquad t \in I,$$

donde I es un cierto intervalo y $f(t,x)$ es una función a valores reales la cual es integrable en la variable x para cada $t \in I$ fijo. En este caso, con frecuencia es deseable saber si:

- ¿Existe $\lim\limits_{t \to t_0} F(t)$? En caso afirmativo, ¿es ese valor igual a $\displaystyle\int_X \lim_{t \to t_0} f(t,x)\, dx$?
- ¿Cuándo es continua F en puntos de I?
- Si F es diferenciable, ¿cuál es su derivada?

La siguiente versión "continua" del Teorema de la Convergencia Dominada, la cual da respuestas a dichas interrogantes, es de mucha utilidad. En lo que sigue, asumiremos que X es un conjunto medible, I un intervalo no-degenerado y t_0 un punto de acumulación de I.

Teorema 10.1.46 (TCD Generalizada). *Sea $\{f_t : t \in I\}$ una familia de funciones medibles definidas sobre X y suponga que:*

(a) $\lim\limits_{t \to t_0} f_t(x)$ existe para ***casi-todo*** *$x \in X$, y*

(b) existe una ***función Lebesgue integrable*** *g tal que $|f_t(x)| \leq g(x)$ para todo $t \in I$ y casi-todo $x \in X$.*

Entonces, tanto $\lim\limits_{t\to t_0} f_t$ así como cada f_t son ***Lebesgue integrables****, y*

$$\int_X \lim_{t\to t_0} f_t\, d\mu \;=\; \lim_{t\to t_0} \int_X f_t\, d\mu.$$

Prueba. La desigualdad $|f_t(x)| \leq g(x)$ para casi-todo $x \in X$ y todo $t \in I$, implica que f_t es Lebesgue integrable y, entonces, haciendo que $t \to t_0$, vemos que $|f| \leq g$ casi-siempre, donde $f = \lim_{t\to t_0} f_t$, de modo que f también es Lebesgue integrable. Para demostrar la última parte, usemos el hecho de que t_0 es un punto de acumulación de I para hallar una sucesión $(t_n)_{n=1}^{\infty}$ en $I \setminus \{t_0\}$ tal que $t_n \to t_0$. Puesto que $|f_{t_n}| \leq g$ casi-siempre y $f = \lim_{n\to\infty} f_{t_n}$ casi-siempre, se sigue del Teorema de la Convergencia Dominada que

$$\int_X f\, d\mu \;=\; \lim_{n\to\infty} \int_X f_{t_n}\, d\mu$$

y termina la prueba. ■

El resultado anterior puede ser utilizado para establecer condiciones que aseguren la continuidad o diferenciabilidad de integrales que dependan continuamente de un parámetro.

Teorema 10.1.47 (Continuidad de la Integral). *Sea $f : I \times X \to \mathbb{R}$ una función tal que para cada $t \in I$, $f(t,x)$ es una función medible de x. Suponga que*

(a) $f(t,x)$ es una ***función continua*** *de $t \in I$ para casi-todo $x \in X$, y*

(b) existe una ***función Lebesgue integrable*** *$g : X \to [0, +\infty)$ tal que $|f(t,x)| \leq g(x)$ para casi-todo $x \in X$ y para todo $t \in I$.*

Entonces

$$F(t) \;=\; \int_X f(t,x)\, d\mu$$

es una función continua de $t \in I$.

Prueba. Para cada $t \in I$, defina $f_t : X \to \mathbb{R}$ por $f_t(x) = f(t,x)$. Fijemos un $t_0 \in I$. Entonces, para cada casi-todo $x \in X$ y todo $t \in I$ se tiene que $|f_t(x)| \leq g$. Más aun, $f(t_0, x) = \lim_{t\to t_0} f_t(x)$ para casi-todo $x \in X$. Por el Teorema de la Convergencia Dominada Continua, Teorema 10.1.46, f_{t_0} y f_t son Lebesgue integrables para todo $t \in I$ y se cumple que

$$\int_X f(t_0,x)\, d\mu \;=\; \lim_{t\to t_0} \int_X f_t\, d\mu,$$

el cual es otro modo de escribir, $F(t_0) = \lim_{t\to t_0} F(t)$. Esta igualdad nos indica que F es continua en $t = t_0$ y, en consecuencia, continua sobre I. ■

Finalizamos esta sección con otro resultado que provee condiciones muy generales bajo la cual uno puede diferenciar una función bajo el signo integral. Este método, originalmente concebido por Leibniz, trata de integrales que dependen de un parámetro, tal como $\int_0^1 xe^{-tx}dx$, donde t es el parámetro y se requiere tener un método general para derivar, respecto a t, tales expresiones.

Teorema 10.1.48 (Diferenciación bajo el signo integral). *Sea $f : I \times X \to \mathbb{R}$ una función tal que para cada $t \in I$, $f(t,x)$ es* ***Lebesgue integrable****. Suponga, además, que*

(a) para casi-todo $x \in X$, la ***derivada parcial*** *$\partial_t f(t,x)$ existe para todo $t \in I$, y*

(b) existe una ***función Lebesgue integrable*** *$g : X \to [0, +\infty)$ tal que $|\partial_t f(t,x)| \leq g(x)$ para todo $t \in I$ y para casi-todo $x \in X$.*

Entonces

$$F(t) \;=\; \int_X f(t,x)\, d\mu$$

es ***diferenciable*** *en cada punto $t \in I$ y*

$$F'(t) \;=\; \int_X \frac{\partial}{\partial t} f(t,x)\, d\mu.$$

Prueba. Fijemos $t_0 \in I$ y para cada $t \in I \setminus \{t_0\}$ defina la función $f_t : X \to \mathbb{R}$ por

$$f_t(x) \;=\; \frac{f(t,x) - f(t_0,x)}{t - t_0}.$$

Observe que para que $F'(t_0)$ exista se debe cumplir

$$\begin{aligned} F'(t_0) \;&=\; \lim_{t to t_0} \frac{F(t) - F(t_0)}{t - t_0} \;=\; \lim_{t\to t_0} \frac{1}{t - t_0}\left(\int_X f(t,x)\, d\mu - \int_X f(t_0,x)\, d\mu \right) \\ &=\; \lim_{t\to t_0} \int_X f_t(x)\, d\mu, \end{aligned}$$

de modo que todo lo que tenemos que demostrar es que

$$\lim_{t\to t_0} \int_X f_t(x)\, d\mu \quad \text{existe y es igual a} \quad \int_X \frac{\partial}{\partial t} f(t,x)\, d\mu.$$

Para ver esto, observe que por (a) se tiene que $\partial_t f(t_0,x) = \lim_{t\to t_0} f_t(x)$ para casi-todo $x \in X$ y así, gracias al Teorema del Valor Medio, para cada $t \in I$, existe un número θ_t entre t y t_0 tal que

$$|f_t(x)| \;=\; \left| \frac{f(t,x) - f(t_0,x)}{t - t_0} \right| \;=\; |\partial_t f(\theta_t, x)|.$$

Por otro lado, por (b) se sigue que $|f_t(x)| \leq g$ para casi-todo $x \in X$ y todo $t \in I$ y, por consiguiente, gracias al Teorema de la Convergencia Dominada Continua, $\partial_t f(t_0,x)$ y cada f_t son Lebesgue integrables, y

$$\lim_{t\to t_0} \int_X f_t(x)\, d\mu \;=\; \int_X \lim_{t\to t_0} f_t(x)\, d\mu \;=\; \int_X \frac{\partial}{\partial t} f(t,x)\, d\mu.$$

La prueba es completa. ■

10.1.6. La Integral de Lebesgue sin Medida

La década de los años 20 del siglo XX se vio penetrada por un enfoque diferente del concepto de integral sustentado, hasta ese momento, en la noción de función de conjuntos (teoría de la medida) en el espacio $\mathbb{R}^n$ Con el surgimiento del Análisis Funcional, se requería de un concepto de integral que no dependiese de la Teoría de la Medida pero sí de las propiedades de los elementos y subconjuntos del espacio considerado. Esta nueva tendencia iba a ser liderizada por el matemático P. J. Daniell, quien, después de estudiar en profundidad la integral de Lebesgue, pensó la nueva integral como un funcional lineal sobre ciertos espacios de funciones más generales que los $\mathbb{R}^n$. Lo maravilloso de ese enfoque es que la integral de Lebesgue en $\mathbb{R}$ se obtiene como un caso particular de la integral Daniell inducida por la integral de Riemann sobre el espacio $C_c(\mathbb{R})$.

Las dos definiciones de integrales que ofrecemos en lo que sigue son equivalentes a la de Lebesgue con una ventaja: ninguna de ellas usa la Teoría de la Medida es su construcción. En el capítulo final de estas notas presentaremos otra integral, muy parecida a la integral de Riemann, que también es equivalente a la de Lebesgue y que tampoco se apoya en la Teoría de la Medida para su edificación.

(I) La Integral de Daniell. Percy John Daniell (1889-1946) nace en Valparaiso, Chile, en el año de 1889. Seis años después sus padres regresan a Inglaterra donde Daniell realiza sus estudios. En el año 1918 Daniell publica un artículo titulado: *A General form of Integral* en Annals of Mathematics, 10(1918), 279-294, donde desarrolla una integral equivalente a la integral de Lebesgue sin usar la Teoría de la Medida. La integral de Daniell es particularmente útil en Análisis Funcional. La idea de Daniell comienza observando que $\mathcal{L}_1(\mathbb{R},\mu)$ es un retículo vectorial y que si definimos la aplicación $\Phi:\mathcal{L}_1(\mathbb{R},\mu)\to\mathbb{R}$ por

$$\Phi(f) \;=\; \int_{\mathbb{R}} f\,d\mu, \quad f\in\mathcal{L}_1(\mathbb{R},\mu)$$

entonces

(a) Φ es un funcional lineal,

(b) Φ es no-negativo, es decir, $\Phi(f)\geq 0$ si $f\geq 0$, y

(c) $\Phi(f_n)\to 0$ si $(f_n)_{n=1}^{\infty}$ es una sucesión decreciente en $\mathcal{L}_1(\mathbb{R},\mu)$ convergiendo puntualmente a cero.

Pues bien, en la Teoría de la Integral de Daniell se procede a la inversa: se parte de un retículo vectorial y un funcional lineal definido sobre él al cual se le exige que cumpla la tres condiciones anteriores para luego obtener la integral. Desarrollemos, brevemente, esas ideas.

Definición 10.1.49. *Una familia de funciones $\mathcal{H}$ definidas sobre un espacio métrico (X,d) y con valores en $\mathbb{R}$ se dice que es un **espacio de Stone**, o un **retículo vectorial real**, si ella satisface los siguientes requerimientos:*

(1) *$\mathcal{H}$ es un espacio vectorial sobre $\mathbb{R}$.*

(2) *Si $h\in\mathcal{H}$, entonces $|h|\in\mathcal{H}$.*

Observe que si $f,g\in\mathcal{H}$, entonces $f\vee g$ y $f\wedge g$ ambas pertenecen a $\mathcal{H}$, donde

$$f\vee g \;=\; \frac{1}{2}\big(f+g+|f-g|\big) \qquad \text{y} \qquad f\wedge g \;=\; \frac{1}{2}\big(f+g-|f-g|\big).$$

Definición 10.1.50. *Sea $\mathcal{H}$ un espacio de Stone. Una* ***integral elemental de Daniel*** *es una función $\mathfrak{I} : \mathcal{H} \to \mathbb{R}$ verificando las tres condiciones siguientes:*

(a) $\mathfrak{I}$ es ***lineal****, es decir, si $f, g \in \mathcal{H}$ y $\alpha, \beta \in \mathbb{R}$, entonces $\mathfrak{I}(\alpha f + \beta g) = \alpha \mathfrak{I}(f) + \beta \mathfrak{I}(g)$,*

(b) $\mathfrak{I}$ es ***no-negativa****, esto es, si $f \in \mathcal{H}$ y $f \geq 0$, entonces $\mathfrak{I}(f) \geq 0$.*

(c) $\mathfrak{I}$ es ***decrecientemente continua*** *en el origen, es decir, si $(f_n)_{n=1}^{\infty}$ es una sucesión decreciente y $\lim\limits_{n\to\infty} f_n = 0$, entonces*

$$\lim_{n\to\infty} \mathfrak{I}(f_n) \ = \ 0.$$

Se sigue de (a) y (b) de la definición anterior que si $f, g \in \mathcal{H}$, entonces

$$f \ \leq \ g \quad \Rightarrow \quad \mathfrak{I}(f) \ \leq \ \mathfrak{I}(g).$$

En particular, si $f \in \mathcal{H}$, entonces

$$\mathfrak{I}(f) \ \leq \ \mathfrak{I}(f^+) \ \leq \ \mathfrak{I}(|f|), \qquad \text{y} \qquad |\mathfrak{I}(f)| \leq \mathfrak{I}(|f|).$$

El símbolo $f_n \nearrow f$ significa que la sucesión $(f_n)_{n=1}^{\infty}$ es creciente y converge puntualmente a f. Similar significado le asignaremos al símbolo $f \searrow f$.

Ejemplo 10.1.2. *Tomando $\mathcal{H} = C([a,b])$ y definiendo la aplicación $\mathfrak{I} : C([a,b]) \to \mathbb{R}$ por*

$$\mathfrak{I}(f) \ = \ (\mathrm{R})\int_a^b f\,dt, \quad f \in C([a,b])$$

entonces, $\mathfrak{I}$ es una integral elemental de Daniell.

Prueba. En efecto, ya hemos visto que $\mathfrak{I}$ es lineal y positiva. El hecho de que $\mathfrak{I}$ es decrecientemente continua en el origen es consecuencia del Teorema de Dini, página 389 y el Teorema de la Convergencia Uniforme para la integral de Riemann, Teorema 8.3.25. ■

En general, si $\mathcal{H} = C_c(\mathbb{R})$ y si $f \in C_c(\mathbb{R})$, entonces por la Observación 8.3.2, página 418, sabemos que f es Riemann integrable y, por consiguiente, podemos definir $\mathfrak{I} : C_c(\mathbb{R}) \to \mathbb{R}$ por

$$\mathfrak{I}(f) \ = \ (\mathrm{R})\int_a^b f\,dt,$$

donde $[a,b]$ es un intervalo compacto que depende de f y fuera del cual f se anula. Claramente, $\mathfrak{I}$ es lineal. Veamos que $\mathfrak{I}$ es decrecientemente continua en el origen. En efecto, sea $(f_n)_{n=1}^{\infty}$ una sucesión en $C_c(\mathbb{R})$ decreciente y convergiendo puntualmente a cero. Puesto que $f_n \leq f_1$ se tiene que $\mathrm{sop}(f_n) \subseteq \mathrm{sop}(f_1)$ para todo $n \in \mathbb{N}$. Sea $[a,b]$ un intervalo compacto tal que $\mathrm{sop}(f_1) \subseteq [a,b]$. Entonces $\mathrm{sop}(f_n) \subseteq [a,b]$ para todo $n \geq 1$ y se sigue del Teorema de Dini que $f_n \to 0$ uniformemente sobre $[a,b]$. Un llamado al Teorema de la Convergencia Uniforme, página 438, nos indica que $\mathfrak{I}(f_n) \to 0$.

El plan, para obtener la integral de Daniell, es extender la integral elemental de Daniell a "espacios más grandes". Sea entonces

$$\mathcal{H}^+ \ = \ \Big\{ f \ : \ f_n \nearrow f \text{ para alguna sucesión creciente } (f_n)_{n=1}^{\infty} \text{ en } \mathcal{H} \Big\}.$$

Observe que $\mathcal{H} \subseteq \mathcal{H}^+$ y que las funciones en $\mathcal{H}^+$ pueden tomar el valor $+\infty$ pero nunca el valor $-\infty$. Más aun, $\mathcal{H}^+$ satisface las siguientes propiedades:

(a) Si $f, g \in \mathcal{H}^+$, entonces $f + g \in \mathcal{H}^+$.

(b) Si $f \in \mathcal{H}^+$ y $\alpha \geq 0$, entonces $\alpha f \in \mathcal{H}^+$.

(c) Si $f, g \in \mathcal{H}^+$, entonces $f \vee g$, $f \wedge g \in \mathcal{H}^+$.

Suponga ahora que $f \in \mathcal{H}^+$ y seleccione una sucesión creciente $(f_n)_{n=1}^\infty$ en $\mathcal{H}$ de modo que $f_n \to f$ puntualmente. Puesto que la sucesión de números reales $(\mathfrak{I}(f_n))_{n=1}^\infty$ es creciente en $\mathbb{R}$, resulta que ella converge en $\overline{\mathbb{R}}$. Este análisis nos permite extender la aplicación $\mathfrak{I}$ al conjunto $\mathcal{H}^+$ y que seguiremos denotando del mismo modo, declarando que

$$\mathfrak{I}(f) \;=\; \lim_{n\to\infty} \mathfrak{I}(f_n)$$

donde $f \in \mathcal{H}^+$ y $(f_n)_{n=1}^\infty$ es cualquier sucesión creciente en $\mathcal{H}$ convergiendo puntualmente a f. Queda por verificar que el número $\mathfrak{I}(f)$ no depende de la elección de la sucesión creciente $(f_n)_{n=1}^\infty$ convergiendo puntualmente a f. Este hecho será demostrado una vez que probemos el siguiente resultado.

Teorema 10.1.51 (Stone). *Sean $f, g \in \mathcal{H}^+$ con $f \leq g$. Si $(f_n)_{n=1}^\infty$ y $(g_n)_{n=1}^\infty$ son sucesiones crecientes en $\mathcal{H}$ que convergen puntualmente a f y g respectivamente, entonces*

$$\lim_{n\to\infty} \mathfrak{I}(f_n) \;\leq\; \lim_{n\to\infty} \mathfrak{I}(g_n). \tag{1}$$

Prueba. Suponga que $f, g \in \mathcal{H}^+$ con $f \leq g$ y sean $(f_n)_{n=1}^\infty$ y $(g_n)_{n=1}^\infty$ sucesiones crecientes en $\mathcal{H}$ que convergen puntualmente a f y g respectivamente. Para cada $k \in \mathbb{N}$, considere la función $h_n = g_n \wedge f_k$ para cada $n \in \mathbb{N}$. Resulta que la sucesión $(h_n)_{n=1}^\infty$ es creciente y se cumple que

$$\lim_{n\to\infty} h_n \;=\; g \wedge f_k \;=\; f_k,$$

es decir, $h_n \nearrow f_k$ y, por consiguiente, $(f_k - h_n) \searrow 0$ cuando $n \to \infty$. Se sigue entonces de (c) de la Definición 10.1.50 que $\lim_{n\to\infty} \mathfrak{I}(h_n) = \mathfrak{I}(f_k)$. Por otro lado, como $g_n \geq h_n$, obtenemos que $\mathfrak{I}(g_n) \geq \mathfrak{I}(h_n)$ y, en consecuencia,

$$\lim_{n\to\infty} \mathfrak{I}(g_n) \;\geq\; \mathfrak{I}(f_k) \quad \text{para todo } k \in \mathbb{N}.$$

Por esto, $\lim\limits_{n\to\infty} \mathfrak{I}(f_n) \leq \lim\limits_{n\to\infty} \mathfrak{I}(g_n)$. ■

Veamos ahora que $\mathfrak{I}(f)$ está unívocamente definida. En efecto, sea $(g_n)_{n=1}^\infty$ otra sucesión creciente en $\mathcal{H}$ convergiendo puntualmente a f. Tomando $g = f$ y aplicando el resultado anterior se concluye que

$$\mathfrak{I}(f) \;=\; \lim_{n\to\infty} \mathfrak{I}(f_n) \;\leq\; \lim_{n\to\infty} \mathfrak{I}(g_n) \;=\; \mathfrak{I}(f).$$

Nótese que la aplicación $\mathfrak{I} : \mathcal{H}^+ \to \overline{\mathbb{R}}$ satisface las siguientes propiedades:

(a) $\mathfrak{I}(f + g) = \mathfrak{I}(f) + \mathfrak{I}(g)$,

(b) $\mathfrak{I}(\alpha f) = \alpha \mathfrak{I}(f)$ si $\alpha \geq 0$,

(c) $\mathfrak{I}(f) \leq \mathfrak{I}(g)$ si $f \leq g$, y

(d) si $f_n \in \mathcal{H}^+$ y $f_n \nearrow f$, entonces $f \in \mathcal{H}^+$ y $\mathfrak{I}(f_n) \to \mathfrak{I}(f)$.

Para completar la construcción de la integral de Daniell, vamos a extender la aplicación $\mathfrak{I}$ a una clase mucho más grande. Suponga que $f : X \to \mathbb{R}$ es una función arbitraria. Si existe $h \in \mathcal{H}^+$ tal que $f \leq h$, definimos la **integral superior** de f por

$$\bar{\mathrm{I}}(f) \;=\; \text{ínf}\left\{\mathfrak{I}(h) \,:\, h \in \mathcal{H}^+,\, f \leq h\right\}.$$

Pongamos $\mathcal{H}^- = -\mathcal{H}^+$ y suponga ahora que existe una función $g \in \mathcal{H}^-$ tal que $g \leq f$. En este caso, definimos **integral inferior** de f por

$$\underline{\mathrm{I}}(f) \;=\; \sup\left\{\mathfrak{I}(g) \,:\, g \in \mathcal{H}^-,\, g \leq f\right\}.$$

Se demuestra que si las integrales superior e inferior de f existen y ambas son finitas, entonces $\underline{\mathrm{I}}(f) \leq \bar{\mathrm{I}}(f)$. Una función $f : X \to \mathbb{R}$ se llama **Daniell integrable** si $\bar{\mathrm{I}}(f)$ y $\underline{\mathrm{I}}(f)$ están definidas, son ambas finitas e iguales. En este caso, la **integral de Daniell** de f se define como

$$\mathfrak{D}(f) \;=\; \bar{\mathrm{I}}(f) \;=\; \underline{\mathrm{I}}(f).$$

Escribiremos $\mathcal{L}_1(\mathcal{H})$ para denotar el conjunto de todas las funciones $f : X \to \mathbb{R}$ que son Daniell integrables. Se demuestra que $\mathcal{L}_1(\mathcal{H})$ contiene a al conjunto $\mathcal{H}^+ \cup \mathcal{H}^-$ y que también puede contener elementos que no están en esa unión. Las propiedades algebraicas de la clase $\mathcal{L}_1(\mathcal{H})$ y la aplicación $\mathfrak{D}$ se muestran a continuación:

Sean $f, g \in \mathcal{L}_1(\mathcal{H})$ y $\alpha \in \mathbb{R}$. Entonces:

(a) $f + g \in \mathcal{L}_1(\mathcal{H})$ y $\mathfrak{D}(f+g) = \mathfrak{D}(f) + \mathfrak{D}(g)$.

(b) $\alpha f \in \mathcal{L}_1(\mathcal{H})$ y $\mathfrak{D}(\alpha f) = \alpha\mathfrak{D}(f)$.

(c) $f \vee g,\, f \wedge g \in \mathcal{L}_1(\mathcal{H})$.

(d) Si $f \leq g$, entonces $\mathfrak{D}(f) \leq \mathfrak{D}(g)$.

(e) $|f| \in \mathcal{L}_1(\mathcal{H})$ y $|\mathfrak{D}(f)| \leq \mathfrak{D}(|f|)$.

Veamos, finalmente, que si $\mathcal{H} = C_c(\mathbb{R})$, entonces las clases $\mathcal{L}_1(\mathcal{H})$ y $\mathcal{L}_1(\mathbb{R}, \mu)$ coinciden con idénticas integrales. Para lograr tal propósito, tenemos que considerar varias hechos importantes pero a vuelo de pájaro, es decir, que mencionaremos brevemente pero no demostraremos.

(1^o) Un conjunto $E \subseteq \mathbb{R}$ se llama D-**nulo** si $\chi_E \in \mathcal{L}_1(\mathcal{H})$ y $\mathfrak{D}(\chi_E) = 0$. Denote por $\mathcal{N}_0$ la familia de todos los conjuntos D-nulos.

(2^o) Una función $f \in C_c(\mathbb{R})$ se llama D-**nula** si $|f| \in \mathcal{L}_1(\mathcal{H})$ y $\mathfrak{I}(|f|) = 0$.

(3^o) Dos funciones $f, g \in C_c(\mathbb{R})$ son **iguales** D-**casi siempre**, si $\{x \in \mathbb{R} : f(x) \neq g(x)\}$ es D-nulo, o equivalentemente, $f - g$ es una función D-nula. Si $f, g \in \mathcal{L}_1(\mathcal{H})$, escribiremos $\sim_D$ si $f = g$ D-casi siempre. Resulta que $\sim_D$ es una relación de equivalencia sobre $\mathcal{L}_1(\mathcal{H})$. Observe que si $f \sim_D g$ y $f \in \mathcal{L}_1(\mathcal{H})$, entonces $g \in \mathcal{L}_1(\mathcal{H})$ y se cumple que $\mathfrak{I}(f) = \mathfrak{I}(g)$. Identificando funciones iguales D-casi siempre, escribiremos cada clase de equivalencia $[f]$ como f. Hecha esta suposición se tiene que si $f \in \mathcal{L}_1(\mathcal{H})$, entonces

$$\|\, f \,\| \;=\; \mathfrak{I}(|f|),$$

define una norma sobre $\mathcal{L}_1(\mathcal{H})/\sim_D$ la cual lo convierte en un espacio de Banach. Al espacio de Banach $\mathcal{L}_1(\mathcal{H})/\sim_D$ lo denotaremos por $\mathrm{DL}_1(\mathcal{H})$.

(4^o) Una función $f \in C_c(\mathbb{R})$ se llama D-**medible** si la función $F = g \vee (f \wedge h) \in \mathcal{L}_1(\mathcal{H})$ para cualquier par de funciones $g, h \in \mathcal{L}_1(\mathcal{H})$ tales que $g \leq 0 \leq h$. Observe que

$$F(x) = \begin{cases} f(x) & \text{si } g(x) \leq f(x) \leq h(x) \\ h(x) & \text{si } h(x) < f(x) \\ g(x) & \text{si } f(x) < g(x). \end{cases}$$

Denote por $\mathcal{M}_\mathrm{D}$ el conjunto de todas las funciones D-medibles.

(5^o) Un conjunto $E \subseteq \mathbb{R}$ se llama D-**medible** si $\chi_E \in \mathcal{M}_\mathrm{D}$. Si $\mathfrak{M}_D$ denota la familia de todos los conjuntos D-medibles, entonces $\mathfrak{M}_D$ es una σ-álgebra y la función de conjuntos $\mu_\mathrm{D} : \mathfrak{M}_D \to [0, +\infty]$ definida por

$$\mu_\mathrm{D}(E) = \begin{cases} \mathfrak{D}(\chi_E) & \text{si } \chi_E \in \mathcal{L}_1(\mathcal{H}), \\ +\infty & \text{si } \chi_E \in \mathcal{M}_\mathrm{D} \setminus \mathcal{L}_1(\mathcal{H}) \end{cases}$$

es una media completa sobre $\mathfrak{M}_D$.

(6^o) $\mathfrak{M}_D = \mathfrak{M}_\mu(\mathbb{R})$, $\mu_\mathrm{D} = \mu$ y $\mathcal{L}_1(\mathcal{H}) = \mathcal{L}_1(\mathbb{R}, \mu)$ con

$$\mathfrak{D}(f) = (\mathrm{L})\int_{\mathbb{R}} f \, d\mu$$

para cualquier $f \in \mathcal{L}_1(\mathcal{H})$. Los detalles de todos estos hechos, y mucho más, lo puede ver el lector en [129], al que se le pueden añadir, por ejemplo, [121, 138], etc.

(II) La Integral de Mikusińki. La siguiente presentación, también a vuelo de pájaro, es una forma alternativa de introducir la integral de Lebesgue al estilo Mikusińki sin utilizar la Teoría de la Medida (véase, por ejemplo, [99]).

Una función $\psi : \mathbb{R} \to \mathbb{R}$ se llama **elemental** si existe un intervalo de longitud finita $I = [a, b)$ tal que $\psi = \chi_{[a,b)}$. Al intervalo $[a, b)$ se le llama el **soporte** de ψ. En este caso, la integral elemental de ψ se define como

$$\int \psi = \ell([a, b)) = b - a.$$

Extendamos esta noción elemental de integral del modo siguiente. Una función $\varphi : \mathbb{R} \to \mathbb{R}$ se dice que es una **función en escalera elemental** si existen escalares $a_1, \ldots, a_n$ y funciones elementales $\psi_1, \ldots, \psi_n$ tales que

$$\varphi(x) = \sum_{i=1}^{n} a_i \psi_i(x) \quad \text{para todo } x \in \mathbb{R}.$$

La integral de φ, que seguiremos denotando como $\int \varphi$, se define por:

$$\int \varphi = a_1 \int \psi_1 + \cdots + a_n \int \psi_n.$$

Definición 10.1.52. *Una función $f : \mathbb{R} \to \mathbb{R}$ se llama **Mikusiński integrable** sobre $\mathbb{R}$ si existe una sucesión $(\varphi_n)_{n=1}^{\infty}$ de funciones en escaleras elementales definidas sobre $\mathbb{R}$ tal que:*

(a) $\displaystyle\sum_{n=1}^{\infty} \int |\varphi_n| < +\infty$ *y*

(b) $\displaystyle f(x) = \sum_{n=1}^{\infty} \varphi_n(x)$ *para todo $x \in \mathbb{R}$ para el cual* $\displaystyle\sum_{n=1}^{\infty} |\varphi_n(x)| < +\infty.$

Si f es Mikusiński integrable sobre $\mathbb{R}$, definimos la **integral de Mikusiński** de f como

$$\int f = \sum_{n=1}^{\infty} \varphi_n,$$

donde $(\varphi_n)_{n=1}^{\infty}$ es cualquier sucesión de funciones en escaleras elementales definidas sobre $\mathbb{R}$ usadas para representar a f según la definición anterior. Es importante destacar que la integral de Mikusiński de f no depende de la elección particular de la sucesión de funciones en escaleras elementales usadas para representarla. Denote por $\mathcal{M}\text{ik}(\mathbb{R})$ el conjunto de todas las funciones $f : \mathbb{R} \to \overline{\mathbb{R}}$ que son Mikusiński integrables. Se comprueba que $\mathcal{M}\text{ik}(\mathbb{R})$ es un espacio vectorial sobre $\mathbb{R}$. Además,

(i) $f \in \mathcal{M}\text{ik}(\mathbb{R})$ implica que $|f| \in \mathcal{M}\text{ik}(\mathbb{R})$.

(ii) $f \in \mathcal{M}\text{ik}(\mathbb{R})$ si, y sólo si, $f \in \mathcal{L}_1(\mathbb{R}, \mu)$ y

$$\int f = (\mathrm{L})\int_{\mathbb{R}} f\, d\mu.$$

10.2. El Espacio $\mathrm{L}_1(X, \mu)$

La teoría de integración hasta ahora desarrollada nos va permitir definir los espacios de funciones $\mathrm{L}_p(X, \mu)$ para cualquier $p \in [1, +\infty]$ los cuales, como veremos en el transcurso de estas notas, poseen notables propiedades y que, por supuesto, constituyen una pieza fundamental y de enorme importancia tanto en el Análisis Real, así como en el Análisis Funcional. El enfoque que seguiremos para presentar los espacios $\mathrm{L}_p(\mu)$ es el siguiente: en primer lugar, definiremos el espacio $\mathrm{L}_1(X, \mu)$ y estudiaremos algunas propiedades que destacan por su versatilidad y un enorme caudal de aplicaciones. Seguidamente analizaremos los espacios $\mathrm{L}_p(X, \mu)$, donde $p \in (1, +\infty)$ y, posteriormente, el espacio correspondiente a $p = +\infty$, esto es, $\mathrm{L}_\infty(X, \mu)$ el cual demanda un tratamiento muy especial. ¿Qué ocurre con los números $p \in (0, 1)$? La razón fundamental para no considerar los espacios $\mathrm{L}_p(X, \mu)$ para tales números en esta exposición es que, si bien cada $\mathrm{L}_p(X, \mu)$ es un espacio vectorial, $\|\cdot\|_p$ no es una norma en dicho espacio y, por lo tanto, tales conjuntos no son de interés en estas notas. Sin embargo, es necesario decirlo, ellos son importantes en la Teoría de los Espacios Vectoriales Topológicos y sirven como una fuente de contraejemplos.

En lo que sigue, y sólo para facilitar un poco las cosas, asumiremos que nuestro espacio medible X es todo $\mathbb{R}$. Ya hemos visto cómo se construyó el espacio vectorial $\mathcal{L}_1(\mu) = \mathcal{L}_1(\mathbb{R}, \mathfrak{M}_\mu(\mathbb{R}), \mu)$. Lo queremos ahora es introducir una norma natural sobre dicho espacio que lo convierta en un espacio de Banach. Para que tal cosa sea posible, debemos introducir una relación de equivalencia sobre $\mathcal{L}_1(\mathbb{R}, \mathfrak{M}_\mu(\mathbb{R}), \mu)$ identificando las funciones medibles que son iguales casi-siempre y pasar luego al espacio cociente al que denotaremos por $\mathrm{L}_1(\mu) := \mathrm{L}_1(\mathbb{R}, \mathfrak{M}_\mu(\mathbb{R}), \mu)$. Finalmente se podrá introducir una norma en dicho espacio que lo convertirá en un espacio de Banach.

Comencemos con el siguiente resultado, el cual es fundamental para construir la "norma" que deseamos sobre $\mathrm{L}_1(\mu)$.

Lema 10.2.1. *Sean* $f, g \in \mathcal{L}_1(\mu)$. *Son equivalentes:*

(1) $f = g \quad \mu -$ c.s.

(2) $\displaystyle\int_{\mathbb{R}} |f - g|\, d\mu = 0$.

(3) $\displaystyle\int_E f\, d\mu = \int_E g\, d\mu$ *para cualquier* $E \in \mathfrak{M}_\mu(\mathbb{R})$.

Prueba. Las equivalencias (1) y (2) son similares a las del Corolario 10.1.15 y se omiten. Para demostrar que (2) $\Rightarrow$ (3), observe que para cualquier conjunto medible E,

$$\left| \int_E f\, d\mu - \int_E g\, d\mu \right| \le \int_E |f - g|\, d\mu \le \int_{\mathbb{R}} |f - g|\, d\mu = 0,$$

de donde sigue (3). Finalmente, si definimos $E = \{x \in \mathbb{R} : f - g \geq 0\}$, resulta de (3) que

$$\begin{aligned} \int_{\mathbb{R}} |f - g|\, d\mu &= \int_E |f - g|\, d\mu + \int_{\mathbb{R}\setminus E} |f - g|\, d\mu \\ &= \int_E (f - g)\, d\mu + \int_{\mathbb{R}\setminus E} (g - f)\, d\mu = 0, \end{aligned}$$

lo cual prueba (2). La prueba es completa. ■

Definamos la aplicación $\|\cdot\|_1 : \mathcal{L}_1(\mu) \to [0, +\infty)$ por

$$\|f\|_1 = \int_{\mathbb{R}} |f|\, d\mu. \tag{1}$$

Se sigue del Corolario 10.1.28 que la afirmación "$\|f\|_1 = 0 \;\Rightarrow\; f = 0$" no es, en general, válida: *sólo podemos asegurar que* $f = 0$ *casi-siempre*. Este hecho nos convence que la igualdad (1) no produce una norma sobre $\mathcal{L}_1(\mu)$. Para obtener, realmente, una norma debemos *identificar a las funciones en* $\mathcal{L}_1(\mu)$ *que son iguales casi-siempre*; es decir, definimos la siguiente relación de equivalencia $\sim$ entre los elementos de $\mathcal{L}_1(\mu)$: si $f, g \in \mathcal{L}_1(\mu)$,

$$f \sim g \quad \Leftrightarrow \quad f = g \quad \mu - \text{c.s.}$$

Resulta que $\sim$ es una **relación de equivalencia** y, en consecuencia, el espacio cociente

$$\mathcal{L}_1(\mu)/\!\sim\; = \Big\{ [f] : f \in \mathcal{L}_1(\mu) \Big\},$$

donde

$$[f] = \{ h \in \mathcal{L}_1(\mu) : f \sim h \}$$

es un espacio vectorial con las operaciones usuales, es decir, $[f+g]=[f]+[g]$ y $[a\cdot f]=a\cdot[f]$ cualesquiera sean $f,g\in\mathcal{L}_1(\mu)$ y $a\in\mathbb{R}$. Más aun, $[f]\leq[g]$ siempre que $f\leq g$ $\mu-$c.s.

Denotaremos, en lo sucesivo, al espacio vectorial $\mathcal{L}_1(\mu)/\sim$ por $\mathrm{L}_1(\mu)$. Observe que, gracias al Lema 10.2.1,

$$\int_{\mathbb{R}} f\,d\mu \;=\; \int_{\mathbb{R}} h\,d\mu, \quad \text{para cualquier} \quad h\in[f], \tag{1}$$

de modo que todos los elementos pertenecientes a $[f]$ poseen la misma integral y, por lo tanto, sólo es necesario elegir un único elemento de $[f]$ para computar su integral. Este argumento permite utilizar la integral de Lebesgue para obtener una "*integral*" para todos los elementos de $\mathrm{L}_1(\mu)$, que seguiremos llamando **integral de Lebesgue**, y que se define como

$$\int_{\mathbb{R}}[f]\,d\mu \;:=\; \int_{\mathbb{R}} f\,d\mu.$$

Nota Adicional 10.2.8 De la igualdad

$$\int_{[a,b]} f\,d\mu \;=\; \int_{[a,b]} h\,d\mu, \quad \text{para cualquier} \quad h\in[f],$$

es probable que el lector acaricie la idea de que tal vez no sea necesaria la intervención de la integral de Lebesgue para definir $\mathrm{L}_1([a,b])$, ya que si logramos encontrar, en cada clase de equivalencia $[f]$, una función Riemann integrable g, entonces tendríamos que

$$\int_{[a,b]} f\,d\mu \;=\; \int_{[a,b]} g\,d\mu \;=\; (\mathcal{R})\int_a^b g\,dx$$

y así, la integral de Riemann sería suficiente para definir $\mathrm{L}_1([a,b])$. Sin embargo, el siguiente hecho nos revela que tal sutileza no es completamente viable: existe al menos una función $f\in\mathcal{L}_1([0,1])$ tal que $[f]$ no contiene a ninguna función que sea Riemann integrable. En efecto, sea G un conjunto abierto conteniendo a todos los racionales de $[0,1]$ tal que $[0,1]\nsubseteq G$ con $\mu(G)<1$ (véase el Teorema 6.4.2, página 304, para ver la construcción de un tal conjunto) y defina $f=\chi_G$. Claramente, $f\in\mathcal{L}_1([0,1])$. Suponga que $g\in[f]$ y veamos que g no puede ser Riemann integrable. Puesto que $g\sim f$, entonces $\mu(N)=0$, donde $N=\{x\in[0,1]: f(x)\neq g(x)\}$ y, por lo tanto,

$$g(x) \;=\; \begin{cases} 1 & \text{si} \quad x\in G\setminus N, \\ 0 & \text{si} \quad x\notin G\cup N. \end{cases}$$

Ahora bien, como G es un conjunto abierto conteniendo a los racionales de $[0,1]$, resulta que $G\setminus N$ es denso en $[0,1]$ y, en consecuencia, $g=1$ sobre un conjunto denso. Por esto, g es discontinua en cualquier punto x tal que $g(x)\neq 1$. En particular, g es discontinua fuera del conjunto $G\cup N$ el cual posee medida positiva. Se sigue del Teorema de Vitali-Lebesgue que g no es Riemann integrable.

Establecido lo anterior, observe que la aplicación $\|\cdot\|_1:\mathrm{L}_1(\mu)\to[0,+\infty)$ definida por

$$\|\,[f]\,\|_1 \;=\; \int_{\mathbb{R}}|f|\,d\mu$$

ahora *sí es una norma sobre* $L_1(\mu)$, es decir,

(1) $\|[f]\|_1 \geq 0$ para toda $[f] \in L_1(\mu)$.

(2) $\|[f]\|_1 = 0$ si, y sólo si, $[f] = [0]$.

(3) $\|a \cdot [f]\|_1 = a \cdot \|[f]\|_1$ para toda $[f] \in L_1(\mu)$ y todo $a \in \mathbb{R}$ y

(4) $\|[f] + [g]\|_1 \leq \|[f]\|_1 + \|[g]\|_1$ para toda $[f], [g] \in L_1(\mu)$.

Debido a esto, y sólo por razones prácticas, ignoraremos la distinción entre una función y su clase de equivalencia, de modo que a los elementos de $L_1(\mu)$ los trataremos como funciones normales y corrientes, en otras palabras, *cuando escribamos "$f \in L_1(\mu)$", lo que realmente estamos diciendo es: "sea f un representante de alguna clase de equivalencia en $L_1(\mu)$, la cual podemos suponer que es $[f]$".* Este acuerdo lo mantendremos a partir de este momento. Si bien es cierto que este convenio nos ahorra la molestia de sobrecargar la notación, debemos estar alerta y ser cuidadoso cuando se interpretan ciertas declaraciones o afirmaciones en $L_1(\mu)$.

Una consecuencia del Teorema de la Convergencia Monótona combinada con el hecho de que μ es invariante por traslación es el siguiente resultado, el cual afirma que la integral de Lebesgue también es invariante por traslación, es decir:

Teorema 10.2.2. *Sea $f \in L_1(\mu)$. Para cada $t \in \mathbb{R}$, la función $\varphi_t : \mathbb{R} \to \mathbb{R}$ definida por $\varphi_t(x) = f(x-t)$ es **Lebesgue integrable** y*

$$\int_{\mathbb{R}} f(x-t)\,d\mu = \int_{\mathbb{R}} f(x)\,d\mu.$$

Prueba. Fijemos $t \in \mathbb{R}$ y suponga que $f(x) = \chi_E(x)$ para algún conjunto medible $E \subseteq \mathbb{R}$. Entonces, como $f(x-t) = \chi_{E+t}(x)$ y $\mu(E+t) = \mu(E)$ resulta que

$$\int_{\mathbb{R}} f(x)\,d\mu = \mu(E) = \mu(E+t) = \int_{\mathbb{R}} f(x-t)\,d\mu.$$

La igualdad anterior también es válida si f es cualquier función simple no negativa. Si f es no-negativa, entonces el Teorema de Aproximación por Funciones Simples nos garantiza la existencia de una sucesión creciente $(f_n)_{n=1}^{\infty}$ de funciones simples no-negativas que converge puntualmente a f. Se sigue del Teorema de la Convergencia Monótona que

$$\begin{aligned}\int_{\mathbb{R}} f(x)\,d\mu &= \lim_{n\to\infty} \int_{\mathbb{R}} f_n(x)\,d\mu \\ &= \lim_{n\to\infty} \int_{\mathbb{R}} f_n(x-t)\,d\mu \\ &= \int_{\mathbb{R}} f(x-t)\,d\mu.\end{aligned}$$

Finalmente, escribiendo $f = f^+ - f^-$ y utilizando lo anterior se termina la prueba. ■

De modo similar, si c es cualquier constante no-nula, entonces la aplicación $x \to f(cx)$ es Lebesgue-integrable y

$$\int_{\mathbb{R}} f(x)\,d\mu = \frac{1}{c}\int_{\mathbb{R}} f(cx)\,d\mu.$$

Los detalles se dejan a cargo del lector.

Definición 10.2.3. *Una sucesión* $(f_n)_{n=1}^{\infty}$ *en* $\mathrm{L}_1(\mu)$ *se dice que* ***converge en la norma*** *de* $\mathrm{L}_1(\mu)$ *si existe una función* $f \in \mathrm{L}_1(\mu)$ *tal que*

$$\lim_{n\to\infty} \| f_n - f \|_1 \;=\; 0.$$

Observe que $\lim\limits_{n\to\infty} \| f_n - f \|_1 = 0$, significa que

$$\lim_{n\to\infty} \int_{\mathbb{R}} |f_n - f| \, d\mu \;=\; 0.$$

Lema 10.2.4. *Sean* $f, f_n \in \mathrm{L}_1(\mu)$ *para* $n \geq 1$. *Si* $\lim\limits_{n\to\infty} \| f_n - f \|_1 = 0$, *entonces*

$$\lim_{n\to\infty} \int_E f_n \, d\mu \;=\; \int_E f \, d\mu \qquad \text{y} \qquad \lim_{n\to\infty} \int_E |f_n| \, d\mu \;=\; \int_E |f| \, d\mu$$

para todo $E \in \mathfrak{M}_\mu(\mathbb{R})$.

Prueba. Fijemos un conjunto medible arbitrario $E \subseteq \mathbb{R}$. Entonces, para todo $n \geq 1$ se cumple que

$$\left| \int_E (f_n - f) \, d\mu \right| \;\leq\; \int_E |f_n - f| \, d\mu \;\leq\; \int_{\mathbb{R}} |f_n - f| \, d\mu,$$

y, en consecuencia, $\lim\limits_{n\to\infty} \int_E f_n \, d\mu = \int_E f \, d\mu$. Por otro lado, como $\big| |f_n| - |f| \big| \leq |f_n - f|$ para todo $n \geq 1$, se tiene que

$$\left| \int_E (|f_n| - |f|) \, d\mu \right| \;\leq\; \int_E \big| |f_n| - |f| \big| \, d\mu \;\leq\; \int_{\mathbb{R}} |f_n - f| \, d\mu \;=\; 0,$$

de donde se sigue que $\lim\limits_{n\to\infty} \int_E |f_n| \, d\mu \;=\; \int_E |f| \, d\mu$ y termina la prueba. ■

Hemos visto que si $f_n \to f$ en la norma de L_1, entonces

$$\lim_{n\to\infty} \int_E f_n \, d\mu \;=\; \int_E f \, d\mu \qquad \text{y} \qquad \lim_{n\to\infty} \int_E |f_n| \, d\mu \;=\; \int_E |f| \, d\mu$$

se cumplen para todo $E \in \mathfrak{M}_\mu(\mathbb{R})$. ¿Qué otra conclusión podemos deducir de la sucesión $(f_n)_{n=1}^{\infty}$? Por ejemplo, ¿converge dicha sucesión a f casi-siempre? La respuesta es que, en general, tal conclusión es falsa. Sin embargo, la siguiente afirmación es válida: para cada $\varepsilon > 0$,

$$\lim_{n\to\infty} \mu\big(\{x \in \mathbb{R} : \; |f_n(x) - f(x)| \geq \varepsilon\}\big) \;=\; 0.$$

En efecto, por la Desigualdad de Chebyshev tenemos que

$$\mu\big(\{x \in \mathbb{R} : \; |f_n(x) - f(x)| \geq \varepsilon\}\big) \;\leq\; \frac{1}{\varepsilon} \int_{\mathbb{R}} |f_n - f| \, d\mu \;=\; \frac{1}{\varepsilon} \| f_n - f \|_1 \;\to\; 0.$$

Esto nos muestra que la sucesión $(f_n)_{n=1}^{\infty}$ *converge en medida* a f. Más aun, del Teorema de Riesz, Teorema 7.2.26, se sigue que existe una subsucesión $(f_{n_k})_{k=1}^{\infty}$ de $(f_n)_{n=1}^{\infty}$ tal que $f_{n_k} \to f$ casi-siempre. Con esto hemos demostrado el siguiente corolario.

Corolario 10.2.5. *Sea $(f_n)_{n=1}^{\infty}$ una sucesión en $\mathrm{L}_1(\mu)$ tal que $\lim_{n\to\infty}\|f_n - f\|_1 = 0$ para alguna función Lebesgue integrable f. Entonces*

*(a) $(f_n)_{n=1}^{\infty}$ **converge en medida** a f.*

*(b) Existe una **subsucesión** $(f_{n_k})_{k=1}^{\infty}$ de $(f_n)_{n=1}^{\infty}$ tal que $f_{n_k}\to f$ **casi-siempre**.*

La importancia de la norma $\|\cdot\|_1$ sobre $\mathrm{L}_1(\mu)$ es que ella es completa, es decir, el espacio normado $(\mathrm{L}_1(\mu),\|1\|)$ es, en realidad, un espacio de Banach.

Teorema 10.2.6. *$(\mathrm{L}_1(\mu),\|\cdot\|_1)$ es un **espacio de Banach**.*

Prueba. Sea $(f_n)_{n=1}^{\infty}$ una sucesión en $\mathrm{L}_1(\mu)$ y suponga que la serie $\sum_{n=1}^{\infty}\|f_n\|_1$ es absolutamente sumable, esto es,

$$\sum_{n=1}^{\infty}\|f_n\|_1 = \sum_{n=1}^{\infty}\int_{\mathbb{R}}|f_n|\,d\mu < +\infty.$$

Se sigue del Teorema de Beppo Levi que la serie $\sum_{n=1}^{\infty} f_n$ converge puntualmente a una función Lebesgue integrable, digamos, f. Una nueva aplicación del Teorema de Beppo Levi, pero ahora a la serie $\sum_{n=N+1}^{\infty} f_n$, nos conduce a

$$\begin{aligned}\lim_{N\to\infty}\left\| f - \sum_{n=1}^{N} f_n\right\|_1 &= \lim_{N\to\infty}\int_{\mathbb{R}}\left| f - \sum_{n=1}^{N} f_n\right| d\mu \\ &= \lim_{N\to\infty}\int_{\mathbb{R}}\left|\sum_{n=N+1}^{\infty} f_n\right| d\mu \\ &\leq \lim_{N\to\infty}\sum_{n=N+1}^{\infty}\int_{\mathbb{R}}|f_n|\,d\mu = 0,\end{aligned}$$

es decir, la serie $\sum_{n=1}^{\infty} f_n$ converge, en la norma-L_1, a la función f. Se sigue del Lema 4.1.13, página 203, que $(\mathrm{L}_1(\mu),\|\cdot\|_1)$ es un espacio de Banach. ■

El siguiente resultado establece, en términos generales, que cualquier función Lebesgue integrable con dominio pequeño su integral también es pequeña; expresado en otras palabras, toda función Lebesgue integrable no-negativa genera una medida. Este hecho nos será de gran utilidad en la próxima sección.

Teorema 10.2.7 (Absoluta continuidad). *Sea $f\in\mathrm{L}_1(\mu)$. Dado $\varepsilon>0$, existe un número $\delta>0$ tal que para cada conjunto medible A con $\mu(A)<\delta$ se tiene que*

$$\int_A |f|\,d\mu < \varepsilon.$$

Prueba. Puesto que $|f|$ es una función no-negativa Lebesgue integrable, el Teorema de la Densidad de las Funciones Simples, Teorema 7.2.10, nos garantiza la existencia de una sucesión creciente de funciones simples no-negativas $(\varphi_n)_{n=1}^{\infty}$ convergiendo puntualmente a $|f|$. Se sigue del Teorema de la Convergencia Monótona que

$$\int_{\mathbb{R}}|f|\,d\mu = \lim_{n\to\infty}\int_{\mathbb{R}}\varphi_n\,d\mu.$$

De allí que, dado $\varepsilon > 0$, existe un $N \in \mathbb{N}$ de modo que

$$\int_{\mathbb{R}} (|f| - \varphi_N)\, d\mu \;=\; \int_{\mathbb{R}} |f|\, d\mu \;-\; \int_{\mathbb{R}} \varphi_N\, d\mu \;<\; \frac{\varepsilon}{2}.$$

Claramente $\varphi_N \leq M$ para alguna constante $M > 0$. Si ahora definimos $\delta = \varepsilon/(2M)$, resulta que si A es cualquier conjunto medible con $\mu(A) < \delta$, entonces

$$\begin{aligned}
\int_A |f|\, d\mu \;&=\; \int_A (|f| - \varphi_N)\, d\mu \;+\; \int_A \varphi_N\, d\mu \\
&\leq\; \int_{\mathbb{R}} (|f| - \varphi_N)\, d\mu \;+\; \int_A M\, d\mu \\
&<\; \frac{\varepsilon}{2} + \frac{\varepsilon}{2M} \cdot M \;=\; \varepsilon
\end{aligned}$$

lo que prueba el teorema. ■

Una manera abreviada de expresar el resultado anterior es por medio de la igualdad

$$\lim_{\mu(A)\to 0} \nu_f(A) \;=\; 0,$$

donde $\nu_f(E) = \int_E |f|\, d\mu$ para cada $E \in \mathfrak{M}_\mu(\mathbb{R})$. Observe que si $\mu(A) = 0$, entonces $\nu_f(A) = 0$. Esto último también se escribe en la forma $\nu_f \ll \mu$.

Definición 10.2.8. *Si ν es cualquier otra medida definida sobre $\mathfrak{M}_\mu(\mathbb{R})$, diremos que ν es **absolutamente continua con respecto a** μ, y se denotará por $\nu \ll \mu$, si para cada $A \in \mathfrak{M}_\mu(\mathbb{R})$ con $\mu(A) = 0$ se cumple que $\nu(A) = 0$.*

Observe que si $\mathfrak{V}$ es una colección finita o infinita de medidas definidas sobre $\mathfrak{M}_\mu(\mathbb{R})$ tal que $\nu \leq \mu$ para cualquier $\nu \in \mathfrak{V}$, entonces $\nu \ll \mu$ para toda $\nu \in \mathfrak{V}$. Medidas que son absolutamente continuas con respecto a la medida de Lebesgue se caracterizan por medio del siguiente resultado.

Teorema 10.2.9. *Sea ν una **medida finita** definida sobre $\mathfrak{M}_\mu(\mathbb{R})$. Las siguientes condiciones son equivalentes:*

(1) $\nu \ll \mu$.

(2) $\lim_{\mu(A)\to 0} \nu(A) = 0$.

Prueba. (1) $\Rightarrow$ (2). Para generar una contradicción, suponga que (1) se cumple pero que (2) es falso. Esto, por supuesto, significa que existe un extraño $\varepsilon > 0$ con la siguiente propiedad: para cada $\delta > 0$, existe un conjunto medible A_δ tal que

$$\mu(A_\delta) \;<\; \delta \quad \text{pero} \quad \nu(A_\delta) \;\geq\; \varepsilon.$$

En particular, si para cada $n \in \mathbb{N}$ tomamos $\delta = 1/2^n$, lo anterior permite construir una sucesión $(A_n)_{n=1}^{\infty}$ de conjuntos medibles tal que para cada $n \geq 1$,

$$\mu(A_n) \;<\; \frac{1}{2^n} \quad \text{pero} \quad \nu(A_n) \;\geq\; \varepsilon.$$

Definamos

$$A \;=\; \limsup_{n\to\infty} A_n \;=\; \bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty} A_m.$$

Observe que, para cada $n \geq 1$, $A \subseteq \bigcup_{m=n}^{\infty} A_m$ y, por lo tanto,

$$\mu(A) \;\leq\; \sum_{m=n}^{\infty}\mu(A_m) \;<\; \sum_{m=n}^{\infty}\frac{1}{2^m} \;=\; \frac{1}{2^{n-1}}.$$

Puesto que esto es válido para todo $n \geq 1$, se sigue que $\mu(A) = 0$. Por otro lado, como ν es una medida finita resulta que $\nu(\bigcup_{n=1}^{\infty} A_n) < +\infty$ y, en consecuencia, por el Teorema 6.3.48 (b), tenemos que

$$\nu(A) \;=\; \nu\Big(\limsup_{n\to\infty} A_n\Big) \;\geq\; \limsup_{n\to\infty}\nu(A_n) \;\geq\; \varepsilon \;>\; 0.$$

Con esto, hemos demostrado la existencia de un conjunto medible A tal que que $\mu(A) = 0$ pero que $\nu(A) > 0$, lo cual nos dice que ν no es absolutamente continua con respecto a μ. Esta contradicción confirma que $\lim_{\mu(A)\to 0}\nu(A) = 0$.

(2) $\Rightarrow$ (1). Sea $\varepsilon > 0$ elegido de forma arbitraria. La validez de (2) significa que existe un $\delta > 0$ tal que $\nu(E) < \varepsilon$ siempre que $\mu(E) < \delta$. Sea $E \in \mathfrak{M}_\mu(\mathbb{R})$ tal que $\mu(E) = 0$. En particular, $\mu(E) < \delta$ y así, $\nu(E) < \varepsilon$. Puesto que $\varepsilon > 0$ fue elegido de modo arbitrario, se tiene que $\nu(E) = 0$ y termina la prueba. ■

El resultado anterior permite restablecer el Teorema 10.2.7 en la siguiente forma:

Para cada $f \in L_1(\mu)$, $\nu_f \ll \mu$.

Existe un viejito, pero poderoso y fascinate resultado conocido como el Teorema de Radon-Nikodým, que establece que cualquier medida ν satisfaciendo $\nu \ll \mu$, es de la forma $\nu = \nu_f$ para una única $f \in L_1(\mu)$.

El siguiente resultado es la pieza fundamental para generar los conjuntos uniformemente integrables los cuales son de vital importancia en Probabilidad y, por supuesto, en la Teoría de la Medida y la Teoría de los Espacios de Banach.

Corolario 10.2.10. *Sea* $f \in L_1(\mu)$. *Dado* $\varepsilon > 0$, *existe una constante* $c > 0$ *tal que*

$$\int_{[\,|f|>c\,]} |f|\,d\mu \;<\; \varepsilon.$$

donde $[\,|f| > c\,] = \{x \in \mathbb{R} : |f(x)| > c\}$.

Prueba. Escoja un $\delta > 0$ satisfaciendo la conclusión del Teorema 10.2.7. Por la Desigualdad de Chevyshev sabemos que

$$c\cdot\{x \in \mathbb{R} : |f(x)| > c\} \;\leq\; \int_{\mathbb{R}} |f|\,d\mu \;<\; +\infty.$$

Si hacemos $\int_{\mathbb{R}} |f|\,d\mu = M$ y elegimos un $c > 0$ de modo que $M/c < \delta$, entonces de la desigualdad anterior se tiene que $\mu([\,|f| > c\,]) < \delta$ y, en consecuencia, se cumple que $\int_{[\,|f|>c\,]} |f|\,d\mu < \varepsilon$. Esto termina la prueba. ■

Un hecho que también resulta ser de gran utilidad en ciertas ocasiones es el siguiente resultado, el cual estable que cada función Lebesgue integrable, digamos f, da origin a la existencia de un conjunto medible K tal que la diferencia entre los números $\nu_f(E)$ y $\nu_f(\mathbb{R})$ puede hacerse tan pequeña como se desee.

Teorema 10.2.11 (Vitali). *Sea $f \in \mathrm{L}_1(\mu)$. Para cada $\varepsilon > 0$, existe un conjunto $E \in \mathfrak{M}_\mu(\mathbb{R})$ con $\mu(E) < +\infty$ tal que*

$$\int_{\mathbb{R}\setminus E} |f|\, d\mu \;<\; \varepsilon.$$

Más aun,

$$\left| \int_{\mathbb{R}} f\, d\mu - \int_E f\, d\mu \right| \;<\; \varepsilon.$$

Prueba. Sea $\varepsilon > 0$. Para cada $n \geq 1$, defina $f_n = |f| \cdot \chi_{[-n,n]}$. Resulta que la sucesión de funciones medibles $(f_n)_{n=1}^\infty$ es no-negativa, creciente y converge puntualmente a $|f|$. Un llamado al Teorema de la Convergencia Monótona nos muestra que

$$\lim_{n\to\infty} \int_{\mathbb{R}} f_n\, d\mu \;=\; \int_{\mathbb{R}} |f|\, d\mu,$$

de donde se sigue la existencia de un $N \in \mathbb{N}$ tal que

$$\int_{\mathbb{R}} |f|\, d\mu \;-\; \int_{\mathbb{R}} f_N\, d\mu \;<\; \varepsilon.$$

Puesto que

$$\int_{\mathbb{R}} f_N\, d\mu \;=\; \int_{[-N,N]} |f|\, d\mu,$$

resulta, por la integrabilidad de f, que

$$\int_{\mathbb{R}\setminus[-N,N]} |f|\, d\mu \;=\; \int_{\mathbb{R}} |f|\, d\mu - \int_{[-N,N]} |f|\, d\mu \;=\; \int_{\mathbb{R}} |f|\, d\mu \;-\; \int_{\mathbb{R}} f_N\, d\mu \;<\; \varepsilon,$$

de modo que, tomando $E = [-N, N]$, se termina la prueba. ■

Teorema 10.2.12 (Riesz). *Sea $(f_n)_{n=1}^\infty$ una sucesión en $\mathrm{L}_1(\mu)$ y suponga que $f_n \to f$ **casi-siempre** para alguna función $f \in \mathrm{L}_1(\mu)$. Las siguientes declaraciones son equivalentes:*

(1) $\lim_{n\to\infty} \| f_n - f \|_1 = 0.$

(2) $\lim_{n\to\infty} \| f_n \|_1 = \| f \|_1.$

Prueba. Ya hemos visto, véase el Lema 10.2.4, que $(1) \Rightarrow (2)$. Para ver que $(2) \Rightarrow (1)$ suponga, en primer lugar, que E es un conjunto medible con $0 < \mu(E) < +\infty$ y sea $\varepsilon > 0$. La absoluta continuidad de la integral, Teorema 10.2.7, nos garantiza la existencia de un $\delta > 0$ con la siguiente propiedad: para cualquier conjunto medible $A \subseteq E$ con $\mu(A) < \delta$, se cumple que

$$\int_A |f|\, d\mu \;<\; \varepsilon/4$$

Fijemos un conjunto medible $A \subseteq E$ con $\mu(A) < \delta$. De (2) se sigue

$$\lim_{n\to\infty} \int_A |f_n|\, d\mu \;=\; \int_A |f|\, d\mu,$$

y, en consecuencia, podemos seleccionar un $N_1 \in \mathbb{N}$ de modo tal que, para todo $n \geq N_1$,

$$\int_A |f_n|\, d\mu \;<\; \int_A |f|\, d\mu \;+\; \varepsilon/4.$$

Como $\mu(E) < +\infty$ y $f_n \to f$ casi-siempre, el Teorema de Severini-Egoroff nos permite determinar un conjunto medible $A \subseteq E$ tal que

$$\mu(A) \;<\; \delta \qquad\qquad \text{y} \qquad\qquad f_n \to f \ \text{ uniformemente sobre } E \setminus A.$$

Elija un $N_2 \in \mathbb{N}$ de modo que $|f_n(x) - f(x)| < \varepsilon/4\mu(E)$ para todo $x \in E \setminus B$ y todo $n \geq N_2$. Se sigue de lo anterior que si $n \geq N = \max\{N_1, N_2\}$, entonces

$$\begin{aligned}
\int_E |f_n - f|\, d\mu \;&=\; \int_{E\setminus A} |f_n - f|\, d\mu \;+\; \int_A |f_n - f|\, d\mu \\
&\leq\; \int_{E\setminus A} |f_n - f|\, d\mu \;+\; \int_A |f_n|\, d\mu \;+\; \int_A |f|\, d\mu \\
&\leq\; \varepsilon/4 \;+\; 2\varepsilon/4 \;+\; \varepsilon/4 \;=\; \varepsilon.
\end{aligned}$$

Esto termina la prueba para el caso finito. Para el caso general, sea $\varepsilon > 0$ y seleccione, usando el Teorema 10.2.11, un conjunto medible $E \subseteq \mathbb{R}$ con $\mu(E) < +\infty$ tal que

$$\int_{\mathbb{R}\setminus E} |f|\, d\mu \;<\; \varepsilon/4.$$

Como $\| f_n \|_1 \to \| f \|_1$, resulta del Lema 10.2.4 que

$$\lim_{n\to\infty} \int_{\mathbb{R}\setminus E} |f_n|\, d\mu \;=\; \int_{\mathbb{R}\setminus E} |f|\, d\mu.$$

Elija un $N_1 \in \mathbb{N}$ tal que

$$\int_{\mathbb{R}\setminus E} |f_n|\, d\mu \;<\; \int_{\mathbb{R}\setminus E} |f|\, d\mu \;+\; \varepsilon/4.$$

para todo $n \geq N_1$. Puesto que $\mu(E) < +\infty$ podemos escoger, por la primera parte, un $N_2 \in \mathbb{N}$ de modo tal que

$$\int_E |f_n - f|\, d\mu \;<\; \varepsilon/4$$

para todo $n \geq N_2$. Finalmente, si $n \geq N = \max\{N_1, N_2\}$, entonces tenemos que,

$$\begin{aligned}
\int_{\mathbb{R}} |f_n - f|\, d\mu \;&=\; \int_{\mathbb{R}\setminus E} |f_n - f|\, d\mu \;+\; \int_E |f_n - f|\, d\mu \\
&\leq\; 2\int_{\mathbb{R}\setminus E} |f|\, d\mu \;+\; \varepsilon/4 \;+\; \int_E |f_n - f|\, d\mu \\
&<\; \varepsilon/2 + \varepsilon/2 \;=\; \varepsilon
\end{aligned}$$

para todo $n \geq N$. La prueba es completa. ■

10.2.1. Densidad en el espacio $L_1([a,b])$

Sea $(X, \|\cdot\|)$ un espacio normado y sea $D \subseteq X$. Recordemos que D es norma-denso en X si $\overline{D} = X$, en otras palabras, si para cada $x \in X$ y cada $\varepsilon > 0$, existe un $z \in D$ tal que $\|x - z\| < \varepsilon$ o, de modo equivalente, si para cada $x \in X$ existe una sucesión $(x_n)_{n=1}^{\infty}$ en D tal que $\|x_n - x\| \to 0$ cuando $n \to \infty$. Además, si D y F son subconjuntos de X tales que: $D \subseteq F$ y D es denso en X, entonces F es denso en X. Igualmente, si D es denso en F y F es denso en X, entonces D es denso en X. En el siguiente resultado veremos que los espacios vectoriales $\mathcal{S}_\mu([a,b])$ y $C([a,b])$ constituidos, respectivamente, por todas las funciones simples y todas las funciones continuas definidas sobre $[a,b]$ son subconjuntos norma-densos en $L_1([a,b])$. En particular, $\mathcal{R}([a,b])$ también lo es.

Teorema 10.2.13 (Densidad en L_1). *Cada uno de los siguientes subespacios:*

*(a) $\mathcal{S}_\mu([a,b])$, el conjunto de las **funciones simples** sobre $[a,b]$;*

*(b) $C([a,b])$, el conjunto de las **funciones continuas** sobre $[a,b]$;*

*(c) $\mathcal{R}([a,b])$, el conjunto de las **funciones Riemann-integrables** sobre $[a,b]$;*

*es **denso** en el espacio $\big(L_1([a,b]), \|\cdot\|_1\big)$.*

Prueba. (a) **Densidad de** $\mathcal{S}_\mu([a,b])$. Sea $f \in L_1([a,b])$ y fijemos un $\varepsilon > 0$. Por el Corolario 7.2.11, página 385, sabemos que existe una sucesión $(\varphi_n)_{n=1}^{\infty}$ de funciones simples en $\mathcal{S}_\mu([a,b])$ tal que $\lim_{n\to\infty} \varphi_n = f$ puntualmente y $|\varphi_n| \leq |f|$. Puesto que $|f|$ es Lebesgue integrable, se sigue que φ_n es Lebesgue integrable para todo $n \geq 1$. Un llamado al Teorema de la Convergencia Dominada nos revela que

$$\lim_{n\to\infty} \int_a^b |\varphi_n - f|\, d\mu \;=\; \int_a^b \lim_{n\to\infty} |\varphi_n - f|\, d\mu \;=\; 0,$$

es decir, $\lim\limits_{n\to\infty} \|\varphi_n - f\|_1 = 0$. Esto prueba que $\mathcal{S}_\mu([a,b])$ es norma-denso en $L_1([a,b])$.

(b) **Densidad de** $C([a,b])$. Para demostrar la densidad de las funciones continuas, en primer lugar vamos a verificar que si $\varphi = \chi_A$ es la función característica de algún conjunto medible $A \subseteq [a,b]$, entonces, para cada $\varepsilon > 0$, se puede determinar una función en escalera $\psi : [a,b] \to \mathbb{R}$ tal que

$$\|\varphi - \psi\|_1 \;=\; \int_a^b |\varphi - \psi|\, d\mu \;<\; \varepsilon.$$

En efecto, se $A \subseteq [a,b]$ un conjunto medible y fijemos un $\varepsilon > 0$. Como A es medible, existe un conjunto abierto $G \subseteq [a,b]$ tal que

$$A \subseteq G \qquad \text{y} \qquad \mu(G \setminus A) \;<\; \varepsilon/2.$$

En particular, $\mu(G) < \mu(A) + \varepsilon/2$. Por otro lado, siendo G abierto, existe una sucesión disjunta $((a_n, b_n))_{n=1}^{\infty}$ de intervalos abiertos tal que $G = \bigcup_{n=1}^{\infty}(a_n, b_n)$ y, por lo tanto,

$$\mu(G) \;=\; \sum_{n=1}^{\infty}(b_n - a_n) \;<\; \mu(A) + \varepsilon/2.$$

Seleccione un $N \in \mathbb{N}$ de modo que $\sum_{n=N+1}^{\infty}(b_n - a_n) < \varepsilon/2$ y sea ψ la función característica del conjunto $[a,b] \cap \bigcup_{n=1}^{N}(a_n,b_n)$. Resulta que ψ es una función en escalera. Sea h la función característica de $[a,b] \cap \bigcup_{n=1}^{\infty}(a_n,b_n)$ y observe que

$$\begin{aligned} \|\varphi - \psi\|_1 &\leq \|\varphi - h\|_1 + \|h - \psi\|_1 \\ &\leq \mu\Big(\bigcup_{n=1}^{\infty}(a_n,b_n) \setminus A\Big) + \mu\Big(\bigcup_{n=N+1}^{\infty}(a_n,b_n)\Big) \\ &< \varepsilon/2 + \varepsilon/2 = \varepsilon. \end{aligned}$$

Finalmente, si $\varphi = \sum_{i=1}^{n} a_i \cdot \chi_{A_i}$ es una función simple arbitraria, se sigue de la parte de arriba que, por cada $i = 1,\ldots,n$, existe una función en escalera ψ_i tal que $\|\chi_{A_i} - \psi_i\|_1 < \varepsilon/|a_i|n$. Si ahora definimos $\psi = \sum_{i=1}^{n} a_i\psi_i$, resulta que ψ es una función en escalera y se cumple que

$$\|\varphi - \psi\|_1 = \sum_{i=1}^{n}|a_i|\,\|\chi_{A_i} - \psi_i\|_1 < \sum_{i=1}^{n}|a_i|\frac{\varepsilon}{|a_i|n} = \varepsilon.$$

Para terminar la prueba, usemos el Ejercicio 8.4.2, página 447, para hallar una función continua $g : [a,b] \to \mathbb{R}$ tal que

$$\|\psi - g\|_1 < \varepsilon.$$

Sea ahora $f \in L_1([a,b])$. Por la primera parte, existe una función simple φ tal que $\|f - \varphi\|_1 < \varepsilon/3$ y por, lo anterior, existe una función en escalera ψ y también una función continua g tales que $\|\varphi - \psi\|_1 < \varepsilon/3$ y $\|\psi - g\|_1 < \varepsilon/3$. Por esto

$$\|f - g\|_1 \leq \|f - \varphi\|_1 + \|\varphi - \psi\|_1 + \|\psi - g\|_1 < \varepsilon.$$

De aquí se concluye que $C([a,b])$ es norma-denso en $L_1([a,b])$.

(c) **Densidad de** $\mathcal{R}([a,b])$. El hecho de que $\mathcal{R}([a,b])$ es denso en $\big(L_1([a,b],\mu,\|\cdot\|_1\big)$ sigue de (b) ya que

$$C([a,b]) \subseteq \mathcal{R}([a,b]) \subseteq L_1([a,b]).$$

La prueba es completa. ■

El lector atento habrá observado que en la demostración del teorema anterior se obtuvo el siguiente resultado.

Corolario 10.2.14. *El conjunto* $\mathrm{Esc}([a,b])$ *constituido por todas las* ***funciones en escaleras*** *definidas sobre* $[a,b]$ *es* ***norma-denso*** *en* $L_1([a,b])$.

Cuando el dominio de nuestras funciones no es un intervalo compacto sino $\mathbb{R}$, o cualquier subconjunto localmente compacto de $\mathbb{R}$, podemos reemplazar funciones continuas por funciones continuas a soporte compacto para aproximar cualquier función en $L_1(\mathbb{R},\mu)$. La prueba de esto requiere de dos resultados. El primero es el siguiente:

$$C_c(\mathbb{R}) \subseteq L_1(\mathbb{R},\mu).$$

En efecto, sea $f \in C_c(\mathbb{R})$. Puesto que $f(x) = 0$ para todo $x \in \mathrm{sop}(f)^c$ y $\mu(\mathrm{sop}(f)) < +\infty$, resulta que

$$\begin{aligned}\int_{\mathbb{R}} |f|\,d\mu &= \int_{\mathrm{sop}(f)^c} |f|\,d\mu + \int_{\mathrm{sop}(f)} |f|\,d\mu \\ &= \int_{\mathrm{sop}(f)} |f|\,d\mu \le \|f\|_\infty\,\mu(\mathrm{sop}(f)) < +\infty,\end{aligned}$$

es decir, $f \in L_1(\mathbb{R}, \mu)$. El segundo resulta que necesitaremos es el siguiente:

Lema 10.2.15. *Sea $g : [a, b] \to \mathbb{R}$ una **función continua** sobre $[a, b]$. Para $\varepsilon > 0$, existe un **intervalo compacto** $K \supseteq [a, b]$ y una función $f \in C_c(\mathbb{R})$ tal que:*

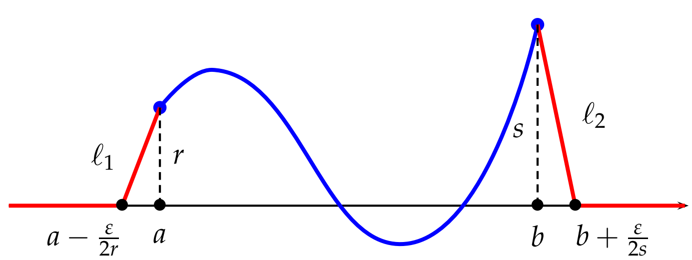
$$(\alpha) \quad f(x) = g(x) \ \text{para todo} \ x \in [a, b], \quad y$$

$$(\beta) \quad \int_{K\setminus[a,b]} |f(x)|\,dx < \varepsilon.$$

Prueba. Extienda a g a todo $\mathbb{R}$ haciendo que $g(x) = 0$ para todo $x \in \mathbb{R} \setminus [a, b]$. Considere la función $f : \mathbb{R} \to \mathbb{R}$ definida por

$$f(x) = \begin{cases} g(x) & \text{si} \quad x \in [a,\, b] \\ \ell_1(x) & \text{si} \quad x \in [a - \varepsilon/2r,\, a) \\ \ell_2(x) & \text{si} \quad x \in (b,\, b + \varepsilon/2s] \\ 0 & \text{si} \quad x \notin [a - \varepsilon/2r,\, b + \varepsilon/2s] \end{cases}$$

donde $r = g(a)$, $s = g(b)$, ℓ_1 es el segmento de recta que une a los puntos $(a - \varepsilon/2r,\, 0)$ y (a, r) y ℓ_2 es el segmento de recta que une a los puntos (b, s) y $(b + \varepsilon/2s,\, 0)$.



Tomando $K = [a - \varepsilon/2r, b + \varepsilon/2s]$, resulta que $[a, b] \subseteq J$ y el área de los dos triángulos mostrados en la figura es:

$$\int_{\mathbb{R}} |f(x) - g(x)|\,dx = \int_{K\setminus[a,b]} |f(x)|\,dx = \frac{1}{2}\cdot\left[\frac{\varepsilon}{2r}\cdot r + \frac{\varepsilon}{2s}\cdot s\right] = \frac{\varepsilon}{2} < \varepsilon.$$

La prueba es completa. ■

Teorema 10.2.16. *$C_c(\mathbb{R})$ es **norma-denso** en $L_1(\mathbb{R}, \mu)$.*

Prueba. Fijemos $f \in L_1(\mathbb{R},\mu)$ y sea $\varepsilon > 0$ arbitrario. Queremos demostrar la existencia de una función $g \in C_c(\mathbb{R})$ tal que $\| f - g \|_1 < 1$. Para ello, considere la sucesión $(f_n)_{n=1}^{\infty}$ incluida en $L_1(\mathbb{R},\mu)$ definida por $f_n = f \cdot \chi_{[-n,n]}$ para cada $n \geq 1$, esto es,

$$f_n(x) \;=\; \begin{cases} f(x) & \text{si } \; |x| \leq n \\ 0 & \text{si } \; |x| > n. \end{cases}$$

Puesto que $f_n \to f$ puntualmente y $|f_n| \leq |f|$ para todo $n \geq 1$, el Teorema de la Convergencia Dominada nos garantiza que

$$\int_{\mathbb{R}} |f|\, d\mu \;=\; \lim_{n\to\infty} \int_{\mathbb{R}} |f_n|\, d\mu.$$

Escojamos un $m \in \mathbb{N}$ de modo tal que

$$\| f_m - f \|_1 \;<\; \frac{\varepsilon}{3},$$

y sea g_m la restricción de f al intervalo $[-m,m]$. Puesto que $g_m \in L_1([-m,m],\mu)$ seleccione, usando el Teorema 10.2.13, una función $h \in C([-m,m])$ tal que

$$\int_{[-m,m]} |g_m - h|\, d\mu \;<\; \frac{\varepsilon}{3}.$$

Extienda h a todo $\mathbb{R}$ declarando que

$$\widetilde{h}(x) \;=\; \begin{cases} h(x) & \text{si } \; x \in [-m,m] \\ 0 & \text{si } \; x \in \mathbb{R} \setminus [-m,m]. \end{cases}$$

Observe que $\widetilde{h} \in L_1(\mathbb{R},\mu)$ y

$$\begin{aligned} \left\| f_m - \widetilde{h} \right\|_1 &= \int_{\mathbb{R}} |f - \widetilde{h}| \cdot \chi_{[-m,m]}\, d\mu \\ &= \int_{[-m,m]} |g_m - h|\, d\mu + \int_{\mathbb{R}\setminus[-m,m]} |0 - 0|\, d\mu \\ &< \frac{\varepsilon}{3}. \end{aligned}$$

Invoque ahora el Lema 10.2.15, para elegir una función $g \in C_c(\mathbb{R})$ y un intervalo compacto $K \supseteq [-m,m]$ tal que $g(x) = \widetilde{h}(x)$ para todo $x \in [-m,m]$ y

$$\left\| \widetilde{h} - g \right\|_1 \;=\; \int_{\mathbb{R}} |g(x) - \widetilde{h}(x)|\, dx \;=\; \int_{K\setminus[-m,m]} |g|\, d\mu \;<\; \frac{\varepsilon}{3}$$

Combinando las desigualdades anteriores se obtiene

$$\| f - g \|_1 \;\leq\; \| f - f_m \|_1 + \left\| f_m - \widetilde{h} \right\|_1 + \left\| \widetilde{h} - g \right\|_1 \;<\; \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \;=\; \varepsilon$$

que era lo que queríamos demostrar. ■

10.2.2. El Lema de Riemann-Lebesgue

Un famoso resultado, con amplias aplicaciones en el Análisis de Fourier y conocido tradicionalmente como el Lema de Riemann-Lebesgue, puede ser demostrado usando el Teorema de Densidad en L_1.

Lema 10.2.17 (Riemann-Lebesgue). *Sea* $f \in \mathrm{L}_1([0, 2\pi])$. *Entonces*

$$\lim_{n\to\infty} \int_0^{2\pi} f(t)\varphi(nt)\, d\mu \;=\; 0,$$

para cualquier función $\varphi : \mathbb{R} \to \mathbb{R}$ ***continua*** *sobre* $\mathbb{R}$ *y* ***periódica*** *de período* 2π *con* $\displaystyle\int_0^{2\pi} \varphi\, d\mu = 0$.

Prueba. La conclusión es trivial si $\varphi = 0$. Suponga que $\varphi \neq 0$. No se pierde generalidad en suponer que φ tiene período 1. Nuestro primera tarea consistirá en demostrar el resultado para cualquier función continua sobre $[0, 1]$. Suponga entonces que $g : [0, 1] \to \mathbb{R}$ es una función continua y pongamos $K = \int_{[0,1]} |\varphi|\, d\mu$. Dado $\varepsilon > 0$ seleccione, usando el hecho de que g es uniformemente continua sobre $[0, 1]$, un $N \in \mathbb{N}$ de modo tal que para todo $x, y \in [0, 1]$ con $|x - y| < 1/N$, se cumpla que $|g(x) - g(y)| < \varepsilon$. De esto se sigue que, para cualquier entero $n \geq N$, cualquier $1 \leq k \leq n$ y cada $t \in [k - 1, k]$, se verifica la desigualdad

$$\left| g\Big(\frac{t}{n}\Big) - g\Big(\frac{k}{n}\Big) \right| \;<\; \varepsilon/2K.$$

El hecho de que φ es periódica de período 1 y $K = \int_{[0,1]} |\varphi|\, d\mu$ nos revela que

$$\int_{k-1}^{k} g\Big(\frac{k}{n}\Big)\varphi(t)\, d\mu \;=\; 0 \qquad\qquad \text{y} \qquad\qquad K \;=\; \int_{k-1}^{k} |\varphi|\, d\mu$$

para cada $1 \leq k \leq n$, de donde resulta que

$$\begin{aligned}
\left| \int_{k-1}^{k} g\Big(\frac{t}{n}\Big)\varphi(t)\, d\mu \right| \;&=\; \left| \int_{k-1}^{k} \Big(g\Big(\frac{t}{n}\Big) - g\Big(\frac{k}{n}\Big)\Big)\varphi(t)\, d\mu \right| \\
&\leq\; \int_{k-1}^{k} \left| g\Big(\frac{t}{n}\Big) - g\Big(\frac{k}{n}\Big) \right| |\varphi(t)|\, d\mu \\
&<\; \varepsilon/2
\end{aligned}$$

y, en consecuencia, para cualquier $n \geq N$,

$$\left| \int_0^1 g(t)\varphi(nt)\, d\mu \right| \;\leq\; \frac{1}{n}\sum_{k=1}^{n} \left| \int_{k-1}^{k} g\Big(\frac{t}{n}\Big)\varphi(t)\, d\mu \right| \;<\; \varepsilon/2. \tag{1}$$

Para el caso general, sea $f \in \mathrm{L}_1([0, 1])$ y para el $\varepsilon > 0$ dado seleccione, invocando el Teorema 10.2.13, una función continua $g \in C([0, 1])$ tal que

$$\| f - g \|_1 \;=\; \int_0^1 |f(t) - g(t)|\, d\mu \;<\; \frac{\varepsilon}{\| \varphi \|_\infty}$$

y, además, cumpla con la desigualdad (1). Entonces

$$\begin{aligned}\left|\int_0^1 f(t)\varphi(nt)\,d\mu\right| &= \left|\int_0^1 (f(t)-g(t))\varphi(nt)\,d\mu + \int_0^1 g(t)\varphi(nt)\,d\mu\right| \\ &\leq \|\varphi\|_\infty \int_0^1 |f(t)-g(t)|\,d\mu + \left|\int_0^1 g(t)\varphi(nt)\,d\mu\right| \\ &< \varepsilon\end{aligned}$$

para todo $n \geq N$. Esto termina la prueba. ■

Como un caso particular del resultado anterior se tiene la comúnmente conocida igualdad de Riemann-Lebesgue.

Corolario 10.2.18 (Riemann-Lebesgue). *Si* $f \in L_1([0,2\pi])$, *entonces*

$$\lim_{n\to\infty} \frac{1}{2\pi}\int_0^{2\pi} f(t)\,\mathrm{sen}(nt)\,d\mu = 0 \qquad y \qquad \lim_{n\to\infty} \frac{1}{2\pi}\int_0^{2\pi} f(t)\cos(nt)\,d\mu = 0.$$

Otra aplicación importante en la Teoría de las Series de Fourier la cual se deriva del Lema de Riemann-Lebesgue es el siguiente:

Teorema 10.2.19 (Cantor-Lebesgue). *Si una serie trigonométrica*

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nt + b_n \,\mathrm{sen}\, nt) \tag{2}$$

converge sobre un conjunto medible E *con* $\mu(E) > 0$, *entonces*

$$\lim_{n\to\infty} a_n = 0 \qquad y \qquad \lim_{n\to\infty} b_n = 0.$$

Prueba. Sin perder generalidad, asumiremos que E tiene medida finita. Es fácil ver que la serie en (2) se puede reescribir en la forma

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} d_n \cos(nt - \varphi_n). \tag{3}$$

En efecto, basta definir, para cada $n \in \mathbb{N}$

$$a_n = d_n \cos\varphi_n \qquad \text{y} \qquad b_n = d_n \,\mathrm{sen}\,\varphi_n,$$

donde

$$d_n = \sqrt{a_n^2 + b_n^2} \qquad \text{y} \qquad \varphi_n = \arccos(a_n/d_n).$$

Con esta redefinición de nuestra serie original, todo lo que tenemos que demostrar es que la sucesión $(d_n)_{n=1}^{\infty}$ converge a cero. Veamos esto. Puesto que la serie en (3) converge, se tiene que

$$\lim_{n\to\infty} d_n \cos(nt - \varphi_n) = 0 \quad \text{para todo} \quad t \in E. \tag{4}$$

Suponga ahora, para construir una contradicción, que la sucesión $(d_n)_{n=1}^{\infty}$ no converge a cero. Entonces, existe un $\varepsilon > 0$ y una subsucesión $(n_k)_{k=1}^{\infty}$ tal que

$$d_{n_k} \;\geq\; \varepsilon \;>\; 0 \quad \text{para todo} \quad k \geq 1.$$

De aquí se sigue, teniendo en cuenta (4), que

$$\lim_{k\to\infty} \cos(n_k t - \varphi_{n_k}) \;=\; 0 \quad \text{para todo} \quad t \in E.$$

Puesto que $\mu(E) < +\infty$ podemos invocar el Teorema de la Convergencia Dominada para obtener

$$\lim_{k\to\infty} \int_E \cos^2(n_k t - \varphi_{n_k})\, dt \;=\; 0. \tag{5}$$

Por otro lado, como $\chi_E \in \mathrm{L}_1(\mathbb{R}, \mu)$ y

$$\cos(2(n_k t - \varphi_{n_k})) \;=\; \cos(2\varphi_{n_k})\cos(2n_k t) \;+\; \operatorname{sen}(2\varphi_{n_k})\operatorname{sen}(2n_k t),$$

se tiene que

$$\begin{aligned}\int_E \cos(2(n_k t - \varphi_{n_k}))\, dt \;&=\; \int_{\mathbb{R}} \chi_E(t)\cos(2(n_k t - \varphi_{n_k}))\, dt \\ &=\; \cos(2\varphi_{n_k})\int_{\mathbb{R}} \chi_E(t)\cos(2n_k t)\, dt \;+\; \operatorname{sen}(2\varphi_{n_k})\int_{\mathbb{R}} \chi_E(t)\operatorname{sen}(2n_k t)\, dt.\end{aligned}$$

Un llamado al Teorema de Riemann-Lebesgue nos muestra que las dos integrales del lado derecho de la igualdad anterior tienden a cero cuando $k \to \infty$. Finalmente, puesto que $\cos^2(n_k t - \varphi_{n_k}) = \frac{1}{2}\big[1 + \cos(2(n_k t - \varphi_{n_k}))\big]$, resulta que

$$\begin{aligned}0 \;=\; \lim_{k\to\infty} \int_E \cos^2(n_k t - \varphi_{n_k})\, dt \;&=\; \lim_{k\to\infty} \int_E \frac{1}{2}\big[1 + \cos(2(n_k t - \varphi_{n_k}))\big] \\ &=\; \frac{\mu(E)}{2} \;>\; 0\end{aligned}$$

Esta contradicción establece $d_n \to 0$ y, por lo tanto, $a_n \to 0$ y $b_n \to 0$. ■

10.2.3. La Completación del Espacio $\mathcal{R}([a, b])$

Recordemos que, dado un espacio métrico arbitrario (X, d), si dicho espacio no es completo, entonces *siempre se puede construir un espacio métrico completo* $\big(\widehat{X}, \widehat{d}\,\big)$ y *una aplicación* $\varphi : X \to \widehat{X}$ con las siguientes propiedades:

*(a) la aplicación φ es una **isometría de** X **sobre** $\varphi(X)$ y $\varphi(X)$ es **denso en** $\widehat{X}$. En este caso, al par $\big(\big(\widehat{X}, \widehat{d}\,\big), \varphi\big)$ se le llama una **completación** de (X, d)* y

*(b) el **espacio métrico completo** $\big(\widehat{X}, \widehat{d}\,\big)$ es, salvo isometría, **único**; es decir, si $\big(\big(\widetilde{X}, \widetilde{d}\,\big), \psi\big)$ es otra completación de (X, d), entonces existe una única isometría $\Phi : \widehat{X} \to \widetilde{X}$ tal que $\Phi \circ \varphi = \psi$.*

En la práctica, casi siempre ocultamos la isometría φ, identificamos a X con su imagen $\varphi(X)$ y simplemente se dice que $\big(\widehat{X}, \widehat{d}\,\big)$ es **la completación de** (X, d). En este caso, $\widehat{d}$ coincide con d sobre $X \times X$.

Siendo $\mathcal{R}([a,b])$ un subconjunto de $L_1([a,b])$, podemos dotar a $\mathcal{R}([a,b])$ de la norma $\|\cdot\|_1$. Es importante tener en cuenta que $\mathcal{R}([a,b])$ no es **completo** en la norma $\|\cdot\|_1$, así como tampoco lo es $C([a,b])$. Sin embrago, si consideramos la aplicación inclusión

$$\mathrm{Id} : \big(\mathcal{R}([a,b]),\, \|\cdot\|_1\big) \;\to\; \big(L_1([a,b]),\, \|\cdot\|_1\big)$$

dada por

$$\mathrm{Id}(f) \;=\; f, \quad f \in \mathcal{R}([a,b])$$

resulta que Id es una isometría de $\mathcal{R}([a,b])$ sobre $\mathrm{Id}(\mathcal{R}([a,b]))$ y como

$$\overline{\mathrm{Id}(\mathcal{R}([a,b]))}^{\,\|\cdot\|_1} \;=\; \overline{\mathcal{R}([a,b])}^{\,\|\cdot\|_1} \;=\; L_1([a,b])$$

gracias al Teorema 10.2.13, se tiene que $\big(L_1([a,b]), \|\cdot\|_1\big)$ es la **completación** de $\big(\mathcal{R}([a,b]), \|\cdot\|_1\big)$. Igual consideración se aplica al espacio $\big(C([a,b]),\, \|\cdot\|_1\big)$; es decir, $\big(L_1([a,b]), \|\cdot\|_1\big)$ también es la completación del espacio métrico no-completo $\big(C([a,b]),\, \|\cdot\|_1\big)$.

10.2.4. Conjuntos Uniformemente Integrables

La noción de conjuntos uniformemente integrables, o equi-integrables, como también se les conoce, es de suma importancia en la Teoría de Probabilidades, fundamentalmente en la Teoría de Martingalas, ya que permite caracterizar a las martingalas que son de un cierto tipo, pero además, en los espacios $L_p(\mu)$ con $1 < p < +\infty$ tales conjuntos se caracterizan por ser norma-acotados. Pero es en el espacio $L_1(\mu)$ donde tales conjuntos adquieren una relevancia fundamental: ellos se caracterizan, precisamente, por ser *relativamente débilmente compactos*, resultado mejor conocido como el Teorema de Dunford-Pettis. Constituyen, de hecho, una herramienta de primer orden para determinar cuándo convergencia casi-siempre se convierte en convergencia en la norma.

En lo que sigue, asumiremos que $X \subseteq \mathbb{R}$ es un **conjunto medible** con $\mu(X) = 1$. La suposición $\mu(X) = 1$ es sólo para que la presentación de los resultados sea más sencilla y elegante, pero todos ellos son válidos para cualquier conjunto medible de medida finita. También, para minimizar y agilizar un poco la notación, escribiremos, para cada función medible $f : X \to \mathbb{R}$,

$$[\,|f| \,>\, c\,] \;=\; \big\{x \in X \,:\, |f(x)| > c\big\},$$

donde $c \geq 0$. El Corolario 10.2.10 es una excelente motivación para la siguiente definición.

Definición 10.2.20. *Un conjunto no vacío* $\mathcal{U} \subseteq L_1(X,\mu)$ *se dice que es* ***uniformemente integrable*** *si*

$$\lim_{c\to\infty} \sup_{f\in\mathcal{U}} \int_{[\,|f|\,>\,c\,]} |f|\, d\mu \;=\; 0. \tag{UI$_1$}$$

Observe que la igualdad en $(\mathrm{UI})_1$ significa que: para cada $\varepsilon > 0$ siempre es posible elegir una constante $c_\varepsilon > 0$ tal que

$$\int_{[\,|f|\,>\,c_\varepsilon\,]} |f|\, d\mu \;\leq\; \varepsilon \quad \text{para toda} \quad f \in \mathcal{U}, \tag{UI$_2$}$$

o, tomando $\varepsilon = 1/2^n$ para cada $n \in \mathbb{N}$: existe una sucesión estrictamente creciente $(c_n)_{n=1}^{\infty}$ de enteros positivos tal que

$$\sup_{f\in\mathcal{U}} \int_{[\,|f|>c_n\,]} |f|\,d\mu \;<\; \frac{1}{2^n} \quad \text{para todo} \quad n \geq 1. \tag{UI$_3$}$$

Diremos que *una sucesión* $(f_n)_{n=1}^{\infty}$ *en* $\mathrm{L}_1(X,\mu)$ *es **uniformemente integrable** si el conjunto* $\mathcal{U} = \{f_1, f_2, \ldots\}$ *es uniformemente integrable*. En ocasiones escribiremos **UI** para abreviar uniformemente integrable.

Un concepto que está íntimamente relacionado al concepto de integrabilidad uniforme es el absoluta continuidad uniforme. Recordemos que cada función f en $\mathrm{L}_1(X,\mu)$ da origen a una medida finita ν_f definida por

$$\nu_f(A) \;=\; \int_A |f|\,d\mu \quad \text{para todo conjunto medible } A \subseteq X,$$

tal que $\nu_f \ll \mu$. Por consiguiente, cada conjunto $\mathcal{U} \subseteq \mathrm{L}_1(X,\mu)$ genera un conjunto de medidas finitas $\mathcal{M}_\mu = \{\nu_f : f \in \mathcal{U}\}$ tal que $\nu_f \ll \mu$ para cada $f \in \mathcal{U}$. Cuando la relación anterior se da manera uniforme se obtiene la siguiente definición.

Definición 10.2.21. *Una familia* $\mathcal{U} \subseteq \mathrm{L}_1(X,\mu)$ *se llama **uniformemente absolutamente continua** si* $\nu_f \ll \mu$ *uniformemente en* $f \in \mathcal{U}$, *es decir, si dado* $\varepsilon > 0$, *existe un* $\delta > 0$ *tal que para cada conjunto medible* $E \subseteq X$ *que verifique* $\mu(E) < \delta$, *se cumple que*

$$\sup_{f\in\mathcal{U}} \int_E |f|\,d\mu \;<\; \varepsilon.$$

Como antes, diremos que *una sucesión* $(f_n)_{n=1}^{\infty}$ *en* $\mathrm{L}_1(X,\mu)$ *es **uniformemente absolutamente continua** si el conjunto* $\mathcal{U} = \{f_1, f_2, \ldots\}$ *es uniformemente absolutamente continuo*. Con frecuencia se utiliza la expresión: $\mathcal{U}$ es **uniformemente** μ**-continua**, en lugar de $\mathcal{U}$ es uniformemente absolutamente continua.

Recordemos que un conjunto $\mathcal{U} \subseteq \mathrm{L}_1(X,\mu)$ es **norma-acotado**, o simplemente, **acotado**, si existe una constante $M > 0$ tal que

$$\sup_{f\in\mathcal{U}} \|f\|_1 \;\leq\; M \qquad\qquad \text{o} \qquad\qquad \int_X |f|\,d\mu \;\leq\; M \quad \text{para toda} \quad f \in \mathcal{U}.$$

El siguiente ejemplo ilustra el caso de una sucesión de funciones en $\mathrm{L}_1([0,1])$ que es norma-acotada, converge puntualmente a cero y, sin embargo, no es uniformemente integrable.

Ejemplo 10.2.1. *Para cada* $n \in \mathbb{N}$, *considere la función* $f_n = n\chi_{[0,1/n]}$. *Resulta que*

$$\|f_n\|_1 \;=\; \int_{[0,1]} |f_n|\,d\mu \;=\; 1$$

y, por lo tanto, la sucesión $(f_n)_{n=1}^{\infty}$ *es **acotada** en* $\mathrm{L}_1([0,1])$. *Claramente,* $f_n \to 0$ *y, sin embargo, para cada* $c > 0$ *y cualquier* $n \geq c$ *se cumple que*

$$\int_{[\,|f_n|>c\,]} |f_n|\,d\mu \;=\; n \cdot \mu([0,1/n]) \;=\; 1.$$

Esto prueba que $\mathcal{U} = \{f_1, f_2, \ldots\}$ *no puede ser **uniformemente integrable**.* ■

El siguiente resultado establece que se requiere algo más que el acotamiento de una colección en $L_1([0,1])$ para garantizar la integrabilidad uniforme. Es interesante destacar que el concepto de conjunto uniformemente integrable, el cual es preferido por probabilistas, fue redescubierto por J. L. Doobs 24 años después de su aparición y usado extensivamente por él en el estudio de las martingalas (véase [46]). Debemos advertir que sólamente la equivalencia $(1) \Leftrightarrow (3)$ del resultado que sigue fue demostrado por de la Vallée-Poussin en [132], pero que, por comodidad, a dicho teorema lo seguiremos nombrando del mismo modo.

Teorema 10.2.22 (Criterio de la Vallée-Poussin). *Sea* $\mathcal{U} \subseteq L_1(X,\mu)$. *Las siguientes condiciones son equivalentes:*

(1) $\mathcal{U}$ *es* ***uniformemente integrable***.

(2) $\mathcal{U}$ *es* ***acotado*** *y* ***uniformemente absolutamente continua***.

(3) *Existe una* ***función creciente*** $\Phi : [0,+\infty) \to [0,+\infty)$ *tal que*

$$\lim_{t\to\infty} \frac{\Phi(t)}{t} = +\infty \qquad y \qquad \sup_{f\in\mathcal{U}} \int_X \Phi(|f|)\, d\mu < +\infty.$$

Prueba. $(1) \Rightarrow (2)$. Suponga que (1) se cumple. Entonces, con $\varepsilon = 1$, seleccione un $c_1 > 0$ tal que

$$\int_{[\,|f|>c_1\,]} |f|\, d\mu \le 1 \quad \text{para todo} \quad f \in \mathcal{U}.$$

Observe ahora que

$$\begin{aligned} \|f\|_1 = \int_X |f|\, d\mu &= \int_{[\,|f|\le c_1\,]} |f|\, d\mu + \int_{[\,|f|>c_1\,]} |f|\, d\mu \\ &\le c_1 \int_X d\mu + \int_{[\,|f|>c_1\,]} |f|\, d\mu \\ &< c_1 \cdot \mu(X) + \int_{[\,|f|>c_1\,]} |f|\, d\mu \qquad (10.2.1) \\ &\le c_1 + 1 \end{aligned}$$

para todo $f \in \mathcal{U}$. Esto prueba que $\mathcal{U}$ es norma-acotado en $L_1(X,\mu)$. Para verificar la validez de que $\mathcal{U}$ es uniformemente absolutamente continua, tome un $\varepsilon > 0$ arbitrario y escoja, usando de nuevo el hecho de que $\mathcal{U}$ es uniformemente integrable, un $c_1 > 0$ tal que

$$\int_{[\,|f|>c_1\,]} |f|\, d\mu \le \varepsilon/2 \quad \text{para toda} \quad f \in \mathcal{U}.$$

Tomando $\delta = \varepsilon/2c_1$ resulta lo siguiente: si E es un subconjunto medible de X con $\mu(E) < \delta$, entonces, por la desigualdad (10.2.1), se tiene que

$$\int_E |f|\, d\mu \le c_1 \cdot \mu(E) + \int_{[\,|f|>c_1\,]} |f|\, d\mu \le \varepsilon$$

para toda $f \in \mathcal{U}$. Esto termina la prueba de (2).

(2) $\Rightarrow$ (1). Sea $M = \sup\{\| f \|_1 : f \in \mathcal{U}\}$ y sea $\varepsilon > 0$. Elija, usando el hecho de que $\mathcal{U}$ es uniformemente absolutamente continua, un $\delta > 0$ tal que

$$\sup_{f\in\mathcal{U}} \int_A |f|\, d\mu \;<\; \varepsilon.$$

para cada conjunto medible $A \subseteq X$ con $\mu(A) < \delta$. Si ahora escojemos $c > 0$ de modo que $M/c < \delta$ resultará, por la Desigualdad de Chevyshev, que

$$\mu\big([\,|f| \,>\, c\,]\big) \;\le\; \frac{1}{c}\int_{[\,|f|>c\,]} |f|\, d\mu \;\le\; \frac{1}{c}\,\| f \|_1 \;\le\; \frac{M}{c} \;<\; \delta$$

para cualquier $f \in \mathcal{U}$ y entonces tendremos que

$$\sup_{f\in\mathcal{U}} \int_{[\,|f|>c\,]} |f|\, d\mu \;<\; \varepsilon.$$

Esto prueba (1).

(1) $\Rightarrow$ (3). Nuestra primera tarea es construir una función creciente $g : [0, +\infty) \to [0, +\infty)$ que sea constante en cada uno de los subintervalos $[n, n+1)$ para $n = 0, 1, 2, \ldots$ y tal que $\lim_{t\to\infty} g(t) = +\infty$. Una vez obtenida la función g se define $\Phi : [0, +\infty) \to [0, +\infty)$ por

$$\Phi(x) \;=\; \int_0^x g(t)\, dt$$

y se prueba que ella posee la propiedades establecidas en (3). Veamos entonces cómo se construye g. Puesto que $\mathcal{U}$ es uniformemente integrable, existe una subsucesión estrictamente creciente $(c_n)_{n=1}^{\infty}$ de enteros positivos tal que

$$\sup_{f\in\mathcal{U}} \int_{[\,|f|>c_n\,]} |f|\, d\mu \;<\; \frac{1}{2^n} \quad \text{para todo} \quad n \ge 1. \tag{UI$_3$}$$

Puesto que $[\,|f| \,\ge\, m\,] = \bigcup_{k=m}^{\infty} [\,k \,\le\, |f| \,<\, k+1\,]$, resulta que

$$\mu\big([\,|f| \,\ge\, m\,]\big) \;=\; \sum_{k=m}^{\infty} \mu\big([\,k \,\le\, |f| \,<\, k+1\,]\big)$$

y, en consecuencia,

$$\begin{aligned}
\int_{[\,|f|\ge c_n\,]} |f|\, d\mu \;&=\; \sum_{m=c_n}^{\infty} \int_{[\,m\le|f|<m+1\,]} |f|\, d\mu \\
&\ge\; \sum_{m=c_n}^{\infty} m\cdot\mu\big([\,m \,\le\, |f| \,<\, m+1\,]\big) \\
&\ge\; \sum_{m=c_n}^{\infty} \mu\big([\,m \,\le\, |f| \,<\, m+1\,]\big) \\
&=\; \sum_{m=c_n}^{\infty} \mu\big([\,|f| \,\ge\, m\,]\big),
\end{aligned}$$

para cada $f \in \mathcal{U}$ y cualquier $n \geq 1$. Se sigue entonces de $(\mathrm{UI})_3$ que

$$\sum_{n=1}^{\infty} \sum_{m=c_n}^{\infty} \mu\big([\,|f| \geq m]\big) \;\leq\; \sum_{n=1}^{\infty} \frac{1}{2^n} \;=\; 1 \quad \text{para cada} \quad f \in \mathcal{U}. \tag{$*$}$$

Defina, para cada $m = 1, 2, \ldots$

$$q_m \;=\; \begin{cases} 0 & \text{si} \quad m < c_1, \\ \text{máx}\{k \in \mathbb{N} \,:\, c_k \leq m\} & \text{si} \quad m \geq c_1. \end{cases}$$

Entonces $\text{lím}_{m\to\infty} q_m = +\infty$ y se cumple, además, que

$$\sum_{n=1}^{\infty} \sum_{m=c_n}^{\infty} \mu\big([\,|f| \geq m]\big) \;=\; \sum_{m=1}^{\infty} q_m \cdot \mu\big([\,|f| \geq m]\big). \tag{$**$}$$

Considere la función $g : [0,+\infty) \to [0,+\infty)$ dada por

$$g(t) \;=\; q_m \quad \text{si} \quad t \in [m, m+1) \quad \text{para} \quad m = 0, 1, 2, \ldots$$

y observe que g es creciente, no-negativa y

$$\int_m^{m+1} g\,dt \;=\; q_m \quad \text{para} \quad m = 0, 1, 2, \ldots$$

Con esto se termina la construcción de nuestra función g. Si ahora definimos

$$\Phi(x) \;=\; \int_0^x g(t)\,dt \quad \text{para todo} \quad x \geq 0,$$

resulta que Φ también es creciente y no-negativa. Más aun, gracias a que g es creciente, se tiene que $g(t) \geq g(x/2)$ para cualquier $t \in [x/2, x]$ y, por lo tanto,

$$\Phi(x) \;=\; \int_0^x g(t)\,dt \;\geq\; \int_{x/2}^x g(t)\,dt \;\geq\; \int_{x/2}^x g(x/2)\,dt \;=\; \frac{x}{2} \cdot g\Big(\frac{x}{2}\Big).$$

De esta última desigualdad se concluye que

$$\lim_{x\to\infty} \frac{\Phi(x)}{x} \;\geq\; \frac{1}{2} \lim_{x\to\infty} g\Big(\frac{x}{2}\Big) \;=\; +\infty.$$

Finalmente, observe que si $n \leq |f(t)| < n+1$, entonces

$$\begin{aligned} \Phi(|f|) \;\leq\; \Phi(n+1) \;&=\; \int_0^{n+1} g\,dt \\ &=\; \int_0^1 g\,dt + \int_1^2 g\,dt + \cdots + \int_n^{n+1} g\,dt \\ &=\; q_0 + q_1 + \cdots + q_n, \end{aligned}$$

de donde resulta, usando $(**)$ y $(*)$, que para cada $f \in \mathfrak{U}$,

$$\begin{aligned}
\int_X \Phi(|f|)\, d\mu &= \sum_{n=0}^{\infty} \int_{[\, n \le |f| < n+1\,]} \Phi(|f|)\, d\mu \\
&\le \sum_{n=0}^{\infty} \int_{[\, n \le |f| < n+1\,]} \Phi(n+1)\, d\mu \\
&= \sum_{n=0}^{\infty} \Big(\sum_{m=0}^{n} q_m \Big) \cdot \mu\big([\, n \le |f| < n+1\,]\big) \\
&= q_0 \cdot \mu\big([\, 0 \le |f| < 1\,]\big) + (q_0 + q_1) \cdot \mu\big([\, 1 \le |f| < 2\,]\big) + \cdots \\
&= \sum_{n=1}^{\infty} q_n \cdot \mu\big([\, |f| \ge n\,]\big) \le 1,
\end{aligned}$$

lo que conduce a que

$$\sup_{f \in \mathfrak{U}} \int_{[0,1]} \Phi(|f|)\, d\mu < +\infty.$$

Esto prueba (3).

(3) $\Rightarrow$ (1). Suponga que (3) se cumple y sea $\varepsilon > 0$. Pongamos

$$M = \sup_{f \in \mathfrak{U}} \int_X \Phi(|f|)\, d\mu < +\infty$$

y use el hecho de que $\Phi(t)/t \to +\infty$ cuando $t \to +\infty$, para encontrar un $t_0 > 0$ tal que $\Phi(t)/t \ge M/\varepsilon$ para todo $t \ge t_0$. Se sigue de esto que, para toda $f \in \mathfrak{U}$, $|f(t)| \le (\varepsilon/M)\Phi(|f(t)|)$ siempre que $|f(t)| \ge t_0$ y, por lo tanto,

$$\int_{[\, |f| > t_0\,]} |f|\, d\mu \le \frac{\varepsilon}{M} \int_{[\, |f| > t_0\,]} \Phi(|f|)\, d\mu \le \varepsilon.$$

Esto termina la prueba. ■

Algunos subconjuntos de $L_1(X, \mu)$ pueden ser fácilmente detectados como uniformemente integrables mediante el siguiente resultado.

Teorema 10.2.23. *Sea $\mathfrak{U} \subseteq L_1(X, \mu)$.*

*(a) Si $\mathfrak{U}$ es un **conjunto finito**, entonces $\mathfrak{U}$ es **uniformemente integrable**.*

*(b) Si $\mathfrak{U} = \{f \in L_1(X, \mu) : |f| \le g\}$, donde $g \in L_1(X, \mu)$, entonces $\mathfrak{U}$ es **uniformemente integrable**.*

Prueba. (a) Suponga que $\mathfrak{U} = \{f_1, \ldots, f_m\}$ y sea $\varepsilon > 0$. Por el Teorema 10.2.7, por cada $j = 1, \ldots, n$, existe un $\delta_j > 0$ tal que

$$\int_E |f_j|\, d\mu < \varepsilon$$

siempre que $E \subseteq X$ sea un conjunto medible con $\mu(E) < \delta_j$. Si ahora seleccionamos un $\delta > 0$ de modo que $\delta < \text{mín}\{\delta_1, \ldots, \delta_n\}$, resultará que

$$\sup_{1 \leq j \leq n} \int_E |f_j|\, d\mu \;<\; \varepsilon$$

para cualquier conjunto medible $E \subseteq X$ con $\mu(E) < \delta$. Como $\mathcal{U}$ es claramente acotado, el Criterio de la Vallée-Poussin nos dice que (a) se cumple.

(b) Sea $\varepsilon > 0$. Puesto que $g \in L_1(X,\mu)$, el Teorema 10.2.7 nos asegura la existencia de un $\delta > 0$ tal que

$$\int_E g\, d\mu \;<\; \varepsilon$$

para cualquier conjunto medible $E \subseteq X$ con $\mu(E) < \delta$. Se sigue entonces que para cualquier $f \in \mathcal{U}$,

$$\int_E |f|\, d\mu \;\leq\; \int_E g\, d\mu \;<\; \varepsilon$$

para todo conjunto medible $E \subseteq X$ satisfaciendo $\mu(E) < \delta$. Esto prueba que la familia $\mathcal{U}$ es uniformemente absolutamente continua. Finalmente, puesto que es válida la desigualdad $\sup_{f \in \mathcal{U}} \| f \|_1 \leq \| g \|_1$, entonces el Criterio de la Vallée-Poussin, Teorema 10.2.22, finaliza la prueba de (b). ■

Ejemplo 10.2.2. *Si* $\mathcal{U} \subseteq L_1([0,1])$ *es* ***uniformemente integrable****, entonces*

$$\sup_{f \in \mathcal{U}} \int_{[0,1]} |f| \ln(|f|)\, d\mu \;<\; +\infty. \tag{1}$$

Prueba. En efecto, la función $\Phi : [0,+\infty) \to [0,+\infty)$ definida por

$$\Phi(t) \;=\; \begin{cases} t\ln(t) & \text{si } \; t \geq 1, \\ 0 & \text{si } \; 0 \leq t < 1, \end{cases}$$

es creciente y satisface $\text{lím}_{t\to\infty} \Phi(t)/t = +\infty$. Se sigue del Criterio de la Vallée-Poussin que (1) se cumple. ■

Es importante destacar que la función Φ definida en el Criterio de la Vallée-Poussin es **convexa**. En efecto, sean $x_1, x_2 \geq 0$. Entonces

$$\begin{aligned}
\Phi\Big(\frac{1}{2}x_1 + \frac{1}{2}x_2\Big) \;&=\; \int_0^{\frac{1}{2}(x_1+x_2)} g\, dt \;=\; \int_0^{x_1} g\, dt \;+\; \int_{x_1}^{\frac{1}{2}(x_1+x_2)} g\, dt \\
&\leq\; \int_0^{x_1} g\, dt \;+\; \frac{1}{2}\int_{x_1}^{\frac{1}{2}(x_1+x_2)} g\, dt \;+\; \frac{1}{2}\int_{\frac{1}{2}(x_1+x_2)}^{x_2} g\, dt \\
&=\; \frac{1}{2}\int_0^{x_1} g\, dt \;+\; \frac{1}{2}\int_0^{x_2} g\, dt \\
&=\; \frac{1}{2}\Phi(x_1) \;+\; \frac{1}{2}\Phi(x_2).
\end{aligned}$$

Denote por

$$\mathfrak{S} \;=\; \left\{\Phi : [0,+\infty) \to [0,+\infty) \;:\; \Phi \text{ es convexa} \quad \text{y} \quad \lim_{t\to\infty} \frac{\Phi(t)}{t} = +\infty\right\}$$

y considere, para cada $\Phi \in \mathfrak{S}$, el conjunto

$$\mathrm{L}^{\Phi}(\mu) \;=\; \left\{f \in \mathcal{F}_{\mu}(X) \;:\; \int_X \Phi(|f|)\, d\mu < +\infty\right\}.$$

Un hecho que resulta interesante es el siguiente resultado.

Corolario 10.2.24. $\mathrm{L}_1(X,\mu) \;=\; \bigcup_{\Phi\in\mathfrak{S}} \mathrm{L}^{\Phi}(\mu).$

Prueba. Sea $f \in \bigcup_{\Phi\in\mathfrak{S}} \mathrm{L}^{\Phi}(\mu)$. Entonces existe $\Phi \in \mathfrak{S}$ tal que $\int_X \Phi(|f|)\, d\mu < +\infty$. Ahora bien, como Φ es convexa, existen $a, b \in \mathbb{R}$ tal que $\Phi(x) \geq ax + b$ para todo $x \geq 0$. De esto se sigue que

$$a\int_X |f|\, d\mu \,+\, b \;=\; \int_X (a|f| + b)\, d\mu \;\leq\; \int_X \Phi(|f|)\, d\mu \;<\; +\infty,$$

de donde se concluye que $\int_X |f|\, d\mu < +\infty$, es decir, $f \in \mathrm{L}_1(X,\mu)$. Así, $\bigcup_{\Phi\in\mathfrak{S}} \mathrm{L}^{\Phi}(\mu) \subseteq \mathrm{L}_1(X,\mu)$. Para demostrar la otra inclusión, suponga que $f \in \mathrm{L}_1(X,\mu)$. Por el Teorema 10.2.7, $\{f\}$ es uniformemente integrable y, así, por el Criterio de la Vallée-Poussin, existe $\Phi_0 \in \mathfrak{S}$ tal que $\int_X \Phi_0(|f|)\, d\mu < +\infty$. Esto nos dice que $f \in \bigcup_{\Phi\in\mathfrak{S}} \mathrm{L}^{\Phi}(\mu)$ y termina la prueba. ■

Una desigualdad que es importante por sus muchas aplicaciones en el Análisis es la siguiente:

Corolario 10.2.25 (Desigualdad de Jensen). *Si $\Phi \in \mathfrak{S}$, entonces para toda $f \in \mathrm{L}^{\Phi}(\mu)$*

$$\Phi\left(\int_X |f|\, d\mu\right) \;\leq\; \int_X \Phi(|f|)\, d\mu.$$

Prueba. Si $f \in \mathrm{L}^{\Phi}(\mu)$, se sigue del resultado anterior que $f \in \mathrm{L}_1(X,\mu)$. Pongamos $x_0 = \int_X |f|\, d\mu \in \mathbb{R}^+$. Como Φ es convexa, el Lema 4.1.2, página 194, nos garantiza la existencia de un número real a tal que

$$\Phi(x) \geq \Phi(x_0) + a(x - x_0).$$

Sustituyendo x por $|f(t)|$ en la desigualdad anterior e integrando, resulta que

$$\Phi\left(\int_X |f|\, d\mu\right) \;\leq\; \int_X \Phi(|f|)\, d\mu,$$

ya que $\int_X (x - x_0)\, d\mu = \int_X |f|\, d\mu - \int_X |f|\, d\mu = 0$. La prueba es completa. ■

Nota Adicional 10.2.9 Si $(X, \|\cdot\|)$ es un espacio normado sobre $\mathbb{K}$, entonces su norma $\|\cdot\|$ genera una topología natural, denotada por $\mathcal{J}_{\|\cdot\|}$, bajo la cual $(X, \mathcal{J}_{\|\cdot\|})$ es un espacio lineal topológico localmente convexo. Existen, sin embargo, otras topologías interesantes sobre X las cuales permiten obtener información valiosa sobre la estructura de esos espacios. Una de tales topologías, la llamada **topología débil**, se genera através de *todos* los funcionales lineales continuos sobre X. Si denotamos por X^* el dual de X, es decir, el espacio de Banach constituido por todos los funcionales lineales continuos $x^* : X \to \mathbb{K}$ provisto de la norma dual, es to es,

$$\| x^* \| = \sup_{\| x \| \leq 1} |x^*(x)|,$$

entonces la **topología débil** sobre X, denotada por $\sigma(X, X^*)$, es la topología inicial asociada a la familia X^*, esto, por supuesto, significa que $\sigma(X, X^*)$ es la topología más débil bajo la cual cada funcional lineal $x^* : X \to \mathbb{K}$ es $\sigma(X, X^*)$-continuo. No es difícil establecer que $(X, \sigma(X, X^*))$ es un espacio vectorial topológico localmente convexo Hausdorff y que cada entorno básico del 0 en $(X, \sigma(X, X^*))$ es de la forma

$$V(0; x_1^*, \ldots, x_n^*, \varepsilon) = \bigcap_{i=1}^{n} \{x \in X \,:\, |x_i^*(x)| < \varepsilon\}$$

donde $x_1^*, \ldots, x_n^* \in X^*$ para todo $n \in \mathbb{N}$ y $\varepsilon > 0$.

En el caso especial de nuestro espacio de Banach $(L_1(X, \mu), \| \cdot \|_1)$ se tiene el siguiente resultado:

Teorema 10.2.26 (Dunford-Pettis). *Un subconjunto K de $L_1(X, \mu)$ es **relativamente débilmente compacto** si, y sólo si K es **uniformemente integrable**.*

Una elegante y hermosa prueba de este resultado se puede ver en el artículo de J. Diestel [44].

10.2.5. Los Teoremas de Convergencia de Vitali

El Teorema de la Convergencia Dominada de Lebesgue es un resultado formidable, poderoso y con amplias aplicaciones. Pero, ¿qué ocurre si no se dispone de una función dominante g en el enunciado de dicho resultado? Pues bien, el Teorema de la Convergencia de Vitali es un buen sustituto a dicho teorema siempre y cuando la medida del espacio sea finita.

Como ya ha sido observado, el Teorema de la Convergencia Dominada de Lebesgue da condiciones *suficientes* para el intercambio entre la integral y el límite. La siguiente pregunta es natural: ¿también existen condiciones *necesarias*? La respuesta es afirmativa y fue dada a conocer por Giuseppe Vitali [135] un año antes de la aparición del Teorema de la Convergencia Dominada de Lebesgue. Este resultado de Vitali es, en varios aspectos, más fuerte que los Teoremas de Convergencias ya vistos y, de hecho, cuando $\mu(X)$ es finita, el Teorema de la Convergencia de Vitali implica cada uno de los anteriores.

Una de las razones fundamentales para considerar familias de funciones uniformemente integrables es que la convergencia puntual de sucesiones de funciones uniformemente integrables definidas sobre un conjunto medible de medida finita permite, sin más restricción, el intercambio del límite. En este sentido, y mirando la parte (b) del Teorema 10.2.23, vemos que el Teorema de la Convergencia de Vitali es de mayor alcance que el Teorema de la Convergencia Dominada de Lebesgue.

Teorema 10.2.27 (Convergencia de Vitali). *Sea X un conjunto medible con $\mu(X) < +\infty$ y suponga que $(f_n)_{n=1}^{\infty}$ es una sucesión en $\mathrm{L}_1(X, \mu)$ tal que $f_n \to f$* ***casi-siempre*** *para alguna función medible $f : X \to \overline{\mathbb{R}}$. Las siguientes condiciones son equivalentes:*

(1) $f \in \mathrm{L}_1(X, \mu)$ y $\lim\limits_{n\to\infty} \| f_n - f \|_1 = 0$.

(2) $(f_n)_{n=1}^{\infty}$ es ***uniformemente integrable****.*

Prueba. $(1) \Rightarrow (2)$. Suponga que (1) se cumple. Para demostrar que $(f_n)_{n=1}^{\infty}$ es uniformemente integrable observe, en primer lugar, que como la sucesión $(f_n)_{n=1}^{\infty}$ converge (en la norma), ella es acotada y, en consecuencia, el conjunto $\mathcal{U} = \{f_1, f_2, \dots\}$ es acotado. Sea $\varepsilon > 0$ y escojamos, por (1), un $N \in \mathbb{N}$ tal que

$$\int_X |f_n - f| \, d\mu \;=\; \| f_n - f \|_1 \;<\; \varepsilon/2 \quad \text{para todo} \quad n \geq N.$$

Por el Teorema 10.2.23 sabemos que el conjunto $\{f, f_1, \dots, f_{N-1}\}$ es uniformemente integrable, en particular, uniformemente absolutamente continuo. Esto significa que existe un $\delta > 0$ con la siguiente propiedad: para cualquier conjunto medible $E \subseteq X$ con $\mu(E) < \delta$, se cumple que

$$\int_E |f| \, d\mu \;<\; \varepsilon/2 \qquad \text{y} \qquad \int_E |f_n| \, d\mu \;<\; \varepsilon/2 \quad \text{para} \quad n = 1, \dots, N-1.$$

Finalmente, si $n \geq N$ y E es un conjunto medible arbitrario con $\mu(E) < \delta$, entonces

$$\begin{aligned} \int_E |f_n| \, d\mu \;&\leq\; \int_E |f| \, d\mu \;+\; \int_E |f_n - f| \, d\mu \\ &\leq\; \int_E |f| \, d\mu \;+\; \int_X |f_n - f| \, d\mu \\ &<\; \varepsilon/2 + \varepsilon/2 \;=\; \varepsilon. \end{aligned}$$

Esto prueba que $\mathcal{U} = \{f_n : n \in \mathbb{N}\}$ es uniformemente absolutamente continuo y entonces, gracias al Teorema 10.2.22, $\mathcal{U} = \{f_n : n \in \mathbb{N}\}$ es uniformemente integrable

$(2) \Rightarrow (1)$. Suponga que (2) se cumple y sea $\varepsilon > 0$ elegido de modo arbitrario. Por la uniforme integrabilidad de la sucesión $(f_n)_{n=1}^{\infty}$, existe un $\delta > 0$ tal que si $B \subseteq X$ es medible con $\mu(B) < \delta$, entonces

$$\int_B |f_n| \, d\mu \;<\; \varepsilon/3, \quad \text{para todo} \quad n \geq 1. \tag{β_1}$$

La desigualdad anterior combinada con el Lema de Fatou nos revela que

$$\int_B |f| \, d\mu \;\leq\; \liminf_{n\to\infty} \int_B |f_n| \, d\mu \;\leq\; \varepsilon/3. \tag{β_2}$$

para cualquier conjunto medible B con $\mu(B) < \delta$. Puesto que $\mu(X) < +\infty$, podemos usar el Teorema de Severini-Egoroff para obtener un conjunto medible $B \subseteq X$ con $\mu(B) < \delta$ tal que $f_n \to f$ uniformemente sobre $X \setminus B$. De aquí se sigue que f es integrable sobre $X \setminus B$, es decir, $\int_{X \setminus B} |f| \, d\mu < +\infty$. Se sigue entonces de (β_2) y del hecho de que

$$\int_X |f| \, d\mu \;=\; \int_B |f| \, d\mu + \int_{X \setminus B} |f| \, d\mu,$$

que $f \in L_1(X,\mu)$. Seleccione, aprovechando el hecho de que $f_n \to f$ uniformemente sobre $X \setminus B$, un $N \in \mathbb{N}$ de modo que la desigualdad

$$|f_n(x) - f(x)| \; < \; \frac{\varepsilon}{3\mu(X \setminus B)} \quad \text{se cumpla para todo } x \in X \setminus B \text{ y todo } n \geq N. \qquad (\beta_3)$$

Combinando la desigualdades (β_1), (β_2) y (β_3) se obtiene que

$$\begin{aligned} \| f_n - f \|_1 &= \int_X |f_n - f| \, d\mu \\ &= \int_{X \setminus B} |f_n - f| \, d\mu + \int_B |f_n - f| \, d\mu \\ &\leq \int_{X \setminus B} |f_n - f| \, d\mu + \int_B |f_n| \, d\mu + \int_B |f| \, d\mu \\ &< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon \end{aligned}$$

para todo $n \geq N$. Con esto queda establecido que $\lim\limits_{n\to\infty} \| f_n - f \|_1 = 0$ y (1) queda demostrado. Fin de la prueba. ∎

Nota Adicional 10.2.10 Dos observaciones son pertinentes respecto al Teorema de Convergencia de Vitali. En primer lugar, la implicación $(2) \Rightarrow (1)$ nos indica que el Teorema de la Convergencia Dominada de Lebesgue es consecuencia inmediata del Teorema de la Convergencia de Vitali cuando $\mu(X) < +\infty$. En efecto, por el Teorema 10.2.23 (a) sabemos que toda sucesión en $L_1(X,\mu)$, digamos $(f_n)_{n=1}^{\infty}$, dominada por una función integrable no-negativa es uniformemente integrable. Por consiguiente, si dicha sucesión también converge casi-siempre a una función medible f, entonces el Teorema de Convergencia de Vitali nos garantiza que f es integrable y

$$\lim_{n\to\infty} \int_X f_n \, d\mu = \int_X f \, d\mu.$$

La forma original y más conocida de expresar el Teorema de la Convergencia de Vitali es como sigue:

Teorema 10.2.28 (Convergencia de Vitali). *Sea X un conjunto medible con $\mu(X) < +\infty$ y suponga que $(f_n)_{n=1}^{\infty}$ es una sucesión en $L_1(X,\mu)$ que es* ***uniformemente integrable****. Si $f_n \to f$* ***casi-siempre*** *para alguna función medible $f : X \to \overline{\mathbb{R}}$, entonces $f \in L_1(X,\mu)$ y*

$$\lim_{n\to\infty} \int_X f_n \, d\mu = \int_X f \, d\mu.$$

Es importante observar que la conclusión del resultado anterior deja de ser cierta si $\mu(X) = +\infty$. En efecto, por cada número natural n, considere la función $f_n : \mathbb{R} \to \mathbb{R}$ definida por

$$f_n = \chi_{[n,n+1]}$$

y sea $f \equiv 0$ sobre $\mathbb{R}$. Entonces $(f_n)_{n=1}^{\infty}$ es uniformemente integrable sobre $\mathbb{R}$ y converge puntualmente a f sobre $\mathbb{R}$. Sin embargo,

$$\lim_{n\to\infty} \int_{\mathbb{R}} f_n \, d\mu = 1 \neq 0 = \int_{\mathbb{R}} \lim_{n\to\infty} f_n \, d\mu = \int_{\mathbb{R}} f \, d\mu.$$

La segunda observación tiene que ver con la demostración de la implicación $(1) \Rightarrow (2)$ del Teorema 10.2.27. En efecto, en dicha demostración no se utilizó en ningún momento la hipótesis de que $\mu(X)$ fuese finita, así como tampoco que la sucesión $(f_n)_{n=1}^{\infty}$ converge casi-siempre a f. Dicha convergencia sólo se requirió para demostrar la implicación $(2) \Rightarrow (1)$, de modo que el siguiente resultado es válido:

Corolario 10.2.29. *Sea X un conjunto medible de medida arbitraria. Sea $(f_n)_{n=1}^{\infty}$ una sucesión de funciones en $\mathrm{L}_1(X,\mu)$ tal que $\lim_{n\to\infty} \| f_n - f \|_1 = 0$ para alguna función $f \in \mathrm{L}_1(X,\mu)$. Entonces $(f_n)_{n=1}^{\infty}$ es **uniformemente integrable**.*

Podemos afinar un poco más el Teorema de la Convergencia de Vitali en el siguiente sentido: como en espacios de medida finita la convergencia puntual de sucesiones de funciones medibles implica la convergencia en medida, Teorema 7.2.25, página 395, el siguiente resultado es más general que el anterior.

Teorema 10.2.30 (Convergencia de Vitali). *Sea X un conjunto medible con $\mu(X) < +\infty$. Sean $f, f_n \in \mathrm{L}_1(X,\mu)$ para todo $n \geq 1$. Las siguientes condiciones son equivalentes:*

(1) $\lim_{n\to\infty} \| f_n - f \|_1 = 0$.

*(2) (i) $f_n \to f$ **en medida** y*

*(ii) $\mathcal{U} = \{f_n : n \in \mathbb{N}\}$ es **uniformemente integrable**.*

Prueba. $(1) \Rightarrow (2)$ sigue del Corolario 10.2.5, página 581 y el Teorema 10.2.27.

$(2) \Rightarrow (1)$ Suponga ahora que (2) se cumple y sea $\varepsilon > 0$. Como $\mathcal{U} = \{f_n : n \in \mathbb{N}\}$ es uniformemente integrable, también lo es $\mathcal{U} \cup \{f\}$, y entonces existe un $\delta > 0$ tal que para cualquier conjunto medible $E \subseteq [0,1]$ con $\mu(E) < \delta$, se cumple que

$$\int_E |f|\, d\mu \; < \; \varepsilon/3 \qquad \text{y} \qquad \sup_{n\geq 1} \int_E |f_n|\, d\mu \; < \; \varepsilon/3. \tag{$\star$}$$

Por otro lado, como $f_n \to f$ en medida, tenemos que

$$\lim_{n\to\infty} \mu([\,|f_n - f| \; > \; \varepsilon\,]) \; = \; 0,$$

y, por consiguiente, podemos elegir un $N \in \mathbb{N}$ tal que

$$\mu([\,|f_n - f| \; > \; \varepsilon\,]) \; < \; \delta \quad \text{para todo} \quad n \geq N.$$

Se sigue de $(\star)$ que

$$\int_{[\,|f_n - f| > \varepsilon\,]} |f|\, d\mu \; < \; \varepsilon/3 \qquad \text{y} \qquad \int_{[\,|f_n - f| > \varepsilon\,]} |f_n|\, d\mu \; < \; \varepsilon/3 \quad \text{para todo} \quad n \geq N$$

y, en consecuencia, si $n \geq N$ resulta que

$$\begin{aligned}
\int_X |f_n - f|\, d\mu \; &= \; \int_{[\,|f_n - f| \leq \varepsilon\,]} |f_n - f|\, d\mu \; + \; \int_{[\,|f_n - f| > \varepsilon\,]} |f_n - f|\, d\mu \\
&\leq \; \int_{[\,|f_n - f| \leq \varepsilon\,]} |f_n - f|\, d\mu \; + \; \int_{[\,|f_n - f| > \varepsilon\,]} |f_n|\, d\mu \; + \; \int_{[\,|f_n - f| > \varepsilon\,]} |f|\, d\mu \\
&< \; \varepsilon/3 \; + \; \varepsilon/3 \; + \; \varepsilon/3 \; = \; \varepsilon.
\end{aligned}$$

Esto nos dice que $\lim_{n\to\infty} \| f_n - f \|_1 = 0$ y termina la prueba. ■

Otra forma de expresar el resultado anterior y que puede resultar más conveniente y práctico a otros es que en presencia de una sucesión uniformemente integrable la convergencia en norma y la convergencia en medida son indistinguibles.

Corolario 10.2.31 (Convergencia de Vitali). *Sea X un conjunto medible con $\mu(X) < +\infty$. Sean $f, f_n \in L_1(X,\mu)$ para todo $n \geq 1$. Suponga que el conjunto $\mathfrak{U} = \{f_1, f_2, \ldots\}$ es **uniformemente integrable**. Las siguientes condiciones son equivalentes:*

(1) $\lim_{n\to\infty} \| f_n - f \|_1 = 0$.

(2) $\lim_{n\to\infty} f_n = f$ ***en medida***.

De nuevo, fijemos una sucesión $(f_n)_{n=1}^{\infty}$ en $L_1(X,\mu)$, donde X es cualquier subconjunto medible de $\mathbb{R}$. Sabemos que si $\lim_{n\to\infty} \| f_n - f \|_1 = 0$ para alguna función $f \in L_1(X,\mu)$, entonces

$$\lim_{n\to\infty} \int_X f_n \, d\mu \; = \; \int_X f \, d\mu.$$

Sin embargo, la dirección opuesta no es válida. El siguiente resultado constituye, bajo ciertas restricciones, la validez de dicha dirección.

Corolario 10.2.32 (Vitali-Scheffé). *Sea X un conjunto medible con $\mu(X) < +\infty$. Suponga que $f, f_n \in L_1(X,\mu)$ con $f, f_n \geq 0$ para todo $n \geq 1$ y, además, que $f_n \to f$ **casi-siempre**. Las siguientes condiciones son equivalentes:*

(1) $\lim_{n\to\infty} \int_X f_n \, d\mu \; = \; \int_X f \, d\mu$.

(2) $\lim_{n\to\infty} \| f_n - f \|_1 \; = \; 0$.

(3) $(f_n)_{n=1}^{\infty}$ *es **uniformemente integrable***.

Prueba. (1) $\Rightarrow$ (2). Observe que, para cada $n \in \mathbb{N}$,

$$f_n + f \; = \; \min\{f_n, f\} \; + \; \max\{f_n, f\}. \tag{II}$$

Puesto que $\min\{f_n, f\} \leq f$ y $\min\{f_n, f\} \to f$ casi-siempre, el Teorema de la Convergencia Dominada de Lebesgue nos asegura que

$$\lim_{n\to\infty} \int_X \min\{f_n, f\} \, d\mu \; = \; \int_X f \, d\mu.$$

Integrando a ambos lados de (II) y luego tomando el límite cuando $n \to \infty$ se obtiene, después de aplicar (1) y la igualdad anterior, que

$$\lim_{n\to\infty} \int_X \max\{f_n, f\} \, d\mu \; = \; \int_X f \, d\mu.$$

Por otro lado, como

$$|f_n - f| \; = \; \max\{f_n, f\} \; - \; \min\{f_n, f\},$$

resulta, después de integrar a ambos de esta última igualdad y posteriormente tomar el límite cuando $n \to \infty$, que

$$\lim_{n\to\infty} \| f_n - f \|_1 \;=\; \lim_{n\to\infty} \int_X |f_n - f|\, d\mu \;=\; 0.$$

$(2) \Rightarrow (3)$ Puesto que estamos asumiendo que $f \in \mathrm{L}_1(X,\mu)$, el Teorema de Convergencia de Vitali, Teorema 10.2.27, nos muestra que $(f_n)_{n=1}^{\infty}$ es uniformemente integrable.

$(3) \Rightarrow (1)$ es consecuencia del Teorema de Convergencia de Vitali, Teorema 10.2.27. ■

Ya hemos observado que el Teorema de Convergencia de Vitali impone una limitación para su validez: que el conjunto medible X sea de medida finita. ¿Bajo qué condiciones, si es que existen, se puede obtener un teorema tipo-Vitali donde la medida del conjunto X sea infinita? La siguiente propiedad, la cual es válida sobre cualquier conjunto de medida infinita y cuya motivación sigue de del Teorema 10.2.11, página 584, constituye un buen aliado de la integrabilidad uniforme para garantizar el paso al límite bajo el signo de la integral de sucesiones de funciones integrables sobre dominios de medida infinita.

Definición 10.2.33. *Sea X un conjunto medible con $\mu(X) = +\infty$. Un conjunto no vacío $\mathfrak{U} \subseteq \mathrm{L}_1(X,\mu)$ se dice que es* ***uniformemente Vitali-pequeño*** *si, para cada $\varepsilon > 0$, existe un conjunto medible $E_0 \subseteq X$ con $\mu(E_0) < +\infty$ tal que*

$$\int_{X\setminus E_0} |f|\, d\mu \;<\; \varepsilon \quad \text{para toda} \quad f \in \mathfrak{U}.$$

En la literatura sobre el tema, a los conjuntos uniformemente Vitali-pequeños también se les suelen llamar **conjuntos tensos** (en inglés, **tight**). Similar al caso del Teorema 10.2.23 se tiene que:

(a_1) *Si $\mathfrak{U} = \{f_1, \dots, f_n\}$ es un* ***conjunto finito*** *en $\mathrm{L}_1(X,\mu)$, entonces $\mathfrak{U}$ es* ***uniformemente Vitali-pequeño***.

En efecto, fijemos un $\varepsilon > 0$. Por el Teorema 10.2.11 podemos elegir, por cada $j = 1, \dots, n$, un conjunto medible $E_j \subseteq X$ con $\mu(E_j) < +\infty$ tal que

$$\int_{\mathbb{R}\setminus E_j} |f_j|\, d\mu \;<\; \varepsilon.$$

Sea $E_0 = E_1 \cup \cdots \cup E_n$. Entonces E_0 es medible, $\mu(E_0) < +\infty$ y como $X \setminus E_0 \subseteq \mathbb{R} \setminus E_j$ para cada $j = 1, \dots, n$, se tiene que

$$\int_{X\setminus E_0} |f_j|\, d\mu \;\le\; \int_{\mathbb{R}\setminus E_j} |f_j|\, d\mu \;<\; \varepsilon,$$

ya que $E \to \int_E |f_j|\, d\mu$ es una medida para cada $j = 1, \dots, n$.

(b_1) *Sea $g \in \mathrm{L}_1(X,\mu)$ con $g \geq 0$. Si $(f_n)_{n=1}^{\infty}$ es una sucesión de funciones medibles definidas sobre $\mathbb{R}$ tal que $|f_n| \leq g$ para todo $n \geq 1$, entonces $\mathfrak{U} = \{f_1, f_2, \dots\}$ es* ***uniformemente Vitali-pequeño***.

La demostración de este hecho es sencilla y, por lo tanto, se omite.

El siguiente resultado *es la versión del Teorema de la Convergencia de Vitali sobre conjuntos de medida infinita*. En este caso, uno reemplaza la convergencia en medida de la sucesión por la condición de que dicha sucesión sea uniformemente Vitali-pequeño en el Teorema 10.2.30 para obtener:

Teorema 10.2.34 (Convergencia de Vitali II). *Sea X un conjunto medible con $\mu(X) = +\infty$. Sean $f, f_n \in L_1(X,\mu)$ para todo $n \geq 1$ y suponga que $f_n \to f$ **casi-siempre**. Las siguientes condiciones son equivalentes:*

(1) $\lim\limits_{n\to\infty} \| f_n - f \|_1 = 0.$

(2) $\mathcal{U} = \{f_1, f_2, \ldots\}$ *es **uniformemente Vitali-pequeño** y **uniformemente integrable**.*

Prueba. (1) $\Rightarrow$ (2). Suponga que (1) se cumple. En vista del Teorema 10.2.27 sólo es suficiente demostrar que $\mathcal{U}$ es uniformemente Vitali-pequeño. Sea $\varepsilon > 0$ y seleccione un $N \in \mathbb{N}$ tal que

$$\| f_n - f \|_1 \;=\; \int_X |f_n - f|\, d\mu \;<\; \varepsilon \quad \text{para todo} \quad n \geq N. \tag{I}$$

Por (a_1) sabemos que $\{f, f_1, \ldots, f_{N-1}\}$ es uniformemente Vitali-pequeño, por lo que existe un conjunto medible E_0 con $\mu(E_0) < +\infty$ tal que

$$\int_{X\setminus E_0} |f|\, d\mu \;<\; \varepsilon \qquad \text{y} \qquad \int_{X\setminus E_0} |f_j|\, d\mu \;<\; \varepsilon \quad \text{para} \quad j = 1, \ldots, N-1.$$

Por supuesto, esto último combinado con (I) nos dice que

$$\begin{aligned}
\int_{X\setminus E_0} |f_n|\, d\mu \;&\leq\; \int_{X\setminus E_0} |f_n - f|\, d\mu \;+\; \int_{X\setminus E_0} |f|\, d\mu \\
&\leq\; \int_X |f_n - f|\, d\mu \;+\; \int_{X\setminus E_0} |f|\, d\mu \\
&<\; 2\varepsilon \quad \text{para todo} \quad n \geq N.
\end{aligned}$$

Esto prueba que $\mathcal{U} = \{f_1, f_2, \ldots\}$ es uniformemente Vitali-pequeño y termina la prueba de la implicación (1) $\Rightarrow$ (2).

(2) $\Rightarrow$ (1). Suponga que (2) se cumple y sea $\varepsilon > 0$. Puesto que $\mathcal{U} = \{f_1, f_2, \ldots\}$ es uniformemente absolutamente continuo, también lo es $\mathcal{U} - f = \{f_n - f : n = 1, 2, \ldots\}$ (véase Ejercicios 10.4.4). Seleccione entonces un $\delta > 0$ tal que para cualquier conjunto medible B con $\mu(B) < \delta$ se cumpla que

$$\int_B |f_n - f|\, d\mu \;<\; \varepsilon/3 \quad \text{para todo } n \geq 1.$$

También, como $\mathcal{U} - f$ es uniformemente Vitali-pequeño, existe un conjunto medible E_0 con $\mu(E_0) < +\infty$ tal que

$$\int_{X\setminus E_0} |f_n - f|\, d\mu \;<\; \varepsilon/3 \quad \text{para todo} \quad n \geq 1.$$

Finalmente, puesto que $f_n \to f$ c.s. sobre E_0, el Teorema de Severini-Egoroff nos garantiza la existencia de un conjunto medible $B \subseteq E_0$ tal que

$$\mu(B) \;<\; \delta \qquad \text{y} \qquad f_n \to f \quad \text{uniformemente sobre} \quad E_0 \setminus B.$$

Seleccione un $N \in \mathbb{N}$ tal que $|f_n(x) - f(x)| < \varepsilon/3\mu(E_0 \setminus B)$ para todo $n \geq N$ y todo $x \in E_0 \setminus B$. Combinando todos estos hechos resulta que si $n \geq N$, entonces

$$\begin{aligned}\int_X |f_n - f|\,d\mu &= \int_{X\setminus E_0} |f_n - f|\,d\mu + \int_{E_0} |f_n - f|\,d\mu \\ &= \int_{X\setminus E_0} |f_n - f|\,d\mu + \int_{E_0\setminus B} |f_n - f|\,d\mu + \int_B |f_n - f|\,d\mu \\ &< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon.\end{aligned}$$

Con esto queda demostrado que $\lim_{n\to\infty} \| f_n - f \|_1 = 0$ y termina la prueba. ■

10.2.6. El Teorema Fundamental del Cálculo

El Primer y Segundo Teorema Fundamental del Cálculo para la integral de Riemann establecen que cada $f \in C([a,b])$ posee una primitiva F y se cumple que

$$F(x) = F(a) + \int_a^x F'\,dt, \quad \text{para todo} \quad x \in [a,b],$$

es decir, uno puede recuperar *primitivas de funciones continuas* por medio de la integral de Riemann. De hecho, la integral de Riemann es capaz de hacer algo más interesante: ella *recaptura la primitiva de cualquier derivada acotada la cual es continua casi-siempre*. Estos resultados permiten que uno pueda formularse una pregunta más general: ¿se puede recuperar cualquier función derivada que sea Riemann integrable?, esto es, si $g : [a,b] \to \mathbb{R}$ es una función diferenciable con g' Riemann integrable, ¿será cierto que

$$g(x) = \int_a^x g'\,dt\,?$$

La función de Cantor φ_Γ provee una respuesta negativa a tal interrogante. En efecto, sabemos que φ_Γ es Riemann integrable y como ella es creciente, el Teorema de Diferenciabilidad de Lebesgue nos asegura que $\varphi'_\Gamma = 0$ $\mu -$c.s. Finalmente, el Teorema de Vitali-Lebesgue nos revela que

$$\varphi_\Gamma(x) \neq 0 = \int_0^x \varphi'_\Gamma\,d\mu$$

para cualquier $x \in [0,1]$. El problema con la función de Cantor es que ella no es absolutamente continua y, en consecuencia, la integral de Riemann no es lo suficientemente buena para recapturar derivadas. La integral de Lebesgue es, una vez más y en este aspecto, mejor que la integral de Riemann ya que, como veremos en esta sección, bajo ciertas hipótesis adicionales se pueden recuperar derivadas de funciones Lebesgue integrables.

Los preparativos para la agenda del Teorema Fundamental del Cálculo para la integral de Lebesgue comienzan con un hecho simple, pero importante.

Lema 10.2.35. *Si $f : [a,b] \to \mathbb{R}$ es una **función creciente** sobre $[a,b]$, entonces f' es **Lebesgue integrable** sobre $[a,b]$ y*

$$\int_a^b f'\,d\mu \leq f(b) - f(a). \tag{1}$$

*En particular, si $f:[a,b]\to\mathbb{R}$ es de **variación acotada** sobre $[a,b]$, entonces $f'\in L_1([a,b])$ y*

$$\int_a^b |f'|\,d\mu \;\le\; V(f,[a,b]).$$

Prueba. Tal como se ha hecho anteriormente, extienda f al intervalo $[a,b+1]$ definiendo $f(x)=f(b)$ para todo $x\in[b,b+1]$. Por el Teorema de Diferenciabilidad de Lebesgue sabemos que f' existe casi-siempre y que $f'(x)\ge 0$ para todo $x\in[a,b]$. Si definimos

$$f_n(x) \;=\; n\big(f(x+1/n)-f(x)\big), \quad \text{para todo} \;\; x\in[a,b],$$

se tiene que $f'=\lim_{n\to\infty} f_n$ casi-siempre. Una aplicación del Lema de Fatou nos revela entonces que

$$\begin{aligned}
\int_a^b f'\,d\mu \;&=\; \int_a^b \lim_{n\to\infty} f_n\,d\mu\\
&\le\; \liminf_{n\to\infty}\Big(n\int_a^b f(x+1/n)\,d\mu \;-\; n\int_a^b f(x)\,d\mu\Big)\\
&=\; \liminf_{n\to\infty}\Big(n\int_{a+1/n}^{b+1/n} f(x)\,d\mu \;-\; n\int_a^b f(x)\,d\mu\Big)\\
&=\; \liminf_{n\to\infty}\Big(n\int_b^{b+1/n} f(x)\,d\mu \;-\; n\int_a^{a+1/n} f(x)\,d\mu\Big)\\
&\le\; f(b)\;-\;f(a).
\end{aligned}$$

Observe que el "cambio de variable" aplicado en el paso anterior queda justificado por el hecho de que, al ser f creciente, cada una de las integrales que aparecen en dicha desigualdades son integrales de Riemann. Esto termina la prueba de la primera parte.

Suponga ahora que $f\in BV([a,b])$. Por el Teorema de Jordan, Teorema 9.1.22, f se puede expresar como la diferencia de dos funciones crecientes, digamos, $f=f_1-f_2$, donde $f_1=V_f$ y $f_2=V_f-f$. Por lo tanto, $f'=f_1'-f_2'$ existe casi-siempre y es, por la primera parte, Lebesgue integrable. Más aun, de las propiedades de la función variación, Corolario 9.1.21, página 467, sabemos que si $x<y$, entonces

$$|f(y)-f(x)| \;\le\; V(f,[x,y]) \;=\; V_f(y)-V_f(x)$$

de donde se sigue que $|f'|\le V_f'$ casi-siempre y, en consecuencia, por (1) se concluye que

$$\int_a^b |f'|\,d\mu \;\le\; \int_a^b V_f'\,d\mu \;\le\; V_f(b)-V_f(a) \;=\; V(f,[a,b]).$$

La prueba es completa. ■

Recordemos que si $f:[a,b]\to\mathbb{R}$ es una función arbitraria, entonces el Lema 9.1.57, página 507, nos asegura que si E es un subconjunto medible de $[a,b]$ tal que $|f'|\le M$ sobre E, entonces $\mu^*(f(E))\le M\mu(E)$. El siguiente resultado es, a todas luces, más general y si se combina con el Lema 10.2.35 puede ser usado para dar otra demostración del Teorema de Banach-Zarecki.

Lema 10.2.36. *Sea $f : [a,b] \to \mathbb{R}$ una **función medible**. Suponga que E es un subconjunto medible de $[a,b]$ tal que $f'(x)$ es **finita** para cualquier $x \in E$. Entonces*

$$\mu^*(f(E)) \leq \int_E |f'|\, d\mu.$$

Prueba. Observe que la existencia de f' sobre el conjunto medible E implica que f' es medible según Lebesgue. Para demostrar nuestro lema, vamos a considerar dos casos según sea f' acotada o no sobre E. Suponga, en primer lugar, que f' es acotada sobre E y pongamos $g = |f'|$. Sea $N \in \mathbb{N}$ tal que $g(x) \leq N$ para todo $x \in E$. Para cada entero $n \geq 1$ y cada $k = 1, \ldots, N2^n$, considere el conjunto

$$\begin{aligned} E_{n,k} &= g^{-1}\left(\left[\frac{k-1}{2^n}, \frac{k}{2^n}\right)\right) \cap E \\ &= \left\{x \in E : \frac{k-1}{2^n} \leq |f'(x)| < \frac{k}{2^n}\right\}. \end{aligned}$$

Defina ahora

$$g_n = \sum_{k=1}^{N2^n} \frac{k-1}{2^n} \chi_{E_{n,k}}.$$

Entonces $(g_n)_{n=1}^{\infty}$ es una sucesión creciente de funciones simples no-negativas convergiendo (uniformemente) a g (véase el Teorema 7.2.10, página 383). Se sigue del Teorema de la Convergencia Monótona que

$$\lim_{n\to\infty} \int_E g_n\, d\mu = \int_E g\, d\mu.$$

Por otro lado, como $g(x) = |f'(x)| < k/2^n$ para todo $x \in E_{n,k}$, se sigue del Lema 9.1.57, página 507, que

$$\mu^*(f(E_{n,k})) \leq \frac{k}{2^n} \cdot \mu(E_{n,k}), \quad \text{para} \quad k = 1, \ldots, N2^n$$

y, en consecuencia,

$$\begin{aligned} \mu^*(f(E)) &= \mu^*\left(f\left(\bigcup_{k=1}^{N2^n} E_{n,k}\right)\right) = \mu^*\left(\bigcup_{k=1}^{N2^n} f(E_{n,k})\right) \\ &\leq \sum_{k=1}^{N2^n} \frac{k}{2^n} \cdot \mu(E_{n,k}) \\ &= \sum_{k=1}^{N2^n} \frac{k-1}{2^n} \cdot \mu(E_{n,k}) + \sum_{k=1}^{N2^n} \frac{1}{2^n} \cdot \mu(E_{n,k}) \\ &= \int_E g_n\, d\mu + \frac{1}{2^n} \cdot \mu(E). \end{aligned}$$

De esto se concluye que

$$\mu^*(f(E)) \leq \lim_{n\to\infty} \left(\int_E g_n\, d\mu + \frac{1}{2^n} \cdot \mu(E)\right) = \int_E g\, d\mu$$

y finaliza la prueba para el caso acotado.

Consideremos ahora el caso general, es decir, cuando f' no es acotada. En esta circunstancia, sea $E_n = \{x \in E : n-1 \leq |f'(x)| < n\}$ para $n = 1, 2, \ldots$ Entonces $(E_n)_{n=1}^{\infty}$ es una sucesión disjunta de conjuntos medibles tal que $E = \bigcup_{n=1}^{\infty} E_n$ y, por supuesto, $f(\bigcup_{n=1}^{\infty} E_n) = \bigcup_{n=1}^{\infty} f(E_n)$. Por la primera parte se tiene que

$$\mu^*(f(E_n)) \leq \int_{E_n} |f'|\,d\mu, \quad \text{para todo } n \geq 1.$$

Finalmente, la subaditividad numerable de μ^* y de la integral nos revelan que

$$\mu^*(f(E)) \leq \sum_{n=1}^{\infty} \mu^*(f(E_n)) \leq \sum_{n=1}^{\infty} \int_{E_n} |f'|\,d\mu = \int_E |f'|\,d\mu.$$

Esto completa la prueba. ■

Con los resultados antes expuestos en esta sección podemos acceder a otra demostración del Teorema de Banach-Zarecki, Teorema 9.1.68, página 517, usando la integral de Lebesgue.

Teorema 10.2.37 (Banach-Zarecki). *Sea $f : [a,b] \to \mathbb{R}$ una función arbitraria. Las siguientes condiciones son equivalentes:*

(1) f es ***absolutamente continua*** *sobre $[a,b]$.*

(2) f es ***continua, de variación acotada*** *y satisface* ***la condición (N) de Lusin*** *sobre $[a,b]$.*

Prueba. Sólo demostraremos $(2) \Rightarrow (1)$ ya que la otra implicación fue probada con anterioridad. Suponga entonces que (2) se cumple. Puesto que f es de variación acotada sobre $[a,b]$, su derivada f' existe casi-siempre gracias al Teorema de Diferenciación de Lebesgue. Más aun, por el Lema 10.2.35, $f' \in L_1([a,b])$ y entonces, la absoluta continuidad de la integral, Teorema 10.2.7, implica que para cada $\varepsilon > 0$ elegido de modo arbitrario, existe un $\delta > 0$ tal que

$$\int_E |f'|\,d\mu < \varepsilon \quad \text{para cualquier medible } E \subseteq [a,b] \text{ con } \mu(E) < \delta. \tag{1}$$

Sea $\{I_k = [a_k, b_k] : k = 1, \ldots, n\}$ una colección finita de intervalos cerrados no-superpuestos incluidos en $[a,b]$ tal que $\sum_{k=1}^{n}(b_k - a_k) < \delta$ y para $k = 1, \ldots, n$, pongamos

$$Z_k = \{x \in I_k : f'(x) \text{ existe}\}.$$

Observe que cada uno de los conjuntos $D_k = I_k \setminus Z_k$ es de medida nula y como f satisface la condición (N) de Lusin, se tiene que $\mu^*(f(D_k)) = 0$ para $k = 1, \ldots, n$. Nótese también que, como $I_k = Z_k \cup D_k$ y $f(I_k) = f(Z_k) \cup f(D_k)$, entonces el Lema 10.2.36 nos garantiza que

$$\mu^*(f(I_k)) \leq \mu^*(f(Z_k)) + \mu^*(f(D_k)) = \mu^*(f(Z_k)) \leq \int_{Z_k} |f'|\,d\mu.$$

Por otro lado, puesto que f es continua y el intervalo I_k es compacto, entonces $f(I_k)$ es compacto. Más aun, por el Teorema del Valor Intermedio sabemos que $f(I_k)$ es un intervalo y, en consecuencia, $f(I_k)$ es un intervalo compacto, es decir, $f(I_k) = [\text{mín}_{I_k} f, \text{máx}_{I_k} f]$. De esto y la desigualdad anterior se obtiene que

$$|f(b_k) - f(a_k)| \leq \min_{I_k} f - \max_{I_k} f = \mu(f(I_k)) \leq \int_{Z_k} |f'|\,d\mu \quad \text{para } k = 1, \ldots, n,$$

y así, sumando desde $k = 1$ hasta n y usando (1) concluimos que

$$\sum_{k=1}^{n} |f(b_k) - f(a_k)| \leq \sum_{k=1}^{n} \int_{Z_k} |f'|\, d\mu = \int_{\bigcup_{k=1}^{n} Z_k} |f'|\, d\mu < \varepsilon$$

ya que

$$\mu\Big(\bigcup_{k=1}^{n} Z_k\Big) = \sum_{k=1}^{n} \mu(Z_k) = \sum_{k=1}^{n} \mu(I_k) = \sum_{k=1}^{n} (b_k - a_k) < \delta.$$

Esto termina la prueba. ■

Observe que en la demostración de la implicación $(2) \Rightarrow (1)$ en el Teorema de Banach-Zarecki, la exigencia de que f fuese de *variación acotada* sólo se utilizó para garantizar que su derivada $f' \in \mathrm{L}_1([a,b])$. En consecuencia, uno puede omitir el requerimiento en esa implicación de que "*f es de variación acotada*" por esta otra:

Corolario 10.2.38. *Sea $f : [a,b] \to \mathbb{R}$ una función **diferenciable** sobre $[a,b]$. Si $f' \in \mathrm{L}_1([a,b])$, entonces $f \in \mathrm{AC}([a,b])$.*

El hecho de que toda función de variación acotada es diferenciable casi-siempre y su integral es Lebesgue integrable, nos permite considerar el Lema 10.2.35 como una "buena aproximación" al Teorema Fundamental del Cálculo para la integral de Lebesgue, aunque que ello no es, aun, suficiente para lograr que la fórmula $f(b) - f(a) = \int_{[a,b]} f'\, d\mu$ se cumpla. ¿Qué otra condición adicional se requiere para obtener tal igualdad? Vamos adelantar la respuesta: ¡se requiere que f sea *absolutamente continua*!

Teorema 10.2.39 (Fundamental del Cálculo 2). *Si $f \in \mathrm{L}_1([a,b])$, entonces su **integral indefinida***

$$F(x) = \int_a^x f(t)\, d\mu, \quad x \in [a,b]$$

*es **absolutamente continua** sobre $[a,b]$. En particular, F es **diferenciable casi-siempre** sobre $[a,b]$ y $F'(x) = f(x)$ para **casi-todo** $x \in [a,b]$.*

Prueba. Sea $\varepsilon > 0$ y seleccione, usando el Teorema de Densidad en L_1, Teorema 10.2.13, una función simple φ tal que

$$\int_a^b |f - \varphi|\, d\mu < \frac{\varepsilon}{2}.$$

Puesto que φ es simple, ella es acotada, digamos $|\varphi| \leq M$ para alguna constante $M > 0$. Defina $\delta = \varepsilon/2M$ y suponga que $\{[a_i, b_i] : i = 1, \ldots, n\}$ es una colección finita arbitraria de intervalos no-superpuestos en $[a,b]$ satisfaciendo $\sum_{i=1}^{n}(b_i - a_i) < \delta$. Si hacemos $E = \bigcup_{i=1}^{n}[a_i, b_i]$, resulta que $\mu(E) < \varepsilon/2M$ y, en consecuencia,

$$\begin{aligned}
\sum_{i=1}^{n} \big|F(b_i) - F(a_i)\big| \;&=\; \sum_{i=1}^{n} \bigg| \int_{[a_i,\,b_i]} f(t)\,d\mu \bigg| \\
&\leq\; \sum_{i=1}^{n} \int_{[a_i,\,b_i]} |f(t)|\,d\mu \;=\; \int_E |f(t)|\,d\mu \\
&\leq\; \int_E |f-\varphi|\,d\mu \;+\; \int_E |\varphi(t)|\,d\mu \\
&\leq\; \int_{[a,b]} |f-\varphi|\,d\mu \;+\; \int_E |\varphi(t)|\,d\mu \\
&<\; \frac{\varepsilon}{2} \;+\; \frac{\varepsilon}{2M}\cdot M \;=\; \varepsilon.
\end{aligned}$$

Esto prueba que F es absolutamente continua y, así, gracias al Teorema 9.1.49, página 501, tenemos que F' existe casi-siempre. Para demostrar la última parte, observe que como

$$F(x) \;=\; \int_a^x f^+(t)\,d\mu \;-\; \int_a^x f^-(t)\,d\mu,$$

entonces trabajando con cada una de las funciones f^+ y f^- separadamente, podemos suponer, y así lo haremos, que $f \geq 0$. Para simplificar la presentación también supondremos, en este caso, que f es acotada, digamos, $|f(x)| \leq M'$ para alguna constante $M' > 0$ y todo $x \in [a,b]$. Hechas estas suposiciones, vamos a proceder de modo similar a como lo hicimos en el lema anterior, es decir, extendamos F al intervalo $[a, b+1]$ poniendo $F(x) = F(b)$ para todo $x \in [b, b+1]$ y defina, para cada $n \geq 1$, la función $F_n : [a,b] \to \mathbb{R}$ por

$$F_n(x) \;=\; n\big(F(x+1/n) - F(x)\big).$$

Resulta que $\lim_{n\to\infty} F_n(x) = F'(x)$ para todo $x \in [a,b]$ y, además, como

$$0 \;\leq\; F_n(x) = n\int_x^{x+1/n} f(t)\,d\mu \;\leq\; M',$$

entonces el Teorema de la Convergencia Dominada de Lebesgue combinado con el Ejemplo 8.4.7, página 450, nos garantizan que

$$\begin{aligned}
\int_{[a,x]} F'\,d\mu \;&=\; \lim_{n\to\infty} \int_a^x F_n(t)\,d\mu \\
&=\; \lim_{n\to\infty} \left(n\int_a^x F(t+1/n)\,d\mu \;-\; n\int_a^x F(t)\,d\mu \right) \\
&=\; \lim_{n\to\infty} \left(n\int_x^{x+1/n} F(t)\,d\mu \;-\; n\int_a^{1/n} F\,d\mu \right) \\
&=\; F(x) - F(a) \;=\; \int_a^x f(t)\,d\mu.
\end{aligned}$$

Finalmente, del Lema 10.2.1, página 577, se sigue que $F' = f$ casi-siempre concluyendo, de este modo, la prueba para el caso en que f es no-negativa, Lebesgue integrable y acotada.

Para demostrar el caso general, es decir, suponiendo que f es no-negativa, Lebesgue integrable pero no necesariamente acotada, considere la sucesión $(f_n)_{n=1}^{\infty}$ de funciones no-negativas, Lebesgue integrables y acotadas definidas por $f_n = f \cdot \chi_{[f\leq n]}$. Observe que $\lim_{n\to\infty} f_n = f$. Si para cada $n \geq 1$ definimos

$$F_n(x) \;=\; \int_0^x f_n(t)\, d\mu, \quad \text{para todo} \quad x \in [a,b]$$

resulta, por la primera parte, que $F_n' = f_n$ casi-siempre. Nótese también que para todo $n \geq 1$, se cumple que $F = (F - F_n) + F_n$ y como $F - F_n$ es creciente y no-negativa, el Teorema de Diferenciabilidad de Lebesgue nos muestra que $(F - F_n)' \geq 0$ y, por lo tanto,

$$F' \;=\; (F - F_n)' + F_n' \;\geq\; F_n' \;=\; f_n \;\to\; f \quad \mu - \text{c.s.},$$

esto es, $F' \geq f$ $\mu -$ c.s. Para finalizar, observe que como F es creciente, el Lema 10.2.35 nos asegura que

$$F(x) \;=\; F(x) - F(a) \;\geq\; \int_a^x F'\, d\mu \;\geq\; \int_a^x f\, d\mu \;=\; F(x)$$

para todo $x \in [a,b]$. En particular, $\int_{[a,b]} F'\, d\mu = \int_{[a,b]} f\, d\mu$ y, de nuevo, por una aplicación del Lema 10.2.1 se obtiene que $F' = f$. La prueba es completa. ■

Del Teorema 10.2.39 vimos que si $f \in \mathrm{L}_1([a,b])$, entonces la integral indefinida de f es absolutamente continua, en particular, de variación acotada sobre $[a,b]$. Sin embargo, éste último hecho se puede demostrar de un modo sorprendentemente fácil por medio del siguiente argumento. Puesto que

$$F(x) \;=\; \int_a^x f\, d\mu \;=\; \int_a^x f^+\, d\mu \;-\; \int_a^x f^-\, d\mu$$

y cada una de las funciones $f_1(x) = \int_a^x f^+\, d\mu$ y $f_2(x) = \int_a^x f^-\, d\mu$ es creciente, el Teorema de Jordan nos regala la conclusión: F *es de variación acotada* sobre $[a,b]$.

Una fácil consecuencia de los dos resultados anteriores permite obtener otra interesante descomposición de las funciones de variación acotada debida a Lebesgue.

Corolario 10.2.40 (Descomposición de Lebesgue). *Si $f \in \mathrm{BV}([a,b])$, entonces existen funciones $g, h : [a,b] \to \mathbb{R}$ tales que: g es **absolutamente continua**, h es **singular** y*

$$f \;=\; g + h.$$

Prueba. Sea $f \in \mathrm{BV}([a,b])$. Por el Lema 10.2.35, $f' \in \mathrm{L}_1([a,b])$. Si ahora definimos

$$g(x) \;=\; \int_a^x f'\, d\mu, \quad \text{para todo} \quad x \in [a,b],$$

se sigue entonces del Teorema 10.2.39 que $g \in \mathrm{AC}([a,b])$ y $g' = f'$ casi-siempre. Definiendo $h = f - g$, se tiene que $h' = 0$ casi-siempre y claramente $f = g + h$. Esto termina la prueba. ■

Observe que la condición $f \in BV([a,b])$, en el corolario anterior, sólo se usó para garantizar que $f' \in L_1([a,b])$ por lo que, si hacemos uso del Lema 10.2.35, podemos requerir que f sólo sea creciente y obtener la misma conclusión.

Otra consecuencia sencilla del Teorema 10.2.39 es la demostración del Teorema de Densidad de Lebesgue, Teorema 9.1.33, página 480. En efecto, observe que un punto $x \in \mathbb{R}$ es *punto de densidad de Lebesgue* de un conjunto medible E si

$$\lim_{h\to 0^+} \frac{\mu(E\cap[x-h,\, x+h])}{2h} \;=\; 1.$$

El Teorema de Densidad de Lebesgue establece que casi todos los puntos de un conjunto medible E son puntos de densidad de Lebesgue de E. Para verificar por qué esto es así suponga, en primer lugar, que E es acotado, es decir, $E \subseteq [a,b]$ para algún intervalo compacto $[a,b]$ y considere la función $f \in L_1([a,b])$ definida por $f = \chi_E$. Se sigue del Teorema 10.2.39 que

$$F(x) \;=\; \int_{[a,x]} \chi_E\, d\mu, \qquad x \in [a,b]$$

es diferenciable casi-siempre y que $F' = \chi_E$ μ-c.s. Esto significa que existe un conjunto $D \subseteq [a,b]$ con $\mu(D) = 0$ tal que $F'(x) = 1$ para todo $x \in [a,b]\setminus D$. Por lo tanto, para todo $x \in [a,b]\setminus D$,

$$\begin{aligned} 1 \;&=\; F'(x) \;=\; \lim_{h\to 0} \frac{F(x+h)-F(x-h)}{h} \;=\; \lim_{h\to 0}\frac{1}{2h}\int_{[x-h,x+h]} \chi_E(t)\, d\mu \\ &=\; \lim_{h\to 0} \frac{\mu(E\cap[x-h,\, x+h])}{2h}. \end{aligned}$$

El caso general sigue del hecho de que $E = \bigcup_{n=1}^{\infty}(E\cap[-n,n])$ y cada $E\cap[-n,n]$ es acotado. ■

Lo bueno se hace esperar. Estamos listo para presentar la versión general del Teorema Fundamental del Cálculo (TFC) para la integral de Lebesgue. Recuerde que ya habíamos demostrado dicho resultado para el caso cuando f era diferenciable con derivada acotada, Teorema 10.1.23, página 545.

Teorema 10.2.41 (Fundamental del Cálculo). *Sea $f : [a,b] \to \mathbb{R}$ una función. Las siguientes condiciones son equivalentes:*

(1) $f \in AC([a,b])$.

(2) f *es* ***diferenciable*** *casi-siempre en* $[a,b]$, $f' \in L_1([a,b])$ *y*

$$f(x) \;=\; f(a) + \int_{[a,x]} f'\, d\mu \quad \textit{para todo} \quad x \in [a,b]. \tag{TFC1}$$

Prueba. (1) ⇒ (2). Sea $f \in AC([a,b])$. Puesto que $f \in BV([a,b])$, tenemos que f es diferenciable casi-siempre, $f' \in L_1([a,b])$ y entonces, gracias al Teorema 10.2.39, la función $F : [a,b] \to \mathbb{R}$ definida por

$$F(x) \;=\; \int_{[a,x]} f'(t)\, d\mu, \quad \text{para todo} \quad x \in [a,b],$$

satisface $F \in AC([a,b])$ y $F' = f'$ casi-siempre. Teniendo en cuenta que $f - F$ es absolutamente continua y $(f-F)' = 0$ casi-siempre, se sigue del Corolario 9.1.62, página 510, que existe una constante c tal que $f(x) - F(x) = c$ para todo $x \in [a,b]$. Finalmente, puesto que $F(a) = 0$, resulta que $f(a) = c$ y, por lo tanto, $f(x) = f(a) + F(x)$ para todo $x \in [a,b]$, es decir,

$$f(x) \;=\; f(a) + \int_{[a,x]} f'\,d\mu \quad \text{para todo} \quad x \in [a,b].$$

$(2) \Rightarrow (1)$. Esta implicación es el Teorema 10.2.39. ■

Uno puede abreviar el Teorema Fundamental del Cálculo para la integral de Lebesgue del modo siguiente:

Teorema Fundamental del Cálculo. *Una función $f : [a,b] \to \mathbb{R}$ es **absolutamente continua** sobre $[a,b]$ si, y sólo si, existe una función $g \in \mathrm{L}_1([a,b])$ tal que*

$$f(x) - f(a) \;=\; \int_a^x g\,d\mu, \quad \text{para todo} \quad x \in [a,b].$$

Nota Adicional 10.2.11 Recordemos que el Primer Teorema Fundamental del Cálculo para la integral de Riemann demanda, para que la igualdad

$$\int_a^x f'\,d\mu \;=\; f(x) - f(a) \quad \text{para todo} \quad x \in [a,b] \tag{1}$$

sea válida, que f' sea Riemann integrable. Una restricción similar se requiere en el Teorema Fundamental del Cálculo para la integral de Lebesgue, es decir, la validez de la fórmula (TFC1) exige que f' sea Lebesgue integrable. Esto, por supuesto, es una seria limitación para la integral de Lebesgue. El siguiente ejemplo revela la existencia de una función diferenciable $f : [0,1] \to \mathbb{R}$ cuya derivada $f' \notin \mathrm{L}_1([0,1],\mu)$.

Ejemplo 10.2.3. *La función $f : [0,1] \to \mathbb{R}$ definida por*

$$f(x) \;=\; \begin{cases} x^2 \operatorname{sen}\left(\dfrac{1}{x^2}\right) & \text{si} \quad x \in (0,1] \\ 0 & \text{si} \quad x = 0 \end{cases}$$

es diferenciable sobre $[0,1]$ con derivada no acotada

$$f'(x) \;=\; \begin{cases} 2x \cos\left(\dfrac{\pi}{x^2}\right) + \dfrac{2\pi}{x^2} \operatorname{sen}\left(\dfrac{\pi}{x^2}\right) & \text{si} \;\; 0 < x \leq 1, \\ 0 & \text{si} \;\; x = 0. \end{cases}$$

*Sin embargo, f' **no es Lebesgue integrable**.*

En efecto, si f' fuese Lebesgue integrable, entonces el Teorema Fundamental del Cálculo, Teorema 10.2.41, nos diría que f es absolutamente continua lo cual es imposible por el Ejemplo 9.1.7, página 500. Por consiguiente, no es aplicable el TFC a una tal función. Esto nos indica que

$$\mathcal{N}([a,b]) \;\nsubseteq\; L_1([a,b]).$$

Lo ideal sería, entonces, disponer de una integral donde la fórmula de Newton-Leibniz (1) fuese verdadera para todas las derivadas. Por fortuna, una tal integral existe: la integral de Henstock-Kurzweil, la cual es, en ciertos aspectos, más general que la de Lebesgue.

El TFC para la integral de Lebesgue en combinación con el Lema 10.2.35 conducen a la obtención del siguiente resultado.

Corolario 10.2.42 (Tonelli). *Sea* $f \in BV([a,b])$. *Las siguientes condiciones son equivalentes:*

(1) $f \in AC([a,b])$.

(2) $V(f,[a,b]) \;=\; \int_{[a,b]} |f'|\,d\mu$.

Prueba. (1) $\Rightarrow$ (2). Como f es de variación acotada, el Lema 10.2.35 nos dice que $f' \in L_1([a,b])$ y

$$\int_a^b |f'|\,d\mu \;\leq\; V(f,[a,b]).$$

Para demostrar la otra dirección de la desigualdad anterior, sea $\mathcal{P} = \{a_0, a_1, \ldots, a_n\}$ una partición de $[a,b]$. Puesto que $f \in AC([a,b])$, podemos usar el Teorema Fundamental del Cálculo para obtener

$$\begin{aligned}\sum_{i=1}^{n} |f(a_i) - f(a_{i-1})| \;&=\; \sum_{i=1}^{n} \left| \int_{[a_{i-1},a_i]} f'\,d\mu \right| \\ &\leq\; \sum_{i=1}^{n} \int_{[a_{i-1},a_i]} |f'|\,d\mu \\ &=\; \int_{[a,b]} |f'|\,d\mu \;<\; +\infty.\end{aligned}$$

De aquí se sigue

$$V(f,[a,b]) \;\leq\; \int_{[a,b]} |f'|\,d\mu$$

y (2) se cumple.

(2) $\Rightarrow$ (1). De nuevo, como f es de variación acotada podemos suponer, gracias al Teorema de Jordan, Teorema 9.1.22, que f es creciente. En este caso, $f' \geq 0$ y así, (2) se convierte en

$$\int_a^b f'\,d\mu \;=\; f(b) - f(a). \qquad (\star)$$

Pongamos

$$F(x) \;=\; \int_a^x f'd\mu, \quad x \in [a,b].$$

Por el Lema 10.2.35 se sigue que, para cualesquiera $x, y \in [a,b]$ con $x < y$,

$$F(y) - F(x) \;=\; \int_x^y f'd\mu \;\leq\; f(y) - f(x),$$

lo cual muestra que la función $f - F$ es creciente en $[a,b]$. Más aun, usando $(\star)$, vemos que

$$F(b) - F(a) \;=\; \int_a^b f'\,d\mu \;=\; f(b) - f(a)$$

y, por lo tanto, $(f-F)(a) = (f-F)(b)$. Esto prueba que la función creciente $f-F$ debe ser constante sobre $[a,b]$, digamos $f-F = c$ para alguna constante c. Finalmente, por el Teorema 10.2.39, tenemos que F es absolutamente continua, de donde resulta que $f = F + c$ también es absolutamente continua y termina la prueba. ■

Combinando el Teorema de Banach-Zarecki, Teorema 9.1.68, página 517, con el Teorema Fundamental del Cálculo para la integral de Lebesgue, se obtienen las siguientes equivalencias para cualquier función $f : [a,b] \to \mathbb{R}$ que es absolutamente continua:

Corolario 10.2.43. *Para cualquier función $f : [a,b] \to \mathbb{R}$, las siguientes condiciones son equivalentes:*

(1) $f \in \mathrm{AC}[a,b]$.

(2) Existe una función $g \in \mathrm{L}_1([a,b])$ tal que

$$f(x) - f(a) = \int_a^x g\, d\mu \quad \text{para todo} \quad x \in [a,b].$$

*(3) f es **continua**, de **variación acotada** y satisface la **condición** (N) **de Luzin** sobre $[a,b]$.*

Uno puede usar el TFC para la integral de Lebesgue para dar una definición alternativa, aunque descriptiva, de la integral de Lebesgue para funciones definidas sobre un intervalo compacto $[a,b]$. Esta manera descriptiva de pensar a la integral de Lebesgue guarda cierta similitud con la definición de la integral de Newton.

Corolario 10.2.44. *Sea $f : [a,b] \to \mathbb{R}$ una **función medible**. Las siguientes condiciones son equivalentes:*

(1) $f \in \mathrm{L}_1([a,b])$.

(2) Existe $F \in \mathrm{AC}([a,b])$ tal que $F' = f$ casi-siempre sobre $[a,b]$.

Prueba. (1) $\Rightarrow$ (2). Sea $f \in \mathrm{L}_1([a,b])$. Si definimos

$$F(x) = \int_{[a,x]} f\, d\mu \quad \text{para todo } x \in [a,b],$$

entonces el Teorema 10.2.39 nos revela que $F \in \mathrm{AC}([a,b])$ y $F' = f$ $\mu - c.s.$ sobre $[a,b]$.

(2) $\Rightarrow$ (1). Suponga que (2) se cumple. Puesto que $F \in \mathrm{AC}([a,b])$, el Teorema Fundamental del Cálculo para la integral de Lebesgue nos muestra que F' es Lebesgue integrable y como $F' = f$ casi-siempre sobre $[a,b]$, resulta que f también es Lebesgue integrable sobre $[a,b]$ y termina la prueba. ■

La siguiente caracterización de las funciones Lipschitz es una interesante aplicación del Teorema Fundamental del Cálculo, el Teorema de Lusin y el Teorema de la Convergencia Dominada.

Teorema 10.2.45. *Sea $f : [a,b] \to \mathbb{R}$ una función arbitraria. Las siguientes condiciones son equivalentes:*

*(1) f es M-**Lipschitz**.*

*(2) Existe una sucesión $(f_n)_{n=1}^{\infty}$ de **funciones continuamente diferenciables** definidas sobre $[a,b]$ tal que*

(i) $|f_n'(x)| \leq M$ para todo $x \in [a,b]$ y todo $n \geq 1$.

(ii) $f(x) = \lim_{n\to\infty} f_n(x)$ para cada $x \in [a,b]$.

Prueba. (1) $\Rightarrow$ (2). Suponga que f es M-Lipschitz. Entonces f es absolutamente continua sobre $[a,b]$ y así, por el Teorema Fundamental del Cálculo para la integral de Lebesgue, f' existe casi-siempre sobre $[a,b]$, $f' \in L_1([a,b])$ y

$$f(y) - f(x) \ = \ \int_x^y f'\,d\mu$$

para todo $x,y \in [a,b]$. Observe que como f es M-Lipschitz, $|f'(x)| \leq M$ para todo $x \in [a,b]$. Puesto que f' es medible, el Teorema de Fréchet, Teorema 7.2.21, página 392, nos garantiza la existencia de una sucesión $(g_n)_{n=1}^{\infty}$ de funciones continuas definidas sobre $[a,b]$ que converge a f' casi-siempre y con $|g_n| \leq M$ sobre $[a,b]$. Defina, para cada $n \geq 1$,

$$f_n(x) \ = \ f(a) \ + \int_a^x g_n(t)\,dt \quad x \in [a,b].$$

Se sigue del Teorema Fundamental del Cálculo para la integral de Riemann que cada f_n es continuamente diferenciable y se cumple que $|f_n'(x)| = |g_n(x)| \leq M$ para todo $x \in [a,b]$ y todo $n \geq 1$. Esto prueba (i). Para ver que (ii) también se cumple, observe que como

$$|f_n(x) - f(x)| \ = \ \left| \int_a^x (g_n(t) - f'(t))\,dt \right| \ \leq \ \int_a^x |g_n(t) - f'(t)|\,dt,$$

entonces el Teorema de la Convergencia Dominada nos revela que

$$f(x) \ = \ \lim_{n\to\infty} f_n(x)$$

para cada $x \in [a,b]$.

(2) $\Rightarrow$ (1). Suponga que la sucesión $(f_n)_{n=1}^{\infty}$ existe y satisface (i) y (ii). Entonces, para cualesquiera $x,y \in [a,b]$ se cumple que

$$|f(x) - f(y)| \ = \ \lim_{n\to\infty} |f_n(x) - f_n(y)| \ \leq \ \overline{\lim_{n\to\infty}} \int_x^y |f_n'(t)|\,d\mu(t) \ \leq \ M\,|x-y|,$$

es decir, f es M-Lipschitz y termina la prueba. ■

10.2.7. El Teorema de Vitali-Carathéodory

El siguiente resultado permite aproximar cualquier función Lebesgue integrable sobre $[a,b]$ por funciones semi-continuas, el cual puede ser usado como un ingrediente fundamental para introducir la integral de Perron, una integral que es equivalente a la integral de Henstock-Kurzweil y donde la fórmula de Newton-Leibniz se cumple sin ningún requerimiento adicional sobre la diferenciabilidad de la función.

Teorema 10.2.46 (Vitali-Carathéodory). *Sea $f : [a,b] \to \overline{\mathbb{R}}$ una **función medible**. Las siguientes condiciones son equivalentes:*

(1) $f \in L_1([a,b])$.

(2) Para cada $\varepsilon > 0$, existen funciones $u, v : [a,b] \to \overline{\mathbb{R}}$ tales que $u \in \mathrm{Sci}([a,b])$, $v \in \mathrm{Scs}([a,b])$,

$$v \leq f \leq u \qquad y \qquad \int_{[a,b]} (u - v)\,d\mu \ < \ \varepsilon.$$

Prueba. (1) $\Rightarrow$ (2) Por simetría, es claramente suficiente, dado $\varepsilon > 0$, demostrar la existencia de una función semicontinua inferiormente $u : [a, b] \to \overline{\mathbb{R}}$ tal que

$$f \leq u \qquad \text{y} \qquad \int_{[a,b]} (u - f)\, d\mu < \varepsilon.$$

Para establecer la existencia de una tal función u con las propiedades deseadas, vamos a suponer, en primer lugar, que f es acotada y no-negativa sobre $[a, b]$. Sea $\varepsilon > 0$ y seleccione un $M > 0$ tal que $0 \leq f(x) < M$ para todo $x \in [a, b]$. Pongamos $\delta = \varepsilon/(b - a + 1)$ y escoja un $N \in \mathbb{N}$ de modo que $N\delta > M$. Para cada k en $\{1, \dots, N\}$, sea

$$E_k = \big\{x \in [a, b] : (k-1)\,\delta \leq f(x) < k\,\delta\big\}.$$

Puesto que E_k es medible, podemos elegir un conjunto abierto G_k tal que

$$E_k \subseteq G_k \qquad \text{y} \qquad \mu(G_k) < \mu(E_k) + \frac{1}{k\,2^k}$$

Si ahora consideramos, por cada $k \in \{1, \dots, N\}$, el conjunto $O_k = G_k \cap [a, b]$ resulta que O_k es abierto en $[a, b]$ y, en consecuencia, gracias al Corolario 3.1.49, página 185, se tiene que χ_{O_k} es semicontinua inferiormente sobre $[a, b]$ y, por lo tanto,

$$u = \sum_{k=1}^{N} k\,\delta \cdot \chi_{O_k}$$

es semicontinua inferiormente sobre $[a, b]$. Además, como $u(x) \geq k\,\delta > f(x)$ para todo $x \in E_k$, entonces

$$\begin{aligned}
\int_{[a,b]} u\, d\mu &= \sum_{k=1}^{N} k\,\delta \cdot \mu(O_k) < \sum_{k=1}^{N} k\,\delta\Big(\mu(E_k) + \frac{1}{k\,2^k}\Big) \\
&= \sum_{k=1}^{N} (k-1)\,\delta\,\mu(E_k) + \sum_{k=1}^{N} \delta\,\mu(E_k) + \sum_{k=1}^{N} \frac{\delta}{2^k} \\
&< \sum_{k=1}^{N} \int_{E_k} f\, d\mu + \delta(b-a) + \delta \\
&= \int_{[a,b]} f\, d\mu + \varepsilon.
\end{aligned}$$

Esto termina la prueba para cuando f es acotada y no-negativa sobre $[a, b]$.

Suponga ahora el caso en que f es no-negativa y, para cada $n \in \mathbb{N}$, considere la función truncada $g_n : [a, b] \to \mathbb{R}$ dada por

$$g_n(x) = \min\big\{f(x),\, n\big\}, \quad x \in [a, b].$$

Si definimos $f_1 = g_1$ y $f_n = g_n - g_{n-1}$ para $n \geq 2$, resulta que cada f_n es no-negativa, acotada y $f = \sum_{n=1}^{\infty} f_n$. Por la primera parte de la prueba, por cada entero $n \geq 1$, existe una función u_n la cual es semicontinua inferiormente sobre $[a, b]$ tal que $u_n \geq f$ sobre $[a, b]$ y

$$\int_{[a,b]} u_n\, d\mu < \int_{[a,b]} f_n\, d\mu + \varepsilon\, 2^{-n}.$$

La función $u = \sum_{n=1}^{\infty} u_n$ es semicontinua inferiormente sobre $[a,b]$ gracias al Teorema 3.1.58, página 189 y es claro que $u \geq f$. Una doble aplicación del Teorema de Beppo Levi nos revela que

$$\int_{[a,b]} u\,d\mu \;=\; \sum_{n=1}^{\infty}\int_{[a,b]} u_n\,d\mu \;<\; \sum_{n=1}^{\infty}\bigg(\int_{[a,b]} f_n\,d\mu + \frac{\varepsilon}{2^n}\bigg) \;=\; \int_{[a,b]} f\,d\mu + \varepsilon.$$

Finalmente, suponga que $f:[a,b]\to\mathbb{R}$ es una función Lebesgue integrable arbitraria. Por cada $n\in\mathbb{N}$, defina $f_n(x) = \text{máx}\{f(x),-n\}$. Puesto que $|f_n| \leq |f|$ para todo $n\geq 1$ y como $f_n \to f$ puntualmente, el Teorema de la Convergencia Dominada es el encargado de garantizarnos que

$$\lim_{n\to\infty}\int_{[a,b]} f_n\,d\mu \;=\; \int_{[a,b]} f\,d\mu,$$

lo cual significa que podemos elegir un $N\in\mathbb{N}$ de modo que

$$\int_{[a,b]} f_N\,d\mu \;<\; \int_{[a,b]} f\,d\mu + \frac{\varepsilon}{2}.$$

Aplicando el razonamiento anterior a la función no negativa $f_N + N$, se obtiene una función semicontinua inferiormente u_N tal que

$$u_N \;\geq\; f_N + N \qquad \text{y} \qquad \int_{[a,b]} u_N\,d\mu \;<\; \int_{[a,b]} (f_N+N)\,d\mu + \frac{\varepsilon}{2}.$$

Resulta que la función $u = u_N - N$ tiene las propiedades requeridas. Por simetría, existe una función semicontinua superiormente $v:[a,b]\to\overline{\mathbb{R}}$ tal que

$$v \;\leq\; f \qquad \text{y} \qquad \int_{[a,b]} (f-v)\,d\mu \;<\; \varepsilon$$

y, en consecuencia,

$$\int_{[a,b]} (u-v)\,d\mu \;=\; \int_{[a,b]} (u-f)\,d\mu + \int_{[a,b]} (f-v)\,d\mu \;<\; 2\varepsilon.$$

Para ver que (2) $\Rightarrow$ (1) seleccione, por cada $n\in\mathbb{N}$, una función $u_n\in \mathrm{Sci}([a,b])$ y una función $v_n \in \mathrm{Scs}([a,b])$ tales que

$$v_n \;\leq\; f \;\leq\; u_n \qquad \text{y} \qquad \int_{[a,b]} (u_n - v_n)\,d\mu \;<\; \frac{1}{n}.$$

Sin perder generalidad, asumiremos que la sucesión $(u_n)_{n=1}^{\infty}$ es decreciente (en caso contrario, reemplace cada u_n por $\text{mín}\{u_1,u_2,\ldots,u_n\}$) y la sucesión $(v_n)_{n=1}^{\infty}$ es creciente. Por el Teorema de la Convergencia Dominada se obtiene que

$$\int_{[a,b]} \lim_{n\to\infty}(u_n - v_n)\,d\mu \;=\; 0.$$

Se sigue del Corolario 10.1.28, página 551, que $\lim_{n\to\infty}(u_n - v_n) = 0$ casi-siempre, es decir, $\lim_{n\to\infty} v_n = f = \lim_{n\to\infty} u_n$ casi-siempre y claramente $f\in \mathrm{L}_1([a,b])$ pues $|f|\leq |u_1| + |v_1|$. ∎

Nota Adicional 10.2.12 Observe que las funciones semicontinuas u y v en el Teorema de Vitali-Carathéodory son Lebesgue integrables y, además, se pueden elegir de modo que $v < f < u$. Más aun,

$$-\infty < u(x) \qquad \text{y} \qquad v(x) < +\infty$$

para todo $x \in [a,b]$. Si ahora definimos, para cada $x \in [a,b]$,

$$U(x) = \int_{[a,x]} u\,d\mu \qquad \text{y} \qquad V(x) = \int_{[a,x]} v\,d\mu$$

resulta, del Segundo Teorema Fundamental del Cálculo para la integral de Lebesgue, que U y V son absolutamente continuas sobre $[a,b]$. Afirmamos que

$$u(x) \leq \underline{D}U(x) \qquad \text{y} \qquad \overline{D}V(x) \leq v(x)$$

para todo $x \in [a,b]$. Para ver esto, sea $c \in [a,b]$ y nótese que $-\infty < u(c)$. Sea α un número real tal que $\alpha < u(c)$ y escoja un $\delta > 0$ de modo $u(x) \geq \alpha$ para todo $x \in (c-\delta, c+\delta) \cap [a,b]$. Para cada tal $x \neq c$ se cumple que

$$\frac{U(x) - U(c)}{x - c} = \frac{1}{x-c}\int_{[c,x]} u\,d\mu \geq \frac{1}{x-c}\int_{[c,x]} \alpha\,d\mu = \alpha,$$

de donde se sigue que $\underline{D}U(c) \geq \alpha$. Puesto que $\alpha < u(c)$ fue arbitraria, una aplicación del Teorema 2.1.30, página 97, nos conduce a que $\underline{D}U(x) \geq u(x)$. La prueba de la otra desigualdad es similar y, por lo tanto, se omite.

Corolario 10.2.47. *Si $f : [a,b] \to \mathbb{R}$ es una **función Lebesgue integrable** sobre $[a,b]$, entonces*

$$\begin{aligned}\int_{[a,b]} f\,d\mu &= \inf\left\{\int_{[a,b]} u\,d\mu \;:\; u \in \mathrm{Sci}([a,b]),\; f \leq u\right\}\\ &= \sup\left\{\int_{[a,b]} v\,d\mu \;:\; v \in \mathrm{Scs}([a,b]),\; f \geq v\right\}.\end{aligned}$$

Prueba. Es consecuencia inmediata del Teorema de Vitali-Carathéodory. ■

Otro resultado de interés lo constituye la siguiente teorema que aproxima funciones no-negativas en $\mathrm{Sci}(\mathbb{R})$ por funciones continuas con soporte compacto.

Teorema 10.2.48. *Sea $f \in \mathrm{Sci}(\mathbb{R})$ **no-negativa**. Entonces*

$$\int_{\mathbb{R}} f\,d\mu = \sup\left\{\int_{\mathbb{R}} g\,d\mu \;:\; g \in C_c(\mathbb{R}),\; 0 \leq g \leq f\right\}.$$

Prueba. Puesto que $f \in L_1(\mathbb{R}, \mu)$ y también $g \in L_1(\mathbb{R}, \mu)$ para cualquier $g \in C_c(\mathbb{R})$, resulta que

$$\int_{\mathbb{R}} f\,d\mu \geq \sup\left\{\int_{\mathbb{R}} g\,d\mu \;:\; g \in C_c(\mathbb{R}),\; 0 \leq g \leq f\right\}.$$

Para demostrar la otra desigualdad, tome cualquier número $a \in \mathbb{R}$ tal que $a < \int_{\mathbb{R}} f\,d\mu$. Nuestro objetivo será determinar una función $g \in C_c(\mathbb{R})$ tal que $0 \leq g \leq f$ y $a < \int_{\mathbb{R}} g\,d\mu$.

Para $n, j \in \mathbb{N}$, defina el conjunto

$$G_{n,j} \;=\; f^{-1}\Big(\Big(\frac{j}{2^n}, +\infty\Big)\Big).$$

Puesto que f es inferiormente semicontinua, resulta que $G_{n,j}$ es abierto para todo $n, j \in \mathbb{N}$. Defina ahora la función $s_n : \mathbb{R} \to \mathbb{R}$ por

$$s_n \;=\; 2^{-n} \sum_{j=1}^{n} \chi_{G_{n,j}}.$$

Observe que

$$s_n(x) \;=\; \begin{cases} j\,2^{-n} & \text{si } x \in G_{n,j} \setminus G_{n,j+1}, \quad 1 \leq j < 2^{2n} \\ 2^n & \text{si } x \in G_{n,2^{2n}} \\ 0 & \text{en otro caso.} \end{cases}$$

Usando el hecho de que $G_{n-1,j} = G_{n,2j}$ para todo $n > 1$, resulta que la sucesión $(s_n)_{n=1}^{\infty}$ es creciente y, además, converge puntualmente a f. En efecto, si $x \in f^{-1}((2^{-n}, +\infty))$, entonces $0 \leq f - s_n \leq 2^{-n}$ y el resultado sigue. Invocando el Teorema de la Convergencia Monótona de Lebesgue se tiene que

$$\int_{\mathbb{R}} f\, d\mu \;=\; \lim_{n\to\infty} \int_{\mathbb{R}} s_n\, d\mu.$$

Seleccionemos ahora un $m \in \mathbb{N}$ de modo que

$$2^{-m} \sum_{j=1}^{2^{2m}} \mu(G_{m,j}) \;=\; \int_{\mathbb{R}} s_m\, d\mu \;>\; a.$$

Fijemos un $b \in \mathbb{R}$ tal que $2^{-m} \sum_{j=1}^{2^{2m}} \mu(G_{m,j}) > b > a$. La regularidad de μ nos garantiza que, por cada $1 \leq j \leq 2^{2m}$, existe un conjunto compacto $K_j \subseteq G_{m,j}$ tal que

$$\mu(K_j) \;>\; \mu(G_{m,j}) - (b-a)2^{-m}.$$

Por esto,

$$2^{-m} \sum_{j=1}^{2^{2m}} \mu(K_j) \;>\; 2^{-m} \sum_{j=1}^{2^{2m}} \Big(\mu(G_{m,j}) \;-\; \frac{b-a}{2^m}\Big) \;>\; a.$$

Usemos ahora el Lema de Urysohn para seleccionar, por cada $1 \leq j \leq 2^{2m}$, una función g_j en $C_c(\mathbb{R})$ que verifique:

(a) $0 \leq g_j \leq 1$,

(b) $g_j(x) = 1$ si $x \in K_j$, y

(c) $g_j(x) = 0$ si $x \in G_{m,j}^c$.

De esto último resulta que la función $g = 2^{-m} \sum_{j=1}^{2^{2m}} g_j$ pertenece a $C_c(\mathbb{R})$ y claramente se cumple que $g \leq s_m \leq f$ y

$$\int_{\mathbb{R}} g\, \mu \;\geq\; 2^{-m} \sum_{j=1}^{2^{2m}} \mu(K_j) \;>\; a.$$

Esto completa la prueba. ∎

10.2.8. Regla de la Cadena e Integración por Partes

En esta corta sección discutiremos la validez de la regla de cadena y la integración por partes para funciones absolutamente continuas. El siguiente resultado establece una regla de la cadena bajo hipótesis muy débiles.

Teorema 10.2.49 (Regla de la Cadena). *Sean $f : [a, b] \to \mathbb{R}$ y $g : [c, d] \to [a, b]$ funciones arbitrarias y suponga que f, g y $f \circ g$ son **diferenciables casi-siempre**. Si f satisface la **condición** $(\mathbf{N})$ **de Lusin**, entonces*

$$(f \circ g)'(x) \;=\; f'(g(x)) \cdot g'(x)$$

para casi-todo $x \in [a, b]$.

Prueba. Sean

$$\begin{aligned} Z &= \{x \in [a, b] \,:\, f \text{ no es diferenciable en } x\}, \\ N &= \{z \in [c, d] \,:\, g \text{ ó } f \circ g \text{ no es diferenciable en } z\}. \end{aligned}$$

Por hipótesis, $\mu(Z) = \mu(N) = 0$. Sea

$$E \;=\; \{x \in [a, b] \setminus N \,:\, g(x) \in Z\}.$$

Puesto que $g(E) \subseteq Z$, se tiene que $\mu(g(E)) = 0$ y como f satisface la condición (N) de Lusin, resulta que $\mu(f(g(E))) = 0$. Aplicando el Corolario 9.1.58, página 508, a las funciones g y $f \circ g$, se concluye que

$$g'(x) \;=\; (f \circ g)'(x) \;=\; 0 \quad \text{casi-siempre.}$$

Por otro lado, si $x \in [a, b] \setminus N$ y $g(x) \notin Z$, entonces podemos aplicar la estándar regla de la cadena para concluir que $f \circ g$ es diferenciable en x con $(f \circ g)'(x) = f'(g(x)) \cdot g'(x)$. La prueba es completa. ■

La fórmula de la integración por partes para la integral de Lebesgue se expresa del modo siguiente.

Corolario 10.2.50 (Integración por parte). *Sean $f, g \in \mathrm{L}_1([a, b])$ y sean F y G, respectivamente, sus **integrales indefinidas**, esto es,*

$$F(x) \;=\; \int_a^x f(t)\, d\mu \qquad \text{y} \qquad G(x) \;=\; \int_a^x g(t)\, d\mu$$

para todo $x \in [a, b]$. Entonces

$$\int_a^b fG\, d\mu \;+\; \int_a^b gF\, d\mu \;=\; F(b)G(b) \;-\; F(a)G(a).$$

Prueba. Por el Teorema 10.2.39 sabemos que $F, G \in \mathrm{AC}([a, b])$ y, por lo tanto, $FG \in \mathrm{AC}([a, b])$. En particular, FG es diferenciable casi-siempre y $(FG)' = F'G + FG'$ μ-c.s. Puesto que $F' = f$ y $G' = g$ casi-siempre, entonces

$$\int_a^b (FG)'\, d\mu \;=\; \int_a^b fG\, d\mu \;+\; \int_a^b gF\, d\mu.$$

Por otro lado, usando de nuevo el hecho de que $FG \in AC([a,b])$, entonces el Teorema Fundamental del Cálculo, Teorema 10.2.41, nos muestra que

$$\int_a^b (FG)'\,d\mu \;=\; F(b)G(b) \;-\; F(a)G(a)$$

y termina la prueba. ■

Para funciones absolutamente continuas, la fórmula de la integración por partes para la integral de Lebesgue toma su forma habitual.

Corolario 10.2.51 (Integración por parte). *Sean* $f,g \in AC([a,b])$. *Entonces*

$$\int_a^b fg'\,d\mu \;+\; \int_a^b f'g\,d\mu \;=\; f(b)g(b) \;-\; f(a)g(a).$$

Prueba. Puesto que $(fg)' = f'g + fg'$ y fg es absolutamente continua, el Teorema Fundamental del Cálculo para la integral de Lebesgue finaliza la prueba. ■

10.2.9. Cambio de Variable para la Integral de Lebesgue

Del Lema 10.2.36 sabemos que si $f:[a,b]\to\mathbb{R}$ es una función arbitraria, entonces la estimación

$$\mu^*(f(E)) \;\leq\; \int_E |f'|\,d\mu.$$

es válida para cualquier conjunto medible E en donde f' exista y sea finita sobre E. El siguiente resultado mejora el anterior bajo algunas hipótesis adicionales sobre la función f.

Teorema 10.2.52. *Sea* $f:[a,b]\to\mathbb{R}$ *una función* ***estrictamente creciente*** *y* ***absolutamente continua*** *sobre* $[a,b]$. *Entonces:*

(a) $\mu(f(E)) = \int_E f'\,d\mu$ *para cualquier conjunto medible* $E\subseteq[a,b]$.

(b) *Si* $E=\{x\in[a,b] : f'(x)\neq 0\}$ *y* N *es un subconjunto de* $[f(a),f(b)]$ *con* $\mu(N)=0$, *entonces* $f^{-1}(N)\cap E$ *tiene medida cero.*

(c) *Si* B *es un subconjunto medible de* $[f(a),f(b)]$, *entonces* $C=f^{-1}(B)\cap E$ *es medible y*

$$\mu(B) \;=\; \int_C f'\,d\mu \;=\; \int_a^b \chi_B(f(t))f'(t)\,d\mu.$$

Prueba. (a) Puesto que $f\in AC([a,b])$, tenemos que f' es Lebesgue integrable. Sea $E\subseteq[a,b]$ un conjunto medible. Se sigue del Teorema 9.1.51 que $A=f(E)$ es medible. Más aun, como f es continua, estrictamente creciente y $[a,b]$ es compacto, se tiene que f es un homeomorfismo de $[a,b]$ sobre $[f(a),f(b)]$. Por esto, si A es abierto, entonces E también lo es y, en consecuencia, existe una sucesión disjunta de intervalos abiertos $((a_n,b_n))_{n=1}^{\infty}$ en $[a,b]$ tal que $E=\bigcup_{n=1}^{\infty}(a_n,b_n)$.

De esto se deduce que $A = \bigcup_{n=1}^{\infty}(c_n, d_n)$, donde $c_n = f(a_n)$, $d_n = f(b_n)$. La numerabilidad aditiva de μ y el Teorema Fundamental del Cálculo nos muestran finalmente que

$$\begin{aligned}\mu(A) &= \sum_{n=1}^{\infty}(d_n - c_n) = \sum_{n=1}^{\infty}\big(f(b_n) - f(a_n)\big)\\ &= \sum_{n=1}^{\infty}\int_{a_n}^{b_n} f'\,d\mu = \int_E f'\,d\mu.\end{aligned}$$

Si A es cerrado, entonces $[f(a), f(b)] \setminus A = f([a,b] \setminus E)$ es un conjunto abierto y se sigue de la primera parte que

$$\mu([f(a), f(b)] \setminus A) = \mu(f([a,b] \setminus E)) = \int_{[a,b]\setminus E} f'\,d\mu.$$

Pero $\mu([f(a), f(b)] \setminus A) = (f(b) - f(a)) - \mu(A)$ y, de nuevo, por el Teorema Fundamental del Cálculo,

$$\int_{[f(a),f(b)]\setminus A} f'\,d\mu = \int_{[a,b]} f'\,d\mu - \int_E f'\,d\mu = (f(b) - f(a)) - \int_E f'\,d\mu$$

de donde se obtiene que $\mu(A) = \int_E f'\,d\mu$. Para demostrar el caso general, sea $\varepsilon > 0$ y usemos la regularidad de μ, Corolario 6.3.66, página 302, para hallar un conjunto abierto G y un conjunto cerrado F tales que $F \subseteq A \subseteq G$ y

$$\mu(G) < \mu(A) + \varepsilon \qquad \text{y} \qquad \mu(A) - \varepsilon < \mu(F)$$

Puesto que $f' \geq 0$, resulta que

$$\int_{f^{-1}(F)} f'\,d\mu \leq \int_E f'\,d\mu \leq \int_{f^{-1}(G)} f'\,d\mu$$

lo cual implica que

$$\mu(A) - \varepsilon < \mu(F) \leq \int_E f'\,d\mu \leq \mu(G) < \mu(A) + \varepsilon,$$

es decir,

$$\left|\mu(A) - \int_E f'\,d\mu\right| < \varepsilon$$

y como $\varepsilon > 0$ fue elegido de modo arbitrario, concluimos que $\mu(A) = \int_E f'\,d\mu$.

(b) Puesto que $\mu(N) = 0$, podemos elegir una sucesión decreciente $\big(G_n\big)_{n=1}^{\infty}$ de conjuntos abiertos tales que

$$N \subseteq G_n \qquad \text{y} \qquad \mu(G_n) < \frac{1}{n}, \quad n = 1, 2, \ldots$$

Pongamos $G = \bigcap_{n=1}^{\infty} G_n$. Entonces $N \subseteq G$ y se sigue del Teorema 6.3.47, página 287, que $\mu(G) = \lim_{n\to\infty} \mu(G_n) = 0$. Observe que $f^{-1}(G) = \bigcap_{n=1}^{\infty} f^{-1}(G_n)$ es medible ya que, por la continuidad de f, cada uno de los conjuntos $f^{-1}(G_n)$ es abierto y, por lo tanto, medible. Por la parte (a) se tiene entonces que

$$0 = \mu(G) = \int_{f^{-1}(G)} f'\,d\mu = \int_{f^{-1}(G)\cap E} f'\,d\mu$$

y como $f' > 0$ sobre E, resulta que $\mu(f^{-1}(G)\cap E) = 0$. Finalmente, puesto que $f^{-1}(N) \subseteq f^{-1}(G)$, entonces $\mu(f^{-1}(N)\cap E) = 0$.

(c) Por el Criterio de Medibilidad, Teorema 6.3.55, página 293, existe un conjunto G_δ, llamémoslo G, y un conjunto N de media cero tal que $B = G\cup N$. Entonces $f^{-1}(B) = f^{-1}(G)\cup f^{-1}(N)$ y puesto que f es un homeomorfismo, $f^{-1}(G)$ es un conjunto G_δ, en particular, medible. Se sigue de la parte (b) que $\mu(f^{-1}(N)\cap E) = 0$ y, por lo tanto, $f^{-1}(B)\cap E$ es medible. Más aun, haciendo uso de la parte (a) se tiene que

$$\int_C f'\,d\mu \;=\; \int_{f^{-1}(G)\cap E} f'\,d\mu \;=\; \int_{f^{-1}(G)} f'\,d\mu \;=\; \mu(G) \;=\; \mu(B).$$

Por último,

$$\begin{aligned}\int_C f'\,d\mu \;&=\; \int_{f^{-1}(B)\cap E} f'\,d\mu \;=\; \int_a^b \chi_{f^{-1}(B)\cap E}(t) f'(t)\,d\mu\\ &=\; \int_a^b \chi_B(f(t))f'(t)\,d\mu.\end{aligned}$$

Esto finaliza la prueba. ■

El resultado anterior es fundamental para obtener el siguiente cambio de variable para la integral de Lebesgue.

Teorema 10.2.53 (Cambio de variable). *Sea $g : [a,b] \to \mathbb{R}$ una función **estrictamente creciente** y **absolutamente continua** sobre $[a,b]$. Si $f \in L_1([c,d],\mu)$, donde $c = g(a)$ y $d = g(b)$, entonces*

$$\int_{g(a)}^{g(b)} f\,d\mu \;=\; \int_a^b f(g(t))g'(t)\,d\mu.$$

Prueba. Vamos a suponer, en primer lugar, que f es una función simple, digamos, $f(x) = \sum_{i=1}^n c_i\cdot\chi_{E_i}$, donde $[c,d] = \bigcup_{i=1}^n E_i$. Por el Teorema 10.2.52 (c),

$$\begin{aligned}\int_{g(a)}^{g(b)} f\,d\mu \;&=\; \sum_{i=1}^n c_i\cdot\mu(E_i) \;=\; \sum_{i=1}^n c_i \int_a^b \chi_{E_i}(g(t))g'(t)\,d\mu\\ &=\; \int_a^b f(g(t))g'(t)\,d\mu.\end{aligned}$$

Suponga ahora que f es no-negativa. Por la densidad de las funciones simples, Teorema 7.2.10, existe una sucesión creciente $(\varphi_n)_{n=1}^\infty$ de funciones simples no-negativas convergiendo a f y así, la primera parte y por el Teorema de la Convergencia Monótona aplicado dos veces, se tiene que

$$\begin{aligned}\int_{g(a)}^{g(b)} f\,d\mu \;&=\; \lim_{n\to\infty}\int_{g(a)}^{g(b)} \varphi_n\,d\mu\\ &=\; \lim_{n\to\infty}\int_a^b \varphi_n(g(t))g'(t)\,d\mu\\ &=\; \int_a^b f(g(t))g'(t)\,d\mu.\end{aligned}$$

Finalmente, para demostrar el caso general, observe que si $f \in L_1([c,d],\mu)$, entonces

$$\int_{g(a)}^{g(b)} f\,d\mu \;=\; \int_{g(a)}^{g(b)} f^{+}\,d\mu \;-\; \int_{g(a)}^{g(b)} f^{-}\,d\mu$$

y el resultado sigue del caso anterior. La prueba es completa. ■

10.2.10. Puntos de Lebesgue de una Función Integrable

Del Segundo Teorema Fundamental del Cálculo para la integral de Lebesgue, Teorema 10.2.39, página 612, sabemos que cada función $f \in L_1([a,b])$ posee una primitiva casi-siempre en $[a,b]$, es decir, la función

$$F(x) \;=\; \int_a^x f\,d\mu, \quad x \in [a,b]$$

es diferenciable casi-siempre en $[a,b]$ y $F(x)' = f(x)$ para casi-todo $x \in [a,b]$. Observe que si $N = \{x \in [a,b] : F'(x) \neq f(x)\}$, entonces $\mu(N) = 0$ y para todo $x \in [a,b] \setminus N$,

$$F'(x) \;=\; \lim_{h\to 0} \frac{F(x+h)-F(x)}{h} \;=\; \lim_{h\to 0} \frac{1}{h}\int_x^{x+h} f(t)\,d\mu \;=\; f(x). \tag{1*}$$

Además, siempre se cumple la desigualdad

$$\begin{aligned} |F'(x) - f(x)| \;&=\; \lim_{h\to 0} \left| \frac{F(x+h)-F(x)}{h} - f(x) \right| \\ &\leq\; \lim_{h\to 0} \frac{1}{h}\int_x^{x+h} \big|f(t) - f(x)\big|\,d\mu. \end{aligned} \tag{2}$$

Esto justifica la siguiente definición

Definición 10.2.54. *Sea $f \in L_1([a,b])$. Diremos que un punto $x \in (a,b)$ es un* ***punto de Lebesgue de*** *f si $f(x) \neq \pm\infty$ y*

$$\lim_{h\to 0} \frac{1}{h}\int_{[x,x+h]} \big|f(t) - f(x)\big|\,d\mu \;=\; 0.$$

Nuestro próximo objetivo es demostrar que casi todos los puntos de $[a,b]$ son puntos de Lebesgue de cualquier $f \in L_1([a,b])$. Antes, observe que:

(a) Si $f \in L_1([a,b])$ y $x \in [a,b]$ es un ***punto de Lebesgue*** *de f, entonces la primitiva de f*

$$F(t) \;=\; \int_{[a,t]} f\,d\mu$$

es ***diferenciable*** *en x y $F'(x) = f(x)$.*

Esto es consecuencia inmediata de (2). ■

(b) Sea $f \in L_1([a,b])$. Si f es ***continua*** *en $x \in [a,b]$, entonces x es un* ***punto de Lebesgue*** *de f.*

En efecto, dado $\varepsilon > 0$, usemos el hecho de que f es continua en x para seleccionar un $\delta > 0$ tal que $|f(t) - f(x)| < \varepsilon$ siempre que $0 < |t - x| < \delta$. Ahora, si $|h| < \delta$, entonces

$$\frac{1}{h}\int_x^{x+h} |f(t) - f(x)|\, d\mu \;<\; \varepsilon$$

lo cual equivale a decir que x es un punto de Lebesgue de f. ■

Teorema 10.2.55 (Diferenciación de Lebesgue). *Sea $f \in L_1([a,b])$. Entonces existe un conjunto medible $N \subseteq [a,b]$ con $\mu(N) = 0$ tal que*

$$\lim_{h\to 0}\frac{1}{h}\int_x^{x+h} |f(t) - f(x)|\, d\mu \;=\; 0$$

para todo $x \in [a,b] \setminus N$.

Prueba. Sea $\mathbb{Q} = \{q_n : n = 1, 2, \ldots\}$ una enumeración de los números racionales en $\mathbb{R}$ y, para cada $n \geq 1$, defina el elemento $g_n \in L_1([a,b])$ por

$$g_n(x) \;=\; |f(x) - q_n| \quad \text{para todo } x \in [a,b]$$

Por el Teorema 10.2.39 existe, por cada entero $n \geq 1$, un conjunto medible N_n con $\mu(N_n) = 0$ tal que (véase la igualdad dada en (1^*) al comienzo de esta sección):

$$\lim_{h\to 0}\frac{1}{h}\int_x^{x+h} |f(t) - q_n|\, d\mu \;=\; |f(x) - q_n|$$

para todo $x \in [a,b] \setminus N_n$. Por otro lado, puesto que $f \in L_1([a,b])$, entonces f es finita casi-siempre en $[a,b]$ y, en consecuencia, $\mu(N_0) = 0$, donde $N_0 = \{x \in [a,b] : f(x) = \pm\infty\}$. Por esto, el conjunto

$$N \;=\; \bigcup_{n=0}^{\infty} N_n$$

es de medida cero. Nuestra tarea consistirá en demostrar que cualquier $x \in [a,b] \setminus N$ es un punto de Lebesgue de f. Para ver esto, fijemos un $x_0 \in [a,b] \setminus N$ elegido de modo arbitrario y sea $\varepsilon > 0$. Por la densidad de $\{q_n : n = 1, 2, \ldots\}$ en $\mathbb{R}$, existe un n tal que $|f(x_0) - q_n| < \varepsilon/3$ y de la desigualdad

$$\big||f(t) - q_n| - |f(t) - f(x_0)|\big| \;\leq\; |f(x_0) - q_n|$$

se tiene que

$$\left|\frac{1}{h}\int_{x_0}^{x_0+h} |f(t) - q_n|\, d\mu \;-\; \frac{1}{h}\int_{x_0}^{x_0+h} |f(t) - f(x_0)|\, d\mu\right| \;\leq\; \frac{1}{h}\int_{x_0}^{x_0+h} |f(x_0) - q_n|\, d\mu \;<\; \frac{\varepsilon}{3}.$$

Puesto que $x \notin N$, existe un $\delta > 0$ tal que si $|h| < \delta$, entonces

$$\left|\frac{1}{h}\int_{x_0}^{x_0+h} |f(t) - q_n|\, d\mu \;-\; |f(x_0) - q_n|\right| \;<\; \frac{\varepsilon}{3},$$

lo cual implica que

$$\frac{1}{h}\int_{x_0}^{x_0+h} |f(t) - q_n|\, d\mu \;<\; \frac{2}{3}\varepsilon.$$

Se sigue de esto que si $|h| < \delta$, entonces

$$\frac{1}{h}\int_{x_0}^{x_0+h} |f(t) - f(x_0)|\, d\mu \;<\; \varepsilon,$$

que era lo que queríamos demostrar. Fin de la prueba. ■

Nota Adicional 10.2.13 Nótese que el Teorema 10.2.55 es, en realidad, un resultado sobre diferenciabilidad: la derivada de la primitiva de f. Este resultado se puede también generalizar para funciones integrables en $\mathbb{R}^n$, véase el Teorema 12.3.9, página 827.

10.3. Los Espacios $\mathrm{L}_p(X,\mu)$, $1 < p \leq +\infty$

El objetivo de esta sección es extender nuestra discusión del espacio $\mathrm{L}_1(X,\mu)$. En concreto, introduciremos un continuum de espacios similares al anterior: los imprescindibles espacios $\mathrm{L}_p(\mu)$, donde $p \in (1,+\infty]$. En cada uno de estos espacios se introduce una norma muy especial, llamada la L_p-norma y denotada por $\|\cdot\|_p$, que lo convierte en un espacio de Banach. Algunas de las propiedades relevantes de esos espacios serán mostradas en el transcurso de estas notas.

Definición 10.3.1. *Fijemos un número $p \in (1,+\infty)$. El espacio $\mathrm{L}_p(X,\mathfrak{M}_\mu(X),\mu)$ se define como* ***la colección de todas las clases de equivalencias, bajo la igualdad casi-siempre, de funciones medibles $f : X \to \overline{\mathbb{R}}$ tal que*** *$|f|^p \in \mathbf{L}_1(X,\mathfrak{M}_\mu(X),\mu)$, es decir,*

$$\int_X |f|^p\, d\mu \;<\; +\infty.$$

Cada $f \in \mathrm{L}_p(X,\mathfrak{M}_\mu(X),\mu)$ será llamada una función p-**integrable**. Por supuesto, como en el caso del espacio $\mathrm{L}_1(X,\mathfrak{M}_\mu(X),\mu)$, pensaremos a las clases de equivalencias de $\mathrm{L}_p(X,\mathfrak{M}_\mu(X),\mu)$ como funciones normales con las precauciones de rigor y, por comodidad, a dicho espacio lo escribiremos como $\mathrm{L}_p(X,\mu)$ o $\mathrm{L}_p(X)$ si no hay otra medida distinta a μ.

10.3.1. Las Desigualdades de Hölder y Minkowski en $\mathrm{L}_p(X,\mu)$

Se obtiene una **norma** sobre $\mathrm{L}_p(X,\mu)$ si definimos

$$\|f\|_p \;=\; \left(\int_X |f|^p\, d\mu\right)^{1/p}$$

para cada $f \in \mathrm{L}_p(X,\mu)$. Para verificar que $\|\cdot\|_p$ es, en realidad, una norma debemos mostrar, en primer lugar, que $\mathrm{L}_p(X,\mu)$ es un espacio vectorial sobre $\mathbb{R}$ y luego que $\|\cdot\|_p$ satisface las propiedades $(1)-(4)$ que definen una norma. Las primeras tres propiedades son inmediatas, mientras que la desigualdad triangular, conocida como la Desigualdad de Minkowski, requiere cierto trabajo para su demostración.

Teorema 10.3.2. *$L_p(X,\mu)$ es un **espacio vectorial** sobre $\mathbb{R}$ para cada $p \in [1,+\infty)$.*

Prueba. Sean $f, g \in L_p(X,\mu)$ y sea $a \in \mathbb{R}$. Claramente $a \cdot f \in L_p(X,\mu)$. Para demostrar que $f+g \in L_p(X,\mu)$, observe que

$$\begin{aligned}
|f+g|^p &\leq \big(|f|+|g|\big)^p \\
&\leq \big(\max\{|f|,|g|\} + \max\{|f|,|g|\}\big)^p \\
&= 2^p\big(\max\{|f|,|g|\}\big)^p \\
&= 2^p \max\{|f|^p, |g|^p\} \\
&\leq 2^p\big(|f|^p + |g|^p\big),
\end{aligned}$$

de donde se sigue que

$$\int_X |f+g|^p \, d\mu \leq 2^p \int_X |f|^p \, d\mu + 2^p \int_X |g|^p \, d\mu < +\infty,$$

lo cual prueba que $f+g \in L_p(X,\mu)$. ■

Para demostrar que $\|\cdot\|_p$ es una norma sobre $L_p(X,\mu)$ es imprescindible definir lo que son exponentes conjugados. Comencemos por notar que para cada número real $p > 1$, siempre es posible determinar otro número real único, digamos $q > 1$, tal que $1/p + 1/q = 1$. Para ello sólo es necesario definir $q = p/(p-1)$. Esto permite la siguiente definición.

Definición 10.3.3. *Fijemos un par de números reales $p, q > 1$. Diremos que p y q son **exponentes conjugados** si*

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Observe que si $p, q > 1$ son exponentes conjugados, entonces todas las igualdades

$$pq = p + q, \qquad (p-1)q = p \qquad \text{y} \qquad p = \frac{q}{q-1}$$

son equivalentes. Los exponentes conjugados son los números bajo los cuales es posible establecer la famosa Desigualdad de Hölder en los espacios $L_p(X,\mu)$ para $p > 1$ y, en consecuencia, probar que todos ellos son espacios de Banach. Aunque Otto Hölder (1859-1937) demostró la desigualdad que lleva su nombre en el año 1889, el primero en descubrirla en el año 1888 fue el matemático británico Leonard James Rogers (1862-1933). De hecho, también fue Rogers quien primero descubrió la famosa identidad llamada la **identidad de Rogers–Ramanujan**.

Teorema 10.3.4 (Desigualdad de Hölder en L_p). *Sean $p, q > 1$ **exponentes conjugados** y suponga que $f \in L_p(X,\mu)$ y $g \in L_q(X,\mu)$. Entonces*

(i) $fg \in L_1(X,\mu)$ *y*

(ii) $\|fg\|_1 = \int_X |fg| \, d\mu \leq \|f\|_p \|g\|_q$.

Prueba. Si $f = 0$ o $g = 0$ el resultado es trivial. Suponga que f y g son no-nulas. Por la Desigualdad de Young, Teorema 4.1.8, página 199, sabemos que

$$\frac{|f(x)|}{\| f \|_p} \frac{|g(x)|}{\| g \|_q} \leq \frac{1}{p} \frac{|f(x)|^p}{\| f \|_p^p} + \frac{1}{q} \frac{|g(x)|^q}{\| g \|_q^q}$$

para todo $x \in X$. Integrando a ambos lados de esta desigualdad se obtiene que

$$\begin{aligned} \frac{1}{\| f \|_p \| g \|_q} \int_X |fg|\, d\mu \ &\leq \ \frac{1}{p \| f \|_p^p} \int_X |f|^p \, d\mu + \frac{1}{q \| g \|_q^q} \int_X |q|^q \, d\mu \\ &= \frac{1}{p} + \frac{1}{q} = 1, \end{aligned}$$

la cual es la prueba de (i) y (ii). ■

El siguiente resultado da una descripción alternativa de la norma $\| \cdot \|_p$ cuando $p \in (1, +\infty)$.

Teorema 10.3.5. *Sean $p, q \in (1, +\infty)$ **exponentes conjugados**. Para cada $f \in \mathrm{L}_p(X, \mu)$ se cumple que*

$$\| f \|_p = \sup \left\{ \left| \int_X fg \, d\mu \right| : g \in \mathrm{L}_q(X, \mu), \ \| g \|_q \leq 1 \right\}. \tag{1}$$

*Más aun, dicho **supremo se alcanza**; en otras palabras, existe $g \in \mathrm{L}_q(X, \mu)$ con $\| g \|_q \leq 1$ tal que*

$$\| f \|_p = \int_X fg \, d\mu.$$

Prueba. Fijemos una función $f \in \mathrm{L}_p(X, \mu)$ y denote por $a(f)$ el lado derecho de (1). De la Desigualdad de Hölder se sigue que, para cada $g \in L_q(X, \mu)$,

$$\left| \int_X fg \, d\mu \right| \leq \int_X |fg| \, d\mu \leq \| f \|_p \| g \|_q$$

y, por lo tanto, $a(f) \leq \| f \|_p$. Para demostrar la otra desigualdad, asumiremos que $\| f \|_p > 0$, pues, en caso contrario, la desigualdad $a(f) \leq \| f \|_p$ conlleva a que $a(f) = 0$ y terminaría la prueba. Defina ahora

$$h(x) = \begin{cases} \dfrac{|f(x)|^p}{f(x)} & \text{si } \ f(x) \neq 0 \\ 0 & \text{si } \ f(x) = 0. \end{cases}$$

Claramente, h es medible y $|h| = |f|^{p-1}$. Más aun, como $pq = p + q$ resulta que $|h|^q = |f|^{pq-q} = |f|^p$ y, en consecuencia, $h \in \mathrm{L}_q(X, \mu)$. De aquí se sigue que

$$\| h \|_q = \left(\int_X |h|^q \, d\mu \right)^{1/q} = \left(\int_X |f|^p \, d\mu \right)^{1/q} = \| f \|_p^{p/q}.$$

Haciendo $g = \| f \|_p^{-p/q} h$, resulta que $\| g \|_q = 1$ y, por consiguiente,

$$a(f) \geq \left| \int_X fg \, d\mu \right| = \frac{1}{\| f \|_p^{p/q}} \left| \int_X fh \, d\mu \right|.$$

Puesto que $fh = |f|^p$, continuando con la desigualdad anterior se obtiene que

$$\begin{aligned} a(f) \; &\geq \; \frac{1}{\|\,f\,\|_p^{p/q}}\left|\int_X fh\,d\mu\right| \; = \; \frac{1}{\|\,f\,\|_p^{p/q}}\int_X |f|^p\,d\mu \\ &= \; \frac{\|\,f\,\|_p^p}{\|\,f\,\|_p^{p/q}} \; = \; \|\,f\,\|_p\,. \end{aligned}$$

Esto termina la prueba. ■

Cuando $p = 2 = q$, a la desigualdad de Hölder se le conoce comúnmente como **Desigualdad de Cauchy-Schwarz** o **Cauchy-Buniakovski-Schwarz**. Con la Desigualdad de Hölder a la mano podemos demostrar que $\|\cdot\|_p$ es una norma sobre $L_p(X,\mu)$ para cada $p \geq 1$. Como ya hemos observado, la única dificultad de esta afirmación es comprobar la desigualdad triangular para $p > 1$, la cual recibe un nombre muy especial.

Teorema 10.3.6 (Desigualdad de Minkowski en L_p). *Sea $p \in [1,+\infty)$. Entonces*

$$\|\,f+g\,\|_p \; \leq \; \|\,f\,\|_p + \|\,g\,\|_p\,,$$

cualesquiera sean $f,g \in L_p(X,\mu)$.

Prueba. Sean $f,g \in L_p(X,\mu)$. La desigualdad es trivial si $f+g = 0$. Suponga que $f+g$ es no-nula. Puesto que $(p-1)q = p$, resulta que $(f+g)^{(p-1)} \in L_q(X,\mu)$ y, en consecuencia, por la Desigualdad de Hölder se tiene que

$$\begin{aligned} \|\,f+g\,\|_p^p \; &= \; \int_X |f+g|\,|f+g|^{(p-1)}\,d\mu \\ &\leq \; \int_X |f|\,|f+g|^{(p-1)}\,d\mu \; + \; \int_X |g|\,|f+g|^{(p-1)}\,d\mu \\ &\leq \; \left(\int_X |f|^p\,d\mu\right)^{1/p}\left(\int_X |f+g|^{(p-1)q}\,d\mu\right)^{1/q} + \\ &+ \; \left(\int_X |g|^p\,d\mu\right)^{1/p}\left(\int_X |f+g|^{(p-1)q}\,d\mu\right)^{1/q} \\ &= \; \big(\|\,f\,\|_p + \|\,g\,\|_p\big)\,\|\,f+g\,\|_p^{\frac{p}{q}}\,, \end{aligned}$$

de donde se sigue, multiplicando a ambos lados de esta desigualdad por el número $\|\,f+g\,\|_p^{-\frac{p}{q}}$, que

$$\|\,f+g\,\|_p \; = \; \|\,f+g\,\|_p^{p-\frac{p}{q}} \; \leq \; \|\,f\,\|_p + \|\,g\,\|_p\,.$$

La prueba es completa. ■

La información obtenida de la Desigualdad de Minkowski permite concluir que: *para cada número real $p \in [1,+\infty)$, $(L_p(X,\mu), \|\cdot\|_p)$ es un espacio normado* y, de hecho, un espacio de Banach. Una prueba de esta afirmación sigue, casi sin cambio, de la que ofrecimos para el espacio $(L_1(X,\mu), \|\cdot\|_1)$. Sin embargo, aquí está otra forma de demostrar ese resultado.

Teorema 10.3.7 (Riesz-Fischer). *Para cada $p \in [1, +\infty)$, $(\mathrm{L}_p(X, \mu), \|\cdot\|_p)$ es un* ***espacio de Banach****.*

Prueba. Ya hemos visto que $(\mathrm{L}_1(X, \mu), \|\cdot\|_1)$ es de Banach, de modo que sólo es necesario considerar el caso $p > 1$. Sea $(f_n)_{n=1}^{\infty}$ una sucesión de Cauchy en $\mathrm{L}_p(X, \mu)$, y considere una subsucesión $(f_{n_k})_{k=1}^{\infty}$ de $(f_n)_{n=1}^{\infty}$ con la propiedad de que

$$\| f_{n_{k+1}} - f_{n_k} \|_p < \frac{1}{2^k}, \quad \text{para todo } k \geq 1. \tag{2}$$

Queremos demostrar la existencia de una $f \in \mathrm{L}_p(X, \mu)$ tal que $\| f_n - f \|_p \to 0$. Para lograr tal objetivo, comencemos por definir

$$g_k = |f_{n_1}| + |f_{n_2} - f_{n_1}| + \cdots + |f_{n_{k+1}} - f_{n_k}|, \quad \text{para } k = 1, 2, \ldots$$

Se sigue de la Desigualdad de Minkowski y de (2) que

$$\begin{aligned} \| g_k \|_p^p &= \Big(\| \, |f_{n_1}| + |f_{n_2} - f_{n_1}| + \cdots + |f_{n_{k+1}} - f_{n_k}| \, \|_p \Big)^p \\ &\leq \left(\| f_{n_1} \|_p + \sum_{j=1}^{k} \Big\| f_{n_{j+1}} - f_{n_j} \Big\|_p \right)^p \\ &\leq \left(\| f_{n_1} \|_p + 1 \right)^p < +\infty \end{aligned}$$

para todo $k \geq 1$. Esto prueba que cada $g_k \in \mathrm{L}_p(X, \mu)$ y, además, como la sucesión $(g_n)_{n=1}^{\infty}$ es creciente, resulta que $g = \lim_{k\to\infty} g_k$ existe. El hecho de que $g \in \mathrm{L}_p(X, \mu)$ sigue de la estimación anterior y del Teorema de la Convergencia Monótona, esto es,

$$\int_X g^p \, d\mu = \int_X \lim_{k\to\infty} g_k^p \, d\mu = \lim_{k\to\infty} \int_X g_k^p \, d\mu < +\infty.$$

Puesto que $(\mathbb{R}, |\cdot|)$ es un espacio de Banach y la serie $\sum_{k=1}^{\infty} g_k$ converge absolutamente, el Lema 4.1.13, página 203, nos asegura que la serie

$$f_{n_1}(x) + \sum_{k=1}^{\infty} \big(f_{n_{k+1}}(x) - f_{n_k}(x) \big)$$

converge $\mu -$ c.s. a una función f. Observe que para cada $k \geq 1$,

$$f_{n_1}(x) + \sum_{j=1}^{k} \big(f_{n_{j+1}}(x) - f_{n_j}(x) \big) = f_{n_{k+1}}(x)$$

y por lo tanto, la subsucesión $(f_{n_k})_{k=1}^{\infty}$ converge $\mu -$ c.s. a f. Veamos que $f \in \mathrm{L}_p(X, \mu)$ y $\| f_n - f \|_p \to 0$. En efecto, sea $\varepsilon > 0$ y seleccione, usando el hecho de que $(f_n)_{n=1}^{\infty}$ es de Cauchy, un entero l lo suficientemente grande de modo que

$$\| f_m - f_n \| < \varepsilon \quad \text{para } m, n \geq n_l.$$

Entonces, si $k \geq l$ y $m \geq n_l$, se tiene que

$$\| f_m - f_{n_k} \|_p < \varepsilon$$

y así, gracias al Lema de Fatou, resulta que

$$\begin{aligned}\int_X |f_m - f|^p\, d\mu &= \int_X \liminf_{k\to\infty} |f_{n_k} - f_m|^p\, d\mu \\ &\le \liminf_{k\to\infty} \int_X |f_{n_k} - f_m|^p\, d\mu \\ &\le \varepsilon^p.\end{aligned}$$

para todo $m \ge n_l$. Esto prueba, en particular, que la función $f - f_{n_l} \in L_p(X,\mu)$ y, por lo tanto, $f = (f - f_{n_l}) + f_{n_l} \in L_p(X,\mu)$. Además,

$$\lim_{n\to\infty} \| f_n - f \|_p = 0.$$

La prueba es completa. ■

10.3.2. Convergencia Fuerte y Débil en L_p, $1 < p < +\infty$

Una vez establecida que $\|\cdot\|_p$ es una norma completa sobre $L_p(X,\mu)$, podemos hablar de un nuevo tipo de convergencia en $L_p(X,\mu)$: la convergencia en la p-norma, o simplemente, convergencia en la norma.

Definición 10.3.8. *Sea $p \in [1,+\infty)$. Una sucesión $(f_n)_{n=1}^\infty$ en $L_p(X,\mu)$ se dice que* ***converge fuertemente, o en la p-norma****, a una función $f \in L_p(X,\mu)$ si*

$$\lim_{n\to\infty} \| f_n - f \|_p = 0.$$

Similar al caso de la convergencia en la norma de $L_1(\mu)$, se tiene que:

Si $\| f_n - f \|_p \to 0$, entonces existe una subsucesión $(f_{n_k})_{k=1}^\infty$ de $(f_n)_{n=1}^\infty$ que converge a f $\mu -$ c.s..

En efecto, como la sucesión $(f_n)_{n=1}^\infty$ es convergente en la p-norma, resulta que ella es de Cauchy, y entonces, como en la demostración del Teorema de Riesz-Fischer, existe una subsucesión $(f_{n_k})_{k=1}^\infty$ de $(f_n)_{n=1}^\infty$ que converge $\mu -$ c.s. a f.

El siguiente resultado, el Teorema de la Convergencia Dominada en L_p, garantiza que bajo hipótesis muy sencillas convergencia casi-siempre implica convergencia en la p-norma.

Teorema 10.3.9 (TCD en L_p). *Fijemos $p \in [1,+\infty)$ y sea $(f_n)_{n=1}^\infty$ una sucesión en $L_p(X,\mu)$ tal que $f_n \to f$* ***casi-siempre****. Si existe una función $g \in L_p(X,\mu)$ tal que $|f_n| \le g$ casi-siempre para todo $n \ge 1$, entonces $f \in L_p(X,\mu)$ y*

$$\lim_{n\to\infty} \| f_n - f \|_p = 0.$$

Prueba. Es suficiente suponer que $|f_n| \le g$ y $f_n \to f$ puntualmente. Entonces $|f| \le g$ y, en particular, $|f|^p \le g^p$. De esto se sigue que $f \in L_p(X,\mu)$. Por otro lado, como $|f_n - f| \le 2g$ y $g^p \in L_1(X,\mu)$, entonces

$$|f_n - f|^p \le 2^p g^p \qquad \text{y} \qquad \lim_{n\to\infty} |f_n - f|^p = 0$$

y el Teorema de la Convergencia Dominada en $L_1(X,\mu)$ nos muestra que

$$\lim_{n\to\infty} \| f_n - f \|_p^p \;=\; \lim_{n\to\infty} \int_X |f_n - f|^p \, d\mu \;=\; 0.$$

La prueba es completa. ■

Otro resultado interesante que se obtiene directamente del Teorema de la Convergencia Dominada en L_p es el siguiente.

Teorema 10.3.10. *Sea $p \in (1, +\infty)$ y suponga que $(f_n)_{n=1}^{\infty}$ es una sucesión en $L_p(X,\mu)$ tal que $\| f_n - f \|_p \to 0$ para alguna $f \in L_p(X,\mu)$. Si $(g_n)_{n=1}^{\infty}$ es una sucesión de funciones medibles tal que $|g_n| \leq M$ para todo $n \in \mathbb{N}$ y $g_n \to g$ casi-siempre, entonces $fg \in L_p(X,\mu)$ y*

$$\lim_{n\to\infty} \| f_n g_n - fg \|_p \;=\; 0.$$

Prueba. Observe que como $|g_n| \leq M$, entonces $|fg_n|^p \leq M|f|^p$ para todo entero $n \geq 1$. Esto prueba que $fg_n \in L_p(X,\mu)$ y, además, $fg \in L_p(X,\mu)$ ya que $g_n \to g$ casi-siempre. Se sigue del Teorema de la Convergencia Dominada en L_p que

$$\lim_{n\to\infty} \| fg_n - fg \|_p \;=\; 0.$$

Finalmente, de la desigualdad

$$\begin{aligned} \| f_n g_n - fg \|_p^p &= \int_X |g_n f_n - gf|^p \, d\mu \\ &\leq 2^p \Big(\int_X |g_n|^p |f_n - f|^p \, d\mu \;+\; \int_{[a,b]} |fg_n - fg|^p \, d\mu \Big) \\ &\leq 2^p \Big(M^p \| f_n - f \|_p^p \;+\; \| fg_n - fg \|_p^p \Big) \end{aligned}$$

se tiene que $\lim_{n\to\infty} \| f_n g_n - fg \|_p = 0$ y termina la prueba. ■

Un resultado similar al Teorema de Riesz, Teorema 10.2.12, página 584, para funciones en $L_p(X,\mu)$ con $1 < p < +\infty$ se muestra a continuación.

Teorema 10.3.11 (Riesz para L_p). *Sea $1 \leq p < +\infty$ un número fijo. Sea $(f_n)_{n=1}^{\infty}$ una sucesión de funciones en $L_p(X,\mu)$ que **converge casi-siempre** a una función $f \in L_p(X,\mu)$. Las siguientes condiciones son equivalentes:*

(1) $\lim_{n\to\infty} \| f_n - f \|_p = 0$.

(2) $\lim_{n\to\infty} \| f_n \|_p = \| f \|_p$.

Prueba. La demostración es similar al Teorema 10.2.12 y, por consiguiente, se omite. ■

Además de las ya conocidas convergencia puntual, en medida y la p-norma, existe otro tipo de convergencia en los espacios $L_p(X,\mu)$ que juega un papel de primer orden en la Teoría de la Dualidad de dichos espacios: ella es la convergencia débil.

Definición 10.3.12. *Sea* $p \in (1,+\infty)$. *Una sucesión* $(f_n)_{n=1}^{\infty}$ *en* $L_p(X,\mu)$ *se dice que* ***converge débilmente*** *en* $L_p(X,\mu)$, *si existe una función* $f \in L_p(X,\mu)$ *tal que*

$$\lim_{n\to\infty} \int_X f_n \cdot g \, d\mu \;=\; \int_X f \cdot g \, d\mu.$$

para cualquier $g \in L_q(X,\mu)$, *donde* $1/p + 1/q = 1$.

Los siguientes dos ejemplos muestra que convergencia débil y convergencia puntual casi-siempre son conceptos distintos.

Ejemplo 10.3.1. *Para cada entero* $n \geq 1$, *considere la función* $f_n(x) = \cos(2n\pi x)$ *para todo* $x \in [0,1]$. *Entonces* $f_n \to 0$ *débilmente en* $L_p([0,1],\mu)$ *pero no puntualmente sobre* $[0,1]$.

Prueba. Observe, en primer lugar, que $(f_n)_{n=1}^{\infty} \subseteq L_p([0,1],\mu)$ para cada $p > 1$. Una aplicación del Lema de Riemann-Lebesgue nos revela que $f_n \to 0$ débilmente en $L_p(X,\mu)$ y, por supuesto, $f_n \not\to 0$ puntualmente. ■

Ejemplo 10.3.2. *La sucesión* $(f_n)_{n=1}^{\infty}$, *donde* $f_n = n\chi_{[n,1/n]}$ *para cada* $n \geq 1$ *converge* $\mu -$ c.s. *a* 0 *pero no débilmente en* $L_p([0,1],\mu)$.

Prueba. Ya sabemos que $f_n \to 0$ $\mu -$ c.s. Por otro lado, si tomamos $g = \chi_{[0,1]}$ vemos que $f_n \not\to 0$ débilmente en $L_p([0,1],\mu)$. ■

A pesar de estos resultados negativos, uno obtiene convergencia débil en los siguientes casos.

Teorema 10.3.13. *Sea* $p \in (1,+\infty)$ *y sea* $(f_n)_{n=1}^{\infty}$ *una sucesión en* $L_p(X,\mu)$. *Si*

$$\lim_{n\to\infty} \| f_n - f \|_p \;=\; 0 \quad \text{para alguna } f \in L_p(X,\mu),$$

entonces $f_n \to f$ ***débilmente*** *en* $L_p(X,\mu)$.

Prueba. Sea $g \in L_q(X,\mu)$, donde $1/p + 1/q = 1$. Por la Desigualdad de Hölder se sigue que

$$\left| \int_X (f_n - f) \cdot g \, d\mu \right| \;\leq\; \int_X |(f_n - f) \cdot g| \, d\mu \;\leq\; \| f_n - f \|_p \, \| g \|_q \;\to\; 0$$

cuando $n \to \infty$. Esto termina la prueba. ■

En el ejemplo anterior vimos que convergencia casi-siempre no implica, en general, convergencia débil. Sin embargo, bajo ciertas hipótesis adicionales, como lo muestra el siguiente resultado, es posible obtener convergencia débil.

Teorema 10.3.14. *Sea* $1 < p < +\infty$ *y sea* $f \in L_p(X,\mu)$. *Suponga que* $(f_n)_{n=1}^{\infty}$ *es una sucesión en* $L_p(X,\mu)$ *tal que*

(a) $f_n \to f$ ***casi-siempre*** *y*

(b) $\sup_{n\geq 1} \| f_n \|_p < +\infty$.

Entonces $f_n \to f$ ***débilmente*** *en* $L_p(X,\mu)$.

Prueba. Pongamos $M = \sup_{n\geq 1} \| f_n \|_p$. Por el Lema de Fatou se tiene que

$$\| f \|_p^p \;=\; \int_X |f|^p \, d\mu \;=\; \int_X \lim_{n\to\infty} |f_n|^p \, d\mu \;\leq\; \liminf_{n\to\infty} \int_X |f_n|^p \, d\mu \;\leq\; M^p$$

y, en consecuencia, por la Desigualdad de Minkowski

$$\| f_n - f \|_p \;\leq\; \| f_n \|_p + \| f \|_p \;\leq\; 2M \quad \text{para todo} \quad n \geq 1.$$

Sea $\varepsilon > 0$ y tomemos cualquier $g \in \mathrm{L}_q(X, \mu)$ donde $1/p + 1/q = 1$. Usemos el Teorema 10.2.7 para obtener un $\delta > 0$ tal que

$$\left(\int_E |g|^q \, d\mu \right)^{1/q} \;<\; \frac{\varepsilon}{6M}$$

para cualquier conjunto medible $E \subseteq X$ para el cual $\mu(E) < \delta$. Seleccione ahora, aplicando el Teorema 10.2.11, un conjunto medible $K \subseteq X$ con $\mu(K) < +\infty$ tal que

$$\left(\int_{X\setminus K} |g|^q \, d\mu \right)^{1/q} \;<\; \frac{\varepsilon}{6M}.$$

Hagamos uso del Teorema de Severini-Egoroff para obtener un conjunto medible $B \subseteq K$ tal que $\mu(K \setminus B) < \delta$ y $f_n \to f$ uniformemente sobre B. Esto último permite la existencia de un $n_0 \in \mathbb{N}$ con la propiedad de que cualquier $n \geq n_0$ conduce a que

$$|f_n(x) - f(x)| \;<\; \frac{\varepsilon}{3(\| g \|_q + 1)(\mu(B)^{1/p} + 1)}$$

para todo $x \in B$. En consecuencia, $n \geq n_0$ implica que

$$\left(\int_B |f_n - f|^p \, d\mu \right)^{1/p} \;<\; \frac{\varepsilon}{3(\| g \|_q + 1)}.$$

Combinando todos estos hechos y usando las desigualdades de Hölder y Minkowski obtenemos

$$\begin{aligned}
\left| \int_X (f_n - f) \cdot g \, d\mu \right| \;&\leq\; \int_X |f_n - f||g| \, d\mu \\
&=\; \int_{X\setminus K} |f_n - f||g| \, d\mu \;+\; \int_{K\setminus B} |f_n - f||g| \, d\mu \;+\; \int_B |f_n - f||g| \, d\mu \\
&<\; \| f_n - f \|_p \left(\int_{X\setminus K} |g|^q \, d\mu \right)^{1/q} \;+\; \| f_n - f \|_p \left(\int_{K\setminus B} |g|^q \, d\mu \right)^{1/q} \\
&\quad\; + \left(\int_B |f_n - f|^p \, d\mu \right)^{1/p} \| g \|_q \\
&<\; \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \;=\; \varepsilon.
\end{aligned}$$

para todo $n \geq n_0$. Esto termina la prueba. ∎

Uno puede reemplazar convergencia casi-siempre por convergencia en medida en el Teorema 10.3.14 para obtener la misma conclusión.

Corolario 10.3.15. *Sea* $1 < p < +\infty$ *y sea* $f \in \mathrm{L}_p(X,\mu)$*. Suponga que* $(f_n)_{n=1}^{\infty}$ *es una sucesión en* $\mathrm{L}_p(X,\mu)$ *tal que*

(a) $f_n \to f$ *en* ***medida*** *y*

(b) $\sup\limits_{n\ge 1} \| f_n \|_p < +\infty.$

Entonces $f_n \to f$ ***débilmente*** *en* $\mathrm{L}_p(X,\mu)$*.*

Prueba. Suponga que la sucesión $(f_n)_{n=1}^{\infty}$ no converge débilmente a f. Esto significa que existe una función $g \in \mathrm{L}_q(X,\mu)$ tal que

$$\limsup_{n\to\infty} \left| \int_X f_n g \, d\mu \;-\; \int_X f g \, d\mu \right| \;=\; \alpha \;>\; 0.$$

Por la definición de límite superior existe una subsucesión $(f_{n_k})_{k=1}^{\infty}$ de $(f_n)_{n=1}^{\infty}$ tal que

$$\lim_{k\to\infty} \left| \int_X (f_{n_k} - f)\cdot g \, d\mu \right| \;=\; \alpha. \tag{3}$$

Usemos ahora el Teorema de Riesz, Teorema 7.2.26, página 396, para encontrar una subsucesión $(f_{n_{k_j}})_{j=1}^{\infty}$ de $(f_{n_k})_{k=1}^{\infty}$ tal que $f_{n_{k_j}} \to f$ $\mu -$ c.s. Se sigue del Teorema 10.3.14 que

$$\lim_{j\to\infty} \left| \int_X (f_{n_{k_j}} - f)\cdot g \, d\mu \right| \;=\; 0$$

lo cual es inconsistente con (3). La prueba es completa. ■

El siguiente resultado establece una desigualdad que es importante cuando se tiene convergencia débil.

Teorema 10.3.16. *Sea* $p \in (1,+\infty)$ *y suponga que* $(f_n)_{n=1}^{\infty}$ *es una sucesión en* $\mathrm{L}_p(X,\mu)$ ***convergiendo débilmente*** *a una función* $f \in \mathrm{L}_p(X,\mu)$*. Entonces*

$$\| f \|_p \;\le\; \varliminf_{n\to\infty} \; \| f_n \|_p \, .$$

Prueba. Sea $q > 1$ el exponente conjugado de p. Por el Teorema 10.3.5 sabemos que existe $g \in \mathrm{L}_q(X,\mu)$ con $\| g \|_q = 1$ tal que

$$\| f \|_p \;=\; \int_X f g \, d\mu.$$

Se sigue de la definición de convergencia débil que, para esta g, y usando la Desigualdad de Hölder se obtiene que

$$\begin{aligned} \| f \|_p \;&=\; \int_X f g \, d\mu \;=\; \lim_{n\to\infty} \int_X f_n g \, d\mu \\ &\le\; \varliminf_{n\to\infty} \; \| f_n \|_p \, \| g \|_q \;=\; \varliminf_{n\to\infty} \; \| f_n \|_p \end{aligned}$$

pues $\| g \|_q = 1$. La prueba es completa. ■

Finalizamos esta sección con una elegante caracterización de la convergencia débil en $\mathrm{L}_p(X,\mu)$.

Teorema 10.3.17. *Fijemos un $p \in (1, +\infty)$. Sea $f \in \mathrm{L}_p(X, \mu)$ y suponga $(f_n)_{n=1}^{\infty}$ es una sucesión en $\mathrm{L}_p(X, \mu)$. Las siguiente condiciones son equivalentes:*

(1) $f_n \to f$ ***débilmente*** *en* $\mathrm{L}_p(X, \mu)$.

(2) (a) $\sup_{n \in \mathbb{N}} \| f_n \|_p < \infty$.

(b) $\lim\limits_{n \to \infty} \int_E f_n \, d\mu = \int_E f \, d\mu$ *para cualquier conjunto medible* $E \subseteq X$ *con* $\mu(E) < +\infty$.

Prueba. (1) $\Rightarrow$ (2). En primer lugar, veamos que (a) se cumple. Sea $q \in (1, +\infty)$ el exponente conjugado de p. Puesto que $f_n \to f$ débilmente,

$$\lim_{n \to \infty} \int_X f_n g \, d\mu \;=\; \int_X f g \, d\mu.$$

para cualquier $g \in \mathrm{L}_q(X, \mu)$. Considere, por cada $n \in \mathbb{N}$, el funcional lineal $L_n : \mathrm{L}_q(X, \mu) \to \mathbb{R}$ definido por

$$L_n(g) \;=\; \int_X f g \, d\mu \quad \text{para cualquier } g \in \mathrm{L}_q(X, \mu).$$

Resulta que L_n es continuo, es decir, $L_n \in \mathrm{L}_q(X, \mu)^*$ y entonces, por el Teorema 12.2.2, página 811, se tiene que $\| L_n \| = \| f_n \|_p$. Más aun, para cada $n \geq 1$,

$$|L_n(g)| \;\leq\; \| L_n \| \, \| g \|_q$$

Esto nos deice que la colección $\mathcal{F} = \{L_n : n \in \mathbb{N}\}$ está puntualmente acotada y entonces, gracias al Teorema de Acotación Uniforme, Teorema 2.2.49, página 145, existe una constante $M \geq 0$ tal que

$$\sup_{n \in \mathbb{N}} \| f_n \|_p \;=\; \sup_{n \in \mathbb{N}} \| L_n \| \;\leq\; M.$$

Para verificar (b), tomemos un conjunto medible arbitrario $E \subseteq X$ con $\mu(E) < +\infty$. Entonces

$$\int_X (\chi_E)^q \, d\mu \;=\; \int_X \chi_E \, d\mu \;=\; \mu(E) \;<\; \infty$$

lo cual nos indica que $\chi_E \in \mathrm{L}_q(X, \mu)$. Por esto, y usando el hecho de que $f_n \to f$ débilmente en $\mathrm{L}_p(X, \mu)$, se tiene que

$$\lim_{n \to \infty} \int_X f_n \cdot \chi_E \, d\mu \;=\; \int_X f \cdot \chi_E \, d\mu,$$

en otras palabras,

$$\lim_{n \to \infty} \int_E f_n \, d\mu \;=\; \int_E f \, d\mu.$$

Dejamos a cargo del lector demostrar que (2) $\Rightarrow$ (1). ■

10.3.3. La inclusión $\mathrm{L}_q \subseteq \mathrm{L}_p$ para $1 \leq p < q$

Recordemos que si X es conjunto medible con $\mu(X) < +\infty$, entonces las funciones constantes son p-integrables y, en consecuencia, todas las funciones medibles y acotadas definidas sobre X son p-integrables. En particular, si $g(x) = 1$ para todo $x \in X$, entonces $\int_X 1^p \, d\mu = \mu(X)$.

Teorema 10.3.18. *Suponga que* $\mu(X) < +\infty$ *y sean* $p, q \in [1, +\infty)$ *con* $p < q$. *Entonces*

$$L_q(X,\mu) \subseteq L_p(X,\mu) \qquad y \qquad \|f\|_p \leq \|f\|_q \cdot \mu(X)^{\frac{1}{p}-\frac{1}{q}}$$

para toda $f \in L_q(X,\mu)$.

Prueba. Sea $r = q/p > 1$ y seleccione un número $s > 1$ tal que $1/r + 1/s = 1$. Para cada función $f \in L_q(X,\mu)$, resulta que $|f|^p \in L_r(X,\mu)$ y como $1 \in L_s(X,\mu)$ (aquí es donde usamos el hecho de que $\mu(X) < +\infty$) entonces la Desigualdad de Hölder nos muestra que

$$\begin{aligned}\int_X |f|^p\, d\mu &\leq \left(\int_X (|f|^p)^r\, d\mu\right)^{1/r}\left(\int_X 1^s\, d\mu\right)^{1/s}\\ &= \left(\int_X |f|^q\, d\mu\right)^{p/q} \mu(X)^{1/s}\\ &= \|f\|_q^p\, \mu(X)^{1/s} < +\infty.\end{aligned}$$

De esto se sigue que $f \in L_p(X,\mu)$ y $\|f\|_p \leq \|f\|_q\, \mu(X)^{\frac{1}{p}-\frac{1}{q}}$ ya que

$$\frac{1}{s} = 1 - \frac{1}{r} = 1 - \frac{p}{q} \quad \text{y, por lo tanto,} \quad \frac{1}{ps} = \frac{1}{p} - \frac{1}{q}.$$

La prueba es completa. ■

El ejemplo que sigue muestra que si $\mu(X) = +\infty$ y $p, q \in (1, +\infty)$ con $p < q$, entonces la conclusión dada en el Teorema 10.3.18 puede ser falsa.

Ejemplo 10.3.3. *Para cada* $p, q \in (1, +\infty)$ *con* $p < q$, *existe una función medible* $f : \mathbb{R} \to \mathbb{R}$ *tal que* $f \in L_q(\mathbb{R},\mu)$ *pero* $f \notin L_p(\mathbb{R},\mu)$; *es otras palabras,*

$$L_q(\mathbb{R},\mu) \nsubseteq L_p(\mathbb{R},\mu).$$

Prueba. Fijemos un $p \in (1, +\infty)$ y sea $q > p$. Seleccione un α tal que $\frac{1}{q} < \alpha < \frac{1}{p}$ y entonces defina $f_\alpha : \mathbb{R} \to \mathbb{R}$ por

$$f_\alpha(x) = \frac{1}{(1+|x|)^\alpha}, \quad x \in \mathbb{R}.$$

Puesto que $q\alpha > 1$, la función $f_\alpha \in L_q(\mathbb{R},\mu)$ pero no pertenece a $L_p(\mathbb{R},\mu)$ ya que $p\alpha < 1$. ■

Del Teorema 10.3.18 se sigue que si $\mu(X) < +\infty$, entonces para cada $p \in [1, +\infty)$, se cumple que

$$\bigcap_{q>p} L_q(X,\mu) \subseteq \bigcup_{q>p} L_q(X,\mu) \subseteq L_p(X,\mu). \tag{4}$$

Además, si $\mu(X) = 1$, entonces para todo $1 \leq p < q$

$$\|f\|_p \leq \|f\|_q\,.$$

Es importante destacar que las inclusiones en (4) pueden ser estrictas. En efecto, en el siguiente ejemplo se muestra una función que está en $L_p([0,1],\mu)$ pero no en $\bigcup_{q>p} L_q([0,1],\mu)$. Para ello se requiere recordar que, *para cada* $\alpha, \beta \in \mathbb{R}$, *la serie*

$$\zeta(\alpha,\beta) \;=\; \sum_{n=1}^{\infty} \frac{1}{n^{\alpha}\log^{\beta}(n+1)}$$

converge si, y sólamente si, $\alpha > 1$ *y* β *es arbitrario (pudiendo ser* $\beta = 0$*), o* $\alpha = 1$ *y* $\beta = 1$.

Ejemplo 10.3.4. *Para cada* $1 \leq p < +\infty$, *existe una función*

$$f \in L_p([0,1],\mu) \setminus \bigcup_{q>p} L_q([0,1],\mu).$$

Prueba. Defina la función $f : [0,1] \to \mathbb{R}$ por

$$f(x) \;=\; \begin{cases} 0 & \text{si } x \notin \left(\dfrac{1}{n+1}, \dfrac{1}{n}\right] \\[2ex] \dfrac{n^{1/p}}{\log^{2}(n+1)} & \text{si } \dfrac{1}{n+1} < x \leq \dfrac{1}{n}. \end{cases}$$

Entonces

$$\int_0^1 |f(x)|^p\,dx \;=\; \sum_{n=1}^{\infty} \int_{\frac{1}{n+1}}^{\frac{1}{n}} \frac{n}{\log^{2p}(n+1)}\,dx \;\leq\; \zeta(1,2p) \;<\; +\infty$$

por lo que $f \in L_p([0,1],\mu)$. Por otro lado, si $q > p$, entonces

$$\int_0^1 |f(x)|^q\,dx \;=\; \sum_{n=1}^{\infty} \int_{\frac{1}{n+1}}^{\frac{1}{n}} \frac{n^{q/p}}{\log^{2q}(n+1)}\,dx \;=\; \sum_{n=1}^{\infty} \frac{n^{(q-p)/p}}{(n+1)\log^{2q}(n+1)} \;=\; +\infty,$$

lo cual significa que $f \notin L_q([0,1],\mu)$ para cualquier $q > p$. ■

Teorema 10.3.19. *Suponga que* $\mu(X) < +\infty$ *y sean* $1 \leq p_1 < p_2$. *Si* $f \in L_{p_1}(X,\mu) \cap L_{p_2}(X,\mu)$, *entonces* $f \in L_p(X,\mu)$ *para todo* $p \in [p_1, p_2]$ *y la función* $\Phi_f : [p_1,p_2] \to [0,+\infty)$ *definida por*

$$\Phi_f(p) \;=\; \| f \|_p$$

es ***continua****.*

Prueba. Fijemos $p \in [p_1, p_2]$ y observe que

$$\begin{aligned} |f|^p \;&\leq\; |f|^{p_2} \quad \text{si } |f| > 1 \\ |f|^p \;&\leq\; |f|^{p_1} \quad \text{si } |f| < 1. \end{aligned}$$

De allí que

$$|f|^p \;\leq\; |f|^{p_1} + |f|^{p_2}.$$

Por otro lado, por nuestra hipótesis, tenemos que $|f|^{p_1}, |f|^{p_2} \in L_1(X,\mu)$, de donde se sigue que

$$f \in L_p(X,\mu) \quad \text{para todo } p \in [p_1, p_2].$$

Para demostrar la continuidad de Φ_f, tomemos una sucesión arbitraria $(q_n)_{n=1}^{\infty}$ en $[p_1, p_2]$ tal que $q_n \to p \in [p_1, p_2]$. Entonces $|f|^{q_n} \to |f|^p$ y puesto que

$$|f|^{q_n} \leq |f|^{p_1} + |f|^{p_2} \quad \text{para todo } n \geq 1,$$

entonces el Teorema de la Convergencia Dominada de Lebesgue permite concluir que

$$\lim_{n\to\infty} \Phi(q_n) = \lim_{n\to\infty} \| f \|_{q_n} = \| f \|_p = \Phi(p).$$

La prueba es completa. ■

10.3.4. Conjuntos Uniformemente Integrables en L_p para $p > 1$

A diferencia de lo que ocurre en $L_1(X,\mu)$ donde el acotamiento en la norma de una familia de funciones no es suficiente para garantizar la integrabilidad uniforme de la misma, véase el Ejemplo 10.2.1, en cambio, en $L_p(X,\mu)$ con $p \in (1, +\infty)$, dicha condición si es suficiente; en otras palabras, para determinar la integrabilidad uniforme de una familia de funciones en $\mathcal{U} \subseteq L_p(X,\mu)$ sólo se requiere que ella sea acotada en la norma siempre y cuando $\mu(X) < +\infty$.

Corolario 10.3.20. *Sean* $p \in (1, +\infty)$ *y* $X \in \mathfrak{M}_\mu(\mathbb{R})$ *con* $\mu(X) < +\infty$. *Una familia* $\mathcal{U}$ *de funciones en* $L_p(X,\mu)$ *es* ***uniformemente Integrable*** *si ella es* ***acotada en la norma***, *es decir, si*

$$\sup_{f\in\mathcal{U}} \| f \|_p < \infty.$$

Prueba. Sin perder generalidad, asumiremos que $\mu(X) = 1$. Sean $M = \sup_{f\in\mathcal{U}} \| f \|_p$ y $c > 0$. Por el Teorema 10.3.18, sabemos que

$$L_p(X,\mu) \subseteq L_1(X,\mu) \qquad \text{y} \qquad \| f \|_1 \leq \| f \|_p \quad \text{para cualquier } f \in L_p(X,\mu).$$

Puesto que $c \cdot \mu([|f| > c]) \leq |f| \cdot \chi_{[|f|>c]}$ para cada $f \in L_p(X,\mu)$, se sigue entonces que

$$c \cdot \mu([|f| > c]) \leq \int_X |f| \cdot \chi_{[|f|>c]}\, d\mu = \int_{[|f|>c]} |f|\, d\mu \leq \| f \|_1 \leq \| f \|_p \leq M.$$

Esto prueba que $\mu([|f| > c]) \leq M/c$ y de la Desigualdad de Hölder se tiene que si $f \in \mathcal{U}$, entonces

$$\int_{[|f|>c]} |f|\, d\mu \leq \| f \|_p \cdot \left\| \chi_{[|f|>c]} \right\|_q \leq M \cdot \big(\mu([|f| > c])\big)^{1/q}.$$

De esto último se concluye que

$$\lim_{c\to\infty} \sup_{f\in\mathcal{U}} \int_{[|f|>c]} |f|\, d\mu = 0$$

y termina la prueba. ■

El Teorema de la Convergencia Dominada puede ser mejorado en los espacios L_p del modo siguiente.

Teorema 10.3.21 (Convergencia de Vitali en L_p). *Sean $p \in (1, +\infty)$ y $X \subseteq \mathbb{R}$ un conjunto medible con $\mu(X) < +\infty$. Sea $(f_n)_{n=1}^{\infty}$ una sucesión en $\mathrm{L}_p(X, \mu)$ que* ***converge casi-siempre*** *a una función medible $f : X \to \mathbb{R}$. Las siguientes condiciones son equivalentes:*

(1) *La sucesión $(|f_n|^p)_{n=1}^{\infty}$ es* ***uniformemente integrable****.*

(2) $\lim_{n\to\infty} \| f_n - f \|_p = 0.$

(3) $\lim_{n\to\infty} \| f_n \|_p = \| f \|_p.$

Prueba. (1) $\Rightarrow$ (2). Suponga que (1) se cumple. En primer lugar, vamos a demostrar que $f \in \mathrm{L}_p(X, \mu)$. En efecto, puesto que $f_n \to f$ casi-siempre, el Lema de Fatou y el hecho de $\sup_n \| f_n \|_p \leq M$ para alguna constante $M > 0$, conducen a la desigualdad:

$$\int_X |f|^p \, d\mu \ \leq \ \liminf_{k\to\infty} \int_X |f_{n_k}|^p \, d\mu \ \leq \ M^p.$$

Esto prueba que $f \in \mathrm{L}_p(X, \mu)$.

Fijemos un $\varepsilon > 0$ y usemos la condición de que $(|f_n|^p)_{n=1}^{\infty}$ es uniformemente integrable para determinar un número $\delta > 0$ de modo que para cualquier conjunto medible $E \subseteq X$ con $\mu(E) < \delta$ se cumpla simultáneamente que

$$\int_E |f_n|^p \, d\mu \ < \ \varepsilon \qquad \text{y} \qquad \int_E |f|^p \, d\mu \ < \ \varepsilon.$$

Con el δ hallado, seleccione un entero positivo N, aquí aplicamos el hecho de que $f_n \to f$ en medida ya que $\mu(X) < +\infty$, tal que

$$\mu\big(x \in X : |f_n(x) - f(x)| > \varepsilon^{1/p}\big) \ < \ \delta \quad \text{para todo } n \geq N.$$

Finalmente, puesto que $|f_n - f|^p \leq 2^p(|f_n|^p + |f|^p)$, resulta que para cualquier $n \geq N$ se tiene que

$$\begin{aligned}
\| f_n - f \|_p^p \ &= \ \int_X |f_n - f|^p \, d\mu \\
&= \ \int_{[|f_n - f|^p > \varepsilon]} |f_n - f|^p \, d\mu \ + \ \int_{[|f_n - f|^p \leq \varepsilon]} |f_n - f|^p \, d\mu \\
&\leq \ 2^p \left(\int_{[|f_n - f|^p > \varepsilon]} |f_n|^p \, d\mu \ + \ \int_{[|f_n - f|^p > \varepsilon]} |f|^p \, d\mu \right) \ + \ \varepsilon \cdot \mu\big([\, |f_n - f|^p \leq \varepsilon\,]\big) \\
&< \ \varepsilon(2^p + 2^p + \mu(X)).
\end{aligned}$$

En otras palabras, $\lim_{n\to\infty} \| f_n - f \|_p = 0.$

(2) $\Leftrightarrow$ (3) es el Teorema de Riesz, Teorema 10.3.11.

(3) $\Rightarrow$ (1). Para cada $n \in \mathbb{N}$, considere la función $g_n = |f_n|^p$. Entonces $g_n \to g = |f|^p$ casi-siempre y, en consecuencia, gracias a la condición (3) se tiene que:

$$\lim_{n\to\infty} \int_X g_n \, d\mu \ = \ \int_X g_n \, d\mu.$$

Un llamado al Teorema de Vitali-Scheffé, Teorema 10.2.32, página 605, permite concluir que la sucesión $(|f_n|^p)_{n=1}^{\infty}$ es uniformemente integrable. ■

Cuando $\mu(X) = +\infty$ se obtiene el siguiente resultado, cuya prueba se deja a cargo del lector.

Corolario 10.3.22 (Convergencia de Vitali en L_p). *Sean $p \in (1, +\infty)$ y $X \subseteq \mathbb{R}$ un conjunto medible con $\mu(X) = +\infty$. Sea $(f_n)_{n=1}^{\infty}$ una sucesión en $\mathrm{L}_p(X,\mu)$ que **converge casi-siempre** a una función $f \in \mathrm{L}_p(X,\mu)$. Las siguientes condiciones son equivalentes:*

(1) $\lim\limits_{n\to\infty} \| f_n - f \|_p = 0$.

*(2) La sucesión $(|f_n|^p)_{n=1}^{\infty}$ es **uniformemente integrable y uniformemente Vitali-pequeño**.*

10.3.5. Densidad en los Espacios $\mathrm{L}_p(X,\mu)$

Similar al caso $p = 1$, en las siguientes líneas mostraremos algunos conjuntos que son norma-densos en $\mathrm{L}_p(X,\mu)$, donde $p \in [1, +\infty)$. Como siempre, si $X \subseteq \mathbb{R}$ es un conjunto medible, denotaremos por $\mathcal{S}_\mu(X)$ *el conjunto de las funciones simples definidas sobre* X.

Teorema 10.3.23 (Densidad en L_p). *Si $X \subseteq \mathbb{R}$ es medible con $\mu(X) < +\infty$, entonces $\mathcal{S}_\mu(X)$ es **norma-denso** en $\mathrm{L}_p(X,\mu)$ para cada $1 \leq p < +\infty$.*

Prueba. Fijemos un $p \in [1, +\infty)$. Como $\mu(X) < +\infty$, entonces $\mathcal{S}_\mu(X) \subseteq \mathrm{L}_p(X,\mu)$. Sea $f \in \mathrm{L}_p(X,\mu)$. Por la densidad de las funciones simples, Corolario 7.2.11, página 385, existe una sucesión de funciones simples $(\psi_n)_{n=1}^{\infty}$ tal que $\psi_n \to f$ puntualmente y $|\psi_n| \leq |f|$ para todo $n \geq 1$. Puesto que $|\psi_n - f|^p \to 0$ y $|\psi_n - f|^p \leq 2^{p+1}|f|^p$, entonces el Teorema de la Convergencia Dominada nos revela que

$$\lim_{n\to\infty} \| \psi_n - f \|_p \;=\; \lim_{n\to\infty} \left(\int_{[a,b]} |\psi_n - f|^p \, d\mu \right)^{1/p} \;=\; 0$$

y termina la prueba. ■

El resultado anterior es válido en el siguiente contexto: Si $\mu(X) = +\infty$ y $\mathcal{S}_\mu^\infty(X)$ consiste de todas *las funciones simples que se anulan fuera de un conjunto medible de medida finita*, entonces $\mathcal{S}_\mu^\infty(X)$ es denso en $\mathrm{L}_p(X,\mu)$. La prueba es similar a la del teorema anterior y, por lo tanto, se omite.

Teorema 10.3.24. *Sea X un subconjunto **compacto** de $\mathbb{R}$. Entonces $C(X)$ es **denso** en el espacio $(\mathrm{L}_p(X,\mu), \|\cdot\|_p)$ para cada $p \in [1, +\infty)$.*

Prueba. Fijemos $p \in [1, +\infty)$. Puesto que $\mu(X) < +\infty$ será suficiente, en vista del resultado anterior, demostrar que cada función simple $\varphi : X \to \mathbb{R}$ se puede aproximar por una función continua definida sobre X. Para ver esto último, sea $\varepsilon > 0$ y tomemos cualquier subconjunto medible E de X. Se sigue de la regularidad de μ que

$$\mu(E) \;=\; \sup\{\mu(K) \,:\, K \subseteq E,\ K \text{ compacto}\},$$

de donde podemos elegir un subconjunto compacto K de X tal que $\mu(E \setminus K) < \varepsilon/2$. De aquí se sigue que

$$\| \chi_E - \chi_K \|_p \;=\; \mu(E \setminus K) \;<\; \frac{\varepsilon}{2}.$$

Para este compacto K defina, para cada $n \in \mathbb{N}$, la función $f_n : X \to \mathbb{R}$ por

$$f_n(x) \;=\; \frac{1}{1 + n \cdot \operatorname{dist}(x, K)}.$$

Claramente:

(a) $f_n \in C(X)$,

(b) $0 \leq f_n \leq 1$,

(c) $f_1 \geq f_2 \geq \cdots$ y $\lim_{n\to\infty} f_n = \chi_K$ puntualmente.

Se sigue del Teorema de la Convergencia Monótona que

$$\lim_{n\to\infty} \| f_n - \chi_K \|_p \;=\; 0.$$

Seleccione un $n_0 \in \mathbb{N}$ lo suficientemente grande de modo que

$$\| f_{n_0} - \chi_K \|_p \;<\; \frac{\varepsilon}{2}.$$

Entonces

$$\| f_{n_0} - \chi_E \|_p \;\leq\; \| f_{n_0} - \chi_K \|_p + \| \chi_E - \chi_K \|_p \;<\; \frac{\varepsilon}{2} + \frac{\varepsilon}{2},$$

de donde se sigue el resultado. ■

La prueba del siguiente corolario es casi idéntica a la del caso $p = 1$ y, en consecuencia, se omite.

Corolario 10.3.25. *$C_c(\mathbb{R})$ es **norma-denso** en $\mathrm{L}_p(\mathbb{R}, \mu)$ para cada $p \in [1, +\infty)$.*

10.3.6. Separabilidad de los Espacios $\mathrm{L}_p(\mathbb{R}, \mu)$, $p \in [1, +\infty)$

Recordemos que un espacio métrico (X, d) es separable si existe un subconjunto $F \subseteq X$ que es numerable y denso en X. Esta pequeña sección está dedicada a mostrar que $\mathrm{L}_p(\mathbb{R}, \mu)$ es separable para cualquier $p \in [1, +\infty)$.

Teorema 10.3.26 (Separabilidad de $\mathrm{L}_p(\mathbb{R})$). *Para cada $p \in [1, +\infty)$, el espacio $(\mathrm{L}_p(\mathbb{R}, \mu), \| \cdot \|_p)$ es **separable**.*

Prueba. Fijemos $p \in [1, +\infty)$ y considere el conjunto $\mathrm{Esc}(\mathbb{Q})$ formado por todas las funciones en escaleras $\varphi : \mathbb{R} \to \mathbb{R}$ que son de la forma

$$\varphi = \sum_{i=1}^{n} q_i \chi_{I_i},$$

donde $n \in \mathbb{N}$ es arbitrario, cada intervalo $I_i = [r_i, s_i)$ posee extremos racionales y los números $q_1, \ldots, q_n$ son racionales. Es claro que $\mathrm{Esc}(\mathbb{Q})$ es un conjunto numerable. Veamos que $\mathrm{Esc}(\mathbb{Q})$ es denso en $\mathrm{L}_p(\mathbb{R}, \mu)$. En efecto, sea $f \in \mathrm{L}_p(\mathbb{R}, \mu)$ y sea $\varepsilon > 0$. Según el Corolario 10.3.25, existe una función $g \in C_c(\mathbb{R})$ tal que $\| f - g \|_p < \varepsilon/2$. Sea $m \in \mathbb{N}$ lo suficientemente grande de modo que $\mathrm{sop}(g) \subseteq [-m, m)$. Usemos ahora el hecho de que g es uniformemente continua sobre $\mathbb{R}$ para hallar un $\delta > 0$ tal que

$$|g(x) - g(y)| \;\leq\; \frac{\varepsilon}{(2m)^{1/p}} \quad \text{siempre que } |x - y| < \delta.$$

Divida el intervalo $[-m, m)$ en n intervalos disjuntos (semi-cerrados) $I_1, \ldots, I_n$ con $\ell(I_i) \leq \delta$ y defina

$$\varphi \;=\; \sum_{i=1}^{n} q_i \chi_{I_i},$$

donde cada q_i es un número racional perteneciente al intervalo abierto $(\text{ínf}_{I_i} g, \frac{\varepsilon}{(2m)^{1/p}} + \text{ínf}_{I_i} g)$. Entonces

$$\varphi \in \mathrm{Esc}(\mathbb{Q}) \qquad \text{y} \qquad \|\, g - \varphi \,\|_\infty \;<\; \frac{\varepsilon}{2(2m)^{1/p}}.$$

De esto se concluye que

$$\|\, f - \varphi \,\|_p \;\leq\; \|\, f - g \,\|_p + \|\, g - \varphi \,\|_p \;<\; \frac{\varepsilon}{2} + (2m)^{1/p} \,\|\, g - \varphi \,\|_\infty \;\leq\; \varepsilon$$

y termina la prueba. ■

Nota Adicional 10.3.14 El Teorema de Aproximación, Teorema 6.3.55, página 293, nos dice que cualquier conjunto $E \in \mathfrak{M}_\mu(\mathbb{R})$ se puede expresar en la forma $E = B \cup N$, donde $B \in \mathfrak{Bo}(\mathbb{R})$, $N \in \mathfrak{N}_\mu(\mathbb{R})$ y $B \cap N = \varnothing$. De esto se sigue que si $f \in L_p(\mathbb{R}, \mu)$ para $p \in [1, +\infty)$, entonces

$$\int_E |f|^p \, d\mu \;=\; \int_B |f|^p \, d\mu + \int_N |f|^p \, d\mu \;=\; \int_B |f|^p \, d\mu,$$

de modo que, desde el punto de vista de la integración, los conjuntos nulos son totalmente despreciables y, en consecuencia, podemos considerar el espacio de medida $(\mathbb{R}, \mathfrak{Bo}(\mathbb{R}), \mu)$ y trabajar con $L_p(\mathbb{R}, \mathfrak{Bo}(\mathbb{R}), \mu)$ en lugar de $L_p(\mathbb{R}, \mathfrak{M}_\mu(\mathbb{R}), \mu)$.

10.3.7. El Espacio $L_\infty(X, \mu)$

Fijemos un conjunto medible X. Denotemos por $\mathcal{B}_\infty(X, \mu)$ *el conjunto de todas las funciones* $g : X \to \mathbb{R}$ *que son* **medibles** *y* **acotadas**. Recordemos que para cada $g \in \mathcal{B}_\infty(X, \mu)$, hemos definido la norma sup de g como:

$$\|\, g \,\|_\infty \;=\; \sup_{x \in X} |g(x)|.$$

También observe que si $\mu(X) < +\infty$, entonces

$$\mathcal{S}_\mu(X) \subseteq \mathcal{B}_\infty(X, \mu) \subseteq L_p(X, \mu) \quad \text{para cada } p \in [1, +\infty).$$

Se sigue del Teorema 10.3.23 que

Corolario 10.3.27. *Si $\mu(X) < +\infty$, entonces $\mathcal{B}_\infty(X, \mu)$ es* ***norma-denso*** *en $(L_p(X, \mu), \|\cdot\|_p)$ para cada $p \in [1, +\infty)$.*

Teniendo en cuenta el Teorema 10.3.18, resulta que si $\mu(X) < +\infty$, entonces

$$\mathcal{B}_\infty(X,\mu) \subseteq \bigcap_{p\geq 1} \mathrm{L}_p(X,\mu).$$

Hasta ahora hemos obtenido y analizado algunas de las propiedades de los espacios $\mathrm{L}_p(X,\mu)$ para cualquier $p \in [1,+\infty)$. ¿Es posible extender esta noción para $p = +\infty$? Uno puede lograr tal cometido mirando entre los espacios $\mathcal{B}_\infty(X,\mu)$ y $\bigcap_{p\geq 1} \mathrm{L}_p(X,\mu)$, es decir, en lugar de considerar sólo funciones medibles y acotadas $f : X \to \mathbb{R}$, vamos a considerar funciones cuya diferencia con una función medible y acotada sea un conjunto nulo, es decir, funciones $f : X \to \mathbb{R}$ para las cuales existe una constante $M > 0$ tal que

$$\mu\big(\{x \in X : |f(x)| > M\}\big) \;=\; 0.$$

Definición 10.3.28. *Sea X un conjunto medible. Una función medible $f : X \to \mathbb{R}$ se llama **esencialmente acotada** sobre X, si existe una función medible y acotada $g : X \to \mathbb{R}$ tal que $f = g$ casi-siempre en X.*

Designaremos por $\mathrm{L}_\infty(X,\mu)$ *el conjunto de todas las clases de equivalencias de funciones $f : X \to \mathbb{R}$ que son esencialmente acotadas sobre X.* Tal y como hicimos con los espacios L_p, no haremos distinción alguna entre una función esencialmente acotada y su clase de equivalencia. Bajo esta suposición, resulta claro que con las operaciones usuales $\mathrm{L}_\infty(X,\mu)$ es un espacio vectorial sobre $\mathbb{R}$.

Lema 10.3.29. *Si $f \in \mathrm{L}_\infty(X,\mu)$, entonces existe una constante $M > 0$ tal que*

$$\mu\big(\{x \in X : |f(x)| > M\}\big) \;=\; 0.$$

Prueba. Sea f una función esencialmente acotada sobre X y seleccione una función medible y acotada $g : X \to \mathbb{R}$ de modo tal que $\mu(\{x \in X : f(x) \neq g(x)\}) = 0$. Sea M una constante positiva para la cual $|g(x)| \leq M$ para todo $x \in X$ y note que

$$\big\{x \in X : |f(x)| > M\big\} \;\subseteq\; \big\{x \in X : f(x) \neq g(x)\big\}.$$

En efecto, si $z \in \big\{x \in X : |f(x)| > M\big\}$, entonces $|f(z)| > M$ y como $|g(z)| \leq M$ resulta que $|f(z)| > |g(z)|$ y, por lo tanto, $f(z) \neq g(z)$, es decir, $z \in \big\{x \in X : f(x) \neq g(x)\big\}$. De la inclusión anterior se sigue

$$\mu\big(\{x \in X : |f(x)| > M\}\big) \;=\; 0$$

y termina la prueba. ■

La información obtenida en el resultado anterior permite que podamos definir, para cada función $f \in \mathrm{L}_\infty(X,\mu)$, el número:

$$\begin{aligned}
\|f\|_{\mathrm{ess}} \;&=\; \operatorname{ínf}\Big\{\,\|g\|_\infty \,:\; g \in \mathcal{B}_\infty(X,\mu),\;\; f \,=\, g \;\; \mu-\text{c.s.}\Big\}\\
&=\; \operatorname{ínf}\Big\{M > 0 \,:\; \mu\big(\{x \in X : |f(x)| > M\}\big) = 0\Big\}\\
&=\; \operatorname{ínf}\Big\{M > 0 \,:\; |f(x)| \leq M \;\; \mu-\text{c.s. sobre } \; X\Big\}.
\end{aligned}$$

A $\| f \|_{\text{ess}}$ también se le llama el **supremo esencial** de f y se denota frecuente como

$$\| f \|_{\text{ess}} \;=\; \text{ess sup}\{|f(x)| \;:\; x \in X\}.$$

Observe que, por cada $f \in \mathrm{L}_\infty(X,\mu)$ existe, por las propiedades del ínfimo, una sucesión decreciente $(M_n)_{n=1}^{\infty}$ de números reales positivos tal que

$$\| f \|_{\text{ess}} \;=\; \lim_{n\to\infty} M_n \qquad \text{y} \qquad |f(x)| \;\leq\; M_n \;\; \mu - \text{c.s.} \quad n = 1,2,\ldots$$

El hecho de que $\| \cdot \|_{\text{ess}}$ es una norma sobre $\mathrm{L}_\infty(X,\mu)$ puede ser deducido del siguiente resultado.

Teorema 10.3.30. *Para cada* $f \in \mathrm{L}_\infty(X,\mu)$, *se cumple que*

$$|f| \;\leq\; \| f \|_{\text{ess}} \;\; \mu - \text{c.s.} \tag{a_∞^1}$$

Prueba. Sea $f \in \mathrm{L}_\infty(X,\mu)$ y seleccione una sucesión decreciente $(M_n)_{n=1}^{\infty}$ de números reales positivos tal que

$$\| f \|_{\text{ess}} \;=\; \lim_{n\to\infty} M_n \qquad \text{y} \qquad |f(x)| \;\leq\; M_n \;\; \mu - \text{c.s.} \quad n = 1,2,\ldots$$

Si para cada $n \geq 1$, hacemos $A_n = \{x \in X : |f(x)| > M_n\}$, resulta que $\mu(A_n) = 0$ y, en consecuencia, el conjunto $A = \bigcup_{n=1}^{\infty} A_n$ también tiene medida cero. Más aun, si $x \in X \setminus A$, entonces $|f(x)| \leq M_n$ para todo $n \geq 1$ y, por lo tanto,

$$|f(x)| \;\leq\; \lim_{n\to\infty} M_n \;=\; \| f \|_{\text{ess}}\,.$$

Esto termina la prueba. ∎

Observe que, gracias al resultado anterior, se tiene que

$$\mu\big(\{x \in \mathrm{X} : |f(x)| > \| f \|_{\text{ess}}\}\big) \;=\; 0$$

para cualquier $f \in \mathrm{L}_\infty(X,\mu)$. En particular, si $a \in \mathbb{R}$,

$$a \;<\; \| f \|_{\text{ess}} \quad \Rightarrow \quad \mu\big(\{x \in X : |f(x)| > a\}\big) > 0. \tag{a_∞^2}$$

Es importante hacer la siguiente aclaratoria: en la mayoría de los textos sobre Teoría de la Medida donde se estudia el espacio $\mathrm{L}_\infty(X,\mu)$ casi siempre se utiliza la notación $\| f \|_\infty$ en lugar de $\| f \|_{\text{ess}}$. La razón para esto es simple. Suponga que $f \in \mathrm{L}_\infty(X,\mu)$ y consider el conjunto $E = \{x \in X : |f(x)| > \| f \|_{\text{ess}}\}$. Como $\mu(E) = 0$, la función $g_0 : X \to \mathbb{R}$ definida por

$$g_0(x) \;=\; \begin{cases} f(x) & \text{si} \quad x \in X \setminus E \\ 0 & \text{si} \quad x \in E \end{cases}$$

satisface las siguientes propiedades: g_0 es medible, acotada y $f = g_0$ casi-siempre en E. De esto se sigue que $\| g_0 \|_\infty \leq \| f \|_{\text{ess}}$. Por otro lado, haciendo uso de la definición de $\| f \|_{\text{ess}}$:

$$\| f \|_{\text{ess}} \;=\; \text{ínf}\Big\{ \| g \|_\infty \;:\; g \in \mathcal{B}_\infty(X,\mu), \;\; g \;=\; f \;\; \mu - \text{c.s.}\Big\} \tag{1}$$

se tiene que $\| f \|_{\text{ess}} \leq \| g_0 \|_\infty$. En consecuencia,

$$\| f \|_{\text{ess}} = \| g_0 \|_\infty .$$

Esta igualdad nos indica que *el ínfimo en* (1) *siempre se alcanza*. Similarmente, la desigualdad obtenida en el Teorema 10.3.30 nos revela que $\| f \|_{\text{ess}} \in \{M > 0 : |f(x)| \leq M \ \ \mu-\text{c.s. sobre } X\}$. Estos argumentos permiten que podamos utilizar indistintamente la notación $\| f \|_\infty$ o $\| f \|_{\text{ess}}$ para cualquier función $f \in \mathrm{L}_\infty(X, \mu)$.

Teniendo en cuenta que la integración no se afecta sobre conjuntos de medida cero tenemos, como una aplicación del Teorema 10.3.30, el siguiente:

Corolario 10.3.31. *Si* $\mu(X) < +\infty$, *entonces para todo* $p \in \mathbb{R}$ *con* $1 \leq p < +\infty$,

$$\mathrm{L}_\infty(X, \mu) \subseteq \mathrm{L}_p(X, \mu) \subseteq \mathrm{L}_1(X, \mu).$$

Prueba. Fijemos un $p \in [1, +\infty)$ y sea $f \in \mathrm{L}_\infty(X, \mu)$. Por el Teorema 10.3.30 tenemos que $|f(x)| \leq \| f \|_\infty$ para todo $x \in X \setminus E$, donde $E = \{x \in X : |f(x)| > \| f \|_\infty\}$ y, por consiguiente,

$$\int_X |f|^p \, d\mu \leq \int_{X \setminus E} \| f \|_\infty^p \, d\mu = \| f \|_\infty^p \cdot \mu(X \setminus E) < +\infty.$$

La prueba es completa. ■

El resultado anterior nos revela que para cualquier conjunto medible X de **medida finita**, $\mathrm{L}_\infty(X, \mu)$ es un espacio "muy pequeño" en comparación con los restantes $\mathrm{L}_p(X, \mu)$. De hecho,

$$\mathrm{L}_\infty(X, \mu) \subseteq \bigcap_{p \geq 1} \mathrm{L}_p(X, \mu)$$

y dicha inclusión es, en general, propia. En efecto, considere la función $f : [0, 1] \to \mathbb{R}$ definida por $f(x) = \ln(x)$. Entonces se tiene que, para todo $n \in \mathbb{N}$,

$$\int_0^1 |\ln(x)|^n \, dx = n!$$

y así, $f \in \mathrm{L}_n([0, 1], \mu)$ para todo $n \in \mathbb{N}$. Se sigue del Teorema 10.3.18 que $f \in \mathrm{L}_p([0, 1], \mu)$ para todo $p \geq 1$. Por supuesto, $f \notin \mathrm{L}_\infty([0, 1], \mu)$.

Usando (a_∞^1) es fácil verificar que:

Teorema 10.3.32. $\| \cdot \|_\infty$ *es una* ***norma*** *sobre* $\mathrm{L}_\infty(X, \mu)$.

Prueba. Sean $f, g \in \mathrm{L}_\infty(X, \mu)$ y $a \in \mathbb{R}$.

(i) $\| f \|_\infty = 0 \Rightarrow f = 0$. En efecto, por (a_∞^1) se tiene que $|f| \leq \| f \|_\infty = 0$ casi-siempre y, así, $f = 0$ (como clase de equivalencia).

(ii) $\| a \cdot f \|_\infty = |a| \, \| f \|_\infty$. Si $a = 0$ no hay nada que probar. Suponga entonces que $a \neq 0$. De nuevo, por (a_∞^1), $|a \cdot f| = |a||f| \leq |a| \, \| f \|_\infty$ casi-siempre y, por consiguiente,

$$\| a \cdot f \|_\infty \leq |a| \, \| f \|_\infty .$$

Usando este hecho se obtiene que $\|f\|_\infty = \left\|\frac{1}{a}(a\cdot f)\right\|_\infty \leq \frac{1}{|a|}\|a\cdot f\|_\infty$ y, en consecuencia,

$$|a|\,\|f\|_\infty \;\leq\; \|a\cdot f\|_\infty.$$

Por esto, $\|a\cdot f\|_\infty = |a|\,\|f\|_\infty$.

(iii) $\|f+g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$. Puesto que $|f| \leq \|f\|_\infty$ casi-siempre y $|g| \leq \|g\|_\infty$ casi-siempre, entonces

$$|f+g| \;\leq\; |f|+|g| \;\leq\; \|f\|_\infty + \|g\|_\infty \quad \mu-\text{c.s.}$$

de donde se obtiene que $\|f+g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$ y termina la prueba. ■

Teorema 10.3.33. *$(L_\infty(X,\mu), \|\cdot\|_\infty)$ es un **espacio de Banach**.*

Prueba. Sea $(f_n)_{n=1}^\infty$ una sucesión de Cauchy en $(L_\infty(X,\mu), \|\cdot\|_\infty)$ y sea $\varepsilon > 0$. Seleccionemos un $N \in \mathbb{N}$ tal que $\|f_m - f_n\|_\infty < \varepsilon$ para todo $m,n \geq N$. Del Teorema 10.3.30 sabemos que existe un conjunto medible E con $\mu(X\setminus E) = 0$ tal que, para todo $x \in E$,

$$|f_m(x) - f_n(x)| \;\leq\; \|f_m - f_n\|_\infty \;<\; \varepsilon \quad \text{para todo} \quad m,n \geq N.$$

Esto nos indica que, para cada $x \in E$, $(f_n(x))_{n=1}^\infty$ es una sucesión de Cauchy en $\mathbb{R}$ y, en consecuencia, converge en $\mathbb{R}$. Defina $f : X \to \mathbb{R}$ por

$$f(x) \;=\; \begin{cases} \lim\limits_{n\to\infty} f_n(x) & \text{si} \quad x \in E, \\ 0 & \text{si} \quad x \in X\setminus E. \end{cases}$$

Observe que si $n \geq N$, entonces para cualquier $x \in E$,

$$|f_n(x) - f(x)| \;=\; \lim_{m\to\infty} |f_m(x) - f_n(x)| \;\leq\; \lim_{m\to\infty} \|f_m - f_n\|_\infty \;<\; \varepsilon \tag{2}$$

Esto nos dice que la sucesión $(f_n)_{n=1}^\infty$ converge uniformemente a f sobre E. Veamos ahora que $f \in L_\infty(X,\mu)$. Para ver esto último, observe que para cualquier $x \in E$,

$$|f(x)| \;\leq\; |f(x) - f_N(x)| + |f_N(x)| \;<\; \varepsilon + \|f_N\|_\infty$$

y como $|f(x)| = 0 \leq \varepsilon + \|f_N\|_{\text{ess}}$ para todo $x \in X\setminus E$, resulta que $f \in L_\infty(X,\mu)$. Finalmente, usando (2) y el hecho de que $f_n - f \in L_\infty(X,\mu)$ para todo $n \geq 1$ se tiene que $\|f_n - f\|_\infty < \varepsilon$ para todo $n \geq N$. La prueba es completa. ■

Nótese que si $1 \leq p < q \leq +\infty$ y $\mu(X) < +\infty$, entonces el Teorema 10.3.18 nos muestra que

$$L_\infty(X,\mu) \;\subseteq\; L_q(X,\mu) \;\subseteq\; L_p(X,\mu) \;\subseteq\; L_1(X,\mu).$$

y cuando $\mu(X) = 1$, se tiene que

$$\|f\|_1 \;\leq\; \|f\|_p \;\leq\; \|f\|_q \;\leq\; \|f\|_\infty.$$

En vista de estas desigualdades, la primera igualdad en el siguiente resultado no debería representar ninguna sorpresa. Antes de demostrarlo, recuerde que si $a > 0$, entonces

$$\lim_{p\to\infty} a^{1/p} \;=\; 1.$$

Teorema 10.3.34. *Sea X un **conjunto medible** con $0 < \mu(X) < +\infty$. Entonces, para cualquier función $f \in \mathrm{L}_\infty(X, \mu)$,*

$$\| f \|_\infty \;=\; \lim_{p\to\infty} \| f \|_p . \tag{4}$$

Más aun, si para cada entero $n \geq 1$, definimos $p_n = \int_X |f|^n \, d\mu$, entonces

$$\| f \|_\infty \;=\; \lim_{n\to\infty} \frac{p_{n+1}}{p_n} \;=\; \lim_{n\to\infty} \sqrt[n]{p_n}. \tag{5}$$

Prueba. Si $\| f \|_\infty = 0$ el resultado es trivial. Suponga que $\| f \|_\infty > 0$. Puesto que $|f| \leq \| f \|_\infty$ casi-siempre, $0 < \mu(X) < +\infty$ y $\lim_{p\to\infty} |a|^{1/p} = 1$, resulta que

$$\overline{\lim_{p\to\infty}} \left(\int_X |f|^p \, d\mu \right)^{1/p} \;\leq\; \| f \|_\infty \lim_{p\to\infty} \mu(X)^{1/p} \;=\; \| f \|_\infty .$$

Por otro lado, dado $0 < \varepsilon < \| f \|_\infty$, de las propiedades que definen a $\| f \|_\infty$, (a_∞^2), resulta que

$$\mu\big(\{x \in X \,:\, |f(x)| > \| f \|_\infty - \varepsilon\}\big) \;=\; \delta \;>\; 0,$$

y, en consecuencia,

$$\int_X |f|^p \, d\mu \;\geq\; \int_{[\,|f| > \| f \|_\infty - \varepsilon\,]} |f|^p \, d\mu \;>\; \big(\| f \|_\infty - \varepsilon\big)^p \cdot \delta,$$

lo cual implica que

$$\underline{\lim_{p\to\infty}} \left(\int_X |f|^p \, d\mu \right)^{1/p} \;\geq\; \| f \|_\infty \;-\; \varepsilon$$

Haciendo que $\varepsilon \to 0$, se obtiene la primera igualdad (4).

Para demostrar las igualdades en (5), observe que gracias a que $|f| \leq \| f \|_\infty$ casi-siempre y $\mu(X) < +\infty$, entonces $p_n \leq (\| f \|_\infty)^n \mu(X)$. De esto se sigue que

$$\overline{\lim_{n\to\infty}} \sqrt[n]{p_n} \;\leq\; \| f \|_\infty \lim_{n\to\infty} \mu(X)^{1/n} \;=\; \| f \|_\infty . \tag{6}$$

Por otro lado, sea $\varepsilon > 0$ y hagamos $E = \{x \in X \,:\, |f(x)| > \| f \|_\infty - \varepsilon\}$. Puesto que $\mu(E) > 0$, resulta que

$$\sqrt[n]{p_n} \;\geq\; \left(\int_E |f|^n \, d\mu \right)^{1/n} \;\geq\; (\| f \|_\infty - \varepsilon)\mu(E)^{1/n}$$

y, en consecuencia,

$$\underline{\lim_{n\to\infty}} \sqrt[n]{p_n} \;\geq\; \| f \|_\infty - \varepsilon.$$

Haciendo que $\varepsilon \to 0$, resulta que $\underline{\lim_{n\to\infty}} \sqrt[n]{p_n} \;\geq\; \| f \|_\infty$ lo que junto con (6) nos da que

$$\lim_{n\to\infty} \sqrt[n]{p_n} \;=\; \| f \|_\infty . \tag{7}$$

Para demostrar la igualdad restante, nótese que $p_{n+1} \leq \| f \|_\infty \, p_n$ para todo $n \geq 1$, de donde se sigue que

$$\overline{\lim_{n\to\infty}} \frac{p_{n+1}}{p_n} \;\leq\; \| f \|_\infty . \tag{8}$$

Puesto que los números $p = (n+1)/n$ y $q = n+1$ son exponentes conjugados, la Desigualdad de Hölder nos muestra que

$$p_n \;=\; \int_X |f|^n \, d\mu \;\leq\; \Big(\int_X |f|^{n+1} \, d\mu\Big)^{\frac{n}{n+1}} \Big(\int_X 1 \, d\mu\Big)^{\frac{1}{n+1}} \;=\; p_{n+1}^{\frac{n}{n+1}} \cdot \mu(X)^{\frac{1}{n+1}},$$

y, en consecuencia, $p_n^{n+1} \leq p_{n+1}^n \cdot \mu(X)$. Dividiendo por $\mu(X)$ y luego tomando raíz n-ésima en esta desigualdad resulta que $p_{n+1}/p_n \geq (p_n/\mu(X))^{1/n}$. Esto último, combinado con (7), produce la desigualdad

$$\underline{\lim_{n\to\infty}} \frac{p_{n+1}}{p_n} \;\geq\; \underline{\lim_{n\to\infty}} \sqrt[n]{p_n} \;=\; \| f \|_\infty .$$

que junto a (8) da la igualdad requerida. ■

A diferencia de los espacios $\mathrm{L}_p(X,\mu)$, $1 \leq p < +\infty$, *ni las funciones en escaleras, ni las funciones continuas definidas sobre* X son densas en $\mathrm{L}_\infty(X,\mu)$.

Ejemplo 10.3.5. *Existe una función* $f \in \mathrm{L}_\infty([a,b],\mu)$ *tal que*

$$\| f - \varphi \|_\infty \;\geq\; \frac{1}{2}$$

para toda función en escalera $\varphi : [a,b] \to \mathbb{R}$.

Prueba. Sea $E \subseteq [a,b]$ el conjunto medible obtenido en el Ejercicio 6.8.9, página 361, que posee la siguiente propiedad:

$$\mu(E \cap J) \;>\; 0 \qquad\qquad \text{y} \qquad\qquad \mu(E^c \cap J) \;>\; 0$$

para todo intervalo abierto $J \subseteq [a,b]$. Entonces, la función $f = \chi_E$ cumple lo establecido. En efecto, si φ es una función en escalera que es igual a c sobre $J = (\alpha,\beta) \subseteq [a,b]$, entonces

$$\| f - \varphi \|_\infty \;\geq\; \text{máx}\,\big\{|c|,\ |1-c|\big\} \;\geq\; \frac{1}{2}.$$

Esto finaliza la prueba. ■

Ejemplo 10.3.6. *Existe una función* $f \in \mathrm{L}_\infty([a,b],\mu)$ *tal que*

$$\| f - \varphi \|_\infty \;\geq\; \frac{1}{2}$$

para toda función continua $g : [a,b] \to \mathbb{R}$.

Prueba. Recordemos que $C([a,b])$ representa el espacio de las funciones continuas definidas sobre $[a,b]$ y veamos, como en el ejemplo anterior, que existe una función $f \in \mathrm{L}_\infty([a,b],\mu)$ tal que

$$\| f - g \|_\infty \;\geq\; \frac{1}{2}$$

para toda función $g \in C([a,b])$. Fijemos un $c \in (a,b)$ y considere la función $f = \chi_{[a,c]}$. Si $g : [a,b] \to \mathbb{R}$ es cualquier función continua, entonces $\lim\limits_{x\to c^+} g(x) = \lim\limits_{x\to c^-} g(x) = g(c)$ y, por consiguiente,

$$\| f - g \|_\infty \;\geq\; \text{máx}\,\big\{|g(c)|,\ |1-g(c)|\big\} \;\geq\; \frac{1}{2}.$$

La prueba es completa. ■

A pesar de los resultados negativos obtenidos en los dos ejemplos anteriores, las funciones simples son una excepción.

Teorema 10.3.35 (Densidad en $\mathrm{L}_\infty(X,\mu)$). *$\mathcal{S}_\mu(X)$ es **denso** en $(\mathrm{L}_\infty(X,\mu), \|\cdot\|_\infty)$.*

Prueba. Sea $f \in \mathrm{L}_\infty(X,\mu)$ y sea $\varepsilon > 0$. Vamos a demostrar la existencia de una función simple s tal que $\|f - s\|_\infty < \varepsilon$. Como $f \in \mathrm{L}_\infty(X,\mu)$, existe una función $g : X \to \mathbb{R}$ medible y acotada tal que $f = g$ casi-siempre y $\|f\|_\infty = \|g\|_\infty$. Siendo g acotada, el conjunto $g(X)$ también lo es y, por consiguiente, existe un intervalo cerrado y acotado $[a,b]$ tal que $g(X) \subseteq [a,b]$. Si denotamos $\mathcal{V}$ la colección de todos los intervalos abiertos de longitud ε cuyos centros son los puntos de $[a,b]$, resulta que $\mathcal{V}$ es un cubrimiento abierto de $[a,b]$ el cual se reduce a un subcubrimiento finito por compacidad, digamos $I_1,\ldots,I_n$. Denote los centros de tales intervalos por $a_1,\ldots,a_n$. Seleccione ahora subconjuntos medibles y disjuntos $A_1,\ldots,A_n$ tales que $A_i \subseteq I_i$ para $i = 1,\ldots,n$ y $g(X) \subseteq \bigcup_{i=1}^n A_i$. Esta elección nos provee de una partición medible de X si tomamos $E_i = g^{-1}(A_i)$, para $i = 1,\ldots,n$. Finalmente, si definimos $s : X \to \mathbb{R}$ por

$$s(x) = \sum_{i=1}^{n} a_i \cdot \chi_{E_i}$$

tendremos que s es simple y

$$\|f - s\|_\infty = \|g - s\|_\infty \le \sup_{x\in X} |g(x) - s(x)| < \varepsilon.$$

La prueba es completa. ■

Los espacios $\mathrm{L}_p(\mathbb{R},\mu)$, para cualquier $p \in [1,+\infty)$, son separables, sin embargo $\mathrm{L}_\infty(\mathbb{R},\mu)$ no disfruta de esa propiedad.

Teorema 10.3.36. *$(\mathrm{L}_\infty(\mathbb{R},\mu), \|\cdot\|_\infty)$ **no es separable**.*

Prueba. Considere la siguiente colección en $\mathrm{L}_\infty(\mathbb{R},\mu)$:

$$\mathfrak{R} = \Big\{ f_a = \chi_{[0,a]} \;:\; a \in \mathbb{R}^+ \Big\}.$$

Observe que si $a, b \in \mathbb{R}^+$ con $a \neq b$, entonces

$$\Big\| \chi_{[0,a]} - \chi_{[0,b]} \Big\|_\infty = 1$$

y, en consecuencia, todas las bolas abiertas $U(f_a, 1/2)$ con centro en f_a y radio $1/2$ son disjuntas. Suponga, para generar una contradicción, que $\mathrm{L}_\infty(\mathbb{R},\mu)$ es separable y seleccione un conjunto numerable y denso, digamos $\mathcal{D} \subseteq \mathrm{L}_\infty(\mathbb{R},\mu)$. La densidad de $\mathcal{D}$ nos garantiza que $\mathcal{D}_a = \mathcal{D} \cap U(f_a, 1/2) \neq \varnothing$ para todo $a \in \mathbb{R}^+$. Usemos el Axioma de Elección para elegir, de cada cada uno de los conjuntos $\mathcal{D}_a$, un único punto x_a y con ellos formemos el conjunto $\{x_a : a \in \mathbb{R}^+\}$. Ahora bien, como la familia $\{U(f_a, 1/2) : a \in \mathbb{R}^+\}$ es disjunta, resulta que $x_a \neq x_b$ cualesquiera sean $a, b \in \mathbb{R}^+$ con $a \neq b$ lo cual nos muestra que $\{x_a : a \in \mathbb{R}^+\}$ es un subconjunto no-numerable del conjunto numerable $\mathcal{D}$. Esta contradicción nos dice que $\mathrm{L}_\infty(\mathbb{R},\mu)$ no puede ser separable y termina la prueba. ■

Uno puede extender la noción de exponentes conjugados al caso cuando $p = 1$ o $p = +\infty$, del modo siguiente:

Definición 10.3.37. *Si* $p = 1$, *su* ***exponente conjugado*** *es el número real extendido* $q = +\infty$, *mientras que si* $p = +\infty$, *definimos* $q = 1$.

Con esta definición, la Desigualdad de Hölder y el Teorema 10.3.5 también son válidos para $p \in \{1, +\infty\}$, esto es:

Teorema 10.3.38 (Desigualdad de Hölder). *Sea* $g \in L_\infty(X,\mu)$. *Entonces,* $fg \in L_1(X,\mu)$ *para cualquier* $f \in L_1(X,\mu)$ *y, además, se cumple que* $\| fg \|_1 \leq \| f \|_1 \| g \|_\infty$. *Más aun,*

$$\| g \|_\infty \;=\; \sup\left\{ \left| \int_X fg\, d\mu \right| \;:\; f \in L_1(X,\mu),\ \| f \|_1 \leq 1 \right\}.$$

Prueba. Observe que

$$\| fg \|_1 \;=\; \int_X |fg|\, d\mu \;\leq\; \| g \|_\infty \int_X |f|\, d\mu \;=\; \| g \|_\infty \| f \|_1 .$$

Esto prueba la primera parte. Como antes, sea

$$a_g \;=\; \sup\left\{ \left| \int_X fg\, d\mu \right| \;:\; f \in L_1(X,\mu),\ \| f \|_1 \leq 1 \right\}.$$

Por la desigualdad anterior tenemos

$$\left| \int_X fg\, d\mu \right| \;\leq\; \| f \|_1 \| g \|_\infty$$

y, en consecuencia, $a_g \leq \| g \|_\infty$. Para demostrar la otra desigualdad, nótese que si $\| g \|_\infty = 0$, entonces la igualdad requerida es obvia. Suponga que $\| g \|_\infty > 0$ y sea ε un número real arbitrario tal que $0 < \varepsilon < \| g \|_\infty$. Por (a^2_∞) el conjunto $E = \{x \in X : |g(x)| > \| g \|_\infty - \varepsilon\}$ es de medida positiva y, en consecuencia, la función $f : X \to \mathbb{R}$ definida por

$$f \;=\; \frac{1}{\mu(E)}\, \chi_E \cdot \operatorname{sign}(g).$$

pertenece a $L_1(X,\mu)$ y claramente $\| f \|_1 \leq 1$. Más aun, como $\operatorname{sign}(g) \cdot g = |g|$, resulta que

$$\left| \int_X fg\, d\mu \right| \;\geq\; \int_X fg\, d\mu \;=\; \frac{1}{\mu(E)} \int_E |g|\, d\mu \;\geq\; \| g \|_\infty - \varepsilon.$$

Como $\varepsilon > 0$ es arbitrario, se concluye que $a_g \geq \| g \|_\infty$ y termina la prueba. ■

10.3.8. Convolución en $L_p(\mathbb{R},\mu)$

La Desigualdad de Hölder establece que si p y q son exponentes conjugados reales, entonces el producto $fg \in L_1(X,\mu)$ siempre que $f \in L_p(\mathbb{R},\mu)$ y $g \in L_q(\mathbb{R},\mu)$. Así mismo, si $p = 1$ y $q = +\infty$, también se cumple que $fg \in L_1(X,\mu)$. Sin embargo, si elegimos $f, g \in L_1(\mathbb{R},\mu)$, entonces no siempre ocurre que $fg \in L_1(\mathbb{R},\mu)$, como lo muestra el siguiente ejemplo: la funciones $f, g : \mathbb{R} \to \mathbb{R}$ definidas por

$$f(x) \;=\; g(x) \;=\; \begin{cases} \dfrac{1}{\sqrt{x}} & \text{si } x \in (0,1], \\ 0 & \text{si } x \in \mathbb{R} \setminus (0,1] \end{cases}$$

están en $L_1(\mathbb{R},\mu)$, pero $fg \notin L_1(\mathbb{R},\mu)$ por lo que la operación de multiplicación de funciones en $L_1(\mathbb{R},\mu)$ no es cerrada. Esta desafortunada (o, tal vez, afortunada) circunstancia permitió estudiar otro tipo diferente de "multiplicación" en $L_1(\mathbb{R},\mu)$ que posee el encanto de ser cerrada con respecto a dicha multiplicación y que es de gran importancia en el Análisis de Fourier. Nos referimos a la noción de *convolución de funciones*. Para beneficio del lector desarrollaremos algunas herramientas que hacen más atractiva su presentación.

Definición 10.3.39. *Sea $f:\mathbb{R}\to\mathbb{R}$ una función arbitraria. Para cada $h\in\mathbb{R}$, definimos la* ***traslación de f por h*** *como la función $\tau_h f:\mathbb{R}\to\mathbb{R}$ dada por*

$$(\tau_h f)(x) = f(x-h), \quad x\in\mathbb{R}.$$

Nótese que $(\tau_h f)(x) = (f\circ g)(x)$, donde g es la función continua $g(x)=x-h$ para todo $x\in\mathbb{R}$. Observe, además, que $\tau_h:\mathbb{R}^{\mathbb{R}}\to\mathbb{R}^{\mathbb{R}}$ es una aplicación lineal, esto es,

$$\tau_h(\alpha f_1+\beta f_2) = \alpha\,\tau_h f_1 + \beta\,\tau_h f_2$$

para toda $f_1,f_2\in\mathbb{R}^{\mathbb{R}}$ y todo $\alpha,\beta\in\mathbb{R}$. Algunas propiedades simples de $\tau_h f$ se muestran a continuación.

(a) *Si f es* ***medible***, *entonces $\tau_h f = f\circ g$ también lo es*.

En efecto, como g es un homeomorfismo, $g^{-1}(E)$ es medible para cualquier conjunto medible E. Por esto, para cada $a\in\mathbb{R}$,

$$\{x\in\mathbb{R}:(f\circ g)(x)>a\} = (f\circ g)^{-1}\big((a,+\infty)\big) = g^{-1}\big(f^{-1}\big((a,+\infty)\big)\big)$$

es un conjunto medible ya que $f^{-1}\big((a,+\infty)\big)$ lo es.

(b) *Si f es* ***continua***, *entonces $\tau_h f = f\circ g$ también es continua.* Además,

$$\lim_{h\to 0}(\tau_h f)(x) = f(x) \quad \text{si, y sólo si, } f \text{ es continua en } x.$$

(c) *Si $f\in C_c(\mathbb{R})$, entonces $\tau_h f\in C_c(\mathbb{R})$ y se cumple que $\|\tau_h f\|_\infty = \|f\|_\infty$*.

Esto sigue del hecho de la igualdad $\mathrm{sop}(\tau_h f) = \mathrm{sop}(f)+h$.

Una propiedad muy importante del espacio de Banach $L_1(\mathbb{R},\mu)$ es su invarianza por traslación.

Teorema 10.3.40 (Traslación invariante de $L_1(\mu)$). *Sea $f\in L_1(\mathbb{R},\mu)$. Para cada $h\in\mathbb{R}$, la función $\tau_h f\in L_1(\mathbb{R},\mu)$ y se cumple que*

$$\|f\|_1 = \|\tau_h f\|_1,$$

para todo $h\in\mathbb{R}$.

Prueba. Suponga, en primer lugar, que $f = \chi_E$, donde E es un conjunto medible arbitrario incluido en $\mathbb{R}$. Puesto que μ es invariante por traslación, se tiene que $\mu(E+h) = \mu(E)$ y como $\chi_{E+h}(x) = \chi_E(x-h)$, resulta que

$$\begin{aligned}\int_{\mathbb{R}} f(x)\,\mu(x) &= \int_{\mathbb{R}} \chi_E(x)\,d\mu(x) = \mu(E) = \mu(E+h)\\ &= \int_{\mathbb{R}} \chi_{E+h}(x)\,d\mu(x) = \int_{\mathbb{R}} \chi_E(x-h)\,d\mu(x)\\ &= \int_{\mathbb{R}} f(x-h)\,d\mu(x).\end{aligned}$$

La linealidad de la integral nos muestra que si $f = \sum_{j=1}^{n} c_j\chi_{E_j}$ es una función simple, entonces

$$\begin{aligned}\int_{\mathbb{R}} f(x)\,\mu(x) &= \sum_{j=1}^{n} c_j \int_{\mathbb{R}} \chi_{E_j}\,d\mu\\ &= \sum_{j=1}^{n} c_j \int_{\mathbb{R}} \chi_{E_j}(x-h)\,d\mu\\ &= \int_{\mathbb{R}} f(x-h)\,d\mu(x).\end{aligned}$$

Suponga ahora que $f \geq 0$ y use el Teorema de Densidad de las funciones simples para determinar la existencia de una *sucesión creciente* de funciones simples no-negativas $(\varphi_n)_{n=1}^{\infty}$ tal que $\varphi_n \to f$ puntualmente. De lo anterior y el Teorema de la Convergencia Monótona se deduce que

$$\int_{\mathbb{R}} f(x)\,d\mu = \lim_{n\to\infty} \int_{\mathbb{R}} \varphi_n(x)\,d\mu = \lim_{n\to\infty} \int_{\mathbb{R}} \varphi_n(x-h)\,d\mu = \int_{\mathbb{R}} f(x-h)\,d\mu.$$

Finalmente, si f es una función Lebesgue integrable arbitraria, entonces

$$\begin{aligned}\int_{\mathbb{R}} f\,d\mu &= \int_{\mathbb{R}} f^+\,d\mu - \int_{\mathbb{R}} f^-\,d\mu\\ &= \int_{\mathbb{R}} f^+(x-h)\,d\mu - \int_{\mathbb{R}} f^-(x-h)\,d\mu\\ &= \int_{\mathbb{R}} f(x-h)\,d\mu.\end{aligned}$$

En particular, si reemplazamos f por su valor absoluto $|f|$ en la igualdad anterior, se obtiene que

$$\|f\|_1 = \int_{\mathbb{R}} |f(x)|\,d\mu(x) = \int_{\mathbb{R}} |f(x-h)|\,d\mu(x) = \|\tau_h f\|_1 \qquad (\mathrm{IPT}_1)$$

para todo $h \in \mathbb{R}$. ■

El resultado anterior se puede ampliar a $L_p(\mathbb{R},\mu)$ para cualquier $1 \leq p < +\infty$ como se muestra a continuación.

Corolario 10.3.41 (Invariante por Traslación). *Sea $p \in [1, +\infty)$. Si $f \in L_p(\mathbb{R}, \mu)$, entonces para cada $h \in \mathbb{R}$ se tiene que $\tau_h f \in L_p(\mathbb{R}, \mu)$ y se cumple que*

$$\| f \|_p = \| \tau_h f \|_p .$$

Prueba. Sólo hace falta considerar el caso en que $p > 1$. Suponga entonces que $p > 1$ y sea $f \in L_p(\mathbb{R}, \mu)$. Entonces $|f(x)|^p \in L_1(\mathbb{R}, \mu)$ y, por consiguiente, por el Teorema 10.3.40, resulta que $|f(x-h)|^p \in L_1(\mathbb{R}, \mu)$ y

$$\int_{\mathbb{R}} |f(x)|^p \, d\mu(x) = \int_{\mathbb{R}} |f(x-h)|^p \, d\mu(x).$$

De esto se sigue que $\tau_h f \in L_p(\mathbb{R}, \mu)$ y $\| f \|_p = \| \tau_h f \|_p$. ■

Fijemos $a \in \mathbb{R}$ con $a \neq 0$ y sea $g_a : \mathbb{R} \to \mathbb{R}$ definida por $g_a(x) = ax$. Si ahora se considera la función

$$(\tau_a^* f)(x) = f(ax)$$

para cada $f : \mathbb{R} \to \mathbb{R}$, resulta que $\tau_a^* f$ posee propiedades similares a la de la aplicación $\tau_h f$. Recuerde que $\mu(aE) = |a|\mu(E)$ para cualquier conjunto medible E. Más aun, si $f \in L_1(\mathbb{R}, \mu)$, entonces $\tau_a^* f \in L_1(\mathbb{R}, \mu)$, y se cumple que

$$\int_{\mathbb{R}} f(x) \, d\mu(x) = |a| \int_{\mathbb{R}} f(ax) \, d\mu(x).$$

En particular, si $a = -1$ y reemplazando f por $|f|^p$ en la igualdad anterior, se obtiene:

$$\int_{\mathbb{R}} |f(x)|^p \, d\mu(x) = \int_{\mathbb{R}} |f(-x)|^p \, d\mu(x) \tag{IPT$_2$}$$

para cualquier $p \geq 1$. Puesto que la demostración de estos hecho es casi una copia al carbón del resultado anterior, la omitiremos.

Gracias al Corolario 10.3.41, $\tau_h f \in L_p(\mathbb{R}, \mu)$ para cada $f \in L_p(\mu)$ y cualquier $h \in \mathbb{R}$. Este hecho permite que podamos asociar, a cada $h \in \mathbb{R}$, la aplicación

$$\Phi_h : L_p(\mathbb{R}, \mu) \to L_p(\mathbb{R}, \mu) \quad \text{definida por} \quad \Phi_h(f) = \tau_h f.$$

Observe que, gracias al Teorema 10.3.41, Φ_h es una aplicación lineal que es una isometría, esto es,

$$\| \Phi_h(f) \|_p = \| f \|_p$$

para cada $f \in L_p(\mathbb{R}, \mu)$. Nuestro siguiente objetivo es demostrar el siguiente resultado:

Teorema 10.3.42. *Fijemos $f \in L_p(\mathbb{R}, \mu)$, donde $1 \leq p < +\infty$ y sea $h \in \mathbb{R}$. Si $(h_n)_{n=1}^{\infty}$ es una sucesión en $\mathbb{R}$ convergiendo a h, entonces*

$$\lim_{n\to\infty} \| \tau_{h_n} f - \tau_h f \|_p = \lim_{n\to\infty} \| \Phi_{h_n}(f) - \Phi_h(f) \|_p = 0.$$

Prueba. Observe que es suficiente demostrar el teorema tomando $h = 0$. El ingrediente clave en la demostración de este resultado es suponer, en primer lugar, que $f = \chi_I$, donde I es un intervalo acotado, el cual podemos asumir es de la forma, $I = [a,b]$. Observe que para cada $h \in \mathbb{R}$, $\tau_h f(x) = \chi_{I_h}(x)$ donde hemos puesto $I_h = [a+h, b+h]$. Como $\mu(I_h \Delta I) = 2|h|$, resulta que

$$\begin{aligned} \| \tau_h f(x) - f(x) \|_p &= \left(\int_{\mathbb{R}} |\tau_h f(x) - f(x)|^p \, d\mu(x) \right)^{1/p} \\ &\leq \left(\int_{\mathbb{R}} \chi_{I_h \, \Delta \, I}(x) \, d\mu(x) \right)^{1/p} \\ &= (2|h|)^{1/p}. \end{aligned}$$

Esto prueba (1) para este caso especial de f. Finalmente, si f es una función en escalera, la linealidad de la integral permite obtener la misma conclusión.

Considere ahora el conjunto

$$\mathfrak{X} = \Big\{ f \in \mathrm{L}_p(\mu) \, : \, \lim_{h \to 0} \| \tau_h f - f \|_p = 0 \Big\}.$$

Veamos que $\mathfrak{X}$ es un subespacio vectorial cerrado de $\mathrm{L}_p(\mu)$.

(a) $\mathfrak{X}$ *es un espacio vectorial*. En efecto, sean $f_1, f_2 \in \mathfrak{X}$. Si $\alpha, \beta \in \mathbb{R}$, defina $f = \alpha f_1 + \beta f_2$. De la Desigualdad de Minkowski se tiene que

$$\| \tau_h(f) - f \|_p \leq |\alpha| \, \| \tau_h f_1 - f_1 \|_p + |\beta| \, \| \tau_h f_2 - f_2 \|_p$$

de donde se sigue, tomando límite cuando $h \to 0$, que $f = \alpha f_1 + \beta f_2 \in \mathfrak{X}$.

(b) $\mathfrak{X}$ *es cerrado en* $\mathrm{L}_p(\mu)$. Para ver esto, sea $g \in \overline{\mathfrak{X}}^{\|\cdot\|_p}$. Esto significa que, dado cualquier $\varepsilon > 0$, existe una función $f \in \mathfrak{X}$ tal que $\| g - f \|_p < \varepsilon/4$. Usando el Corolario 10.3.41 se obtiene que

$$\begin{aligned} \| \tau_h g - g \|_p &= \| \tau_h g - \tau_h f + \tau_h f - f + f - g \|_p \\ &\leq \| \tau_h g - \tau_h f \|_p + \| \tau_h f - f \|_p + \| f - g \|_p \\ &= 2 \| f - g \|_p + \| \tau_h f - f \|_p \\ &< \varepsilon/2 + \| \tau_h f - f \|_p, \end{aligned}$$

Ahora bien, puesto que $f \in \mathfrak{X}$, existe un $\delta > 0$ tal que $\| \tau_h f - f \|_p < \varepsilon/2$ siempre que $|h| \leq \delta$. De esto se sigue que si $|h| \leq \delta$, entonces $\| \tau_h g - g \|_p < \varepsilon$, lo cual significa que $g \in \mathfrak{X}$ y, por lo tanto, $\mathfrak{X}$ es cerrado.

Finalmente, como el conjunto de las funciones en escaleras $\mathrm{Esc}(\mathbb{R})$ es norma-denso en $\mathrm{L}_p(\mu)$ y $\mathrm{Esc}(\mathbb{R}) \subseteq \mathfrak{X}$, gracias a la primera parte, se tiene que

$$\mathfrak{X} \subseteq \mathrm{L}_p(\mathbb{R}, \mu) = \overline{\mathrm{Esc}(\mathbb{R})} \subseteq \overline{\mathfrak{X}} = \mathfrak{X},$$

es decir, $\mathfrak{X} = \mathrm{L}_p(\mathbb{R}, \mu)$ y termina la prueba. ■

Ya hemos visto la definición de soporte para funciones continuas definidas sobre $\mathbb{R}$. En lo que sigue, extenderemos dicha definición a funciones medibles las cuales, como ya sabemos, no necesariamente son finitas en todos los puntos de $\mathbb{R}$. Este hecho es importante. Por ejemplo, si $f = \chi_{\mathbb{Q}}$, entonces $\{x \in \mathbb{R} : f(x) \neq 0\} = \mathbb{Q}$ y, por lo tanto, $\mathrm{sop}(f) = \mathbb{R}$. Por otro lado, si tomamos $g = 0$, resulta que $f = g$ casi-siempre y, en consecuencia, deberíamos tener, mirando a f como un elemento de $\mathrm{L}_1(\mu)$ que $\mathrm{sop}(f) = \mathrm{sop}(g)$ lo cual es falso. Esta observación sugiere tener cierta precaución con la definición de $\mathrm{sop}(f)$ si f es finita casi-siempre.

Definición 10.3.43. *Sea $f : \mathbb{R} \to \mathbb{R}$ una función medible. Sea $\mathcal{V} = \{V_\alpha : \alpha \in J\}$ la familia de todos subconjuntos abiertos de $\mathbb{R}$ tal que para cada $\alpha \in J$, $f = 0$ casi-siempre en V_α. Definimos el **soporte** de f como el conjunto cerrado*

$$\mathrm{sop}(f) \;=\; \mathbb{R} \setminus \bigcup_{\alpha} V_\alpha.$$

Observe que $V_0 = \bigcup_\alpha V_\alpha$ es el conjunto abierto más grande sobre el cual $f = 0$ casi-siempre y que si $f = g$ casi-siempre, entonces $\mathrm{sop}(f) = \mathrm{sop}(g)$.

Ya es hora de definir la convolución de funciones medibles. Usualmente, la convolución de dos funciones f y g, que será representada por el símbolo $f * g$, posee propiedades “refinadas” y “agradables” aun si ninguna de ellas las posee. Por ejemplo, veremos que si $f \in \mathrm{L}_p(\mu)$ y $g \in \mathrm{L}_q(\mu)$ donde p, q son exponentes conjugados reales, entonces $f * g$ es continua aun si f, g son ambas discontinuas. De hecho, esta noción suele ser usada, entre otras cosas, para aproximar funciones que pueden tener un “*mal comportamiento*” respecto a alguna propiedad.

Definición 10.3.44. *Sean $f, g : \mathbb{R} \to \mathbb{R}$ funciones medibles. La **convolución** de f y g se define como la función $f * g : \mathbb{R} \to \mathbb{R}$ dada por*

$$(f * g)(x) \;=\; \int_{\mathbb{R}} f(x-y)g(y)\, d\mu(y)$$

para todo $x \in \mathbb{R}$ para el cual la integral exista.

Es importante insistir en que la expresión $(f * g)(x)$ puede no existir para todo $x \in \mathbb{R}$. En efecto, en el siguiente ejemplo se muestra una función $f \in \mathrm{L}_1(\mu)$ tal que $(f * f)(x)$ no existe para ciertos valores $x \in \mathbb{R}$.

Ejemplo 10.3.7. *Elija, usando el Ejemplo 10.3.4, una función $f \in \mathrm{L}_1(\mathbb{R}, \mu) \setminus \mathrm{L}_2(\mathbb{R}, \mu)$ y suponga que f es par, es decir, $f(x) = f(-x)$ para todo $x \in \mathbb{R}$. Como se prueba un poco más abajo, la condición $f \in \mathrm{L}_1(\mathbb{R}, \mu)$ implica que $(f * f)(x)$ existe para casi-todo $x \in \mathbb{R}$. Sin embargo, como f es par y $f \notin \mathrm{L}_2(\mathbb{R}, \mu)$, entonces*

$$(f * f)(0) \;=\; \int_{\mathbb{R}} f(-y)f(y)\, d\mu(y) \;=\; \int_{\mathbb{R}} |f(y)|^2\, d\mu(y) \;=\; +\infty.$$

Otra observación que puede ser de interés es la siguiente: si f y g son funciones medibles tales $f * g$ existe, entonces

$$\mathrm{sop}(f * g) \;\subseteq\; \overline{\mathrm{sop}(f) \,+\, \mathrm{sop}(g)}.$$

En efecto, suponga que $x \notin \text{sop}(f) + \text{sop}(g)$. Entonces, para todo $y \in \text{sop}(g)$ se tiene que $(x-y) \notin \text{sop}(f)$ y, por consiguiente, $f(x-y) = 0$. De esto se sigue que

$$(f * g)(x) \;=\; \int_{\mathbb{R}} f(x-y)g(y)\,d\mu(y) \;=\; 0.$$

Puesto que el soporte de una función es el complemento del abierto más grande sobre el cual ella se anula, se deduce que el interior de $(\text{sop}(f) + \text{sop}(g))^c$ está incluido en el complemento de $\text{sop}(f * g)$ y, por lo tanto,

$$\text{sop}(f * g) \;\subseteq\; \overline{\text{sop}(f) + \text{sop}(g)}.$$

Nuestro objetivo será investigar para qué clases de funciones medibles a las cuales pertenecen f, g, la expresión $(f * g)(x)$ existe casi-siempre. Una primera aproximación lo constituye el siguiente resultado.

Teorema 10.3.45 (Convolución en $L_\infty(\mathbb{R},\mu)$). *Sean $p, q \in [1, +\infty]$ exponentes conjugados y suponga que $f \in L_p(\mathbb{R},\mu)$ y $g \in L_q(\mathbb{R},\mu)$. Entonces*

(a) $f * g$ es continua en $\mathbb{R}$,

(b) $f * g \in L_\infty(\mathbb{R},\mu)$ y

$$\| f * g \|_\infty \;\leq\; \| f \|_p \, \| g \|_q . \tag{2}$$

Prueba. (a) Fijemos $x \in \mathbb{R}$ y sea $(x_n)_{n=1}^\infty$ una sucesión que converge a x. Entonces, el cambio de variable $z \leftrightarrow -z$, la Desigualdad de Hölder y el Teorema 10.3.42 nos muestran que cuando $n \to \infty$,

$$\begin{aligned} |(f * g)(x_n) - (f * g)(x)| \;&=\; \left| \int_{\mathbb{R}} f(x_n - y)g(y)\,dy \;-\; \int_{\mathbb{R}} f(x-y)g(y)\,dy \right| \\ &\leq\; \int_{\mathbb{R}} |f(y - x_n) - f(y-x)|\,|g(-y)|\,dy \\ &\leq\; \| \Phi_{x_n}(f) - \Phi_x(f) \|_p \, \| g \|_q \;\to\; 0 \end{aligned}$$

Esto termina la prueba de que $f * g$ es continua en $\mathbb{R}$.
(b) La Desigualdad de Hölder combinada con el Corolario 10.3.41 conducen a

$$|(f * g)(x)| \;\leq\; \int_{\mathbb{R}} |f(x-y)g(y)|\,d\mu(y) \;\leq\; \| f \|_p \, \| g \|_q$$

para todo $x \in \mathbb{R}$. Esto prueba que $f * g \in L_\infty(\mathbb{R},\mu)$ y (2) se cumple. ■

El resultado anterior puede ser mejorado. Antes de hacerlo, probaremos el resultado fundamental acerca de la convolución en $L_1(\mathbb{R},\mu)$, es decir, que dicha operación es cerrada en ese espacio.

Teorema 10.3.46 (Convolución en $L_1(\mathbb{R},\mu)$). *Si $f, g \in L_1(\mathbb{R},\mu)$, entonces se cumple que:*

(a) La aplicación $H : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ definida por $H(x,y) = f(x-y)g(y)$ para todo $(x,y) \in \mathbb{R} \times \mathbb{R}$, es ***medible*** *sobre $\mathbb{R} \times \mathbb{R}$.*

*(b) Para casi-todo $x \in \mathbb{R}$, la aplicación $y \to f(x-y)g(y)$ es **medible** y **Lebesgue integrable** sobre $\mathbb{R}$ y, en consecuencia, $(f * g)(x)$ está bien definida para casi-todo $x \in \mathbb{R}$.*

(c) $f * g \in \mathrm{L}_1(\mathbb{R}, \mu)$ y

$$\| f * g \|_1 \;\leq\; \| f \|_1 \, \| g \|_1 .$$

Prueba. (a) Si definimos $h(x,y) = f(x)$ para todo $(x,y) \in \mathbb{R} \times \mathbb{R}$, resulta que

$$h^{-1}((\alpha, +\infty)) \;=\; \big\{(x,y) : h(x,y) > \alpha\big\} \;=\; \big\{(x,y) : f(x) > \alpha\big\} \;=\; f^{-1}((\alpha,+\infty)) \times \mathbb{R}$$

el cual es un subconjunto medible de $\mathbb{R} \times \mathbb{R}$ ya que $f^{-1}((\alpha,+\infty))$ y $\mathbb{R}$ son medibles. Similarmente, la función $k(x,y) = g(y)$ es medible. Puesto que el producto de funciones medibles es medible, se concluye que $F(x,y) = f(x)g(y)$ es medible. Más aun, si consideramos la aplicación lineal $T(x,y) = (x-y, y)$, resulta que ella es medible ya que es continua y, por consiguiente,

$$H(x,y) \;=\; (F \circ T)(x,y) \;=\; F(x-y, y) \;=\; f(x-y)g(y)$$

es medible.

(b) En primer lugar, observe que gracias al Teorema de Fubini-Tonelli y el Teorema 10.3.40, se tiene que

$$\begin{aligned}
\iint_{\mathbb{R}\times\mathbb{R}} |H(x,y)|\, dx\, dy \;&=\; \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x-y)|\, dx \right) |g(y)|\, dy \\
&=\; \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x)|\, dx \right) |g(y)|\, dy \\
&=\; \| f \|_1 \, \| g \|_1 ,
\end{aligned}$$

es decir, $H(x,y) = f(x-y)g(y) \in \mathrm{L}_1(\mathbb{R} \times \mathbb{R})$. Un llamado al Teorema de Fubini nos revela que $\int_{\mathbb{R}} f(x-y)g(y)\, dy$ existe y es integrable para casi-todo $x \in \mathbb{R}$.

(c) Usando la parte (b), se tiene que

$$\| f * g \|_1 \;=\; \int_{\mathbb{R}} |(f*g)(x)|\, dx \;=\; \int_{\mathbb{R}} \left| \int_{\mathbb{R}} f(x-y)g(y)\, dy \right| dx \;\leq\; \| f \|_1 \, \| g \|_1 .$$

La prueba es completa. ■

El resultado anterior establece que la operación de convolución en $\mathrm{L}_1(\mathbb{R}, \mu)$ es cerrada, lo que escribiremos brevemente por

$$\mathrm{L}_1(\mathbb{R}, \mu) * \mathrm{L}_1(\mathbb{R}, \mu) \;\subseteq\; \mathrm{L}_1(\mathbb{R}, \mu).$$

Esa operación, sin embargo, hace que $\mathrm{L}_1(\mathbb{R}, \mu)$ se comporte como un espacio privilegiado ya que no es verdad que para $p > 1$ dicha operación sea cerrada en $\mathrm{L}_p(\mathbb{R}, \mu)$. Esto último se podrá constatar en el siguiente resultado, el cual extiende el Teorema 10.3.45.

Teorema 10.3.47 (Desigualdad de Young para convoluciones). *Si $1 \leq p \leq +\infty$, entonces*

$$\mathrm{L}_p(\mathbb{R}, \mu) * \mathrm{L}_1(\mathbb{R}, \mu) \;\subseteq\; \mathrm{L}_p(\mathbb{R}, \mu)$$

y se cumple que

$$\| f * g \|_p \;\leq\; \| f \|_p \, \| g \|_1 \tag{3}$$

para toda $f \in \mathrm{L}_p(\mathbb{R}, \mu)$ y toda $g \in \mathrm{L}_1(\mathbb{R}, \mu)$.

Prueba. Los casos $p = 1$ y $p = +\infty$ ya fueron considerados en los Teoremas 10.3.46 y 10.3.45 respectivamente, por lo que sólo resta demostrar el resultado para $1 < p < +\infty$. Sea q el exponente conjugado de p y considere funciones arbitrarias $f \in L_p(\mathbb{R},\mu)$ y $g \in L_1(\mathbb{R},\mu)$. Ahora, por la Desigualdad de Hölder se tiene que

$$\begin{aligned}
|(f*g)(x)| &\le \int_{\mathbb{R}} |f(x-y)|\,|g(y)|\,dy \\
&= \int_{\mathbb{R}} |f(x-y)|\,|g(y)|^{1/p}\,|g(y)|^{1/q}\,dy \\
&\le \left(\int_{\mathbb{R}} |f(x-y)|^p\,|g(y)|\,dy\right)^{1/p}\left(\int_{\mathbb{R}} |g(y)|\,dy\right)^{1/q} \\
&= \left(\int_{\mathbb{R}} |f(x-y)|^p\,|g(y)|\,dy\right)^{1/p} \|g\|_1^{1/q}
\end{aligned}$$

y así, gracias al Teorema de Fubini y al Corolario 10.3.41 se logra finalmente que

$$\begin{aligned}
\|f*g\|_p^p &= \int_{\mathbb{R}} |(f*g)(x)|^p\,dx \\
&\le \|g\|_1^{p/q} \int_{\mathbb{R}}\int_{\mathbb{R}} |f(x-y)|^p\,|g(y)|\,dy\,dx \\
&= \|g\|_1^{p/q} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x-y)|^p\,dx\right)|g(y)|\,dy \\
&= \|g\|_1^{p/q}\,\|f\|_p^p\,\|g\|_1 \;=\; \|g\|_1^p\,\|f\|_p^p
\end{aligned}$$

de donde se obtiene (3) y finaliza la prueba. ■

Nota Adicional 10.3.15 La continuidad de la operación de convolución dada en el Teorema 10.3.45 (a) en combinación con el Teorema 10.3.46 permiten, una vez más, deleitarnos con una nueva prueba del Teorema de Steinhaus:

Teorema de Steinhaus. *Sean $E, F \in \mathfrak{M}_\mu(\mathbb{R})$ ambos de medida positiva. Entonces $E + F$ contiene un intervalo abierto. En particular, tomando $F = -E$ se tiene que $(-\delta,\delta) \subseteq E - E$ para algún $\delta > 0$.*

Prueba. Puesto que $E = \bigcup_{n=1}^{\infty} \big([-n,n] \cap E\big)$, entonces no se pierde generalidad en suponer que E está acotado y, por consiguiente, $0 < \mu(E) < +\infty$. Similarmente, asumiremos que F es de medida finita. Defina ahora $f = \chi_E$ y $g = \chi_F$. Puesto que E, F son de medida finita, resulta que $f, g \in L_1(\mathbb{R},\mu)$ y, por lo tanto, $f * g \in L_1(\mathbb{R},\mu)$. Más aun, el Teorema de Fubini y el hecho de que μ es invariante por traslación, nos aseguran que

$$\int_{\mathbb{R}} (f*g)(x)\,dx \;=\; \mu(E)\cdot\mu(F) \;>\; 0.$$

De esto se sigue que $f * g$ no es idénticamente cero, por lo que podemos seleccionar un $x_0 \in \mathbb{R}$ tal que $(f*g)(x_0) \neq 0$. Usemos ahora el hecho de que $f * g$ es continua en $\mathbb{R}$, para hallar un intervalo abierto J conteniendo a x_0 tal que $(f * g)(x) \neq 0$ para todo $x \in J$. Veamos que J es el intervalo indicado. En efecto, si $x \in J$, entonces

$$(f * g)(x) = \int_{\mathbb{R}} \chi_E(x-y)\chi_F(y)\, dy \neq 0$$

y, por lo tanto, $\chi_E(x-y)\chi_F(y) \neq 0$ para algún $y \in \mathbb{R}$. Para este y se tiene que $x = (x-y) + y \in E + F$ y, por lo tanto, $I \subseteq E + F$. Fin de la prueba. ■

Algunas otras propiedades de la convolución que son de utilidad se muestran el siguiente resultado.

Teorema 10.3.48. *Sean $f, g, h \in \mathrm{L}_1(\mathbb{R}, \mu)$. Entonces:*

(a) $f * g = g * f$.

(b) $f * (g * h) = (f * g) * h$.

(c) $(a \cdot f + b \cdot g) * h = a \cdot (f * h) + b \cdot (g * h)$ para todo $a, b \in \mathbb{R}$

Prueba. (a) El cambio de variable $t = x - y$ conduce a

$$(f * g)(x) = \int_{\mathbb{R}} f(x-y)g(y)\, d\mu(y) = \int_{\mathbb{R}} g(x-t)f(t)\, d\mu(t) = (g * f)(x).$$

Las otras pruebas, por ser muy similares, se dejan a cargo del lector. ■

Nos ocuparemos ahora de una aplicación muy importante del producto de convoluciones: la así llamada *solución fundamental* para un operador diferencial con coeficientes constantes. Recordemos que si $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$, la derivada parcial

$$(D^\alpha f)(x) = \frac{\partial^{|\alpha|} f(x)}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}}$$

donde $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n$, se usa para definir el operador $T : C_c^k(\mathbb{R}^n) \to \mathrm{L}_1(\mathbb{R}^n, \mu_n)$ por

$$T(f) = \sum_{|\alpha| \leq k} c_\alpha D^\alpha f. \tag{3}$$

No es difícil establecer que:

(a) $T(f) \in \mathrm{L}_1(\mathbb{R}^n, \mu_n)$ y

(b) $T(f * g) = T(f) * g$ para cualquier $f \in C_c^k(\mathbb{R}^n)$ y cualquier $g \in \mathrm{L}_1(\mathbb{R}^n, \mu_n)$.

Aquí está el punto clave: si existiese una función $g \in \mathrm{L}_1(\mathbb{R}^n, \mu_n)$ tal que $g * f = f$ para cada $f \in \mathrm{L}_1(\mathbb{R}^n, \mu_n)$, entonces tendríamos un método para resolver, en $\mathbb{R}^n$, el problema diferencial del tipo $Tu = f$, donde T es el operador dado en (3). Sin embrago, como veremos de inmediato, una tal función g nunca existe, es decir, en $\mathrm{L}_1(\mathbb{R}^n, \mu_n)$ no existe *la unidad* con la operación de convolución.

Teorema 10.3.49. *No existe **ninguna** función $g \in \mathrm{L}_1(\mathbb{R}, \mu)$ tal que $g * f = f$ para cada $f \in \mathrm{L}_1(\mathbb{R}, \mu)$.*

Prueba. Suponga, para generar una contradicción, que existe $g \in L_1(\mathbb{R}, \mu)$ tal que $g * f = f$ para cada $f \in L_1(\mathbb{R}, \mu)$. Por el Teorema 10.2.7, página 581, existe un $\delta > 0$ tal que

$$\int_E |g| \, d\mu \ < \ 1$$

para cada conjunto medible E tal que $\mu(E) < \delta$. Sea $F = [-\delta/4, \delta/4]$ y considere la función $f = \chi_F$. Observe ahora que

$$f(x) \ = \ (g * f)(x) \ = \ \int_{\mathbb{R}} g(x-y) f(y) \, dy \ = \ \int_F g(x-y) \, dy \ = \ \int_F g(y) \, dy$$

y como $\mu(F) < \delta$, resulta entonces que para todo $x \in F$,

$$|f(x)| \ = \ |(g * f)(x)| \ \leq \ \int_F |g(y)| \, dy \ < \ 1$$

lo cual es absurdo ya que $f(x) = 1$ para todo $x \in F$. Esto termina la prueba. ■

10.4. Ejercicios Resueltos

Ejercicio 10.4.1. *Sea E un subconjunto medible de $\mathbb{R}$ con $\mu(E) < +\infty$ y sea $\varepsilon > 0$. Demuestre que existen una* ***función en escalera*** *$\varphi : \mathbb{R} \to \mathbb{R}$ y una función $\psi \in C_c(\mathbb{R})$ tales que*

$$\int_{\mathbb{R}} |\varphi - \chi_E| \, d\mu \ < \ \varepsilon \qquad \text{y} \qquad \int_{\mathbb{R}} |\psi - \chi_E| \, d\mu \ < \ \varepsilon.$$

Prueba. Como $\mu(E) < +\infty$ podemos usar el *Primer Principio de Littlewood*, Teorema 6.3.62, página 298, para hallar un conjunto U, que es unión finita de intervalos abiertos y disjuntos dos a dos, tal que $\mu(E \Delta U) < \varepsilon$. Si ahora definimos $\varphi = \chi_U$, resulta que $\varphi \in \mathrm{Esc}(\mathbb{R})$ y se cumple que

$$\| \varphi - \chi_E \|_1 \ = \ \int_{\mathbb{R}} |\chi_U - \chi_E| \, d\mu \ = \ \mu(E \Delta U) \ < \ \varepsilon.$$

Esto prueba la primera parte. Para demostrar la segunda parte, usemos la regularidad de μ para elegir un conjunto compacto K y un conjunto abierto G tales que

$$K \subseteq E \subseteq G \qquad \text{y} \qquad \mu(G \setminus K) \ < \ \varepsilon.$$

Por el Lema de Urysohn, página 159, existe una función $\psi \in C_c(\mathbb{R})$ tal que

$$0 \ \leq \ \psi \ \leq \ 1, \qquad \psi \ = \ 1 \quad \text{sobre } K \qquad \text{y} \qquad \psi \ = \ 0 \quad \text{fuera de } G.$$

De esto se sigue que $|\psi - \chi_E|$ se anula fuera de $G \setminus K$ y que $|\psi - \chi_E| \leq 1$ sobre $G \setminus K$. De allí que

$$\int_{\mathbb{R}} |\psi - \chi_E| \, d\mu \ = \ \int_{G \setminus K} |\psi - \chi_E| \, d\mu \ \leq \ \mu(G \setminus K) \ < \ \varepsilon.$$

Fin de la prueba. ■

Ejercicio 10.4.2. *Sea* $f \in \mathrm{L}_1(\mu)$*. Si* $\left(E_n\right)_{n=1}^{\infty}$ *es una **sucesión creciente** de conjuntos medibles, pruebe que*

$$\lim_{n\to\infty} \int_{E_n} f\,d\mu$$

existe.

Prueba. Sea $E = \bigcup_{n=1}^{\infty} E_n$. Como $f \in \mathrm{L}_1(\mu)$, entonces para todo $n \in \mathbb{N}$

$$\left| \int_E f\,d\mu - \int_{E_n} f\,d\mu \right| = \left| \int_{E\setminus E_n} f\,d\mu \right| \le \int_{E\setminus E_n} |f|\,d\mu < +\infty.$$

Por otro lado, puesto que $\nu(A) = \int_A |f|\,d\mu$ es una medida y la sucesión $(E \setminus E_n)_{n=1}^{\infty}$ es decreciente y $\nu(E \setminus E_1) \le \nu(\mathbb{R}) = \int_{\mathbb{R}} |f|\,d\mu < \infty$, entonces

$$\lim_{n\to\infty} \left| \int_E f\,d\mu - \int_{E_n} f\,d\mu \right| \le \lim_{n\to\infty} \int_{E\setminus E_n} |f|\,d\mu = \lim_{n\to\infty} \nu(E \setminus E_n) = 0,$$

y termina la prueba. ■

Ejercicio 10.4.3. *Sea* $\mathcal{U} \subseteq \mathrm{L}_1(X,\mu)$*. Son equivalentes:*

(1) $\mathcal{U}$ *es **uniformemente integrable**.*

(2) *Para cada* $\varepsilon > 0$*, existe un* $\delta > 0$ *tal que para cada* $f \in \mathcal{U}$*, si* V *es un subconjunto abierto de* $\mathbb{R}$ *y* $\mu(X \cap V) < \delta$*, entonces*

$$\int_{X\cap V} |f|\,d\mu < \varepsilon.$$

Prueba. Claramente, si $\mathcal{U}$ es uniformemente integrable, entonces (2) se cumple. Suponga ahora que la condición (2) es cierta y sea $\varepsilon > 0$. Seleccione un $\delta > 0$ tal que para cada $f \in \mathcal{U}$, si V es un conjunto abierto y $\mu(X \cap V) < \delta$, entonces $\int_{X\cap V} |f|\,d\mu < \varepsilon$. Sea $A \subseteq X$ medible con $\mu(A) < \delta$. Escojamos un conjunto abierto V conteniendo a A tal que $\mu(V \setminus A) < \delta - \mu(A)$. Entonces

$$\mu(X \cap V) \le \mu(V) = \mu(A) + \mu(V \setminus A) < \delta$$

y como $A \subseteq X \cap V$, resulta que

$$\int_A |f|\,d\mu < \int_{X\cap V} |f|\,d\mu < \varepsilon.$$

La prueba es completa. ■

Ejercicio 10.4.4. *Sea* $\mathcal{U} \subseteq \mathrm{L}_1([0,1],\mu)$*. Son equivalentes:*

(1) $\mathcal{U}$ *es **uniformemente integrable**.*

(2) *Para cada* $\varepsilon > 0$*, existe un* $c_0 > 0$ *tal que para todo* $c > c_0$*,*

$$\left| \int_{[\,|f|>c\,]} f\,d\mu \right| < \varepsilon \quad \textit{para toda} \quad f \in \mathcal{U}.$$

Prueba. Puesto que $|\int_E f\,d\mu| \leq \int_E |f|\,d\mu$ para cualquier conjunto medible E, resulta que (1) implica (2). Que (2) $\Rightarrow$ (1) sigue de la observación siguiente:

$$\begin{aligned}\int_E |f|\,d\mu \;&\leq\; \int_{E\cap[f\leq 0]}(-f)\,d\mu \;+\; \int_{E\cap[f>0]} f\,d\mu\\ &=\; \Big|\int_{E\cap[f\leq 0]} f\,d\mu\Big| \;+\; \Big|\int_{E\cap[f>0]} f\,d\mu\Big|.\end{aligned}$$

Ejercicio 10.4.5. *Sean* $\mathcal{U},\mathcal{V}\subseteq \mathrm{L}_1([0,1],\mu)$ *conjuntos uniformemente integrables. Demuestre que:*

(a) $\overline{\mathcal{U}}$ *y* $\mathcal{U}\cup\mathcal{V}$ *son **uniformemente integrables**.*

(b) $\mathrm{co}(\mathcal{U})$ *es **uniformemente integrable**, donde*

$$\mathrm{co}(\mathcal{U}) \;=\; \big\{t_1f_1+\cdots+t_nf_n \,:\, f_i\in\mathcal{U},\ t_i\geq 0,\ t_1+\cdots+t_n=1,\ n\in\mathbb{N}\big\}.$$

(c) $\mathcal{U}+\mathcal{V}$ y $\lambda\cdot\mathcal{U}$ *son **uniformemente integrables**.*

Prueba. (a) es consecuencia inmediata de la definición. Para demostrar (b), sea $\varepsilon>0$ y seleccione un $\delta>0$ tal que

$$\sup_{f\in\mathcal{U}}\int_A |f|\,d\mu \;<\; \varepsilon$$

para cualquier conjunto medible $A\subseteq[0,1]$ con $\mu(A)<\delta$. Sean $f_1,\ldots,f_n\in\mathcal{U}$, $t_1,\ldots,t_n$ números reales no-negativos tal que $t_1+\cdots+t_n=1$ y A un conjunto medible con $\mu(A)<\delta$. Entonces

$$\int_A |t_1f_1+\cdots+t_nf_n|\,d\mu \;\leq\; t_1\int_A|f_1|\,d\mu+\cdots+t_n\int_A|f_n|\,d\mu \;\leq\; \varepsilon.$$

Esto prueba que $\mathrm{co}(\mathcal{U})$ es uniformemente absolutamente continuo y como él es claramente acotado, tal conjunto es uniformemente integrable.

(c) Una demostración simple de que $\mathcal{U}+\mathcal{V}$ es uniformemente integrable es como sigue: puesto que $\mathcal{U}\cup\mathcal{V}$ es uniformemente integrable, entonces (b) nos asegura que su cápsula convexa $\mathrm{co}(\mathcal{U}\cup\mathcal{V})$ también es uniformemente integrable y, en consecuencia, de la inclusión

$$\frac{1}{2}(\mathcal{U}+\mathcal{V}) \;\subseteq\; \mathrm{co}(\mathcal{U}\cup\mathcal{V})$$

se sigue que $\mathcal{U}+\mathcal{V}$ es uniformemente integrable. El hecho de que $\lambda\cdot\mathcal{U}$ es uniformemente integrable es trivial. ■

Ejercicio 10.4.6. *Sea* Γ *el conjunto ternario de Cantor en* $[0,1]$ *y considere la función* $f:[0,1]\to\mathbb{R}$ *dada por*

$$f(x) \;=\; \begin{cases}0 & si\quad x\in\Gamma,\\ n & si\quad x\in J_n,\quad n=1,2,\ldots,\end{cases}$$

donde cada J_n *es la unión disjunta de los* 2^n-1 *intervalos abiertos borrados en el* n*-ésimo paso en la construcción de* Γ*. Demuestre que*

$$\int_{[0,1]} f\,d\mu \;=\; 3.$$

Prueba. Puesto que $\mu(\Gamma) = 0$ y $f \geq 0$, el Corolario 10.1.29 nos asegura que

$$\int_\Gamma f \, d\mu \; = \; 0.$$

Esto, combinado con el hecho de que la medida de cada uno de los conjuntos J_n es $2^{n-1}(1/3^n)$, más el Teorema de Beppo Levi conduce a

$$\int_{[0,1]} f \, d\mu \; = \; \int_{[0,1]\setminus\Gamma} f \, d\mu \; = \; \int_{\bigcup_{n=1}^\infty J_n} f \, d\mu \; = \; \sum_{n=1}^{\infty} \int_{J_n} n \, d\mu \; = \; \sum_{n=1}^{\infty} n \frac{2^{n-1}}{3^n} \; = \; 3.$$

Ejercicio 10.4.7. *Sea X un conjunto medible con $\mu(X) < +\infty$ y para cada par de funciones medibles $f, g : X \to \mathbb{R}$ defina*

$$\rho(f,g) \; = \; \int_X \frac{|f-g|}{1+|f-g|} \, d\mu.$$

Identificando funciones que coinciden casi-siempre resulta que

(a) ρ es una métrica sobre $\mathcal{F}_\mu(X)$.

(b) Pruebe que la convergencia en $\mathcal{F}_\mu(X)$ respecto a la métrica ρ es equivalente a la convergencia en medida, es decir, si $(f_n)_{n=1}^\infty$ es una sucesión en $\mathcal{F}_\mu(X)$ y $f \in \mathcal{F}_\mu(X)$, entonces $\rho(f_n, f) \to 0$ si, y sólo si, $f_n \to f$ en medida.

Prueba. (a) Es fácil establecer que la relación

$$f \sim g \qquad \Leftrightarrow \qquad f \; = \; g \text{ c.s.}$$

es una relación de equivalencia sobre $\mathcal{F}_\mu(X)$. Si identificamos cada función f con su clase de equivalencia $[f]$, entonces ρ es una métrica sobre $\mathcal{F}_\mu(X)$.

(i) $\rho(f,g) \geq 0$ para todo $f, g \in \mathcal{F}_\mu(X)$ sigue de la definición.

(ii) $\rho(f,g) = 0 \Leftrightarrow f = g$. En efecto,

$$\begin{aligned}
\rho(f,g) \; = \; 0 \quad &\Leftrightarrow \quad \int_X \frac{|f-g|}{1+|f-g|} \, d\mu \; = \; 0 \\
&\Leftrightarrow \quad \frac{|f-g|}{1+|f-g|} \; = \; 0 \\
&\Leftrightarrow \quad f \; = \; g
\end{aligned}$$

(iii) $\rho(f,g) = \rho(g,f)$ pues

$$\rho(f,g) \; = \; \int_X \frac{|f-g|}{1+|f-g|} \, d\mu \; = \; \int_X \frac{|g-f|}{1+|g-f|} \, d\mu \; = \; \rho(g,f).$$

(iv) $\rho(f,h) \leq \rho(f,g) + \rho(g,h)$ para toda $f, g, h \in \mathcal{F}_\mu(X)$. Para ver esto, considere la función $u : [0, +\infty) \to \mathbb{R}$ definida por

$$u(t) \; = \; \frac{t}{1+t} \quad \text{para todo } t \geq 0.$$

En primer lugar observe que:

(1) $u(t) \leq 1$ para todo $t \geq 0$, y

(2) u es creciente y no-negativa sobre $[0, +\infty)$.

Puesto que $u'(t) = 1/(1+t)^2$ para todo $t \geq 0$ se sigue lo afirmado en (2). Usando este hecho vemos que para todo $x \in X$,

$$u(|f(x) - h(x)|) \leq u\big(|f(x) - g(x)| + |g(x) - h(x)|\big),$$

es decir,

$$\begin{aligned}
\frac{|f-h|}{1+|f-h|} &\leq \frac{|f-g| + |g-h|}{1+|f-g|+|g-h|} \\
&= \frac{|f-g|}{1+|f-g|+|g-h|} + \frac{|g-h|}{1+|f-g|+|g-h|} \\
&\leq \frac{|f-g|}{1+|f-g|} + \frac{|g-h|}{1+|g-h|}.
\end{aligned}$$

Integrando sobre X se obtiene que $\rho(f,h) \leq \rho(f,g) + \rho(g,h)$. Esto termina la prueba de (a).

Para demostrar (b) no se pierde generalidad si se asume que $\mu(X) = 1$. Suponga, en primer lugar, que $f_n \to f$ en medida y sea $\varepsilon > 0$. Entonces existe un $N \in \mathbb{N}$ tal que para todo $n \geq N$ se cumple que

$$\mu\big(\{x \in X : |f_n(x) - f(x)| \geq \varepsilon\}\big) < \varepsilon.$$

Para cada $n \geq 1$, consider el conjunto $A_\varepsilon^n = \{x \in X : |f_n(x) - f(x)| \geq \varepsilon\}$. Entonces, usando el hecho de que $u(t) \leq 1$ para todo $t \geq 0$ y es creciente, se obtienen las siguientes desigualdades:

(α) $x \in A_\varepsilon^n \Rightarrow u(|f_n(x) - f(x)|) = \dfrac{|f_n(x) - f(x)|}{1+|f_n(x) - f|(x)} \leq 1.$

(β) $x \in X \setminus A_\varepsilon^n \Rightarrow |f_n(x) - f(x)| < \varepsilon \Rightarrow u(|f_n(x) - f(x)|) \leq u(\varepsilon)$, esto es,

$$\frac{|f_n(x) - f(x)|}{1+|f_n(x) - f|(x)} \leq \frac{\varepsilon}{1+\varepsilon} < \varepsilon.$$

De allí que, para cualquier $n \geq N$,

$$\begin{aligned}
\rho(f_n, f) &= \int_X \frac{|f_n - f|}{1+|f_n - f|}\, d\mu \\
&= \int_{A_\varepsilon^n} \frac{|f_n - f|}{1+|f_n - f|}\, d\mu + \int_{X\setminus A_\varepsilon^n} \frac{|f_n - f|}{1+|f_n - f|}\, d\mu \\
&\leq \int_{A_\varepsilon^n} 1 \cdot d\mu + \int_{X\setminus A_\varepsilon^n} \frac{\varepsilon}{1+\varepsilon}\, d\mu \\
&= \mu(A_\varepsilon^n) + \frac{\varepsilon}{1+\varepsilon}\mu(X \setminus A_\varepsilon^n) \\
&< \varepsilon + \varepsilon = 2\varepsilon
\end{aligned}$$

con lo cual queda demostrado que $\rho(f_n, f) \to 0$.

Recíprocamente, suponga que $\rho(f_n, f) \to 0$ y tomemos un $\varepsilon > 0$ arbitrario. Por la monotonía de u se tiene que si $x \in A^n_\varepsilon$, entonces $\varepsilon \leq |f_n(x) - f(x)|$ y, por lo tanto, $u(\varepsilon) \leq u(|f_n(x) - f(x)|)$, es decir,

$$\frac{\varepsilon}{1+\varepsilon} \leq \frac{|f_n(x) - f(x)|}{1 + |f_n(x) - f(x)|}.$$

Integrando ésta última expresión sobre A^n_ε se obtiene

$$\begin{aligned}
\frac{\varepsilon}{1+\varepsilon}\mu(A^n_\varepsilon) &= \int_{A^n_\varepsilon} \frac{\varepsilon}{1+\varepsilon}\, d\mu \\
&\leq \int_{A^n_\varepsilon} \frac{|f_n(x) - f(x)|}{1 + |f_n(x) - f(x)|}\, d\mu \\
&\leq \int_X \frac{|f_n(x) - f(x)|}{1 + |f_n(x) - f(x)|}\, d\mu \\
&= \rho(f_n, f),
\end{aligned}$$

de donde resulta que

$$\mu(A^n_\varepsilon) \leq \frac{1+\varepsilon}{\varepsilon}\rho(f_n, f) \to 0,$$

cuando $n \to \infty$ y termina la prueba. ■

Ejercicio 10.4.8. *Sea $f \in \mathrm{L}_1([a, b])$. Pruebe que*

(a) $\displaystyle\lim_{n\to\infty} \int_{E_n} |f|\, d\mu = 0$ *y*

(b) $\displaystyle\lim_{n\to\infty} n \cdot \mu(E_n) = 0$, *donde* $E_n = \{x \in [a, b] : |f| \geq n\}$ *para cada* $n \geq 1$.

Prueba. Para cada subconjunto medible $E \subseteq [a, b]$, defina

$$\nu_{|f|}(E) = \int_E |f|\, d\mu.$$

Como $E^\infty_f = \{x \in [a, b] : |f(x)| = +\infty\} = \bigcap_{n=1}^\infty E_n$ y $\nu_{|f|}$ es una medida, resulta que

$$\nu_{|f|}(E^\infty_f) = \nu_{|f|}\left(\bigcap_{n=1}^\infty E_n\right) = \lim_{n\to\infty} \nu_{|f|}(E_n) = \lim_{n\to\infty} \int_{E_n} |f|\, d\mu, \tag{1}$$

ya que $E_1 \supseteq E_2 \supseteq E_3 \supseteq \cdots$ y $\nu_{|f|}(E_1) < +\infty$. Por la Desigualdad de Chevyshev sabemos que

$$n \cdot \mu(E_n) \leq \int_{E_n} |f|\, d\mu \leq \int_{[a,b]} |f|\, d\mu = M \tag{2}$$

para todo $n \geq 1$, por lo que $\lim_{n\to\infty} \mu(E_n) = 0$ y, en consecuencia,

$$\mu(E^\infty_f) = \mu\left(\bigcap_{n=1}^\infty E_n\right) = \lim_{n\to\infty} \mu(E_n) = 0.$$

El Corolario 10.1.29 combinado con (1) conduce a la igualdad

$$0 \;=\; \int_{E_f^\infty} |f|\, d\mu \;=\; \lim_{n\to\infty} \int_{E_n} |f|\, d\mu$$

lo cual prueba (a). Aplicando el resultado anterior a la primera desigualdad de (2) se obtiene (b). ■

Ejercicio 10.4.9. *Sea $f : [a,b] \to \mathbb{R}$ una función **medible no-negativa** y para cada entero $n \geq 0$, sea $E_n = \{x \in [a,b] : n \leq f(x) < n+1\}$. Demuestre que las siguientes afirmaciones son equivalentes:*

(1) $f \in \mathrm{L}_1([a,b])$

(2) $\sum_{n=1}^{\infty} n \cdot \mu(E_n)$ *converge.*

(3) $\sum_{n=1}^{\infty} \mu(F_n)$ *converge, donde* $F_n = \{x \in [a,b] : f(x) \geq n\}$, $n = 0, 1, \ldots$

Prueba. (1) $\Rightarrow$ (2). Suponga que la serie $\sum_{n=1}^{\infty} n \cdot \mu(E_n)$ no converge y defina la sucesión de funciones $(f_n)_{n=1}^{\infty}$ por

$$f_n(x) \;=\; \begin{cases} f(x) & \text{si} \quad f(x) < n, \\ 0 & \text{si} \quad f(x) \geq n. \end{cases}$$

Puesto que $f_n \leq f$, resulta entonces que

$$\int_{[a,b]} f\, d\mu \;\geq\; \int_{[a,b]} f_n\, d\mu \;\geq\; \sum_{k=0}^{n-1} \int_{E_k} f\, d\mu \;\geq\; \sum_{k=0}^{n-1} k \cdot \mu(E_k).$$

Por otro lado, como $\lim_{n\to\infty} \sum_{k=0}^{n-1} k \cdot \mu(E_k) = +\infty$, vemos que $\int_{[a,b]} f\, d\mu = +\infty$. Esta contradicción establece que la serie $\sum_{n=1}^{\infty} n \cdot \mu(E_n)$ converge.

(2) $\Rightarrow$ (3). Observe, en primer lugar, que $F_k = E_k \cup E_{k+1} \cup \cdots$ y puesto que $E_i \cap E_j = \varnothing$ para todo $j \neq j$ se obtiene que $\mu(F_k) = \sum_{n=k}^{\infty} \mu(E_k)$. Ahora bien, como

$$\sum_{k=1}^{n} \mu(F_k) \;=\; \mu(F_1) + \mu(F_2) + \cdots + \mu(F_n) \;=\; \sum_{k=0}^{n} k \cdot \mu(E_k) + \sum_{k=n+1}^{\infty} n \cdot \mu(E_k). \tag{1}$$

y ya que por hipótesis la serie $\sum_{k=1}^{\infty} k \cdot \mu(E_k)$ converge, resulta que

$$\sum_{k=n+1}^{\infty} n \cdot \mu(E_k) \;\leq\; \sum_{k=n+1}^{\infty} k \cdot \mu(E_k) \;\xrightarrow[n\to\infty]{}\; 0.$$

Se sigue de (1) que la serie $\sum_{n=1}^{\infty} \mu(F_n)$ converge.

(3) $\Rightarrow$ (1). Asuma que la serie $\sum_{n=1}^{\infty} \mu(F_n)$ converge. De la desigualdad (1) vemos que

$$\sum_{k=0}^{n} k \cdot \mu(E_k) \;=\; \sum_{k=1}^{n} \mu(F_k) \;-\; \sum_{k=n+1}^{\infty} n \cdot \mu(E_k) \;\leq\; \sum_{n=1}^{\infty} \mu(F_n) \;<\; +\infty$$

de donde se deduce que $\sum_{k=1}^{\infty} k \cdot \mu(E_k)$ converge y, en consecuencia, también converge la serie $\sum_{k=1}^{\infty}(k+1)\cdot\mu(E_k)$. Más aun, si ponemos $g(x) = k+1$ para cada cada $x \in E_k$, tendremos que $0 \leq f(x) \leq g(x)$ para todo $x \in [a,b]$ y, por lo tanto,

$$\int_{[a,b]} f\,d\mu \;\leq\; \int_{[a,b]} g\,d\mu \;\leq\; \sum_{k=0}^{\infty}(k+1)\cdot\mu(E_k) \;<\; +\infty.$$

La prueba es completa. ■

Ejercicio 10.4.10. *Este ejercicio proporciona un criterio para que una función continua $f : [a,b] \to \mathbb{R}$ sea de variación acotada.*

Definición 10.4.1. *Sea $f : [a,b] \to \mathbb{R}$ una función continua y para cada $y \in \mathbb{R}$ sea*

$$R_y \;=\; f^{-1}(\{y\}) \;=\; \{x \in [a,b] \,:\, f(x) = y\}.$$

*La **indicatriz de Banach** es la función $i_B : \mathbb{R} \to [1, +\infty]$ definida por*

$$i_B(y) \;=\; \operatorname{card}(R_y) \qquad \text{para todo} \quad y \in \mathbb{R}.$$

Teorema 10.4.2 (Banach-Vitali). *Sea $f \in C([a,b])$ una **función continua**. Pruebe que:*

*(a) i_B es **medible** y*

$$\int_c^d i_B(y)\,dy \;=\; V(f,[a,b]),$$

donde $c = \inf\{f(x) : x \in [a,b]\}$ y $d = \sup\{f(x) : x \in [a,b]\}$.

(b) $f \in V(f,[a,b])$ si, y sólo si, $i_B \in L_1([a,b])$.

(c) Si $f \in V(f,[a,b])$, entonces $\mu\big(y \in \mathbb{R} : i_B(y) \geq \aleph_0\big) = 0$.

Prueba. Para cada $n \in \mathbb{N}$, considere la partición $P_n = \{a = x_0, x_1, \ldots, x_n = b\}$ de $[a,b]$, donde $x_j = a + j(b-a)/2^n$ para $j = 1,2,\ldots,n$. En este caso, la norma de P_n es: $\|P_n\| = (b-a)/2^n$. Para cada $j = 1,2,\ldots,n$ defina $S_j : \mathbb{R} \to [0,1]$ por

$$S_j(y) \;=\; \begin{cases} 1 & \text{si existe } x \in (x_{j-1}, x_j] \text{ tal que } f(x) = y \\ 0 & \text{si para todo } x \in (x_{j-1}, x_j] \text{ se cumple que } f(x) \neq y. \end{cases}$$

Observe que S_j es una función acotada y medible Lebesgue con a lo sumo dos puntos de discontinuidad. Además,

$$\int_c^d S_j(y)\,dy \;=\; \operatorname{osc}(f,J_j) \;=\; \sup\{f(x) : x \in J_j\} - \inf\{f(x) : x \in J_j\},$$

donde $J_j = (x_{j-1}, x_j]$. Defina ahora $B_n = \sum_{j=1}^{2^n} S_j$. Por el Teorema 9.1.26, página 470, tenemos que

$$\lim_{\|P_n\|\to 0} \sum_{j=1}^{2^n} \operatorname{osc}(f,J_j) \;=\; V(f,[a,b])$$

y, por consiguiente,

$$\lim_{n\to\infty}\int_c^d B_n(y)\,dy \;=\; V(f,[a,b]).$$

Puesto que $\big(B_n\big)_{n=1}^{\infty}$ es una sucesión no negativa y creciente de funciones medibles, el Teorema de la Convergencia Dominada de Lebesgue nos dice que

$$\int_c^d \lim_{n\to\infty} B_n(y)\,dy \;=\; \lim_{n\to\infty}\int_c^d B_n(y)\,dy.$$

Sea $D = \lim_{n\to\infty} B_n$. Claramente, $D \leq i_B$ ya que $B_n \leq i_B$ para todo $n \geq 1$. Para completar la prueba de la parte (a) sólo nos resta verificar que $i_B \leq D$.

Sea $m \in \mathbb{N}$ tal que $m \leq i_B(y)$ y escoja $z_1 < z_2 < \cdots < z_m$ tal que $f(z_k) = y$ para $k = 1, 2, \ldots, m$. Tome n suficientemente grande de modo que

$$\frac{b-a}{2^n} \;<\; z_k - z_{k-1} \quad \text{para } k = 1, 2, \ldots, m.$$

Por supuesto, esta elección de n impide que dos z_k's estén en el mismo intervalo $[x_{j-1},\, x_j]$ de la partición P_n. Entonces $B_n(y) \geq m$ y, en consecuencia, $D(y) \geq m$. Puesto que esto es cierto para cada $m \leq i_B(y)$, se sigue del Teorema 2.1.30, página 97, que $i_B(y) \leq D(y)$.

(b) Es consecuencia inmediata de la parte (a), mientras que (c) sigue de (b). ■

Ejercicio 10.4.11. *Sea $(f_n)_{n=1}^{\infty}$ una sucesión de funciones medibles definidas sobre $[0,1]$ tal que $|f_n| \leq g$, donde $g \in \mathrm{L}_1([0,1])$. Demuestre que $f_n \in \mathrm{L}_1([0,1],\mu)$ para todo $n \geq 1$ y*

$$\begin{aligned}
\int_0^1 \liminf_{n\to\infty} |f_n|\,d\mu &\leq \liminf_{n\to\infty}\int_0^1 |f_n|\,d\mu \\
&\leq \limsup_{n\to\infty}\int_0^1 |f_n|\,d\mu \\
&\leq \int_0^1 \limsup_{n\to\infty} |f_n|\,d\mu.
\end{aligned}$$

Prueba. Que $f_n \in \mathrm{L}_1([0,1],\mu)$ para todo $n \geq 1$, sigue de la hipótesis. Para demostrar el conjunto de desigualdades, observe que como las sucesiones $(f_n+g)_{n=1}^{\infty}$ y $(g-f_n)_{n=1}^{\infty}$ son no-negativas y medibles, el Lema de Fatou, combinado con la aditividad de la integral más las propiedades de los límites superiores e inferiores nos revelan que

$$\begin{aligned}
\int_0^1 g\,d\mu \;+\; \int_{[0,1]} \liminf_{n\to\infty} f_n\,d\mu &= \int_0^1 \liminf_{n\to\infty}(g+f_n)\,d\mu \\
&\leq \liminf_{n\to\infty}\int_0^1 (g+f_n)\,d\mu \\
&= \int_0^1 g\,d\mu + \liminf_{n\to\infty}\int_{[0,1]} f_n\,d\mu
\end{aligned}$$

y

$$\begin{aligned}\int_0^1 g\,d\mu \;-\; \int_0^1 \limsup_{n\to\infty} f_n\,d\mu \;&=\; \int_0^1 g\,d\mu \;+\; \int_0^1 \liminf_{n\to\infty}\,(-f_n)\,d\mu\\ &=\; \int_0^1 \liminf_{n\to\infty}(g-f_n)\,d\mu\\ &\le\; \liminf_{n\to\infty}\int_0^1 (g-f_n)\,d\mu\\ &=\; \int_0^1 g\,d\mu - \limsup_{n\to\infty}\int_0^1 f_n\,d\mu\end{aligned}$$

Cancelando la expresión $\int_0^1 g\,d\mu$ en ambas desigualdades se obtiene el resultado deseado. ■

Ejercicio 10.4.12. *Sea* $f \in L_1([0,+\infty),\mu)$ *y suponga que*

$$\int_0^x f\,d\mu \;=\; 0 \quad \text{para todo } x \ge 0.$$

Pruebe que $f = 0$ *casi-siempre.*

Prueba. Sean $0 \le a < b$. Entonces

$$\int_a^b f\,\mu \;=\; \int_0^b f\,d\mu \;-\; \int_0^a f\,d\mu \;=\; 0.$$

Sea V un conjunto abierto no vacío incluido en $[0,+\infty)$ y seleccione una sucesión disjunta $(I_n)_{n=1}^{\infty}$ de intervalos abiertos todos de longitud finita tal que $V = \bigcup_{n=1}^{\infty} I_n$. Se sigue de lo anterior que

$$\int_V f\,d\mu \;=\; \sum_{n=1}^{\infty}\int_{I_n} f\,d\mu \;=\; 0$$

y, por consiguiente,

$$\int_E f\,d\mu \;=\; 0 \quad \text{para todo medible } E \subseteq [0,+\infty).$$

Del Ejercicio 10.5 (2) se tiene que $f = 0$ casi-siempre en $[0,+\infty)$. ■

10.5. Problemas

(1) Para cada $n \ge 1$, sea $f_n : (0,1] \to \mathbb{R}$ una función Lebesgue integrable tal que $0 \le f_n(x) \le 1$ para todo $x \in (0,1]$. Demuestre que:

$$\text{si} \quad \lim_{n\to\infty}\int_{(0,1]} f_n\,d\mu \;=\; 0, \qquad \text{entonces} \qquad \lim_{n\to\infty} f_n \;=\; 0 \text{ casi-siempre.}$$

(2) Sea $f \in \mathcal{L}_1(\mu)$. Demuestre que si

$$\int_E f \, d\mu \;=\; 0 \quad \text{para todo} \quad E \in \mathfrak{M}_\mu(\mathbb{R}),$$

entonces $f = 0 \;\; \mu - \text{c.s.}$

(3) Sea $f \in \mathcal{L}_1([a,b],\mu)$ y defina $g : [a,b] \to \overline{\mathbb{R}}$ por

$$g(x) \;=\; \int_a^x f \, d\mu \quad \text{para todo} \quad x \in [a,b].$$

Pruebe que g es continua sobre $[a,b]$.

(4) Sea $(f_n)_{n=1}^{\infty}$ una sucesión de funciones medibles definidas sobre $[a,b]$ convergiendo puntualmente a una función f. Suponga que existe una función medible y acotada casi-siempre $g : [a,b] \to \overline{\mathbb{R}}$ tal que $|f_n| \leq g$ para todo $n \geq 1$. Pruebe que

$$\lim_{n\to\infty} \int_a^b f_n \cdot g \, d\mu \;=\; \int_a^b f \cdot g \, d\mu.$$

(5) ***Recíproco del* TCD**. Sea $(f_n)_{n=1}^{\infty}$ una sucesión **creciente** de funciones medibles definidas sobre X y suponga que existe una función $f \in \mathcal{L}_1(X,\mu)$ tal que $|f_n| \leq f$ para todo $n \geq 1$. Pruebe que si

$$\lim_{n\to\infty} \int_X f_n \, d\mu \;=\; \int_X f \, d\mu,$$

entonces $\lim\limits_{n\to\infty} f_n = f \quad \mu - c.s.$

(6) Sea E un conjunto medible con $\mu(E) > 0$. Demuestre que si $\int_A f \, d\mu = 0$ para cualquier conjunto medible $A \subseteq E$, entonces $f = 0 \;\; \mu - c.s.$

(7) Pruebe que si $f \in \mathcal{L}_1(X,\mu)$ y

$$\left| \int_X f \, d\mu \right| \;=\; \int_X |f| \, d\mu,$$

entonces o $f \geq 0$ casi-siempre, o $f \leq 0$ casi-siempre en X.

(8) De un ejemplo de una medida no finita ν tal que la condición $\nu \ll \mu$ no es equivalente a $\lim_{\mu\to 0} \nu = 0$.

(9) Sea $f : [a,b] \to \mathbb{R}$ una función continua y estrictamente creciente. Pruebe que $f \in \mathrm{AC}([a,b])$ si, y sólo si, $\mu(E_\infty) = 0$, donde

$$E_\infty \;=\; \{x \in [a,b] \,:\, f'(x) = +\infty\}.$$

(10) Sea $f : [a,b] \to \mathbb{R}$ una función continua y estrictamente creciente. Pruebe que su inversa $f^{-1} : [f(a), f(b)] \to [a,b]$ es absolutamente continua si, y sólo si, $\mu(N) = 0$, donde

$$N \;=\; \{x \in [a,b] \,:\, f'(x) = 0\}.$$

(11) Sea $f : [a,b] \to \mathbb{R}$ una función continua y diferenciable en $[a,b]$, excepto sobre un conjunto a lo más numerable. Pruebe que si $f' \in \mathrm{L}_1([a,b])$, entonces $f \in \mathrm{AC}([a,b])$.

(12) Sea $f : [a,b] \to \mathbb{R}$ una función monótona. Pruebe que $f \in \mathrm{AC}([a,b])$ si, y sólo si,

$$\int_a^b |f'|\, d\mu \;=\; |f(b) - f(a)|.$$

(13) Sea $f \in \mathrm{L}_1(\mu)$ y defina $\phi(x) = \sum_{n=1}^{\infty} f(2^n x + 1/n)$. Demuestre que $\phi \in \mathrm{L}_1(\mu)$ y que

$$\int_{\mathbb{R}} f\, d\mu \;=\; \int_{\mathbb{R}} \phi\, d\mu.$$

(14) Sea $f \in \mathrm{L}_1(\mu)$. Demuestre que para cada $0 \neq a \in \mathbb{R}$, la función $\varphi_a : \mathbb{R} \to \mathbb{R}$ definida por $\varphi_a(x) = f(ax)$ es Lebesgue integrable y se cumple, además, que

$$\int_{\mathbb{R}} f(ax)\, d\mu \;=\; \frac{1}{|a|} \int_{\mathbb{R}} f(x)\, d\mu.$$

(15) Construya una sucesión $(f_n)_{n=1}^{\infty}$ en $\mathrm{L}_1(\mu)$ tal que

$$\lim_{n\to\infty} \int_{\mathbb{R}} |f_n|\, d\mu \;=\; 0 \quad \text{pero} \quad \lim_{n\to\infty} f_n \;\neq\; 0 \;\; \mu - c.s.$$

(16) Sea $f \in \mathrm{L}_1([0,1])$. Demuestre que $x^n f \in \mathrm{L}_1([0,1])$ para todo $n \geq 1$ y

$$\lim_{n\to\infty} \int_{[0,1]} x^n f(x)\, d\mu \;=\; 0.$$

Si $\lim_{x\uparrow 1} f(x) = a$ para algún $a \in \mathbb{R}$, entonces

$$\lim_{n\to\infty} n \int_{[0,1]} x^n f(x)\, d\mu \;=\; a.$$

(17) Suponga que $E \subseteq [0, 2\pi]$ es medible y que

$$\int_E x^n \cos(x)\, d\mu \;=\; 0 \quad \text{para} \quad n = 0, 1, 2, \ldots$$

Pruebe que $\mu(E) = 0$.

(18) Sea $f : \mathbb{R} \to \mathbb{R}$ una función medible. Demuestre que las siguientes condiciones son equivalentes:

(a) $f \in \mathrm{L}_1(\mu)$.

(b) $\sum_{n=-\infty}^{\infty} 2^n \cdot \mu\big(\{x \in \mathbb{R} : |f(x)| > 2^n\}\big) < \infty.$

(19) Sea $(f_n)_{n=1}^{\infty}$ una sucesión **decreciente** de funciones medibles no-negativas tal que

(a) $\lim_{n\to\infty} f_n = f$ y

(b) $\int_E f_k \, d\mu < +\infty$ para algún $k \geq 1$.

Pruebe que

$$\int_E f \, d\mu = \lim_{n\to\infty} \int_E f_n \, d\mu.$$

(20) Sea $f : \mathbb{R} \to \overline{\mathbb{R}}$ una función medible no-negativa. Demuestre que

$$\int_{\mathbb{R}} f \, d\mu = \lim_{n\to\infty} \int_{[-n,n]} f \, d\mu = \lim_{n\to\infty} \int_{E_n} f \, d\mu,$$

donde $E_n = \{x \in \mathbb{R} : f(x) \geq 1/n\}$ para cada $n \geq 1$.

(21) Si $f : \mathbb{R} \to \overline{\mathbb{R}}$ es una función medible no-negativa y Lebesgue integrable, demuestre que

$$\int_{\mathbb{R}} f \, d\mu = \lim_{n\to\infty} \int_{E_n} f \, d\mu,$$

donde $E_n = \{x \in \mathbb{R} : f(x) \leq n\}$ para cada $n \geq 1$.

(22) Sea $(f_n)_{n=1}^{\infty}$ una sucesión de funciones medibles no-negativas y $f \geq 0$ tales que

(a) $\lim_{n\to\infty} f_n = f$ y

(b) $\lim_{n\to\infty} \int_{\mathbb{R}} f_n \, d\mu = \int_{\mathbb{R}} f \, d\mu < \infty$.

Pruebe que

$$\lim_{n\to\infty} \int_E f_n \, d\mu = \int_E f \, d\mu$$

para cualquier conjunto medible E. Exhiba un ejemplo para mostrar que la conclusión es falsa si $\lim_{n\to\infty} \int_{\mathbb{R}} f_n \, d\mu = \int_{\mathbb{R}} f \, d\mu = +\infty$

(23) ***Teorema de Severini-Egoroff Generalizado***. Sea $(f_n)_{n=1}^{\infty}$ una sucesión de funciones medibles con $|f_n| \leq g$ para todo $n \geq 1$, donde $g \in L_1(\mu)$. Demuestre que si $\lim_{n\to\infty} f_n = f$ $\mu - c.s.$, entonces $f_n \to f$ casi-uniformemente.

(24) Pruebe que $\mathcal{P}(\mathbb{Q})$, el conjuntos de todos los polinomios con coeficientes racionales, es denso en $L_1(\mu)$. Deduzca de esto que $L_1(\mu)$ es separable.

(25) Sea $C^{\infty}(\mathbb{R})$ el espacio de todas las funciones $f : \mathbb{R} \to \mathbb{R}$ que son infinitamente diferenciables. Pruebe que para cada $f \in L_1(\mu)$ y cualquier $\varepsilon > 0$, existe una función $\phi \in C^{\infty}(\mathbb{R})$ tal que $\phi = 0$ fuera de algún intervalo compacto y

$$\int_{\mathbb{R}} |f - \phi| \, d\mu < \varepsilon.$$

(26) Pruebe que $\mathcal{U} \subseteq \mathrm{L}_1([0,1])$ es uniformemente integrable si, y sólo si,

$$\sum_{m=1}^{\infty} m \cdot \mu\big([\, m \;\leq\; |f| \;<\; m+1\,]\big)$$

converge uniformemente para toda $f \in \mathcal{U}$.

(27) Pruebe que una familia $\mathcal{U} \subseteq \mathrm{L}_1(X,\mu)$ es **uniformemente absolutamente continua** si, y sólo si, para cada $\varepsilon > 0$, existe un $\delta > 0$ tal que

$$\sup_{f\in\mathcal{U}} \left| \int_E f \, d\mu \right| \;<\; \varepsilon$$

para cualquier conjunto medible E que satisfaga $\mu(E) < \delta$.

(28) Sea $\big(E_n\big)_{n=1}^{\infty}$ una sucesión de conjuntos medibles tal que $\lim_{n\to\infty} \mu(E_n) = 0$. Demuestre que

$$\lim_{n\to\infty} \int_{E_n} f \, d\mu \;=\; 0 \quad \text{para toda} \quad f \in L_1(\mu).$$

(29) Sea $f : [a,b] \to \mathbb{R}$ una función arbitraria. Pruebe que las siguientes condiciones son equivalentes:

(1) $f \in \mathrm{Lip}([a,b])$.

(2) f se puede escribir en la forma

$$f(x) \;=\; f(a) \;+\; \int_a^x g(t)\, dt,$$

para alguna función $g \in \mathrm{L}_\infty([a,b])$.

CAPÍTULO 11

Medida e Integración Abstracta

Una vez establecida que la familia $\mathfrak{M}_\mu(\mathbb{R})$ de los subconjuntos de $\mathbb{R}$ que son medibles según Lebesgue es una σ-álgebra y que μ, la restricción de μ^* a $\mathfrak{M}_\mu(\mathbb{R})$, es una medida, el siguiente paso es, racionalmente, necesario: generalizar dichas nociones a conjuntos abstractos e intentar obtener resultados importantes, similares a los anteriores, en este nuevo contexto. Como ocurre en casi toda generalización, algunas cosas se pierden, otras se mantienen y algunas nuevas surgen. En esta sección haremos un breve exposición de la noción abstracta de medida y de integral, es decir, introduciremos una forma general de definir diferentes medidas sobre un conjunto y sus integrales asociadas al estilo Lebesgue. De hecho, veremos que muchas de las propiedades que se cumplen para la medida e integral de Lebesgue en $\mathbb{R}$ también son válidas en este contexto abstracto. Sin embargo, la mayoría de las demostraciones serán omitidas debido a que las mismas son casi idénticas a las de sus contrapartes en $\mathbb{R}$.

11.1. Espacios de Medidas

Recordemos que una familia $\mathcal{M}$ de subconjuntos de un conjunto no vacío X se dice que es una **σ-álgebra** sobre X si:

(σ_1) $X \in \mathcal{M}$,

(σ_2) $X \setminus E \in \mathcal{M}$ siempre que $E \in \mathcal{M}$, y

(σ_3) $\bigcup_{n=1}^{\infty} E_n \in \mathcal{M}$ para cualquier colección numerable $\{E_n : n = 1, 2, \ldots\} \subseteq \mathcal{M}$.

$\|$ ► **Algunos Ejemplos de σ-álgebras**.

(a) Si $\mathcal{M}$ es una σ-álgebra sobre algún conjunto X y si E es un subconjunto no vacío de X, entonces

$$\mathcal{M}_E = \{A \cap E : A \in \mathcal{M}\}$$

es una σ-álgebra sobre E.

(b) Si $(\mathcal{M}_\alpha)_{\alpha\in I}$ es una familia de σ-álgebras sobre un conjunto X, entonces $\bigcap_{\alpha\in I}\mathcal{M}_\alpha$ también es una σ-álgebra. En particular, si $\mathcal{E}$ es cualquier colección de subconjuntos de X y si consideramos a la familia de todas las σ-álgebras que contienen a $\mathcal{E}$, esto es,

$$\mathcal{P} = \big\{\mathcal{M} \subseteq \mathscr{P}(X) : \mathcal{E} \subseteq \mathcal{M},\ \mathcal{M} \text{ es una } \sigma\text{-álgebra}\big\},$$

entonces la σ-álgebra $\sigma(\mathcal{E}) = \bigcap_{\mathcal{M}\in\mathcal{P}}\mathcal{M}$ será llamada la ***σ-álgebra generada*** por $\mathcal{E}$.

(c) Si X, Y son conjuntos no vacíos, $\mathcal{N}$ una σ-álgebra sobre Y y $f : X \to Y$ una función arbitraria, entonces

$$f^{-1}(\mathcal{N}) \;=\; \big\{f^{-1}(N) \subseteq X \,:\, N \in \mathcal{N}\big\}$$

es una σ-álgebra sobre X.

(d) Sea X un conjunto infinito. Entonces la colección

$$\mathcal{M} \;=\; \big\{A \subseteq X \,:\, A \text{ o } A^c \text{ es numerable}\big\}$$

es una σ-álgebra.

Fijemos un conjunto no vacío X y sea $\mathcal{M}$ una σ-álgebra de subconjuntos de X. Al par $(X, \mathcal{M})$ se le denomina un **espacio medible** y a los elementos de $\mathcal{M}$ se les llaman **conjuntos medibles** o $\mathcal{M}$-**medibles** si se desea enfatizar su procedencia.

Definición 11.1.1. *Sean X un conjunto no-vacío y $\mathcal{M}$ una σ-álgebra de subconjuntos de X. Una función de conjuntos $\lambda : \mathcal{M} \to [0, +\infty]$ se llama una **medida numerablemente aditiva**, o simplemente, una **medida**, sobre X si*

(a) $\lambda(\varnothing) = 0$ *y*

(b) $\lambda\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \sum_{n=1}^{\infty}\lambda(A_n)$, *para toda sucesión disjunta* $(A_n)_{n=1}^{\infty}$ *de* $\mathcal{M}$.

Si la condición (b) de la definición anterior es reemplazada por

$$\lambda\Big(\bigcup_{n=1}^{k} A_n\Big) \;=\; \sum_{n=1}^{k}\lambda(A_n)$$

para cualquier colección finita y disjunta de conjuntos $A_1, \ldots, A_n \in \mathcal{M}$, entonces diremos que λ es una **medida finitamente aditiva**. Por supuesto, toda medida numerablemente aditiva es automáticamente finitamente aditiva.

Toda tripleta de la forma $(X, \mathcal{M}, \lambda)$, donde X es un conjunto no-vacío, $\mathcal{M}$ es una σ-álgebra de X y λ es una medida sobre X, se le denomina un **espacio de medida**. Cada elemento $E \in \mathcal{M}$ se llama un conjunto **medible** y el valor $\lambda(E)$ la **medida** de E. Por supuesto, si se consideran varias σ-álgebras y varias medidas asociadas a un conjunto X, es necesario especificar a cuáles de ellas se refieren la palabras medible y medida. De darse esta situación emplearemos expresiones como "conjunto $\mathcal{M}$-medible" y "λ-medida", etc.

Si X es un espacio topológico y $\mathcal{O}_X$ es la colección de todos los subconjuntos abiertos de X, entonces la σ-álgebra de Borel es la σ-álgebra generada por $\mathcal{O}_X$ y denotada, como siempre, por $\mathfrak{Bo}(X)$. Cualquier medida $\lambda : \mathfrak{Bo}(X) \to [0, +\infty]$ es llamada una **medida de Borel**.

Con estas definiciones se puede demostrar que gran parte de las propiedades que posee la medida de Lebesgue en $\mathbb{R}$ se transfieren, casi sin cambio, a cualquier espacio de medida. Sin embargo, hay excepciones. Una de ellas es la noción de completitud que posee la medida de Lebesgue. Esta propiedad no es, en general, válida para cualquier otro espacio de medida. Por ejemplo, como se demostró en el Ejercicio 6.7.1 (e), página 354, la medida de Borel μ_0 no es completa. Por otro lado, en este contexto abstracto, si nuestro conjunto está desprovisto de topología alguna, entonces las propiedades topológicas de la medida de Lebesgue en $\mathbb{R}$ no se transfieren a la del espacio de medida en cuestión.

Definición 11.1.2. *Sea $(X, \mathcal{M}, \lambda)$ un espacio de medida. Si $\lambda(X) < +\infty$ diremos que λ es una **medida finita**, o también, que $(X, \mathcal{M}, \lambda)$ es un **espacio de medida finita**. En particular, si $\lambda(X) = 1$ se dice que λ es una **medida de probabilidad**, o que $(X, \mathcal{M}, \lambda)$ es un **espacio de probabilidad**.*

Por ejemplo, si $X = [0, 1]$, entonces $(X, \mathfrak{M}_\mu(X), \mu)$ es un espacio de probabilidad.

Definición 11.1.3. *Sea $(X, \mathcal{M}, \lambda)$ un espacio de medida. Si existe una sucesión $(A_n)_{n=1}^\infty$ en $\mathcal{M}$ tal que para todo $n \in \mathbb{N}$ se cumple que $\lambda(A_n) < \infty$ y $X = \bigcup_{n=1}^\infty A_n$, entonces diremos que $(X, \mathcal{M}, \lambda)$ es un espacio de medida σ-**finita** o, simplemente, que la medida λ es σ-finita.*

Es importante destacar que la sucesión $(A_n)_{n=1}^\infty$ en la definición de medida σ-finita se puede elegir *creciente*, o *disjunta*. Por ejemplo, $(\mathbb{R}, \mathfrak{M}_\mu(\mathbb{R}), \mu)$ constituye un espacio de medida σ-finita y a $\mathbb{R}$ lo podemos escribir como

$$\mathbb{R} = \bigcup_{n=1}^\infty [-n, n] = \bigcup_{n \in \mathbb{Z}} (n, n+1].$$

Ejemplo 11.1.1. (i) *Sea $\mathcal{M} = \mathscr{P}(\mathbb{N})$ y defina $\nu : \mathcal{M} \to [0, +\infty]$ por*

$$\nu(A) = \begin{cases} \mathrm{card}(A) & \text{si } A \text{ es finito} \\ +\infty & \text{si } A \text{ es infinito.} \end{cases}$$

*Entonces $(\mathbb{N}, \mathcal{M}, \nu)$ es un espacio de medida. A ν se le llama la **medida de conteo**.* Observe que $(\mathbb{N}, \mathcal{M}, \nu)$ es un espacio de medida σ-finita ya que, por ejemplo,

$$\mathbb{N} = \bigcup_{n=1}^\infty E_n,$$

donde $E_n = \{1, 2, \ldots, n\}$. Sin embargo, si en lugar de trabajar con $\mathbb{N}$ lo hacemos con el intervalo $[0, 1]$ resulta que $([0, 1], \mathscr{P}([0, 1]), \nu)$ es un espacio de medida que no es σ-finita. Que ello sea cierto es consecuencia inmediata del Teorema de Cantor sobre la no-numerabilidad del conjunto $[0, 1]$.

(ii) *Sea X un conjunto no-numerable y considere la σ-álgebra*

$$\mathcal{M} = \{A \subseteq X : A \text{ o } A^c \text{ es numerable}\}.$$

Definiendo

$$\nu(A) = \begin{cases} 0 & \text{si } A \text{ es numerable} \\ +\infty & \text{si } A^c \text{ es numerable,} \end{cases}$$

se prueba que $(X, \mathcal{M}, \lambda)$ *es un espacio de medida.*

(iii) *Sea* X *un conjunto no vacío y sea* $x \in X$. *Si definimos* $\delta_x : \mathscr{P}(X) \to [0, +\infty)$ *por*

$$\delta_x(A) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A, \end{cases}$$

entonces $(X, \mathscr{P}(X), \delta_x)$ *es un espacio de medida. A* δ_x *se le llama la* ***medida de Dirac*** *sobre* X.

(iv) *Sea* $(X, \mathcal{M}, \lambda)$ *un espacio de medida, y sea* Y *un conjunto arbitrario. Si* $f : X \to Y$ *es una función, entonces definiendo* $\lambda_f : \mathcal{M}_f \to [0, +\infty]$ *por*

$$\lambda_f(B) = \lambda(f^{-1}(B))$$

para cualquier $B \in \mathcal{M}_f = \{f(A) : A \in \mathcal{M}\}$, *resulta que* $(Y, \mathcal{M}_f, \lambda_f)$ *es un espacio de medida y a* λ_f *la llamaremos la* ***medida imagen***.

‖ ► **Algunas Propiedades de la Medida Abstracta**.

En el siguiente teorema veremos algunas de las propiedades que se cumplen en cualquier espacio de medida abstracta y que son idénticas a las de la medida de Lebesgue. Para demostrar dichos resultados, comencemos por fijar un espacio de medida $(X, \mathcal{M}, \lambda)$ y observe que si $(E_n)_{n=1}^{\infty}$ es una sucesión arbitraria de conjuntos medibles, entonces *existe una sucesión disjunta* $(F_n)_{n=1}^{\infty}$ en $\mathcal{M}$ tal que $F_n \subseteq E_n$ para todo $n \in \mathbb{N}$ y

$$\bigcup_{n=1}^{\infty} E_n = \bigcup_{n=1}^{\infty} F_n.$$

En efecto, pongamos $F_1 = E_1$ y para $n \geq 2$, defina

$$F_n = E_n \setminus \bigcup_{j=1}^{n-1} E_j.$$

Es claro que la sucesión $(F_n)_{n=1}^{\infty}$ cumple con los requerimientos exigidos.

En el siguiente resultado se establece que ningún conjunto medible de medida finita y estrictamente positivo puede ser representado por una unión no-numerable y disjunta de conjuntos medibles todos de medidas estrictamente positivas.

Lema 11.1.4. *Sea* $(X, \mathcal{M}, \lambda)$ *un* ***espacio de medida*** *y sea* $E \in \mathcal{M}$ *tal que* $0 < \lambda(E) < +\infty$. *Si* $\mathcal{F} = \{E_\alpha \subseteq E : \alpha \in \Lambda\}$ *es una colección* ***disjunta*** *de subconjuntos medibles de* E *tal que*

$$E = \bigcup_{\alpha \in \Lambda} E_\alpha \qquad y \qquad \lambda(E_\alpha) > 0 \quad \forall \alpha \in \Lambda,$$

entonces Λ *es numerable.*

Prueba. Como siempre, denote por $\mathcal{P}_{\text{fin}}(\Lambda)$ la colección de todos los subconjuntos finitos de Λ. Puesto que λ es finitamente aditiva se tiene que

$$\sum_{i\in F}\lambda(E_i) \;=\; \lambda\Big(\bigcup_{i\in F}E_i\Big) \;\le\; \lambda(E)$$

para cualquier conjunto $F\in\mathcal{P}_{\text{fin}}(\Lambda)$. Más aun, como $\lambda(E)<+\infty$, resulta que

$$\sup\Big\{\sum_{i\in F}\lambda(E_i) \,:\, F\in\mathcal{P}_{\text{fin}}(\Lambda)\Big\} \;\le\; \lambda(E) \;<\; +\infty$$

y, en consecuencia, la colección $(E_\alpha)_{\alpha\in\Lambda}$ es sumable. Se sigue del Corolario 2.1.56, página 116, que $\{\alpha\in\Lambda : \lambda(E_\alpha)>0\}=\Lambda$ es numerable. ■

Teorema 11.1.5. *Sea $(X,\mathcal{M},\lambda)$ un espacio de medida. Las siguientes propiedades se cumplen:*

*(a) (λ **es monótona**). Para todo $E,F\in\mathcal{M}$ con $E\subseteq F$ se tiene que $\lambda(E)\le\lambda(F)$. En particular, si $\lambda(E)<+\infty$, entonces*

$$\lambda(F\setminus E) \;=\; \lambda(F)\,-\,\lambda(E).$$

*(b) (λ **es subaditiva**). Si $(E_n)_{n=1}^{\infty}$ es cualquier sucesión en $\mathcal{M}$, entonces*

$$\lambda\Big(\bigcup_{n=1}^{\infty}E_n\Big) \;\le\; \sum_{n=1}^{\infty}\lambda(E_n).$$

*(c) (λ **es continua**). Si $(E_n)_{n=1}^{\infty}$ es cualquier **sucesión creciente** en $\mathcal{M}$, entonces*

$$\lambda\Big(\bigcup_{n=1}^{\infty}E_n\Big) \;=\; \lim_{n\to\infty}\lambda(E_n).$$

*(d) (λ **es continua**). Si $(E_n)_{n=1}^{\infty}$ es cualquier **sucesión decreciente** en $\mathcal{M}$ con $\lambda(E_1)<+\infty$, entonces*

$$\lambda\Big(\bigcap_{n=1}^{\infty}E_n\Big) \;=\; \lim_{n\to\infty}\lambda(E_n).$$

Prueba. (a) Puesto que $F=E\cup(F\setminus E)$ y λ es finitamente aditiva, se tiene que

$$\lambda(F) \;=\; \lambda(E)+\lambda(F\setminus E) \;\ge\; \lambda(E).$$

(b) Sea $(E_n)_{n=1}^{\infty}$ una sucesión en $\mathcal{M}$ y seleccione una sucesión disjunta $(F_n)_{n=1}^{\infty}$ en $\mathcal{M}$ tal que $F_n\subseteq E_n$ para todo $n\in\mathbb{N}$ y con $\bigcup_{n=1}^{\infty}E_n=\bigcup_{n=1}^{\infty}F_n$. De esto se sigue que

$$\lambda\Big(\bigcup_{n=1}^{\infty}E_n\Big) \;=\; \lambda\Big(\bigcup_{n=1}^{\infty}F_n\Big) \;=\; \sum_{n=1}^{\infty}\lambda(F_n) \;\le\; \sum_{n=1}^{\infty}\lambda(E_n).$$

Las pruebas de (c) y (d) son idénticas al caso de la medida de Lebesgue (Teorema 6.3.45 y Teorema 6.3.47) y, por lo tanto, se omiten. ■

Es claro que toda medida numerablemente aditiva es finitamente aditiva; sin embargo, el recíproco no es necesariamente cierto. El siguiente criterio provee una condición necesaria y suficiente para que una medida finitamente aditiva a valores reales sea numerablemente aditiva, es decir, una medida.

Teorema 11.1.6. *Sea $(X, \mathcal{M})$ un **espacio medible** y suponga que $\lambda : \mathcal{M} \to [0, +\infty)$ es una **medida finitamente aditiva**. Las siguientes condiciones son equivalentes:*

(1) *λ es **numerablemente aditiva**.*

(2) *λ es **numerablemente subaditiva**; es decir, para cualquier sucesión $(E_n)_{n=1}^{\infty}$ en $\mathcal{M}$,*

$$\lambda\Big(\bigcup_{n=1}^{\infty} E_n\Big) \leq \sum_{n=1}^{\infty} \lambda(E_n).$$

(3) *$\lim_{n\to\infty} \lambda(E_n) = 0$ para cualquier **sucesión decreciente** $(E_n)_{n=1}^{\infty} \subseteq \mathcal{M}$ con $\bigcap_{n=1}^{\infty} E_n = \varnothing$.*

Prueba. (1) $\Rightarrow$ (2) En primer lugar, observe que como λ es finitamente aditiva, ella es monótona, es decir, si $E, F \in \mathcal{M}$ con $E \subseteq F$, entonces $\lambda(E) \leq \lambda(F)$. En efecto, si $E \subseteq F$, entonces

$$F = E \cup (F \setminus E) \qquad \text{y} \qquad E \cap (F \setminus F) = \varnothing.$$

El hecho de que λ es finitamente aditiva nos garantiza que

$$\lambda(F) = \lambda(E) + \lambda(F \setminus E) \geq \lambda(E).$$

Sea ahora $(E_n)_{n=1}^{\infty}$ una sucesión arbitraria en $\mathcal{M}$. Entonces existe una sucesión disjunta $(F_n)_{n=1}^{\infty}$ en $\mathcal{M}$ tal que

$$F_n \subseteq E_n \quad \text{para todo } n \geq 1 \qquad \text{y} \qquad \bigcup_{n=1}^{\infty} E_n = \bigcup_{n=1}^{\infty} F_n.$$

Por esto,

$$\lambda\Big(\bigcup_{n=1}^{\infty} E_n\Big) = \lambda\Big(\bigcup_{n=1}^{\infty} F_n\Big) = \sum_{n=1}^{\infty} \lambda(F_n) \leq \sum_{n=1}^{\infty} \lambda(E_n).$$

(2) $\Rightarrow$ (1) Sea $(E_n)_{n=1}^{\infty}$ una sucesión disjunta en $\mathcal{M}$. Fijemos cualquier $k \in \mathbb{N}$. Entonces, por la monotonicidad de λ se tiene que

$$\lambda\Big(\bigcup_{n=1}^{\infty} E_n\Big) \geq \lambda\Big(\bigcup_{n=1}^{k} E_n\Big) = \sum_{n=1}^{k} \lambda(E_n)$$

de donde se obtiene que

$$\lambda\Big(\bigcup_{n=1}^{\infty} E_n\Big) \geq \sum_{n=1}^{\infty} \lambda(E_n).$$

La otra desigualdad sigue directamente de la hipótesis y con ella la numerabilidad aditiva de λ.

(1) $\Rightarrow$ (3) Suponga que λ es numerablemente aditiva y sea $(E_n)_{n=1}^{\infty}$ una sucesión decreciente de conjuntos medibles con $\bigcap_{n=1}^{\infty} E_n = \varnothing$. Puesto que E_1 se puede escribir en la forma $E_1 = \bigcup_{n=1}^{\infty}(E_n \setminus E_{n+1})$, la numerabilidad aditiva de λ combinada con el hecho de que $\lambda(E_n)$ es finito

para todo $n \geq 1$, nos muestra que

$$\begin{aligned}
\lambda(E_1) &= \sum_{n=1}^{\infty} \lambda(E_n \setminus E_{n+1}) \;=\; \lim_{k\to\infty} \sum_{n=1}^{k} \lambda(E_n \setminus E_{n+1}) \\
&= \lim_{k\to\infty} \big[\lambda(E_1 \setminus E_2) + \cdots + \lambda(E_k \setminus E_{k+1})\big] \\
&= \lim_{k\to\infty} \big[\lambda(E_1) - \cancel{\lambda(E_2)} + \cancel{\lambda(E_2)} - \cancel{\lambda(E_3)} + \cdots + \cancel{\lambda(E_k)} - \lambda(E_{k+1})\big] \\
&= \lambda(E_1) - \lim_{k\to\infty} \lambda(E_{k+1}).
\end{aligned}$$

Más aun, como $\lambda(E_1)$ es finito, se tiene entonces que $\lim_{n\to\infty} \lambda(E_n) = 0$ y termina la prueba de esta implicación.

(3) $\Rightarrow$ (1) Sea $(E_n)_{n=1}^{\infty}$ una sucesión disjunta de conjuntos medibles y sea $E = \bigcup_{n=1}^{\infty} E_n$. Para cada $n \geq 1$, defina

$$F_n \;=\; E \setminus (E_1 \cup \cdots \cup E_n) \;=\; \bigcup_{k=n+1}^{\infty} E_k.$$

Resulta que $(F_n)_{n=1}^{\infty}$ es una sucesión decreciente de conjuntos medibles con $\bigcap_{n=1}^{\infty} F_n = \varnothing$ y entonces, por hipótesis, $\lim_{n\to\infty} \lambda(F_n) = 0$. Ahora bien, como λ es finitamente aditiva,

$$\lambda(E) \;=\; \sum_{k=1}^{n} \lambda(E_k) + \lambda(F_n), \qquad n = 1, 2, \ldots$$

y, en consecuencia, tomando límite cuando $n \to \infty$ se obtiene que

$$\lambda(E) \;=\; \lim_{n\to\infty} \sum_{k=1}^{n} \lambda(E_k) + \lim_{n\to\infty} \lambda(F_n) \;=\; \sum_{k=1}^{\infty} \lambda(E_k)$$

y termina la prueba. ■

Suponga, de nuevo, que $(X, \mathcal{M}, \lambda)$ es un **espacio de medida finitamente aditiva**.

(1) *Si para cualquier* ***sucesión creciente*** $(E_n)_{n=1}^{\infty}$ *en* $\mathcal{M}$ *se cumple que*

$$\lambda\Big(\bigcup_{n=1}^{\infty} E_n\Big) \;=\; \lim_{n\to\infty} \lambda(E_n)$$

entonces λ *es* ***numerablemente aditiva***.

Prueba. En efecto, sea $(E_n)_{n=1}^{\infty}$ una sucesión disjunta en $\mathcal{M}$. Tomando $F_n = \bigcup_{j=1}^{n} E_j$ para cada $n \in \mathbb{N}$, resulta que $(F_n)_{n=1}^{\infty}$ es una sucesión creciente en $\mathcal{M}$, $\bigcup_{n=1}^{\infty} E_n = \bigcup_{n=1}^{\infty} F_n$ y entonces, por hipótesis, se tiene que

$$\lambda\Big(\bigcup_{n=1}^{\infty} E_n\Big) \;=\; \lim_{n\to\infty} \lambda(F_n) \;=\; \lim_{n\to\infty} \sum_{j=1}^{n} \lambda(E_j) \;=\; \sum_{j=1}^{\infty} \lambda(E_j).$$

(2) *Si para cualquier* ***sucesión decreciente*** $(E_n)_{n=1}^{\infty}$ *en* $\mathcal{M}$ *se cumple que* $\lambda(E_n) < +\infty$ *para algún* $n \in \mathbb{N}$ *y*

$$\lambda\Big(\bigcap_{n=1}^{\infty} E_n\Big) \;=\; \lim_{n\to\infty} \lambda(E_n)$$

entonces λ *es* ***numerablemente aditiva***.

Prueba. La prueba es similar a la anterior. En efecto, si $(E_n)_{n=1}^{\infty}$ es una sucesión disjunta en $\mathcal{M}$, sólo hay que definir $F_n = \bigcap_{k=1}^{\infty} E_k \setminus (\bigcup_{j=1}^{n} E_j)$ y verificar que la sucesión $(F_n)_{n=1}^{\infty}$ es decreciente, $\bigcap_{n=1}^{\infty} F_n = \varnothing$ y usar el hecho de que $\lambda(E_n) < +\infty$ para algún $n \geq 1$ para obtener que

$$\lambda\Big(\bigcup_{n=1}^{\infty} E_n\Big) \;=\; \sum_{j=1}^{\infty} \lambda(E_j).$$

11.1.1. Medidas sin Átomos

La medida de Lebesgue en $\mathbb{R}$ posee una importantísima propiedad: *carece de átomos*. ¿Qué significa esto?

Definición 11.1.7. *Sea* $(X, \mathcal{M}, \lambda)$ *un espacio de medida. Un conjunto* $A \in \mathcal{M}$ *es un* ***átomo*** *para* λ *si se cumplen las siguientes condiciones:*

(a) $\lambda(A) > 0$ *y*

(b) *para cualquier* $B \in \mathcal{M}$ *con* $B \subseteq A$, *ocurre sólo una de las siguientes alternativas:*

$$\lambda(B) \;=\; 0 \qquad\qquad o \qquad\qquad \lambda(B) \;=\; \lambda(A).$$

Si λ *posee al menos un átomo, diremos que* λ *es una* ***medida atómica***. *En caso contrario se dirá que* λ *es* ***no-atómica*** *o* ***difusa***.

Algunas observaciones con respecto a la noción de átomos son las siguientes:

(1) Si $(X, \mathcal{M}, \lambda)$ es un espacio de **medida finita**, entonces la condición (b) en la definición anterior puede ser reemplazada por esta otra:

(b') *para cualquier* $B \in \mathcal{M}$ *con* $B \subseteq A$, *ocurre que* $\lambda(B) = 0$, *o bien,* $\lambda(A \setminus B) = 0$.

(2) *Si* A_1 *y* B_1 *son átomos, entonces* $\lambda(A_1 \cap A_2) = 0$ o $\lambda(A_1 \triangle A_2) = 0$.

(3) *Si* $(X, \mathcal{M}, \lambda)$ *es un espacio de* ***medida finita*** *y si* $\mathcal{A}_\lambda$ *denota* ***cualquier colección disjunta de átomos*** *de* λ, *entonces* $\mathcal{A}_\lambda$ *es* ***a lo más numerable***. En efecto, como λ es finita y finitamente aditiva, resulta que la familia $(\lambda(A))_{A \in \mathcal{A}_\lambda}$ es sumable y, en consecuencia, por el Corolario 2.1.56, página 116, $\mathcal{A}_\lambda$ es a lo más numerable. Esto último también es válido si λ es σ-finita.

(4) *Si* λ *no posee átomos y* $E \in \mathcal{M}$ *es tal que* $0 < \lambda(E) < +\infty$, *entonces existen conjuntos medibles y disjuntos, digamos* A *y* B, *tales que:* $\lambda(A) > 0$, $\lambda(B) > 0$ *y* $C = A \cup B$. Para ver esto, observe que como λ no es un atómica, entonces existe algún conjunto medible $A \subsetneq E$ con $0 < \lambda(A) < \lambda(E)$. Defina entonces $B = E \setminus A$ y nótese que $\lambda(B) > 0$. Observe que uno de ellos, digamos B, satisface $\lambda(B) \in (0, \frac{1}{2}\lambda(E)]$ y el otro, $\lambda(A) \in [\frac{1}{2}\lambda(E), \lambda(E))$.

Teorema 11.1.8 (Saks). *Sea $(X, \mathcal{M}, \lambda)$ un espacio de **medida no-atómica**. Suponga que $E \in \mathcal{M}$ satisface $0 < \lambda(E) < +\infty$. Entonces, para cada $\varepsilon > 0$, existe $A_\varepsilon \in \mathcal{M}$ con $A \subseteq E$ tal que*

$$0 \ < \ \lambda(A) \ < \ \varepsilon.$$

Prueba. Puesto que λ no es atómica, la observación (4) del parágrafo anterior nos permite descomponer a E como una unión disjunta de dos conjuntos medibles, digamos A_1 y B_1, cada uno de medida estrictamente positiva. Sin perder generalidad, asumiremos que $\lambda(A_1) \geq \lambda(B_1)$. Como $\lambda(E) = \lambda(A_1) + \lambda(B_1) \geq 2\lambda(B_1)$, se tiene que $0 < \lambda(B_1) \leq \lambda(E)/2$. De nuevo, como $\lambda(A_1) > 0$ y λ es no-atómica, podemos representar a A_1 como la unión disjunta de dos conjuntos medibles ambos de medidas estrictamente positivas, digamos A_2 y B_2 y tal que $\lambda(B_2) \leq \lambda(B_1)/2$. Continuando inductivamente con este procedimiento se obtiene una sucesión de conjuntos medibles $(B_n)_{n=1}^\infty$ tal que, para todo $n \geq 1$, se cumple que:

$$\lambda(B_n) \ \leq \ \lambda(B_{n-1})/2 \ \leq \ \cdots \ \leq \ \frac{\lambda(E)}{2^n}.$$

Puesto que $\lambda(E) < +\infty$, entonces podemos seleccionar un $N \in \mathbb{N}$ lo suficientemente grande de modo que $\lambda(E)/2^N < \varepsilon$. La elección $A_\varepsilon = B_N$ termina la prueba. ■

Observe que en la demostración del resultado anterior la sucesión $(B_n)_{n=1}^\infty$ es disjunta, mientras que la sucesión $(A_n)_{n=1}^\infty$ es decreciente y

$$\lambda(E) \ > \ \lambda(A_1) \ > \ \lambda(A_2) \ > \ \lambda(A_3) \ > \ \cdots \ > \ 0.$$

El siguiente teorema muestra, efectivamente, que existe un continuum de tales subconjuntos medibles incluidos en E. Recordemos, véase el Teorema 6.4.8, página 308, que si E es un conjunto medible con $\mu(E) > 0$ y si $0 < q < \mu(E)$, entonces existe un conjunto medible $F \subseteq E$ tal que $\mu(F) = q$. La razón de por qué ello es así se debe a que μ es no-atómica.

Teorema 11.1.9 (Fichtenholz-Sierpiński). *Sea $(X, \mathcal{M}, \lambda)$ un espacio de **medida no-atómica** y suponga que $E \in \mathcal{M}$ es tal que $0 < \lambda(E) < +\infty$. Si $0 < q < \lambda(E)$, entonces existe $F \in \mathcal{M}$ con $F \subseteq E$ tal que $\lambda(F) = q$.*

Prueba. Nuestra primer objetivo es demostrar la siguiente:

Afirmación: *Existe un conjunto medible $S \subseteq E$ tal que*

$$\frac{q}{2} \ < \ \lambda(S) \ < \ q. \tag{1}$$

Suponga que tal afirmación es falsa; es decir, para cualquier conjunto medible $S \subseteq E$ se verifica que

$$\lambda(S) \ \leq \ \frac{q}{2} \qquad \text{o} \qquad \lambda(S) \ \geq \ q. \tag{2}$$

Considere ahora la siguiente familia de subconjuntos medibles de E:

$$\mathcal{S} \ = \ \big\{S \ : \ S \ \text{ es medible}, \ \ S \subseteq E \ \ \text{y} \ \ 0 < \lambda(S) \leq q\big\}.$$

Observe que, gracias al Teorema 11.1.8, nuestro conjunto $\mathcal{S}$ es no-vacío y como estamos asumiendo la validez de (2) para todo conjunto medible incluido en E, resulta entonces que

$$\alpha \ = \ \sup\big\{\lambda(S) \ : \ S \in \mathcal{S}\big\} \ \leq \ \frac{q}{2}.$$

Nótese, además, que si $S_1, \ldots, S_n \in \mathcal{S}$, entonces $\bigcup_{j=1}^n S_j \in \mathcal{S}$. En efecto, un razonamiento inductivo nos muestra que es suficiente verificar que $S_1 \cup S_2 \in \mathcal{S}$, lo cual es cierto ya que

$$\lambda(S_1 \cup S_2) \leq \lambda(S_1) + \lambda(S_2) \leq \alpha + \alpha \leq q.$$

Por otro lado, como $\alpha = \sup_{S \in \mathcal{M}} \lambda(S)$, existe una sucesión $(S_n)_{n=1}^\infty$ en $\mathcal{S}$ tal que $\alpha - 1/n < \lambda(S_n)$ para todo $n \in \mathbb{N}$ y, por consiguiente,

$$\alpha - 1/n < \lambda(S_n) \leq \lambda\Big(\bigcup_{j=1}^n S_j\Big).$$

Teniendo en cuenta que $\bigcup_{j=1}^n S_j \in \mathcal{S}$, resulta entonces que

$$\alpha - 1/n < \lambda\Big(\bigcup_{j=1}^n S_j\Big) \leq \alpha$$

y como la sucesión $(\bigcup_{j=1}^n S_j \in \mathcal{S})_{n=1}^\infty$ es creciente, la continuidad de λ nos revela que

$$\lambda\Big(\bigcup_{j=1}^\infty S_j\Big) = \alpha \leq \frac{q}{2}.$$

Puesto que $E \setminus \bigcup_{j=1}^\infty S_j$ es un conjunto medible de medida positiva:

$$\lambda\Big(E \setminus \bigcup_{j=1}^\infty S_j\Big) = \lambda(E) - \lambda\Big(\bigcup_{j=1}^\infty S_j\Big) > q - \frac{q}{2} = \frac{q}{2} > 0,$$

el Teorema 11.1.8 nos garantiza la existencia de un conjunto medible $F \subseteq E \setminus \bigcup_{j=1}^\infty S_j$ tal que $0 < \lambda(F) < q/2$. Si ahora definimos $R = F \cup \bigcup_{j=1}^\infty S_j \in \mathcal{S}$, resulta que

$$\lambda(R) = \lambda(F) + \lambda\Big(\bigcup_{j=1}^\infty S_j\Big) > \lambda\Big(\bigcup_{j=1}^\infty S_j\Big) = \alpha$$

lo cual contradice la definición de α. Esta contradicción establece que nuestra afirmación es verdadera.

Sea $q_1 = q$. Aplique nuestra **Afirmación** para obtener un conjunto medible $S_1 \subseteq E := E_1$ tal que $q_1/2 < \lambda(S_1) < q_1$. Sean

$$E_2 = E_1 \setminus S_1 \qquad \text{y} \qquad q_2 = q_1 - \lambda(S_1).$$

Como $q_2 > 0$, podemos, usando de nuevo la **Afirmación**, obtener un conjunto medible $S_2 \subseteq E_2$ de modo que $q_2/2 < \lambda(S_2) < q_2$. Hagamos

$$E_3 = E_2 \setminus S_2 \qquad \text{y} \qquad q_3 = q_2 - \lambda(S_2).$$

Continuando indefinidamente con este procedimiento se obtiene una sucesión $(S_n)_{n=1}^\infty$ de conjuntos medibles disjuntos dos a dos y una sucesión (q_n) de números positivos tales que

$$\frac{q_n}{2} < \lambda(S_n) < q_n \qquad \text{y} \qquad q_n = q_{n-1} - \lambda(S_{n-1}).$$

Puesto que $q_n = q_{n-1} - \lambda(S_{n-1}) = \cdots = q_{n-2} - \lambda(S_{n-2}) - \lambda(S_{n-1}) = \cdots = q_1 - \sum_{j=1}^{n-1} \lambda(S_j)$, entonces la desigualdad anterior se puede escribir en la forma

$$\frac{1}{2} \cdot \left(q - \sum_{j=1}^{n-1} \lambda(S_j) \right) = \frac{1}{2} \cdot q_n < \lambda(S_n) \leq q_n = q - \sum_{j=1}^{n-1} \lambda(S_j)$$

para todo $n \in \mathbb{N}$ y, así,

$$q - \lambda(S_n) < \lambda(S_1) + \lambda(S_2) + \cdots + \lambda(S_n) < q.$$

Se deduce de esto último que la sucesión no-negativa $(\sum_{j=1}^{n} \lambda(S_n))_{n=1}^{\infty}$ es creciente y acotada y, en consecuencia, converge. En particular, $\lim_{n\to\infty} \lambda(S_n) = 0$ y, además se cumple que:

$$\lambda\left(\bigcup_{j=1}^{\infty} S_j \right) = \sum_{j=1}^{\infty} \lambda(S_j) = q.$$

Tomando $F = \bigcup_{j=1}^{\infty} S_j$ se finaliza la prueba. ∎

Nota Adicional 11.1.1 Si $(X, \mathcal{M}, \lambda)$ es un espacio de medida arbitrario, definimos el **rango de** λ como

$$\operatorname{Rang}(\lambda) = \{\lambda(E) : E \in \mathcal{M}\}.$$

El Teorema de Fichtenholz-Sierpiński establece que si $(X, \mathcal{M}, \lambda)$ es **finita** y **no-atómica**, entonces

$$\operatorname{Rang}(\lambda) = [0, \lambda(X)].$$

En general, dicho resultado es válido para *cualquier medida no-atómica tomando valores en un espacio de Banach de dimensión finita*. En este caso, $\operatorname{Rang}(\lambda)$ es un *conjunto convexo* y *compacto*. A este resultado se le conoce como el **Teorema de Convexidad de Liapounoff** y, en general, es falso para medidas no-atómicas con valores en espacios de Banach de dimensión infinita (véase, por ejemplo, [48], pág. 261-267).

11.1.2. Completación de una Medida

Los conjuntos de medida cero son muy importantes en la Teoría de de la Medida y aunque pueden ser, en cierto modo, ignorados en la Teoría de Integración, ellos también constituyen una pieza fundamental en dicha teoría. En el caso particular de la medida de Lebesgue, ésta posee una propiedad que es envidiable: *cualquier subconjunto de un conjunto de* ***medida cero*** *es* ***medible***. A diferencia de lo que ocurre con la medida de Lebesgue, existen espacios de medidas donde no todos los subconjuntos de un conjunto medible de medida nula son medibles. Por ejemplo, si tomamos $X = \{a, b, c\}$, $\mathcal{M} = \{\varnothing, \{a\}, \{b, c\}, X\}$ y la medida $\lambda : \mathcal{M} \to [0, +\infty]$ se define por $\lambda(\{a\}) = 1$ y $\lambda(\{b, c\}) = 0$, entonces $N = \{b, c\}$ es un conjunto medible de medida cero, mientras que los subconjuntos $\{b\}$ y $\{c\}$ de N no son medibles. Sin embargo, siempre existe un modo de extender la familia $\mathcal{M}$ para obtener una nueva familia $\mathcal{M}'$ y una nueva medida λ' definida sobre $\mathcal{M}'$, en donde los subconjuntos de los conjuntos de λ'-medida cero permanecen en $\mathcal{M}'$.

Definición 11.1.10. *Sea $(X, \mathcal{M}, \lambda)$ un espacio de medida. Se dice que λ es una* ***medida completa*** *si para todo $N \in \mathcal{M}$ con $\lambda(N) = 0$, la condición $E \subseteq N$ implica que E es medible, es decir, $E \in \mathcal{M}$.*

En otras palabras, una medida λ *es completa si* ***todos los subconjuntos*** *de cualquier conjunto medible de medida cero son medibles*; brevemente, si $N \in \mathcal{M}$, entonces

$$\lambda(N) = 0 \quad \Rightarrow \quad \mathscr{P}(N) \subseteq \mathcal{M}.$$

En lo que sigue veremos que todo espacio de medida $(X, \mathcal{M}, \lambda)$ que no es completo se puede completar, esto es:

Teorema 11.1.11 (Completación de una Medida). *Sea $(X, \mathcal{M}, \lambda)$ un* ***espacio de medida no completo****. Existe un* ***espacio de medida completo****, digamos $(X, \widetilde{\mathcal{M}}, \widetilde{\lambda})$, tal que:*

$$\mathcal{M} \subseteq \widetilde{\mathcal{M}} \qquad y \qquad \widetilde{\lambda}(E) = \lambda(E)$$

para todo $E \in \mathcal{M}$.

Prueba. (1) **Construcción de** $\widetilde{\mathcal{M}}$. Defina la colección $\widetilde{\mathcal{M}}$ del modo siguiente:

$$\widetilde{\mathcal{M}} = \big\{E \subseteq X : \text{ existen } A, B \in \mathcal{M} \text{ con } A \subseteq E \subseteq B \text{ tal que } \lambda(B \setminus A) = 0\big\}.$$

Vamos a demostrar que $\widetilde{\mathcal{M}}$ es una σ-álgebra conteniendo a $\mathcal{M}$. En efecto:

(a) Claramente $X \in \widetilde{\mathcal{M}}$ ya que $X \in \mathcal{M}$ y $\lambda(X \setminus X) = 0$.

(b) Sea $E \in \widetilde{\mathcal{M}}$. Entonces existen $A, B \in \mathcal{M}$ con $A \subseteq E \subseteq B$ tal que $\lambda(B \setminus A) = 0$. Puesto que $\mathcal{M}$ es una σ-álgebra, los conjuntos A^c y B^c están en $\mathcal{M}$ y, además, se verifica que $B^c \subseteq E^c \subseteq A^c$ y $\lambda(A^c \setminus B^c) = 0$ ya que $A^c \setminus B^c = B \setminus A$. De esto resulta que $E^c \in \widetilde{\mathcal{M}}$.

(c) Finalmente, suponga que $(E_n)_{n=1}^{\infty}$ es una sucesión en $\widetilde{\mathcal{M}}$ y entonces seleccione sucesiones $(A_n)_{n=1}^{\infty}$ y $(B_n)_{n=1}^{\infty}$ en $\mathcal{M}$ tales que $A_n \subseteq E_n \subseteq B_n$ y $\lambda(B_n \setminus A_n) = 0$ para todo $n \geq 1$. Claramente

$$\bigcup_{n=1}^{\infty} A_n \subseteq \bigcup_{n=1}^{\infty} E_n \subseteq \bigcup_{n=1}^{\infty} B_n$$

y como $\mathcal{M}$ es una σ-álgebra, los conjuntos $\bigcup_{n=1}^{\infty} A_n$ y $\bigcup_{n=1}^{\infty} B_n$ están en $\mathcal{M}$ y se cumple que

$$\bigcup_{n=1}^{\infty} B_n \setminus \bigcup_{n=1}^{\infty} A_n \subseteq \bigcup_{n=1}^{\infty} (B_n \setminus A_n).$$

La subaditividad de λ garantiza entonces que $\bigcup_{n=1}^{\infty} E_n \in \widetilde{\mathcal{M}}$. Esto prueba que $\widetilde{\mathcal{M}}$ es una σ-álgebra.

(c) Por supuesto, $\mathcal{M} \subseteq \widetilde{\mathcal{M}}$.

(2) **Construcción de** $\widetilde{\lambda}$. Sea $E \in \widetilde{\mathcal{M}}$ y seleccione, usando la definición, conjuntos $A, B \in \mathcal{M}$ con $A \subseteq E \subseteq B$ tal que $\lambda(B \setminus A) = 0$. Defina ahora

$$\widetilde{\lambda}(E) = \lambda(A).$$

Veamos que $\widetilde{\lambda}$ está bien definida. En efecto, sean A_1, B_1 otro par de conjuntos en $\mathcal{M}$ con $A_1 \subseteq E \subseteq B_1$ satisfaciendo $\lambda(B_1 \setminus A_1) = 0$. Entonces

$$A \setminus A_1 \;\subseteq\; E \setminus A_1 \;\subseteq\; B_1 \setminus A_1,$$

de donde se obtiene que $\lambda(A \setminus A_1) = 0$. Como $A = (A \cap A_1) \cup (A \setminus A_1)$, resulta que $\lambda(A) = \lambda(A \cap A_1)$ y por simetría, $\lambda(A_1) = \lambda(A \cap A_1)$. De esto se concluye que $\lambda(A) = \lambda(A_1)$, quedando demostrado que $\widetilde{\lambda}$ está bien definida. Observe, además, que

$$\widetilde{\lambda}(E) \;=\; \lambda(E) \tag{2}$$

para cualquier $E \in \mathcal{M}$.

Nos ocuparemos ahora de demostrar que $\widetilde{\lambda}$ es una medida. En efecto, como $\varnothing \in \mathcal{M}$, resulta que

$$\widetilde{\lambda}(\varnothing) \;=\; \lambda(\varnothing) \;=\; 0.$$

Suponga que $(E_n)_{n=1}^{\infty}$ es una sucesión disjunta de conjuntos en $\widetilde{\mathcal{M}}$ y como antes, seleccione sucesiones $(A_n)_{n=1}^{\infty}$ y $(B_n)_{n=1}^{\infty}$ en $\mathcal{M}$ tales que $A_n \subseteq E_n \subseteq B_n$ y $\lambda(B_n \setminus A_n) = 0$ para todo $n \geq 1$. Claramente la sucesión $(A_n)_{n=1}^{\infty}$ es disjunta y $\bigcup_{n=1}^{\infty} A_n \subseteq \bigcup_{n=1}^{\infty} E_n \subseteq \bigcup_{n=1}^{\infty} B_n$. Se sigue de la definición de $\widetilde{\lambda}$ que

$$\begin{aligned} \widetilde{\lambda}\Big(\bigcup_{n=1}^{\infty} E_n\Big) \;&=\; \lambda\Big(\bigcup_{n=1}^{\infty} A_n\Big) \\ &=\; \sum_{n=1}^{\infty} \lambda(A_n) \\ &=\; \sum_{n=1}^{\infty} \widetilde{\lambda}(E_n). \end{aligned}$$

Esto nos revela que $\widetilde{\lambda}$ es una medida que es una extensión de λ.

(3) $\widetilde{\lambda}$ **es completa**. En efecto, sea $N \in \widetilde{\mathcal{M}}$ con $\widetilde{\lambda}(N) = 0$ y sea $E \subseteq N$. Escoja $A, B \in \mathcal{M}$ con $A \subseteq N \subseteq B$ tal que $\lambda(B \setminus A) = 0$. Puesto $\varnothing \subseteq E \subseteq A$ y $\lambda(A) = \widetilde{\lambda}(N) = 0$ se tiene que $E \in \widetilde{\mathcal{M}}$ y termina la prueba. ■

Definición 11.1.12. *El espacio de medida $(X, \widetilde{\mathcal{M}}, \widetilde{\lambda})$ del teorema anterior se le llama la **completación** de $(X, \mathcal{M}, \lambda)$ y a $\widetilde{\lambda}$ la **extensión completa** de λ.*

Es un ejercicio sencillo verificar que $\widetilde{\mathcal{M}}$ también se puede escribir en la forma

$$\widetilde{\mathcal{M}} \;=\; \big\{E \cup N \,:\, E \in \mathcal{M},\, N \in \mathcal{N}\big\},$$

donde $\mathcal{N} = \{N \in \mathscr{P}(X) : N \subseteq M,\ M \in \mathcal{M},\ \lambda(M) = 0\}$. Esta descripción nos dice que para obtener una extensión completa de un espacio de medida no completa, todo lo que debemos hacer es *añadir a $\mathcal{M}$ todos los subconjuntos de conjuntos medibles de λ-medida cero*.

En relación con el teorema anterior conviene observar que, gracias al Teorema 6.3.55, página 293, cualquier conjunto medible E según Lebesgue se puede escribir en la forma $E = B \cup N$, donde $B \in \mathfrak{Bo}(\mathbb{R})$ y N es un conjunto μ-nulo. Esto nos muestra que

Corolario 11.1.13. *El espacio de medida de Lebesgue* $(\mathbb{R}, \mathfrak{M}_\mu(\mathbb{R}), \mu)$ *es la* ***completación*** *del espacio de* ***medida no completo*** $(\mathbb{R}, \mathfrak{Bo}(\mathbb{R}), \mu_0)$.

Definición 11.1.14. *Sea* $(X, \mathcal{M}, \lambda)$ *un espacio de medida. Una propiedad* $\mathbf{P}$ *se dice que se cumple* ***casi-siempre respecto a*** λ, *en notación,* λ-c.s., *si*

$$\lambda\big(\{x \in X : \mathbf{P}(x) \text{ es falso}\}\big) = 0.$$

En este caso también se dice que $\mathbf{P}(x)$ se cumple para λ-casi-todo $x \in X$ o, para casi-todo $x \in X$, si no existe ninguna otra medida en el contexto.

Nota Adicional 11.1.2 Si bien es cierto que el proceso de completar un espacio de medida tiene sus ventajas, también es cierto que ello nos conduce a otros problemas. Por ejemplo: la σ-álgebra completa $\widetilde{\mathcal{M}}$ es, a menudo, mucho más complicada que la original. Por otro lado, al ser $\widetilde{\lambda}$ una extensión de λ ella no es, en general, única, lo cual nos indica que dos extensiones distintas pueden dar origen, en términos generales, a dos σ-álgebras completas diferentes. Estos hechos justifican el por qué en el estudio de las propiedades básicas de las medidas abstractas se evita que los argumentos dependan de la completación de la medida.

11.1.3. El Teorema de Extensión de Carathéodory

En esta sección presentamos el Teorema de Extensión de Carathéodory el cual es considerado, por muchos expertos, como el más importante resultado en la Teoría de la Medida. En efecto, usando dicho resultado se puede construir casi todas las medidas importantes incluyendo, por supuesto, las medidas de Lebesgue y de Lebesgue-Stieltjes sobre $\mathbb{R}$. Algunos ingredientes son necesarios para obtener dicho resultado. Comencemos.

Definición 11.1.15. *Sea* X *un conjunto arbitrario no vacío. Una* ***medida exterior*** *sobre* X *es una función de conjuntos*

$$\lambda^* : \mathscr{P}(X) \to [0, +\infty]$$

verificando las siguientes propiedades:

(a) $\lambda^*(\varnothing) = 0$,

(b) $\lambda^*(A) \leq \lambda^*(B)$ *si* $A \subseteq B$, *y*

(c) $\lambda^*\Big(\bigcup_{n=1}^{\infty} A_n\Big) \leq \sum_{n=1}^{\infty} \lambda^*(A_n)$ *para cualquier sucesión* $(A_n)_{n=1}^{\infty}$ *en* $\mathscr{P}(X)$.

Un modo estándar de obtener una medida exterior sobre un conjunto X consiste en encontrar, primeramente, una familia adecuada $\mathcal{G}$ de "subconjuntos elementales" sobre los cuales se define una cierta noción de medida (por ejemplo: su longitud, su masa, su volumen, etc) y entonces aproximar conjuntos arbitrarios "por fuera" por uniones numerables de miembros de $\mathcal{G}$. Aquí está la idea exacta de lo que queremos hacer.

Teorema 11.1.16. *Sean X un conjunto no vacío, $\mathcal{G} \subseteq \mathscr{P}(X)$ tal que $\varnothing, X \in \mathcal{G}$ y $\lambda : \mathcal{G} \to [0, +\infty]$ una función de conjuntos verificando $\lambda(\varnothing) = 0$. Entonces la aplicación $\lambda^* : \mathscr{P}(X) \to [0, +\infty]$ definida por*

$$\lambda^*(A) \;=\; \text{ínf}\left\{ \sum_{n=1}^{\infty} \lambda(E_n) \;:\; A \subseteq \bigcup_{n=1}^{\infty} E_n,\; (E_n)_{n=1}^{\infty} \subseteq \mathcal{G} \right\}$$

*para todo $A \subseteq X$, es una **medida exterior**.*

Prueba. Claramente $\lambda^*(\varnothing) = 0$ y si $A \subseteq B \subseteq X$, entonces $\lambda(A) \le \lambda(B)$. Para verificar que λ^* es numerablemente subaditiva, tome cualquier sucesión $(A_n)_{n=1}^{\infty}$ en $\mathscr{P}(X)$ y fijemos un $\varepsilon > 0$. Entonces, para cada $n \in \mathbb{N}$, existe una sucesión $(E_{n,k})_{k=1}^{\infty} \subseteq \mathcal{G}$ tal que

$$A_n \;\subseteq\; \bigcup_{k=1}^{\infty} E_{n,k} \qquad \text{y} \qquad \sum_{k=1}^{\infty} \lambda(E_{n,k}) \;<\; \lambda^*(A_n) \;+\; \frac{\varepsilon}{2^n}.$$

Finalmente, si hacemos $A = \bigcup_{n=1}^{\infty} A_n$, resulta que

$$A \;\subseteq\; \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{\infty} E_{n,k} \qquad \text{y} \qquad \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \lambda(E_{n,k}) \;<\; \sum_{k=1}^{\infty} \lambda^*(A_n) \;+\; \varepsilon,$$

de donde se sigue que $\lambda^*(A) \le \sum_{k=1}^{\infty} \lambda^*(A_n) + \varepsilon$. Puesto que $\varepsilon > 0$ es arbitrario, se concluye que $\lambda^*(A) \le \sum_{k=1}^{\infty} \lambda^*(A_n)$ y termina la prueba. ■

Siguiendo a Carathéodory podemos introducir la siguiente definición.

Definición 11.1.17. *Sea $\lambda^* : \mathscr{P}(X) \to [0, +\infty]$ una medida exterior. Diremos que un conjunto $E \subseteq X$ es λ^*-**medible** si*

$$\lambda^*(A) \;=\; \lambda^*(A \cap E) \;+\; \lambda^*(A \cap E^c),$$

para cualquier conjunto $A \subseteq X$.

Como $A = (A \cap E) \cup (A \cap E^c)$ es una unión disjunta, la subaditividad de λ^* siempre garantiza que $\lambda^*(A) \le \lambda^*(A \cap E) + \lambda^*(A \cap E^c)$ de modo que, para verificar que un conjunto E es λ^*-medible sólo es necesario demostrar que

$$\lambda^*(A) \;\ge\; \lambda^*(A \cap E) \;+\; \lambda^*(A \cap E^c) \tag{α}$$

para todo $A \subseteq \mathbb{R}$. Más aun, puesto que esta última desigualdad se cumple trivialmente si $\lambda^*(A) = +\infty$, entonces se tiene que:

Corolario 11.1.18. *Un conjunto $E \subseteq X$ es λ^*-**medible** si, y sólo si,*

$$\lambda^*(A) \;\ge\; \lambda^*(A \cap E) \;+\; \lambda^*(A \cap E^c)$$

para todo $A \subseteq \mathbb{R}$ con $\lambda^*(A) < +\infty$.

Fijemos ahora una medida exterior $\lambda^* : \mathscr{P}(X) \to [0, +\infty]$ y considere la colección

$$\mathcal{M}_{\lambda^*}(X) \;=\; \big\{ E \subseteq X \;:\; E \;\; \text{es } \lambda^*\text{-medible} \big\}.$$

Puesto que la noción de λ^*-medibilidad es simétrica respecto al complemento, resulta entonces que

$$E \in \mathcal{M}_{\lambda^*}(X) \quad \Leftrightarrow \quad E^c \in \mathcal{M}_{\lambda^*}(X).$$

Claramente, $\varnothing \in \mathcal{M}_{\lambda^*}(X)$ y, por lo tanto, $X \in \mathcal{M}_{\lambda^*}(X)$. Nuestro siguiente paso es demostrar que λ^* es **finitamente aditiva** sobre $\mathcal{M}_{\lambda^*}(X)$, es decir, si $E_1, \ldots, E_n$ son conjuntos disjuntos y λ^*-medibles, entonces $\bigcup_{j=1}^n E_j$ es λ^*-medible y

$$\lambda^*\left(\bigcup_{j=1}^n E_j\right) = \sum_{j=1}^n \lambda^*(E_j)$$

Para verificar dicha igualdad, es suficiente demostrar que

$$\lambda^*\left(A \cap \left[\bigcup_{j=1}^n E_j\right]\right) = \sum_{j=1}^n \lambda^*(A \cap E_j)$$

para cualquier conjunto $A \subseteq X$. La prueba es por inducción. Claramente el resultado se cumple para $n = 1$. Suponga ahora que la igualdad anterior es válida para cualquier colección finita y disjunta con $n-1$ conjuntos λ^*-medibles y sea $\{E_1, \ldots, E_n\}$ una colección arbitraria, finita y disjunta de conjuntos λ^*-medibles. Puesto que

$$\left(A \cap \left[\bigcup_{j=1}^n E_j\right]\right) \cap E_n = A \cap E_n \qquad \text{y} \qquad \left(A \cap \left[\bigcup_{j=1}^n E_j\right]\right) \cap E_n^c = A \cap \left[\bigcup_{j=1}^{n-1} E_j\right]$$

resulta de la λ^*-medibilidad de E_n y la hipótesis inductiva que

$$\begin{aligned}
\lambda^*\left(A \cap \left[\bigcup_{j=1}^n E_j\right]\right) &= \lambda^*\left(\left(A \cap \left[\bigcup_{j=1}^n E_j\right]\right) \cap E_n\right) + \lambda^*\left(\left(A \cap \left[\bigcup_{j=1}^n E_j\right]\right) \cap E_n^c\right) \\
&= \lambda^*(A \cap E_n) + \lambda^*\left(A \cap \left[\bigcup_{j=1}^{n-1} E_j\right]\right) \\
&= \lambda^*(A \cap E_n) + \sum_{j=1}^{n-1} \lambda^*(A \cap E_j) = \sum_{j=1}^{n} \lambda^*(A \cap E_j).
\end{aligned}$$

En particular, tomando $A = X$ se tiene que λ^* es finitamente aditiva. ∎

Teorema 11.1.19 (Carathéodory). *La colección $\mathcal{M}_{\lambda^*}(X)$ es una σ-**álgebra** y la función de conjuntos $\lambda : \mathcal{M}_{\lambda^*}(X) \to [0, +\infty]$ dada por*

$$\lambda(E) = \lambda^*(E) \quad \textit{para todo } E \in \mathcal{M}_{\lambda^*}(X)$$

*es una **medida** que, además, es **completa**.*

Prueba. Para demostrar que $\mathcal{M}_{\lambda^*}(X)$ es una σ-álgebra, sólo nos resta demostrar que si $\left(E_n\right)_{n=1}^{\infty}$ es una sucesión en $\mathcal{M}_{\lambda^*}(X)$, entonces $\bigcup_{n=1}^{\infty} E_n$ también está en $\mathcal{M}_{\lambda^*}(X)$. Si se considera la sucesión $\left(F_n\right)_{n=1}^{\infty}$ donde

$$F_1 \;=\; E_1 \qquad\qquad \text{y} \qquad\qquad F_n \;=\; E_n \setminus \bigcup_{j=1}^{n-1} E_j, \quad n \geq 2$$

resulta que ella disjunta, λ^*-medible y se cumple que $\bigcup_{n=1}^{\infty} E_n = \bigcup_{n=1}^{\infty} F_n$. Se sigue de esto que es suficiente suponer, y así convenimos, que nuestra sucesión $\left(E_n\right)_{n=1}^{\infty}$ es disjunta. Aceptado esto último, defina, para cada $n \geq 1$,, el conjunto $G_n \;=\; \bigcup_{j=1}^{n} E_j$. Por lo visto anteriormente, $G_n \in \mathcal{M}_{\lambda^*}(X)$ y como

$$G_n \;=\; \bigcup_{j=1}^{n} E_j \;\subseteq\; \bigcup_{j=1}^{\infty} E_j$$

resulta que $(\bigcup_{j=1}^{\infty} E_j)^c \subseteq G_n^c$ y, por lo tanto,

$$\begin{aligned}\lambda^*(A) \;&=\; \lambda^*(A \cap G_n) \,+\, \lambda^*(A \cap G_n^c)\\ &\geq\; \sum_{j=1}^{n} \lambda^*(A \cap E_j) \,+\, \lambda^*\left(A \cap \left[\bigcup_{n=1}^{\infty} E_n\right]^c\right)\end{aligned}$$

Tomando limite cuando $n \to \infty$ en la desigualdad anterior y usando el hecho de λ^* es numerablemente subaditiva, se obtiene que

$$\begin{aligned}\lambda^*(A) \;&\geq\; \sum_{j=1}^{\infty} \lambda^*(A \cap E_j) \,+\, \lambda^*\left(A \cap \left[\bigcup_{n=1}^{\infty} E_n\right]^c\right)\\ &\geq\; \lambda^*\left(A \cap \left[\bigcup_{n=1}^{\infty} E_n\right]\right) \,+\, \lambda^*\left(A \cap \left[\bigcup_{n=1}^{\infty} E_n\right]^c\right)\\ &\geq\; \lambda^*(A).\end{aligned}$$

Esto prueba que $\bigcup_{n=1}^{\infty} E_n \in \mathcal{M}_{\lambda^*}(X)$ y, por lo tanto, $\mathcal{M}_{\lambda^*}(X)$ es una σ-álgebra. En particular,

$$\lambda^*(A) \;=\; \sum_{j=1}^{\infty} \lambda^*(A \cap E_j) \,+\, \lambda^*\left(A \cap \left[\bigcup_{n=1}^{\infty} E_n\right]^c\right) \tag{β}$$

para todo conjunto $A \subseteq X$.

Para verificar que λ^* es una medida sobre $\mathcal{M}_{\lambda^*}(X)$, sea $\left(E_n\right)_{n=1}^{\infty}$ una sucesión disjunta en $\mathcal{M}_{\lambda^*}(X)$ y sustituya A por $\bigcup_{n=1}^{\infty} E_n$ en (β) para obtener

$$\lambda^*\left(\bigcup_{j=1}^{\infty} E_j\right) \;=\; \sum_{j=1}^{\infty} \lambda^*(E_j).$$

Resta demostrar que λ es una medida completa, sea $N \subseteq X$ con $\lambda^*(N) = 0$. Para cualquier conjunto $A \subseteq X$ tenemos que $A \cap N \subseteq N$ y, por lo tanto, $\lambda^*(A \cap N) = 0$ ya que $0 \leq \lambda^*(A \cap N) \leq \lambda^*(N) = 0$. Finalmente, como $N \supseteq A \cap N^c$, entonces

$$\lambda^*(A) \;\geq\; \lambda^*(A \cap N^c) \;=\; \lambda^*(A \cap N) + \lambda^*(A \cap N^c).$$

Usando (α) se concluye que $N \in \mathcal{M}_{\lambda^*}(X)$ y termina la prueba. ■

Observe que medida λ obtenida en el teorema anterior no es otra cosa que λ^* restringida a $\mathcal{M}_{\lambda^*}(X)$.

Sea $\mathcal{A}$ una álgebra sobre conjunto no vacío X. En general, si $(A_n)_{n=1}^{\infty}$ es una sucesión en $\mathcal{A}$, entonces no hay garantía de que su unión $\bigcup_{n=1}^{\infty} A_n$ pertenezca a $\mathcal{A}$. Por esta razón, definimos:

Definición 11.1.20. *Sea $\mathcal{A}$ una álgebra de subconjuntos sobre conjunto no vacío X. Una función de conjuntos $\nu : \mathcal{A} \to [0, +\infty]$ se dice que es una **pre-medida** sobre $\mathcal{A}$ si*

(a) $\lambda(\varnothing) = 0$ y

*(b) λ es **numerablemente aditiva sobre** $\mathcal{A}$, es decir, si $(A_n)_{n=1}^{\infty}$ es una sucesión disjunta en $\mathcal{A}$ tal que $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$, entonces*

$$\nu\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \sum_{n=1}^{\infty} \nu(A_n).$$

Observe que toda pre-medida $\nu : \mathcal{A} \to [0, +\infty]$ es **numerablemente subaditiva**. En efecto, sea $(A_n)_{n=1}^{\infty}$ una sucesión en $\mathcal{A}$ tal que $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$ y considere la sucesión $(B_n)_{n=1}^{\infty}$ definida por

$$B_1 = A_1 \qquad \text{y} \qquad B_n = A_n \setminus \bigcup_{j=1}^{n-1} A_j, \quad n \geq 2.$$

Resulta entonces que $(B_n)_{n=1}^{\infty}$ es una sucesión disjunta en $\mathcal{A}$ tal que $B_n \subseteq A_n$ para todo $n \in \mathbb{N}$ y $\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$. Se sigue de esto que $\bigcup_{n=1}^{\infty} B_n \in \mathcal{A}$ y

$$\nu\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \nu\Big(\bigcup_{n=1}^{\infty} B_n\Big) = \sum_{n=1}^{\infty} \nu(B_n) \leq \sum_{n=1}^{\infty} \nu(A_n).$$

Tenemos ahora todas las herramientas en las manos para demostrar el Teorema de Extensión de Carathéodory el cual se expresa en los siguientes términos.

Teorema 11.1.21 (Extensión de Carathéodory). *Sea $\mathcal{A}$ una **álgebra** sobre X y sea $\lambda : \mathcal{A} \to [0, +\infty]$ una **pre-medida** sobre $\mathcal{A}$. Entonces existe una **medida** $\nu : \sigma(\mathcal{A}) \to [0, +\infty]$ tal que*

$$\nu(A) = \lambda(A) \quad \textit{para todo } A \in \mathcal{A}.$$

En particular, si $\lambda(X) < +\infty$, entonces $\nu(X) < +\infty$.

Prueba. Por el Teorema 11.1.16 la aplicación $\lambda^* : \mathscr{P}(X) \to [0, +\infty]$ definida por

$$\lambda^*(A) = \inf\Big\{ \sum_{n=1}^{\infty} \lambda(E_n) : A \subseteq \bigcup_{n=1}^{\infty} E_n,\ E_n \in \mathcal{A} \Big\}$$

para todo $A \subseteq X$, es una medida exterior. Por el Teorema 11.1.19 la colección

$$\mathcal{M}_{\lambda^*}(X) = \big\{ E \subseteq X : E \text{ es } \lambda^*\text{-medible} \big\}.$$

es una σ-álgebra y la restricción de λ^* sobre $\mathcal{M}_{\lambda^*}(X)$ es una medida. Sea $\widehat{\nu} = \lambda^*|_{\mathcal{M}_{\lambda^*}(X)}$. Observe que ahora tenemos tres funciones de conjuntos: la original λ, la medida exterior λ^* y $\widehat{\nu}$, la restricción de λ^* sobre $\mathcal{M}_{\lambda^*}(X)$.

Afirmación (1). $\mathcal{A} \subseteq \mathcal{M}_{\lambda^*}(X)$. Para ver esto, sea $E \in \mathcal{A}$. En vista del Corolario 11.1.18 será suficiente demostrar que

$$\lambda^*(A) \;\geq\; \lambda^*(A \cap E) \;+\; \lambda^*(A \cap E^c) \quad \text{para todo } A \subseteq \mathbb{R} \text{ con } \lambda^*(A) < +\infty.$$

Sea $\varepsilon > 0$ y escojamos un conjunto $A \subseteq \mathbb{R}$ con $\lambda^*(A) < +\infty$. Por la definición de λ^*, existe una sucesión de conjuntos $(E_n)_{n=1}^{\infty}$ en $\mathcal{A}$ tal que

$$A \;\subseteq\; \bigcup_{n=1}^{\infty} E_n \qquad \text{y} \qquad \sum_{n=1}^{\infty} \lambda^*(E_n) \;<\; \lambda^*(A) + \varepsilon.$$

Para cada $n \in \mathbb{N}$, sea $B_n = E_n \cap E$ y observe que $B_n \in \mathcal{A}$, $B_n \subseteq E_n$ y

$$A \cap E \;\subseteq\; \Big(\bigcup_{n=1}^{\infty} E_n\Big) \cap E \;=\; \bigcup_{n=1}^{\infty} E \cap E_n \;=\; \bigcup_{n=1}^{\infty} B_n,$$

$$A \cap E^c \;\subseteq\; \Big(\bigcup_{n=1}^{\infty} E_n\Big) \cap E^c \;=\; \bigcup_{n=1}^{\infty} E_n \cap E^c \;=\; \bigcup_{n=1}^{\infty} (E_n \setminus E).$$

Por otro lado, como $E_n = B_n \cup (E_n \setminus E)$ para todo $n \in \mathbb{N}$, resulta que

$$(A \cap E) \cup (A \cap E^c) \;\subseteq\; \bigcup_{n=1}^{\infty} B_n \;\cup\; \bigcup_{n=1}^{\infty} (E_n \setminus E) \;=\; \bigcup_{n=1}^{\infty} B_n \cup (E_n \setminus E) \;=\; \bigcup_{n=1}^{\infty} E_n$$

y, por lo tanto,

$$\lambda^*(A \cap E) \;+\; \lambda^*(A \cap E^c) \;\leq\; \lambda^*\Big(\bigcup_{n=1}^{\infty} E_n\Big) \;\leq\; \sum_{n=1}^{\infty} \lambda^*(E_n) \;<\; \lambda^*(A) + \varepsilon.$$

Puesto que $\varepsilon > 0$ es arbitrario se tiene que $\lambda^*(A \cap E) + \lambda^*(A \cap E^c) \leq \lambda^*(A)$ y, en consecuencia, $E \in \mathcal{M}_{\lambda^*}(X)$. Esto prueba nuestra afirmación (1).

Afirmación (2). $\sigma(\mathcal{A}) \subseteq \mathcal{M}_{\lambda^*}(X)$. En efecto, como $\mathcal{M}_{\lambda^*}(X)$ es una σ-álgebra conteniendo a $\mathcal{A}$ y $\sigma(\mathcal{A})$ es la σ-álgebra más pequeña conteniendo a $\mathcal{A}$, nuestra afirmación (2) queda probada.

Afirmación (3). $\lambda(A) = \lambda^*(A)$ para todo $A \in \mathcal{A}$. Sea $A \in \mathcal{A}$. Por definición, $\lambda^*(A) \leq \lambda(A)$ de modo que sólo es necesario verificar la otra desigualdad. Observe que si $\lambda^*(A) = +\infty$, entonces se cumple que $\lambda(A) = +\infty$. Suponga ahora que $\lambda^*(A) < +\infty$ y sea $(E_n)_{n=1}^{\infty}$ cualquier sucesión en $\mathcal{A}$ tal que $A \subseteq \bigcup_{n=1}^{\infty} E_n$. La numerabilidad subaditiva de λ nos muestra que

$$\lambda(A) \;=\; \lambda\Big(\bigcup_{n=1}^{\infty} (A \cap E_n)\Big) \;\leq\; \sum_{n=1}^{\infty} \lambda(A \cap E_n) \;\leq\; \sum_{n=1}^{\infty} \lambda(E_n).$$

Tomando ínfimo sobre todas las colecciones $(E_n)_{n=1}^{\infty}$ incluidas en $\mathcal{A}$ y que cubren a A se deduce que $\lambda(A) \leq \lambda^*(A)$ y así, $\lambda(A) = \lambda^*(A)$ para todo $A \in \mathcal{A}$. Esto termina la prueba de la afirmación (3).

Finalmente, como $\widehat{\nu}$ es una medida, resulta que $\nu = \widehat{\nu}|_{\sigma(\mathcal{A})}$ también es una medida que coincide con λ sobre $\mathcal{A}$. En particular, si $\lambda(X) < +\infty$, entonces $\nu(X) = \lambda(X) < +\infty$ y termina la prueba. ■

¿Es la extensión obtenida en el Teorema de Extensión de Carathéodory única? La respuesta es afirmativa siempre que la pre-medida λ sea σ-finita. Para demostrar esta afirmación debemos introducir una nueva definición.

Definición 11.1.22. *Sea $\mathcal{D}$ una familia de subconjuntos de X. Diremos que $\mathcal{D}$ es un* ***sistema de Dynkin****, o un* D-***sistema****, sobre X si:*

(a) $X \in \mathcal{D}$,

(b) $A \setminus B \in \mathcal{D}$ para todo $A, B \in \mathcal{D}$ con $B \subseteq A$, y

(c) $\displaystyle\bigcup_{n=1}^{\infty} A_n \in \mathcal{D}$ para cualquier sucesión disjunta $(A_n)_{n=1}^{\infty} \subseteq \mathcal{D}$.

Nótese que la condición (b) de la definición anterior es equivalente a esta otra:

(b') $A^c \in \mathcal{D}$ siempre que $A \in \mathcal{D}$.

En efecto, claramente (b) implica (b') tomando $A = X$. Recíprocamente, suponga (b') se cumple y sean $A, B \in \mathcal{D}$ con $B \subseteq A$. Como $X \setminus A$ y B son disjuntos y $(A \setminus B)^c = (X \setminus A) \cup B$, resulta de (c) que $(A \setminus B)^c \in \mathcal{D}$ y, en consecuencia, $A \setminus B = X \setminus (A \setminus B)^c \in \mathcal{D}$.

Por supuesto, cualquier σ-álgebra es un D-sistema, pero el recíproco no es, en general, verdadero. Nótese que la diferencia entre una σ-álgebra y un D-sistema, es que la colección numerable en la primera no tiene por qué ser disjunta. Sin embargo, se tiene el siguiente resultado.

Teorema 11.1.23. *Sea $\mathcal{D}$ un D-**sistema** sobre X. Las siguientes condiciones son equivalentes:*

*(1) $\mathcal{D}$ es una **σ-álgebra**.*

*(2) $\mathcal{D}$ es **estable bajo intersección**, es decir, $A \cap B \in \mathcal{D}$ cualesquiera sean $A, B \in \mathcal{D}$.*

Prueba. $(1) \Rightarrow (2)$ es obvia. Suponga ahora que $\mathcal{D}$ es un D-sistema sobre X tal que $A \cap B \in \mathcal{D}$ siempre que $A, B \in \mathcal{D}$. En primer lugar, observe que por definición, si $A \in \mathcal{D}$, entonces $A^c \in \mathcal{D}$. Por otro lado, de las igualdades

$$A \cup B = A \cup (B \setminus (A \cap B)) \qquad \text{y} \qquad A \cap (B \setminus (A \cap B)) = \varnothing$$

se sigue que $A \cup B \in \mathcal{D}$. Por inducción, para cualquier colección finita $\{A_1, \ldots, A_n\}$ incluida en $\mathcal{D}$, se cumple que $\bigcup_{j=1}^{n} A_j \in \mathcal{D}$. Considere ahora cualquier sucesión $(A_n)_{n=1}^{\infty}$ en $\mathcal{D}$. Si definimos la sucesión $(B_n)_{n=1}^{\infty}$ por

$$B_1 = A_1 \qquad \text{y} \qquad B_n = A_n \setminus (A_1 \cup A_2 \cup \cdots \cup A_{n-1}), \quad n \geq 2$$

resulta que ella es disjunta en $\mathcal{D}$ y $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n \in \mathcal{D}$. La prueba es completa. ■

Nota Adicional 11.1.3 En muchos textos sobre Teoría de la Medida se usan con más frecuencia las expresiones: λ-**sistema** y π-**sistema** en lugar de "D-sistema" y "estable bajo intersección", respectivamente.

Similar al caso de la σ-álgebra generada por una familia de conjuntos, si $\mathcal{C}$ es cualquier colección de subconjuntos de X y si $\mathfrak{D}$ es la familia de todos los D-sistemas sobre X cada uno de los cuales contiene a $\mathcal{C}$, entonces

$$\delta(\mathcal{C}) \;=\; \bigcap_{\mathcal{D}\in\mathfrak{D}} \mathcal{D}$$

es el D-sistema sobre X más pequeño conteniendo a $\mathcal{C}$. Claramente, $\delta(\mathcal{C}) \subseteq \sigma(\mathcal{C})$. El siguiente resultado establece que si $\mathcal{C}$ es un π-sistema contenido en un λ-sistema, entonces el λ-sistema más pequeño que contiene a $\mathcal{C}$ es precisamente $\sigma(\mathcal{C})$.

Teorema 11.1.24. *Si $\mathcal{C} \subseteq \mathcal{P}(X)$ es un π-**sistema**, entonces $\delta(\mathcal{C}) = \sigma(\mathcal{C})$.*

Prueba. Sólo hay que demostrar que $\delta(\mathcal{C}) \supseteq \sigma(\mathcal{C})$. Para ello será suficiente verificar que $\delta(\mathcal{C})$ es una σ-álgebra. Sin embargo, gracias al Teorema 11.1.23, sólo debemos comprobar que $\delta(\mathcal{C})$ es estable bajo intersecciones. Defina

$$\mathcal{D} \;=\; \big\{B \in \delta(\mathcal{C}) \,:\, B \cap C \,\in\, \delta(\mathcal{C}) \text{ para todo } C \in \mathcal{C}\big\}$$

y veamos que $\mathcal{D} = \delta(\mathcal{C})$. En efecto, por definición, $\mathcal{D} \subseteq \delta(\mathcal{C})$. Para verificar $\delta(\mathcal{C}) \subseteq \mathcal{D}$, es suficiente demostrar que $\mathcal{D}$ es un D-sistema conteniendo a $\mathcal{C}$. El hecho de que $\mathcal{C}$ es estable bajo intersección implica que $\mathcal{C} \subseteq \mathcal{D}$. Veamos ahora que $\mathcal{D}$ es un D-sistema. Claramente $X \in \mathcal{D}$. Sean $A, B \in \mathcal{D}$ con $B \subseteq A$ y tomemos cualquier $C \in \mathcal{C}$. Entonces

$$(A \setminus B) \cap C \;=\; A \cap C \cap (B^c \cup C^c) \;=\; (A \cap C) \setminus (B \cap C) \;\in\; \delta(\mathcal{C})$$

y, en consecuencia, $A \setminus B \in \mathcal{D}$. Finalmente, sea $(B_n)_{n=1}^{\infty}$ una sucesión disjunta en $\mathcal{D}$ y sea $C \in \mathcal{C}$. Entonces $(B_n \cap C)_{n=1}^{\infty}$ es una sucesión disjunta en $\delta(\mathcal{C})$ y, por lo tanto,

$$\left(\bigcup_{n=1}^{\infty} B_n\right) \cap C \;=\; \bigcup_{n=1}^{\infty} (B_n \cap C) \;\in\; \delta(\mathcal{C}).$$

Esto nos indica que $\bigcup_{n=1}^{\infty} B_n \in \mathcal{D}$ y, por consiguiente, $\mathcal{D}$ es un D-sistema. De allí que $\mathcal{D} = \delta(\mathcal{C})$. Considere ahora la clase

$$\mathcal{D}_1 \;=\; \big\{A \in \delta(\mathcal{C}) \,:\, A \cap B \,\in\, \delta(\mathcal{C}) \text{ para todo } C \in \delta(\mathcal{C})\big\}.$$

De nuevo, como $\mathcal{C}$ es estable bajo intersección, entonces $\mathcal{C} \subseteq \mathcal{D}_1$ y, como en la demostración anterior, $\mathcal{D}_1 = \delta(\mathcal{C})$. Esto último nos revela que $\delta((\mathcal{C}))$ es estable bajo intersección y, por lo tanto, $\delta(\mathcal{C}) = \sigma(\mathcal{C})$. ■

Usualmente, el resultado anterior se expresa en la siguiente forma:

Teorema 11.1.25 (Dynkin π-λ). *Si $\mathcal{C} \subseteq \mathcal{P}(X)$ es un π-**sistema** y $\mathcal{D}$ es un λ-**sistema** conteniendo a $\mathcal{C}$, entonces $\mathcal{D} \supseteq \sigma(\mathcal{C})$.*

Corolario 11.1.26 (Unicidad de la Extensión de Carathéodory). *Si en el Teorema de Extensión de Carathéodory la pre-medida λ es σ-**finita**, entonces la extensión ν es **única**.*

Prueba. Suponga, primeramente, que λ es finita y que $\nu_1 : \sigma(\mathcal{A}) \to [0, +\infty)$ es otra extensión de λ en el Teorema de Extensión de Carathéodory. Como $X \in \mathcal{A}$ y $\nu(X) = \lambda(X) = \nu_1(X)$ resulta que ambas medidas, ν y ν_1, son finitas. Considere ahora la familia

$$\mathcal{D} \;=\; \big\{A \in \sigma(\mathcal{A}) \,:\, \nu(A) = \nu_1(A)\big\}.$$

Observe que $\mathcal{A} \subseteq \mathcal{D}$ ya que $\nu = \nu_1$ sobre $\mathcal{A} \subseteq \sigma(\mathcal{A})$ y por definición, $\mathcal{D} \subseteq \sigma(\mathcal{A})$. Veamos ahora que $\mathcal{D}$ es un D-sistema sobre X. En efecto, es claro que $X \in \mathcal{D}$ y como ν y ν_1 son medidas finitas sobre $\mathcal{A}$, resulta que si $A, B \in \mathcal{D}$ con $B \subseteq A$, entonces

$$\nu_1(A \setminus B) \;=\; \nu_1(A) - \nu_1(B) \;=\; \nu(A) - \nu(B) \;=\; \nu(A \setminus B)$$

lo cual prueba que $A \setminus B \in \mathcal{D}$. Finalmente, si $\big(A_n\big)_{n=1}^{\infty}$ es una sucesión disjunta en $\mathcal{D}$, entonces

$$\nu_1\Big(\bigcup_{n=1}^{\infty} A_n\Big) \;=\; \sum_{n=1}^{\infty} \nu_1(A_n) \;=\; \sum_{n=1}^{\infty} \nu(A_n) \;=\; \nu\Big(\bigcup_{n=1}^{\infty} A_n\Big).$$

Esto prueba que $\mathcal{D}$ es un D-sistema conteniendo a $\mathcal{A}$. Se sigue del Teorema 11.1.25 que $\sigma(\mathcal{A}) \subseteq \mathcal{D}$ y, así, $\mathcal{D} = \sigma(\mathcal{A})$ y termina la prueba para el caso en que λ es finita.

Suponga ahora que λ es σ-finita y elija una sucesión disjunta $\big(X_n\big)_{n=1}^{\infty}$ en $\mathcal{A}$ tal que $X = \bigcup_{n=1}^{\infty} X_n$. Entonces, para cualquier $A \in \sigma(\mathcal{A})$ se tiene que $\nu(A \cap X_n) = \nu_1(A \cap X_n)$ para todo $n \geq 1$ y, por lo tanto,

$$\begin{aligned} \nu_1(A) \;&=\; \nu_1\Big(A \cap \bigcup_{n=1}^{\infty} X_n\Big) \;=\; \sum_{n=1}^{\infty} \nu_1(A \cap X_n) \\ &=\; \sum_{n=1}^{\infty} \nu(A \cap X_n) \;=\; \nu(A). \end{aligned}$$

Esto nos indica que $A \in \mathcal{D}$ y termina la prueba. ■

11.1.4. La Medida de Lebesgue-Stieltjes en $\mathbb{R}$

La medida de Lebesgue-Stieltjes que describiremos en esta sección se obtiene identificando medidas de Borel con funciones crecientes continuas por la derecha. El Teorema de Extensión de Carathéodory juega, en esta identificación, un papel de primer orden. Una vez logrado lo anterior, la medida de Lebesgue en $\mathbb{R}$ se obtiene de manera inmediata. La idea consiste en asociar, a cada función creciente y continua por la derecha $F : \mathbb{R} \to \mathbb{R}$, una pre-medida ν_F definida sobre una cierta álgebra $\mathcal{A}$ y entonces aplicar el Teorema de Extensión de Carathéodory para obtener la medida de Lebesgue-Stieltjes sobre $\sigma(\mathcal{A})$. Lo interesante, en este caso, es que $\sigma(\mathcal{A})$ es exactamente $\mathfrak{Bo}(\mathbb{R})$, la σ-álgebra de Borel. Antes de andar este camino, comencemos recorriendo el camino opuesto.

Recordemos que, gracias al Teorema 3.1.34, página 173, si $F : \mathbb{R} \to \mathbb{R}$ es una *función creciente* sobre $\mathbb{R}$, entonces para cualquier $x \in \mathbb{R}$ los límites

$$\lim_{t \to x^-} F(t) \;=\; F(x^-) \qquad \text{y} \qquad \lim_{t \to x^+} F(t) \;=\; F(x^+)$$

existen y se cumple que

$$-\infty \ < \ F(x^-) \ \leq \ F(x) \ \leq \ F(x^+) \ < \ +\infty.$$

Si, además, F es acotada sobre $\mathbb{R}$, entonces

$$F(-\infty) \ = \ \lim_{t\to-\infty} F(t) \qquad \text{y} \qquad F(+\infty) \ = \ \lim_{t\to+\infty} F(t).$$

son números reales. Si F no es acotada sobre $\mathbb{R}$, entonces $F(-\infty) = -\infty$ o $F(+\infty) = +\infty$ y, en consecuencia, $-\infty \ \leq \ F(-\infty) \ \leq \ F(+\infty) \ \leq \ +\infty$.

Por otro lado, asumiendo de nuevo que F es creciente, del Corolario 3.1.35, página 174, sabemos que el conjunto de todos los puntos donde F no es continua, $\mathrm{Disc}(F)$, es a lo más numerable. Recordemos también que una función $F : \mathbb{R} \to \mathbb{R}$ es *continua por la derecha* en $x \in \mathbb{R}$ si $F(x) = F(x^+)$ y que si F es creciente sobre $\mathbb{R}$, entonces la función asociada $F_+ : \mathbb{R} \to \mathbb{R}$ definida por $F_+(x) = F(x^+)$ es continua por la derecha para todo $x \in \mathbb{R}$, Corolario 3.1.37, página 175. Esta función asociada F_+ posee las siguientes propiedades:

(a) F_+ es creciente sobre $\mathbb{R}$.

(b) $F_+ = F$ excepto sobre un subconjunto a lo más numerable de $\mathbb{R}$.

(c) F_+ es continua en $x \in \mathbb{R}$ si, y sólo si, F es continua en x. Además, si F es continua en $x \in \mathbb{R}$, entonces $F(x) = F_+(x)$.

(d) Existe un conjunto $E \subseteq \mathbb{R}$ con $\mu(E) = 0$ tal que las derivadas $F'(x)$ y $F'_+(x)$ existen y $F'(x) = F'_+(x)$ para todo $x \in \mathbb{R} \setminus E$.

Por consiguiente, dada cualquier función creciente $F : \mathbb{R} \to \mathbb{R}$ siempre podemos asumir que ella es continua por la derecha sobre $\mathbb{R}$.

Recuerde también que toda medida ν cuyo dominio es $\mathfrak{Bo}(\mathbb{R})$, la σ-álgebra de Borel sobre $\mathbb{R}$, es llamada una **medida de Borel**. Denotemos por $\mathcal{M}_{\mathrm{bo}}(\mathbb{R})$ *la familia de todas las medidas de Borel que son **finitas** sobre* $\mathbb{R}$.

Teorema 11.1.27. *Sea $\nu \in \mathcal{M}_{\mathrm{bo}}(\mathbb{R})$. La función $F_\nu : \mathbb{R} \to \mathbb{R}$ definida por*

$$F_\nu(x) \ = \ \nu\big((-\infty, x]\big) \quad \text{para todo } x \in \mathbb{R}$$

*Entonces F_ν es **creciente** y **continua por la derecha** sobre $\mathbb{R}$.*

Prueba. Sean $x_1, x_2 \in \mathbb{R}$ con $x_1 < x_2$. Como $(-\infty, x_1] \subseteq (-\infty, x_2]$ y ν es una medida, se sigue que $F_\nu(x_1) \leq F_\nu(x_2)$. Veamos ahora que F_ν es continua por la derecha en x, para cualquier $x \in \mathbb{R}$. En efecto, sea $x \in \mathbb{R}$ y elija cualquier sucesión positiva y decreciente $(x_n)_{n=1}^\infty$ tal que $\lim_{n\to\infty} x_n = x$. Entonces $\bigcap_{n=1}^\infty (-\infty, x_n] = (-\infty, x]$ y como ν es finita, resulta del Teorema 11.1.5 que

$$\nu((-\infty, x]) \ = \ \lim_{n\to\infty} \nu\big((-\infty, x_n]\big),$$

en otras palabras, $F_\nu(x) = \lim\limits_{n\to\infty} F_\nu(x_n)$ y termina la prueba. ■

La función creciente y continua por la derecha F_ν obtenida en el teorema anterior se llama una **función de distribución** asociada a ν. Nótese que, en este caso, F_ν es acotada sobre X. En particular, si ν es una **medida de probabilidad**, entonces,

$$F_\nu(-\infty) \ = \ \lim_{x\to-\infty} F_\nu(x) \ = \ 0 \qquad \text{y} \qquad F_\nu(+\infty) \ = \ \lim_{x\to+\infty} F_\nu(x) \ = \ 1.$$

Para obtener la medida de Lebesgue-Stieltjes hay que ir en la otra dirección, es decir, se comienza con una función arbitraria F creciente y continua por la derecha y a través de un proceso sencillo y elegante se construye una medida de Borel. Para este fin, considere la colección $\mathcal{C}$ de todos los intervalos semi-cerrados más el conjunto vacío, esto es,

$$\mathcal{C} \;=\; \big\{(a,b] \,:\, a,b \in \mathbb{R},\; a < b\big\} \cup \{\varnothing\},$$

La elección de los intervalos del tipo $(a,b]$ no sólo es una cuestión de conveniencia técnica: ella es más adecuada y práctica que otras elecciones: por ejemplo, una ventaja sobre los intervalos abiertos es que la unión de dos intervalos consecutivos, digamos, $(a,b]$ y $(b,c]$ en $\mathcal{C}$ es, de nuevo, un miembro de $\mathcal{C}$. También, si $(a,b], (c,d]$ son dos intervalos arbitrarios en $\mathcal{C}$, entonces su intersección es, o bien vacía, o es de la forma $(a',b']$. En cualquier caso, $(a,b] \cap (c,d]$ sigue siendo un elemento de $\mathcal{C}$. Sin embargo, la diferencia $(a,b] \setminus (c,d]$ puede, en algunos casos, ser un miembro de $\mathcal{C}$ si ocurre, por ejemplo, que $a < c < b < d$, pero, hay casos donde tal cosa no sucede: por ejemplo, si $a < c < d < b$, entonces $(a,b] \setminus (c,d] = (a,c] \cup (d,b]$ es la unión disjunta de dos intervalos de $\mathcal{C}$. Aunque la clase $\mathcal{C}$ no es una álgebra ella satisface, gracias al Ejemplo 6.3.1, página 266, que

$$\sigma(\mathcal{C}) \;=\; \mathfrak{Bo}(\mathbb{R}).$$

La idea es ir ampliando la clase $\mathcal{C}$ hasta obtener una álgebra. Para lograr eso, considere, en primer lugar, la clase

$$\mathcal{A}_0 \;=\; \mathcal{C} \cup \big\{(-\infty,b] \,:\, b \in \mathbb{R}\big\} \cup \big\{(a,+\infty) \,:\, a \in \mathbb{R}\big\}$$

y observe que $A \in \mathcal{A}_0$ si, y sólo si, $A^c \in \mathcal{A}_0$. Por supuesto, $\mathcal{A}_0$ no es una álgebra ya que la unión de dos miembros disjuntos de $\mathcal{A}$ no pertenece a dicha familia, pero estamos cerca. Si ahora definimos $\mathcal{A}$ como la familia formada por *todas las **uniones finitas** y **disjuntas** de elementos de* $\mathcal{A}_0$, resulta que $\mathcal{A}_0$ es una álgebra que satisface la condición

$$\sigma(\mathcal{A}) \;=\; \mathfrak{Bo}(\mathbb{R}).$$

Dejamos a cargo del lector verificar que $\mathcal{A}$ es una álgebra. Nótese, sin embargo, que $\mathcal{A}$ no es una σ-álgebra pues la sucesión $((-1/n,1])_{n=1}^{\infty}$ pertenece a $\mathcal{A}$ pero no así su intersección:

$$\bigcap_{n=1}^{\infty}(-1/n,1] \;=\; [0,1] \notin \mathcal{A}.$$

Observe que cualquier elemento $A \in \mathcal{A}$ se puede expresar en la forma $A = \bigcup_{n=1}^{\infty}(a_n,\, b_n]$, donde la colección $\big((a_n,\, b_n]\big)_{n=1}^{\infty}$ es disjunta e incluida en $\mathcal{C}$.

El plan ahora es definir una cierta función de conjuntos ν sobre la clase $\mathcal{C}$ y luego extenderla sobre la álgebra $\mathcal{A}$ de modo que dicha extensión sea una pre-medida y finalmente aplicar el Teorema de Extensión de Carathéodory.

Fijemos una función $F : \mathbb{R} \to \mathbb{R}$ creciente y continua por la derecha en cada punto $x \in \mathbb{R}$ y defina $\nu : \mathcal{C} \to [0,+\infty)$ por

$$\nu(\varnothing) \;=\; 0 \qquad\qquad \text{y} \qquad\qquad \nu\big((a,b]\big) \;=\; F(b) - F(a).$$

Observe que:

(a) $\nu \geq 0$.

(b) $\lim\limits_{b \to a+} \nu\big((a,b]\big) = 0$. Sigue de la continuidad por la derecha de F.

(c) Si $\mathcal{P} = \{a = a_1, b_1, a_2, b_2 \ldots, a_n, b_n = b\}$ es una partición finita de $(a,b]$, entonces

$$\nu((a,b]) \;=\; F(b) - F(a) \;=\; \sum_{i=1}^{n} F(b_i) - F(a_i) \;=\; \sum_{i=1}^{n} \nu((a_i, b_i]).$$

(d) $\nu((a_1,b_1]) \leq \nu((a_2,b_2])$ si $(a_1,b_1] \subseteq (a_2,b_2]$. Puesto que $(a_2,b_2] = (a_2,a_1] \cup (a_1,b_1] \cup (b_1,b_2]$, se sigue de (c) y (a) que

$$\nu((a_2,b_2]) \;=\; \nu((a_2,a_1]) + \nu((a_1,b_1]) + \nu((b_1,a_2]) \;\geq\; \nu((a_1,b_1]).$$

Lema 11.1.28. *Si $F : \mathbb{R} \to \mathbb{R}$ es una* ***función creciente*** *y* ***continua por la derecha*** *sobre $\mathbb{R}$, entonces la función de conjuntos $\nu : \mathcal{C} \to [0, +\infty)$ definida por*

$$\nu(\varnothing) \;=\; 0 \qquad y \qquad \nu\big((a,b]\big) \;=\; F(b) - F(a). \tag{1}$$

es ***numerablemente aditiva*** *sobre $\mathcal{C}$.*

Prueba. Suponga que $(a,b] = \bigcup_{i=1}^{\infty}(a_i,b_i]$, donde la colección $((a_i,b_i])_{i=1}^{\infty}$ es disjunta. Considere n intervalos arbitrarios de la colección $((a_i,b_i])_{i=1}^{\infty}$ los cuales se pueden arreglar en la forma $(a_{i1},b_{i1}], (a_{i2},b_{i2}], \ldots, (a_{in},b_{in}]$ donde

$$a_{i1} \;<\; b_{i1} \;\leq\; a_{i2} \;<\; b_{i2} \;\leq\; \cdots \;\leq\; a_{in} \;<\; b_{in}.$$

Usando la propiedad (d) vista anteriormente, se tiene que

$$\sum_{j=1}^{n} \nu\big((a_{ij},b_{ij}]\big) \;=\; \nu\Big(\bigcup_{j=1}^{n}(a_{ij},b_{ij}]\Big) \;\leq\; \nu\big((a_{i1},b_{in}]\big) \;\leq\; \nu\big((a,\, b]\big)$$

y, así, como n fue elegido de modo arbitrario, se concluye que $\sum_{i=1}^{\infty} \nu\big((a_i,b_i]\big) \leq \nu\big((a,\, b]\big)$. Para establecer la otra desigualdad, es decir,

$$\sum_{i=1}^{\infty} \nu\big((a_i,b_i]\big) \;\geq\; \nu\big((a,\, b]\big) \tag{2}$$

escoja un $\varepsilon > 0$ de modo que $\varepsilon < b - a$ y observe que $a < a + \varepsilon < b$. Miremos ahora el intervalo $[a+\varepsilon, b] \subseteq (a,b]$. Como $\lim_{\delta \downarrow 0} \nu\big((b_i, b_i + \delta]\big) = 0$ para cada $i \in \mathbb{N}$, seleccione un $\delta_i > 0$ tal que

$$\nu\big((b_i, b_i + \delta_i]\big) \;<\; \frac{\varepsilon}{2^i}.$$

De esto se sigue que

$$\begin{aligned}\nu\big((a_i, b_i + \delta_i]\big) \;&=\; \nu\big((a_i,b_i]\big) + \nu\big((b_i, b_i+\delta_i]\big)\\ &<\; \nu\big((a_i,b_i]\big) + \frac{\varepsilon}{2^i}\end{aligned}$$

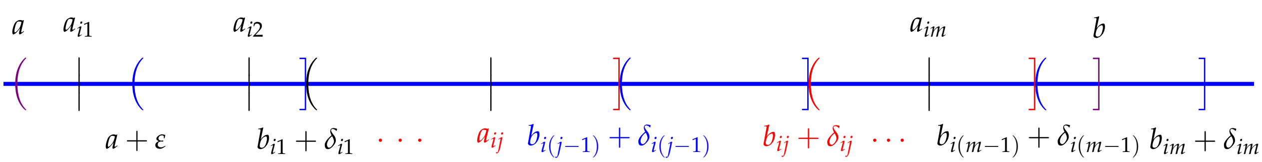
Observe que los intervalos abiertos $\{(a_i, b_i + \delta_i) : i = 1, 2, \ldots\}$ cubren al intervalo compacto $[a + \varepsilon, b]$, de modo que dicho cubrimiento, gracias al Teorema de Heine-Borel, se reduce a un subcubrimiento finito, es decir, existe un $n_0 \in \mathbb{N}$ tal que

$$[a + \varepsilon, b] \subseteq \bigcup_{j=1}^{n_0} (a_j, b_j + \delta_j).$$

De estos n_0 intervalos, seleccionemos m de ellos con las siguientes propiedades:

$$a_{i1} < a + \varepsilon, \qquad b_{im} + \delta_{im} > b,$$

$$a_{i(j-1)} < a_{ij} < b_{i(j-1)} + \delta_{i(j-1)} < b_{ij} + \delta_{ij}, \quad j = 2, \ldots, m.$$



Si consideramos ahora los intervalos:

$$\big(a_{i1}, b_{i1} + \delta_{i1}\big], \quad \big(b_{i1} + \delta_{i1}, b_{i2} + \delta_{i2}\big], \ldots, \big(b_{i(m-1)} + \delta_{i(m-1)}, b_{im} + \delta_{im}\big]$$

se observa que ellos son disjuntos y su unión es el intervalo $\big(a_{i1}, b_{im} + \delta_{im}\big]$. Por esto, $(a + \varepsilon, b] \subseteq \big(a_{i1}, b_{im} + \delta_{im}\big]$ y entonces

$$\begin{aligned}
\nu\big((a + \varepsilon, b]\big) &\leq \nu\Big(\big(a_{i1}, b_{im} + \delta_{im}\big]\Big) \\
&= \nu\Big(\big(a_{i1}, b_{i1} + \delta_{i1}\big]\Big) + \sum_{j=2}^{m} \nu\Big(\big(b_{i(j-1)} + \delta_{i(j-1)}, b_{ij} + \delta_{ij}\big]\Big) \\
&\leq \sum_{j=1}^{m} \nu\Big(\big(a_{ij}, b_{ij} + \delta_{ij}\big]\Big),
\end{aligned}$$

ya que $\big(b_{i(j-1)} + \delta_{i(j-1)}, b_{ij} + \delta_{ij}\big] \subseteq \big(a_{ij}, b_{ij} + \delta_{ij}\big]$. Es decir,

$$\nu\big((a + \varepsilon, b]\big) \leq \sum_{j=1}^{m} \nu\Big(\big(a_{ij}, b_{ij} + \delta_{ij}\big]\Big).$$

Por otro lado, como $\big(a_{ij}, b_{ij} + \delta_{ij}\big]$, $j = 1, \ldots, m$ es sólo una cantidad finita de la colección de intervalos $\{(a_j, b_j + \delta_j] : j = 1, 2, \dots\}$, resulta que

$$\begin{aligned}
\sum_{j=1}^{m} \nu\big((a_{ij}, b_{ij} + \delta_{ij}]\big) &\le \sum_{j=1}^{\infty} \nu\big((a_j, b_j + \delta_j]\big) \\
&= \sum_{j=1}^{\infty} \Big[\nu\big((a_j, b_j]\big) + \nu\big((b_j, b_j + \delta_j]\big)\Big] \\
&< \sum_{j=1}^{\infty} \Big[\nu\big((a_j, b_j]\big) + \frac{\varepsilon}{2^j}\Big] = \sum_{j=1}^{\infty} \nu\big((a_j, b_j]\big) + \varepsilon.
\end{aligned}$$

Más aun, puesto que $\nu\big((a + \varepsilon, b]\big) = F(b) - F(a+\varepsilon)$, entonces

$$F(b) - F(a+\varepsilon) \le \sum_{j=1}^{\infty} \nu\big((a_j, b_j]\big) + \varepsilon.$$

Si hacemos que $\varepsilon \to 0+$, entonces la continuidad por la derecha de F nos revela que

$$\nu\big((a, b]\big) = F(b) - F(a) \le \sum_{j=1}^{\infty} \nu\big((a_j, b_j]\big)$$

que era lo que queríamos demostrar. ■

Teorema 11.1.29 (Lebesgue-Stieltjes). *Sea $F : \mathbb{R} \to \mathbb{R}$ una **función creciente** y **continua por la derecha** sobre $\mathbb{R}$. Entonces existe una **única medida de Borel** $\nu_F : \mathfrak{Bo}(\mathbb{R}) \to [0, +\infty)$ tal que*

$$\nu_F\big((a, b]\big) = F(b) - F(a)$$

para todo $(a, b] \in \mathcal{C}$.

Prueba. Por el Lema 11.1.28 sabemos que la función de conjuntos $\widetilde{\nu} : \mathcal{C} \to [0, +\infty)$ definida por

$$\widetilde{\nu}(\varnothing) \qquad \text{y} \qquad \widetilde{\nu}\big((a, b]\big) = F(b) - F(a)$$

para todo $(a, b] \in \mathcal{C}$ es numerablemente aditiva sobre $\mathcal{C}$. Considere la álgebra $\mathcal{A}$ formada por la colección de todas las uniones finitas y disjuntas de elementos de $\mathcal{A}_0$, donde

$$\mathcal{A}_0 = \mathcal{C} \cup \big\{(-\infty, b] : b \in \mathbb{R}\big\} \cup \big\{(a, +\infty) : a \in \mathbb{R}\big\}.$$

Ya hemos visto que todo elemento $A \in \mathcal{A}$ se puede expresar en la forma $A = \bigcup_{n=1}^{\infty}(a_n, b_n]$ para alguna colección disjunta $\big((a_n, b_n]\big)_{n=1}^{\infty}$ incluida en $\mathcal{C}$. Defina entonces $\nu : \mathcal{A} \to [0, +\infty]$ por

$$\nu(A) = \sum_{n=1}^{\infty} \widetilde{\nu}\big((a_n, b_n]\big)$$

para cada $A \in \mathcal{A}$ representada del modo anterior. Veamos que ν está bien definida, es decir, no depende de la colección particular escogida para representar al conjunto A. En efecto, suponga que $\big((a_n,\, b_n]\big)_{n=1}^{\infty}$ y $\big((c_m,\, d_m]\big)_{m=1}^{\infty}$ son colecciones en $\mathcal{C}$, cada una disjunta, tal que $A = \bigcup_{n=1}^{\infty}(a_n,\, b_n] = \bigcup_{n=1}^{\infty}(c_n,\, d_n]$ y veamos que

$$\sum_{n=1}^{\infty} \widetilde{\nu}((a_n,\, b_n]) \;=\; \sum_{m=1}^{\infty} \widetilde{\nu}((c_m,\, d_m]).$$

Claramente $(a_n,\, b_n] \cap (c_m,\, d_m] \in \mathcal{C}$ para todo $n, m \in \mathbb{N}$. Puesto que

$$(a_n,\, b_n] \;=\; (a_n,\, b_n] \cap A \;=\; (a_n,\, b_n] \cap \bigcup_{m=1}^{\infty} (c_m,\, d_m] \;=\; \bigcup_{m=1}^{\infty} (a_n,\, b_n] \cap (c_m,\, d_m],$$

para todo $n \in \mathbb{N}$, resulta que al ser $\widetilde{\nu}$ una medida sobre $\mathcal{C}$ se cumple que $\widetilde{\nu}((a_n,\, b_n]) = \sum_{m=1}^{\infty} \widetilde{\nu}\big((a_n,\, b_n] \cap (c_m,\, d_m]\big)$. De forma similar, $\widetilde{\nu}((c_m,\, d_m]) = \sum_{n=1}^{\infty} \widetilde{\nu}\big((a_n,\, b_n] \cap (c_m,\, d_m]\big)$ para todo $m \in \mathbb{N}$. Por esto,

$$\begin{aligned}
\sum_{n=1}^{\infty} \widetilde{\nu}\big((a_n,\, b_n]\big) &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \widetilde{\nu}\big((a_n,\, b_n] \cap (c_m,\, d_m]\big) \\
&= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \widetilde{\nu}\big((a_n,\, b_n] \cap (c_m,\, d_m]\big) \\
&= \sum_{m=1}^{\infty} \widetilde{\nu}\big((c_m,\, d_m]\big).
\end{aligned}$$

Es inmediato que ν coincide con $\widetilde{\nu}$ sobre $\mathcal{C}$. Para poder aplicar el Teorema de Extensión de Carathéodory sólo nos falta verificar que ν es numerablemente aditiva sobre $\mathcal{A}$. Suponga entonces que $\big(A_n\big)_{n=1}^{\infty}$ es una sucesión disjunta en $\mathcal{A}$ tal que $A = \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$. Puesto que $A_n \in \mathcal{A}$ para cada $n \in \mathbb{N}$, resulta que tal conjunto se puede representar en la forma $A_n = \bigcup_{m=1}^{\infty} I_{nm}$ donde $(I_{nm})_{m=1}^{\infty}$ es una colección disjunta en $\mathcal{C}$ y, en consecuencia, $\nu(A_n) = \sum_{m=1}^{\infty} \widetilde{\nu}(I_{nm})$. Claramente $\{I_{nm} : n, m \in \mathbb{N}\}$ es una colección disjunta cuya unión cubre al conjunto A y, por lo tanto,

$$\begin{aligned}
\nu(A) \;=\; \sum_{n,m=1}^{\infty} \widetilde{\nu}\big(I_{nm}\big) \;&=\; \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \widetilde{\nu}(I_{nm}) \\
&= \sum_{n=1}^{\infty} \nu(A_n).
\end{aligned}$$

Hemos demostrado que ν es una pre-medida sobre la álgebra $\mathcal{A}$. Más aun, como

$$\mathbb{R} \;=\; \bigcup_{n=1}^{\infty} (-n, n] \qquad \text{y} \qquad \nu\big((-n, n]\big) \;=\; F(n) - F(-n) \;<\; +\infty$$

se sigue que ν es σ-finita. Un llamado al Teorema de Extensión de Carathéodory nos garantiza la existencia de una única medida sobre $\sigma(\mathcal{A}) = \mathfrak{Bo}(\mathbb{R})$, que denotaremos por ν_F, tal que $\nu_F((a, b]) = F(b) - F(a)$ para todo $(a, b] \in \mathcal{C}$. Esto termina la prueba. ■

La medida ν_F obtenida en el Teorema 11.1.29 se llama la **medida de Lebesgue-Stieltjes** asociada a la función creciente y continua por la derecha F. Algunos autores llaman medida de Lebesgue-Stieltjes a la completación de ν_F. La integral respecto a esta medida, véase la Sección 11.1.7, página 712, se denomina la **integral de Lebesgue-Stieltjes** y se denota por

$$\text{(LS)}\int_a^b f(t)\, d\nu_F(t).$$

Si en los resultados anteriores reemplazamos la familia $\mathcal{C}$ por $\mathcal{C}' = \{[a,b) : a,b \in \mathbb{R},\, a < b\}$ y la función creciente y continua por la derecha F por cualquier **función decreciente** y **continua por la izquierda**, se desarrolla una teoría enteramente similar.

Nota Adicional 11.1.4 (i) Nótese que si ν_F es la medida de Lebesgue-Stieltjes asociada a una función creciente y continua por la derecha F, entonces para cada $x \in \mathbb{R}$,

$$\nu_F(\{x\}) \;=\; \lim_{n\to\infty} \nu_F\big((x - 1/n,\, x]\big) \;=\; F(x) - F(x^-),$$

y como F es continua por la derecha en x, resulta que $F(x) = F(x^+)$ y, en consecuencia,

$$F \text{ es } \textbf{continua} \text{ en } x \quad \Leftrightarrow \quad F(x^-) = F(x) = F(x^+) \quad \Leftrightarrow \quad \nu_F(\{x\}) \;=\; 0.$$

Además, si F es **acotada**, podemos expresar la medida de cualquier subintervalo de $\mathbb{R}$ en función de la función de distribución F. En efecto,

$$\nu_F\big((a,\, b)\big) \;=\; \lim_{n\to\infty} \nu_F\big((a,\, b - 1/n]\big) \;=\; F(b^-) - F(a),$$

$$\nu_F\big([a,\, b)\big) \;=\; \nu_F(\{x\}) + \nu_F\big((a,\, b)\big) \;=\; F(b^-) - F(a^-),$$

$$\nu_F\big([a,\, +\infty)\big) \;=\; F(+\infty) - F(a^-) \qquad \text{y} \qquad \nu_F\big((-\infty,\, +\infty)\big) \;=\; F(+\infty) - F(-\infty).$$

(ii) Algunos casos particularmente importantes de medidas de Lebesgue-Stieltjes se obtienen cuando la función F se elige adecuadamente. Por ejemplo:

(a) Si $F(x) = x$ para todo $x \in \mathbb{R}$, entonces $\nu_F = \mu$ sobre $\mathfrak{Bo}(\mathbb{R})$. En efecto, por el Teorema de Lebesgue-Stieltjes existe una única medida de Borel ν_F tal que $\nu_F((a,b]) = b - a$ para todo intervalo $(a,b] \in \mathcal{C}$. Por otro lado, como la medida de Lebesgue también cumple que $\mu((a,b]) = b - a$, entonces la unicidad nos asegura que $\nu_F = \mu$ sobre $\mathfrak{Bo}(\mathbb{R})$. Este método constituye, por supuesto, otro enfoque de presentar la medida de Lebesgue sobre $\mathfrak{Bo}(\mathbb{R})$.

(b) Sea $f : \mathbb{R} \to [0, +\infty)$ una **función continua** sobre $\mathbb{R}$ y defina $F : \mathbb{R} \to \mathbb{R}$ por

$$F(x) \;=\; \begin{cases} \displaystyle\int_0^x f(t)\, dt & \text{si } x > 0, \\ 0 & \text{si } x = 0, \\ \displaystyle -\int_0^x f(t)\, dt & \text{si } x < 0. \end{cases}$$

Puesto que F es creciente y continua, entonces la medida de Lebesgue-Stieltjes ν_F asociada a F satisface

$$\nu_F\big((a,\, b]\big) \;=\; \int_a^b f(t)\, dt.$$

(c) Para cada $x \in \mathbb{R}$, definamos la **medida de Dirac** $\delta_x : \mathfrak{Bo}(\mathbb{R}) \to [0, +\infty)$ por

$$\delta_x(E) = \chi_E(x) = \begin{cases} 1 & \text{si } x \in E \\ 0 & \text{si } x \notin E \end{cases}$$

para todo $E \in \mathfrak{Bo}(\mathbb{R})$. Si definimos la función $F : \mathbb{R} \to \mathbb{R}$ por

$$F(t) = \begin{cases} \delta_x\big((0, t]\big) & \text{si } t \geq 0 \\ -\delta_x\big((t, 0]\big) & \text{si } t < 0 \end{cases}$$

se tiene que F es creciente y continua por la derecha sobre $\mathbb{R}$.

Sea $(x_n)_{n=1}^{\infty}$ una sucesión en $\mathbb{R}$ y suponga que $(a_n)_{n=1}^{\infty}$ es una sucesión no-negativa de números reales tal que $\sum_{n=1}^{\infty} a_n < +\infty$. Entonces $F = \sum_{n=1}^{\infty} a_n \chi_{[x_n, +\infty)}$ es la función de distribución de la medida finita $\nu_F = \sum_{n=1}^{\infty} a_n \delta_{x_n}$.

(iii) Denotemos por $\mathcal{F}_{\text{accd}}(\mathbb{R})$ el conjunto de todas las funciones $F : \mathbb{R} \to \mathbb{R}$ que son acotadas, crecientes y continuas por la derecha sobre $\mathbb{R}$. Sean $F, G \in \mathcal{F}_{\text{accd}}(\mathbb{R})$. Si las medidas de Lebesgue-Stieltjes asociadas a ellas a F y G satisfacen $\nu_F = \nu_G$, entonces $F - G$ es una constante. En efecto, sean $a, b \in \mathbb{R}$ con $a < b$. Entonces

$$\nu_F((a, b]) = \nu_G((a, b]) \quad \Rightarrow \quad F(b) - F(a) = G(b) - G(a) \quad \Rightarrow \quad (F - G)(b) = (F - G)(a)$$

y como a y b son arbitrarios se concluye que $F - G$ es constante. Sobre $\mathcal{F}_{\text{accd}}(\mathbb{R})$ introduzcamos la siguiente relación:

$$F \sim G \quad \Leftrightarrow \quad F - G \quad \text{es una constante.}$$

Resulta que $\sim$ es una relación de equivalencia sobre $\mathcal{F}_{\text{accd}}(\mathbb{R})$. Identificando cada función $F \in \mathcal{F}_{\text{accd}}(\mathbb{R})$ con su clase de equivalencia, podemos concluir que la aplicación

$$\Psi : \mathcal{M}_{\text{bo}}(\mathbb{R}) \to \mathcal{F}_{\text{accd}}(\mathbb{R})$$

que asocia a cada $\nu \in \mathcal{M}_{\text{bo}}(\mathbb{R})$ la función de distribución F_ν es una biyección. De este resultado se concluye que:

> *Cualquier* ***medida de Borel finita*** *es una* ***medida de Lebesgue-Stieltjes*** *para alguna función de distribución acotada.*

A diferencia de lo que ocurre con la medida de Lebesgue, la medida de Lebesgue-Stieltjes ν_F asociada a una función $F \in \mathcal{F}_{\text{accd}}(\mathbb{R})$ no cumple, necesariamente, la propiedad de que $\nu_F(\{a\}) = 0$ para cualquier $a \in \mathbb{R}$, es decir, ν_F puede tener un átomo en $\{a\}$. En efecto, puesto que $\nu_F(\{a\}) = 0$ si, y sólo si, F es continua en a, resulta que si F no es continua en a, entonces el Teorema 3.1.34, página 173 nos garantiza que F posee una discontinuidad de salto en a y, por lo tanto, $\nu_F(\{a\}) = F(a) - F(a^-) > 0$. Puesto que el número de discontinuidades de F es a lo más numerable, Corolario 3.1.35, página 174, se sigue que el número de átomos de ν_F es a lo más numerable.

11.1.5. Funciones de Variación Acotada sobre $\mathbb{R}$

En esta sección presentaremos dos resultados que vinculan a μ, la medida de Lebesgue sobre $\mathfrak{Bo}(\mathbb{R})$, con ν_F la medida de Lebesgue-Stieltjes asociada a una función F absolutamente continua sobre $\mathbb{R}$. Antes, debemos precisar algunas nociones sobre funciones de variación acotada sobre $\mathbb{R}$. Denote por $\mathcal{F}_{\mathrm{IC}}(\mathbb{R})$ *la colección de todos los intervalos cerrados* $[a,b]$ incluidos en $\mathbb{R}$.

Definición 11.1.30. *Una función $f:\mathbb{R}\to\mathbb{R}$ es **localmente de variación acotada** en $\mathbb{R}$ si para todo intervalo $[a,b]\subseteq\mathbb{R}$ se cumple que $V(f,[a,b])<+\infty$. Si ocurre que*

$$V(f,\mathbb{R}) \;=\; \sup_{[a,b]\in\mathcal{F}_{\mathrm{IC}}(\mathbb{R})} V(f,[a,b]) \;<\; +\infty$$

*entonces se dice que f es de **variación acotada sobre** $\mathbb{R}$. En este caso escribiremos $f\in\mathrm{BV}(\mathbb{R})$.*

Si $f\in\mathrm{BV}([a,b])$, entonces podemos extender f a una función $\widetilde{f}:\mathbb{R}\to\mathbb{R}\in\mathrm{BV}(\mathbb{R})$ declarando que

$$\widetilde{f}(x) \;=\; \begin{cases} f(a) & \text{si } x<a \\ f(x) & \text{si } x\in[a,b] \\ f(b) & \text{si } b<x. \end{cases}$$

Observe que $\widetilde{f}$ es acotada sobre $\mathbb{R}$ y, obviamente, se cumple que $V(\widetilde{f},\mathbb{R})=V(f,[a,b])<+\infty$. Recíprocamente, si $f\in\mathrm{BV}(\mathbb{R})$, entonces la restricción de f a $[a,b]$, $f|_{[a,b]}$, está en $\mathrm{BV}([a,b])$ para todo intervalo $[a,b]\subseteq\mathbb{R}$.

Una diferencia importante entre los espacios $\mathrm{BV}([a,b])$ y $\mathrm{BV}(\mathbb{R})$ es la siguiente. Ya hemos visto que si $f\in\mathrm{BV}(\mathbb{R})$, entonces $f\in\mathrm{BV}([a,b])$ para todo intervalo $[a,b]\subseteq\mathbb{R}$. Sin embargo, lo contrario no es cierto, es decir, el hecho de que $f\in\mathrm{BV}([a,b])$ para todo intervalo $[a,b]\subseteq\mathbb{R}$, no garantiza que $f\in\mathrm{BV}(\mathbb{R})$. En efecto, si tomamos $f:\mathbb{R}\to\mathbb{R}$ definida por $f(x)=\mathrm{sen}(x)$, resulta que $f\in\mathrm{BV}([a,b])$ para todo intervalo $[a,b]\subseteq\mathbb{R}$, pero $f\notin\mathrm{BV}(\mathbb{R})$.

Considera ahora cualquier función $f\in\mathrm{BV}(\mathbb{R})$ y defina la función $V_f:\mathbb{R}\to\mathbb{R}$ por

$$V_f(x) \;=\; \begin{cases} f(0)\,+\,V(f,[0,x]) & \text{si } x\geq 0 \\ f(0)\,-\,V(f,[x,0]) & \text{si } x<0. \end{cases}$$

Se comprueba sin dificultad que V_f es creciente y satisface $V_f(b)-V_f(a) \;=\; V(f,[a,b])$ para todo intervalo $[a,b]\subseteq\mathbb{R}$. Más aun, si hacemos $f_1=\frac{1}{2}(V_f+f)$ y $f_2=\frac{1}{2}(V_f-f)$, entonces

$$f \;=\; f_1 \,-\, f_2,$$

lo cual nos revela que cada $f\in\mathrm{BV}(\mathbb{R})$ se puede descomponer como la diferencia de dos funciones crecientes.

Definición 11.1.31. *Una función $f:\mathbb{R}\to\mathbb{R}$ se llama **localmente absolutamente continua** sobre $\mathbb{R}$ si $f\in\mathrm{AC}([a,b])$ para todo intervalo cerrado $[a,b]\subseteq\mathbb{R}$. Diremos que f es **absolutamente continua** sobre $\mathbb{R}$, en notación $f\in\mathrm{AC}(\mathbb{R})$, si dado $\varepsilon>0$, existe un $\delta>0$ tal que*

$$\sum_{n=1}^{\infty}|f(b_n)-f(a_n)| \;<\; \varepsilon$$

siempre que $\{(a_n,b_n):n=1,2,\ldots\}$ es cualquier colección numerable de intervalos abiertos y disjuntos incluidos en $\mathbb{R}$ satisfaciendo la desigualdad $\sum_{n=1}^{\infty}(b_n-a_n)<\delta$.

Claramente $\mathrm{AC}(\mathbb{R}) \subseteq \mathrm{C}(\mathbb{R}) \subseteq \mathrm{BV}(\mathbb{R})$ y así, toda función $f \in \mathrm{AC}(\mathbb{R})$ se puede expresar como la diferencia de dos funciones crecientes y localmente absolutamente continuas. Escriba $f = F_1 - F_2$, donde F_1 y F_2 son crecientes y localmente absolutamente continuas. Si ν_1 y ν_2 son las medidas de Lebesgue-Stieltjes asociadas a F_1 y F_2 respectivamente, entonces $\nu_F = \nu_1 - \nu_2$ es una medida de Borel asociada a F.

Teorema 11.1.32. *Sea $F : \mathbb{R} \to \mathbb{R}$ una función* ***absolutamente continua sobre*** *$\mathbb{R}$. Si ν_F es la medida de Lebesgue-Stieltjes asociada a F, entonces $\nu_F \ll \mu$.*

Prueba. Sea $\varepsilon > 0$. Puesto que F es absolutamente continua sobre $\mathbb{R}$, existe un $\delta > 0$ tal que

$$\sum_{n=1}^{\infty} |F(b_n) - F(a_n)| \ < \ \varepsilon$$

siempre que $\{(a_n, b_n) : n = 1, 2, \ldots\}$ es una colección numerable de intervalos abiertos y disjuntos satisfaciendo la desigualdad $\sum_{n=1}^{\infty}(b_n - a_n) < \delta$. Si hacemos $A = \bigcup_{n=1}^{\infty}(a_n, b_n)$, se tiene que $\nu_F(A) < \varepsilon$ siempre que $\mu(A) < \delta$.

Suponga ahora que $E \in \mathfrak{M}_\mu(\mathbb{R})$ con $\mu(E) = 0$ y escoja una sucesión $((a_n, b_n))_{n=1}^{\infty}$ de intervalos abiertos y disjuntos tales que

$$E \subseteq \bigcup_{n=1}^{\infty}(a_n, b_n) \qquad \text{y} \qquad \sum_{n=1}^{\infty}(b_n - a_n) \ < \ \mu(E) + \delta \ = \ \delta.$$

Por lo probado anteriormente resulta que $\nu_F(E) \leq \varepsilon$ y como $\varepsilon > 0$ es arbitrario, se concluye que $\nu_F(E) = 0$. ■

11.1.6. Funciones Medibles

Sea $(X, \mathcal{M})$ un espacio medible. Una función $f : X \to \overline{\mathbb{R}}$ se llama $\mathcal{M}$-**medible** si para cada $a \in \mathbb{R}$ el conjunto

$$f^{-1}((a, +\infty]) \ = \ \{x \in X \ : \ f(x) > a\} \ \in \ \mathcal{M}.$$

Es fácil ver que las condiciones siguientes son equivalentes para cualquier función $f : X \to \overline{\mathbb{R}}$:

(1) f es $\mathcal{M}$-medible.

(2) $\{x \in X \ : \ f(x) \geq a\} \in \mathcal{M}$ para todo $a \in \mathbb{R}$.

(3) $\{x \in X \ : \ f(x) < a\} \in \mathcal{M}$ para todo $a \in \mathbb{R}$.

(4) $\{x \in X \ : \ f(x) \leq a\} \in \mathcal{M}$ para todo $a \in \mathbb{R}$.

Si denotamos por $\mathcal{F}(X)$, *la colección de todas las funciones $f : X \to \overline{\mathbb{R}}$ que son $\mathcal{M}$-medibles*, es fácil establecer que $\mathcal{F}(X)$ es un espacio vectorial sobre $\mathbb{R}$. Más aun, si $f, g \in \mathcal{F}(X)$ y $(f_n)_{n=1}^{\infty}$ es una sucesión de funciones $\mathcal{M}$-medibles, entonces:

$$f \cdot g, \quad \text{máx}\{f, g\}, \quad \text{mín}\{f, g\}, \quad \sup_n f_n, \quad \inf_n f_n, \quad \text{lím}\sup_n f_n, \quad \text{lím}\inf_n f_n,$$

todas pertenecen a $\mathcal{F}(X)$. Sino existe posibilidad alguna de confundir la notación, diremos que f es medible, en lugar de $\mathcal{M}$-medible.

Similar al caso real, una función $\varphi : X \to \mathbb{R}$ se llama $\mathcal{M}$-**simple**, o simplemente **simple**, si el rango de φ es un conjunto finito de números reales y si $\{x \in X : \varphi(x) = a\} \in \mathcal{M}$ siempre que $a \neq 0$. En particular, la función característica χ_E de cualquier conjunto $E \in \mathcal{M}$ es simple. Como antes, observe que si φ es simple, entonces ella es siempre se puede representar en la forma

$$\varphi \;=\; \sum_{j=1}^{n} a_j \cdot \chi_{E_j}$$

donde los conjuntos $E_1, \ldots, E_n \in \mathcal{M}$ son disjuntos dos a dos y los números reales $a_1, \ldots, a_n$ son distintos dos a dos. Por supuesto, toda función simple es medible. Denotaremos por $\mathcal{S}(X)$ *el conjuntos de todas las funciones* $\varphi : X \to \mathbb{R}$ *que son simples*. En un ejercicio sencillo establecer que $\mathcal{S}(X)$ es un espacio vectorial sobre $\mathbb{R}$. Además, $\mathcal{S}(X) \subseteq \mathcal{F}(X)$. Más aun, si $\varphi_1, \varphi_2 \in \mathcal{S}(X)$, entonces

$$\varphi_1 \cdot \varphi_2, \quad \text{máx}\{\varphi_1, \varphi_2\} \quad \text{y} \quad \text{mín}\{\varphi_1, \varphi_2\}$$

también pertenecen a $\mathcal{S}(X)$.

Teorema 11.1.33 (Teorema de Aproximación). *Sea $(X, \mathcal{M}, \lambda)$ un espacio de medida. Para cada **función medible** $f : X \to [0, +\infty]$, existe una **sucesión creciente** $(\varphi_n)_{n=1}^{\infty}$ de funciones simples no-negativas tal que $\varphi_n(x) \to f(x)$ para casi-todo $x \in X$.*

Puesto que la prueba es idéntica a la del Teorema 7.2.10, página 383, la omitiremos. De hecho, si $f : X \to \overline{\mathbb{R}}$ es medible, entonces existe una sucesión de funciones simples $(\psi_n)_{n=1}^{\infty}$ tal que

(a) $|\psi_n| \leq |f|$ para todo $n \geq 1$ y

(b) $\lim\limits_{n\to\infty} \psi_n(x) = f(x)$ para casi-todo $x \in X$.

Suponga que $(X, \mathcal{M})$ es un espacio medible y que $f, g : X \to \mathbb{R}$ son funciones medibles. Si $\Phi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ es una función continua, entonces la función $F_\Phi : X \to \mathbb{R}$ definida por

$$F_\Phi(x) \;=\; \Phi(f(x), g(x)), \quad x \in X$$

es medible. En efecto, observe en primer lugar que la función $\phi : X \to \mathbb{R} \times \mathbb{R}$ dada por $\phi(x) = (f(x), g(x))$ es medible. Para ver esto sólo hay que tener en cuenta que todo conjunto abierto de $\mathbb{R} \times \mathbb{R}$ es unión numerable de conjuntos de la forma $(a_1, b_1] \times (a_2, b_2]$ y que

$$\phi^{-1}\big((a_1, b_1] \times (a_2, b_2]\big) \;=\; f^{-1}((a_1, b_1]) \cap g^{-1}((a_2, b_2]) \in \mathcal{M}.$$

De esto se sigue que $F_\Phi = \Phi \circ \phi$ es medible.

Teorema 11.1.34 (Severini-Egoroff). *Sea $(X, \mathcal{M}, \lambda)$ un espacio de **medida finita** y sea $(f_n)_{n=1}^{\infty}$ una sucesión de funciones medibles a valores reales definidas sobre X. Si $(f_n)_{n=1}^{\infty}$ **converge puntualmente** a una función $f : X \to \mathbb{R}$, entonces para cada $\varepsilon > 0$, existe un conjunto $E_\varepsilon \in \mathcal{M}$ tal que*

$$\lambda(X \setminus E_\varepsilon) \;<\; \varepsilon \qquad \text{y} \qquad f_n \to f \;\; \textbf{uniformemente} \text{ sobre } E_\varepsilon.$$

La demostración, por ser enteramente similar a la del Teorema de Severini-Egoroff, Teorema 7.2.12, página 386, la omitiremos.

11.1.7. Funciones Integrables

Sea $(X, \mathcal{M}, \lambda)$ un espacio de medida y sea $\varphi : X \to [0, +\infty)$ una función simple. Definimos la **integral** de φ como

$$\int_X \varphi \, d\lambda \;=\; \sum_{j=1} a_j \cdot \lambda(E_j).$$

Diremos que φ es **integrable** si $\int_X \varphi \, d\lambda < +\infty$. Observe que φ es integrable es equivalente a exigir que $\lambda(E_j) < +\infty$ para $a_j > 0$. Por supuesto, estamos usando lo que convenimos anteriormente: $0 \cdot \infty = 0$.

Definición 11.1.35. *Sea $(X, \mathcal{M}, \lambda)$ un espacio de medida y sea $f : X \to [0, +\infty]$ una función medible. Definimos la **integral** de f por*

$$\int_X f \, d\lambda \;=\; \sup\left\{ \int_X \varphi \, d\lambda \;:\; 0 \le \varphi \le f, \;\; \varphi \in \mathcal{S}(X) \right\}.$$

*Diremos que f es λ-**integrable**, o simplemente **integrable**, si $\int_X f \, d\lambda < +\infty$. Además, si $E \in \mathcal{M}$, pondremos*

$$\int_E f \, d\lambda \;=\; \int_X f \cdot \chi_E \, d\lambda.$$

Se puede demostrar que si $f, g : X \to [0, +\infty]$ son funciones medibles, $E, F \in \mathcal{M}$ y $a \ge 0$, entonces

(a) $\displaystyle\int_X (f+g) \, d\lambda = \int_X f \, d\lambda + \int_X g \, d\lambda.$

(b) $\displaystyle\int_X a \cdot f \, d\lambda \;=\; a \int_X f \, d\lambda.$

(c) $\displaystyle\int_X f \, d\lambda \le \int_X g \, d\lambda$ si $f \le g$.

(d) $\displaystyle\int_E f \, d\lambda \le \int_F f \, d\lambda$ si $E \subseteq F$.

Teorema 11.1.36 (Convergencia Monótona). *Sea $(X, \mathcal{M}, \lambda)$ un espacio de medida y sea $(f_n)_{n=1}^{\infty}$ una **sucesión creciente** de funciones medibles no-negativas. Si $\lim_{n\to\infty} f_n(x) = f(x)$ para todo $x \in X$, entonces*

$$\int_X f \, d\lambda \;=\; \lim_{n\to\infty} \int_X f_n \, d\lambda.$$

Muchas consecuencias se derivan del Teorema de la Convergencia Monótona. Por ejemplo, los siguientes dos corolarios siguen de ese resultado.

Corolario 11.1.37 (Beppo Levi). *Sea $(X, \mathcal{M}, \lambda)$ un espacio de medida y suponga que $(f_n)_{n=1}^{\infty}$ es una sucesión no-negativa de funciones medibles. Entonces*

$$\int_X \left(\sum_{n=1}^{\infty} f_n \right) d\lambda \;=\; \sum_{n=1}^{\infty} \int_X f_n \, d\lambda.$$

Prueba. Aplique el Teorema de la Convergencia Monótona a la sucesión creciente $(g_n)_{n=1}^{\infty}$, donde $g_n = f_1 + \cdots + f_n$ para cada $n \geq 1$. ■

Corolario 11.1.38 (Fatou). *Sea $(X, \mathcal{M}, \lambda)$ un espacio de medida y suponga que $(f_n)_{n=1}^{\infty}$ es una sucesión no-negativa de funciones medibles. Entonces*

$$\int_X \big(\varliminf_{n\to\infty} f_n\big)\, d\lambda \;\leq\; \varliminf_{n\to\infty} \int_X f_n\, d\lambda.$$

Prueba. Basta con definir $g_n = \inf\{f_k : k \geq n\}$ y observar que como $g_n \leq g_{n+1}$ y $g_n \leq f_n$ para todo $n \in \mathbb{N}$, entonces

$$\begin{aligned} \int_X \big(\varliminf_{n\to\infty} f_n\big)\, d\lambda \;&=\; \int_X \big(\lim_{n\to\infty} g_n\big)\, d\lambda \;=\; \lim_{n\to\infty} \int_X g_n\, d\lambda \\ &=\; \varliminf_{n\to\infty} \int_X g_n\, d\lambda \;\leq\; \varliminf_{n\to\infty} \int_X f_n\, d\lambda, \end{aligned}$$

y termina la prueba. ■

Sea $(X, \mathcal{M}, \lambda)$ un espacio de medida y sea $f : X \to [0, +\infty]$ una función medible. Si definimos $\nu : \mathcal{M} \to [0, +\infty]$ por

$$\nu(E) \;=\; \int_E f\, d\lambda, \quad E \in \mathcal{M},$$

resulta que

(1) ν es una medida.

(2) ν es finita si, y sólo si, f es integrable.

(3) $\lambda(E) = 0 \;\Rightarrow\; \nu(E) = 0$. En este caso se dice que ν es **absolutamente continua** con respecto a λ, lo cual escribiremos como $\nu \ll \lambda$.

*En general, dado un espacio de medida $(X, \mathcal{M}, \lambda)$, una función medible $f : X \to \overline{\mathbb{R}}$ se dice λ-**integrable** si*

$$\int_X f^+\, d\lambda \;<\; +\infty \qquad \text{y} \qquad \int_X f^-\, d\lambda \;<\; +\infty$$

En este caso, se define

$$\int_X |f|\, d\lambda \;=\; \int_X f^+\, d\lambda \;+\; \int_X f^-\, d\lambda.$$

Si una de las integrales $\int_X f^+\, d\lambda$ o $\int_X f^-\, d\lambda$ es finita, definimos

$$\int_X f\, d\lambda \;=\; \int_X f^+\, d\lambda \;-\; \int_X f^-\, d\lambda$$

y si f es integrable, entonces se cumple la desigualdad

$$\left| \int_X f\, d\lambda \right| \;\leq\; \int_X |f|\, d\lambda.$$

Denotaremos por $\mathrm{L}_1(X, \mathcal{M}, \lambda)$ *el espacio vectorial sobre* $\mathbb{R}$ *de todas las clases de equivalencia funciones medibles* $f : X \to \overline{\mathbb{R}}$ *que son* λ-***integrables***. Con frecuencia, escribiremos simplemente $\mathrm{L}_1(X)$ en lugar de $\mathrm{L}_1(X, \mathcal{M}, \lambda)$. Si para cada $f \in \mathrm{L}_1(X, \mathcal{M}, \lambda)$ definimos

$$\| f \|_1 = \int_X |f| \, d\lambda,$$

entonces $(\mathrm{L}_1(X, \mathcal{M}, \lambda), \| \cdot \|_1)$ es un espacio de Banach.

Teorema 11.1.39 (Chebyshev). *Sea* $(X, \mathcal{M}, \lambda)$ *un espacio de medida. Para cualquier función* ***medible no-negativa*** $f : X \to \overline{\mathbb{R}}$ *y para cualquier* $c > 0$, *se cumple que*

$$\mu([f(x) > c]) \leq \frac{1}{c} \int_X f \, d\lambda.$$

Prueba. Sea $A = [f(x) > c]$. Puesto que $c \cdot \chi_A \leq f \cdot \chi_A$ y $f \cdot \chi_A \leq f$ tenemos que $c \cdot \chi_A \leq f$. El hecho de que la integral es monótona nos dice que

$$c \cdot \lambda(A) = c \cdot \int_X \chi_A \, d\lambda \leq \int_X f \, d\lambda.$$

Esto finaliza la prueba. ■

Otros resultados útiles y cuyas pruebas omitiremos por ser similares a las de la integral de Lebesgue, véase el Corolario 10.1.28, son los siguientes:

Teorema 11.1.40. *Sea* $(X, \mathcal{M}, \lambda)$ *un espacio de medida y sea* $f : X \to [0, +\infty]$ *una función medible.*

(a) Si $\int_X f \, d\lambda = 0$, *entonces* $f = 0$ *casi-siempre.*

(b) Si $f > 0$ *y* $\int_E f \, d\lambda = 0$, *entonces* $\lambda(E) = 0$.

(c) Si $\int_X f \, d\lambda < +\infty$, *entonces* $f(x) \in \mathbb{R}$ *casi-siempre.*

Teorema 11.1.41 (Convergencia Dominada). *Sea* $(X, \mathcal{M}, \lambda)$ *un espacio de medida y suponga* $(f_n)_{n=1}^{\infty}$ *es una sucesión de funciones medibles tal que* $f(x) = \lim_{n\to\infty} f_n(x)$ *para casi-todo* $x \in X$. *Si existe una función* λ-***integrable*** $g \geq 0$ *tal que* $|f_n| \leq g$ *para todo* $n \geq 1$, *entonces* f, *y todas las* f_n, *son* λ-*integrables y se cumple que*

$$\lim_{n\to\infty} \int_X f_n \, d\lambda = \int_X f \, d\lambda.$$

En particular, $\lim_{n\to\infty} \int_X |f_n - f| \, d\lambda = 0.$

Como antes, omitimos la demostración por ser muy similar al Teorema de la Convergencia Dominada de Lebesgue, Teorema 10.1.40, página 561.

Corolario 11.1.42. *Sea* $(X, \mathcal{M}, \lambda)$ *un espacio de medida y suponga* $(f_n)_{n=1}^{\infty}$ *es una sucesión de funciones medibles no-negativas tal que* $f(x) = \lim_{n\to\infty} f_n(x)$ *para todo* $x \in X$ *y*

$$\lim_{n\to\infty} \int_X f_n \, d\lambda \; = \; \int_X f \, d\lambda$$

con $\int_X f \, d\lambda < +\infty$*. Entonces, todas las* f_n *son* λ*-integrables y*

$$\lim_{n\to\infty} \int_X |f_n - f| \, d\lambda \; = \; 0.$$

Prueba. Para cada $n \in \mathbb{N}$, considere la función $g_n = f - f_n$. De nuestra hipótesis se tiene que

$$\lim_{n\to\infty} g_n \; = \; 0 \qquad \text{y} \qquad \lim_{n\to\infty} \int_X g_n \, d\lambda \; = \; 0.$$

De aquí se sigue que

$$\lim_{n\to\infty} g_n^+ \; = \; 0 \qquad \text{y} \qquad \lim_{n\to\infty} g_n^- \; = \; 0$$

y puesto que $0 \leq g_n^+ \leq f$ y $\int_X f \, d\lambda < +\infty$, entonces el Teorema de la Convergencia Dominada nos indica que

$$\lim_{n\to\infty} \int_X g_n^+ \, d\lambda \; = \; 0.$$

Por esto,

$$\lim_{n\to\infty} \int_X g_n^- \, d\lambda \; = \; \lim_{n\to\infty} \int_X (g_n^+ - g_n) \, d\lambda \; = \; \lim_{n\to\infty} \int_X g_n^+ \, d\lambda \; - \; \lim_{n\to\infty} \int_X g_n \, d\lambda \; = \; 0.$$

Finalmente,

$$\lim_{n\to\infty} \int_X |f_n - f| \, d\lambda \; = \; \lim_{n\to\infty} \int_X |g_n| \, d\lambda \; = \; \lim_{n\to\infty} \int_X g_n^+ \, d\lambda \; + \; \lim_{n\to\infty} \int_X g_n^- \, d\lambda \; = \; 0.$$

La prueba es completa. ■

11.2. Medida Producto y el Teorema de Fubini

11.2.1. Clases Monótonas

El objetivo de esta sección es mostrar un resultado fundamental debido a Paul R. Halmos el cual establece que la diferencia entre una σ-álgebra y una álgebra es la propiedad de monotonía que posee la primera, dicho de otro modo, toda álgebra que posea la propiedad de monotonía es una σ-álgebra. De este hecho se deriva otro resultado interesante que permite determinar con precisión cómo son las σ-álgebras generadas por álgebras. Éste resultado es una pieza clave en la obtención de σ-álgebras productos y, por consiguiente, en la definición de una integral doble en el sentido de Lebesgue.

Definición 11.2.1. *Sea X un conjunto no vacío y $\mathcal{M} \subseteq \mathscr{P}(X)$. Diremos que $\mathcal{M}$ es una **clase monótona**, o que posee la **propiedad de monotonía**, si ella satisface las siguientes dos condiciones:*

*(i) $\bigcup_{n=1}^{\infty} E_n \in \mathcal{M}$ para cualquier **sucesión creciente** $(E_n)_{n=1}^{\infty}$ en $\mathcal{M}$, y*

*(ii) $\bigcap_{n=1}^{\infty} F_n \in \mathcal{M}$ para cualquier **sucesión decreciente** $(F_n)_{n=1}^{\infty}$ en $\mathcal{M}$.*

Observe que $\mathscr{P}(X)$ es, trivialmente, una clase monótona. Similar a la demostración de la existencia de σ-álgebras minimales, el siguiente resultado establece cómo se genera, a partir de cualquier colección de conjuntos dada, una clase monótona minimal.

Lema 11.2.2. *Sea X un conjunto no vacío y sea $\mathcal{A}$ una colección de subconjuntos de X. Entonces existe una familia $\mathcal{C}_{\text{mon}}(\mathcal{A})$ de subconjuntos de X que posee las siguientes dos propiedades:*

*(a) $\mathcal{C}_{\text{mon}}(\mathcal{A})$ es una **clase monótona** conteniendo a $\mathcal{A}$.*

(b) Si $\mathcal{M}$ es otra clase monótona conteniendo a $\mathcal{A}$, entonces $\mathcal{C}_{\text{mon}}(\mathcal{A}) \subseteq \mathcal{M}$.

Prueba. Considere la colección

$$\mathfrak{P} = \{\mathcal{G} \subseteq \mathscr{P}(X) : \mathcal{G} \text{ es una clase monótona}, \ \mathcal{A} \subseteq \mathcal{G}\}.$$

Observe que $\mathfrak{P}$ es no vacía ya que $\mathscr{P}(X) \in \mathfrak{P}$. Es un ejercicio sencillo verificar que

$$\mathcal{C}_{\text{mon}}(\mathcal{A}) = \bigcap_{\mathcal{G} \in \mathfrak{P}} \mathcal{G}$$

es la clase monótona que posee las dos propiedades requeridas. ■

La clase monótona $\mathcal{C}_{\text{mon}}(\mathcal{A})$ obtenida en el lema anterior es, por construcción, la clase monótona más pequeña conteniendo a la colección $\mathcal{A}$ y se le suele llamar la **clase monótona generada** por $\mathcal{A}$. También es claro que *cualquier σ-álgebra es una clase monótona*, pero no recíprocamente. Sin embargo, como veremos a continuación, toda álgebra que es una clase monótona es una σ-álgebra. Recordemos que una colección no vacía $\mathcal{A}$ de subconjuntos de un conjunto X es dice que es una **álgebra** si ella satisface las siguientes condiciones:

(a_1) $X \in \mathcal{A}$.

(a_2) $A^c \in \mathcal{A}$ para cualquier $A \in \mathcal{A}$, y

(a_3) $\bigcup_{j=1}^{n} A_n \in \mathcal{A}$ para cualquier colección finita $\{A_1, \ldots, A_n\} \subseteq \mathcal{A}$.

Observe que si $\mathcal{A}$ es una álgebra y $A, B \in \mathcal{A}$, entonces $A \cap B \in \mathcal{A}$ y $A \setminus B \in \mathcal{A}$. En efecto,

$$A \cap B = (A^c \cup B^c)^c \in \mathcal{A} \qquad \text{y} \qquad A \setminus B = A \cap B^c \in \mathcal{A}.$$

El siguiente resultado revela cómo obtener una álgebra bajo mínimas condiciones.

Lema 11.2.3. *Sea $\mathcal{A}$ una familia de subconjuntos de un conjunto X con las siguientes propiedades:*

(a) $X \in \mathcal{A}$, y

(b) $A \setminus B \in \mathcal{A}$ cualesquiera sean $A, B \in \mathcal{A}$.

*Entonces $\mathcal{A}$ es una **álgebra**.*

Prueba. Sea $A \in \mathcal{A}$. Como $X \in \mathcal{A}$, se sigue entonces de (b) que $A^c = X \setminus A \in \mathcal{A}$ lo cual prueba la validez de (a_2). Para verificar que (a_3) se cumple, primero observe que si $A, B \in \mathcal{A}$, entonces $A \cap B = A \setminus B^c \in \mathcal{A}$ por la primera parte. Luego, $A \cup B = A^c \cap B^c \in \mathcal{A}$. Fin de la prueba. ■

Por ejemplo, del lema anterior se sigue que, si X es un espacio topológico y $\mathcal{A} = \mathcal{O} \cup \mathcal{F}$, entonces $\mathcal{A}$ es una álgebra, donde $\mathcal{O}$ y $\mathcal{F}$ son, respectivamente, las familias de todos los subconjuntos abiertos y cerrados de X.

Teorema 11.2.4 (Halmos). *Sea $\mathcal{A}$ una **álgebra** de subconjuntos de un conjunto X. Si $\mathcal{A}$ es una **clase monótona**, entonces $\mathcal{A}$ es una σ-**álgebra**.*

Prueba. Puesto que $\mathcal{A}$ es una álgebra, para demostrar que es una σ-álgebra sólo hace falta verificar que ella es cerrada bajo uniones numerables, es decir, suponga que $(A_n)_{n=1}^{\infty}$ es una sucesión arbitraria en $\mathcal{A}$ y veamos que $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$. Para cada $n \in \mathbb{N}$, considere el conjunto $F_n = \bigcup_{k=1}^{n} A_k$. Como $\mathcal{A}$ es una álgebra, resulta que $F_n \in \mathcal{A}$. Además, siendo $(F_n)_{n=1}^{\infty}$ una sucesión creciente en $\mathcal{A}$, se tiene que $\bigcup_{n=1}^{\infty} F_n \in \mathcal{A}$ gracias a que $\mathcal{A}$ es una clase monótona. Finalmente, puesto $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} F_n$ tenemos que $\bigcup_{n=1}^{\infty} E_n \in \mathcal{A}$ y termina la prueba. ■

Se sigue del resultado anterior que *si X es un **conjunto finito** y $\mathcal{A}$ es una álgebra sobre X, entonces $\mathcal{A}$ es una σ-álgebra*. Por lo tanto, la distinción entre las nociones de una álgebra y una σ-álgebra desaparece cuando X es finito.

Lo que resulta altamente agradable es el siguiente resultado debido, también, a P. Halmos.

Teorema 11.2.5 (Teorema de la Clase Monótona). *Sea X un conjunto no vacío y sea $\mathcal{A}$ una **álgebra** de subconjuntos de X. Entonces*

$$\mathcal{C}_{\text{mon}}(\mathcal{A}) \;=\; \sigma(\mathcal{A}).$$

Prueba. Puesto que cualquier σ-álgebra es una clase monótona y $\mathcal{A} \subseteq \sigma(\mathcal{A})$, resulta, por la definición de $\mathcal{C}_{\text{mon}}(\mathcal{A})$, que

$$\mathcal{A} \subseteq \mathcal{C}_{\text{mon}}(\mathcal{A}) \subseteq \sigma(\mathcal{A})$$

Para demostrar que $\sigma(\mathcal{A}) \subseteq \mathcal{C}_{\text{mon}}(\mathcal{A})$ es suficiente probar que $\mathcal{C}_{\text{mon}}(\mathcal{A})$ es una σ-álgebra. Sin embargo, en vista del Teorema 11.2.4, es suficiente verificar que $\mathcal{C}_{\text{mon}}(\mathcal{A})$ es una álgebra. Observe que para lograr tal objetivo sólo debemos comprobar, gracias al Lema 11.2.3, que

$$E \setminus F \in \mathcal{C}_{\text{mon}}(\mathcal{A}) \qquad \text{y} \qquad E \cup F \in \mathcal{C}_{\text{mon}}(\mathcal{A})$$

para todo $E, F \in \mathcal{C}_{\text{mon}}(\mathcal{A})$. En efecto, suponga que $E \setminus F \in \mathcal{C}_{\text{mon}}(\mathcal{A})$ para cualquier par de conjuntos $E, F \in \mathcal{C}_{\text{mon}}(\mathcal{A})$. Entonces, como $X \in \mathcal{A} \subseteq \mathcal{C}_{\text{mon}}(\mathcal{A})$, se tiene que $E^c = X \setminus E \in \mathcal{C}_{\text{mon}}(\mathcal{A})$, lo cual muestra que $\mathcal{C}_{\text{mon}}(\mathcal{A})$ es cerrado bajo uniones finitas y también tomando complementos.

Para cada $E \in \mathcal{C}_{\text{mon}}(\mathcal{A})$, considere la familia,

$$\mathcal{M}_E \;=\; \big\{F \in \mathcal{C}_{\text{mon}}(\mathcal{A}) : E \cup F \in \mathcal{C}_{\text{mon}}(\mathcal{A}),\; E \setminus F \in \mathcal{C}_{\text{mon}}(\mathcal{A}),\; F \setminus E \in \mathcal{C}_{\text{mon}}(\mathcal{A})\big\}.$$

Observe que cualesquiera sean $E, F \in \mathcal{C}_{\text{mon}}(\mathcal{A})$ se cumple que

$$F \in \mathcal{M}_E \quad \text{si, sólo si} \quad E \in \mathcal{M}_F. \tag{1}$$

Más aun, $\varnothing, E \in \mathcal{M}_E$ para cualquier $E \in \mathcal{C}_{\text{mon}}(\mathcal{A})$. Veamos ahora que:

(i) *$\mathcal{M}_E$ es una clase monótona para cualquier $E \in \mathcal{C}_{\rm mon}(\mathcal{A})$.* En efecto, sea $E \in \mathcal{C}_{\rm mon}(\mathcal{A})$ y suponga que $(F_n)_{n=1}^{\infty}$ una sucesión monótona en $\mathcal{M}_E$. Como siempre, escribiremos $\lim_{n\to\infty} F_n$ para el conjunto $\bigcup_{n=1}^{\infty} F_n$ o $\bigcap_{n=1}^{\infty} F_n$ según sea $(F_n)_{n=1}^{\infty}$ una sucesión creciente o decreciente respectivamente. Entonces

$$E \cup \big(\lim_{n\to\infty} F_n\big) = \lim_{n\to\infty} (E \cup F_n) \in \mathcal{M}_E,$$

$$E \cap \big(\lim_{n\to\infty} F_n\big)^c = E \cup \big(\lim_{n\to\infty} F_n^c\big) = \lim_{n\to\infty} (E \cup F_n^c) \in \mathcal{M}_E,$$

$$\big(\lim_{n\to\infty} F_n\big) \cap E^c = \lim_{n\to\infty} \big(F_n \cap E^c\big) \in \mathcal{M}_E.$$

Esto prueba nuestra afirmación.

(ii) *$\mathcal{C}_{\rm mon}(\mathcal{A}) \subseteq \mathcal{M}_E$ para cualquier $E \in \mathcal{A}$.* Fijemos un $E \in \mathcal{A}$. Si $F \in \mathcal{A}$, entonces, puesto que $\mathcal{A}$ es una álgebra, $E \cup F$, $E \setminus F$, $F \setminus E \in \mathcal{A} \subseteq \mathcal{C}_{\rm mon}(\mathcal{A})$. Esto muestra que $F \in \mathcal{M}_E$ y, por lo tanto, $\mathcal{A} \subseteq \mathcal{M}_E$. Se sigue de (i) que $\mathcal{A} \subseteq \mathcal{C}_{\rm mon}(\mathcal{A}) \subseteq \mathcal{M}_E$.

(iii) *$\mathcal{C}_{\rm mon}(\mathcal{A}) \subseteq \mathcal{M}_F$ para todo $F \in \mathcal{C}_{\rm mon}(\mathcal{A})$.* Sea $F \in \mathcal{C}_{\rm mon}(\mathcal{A})$. Por (ii), $F \in \mathcal{M}_E$ para cualquier $E \in \mathcal{A}$, lo cual implica, por (1), que $E \in \mathcal{M}_F$ para cualquier $E \in \mathcal{A}$, es decir, $\mathcal{A} \subseteq \mathcal{M}_F$. De nuevo, haciendo uso de (i) y la definición de $\mathcal{C}_{\rm mon}(\mathcal{A})$, se concluye $\mathcal{C}_{\rm mon}(\mathcal{A}) \subseteq \mathcal{M}_F$ para cualquier $F \in \mathcal{C}_{\rm mon}(\mathcal{A})$.

(iv) *$\mathcal{C}_{\rm mon}(\mathcal{A})$ es una álgebra.* Sean $E, F \in \mathcal{C}_{\rm mon}(\mathcal{A})$. Por (iii) $E \in \mathcal{M}_F$ y, en consecuencia, $E \cup F$, $E \setminus F$, $F \setminus E \in \mathcal{C}_{\rm mon}(\mathcal{A})$. Esto prueba que $\mathcal{C}_{\rm mon}(\mathcal{A})$ es una álgebra y termina la prueba. ■

11.2.2. Medida Producto y el Teorema de Fubini

En esta sección, siempre supondremos que $(X, \mathcal{M}, \lambda)$ y $(Y, \mathcal{N}, \nu)$ son espacios de medidas. Lo que queremos en esta parte es definir un producto entre dichos espacios, al que llamaremos espacio de medida producto y que denotaremos por

$$(X \times Y,\ \mathcal{M} \otimes \mathcal{N},\ \lambda \otimes \mu),$$

donde $\mathcal{M} \otimes \mathcal{N}$ es una σ-álgebra muy particular de subconjuntos de $X \times Y$ y $\lambda \otimes \mu$ es una medida sobre $\mathcal{M} \otimes \mathcal{N}$ para la cual se requiere que cumpla la condición

$$(\lambda \otimes \mu)(E \times F) = \lambda(E) \cdot \nu(F)$$

cualesquiera sean $E \in \mathcal{M}$ y $F \in \mathcal{N}$. Este enfoque generaliza, en primer lugar, la noción geométrica usual del área de un rectángulo. En segundo lugar, ello permitirá definir una integral respecto al espacio de medida ya establecido de modo que la igualdad

$$\int_{X\times Y} f\, d(\lambda \otimes \mu) = \int_X \bigg(\int_Y f\, d\nu\bigg)\, d\lambda = \int_Y \bigg(\int_X f\, d\lambda\bigg)\, d\nu$$

se cumpla para una clase razonablemente amplia de funciones medibles f definidas sobre $X \times Y$. Existen varias estrategias para alcanzar los objetivos delineados anteriormente. Una de ellas es la que a continuación se muestra.

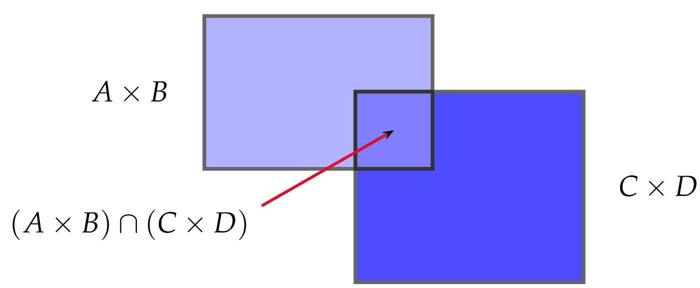
Definición 11.2.6. *Un subconjunto $R \subseteq X \times Y$ se dice que es un* ***rectángulo medible*** *si existen $A \in \mathcal{M}$ y $B \in \mathcal{N}$ tal que $R = A \times B$.*

El conjunto de todos los rectángulos medibles en $X \times Y$ será denotado por

$$\mathcal{M} \times \mathcal{N} = \{A \times B : A \in \mathcal{M}, B \in \mathcal{N}\}.$$

Observe que el producto cartesiano $\mathcal{M} \times \mathcal{N}$ de todos los rectángulos medibles no es una σ-álgebra ya que, por ejemplo, el complemento de un rectángulo, no es, en general, un rectángulo: en efecto, $[0,1] \times [0,1] \in \mathfrak{M}_\mu(\mathbb{R}) \times \mathfrak{M}_\mu(\mathbb{R})$, pero su complemento no es un rectángulo en $\mathbb{R} \times \mathbb{R}$. Esto nos muestra que *el producto cartesiano de dos* σ*-álgebras no es necesariamente una* σ*-álgebra*. Sin embargo, la intersección de dos rectángulos medibles sigue siendo un rectángulo medible ya que

$$(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D). \tag{1}$$



Por otro lado, el complemento de un rectángulo medible siempre se puede representar, aunque no de manera única, como una unión finita y disjunta de rectángulos medibles. Por ejemplo:

$$(A \times B)^c = (A^c \times B) \cup (A \times B^c) \cup (A^c \times B^c).$$

Teorema 11.2.7. *Sea* $\mathcal{A}_0$ *la colección de todas las* ***uniones finitas*** *y* ***disjuntas*** *de* ***rectángulos medibles*** *de* $\mathcal{M} \times \mathcal{N}$*. Entonces* $\mathcal{A}_0$ *es una álgebra.*

Prueba. Claramente $X \times Y \in \mathcal{A}_0$. Sean

$$E = \bigcup_{i=1}^{n} (A_i \times B_i) \qquad \text{y} \qquad F = \bigcup_{j=1}^{m} (C_j \times D_j)$$

elementos de $\mathcal{A}_0$. Entonces, por (1), se tiene que

$$E \cap F = \bigcup_{i=1}^{n} \bigcup_{j=1}^{m} \big[(A_i \cap C_j) \times (B_i \cap D_j)\big]$$

y, en consecuencia, $E \cap F \in \mathcal{A}_0$. Similarmente, $E^c = \bigcap_{i=1}^{n} (A_i \times B_i)^c \in \mathcal{A}_0$. De aquí se deduce que $\mathcal{A}_0$ es una álgebra. ■

Una aplicación del Teorema de la Clase Monótona, Teorema 11.2.5, en combinación con el resultado anterior nos revela que:

Corolario 11.2.8. *Si* $\mathcal{A}_0$ *es la colección de todas las* ***uniones finitas*** *y* ***disjuntas de rectángulos medibles*** *de* $\mathcal{M} \times \mathcal{N}$*, entonces*

$$\mathcal{C}_{\text{mon}}(\mathcal{A}_0) = \sigma(\mathcal{A}_0).$$

Definición 11.2.9. *Sean $(X, \mathcal{M}, \lambda)$ y $(Y, \mathcal{N}, \nu)$ espacios de medidas. La σ-álgebra generada por los rectángulos medibles, que denotaremos por*

$$\mathcal{M} \otimes \mathcal{N} = \sigma(\mathcal{M} \times \mathcal{N}),$$

se llama la ***σ-álgebra producto*** *de $\mathcal{M}$ y $\mathcal{N}$.*

Puesto que $\mathcal{M} \times \mathcal{N} \subseteq \mathcal{A}_0 \subseteq \mathcal{M} \otimes \mathcal{N}$, resulta del corolario anterior que

$$\mathcal{M} \otimes \mathcal{N} = \sigma(\mathcal{A}_0) = \mathcal{C}_{\text{mon}}(\mathcal{A}_0).$$

Definición 11.2.10. *Sea $E \subseteq X \times Y$. Para cada $x \in X$ y cada $y \in Y$, los conjuntos*

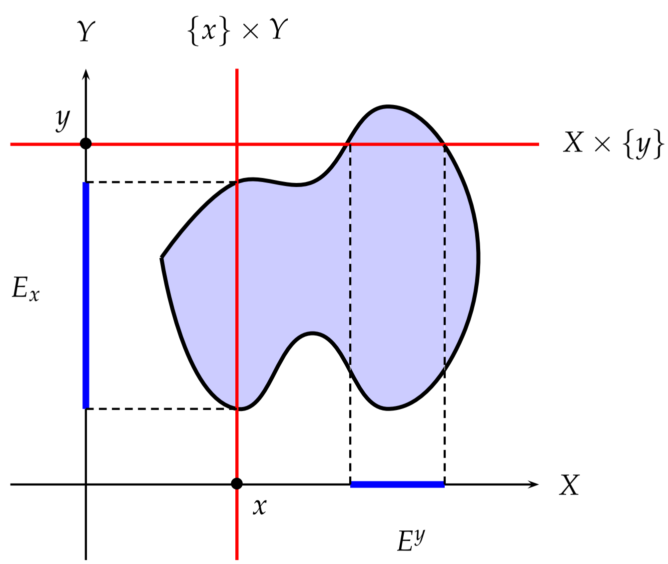
$$E_x = \{y \in Y : (x, y) \in E\} \qquad y \qquad E^y = \{x \in X : (x, y) \in E\}$$

son llamados, respectivamente, la ***x-sección*** *y la* ***y-sección*** *de E.*

Nótese que E_x no es otra cosa que el resultado de intersectar a E con la línea vertical $\{x\} \times Y$ y luego "proyectar" esa intersección sobre Y:

$$E_x = \text{pr}_Y(E \cap (\{x\} \times Y)).$$

Otra forma alternativa de describir E_x es considerar la aplicación $i_x : Y \to X \times Y$ dada por $i_x(y) = (x, y)$ y notar que $E_x = i_x^{-1}(E)$.



Similarmente, E^y es la proyección sobre el eje X del conjunto $E \cap (X \times \{y\})$:

$$E^y = \text{pr}_X(E \cap (X \times \{y\})) = (i^y)^{-1}(E),$$

donde $i^y : X \to X \times Y$ se define por $i^y(x) = (x, y)$. Por ejemplo, si $E = A \times B$, entonces

$$E_x = (A \times B)_x = \begin{cases} B & \text{si } x \in A \\ \varnothing & \text{si } x \notin A. \end{cases}$$

En efecto,

$$(A\times B)_x \;=\; \mathrm{pr}_Y\big((A\times B)\cap(\{x\}\times\mathbb{R})\big) \;=\; \mathrm{pr}_Y\big((A\cap\{x\})\times B\big)$$

y como $A\cap\{x\}=\{x\}$ o $\varnothing$ dependiendo si $x\in A$ o $x\notin A$, se obtiene el resultado. Con un argumento análogo se obtiene que

$$E^y \;=\; (A\times B)^y \;=\; \begin{cases} A & \text{si } \; y\in B \\ \varnothing & \text{si } \; y\notin B. \end{cases}$$

Observe que si $(E_n)_{n=1}^{\infty}$ es una sucesión de subconjuntos de $\mathbb{R}\times\mathbb{R}$, entonces

$$\Big(\bigcup_{n=1}^{\infty}E_n\Big)_x \;=\; \bigcup_{n=1}^{\infty}(E_n)_x \qquad\text{y}\qquad \Big(\bigcap_{n=1}^{\infty}E_n\Big)_x \;=\; \bigcap_{n=1}^{\infty}(E_n)_x$$

para todo $x\in\mathbb{R}$. En efecto,

$$\begin{aligned} y\in\Big(\bigcup_{n=1}^{\infty}E_n\Big)_x \;&\Leftrightarrow\; (x,y)\in\bigcup_{n=1}^{\infty}E_n \\ &\Leftrightarrow\; (x,y)\in E_m \;\text{ para algún }\; m\in\mathbb{N} \\ &\Leftrightarrow\; y\in(E_m)_x \;\text{ para algún }\; m\in\mathbb{N} \\ &\Leftrightarrow\; y\in\bigcup_{n=1}^{\infty}(E_n)_x \end{aligned}$$

y una prueba similar trabaja para la intersección. Igualdades semejantes se cumplen para la y-sección.

El siguiente resultado establece que todas las secciones de un conjunto medible son medibles. Sin embargo, puede ocurrir que todas las secciones de un conjunto sean medibles sin que dicho conjunto sea medible.

Teorema 11.2.11. *Sea $E\in\mathcal{M}\otimes\mathcal{N}$. Entonces:*

(a) $E_x\in\mathcal{N}$ para todo $x\in X$ y

(b) $E^y\in\mathcal{M}$ para todo $y\in Y$.

Prueba. Es suficiente dar la prueba para el caso (a); es decir, para las x-secciones. Considere la colección

$$\mathfrak{B} \;=\; \big\{E\subseteq X\times Y \,:\, E_x\in\mathcal{N} \text{ para todo } x\in X\big\}.$$

Nuestra tarea consistirá en demostrar que $\mathfrak{B}$ es una σ-álgebra conteniendo a los rectángulos medibles. De allí que, al ser $\mathcal{M}\otimes\mathcal{N}$ la σ-álgebra más pequeña conteniendo a los rectángulos medibles, tendremos que

$$\mathcal{M}\otimes\mathcal{N} \;\subseteq\; \mathfrak{B}$$

y entonces quedará demostrado que cualesquiera sea el conjunto $E \in \mathcal{M} \otimes \mathcal{N}$, se tendrá que $E_x \in \mathcal{N}$ para todo $x \in X$. Comencemos. Para cualquier rectángulo medible $A \times B$ y cualquier $x \in X$, se tiene que

$$(A \times B)_x = \begin{cases} B & \text{si } \; x \in A \\ \varnothing & \text{si } \; x \notin A. \end{cases}$$

Esto muestra que $A \times B \in \mathfrak{B}$ y, por consiguiente, $\mathfrak{B}$ contiene a todos los rectángulos medibles. Veamos ahora que $\mathfrak{B}$ es una σ-álgebra.

(i) Claramente $X \times Y \in \mathfrak{B}$.

(ii) Sea $E \in \mathfrak{B}$. Entonces $E_x \in \mathcal{N}$ y como $\mathcal{N}$ es una σ-álgebra, resulta que $(E_x)^c \in \mathcal{N}$. Finalmente, puesto que $(E^c)_x = (E_x)^c$ se tiene que $E^c \in \mathfrak{B}$.

(iii) Sea $(E_n)_{n=1}^{\infty}$ una sucesión en $\mathfrak{B}$. Entonces, como vimos anteriormente, para todo $x \in X$,

$$\Big(\bigcup_{n=1}^{\infty} E_n\Big)_x = \bigcup_{n=1}^{\infty} (E_n)_x \in \mathcal{N}.$$

Esto prueba que $\bigcup_{n=1}^{\infty} E_n \in \mathfrak{B}$ y termina la demostración. ■

Dos consecuencias se derivan inmediatamente del resultado anterior:

(1) *Si A y B son conjuntos no vacíos tal que $A \times B \in \mathcal{M} \otimes \mathcal{N}$, entonces $A \in \mathcal{M}$ y $B \in \mathcal{N}$.* En efecto, seleccionando cualquier $x \in A$ y cualquier $y \in B$ se obtiene, por el resultado precedente, que

$$(A \times B)_x = B \in \mathcal{N} \qquad \text{y} \qquad (A \times B)^y = A \in \mathcal{M}.$$

Por supuesto, si $A = \varnothing$, entonces $\varnothing \times B = \varnothing \in \mathcal{M} \otimes \mathcal{N}$ sin importar si B es, o no, medible. Similarmente, $A \times \varnothing = \varnothing \in \mathcal{M} \otimes \mathcal{N}$ para cualquier conjunto $A \subseteq X$.

(2) *Si $V \subseteq \mathbb{R}$ es un conjunto* ***no-medible****, entonces $V \times V$ es* ***no-medible*** *en $\mathbb{R}^2$ y, por consiguiente,*

$$\mathfrak{M}_\mu(\mathbb{R}) \otimes \mathfrak{M}_\mu(\mathbb{R}) \subsetneqq \mathscr{P}(\mathbb{R} \times \mathbb{R}).$$

Definición 11.2.12. *Una función $f : X \times Y \to \overline{\mathbb{R}}$ se dice $\mathcal{M} \otimes \mathcal{N}$-**medible** si*

$$f^{-1}\big((a, +\infty]\big) = \big\{(x, y) \in X \times Y \,:\, f(x, y) > a\big\} \in \mathcal{M} \otimes \mathcal{N}$$

para cualquier $a \in \mathbb{R}$.

Como antes, escribiremos f es *medible*, en lugar de $\mathcal{M} \otimes \mathcal{N}$-medible. Similar al caso de funciones de una sola variable, se tiene que f es medible si, y sólo si:

(1) $\big\{(x, y) \in X \times Y \,:\, f(x, y) \geq a\big\} \in \mathcal{M} \otimes \mathcal{N}$ para cada $a \in \mathbb{R}$.

(2) $\big\{(x, y) \in X \times Y \,:\, f(x, y) < a\big\} \in \mathcal{M} \otimes \mathcal{N}$ para cada $a \in \mathbb{R}$.

(3) $\big\{(x, y) \in X \times Y \,:\, f(x, y) \leq a\big\} \in \mathcal{M} \otimes \mathcal{N}$ para cada $a \in \mathbb{R}$.

También, si $f, g : X \times Y \to \overline{\mathbb{R}}$ son medibles, entonces:

(4) $\alpha f + \beta g$, $f \cdot g$ son medibles para todo $\alpha, \beta \in \mathbb{R}$.

(5) $\lim\limits_{n \to \infty} h_n = h$ es medible siempre que $(h_n)_{n=1}^{\infty}$ sea una sucesión convergente de funciones medibles.

Definición 11.2.13. *Sea $f : X \times Y \to \overline{\mathbb{R}}$ una función. Para cada $x \in X$, la x-**sección** de f es la función $f_x : Y \to \overline{\mathbb{R}}$ definida por*

$$f_x(y) = f(x,y) \quad \text{para todo } y \in Y.$$

*Similarmente, para cada $y \in Y$, la y-**sección** de f es la función $f^y : X \to \overline{\mathbb{R}}$ definida por*

$$f^y(x) = f(x,y) \quad \text{para todo } x \in X.$$

Teorema 11.2.14. *Si $f : X \times Y \to \overline{\mathbb{R}}$ es una **función medible**, entonces*

(a) f_x es $\mathcal{N}$-medible para todo $x \in X$, y

(b) f^y es $\mathcal{M}$-medible para todo $y \in Y$.

Prueba. Fijemos $a \in \mathbb{R}$ y sea $E = \{(x,y) \in X \times Y : f(x,y) > a\}$. Como f es medible, resulta que E es medible y entonces, para cada $x \in X$, el conjunto

$$E_x = \{y \in Y : (x,y) \in E\} = \{y \in Y : f(x,y) > a\} = \{y \in Y : f_x(y) > a\}$$

es medible en Y gracias al Teorema 11.2.14. Eso prueba que f_x es medible. Un argumento enteramente similar trabaja para f^y. ■

En vista del resultado anterior uno puede preguntarse, por ejemplo, lo siguiente: suponga que $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *es una función que es continua en cada variable separadamente*, es decir, para cada $x \in X$, f_x es continua sobre Y y para cada $y \in Y$, f^y es continua sobre X: ¿es necesariamente f medible? Este problema fue analizado por primera vez Henry Lebesgue quien demostró que la respuesta es afirmativa. Sin embrago, en dicha conclusión no se puede cambiar "medibilidad" por "continuidad". El ejemplo clásico de una tal función es:

$$f(x,y) = \begin{cases} \dfrac{2xy}{x^2+y^2} & \text{si} \quad (x,y) \neq (0,0) \\ 0 & \text{si} \quad (x,y) = (0,0) \end{cases}$$

la cual es continua en cada variable separadamente, pero deja de ser continua en $(0,0)$.

Teorema 11.2.15 (Lebesgue). *Sea $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ una función la cual es **continua en cada variable separadamente**. Entonces f es **medible**.*

Prueba. Para cada $n \in \mathbb{N}$, considere el conjunto

$$E_n = \big\{(k/n, y) : k \in \mathbb{Z},\, y \in \mathbb{R}\big\}$$

y defina la función $f_n : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ declarando que ella sea igual a f sobre E_n, y para cada (x,y) tal que $k/n < x < (k+1)/n$ para algún $k \in \mathbb{Z}$, $f_n(x,y)$ es un punto en el segmento lineal que une a $f(k/n, y)$ con $f((k+1)/n, y)$. Con esta definición, resulta que cada f_n es continua sobre $\mathbb{R} \times \mathbb{R}$ ya que f es continua en y para cada $x \in \mathbb{R}$. Por otro lado, puesto que f es continua en x para cada $y \in \mathbb{R}$, se tiene que

$$f(x,y) = \lim_{n\to\infty} f_n(x,y).$$

De esta forma f es medible por ser el límite de una sucesión de funciones medibles. ■

El siguiente resultado es una pieza fundamental para definir una medida sobre la σ-álgebra producto $\mathcal{M} \otimes \mathcal{N}$.

Teorema 11.2.16. *Sean $(X, \mathcal{M}, \nu)$ y $(Y, \mathcal{N}, \lambda)$ espacios de **medidas σ-finitas** y sea $f : X \times Y \to \overline{\mathbb{R}}$ una **función medible**. Entonces, para cualquier $E \in \mathcal{M} \otimes \mathcal{N}$, se cumple que:*

(a) la función $h : X \to [0, +\infty]$ dada por $h(x) = \nu(E_x)$ es $\mathcal{M}$-medible,

(b) la función $g : Y \to [0, +\infty]$ dada por $g(y) = \lambda(E^y)$ es $\mathcal{N}$-medible,

(c) $\displaystyle\int_Y \nu(E_x)\, d\lambda(x) = \int_X \lambda(E^y)\, d\nu(y)$.

Prueba. Sea $\mathcal{F}$ la familia de todos los conjuntos $E \in \mathcal{M} \otimes \mathcal{N}$ para los cuales (a), (b) y (c) se cumplen. Si logramos demostrar que $\mathcal{F}$ es una clase monótona conteniendo a $\mathcal{A}_0$, la colección de todas las uniones finitas y disjuntas de rectángulos medibles, entonces, por el Teorema 11.2.8, tendremos que

$$\mathcal{F} = \sigma(\mathcal{A}_0) = \mathcal{M} \otimes \mathcal{N}$$

y el teorema quedará demostrado. Suponga que $E = A \times B$ es un rectángulo medible. Puesto que

$$(A \times B)_x = \begin{cases} B & \text{si } x \in A \\ \varnothing & \text{si } x \notin A. \end{cases} \qquad \text{y} \qquad (A \times B)^y = \begin{cases} B & \text{si } y \in B \\ \varnothing & \text{si } y \notin B. \end{cases}$$

resulta que $\lambda(E_x) = \lambda(B) \cdot \chi_A(x)$ y $\nu(E^y) = \nu(A) \cdot \chi_B(y)$, de donde se tiene que

$$\begin{aligned} \int_X \nu(E_x)\, d\lambda(x) &= \int_X \nu(B) \cdot \chi_A\, d\lambda = \nu(B) \cdot \lambda(A) \qquad (11.2.1) \\ &= \int_Y \lambda(A) \cdot \chi_B\, d\nu = \int_Y \lambda(E^y)\, d\nu(y). \end{aligned}$$

Así, para este E, la igualdad (c) se cumple y, en consecuencia, $E \in \mathcal{F}$. Sea $\{E_1, \ldots, E_k\}$ una colección finita y disjunta de rectángulos medibles. Puesto que

$$\left(\bigcup_{n=1}^{k} E_n\right)_x = \bigcup_{n=1}^{k} (E_n)_x \qquad \text{y} \qquad \left(\bigcup_{n=1}^{k} E_n\right)^y = \bigcup_{n=1}^{k} (E_n)^y$$

para todo $x, y \in \mathbb{R}$, entonces la σ-aditividad de μ nos garantiza que

$$h(x) = \nu\left(\left(\bigcup_{n=1}^{k} E_n\right)_x\right) = \nu\left(\bigcup_{n=1}^{k} (E_n)_x\right) = \sum_{n=1}^{k} \nu\big((E_n)_x\big)$$

lo cual muestra que h es medible por ser suma de funciones medibles. Una conclusión similar se obtiene para la función g que corresponde a la y-sección. Usando la propiedad de linealidad de la integral y lo demostrado en la primera parte se obtiene finalmente que

$$\begin{aligned} \int_X \nu\left(\left(\bigcup_{n=1}^{k} E_n\right)_x\right) d\lambda(x) &= \int_X \sum_{n=1}^{k} \nu\big((E_n)_x\big)\, d\lambda(x) = \sum_{n=1}^{k} \int_X \nu\big((E_n)_x\big)\, d\lambda(x) \\ &= \sum_{n=1}^{k} \int_Y \lambda\big((E_n)^y\big)\, d\nu(y) = \int_Y \lambda\left(\left(\bigcup_{n=1}^{k} E_n\right)^y\right) d\nu(y). \end{aligned}$$

Esto prueba que $\bigcup_{n=1}^{k} E_n \in \mathcal{F}$ y, por lo tanto, $\mathcal{A}_0 \subseteq \mathcal{F}$. Suponga que $(E_n)_{n=1}^{\infty}$ es una sucesión creciente en $\mathcal{F}$ y que $E = \bigcup_{n=1}^{\infty} E_n$. Se sigue de la continuidad de ν, Teorema 11.1.5, página 683, que

$$h(x) \;=\; \nu(E_x) \;=\; \lim_{n\to\infty} \nu\big((E_n)_x\big),$$

para todo $x \in \mathbb{R}$, lo cual nos indica que h, por ser límite de una sucesión de funciones medibles, es una función medible. El Teorema de la Convergencia Monótona, Teorema 11.1.36, nos garantiza entonces que

$$\begin{aligned}\int_X \nu(E_x)\, d\lambda(x) \;&=\; \lim_{n\to\infty} \int_X \nu\big((E_n)_x\big)\, d\lambda(x)\\ &=\; \lim_{n\to\infty} \int_Y \lambda\big((E_n)^y\big)\, d\nu(y)\\ &=\; \int_Y \lambda(E^y)\, d\nu(y),\end{aligned}$$

de donde se concluye que $\bigcup_{n=1}^{\infty} E_n \in \mathcal{F}$.

Para terminar la demostración debemos probar que si $(F_n)_{n=1}^{\infty}$ es una sucesión decreciente en $\mathcal{F}$, entonces $F = \bigcap_{n=1}^{\infty} F_n$ pertenece a $\mathcal{F}$. Suponga, en primer lugar, que $(F_n)_{n=1}^{\infty}$ es una sucesión decreciente en $\mathcal{F}$ y que $F_1 \subseteq A \times B$, donde $A \in \mathcal{M}$ y $B \in \mathcal{N}$ son ambos de medida finita. Para cada $x \in X$ se tiene que $(F_1)_x \subseteq B$, y así, $\lambda\big((F_1)_x\big) < +\infty$. De nuevo, por el Teorema 11.1.5, tenemos que

$$\nu(F_x) \;=\; \lim_{n\to\infty} \nu\big((F_n)_x\big)$$

para cualquier $x \in X$. Esto nos revela que la aplicación $x \to \lambda(F_x)$ es medible por ser el límite de una sucesión de funciones medibles. Puesto que

$$\int_X \nu\big((F_1)_x\big)\, d\lambda(x) \;\leq\; \int_X \nu(B)\cdot \chi_A\, d\lambda \;<\; +\infty$$

y

$$\int_Y \lambda\big((F_1)^y\big)\, d\nu(y) \;\leq\; \int_Y \lambda(A)\cdot \chi_B\, d\nu \;<\; +\infty,$$

el Teorema de la Convergencia Dominada implica que

$$\begin{aligned}\int_X \nu(F_x)\, d\lambda(x) \;&=\; \lim_{n\to\infty} \int_X \nu\big((F_n)_x\big)\, d\lambda(x)\\ &=\; \lim_{n\to\infty} \int_Y \lambda\big((F_n)^y\big)\, d\nu(y)\\ &=\; \int_Y \lambda(F^y)\, d\nu(y),\end{aligned}$$

de donde se sigue que $F \in \mathcal{F}$. Lo que acabamos de demostrar nos permite considerar la colección no vacía

$$\mathcal{E} \;=\; \big\{E \in \mathcal{M}\otimes\mathcal{N} : \, E\cap(A_k \times B_k) \in \mathcal{F} \text{ para todo } k \in \mathbb{N}\big\},$$

donde hemos usado el hecho de λ y ν son ambas σ-finitas para hallar sucesiones crecientes $(A_n)_{n=1}^{\infty}$ en $\mathcal{M}$ y $(B_n)_{n=1}^{\infty}$ en $\mathcal{N}$ tales que $\mu(A_k) < \infty$, $\mu(B_k) < \infty$ para todo $k \in \mathbb{N}$ y

$$\bigcup_{n=1}^{\infty} A_n \;=\; X \qquad \text{y} \qquad \bigcup_{n=1}^{\infty} B_n \;=\; Y.$$

Puesto que la familia $\mathcal{A}_0 \subseteq \mathcal{F}$, resulta que $\mathcal{A}_0 \subseteq \mathcal{E}$. Si $(E_n)_{n=1}^{\infty}$ es una sucesión creciente en $\mathcal{E}$, entonces

$$\left(\bigcup_{n=1}^{\infty} E_n\right) \cap (A_k \times B_k) \;=\; \bigcup_{n=1}^{\infty} \big(E_n \cap (A_k \times B_k)\big) \;\in\; \mathcal{F}$$

ya que $\mathcal{F}$ es cerrada bajo la formación de uniones de sucesiones crecientes. Esto prueba que $\bigcup_{n=1}^{\infty} E_n \in \mathcal{E}$. Similarmente, si $(E_n)_{n=1}^{\infty}$ es una sucesión decreciente en $\mathcal{E}$, entonces tal como se demostró en el parágrafo anterior, se tiene que

$$\left(\bigcap_{n=1}^{\infty} E_n\right) \cap (A_k \times B_k) \;=\; \bigcap_{n=1}^{\infty} \big(E_n \cap (A_k \times B_k)\big) \;\in\; \mathcal{F}$$

por lo que, $\bigcap_{n=1}^{\infty} E_n \in \mathcal{E}$. Se sigue del Teorema 11.2.8 que

$$\mathcal{E} \;=\; \mathcal{M} \otimes \mathcal{N}.$$

Consideremos de nuevo una sucesión decreciente $(F_n)_{n=1}^{\infty}$ en $\mathcal{F}$ y sea $F = \bigcap_{n=1}^{\infty} F_n$. Veamos que $F \in \mathcal{F}$. En efecto, puesto que $F \in \mathcal{E}$, se sigue de la definición de $\mathcal{E}$ que $F \cap (A_k \times B_k) \in \mathcal{F}$ para todo $k \in \mathbb{N}$ y entonces

$$F \;=\; \bigcup_{k=1}^{\infty} \big(F \cap (A_k \times B_k)\big) \;\in\; \mathcal{F}.$$

Con esto se concluye que $\mathcal{F}$ es una clase monótona conteniendo a $\mathcal{A}_0$ y, por lo tanto, $\mathcal{F} = \mathcal{M} \otimes \mathcal{N}$. Fin de la prueba. ■

Nota Adicional 11.2.5 El hecho de que ambas medidas en el Teorema 11.2.16 sean σ-finitas es una condición necesaria para la validez de (c). En efecto, tome $X = Y = [0,1]$ y considere $\lambda = \mu$ la medida de Lebesgue sobre X y sea ν la medida de conteo en $[0,1]$. Si $E = \{(x,x) : x \in [0,1]\}$, entonces para todo $x, y \in [0,1]$ resulta que

$$E_x \;=\; \{x\} \qquad \text{y} \qquad E^y \;=\; \{y\},$$

de donde se sigue que $\nu(E_x) = 1$ y $\lambda(E^y) = 0$. Por esto,

$$1 \;=\; \int_X \nu(E_x)\, d\mu \;\neq\; \int_Y \lambda(E^y)\, d\lambda \;=\; 0.$$

Por supuesto, la igualdad en (c) falla debido a que ν no es σ-finita.

Ya tenemos todos los ingredientes necesarios para definir una medida sobre la σ-álgebra producto $\mathcal{M} \otimes \mathcal{N}$.

Teorema 11.2.17 (Principio de Cavalieri). *La función de conjuntos* $\lambda \otimes \nu : \mathcal{M} \otimes \mathcal{N} \to [0, +\infty]$ *dada por*

$$(\lambda \otimes \nu)(E) \;=\; \int_X \nu(E_x)\, d\lambda(x) \;=\; \int_Y \lambda(E^y)\, d\nu(y)$$

es una ***medida*** *sobre* $\mathcal{M} \otimes \mathcal{N}$. *Más aun,*

$$(\lambda \otimes \nu)(A \times B) \;=\; \lambda(A) \cdot \nu(B)$$

para todo rectángulo medible $A \times B$. *Además,* $\lambda \otimes \nu$ *es* σ*-finita.*

Prueba. Claramente $\lambda\otimes\nu \geq 0$ y $(\lambda\otimes\nu)(\varnothing) = 0$. Para ver que $\lambda\otimes\nu$ es numerablemente aditiva, considere cualquier sucesión disjunta $(E_n)_{n=1}^{\infty}$ de conjuntos en $\mathcal{M}\otimes\mathcal{N}$. Entonces

$$\begin{aligned}
(\lambda\otimes\nu)\Big(\bigcup_{n=1}^{\infty} E_n\Big) &= \int_X \nu\bigg(\Big(\bigcup_{n=1}^{\infty} E_n\Big)_x\bigg)\,d\lambda(x) \\
&= \int_X \sum_{n=1}^{\infty} \nu\big((E_n)_x\big)\,d\lambda(x) \\
&= \sum_{n=1}^{\infty} \int_X \nu\big((E_n)_x\big)\,d\lambda(x) \\
&= \sum_{n=1}^{\infty} (\lambda\otimes\nu)(E_n)
\end{aligned}$$

Con esto queda demostrado que $\lambda\otimes\nu$ es una medida sobre $\mathcal{M}\otimes\mathcal{N}$. El hecho de que

$$(\lambda\otimes\nu)(A\times B) = \lambda(A)\cdot\nu(B) \tag{1}$$

para todo rectángulo medible $A\times B$, sigue de (11.2.1) en la prueba del Teorema 11.2.16. Finalmente, como $X\times Y = \bigcup_{k=1}^{\infty} A_k\times B_k$ donde las sucesiones $(A_k)_{k=1}^{\infty}$ y $(B_k)_{k=1}^{\infty}$ se eligen en $\mathcal{M}$ y $\mathcal{N}$ respectivamente tales que

$$X = \bigcup_{k=1}^{\infty} A_k, \quad \lambda(A_k) < \infty \qquad \text{y} \qquad Y = \bigcup_{k=1}^{\infty} B_k, \quad \nu(B_k) < \infty$$

para todo $k\in\mathbb{N}$, entonces de (1) se obtiene que $(\lambda\otimes\nu)(A_k\times B_k) < +\infty$ para todo $k\in\mathbb{N}$. Por esto $\lambda\otimes\nu$ es σ-finita. ■

Los resultados expuestos con anterioridad justifican la siguiente definición.

Definición 11.2.18. *Sean $(X,\mathcal{M},\lambda)$ y $(Y,\mathcal{N},\nu)$ dos espacios de medidas σ-finitas. La función de conjuntos $\lambda\otimes\nu : \mathcal{M}\otimes\mathcal{N} \to [0,+\infty]$ dada por*

$$(\lambda\otimes\nu)(E) = \int_X \nu(E_x)\,d\lambda(x) = \int_Y \lambda(E^y)\,d\nu(y)$$

*para todo $E\in\mathcal{M}\otimes\mathcal{N}$ es llamada la **medida producto** de λ y ν sobre $\mathcal{M}\otimes\mathcal{N}$.*

Como siempre, usaremos la notación $\mathrm{L}_1(M\times N, \lambda\otimes\nu)$ para designar *el espacio de todas las clases de equivalencias de funciones medibles $f : X\times Y \to \overline{\mathbb{R}}$ que son $\lambda\otimes\nu$-integrables*, es decir, tales que

$$\int_{X\times Y} |f|\,d(\lambda\otimes\nu) < +\infty.$$

Un adelanto de lo que será la demostración del Teorema de Fubini-Tonelli consiste en observar que si $E\in\mathcal{M}\otimes\mathcal{N}$ y $f = \chi_E$, entonces

$$(\chi_E)_x = \chi_{E_x} \qquad \text{y} \qquad (\chi_E)^y = \chi_{E^y}.$$

En efecto,

$$\chi_E(x,y) = 1 \quad\Leftrightarrow\quad (x,y) \in E \quad\Leftrightarrow\quad y \in E_x \quad\Leftrightarrow\quad \chi_{E_x}(y) = 1. \tag{1}$$

De modo similar, $(\chi_E)^y = \chi_{E^y}$. De esto se sigue que

$$(\lambda \otimes \nu)(E) = \int_X \nu(E_x)\, d\lambda(x) = \int_X \left(\int_Y \chi_{E_x}(y)\, d\nu(y) \right) d\lambda(x),$$

y también

$$(\lambda \otimes \nu)(E) = \int_Y \lambda(E^y)\, d\nu(y) = \int_Y \left(\int_X \chi_{E^y}(x)\, d\lambda(x) \right) d\nu(y),$$

de donde se obtiene que

$$\int_X \left(\int_Y \chi_{E_x}(y)\, d\nu(y) \right) d\lambda(x) = \int_Y \left(\int_X \chi_{E^y}(x)\, d\lambda(x) \right) d\nu(y) \tag{2}$$

Los siguientes dos teoremas extienden la igualdad obtenida en (2) para funciones medibles no-negativas y, también, si tales funciones son integrables, es decir, ellos dan condiciones bajo la cual es posible calcular una integral doble usando integrales iteradas. Pero, además, permiten intercambiar el orden de integración de las integrales iteradas. El primer resultado que vamos a presentar es una versión de uno atribuido tanto a Leonida Tonelli (1885-1946), así como a Guido Fubini (1879-1943) y comúnmente llamado el Teorema de Tonelli o Teorema de Fubini-Tonelli, mientras que el segundo resultado se debe fundamentalmente a Fubini.

Teorema 11.2.19 (Fubini-Tonelli). *Sean $(X, \mathcal{M}, \lambda)$ y $(Y, \mathcal{N}, \nu)$ espacios de **medidas σ-finitas** y suponga que $f : X \times Y \to \overline{\mathbb{R}}$ es una función $\mathcal{M} \otimes \mathcal{N}$-**medible** y **no-negativa**. Entonces:*

*(a) la función $x \to \int_Y f_x(y)\, d\nu(y)$ es $\mathcal{M}$-**medible**,*

*(b) la función $y \to \int_X f^y(x)\, d\lambda(x)$ es $\mathcal{N}$-**medible**, y*

(c) se cumplen las igualdades

$$\begin{aligned}\int_{X\times Y} f(x,y)\, d(\lambda \otimes \nu)(x,y) &= \int_X \left(\int_Y f_x(y)\, d\nu(y) \right) d\lambda(x) \\ &= \int_Y \left(\int_X f^y(x)\, d\lambda(x) \right) d\nu(y).\end{aligned}$$

Prueba. Del Teorema 11.2.14 sabemos que f_x y f^y son, respectivamente, $\mathcal{N}$-medible y $\mathcal{M}$-medible para todo $(x,y) \in X \times Y$. Para demostrar (a), (b) y (c) suponga que $f = \chi_E$ para algún $E \in \mathcal{M} \otimes \mathcal{N}$. Se sigue de (2) que

$$\begin{aligned}\int_{X\times Y} \chi_E(x,y)\, d(\lambda \otimes \nu)(x,y) &= (\lambda \otimes \nu)(E) \\ &= \int_X \left(\int_Y \chi_{E_x}(y)\, d\nu(y) \right) d\lambda(x) \\ &= \int_Y \left(\int_X \chi_{E^y}(x)\, d\lambda(x) \right) d\nu(y).\end{aligned}$$

Si f es una función simple $\mathcal{M} \otimes \mathcal{N}$-medible, entonces la linealidad de la integral nos revela que (a), (b) y (c) también se cumplen. Finalmente, si $f : X \times Y \to \overline{\mathbb{R}}$ es una función $\mathcal{M} \otimes \mathcal{N}$-medible y no-negativa, entonces existe una sucesión creciente no-negativa de funciones simples $\mathcal{M} \otimes \mathcal{N}$-medibles, digamos $(\varphi_n)_{n=1}^{\infty}$, tal que $\varphi_n(x,y) \to f(x,y)$ para cada $(x,y) \in X \times Y$. Para cada $y \in Y$, se tiene que

$$\varphi_n(x,y) \longrightarrow f(x,y) = f^y(x)$$

y, entonces, el Teorema de la Convergencia Monótona nos muestra que

$$\int_X f(x,y)\, d\lambda(x) = \lim_{n\to\infty} \int_X \varphi_n(x,y)\, d\lambda(x) = \int_X f^y(x)\, d\lambda(x)$$

y, por lo tanto, la función $y \to \int_X f^y(x)\, d\lambda(x)$ es $\mathcal{N}$-medible por ser el límite de una sucesión de funciones $\mathcal{N}$-medibles. Esto prueba (b). De modo enteramente similar se verifica que la función $x \to \int_Y f_x(y)\, d\lambda(y)$ es $\mathcal{M}$-medible de donde se sigue que (a) también se cumple. Para demostrar (c) usemos, una vez más, el Teorema de la Convergencia Monótona para obtener

$$\begin{aligned}
\int_{X\times Y} f(x,y)\, d(\lambda\otimes\nu)(x,y) &= \lim_{n\to\infty} \int_{X\times Y} \varphi_n(x,y)\, d(\lambda\otimes\nu)(x,y) \\
&= \lim_{n\to\infty} \int_Y \left(\int_X \varphi_n(x,y)\, d\lambda(x) \right) d\nu(y) \\
&= \int_Y \left(\lim_{n\to\infty} \int_X \varphi_n(x,y)\, d\lambda(x) \right) d\nu(y) \\
&= \int_Y \left(\int_X f^y(x)\, d\lambda(x) \right) d\nu(y).
\end{aligned}$$

Con una comprobación similar se establece la igualdad

$$\int_{X\times Y} f(x,y)\, d(\lambda\otimes\nu)(x,y) = \int_X \left(\int_Y f_x(y)\, d\nu(y) \right) d\lambda(x)$$

y termina la prueba. ■

El siguiente resultado, el cual es fundamental para calcular integrales en espacios productos, es una versión de un poderoso teorema atribuido a Fubini y conocido como el Teorema de Fubini.

Teorema 11.2.20 (Fubini). *Sean $(X,\mathcal{M},\lambda)$ y $(Y,\mathcal{N},\nu)$ espacios de **medidas σ-finitas** y suponga que $f : X \times Y \to \overline{\mathbb{R}}$ es una función $\mathcal{M} \otimes \mathcal{N}$-**medible** y $\lambda\otimes\nu$-**integrable**. Entonces:*

(a) $f_x \in \mathrm{L}_1(Y,\nu)$ para λ-casi-todo $x \in X$ y $\int_Y f_x(y)\, d\nu(y) \in \mathrm{L}_1(X,\lambda)$,

(b) $f^y \in \mathrm{L}_1(X,\lambda)$ para ν-casi-todo $y \in Y$ y $\int_Y f^y(x))\, d\lambda(x) \in \mathrm{L}_1(Y,\nu)$, y

(c) se cumplen las igualdades

$$\begin{aligned}
\int_{X\times Y} f(x,y)\, d(\lambda\otimes\nu)(x,y) &= \int_X \left(\int_Y f_x(y)\, d\nu(y) \right) d\lambda(x) \\
&= \int_Y \left(\int_X f^y(x)\, d\lambda(x) \right) d\nu(y).
\end{aligned}$$

Prueba. Que f sea $\lambda\otimes\nu$-integrable significa que

$$\int_{X\times Y}|f(x,y)|\,d(\lambda\otimes\nu)(x,y)\ <\ +\infty.$$

Puesto que $|f|_x=|f_x|$ y $|f|^y=|f^y|$, el Teorema de Fubini-Tonelli nos garantiza entonces que

$$\begin{aligned}\int_X\left(\int_Y|f_x(y)|\,d\nu(y)\right)d\lambda(x)\ &=\ \int_Y\left(\int_X|f^y(x)|\,d\lambda(x)\right)d\nu(y)\\ &=\ \int_{X\times Y}|f(x,y)|\,d(\lambda\otimes\nu)(x,y)\ <\ +\infty\end{aligned}\tag{11.2.2}$$

Consideremos las funciones $g:X\to\overline{\mathbb{R}}$ y $h:Y\to\overline{\mathbb{R}}$ definidas por

$$g(x)\ =\ \int_Y f_x(y)\,d\nu(y)\qquad\text{y}\qquad h(y)\ =\ \int_X f^y(x)\,d\lambda(x)$$

De (11.2.2) tenemos que

$$\int_X|g(x)|\,d\lambda(x)\ =\ \int_X\left|\int_Y f_x(y)\,d\nu(y)\right|d\lambda(x)\ \le\ \int_X\left(\int_Y|f_x(y)|\,d\nu(y)\right)d\lambda(x)\ <\ +\infty,\tag{1}$$

y puesto que toda *función integrable es finita casi-siempre*, Teorema 11.1.40 (c), página 714, resulta que $|g(x)|$ es finita casi-siempre, de donde se deduce que $f_x\in \mathrm{L}_1(Y,\nu)$ para casi-todo $x\in X$. También, de (1) se sigue que

$$\int_Y f_x(y)\,d\nu(y)\ \in\ \mathrm{L}_1(X,\lambda)$$

y termina la prueba de (a). Un argumento enteramente similar nos muestra que (b) se cumple.

Para demostrar la parte (c), pongamos $f=f^+-f^-$. Puesto que f^+ y f^- son funciones medibles no-negativas y, además, $(f^\pm)_x=(f_x)^\pm$ y $(f^\pm)^y=(f^y)^\pm$, entonces del Teorema de Fubini-Tonelli se obtiene

$$\int_X\left(\int_Y(f_x)^\pm\,d\nu\right)d\lambda\ =\ \int_{X\otimes Y}f^\pm\,d(\lambda\times\nu)\ =\ \int_Y\left(\int_X(f^y)^\pm\,d\lambda\right)d\nu\tag{2}$$

Puesto que $f_x=(f_x)^+-(f_x)^-$ y $f^y=(f^y)^+-(f^y)^-$, si se suman las integrales que corresponden a cada uno de los signos $+$ y $-$ se obtiene (c). Por ejemplo,

$$\begin{aligned}\int_{X\times Y}f(x,y)\,d(\lambda\otimes\nu)\ &=\ \int_{X\times Y}f^+(x,y)\,d(\lambda\otimes\nu)\ -\ \int_{X\times Y}f^-(x,y)\,d(\lambda\otimes\nu)\\ &=\ \int_X\left(\int_Y f_x^+(y)\,\nu(y)\right)d\lambda(x)\ -\ \int_X\left(\int_Y f_x^-(y)\,\nu(y)\right)d\lambda(x)\\ &=\ \int_X\left(\int_Y f(x,y)\,d\nu(y)\right)d\lambda(x)\end{aligned}$$

Fin de la prueba. ∎

Corolario 11.2.21 (Fubini). *Sean $(X, \mathcal{M}, \lambda)$ y $(Y, \mathcal{N}, \nu)$ espacios de* ***medidas σ-finitas*** *y suponga que $f : X \times Y \to \overline{\mathbb{R}}$ es una función $\mathcal{M} \otimes \mathcal{N}$-****medible****. Las siguientes condiciones son equivalentes:*

(1) $f \in L_1(X \times Y, \lambda \otimes \nu)$.

(2) $\displaystyle\int_X \left(\int_Y |f(x,y)|\, d\nu(y) \right) d\lambda(x) < +\infty \quad \Leftrightarrow \quad \int_Y \left(\int_X |f(x,y)|\, d\lambda(x) \right) d\nu(y) < +\infty.$

Prueba. Sigue inmediatamente del Teorema de Fubini. ■

Nota Adicional 11.2.6 Permítanme explicar qué es exactamente lo que hemos ocultado en la demostración del Teorema de Fubini. En primer lugar, puede suceder que la función $y \to f_x(y)$ no sea integrable para ciertos valores de x lo que, en muchas casos, es completamente normal. Por ejemplo, si tomamos $X = Y = \mathbb{R}$ y $\mu \otimes \mu$ es la medida producto sobre la σ-álgebra $\mathfrak{M}_\mu(\mathbb{R}) \otimes \mathfrak{M}_\mu(\mathbb{R})$, resulta que la función $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ dada por

$$f(x,y) = \begin{cases} \dfrac{1}{(x^2+y^2)^{1/2}} & \text{si } 0 < x^2 + y^2 < 1 \\ 0 & \text{en cualquier otro lugar,} \end{cases}$$

es $\mu \otimes \mu$-integrable. Sin embargo, tomando $x = 0$, se obtiene que $f_0(y) = f(0,y) = 1/|y|$ para $0 < |y| < 1$, la cual no es integrable. Por lo tanto, lo que el Teorema de Fubini afirma es que *existe un conjunto medible N tal que $\lambda(N) = 0$ y la función f_x es ν-integrable para todo* $x \in X \setminus N$. Por lo tanto, si definimos

$$g(x) = \begin{cases} \displaystyle\int_Y f_x(y)\, d\nu(y) & \text{si } x \in X \setminus N \\ 0 & \text{si } x \in N, \end{cases}$$

entonces g es $\mathcal{M}$-medible, integrable con respecto a λ y se cumple que

$$\int_{X\times Y} f(x,y)\, d(\lambda \otimes \nu) = \int_X g(x)\, d\lambda(x) = \int_X \left(\int_Y f_x(y)\, d\nu(y) \right) d\lambda(x)$$

Observe que no importa cómo se define $g(x)$ sobre N, ya que $\lambda(N) = 0$. El análisis para el caso (b) es similar.

Es importante apreciar la diferencia entre el Teorema de Fubini-Tonelli y el de Fubini. Aunque la conclusión en ambos resultados es la misma, ellos difieren en las hipótesis: el Teorema de Fubini-Tonelli dice que la integral doble de cualquier función medible f no-negativa, sea esta finita o no, siempre puede ser evaluada por una integral iterada independientemente si ellas son o no integrables. Por otro lado, en el Teorema de Fubini, a f no se le exige que sea no-negativa pero sí se requiere que sea integrable con respecto a la medida producto $\lambda \otimes \nu$.

Otro aspecto que es tremendamente importante en ambos teoremas es que la condición de que ambas medidas sean σ-fintas no pueden ser omitidas. En los siguientes ejemplos se muestran casos donde el Teorema de Fubini deja de cumplirse por faltar algunas de las hipótesis.

En el siguiente ejemplo haremos uso de la siguiente información: Recordemos que si una serie doble, digamos $\sum_{m,n=1}^{\infty} b_{mn}$, *converge absolutamente*, entonces $\sum_{m=1}^{\infty} b_{mn}$ y $\sum_{n=1}^{\infty} b_{mn}$ también convergen absolutamente y se cumple que

$$\sum_{m=1}^{\infty}\Big(\sum_{n=1}^{\infty} b_{mn}\Big) = \sum_{m,n=1}^{\infty} b_{mn} = \sum_{n=1}^{\infty}\Big(\sum_{m=1}^{\infty} b_{mn}\Big) < +\infty.$$

Ejemplo 11.2.1. *Existencia de una* ***función medible no-integrable*** $f : X \times Y \to \overline{\mathbb{R}}$ *tal que sus integrales iteradas*

$$\int_X \Big(\int_Y f(x,y)\, d\nu(y)\Big)\, d\lambda(x) \qquad \text{y} \qquad \int_Y \Big(\int_X f(x,y)\, d\lambda(x)\Big)\, d\nu(y)$$

existen pero son distintas.

Prueba. Sea $(a_{ij})_{i,j=1}^{\infty}$ una sucesión doble de números reales tal que las series

$$\sum_{j=1}^{\infty} a_{ij}, \quad i = 1, 2, \ldots, \qquad \sum_{i=1}^{\infty} a_{ij}, \quad j = 1, 2, \ldots, \tag{1}$$

$$\sum_{i=1}^{\infty}\Big(\sum_{j=1}^{\infty} a_{ij}\Big) \qquad \text{y} \qquad \sum_{j=1}^{\infty}\Big(\sum_{i=1}^{\infty} a_{ij}\Big) \tag{2}$$

sean todas absolutamente convergentes, pero que

$$\sum_{i=1}^{\infty}\Big(\sum_{j=1}^{\infty} a_{ij}\Big) \neq \sum_{j=1}^{\infty}\Big(\sum_{i=1}^{\infty} a_{ij}\Big).$$

Un modo de construir una tal sucesión es considerar, por ejemplo, para cada $i, j \in \mathbb{N}$, el número

$$a_{ij} = \begin{cases} 1 & \text{si } j = i \\ -1 & \text{si } j = i+1 \\ 0 & \text{en cualquier otro caso.} \end{cases}$$

y observar que todas las series en (1) y (2) sólo poseen un número finito de términos no-nulos por lo que todas ellas son absolutamente convergentes y se cumple que

$$\sum_{i=1}^{\infty}\Big(\sum_{j=1}^{\infty} a_{ij}\Big) = 0 \qquad \text{y} \qquad \sum_{j=1}^{\infty}\Big(\sum_{i=1}^{\infty} a_{ij}\Big) = 1. \tag{3}$$

Definamos ahora la función $f : [0, +\infty) \times [0, +\infty) \to \mathbb{R}$ por

$$f(x,y) = \sum_{i,j=1}^{\infty} a_{ij} \cdot \chi_{[i-1,\, i) \times [j-1,\, j)}(x,y).$$

Fijemos $x \in [0, +\infty)$, y observe que

$$f_x(y) = f(x,y) = \sum_{j=1}^{\infty} a_{ij} \cdot \chi_{[j-1,j)}(y)$$

para todo $y \in [0, +\infty)$ donde $i \geq 1$ es tal que $x \in [i-1, i)$. Como la serie $\sum_{j=1}^{\infty} a_{ij}$ es absolutamente convergente, se sigue que la función $y \to f_x(y)$ es integrable y

$$\int_{[0,+\infty)} f_x(y)\, d\mu(y) = \int_0^{+\infty} f(x,y)\, d\mu(y) = \sum_{j=1}^{\infty} a_{ij}.$$

De esto se obtiene que

$$\int_0^{+\infty} f(x,y)\, d\mu(y) = \sum_{i=1}^{\infty} \Big(\sum_{j=1}^{\infty} a_{ij} \Big) \cdot \chi_{[i-1,i)}(x)$$

para todo $x \in [0, +\infty)$, de donde se concluye que la función $x \to \int_0^{+\infty} f(x,y)\, d\mu(y)$ es integrable y

$$\int_0^{+\infty} \Big(\int_0^{+\infty} f(x,y)\, d\mu(y) \Big)\, d\mu(x) = \sum_{i=1}^{\infty} \Big(\sum_{j=1}^{\infty} a_{ij} \Big).$$

De modo completamente análogo, se muestra que

$$\int_0^{+\infty} \Big(\int_0^{+\infty} f(x,y)\, d\mu(x) \Big)\, d\mu(y) = \sum_{j=1}^{\infty} \Big(\sum_{i=1}^{\infty} a_{ij} \Big)$$

y, por lo tanto, usando (3), se concluye que

$$\int_0^{+\infty} \Big(\int_0^{+\infty} f(x,y)\, d\mu(y) \Big)\, d\mu(x) \neq \int_0^{+\infty} \Big(\int_0^{+\infty} f(x,y)\, d\mu(x) \Big)\, d\mu(y).$$

Esto prueba que la serie doble no converge absolutamente y se sigue entonces del Teorema de Fubini que f no es integrable. ■

Ejemplo 11.2.2. *Sea $([0,1], \mathfrak{Bo}(\mathbb{R}), \mu)$ el espacio de medida de Lebesgue sobre $\mathfrak{Bo}([0,1])$ y λ la medida de conteo sobre $\mathfrak{Bo}([0,1])$. Claramente λ no es σ-finita ya que $[0,1]$ es no-numerable. Considere el conjunto $D = \{(x,x) : x \in [0,1]\}$ y sea $f : [0,1] \times [0,1] \to \mathbb{R}$ dada por $f = \chi_D$. Entonces las integrales iteradas de f existen pero son distintas.*

Prueba. En efecto, fijemos $x \in [0,1]$. Entonces $f_x(y) = 1$ únicamente en el punto $y = x$ y cero en cualquier otro punto de $[0,1] \setminus \{x\}$. Puesto que λ es la medida de conteo, se tiene que $\int_0^1 f_x(y)\, d\lambda(y) = 1$ para todo $x \in [0,1]$ y, por consiguiente,

$$\int_0^1 \Big(\int_0^1 f_x(y)\, d\lambda(y) \Big)\, d\mu(x) = \int_0^1 1 \cdot \mu(x) = 1.$$

Por otro lado, para cada $y \in [0,1]$, y fijo, $f^y(x) = 1$ sólo en $x = y$ y cero en los otros puntos de $[0,1]$. Esto nos dice que la función f^y es igual a cero $\mu-$c.s. y, en consecuencia, $\int_0^1 f^y(x)\, d\mu(x) = 0$. Por esto,

$$\int_0^1 \Big(\int_0^1 f^y(x)\, d\mu(x) \Big)\, d\lambda(y) = \int_0^1 0 \cdot \lambda(y) = 0 \cdot \infty = 0.$$

Ejemplo 11.2.3. *Existencia de un conjunto $E \subseteq \mathbb{R}^2$ tal que los conjuntos E_x y E^y son μ-medibles para todo $x, y \in [0,1]$ y, sin embargo, E no es $\mu \otimes \mu$-medible.*

Prueba. Este ejemplo se sustenta sobre la base de la Hipótesis del Continuo. Considere el conjunto

$$\Omega = \{\alpha \in \text{Ord} : \alpha < \omega_1\},$$

donde ω_1 es el primer ordinal no-numerable. Puesto que $\text{card}(\Omega) = \aleph_1$, la Hipótesis del Continuo nos dice Ω es equipotente, por ejemplo, con el intervalo $[0,1]$. Sea $\Phi : [0,1] \to \Omega$ una biyección y defina

$$E = \{(x,y) \in [0,1] \times [0,1] : \Phi(y) \leq \Phi(x)\}.$$

(1) *$E_x \in \mathfrak{M}_\mu(\mathbb{R})$ para todo $x \in [0,1]$.* En efecto, sea $x \in [0,1]$ y observe que como $\Phi(x) < \omega_1$, entonces $\Phi(x)$ es numerable y, por lo tanto,

$$E_x = \{y \in [0,1] : \Phi(y) \leq \Phi(x)\}$$

es un conjunto numerable. Puesto que $\mu(E_x) = 0$, se tiene que $E_x \in \mathfrak{M}_\mu(\mathbb{R})$ para todo $x \in [0,1]$.

(2) *$E^y \in \mathfrak{M}_\mu(\mathbb{R})$ para todo $y \in [0,1]$.* Para ver esto, nótese que para cada $y \in [0,1]$,

$$E^y = \{x \in [0,1] : \Phi(y) \leq \Phi(x)\}$$

es el complemento de un conjunto numerable y, en consecuencia, él es medible.

(3) *E no es medible.* Suponga, por un momento, que E es medible y considere la función $f : [0,1] \times [0,1] \to [0,1]$ dada por $f = \chi_E$. Resulta que f es una función $\mathfrak{M}_\mu(\mathbb{R}) \otimes \mathfrak{M}_\mu(\mathbb{R})$-medible no-negativa y sus secciones f_x y f^y son μ-medibles ya que ellas son constantes casi-siempre. Pero, además,

$$\int_0^1 f(x,y)\, d\mu(x) = 1 \qquad \text{y} \qquad \int_0^1 f(x,y)\, d\mu(y) = 0.$$

Esto, por supuesto, viola la conclusión del Teorema de Fubini-Tonelli, Teorema 11.2.19, por lo que nuestra suposición de que E era medible es falsa. ■

11.3. La Medida de Lebesgue en $\mathbb{R}^n$

Ya hemos asomado la idea de que la noción de medida de Lebesgue construida en $\mathbb{R}$ se podía llevar a cabo en cualquier $\mathbb{R}^n$ con $n \geq 2$. En esta parte vamos a extender esa noción sobre $\mathbb{R}^n$ para cualquier $n \geq 2$ construyendo una σ-álgebra $\mathfrak{M}(\mathbb{R}^n)$ de subconjuntos de $\mathbb{R}^n$ conteniendo a los conjuntos abiertos y una medida μ^n sobre $\mathfrak{M}(\mathbb{R}^n)$ tal que la medida de cada "rectángulo" R de $\mathbb{R}^n$ sea precisamente el "volumen" de R.

Sean $a_1, \ldots, a_n, b_1, \ldots, b_n$ puntos en $\mathbb{R}$ tales que $a_i \leq b_i$ para $i = 1, 2, \ldots, n$. Cualquier conjunto de la forma

$$\begin{aligned} Q_{a,b} &= [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n] \\ &= \{(\xi_1, \ldots, \xi_n) : a_i \leq \xi_i \leq b_i,\ i = 1, 2, \ldots, n\}. \end{aligned}$$

será llamado un **rectángulo cerrado** n-dimensional de $\mathbb{R}^n$. Si las coordenadas de cada vértice de $Q_{a,b}$ son números racionales, es decir, si todos los a_i, b_i son números racionales, entonces diremos que $Q_{a,b}$ es un ***rectángulo racional***. En general, si $I_1, \ldots, I_n$ son subintervalos en $\mathbb{R}$, lo cuales pueden ser abiertos, cerrados o semi-cerrados, llamaremos **rectángulo n-dimensional** al conjunto

$$Q \;=\; I_1 \times \cdots \times I_n.$$

En este caso, su **volumen** viene expresado en la forma

$$\mathbf{vol}(Q) \;=\; \ell(I_1) \cdot \ldots \cdot \ell(I_n),$$

Observe que si uno de los intervalos, digamos I_j, se reduce a un punto, es decir,

$$Q \;=\; I_1 \times \cdots \times \{t_j\} \times \cdots \times I_n$$

entonces

$$\mathbf{vol}(Q) \;=\; \ell(I_1) \cdot \ell(I_2) \cdot \ldots \cdot 0 \cdot \ldots \cdot \ell(I_n) \;=\; 0,$$

aun si alguno de los intervalos I_k es de longitud infinita ya que, de acuerdo a nuestra convención, $0 \cdot \infty \;=\; 0$. Si la longitud de todos los intervalos $I_1, \ldots, I_n$ del rectángulo $Q \;=\; I_1 \times \cdots \times I_n$ es la misma, digamos $\ell(I_1) \;=\; \cdots \;=\; \ell(I_n) \;=\; \rho$, entonces diremos que Q es un **ρ-cubo** y lo denotaremos por Q_ρ. Pongamos

$$Q^1 \;=\; \underbrace{[0,1] \times \cdots \times [0,1]}_{n \text{ veces}} = [0,1]^n$$

y observe que cualquier ρ-cubo Q_ρ se puede representar en la forma

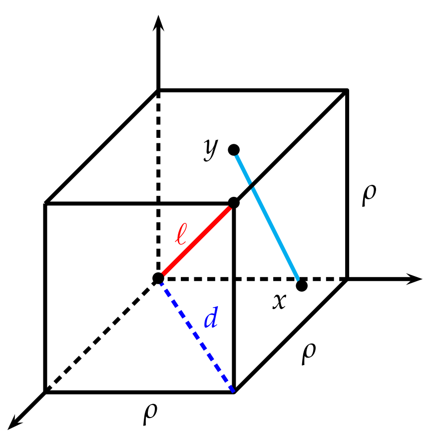
$$Q_\rho \;=\; x_0 + \rho \cdot Q^1,$$

para algún $x_0 \in \mathbb{R}^n$. Nótese también que si $\|\cdot\|$ es la norma euclídea sobre $\mathbb{R}^n$, es decir,

$$\|\, x \,\|_2 \;=\; \left(x_1^2 + \cdots + x_n^2\right)^{1/2} \quad \text{para todo} \quad x = (x_1, \ldots, x_n) \in \mathbb{R}^n$$

resulta que la *longitud de cualquier diagonal* en cualquier ρ-cubo Q_ρ es igual a $\sqrt{n}\,\rho$ y, por lo tanto, si $[x,y] \;=\; \{tx + (1-t)y \;:\; 0 \leq t \leq 1\}$ es cualquier segmento que une dos puntos arbitrarios $x, y \in Q_\rho$, entonces

$$\|\, x - y \,\|_2 \;\leq\; \sqrt{n}\,\rho.$$



$$d \;=\; \sqrt{\rho^2 + \rho^2} \;=\; \sqrt{2}\,\rho$$

$$\ell \;=\; \sqrt{d^2 + \rho^2} \;=\; \sqrt{3}\,\rho$$

Para cada $k \in \mathbb{Z}$ y cada n-upla $m = (m_1, \ldots, m_n) \in \mathbb{Z}^n$, considere el rectángulo $Q_{k,m}$ definido por

$$Q_{k,m} = \left\{ (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n : \frac{m_j}{2^k} < \xi_i \leq \frac{m_j + 1}{2^k}, \; i = 1, 2, \ldots, n \right\}$$
$$= \prod_{i=1}^{n} \left(\frac{m_j}{2^k}, \frac{m_j + 1}{2^k} \right].$$

Cada rectángulo de la forma $Q_{k,m}$ será llamado un ***cubo diádico*** en $\mathbb{R}^n$. Observe que cada vértice del cubo $Q_{k,m}$ es una n-upla en $2^{-k}\mathbb{Z}^n$ y la longitud de cada uno de sus lados es 2^{-k}. Si para cada $k \in \mathbb{Z}$, denotamos por D_k la colección de todos los cubos diádicos $Q_{k,m}$, para todo $m \in \mathbb{Z}^n$, resulta que D_k es una colección numerable y disjunta. Por consiguiente, el conjunto

$$\mathfrak{D}_n = \bigcup_{k \in \mathbb{Z}} D_k,$$

el cual representa la colección de todos los cubos diádicos en $\mathbb{R}^n$, también es numerable. Nótese que cada elemento de D_k tiene volumen $\left(\frac{1}{2^k}\right)^n$. Los cubos diádicos tienen muchas propiedades importantes. Por ejemplo:

(1) *Para cada $k \in \mathbb{Z}$, los cubos de D_k forman una **partición numerable** de $\mathbb{R}^n$, es decir,*

$$\mathbb{R}^n = \bigcup_{Q \in D_k} Q = \bigcup_{m \in \mathbb{Z}^n} Q_{m,k} \qquad \text{con} \qquad Q_{k,m} \cap Q_{k,m'} = \varnothing \text{ si } m \neq m'.$$

(2) *Si $Q, Q' \in \mathfrak{D}_n$, entonces sólo puede ocurrir uno de estos dos casos: o son disjuntos, o uno de ellos está incluido en el otro. En particular, si $Q_1, \ldots, Q_m$ son cubos diádicos y $\bigcap_{j=1}^{m} Q_j \neq \varnothing$, entonces $\bigcup_{j=1}^{m} Q_j \in \mathfrak{D}_n$.*

(3) *Dado $Q \in D_k$, existen 2^n cubos de D_{k+1}, digamos $Q_1, \ldots, Q_{2^n}$, tales que $Q_j \in Q$, $j = 1, \ldots, 2^n$.*

A diferencia de los subconjuntos abiertos de $\mathbb{R}$, los cuales se pueden expresar como una unión numerable de intervalos abiertos disjuntos dos a dos, los conjuntos abiertos en $\mathbb{R}^n$ con $n > 1$ no se pueden representar como una unión numerable de rectángulos abiertos disjuntos dos a dos. Sin embargo, si los rectángulos son cubos diádicos, tal descomposición siempre es posible como se muestra a continuación. Observe que el siguiente resultado ya lo obtuvimos para el caso en que $n = 1$, Lema 2.2.10, página 124.

Teorema 11.3.1 (Caracterización de los abiertos de $\mathbb{R}^n$). *Sea G un **conjunto abierto** en $\mathbb{R}^n$. Entonces existe una sucesión disjunta $(Q_j)_{j=1}^{\infty}$ en $\mathfrak{D}_n$ tal que $G = \bigcup_{j=1}^{\infty} Q_j$.*

Prueba. Sea $x \in G$. Como G es abierto, existe un cubo diádico Q tal que $x \in Q \subseteq G$. Esto permite considerar la siguiente colección:

$$\mathcal{F} = \left\{ Q \in \bigcup_{k=0}^{\infty} D_k : Q \subseteq G \right\}.$$

Nótese que la familia $\mathcal{F}$ consiste de todos los cubos diádicos cuyos lados tienen longitud ≤ 1. Ordene parcialmente a $\mathcal{F}$ por la relación de inclusión $\subseteq$ y sea $\mathcal{C}$ una cadena en $\mathcal{F}$. Por la propiedad (2) de los cubos diádicos se tiene que $\widehat{Q} = \bigcup_{Q \in \mathcal{C}} Q$ es un cubo diádico que contiene a todos los miembros de $\mathcal{C}$, es decir, $\widehat{Q}$ es una cota superior de $\mathcal{C}$. Un llamado al Lema de Zorn nos dice que $\mathcal{F}$ posee al menos un elemento maximal, esto es, un cubo diádico maximal. Sea $\mathcal{F}_{\max}$ el conjunto de todos los cubos diádicos de $\mathcal{F}$ que son maximales. Veamos que:

(a) *$\mathcal{F}_{max}$ es una familia numerable y disjunta.* En efecto, como $\mathcal{F}$ es numerable, también lo es $\mathcal{F}_{\max}$. Para verificar que $\mathcal{F}_{\max}$ es disjunta, seleccione un par arbitrario de cubos $Q_1, Q_2 \in \mathcal{F}_{\max}$. Afirmamos que $Q_1 = Q_2$ o $Q_1 \cap Q_2 = \varnothing$. Para ver esto, suponga que $Q_1 \cap Q_2 \neq \varnothing$. Entonces $Q_1 \cup Q_2$ es un cubo diádico contenido en G y como Q_1 es maximal resulta que $Q_1 \cup Q_2 \subseteq Q_1$. De modo similar se tiene que $Q_1 \cup Q_2 \subseteq Q_2$, de donde se concluye que $Q_1 = Q_2$. Escribamos

$$\mathcal{F}_{\max} = \{Q_1, Q_2, \ldots\}$$

(b) $G = \bigcup_{j=1}^{\infty} Q_j$. Puesto que $Q_j \subseteq G$ para todo $j \geq 1$, sólo nos resta demostrar la inclusión $G \subseteq \bigcup_{j=1}^{\infty} Q_j$. Sea $x \in G$. Entonces existe un $Q \in \mathcal{F}$ tal que $x \in Q$ y, por supuesto, Q está contenido en un cubo diádico maximal Q_{j_0}. Por esto $x \in Q_{j_0} \subseteq \bigcup_{j=1}^{\infty} Q_j$. Esto prueba que $G \subseteq \bigcup_{j=1}^{\infty} Q_j$ y finaliza la demostración. ■

Puesto que la clausura de dos cubos diádicos son casi-disjuntos o uno de ellos incluye al otro, se tiene el siguiente:

Corolario 11.3.2. *Sea G un **conjunto abierto** en $\mathbb{R}^n$. Entonces existe una sucesión $(Q_j)_{j=1}^{\infty}$ de cubos **diádicos cerrados casi-disjuntos** tal que $G = \bigcup_{j=1}^{\infty} Q_j$.*

Una forma más constructiva de demostrar el resultado anterior se obtiene reduciéndolo al caso cuando $n = 2$ ya que en este contexto la demostración es esencialmente la misma a la del caso general. Suponga entonces que $n = 2$. Recordemos que

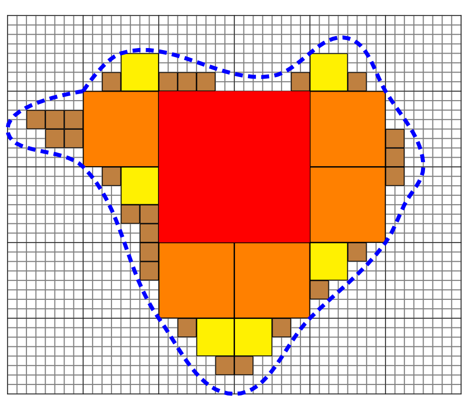
$$D_1 = \{Q_{m,1} : m \in \mathbb{Z}^2\} = \{(m_1, m_1 + 1] \times (m_2, m_2 + 1] : m_1, m_2 \in \mathbb{Z}\}$$

y

$$\mathbb{R}^2 = \bigcup_{m \in \mathbb{Z}^2} Q_{m,1}.$$

Se consideran ahora todos los miembros de D_1 que están contenidos en G, esto es, sea

$$\mathcal{G}_1 = \{Q \in D_1 : Q \subseteq G\}.$$



A continuación elegimos los miembros de D_2 contenidos en G pero no en la unión de los cubos pertenecientes a $\mathcal{G}_1$, es decir, sea

$$\mathcal{G}_2 \;=\; \Big\{Q \in D_2 \,:\, Q \subseteq G,\; Q \not\subseteq \bigcup \mathcal{G}_1\Big\}.$$

y se continua. La colección $\mathcal{G} = \bigcup_{k=1}^{\infty} \mathcal{G}_k$ es la que buscamos. En efecto:

(a) $\mathcal{G}$ es claramente una colección numerable y, por supuesto, disjunta por construcción.

(b) $G = \bigcup \mathcal{G}$. Para ver esto último, sea $x \in G$. Como G es abierto, existe un rectángulo abierto $U \subseteq G$. Dividiendo tal rectángulo en piezas cada vez más pequeñas, se puede hallar un k lo suficientemente grande de modo que algún cubo diádico $Q_k(x) \in \mathcal{G}_k$ satisface que $x \in Q_k(x) \subseteq G$. Con esto queda demostrado (b) y termina la prueba.

No es difícil establecer que si $(Q'_n)_{n=1}^{\infty}$ es otra sucesión disjunta de cubos diádicos en $\mathfrak{D}_n$ cuya unión es G, entonces

$$\sum_{n=1}^{\infty} \mathbf{vol}(Q_n) \;=\; \sum_{n=1}^{\infty} \mathbf{vol}(Q'_n).$$

Este hecho permite definir el **volumen** de G como:

$$\mathbf{vol}(G) \;=\; \sum_{n=1}^{\infty} \mathbf{vol}(Q_n)$$

donde $(Q_n)_{n=1}^{\infty}$ es cualquier sucesión disjunta de cubos semi-cerrados en $\mathfrak{D}_n$ cuya unión cubre a G. Esta observación nos proporciona un modo elegante de definir, similar al caso real, una medida exterior sobre los subconjuntos de $\mathbb{R}^n$ del modo siguiente.

Definición 11.3.3. *La aplicación* $\mu_n^* : \mathscr{P}(\mathbb{R}^n) \to [0, +\infty]$ *definida por*

$$\mu_n^*(A) \;=\; \inf\big\{\mathbf{vol}(G) \,:\, A \subseteq G,\; G \textit{ abierto}\big\}$$

se llama la ***medida exterior de Lebesgue*** *en* $\mathbb{R}^n$.

Similar al caso de la medida exterior en $\mathbb{R}$, la medida exterior de Lebesgue en $\mathbb{R}^n$ goza de las mismas propiedades. Recordemos algunas de ellas:

(1) $0 \leq \mu_n^*(A) \leq +\infty, \qquad \mu_n^*(\varnothing) = 0.$

(2) $\mu_n^*(A) \leq \mu_n^*(B)$ siempre que $A \subseteq B$.

(3) $\mu_n^*\Big(\bigcup_{n=1}^{\infty} A_n\Big) \leq \sum_{n=1}^{\infty} \mu_n^*(A_n)$ para cualquier sucesión $(A_n)_{n=1}^{\infty}$.

(4) $\mu_n^*(R) = \mathbf{vol}(R)$ si R es un cubo n-dimensional.

(5) $\mu_n^*(x + A) = \mu_n^*(A)$ para todo $x \in \mathbb{R}^n$ y cualquier $A \subseteq \mathbb{R}^n$.

Definición 11.3.4. *Un conjunto* $E \subseteq \mathbb{R}^n$ *se dice que* ***medible Lebesgue*** *si para cada* $\varepsilon > 0$, *existe un conjunto abierto* $G \subseteq \mathbb{R}^n$ *tal que*

$$E \subseteq G \qquad \textit{y} \qquad \mu_n^*(G \setminus E) \;<\; \varepsilon.$$

Como antes, se comprueba que la colección $\mathfrak{M}(\mathbb{R}^n)$, *formada por todos los conjuntos* $E \subseteq \mathbb{R}^n$ *que son medibles Lebesgue*, es una σ-álgebra llamada la σ-**álgebra de Lebesgue** en $\mathbb{R}^n$. De la definición anterior resulta que la colección $\mathcal{O}_n$ de todos los subconjuntos abiertos de $\mathbb{R}^n$ habita en $\mathfrak{M}(\mathbb{R}^n)$, esto es,

$$\mathcal{O}_n \subseteq \mathfrak{M}(\mathbb{R}^n).$$

Definición 11.3.5. *La aplicación* $\mu_n : \mathfrak{M}(\mathbb{R}^n) \to [0, +\infty]$ *dada por*

$$\mu_n \;=\; \mu_n^*\big|_{\mathfrak{M}(\mathbb{R}^n)}$$

es llamada la ***medida de Lebesgue*** *en* $\mathbb{R}^n$.

Las equivalencias de las condiciones dadas en el siguiente resultado pueden ser demostradas de modo similar al caso real y por lo tanto se omiten.

Teorema 11.3.6. *Para cada conjunto* $E \subseteq \mathbb{R}^n$ *las siguientes condiciones son equivalentes:*

(1) $E \in \mathfrak{M}(\mathbb{R}^n)$.

(2) *Para cada* $\varepsilon > 0$, *existe un conjunto abierto* $G \supseteq E$ *con* $\mu_n^*(G \setminus E) < \varepsilon$.

(3) *Existe un conjunto* G_δ $G \supseteq E$ *con* $\mu_n^*(G \setminus E) = 0$.

(4) *Para cada* $\varepsilon > 0$, *existe un conjunto cerrado* $F \subseteq E$ *con* $\mu_n^*(E \setminus F) < \varepsilon$.

(5) *Existe un conjunto* F_σ $F \subseteq E$ *con* $\mu_n^*(E \setminus F) = 0$.

El siguiente observación es útil pues permite describir a los elementos de $\mathfrak{M}(\mathbb{R}^n)$ como conjuntos que son muy parecidos a los conjuntos F_σ, es decir, la diferencia entre un conjunto medible y un cierto conjunto F_σ es un conjunto nulo.

Corolario 11.3.7. *Sea* $E \subseteq \mathbb{R}^n$. *Las siguientes condiciones son equivalentes:*

(1) $E \in \mathfrak{M}(\mathbb{R}^n)$.

(2) *Existen* ***conjuntos disjuntos*** B *y* N *en* $\mathbb{R}^n$ *tales que* B *es un* F_σ, $\mu_n(N) = 0$ *y*

$$E \;=\; B \cup N.$$

Prueba. Claramente $(2) \Rightarrow (1)$. Para verificar que $(1) \Rightarrow (2)$ observe que por (3) y (5) del resultado anterior existe un conjunto G_δ $G \supseteq E$ con $\mu_n(G \setminus E) = 0$ y también un conjunto F_σ $F \subseteq E$ tal que $\mu_n(E \setminus F) = 0$. Entonces

$$F \subseteq E \subseteq G \qquad \text{y} \qquad G \setminus F \;=\; (G \setminus E) \cup (E \setminus F)$$

donde $(G \setminus E) \cap (E \setminus F) = \varnothing$. De esto se sigue que

$$\mu_n(G \setminus F) \;=\; \mu_n(G \setminus E) + \mu_n(E \setminus F) \;=\; 0.$$

Nótese que $E = F \cup (E \setminus F)$, de modo que si definimos $B = F$ y $N = E \setminus F$, resulta $N \subseteq G \setminus F$ y, por lo tanto, $\mu_n(N) = 0$ y termina la prueba. ■

Similar a la prueba de la regularidad de la medida de Lebesgue en $\mathbb{R}$, se puede constatar que μ_n también es *regular*, esto es,

Teorema 11.3.8 (Regularidad de μ_n). *Para cada $E \in \mathfrak{M}(\mathbb{R}^n)$,*

$$\begin{aligned}\mu_n(E) &= \inf\{\mu_n(G) : A \subseteq G,\ G \text{ abierto}\}\\ &= \sup\{\mu_n(K) : K \subseteq E,\ K \text{ compacto}\}\end{aligned}$$

Observe que si $E \in \mathfrak{M}(\mathbb{R}^n)$ es de medida finita, entonces la regularidad de μ_n nos garantiza que, por cada $\varepsilon > 0$ fijado arbitrariamente, existe un conjunto compacto K y un conjunto abierto G tales que

$$K \subseteq E \subseteq G \qquad \text{y} \qquad \mu_n(G \setminus K) < \varepsilon.$$

Definición 11.3.9. *La σ-álgebra generada por $\mathcal{O}_n$ se llama la **σ-álgebra de Borel** en $\mathbb{R}^n$ la cual será denotada por $\mathfrak{Bo}(\mathbb{R}^n)$. La restricción de la medida de Lebesgue μ_n a $\mathfrak{Bo}(\mathbb{R}^n)$ se llama la **medida de Borel** en $\mathbb{R}^n$.*

Puesto que cualquier σ-álgebra es cerrada bajo complementos, resulta que $\mathfrak{Bo}(\mathbb{R}^n)$ también es generada por la colección $\mathcal{F}_n$ de todos los subconjuntos cerrados en $\mathbb{R}^n$. Más aun, puesto que $\mathbb{R}^n$ es σ-compacto (esto significa que es unión numerable de compactos), entonces $\mathfrak{Bo}(\mathbb{R}^n)$ es generada por la colección $\mathcal{K}_n$ de los compactos en $\mathbb{R}^n$. En general, existen muchas otras familias de subconjuntos de $\mathbb{R}^n$ que generan a la σ-álgebra de los borelianos. Por ejemplo, la colección $\mathfrak{D}_n$ de todos los cubos diádicos n-dimensionales, satisface

$$\mathfrak{Bo}(\mathbb{R}^n) = \sigma(\mathfrak{D}_n).$$

En efecto, por el Teorema 11.3.1 sabemos que $\mathcal{O}_n \subseteq \sigma(\mathfrak{D}_n)$ y, en consecuencia, por ser $\mathfrak{Bo}(\mathbb{R}^n)$ la σ-álgebra más pequeña conteniendo a $\mathcal{O}_n$ se tiene que $\mathfrak{Bo}(\mathbb{R}^n) \subseteq \sigma(\mathfrak{D}_n)$. Por otro lado, como $\mathfrak{D}_n \subseteq \mathfrak{Bo}(\mathbb{R}^n)$ se obtiene lo deseado.

Teorema 11.3.10. *Sean $m, n \in \mathbb{N}$. Si identificamos $\mathbb{R}^{m+n} = \mathbb{R}^m \times \mathbb{R}^n$, entonces*

$$\mathfrak{Bo}(\mathbb{R}^{m+n}) = \mathfrak{Bo}(\mathbb{R}^m) \otimes \mathfrak{Bo}(\mathbb{R}^n).$$

Prueba. Cualquier rectángulo cerrado $(m+n)$-dimensional es el producto de un rectángulo cerrado m-dimensional con un rectángulo cerrado n-dimensional. De allí que

$$\mathfrak{Bo}(\mathbb{R}^m) \otimes \mathfrak{Bo}(\mathbb{R}^n) \supseteq \mathcal{R}(\mathbb{R}^{m+n}),$$

donde $\mathcal{R}(\mathbb{R}^{m+n})$ denota la colección de todos los rectángulos cerrados en $\mathbb{R}^{m+n}$. Por otro lado, puesto que $\mathcal{R}(\mathbb{R}^{m+n}) \subseteq \mathcal{F}_{m+n}$, donde $\mathcal{F}_{m+n}$ es la familia de todos todos los subconjuntos cerrados de $\mathbb{R}^{m+n}$ la cual, como sabemos, genera a $\mathfrak{Bo}(\mathbb{R}^{m+n})$, resulta que

$$\mathfrak{Bo}(\mathbb{R}^m) \otimes \mathfrak{Bo}(\mathbb{R}^n) \supseteq \mathfrak{Bo}(\mathbb{R}^{m+n}).$$

Para demostrar la otra inclusión, considere la colección

$$\mathcal{M}_m = \{A \subseteq \mathbb{R}^m : A \times \mathbb{R}^n \in \mathfrak{Bo}(\mathbb{R}^{m+n})\}.$$

Puesto que $\mathfrak{Bo}(\mathbb{R}^{m+n})$ es una σ-álgebra, se tiene que $\mathcal{M}_m$ es una σ-álgebra en $\mathbb{R}^m$. Observe, además, que $\mathcal{M}_m$ contiene a todos lo abiertos de $\mathbb{R}^m$ ya que $G \times \mathbb{R}^n$ es abierto en $\mathbb{R}^{m+n}$ si G

es abierto en $\mathbb{R}^m$. De esto se sigue que $\mathfrak{Bo}(\mathbb{R}^m) \subseteq \mathcal{M}_m$. Este hecho nos revela, en particular, que $A \times \mathbb{R}^n \in \mathfrak{Bo}(\mathbb{R}^{m+n})$ para cualquier $A \in \mathfrak{Bo}(\mathbb{R}^m)$, lo cual significa que

$$\mathfrak{Bo}(\mathbb{R}^{m+n}) \supseteq \{A \times \mathbb{R}^n : A \in \mathfrak{Bo}(\mathbb{R}^m)\}.$$

Un argumento enteramente similar conduce a

$$\mathfrak{Bo}(\mathbb{R}^{m+n}) \supseteq \{\mathbb{R}^m \times B : B \in \mathfrak{Bo}(\mathbb{R}^n)\}$$

lo cual implica que

$$\mathfrak{Bo}(\mathbb{R}^{m+n}) \supseteq \mathfrak{Bo}(\mathbb{R}^m) \otimes \mathfrak{Bo}(\mathbb{R}^n).$$

Esto termina la prueba. ■

Fijemos $n \in \mathbb{N}$. Por una aplicación repetida del resultado anterior resulta que

$$\mathfrak{Bo}(\mathbb{R}^n) = \underbrace{\mathfrak{Bo}(\mathbb{R}) \otimes \cdots \otimes \mathfrak{Bo}(\mathbb{R})}_{n-\text{veces}},$$

es decir, la σ-álgebra de Borel en $\mathbb{R}^n$ es el producto de las σ-álgebras de Borel en $\mathbb{R}$. Este hecho permite, entonces, representar a la medida de Borel μ_n sobre $\mathfrak{Bo}(\mathbb{R}^n)$ por

$$\mu_n = \underbrace{\mu \otimes \cdots \otimes \mu}_{n-\text{veces}} \quad \text{sobre} \quad \mathfrak{Bo}(\mathbb{R}^n).$$

Esta afortunada propiedad de la medida de Borel nos revela que cualquier conjunto E en $\mathfrak{Bo}(\mathbb{R}^n)$ se puede expresar en la forma

$$E = B_1 \times B_2 \times \cdots \times B_n, \tag{1}$$

donde $B_j \in \mathfrak{Bo}(\mathbb{R})$ para $j = 1, 2, \ldots, n$ y, por supuesto,

$$\mu_n(E) = \mu(B_1) \cdot \mu(B_2) \cdot \ldots \cdot \mu(B_n). \tag{2}$$

El hecho de que cualquier conjunto $E \in \mathfrak{M}(\mathbb{R}^n)$ se puede escribir en la forma $E = B \cup N$, donde $B \in \mathfrak{Bo}(\mathbb{R}^n)$ y $\mu_n(N) = 0$, permite concluir que

Corolario 11.3.11. *$(\mathbb{R}^n, \mathfrak{M}(\mathbb{R}^n), \mu_n)$ es la* ***completación*** *de $(\mathbb{R}^n, \mathfrak{Bo}(\mathbb{R}^n), \mu_n)$.*

Estos hechos conducen, en muchos casos, a preferir trabajar con $\mathfrak{Bo}(\mathbb{R}^n)$ en lugar de $\mathfrak{M}(\mathbb{R}^n)$ que es un espacio mucho más grande.

Nota Adicional 11.3.7 Es importante destacar, sin embargo, que el Teorema 11.3.10 no es válido para la σ-álgebra de Lebesgue, es decir,

$$\mathfrak{M}(\mathbb{R}^m) \otimes \mathfrak{M}(\mathbb{R}^n) \neq \mathfrak{M}(\mathbb{R}^m \times \mathbb{R}^n).$$

Para ver que esto, sean $(\mathbb{R}^m, \mathfrak{M}(\mathbb{R}^m), \mu_m)$ y $(\mathbb{R}^n, \mathfrak{M}(\mathbb{R}^n), \mu_n)$ los espacios de medidas de Lebesgue en $\mathbb{R}^m$ y $\mathbb{R}^n$ respectivamente. Sea $x \in \mathbb{R}^m$ y elija un conjunto no-medible en $\mathbb{R}^n$, por ejemplo, un conjunto de Vitali $V \subseteq [0,1]^n$, el cual se puede construir de modo similar al conjunto de Vitali no-medible construido en $[0,1]$. Considere ahora el conjunto $E = \{x\} \times V \subseteq \mathbb{R}^m \times \mathbb{R}^n$. Afirmamos que

$$E \in \mathfrak{M}(\mathbb{R}^m \times \mathbb{R}^n).$$

En efecto, como $E = \{x\} \times V \subseteq \{x\} \times [0,1]^n \subseteq \mathbb{R}^m \times \mathbb{R}^n$ y $\mu_{m+n}(\{x\} \times [0,1]^n) = 0$, resulta que $\mu_{m+n}(E) = 0$ y, en consecuencia, $E \in \mathfrak{M}(\mathbb{R}^m \times \mathbb{R}^n)$ ya que μ_{m+n} es una medida completa. Sin embargo, puesto que $E_x = V \notin \mathfrak{M}(\mathbb{R}^n)$, se tiene que $E \notin \mathfrak{M}(\mathbb{R}^m) \otimes \mathfrak{M}(\mathbb{R}^n)$, por lo que las σ-álgebras $\mathfrak{M}(\mathbb{R}^m \times \mathbb{R}^n)$ y $\mathfrak{M}(\mathbb{R}^m) \otimes \mathfrak{M}(\mathbb{R}^n)$ no coinciden. Por esto,

$$\mathfrak{M}(\mathbb{R}^m) \otimes \mathfrak{M}(\mathbb{R}^n) \subsetneqq \mathfrak{M}(\mathbb{R}^m \times \mathbb{R}^n).$$

De paso, éste ejemplo también muestra que $\mu_m \otimes \mu_n$, *la medida producto de dos medidas completas, no es necesariamente completa*. La **completación** del espacio de medida

$$\big(\mathbb{R}^{m+n}, \mathfrak{M}(\mathbb{R}^m) \otimes \mathfrak{M}(\mathbb{R}^n), \mu_m \otimes \mu_n\big)$$

es el espacio de medida de Lebesgue $\big(\mathbb{R}^{m+n}, \mathfrak{M}(\mathbb{R}^{m+n}), \mu_{m+n}\big)$.

Teorema 11.3.12 (Invariancia por traslación). *$\mu_n : \mathfrak{Bo}(\mathbb{R}^n) \to [0,+\infty]$ es **invariante por traslación**, es decir,*

$$\mu_n(x+E) = \mu_n(E) \tag{1}$$

para todo conjunto medible $E \in \mathfrak{Bo}(\mathbb{R}^n)$ y todo $x \in \mathbb{R}^n$.

Prueba. Sea $E = B_1 \times \cdots \times B_n$ con $B_i \in \mathfrak{Bo}(\mathbb{R})$ para cada $i = 1, \ldots, n$. Puesto que μ es invariante por traslación sobre $\mathfrak{Bo}(\mathbb{R})$ se tiene que, para cada $x \in \mathbb{R}^n$,

$$\mu_n(x+E) = \mu(x+B_1) \cdot \ldots \cdot \mu(x+B_n) = \mu(B_1) \cdot \ldots \cdot \mu(B_n) = \mu_n(E)$$

y termina la prueba. ■

Nota Adicional 11.3.8 Recordemos que el conjunto $Q_i = \{a \in (0,1]^n : a_i = 0\}$ donde $i = 1, \ldots, n$, es μ_n-nulo, es decir, $\mu_n(Q_i) = 0$. Usando este hecho y la propiedad de que la medida μ_n es invariante por traslación, resulta que *cualquier subespacio de dimensión $n-1$ de $\mathbb{R}^n$, es decir, cualquier hiperplano, también es μ_n-nulo*. En efecto, si X es un tal subespacio de $\mathbb{R}^n$, entonces existe un $i \in \{1, \ldots, n\}$ tal que $X = \bigcup_{m \in \mathbb{Z}_i^n}(m + Q_i)$, donde $\mathbb{Z}_i^n = \{m \in \mathbb{Z}^n : m_i = 0\}$. Se sigue de esto que $\mu_n(X) = 0$. Otro modo de ver esto es identificar $\mathbb{R}^{n-1}$, visto como un subconjunto de $\mathbb{R}^n$, con el subespacio $\mathbb{R}^{n-1} \times \{0\}$. Puesto que éste último conjunto es un boreliano, resulta que

$$\mu_n(\mathbb{R}^{n-1}) = \mu_n\big(\mathbb{R}^{n-1} \times \{0\}\big) = \big[\mu(\mathbb{R})\big]^{n-1} \cdot \mu(\{0\}) = +\infty \cdot 0 = 0.$$

En general, uno puede demostrar que *cualquier subespacio propio de $\mathbb{R}^n$ posee medida (de Lebesgue) cero.*

Un hecho que es muy útil a la hora de verificar la unicidad de la medida de Borel es el siguiente.

Teorema 11.3.13. *Si $\lambda : \mathfrak{Bo}(\mathbb{R}^n) \to [0,+\infty]$ es una **medida σ-finita** tal que*

$$\lambda(Q) = \mu_n(Q) \quad \text{para todo} \quad Q \in \mathfrak{D}_n,$$

entonces $\lambda = \mu_n$ sobre $\mathfrak{Bo}(\mathbb{R}^n)$.

Prueba. Considere la familia

$$\mathcal{A} = \{E \subseteq \mathbb{R}^n : \lambda(E) = \mu_n(E)\}.$$

Nuestro objetivo será demostrar que $\mathfrak{Bo}(\mathbb{R}^n) \subseteq \mathcal{A}$. Sea $(\mathcal{C}_\alpha)_{\alpha<\omega_1}$ la colección de subconjuntos de $\mathbb{R}^n$ tal como fue definida en el Teorema 6.3.13, página 265, tal que

$$\mathfrak{Bo}(\mathbb{R}^n) = \sigma(\mathcal{C}) = \bigcup_{\alpha<\omega_1} \mathcal{C}_\alpha,$$

y donde $\mathcal{C} = \mathcal{O}_n$.

(a) $\mathcal{O}_n \subseteq \mathcal{A}$. Para ver esto, nótese que, por hipótesis, $\mathfrak{D}_n \subseteq \mathcal{A}$. Sea G un conjunto abierto en $\mathbb{R}^n$. Por el Teorema 11.3.1 existe una sucesión disjunta $(Q_j)_{j=1}^\infty$ en $\mathfrak{D}_n$ tal que $G = \bigcup_{j=1}^\infty Q_j$. Como λ es una medida, resulta que

$$\begin{aligned}\lambda(G) &= \lambda\Big(\bigcup_{j=1}^\infty Q_j\Big) = \sum_{j=1}^\infty \lambda(Q_j) = \sum_{j=1}^\infty \mu_n(Q_j)\\ &= \mu_n\Big(\bigcup_{j=1}^\infty Q_j\Big) = \mu_n(G).\end{aligned}$$

Esto prueba que $G \in \mathcal{A}$ y por lo tanto, $\mathcal{O}_n \subseteq \mathcal{A}$. Veamos ahora que $\mathcal{A}$ contiene cualquier conjunto cerrado. Para ver esto, sea $F \subseteq X$ un conjunto cerrado. Como λ es σ-finita, podemos asumir que E es acotado. En este caso, $\lambda(E) < +\infty$. Puesto que $G = F^c$ es abierto tenemos, por lo anterior, que $G = F^c \in \mathcal{A}$.

(b) Sea $\alpha < \omega_1$ y suponga que hemos demostrado que $\mathcal{C}_\beta \subseteq \mathcal{A}$ para $\beta < \alpha$. Si α es un ordinal límite, entonces $\mathcal{C}_\alpha = \bigcup_{\beta<\alpha} \mathcal{C}_\beta$, de donde se sigue que $\mathcal{C}_\alpha \subseteq \mathcal{A}$. Si α posee un predecesor inmediato, digamos β, entonces $\mathcal{D}_\alpha \subseteq \mathcal{A}$ y, en consecuencia, también $\mathcal{D}_\alpha^c \subseteq \mathcal{A}$. De esto se sigue que $\mathcal{C}_\alpha \subseteq \mathcal{A}$ y el Principio de Inducción Transfinita finaliza la prueba. ■

En lo que sigue, denotaremos por Q^1 el cubo unitario n-dimensional en $\mathbb{R}^n$, esto es,

$$Q^1 = [0,1]^n = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : 0 \leq x_i \leq 1,\ i = 1,2,\ldots,n\}.$$

El próximo resultado es crucial en la demostración de un teorema sobre el cambio de variable en $\mathbb{R}^n$. S. Saks fue el primero en preguntarse si cualquier medida de Borel σ-finita e invariante por traslación es un múltiplo de la medida de Lebesgue. La respuesta es sí siempre que la definición de medida de Borel sea la que hemos dado en estas notas.

Teorema 11.3.14 (Unicidad de la Medida de Lebesgue). *Sea $\lambda : \mathfrak{Bo}(\mathbb{R}^n) \to [0,+\infty]$ una medida **σ-finita, invariante por traslación** y tal que $\lambda(K) < +\infty$ para cualquier compacto $K \subseteq \mathbb{R}^n$. Entonces existe una constante $c \geq 0$ tal que*

$$\lambda(E) = c \cdot \mu_n(E), \tag{3}$$

para cualquier $E \in \mathfrak{Bo}(\mathbb{R}^n)$. En particular, si $a \in \mathbb{R}$, entonces

$$\mu_n(a \cdot E) = |a|^n \mu_n(E), \tag{4}$$

para todo $E \in \mathfrak{Bo}(\mathbb{R}^n)$.

Prueba. Como Q^1 es compacto, se tiene que $c = \lambda(Q^1)$ es un número real no-negativo. Observe que si $c = 0$, entonces $\lambda = 0$. En efecto, haciendo uso del hecho de que λ es invariante por traslación y de que

$$\mathbb{R}^n = \bigcup_{m \in \mathbb{Z}^n} (m + Q^1)$$

resulta que $\lambda(\mathbb{R}^n) = 0$ y, por lo tanto, $0 = \lambda(E) = 0 \cdot \mu_n(E)$ para todo $E \in \mathfrak{Bo}(\mathbb{R}^n)$. Este hecho permite que podamos suponer que $c > 0$. Considere ahora la familia

$$\mathcal{K} \;=\; \big\{E \in \mathfrak{Bo}(\mathbb{R}^n) \;:\; \lambda(E) = c \cdot \mu_n(E)\big\}.$$

Veamos que $\mathfrak{Bo}(\mathbb{R}^n) \subseteq \mathcal{K}$. Para ello, es suficiente demostrar que $\mathcal{K}$ contiene a $\mathfrak{D}_n$, la colección de los cubos semi-cerrados de $\mathbb{R}^n$.

(i) $\mathcal{K} \neq \varnothing$ ya que $\mu_n(Q^1) = 1$ y, por lo tanto, $Q^1 \in \mathcal{K}$.

(ii) Si $E \in \mathcal{K}$ y si existe un conjunto $F \in \mathfrak{Bo}(\mathbb{R}^n)$ tal que $E = \bigcup_{i=1}^n (q_i + F)$, donde los trasladados $q_1 + F, \ldots, q_n + F$ son disjuntos dos a dos, entonces $F \in \mathcal{K}$. En efecto,

$$\begin{aligned} n\big(c \cdot \mu_n(F)\big) \;&=\; c \cdot \sum_{i=1}^n \mu_n(q_i + F) \;=\; c \cdot \mu_n(E) \\ &=\; \lambda(E) \;=\; \sum_{i=1}^n \lambda(q_i + F) \;=\; \sum_{i=1}^n \lambda(F) \;=\; n \cdot \lambda(F) \end{aligned}$$

De allí que $c \cdot \mu_n(F) = \lambda(F)$ y, por lo tanto, $F \in \mathcal{K}$.

(iii) $\mathfrak{D}_n \subseteq \mathcal{K}$. En primer lugar, observe que si

$$H_i \;=\; \big\{(a_1, \ldots, a_n) \in \mathbb{R}^n \;:\; a_i = 0\big\},$$

para cada $i = 1, \ldots, n$, entonces $\lambda(H_i) = 0$ para cada $i = 1, \ldots, n$. Para ver esto, nótese que H_i se puede escribir en la forma

$$H_i \;=\; \bigcup_{m \in \mathbb{Z}_i^n} (m + Q_i) \tag{5}$$

donde $Q_i = \{a \in [0,1]^n : a_i = 0\}$ y $\mathbb{Z}_i^n = \{m \in \mathbb{Z}^n : m_i = 0\}$. Denote por $e_i = (0, \ldots, 0, 1, 0, \ldots)$ el vector en $\mathbb{R}^n$ cuya i-ésima coordenada es 1 y las restantes son 0's. Puesto que la sucesión $\mathcal{F} = \{e_i j^{-1} + Q_i : j \in \mathbb{N}\}$ es una colección disjunta tal que

$$\bigcup_{j=1}^{\infty} \left(\frac{e_i}{j} + Q_i\right) \;\subseteq\; Q^1$$

resulta, usando de nuevo el hecho de que λ es invariante por traslación, que

$$\sum_{j=1}^{\infty} \lambda(Q_i) \;\leq\; \lambda(Q^1) \;=\; c \;<\; \infty,$$

lo cual sólo es posible si $\lambda(Q_i) = 0$. Se sigue de (5) que $\lambda(H_i) = 0$ y, por consiguiente, $H_i \in \mathcal{K}$. En general, si H es cualquier hiperplano paralelo a los ejes de coordenadas, esto es, $H_x = x + H_i$

para cualquier $x \in \mathbb{R}^n$ y algún $i = 1, 2, \ldots, n$, entonces $\lambda(H_x) = 0$. De aquí que $H_x \in \mathcal{K}$ para cualquier $x \in \mathbb{R}^n$. Combinando estos resultados junto con (ii) vemos que $(0,1]^n \in \mathcal{K}$ ya que

$$(0,1]^n = [0,1]^n \setminus (H_1 \cup \cdots \cup H_n).$$

y, por lo tanto, todos los cubos semi-cerrados de la forma

$$Q_{x,k} = \prod_{i=1}^{n} \left(x_i,\, x_i + \frac{1}{2^k} \right]$$

pertenecen a $\mathcal{K}$ para cualquier $x \in \mathbb{R}^n$ y todo $k \in \mathbb{N}$. Finalmente, por el Teorema 11.3.13 tenemos que

$$\lambda(E) = c \cdot \mu_n(E) \quad \text{para todo} \quad E \in \mathfrak{Bo}(\mathbb{R}^n).$$

Esto termina la prueba de (3). Para demostrar (4), fijemos un $a \in \mathbb{R}$. Si $a = 0$, entonces $a \cdot E = \{0\}$ para todo $E \in \mathfrak{Bo}(\mathbb{R}^n)$ y (4) se cumple. Suponga entonces que $a \neq 0$ y defina $\lambda : \mathfrak{Bo}(\mathbb{R}^n) \to [0, +\infty]$ por

$$\lambda(E) = \mu_n(a \cdot E) \quad \text{para todo} \quad E \in \mathfrak{Bo}(\mathbb{R}^n).$$

Es fácil ver que λ es una medida. Veamos que ella es invariante por traslación. En efecto, para cada $x \in \mathbb{R}^n$ y cada $E \in \mathfrak{Bo}(\mathbb{R}^n)$, tenemos, usando el hecho de que μ_n es invariante por traslación, que

$$\lambda(x + E) = \mu_n(a \cdot (x + E)) = \mu_n(a \cdot x + a \cdot E) = \mu_n(a \cdot E) = \lambda(E).$$

Finalmente, observe que $a \cdot Q^1$ es un cubo cerrado cuyos lados tiene longitud $|a|$ y, por lo tanto, $\lambda(Q^1) = \mu_n(a \cdot Q^1) = |a|^n$. Por esto, si $K \subseteq \mathbb{R}^n$ es compacto, entonces existe un $x \in \mathbb{R}$ y una constante $k > 0$ tal que $K \subseteq x + kQ^1$ y, en consecuencia,

$$\lambda(K) \le \lambda(kQ^1) = \mu_n(ka \cdot Q^1) = |ka|^n < +\infty.$$

Un llamado a la primera parte nos revela que

$$\mu_n(a \cdot E) = \lambda(Q^1)\mu_n(E) = |a|^n \mu_n(E)$$

para todo $E \in \mathfrak{Bo}(\mathbb{R}^n)$ y termina la prueba. ■

El resultado anterior establece que, salvo constante, cualquier medida de Borel que es σ-finita y finita sobre compactos, es la medida de Lebesgue sobre $\mathfrak{Bo}(\mathbb{R}^n)$. Toda medida de Borel que es σ-finita y finita sobre compactos se le llama una **medida de Radon**.

Pongamos $\mu_n(U(0,1)) = 1$, donde

$$U(0,1) = \{x \in \mathbb{R}^n : \|x\|_2 < 1\}.$$

Si $U(x,r)$ es la bola abierta con centro en x y radio $r > 0$, entonces como $U(x,r) = x + r \cdot U(0,1)$, se tiene del teorema anterior que

$$\mu_n(U(x,r)) = \mu_n(r \cdot U(0,1)) = r^n.$$

Como una aplicación de la regularidad de la medida de Lebesgue μ_n en $\mathbb{R}^n$, consideremos el espacio de Banach $(\mathrm{L}_1(\mathbb{R}^n, \mathfrak{M}(\mathbb{R}^n), \mu_n), \|\cdot\|_1)$, al que denotaremos brevemente por $\mathrm{L}_1(\mathbb{R}^n)$, y veamos que las funciones continuas a soporte compacto definidas sobre $\mathbb{R}^n$ son densas en $\mathrm{L}_1(\mathbb{R}^n)$. Recordemos que una función continua $f : \mathbb{R}^n \to \mathbb{R}$ tiene **soporte compacto** si el conjunto

$$\mathrm{sop}(f) = \overline{\{x \in \mathbb{R}^n : f(x) \neq 0\}}$$

es compacto. Puesto que $(\overline{A})^c = \mathrm{int}(A^c)$, resulta que

$$\begin{aligned} x \notin \mathrm{sop}(f) \quad &\Leftrightarrow \quad x \in \mathrm{int}\{x \in \mathbb{R}^n : f(x) = 0\} \\ &\Leftrightarrow \quad \text{existe un entorno abierto } V \text{ de } x \text{ tal que } f|_V = 0. \end{aligned}$$

Denotemos por $C_c(\mathbb{R}^n)$ *el espacio vectorial de todas las funciones* $f : \mathbb{R}^n \to \mathbb{R}$ *que tienen soportes compactos*. Observe que $f \in C_c(\mathbb{R}^n)$ si, y sólo si, f se anula fuera de un conjunto acotado.

Teorema 11.3.15 (Aproximación en $\mathrm{L}_1(\mathbb{R}^n)$). *El espacio $C_c(\mathbb{R}^n)$ es* ***norma-denso*** *en $\mathrm{L}_1(\mathbb{R}^n)$. Explícitamente, dado $f \in \mathrm{L}_1(\mathbb{R}^n)$ y $\varepsilon > 0$, existe una función $g \in C_c(\mathbb{R}^n)$ tal que*

$$\| f - g \|_1 < \varepsilon.$$

Prueba. Suponga, en primer lugar, que $f = \chi_E$ para algún conjunto $E \in \mathfrak{M}(\mathbb{R}^n)$ y sea $\varepsilon > 0$. Por la regularidad de μ_n, Teorema 11.3.8, existe un conjunto abierto y acotado G y un conjunto compacto G tal que

$$K \subseteq E \subseteq G \qquad \text{y} \qquad \mu_n(G \setminus K) < \varepsilon.$$

El Lema de Urysohn, Teorema 3.1.11, página 159, nos garantiza entonces la existencia de una función continua $g \in C_c(\mathbb{R}^n)$ tal que

$$g(x) = 1 \quad \text{si} \quad x \in K \qquad \text{y} \qquad g(x) = 0 \quad \text{si} \quad x \in G^c.$$

De esto se sigue que

$$\{x \in \mathbb{R}^n : g(x) \neq \chi_E\} \subseteq G \setminus K$$

y, por lo tanto,

$$\| \chi_E - g \|_1 = \int_{G\setminus K} |\chi_E - g| \, d\mu_n \leq \mu_n(G \setminus K) < \varepsilon.$$

Sea $f \in \mathrm{L}_1(\mathbb{R}^n)$. Descomponiendo a f es su parte positiva y negativa: $f = f^+ - f^-$, podemos asumir, además, que f es no-negativa. En este caso, seleccione una sucesión creciente $(s_k)_{k=1}^{\infty}$ de funciones simples, no-negativas y con soportes compactos tal que $f(x) = \lim_{k\to\infty} s_k(x)$ para cada $x \in \mathbb{R}^n$. Escoja ahora un k lo suficientemente grande de modo que $|s_k(x) - f(x)| < \varepsilon$ para todo $x \in \mathbb{R}^n$. Por la primera parte podemos elegir una función $g \in C_c(\mathbb{R}^n)$ tal que $\| s_k - g \|_1 < \varepsilon/2$, de donde se sigue que

$$\| f - g \|_1 \leq \| f - s_k \|_1 + \| s_k - g \|_1 < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

y termina la prueba. ■

11.3.1. Cambio de Variable: caso lineal

El problema general que abordaremos en ésta y la siguiente sección consiste en lo siguiente: dado cualquier conjunto medible $E \subseteq \mathbb{R}^n$ y cualquier función integrable $f : E \to \mathbb{R}$, sabemos por experiencia que no siempre es fácil calcular su integral directamente; sin embrago, un cierto cambio de variable puede, en algunos casos, resolver dicho problema. En consecuencia, el método del cambio de variable, o difeomorfismo como también se le conoce, permite (cuando ello sea posible) hallar la integral sobre E pero *cambiando* a E por otra región B sobre la cual resulta más fácil integrar. Como un primer paso nos embarcaremos en un cambio de variable sencillo: el caso lineal.

En primer lugar, para cambiar a E por otro conjunto es preciso, en una primera fase, hacerlo a través de una transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^m$ que actúa sobre el conjunto medible E. Para ello necesitaremos recordar algunas de las herramientas básicas del Álgebra Lineal que vamos a requerir.

Sea $\|\cdot\|$ cualquier p-norma sobre $\mathbb{R}$ y sea $T : \mathbb{R}^n \to \mathbb{R}^n$ una transformación lineal arbitraria. Para cada $x \in \mathbb{R}^n$, escribiremos $T(x) = (T_1(x), \ldots, T_n(x))$, donde cada una de las funciones $T_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, n$, es una *aplicación lineal* a las que llamaremos las **coordenadas de** T. Recordemos que T es ***invertible*** si T^{-1} existe y es lineal. Del Álgebra Lineal sabemos que

(i) *T es inyectiva si, y sólo si, T es sobreyectiva*. Por consiguiente, si T es inyectiva, entonces ella es biyectiva y, en consecuencia, T^{-1} existe y es lineal, es decir, T es invertible.

(ii) *T siempre es continua*. Más aun, T ***es Lipschitz***, un resultado que fue demostrado en el Teorema 4.1.12, página 202.

Combinando los dos hechos anteriores vemos que:

(iii) Si T es inyectiva, entonces T^{-1} también es continua. Por esto, *toda transformación lineal invertible $T : \mathbb{R}^n \to \mathbb{R}^n$ es un homeomorfismo*. En consecuencia, T preserva cualquier propiedad topológica y, por supuesto, lineal del espacio $\mathbb{R}^n$. En particular, si E es abierto, (respectivamente, cerrado, compacto, F_σ, etc.), entonces $T(E)$ posee, respectivamente, la misma propiedad. Más aun, como veremos más abajo, si E es medible Lebesgue, también lo es $T(E)$.

Denote por $\mathrm{GL}(\mathbb{R}^n)$ el *grupo lineal de todas las* ***transformaciones lineales invertibles*** $T : \mathbb{R}^n \to \mathbb{R}^n$. Cada $T \in \mathrm{GL}(\mathbb{R}^n)$ se identifica, con respecto a la base canónica de $\mathbb{R}^n$, con una única matriz invertible $n \times n$ que designaremos por M_T.

(iv) $T \in \mathrm{GL}(\mathbb{R}^n)$ si, y sólo si, el determinante de M_T, $\det(M_T)$, es distinto de cero. Esto permite definir $\det(T) = \det(M_T)$. De las propiedades de los determinantes, sabemos que si $T = T_1 \circ T_2$, entonces

$$\det(T) \;=\; \det(T_1) \cdot \det(T_2).$$

La parte (a) del siguiente resultado ya fue demostrada cuando $T : \mathbb{R} \to \mathbb{R}$ era una función continua e inyectiva, Lema 6.3.22, página 272.

Teorema 11.3.16 (Medida de una transformación). *Sea $T \in \mathrm{GL}(\mathbb{R}^n)$. Entonces*

(a) $T(E) \in \mathfrak{Bo}(\mathbb{R}^n)$ para cualquier $E \in \mathfrak{Bo}(\mathbb{R}^n)$ y

(b) $\lambda : \mathfrak{Bo}(\mathbb{R}^n) \to [0, +\infty]$ definida por

$$\lambda(E) \;=\; (\mu_n \circ T)(E) \;=\; \mu_n(T(E)) \quad \textit{para todo } E \in \mathfrak{Bo}(\mathbb{R}^n)$$

es una medida la cual es ***invariante por traslación*** *y tal que $\lambda(K) < +\infty$ para cualquier conjunto compacto $K \subseteq \mathbb{R}^n$.*

Prueba. (a) Considere la familia

$$\mathcal{F} = \{E \subseteq \mathbb{R}^n : T(E) \in \mathfrak{Bo}(\mathbb{R}^n)\}$$

y veamos que $\mathfrak{Bo}(\mathbb{R}^n) \subseteq \mathcal{F}$. Puesto que $T(\mathbb{R}^n) = \mathbb{R}^n \in \mathfrak{Bo}(\mathbb{R}^n)$, se tiene que $\mathbb{R}^n \in \mathcal{F}$. Suponga ahora que $E \in \mathcal{F}$. Entonces $T(E) \in \mathfrak{Bo}(\mathbb{R}^n)$ y puesto que T es biyectiva, resulta que $T(E^c) = (T(E))^c \in \mathfrak{Bo}(\mathbb{R}^n)$, de modo que $E^c \in \mathcal{F}$. Esto prueba que $\mathcal{F}$ es cerrado bajo complementos. Si $(E_n)_{n=1}^{\infty}$ es una sucesión en $\mathcal{F}$, entonces $T(E_n) \in \mathfrak{Bo}(\mathbb{R}^n)$ para todo $n \geq 1$ y como $\mathfrak{Bo}(\mathbb{R}^n)$ es una σ-álgebra tenemos que $T(\bigcup_{n=1}^{\infty} E_n) = \bigcup_{n=1}^{\infty} T(E_n) \in \mathfrak{Bo}(\mathbb{R}^n)$. Esto nos dice que $\bigcup_{n=1}^{\infty} E_n \in \mathcal{F}$. Con esto queda establecido que $\mathcal{F}$ es una σ-álgebra.

Veamos ahora que $\mathcal{O}_n \subseteq \mathcal{F}$. En efecto, sea $G \in \mathcal{O}_n$. Como T es un homeomorfismo, resulta que $T(G)$ es abierto y, por lo tanto, $T(G) \in \mathfrak{Bo}(\mathbb{R}^n)$. Esto nos muestra que $G \in \mathcal{F}$ y así, $\mathcal{O}_n \subseteq \mathcal{F}$. Siendo $\mathfrak{Bo}(\mathbb{R}^n)$ la σ-álgebra más pequeña conteniendo a $\mathcal{O}_n$, se tiene que $\mathfrak{Bo}(\mathbb{R}^n) \subseteq \mathcal{F}$ y entonces $T(E) \in \mathfrak{Bo}(\mathbb{R}^n)$ para cualquier $E \in \mathfrak{Bo}(\mathbb{R}^n)$, es decir,

$$T(\mathfrak{Bo}(\mathbb{R}^n)) \subseteq \mathfrak{Bo}(\mathbb{R}^n).$$

(b) Puesto que $T(\varnothing) = \varnothing$, entonces $\lambda(\varnothing) = \mu_n(\varnothing) = 0$. Para completar la demostración de que λ es una medida, sea $(E_n)_{n=1}^{\infty}$ una sucesión disjunta en $\mathfrak{Bo}(\mathbb{R}^n)$. Por (a) y el hecho de que T es inyectiva tenemos que la sucesión $(T(E_n))_{n=1}^{\infty}$ es disjunta en $\mathfrak{Bo}(\mathbb{R}^n)$ y así,

$$\begin{aligned}\lambda\Big(\bigcup_{n=1}^{\infty} E_n\Big) &= \mu_n\Big(T\Big(\bigcup_{n=1}^{\infty} E_n\Big)\Big) = \mu_n\Big(\bigcup_{n=1}^{\infty} T(E_n)\Big) \\ &= \sum_{n=1}^{\infty} \mu_n(T(E_n)) = \sum_{n=1}^{\infty} \lambda(E_n),\end{aligned}$$

lo cual prueba que λ es numerablemente aditiva y, por consiguiente, λ es una medida. Para verificar que λ es invariante por traslación, sean $E \in \mathfrak{Bo}(\mathbb{R}^n)$ y $x \in \mathbb{R}^n$. Entonces, como que μ_n es invariante por traslación,

$$\lambda(x+E) = \mu_n(T(x+E)) = \mu_n(Tx + T(E)) = \mu_n(T(E)) = \lambda(E).$$

Finalmente, puesto que $T(K)$ es compacto para cualquier conjunto compacto $K \subseteq \mathbb{R}^n$, resulta entonces $\lambda(K) < +\infty$ ya que μ_n posee esa propiedad. ■

Fijemos $T \in \mathrm{GL}(\mathbb{R}^n)$. De las delicias del Álgebra Lineal sabemos que T se puede puede expresar en la forma

$$T = S_1 \circ S_2 \circ \cdots \circ S_k \tag{$*$}$$

donde cada transformación lineal $S_j \in \mathrm{GL}(\mathbb{R}^n)$ es una de los tres tipos siguientes:

1° $S(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) = (x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n)$ para ciertos $i, j = 1, \ldots, n;\ i < j$.

2° $S(x_1, x_2 \ldots, x_n) = (\alpha \cdot x_1, x_2 \ldots, x_n)$ para algún $\alpha \in \mathbb{R},\ \alpha \neq 0$.

3° $S(x_1, x_2 \ldots, x_n) = (x_1 + x_2, x_2 \ldots, x_n)$.

Para la transformación de tipo 1°, M_S es la matriz obtenida intercambiando la fila i por la fila j de la matriz identidad I_n y, por lo tanto, de la Teoría de los Determinantes sabemos que $\det(M_S) = \pm 1$.

Para la transformación de tipo 2°, M_S es la matriz obtenida multiplicando la primera fila de la matriz identidad I_n por α, de modo que $\det(M_S) = \alpha$.

Para la transformación de tipo 3°, M_S es la matriz obtenida añadiendo la segunda fila a la primera fila de la matriz identidad I_n y entonces se cumple que $\det(M_S) = \det(I_n) = 1$.

Teorema 11.3.17 (Medida de una transformación II). *Si $T \in \mathrm{GL}(\mathbb{R}^n)$, entonces,*

$$\mu_n(T(E)) \;=\; |\det(T)| \cdot \mu_n(E) \tag{6}$$

para todo $E \in \mathfrak{Bo}(\mathbb{R}^n)$.

Prueba. Por el Teorema 11.3.16 sabemos que $\lambda = \mu_n \circ T$ es una medida que es σ-finita, invariante por traslación y finita sobre conjuntos compactos y, entonces, gracias al Teorema 11.3.14 existe una constante $c > 0$ tal que $\mu_n(T(E)) = c \cdot \mu_n(E)$ para todo $E \in \mathfrak{Bo}(\mathbb{R}^n)$. Lo que resta demostrar es que $c = |\det(T)|$.

Nuestro primer paso será verificar que $c = |\det(S)|$ si S es de tipo $1^\circ, 2^\circ$ y 3°. Fijemos entonces $E \in \mathfrak{Bo}(\mathbb{R}^n)$ y escribámoslo en la forma $E = \prod_{i=1}^n B_i$ donde $B_i \in \mathfrak{Bo}(\mathbb{R})$ para todo $i = 1, 2, \ldots, n$.

(a) Si S es de tipo 1°, entonces $S(E) = B_1 \times \cdots \times B_j \times \cdots \times B_i \times \cdots B_n$ y, por lo tanto,

$$\mu_n(S(E)) \;=\; \mu(B_1) \cdot \ldots \cdot \mu(B_j) \cdot \ldots \cdot \mu(B_i) \cdot \ldots \cdot \mu(B_n) \;=\; c \cdot \mu_n(E)$$

Puesto que $\det(S) = \pm 1$, entonces $|\det(S)| = 1$ y así $c = |\det(S)|$ en este caso.

(b) Si S es de tipo 2°, entonces $S(E) = \alpha B_1 \times B_2 \times \cdots \times B_n$ y, por lo tanto,

$$\mu_n(S(E)) \;=\; \mu(\alpha B_1) \cdot \mu(B_2) \cdot \ldots \cdot \mu(B_n) \;=\; |\alpha| \mu(B_1) \cdot \mu(B_2) \cdot \ldots \cdot \mu(B_n) \;=\; |\alpha| \mu_n(E) \;=\; c \cdot \mu_n(E)$$

ya que $\mu(\alpha B_1) = |\alpha| \mu(B_1)$ gracias al Teorema 11.3.14 para $n = 1$. Puesto que $\det(S) = \alpha$ vemos de nuevo que $c = |\det(S)|$.

(c) Si S es de tipo 3°, entonces $S(E) = R \times B_3 \times \cdots \times B_n$, donde R es un rectángulo con la misma área que el rectángulo $B_1 \times B_2$ esto es, $\mu_2(R) = \mu(B_1)\mu(B_2)$. Por esto,

$$\mu_n(S(E)) \;=\; \mu_2(R) \cdot \mu(B_2) \cdot \mu(B_3) \cdot \ldots \cdot \mu(B_n) \;=\; \mu(B_1) \cdot \mu(B_2) \cdot \ldots \cdot \mu(B_n) \;=\; c \cdot \mu_n(E).$$

Puesto que $\det(S) = 1$, resulta que $c = |\det(S)|$ también en este caso. Para terminar la prueba, escriba T en la forma

$$T \;=\; S_1 \circ S_2 \circ \cdots \circ S_k$$

donde cada S_i es de tipo 1°, 2° y 3°. Por lo anterior y las propiedades del Determinante,

$$\begin{aligned} c \cdot \mu_n(E) &= \mu_n(T(E)) \\ &= \mu_n(S_1 \circ S_2 \circ \cdots \circ S_k(E)) \;=\; |\det(S_1)| \cdot \mu_n(S_2 \circ S_3 \circ \cdots \circ S_k(E)) \;=\; \cdots \\ &= |\det(S_1)| \cdot \ldots \cdot |\det(S_k)| \cdot \mu_n(E) \;=\; |\det(T)| \cdot \mu_n(E) \end{aligned}$$

de donde se obtiene que $c = \det(T)$ y termina la prueba. ■

Recordemos que el producto interno que usualmente se usa sobre $\mathbb{R}^n$ viene dado por

$$\langle x, y \rangle \;=\; \sum_{i=1}^n x_i y_i \quad \text{para} \quad x = (x_1, \ldots, x_b), \quad y = (y_1, \ldots, y_n) \text{ en } \mathbb{R}^n.$$

Definición 11.3.18. *Una transformación lineal* $T : \mathbb{R}^n \to \mathbb{R}^n$ *se llama* **ortogonal**, *si ella preserva el producto interno, esto es,*

$$\langle T(x), T(y)\rangle \;=\; \langle x, y\rangle \quad \text{para todo} \quad x, y \in \mathbb{R}^n.$$

Que T es ortogonal es equivalente a afirmar que $TT^* = I_n$, donde T^* es la transpuesta de T, lo cual se define por medio de la relación, $\langle T^*(x), y\rangle = \langle x, T(y)\rangle$ para todo $x, y \in \mathbb{R}^n$. Dicho de otro modo, T es ortogonal si T es invertible y $T^* = T^{-1}$. A una tal transformación también se le llama una ***rotación***. Observe que si T es una rotación, entonces $T \in \mathrm{GL}(\mathbb{R}^n)$ y puesto que $TT^* = I_n$ resulta que

$$\det(T) \cdot \det(T^*) \;=\; 1$$

y, en consecuencia, como $\det(T) = \det(T^*)$, se tiene que $\det(T) = \pm 1$. Se sigue entonces del resultado anterior que:

Corolario 11.3.19 (μ_n **es invariante por rotación).** μ_n *es invariante bajo* ***rotación***, *es decir, para cualquier* ***rotación*** $T \in \mathrm{GL}(\mathbb{R}^n)$, *se verifica que*

$$\mu_n(T(E)) \;=\; \mu_n(E)$$

para todo $E \in \mathfrak{Bo}(\mathbb{R}^n)$.

Éste resultado es el que nos va a permitir demostrar que cualquier subespacio propio de $\mathbb{R}^n$ posee medida de Lebesgue cero.

Corolario 11.3.20. *Si* X *es un* ***subespacio vectorial propio*** *de* $\mathbb{R}^n$, *entonces* $\mu_n(X) = 0$.

Prueba. Sea $k = \dim(X)$. Puesto que X es un subespacio propio de $\mathbb{R}^n$, resulta que $k < n$. Observe que si $k = 0$, entonces $X = \{0\}$ y, por supuesto, $\mu_n(X) = 0$. Suponga entonces que $1 \leq k < n$ y sea $\mathcal{B}_0 = \{x_1, \ldots, x_k\}$ una base ortonormal para X. Extienda dicho conjunto a una base ortonormal $\mathcal{B}_1 = \{x_1, \ldots, x_n\}$ para $\mathbb{R}^n$. Sea $\mathcal{B} = \{e_1, \ldots, e_n\}$ la base ortonormal estándar de $\mathbb{R}^n$ y considere la aplicación $T : \mathbb{R}^n \to \mathbb{R}^n$ definida por

$$T(x_i) \;=\; e_i, \quad i = 1, \ldots, n.$$

Entonces T es una transformación lineal ortogonal tal que $T(X) = \mathbb{R}^k$. Se sigue del Corolario 11.3.19 que

$$\mu_n(X) \;=\; \mu_n(T(X)) \;=\; \mu_n(\mathbb{R}^k).$$

Puesto que $\mathbb{R}^k \subseteq \mathbb{R}^n$, resulta que cada elemento $z \in \mathbb{R}^k$ se representa como $(z_1, \ldots, z_k, 0, \ldots, 0)$ y, en consecuencia, $\mathbb{R}^k$ se identifica con $\mathbb{R}^k \times \{0\}^{n-k}$. Por esto, y teniendo en cuenta que nuestro conjunto $\mathbb{R}^k \times \{0\}^{n-k} \in \mathfrak{Bo}(\mathbb{R}^n)$ se tiene que

$$\mu_n(\mathbb{R}^k) \;=\; \big[\mu(\mathbb{R})\big]^k \big[\mu(\{0\})\big]^{n-k} \;=\; +\infty \cdot 0 \;=\; 0.$$

Así, $\mu_n(X) = 0$ y termina la prueba. ■

Usando el corolario anterior podemos verificar, una vez más, que $\mathfrak{Bo}(\mathbb{R}^n) \subsetneqq \mathfrak{M}(\mathbb{R}^n)$. Es suficiente considerar el caso $n = 1$. Veamos esto. Sea $N = V \times \{0\} \subseteq \mathbb{R}^2$, donde V es un subconjunto no-medible en $\mathbb{R}$. Entonces N es un subconjunto del subespacio vectorial propio $X = \{(x, 0) : x \in \mathbb{R}\}$ y se sigue del Corolario 11.3.20 que $\mu_2(N) = 0$. Por esto, $N \in \mathfrak{M}(\mathbb{R}^2)$. Observe ahora que $N \notin \mathfrak{Bo}(\mathbb{R}^n)$. En efecto, si ello fuese verdad, entonces las todas y-secciones de N serían de Borel en $\mathbb{R}$ lo que resulta imposible ya que la 0-sección es V que no es medible según Borel.

Nota Adicional 11.3.9 Si $T : \mathbb{R}^n \to \mathbb{R}^n$ es una transformación que *no es invertible*, entonces $\det(T) = 0$. En particular, T no es sobreyectiva y, por consiguiente, $T(\mathbb{R}^n)$ es un subespacio propio de $\mathbb{R}^n$. Se sigue del Corolario 11.3.20 que $\mu_n(T(\mathbb{R}^n)) = 0$ y así, $\mu_n(T(E)) = 0$ para cualquier conjunto medible Borel $E \subseteq \mathbb{R}^n$. Por esto,

$$\mu_n(T(E)) \;=\; |\det(T)| \cdot \mu_n(E) \tag{6}$$

para cualquier conjunto $E \in \mathfrak{Bo}(\mathbb{R}^n)$. Esto prueba que la conclusión del Teorema 11.3.17 es válida no sólo para transformaciones invertibles sino para cualquier transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^n$. Sin embargo, debemos advertir que si bien la igualdad (6) es idénticamente igual a cero para cualquier conjunto $E \in \mathfrak{Bo}(\mathbb{R}^n)$, aun no sabemos si $T(E)$ es un miembro de $\mathfrak{Bo}(\mathbb{R}^n)$. La única conclusión a la que podemos llegar es que $\mu_n(T(E)) = 0$ y, por lo tanto, $T(E) \in \mathfrak{M}(\mathbb{R}^n)$ gracias al hecho de que μ_n es una medida completa. Por esta razón, muchos autores consideran conveniente trabajar con la colección

$$\mathfrak{Bo}(T) \;=\; \left\{T^{-1}(B) \;:\; B \in \mathfrak{Bo}(\mathbb{R}^n)\right\}$$

la cual es fácil ver es una σ-álgebra en $\mathbb{R}^n$.

Lo que resulta interesante es que la igualdad (6) sigue siendo válida para todo conjunto $E \in \mathfrak{M}(\mathbb{R}^n)$ como se muestra a continuación.

Teorema 11.3.21 (Medida de una Transformacion III). *Si $T : \mathbb{R}^n \to \mathbb{R}^n$ es una **transformación lineal** arbitraria, entonces para todo conjunto $E \in \mathfrak{M}(\mathbb{R}^n)$, se cumple que $T(E) \in \mathfrak{M}(\mathbb{R}^n)$ y*

$$\mu_n(T(E)) \;=\; |\det(T)| \cdot \mu_n(E). \tag{6}$$

*En particular, si T es **ortogonal**, entonces $\mu_n(T(E)) = \mu_n(E)$ para todo $E \in \mathfrak{M}(\mathbb{R}^n)$.*

Prueba. Suponga que $T \in \mathrm{GL}(\mathbb{R}^n)$ y sea $E \in \mathfrak{M}(\mathbb{R}^n)$. Por el Corolario 11.3.7 sabemos que E se puede escribir en la forma $E = B \cup D$, donde $B \cap D = \varnothing$, $B \in \mathfrak{Bo}(\mathbb{R}^n)$ y D es un subconjunto de un conjunto $N \in \mathfrak{Bo}(\mathbb{R}^n)$ con $\mu_n(N) = 0$. Como $B \in \mathfrak{Bo}(\mathbb{R}^n)$, se sigue del Teorema 11.3.16 que $T(B) \in \mathfrak{Bo}(\mathbb{R}^n)$ y entonces, por el Teorema 11.3.17 tenemos que

$$\mu_n(T(B)) \;=\; |\det(T)| \cdot \mu_n(B).$$

Similarmente, $T(N) \in \mathfrak{Bo}(\mathbb{R}^n)$ y

$$\mu_n(T(N)) = |\det(T)| \cdot \mu_n(N) = 0.$$

Ahora bien, puesto que $T(E) = T(B) \cup T(D)$ donde $T(B) \in \mathfrak{Bo}(\mathbb{R}^n)$ y $T(D)$ es un subconjunto de $T(N) \in \mathfrak{Bo}(\mathbb{R}^n)$ con $\mu_n(T(N)) = 0$, tenemos, por el hecho de ser μ_n una medida completa, que $T(D) \in \mathfrak{M}(\mathbb{R}^n)$ y $\mu_n(T(D)) = 0$. De aquí se sigue que $T(E) \in \mathfrak{M}(\mathbb{R}^n)$ y, en consecuencia,

$$\begin{aligned}
\mu_n(T(E)) \;&=\; \mu_n(T(B)) + \mu_n(T(D)) \\
&=\; |\det(T)| \cdot \mu_n(B) + |\det(T)| \cdot \mu_n(D) \\
&=\; |\det(T)| \cdot \mu_n(E).
\end{aligned}$$

Esto prueba la igualdad (6) si T es invertible. Si T no es invertible, entonces $\det(T) = 0$ y el argumento dado anteriormente nos revela que $\mu_n(T(E)) = 0$ para cualquier $E \in \mathfrak{M}(\mathbb{R}^n)$ y, además, que $T(E) \in \mathfrak{M}(\mathbb{R}^n)$. Por esto,

$$\mu_n(T(E)) \;=\; |\det(T)| \cdot \mu_n(E) \;=\; 0$$

para todo $E \in \mathfrak{M}(\mathbb{R}^n)$ y termina la prueba. ■

Ejemplo 11.3.1. *Suponga que ya sabemos que el área de un disco de radio r en el plano es πr^2 y se quiere calcular el área de una elipse cuyos ejes son a y b, con $a > 0$, $b > 0$. Para calcular tal área procedemos del modo siguiente: sean*

$$E = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\} \qquad y \qquad F = \Big\{(x,y) \in \mathbb{R}^2 : \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1\Big\}.$$

Si consideramos la transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T(x,y) = (a\,x, \, b\,y)$ resulta que $T(E) = F$. Puesto que la matriz asociada a T es simplemente

$$M_T = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix},$$

entonces $\det(M_T) = ab$, de donde se sigue, usando resultado anterior, que

$$\mu_2(F) = \mu_2(T(E)) = |\det(T)|\mu_2(E) = \pi \cdot ab.$$

Nota Adicional 11.3.10 Sea a un número real distinto de cero. Una transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^n$ de la forma $T(x) = a \cdot x$ para todo $x \in \mathbb{R}^n$ es llamada una ***dilatación***. Observe que su matriz M_T es igual a $a \cdot I_n$ por lo que $\det(M_T) = a^n \neq 0$. Se sigue de esto que T es invertible, es decir, $T \in \mathrm{GL}(\mathbb{R}^n)$. Un llamado al Teorema 11.3.21 nos muestra que

$$\mu_n(a \cdot E) = \mu_n(T(E)) = |\det(T)| \cdot \mu_n(E) = |a|^n \cdot \mu_n(E)$$

para todo $E \in \mathfrak{M}(\mathbb{R}^n)$. Este hecho ya lo habíamos demostrado en el Teorema 11.3.14, página 743.

El Teorema del Cambio de Variable dado en el Teorema 10.2.53 puede ser tratado de un modo particular en $\mathbb{R}^n$ como sigue

Teorema 11.3.22 (Cambio de variables: caso lineal). *Sea $T \in \mathrm{GL}(\mathbb{R}^n)$ y sea $f : \mathbb{R}^n \to \mathbb{R}$ una **función medible Lebesgue no-negativa** (respectivamente $f \in \mathrm{L}_1(\mathbb{R}^n, \mu_n)$). Entonces $f \circ T$ es **medible Lebesgue** (respectivamente $f \circ T \in \mathrm{L}_1(\mathbb{R}^n, \mu_n)$) y se cumple que*

$$\int_{\mathbb{R}^n} (f \circ T)\, d\mu_n = \frac{1}{|\det(T)|} \cdot \int_{\mathbb{R}^n} f\, d\mu_n. \tag{7}$$

Prueba. Para demostrar que $f \circ T$ es medible, sea G un subconjunto abierto de $\mathbb{R}$. Como f es medible, $f^{-1}(G)$ es medible Lebesgue y entonces

$$(f \circ T)^{-1}(G) = T^{-1}\big(f^{-1}(G)\big)$$

es, por el Teorema 11.3.21, medible Lebesgue ya que $T^{-1} \in \mathrm{GL}(\mathbb{R}^n)$. Para demostrar la igualdad (7), suponga que f es la función característica de un conjunto medible $E \in \mathfrak{M}(\mathbb{R}^n)$. En virtud de la fórmula

$$\big(\chi_E \circ T\big)(x) = \chi_{T^{-1}(E)}(x)$$

y el Teorema 11.3.21 se tiene que

$$\begin{aligned}\int_{\mathbb{R}^n} (f \circ T)\, d\mu_n &= \int_{\mathbb{R}^n} (\chi_E \circ T)\, d\mu_n = \int_{\mathbb{R}^n} \chi_{T^{-1}(E)}\, d\mu_n \\ &= \mu_n(T^{-1}(E)) = |\det(T^{-1})| \cdot \mu(E) \\ &= \frac{1}{|\det(T)|} \cdot \int_{\mathbb{R}^n} \chi_E\, d\mu_n = \frac{1}{|\det(T)|} \cdot \int_{\mathbb{R}^n} f\, d\mu_n.\end{aligned}$$

De esto y la linealidad de la integral se sigue que si f es una función simple, digamos $f = \sum_{j=1}^{k} a_j \cdot \chi_{E_j}$, entonces

$$\int_{\mathbb{R}^n} f\, d\mu_n = \sum_{j=1}^{k} a_j \cdot \int_{\mathbb{R}^n} \chi_{E_j}\, d\mu_n = |\det(T)| \sum_{j=1}^{k} a_j \cdot \int_{\mathbb{R}^n} \chi_{E_j} \circ T\, d\mu_n = |\det(T)| \cdot \int_{\mathbb{R}^n} (f \circ T)\, d\mu_n.$$

Suponga ahora que $f \geq 0$ sobre $\mathbb{R}^n$. Por el Teorema de Aproximación, existe una sucesión creciente de funciones simples no-negativas, digamos $(\varphi_k)_{k=1}^{\infty}$ convergiendo puntualmente a f. Usemos ahora el Teorema de la Convergencia Monótona para obtener

$$\begin{aligned}\int_{\mathbb{R}^n} (f \circ T)\, d\mu_n &= \lim_{k\to\infty} \int_{\mathbb{R}^n} (\varphi_k \circ T)\, d\mu_n \\ &= \frac{1}{|\det(T)|} \cdot \lim_{k\to\infty} \int_{\mathbb{R}^n} \varphi_k\, d\mu_n \\ &= \frac{1}{|\det(T)|} \cdot \int_{\mathbb{R}^n} f\, d\mu_n.\end{aligned}$$

Finalmente, si f es arbitraria, entonces $f = f^+ - f^-$ de donde resulta, usando de nuevo la linealidad de la integral, que (7) se cumple. ■

11.3.2. Cambio de Variable: caso no-lineal

Para obtener un Teorema del Cambio de Variable más general que el anterior, es decir, para transformaciones que no son necesariamente lineales, debemos abordar algunas nociones y resultados (que asumiremos el lector conoce) sobre funciones $f : \mathbb{R}^n \to \mathbb{R}^m$ que son diferenciables sobre un conjunto abierto $U \subseteq \mathbb{R}^n$.

A pesar de que todas las p-normas sobre $\mathbb{R}$ son equivalentes, $1 \leq p \leq +\infty$, la norma que consideraremos en esta sección sobre $\mathbb{R}^n$ es la norma del supremo $\|\cdot\|_\infty$, la cual escribiremos brevemente como $\|\cdot\|$, es decir,

$$\|x\| = \text{máx}\{|x_1|, \ldots, |x_n|\}, \quad x = (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

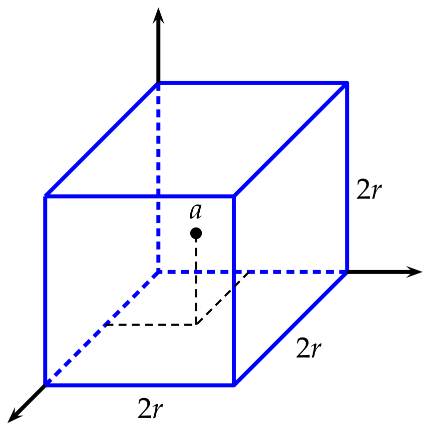
La razón para elegir ésta norma es muy simple: toda **bola cerrada** con centro $a = (a_1, \ldots, a_n) \in \mathbb{R}$ y radio $r > 0$:

$$B_\infty(a, r) = \{x \in \mathbb{R}^n : \|x - a\|_\infty \leq r\}$$

es un ρ-cubo, donde $\rho = 2r$, esto es,

$$B_\infty(a, r) = [a_1 - r, a_1 + r] \times \cdots \times [a_n - r, a_n + r]$$

y, por lo tanto, su volumen viene dado por $\mu_n(B_\infty(a,r)) = (2r)^n$.



$$B_\infty(a,r) = \{x \in \mathbb{R}^n : \| x - a \|_\infty \leq r\}$$

Otra de las ventajas al considerar la norma $\| \cdot \|_\infty$ en $\mathbb{R}^n$ es que la **longitud** de cualquier diagonal D en $B_\infty(a,r)$ es siempre $2r$, esto es:

$$\ell(D) = 2r$$

mientras que si medimos la misma diagonal, por ejemplo, con la norma euclídea $\| \cdot \|_2$, entonces

$$\ell(D) = \sqrt{n}\,2r.$$

La base canónica usual en $\mathbb{R}^n$ será denotada por $\mathcal{B}_n = \{e_1, \ldots, e_n\}$, esto es, para cada $i = 1, \ldots, n$, $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$, donde 1 ocurre en el i-ésimo lugar y todas las restantes componentes son 0. Sin embargo, si trabajamos simultáneamente con los espacios $\mathbb{R}^n$ y $\mathbb{R}^m$, entonces para evitar confusiones escribiremos $\{e'_1, \ldots, e'_m\}$ para la base estándar de $\mathbb{R}^m$.

Recordemos que toda transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^m$ es continua, Teorema 4.1.12, página 202. Denote por $\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ el espacio vectorial real de todas las transformaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^m$. Si $T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ y si $(a_{ij})_{i,j}$ es la matriz asociada a T en sus bases canónicas, entonces definimos

$$\| T \| = \max_{1 \leq i \leq 1} \sum_{j=1}^{m} |a_{ij}|. \tag{1}$$

Resulta que $\| \cdot \|$ es una norma sobre $\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ verificándose que

$$\| T \| = \sup\{\| T(x) \| : \| x \| \leq 1\}$$

para cada $T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$. Además, si $T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$, se cumple entonces que

$$\| T(x) \| \leq \| T \| \, \| x \| \quad \text{para todo } x \in \mathbb{R}^n.$$

Denote por $\mathrm{M}_n(\mathbb{R})$ el conjunto de todas las matrices $n \times n$ con entradas reales. De la Teoría de los Determinantes, sabemos que:

(a) $\det(AB) = \det(A) \cdot \det(B)$ para todo $A, B \in \mathrm{M}_n(\mathbb{R})$.

(b) $\det(A) \neq 0$ si, y sólo si, A es invertible.

(c) $\det(A)$ es un polinomio en las componentes de A y

(d) $\det(A) = \det(A^t)$, donde A^t es la transpuesta de A.

Fijaremos, a partir de este momento y permanentemente en esta sección, un par de subconjuntos abiertos no vacíos U, V en $\mathbb{R}^n$.

Definición 11.3.23. *Sea* $f : U \to \mathbb{R}^m$ *una función. Diremos que* f *es* **Fréchet diferenciable***, o simplemente,* **diferenciable** *en* $x \in U$*, si existe una* transformación lineal *única* $T : \mathbb{R}^n \to \mathbb{R}^m$ *tal que*

$$\lim_{h \to 0} \frac{\| f(x+h) - f(x) - T(h) \|}{\| h \|} = 0.$$

A T se le llama la ***derivada de*** f ***en*** x y se designa por $D_x f$ o $Df(x)$. Observe que si f es diferenciable en $x \in U$, entonces $D_x f : \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal cuya matriz asociada M_T siempre será representada en las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$. A esta matriz se le denomina ***matriz Jacobiana*** de f en x y se indicará por $J_f(x)$. Cuando $m = n$, el determinante de $J_f(x)$, que denotaremos por $\det(D_x f)$ o $|J_f(x)|$, se llama el ***determinante Jacobiano*** de f en x. En este caso es importante recordar que:

$$D_x f \ \text{ es } \textbf{\textit{invertible}} \text{ si, y sólo si, } \ |J_f(x)| \neq 0.$$

En general, diremos que f es diferenciable en U si ella es diferenciable en cada punto de U. Si $U_0 = \{x \in U : f \text{ es diferenciable en } x\}$ entonces la aplicación $Df : U_0 \to \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ que asigna a cada $x \in U_0$ la transformación lineal $Df(x) = D_x f$, se llama la ***diferencial*** de f. Las ***derivadas parciales*** asociadas a f son definidas por

$$\frac{\partial f_i}{\partial x_j}(x) = \frac{d}{dt} f_i(x + te_j)|_{t=0}.$$

Nótese que si $f : U \to \mathbb{R}^m$ es una función, entonces las siguientes condiciones son equivalentes:

(d_1) f es *diferenciable* en $x \in U$.

(d_2) Dado $\varepsilon > 0$, existe un $\delta > 0$ tal que

$$\| f(y) - f(x) - D_x f(y - x) \| < \varepsilon \| y - x \|$$

para cualquier $y \in U$ que satisfaga $0 < \| y - x \| < \delta$.

Varios hechos conocidos sobre la teoría de diferenciabilidad en $\mathbb{R}^n$ se muestran a continuación:

(1) Si f es *diferenciable* en $x \in U$, entonces f es *continua* en x.

(2) Si $f : \mathbb{R}^n \to \mathbb{R}^m$ es una *transformación lineal*, entonces $D_x f = f$

(3) $f = (f_1, \ldots, f_m)$ es *diferenciable* en $x \in U$ si, y sólo si, cada $f_i : U \to \mathbb{R}$ es *diferenciable* en x para todo $i = 1, \ldots, m$. En este caso,

$$D_x f = \begin{pmatrix} D_x f_1 \\ D_x f_2 \\ \vdots \\ D_x f_m \end{pmatrix}$$

donde, para cada $i = 1, \ldots, m$, $D_x f_i$ es la matriz $1 \times n$ dada por

$$D_x f_i = \begin{pmatrix} \dfrac{\partial f_i(x)}{\partial x_1} & \dfrac{\partial f_i(x)}{\partial x_2} & \cdots & \dfrac{\partial f_i(x)}{\partial x_n} \end{pmatrix}$$

(4) Si f es *diferenciable* en $x \in U$, entonces $\partial f_i(x)/\partial x_j$ existe para todo $i = 1, \ldots, m$ y todo $j = 1, \ldots, n$. Además,

$$J_f(x) = \begin{pmatrix} \dfrac{\partial f_1(x)}{\partial x_1} & \dfrac{\partial f_1(x)}{\partial x_2} & \cdots & \dfrac{\partial f_1(x)}{\partial x_n} \\ \dfrac{\partial f_2(x)}{\partial x_1} & \dfrac{\partial f_2(x)}{\partial x_2} & \cdots & \dfrac{\partial f_2(x)}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial f_m(x)}{\partial x_1} & \dfrac{\partial f_m(x)}{\partial x_2} & \cdots & \dfrac{\partial f_m(x)}{\partial x_n} \end{pmatrix}$$

Se sigue de nuestra definición de la norma de una transformación lineal que

$$\| D_x f \| = \max_{1 \le i \le n} \sum_{j=1}^{m} \left| \frac{\partial f_i(x)}{\partial x_j} \right|.$$

Observe que *si* $f : U \to \mathbb{R}^n$ *es de clase* C^1, entonces para todo $i = 1, \ldots, m$ y todo $j = 1, \ldots, n$ las derivadas parciales $\partial f_i(x)/\partial x_j$ existen y son continuas en cada punto $x \in U$. Se sigue de esto y la propiedad (c) de los determinantes que

$\det(D_x\varphi)$ *es un una* ***función continua*** *y, en consecuencia,* ***medible Lebesgue***.

Recíprocamente, si f es *continua* y si sus *derivadas parciales* $\partial f_i/\partial x_j$ *existen* y *son continuas* en cada punto $x \in U$, entonces f es *diferenciable* en cada $x \in U$.

Otros dos resultados que serán recordados son los siguientes:

Regla de la Cadena *Si* $f : U \to \mathbb{R}^n$ *es diferenciable en* $x \in U$ *y* $g : f(U) \to \mathbb{R}^n$ *es diferenciable en* $f(x)$, *entonces* $g \circ f$ *es diferenciable en* x, *y se cumple*

$$D_x(g \circ f) = D_{f(x)} g \circ D_x f.$$

Teorema del Valor Medio. *Sea* $f : U \to \mathbb{R}$ *una* ***función diferenciable*** *en* U. *Si* $x, y \in U$ *son tales que el segmento* $[x, y]$ *está totalmente contenido en* U, *entonces existe un punto* c *en el interior de* $[x, y]$ *tal que*

$$f(y) - f(x) = f'(c)(y - x),$$

o, equivalentemente,

$$f_i(y) - f_i(x) = \sum_{j=1}^{n} \frac{\partial f_i(c)}{\partial x_j} \cdot (y_i - x_i), \qquad i = 1, \ldots, m.$$

Observe que la igualdad obtenida en la conclusión del Teorema del Valor Medio es equivalente a

$$\begin{pmatrix} f_1(y)-f_1(x) \\ f_2(y)-f_2(x) \\ \vdots \\ f_m(y)-f_m(x) \end{pmatrix} = \begin{pmatrix} \dfrac{\partial f_1(c)}{\partial x_1} & \dfrac{\partial f_1(c)}{\partial x_2} & \cdots & \dfrac{\partial f_1(c)}{\partial x_n} \\ \dfrac{\partial f_2(c)}{\partial x_1} & \dfrac{\partial f_2(c)}{\partial x_2} & \cdots & \dfrac{\partial f_2(c)}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial f_m(c)}{\partial x_1} & \dfrac{\partial f_m(c)}{\partial x_2} & \cdots & \dfrac{\partial f_m(c)}{\partial x_n} \end{pmatrix} \cdot \begin{pmatrix} y_1-x_1 \\ y_2-x_2 \\ \vdots \\ y_n-x_n \end{pmatrix},$$

de modo que, para cada $i=1,\dots,m$,

$$f_i(y)-f_i(x) \;=\; \sum_{j=1}^{n} \frac{\partial f_i(c)}{\partial x_j} \cdot (y_j - x_j).$$

Nótese que si $f : U \to \mathbb{R}^n$ es una función de clase C^1 sobre U, entonces para cualquier compacto $Q \subseteq U$ se cumple $Df(Q)$ es un subconjunto acotado en $\mathrm{L}(\mathbb{R}^n, \mathbb{R}^n)$, es decir, existe una constante $K>0$ tal que

$$\sup_{x\in Q} \| D_x f \| \;\leq\; K.$$

Una consecuencia del Teorema del Valor Medio es que toda función diferenciable $f : U \to \mathbb{R}^n$ cuya diferencial es acotada en un compacto Q, es Lipschitz, un hecho que ya habíamos demostrado cuando el dominio era un intervalo de $\mathbb{R}$.

Corolario 11.3.24. *Sea $f : U \to \mathbb{R}^n$ una **función diferenciable**. Si Q es un cubo cerrado contenido en U y si existe una constante $K>0$ tal que $\sup_{x\in Q} \| D_x f \| = K$, entonces f es K-Lipschitz sobre Q, es decir,*

$$\| f(y) - f(x) \| \;\leq\; K \cdot \| y - x \|, \tag{1a}$$

*para todo $x, y \in Q$. Más aun, si f es de clase C^1, entonces f es **uniformemente diferenciable** sobre Q, esto es, dado $\varepsilon > 0$, existe un $\delta > 0$ tal que*

$$\| f(y) - f(x) - D_x f(x-y) \| \;\leq\; \varepsilon \cdot \| x - y \| \tag{1b}$$

para todo $x, y \in Q_\rho$ con $0 < \| x - y \| < \delta$.

Prueba. Observe que como Q es un cubo, él es convexo y, en consecuencia, el segmento $[x,y]$ permanece dentro de Q para cada par de puntos $x, y \in Q$. Por el Teorema del Valor Medio aplicado a cada una de las componentes f_i de f, existen $z_1, \dots, z_n$ en el interior de $[x,y]$ tales que

$$f_i(y) - f_i(x) \;=\; \sum_{j=1}^{n} \frac{\partial f_i(z_i)}{\partial x_j} \cdot (y_j - x_j), \qquad i = 1, \dots, n.$$

Usando las propiedades de la norma, resulta que

$$\begin{aligned} |f_i(y) - f_i(x)| &\leq \sum_{j=1}^{n} \left| \frac{\partial f_i(z_i)}{\partial x_j} \right| \cdot |(y_j - x_j)| \\ &\leq \| y - x \| \cdot \sum_{j=1}^{n} \left| \frac{\partial f_i(z_i)}{\partial x_j} \right| \\ &\leq \| y - x \| \cdot \| D_{z_i} f \| \leq K \cdot \| y - x \| . \end{aligned}$$

para cada $i = 1, \ldots, m,$ y así,

$$\| f(y) - f(x) \| = \text{máx}\{|f_1(y) - f_1(x)|, \ldots, |f_n(y) - f_n(x)|\} \leq K \| y - x \| ,$$

para todo $x, y \in Q$.

Para demostrar la segunda parte, suponga que f es de clase C^1. Esto, por supuesto, nos indica que $Df : U \to \mathrm{L}(\mathbb{R}^n, \mathbb{R}^n)$ es una función continua y como Q es compacto, resulta que Df es uniformemente continua sobre Q, es decir, dado $\varepsilon > 0$, existe un $\delta > 0$ tal que

$$\left\| D_x f - D_y f \right\| \leq \varepsilon$$

para todo $x, y \in Q$ que satisfaga $0 < \| x - y \| < \delta$. Fijemos un par de puntos $x, y \in Q$ tales que $0 < \| x - y \| < \delta$. Por el Teorema del Valor Medio, para cada $i = 1, \ldots, n,$ existe un punto z_i en el interior del segmento $[x, y]$ tal que

$$f_i(y) - f_i(x) = D_{z_i} f_i(y - x)$$

y entonces

$$f_i(y) - f_i(x) - D_x f_i(y - x) = \big(D_{z_i} f_i - D_x f_i\big)(y - x).$$

De aquí se obtiene que

$$\begin{aligned} \| f(y) - f(x) - D_x f(y - x) \| &= \text{máx}\left\{ \big(D_{z_i} f_i - D_x f_i\big)(y - x) \,:\, 1 \leq i \leq n \right\} \\ &\leq \text{máx}\left\{ \| D_{z_i} f - D_x f \| \, \| y - x \| \,:\, 1 \leq i \leq n \right\} \\ &\leq \varepsilon \cdot \| y - x \| \end{aligned}$$

y, por lo tanto, $(1b)$ se cumple siempre que $0 < \| y - x \| < \delta$. Esto termina la prueba. ∎

Definición 11.3.25. *Una función $f : U \to \mathbb{R}^m$ es de* **clase** C^1 *sobre U si f es diferenciable en cada $x \in U$ y su aplicación diferencial $Df : U \to L(\mathbb{R}^n, \mathbb{R}^m)$ es continua sobre U.*

Un hecho importante y muy conocido es el siguiente:

Teorema de la Aplicación Abierta. *Sea $f : U \to \mathbb{R}^n$ una **aplicación inyectiva de clase** C^1 **sobre** U. Si $\det(D_x \varphi) \neq 0$ para todo $x \in U$, entonces f es una* **aplicación abierta**, *es decir, si G es un abierto en U, entonces $f(G)$ es abierto en $f(U)$.*

La noción de difeomorfismo, que recordaremos de inmediato, permite que podamos hacer cambios de variables en integrales de funciones que poseen esa propiedad.

Definición 11.3.26. *Una función $f : U \to V$ se dice que es un* **difeomorfismo** *si*

(a) f *es biyectiva.*

(b) f *es diferenciable sobre* U.

(c) f^{-1} *es diferenciable sobre* V.

Es claro que todo difeomorfismo es un homeomorfismo. Un ***difeomorfismo*** $f : U \to V$ ***es de clase*** C^1 si f y f^{-1} son ambas de clase C^1 sobre U y V respectivamente. En este caso también se dice que φ es un **cambio de variable**. Nótese que:

Si $f : U \to V$ *es un* ***difeomorfismo de clase*** C^1, *entonces* $\det(D_x\varphi) \neq 0$ *para todo* $x \in U$.

Para ver esto, usemos la relación $(f^{-1} \circ f)(x) = x$ y entonces la Regla de la Cadena para obtener $I = D_x(f^{-1} \circ f) = D_{f(x)}f^{-1} \circ D_x f$, lo cual nos indica que $D_x f$ es invertible y, en consecuencia, $\det(D_x\varphi) \neq 0$ para todo $x \in U$. Además, como también $\det(D_x\varphi)$ es una función continua sobre U, resulta que

$$\int_E |\det(D_x\varphi)|\, d\mu_n(x)$$

siempre existe como un elemento de $\overline{\mathbb{R}}$ para cualquier conjunto medible $E \subseteq U$.

El siguiente resultado es una pieza clave para demostrar el Teorema del Cambio de Variable en su versión no-lineal.

Lema 11.3.27 (Desigualdad de Sard). *Si* $\varphi : U \to V$ *es un* ***difeomorfismo de clase*** C^1, *entonces*

$$\mu_n(\varphi(E)) \;\leq\; \int_E |\det(D_x\varphi)|\, d\mu_n(x) \tag{1c}$$

para todo conjunto medible $E \subseteq U$.

Prueba. Sea $Q = B_\infty(a, r)$ una bola cerrada con centro en $a = (a_1, \ldots, a_n) \in U$ y radio $r > 0$ incluida en U, es decir, Q es un ρ-cubo. Por el Corolario 11.3.24 $(1a)$ sabemos que para todo $x \in Q$,

$$\|\, \varphi(x) - \varphi(a)\, \| \;\leq\; \sup_{x\in Q} \|\, D_x\varphi\, \|\, \|\, x - a\, \| \;\leq\; r \cdot \sup_{x\in Q} \|\, D_x\varphi\, \|\,.$$

Esto muestra que $\varphi(Q)$ está contenido en el cubo

$$Q^* \;=\; \Big\{ y \in \mathbb{R}^n \;:\; \|\, y - \varphi(a)\, \| \;\leq\; r \cdot \sup_{x\in Q} \|\, D_x\varphi\, \| \Big\},$$

de donde resulta que

$$\begin{aligned}
\mu_n(\varphi(Q)) \;&\leq\; \mu_n(Q^*) \\
&=\; \big(2r \cdot \sup_{x\in Q} \|\, D_x\varphi\, \|\big)^n \\
&=\; \big(\sup_{x\in Q} \|\, D_x\varphi\, \|\big)^n \cdot \mu_n(Q).
\end{aligned}$$

Si $T \in \mathrm{GL}(\mathbb{R}^n)$, entonces podemos aplicar el resultado anterior a $T^{-1} \circ \varphi$ que, al combinarlo con el Teorema 11.3.21, produce la desigualdad

$$\begin{aligned} \mu_n(\varphi(Q)) &= |\det(T)| \cdot \mu_n(T^{-1}(\varphi(Q))) \\ &\leq |\det(T)| \cdot \Big(\sup_{x \in Q} \Big\| T^{-1} \circ D_x\varphi \Big\| \Big)^n \cdot \mu_n(Q). \end{aligned} \tag{1d}$$

Recordemos que la aplicación diferencial $D\varphi^{-1} : V \to \mathrm{L}(\mathbb{R}^n, \mathbb{R}^n)$ es uniformemente continua sobre el compacto $\varphi(Q)$ y, por lo tanto, $\sup_{y \in \varphi(Q)} \| D_x\varphi^{-1} \|$ es finito. Pongamos

$$K = \sup_{y \in \varphi(Q)} \Big\| D_x\varphi^{-1} \Big\|.$$

Fijemos un $\varepsilon > 0$ y usemos de nuevo la continuidad uniforme sobre Q de la aplicación diferencial $D\varphi$ para seleccionar un $\delta > 0$ de modo que

$$\| D_u\varphi - D_v\varphi \| < \varepsilon$$

siempre que $u, v \in Q$ satisfagan $0 < \| u - v \| < \delta$. Subdivida el cubo Q en N subcubos $Q_1, \ldots, Q_N$ cuyos interiores sean disjuntos dos a dos y cuyos lados sean menores o iguales a δ. Llamando x_i el centro del cubo Q_i y haciendo $y_i = \varphi(x_i)$, $i = 1, \ldots, N$, resulta, aplicando la desigualdad $(1d)$ al cubo Q_i y a la transformación lineal $T = D_{x_i}\varphi$, que

$$\begin{aligned} \mu_n(\varphi(Q)) &\leq \sum_{i=1}^{N} \mu_n(\varphi(Q_i)) \\ &\leq \sum_{i=1}^{N} |\det(D_{x_i}\varphi)| \cdot \Big(\sup_{x \in Q_i} \Big\| D_{x_i}\varphi^{-1} \circ D_x\varphi \Big\| \Big)^n \cdot \mu_n(Q_i) \end{aligned}$$

Observe que gracias a la Regla de la Cadena, $D_{y_i}\varphi^{-1} \circ D_{x_i}\varphi = I$ y, por lo tanto, para cada $x \in Q_i$,

$$\begin{aligned} \Big\| D_{y_i}\varphi^{-1} \circ D_x\varphi \Big\| - 1 &= \Big\| D_{y_i}\varphi^{-1} \circ D_x\varphi \Big\| - \| I \| \\ &\leq \Big\| D_{y_i}\varphi^{-1} \circ D_x\varphi - I \Big\| = \Big\| D_{y_i}\varphi^{-1} \circ \big(D_x\varphi - D_{x_i}\varphi \big) \Big\| \\ &\leq \Big\| D_{y_i}\varphi^{-1} \Big\| \, \| D_x\varphi - D_{x_i}\varphi \| < K\varepsilon. \end{aligned}$$

De esto se sigue que

$$\Big\| D_{y_i}\varphi^{-1} \circ D_x\varphi \Big\| < 1 + K\varepsilon \quad \text{para todo } x \in Q_i$$

y, en consecuencia,

$$\mu_n(\varphi(Q)) \leq (1 + K\varepsilon)^n \sum_{i=1}^{N} |\det(D_{x_i}\varphi)| \cdot \mu_n(Q_i) \leq (1 + K\varepsilon)^n \sum_{i=1}^{N} \int_Q |\det(D_x\varphi)| \cdot \chi_{Q_i}(x)\, dx.$$

Haciendo que ε tienda a cero en la última desigualdad se concluye que

$$\mu_n(\varphi(Q)) \leq \int_Q |\det(D_x\varphi)| \, d\mu_n(x).$$

Afirmamos que la desigualdad anterior es válida si Q se reemplaza por cualquier conjunto medible incluido en U. En efecto, comencemos con un conjunto abierto no vacío $G \subseteq U$ e invoquemos el Teorema 11.3.1 para representar a G en la forma $G = \bigcup_{i=1}^{\infty} Q_i$, donde $(Q_i)_{i=1}^{\infty}$ es una sucesión disjunta de cubos pertenecientes a $\mathfrak{D}_n$. Puesto que φ es un homeomorfismo, resulta que $\varphi(G)$ es abierto y, por lo tanto, medible. Además, como el borde de cada $\overline{Q}_i$ es de medida cero y $\overline{Q}_i$ es compacto e incluido en G, resulta de lo anterior que

$$\mu_n(\varphi(G)) \;=\; \sum_{i=1}^{\infty} \mu_n(\varphi(Q_i)) \;\leq\; \sum_{i=1}^{\infty} \int_{Q_i} |\det(D_x\varphi)|\, d\mu_n(x) \;=\; \int_G |\det(D_x\varphi)|\, d\mu_n(x).$$

Finalmente, suponga que E un conjunto medible Lebesgue incluido en U. Vamos a hacer dos consideraciones: la primera es suponer que $\det(D_x\varphi)$ es acotado sobre U, es decir, existe una constante $M \geq 0$ tal que $\sup_{x\in U} |\det(D_x\varphi)| \leq M$. Hecha esta suposición, también aceptemos que E es acotado. Siendo E medible, se sigue de la regularidad de μ_n que, dado $\varepsilon > 0$, existen un compacto K y un abierto G, que podemos suponer incluido en U, tales que

$$K \subseteq E \subseteq G \qquad \text{y} \qquad \mu_n(G \setminus K) \;<\; \varepsilon/M.$$

De nuevo, como φ es un homeomorfismo, resulta que $\varphi(K)$ es compacto, $\varphi(G)$ es abierto y $\varphi(K) \subseteq \varphi(E) \subseteq \varphi(G)$. Además, como $G \setminus K$ es un abierto incluido en U y $\varphi(G\setminus K) = \varphi(G) \setminus \varphi(K)$ se obtiene, aplicando las dos desigualdades anteriores,

$$\mu_n\big(\varphi(G) \setminus \varphi(K)\big) \;=\; \mu_n\big(\varphi(G\setminus K)\big) \;\leq\; \int_{G\setminus K} |\det(D_x\varphi)|\, d\mu_n(x) \;\leq\; M\frac{\varepsilon}{M} \;=\; \varepsilon.$$

De lo anterior se sigue que $\varphi(E)$ es medible y, además,

$$\begin{aligned} \mu_n(\varphi(E)) \;\leq\; \mu_n(\varphi(G)) \;&\leq\; \int_E |\det(D_x\varphi)|\, d\mu_n(x) + \int_{G\setminus K} |\det(D_x\varphi)|\, d\mu_n(x) \\ &\leq\; \int_E |\det(D_x\varphi)|\, d\mu_n(x) + \varepsilon. \end{aligned}$$

Como ε es arbitrario, concluimos que $(1c)$ es válido para todo conjunto E que esté incluido en U y que sea medible y acotado.

Si E no es acotado, entonces podemos representarlo como $E = \bigcup_{k=1}^{\infty} E_k$, donde $E_k = E \cap B_\infty(0,k)$ es medible y acotado para cada $k \in \mathbb{N}$. Ya que $\varphi(E) = \bigcup_{k=1}^{\infty} \varphi(E_k)$, aplicando la continuidad de μ_n y la conclusión anterior se obtiene la desigualdad deseada:

$$\mu_n(\varphi(E)) \;=\; \lim_{k\to\infty} \mu_n(\varphi(E_k)) \;\leq\; \lim_{k\to\infty} \int_{E_k} |\det(D_x\varphi)|\, d\mu_n(x) \;=\; \int_E |\det(D_x\varphi)|\, d\mu_n(x).$$

La segunda consideración que vamos analizar es suponer que $|\det(D_x\varphi)|$ no está acotado sobre U. En este caso, para cada $m \in \mathbb{N}$, definamos

$$U_m \;=\; \big\{x \in U \,:\, |\det(D_x\varphi)| < m\big\}$$

y observe que $E = \bigcup_{m=1}^{\infty}(E \cap U_m)$. Aplicando la desigualdad anterior a cada uno de los conjuntos medibles $E \cap U_m$ resulta que

$$\mu_n(\varphi(E)) \;=\; \lim_{m\to\infty} \mu_n(\varphi(E\cap U_m)) \;\leq\; \lim_{m\to\infty} \int_{E\cap U_m} |\det(D_x\varphi)|\, d\mu_n(x) \;=\; \int_E |\det(D_x\varphi)|\, d\mu_n(x).$$

La prueba ha concluido. ■

Estamos listo para presentar el Teorema del Cambio de Variable en su versión no-lineal.

Teorema 11.3.28 (Cambio de variable: caso no-lineal). *Sea $\varphi : U \to V$ un* ***difeomorfismo de clase*** C^1*. Si $f : V \to \mathbb{R}$ es una* ***función medible no-negativa*** *o* μ_n***-integrable****, entonces*

$$\int_V f(x)\, d\mu_n(x) \;=\; \int_U (f \circ \varphi)(x) \cdot |\det(D_x\varphi)|\, d\mu_n(x).$$

Prueba. Observe que $(f \circ \varphi) \cdot |\det(D_x\varphi)|$ es una función medible siempre que f lo sea. Suponga, en primer lugar, que f es una función simple no-negativa, digamos $f = \sum_{i=1}^k a_i \chi_{\varphi(E_i)}$ con $U = \bigcup_{i=1}^k E_i$ y donde los E_i son disjuntos doa a dos. Por la Desigualdad de Sard, Lema 11.3.27, tenemos que

$$\begin{aligned}
\int_V f(x)\, d\mu_n(x) \;&=\; \sum_{i=1}^n a_i \mu_n(\varphi(E_i)) \;\leq\; \sum_{i=1}^n a_i \int_{E_i} |\det(D_x\varphi)|\, d\mu_n \\
&= \int_U (f \circ \varphi)(x) \cdot |\det(D_x\varphi)|\, d\mu_n(x).
\end{aligned}$$

Si ahora f es no-negativa, entonces

$$\begin{aligned}
\int_V f(x)\, d\mu_n(x) \;&=\; \sup\left\{ \int_V s(x)\, d\mu_n(x) \;:\; 0 \leq s \leq f,\ s \text{ simple} \right\} \\
&\leq \sup\left\{ \int_U (s \circ \varphi)(x) \cdot |\det(D_x\varphi)|\, d\mu_n(x) \;:\; 0 \leq s \leq f,\ s \text{ simple} \right\} \\
&\leq \int_U (f \circ \varphi)(x) \cdot |\det(D_x\varphi)|\, d\mu_n(x).
\end{aligned}$$

Para verificar la desigualdad en el otro sentido, se trabaja con el cambio de variable φ^{-1} aplicado a la función $f \circ \varphi |\det(D_x\varphi)|$ y al conjunto $\varphi(U) = V$, es decir,

$$\begin{aligned}
\int_U (f \circ \varphi)(x) \cdot |\det(D_x\varphi)|\, d\mu_n(x) \;&\leq\; \int_{\varphi(U)} ((f \circ \varphi) \circ \varphi^{-1}(x) \cdot |\det(D_{\varphi^{-1}(x)}\varphi)|) \cdot |\det(D_x\varphi^{-1})|) \\
&= \int_V f(x)\, d\mu_n(x).
\end{aligned}$$

Esto establece la igualdad requerida para cuando f es cualquier función medible no-negativa. Finalmente, si f es μ_n-integrable sobre V, entonces descomponiendo a f en la forma $f = f^+ - f^-$ y teniendo en cuenta que

$$\big((f \circ \varphi) \cdot |\det(D_x\varphi)|\big)^+ = (f^+ \circ \varphi) \cdot |\det(D_x\varphi)| \quad \text{y} \quad \big((f \circ \varphi) \cdot |\det(D_x\varphi)|\big)^- = (f^- \circ \varphi) \cdot |\det(D_x\varphi)|$$

el resultado sigue de la primera parte aplicado a cada una de éstas funciones. ■

Si aplicamos el Teorema del Cambio de Variable en su versión no-lineal a la función constante $f(x) = 1$ para todo $x \in U$, entonces la Desigualdad de Sard se convierte, en realidad, en una igualdad, es decir:

Corolario 11.3.29 (Preservando volumen). *Sea $\varphi : U \to V$ un **difeomorfismo de clase** C^1. Si $E \subseteq U$ es **medible**, entonces*

$$\mu_n(\varphi(E)) \;=\; \int_E |\det(D_x\varphi)|\, d\mu_n.$$

En particular, si $|\det(D_x\varphi)| = 1$ para todo $x \in E$, entonces $\mu_n(\varphi(E)) = \mu_n(E)$.

Ejemplo 11.3.2 (Coordenadas Polares). En $\mathbb{R}^2$ las coordenada cartesianas (x, y) y las coordenadas polares (r, θ) están relacionadas por las ecuaciones:

$$x \;=\; r\cos(\theta) \qquad y \qquad y \;=\; r\,\mathrm{sen}(\theta)$$

para todo $r > 0$ y todo $\theta \in (-\pi, \pi)$. Considere los subconjuntos abiertos de $\mathbb{R}^2$,

$$U \;=\; (0, +\infty) \times (-\pi, \pi) \qquad \text{y} \qquad V \;=\; \mathbb{R}^2 \setminus Q,$$

donde $Q = (-\infty, 0] \times \{0\}$. La aplicación $\varphi : U \to V$ dada por $\varphi(r, \theta) = (r\cos(\theta), r\,\mathrm{sen}(\theta))$ es un difeomorfismo de clase C^1 cuya matriz Jacobiana es:

$$J_\varphi(r, \theta) \;=\; \begin{pmatrix} \cos(\theta) & -r\,\mathrm{sen}(\theta) \\ \mathrm{sen}(\theta) & r\cos(\theta) \end{pmatrix} \qquad \text{y} \qquad \det(J_\varphi(r, \theta)) \;=\; r \;>\; 0.$$

Si $f : \mathbb{R}^2 \to \mathbb{R}$ es cualquier función medible no-negativa o μ_2-integrable, entonces el Teorema del Cambio de Variable no-lineal y el hecho de que $\mu_2(Q) = 0$, permiten concluir que

$$\begin{aligned}
\int_{\mathbb{R}^2} f(x, y)\, d\mu_2(x, y) \;&=\; \int_{\mathbb{R}^2 \setminus Q} f(x, y)\, d\mu_2(x, y) + \int_Q f(x, y)\, d\mu_2(x, y) \\
&=\; \int_V f(x, y)\, d\mu_2(x, y) \;=\; \int_U (f \circ \varphi)(r, \theta) |\det(J_\varphi(r, \theta))|\, d\mu_2(r, \theta) \\
&=\; \int_{(0, +\infty) \times (-\pi, \pi)} f(r\cos(\theta), r\,\mathrm{sen}(\theta))\, r\, dr\, d\theta.
\end{aligned}$$

(a) **Área de un Disco**. Podemos aplicar el cambio de coordenadas polares para calcular el área del disco $E = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq r^2\}$ con $r > 0$. En efecto, tomando $f(x, y) = 1$ para todo $(x, y) \in \mathbb{R}^2$, resulta

$$\mu_2(E) \;=\; \int_E d\mu_2(x, y) \;=\; \int_{(0, r) \times (-\pi, \pi)} r\, dr\, d\theta \;=\; \int_0^r \left(\int_{-\pi}^{\pi} r\, d\theta \right) dr \;=\; \pi r^2$$

donde se ha utilizado el Teorema de Fubini.

(b) **Integral de Gauss**. Otra aplicación interesante del cambio de coordenadas polares combinado con el Teorema de Fubini permite determinar el valor de la ***integral de Gauss***, esto es:

$$\int_{-\infty}^{\infty} e^{-\varrho x^2}\, dx \;=\; \sqrt{\frac{\pi}{\varrho}}, \quad \varrho > 0.$$

En efecto, si se considera la función $f(x,y) = e^{-\varrho\,(x^2+y^2)}$ para todo $(x,y) \in \mathbb{R}^2$, entonces como $x^2 + y^2 = r^2$,

$$\begin{aligned}\int_{\mathbb{R}^2} f(x,y)\, d\mu_2(x,y) &= \int_{(0,+\infty)\times(-\pi,\pi)} e^{-\varrho\, r^2} r\, d\mu_2(r,\theta) \\ &= \int_{-\pi}^{\pi} \left(\int_0^{\infty} r\, e^{-\varrho\, r^2}\, dr \right) d\theta \\ &= \int_{-\pi}^{\pi} \frac{1}{2\varrho}\, d\theta = \frac{\pi}{\varrho}.\end{aligned}$$

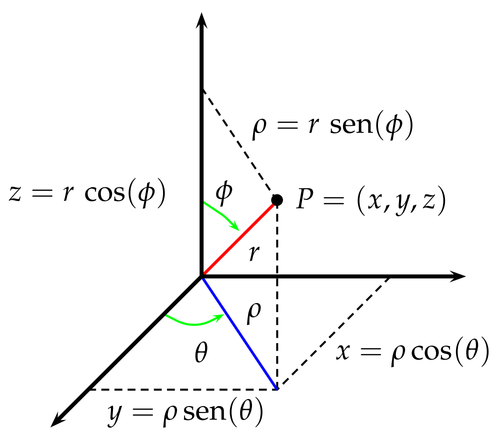
Por una nueva aplicación del Teorema de Fubini tenemos que

$$\int_{\mathbb{R}^2} f(x,y)\, d\mu_2(x,y) = \left(\int_{\mathbb{R}} e^{-\varrho\, x^2}\, dx \right)\left(\int_{\mathbb{R}} e^{-\varrho\, y^2}\, dy \right) = \left(\int_{\mathbb{R}} e^{-\varrho\, t^2}\, dt \right)^2,$$

y, en consecuencia,

$$\int_{\mathbb{R}} e^{-\varrho\, t^2}\, dt = \sqrt{\frac{\pi}{\varrho}}.$$

Ejemplo 11.3.3 (Coordenadas Esféricas). En $\mathbb{R}^3$ las coordenada cartesianas (x,y,z) y las coordenadas esféricas (r,θ,ϕ) están relacionadas por las ecuaciones:



$$\begin{cases} x = r\ \mathrm{sen}(\phi)\cos(\theta) \\ y = r\ \mathrm{sen}(\phi)\,\mathrm{sen}(\theta) \\ z = r\ \cos(\phi) \end{cases}$$

Sean

$$\begin{aligned} U &= \{(r,\phi,\theta) : 0 < r < +\infty,\ 0 < \phi < \pi,\ -\pi < \theta < \pi\} = (0,+\infty)\times(0,\pi)\times(-\pi,\pi) \\ V &= \mathbb{R}^3 \setminus \{(x,0,z) : x \le 0,\ z \in \mathbb{R}\}. \end{aligned}$$

La función $\varphi : U \to V$ dada por $\varphi(r,\phi,\theta) = (r\ \mathrm{sen}(\phi)\cos(\theta), r\ \mathrm{sen}(\phi)\,\mathrm{sen}(\theta), r\ \cos(\phi))$ es un difeomorfismo de clase C^1, donde

$$J_\varphi(r,\phi,\theta) = \begin{pmatrix} \mathrm{sen}(\phi)\cos(\theta) & r\ \cos(\phi)\cos(\theta) & -r\ \mathrm{sen}(\phi)\,\mathrm{sen}(\theta) \\ \mathrm{sen}(\phi)\,\mathrm{sen}(\theta) & r\ \cos(\phi)\,\mathrm{sen}(\theta) & r\ \mathrm{sen}(\phi)\cos(\theta) \\ \cos(\phi) & -r\ \mathrm{sen}(\phi) & 0 \end{pmatrix}$$

y

$$\det(J_\varphi(r,\phi,\theta)) \;=\; r^2 \operatorname{sen}(\phi) \;>\; 0.$$

Si $f:\mathbb{R}^3 \to \mathbb{R}$ es una función medible no-negativa o μ_3-integrable, entonces

$$\int_{\mathbb{R}^3} f(x,y,z)\, d\mu_3(x,y,z) \;=\; \int_U f(r \operatorname{sen}(\phi)\cos(\theta), r \operatorname{sen}(\phi)\operatorname{sen}(\theta), r \cos(\phi))\, r^2 \operatorname{sen}(\phi)\, dr\, d\phi\, d\theta.$$

(a) **Volumen de una Esfera**. Sea $E = \{(x,y,z)\in\mathbb{R}^3 : x^2+y^2+z^2 \leq R^2\}$ la esfera euclidiana con centro $(0,0,0)$ y radio $R>0$. Se trata de calcular el volumen de E. Tomando $f\equiv 1$ sobre $\mathbb{R}^3$ y aplicando el Teorema de Fubini se obtiene:

$$\begin{aligned}
\mu_3(E) &= \int_E d\mu_2 \;=\; \int_{(0,R)\times(0,\pi)\times(-\pi,\pi)} r^2 \operatorname{sen}(\theta)\, dr\, d\phi\, d\theta \\
&= \int_0^R \left(\int_0^\pi \left[r^2 \operatorname{sen}(\phi) \int_{-\pi}^{\pi} d\theta \right] d\phi \right) dr \\
&= \int_0^R \left(\int_0^\pi 2\pi r^2 \operatorname{sen}(\phi)\, d\phi \right) dr \;=\; \int_0^R 4\pi r^2\, dr \;=\; \frac{4}{3}\pi R^3.
\end{aligned}$$

11.3.3. El Teorema de Sard

Sea U un subconjunto abierto de $\mathbb{R}^n$ y suponga que $f:U\to\mathbb{R}^n$ es de clase C^1. Un punto $x\in U$ es un **punto crítico** de f si $D_x f:\mathbb{R}^n\to\mathbb{R}^n$ no es biyectiva. Esto, por supuesto, significa que $D_x f$ no es invertible y, por lo tanto, $\det(J_f(x)) = 0$. Sea

$$\mathrm{PC}_f \;=\; \{x\in U : x \text{ es un punto crítico de } f\}.$$

Este conjunto puede ser tan grande como U. Por ejemplo, si f es una función constante, digamos que $f(x)=c$ para todo $x\in U$, entonces $\mathrm{PC}_f = U$. A pesar de este hecho, un hermoso resultado debido a Arthur Sard (1909-1980) establece que, no importa que tan grande pueda ser PC_f, la medida (de Lebesgue) de su imagen, $\mu_n(f(\mathrm{PC}_f))$, es siempre cero.

Teorema 11.3.30 (Sard). *Sea U un subconjunto abierto de $\mathbb{R}^n$ y sea $f:U\to\mathbb{R}^n$ una **función de clase** C^1. Entonces*

$$\mu_n(f(\mathrm{PC}_f)) \;=\; 0.$$

Prueba. Sea Q_ρ un ρ-cubo cerrado con centro en $a=(a_1,\ldots,a_n)$ incluido en U, esto es,

$$Q_\rho \;=\; \{x\in\mathbb{R}^n : \|x-a\| \leq \rho\}.$$

Como f es **uniformemente diferenciable** sobre Q_ρ, entonces, para cada $\varepsilon>0$, existe un $\delta>0$ tal que

$$\|f(y)-f(x)-D_x f(x-y)\| \;\leq\; \varepsilon\cdot\|x-y\| \tag{1a}$$

para todo $x,y\in Q_\rho$ con $0<\|x-y\|<\delta$. Escoja un $N\in\mathbb{N}$ lo suficientemente grande de modo que $\rho/N<\delta$ y subdivida el cubo Q_ρ en N^n subcubos cerrados, digamos $Q_1,\ldots,Q_{N^n}$, de modo que cada Q_i tenga todos sus lados iguales a ρ/N. Esta elección de los $\frac{\rho}{N}$-cubos Q_i

garantiza que cualesquiera sean $x, y \in Q_i$, la longitud del segmento que los une es menor que la longitud de cualquiera de sus diagonales, es decir,

$$\| y - x \| \leq \sqrt{n}\,\frac{\rho}{N} \quad \text{para todo } x, y \in Q_i.$$

Imaginemos ahora que algún $\frac{\rho}{N}$-cubo, digamos Q_i, intersecta a PC_f y sea $x \in Q_i \cap \mathrm{PC}_f$. Entonces $\det(J_f(x)) = 0$ y, por consiguiente, la aplicación lineal $D_x f : \mathbb{R}^n \to \mathbb{R}^n$ no es sobreyectiva, es decir, $D_x f(\mathbb{R}^n)$ es un subespacio propio de $\mathbb{R}^n$ y, por consiguiente, está incluido en un subespacio X de dimensión $n-1$. En particular, el conjunto

$$A = \{D_x f(y - x) : y \in Q_i\} \subseteq X.$$

Para cada $y \in Q_i$, pongamos $v_y = D_x f(y - x)$. Por $(1a)$ se tiene que

$$\| f(y) - f(x) - v_y \| = \| f(y) - f(x) - D_x f(x - y) \| \leq \varepsilon \cdot \| x - y \| \leq \varepsilon \sqrt{n}\,\frac{\rho}{N}$$

para todo $y \in Q_i$, de donde se sigue que el conjunto

$$B = \{f(y) - f(x) : y \in Q_i\}$$

está situado a una distancia menor que $\varepsilon \sqrt{n}\,(\rho/N)$ de X y, por consiguiente, el conjunto

$$C = \{f(y) : y \in Q_i\}$$

está situado a una distancia menor que $\varepsilon \sqrt{n}\,(\rho/N)$ del hiperplano $n-1$ dimensional $f(x) + X$. Pero por el Corolario 11.3.24, sabemos que existe una constante $M > 0$ tal que

$$\| f(y) - f(x) \| \leq M \| y - x \| \leq M \sqrt{n}\,\frac{\rho}{N},$$

lo cual nos indica que $f(y)$ pertenece a un cubo $\widehat{Q}$ en $\mathbb{R}^n$ cuya base es un "rectángulo" de dimensión $n-1$ donde cada uno de sus lados mide $2M\sqrt{n}\,\frac{\rho}{N}$ y cuya "altura" mide $2\varepsilon\sqrt{n}\,\frac{\rho}{N}$. Así,

$$\boldsymbol{vol}(\widehat{Q}) = \left(2\,M\sqrt{n}\,\frac{\rho}{N}\right)^{n-1} \times \left(2\,\varepsilon\sqrt{n}\,\frac{\rho}{N}\right) = C\,\frac{\varepsilon}{N^n},$$

donde C no depende ni de ε ni de N. En conclusión, hemos demostrado que si $Q_i \cap \mathrm{PC}_f \neq \varnothing$, entonces $f(Q_i \cap \mathrm{PC}_f)$ está contenido en un cubo cuyo volumen es $\frac{C\,\varepsilon}{N^n}$. Pero como existen a lo sumo N^n de tales cubos, resulta que

$$f(Q_\rho \cap \mathrm{PC}_f) = \bigcup_{i=1}^{N^n} f(Q_i \cap \mathrm{PC}_f)$$

y, por lo tanto,

$$\mu_n\big(f(Q_\rho \cap \mathrm{PC}_f)\big) \leq \sum_{i=1}^{N^n} \frac{C\,\varepsilon}{N^n} = C\,\varepsilon.$$

Pues bien, siendo ε arbitrario se concluye que $\mu_n\big(f(Q_\rho \cap \mathrm{PC}_f)\big) = 0$. Para cada $x \in U$, considere un ρ-cubo $Q_\rho^x \subseteq \mathbb{R}^n$ con vértices racionales y tal que $x \in Q_x$. Claramente la colección $\mathcal{Q} = \{Q_\rho^x : x \in U\}$ es un cubrimiento numerable de U y, en consecuencia,

$$\mu_n(f(\mathrm{PC}_f)) \leq \sum_{Q_\rho^x \in \mathcal{Q}} \mu_n\big(f(Q_\rho^x \cap \mathrm{PC}_f)\big) = 0.$$

La prueba es completa. ■

Concluimos esta sección con un resultado que ya ha sido demostrado para aplicaciones Lipschitz definidas sobre $[a, b]$.

Teorema 11.3.31. *Sea $f : U \to \mathbb{R}^n$ una aplicación M-**Lipschitz**. Entonces f satisface la condición (N) de Lusin, es decir, transforma conjuntos nulos en conjuntos nulos.*

Prueba. Sea $Q = B_\infty(a, r)$ una bola cerrada con centro en $a = (a_1, \ldots, a_n) \in U$ y radio $r > 0$ incluida en U, es decir, Q es un r-cubo. Como f es M-Lipschitz, resulta que para todo $x \in Q$,

$$\| f(x) - f(a) \| \leq M \| x - a \| \leq rM.$$

Esto muestra que $f(Q)$ está contenido en el cubo

$$Q^* = \Big\{ y \in \mathbb{R}^n : \| y - f(a) \| \leq rM \Big\},$$

de donde resulta que

$$\begin{aligned} \mu_n(f(Q)) &\leq \mu_n(Q^*) \\ &= (2rM)^n \\ &= M^n \cdot \mu_n(Q). \end{aligned}$$

Suponga ahora que E es un conjunto nulo incluido en U y sea $\varepsilon > 0$. Por definición, existe un conjunto abierto $G \subseteq U$ tal que $E \subseteq G$ y $\mu_n(G) < \varepsilon / M^n$. Exprese a G, usando el Corolario 11.3.2, como una unión casi-disjunta de cubos diádicos cerrados $(Q_j)_{j=1}^\infty$. Se sigue de la desigualdad anterior que

$$\mu_n(f(E)) \leq \mu_n(f(G)) \leq \sum_{j=1}^\infty \mu_n(f(Q_j)) \leq M^n \sum_{j=1}^\infty \mu_n(Q_j) = M^n \mu_n(G) < \varepsilon.$$

Esto termina la prueba. ■

11.4. Medida de Borel sobre el Espacio de Cantor $2^{\mathbb{N}}$

Recordemos que una aplicación $f : X \to Y$, donde $(X, \mathcal{A})$ y $(Y, \mathcal{B})$ son espacios medibles, se dice que es $(\mathcal{A}, \mathcal{B})$-***medible***, o simplemente, ***medible*** si $f^{-1}(B) \in \mathcal{A}$ para cada $B \in \mathcal{B}$. Observe que si $\mathcal{C}$ es una colección de subconjuntos de Y tal que $\mathcal{B} = \sigma(\mathcal{C})$, entonces f es medible si, y sólo si, $f^{-1}(C) \in \mathcal{A}$ para cada $C \in \mathcal{C}$. En efecto, si se considera la familia

$$\mathcal{D} = \{ B \subseteq Y : f^{-1}(B) \in \mathcal{A} \},$$

resulta que, por nuestra hipótesis, $\mathcal{C} \subseteq \mathcal{D}$ y como $\mathcal{D}$ es una σ-álgebra conteniendo a $\mathcal{C}$, se tiene que $\mathcal{B} = \sigma(\mathcal{C}) \subseteq \mathcal{D}$. Por consiguiente, $f^{-1}(B) \in \mathcal{A}$ para cada $B \in \mathcal{B}$, lo cual muestra que f es medible.

Nuestro primer objetivo, en esta sección, es introducir una σ-álgebra en el espacio producto de una familia arbitraria $(X_\alpha, \mathcal{A}_\alpha)_{\alpha \in I}$ de espacios medibles y, posteriormente, definir una medida muy natural sobre el espacio de Cantor $2^{\mathbb{N}}$.

11.4.1. σ-álgebra Producto

Similar a la construcción de la topología producto, se puede obtener una adecuada σ-álgebra sobre el espacio producto relacionada con las proyecciones. Suponga entonces que $(X_\alpha, \mathcal{A}_\alpha)_{\alpha\in I}$ es una familia de espacios medibles y, como antes, sea

$$X = \prod_{\alpha\in I} X_\alpha$$

su espacio producto. Lo que deseamos es construir la σ-álgebra más pequeña sobre X, digamos $\mathcal{S}_\pi$, de modo que cada proyección p_α sea $(\mathcal{S}_\pi, \mathcal{A}_\alpha)$-medible. La construcción de tal σ-álgebra se realiza procediendo del modo siguiente: en primer lugar, se considera la colección

$$\mathcal{S} = \bigcup_{\alpha\in I} p_\alpha^{-1}(\mathcal{A}_\alpha) = \bigcup_{\alpha\in I} \{p_\alpha^{-1}(A_\alpha) : A_\alpha \in \mathcal{A}_\alpha\}.$$

Si bien es cierto que $p_\alpha^{-1}(\mathcal{A}_\alpha)$ es una σ-álgebra sobre X para cada $\alpha \in I$ (la cual hace medible sólo a p_α) resulta que, por lo general, $\mathcal{S}$ no es una σ-álgebra. La **σ-álgebra generada** por $\mathcal{S}$, $\sigma(\mathcal{S})$, la llamaremos la **σ-álgebra producto** sobre X y será indicada por

$$\bigotimes_{\alpha\in I} \mathcal{A}_\alpha.$$

Observe que, por definición, $\bigotimes_{\alpha\in I} \mathcal{A}_\alpha$ es la σ-álgebra más pequeña que permite que cada proyección p_β sea $(\bigotimes_{\alpha\in I} \mathcal{A}_\alpha, \mathcal{A}_\beta)$-medible. Al par $(\prod_{\alpha\in I} X_\alpha, \bigotimes_{\alpha\in I} \mathcal{A}_\alpha)$ lo llamaremos **espacio medible producto**.

Otra forma más practica de obtener $\bigotimes_{\alpha\in I} \mathcal{A}_\alpha$ es la siguiente: considere la colección $\mathcal{S}'$ *formada por todas las intersecciones finitas de elementos de* $\mathcal{S}$. En general, $\mathcal{S}'$ no es una álgebra. Sin embargo, si para cada conjunto finito $F = \{\alpha_1, \ldots, \alpha_n\} \subseteq I$, se define

$$\mathcal{C}_F = \bigcup_{i=1}^{n} p_{\alpha_i}^{-1}(\mathcal{A}_{\alpha_i}),$$

resulta que

$$\mathcal{S}^* = \bigcup_{F\in\mathcal{P}_{\text{fin}}(I)} \mathcal{S}_F,$$

si es una álgebra, donde $\mathcal{S}_F$ es la **álgebra generada por** $\mathcal{C}_F$. Cualquier elemento $C \in \mathcal{S}^*$ se llama un **cilindro**. Con estas notaciones se tiene que:

Teorema 11.4.1. *$\mathcal{S}^* = \bigcup_{F\in\mathcal{P}_{\text{fin}}(I)} \mathcal{S}_F$ es una **álgebra** y se cumple que*

$$\sigma(\mathcal{S}^*) = \bigotimes_{\alpha\in I} \mathcal{A}_\alpha. \tag{1}$$

Prueba. Veamos, en primer lugar, que $\mathcal{S}^*$ es una álgebra.

(i) Si $A \in \mathcal{S}^*$, entonces existe $F \in \mathcal{P}_{\text{fin}}(I)$ tal que $A \in \mathcal{S}_F$. Luego, $A^c \in \mathcal{S}_F \subseteq \mathcal{S}^*$.

(ii) Sean $A_1, A_2 \in \mathcal{S}^*$ y seleccione subconjuntos finitos F_1 y F_2 de I tales que $A_i \in \mathcal{S}_{F_i}$ para $i = 1, 2$. Puesto que $F = F_1 \cup F_2$ es finito y $\mathcal{S}_{F_i} \subseteq \mathcal{S}_F$ se tiene que $A_1 \cup A_2 \in \mathcal{S}_F \subseteq \mathcal{S}^*$.

Para comprobar que (1) se cumple, observe que si $F = \{\alpha\}$ consta de un único elemento, entonces $\mathcal{C}_F = p_\alpha^{-1}(\mathcal{A}_\alpha) = \mathcal{S}_F$ y, por lo tanto,

$$\bigcup_{\alpha \in I} p_\alpha^{-1}(\mathcal{A}_\alpha) \subseteq \bigcup_{F \in \mathcal{P}_{\text{fin}}(I)} \mathcal{S}_F = \mathcal{S}^*.$$

De esto se sigue que

$$\bigotimes_{\alpha \in I} \mathcal{A}_\alpha \subseteq \sigma(\mathcal{S}^*).$$

Para verificar la otra inclusión, observe que para cada $F \in \mathcal{P}_{\text{fin}}(I)$,

$$\mathcal{C}_F \subseteq \sigma\Big(\bigcup_{\alpha \in I} p_\alpha^{-1}(\mathcal{A}_\alpha) \Big) = \bigotimes_{\alpha \in I} \mathcal{A}_\alpha,$$

luego

$$\mathcal{S}_F \subseteq \bigotimes_{\alpha \in I} \mathcal{A}_\alpha \quad \text{para todo } F \in \mathcal{P}_{\text{fin}}(I)$$

y, en consecuencia,

$$\mathcal{S}^* = \bigcup_{F \in \mathcal{P}_{\text{fin}}(I)} \mathcal{S}_F \subseteq \bigotimes_{\alpha \in I} \mathcal{A}_\alpha,$$

de donde se deduce que

$$\sigma(\mathcal{S}^*) \subseteq \bigotimes_{\alpha \in I} \mathcal{A}_\alpha$$

y termina la prueba. ■

A la colección $\mathcal{S}^*$ del teorema anterior la llamaremos la **álgebra de los cilindros**. Observe que cuando I es un conjunto finito, digamos $I = \{1, 2, \ldots, n\}$, entonces

$$p_i^{-1}(A_i) = X_1 \times X_2 \times \cdots \times A_i \times \cdots \times X_n,$$

para cada $i \in I$, por lo que

$$\bigotimes_{i=1}^{n} \mathcal{A}_i = \sigma\big(\{A_1 \times \cdots \times A_n : A_1 \in \mathcal{A}_1, \ldots, A_n \in \mathcal{A}_n\}\big).$$

Suponga ahora que $\operatorname{card}(I) \geq \aleph_0$. Si en el análisis anterior, en lugar de considerar intersecciones finitas trabajamos con intersecciones infinitas, es decir, si tomamos un conjunto infinito numerable arbitrario $F \subseteq I$ y definimos

$$\mathcal{C}_F = \Big\{ \bigcup_{\alpha_k \in F} p_{\alpha_k}^{-1}(E_{\alpha_k}) : E_{\alpha_k} \in \mathcal{A}_{\alpha_k} \text{ para cada } \alpha_k \in F \Big\},$$

y si hacemos $\mathcal{A}_F = \sigma(\mathcal{C}_F)$, resulta entonces que $\bigcup_{F \in I^{\mathbb{N}}} \mathcal{A}_F$ es una σ-álgebra y

$$\bigotimes_{\alpha \in I} \mathcal{A}_\alpha = \bigcup_{F \in I^{\mathbb{N}}} \mathcal{A}_F$$

Este hecho nos indica que cada elemento de la σ-álgebra producto depende sólo de una cantidad numerable de coordenadas y, por lo tanto, $\bigotimes_{\alpha\in I}\mathcal{A}_\alpha$ es, por lo general, muy pequeña y muchas veces no pertenecen a ella conjuntos que resultan ser de gran interés en matemáticas. Por ejemplo, la colección $\mathcal{C}(\mathbb{R})$ de todas las funciones continuas $f:\mathbb{R}\to[0,1]$ puede ser vista como un subconjunto del espacio producto $\prod_{\alpha\in I}\mathbb{R}_\alpha$, donde $I=[0,1]$ y $\mathbb{R}_\alpha=\mathbb{R}$ para todo $\alpha\in I$. Si $\mathfrak{B}_\alpha=\mathfrak{Bo}(\mathbb{R})$ es la σ-álgebra de Borel sobre $\mathbb{R}$ y teniendo en cuenta que toda función continua $f:\mathbb{R}\to[0,1]$ depende sólo de los valores que asume en $[0,1]$, resulta que $\mathcal{C}(\mathbb{R})\notin\bigotimes_{\alpha\in I}\mathfrak{B}_\alpha$.

Definición 11.4.2. *Sea $(X_\alpha,\mathfrak{I}_\alpha)_{\alpha\in I}$ una familia de espacios topológicos Hausdorff y sea $\mathfrak{T}_p$ la topología producto sobre el espacio producto $X=\prod_{\alpha\in I}X_\alpha$. La σ-álgebra generada por $\mathfrak{T}_p$ se llama la **σ-álgebra producto de Borel** del espacio producto X y denotada por $\mathfrak{Bo}(X)$.*

Si para cada $\alpha\in I$, indicamos por $\mathfrak{Bo}(X_\alpha)$ la σ-álgebra de Borel de X_α, entonces la σ-álgebra producto $\bigotimes_{\alpha\in I}\mathfrak{Bo}(X_\alpha)$ y la σ-álgebra producto de Borel $\mathfrak{Bo}(X)$ son, en general, distintas. De hecho, siempre se cumple la inclusión:

$$\bigotimes_{\alpha\in I}\mathfrak{Bo}(X_\alpha)\ \subseteq\ \mathfrak{Bo}(X). \tag{2}$$

Para ver esto último, fijemos un $\alpha\in I$ tomado de modo arbitrario. Por la continuidad de p_α se tiene que

$$p_\alpha^{-1}(G)\in\mathfrak{T}_p\subseteq\mathfrak{Bo}(X)\quad\text{para todo } G\in\mathfrak{I}_\alpha$$

y, en consecuencia, $p_\alpha^{-1}(\mathfrak{I}_\alpha)\subseteq\mathfrak{Bo}(X)$ para todo $\alpha\in I$. Por esto,

$$\bigcup_{\alpha\in I}p_\alpha^{-1}(\mathfrak{I}_\alpha)\subseteq\mathfrak{Bo}(X)$$

y así, $\bigotimes_{\alpha\in I}\mathfrak{Bo}(X_\alpha)\ \subseteq\ \mathfrak{Bo}(X)$. ■

Existen, por supuesto, condiciones bajo las cuales la igualdad en (2) se alcanza. Una de tales condiciones viene dada por el siguiente teorema.

Teorema 11.4.3. *Sea $(X_n,d_n)_{n=1}^\infty$ una sucesión de **espacios métricos separables** y sea $X=\prod_{n=1}^\infty X_n$ su espacio producto. Entonces*

$$\bigotimes_{n=1}^{\infty}\mathfrak{Bo}(X_n)\ =\ \mathfrak{Bo}(X).$$

Prueba. Para cada $n\in\mathbb{N}$, sea $\mathcal{B}_n$ una base numerable para la topología $\mathfrak{I}_n$ y sea

$$\mathcal{E}\ =\ \bigcup_{n=1}^{\infty}p_n^{-1}(\mathcal{B}_n).$$

Observe que $\sigma(\mathcal{B}_n)=\mathfrak{Bo}(X_n)$, por lo que $\sigma(\mathcal{E})\ =\ \bigotimes_{n=1}^\infty\mathfrak{Bo}(X_n)$. Veamos ahora que $p_n^{-1}(\mathfrak{I}_n)\subseteq\sigma(\mathcal{E})$ para cualquier $n\in\mathbb{N}$. En efecto, sea $V\in p_n^{-1}(\mathfrak{I}_n)$. Entonces existe un conjunto abierto $U\in\mathfrak{I}_n$ tal que $p_n^{-1}(U)=V$. Puesto que $\mathcal{B}_n$ es una base para $\mathfrak{I}_n$, existe una sucesión $(B_k)_{k=1}^\infty$ en $\mathcal{B}_n$ tal que $U=\bigcup_{k=1}^\infty B_k$ de donde resulta que $p_n^{-1}(U)=\bigcup_{k=1}^\infty p_n^{-1}(B_k)\in\sigma(\mathcal{E})$ y, así, $p_n^{-1}(\mathfrak{I}_n)\subseteq\sigma(\mathcal{E})$. Como lo anterior es válido para todo $n\in\mathbb{N}$, se tiene que

$$\mathcal{G}\ =\ \bigcup_{n=1}^{\infty}p_n^{-1}(\mathfrak{I}_n)\ \subseteq\ \sigma(\mathcal{E}).$$

De esto se sigue

$$\mathfrak{Bo}(X) \;=\; \sigma(\mathcal{G}) \;\subseteq\; \sigma(\mathcal{E}) \;=\; \bigotimes_{n=1}^{\infty} \mathfrak{Bo}(X_n)$$

lo que, combinado con la inclusión (2), da por finalizada la prueba. ■

11.4.2. Una Métrica sobre $2^{\mathbb{N}}$

En esta sección nos ocuparemos de construir métrica muy particular sobre $2^{\mathbb{N}}$ que genera la topología producto.

Para cada $n \in \mathbb{N}$, sea $X_n = \{0,1\}$ y dotemos a dicho conjunto con la topología discreta $\mathfrak{I}_n = \mathscr{P}(X_n)$. Si d_n es la métrica discreta sobre X_n, esto es,

$$d_n(x,y) \;=\; \begin{cases} 1 & \text{si } x \neq y \\ 0 & \text{si } x = y, \end{cases}$$

resulta que (X_n, d_n) es un espacio métrico separable y, entonces, por el Teorema 11.4.3

$$\mathfrak{Bo}(X) \;=\; \bigotimes_{n=1}^{\infty} \mathfrak{Bo}(X_n),$$

donde $X = \prod_{n=1}^{\infty} X_n$. El espacio producto $\prod_{n=1}^{\infty} X_n$ con su topología producto $\mathfrak{T}_p$ es, por el Teorema de Tychonoff, un espacio compacto de Hausdorff al que denotaremos por $(2^{\mathbb{N}}, \mathfrak{T}_p)$ y llamaremos el ***espacio de Cantor***. El conjunto $2^{\mathbb{N}}$ puede ser identificado con el conjunto de todas las sucesiones $(x_n)_{n=1}^{\infty}$ donde $x_n \in \{0,1\}$ para todo $n \in \mathbb{N}$. El Principio del Buen-Orden en $\mathbb{N}$, nos permite considerar la aplicación $d : 2^{\mathbb{N}} \times 2^{\mathbb{N}} \to [0, +\infty)$ dada por

$$d(x,y) \;=\; \begin{cases} 2^{-\operatorname{mín}\{n \in \mathbb{N} \,:\, x_n \neq y_n\}} & \text{si } x \neq y \\ 0 & \text{si } x = y \end{cases}$$

donde $x = (x_n)_{n=1}^{\infty}$ y $y = (y_n)_{n=1}^{\infty}$ elementos arbitrarios de $2^{\mathbb{N}}$. Lo que vamos a demostrar es que la topología producto $\mathfrak{T}_p$ es, efectivamente, determinada por la métrica d. Veamos antes que:

(1) *d es una métrica sobre* $2^{\mathbb{N}}$. Sólo resta verificar la desigualdad triangular. Sean $x, y, z \in 2^{\mathbb{N}}$. Si $x = y$ o $y = z$, entonces claramente $d(x,z) \leq d(x,y) + d(y,z)$ es cierta. Suponga entonces que $x \neq y$ y $y \neq z$ y defina

$$n_0 \;=\; \operatorname{mín}\{n \in \mathbb{N} \,:\, x_n \neq y_n\} \qquad \text{y} \qquad n_1 \;=\; \operatorname{mín}\{n \in \mathbb{N} \,:\, y_n \neq z_n\}.$$

Observe que para cualquier $i = 1, 2, \ldots, \operatorname{mín}\{n_0, n_1\} - 1$ se tiene que $x_i = z_i$, de donde se sigue que

$$d(x,z) \leq \frac{1}{2^{\operatorname{mín}\{n_0, n_1\}}} \;\leq\; \frac{1}{2^{n_0}} + \frac{1}{2^{n_1}} \;=\; d(x,y) \,+\, d(y,z)$$

y termina la prueba.

Como antes, sea $\bigotimes_{n=1}^{\infty}\mathcal{A}_n$ la σ-álgebra producto. Los **cilindros** en $2^{\mathbb{N}}$ son particularmente fáciles de describir: ellos siempre son de la forma $C = 2^{\mathbb{N}}$ o $C = C_{n_1,\dots,n_k}(a_1,\dots,a_k)$, donde

$$C_{n_1,\dots,n_k}(a_1,\dots,a_k) \;=\; \Big\{(x_n)_{n=1}^{\infty}\in 2^{\mathbb{N}} \;:\; x_{n_1}=a_1,\dots,x_{n_k}=a_k\Big\},$$

para algún $k\in\mathbb{N}$, $n_1<n_2<\cdots<n_k$ y $a_1,\dots,a_k\in\{0,1\}$.

(2) *Cada bola d-abierta es un cilindro*. Para ver esto, sean $x=(x_n)_{n=1}^{\infty}\in 2^{\mathbb{N}}$ y $r>0$.

(a_1) Si $r\geq 1$, entonces

$$U(x,r) \;=\; \{y\in 2^{\mathbb{N}} \;:\; d(x,y)<r\} \;=\; 2^{\mathbb{N}}.$$

(a_2) Suponga ahora que $0<r<1$ y sea $y\in U(x,r)$ con $x\neq y$. Sea n el primer entero no negativo tal que $x_{n+1}\neq y_{n+1}$. Entonces $d(x,y)=2^n<r$ y, en consecuencia, $y\in C_{1,\dots,n}(x_1,\dots,x_n)$. Esto prueba que

$$U(x,r) \;\subseteq\; C_{1,\dots,n}(x_1,\dots,x_n)$$

Por supuesto, como cada cilindro es claramente una bola abierta, (2) queda demostrado.

(3) $(2^{\mathbb{N}},d)$ *es compacto*. La aplicación identidad $\mathrm{Id}:(2^{\mathbb{N}},\mathfrak{T}_p)\to(2^{\mathbb{N}},d)$ es, por supuesto, biyectiva. Puesto que la colección de todas las uniones arbitrarias de cilindros coincide con la topología producto, se sigue de (2) que Id también es continua. Un llamado al Teorema 3.1.18, página 163, nos convence que $(2^{\mathbb{N}},d)$ es compacto.

Combinando todos estos hechos da como resultado el siguiente:

Teorema 11.4.4. *Si denotamos por $\mathcal{O}_d$ la familia de todos los conjuntos d-abiertos en $(2^{\mathbb{N}},d)$, resulta que $\tau(\mathcal{O}_d)$, la topología generada por $\mathcal{O}_d$, coincide con la topología producto $\mathfrak{T}_p$ de $2^{\mathbb{N}}$.*

11.4.3. Una Medida sobre $2^{\mathbb{N}}$

Puede considerarse que los resultados "posibles" al arrojar una moneda simétrica son cara o sello que designaremos por 0 y 1 respectivamente. Desde el punto de vista probabilístico, usualmente uno asigna la misma probabilidad de $1/2$ a la ocurrencia de los sucesos 0 y 1, es decir, el experimento podría interpretarse mediante un espacio de probabilidad $(X,\mathscr{P}(X),\nu_1)$ donde $X=\{0,1\}$ y $\nu_1(\{0\})=\nu_1(\{1\})=\frac{1}{2}$. Si ahora consideramos el lanzamiento de la misma moneda dos veces, los resultados posibles serán $(0,0),(0,1),(1,0),(1,1)$ y la probabilidad asignada a cada suceso es igual a $1/4$. En tal supuesto, la probabilidad de obtener por lo menos una cara (= 0) será $3/4$. Este experimento puede ser interpretado por medio del espacio de probabilidad $(X\times X,\mathscr{P}(X)\otimes\mathscr{P}(X),\nu_2)$ donde

$$\nu_2(A) \;=\; \frac{\mathrm{card}(A)}{2^2} \quad \text{para cada } A\in\mathscr{P}(X)\otimes\mathscr{P}(X).$$

Podemos, por supuesto, continuar indefinidamente con este proceder, pero, ¿a dónde nos conduce tal situación? Pues bien, una respuesta es esta: la Teoría de Probabilidad estudia experimentos que se repiten un número "muy grande" de veces, de modo que resulta razonable, y altamente deseable, disponer de un espacio de probabilidad $(\prod_{n=1}^{\infty}X_n,\mathcal{S}_\pi,\nu)$ que "contenga" a todos los espacios $(\prod_{k=1}^{n}X_k,\bigotimes_{k=1}^{n}\mathscr{P}(X),\nu_n)$ y además, que tales espacios sean, de algún modo, "buenas aproximaciones" de $(\prod_{n=1}^{\infty}X_n,\mathcal{A},\nu)$ para n suficientemente grande.

Hasta este momento sólo hemos podido confeccionar el espacio medible $(2^{\mathbb{N}}, \mathcal{S}_\pi)$. Nos falta construir ν. Esto, por supuesto, es la parte "dura" de la construcción. Sin embargo, ya disponemos de algunos resultados fundamentales que junto con otros que desarrollaremos un poco más abajo nos permitirán lograr nuestro objetivo. Para comenzar, vamos a asignar a cada cilindro C un valor bien definido: esta es la parte fácil. En efecto, puesto que todo cilindro en $2^{\mathbb{N}}$ es de la forma $C = 2^{\mathbb{N}}$ o $C = C_{n_1,\ldots,n_k}(a_1,\ldots,a_k)$, la única posibilidad que tenemos es definir

$$\nu_{\mathcal{C}}(2^{\mathbb{N}}) \;=\; 1 \qquad \text{o} \qquad \nu_{\mathcal{C}}\big(C_{n_1,\ldots,n_k}(a_1,\ldots,a_k)\big) \;=\; \underbrace{\frac{1}{2}\times\frac{1}{2}\times\cdots\times\frac{1}{2}}_{k\text{ veces}} \;=\; \frac{1}{2^k} \tag{2}$$

¿Pero, cómo se extiende esta definición a *todos* los elementos de $\mathcal{S}_\pi$? Por ejemplo, para establecer un resultado de alto calibrador en Probabilidad, conocido con el nombre de la Ley de los Grandes Números, se requiere computar la medida del conjunto

$$A \;=\; \Big\{x \in 2^{\mathbb{N}} \,:\, \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} x_k = \frac{1}{2}\Big\},$$

el cual no es, por supuesto, un cilindro. Sin embargo, como la colección de todos los cilindros forman una álgebra y ella genera la σ-álgebra $\mathcal{S}_\pi$, podemos usar el siguiente resultado, el cual ya fue demostrado, Teorema 11.1.21, página 696.

Teorema 11.4.5 (Extensión de Carathéodory). *Sean Ω un conjunto no vacío y $\mathcal{A}$ una **álgebra de subconjuntos** de Ω. Si $m : \mathcal{A} \to [0,+\infty)$ es una pre-medida, entonces existe una única medida $\widetilde{m} : \sigma(\mathcal{A}) \to [0,+\infty)$ tal que $\widetilde{m}(A) = m(A)$ para todo $A \in \mathcal{A}$.*

En otras palabras, este resultado establece que toda medida numerablemente aditiva definida sobre una álgebra $\mathcal{A}$ se puede extender a una medida numerablemente aditiva sobre la σ-álgebra generada por $\mathcal{A}$. De modo que, para poder aplicar el Teorema de Extensión de Carathéodory necesitaremos verificar que $\nu_{\mathcal{C}}$ es, efectivamente, numerablemente aditiva sobre el álgebra $\mathcal{C}$ de todos los cilindros.

Lema 11.4.6. *$\nu_{\mathcal{C}} : \mathcal{C} \to [0,+\infty)$ es **numerablemente aditiva** sobre $\mathcal{C}$.*

Prueba. Claramente $\nu_{\mathcal{C}}$ es finitamente aditiva. Para demostrar nuestro lema, nos apoyaremos en el Teorema 11.1.6, página 684. Suponga entonces que $(C_n)_{n=1}^{\infty}$ es una sucesión decreciente de conjuntos en $\mathcal{C}$ tal que $\bigcap_{n=1}^{\infty} C_n = \varnothing$. Cada C_n es una union finita de cilindros y puesto que los cilindros son conjuntos cerrados, resulta que C_n es cerrado para todo $n \geq 1$. Pero como $(2^{\mathbb{N}}, \mathfrak{T}_p)$ es compacto, se sigue de la Propiedad de Intersección Finita que $C_{n_1} \cap \ldots \cap C_{n_k} = \varnothing$ para alguna colección finita de índices $n_1,\ldots,n_k$. Sin embargo, como la sucesión $(C_n)_{n=1}^{\infty}$ es decreciente, debe ocurrir que $C_n = \varnothing$, para todo $n \geq n_k$ y, en consecuencia, $\lim_{n\to\infty}\nu_{\mathcal{C}}(C_n) = 0$. Un llamado al Teorema 11.1.6 nos revela que $\nu_{\mathcal{C}}$ es una medida sobre $\mathcal{C}$. ■

Estamos en condiciones de garantizar la existencia de una medida ν sobre el espacio medible $(2^{\mathbb{N}}, \mathcal{S}_\pi)$.

Teorema 11.4.7 (Existencia de ν). *Existe una **única medida** ν sobre $(2^{\mathbb{N}}, \mathcal{S}_\pi)$ que extiende a $\nu_{\mathcal{C}}$. En particular,*

$$\nu\big(C_{n_1,\ldots,n_k}(a_1,\ldots,a_k)\big) \;=\; \nu_{\mathcal{C}}\big(C_{n_1,\ldots,n_k}(a_1,\ldots,a_k)\big) \;=\; 1/2^k$$

para cada cilindro $C_{n_1,\ldots,n_k}(a_1,\ldots,a_k) \in \mathcal{C}$.

Prueba. Por el Lema 11.4.6 sabemos que $\nu_{\mathcal{C}}$ es numerablemente aditiva sobre la álgebra $\mathcal{C}$ de los conjuntos cilindros. El Teorema de Extensión de Carathéodory entonces nos garantiza la existencia de una única medida ν sobre $\sigma(\mathcal{C}) = \mathcal{S}_\pi$ tal que ν y $\nu_{\mathcal{C}}$ coinciden sobre $\mathcal{C}$. ■

11.5. Ejercicios Resueltos

Ejercicio 11.5.1. *Sea X un conjunto no vacío y suponga que $\mathcal{A}$ es una σ-**álgebra infinita** de X. Pruebe que $\mathcal{A}$ es **no-numerable**.*

Prueba. Suponga, para generar una contradicción, que $\operatorname{card}(\mathcal{A}) = \aleph_0$. Para cada $x \in X$, considere la colección

$$\mathcal{A}_x = \{A \in \mathcal{A} : x \in A\}.$$

Claramente $\mathcal{A}_x$ es a lo más numerable y como $\mathcal{A}$ es una σ-álgebra se tiene que

$$E_x = \bigcap_{A \in \mathcal{A}_x} A \in \mathcal{A}.$$

Observe que E_x es el conjunto más pequeño en $\mathcal{A}$ conteniendo a x. Defina sobre X la siguiente relación $\sim$:

$$x \sim y \qquad \Leftrightarrow \qquad E_x = E_y.$$

Es fácil ver que $\sim$ es reflexiva, simétrica y transitiva, es decir, una relación de equivalencia. Por consiguiente, la familia $\mathcal{A}^* = \{E_x : x \in X\}$ de las clases de equivalencias determinadas por $\sim$ forman una partición de X; vale decir, $X = \bigcup_{x \in X} E_x$ y cualesquiera sean $x, y \in X$ se cumple que

$$E_x \cap E_y = \varnothing \qquad \text{o} \qquad E_x = E_y.$$

Suponga ahora que $A \in \mathcal{A}$. Afirmamos que $A = \bigcup_{x \in A} E_x$. En efecto, en primer nótese que como $\{x\} \subseteq E_x$ para cada $x \in A$, resulta que $A = \bigcup_{x \in A} \{x\} \subseteq \bigcup_{x \in A} E_x$. También, por la minimalidad de E_x se tiene que $E_x \subseteq A$ para cualquier $x \in A$, de donde resulta que $\bigcup_{x \in A} E_x \subseteq A$ y así queda verificada nuestra afirmación. Lo anterior nos muestra que cualquier conjunto $A \in \mathcal{A}$ es una unión disjunta de conjuntos en $\mathcal{A}^*$. La familia $\mathcal{A}^*$, por ser una subfamilia (disjunta) de $\mathcal{A}$, también es numerable. Sea $N = \{n_1, n_2, \ldots\}$ un conjunto infinito numerable de índices tal que $\mathcal{A}^* = \{A_{n_k} : n_k \in N\}$. Por lo visto anteriormente, tenemos que si $I \subseteq N$, entonces

$$E_I = \bigcup_{n_k \in I} E_{n_k} \in \mathcal{A}$$

y, además,

$$E_I \neq E_J \quad \text{para cualesquiera } I, J \subseteq N \quad \text{con} \quad I \neq J.$$

Finalmente, la aplicación $\psi : 2^N \to \mathcal{A}$ definida por $\psi(I) = E_I$ es, por la observación anteriormente, una biyección. Esto, por supuesto, es lo que genera la contradicción ya que 2^N es no-numerable, mientras $\mathcal{A}$ es numerable por nuestra suposición. Por esto, $\mathcal{A}$ es no-numerable.■

Ejercicio 11.5.2. *Sea (X, τ) un espacio topológico. Pruebe que la familia*

$$\mathcal{A} = \{A \subseteq X : A \text{ o } X \setminus A \text{ es de \textbf{primera categoría}}\}.$$

*es una σ-**álgebra**.*

Prueba. Observe que $\varnothing, X \in \mathcal{A}$ y que $\mathcal{A}$ es cerrada bajo complementos. Considere ahora una sucesión $(A_n)_{n=1}^{\infty}$ en $\mathcal{A}$. Si cada A_n es de primera categoría, entonces $\bigcup_{n=1}^{\infty} A_n$ también lo es y, así,

$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}.$$

En caso contrario, existe un n_0 tal que $X \setminus A_{n_0}$ es de primera categoría y entonces la inclusión

$$X \setminus \bigcup_{n=1}^{\infty} A_n \subseteq X \setminus A_{n_0}$$

implica que $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$. Esto prueba que $\mathcal{A}$ es una σ-álgebra la cual es llamada la σ-álgebra **generada por todos los subconjuntos nunca-densos de** X. ∎

Ejercicio 11.5.3. *Sea (X, d) un espacio métrico y τ_d la topología generada por la métrica d. Demuestre que si τ_d es una σ-álgebra, entonces X es un espacio discreto.*

Prueba. Si τ_d es una σ-álgebra, entonces para cada $x \in X$ se cumple que

$$\{x\} = \bigcap_{n=1}^{\infty} U(x, 1/n) \in \tau_d$$

donde $U(x, 1/n)$ es la bola abierta con centro en x y radio $1/n$. Esto nos dice que cada conjunto $\{x\}$ es abierto y, por lo tanto, $\tau = \mathscr{P}(X)$. ∎

Ejercicio 11.5.4. *Sea $(X, \mathcal{A})$ un espacio medible y suponga que $\nu_n : \mathcal{A} \to [0, +\infty]$ es, para cada $n \in \mathbb{N}$, una medida satisfaciendo*

$$\nu_n(A) \leq \nu_{n+1}(A) \quad \textit{para todo } A \in \mathcal{A} \quad \textit{y todo} \quad n \geq 1.$$

Pruebe que la función de conjuntos $\nu : \mathcal{A} \to [0, +\infty]$ dada por

$$\nu(A) = \sup\{\nu_n(A) : n \in \mathbb{N}\}, \quad A \in \mathcal{A},$$

es una medida.

Prueba. Claramente $\nu(\varnothing) = 0$. Sea $(A_n)_{n=1}^{\infty}$ una sucesión disjunta en $\mathcal{A}$ y pongamos $A = \bigcup_{n=1}^{\infty} A_n$. Entonces

$$\nu_k(A) = \sum_{n=1}^{\infty} \nu_k(A_n) \leq \sum_{n=1}^{\infty} \nu(A_n)$$

y, por lo tanto,

$$\nu(A) \leq \sum_{n=1}^{\infty} \nu(A_n).$$

Por otro lado, como $\nu_1(A) \leq \nu_2(A) \leq \cdots$ para cada $A \in \mathcal{A}$, entonces la igualdad

$$\sup\{\nu_k(A) : k \in \mathbb{N}\} = \lim_{k \to \infty} \nu_k(A)$$

es válida en $[0, +\infty]$ y, en consecuencia, para cualquier entero $m \geq 1$, se tiene que

$$\sum_{n=1}^{m} \nu(A_n) \;=\; \lim_{k\to\infty} \sum_{n=1}^{m} \nu_k(A_n) \;=\; \lim_{k\to\infty} \nu_k\Big(\bigcup_{n=1}^{m} A_n\Big) \;\leq\; \nu(A).$$

De lo anterior se concluye que $\sum_{n=1}^{\infty} \nu(A_n) \leq \nu(A)$ y termina la prueba. ■

Ejercicio 11.5.5. *Sea (X, d) un espacio métrico y suponga que λ, ν son dos* ***medidas finitas*** *definidas sobre $\mathfrak{Bo}(X)$. Pruebe que si $\lambda(E) = \nu(E)$ para cualquier $E \subseteq X$ que sea abierto o cerrado, entonces $\lambda = \nu$.*

Prueba. Recordemos que la familia $\mathcal{A} = \mathcal{O} \cup \mathcal{F}$ es una álgebra, donde $\mathcal{O}$ y $\mathcal{F}$ son, respectivamente, las colecciones de los conjuntos abiertos y cerrados de X y que $\sigma(\mathcal{A}) = \mathfrak{Bo}(X)$.

Considere ahora la familia

$$\mathcal{C} \;=\; \big\{E \in \mathfrak{Bo}(X) \,:\, \lambda(E) = \nu(E)\big\}.$$

Por hipótesis, $\mathcal{A} \subseteq \mathcal{C}$. Veamos que $\mathcal{C}$ es una clase monótona. Sea $(E_n)_{n=1}^{\infty}$ es una sucesión creciente en $\mathcal{C}$. Entonces $\lambda(E_n) = \nu(E_n)$ para todo $n \geq 1$ y se sigue de la continuidad de λ y ν que

$$\lambda\Big(\bigcup_{n=1}^{\infty} E_n\Big) \;=\; \lim_{n\to\infty} \lambda(E_n) \qquad \text{y} \qquad \nu\Big(\bigcup_{n=1}^{\infty} E_n\Big) \;=\; \lim_{n\to\infty} \nu(E_n)$$

es decir,

$$\lambda\Big(\bigcup_{n=1}^{\infty} E_n\Big) \;=\; \nu\Big(\bigcup_{n=1}^{\infty} E_n\Big).$$

Por esto, $\bigcup_{n=1}^{\infty} E_n \in \mathcal{C}$. De forma similar, si $(E_n)_{n=1}^{\infty}$ es una sucesión decreciente en $\mathcal{C}$, entonces usando el hecho de que ambas medidas son finitas y continuas, se obtiene que

$$\lambda\Big(\bigcap_{n=1}^{\infty} E_n\Big) \;=\; \nu\Big(\bigcap_{n=1}^{\infty} E_n\Big)$$

y, por lo tanto, $\bigcap_{n=1}^{\infty} E_n \in \mathcal{C}$. Se sigue del Teorema de la Clase Monótona que

$$\mathcal{C} \;=\; \mathfrak{Bo}(X)$$

y, en consecuencia, $\lambda = \nu$. ■

Ejercicio 11.5.6. *Sea $(X, \mathcal{A})$ un espacio medible y sea $(\lambda_n)_{n=1}^{\infty}$ una sucesión de* ***medidas finitas***. *Si $(\varepsilon_n)_{n=1}^{\infty}$ es una sucesión de números positivos tal que*

$$\sum_{n=1}^{\infty} \varepsilon_n \lambda_n(X) \;<\; +\infty,$$

pruebe entonces que la función de conjuntos $\lambda : \mathcal{A} \to [0, +\infty]$ definida por

$$\lambda(A) \;=\; \sum_{n=1}^{\infty} \varepsilon_n \lambda_n(A), \quad A \in \mathcal{A}$$

es una medida finita.

Prueba. Claramente $\lambda(\varnothing) = 0$ y para cualquier $A \in \mathcal{A}$ se tiene que

$$\lambda(A) \;\leq\; \lambda(X) \;<\; +\infty.$$

Suponga ahora que $(A_n)_{n=1}^{\infty}$ es una sucesión disjunta en $\mathcal{A}$ y pongamos $A = \bigcup_{n=1}^{\infty} A_n$. Entonces, para cada $n \in \mathbb{N}$, se tiene que

$$\lambda_n(A) \;=\; \sum_{j=1}^{\infty} \lambda_n(A_j)$$

y, por lo tanto,

$$\sum_{n=1}^{m} \varepsilon_n \lambda_n(A) \;=\; \sum_{n=1}^{m} \varepsilon_n \sum_{j=1}^{\infty} \lambda_n(A_j) \quad \text{para todo } m \geq 1.$$

De aquí se sigue que

$$\lambda(A) \;=\; \sum_{n=1}^{\infty} \varepsilon_n \lambda_n(A) \;=\; \sum_{n=1}^{\infty}\sum_{j=1}^{\infty} \varepsilon_n \lambda_n(A_j) \;=\; \sum_{j=1}^{\infty}\sum_{n=1}^{\infty} \varepsilon_n \lambda_n(A_j) \;=\; \sum_{j=1}^{\infty} \lambda(A_j).$$

Esto prueba que λ es numerablemente aditiva. ■

Ejercicio 11.5.7. *Encuentre una medida de Borel* $\lambda : \mathfrak{Bo}(\mathbb{R}) \to [0, +\infty]$ *tal que*

$$\int_{\mathbb{R}} f \, d\lambda \;=\; f(0) \quad \textit{para cualquier } f \in C_c(\mathbb{R}).$$

Prueba. Defina λ como

$$\lambda(A) \;=\; \chi_A(0) \quad \text{para todo } A \in \mathfrak{Bo}(\mathbb{R}).$$

Entonces, para cualquier $f \in C_c(\mathbb{R})$, se tiene que $\int_{\mathbb{R}} f \, d\lambda = f(0)$. ■

Ejercicio 11.5.8. *Sea* $f : [0, +\infty) \to \mathbb{R}$ *una función continua tal que* $\lim\limits_{x\to\infty} f(x) = \theta \in \mathbb{R}$. *Pruebe que*

$$\lim_{n\to\infty} \int_0^a f(nx)\, dx \;=\; a\theta \quad \textit{para todo } a > 0.$$

Prueba. Para cada $n \geq 1$, considere la función $g_n : [0, +\infty) \to \mathbb{R}$ definida por

$$g_n(x) \;=\; f(nx) \quad \text{para todo } x \geq 0.$$

Sea $a > 0$ y observe que, para todo $x \in (0, a]$, se tiene que $\lim_{n\to\infty} g_n(x) = \theta$. Además, como $\lim_{x\to\infty} f(x) = \theta$, podemos determinar un $M > 0$ tal que

$$|f(x)| \;\leq\; |\theta| + 1 \quad \text{para todo } x \geq M.$$

Por otro lado, como $[0, M]$ es un conjunto compacto, la continuidad de f nos dice que ella es acotada sobre $[0, M]$; es decir, existe un $\delta > 0$ tal que

$$|f(x)| \;\leq\; \delta \quad \text{para todo } x \in [0, M].$$

Sea $K = \max\{\delta, |\theta| + 1\}$. Entonces

$$|f(x)| \;\leq\; K \quad \text{para todo } x \in (0, +\infty)$$

y, por consiguiente,

$$|g_n(x)| \;=\; |f(nx)| \;\leq\; K \quad \text{para todo } x \in [0, a].$$

Invocando al Teorema de la Convergencia Dominada se tiene que

$$\lim_{n\to\infty} \int_0^a f(nx)\,dx \;=\; \lim_{n\to\infty} \int_0^a g_n(x)\,dx \;=\; \int_0^a \theta\,dx \;=\; a\theta.$$

Fin de la prueba. ■

Ejercicio 11.5.9. *Sea $(X, \mathcal{A}, \lambda)$ un espacio de medida, $f \in \mathrm{L}_1(X, \lambda)$ con $f > 0$ y suponga que $g : X \to \mathbb{R}$ es una función $\mathcal{A}$-medible tal que, para algún par $a, b \in \mathbb{R}$, $a \leq g(x) \leq b$ para casi-todo $x \in X$. Pruebe que existe $c \in [a, b]$ tal que*

$$\int_X fg\,d\lambda \;=\; c\int_X f\,d\lambda.$$

Prueba. En primer lugar, nótese que $|fg(x)| \leq \max\{|a|, |b|\}|f(x)|$ para casi-todo $x \in X$, de donde se deduce que $fg \in \mathrm{L}_1(X, \lambda)$. Más aun, como

$$a\int_X f\,d\lambda \;\leq\; \int_X fg\,d\lambda \;\leq\; b\int_X f\,d\lambda$$

si definimos $M = \int_X f\,d\lambda$, resulta que $M > 0$ y

$$a \;\leq\; \frac{\displaystyle\int_X fg\,d\lambda}{M} \;\leq\; b.$$

Esto termina la prueba. ■

11.6. Problemas

(1) Sea X un conjunto arbitrario. Un subconjunto A de X se dice que es **co-finito** si su complemento A^c es un subconjunto finito de X. Designe por $\mathcal{A}_{\mathrm{co}}$ la familia de todos los subconjuntos finitos y co-finitos de X. Demuestre que:

   (a) $\mathcal{A}_{\mathrm{co}}$ es una álgebra de subconjuntos de X.

   (b) $\mathcal{A}_{\mathrm{co}}$ es una σ-álgebra si, y sólo si X es un conjunto finito.

(2) Sea X un conjunto infinito y sea $\lambda : \mathcal{A}_{\mathrm{co}} \to [0, 1]$ definida por

$$\lambda(A) \;=\; \begin{cases} 1 & \text{si } A \text{ es finito,} \\ 0 & \text{si } A \text{ es co-finito.} \end{cases}$$

Demuestre que:

(a) λ es finitamente aditiva.

(b) Si X es infinito numerable, entonces λ no es numerablemente aditiva.

(c) Si X es no-numerable, entonces λ es numerablemente aditiva.

(3) Sea $(\mathcal{A}_n)_{n=1}^{\infty}$ una **sucesión creciente** de familias de subconjuntos de un conjunto X.

(a) Pruebe que si cada $\mathcal{A}_n$ es una álgebra, entonces $\bigcup_{n=1}^{\infty} \mathcal{A}_n$ también lo es.

(b) De un ejemplo para mostrar que si cada $\mathcal{A}_n$ es una σ-álgebra, entonces $\bigcup_{n=1}^{\infty} \mathcal{A}_n$ no es una σ-álgebra.

(4) Sea X un conjunto infinito y sea $\lambda : \mathscr{P}(X) \to [0, +\infty]$ la medida de conteo. Demuestre que existe una sucesión decreciente $(E_n)_{n=1}^{\infty}$ en $\mathscr{P}(X)$ tal que

$$\lim_{n\to\infty} E_n \;=\; \varnothing \qquad \text{pero} \qquad \lim_{n\to\infty} \lambda(E_n) \;\neq\; 0.$$

(5) Sea X un conjunto no vacío y λ^* una medida exterior definida sobre $\mathscr{P}(X)$. Pruebe que las siguientes condiciones son equivalentes:

(1) λ^* es finitamente aditiva sobre $\mathscr{P}(X)$.

(2) Cualquier elemento de $\mathscr{P}(X)$ es λ^*-medible; es decir, $\mathcal{M}_{\lambda^*}(X) = \mathscr{P}(X)$.

(3) λ^* es numerablemente aditiva sobre $\mathscr{P}(X)$.

(6) Sea $(X, \mathcal{A})$ un espacio de medida y ν una función de conjuntos aditiva. Defina $\lambda : \mathcal{A} \to [0, +\infty]$ por

$$\lambda(E) \;=\; \inf\Big\{ \lim_{n\to\infty} \nu(E_n) \;:\; (C_n)_{n=1}^{\infty} \in \mathcal{C}_E \Big\},$$

donde $\mathcal{C}_E$ es la familia de todas las **sucesiones crecientes** $(E_n)_{n=1}^{\infty}$ incluidas en $\mathcal{A}$ tal que $E = \bigcup_{n=1}^{\infty} E_n$. Pruebe que λ es una medida.

(7) Sea $(X, \mathcal{A})$ un espacio medible. Demuestre que son equivalentes:

(1) $\mathcal{A}$ es numerablemente generada.

(2) Existe una función $\mathcal{A}$-medible $f : X \to [0,1]$ tal que $\mathcal{A} = f^{-1}(\mathfrak{B}o([0,1]))$.

(8) Suponga que λ es una medida de Borel sobre $\mathbb{R}$ tal que:

(a) $\lambda([0,1]) = 1$ y

(b) $\lambda(x + A) = \lambda(A)$ para todo $A \in \mathfrak{B}o(\mathbb{R})$ y todo $x \in \mathbb{R}$.

Pruebe que $\lambda = \mu$ (μ la medida de Lebesgue sobre $\mathbb{R}$).

(9) Sea X un **conjunto compacto**. Una medida de Borel μ sobre X se llama **puramente irregular** si no existe ninguna medida de Borel regular ν, con $\nu \neq 0$, tal que $0 \leq \nu(E) \leq \mu(E)$ para cualquier conjunto de Borel $E \subseteq X$.

(a) Use el Lema de Zorn para demostrar que cualquier medida de Borel μ sobre X es de la forma $\mu = \lambda + \nu$, donde λ es una medida de Borel regular y ν es puramente irregular.

(b) Pruebe que si μ es una medida de Borel regular y ν es puramente irregular con $0 \leq \nu \leq \mu$, entonces $\nu = 0$.

(c) Use la Descomposición de Jordan para demostrar que la descomposición en el problema (a) es única.

(10) Asumiendo **HC**, pruebe que existe un conjunto $A \subseteq [0,1] \times [0,1]$ tal que:

(i) para cada $x \in [0,1]$, el conjunto

$$A_x = \{y \in [0,1] : (x,y) \in A\} \quad \text{es numerable,}$$

(ii) para cada $y \in [0,1]$, el conjunto

$$A^y = \{x \in [0,1] : (x,y) \in A\} \quad \text{es co-numerable (su complemento es numerable),}$$

pero A no es medible según Lebesgue.

(11) Sea $(X, \mathcal{A}, \lambda)$ un espacio de medida finita y sean $f, g : X \to \mathbb{R}$ funciones $\mathcal{A}$-medibles tal que

$$\int_X f \, d\lambda = \int_X g \, d\lambda.$$

Demuestre que una, y sólo una, de las siguientes alternativas se cumple:

(a) $f = g$ λ-c.s.

(b) Existe $A \in \mathcal{A}$ tal que

$$\int_A f \, d\lambda < \int_X g \, d\lambda.$$

(12) Demuestre que el Teorema de la Convergencia Dominada de Lebesgue no se cumple para redes de funciones.

(13) Sea $f \in L_1(\mu)$. Demuestre que la función $F : \mathbb{R} \to \mathbb{R}$ definida por

$$F(x) = \int_{-\infty}^{x} f(t) \, d\mu(t)$$

es absolutamente continua sobre $\mathbb{R}$ y $\lim\limits_{x \to -\infty} F(x) = 0$.

CAPÍTULO 12

El Teorema de Radon-Nikodým

En la Teoría de los Espacios de Hilbert, cada funcional lineal continuo queda completamente determinado a través de su producto escalar, esto es, si $\mathcal{H}$ es un espacio de Hilbert y $\Lambda : \mathcal{H} \to \mathbb{K}$ es un *funcional lineal continuo*, entonces existe un *único* $x_0 \in \mathcal{H}$ tal que

$$\Lambda(x) \;=\; \langle x, x_0\rangle \quad \text{para todo} \quad x \in \mathcal{H} \tag{1}$$

y recíprocamente, es decir, cada $x_0 \in \mathcal{H}$ da origen a un único funcional lineal continuo sobre $\mathcal{H}$ definido por la ecuación (1). Otro resultado que también identifica a los funcionales lineales continuos en el espacio dado es el siguiente. Sea X un espacio métrico compacto y sea $(C(X), \|\cdot\|_\infty)$ el espacio de Banach formado por todas las funciones continuas $f : X \to \mathbb{R}$. Un famoso resultado atribuido a F. Riesz establece que cada *funcional lineal continuo* $\Lambda : C(X) \to \mathbb{R}$ se identifica con una única *medida con signo* μ por medio de la siguiente identidad:

$$\Lambda(f) \;=\; \int_X f\, d\mu \quad \text{para toda} \quad f \in C(X).$$

A estos dos resultados se les conoce bajo el mismo nombre de **Teorema de Representación de Riesz**. En el caso de la Teoría de la Medida ya habíamos visto que cada función $f \in L_1(\mu)$ determina una *única medida* ν expresada en la forma

$$\nu(E) \;=\; \int_E |f|\, d\mu \quad \text{para todo } E \in \mathfrak{M}_\mu(\mathbb{R}), \tag{2}$$

y que ambas medidas ν y μ son "*comparables*" en el sentido de que $\nu \ll \mu$, en otras palabras, si $E \in \mathfrak{M}_\mu(\mathbb{R})$ y $\mu(E) = 0$, entonces $\nu(E) = 0$. Existe un extraordinario y poderoso resultado al que se le llama **Teorema de Radon-Nikodým**, o también **Teorema de Lebesgue-Radon-Nikodým**, el cual nos concede el privilegio de poder andar en la otra dirección, es decir, cualquier medida σ-finita, digamos ν, que pueda ser comparada con μ, permite que ν se pueda representar de modo único por la igualdad (2) para alguna función no-negativa y μ-integrable f. En general, cualquier par de medidas σ-finitas, digamos ν y μ, definidas sobre un espacio medible $(X, \mathcal{M})$

y comparables en el sentido anterior, quedan relacionadas por una igualdad similar a la anterior. Este hecho es de tal magnitud que todo el desarrollo posterior de la Teoría de la Dualidad de los Espacios $L_p(\mu)$ fue posible gracias a dicho resultado. Pero además, la existencia de la Esperanza Condicional, la cual es una pieza fundamental para el análisis y desarrollo de las Martingalas en la Teoría de Probabilidad, se soporta sobre el Teorema de Radon-Nikodým. Algunos autores prefieren llamar el Teorema de Radon-Nikodým como **El Teorema Fundamental de la Teoría de la Medida e Integración** similar a como existen El Teorema Fundamental del Álgebra, El Teorema Fundamental de la Aritmética y El Teorema Fundamental del Cálculo.

Nuestros esfuerzos, en esta sección, estarán dirigidos en presentar todos los elementos necesarios para demostrar el Teorema de Radon-Nikodým en su versión más simple y práctica. Por supuesto, enmarcado dentro de sus incuestionables aplicaciones, se han realizado, desde diversos puntos de vista, numerosas demostraciones y generalizaciones de dicho teorema.

El programa para obtener el Teorema de Radon-Nikodým pasa por introducir algunas nociones de medidas abstractas sobre un espacio medible $(X, \mathcal{M})$. La más importante es el concepto de medidas con signo. Ésta noción permite clasificar en dos categorías de conjuntos muy especiales a los elementos de la σ-álgebra $\mathcal{M}$, a los que llamaremos conjuntos positivos y conjuntos negativos. A partir de allí se obtiene una descomposición del conjunto X en dos partes: una positiva y la otra negativa. A tal descomposición se le denomina el Teorema de Descomposición de Hahn. Luego se desarrollan algunos resultados adicionales necesarios que aligeran y acortan la demostración de nuestro teorema fundamental.

12.1. Medidas con Signos e Integración

La evidencia más clara de una "medida con signo" se presenta cuando se elige cualquier función $f \in L_1(\mu)$ y entonces se define

$$\nu(E) \;=\; \int_E f\, d\mu \quad \text{para todo } E \in \mathfrak{M}_\mu(\mathbb{R}) \tag{2}$$

Claramente ν es numerablemente aditiva y la distinción entre ν y μ es precisamente el hecho de que ν puede tomar valores negativos. Otra circunstancia que produce la aparición de medidas con valores reales es la siguiente: Imaginemos que μ_1 y μ_2 son dos medidas definidas sobre un espacio medible $(X, \mathcal{M})$. Podemos formar una nueva medida sobre X tomando

$$\mu_3(E) \;=\; a_1\mu_1(E) + a_2\mu_2(E) \quad \text{para todo } E \in \mathcal{A},$$

donde $a_1, a_2 \geq 0$. Pero, ¿qué ocurre, por ejemplo, si $a_1 = 1$ y $a_2 = -1$? En este caso se tiene que $\mu_3(E) = \mu_1(E) - \mu_2(E)$ y, por supuesto, puede ocurrir que μ_3 tome valores negativos si, por ejemplo, $\mu_1(E) < \mu_2(E)$ para algún $E \in \mathcal{M}$. Sin embargo, aquí se presenta otro problema. En efecto, si se da el caso de la existencia de un conjunto $E \in \mathcal{M}$ para el cual $\mu_1(E) = \mu_2(E) = +\infty$, entonces $\mu_3(E) = \infty - \infty$ y, como sabemos, tal expresión carece de sentido. Estos dos hechos permiten que podamos considerar medidas con valores reales extendidos, es decir, medidas con signos, pero se debe imponer una cierta restricción.

En lo que sigue, asumiremos que $(X, \mathcal{M})$ es un espacio medible, es decir, X es un conjunto arbitrario no vacío y $\mathcal{M}$ es una σ-álgebra de subconjuntos de X. Los elementos de $\mathcal{M}$ se llamarán **conjuntos medibles** o $\mathcal{M}$-**medibles** si se desea enfatizar su procedencia.

Definición 12.1.1. *Sea $(X, \mathcal{M})$ un espacio medible. Una función de conjuntos $\nu : M \to [-\infty, +\infty]$ es llamada una **medida con signo** si ella satisface las siguientes condiciones:*

(a) *ν asume **sólo uno** de los valores $+\infty$ o $-\infty$.*

(b) $\nu(\varnothing) = 0$.

(c) *ν es **numerablemente aditiva**, es decir,*

$$\nu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \nu(E_n)$$

para cualquier sucesión disjunta $(E_n)_{n=1}^{\infty}$ incluida en $\mathcal{M}$.

Nótese que la condición (a) garantiza que la medida de la unión de cualquier par de conjuntos medibles y disjuntos está bien definida, es decir, dicha condición prohíbe, de manera expresa, la existencia de dos conjuntos medibles y disjuntos, digamos A y B, tales que $\nu(A) = +\infty$ y $\nu(B) = -\infty$ o viceversa. En efecto, si tales conjuntos existieran, entonces $\nu(A \cup B) = \nu(A) + \nu(B) = +\infty - \infty$ una expresión que, como sabemos, no está definida. En consecuencia, *cualquier medida con signo tiene su rango en $(-\infty, +\infty]$, o bien en $[-\infty, +\infty)$.* Esta observación obliga, cada vez que fijemos una medida con signo, a especificar si su rango está incluido en $(-\infty, +\infty]$ o en $[-\infty, +\infty)$. Nótese que si ν es una medida con signo cuyo rango está incluido en $[-\infty, +\infty)$, entonces $-\nu$ es una medida con signo cuyo rango está incluido en $(-\infty, +\infty]$. Por tal razón, asumiremos, salvo que se indique expresamente lo contrario, que *cualquier medida con signo ν definida sobre $(X, \mathcal{M})$ satisface $-\infty < \nu(E) \leq +\infty$ para cualquier $E \in \mathcal{M}$.* Observe también que cualquier medida no-negativa $\nu : \mathcal{M} \to [0, +\infty]$ es una medida con signo. En este caso, y para evitar cualquier confusión en el lenguaje, diremos que ν es una **medida positiva**. Por supuesto, toda medida con signo es **finitamente aditiva**.

La condición (c) de la definición anterior establece que para cualquier sucesión disjunta $(E_n)_{n=1}^{\infty}$ incluida en $\mathcal{M}$ se tiene que $\nu(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \nu(E_n) \in (-\infty, +\infty]$, de donde se sigue que:

$$\nu\left(\bigcup_{n=1}^{\infty} E_n\right) \in \mathbb{R} \quad \Leftrightarrow \quad \sum_{n=1}^{\infty} \nu(E_n) \in \mathbb{R}.$$

De hecho, podemos afirmar algo más fuerte, lo cual es consecuencia del Teorema 2.1.51, página 110.

Teorema 12.1.2. *Sea ν una **medida con signo** sobre $(X, \mathcal{M})$. Si $(E_n)_{n=1}^{\infty}$ es una **sucesión disjunta** de conjuntos medibles, entonces*

$$\sum_{n=1}^{\infty} \nu(E_n) = \nu\left(\bigcup_{n=1}^{\infty} E_n\right) \in \mathbb{R} \quad \Leftrightarrow \quad \sum_{n=1}^{\infty} |\nu(E_n)| < +\infty.$$

*En otras palabras, la serie $\sum_{n=1}^{\infty} \nu(E_n)$ **converge** si, y sólo si, ella **converge absolutamente**.*

Prueba. Sea $(E_n)_{n=1}^{\infty}$ una sucesión disjunta de conjuntos medibles. Sabemos que en $\mathbb{R}$ (en realidad, en cualquier espacio de Banach) toda serie absolutamente convergente es convergente, Teorema 2.1.48, página 109. De esto resulta que si $\sum_{n=1}^{\infty} |\nu(E_n)| < +\infty$, entonces $\nu(\bigcup_{n=1}^{\infty} E_n) =$

$\sum_{n=1}^{\infty} \nu(E_n) \in \mathbb{R}$. Recíprocamente, suponga que $\nu(\bigcup_{n=1}^{\infty} E_n) \in \mathbb{R}$ y sea $\sum_{n=1}^{\infty} \nu(F_n)$ un reordenamiento de $\sum_{n=1}^{\infty} \nu(E_n) \in \mathbb{R}$, esto es, $F_n = E_{\pi(n)}$ para todo $n \geq 1$, donde π es una permutación de $\mathbb{N}$. Entonces $(F_n)_{n=1}^{\infty}$ es una sucesión disjunta en $\mathcal{M}$ con $\bigcup_{n=1}^{\infty} F_n = \bigcup_{n=1}^{\infty} E_n$ tal que

$$\sum_{n=1}^{\infty} \nu(F_n) \;=\; \nu\Big(\bigcup_{n=1}^{\infty} F_n\Big) \;=\; \nu\Big(\bigcup_{n=1}^{\infty} E_n\Big) \in \mathbb{R}.$$

Esto prueba que la serie $\sum_{n=1}^{\infty} \nu(E_n)$ es incondicionalmente convergente y, así, gracias al Teorema 2.1.51, $\sum_{n=1}^{\infty} |\nu(E_n)| < +\infty$. Fin de la prueba. ■

Puesto que la serie $\sum_{n=1}^{\infty} (-1)^n/n$ es convergente, pero no absolutamente convergente, se sigue del resultado anterior que ***no existe ningún espacio de medida con signo*** $(X, \mathcal{M}, \nu)$ tal que $\nu(E_n) = (-1)^n \frac{1}{n}$ para todo $n \geq 1$, donde $(E_n)_{n=1}^{\infty}$ es una sucesión disjunta de conjuntos medibles.

A diferencia de las medidas positivas, las medidas con signo no son, en general, monótonas, es decir, si ν es una medida con signo sobre $(X, \mathcal{M})$ y si E_1, E_2 son conjuntos medibles con $E_1 \subseteq E_2$, entonces no siempre ocurre que $\nu(E_1) \leq \nu(E_2)$. Por ejemplo, sea $f(x) = \cos(x)$ para todo $x \in [0, \pi]$ y considere la medida con signo

$$\nu(E) \;=\; \int_E f \, d\mu,$$

donde μ es la medida de Lebesgue restringida a los subconjuntos medibles de $[0, \pi]$. Si tomamos $E_1 = [0, \pi/2]$ y $E_2 = [0, \pi]$, resulta que $E_1 \subseteq E_2$, pero

$$1 \;=\; \nu(E_1) \;=\; \int_{E_1} \cos(x) \, d\mu(x) \;>\; \int_{E_2} \cos(x) \, d\mu(x) \;=\; 0.$$

La ausencia de la propiedad de monotonía en una medida con signo puede ser, en cierta medida, recompensada con los siguientes hechos:

Lema 12.1.3. *Sea ν una **medida con signo** sobre $(X, \mathcal{M})$. Sean $E, F \in \mathcal{M}$ con $E \subseteq F$.*

(a) Si $\nu(F) \in \mathbb{R}$, entonces $\nu(E) \in \mathbb{R}$.

(b) Si $\nu(E) = +\infty$, entonces $\nu(F) = +\infty$.

Prueba. (a) Observe que, por la aditividad de ν, se tiene que $\nu(F) = \nu(E) + \nu(F \setminus E)$, de modo que si $\nu(F)$ es finito, entonces tanto $\nu(E)$ así como $\nu(F \setminus E)$ también lo son.

(b) Si $\nu(E) = +\infty$, entonces $\nu(F) \notin \mathbb{R}$ ya que de otro modo tendríamos, por la parte (a), que $\nu(E) \in \mathbb{R}$. Entonces $\nu(F) \in \{-\infty, +\infty\}$. Pero como estamos asumiendo que el rango de ν está incluido en $(-\infty, +\infty]$, resulta que $\nu(F) = +\infty$. Esto prueba (b). ■

Observe que si el rango de ν está incluido en $[-\infty, +\infty)$, entonces la condición (b) del lema anterior debe ser reemplazada por:

$(b)'$ Si $\nu(E) = -\infty$, entonces $\nu(F) = -\infty$.

Similar a la demostración de los Teoremas 6.3.45 y 6.3.47 se obtiene el siguiente resultado para medidas con signo.

Teorema 12.1.4 (Continuidad de ν). *Sea ν una* ***medida con signo*** *sobre un conjunto medible $(X, \mathcal{M})$.*

(a) Si $(E_n)_{n=1}^{\infty}$ es una ***sucesión creciente*** *de conjuntos medibles, entonces*

$$\nu\bigg(\bigcup_{n=1}^{\infty} E_n\bigg) \;=\; \lim_{n\to\infty} \nu(E_n).$$

(b) Si $(E_n)_{n=1}^{\infty}$ es una ***sucesión decreciente*** *de conjuntos medibles con $\nu(E_1) \in \mathbb{R}$, entonces*

$$\nu\bigg(\bigcap_{n=1}^{\infty} E_n\bigg) \;=\; \lim_{n\to\infty} \nu(E_n).$$

Prueba. (a) Sea $(E_n)_{n=1}^{\infty}$ una sucesión creciente de conjuntos medibles y sea $\lim_{n\to\infty} E_n = \bigcup_{n=1}^{\infty} E_n$. Si $\nu(E_{n_0}) = +\infty$ para algún $n_0 \in \mathbb{N}$, entonces puesto que $E_{n_0} \subseteq E_m \subseteq \bigcup_{n=1}^{\infty} E_n$ para cualquier $m \geq n_0$, se sigue del Lema 12.1.3 (a) que $\nu(E_m) = +\infty$, así como también $\nu(\bigcup_{n=1}^{\infty} E_n) = +\infty$. Por esto, $\lim_{n\to\infty} \nu(E_n) = +\infty = \nu(\lim_{n\to\infty} E_n)$.

Si $\nu(E_n) \in \mathbb{R}$ para todo $n \geq 1$, entonces la prueba procede de modo completamente similar a la del Teorema 6.3.45, página 286. La demostración de (b) es también similar a la del Teorema 6.3.47, página 287 y, por lo tanto, se omite. ■

El siguiente criterio, similar a la caracterización dada en el Teorema 11.1.6, página 684, será de mucha utilidad en lo que sigue. Observe la diferencia con el resultado anterior.

Teorema 12.1.5. *Sea $(X, \mathcal{M})$ un espacio medible y $\nu : \mathcal{M} \to \mathbb{R}$ una* ***medida con signo finitamente aditiva****. Las siguientes condiciones son equivalentes:*

(1) ν es ***numerablemente aditiva****.*

(2) Para cualquier ***sucesión creciente*** *$(E_n)_{n=1}^{\infty} \subseteq \mathcal{M}$,*

$$\nu\bigg(\bigcup_{n=1}^{\infty} E_n\bigg) \;=\; \lim_{n\to\infty} \nu(E_n).$$

(3) Para cualquier ***sucesión decreciente*** *$(E_n)_{n=1}^{\infty} \subseteq \mathcal{M}$,*

$$\nu\bigg(\bigcap_{n=1}^{\infty} E_n\bigg) \;=\; \lim_{n\to\infty} \nu(E_n).$$

Prueba. La implicación $(1) \Rightarrow (2)$ es la parte (a) del Teorema 12.1.4 y claramente $(2) \Rightarrow (3)$. Para verificar que $(3) \Rightarrow (1)$ considere una sucesión disjunta $(E_n)_{n=1}^{\infty}$ en $\mathcal{M}$ y sea $E = \bigcup_{n=1}^{\infty} E_n$. Para cada $n \geq 1$, defina

$$F_n \;=\; E \setminus \big(E_1 \cup \cdots \cup E_n\big) \;=\; \bigcup_{k=n+1}^{\infty} E_k.$$

Resulta que $(F_n)_{n=1}^{\infty}$ es una sucesión decreciente de conjuntos medibles con $\bigcap_{n=1}^{\infty} F_n = \varnothing$ y entonces, por hipótesis, $\lim_{n\to\infty} \nu(F_n) = 0$. Ahora bien, como ν toma sólo valores reales y es finitamente aditiva, entonces

$$\nu(E) \;=\; \sum_{k=1}^{n} \nu(E_k) \,+\, \nu(F_n), \qquad n = 1, 2, \ldots$$

y, en consecuencia, tomando límite cuando $n \to \infty$ se obtiene que

$$\nu(E) = \lim_{n\to\infty} \sum_{k=1}^{n} \nu(E_k) + \lim_{n\to\infty} \nu(F_n) = \sum_{k=1}^{\infty} \nu(E_k)$$

y termina la prueba. ■

Definición 12.1.6. *Sea ν una medida con signo definida sobre $(X, \mathcal{M})$. Un* ***conjunto medible*** *$A \subseteq X$ se dice que es* ***ν-positivo*** *si $\nu(E) \geq 0$ para cualquier subconjunto medible $E \subseteq A$.*

Si la medida con signo ν es la única con la que se está trabajando, entonces diremos que A es **positivo** en lugar de ν-positivo. De modo similar, fijada una medida con signo ν, diremos que un conjunto medible $B \subseteq X$ es ν-**negativo**, o simplemente **negativo**, si para cualquier conjunto medible $E \subseteq B$ se tiene que $\nu(E) \leq 0$. Finalmente, un conjunto medible $N \subseteq X$ es ν-**nulo**, o **nulo**, si $\nu(E) = 0$ para cualquier conjunto medible $E \subseteq N$.

Denote por $\mathcal{P}_{\text{os}}(X)$ *la familia de todos los* ***subconjuntos positivos*** *de* X. Observe que:

(i) Cualquier subconjunto medible de un conjunto positivo hereda esa propiedad.

(ii) Si $\{A_1, \ldots, A_n\}$ es una familia finita de de conjuntos positivos, entonces $\bigcup_{i=1}^{n} A_i$ y $\bigcap_{i=1}^{n} A_i$ son ambos conjuntos positivos. Por supuesto, si $(A_n)_{n=1}^{\infty}$ es una sucesión de conjuntos positivos, entonces por la primera parte se tiene que $\bigcap_{n=1}^{\infty} A_n \in \mathcal{P}_{\text{os}}(X)$. El hecho de que $\mathcal{P}_{\text{os}}(X)$ también es estable bajo uniones de colecciones numerables sigue del siguiente resultado.

Lema 12.1.7. *Sea ν una* ***medida con signo*** *sobre un espacio medible $(X, \mathcal{M})$.*

(a) Si $(A_n)_{n=1}^{\infty}$ es una sucesión de ***conjuntos positivos****, entonces $\bigcup_{n=1}^{\infty} A_n$ es* ***positivo****.*

(b) Si $\varsigma = \sup\{\nu(A) : A \in \mathcal{P}_{\text{os}}(X)\}$, entonces dicho supremo se alcanza; es decir, existe $A_0 \in \mathcal{P}_{\text{os}}(X)$ tal que $\nu(A_0) = \varsigma$.

Prueba. (a) Sea E un subconjunto medible de $\bigcup_{n=1}^{\infty} A_n$. Defina $E_1 = E \cap A_1$ y para cada $n \geq 2$ hagamos

$$E_n = E \cap \Big(A_n \setminus \bigcup_{i=1}^{n-1} A_i\Big).$$

Resulta que cada E_n es un conjunto medible incluido en A_n y, por lo tanto, $\nu(E_n) \geq 0$. Puesto que la colección $(E_n)_{n=1}^{\infty}$ es una sucesión disjunta con $E = \bigcup_{n=1}^{\infty} E_n$ y como ν es numerablemente aditiva, se tiene que

$$\nu(E) = \sum_{n=1}^{\infty} \nu(E_n) \geq 0.$$

Esto prueba que $\bigcup_{n=1}^{\infty} A_n$ es positivo.

(b) La definición de ς permite obtener una sucesión $(A_n)_{n=1}^{\infty}$ de conjuntos positivos tal que

$$\varsigma = \lim_{n\to\infty} \nu(A_n).$$

Por supuesto, podemos elegir creciente a la sucesión $(A_n)_{n=1}^{\infty}$ pues, en caso contrario, la sucesión $(B_n)_{n=1}^{\infty}$, donde $B_n = A_1 \cup \ldots \cup A_n$ para cada $n \geq 1$ es creciente, cada B_n es un conjunto positivo y se cumple que $\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$. Teniendo en cuenta esto, se sigue de la parte (a) que $A_0 = \bigcup_{n=1}^{\infty} A_n \in \mathcal{P}_{\rm os}(X)$ y entonces, la continuidad de ν, Teorema 12.1.4, nos revela que

$$\nu(A_0) \;=\; \lim_{n\to\infty} \nu(A_n) \;=\; \varsigma.$$

La prueba es completa. ■

Observe que el conjunto positivo A_0 obtenido en la parte (b) del Lema 12.1.7 es, por definición, **maximal** (en el orden "$\subseteq$") en la clase $\mathcal{P}_{\rm os}(X)$.

El siguiente resultado, combinado con el lema anterior, es fundamental en la demostración del Teorema de Descomposición de Hahn.

Lema 12.1.8. *Sea ν una **medida con signo** definida sobre un conjunto medible $(X, \mathcal{M})$. Si $E \in \mathcal{M}$ es tal que $0 < \nu(E) < +\infty$, entonces existe un **conjunto positivo** $A \subseteq E$ con $\nu(A) > 0$.*

Prueba. Si E es positivo no hay nada que demostrar. Suponga entonces que E no es positivo. Esto, por supuesto, significa que E contiene subconjuntos medibles de medida no-positiva, es decir, existe al menos un conjunto medible $E_1 \subseteq E$ tal que $\nu(E_1) < 0$. El Principio de Arquímedes nos garantiza la existencia de un $n_0 \in \mathbb{N}$ tal que $\nu(E_1) < -1/n_0$ y, entonces, el Principio del Buen Orden en $\mathbb{N}$ permite elegir

$$n_1 \;=\; \min\big\{n \in \mathbb{N} \,:\, E_1 \subseteq E,\ \nu(E_1) < -1/n\big\}.$$

De aquí se sigue que

$$-\frac{1}{n_1 - 1} \;<\; \nu(E_1) \;\leq\; -\frac{1}{n_1}.$$

Observe que, gracias a la aditividad de ν,

$$\nu(E \setminus E_1) \;=\; \nu(E) \;-\; \nu(E_1) \;>\; \nu(E) \;>\; 0.$$

Si $E \setminus E_1$ es un conjunto positivo, nos detenemos. En caso contrario, $E \setminus E_1$ contiene un conjunto medible E_2 con $\nu(E_2) < 0$. Como antes, sea

$$n_2 \;=\; \min\big\{n \in \mathbb{N} \,:\, E_2 \subseteq E \setminus E_1,\ \nu(E_1) < -1/n\big\}.$$

Entonces

$$-\frac{1}{n_2 - 1} \;<\; \nu(E_2) \;\leq\; -\frac{1}{n_2}.$$

De nuevo, la aditividad de ν nos conduce a

$$\nu(E \setminus (E_1 \cup E_2)) \;=\; \nu(E) - \nu(E_1) - \nu(E_2) \;>\; 0.$$

Si $E \setminus (E_1 \cup E_2)$ es un conjunto positivo, paramos; sino, continuamos hasta obtener una sucesión $(E_k)_{k=1}^{\infty}$ en $\mathcal{M}$ tal que para cada $k \in \mathbb{N}$,

$$E_k \subseteq E \setminus \bigcup_{n=1}^{k} E_n \qquad \text{y} \qquad -\frac{1}{n_k - 1} \;<\; \nu(E_k) \;<\; -\frac{1}{n_k}.$$

Defina

$$A_0 \;=\; E \setminus \bigcup_{n=1}^{\infty} E_n.$$

Afirmamos que A_0 es un conjunto positivo. Para verificar esto último, observe en primer lugar que como $E = A_0 \cup (\bigcup_{n=1}^{\infty} E_n)$ y dicha unión es disjunta, entonces la aditividad de ν nos asegura que

$$\nu(E) \;=\; \nu(A_0) + \nu\bigg(\bigcup_{n=1}^{\infty} E_n\bigg).$$

Por otro lado, puesto que $\bigcup_{n=1}^{\infty} E_n \subseteq E$, entonces

$$\sum_{n=1}^{\infty} |\nu(E_n)| \;=\; \bigg|\nu\bigg(\bigcup_{n=1}^{\infty} E_n\bigg)\bigg| \;\leq\; |\nu(E)| \;<\; +\infty,$$

lo cual nos indica que la serie $\sum_{n=1}^{\infty} \nu(E_n)$ converge absolutamente y, en consecuencia,

$$\nu(E) \;=\; \nu(A_0) + \sum_{n=1}^{\infty} \nu(E_n).$$

Puesto que $|\nu(E_k)| = -\nu(E_k) \geq 1/n_k$ y $\sum_{k=1}^{\infty} |\nu(E_k)|$ converge, se tiene que $\sum_{k=1}^{\infty} 1/n_k$ converge. De aquí se sigue que $\frac{1}{n_k} \to 0$ cuando $k \to \infty$, es decir,

$$\lim_{k\to\infty} n_k \;=\; +\infty. \tag{3}$$

Finalmente, suponga, para generar una contradicción, que A_0 contiene un conjunto medible B con $\nu(B) < 0$. Usemos (3) para elegir un entero $k \geq 1$ de modo que $\nu(B) < -(n_k - 1)^{-1}$. Entonces

$$B \subseteq A_0 \subseteq E \setminus \bigcup_{n=1}^{k-1} E_n \qquad \text{y} \qquad \nu(B) \;<\; -\frac{1}{n_k - 1}.$$

Sin embargo, por construcción, n_k es el entero positivo más pequeño para el cual existe un conjunto medible $E_k \subseteq E \setminus \bigcup_{n=1}^{k-1} E_n$ tal que $-1/(n_k - 1) < \nu(E_k) < -1/n_k$. Esta contradicción establece que nuestro conjunto A_0 no puede contener conjuntos de medida negativa por lo que él es un conjunto positivo. ■

Ya tenemos en nuestras manos todos los argumentos necesarios para demostrar el Teorema de Descomposición de Hahn, el cual establece una partición del espacio medible en dos conjuntos medible: uno positivo y el otro negativo.

Teorema 12.1.9 (Descomposición de Hahn). *Sea ν una **medida con signo** sobre $(X, \mathcal{M})$. Entonces existe un **conjunto positivo** A y un **conjunto negativo** B tales que*

$$X \;=\; A \cup B \qquad \text{y} \qquad A \cap B \;=\; \varnothing.$$

Prueba. Reemplazando ν por $-\nu$ si fuese necesario, podemos asumir que ν evita el valor $+\infty$, es decir, supondremos que el rango de ν es $[-\infty, +\infty)$. Sea

$$\varsigma \;=\; \sup\big\{\nu(P) \,:\, P \in \mathcal{P}_{\text{os}}(X)\big\},$$

donde $\mathcal{P}_{\text{os}}(X)$ es la familia de todos los conjuntos positivos de X. Por el Lema 12.1.7 existe un conjunto positivo A tal que $\nu(A) = \varsigma$. Puesto que estamos asumiendo que $-\infty \leq \nu(E) < +\infty$ para todo $E \in \mathcal{M}$, entonces $\varsigma < +\infty$. Defina $B = X \setminus A$. Veamos que B es negativo. Suponga, para producir una contradicción, que B no es negativo. Entonces existe un conjunto medible $E \subsetneq B$ tal que $\nu(E) > 0$. Puesto que $\nu(E) < +\infty$, el Lema 12.1.8 nos garantiza la existencia de un conjunto positivo $P \subseteq E$ con $\nu(P) > 0$. Esto, sin embargo, contradice el siguiente el hecho: como P y A son disjuntos y $A \cup P$ es positivo, entonces

$$\varsigma \;\geq\; \nu(A \cup P) \;=\; \nu(A) + \nu(P) \;=\; \varsigma + \nu(P) \;>\; \varsigma.$$

Por esto, B es un conjunto negativo y termina la prueba. ■

Al par de conjuntos (A, B) obtenidos en el resultado anterior se le llama una ν-**descomposición de Hahn para** X. Es importante destacar que tal descomposición no es única. Por ejemplo, si N es cualquier conjunto ν-nulo, entonces $(A \cup N, B \setminus N)$ es otra ν-descomposición de Hahn para X. Sin embargo, si (A, B) y (A', B') son dos ν-descomposiciones de Hahn para X, entonces

$$\nu(E \cap A) \;=\; \nu(E \cap A') \qquad \text{y} \qquad \nu(E \cap B) \;=\; \nu(E \cap B') \tag{4}$$

para todo $E \in \mathcal{M}$. Para ver esto, fijemos $E \in \mathcal{M}$. Puesto que $A \setminus A'$ es un subconjunto medible tanto de A así como de B', se tiene que $E \cap (A \setminus A')$ también lo es y, en consecuencia, $\nu(E \cap (A \setminus A')) = 0$. Por simetría, $\nu(E \cap (A' \setminus A)) = 0$. De esto se sigue que

$$\begin{aligned}
\nu(E \cap (A \cup A')) \;&=\; \nu(E \cap A \cap (A \cup (A' \setminus A))) \;=\; \nu((E \cap A) \cup (E \cap (A' \setminus A))) \\
&=\; \nu(E \cap A) + \nu(E \cap (A' \setminus A)) \\
&=\; \nu(E \cap A).
\end{aligned}$$

Similarmente, $\nu(E \cap (A \cup A')) \;=\; \nu(E \cap A')$. Argumentando de modo similar se prueba que $\nu(E \cap B) \;=\; \nu(E \cap B')$. Las igualdades dadas en (4) permiten justificar la siguiente definición.

Definición 12.1.10. *Sea ν una **medida con signo** sobre $(X, \mathcal{M})$ y sea (A, B) una ν-descomposición de Hahn sobre X. Defina ν^+, ν^- y $|\nu|$ sobre $\mathcal{M}$ por*

$$\nu^+(E) \;=\; \nu(E \cap A), \qquad \nu^-(E) \;=\; -\nu(E \cap B) \qquad y \qquad |\nu|(E) \;=\; \nu^+(E) + \nu^-(E)$$

*para todo $E \in \mathcal{M}$. Las funciones ν^+, ν^- y $|\nu|$ son llamadas, respectivamente, la **variación positiva**, la **variación negativa** y la **variación total** de ν.*

Es un ejercicio sencillo verificar que ν^+, ν^- y $|\nu|$ son **medidas positivas** definidas sobre $\mathcal{M}$. Además,

$$\nu(E) \;=\; \nu^+(E) - \nu^-(E) \quad \text{para todo } E \in \mathcal{M}.$$

La expresión $\nu = \nu^+ - \nu^-$ se conoce como la **descomposición de Jordan de** ν. Observe que si (A, B) es una ν-descomposición de Hahn para X, entonces

$$\nu^+(B) \;=\; \nu^-(A) \;=\; 0.$$

Definición 12.1.11. *Sean λ y ν dos medidas con signo definidas sobre $(X, \mathcal{M})$. Diremos que λ y ν son **mutuamente singulares**, lo que escribiremos como $\lambda \perp \nu$, si existen conjuntos medibles y disjuntos P y Q tales que*

$$X \;=\; P \cup Q \qquad y \qquad \lambda(P) \;=\; \nu(Q) \;=\; 0.$$

Lo que resulta interesante es el siguiente resultado importante.

Teorema 12.1.12 (Descomposición de Jordan). *Si ν es una **medida con signo** sobre $(X, \mathcal{M})$, entonces existen **medidas positivas únicas** ν^+ y ν^- tales que*

$$\nu = \nu^+ - \nu^- \qquad y \qquad \nu^+ \perp \nu^-.$$

Prueba. Sea (A, B) una ν-descomposición de Hahn para X. Definiendo

$$\nu^+(E) = \nu(E \cap A), \qquad \text{y} \qquad \nu^-(E) = -\nu(E \cap B),$$

tenemos que $\nu = \nu^+ - \nu^-$ y $\nu^+ \perp \nu^-$.

Suponga ahora que $\nu = \lambda_1 - \lambda_2$ es otra descomposición de ν con $\lambda_1 \perp \lambda_2$. Sean P, Q conjuntos medibles y disjuntos tales que $X = P \cup Q$ y $\lambda_1(P) = \lambda_2(Q) = 0$. Afirmamos que (A', B') es otra ν-descomposición de Hahn para X, donde $A' = Q$ y $B' = P$. En efecto, si $E \subseteq A'$ es medible, entonces $\lambda_2(E) \leq \lambda_2(A') = 0$ y, por lo tanto,

$$\nu(E) = \lambda_1(E) - \lambda_2(E) = \lambda_1(E) \geq 0.$$

Esto nos dice que A' es un conjunto ν-positivo. Similarmente, B' es ν-negativo. Puesto que $\lambda_1(P) = 0$ y $\lambda_2(Q) = 0$, se sigue de las desigualdades dadas en (4) que si $E \in \mathcal{M}$, entonces

$$\begin{aligned} \nu^+(E) &= \nu(E \cap A) = \nu(E \cap A') \\ &= \lambda_1(E \cap A') - \lambda_2(E \cap A') \\ &= \lambda_1(E \cap A') = \lambda_1(E). \end{aligned}$$

De modo similar, $\nu^- = \lambda_2$ y termina la prueba. ■

Ejemplo 12.1.1. *Sea $f \in L_1(\mu)$. Si se define $\nu_f : \mathfrak{M}_\mu(\mathbb{R}) \to \mathbb{R}$ por*

$$\nu_f(E) = \int_E f\, d\mu, \quad \textit{para todo } E \in \mathfrak{M}_\mu(\mathbb{R}),$$

entonces

$$\nu_f^+(E) = \int_E f^+\, d\mu, \qquad \nu_f^-(E) = \int_E f^-\, d\mu \qquad y \qquad |\nu_f|(E) = \int_E |f|\, d\mu$$

para todo $E \in \mathfrak{M}_\mu(\mathbb{R})$.

Es importante tener en cuenta que muchas de la propiedades que posee una medida con signo ν las hereda $|\nu|$ y viceversa. Por ejemplo:

(I) *Sea $N \in \mathcal{M}$. Entonces, $\nu(N) = 0$ si, y sólo si, $|\nu|(N) = 0$.* En efecto, sea (A, B) una ν-descomposición de Hahn para X. Puesto que

$$|\nu|(N) = \nu^+(N) + \nu^-(N) = \nu(N \cap A) - \nu(N \cap B)$$

y $N \cap A$, así como $N \cap B$, son subconjuntos medibles de N, se tiene que si N es ν-nulo, entonces $\nu(N \cap A) = 0 = \nu(N \cap B)$, por lo que $|\nu|(N) = 0$. Recíprocamente, si $|\nu|(N) = 0$, las desigualdades $\nu^+(N) \leq |\nu|(N)$ y $\nu^-(N) \leq |\nu|(N)$ implican que $\nu(N) = \nu^+(N) - \nu^-(N) = 0$. ■

(II) *Si* ν *es una medida con signo sobre* $(X, \mathcal{M})$ *y* (A, B) *es una* ν*-descomposición de Hahn para* X*, entonces se cumple que, para todo* $E \in \mathcal{M}$*:*

$$\nu(E) = \int_E (\chi_A - \chi_B)\, d|\nu|.$$

Prueba. Sea (A, B) es una ν-descomposición de Hahn para X. Entonces, para cualquier $E \in \mathcal{M}$,

$$\begin{aligned}
\int_E (\chi_A - \chi_B)\, d|\nu| &= \int_X \chi_{E\cap A}\, d|\nu| - \int_X \chi_{E\cap B}\, d|\nu| \\
&= |\nu|(E\cap A) - |\nu|(E\cap A) \\
&= \big(\nu^+(E\cap A) + \nu^-(E\cap A)\big) - \big(\nu^+(E\cap B) + \nu^-(E\cap B)\big) \\
&= \nu(E\cap A) - \nu(E\cap A\cap B) - \nu(E\cap B\cap A) + \nu(E\cap B) \\
&= \nu(E\cap A) + \nu(E\cap B) = \nu(E).
\end{aligned}$$

La siguiente es una útil y práctica caracterización de la variación total $|\nu|$, la cual también puede ser tomada como su definición. Cuando ν es una medida positiva, entonces $|\nu|$ representa la *medida positiva más pequeña* que domina a ν.

Teorema 12.1.13. *Sea* ν *una* ***medida con signo finitamente aditiva*** *sobre* $(X, \mathcal{M})$*. Entonces, para todo* $E \in \mathcal{M}$*, se cumple que*

$$|\nu|(E) = \sup\left\{\sum_{k=1}^{n} |\nu(E_k)| \,:\, \{E_1, \ldots, E_n\} \text{ es una partición medible de } E\right\}. \tag{5}$$

Prueba. Fijemos un conjunto $E \in \mathcal{M}$ y sea

$$\beta = \sup\left\{\sum_{k=1}^{n} |\nu(E_k)| \,:\, \{E_1, \ldots, E_n\} \text{ es una partición medible de } E\right\}.$$

Para cada partición medible $\{E_1, \ldots, E_n\}$ de E se tiene que

$$\begin{aligned}
\sum_{k=1}^{n} |\nu(E_k)| &= \sum_{k=1}^{n} |\nu^+(E_k) - \nu^-(E_k)| \\
&\leq \sum_{k=1}^{n} \big(\nu^+(E_k) + \nu^-(E_k)\big) \\
&= \sum_{k=1}^{n} |\nu|(E_k) = |\nu|(E)
\end{aligned}$$

y, por lo tanto, $\beta \leq |\nu|(E)$. Para demostrar la desigualdad opuesta, sea (A, B) una ν-descomposición de Hahn para X. Entonces, $\{E\cap A, E\cap B\}$ es una partición medible de E y se sigue de la definición de β que

$$\beta \geq |\nu(E\cap A)| + |\nu(E\cap B)| = \nu^+(E) + \nu^-(E) = |\nu|(E).$$

Esto último finaliza la prueba. ■

En ocasiones, puede ser beneficioso observar que $|\nu|(E)$ también se puede expresar en la forma:

$$|\nu|(E) \;=\; \sup\left\{\sum_{k=1}^{\infty}|\nu(E_k)| \;:\; \{E_1,E_2,\ldots\} \text{ es una partición medible infinita de } E\right\}. \tag{5a}$$

En efecto, para demostrar $(5a)$, pongamos

$$\beta^* \;=\; \sup\left\{\sum_{k=1}^{\infty}|\nu(E_k)| \;:\; \{E_1,E_2,\ldots\} \text{ es una partición medible infinita de } E\right\}.$$

Puesto que toda partición medible finita de E, digamos $\{E_1,\ldots,E_n\}$, es un caso particular de una partición medible infinita de E (basta considerar la colección infinita $\{E_1,\ldots,E_n,\varnothing,\varnothing,\ldots\}$), resulta que $|\nu|(E) \leq \beta^*$. Para demostrar la desigualdad $\beta^* \leq |\nu|(E)$ vamos a considerar los siguientes dos casos:

(1) Existe una partición medible infinita $\{E_1,E_2,\ldots\}$ de E tal que $\sum_{k=1}^{\infty}|\nu(E_k)| = +\infty$. En este caso $\beta^* = +\infty$. Seleccione un número arbitrario $M > 0$. Puesto que $\sum_{k=1}^{\infty}|\nu(E_k)| = +\infty$, existe un $N \in \mathbb{N}$ tal que $\sum_{k=1}^{N}|\nu(E_k)| \geq M$. Sea $E_0 = \bigcup_{n=N+1}^{\infty}E_n$. Resulta que $\{E_1,\ldots,E_N,E_0\}$ es una partición medible finita de E la cual satisface

$$\sum_{k=0}^{N}|\nu(E_k)| \;=\; |\nu(E_0)| + \sum_{k=1}^{N}|\nu(E_k)| \;\geq\; M. \tag{a_1}$$

Esto último, por supuesto, implica que $|\nu|(E) \geq M$. Finalmente, como (a_1) es válida para cualquier elección arbitraria de $M > 0$, se tiene que $|\nu|(E) = +\infty = \beta^*$.

(2) Suponga ahora que para cualquier partición medible infinita $\{E_1,E_2,\ldots\}$ de E se cumple que $\sum_{k=1}^{\infty}|\nu(E_k)| < +\infty$. Fijemos un $\varepsilon > 0$ elegido arbitrariamente y usemos el hecho de que $\sum_{k=1}^{\infty}|\nu(E_k)| < +\infty$ para hallar un $N \in \mathbb{N}$ de modo que

$$\sum_{k=1}^{N}|\nu(E_k)| \;>\; \sum_{k=1}^{\infty}|\nu(E_k)| - \varepsilon.$$

Hagamos $E_0 = \bigcup_{n=N+1}^{\infty}E_n$. Entonces $\{E_1,\ldots,E_N,E_0\}$ es una partición medible finita de E satisfaciendo

$$\sum_{k=0}^{N}|\nu(E_k)| \;=\; |\nu(E_0)| + \sum_{k=1}^{N}|\nu(E_k)| \;>\; \sum_{k=1}^{\infty}|\nu(E_k)| - \varepsilon.$$

La conclusión es clara: hemos demostrado que para cualquier partición medible infinita de E, digamos $\{E_1,E_2,\ldots\}$, para la cual $\sum_{k=1}^{\infty}|\nu(E_k)| < +\infty$ y cualquier $\varepsilon > 0$, existe una partición medible finita $\{E_1,E_2,\ldots,E_N,E_0\}$ de E tal que

$$\sum_{k=1}^{\infty}|\nu(E_k)| - \varepsilon \;<\; \sum_{k=0}^{N}|\nu(E_k)| \;\leq\; |\nu|(E).$$

En otras palabras,

$$\sum_{k=1}^{\infty} |\nu(E_k)| \; < \; |\nu|(E) \, + \, \varepsilon.$$

De aquí se sigue, $\beta^* \leq |\nu|(E) + \varepsilon$ y como $\varepsilon > 0$ fue elegido de modo arbitrario, concluimos que $\beta^* \leq |\nu|(E)$ y termina la prueba. ■

Puesto que $\{E, \varnothing\}$ es una partición medible de E y $|\nu|$ es una medida positiva, entonces la caracterización de $|\nu|$ dada por el Teorema 12.1.13 implica que

$$|\nu(E)| \; \leq \; |\nu|(E) \; \leq \; |\nu|(X), \quad \text{para todo } E \in \mathcal{M}. \tag{6}$$

Teorema 12.1.14. *Sea ν una **medida con signo** sobre $(X, \mathcal{M})$. Las siguientes condiciones son equivalentes:*

(1) $\nu(E) \in \mathbb{R}$ para todo $E \in \mathcal{M}$.

*(2) $|\nu|$ es una **medida finita**.*

Prueba. $(1) \Rightarrow (2)$. Suponga que $\nu(E) \in \mathbb{R}$ para todo $E \in \mathcal{M}$ y sea (A, B) una ν-descomposición de Hahn para X. Puesto que ν omite el valor $+\infty$, resulta que $\nu^+(X) = \nu(A) < +\infty$ de modo que ν^+ es una medida finita y ν está acotada superiormente por $\nu^+(X)$. Similarmente, como ν no alcanza el valor $-\infty$, entonces ν^- es una medida finita y ν está acotada inferiormente por $-\nu^-(X)$. Por esto, $|\nu|(X) = \nu^+(X) + \nu^-(X) < +\infty$.

$(2) \Rightarrow (1)$ Esto es consecuencia inmediata de (6). ■

Del resultado anterior se deduce que si ν es una medida con signo tal que $\nu(E) \in \mathbb{R}$ para todo $E \in \mathcal{M}$, entonces

$$-|\nu|(X) \; \leq \; \nu(E) \; \leq \; |\nu|(X) \quad \text{para todo } E \in \mathcal{M}.$$

Definición 12.1.15. *Un espacio de medida con signo $(X, \mathcal{M}, \nu)$ se dirá que es un **espacio de medida real** si $\nu(E) \in \mathbb{R}$ para todo $E \in \mathcal{M}$ o, equivalentemente, si $|\nu|(X) < +\infty$. En este caso también se dice que ν es una **medida real** o **finita**.*

Teorema 12.1.16. *Si $(X, \mathcal{M}, \nu)$ es un **espacio de medida real**, entonces*

$$\sup\{|\nu(A)| \,:\, A \in \mathcal{M},\, A \subseteq E\} \;\leq\; |\nu|(E) \;\leq\; 2 \cdot \sup\{|\nu(A)| \,:\, A \in \mathcal{M},\, A \subseteq E\}, \tag{7}$$

para todo $E \in \mathcal{M}$.

Prueba. Fijemos $E \in \mathcal{M}$ y sea $\alpha = \sup\{|\nu(A)| \,:\, A \in \mathcal{M},\, A \subseteq E\}$. Observe que $\alpha \in \mathbb{R}$ ya que ν es finita. Tome cualquier $A \in \mathcal{M}$ con $A \subseteq E$. Entonces $\{A, E \setminus A\}$ es una partición medible de E y, en consecuencia,

$$|\nu(A)| \;\leq\; |\nu(A)| + |\nu(E \setminus A)| \;\leq\; |\nu|(E).$$

Tomando el supremo sobre tales A se obtiene la primera desigualdad, esto es, $\alpha \leq |\nu|(E)$. Para comprobar la otra desigualdad en (7), sea (A, B) una ν-descomposición de Hahn para X. Entonces

$$|\nu|(E) \;=\; \nu(E \cap A) - \nu(E \cap B) \;\leq\; |\nu(E \cap A)| + |\nu(E \cap B)| \;\leq\; 2\alpha,$$

que es lo que queríamos establecer. ■

Definición 12.1.17. *Diremos que una medida con signo ν sobre $(X, \mathcal{M})$ es **acotada** si*

$$\sup\{|\nu(A)| \,:\, A \in \mathcal{M},\ A \subseteq E\} \;<\; +\infty.$$

En vista del resultado anterior tenemos que:

Corolario 12.1.18. *Si ν es una **medida real**, entonces ella es **acotada**.*

Denote por $\mathrm{ca}(\mathcal{M}, \mathbb{R})$ *el espacio vectorial* $($*sobre* $\mathbb{R})$ *de todas las **medidas reales** sobre* $(X, \mathcal{M})$. El Teorema 12.1.14 permite dotar al espacio $\mathrm{ca}(\mathcal{M}, \mathbb{R})$ con la **norma de la variación**

$$\|\nu\| \;=\; |\nu|(X) \tag{8}$$

para cada $\nu \in \mathrm{ca}(\mathcal{M}, \mathbb{R})$. Nótese que $0 \leq \|\nu\| < +\infty$ y es rutinario verificar que (8) es, en realidad, una norma sobre $\mathrm{ca}(\mathcal{M}, \mathbb{R})$. Observe que $\mathrm{ca}(\mathcal{M}, \mathbb{R})$ puede también ser equipado con esta otra norma

$$\|\nu\|_\infty \;=\; \sup\big\{|\nu(E)| \,:\, E \in \mathcal{M}\big\},$$

la cual es, por las desigualdades dadas en (7), equivalente a la norma de la variación.

Teorema 12.1.19. $(\mathrm{ca}(\mathcal{M}, \mathbb{R}), \|\cdot\|)$ *es un **espacio de Banach**.*

Prueba. Sea $(\nu_n)_{n=1}^{\infty}$ una sucesión de Cauchy en $\mathrm{ca}(\mathcal{M}, \mathbb{R})$. Entonces, dado $\varepsilon > 0$, existe un $N_\varepsilon \in \mathbb{N}$ tal que $\|\nu_m - \nu_n\| < \varepsilon$ para todo $m, n \geq N_\varepsilon$. Se sigue de la desigualdad (6) que, para cada $E \in \mathcal{M}$,

$$\big|\nu_m(E) - \nu_n(E)\big| \;=\; \big|(\nu_m - \nu_n)(E)\big| \;\leq\; |\nu_m - \nu_n|(E) \;\leq\; \|\nu_m - \nu_n\| \;<\; \varepsilon$$

para todo $m, n \geq N_\varepsilon$, lo cual nos indica que $(\nu_n(E))_{n=1}^{\infty}$ es de Cauchy en $\mathbb{R}$ y, en consecuencia, $\lim_{n\to\infty} \nu_n(E)$ existe para cada $E \in \mathcal{M}$. Defina $\nu : \mathcal{M} \to \mathbb{R}$ por

$$\nu(E) \;=\; \lim_{n\to\infty} \nu_n(E), \quad E \in \mathcal{M}.$$

Afirmamos que $\nu \in \mathrm{ca}(\mathcal{M}, \mathbb{R})$ y que $\|\nu_m - \nu\| \to 0$ cuando $n \to \infty$. Para verificar la primera afirmación, observe, en primer lugar, que ν es finitamente aditiva. En efecto, sean $E_1, E_2 \in \mathcal{M}$ con $E_1 \cap E_2 = \varnothing$. Entonces

$$\begin{aligned}
\nu(E_1 \cup E_2) &= \lim_{n\to\infty} \nu_n(E_1 \cup E_2) \\
&= \lim_{n\to\infty} \big(\nu_n(E_1) + \nu_n(E_2)\big) \\
&= \lim_{n\to\infty} \nu_n(E_1) + \lim_{n\to\infty} \nu_n(E_2) \\
&= \nu(E_1) + \nu(E_2).
\end{aligned}$$

Demostrar que ν es numerablemente aditiva requiere un paso más: a saber, que ν también es continua. Para probar esto, fijemos una sucesión $(F_k)_{k=1}^{\infty}$ de conjuntos medibles tal que

$$F_1 \supseteq F_2 \supseteq \cdots \qquad \text{y} \qquad \bigcap_{k=1}^{\infty} F_k = \varnothing.$$

La continuidad de ν_{N_ε}, Teorema 12.1.4 (c), nos muestra que

$$\lim_{k\to\infty} \nu_{N_\varepsilon}(F_k) \;=\; 0,$$

de donde se sigue la existencia de un entero $J_\varepsilon \geq 1$ tal que

$$\nu_{N_\varepsilon}(F_k) \;<\; \varepsilon \quad \text{para todo} \quad k \geq J_\varepsilon.$$

Ahora, para cualquier $k \geq J_\varepsilon$ se obtiene que

$$\begin{aligned}
|\nu(F_k)| \;&=\; \lim_{n\to\infty} |\nu_n(F_k)| \\
&=\; \lim_{n\to\infty} |\nu_n(F_k) \;-\; \nu_{N_\varepsilon}(F_k) \;+\; \nu_{N_\varepsilon}(F_k)| \\
&\leq\; |\nu_{N_\varepsilon}(F_k)| \;+\; \lim_{n\to\infty} |(\nu_n \;-\; \nu_{N_\varepsilon})(F_k)| \\
&\leq\; |\nu_{N_\varepsilon}(F_k)| \;+\; \limsup_{n\to\infty} \|\, \nu_n \;-\; \nu_{N_\varepsilon} \,\| \\
&<\; 2\varepsilon,
\end{aligned}$$

es decir, $\lim_{k\to\infty} \nu(F_k) = \nu(\bigcap_{k=1}^\infty F_k) = 0$. Estamos listo para demostrar la σ-aditividad de ν. En efecto, sea $\big(E_n\big)_{n=1}^\infty$ una sucesión disjunta de conjuntos medibles. Pongamos, para cada $n \in \mathbb{N}$, $F_n = \bigcup_{k=n+1}^\infty E_k$. Por la aditividad de ν

$$\begin{aligned}
\nu\Big(\bigcup_{n=1}^\infty E_n\Big) \;&=\; \sum_{n=1}^k \nu(E_n) \;+\; \nu\Big(\bigcup_{n=k+1}^\infty E_n\Big) \\
&=\; \sum_{n=1}^k \nu(E_n) \;+\; \nu(F_k),
\end{aligned}$$

para todo $k \in \mathbb{N}$. Puesto que la sucesión $\big(F_n\big)_{n=1}^\infty$ es decreciente y satisface $\bigcap_{n=1}^\infty F_n = \varnothing$, resulta de lo anterior, tomando límite cuando $k \to \infty$, que

$$\nu\Big(\bigcup_{n=1}^\infty E_n\Big) \;=\; \sum_{n=1}^\infty \nu(E_n).$$

Falta comprobar que $\|\, \nu_m - \nu \,\| \;\to\; 0$ cuando $n \;\to\; \infty$. Para ver esto, sea $\{E_1, \ldots, E_k\}$ una partición medible de X. Por el Teorema 12.1.13 se tiene que, para todo $m, n \geq N_\varepsilon$,

$$\sum_{j=1}^k \big|\nu_m(E_j) - \nu_n(E_j)\big| \;\leq\; \|\, \nu_m - \nu_n \,\| \;<\; \varepsilon$$

y, así, usando la continuidad del valor absoluto, resulta que

$$\sum_{j=1}^k \big|\nu(E_j) - \nu_n(E_j)\big| \;=\; \lim_{m\to\infty} \sum_{j=1}^k \big|\nu_m(E_j) - \nu_n(E_j)\big| \;\leq\; \varepsilon.$$

para todo $n \geq N_\varepsilon$. Finalmente, tomando el supremo sobre todas las particiones finitas y medibles de X e invocando de nuevo el Teorema 12.1.13 se descubre que $\| \nu_m - \nu \| < \varepsilon$ para todo $n \geq N_\varepsilon$. Esto termina la prueba. ■

Es importante destacar que cada función $f \in L_1(\mu)$ se identifica con una única medida con signo $\nu_f \in \mathrm{ca}(\mathcal{M}, \mathbb{R})$ tal que $\| \nu_f \| = \| f \|_1$. En efecto, considere la aplicación lineal $T : L_1(\mu) \to \mathrm{ca}(\mathcal{M}, \mathbb{R})$ dada por $T(f) = \nu_f$, donde

$$\nu_f(E) \;=\; \int_E f\, d\mu \quad \text{para todo} \quad E \in \mathfrak{M}_\mu(\mathbb{R}).$$

Del Ejemplo 12.1.1 sabemos que

$$|\nu_f|(E) \;=\; \int_E |f|\, d\mu \quad \text{para todo} \quad E \in \mathfrak{M}_\mu(\mathbb{R}).$$

En particular, $\| \nu_f \| = \| f \|_1$ lo cual significa que T es una isometría de $L_1(\mu)$ en $\mathrm{ca}(\mathcal{M}, \mathbb{R})$. Esto hecho nos revela una conexión interesante: el espacio $(L_1(\mu), \| \cdot \|_1)$ de las funciones integrables según Lebesgue se puede identificar con un **subespacio cerrado** del espacio de Banach $(\mathrm{ca}(\mathcal{M}, \mathbb{R}), \| \cdot \|)$. De hecho, si definimos

$$\mathcal{M}_{\mathrm{ac}}(\mu) \;=\; \{\nu \in \mathcal{M}_{\mathbb{R}}(\mathbb{R}) \,:\, \nu \ll \mu\}$$

entonces, gracias al Teorema de Radon-Nikodým, véase el Teorema 12.1.28, página 804, se tiene que $(L_1(\mu), \| \cdot \|_1)$ y $(\mathcal{M}_{\mathrm{ac}}(\mu), \| \cdot \|)$ son idénticos en el sentido de que la aplicación T es una isometría de $L_1(\mu)$ sobre $\mathcal{M}_{\mathrm{ac}}(\mu)$.

12.1.1. El Teorema de Drewnowski

Esta corta sección está dedicada a establecer un resultado estupendo de Drewnowski el cual afirma que toda medida con signo finitamente aditiva y acotada definida sobre una σ-álgebra es *muy parecida* a una medida con signo numerablemente aditiva.

Sea $\mathcal{M}$ una σ-álgebra de subconjuntos de X y sea $\nu : \mathcal{M} \to \mathbb{R}$ una medida finitamente aditiva y acotada. Si $(E_n)_{n=1}^\infty$ es una sucesión disjunta en $\mathcal{M}$, entonces la serie $\sum_{n=1}^\infty \nu(E_n)$ converge absolutamente. En efecto, como ν es acotada se sigue del Teorema 12.1.16 que para todo $n \geq 1$,

$$\sum_{k=1}^{n} |\nu(E_k)| \;<\; +\infty$$

y así, la serie $\sum_{n=1}^\infty \nu(E_n)$ converge absolutamente. Por esto, si ν no es numerablemente aditiva, entonces debe existir alguna sucesión disjunta $(E_n)_{n=1}^\infty$ en $\mathcal{M}$ tal que la serie $\sum_{n=1}^\infty \nu(E_n)$ no converge al "valor esperado" $\nu(\bigcup_{n=1}^\infty E_n)$.

Teorema 12.1.20 (Drewnowski). *Sea $(X, \mathcal{M})$ un espacio medible y sea $\nu : \mathcal{M} \to \mathbb{R}$ una* ***medida finitamente aditiva*** *y* ***acotada****. Si $(E_n)_{n=1}^\infty$ es una sucesión disjunta en $\mathcal{M}$, entonces existe una subsucesión $(E_{n_j})_{j=1}^\infty$ de $(E_n)_{n=1}^\infty$ tal que ν es* ***numerablemente aditiva*** *sobre la σ-álgebra generada por $\mathcal{E} = \{E_{n_j} : j \in \mathbb{N}\}$.*

Prueba. Usemos el Teorema 1.2.12, página 17, para obtener una partición infinita $(N_k^1)_{k=1}^{\infty}$ de $\mathbb{N}$ donde cada N_k^1 es infinito. Por la observación hecha anteriormente, sabemos que la serie

$$\sum_{k=1}^{\infty} |\nu|\Bigg(\bigcup_{j\in N_k^1} E_j\Bigg) \tag{α}$$

converge. También recordemos que si una serie $\sum_{n=1}^{\infty} a_n$ converge, entonces $\lim_{n\to\infty} a_n = 0$ y, por consiguiente, existe al menos un $n_0 \in \mathbb{N}$ tal que $|a_n| < 1$. Pues bien, usemos éste hecho aplicándolo a la serie (α) para obtener algún N_k^1, al que denotaremos por N_1, tal que

$$|\nu|\Bigg(\bigcup_{j\in N_1} E_j\Bigg) < 1/2.$$

Sea $n_1 = \min N_1$. Particione el conjunto $N_1 \setminus \{n_1\}$ en una sucesión disjunta de conjuntos infinitos $(N_k^2)_{k=1}^{\infty}$. Como antes, existe un índice $k \in \mathbb{N}$ tal que

$$|\nu|\Bigg(\bigcup_{j\in N_k^2} E_j\Bigg) < 1/2^2.$$

Pongamos $N_2 = N_k^2$ y sea $n_2 = \min N_2$. Entonces $n_2 > n_1$ y $N_2 \subseteq N_1$. Continuando con este mecanismo se logra producir una sucesión decreciente $(N_k)_{k=1}^{\infty}$ de subconjuntos infinitos de $\mathbb{N}$ y una sucesión estrictamente creciente de números naturales $(n_k)_{k=1}^{\infty}$ tal que

$$|\nu|\Bigg(\bigcup_{j\in N_k} E_j\Bigg) < 1/2^k \tag{α_1}$$

para todo $k \geq 1$. Para cada $k \in \mathbb{N}$, sea $\mathcal{E}_k = \{E_j : j \in N_k\}$. Entonces $\mathcal{E} = \bigcup_{k=1}^{\infty} \mathcal{E}_k$ es una subsucesión de la sucesión $(E_n)_{n=1}^{\infty}$, la que escribiremos como $\mathcal{E} = \{E_{m_j} : j \in \mathbb{N}\}$. Sea $\mathcal{M}_0$ la σ-álgebra generada por $\mathcal{E}$.

Afirmamos que ν es numerablemente aditiva sobre $\mathcal{M}_0$. En efecto, sea $(G_n)_{n=1}^{\infty}$ una sucesión decreciente en $\mathcal{M}_0$ tal que $\bigcap_{n=1}^{\infty} G_n = \varnothing$. Si existe un $m \in \mathbb{N}$ tal que $G_n = \varnothing$ para todo $n \geq m$, el teorema queda demostrado. Suponga entonces que $G_n \neq \varnothing$ para cualquier $n \in \mathbb{N}$. Puesto $G_n \in \mathcal{M}_0$ por cada $n \in \mathbb{N}$, entonces existe un conjunto $I_n \subseteq \mathbb{N}$ tal que

$$G_n = \bigcup_{j\in I_n} E_{m_j}$$

Claramente $I_{n+1} \subseteq I_n$ y puesto que ninguno de los E_{n_k} está contenido en $\bigcap_{n=1}^{\infty} G_n$ se tiene que $\lim_{n\to\infty} I_n = +\infty$. Para un arbitrario $\varepsilon > 0$, escoja un $k \in \mathbb{N}$ tal que $1/2^k < \varepsilon$. Entonces, para todo $n \geq k$,

$$G_n = \bigcup_{j\in I_n} E_{m_j} \subseteq \bigcup_{j\in N_k} E_{m_j}$$

y, en consecuencia, usando (α_1) se concluye que

$$|\nu|(G_n) \leq |\nu|\Bigg(\bigcup_{j\in N_k} E_j\Bigg) < 1/2^k < \varepsilon$$

para todo $n \geq k$. Esto prueba que $\lim_{n\to\infty} |\nu|(G_n) = 0$ y puesto que $|\nu(E)| \leq |\nu|(E)$ para todo $E \in \mathcal{M}$, resulta que $\lim_{n\to\infty} \nu(G_n) = 0$. Se sigue entonces del Teorema 12.1.5 que ν es numerablemente aditiva y termina la prueba. ■

Lo que resulta interesante es que la conclusión del Teorema de Drewnowski es válida para cualquier colección numerable de medidas con signo finitamente aditivas y acotadas.

Corolario 12.1.21 (Drewnowski). *Sea $(X, \mathcal{M})$ un espacio medible y suponga que para cada $n \in \mathbb{N}$, $\nu_n : \mathcal{M} \to \mathbb{R}$ es una **medida finitamente aditiva** y **acotada**. Si $(E_n)_{n=1}^{\infty}$ es una sucesión disjunta en $\mathcal{M}$, entonces existe una subsucesión $(E_{n_j})_{j=1}^{\infty}$ de $(E_n)_{n=1}^{\infty}$ tal que cada ν_n es **numerablemente aditiva** sobre la σ-álgebra generada por $\mathcal{E} = \{E_{n_j} : j \in \mathbb{N}\}$.*

Prueba. Defina $\nu : \mathcal{M} \to \mathbb{R}$ por

$$\nu(E) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{|\nu_n|(E)}{1 + |\nu|(X)} \quad \text{para todo } E \in \mathcal{M}$$

Resulta que ν es una medida con signo acotada y finitamente aditiva y, entonces, por el Teorema de Drewnowski existe una subsucesión $(E_{n_j})_{j=1}^{\infty}$ de $(E_n)_{n=1}^{\infty}$ tal que ν es numerablemente aditiva sobre $\mathcal{M}_0$, la σ-álgebra generada por $\mathcal{E} = \{E_{n_j} : j \in \mathbb{N}\}$. Sea ahora $(G_j)_{j=1}^{\infty}$ una sucesión decreciente en $\mathcal{M}_0$ tal que $\bigcap_{j=1}^{\infty} G_j = \varnothing$. Entonces

$$\lim_{j\to\infty} \nu(G_j) = 0$$

y, por lo tanto,

$$\lim_{j\to\infty} \nu_n(G_j) = 0$$

para cada $n \in \mathbb{N}$. Se sigue del Teorema 12.1.5 que cada ν_n es numerablemente aditiva y termina la prueba. ■

12.1.2. Integración Respecto a una Medida con Signo

¿Cómo definir $L_1(\nu)$ si ν es una medida con signo? Para hacer esto, fijemos un espacio de medida con signo $(X, \mathcal{M}, \nu)$. Diremos que

(a) $(X, \mathcal{M}, \nu)$ es **completa** si para cualquier conjunto ν-nulo N se cumple que

$$E \subseteq N \quad \Rightarrow \quad E \in \mathcal{M}.$$

(b) $(X, \mathcal{M}, \nu)$ es σ**-finita** si existe una sucesión $(E_n)_{n=1}^{\infty}$ en $\mathcal{M}$ tal que

$$X = \bigcup_{n=1}^{\infty} E_n \qquad \text{y} \qquad \nu(E_n) \in \mathbb{R} \quad \text{para todo } n \in \mathbb{N}.$$

Observe que si definimos

$$F_1 = E_1 \qquad \text{y} \qquad F_n = E_n \setminus \bigcup_{j=1}^{n-1} E_j \quad \text{para todo } n \geq 2,$$

entonces la sucesión $(F_n)_{n=1}^{\infty}$ es medible, disjunta y $\bigcup_{n=1}^{\infty} F_n = \bigcup_{n=1}^{\infty} E_n$. Más aun, puesto que $F_n \subseteq E_n$, se sigue del Lema 12.1.3 que $\nu(F_n) \in \mathbb{R}$ para todo $n \in \mathbb{N}$. Finalmente, si hacemos $G_n = \bigcup_{j=1}^{n} F_j$, entonces $(G_n)_{n=1}^{\infty}$ resulta ser una sucesión creciente en $\mathcal{M}$ tal que

$$\bigcup_{n=1}^{\infty} G_n = \bigcup_{n=1}^{\infty} F_n = X \qquad \text{y} \qquad \nu(G_n) = \sum_{j=1}^{n} \nu(F_j) \quad \text{para todo } n \in \mathbb{N}.$$

Es un ejercicio sencillo, que dejaremos a cargo del lector, establecer lo siguiente:

(III) $(X, \mathcal{M}, \nu)$ es un espacio de **medida con signo completa** si, y sólo si, $(X, \mathcal{M}, |\nu|)$ es **completa**.

(IV) $(X, \mathcal{M}, \nu)$ es un espacio de **medida con signo σ-finita** si, y sólo si $(X, \mathcal{M}, |\nu|)$ es **σ-finita**.

La integración con respecto a una medida con signo ν, la cual se puede asumir que es completa y σ-finita, se define del modo esperado:

$$\mathrm{L}_1(\nu) = \mathrm{L}_1(X, \mathcal{M}, |\nu|),$$

y si $f \in \mathrm{L}_1(\nu)$, entonces

$$\int_X f\, d\nu = \int_X f\, d\nu^+ - \int_X f\, d\nu^-.$$

Observe que de la igualdad

$$\int_X |f|\, d\nu^+ + \int_X |f|\, d\nu^- = \int_X |f|\, d|\nu|$$

se sigue que

$$\mathrm{L}_1(\nu) = \mathrm{L}_1(\nu^+) \cap \mathrm{L}_1(\nu^-).$$

Muchas afirmaciones demostradas en el caso del espacio de Lebesgue $\mathrm{L}_1(\mathbb{R}, \mathfrak{M}_\mu(\mathbb{R}), \mu)$ se trasladan, casi sin cambio, al espacio $\mathrm{L}_1(\nu)$. Por ejemplo:

(a) Si $f \in \mathrm{L}_1(\nu)$, entonces

$$\left| \int_X f\, d\nu \right| \leq \int_X |f|\, d|\nu|.$$

En efecto,

$$\begin{aligned} \left| \int_X f\, d\nu \right| &= \left| \int_X f\, d\nu^+ - \int_X f\, d\nu^- \right| \leq \left| \int_X f\, d\nu^+ \right| + \left| \int_X f\, d\nu^- \right| \\ &\leq \int_X |f|\, d\nu^+ + \int_X |f|\, d\nu^- = \int_X |f|\, d|\nu|. \end{aligned}$$

(b) El Teorema de la Convergencia Dominada vale en $\mathrm{L}_1(\nu)$.

Es importante advertir, sin embargo, que existen afirmaciones que no son ciertas en este contexto. Por ejemplo:

(i) La relación $f \leq g$ no implica, en general, que $\int_X f\, d\nu \leq \int_X g\, d\nu$.

(ii) El Lema de Fatou, así como el Teorema de la Convergencia Monótona, tampoco se cumplen para medidas con signo.

12.1.3. El Teorema de Radon-Nikodým

En la Definición 10.2.8, página 582, introducimos un modo de comparar dos medidas positivas a través de la noción de absoluta continuidad. Dicha noción es susceptible de ser generalizada para medidas con signo del modo siguiente.

Definición 12.1.22. *Sean* λ, $\nu \in \mathrm{ca}(\mathcal{M}, \mathbb{R})$. *Diremos que* λ *es* ***absolutamente continua*** *con respecto a* ν, *en símbolos* $\lambda \ll \nu$, *si* $\lambda(E) = 0$ *para cualquier conjunto medible* E *para el cual* $|\nu|(E) = 0$.

El siguiente resultado reduce, como en el caso de la Definición 10.2.8, a considerar sólo medidas positivas.

Teorema 12.1.23. *Sean* λ, $\nu \in \mathrm{ca}(\mathcal{M}, \mathbb{R})$. *Las siguientes condiciones son equivalentes:*

(1) $\lambda \ll \nu$.

(2) $\lambda^+ \ll \nu$ *y* $\lambda^- \ll \nu$.

(3) $|\lambda| \ll \nu$.

Prueba. (1) $\Rightarrow$ (2). Sea (A, B) una ν-descomposición de Hahn para X y sea $E \in \mathcal{M}$ tal que $|\nu|(E) = 0$. Entonces

$$0 \ \leq \ |\nu|(A \cap E) \ \leq \ |\nu|(E) \ = \ 0,$$

es decir, $|\nu|(A \cap E) = 0$. Similarmente, $|\nu|(B \cap E) = 0$. Puesto que $\lambda \ll \nu$, se sigue de lo anterior que $\lambda(A \cap E) = 0$ y $\lambda(B \cap E) = 0$; esto es, $\lambda^+(E) = 0 = \lambda^-(E)$. Por esto, $\lambda^+ \ll \nu$ y $\lambda^- \ll \nu$.
(2) $\Rightarrow$ (3). Sea $E \in \mathcal{M}$ tal que $|\nu|(E) = 0$. Entonces

$$\lambda^+ \ll \nu \ \Rightarrow \ \lambda^+(E) \ = \ 0 \qquad \text{y} \qquad \lambda^- \ll \nu \ \Rightarrow \ \lambda^-(E) \ = \ 0.$$

De aquí se sigue que $|\lambda|(E) = \lambda^+(E) + \lambda^-(E) = 0$ y, por lo tanto, $|\lambda| \ll \nu$.
(3) $\Rightarrow$ (1). Sea $E \in \mathcal{M}$ tal que $|\nu|(E) = 0$. Entonces $|\lambda|(E) = 0$ y como $0 \leq |\lambda(E)| \leq |\lambda|(E) = 0$, resulta que $\lambda(E) = 0$ y finaliza la prueba. ■

Similar al Teorema 10.2.9 se tiene el siguiente resultado para medidas reales.

Teorema 12.1.24. *Sean* $\lambda, \nu \in \mathrm{ca}(\mathcal{M}, \mathbb{R})$. *Las siguientes condiciones son equivalentes:*

(1) $\lambda \ll \nu$.

(2) *Para cada* $\varepsilon > 0$, *existe un* $\delta > 0$ *con la siguiente propiedad: para cualquier conjunto* $E \in \mathcal{M}$ *satisfaciendo* $|\nu|(E) < \delta$, *se cumple que* $|\lambda(E)| < \varepsilon$.

Prueba. Puesto que λ y ν son medidas reales, se sigue del Teorema 12.1.14 que $|\lambda|$ y $|\nu|$ son medidas finitas. Más aun, por el Teorema 12.1.23 sabemos que

$$\lambda \ll \nu \quad \Leftrightarrow \quad |\lambda| \ll \nu \quad \Leftrightarrow \quad |\lambda| \ll |\nu|.$$

Esta observación permite concluir que es suficiente demostrar el teorema suponiendo que ambas medidas son positivas. Claramente (2) $\Rightarrow$ (1). Para demostrar la otra implicación, suponga que (1) es cierto, pero que (2) no se cumple. Similar a la demostración del Teorema 10.2.9, página 582, para cada $n \in \mathbb{N}$ podemos encontrar un $A_n \in \mathcal{M}$ tal que

$$\nu(A_n) \ < \ \frac{1}{2^n} \quad \text{pero} \quad \lambda(A_n) \ \geq \ \varepsilon.$$

Definamos

$$A = \limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty} A_m.$$

Observe que, para cada $n \geq 1$, $A \subseteq \bigcup_{m=n}^{\infty} A_m$ y, por lo tanto,

$$\nu(A) \leq \sum_{m=n}^{\infty} \nu(A_m) < \sum_{m=n}^{\infty} \frac{1}{2^m} = \frac{1}{2^{n-1}}.$$

Puesto que esto es válido para todo $n \geq 1$, se sigue que $\nu(A) = 0$. Por otro lado, como λ es una medida finita resulta que $\lambda(\bigcup_{n=1}^{\infty} A_n) < +\infty$ y, en consecuencia, si tomamos $B_n = \bigcup_{m=n}^{\infty} A_m$ para cada $n \geq 1$, tenemos que $(B_n)_{n=1}^{\infty}$ es una sucesión decreciente de conjuntos medible tal que

$$\lambda(B_1) < +\infty \qquad \text{y} \qquad \lambda(B_n) \geq \lambda(A_n) \geq \varepsilon \quad \text{para todo } n \in \mathbb{N}.$$

Se sigue entonces del Teorema 12.1.4 que

$$\lambda(A) = \lambda\Big(\bigcap_{n=1}^{\infty} B_n\Big) = \lim_{n\to\infty} \lambda(B_n) \geq \varepsilon > 0.$$

De este modo ha quedado establecida la existencia de un conjunto medible A tal que que $\nu(A) = 0$ pero con $\lambda(A) > 0$, lo cual nos dice que λ no es absolutamente continua con respecto a ν. Esta contradicción termina la prueba. ■

Situaciones similares a los dos resultados anteriores se obtienen reemplazando el símbolo $\ll$ por $\perp$, esto es:

Teorema 12.1.25. *Sean $\lambda, \nu \in \mathrm{ca}(\mathcal{M}, \mathbb{R})$. Las siguientes condiciones son equivalentes:*

(1) $\lambda \perp \nu$.

(2) $|\lambda| \perp \nu$.

(3) $\lambda^+ \perp \nu$ *y* $\lambda^- \perp \nu$.

Prueba. La equivalencia $\lambda \perp \nu \Leftrightarrow |\lambda| \perp \nu$ es consecuencia inmediata de (I); es decir, un conjunto $N \in \mathcal{M}$ es λ-nulo si, y sólo si, $|\lambda|(N) = 0$.
Suponga ahora que $|\lambda| \perp \nu$. Entonces existen conjuntos medibles y disjuntos P, Q tales que

$$X = P \cup Q \qquad \text{y} \qquad |\lambda|(P) = \nu(Q) = 0.$$

Esto implica, en particular, que $\lambda^+(P) = \lambda^-(P) = 0$ y, por consiguiente, $\lambda^+ \perp \nu$ y $\lambda^- \perp \nu$.
Recíprocamente, suponga que $\lambda^+ \perp \nu$ y $\lambda^- \perp \nu$. Seleccione conjuntos medibles P^+, Q^+, P^-, Q^- tales que

$$\begin{aligned} X &= P^+ \cup Q^+, \quad P^+ \cap Q^+ = \varnothing \qquad \text{y} \qquad \lambda^+(P^+) = \nu(Q^+) = 0,\\ X &= P^- \cup Q^-, \quad P^- \cap Q^- = \varnothing \qquad \text{y} \qquad \lambda^-(P^-) = \nu(Q^-) = 0. \end{aligned}$$

Considere ahora los conjuntos medibles P y Q definidos por:

$$P = P^+ \cap P^- \qquad \text{y} \qquad Q = X \setminus P = Q^+ \cup Q^-.$$

Entonces, puesto que $\lambda^+(P^+) = 0$ y λ^+ es una medida positiva, resulta que $\lambda^+(P) = 0$. Similarmente, $\lambda^-(P^+) = 0$, de donde se obtiene que

$$|\lambda|(P) \;=\; \lambda^+(P) + \lambda^-(P) \;=\; 0.$$

Por otro lado, puesto que Q^+ y Q^- son conjuntos ν-nulos, entonces la aditividad de ν nos garantiza que $Q = Q^+ \cup Q^-$ también es ν-nulo. En resumen, hemos hallado dos conjuntos medibles y disjuntos P y Q tales que

$$X \;=\; P \cup Q \qquad \text{y} \qquad |\lambda|(P) \;=\; \nu(Q) \;=\; 0.$$

Esto nos indica que $|\lambda| \perp \nu$ y termina la prueba. ■

Teorema 12.1.26. *Sea ν una **medida positiva** definida sobre $(X, \mathcal{M})$ y sea $\lambda \in \mathrm{ca}(\mathcal{M}, \mathbb{R})$. Las siguientes condiciones son equivalentes:*

(1) $\lambda \perp \nu$.

(2) Dado $\varepsilon > 0$, existe un conjunto $E_\varepsilon \in \mathcal{M}$ tal que

$$\nu(E_\varepsilon) \;<\; \varepsilon \qquad \text{y} \qquad |\lambda|(X \setminus E_\varepsilon) \;<\; \varepsilon.$$

Prueba. $(1) \Rightarrow (2)$. Suponga que (1) se cumple y sea $\varepsilon > 0$. Use el Teorema 12.1.25 (2) para elegir un par de conjuntos medibles y disjuntos P y Q tales que

$$X \;=\; P \cup Q \qquad \text{y} \qquad |\lambda|(P) \;=\; \nu(Q) \;=\; 0.$$

Si ahora definimos $E_\varepsilon = Q$ resulta que (2) se cumple. Para demostrar la otra implicación, aceptemos que (2) se cumple. Con $\varepsilon = 1/2^n$ para cada $n \in \mathbb{N}$, seleccione un conjunto medible E_n tal que

$$\nu(E_n) \;<\; \frac{1}{2^n} \qquad \text{y} \qquad |\lambda|(X \setminus E_n) \;<\; \frac{1}{2^n}.$$

Defina

$$P \;=\; \limsup_{n\to\infty} E_n \;=\; \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} E_n.$$

Por el Lema de Borel-Cantelli, Lema 6.3.49, página 288, se tiene que

$$\nu(P) \;=\; \limsup_{n\to\infty} \nu(E_n) \;=\; 0.$$

Por otro lado, como

$$E^c \;=\; \liminf_{n\to\infty} E_n^c \;=\; \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} E_n^c,$$

entonces, para cualquier $k \in \mathbb{N}$,

$$|\lambda|\Big(\bigcap_{n=k}^{\infty} E_n^c\Big) \;\leq\; |\lambda|(E_k^c) \quad \text{para todo } n \geq k.$$

Puesto que la sucesión $(\bigcap_{n=k}^{\infty} E_n^c)_{k=1}^{\infty}$ es medible y creciente, la continuidad de $|\lambda|$ nos asegura que

$$|\lambda|(E^c) \;=\; \lim_{k\to\infty} |\lambda|\Big(\bigcap_{n=k}^{\infty} E_n^c\Big) \;\leq\; \lim_{k\to\infty} |\lambda|(E_k^c) \;=\; 0.$$

Tomando $Q = E^c$ vemos que $\lambda \perp \nu$ y concluye la prueba. ■

Como afirmamos al comienzo de esta sección, si $(X, \mathcal{M}, \nu)$ es un espacio de medida positiva y $f \in L_1(X, \mathcal{M}, \nu)$, entonces la función de conjuntos $\lambda : \mathcal{M} \to \mathbb{R}$ definida por

$$\lambda(E) = \int_E f\, d\nu \quad \text{para todo } E \in \mathcal{M}$$

es una medida real sobre $(X, \mathcal{M})$. Además, $\lambda \ll \nu$. El Teorema de Radon-Nikodým afirma que si λ y ν son medidas reales σ-finitas con ν positiva y $\lambda \ll \nu$, entonces la medida λ puede ser representa por la igualdad anterior, es decir, dicha igualdad constituye la regla y no la excepción. La demostración será llevada a cabo primero, considerando el caso en que ambas medidas son positivas y finitas, luego el caso en que ambas son σ-finitas y finalmente considerando el caso general, es decir, que ambas son σ-finitas, ν es positiva y λ una medida con signo.

Antes de formular y demostrar el Teorema de Radon-Nikodým en su versión más simple, necesitaremos el siguiente lema auxiliar.

Lema 12.1.27. *Sean $\nu, \lambda : (X, \mathcal{M}) \to [0, +\infty)$ **medidas positivas finitas**. Entonces ocurre una, y sólo una, de las dos condiciones siguientes:*

(a) $\lambda \perp \nu$.

*(b) Existe un $\varepsilon > 0$ y un conjunto $E_\varepsilon \in \mathcal{M}$ tal que $\nu(E_\varepsilon) > 0$ y E_ε es un **conjunto positivo** para $\lambda - \varepsilon \cdot \nu$.*

Prueba. Para cada $n \in \mathbb{N}$, considere la medida real $\nu_n = \lambda - \frac{1}{n} \cdot \nu$ y sea (A_n, B_n) una ν_n-descomposición de Hahn para X. Defina ahora

$$A_0 = \bigcup_{n=1}^{\infty} A_n \qquad \text{y} \qquad B_0 = \bigcap_{n=1}^{\infty} B_n.$$

Puesto que $B_0 \subseteq B_n$ y cada B_n es un conjunto ν_n-negativo, resulta que

$$\lambda(B_0) - \frac{1}{n} \cdot \nu(B_0) = \nu_n(B_0) \leq 0,$$

es decir, $0 \leq \lambda(B_0) \leq \frac{1}{n} \cdot \nu(B_0)$ para cualquier $n \geq 1$. De allí que $\lambda(B_0) = 0$.

(a) Si ocurre que $\nu(A_0) = 0$, entonces se tendrá que $\lambda \perp \nu$.

(b) Si, por el contrario, $\nu(A_0) > 0$, entonces la subaditividad de ν nos revela que

$$0 < \nu(A_0) \leq \sum_{n=1}^{\infty} \nu(A_n),$$

de donde se sigue la existencia de un $n_0 \in \mathbb{N}$ tal que $\nu(A_{n_0}) > 0$. Como A_{n_0} es un conjunto ν_{n_0}-positivo, entonces haciendo $E_\varepsilon = A_{n_0}$ y $\varepsilon = 1/n_0$ se obtiene que $\nu(E_\varepsilon) > 0$ y E_ε es un conjunto positivo para $\lambda - \varepsilon \cdot \nu > 0$. La prueba es completa. ■

Teorema 12.1.28 (Radon-Nikodým - Caso finito). *Sea $(X, \mathcal{M}, \nu)$ un espacio de* ***medida positiva finita*** *con $\nu(X) > 0$. Si λ es cualquier otra* ***medida positiva finita*** *definida sobre $(X, \mathcal{M})$ tal que $\lambda \ll \nu$, entonces existe una función $f \in L_1(X, \mathcal{M}, \nu)$ verificando la igualdad*

$$\lambda(E) = \int_E f\, d\nu \quad \text{para todo } E \in \mathcal{M}. \tag{$*$}$$

La función f es ***única*** *en el siguiente sentido: si $g : X \to [0, +\infty)$ es otra función ν-integrable satisfaciendo la igualdad $(*)$, entonces $f = g$ ν-casi-siempre.*

Prueba. Considere el subconjunto $\mathcal{F}_\lambda$ de $L_1(X, \mathcal{M}, \nu)$ definido del modo siguiente:

$$\mathcal{F}_\lambda = \left\{ f \in L_1(X, \mathcal{M}, \nu) : f \geq 0 \ \text{ y } \ \int_E f\, d\nu \leq \lambda(E) \ \text{ para todo } E \in \mathcal{M} \right\}.$$

Observe que $\mathcal{F}_\lambda$ es no vacío ya que $0 \in \mathcal{F}_\lambda$. También, si $f, g \in \mathcal{F}_\lambda$, entonces $h = \max\{f, g\} \in \mathcal{F}_\lambda$. En efecto, tomando $A = \{x \in X : f(x) > g(x)\}$ se tiene que cualquiera sea el conjunto $E \in \mathcal{M}$,

$$\int_E h\, d\nu = \int_{E\cap A} h\, d\nu + \int_{E\setminus A} h\, d\nu = \int_{E\cap A} f\, d\nu + \int_{E\setminus A} g\, d\nu \leq \lambda(E \cap A) + \lambda(E \setminus A) = \lambda(E).$$

Sea

$$\alpha = \sup\left\{ \int_E f\, d\nu : f \in \mathcal{F}_\lambda \right\}$$

y observe que $0 \leq \alpha \leq \lambda(X) < +\infty$. Nuestro objetivo será demostrar que dicho supremo se alcanza, es decir, existe una función $f \in \mathcal{F}_\lambda$ tal que $\alpha = \int_X f\, d\nu$. Para verificar esto, usemos la definición de α para hallar una sucesión $(f_n)_{n=1}^\infty$ en $\mathcal{F}_\lambda$ de modo que

$$\alpha = \lim_{n\to\infty} \int_E f_n\, d\nu, \quad \text{para cada } E \in \mathcal{M}.$$

Por lo visto anteriormente, y aplicando un simple argumento inductivo, se tiene que la función $h_n = \max\{f_1, \ldots, f_n\} \in \mathcal{F}_\lambda$ para cada $n \geq 1$. Puesto que la sucesión $(h_n)_{n=1}^\infty$ es creciente y acotada superiormente, ella converge puntualmente a una función f. En particular, $0 \leq h_n \leq f$ para todo $n \geq 1$. Veamos que $f \in \mathcal{F}_\lambda$. En efecto, para cada $E \in \mathcal{M}$, el Teorema de la Convergencia Monótona nos revela que

$$\int_E f\, d\nu = \lim_{n\to\infty} \int_E h_n\, d\nu \leq \lambda(E).$$

y, así, $f \in \mathcal{F}_\lambda$. Por otro lado, como

$$\int_X f_n\, d\nu \leq \int_X h_n\, d\nu \leq \int_X f\, d\nu$$

resulta que

$$\int_X f\, d\nu \leq \alpha = \lim_{n\to\infty} \int_X f_n\, d\nu \leq \int_X f\, d\nu,$$

y, por lo tanto, $\alpha = \int_X f\, d\nu$. Veamos ahora que f es el candidato que andamos buscando. Para verificar esto, considere la medida $v_0 : \mathcal{M} \to [0, +\infty)$ dada por

$$v_0(E) = \lambda(E) - \int_E f\, d\nu \quad \text{para todo } E \in \mathcal{M}.$$

Observe que como $f \in \mathcal{F}_\lambda$, entonces v_0 es una medida positiva y finita. Nos proponemos demostrar que $v_0 \perp \nu$. Suponga, para construir una contradicción, que la relación $v_0 \perp \nu$ no es válida. Un llamado al Lema 12.1.27 nos garantiza la existencia de un $\varepsilon > 0$ y un conjunto medible E_ε tal que $\nu(E_\varepsilon) > 0$ y E_ε es un conjunto positivo para $v_0 - \varepsilon \cdot \nu$. Tome cualquier $E \in \mathcal{M}$. Entonces

$$\lambda(E) - \int_E f\,d\nu \;=\; v_0(E) \;\geq\; v_0(E \cap E_\varepsilon) \;\geq\; \varepsilon \cdot \nu(E \cap E_\varepsilon) \;=\; \int_E \varepsilon \cdot \chi_{E_\varepsilon}\,d\nu$$

de donde se sigue que

$$\lambda(E) \;\geq\; \int_E (f + \varepsilon \cdot \chi_{E_\varepsilon})\,d\nu.$$

Esto prueba que $f + \varepsilon \cdot \chi_{E_\varepsilon} \in \mathcal{F}_\lambda$ lo que, por supuesto, genera la siguiente contradicción:

$$\int_X (f + \varepsilon \cdot \chi_{E_\varepsilon})\,d\nu \;=\; \alpha + \varepsilon \cdot \nu(X) \;>\; \alpha.$$

La conclusión es que lo que habíamos supuesto es falso y, por lo tanto, la relación $v_0 \perp \nu$ es verdadera. Seleccione entonces conjuntos medibles y disjuntos P y Q tales que

$$X \;=\; P \cup Q \qquad \text{y} \qquad v_0(P) \;=\; \nu(Q) \;=\; 0.$$

Hasta ahora no hemos utilizado el hecho de que $\lambda \ll \nu$. Llegó el momento. Puesto que $\lambda \ll \nu$ y $\nu(Q) = 0$, resulta que $\lambda(Q) = 0$ y, en consecuencia,

$$v_0(Q) \;=\; \lambda(Q) - \int_Q f\,d\nu \;=\; 0.$$

De allí que $v_0(X) = v_0(P) + v_0(Q) = 0$ y así,

$$\lambda(E) \;=\; \int_E f\,d\nu$$

para todo $E \in \mathcal{M}$.

Falta demostrar la unicidad de f. Para este fin, suponga que g es otra función medible no-negativa tal que $\lambda(E) = \int_E g\,d\nu$ para todo $E \in \mathcal{M}$. Entonces

$$\int_X (f - g)\,d\nu \;=\; 0$$

y se sigue del Teorema 11.1.40, página 714, que $f = g$ ν-casi-siempre. La prueba es completa. ■

La función f satisfaciendo la igualdad $(*)$ en el Teorema de Radon-Nikodým se denomina la **derivada de Radon-Nikodým de λ con respecto a ν** y, por lo general, se le denota por el símbolo

$$f \;=\; \frac{d\lambda}{d\nu}$$

y $(*)$ se escribe en la forma:

$$d\lambda \;=\; f \cdot d\nu \;=\; \frac{d\lambda}{d\nu} \cdot d\nu \qquad \text{o} \qquad \lambda \;=\; f\,d\nu.$$

Nota Adicional 12.1.1 Observe que la construcción de la función f en la prueba del Teorema de Radon-Nikodým nada tuvo que ver con la condición $\lambda \ll \nu$, por consiguiente, podemos usar dicha función para dar una prueba de la siguiente descomposición de λ.

Teorema 12.1.29 (Descomposición de Lebesgue). *Sean λ y ν dos **medidas positivas finitas** definidas sobre $(X, \mathcal{M})$. Entonces existen medidas positivas λ_a y λ_s sobre $(X, \mathcal{M})$ tales que:*

$$\lambda_a \ll \nu, \qquad \lambda_s \perp \nu \qquad y \qquad \lambda = \lambda_a + \lambda_s.$$

Prueba. Sea f_0 la función obtenida en la prueba del Teorema de Radon-Nikodým; en otras palabras, $f_0 \in \mathcal{F}_\lambda$ y satisface

$$\int_E f_0\, d\nu = \sup\left\{ \int_E f\, d\nu \,:\, f \in \mathcal{F}_\lambda \right\}.$$

Defina ahora

$$\lambda_a(E) = \int_E f_0\, d\nu \qquad \text{y} \qquad \lambda_s(E) = \lambda(E) - \lambda_a(E)$$

para todo $E \in \mathcal{M}$. Se tiene entonces que λ_a es una medida positiva y, en consecuencia, λ_s también lo es y ambas satisfacen las conclusiones impuestas en el Teorema de Descomposición de Lebesgue. ■

Nuestro siguiente paso es demostrar el Teorema de Radon-Nikodým asumiendo que ambas medidas λ y ν son σ-finitas.

Teorema 12.1.30 (Radon-Nikodým - Caso σ-finito). *Sea $(X, \mathcal{M}, \nu)$ un espacio de **medida positiva σ-finita**. Si λ es cualquier otra **medida positiva σ-finita** definida sobre $(X, \mathcal{M})$ tal que $\lambda \ll \nu$, entonces existe una única función medible $f : X \to [0, +\infty]$ tal que*

$$\lambda(E) = \int_E f\, d\nu \quad \textit{para todo } E \in \mathcal{M}.$$

Prueba. Puesto que ν y λ son σ-finitas, existe un par de colecciones numerables y disjuntas de conjuntos medibles, digamos, $(E_n)_{n=1}^\infty$ y $(F_n)_{n=1}^\infty$ tales que

$$X = \bigcup_{n=1}^\infty E_n = \bigcup_{n=1}^\infty F_n \qquad \text{y} \qquad \nu(E_n),\ \lambda(F_n) < +\infty, \quad n = 1, 2, \ldots$$

Para cada $m, n \in \mathbb{N}$ defina $G_{mn} = E_m \cap F_n$. Se tiene entonces que

$$E_m = E_m \cap X = E_m \cap \left(\bigcup_{n=1}^\infty F_n \right) = \bigcup_{n=1}^\infty (E_m \cap F_n) = \bigcup_{n=1}^\infty G_{mn},$$

$$X = \bigcup_{m=1}^\infty E_m = \bigcup_{m=1}^\infty \left(\bigcup_{n=1}^\infty (E_m \cap F_n) \right) = \bigcup_{m,n=1}^\infty G_{mn}.$$

Más aun, como $G_{mn} \cap G_{ij} = \varnothing$ si, y sólo si, $m \neq i$ o $n \neq j$, resulta que la colección disjunta de conjuntos medibles $\mathcal{G} = \{G_{mn} : m, n \in \mathbb{N}\}$ es una partición medible de X. Sea $\{G_m : m \in \mathbb{N}\}$ una enumeración de $\mathcal{G}$ y para cada $m \in \mathbb{N}$, defina $\nu_m, \lambda_m : \mathcal{M} \to [0, +\infty)$ por

$$\nu_m(E) = \nu(E \cap G_m) \qquad \text{y} \qquad \lambda_m(E) = \lambda(E \cap G_m).$$

Puesto que $\nu_m(X) \leq \nu(E_m) < +\infty$ y $\lambda_m(X) \leq \lambda(E_m) < +\infty$ se tiene que las medidas ν_m y λ_m son finitas. Veamos ahora que $\lambda_m \ll \nu_m$ para todo $m \in \mathbb{N}$. En efecto, fijemos un $m \in \mathbb{N}$ y sea $E \in \mathcal{M}$ tal que $\nu_m(E) = 0$. Entonces $\nu(E \cap G_m) = 0$ y puesto que $\lambda \ll \nu$, resulta que $\lambda_m(E) = \lambda(E \cap G_m) = 0$. Esto prueba que $\lambda_m \ll \nu_m$ y así, gracias al Teorema de Radon-Nikodým - Caso finito existe, por cada $m \in \mathbb{N}$, una función medible no-negativa f_m tal que

$$\lambda_m(E) = \int_E f_m \, d\nu_m \quad \text{para todo } E \in \mathcal{M}.$$

Nótese que, por cada $m \in \mathbb{N}$, $\nu_m(X \setminus G_m) = \nu(\varnothing) = 0$ y, en consecuencia, como $\lambda_m \ll \nu_m$,

$$\int_{X \setminus G_m} f_m \, d\nu_m = \lambda_m(X \setminus G_m) = 0$$

Se sigue del Teorema 11.1.40, página 714, que $f_m(x) = 0$ para ν_m-casi todo $x \in X \setminus G_m$. Sin perder generalidad, asumiremos que $f_m(x) = 0$ para todo $x \in X \setminus G_m$. Puesto que $(G_m)_{m=1}^{\infty}$ es una sucesión disjunta cubriendo a X resulta que para cada $x \in X$, existe un único $m \in \mathbb{N}$ tal que $x \in G_m$. Defina entonces $f : X \to [0, +\infty)$ por

$$f(x) = \begin{cases} f_m(x) & \text{si } x \in G_m \\ 0 & \text{si } x \notin G_m. \end{cases}$$

Observe que, en este caso, f también se puede representar en la forma $f = \sum_{m=1}^{\infty} f_m$. Resta demostrar que

$$\lambda(E) = \int_E f \, d\nu \quad \text{para todo } E \in \mathcal{M}.$$

En efecto, por una aplicación del Teorema de Beppo Levi, Corolario 11.1.37, página 712 y la definición de λ_m, se tiene que

$$\begin{aligned} \int_E f \, d\nu &= \int_{\bigcup_{m=1}^{\infty}(E \cap G_m)} f \, d\nu = \sum_{m=1}^{\infty} \int_{E \cap G_m} f \, d\nu \\ &= \sum_{m=1}^{\infty} \int_{E \cap G_m} f_m \, d\nu_m = \sum_{m=1}^{\infty} \lambda_m(E \cap G_m) \\ &= \sum_{m=1}^{\infty} \lambda(E \cap G_m) = \lambda(E). \end{aligned}$$

La unicidad de la función f es similar a la del caso finito y se omite. ■

Nota Adicional 12.1.2 Es importante destacar que la condición: "*ν es σ-finita*" en el resultado anterior no se puede eliminar. En efecto, considere el espacio medible $(\mathbb{R}, \mathcal{M})$, donde $\mathcal{M} = \{E \subseteq \mathbb{R} : E \text{ o } E^c \text{ es numerable}\}$. Para cada $E \in \mathcal{M}$, defina

$$\nu(E) = \operatorname{card}(E) \qquad \text{y} \qquad \lambda(E) = \begin{cases} 0 & \text{si } E \text{ es numerable} \\ 1 & \text{si } E \text{ no es numerable.} \end{cases}$$

Puesto que $\mathbb{R}$ es no-numerable, resulta que ν no puede ser σ-finita. También es claro que $\lambda \ll \nu$. Sin embargo, no existe ninguna función medible no-negativa f tal que

$$\lambda(E) = \int_E f \, d\nu \quad \text{para todo } E \in \mathcal{M}. \tag{1}$$

Para ver esto, suponga la existencia de una función medible no-negativa f tal que (1) se cumple. Observe que f no puede ser idénticamente cero, pues si tal cosa fuese cierta, entonces

$$1 = \lambda(\mathbb{R}) = \int_{\mathbb{R}} 0 \cdot d\nu = 0.$$

El hecho de $f \neq 0$ nos indica la existencia de algún $x_0 \in \mathbb{R}$ tal que $f(x_0) > 0$. Tomando $E = \{x_0\}$, resulta que $E \in \mathcal{M}$, $\lambda(E) = 0$, pero

$$0 = \lambda(E) = \int_{\mathbb{R}} f \, d\nu = f(x_0) > 0.$$

Esta contradicción confirma que la condición *ν es σ-finita* no puede ser eliminada en el Teorema de Radon-Nikodým.

Teorema 12.1.31 (Radon-Nikodým - Caso general). *Sea $(X, \mathcal{M}, \nu)$ un espacio de **medida positiva σ-finita**. Si λ es una **medida real σ-finita** definida sobre $(X, \mathcal{M})$ tal que $\lambda \ll \nu$, entonces existe una única función medible y ν-integrable $f : X \to \mathbb{R}$ tal que*

$$\lambda(E) = \int_E f \, d\nu \quad \textit{para todo } E \in \mathcal{M}.$$

Prueba. Si λ es positiva, el resultado sigue del Teorema 12.1.30. Suponga ahora que λ es real y sea $\lambda = \lambda^+ - \lambda^-$ la descomposición de Jordan de λ. Puesto que $\lambda \ll \nu$, el Teorema 12.1.23 nos garantiza que $\lambda^+ \ll \nu$ y $\lambda^- \ll \nu$. Claramente λ^+ y λ^- son medidas positivas finitas y se sigue del Teorema de Radon-Nikodým - Caso σ-finito, que existen funciones medibles únicas, no-negativas $f_1, f_2 \in L_1(X, \mathcal{M}, \nu)$ tales que

$$\lambda^+(E) = \int_E f_1 \, d\nu \qquad \text{y} \qquad \lambda^-(E) = \int_E f_2 \, d\nu$$

para todo $E \in \mathcal{M}$. Si ahora definimos $f = f_1 - f_2$, se tiene que $f \in L_1(X, \mathcal{M}, \nu)$ y

$$\lambda(E) = \lambda^+(E) - \lambda^-(E) = \int_E f_1 \, d\nu - \int_E f_2 \, d\nu = \int_E f \, d\nu$$

para todo $E \in \mathcal{M}$. ■

Como antes, a la función f en el Teorema de Radon-Nikodým - Caso general, se le denomina la **derivada de Radon-Nikodým de λ con respecto a ν** y se denota por

$$f = \frac{d\lambda}{d\nu}.$$

Nota Adicional 12.1.3 El Teorema de Radon-Nikodým fue demostrado, en su fase inicial, por H. Lebesgue en el año 1.910 para el caso en que ν era la medida de Lebesgue en $\mathbb{R}^n$. Posteriormente el asunto fue ampliado por M. Radon en 1.913 para cualquier medida de Borel regular arbitraria y generalizado por O. M. Nikodým en 1.930 para medidas sobre espacios abstractos. Por esta razón, algunos autores prefieren denominar el Teorema de Radon-Nikodým como el **Teorema de Lebesgue-Radon-Nikodým**. Existen, como ya hemos observado, muchas variantes y extensiones de éste resultado. Tal vez un buen comienzo para hurgar en ello sea el artículo de D. Candeloro y A. A. Volčič [106], pág. 251-294. Cuando se sustituye a $\mathbb{R}$ por un espacio de Banach arbitrario y la integral considerada es la integral de Bochner se obtiene un Teorema de Radon-Nikodým que, además de ser fascinante, es impresionante por los extraordinarios resultados que de él se derivan en la Teoría de los Espacios de Banach (véase, por ejemplo, el maravilloso libro de Diestel-Uhl [48]).

La versión del Teorema de Radon-Nikodým dada en el Teorema 12.1.31, sigue siendo válida si λ es una medida con signo σ-finita. Bajo ésta hipótesis sólo hay que considerar los casos $\lambda(X) = +\infty$ o $\lambda(X) = -\infty$. Véase, por ejemplo, la demostración del Teorema 11.16, pág. 234-236 en [141]. Existe, sin embargo, un caso más general para la validez del Teorema de Radon-Nikodým, el cual tiene que ver con la noción de medida descomposable.

Definición 12.1.32. *Sea $(X, \mathcal{M})$ un espacio medible. Una medida positiva ν sobre $(X, \mathcal{M})$ se llama **super σ-finita** o **descomposable**, si existe una familia $\mathcal{F} \subseteq \mathcal{M}$ con las siguientes propiedades:*

(a) $\mathcal{F}$ es una familia disjunta y $X = \bigcup_{F\in\mathcal{F}} F$;

(b) $\nu(F) < +\infty$ para todo $F \in \mathcal{F}$;

(c) si $E \in \mathcal{M}$ y $\nu(E) < +\infty$, entonces

$$\nu(E) = \sum_{F\in\mathcal{F}} \nu(E \cap F); \tag{1}$$

(d) si $E \subseteq X$ y $E \cap F \in \mathcal{M}$ para todo $F \in \mathcal{F}$, entonces $E \in \mathcal{M}$.

A la familia $\mathcal{F}$ le llamaremos una **super-descomposición** de $(X, \mathcal{M}, \nu)$. Puesto que $\mathcal{F}$ puede ser una familia no-numerable, la serie que aparece en (1) debe interpretarse como una serie sumable en el sentido de la Definición 2.1.52. Es fácil ver que toda medida ν super σ-finita es σ-finita. Más aun,

Teorema 12.1.33 (Radon-Nikodým). *Sea $(X, \mathcal{M}, \nu)$ un espacio de **medida super σ-finita** con una super-descomposición $\mathcal{F}$. Si λ es una **medida con signo σ-finita** definida sobre $(X, \mathcal{M})$ tal que $\lambda \ll \nu$, entonces existe una única función medible $f : X \to [0, +\infty]$ tal que*

$$\lambda(E) = \int_E f\, d\nu$$

para todo los $E \in \mathcal{M}$ que son σ-finitos respecto a ν.

Una demostración de este resultado se puede mirar, por ejemplo, en [72], Theorem 19.27, p. 318-319.

12.2. Aplicaciones del Teorema de Radon-Nikodým

Esta sección cubre algunas aplicaciones interesantes del Teorema de Radon-Nikodým.

12.2.1. Una Identidad en $L_1(X, \mathcal{M}, \nu)$

Recordemos que si $\mu(X) < +\infty$ y $f, g \in L_1(X, \mu)$, siendo una de esas funciones acotada, entonces $f \cdot g \in L_1(X, \mu)$. En general, si $f, g \in L_1(X, \mathcal{M}, \nu)$, entonces no siempre ocurre que $f \cdot g \in L_1(X, \mathcal{M}, \nu)$. La siguiente aplicación del Teorema de Radon-Nikodým establece que, bajo ciertas condiciones, el producto $f \cdot g \in L_1(X, \mathcal{M}, \nu)$ para una cierta función f.

Teorema 12.2.1. *Sean ν y λ **medida positivas** y **finitas** definidas sobre $(X, \mathcal{M})$ tal que $\lambda \ll \nu$. Si $f = d\lambda/d\nu$ es la **derivada de Radon-Nikodým** de λ con respecto a ν, entonces para cualquier función $g \in L_1(X, \mathcal{M}, \lambda)$, se cumple que $g \cdot f \in L_1(X, \mathcal{M}, \nu)$ y*

$$\int_X g\, d\lambda = \int_X g \cdot f\, d\nu. \tag{$**$}$$

Prueba. Para comenzar, suponga que $g = \chi_E$ para algún conjunto medible E. Entonces $(**)$ no es otra cosa que la igualdad $(*)$ del Teorema 12.1.28. Por supuesto, si g es una función simple no-negativa, entonces la linealidad de la integral nos conduce, una vez más, a validez de la igualdad $(**)$ para dicha función. Por otro lado, si g es no-negativa, el Teorema de Aproximación, Teorema 11.1.33, página 711, nos garantiza la existencia de una sucesión creciente $(\varphi_n)_{n=1}^{\infty}$ de funciones simples no-negativas convergiendo puntualmente a g y, entonces, el Teorema de la Convergencia Monótona es el encargado de verificar dicha igualdad. Finalmente, como $g = g^+ - g^-$, el argumento anterior, más la linealidad de la integral, nos muestra que $(**)$ también es válida en este caso. Fin de la prueba. ■

12.2.2. El Teorema de Descomposición de Lebesgue

Otra forma de demostrar el Teorema de Descomposición de Lebesgue es a través del Teorema de Radon-Nikodym. En efecto, en primer lugar, considere la medida positiva y finita $\mu_0 = \lambda + \nu$ y observe que como $\lambda \leq \mu_0$ y $\nu \leq \mu_0$, entonces $\lambda \ll \mu_0$ y $\nu \ll \mu_0$. Invocando el Teorema de Radon-Nikodým, podemos hallar un par de funciones medibles no-negativas f y g tales que

$$\lambda(E) = \int_E f\, d\mu_0 \qquad \text{y} \qquad \nu(E) = \int_E g\, d\mu_0$$

para todo $E \in \mathcal{M}$. Hagamos

$$P = \{x \in X : g(x) > 0\} \qquad \text{y} \qquad Q = \{x \in X : g(x) = 0\}.$$

Entonces $X = P \cup Q$, $P \cap Q = \varnothing$ y $\nu(Q) = 0$. Definiendo, para todo $E \in \mathcal{M}$, las medidas positivas

$$\lambda_a(E) = \lambda(E \cap P) \qquad \text{y} \qquad \lambda_s(E) = \lambda(E \cap Q)$$

vemos que $\lambda_s(P) = \nu(Q) = 0$, es decir, $\lambda_s \perp \nu$. Claramente, $\lambda = \lambda_a + \lambda_s$ de modo que sólo falta verificar que $\lambda_a \ll \nu$. Para demostrar esto último, sea $E \in \mathcal{M}$ tal que $\nu(E) = 0$. Entonces

$\int_E g\, d\mu_0 = 0$ y como $g \geq 0$, resulta que $g = 0$ μ_0-casi-siempre. Más aun, el hecho de que $g > 0$ sobre $E \cap P$, obliga a que $\mu_0(E \cap P) = 0$ y, por lo tanto,

$$\lambda_a(E) \;=\; \lambda(E \cap P) \;=\; \int_{E \cap P} f\, d\mu_0 \;=\; 0.$$

Así, $\lambda_a \ll \nu$ y termina la prueba. ■

12.2.3. El Teorema de Representación de Riesz - El Dual de $\mathrm{L}_p(X, \nu)$, $1 \leq p < +\infty$

Recordemos que si $(X, \|\cdot\|_X)$ es un espacio normado, entonces su **dual**, el cual denotaremos por X^*, consiste de todos los **funcionales lineales continuos** $x^* : X \to \mathbb{R}$. En el caso especial de los espacios de Banach $(\mathrm{L}_p(X, \nu), \|\cdot\|_p)$ con $p \in [1, +\infty)$, convenimos en escribir, para cualquier $f \in \mathrm{L}_p(X, \nu)$ y $\varphi \in \mathrm{L}_p(X, \nu)^*$ sus normas como:

$$\|\, f\,\|_{\mathrm{L}_p} \qquad\qquad \text{y} \qquad\qquad \|\, \varphi\,\|_{\mathrm{L}_p^*}\,.$$

Comencemos con el siguiente resultado.

Teorema 12.2.2. *Sea $(X, \mathcal{M}, \nu)$ un **espacio de medida σ-finita** y fijemos un número $p \in (1, +\infty)$. Si $g \in \mathrm{L}_q(X, \nu)$, donde q es el exponente conjugado de p, entonces el funcional lineal $\varphi_g : \mathrm{L}_p(X, \nu) \to \mathbb{R}$ dado por*

$$\varphi_g(f) \;=\; \int_X f \cdot g\, d\nu \quad \textit{para todo } f \in \mathrm{L}_p(X, \nu) \tag{1}$$

*es **continuo**, es decir, $\varphi_g \in \mathrm{L}_p(X, \nu)^*$ y, además, $\|\, \varphi_g\,\|_{\mathrm{L}_p^*} = \|\, g\,\|_{\mathrm{L}_q}$.*

Prueba. Claramente φ_g es una aplicación lineal. Por la Desigualdad de Hölder, Teorema 10.3.4, página 631, tenemos que

$$|\varphi_g(f)| \;\leq\; \|\, f\,\|_{\mathrm{L}_p} \|\, g\,\|_{\mathrm{L}_q}$$

lo cual prueba que $\varphi_g \in \mathrm{L}_p(X, \nu)^*$ y

$$\|\, \varphi_g\,\|_{\mathrm{L}_p^*} \;\leq\; \|\, g\,\|_{\mathrm{L}_q}\,.$$

Para obtener la otra desigualdad recordemos que gracias al Teorema 10.3.5, página 632, existe una función $f \in \mathrm{L}_p(X, \nu)$ con $\|\, f\,\|_p \leq 1$ tal que

$$\|\, g\,\|_{\mathrm{L}_q} \;=\; \left| \int_X f \cdot g\, d\mu \right| \;=\; |\varphi_g(f)| \;\leq\; \|\, \varphi_g\,\|_{\mathrm{L}_p^*}\,.$$

La prueba es completa. ■

Nuestro siguiente objetivo es demostrar que cualquier funcional lineal continuo en $\mathrm{L}_p(X, \nu)^*$ es de la forma φ_g para alguna función $g \in \mathrm{L}_q(X, \nu)$. Esto, por supuesto, nos revela que los espacios de Banach $\mathrm{L}_p(X, \nu)^*$ y $\mathrm{L}_q(X, \nu)$ son, desde el punto de visto algebraico y métrico, indistinguibles. Ese resultado, el cual se le conoce como el Teorema de Representación de Riesz para los espacios L_p es válido para cualquier espacio de medida $(X, \mathcal{M}, \nu)$, donde μ puede ser finita, σ-finita o no σ-finita siempre que $1 < p < +\infty$. Sin embargo, si queremos incluir, además, el caso $p = 1$, es decir, si $1 \leq p < +\infty$, debemos exigir que ν sea σ-finita. Existen varias formas de demostrar el Teorema de Representación de Riesz. En estas notas hemos elegido usar el Teorema de Radon-Nikodym y suponer que el espacio de medida $(X, \mathcal{M}, \nu)$ es σ-finito. Para aligerar un poco la notación, escribiremos $\mathrm{L}_p(X)$ en lugar de $\mathrm{L}_p(X, \nu)$.

Teorema 12.2.3 (Representación de Riesz). *Sea $(X, \mathcal{M}, \nu)$ un* ***espacio de medida σ-finita****. Para cada $1 \leq p < +\infty$ se cumple que*

$$\mathrm{L}_p(X)^* = \mathrm{L}_q(X)$$

donde q es el exponente conjugado de p.

Prueba. Fijemos un $p \in [0, +\infty)$ y sea q el exponente conjugado de p. Suponga, en primer lugar, que $\nu(X) < +\infty$. Esta suposición nos indica que *todas las funciones simples están en* $\mathrm{L}_p(X)$. Considere la aplicación $T : \mathrm{L}_q(X) \to \mathrm{L}_p(X)^*$ dada por $T(g) = \varphi_g$, donde

$$\varphi_g(f) = \int_X f \cdot g \, d\nu, \quad \text{para toda } f \in \mathrm{L}_p(X).$$

Es fácil ver que T es lineal, mientras que el Teorema 12.2.2 nos muestra T es una isometría. La prueba finalizará una vez que logremos demostrar que dicha aplicación es sobreyectiva. Para este fin, sea $\varphi \in \mathrm{L}_p(X)^*$ y defina la función de conjuntos $\lambda : \mathcal{M} \to \mathbb{R}$ por

$$\lambda(E) = \varphi(\chi_E) \quad \text{para todo } E \in \mathcal{M}.$$

Esta función está bien definida ya que $\chi_E \in \mathrm{L}_p(X)$ para todo $E \in \mathcal{M}$. Observe que

$$|\lambda(E)| \leq \|\varphi\|_{\mathrm{L}_p^*} \|\chi_E\|_{\mathrm{L}_p} = c\big(\nu(E)\big)^{1/p}, \tag{Ξ}$$

donde hemos puesto $c = \|\varphi\|_{\mathrm{L}_p^*}$. Veamos que λ es una medida real. En efecto, sean $E, F \in \mathcal{M}$ disjuntos. Puesto que $\chi_{E \cup F} = \chi_E + \chi_F$ se tiene que $\lambda(E \cup F) = \lambda(E) + \lambda(F)$ y entonces la linealidad de φ nos muestra que λ es finitamente aditiva. Considere ahora $(E_n)_{n=1}^{\infty}$ una sucesión disjunta en $\mathcal{M}$ y sea $E = \bigcup_{n=1}^{\infty} E_n$. Con $A_k = \bigcup_{n=1}^{k} E_n$ se observa que

$$\left\| \chi_E - \chi_{A_k} \right\|_{\mathrm{L}_p}^p = \int_{E \setminus A_k} d\nu = \nu(E \setminus A_k).$$

y entonces el Teorema de la Convergencia Dominada nos muestra que $\lim_{k\to\infty} \nu(E \setminus A_k) = 0$. Finalmente, puesto que φ es una función continua sobre $\mathrm{L}_p(X)$, resulta que $\lim_{k\to\infty} \lambda(A_k) = \lambda(E)$, es decir,

$$\lambda(E) = \sum_{n=1}^{\infty} \lambda(E_n).$$

Esto nos dice que λ es una medida real. Por otro lado, la desigualdad (Ξ) nos revela que $\lambda \ll \nu$ de modo que podemos invocar el Teorema de Radon-Nikodým para obtener una única función $g \in \mathrm{L}_1(X)$ tal que

$$\varphi(\chi_E) = \lambda(E) = \int_E g \, d\nu = \int_X g \cdot \chi_E \, d\nu$$

para todo $E \in \mathcal{M}$. Observe que si f es una función simple, digamos $f = \sum_{i=1}^{m} a_i \chi_{E_i}$, donde los conjuntos $E_1, \ldots, E_m$ son medibles y disjuntos dos a dos, entonces la linealidad de φ, así como la de integral, nos conducen a la igualdad

$$\varphi(f) = \int_X f \cdot g \, d\nu. \tag{2}$$

Puesto que estamos asumiendo que $\nu(X) < +\infty$, se tiene que las funciones simples son densas en $\mathrm{L}_p(X)$, Teorema 10.3.23, página 645, de donde resulta que la igualdad (2) es válida para toda $f \in \mathrm{L}_p(X)$. Falta verificar que

$$g \in \mathrm{L}_q(X) \qquad \text{y} \qquad \|\varphi\|_{\mathrm{L}_p^*} = \|g\|_{\mathrm{L}_q}. \tag{3}$$

(a) Suponga que $p = 1$. Para cada $n \in \mathbb{N}$, considere el conjunto $E_n = \{x \in X : g(x) > n\}$. Entonces

$$n \cdot \nu(E_n) \leq \int_X \chi_{E_n} g \, d\nu = \varphi(\chi_{E_n}) \leq \|\varphi\|_{\mathrm{L}_1^*} \left\| \chi_{E_n} \right\|_{\mathrm{L}_1} = \|\varphi\|_{\mathrm{L}_1^*} \nu(E_n).$$

de donde se sigue que $\nu(E_n) = 0$ si $n > \|\varphi\|_{\mathrm{L}_1^*}$. De esto se concluye que $|g^+(x)| \leq \|\varphi\|_{\mathrm{L}_1^*}$ ν-casi-siempre y, así

$$\|g^+\|_{\mathrm{L}_\infty} \leq \|\varphi\|_{\mathrm{L}_1^*}.$$

De modo similar se verifica que $\|g^-\|_{\mathrm{L}_\infty} \leq \|\varphi\|_{\mathrm{L}_1^*}$ y, por lo tanto, $\|g\|_{\mathrm{L}_\infty} \leq \|\varphi\|_{\mathrm{L}_1^*}$. Esto prueba que $g \in \mathrm{L}_\infty(X)$. El hecho de $\|g\|_{\mathrm{L}_\infty} \geq \|\varphi\|_{\mathrm{L}_1^*}$ se obtiene del Teorema 10.3.38, página 655, con lo cual queda establecido (3) cuando $p = 1$.

(b) Suponga ahora que $1 < p < +\infty$ y considere la función $h : X \to \mathbb{R}$ dada por

$$h(x) = \begin{cases} \dfrac{|g(x)|}{g(x)} & \text{si } g(x) \neq 0 \\ 0 & \text{si } g(x) = 0. \end{cases}$$

Defina, para cada $n \in \mathbb{N}$, la función

$$f_n = \chi_{E_n} \cdot |g|^{q-1} h, \qquad \text{donde} \quad E_n = \{x \in X : |g(x)| \leq n\}.$$

Observe que $|h| = 1$ sobre el conjunto $E = \{x \in X : g(x) \leq 0\}$ y como $p(q-1) = q$ resulta que

$$|f_n|^p = \chi_{E_n} \cdot |g|^q \leq n^q.$$

Esta desigualdad nos revela que cada $f_n \in \mathrm{L}_p(X)$ y se sigue de (2) que

$$\varphi(f_n) = \int_X \left(\chi_{E_n} \cdot |g|^{q-1} h\right) g \, d\nu = \int_{E_n} |g|^q \, d\nu$$

y entonces

$$\int_{E_n} |g|^q \, d\nu \leq \|\varphi\|_{\mathrm{L}_p^*} \|f_n\|_{\mathrm{L}_p} = \|\varphi\|_{\mathrm{L}_p^*} \left(\int_{E_n} |g|^q \, d\nu \right)^{1/p},$$

lo cual implica que $g \in \mathrm{L}_q(X)$ y, además,

$$\|g\|_q = \left(\int_{E_n} |g|^q \, d\nu \right)^{1-\frac{1}{p}} \leq \|\varphi\|_{\mathrm{L}_p^*}.$$

Por una nueva aplicación del Teorema 10.3.5, página 632, se obtiene la desigualdad contraria y termina la prueba para el caso cuando ν es finita.

Asuma ahora que ν es σ-finita. Este caso se deduce del anterior pero no es trivial. Hay que trabajar un poco. Fijemos un funcional lineal $\varphi \in \mathrm{L}_p(X)^*$. Como antes, la idea es construir una única función $g \in \mathrm{L}_q(X)$ de modo que $\varphi = \varphi_g$, es decir, que

$$\varphi(f) \;=\; \int_X f \cdot g \, d\nu, \quad \text{para toda } f \in \mathrm{L}_p(X).$$

Pues bien, como ν es σ-finita, existe una sucesión creciente $\left(X_n\right)_{n=1}^{\infty}$ de conjuntos medibles tal que

$$X \;=\; \bigcup_{n=1}^{\infty} X_n \qquad \text{y} \qquad \nu(X_n) < +\infty \quad \text{para todo } n \in \mathbb{N}.$$

Por cada $n \in \mathbb{N}$, considere la medida $\nu_n : \mathcal{M} \to [0, +\infty)$ definida por

$$\nu_n(E) \;=\; \nu(E \cap X_n) \quad \text{para todo } E \in \mathcal{M}.$$

Entonces cada ν_n es una medida finita sobre X. Defina $\varphi_n : \mathrm{L}_p(X, \nu_n) \to \mathbb{R}$ por

$$\varphi_n(f) \;=\; \varphi(f \cdot \chi_{X_n}) \quad \text{para toda } f \in \mathrm{L}_p(X, \nu_n).$$

Observe que $f \in \mathrm{L}_p(X, \nu_n)$ si, y sólo si, $f \cdot \chi_{X_n} \in \mathrm{L}_p(X, \nu_n)$.

(1^{o}) *φ_n es un funcional lineal acotado sobre* $\mathrm{L}_p(X, \nu_n)$ con $\| \varphi_n \| \leq \| \varphi \|$.
Verifiquemos esto. Fijemos $n \in \mathbb{N}$ y sean $f, g \in \mathrm{L}_p(X, \nu_n)$. Usemos la linealidad de φ para obtener

$$\varphi_n(f+g) \;=\; \varphi((f+g) \cdot \chi_{X_n}) \;=\; \varphi(f \cdot \chi_{X_n}) + \varphi(g \cdot \chi_{X_n}) \;=\; \varphi_n(f) + \varphi_n(g)$$

Similarmente, $\varphi_n(af) = a\varphi_n(f)$ para cualquier $a \in \mathbb{R}$. Para cada $f \in \mathrm{L}_p(X, \nu_n)$, denote su norma por $\| f \|_{p,n}$. Entonces

$$\begin{aligned} |\varphi_n(f)| \;=\; |\varphi(f \cdot \chi_{X_n})| \;&\leq\; \| \varphi \| \left\| f \cdot \chi_{X_n} \right\|_{p,n} \\ &=\; \| \varphi \| \left(\int_X \left| f \cdot \chi_{X_n} \right|^p d\nu \right)^{1/p} \\ &=\; \| \varphi \| \left(\int_X |f|^p \, d\nu_n \right)^{1/p} \;=\; \| \varphi \| \, \| f \|_{p,n} . \end{aligned}$$

Esto prueba que φ_n es lineal, acotada y satisface $\| \varphi_n \| \leq \| \varphi \|$.

(2^{o}) P*or el caso finito, existe* $g_n \in L_q(X, \nu_n)$ *tal que*

(a) $\varphi_n(f) = \displaystyle\int_X f g_n \, d\nu_n = \int_X f \cdot \chi_{X_n} \, d\nu$ *para cualquier* $f \in \mathrm{L}_p(X, \nu_n)$.

(b) $\| \varphi_n \| = \| g_n \|_{q,\nu_n}$.

Como el valor de $\varphi_n(f)$ no depende sobre cómo se comporte g_n sobre X_n^c podemos asumir, sin perder generalidad, que:

(c) $g_n = 0$ sobre X_n^c.

(d) g_n *es única salvo un conjunto de* ν_n*-medida cero.*

(3^{o}) *Como* $X_1 \subseteq X_2 \subseteq \cdots$, *entonces* $\nu_1 \leq \nu_2 \leq \cdots$ *y, por lo tanto, para cada entero* $n \geq 1$

$$\int_X |f|^p \, d\nu_n \ \leq \ \int_X |f|^p \, d\nu_{n+1} \quad \text{para cada } f \in \mathrm{L}_p(X, \nu_{n+1}).$$

De allí que

$$\mathrm{L}_p(X, \nu_{n+1}) \subseteq \mathrm{L}_p(X, \nu_n) \quad \text{para todo } n \in \mathbb{N}.$$

Se sigue de la unicidad de las funciones $g_1, g_2, \ldots$ que podemos, por cada $n \in \mathbb{N}$, tomar $g_{n+1} = g_n$ sobre X_n.

(4^{o}) *Finalmente, definiendo* $g : X \to \mathbb{R}$ *por* $g(x) = g_n(x)$ *siempre que* $x \in X_n$, *resulta que*:

(a) $\left\| g \cdot \chi_{X_n} \right\|_q \leq \| \varphi \|$.

En efecto, por la definición de g y el hecho de que $g_n = 0$ sobre X_n^c, se tiene que $g \cdot \chi_{X_n} = g_n$ sobre X y, por lo tanto,

$$\begin{aligned}
\left\| g \cdot \chi_{X_n} \right\|_q \ = \ \| g_n \|_q \ &= \ \left(\int_X |g_n|^q \, d\nu \right)^{1/q} \\
&= \ \left(\int_X |g_n|^q \cdot \chi_{X_n} \, d\nu \right)^{1/q} \\
&= \ \left(\int_X |g_n|^q \, d\nu_n \right)^{1/q} \\
&= \ \| g_n \|_{q,n} \ = \ \| \varphi_n \| \ \leq \ \| \varphi \| \, .
\end{aligned}$$

(b) $g \in \mathrm{L}_q(X, \nu)$ y $\| g \|_q \leq \| \varphi \|$.

Es consecuencia del Teorema de la Convergencia Monótona:

$$\begin{aligned}
\int_X |g|^q \, d\nu \ &= \ \int_X \lim_{n \to \infty} \chi_{X_n} \cdot |g|^q \, d\nu \\
&= \ \lim_{n \to \infty} \int_X \chi_{X_n} \cdot |g|^q \, d\nu \\
&= \ \lim_{n \to \infty} \left\| \chi_{X_n} \cdot g \right\|_q^q \ \leq \ \| \varphi \|^q \, .
\end{aligned}$$

(c) $\varphi(f) = \displaystyle\int_X f g \, d\nu$ *para cualquier* $f \in \mathrm{L}_p(X, \nu)$.

Para ver esto, sea $f \in \mathrm{L}_p(X, \nu)$. Observe que, gracias a la Desigualdad de Hölder, se cumple que $|f \cdot \chi_{X_n} \cdot g| \leq |f||g| \in \mathrm{L}_1(X, \nu)$ y, entonces, por el Teorema de la Convergencia Dominada

$$\begin{aligned}
\int_X f g \, d\nu \ &= \ \int_X \lim_{n \to \infty} f \cdot \chi_{X_n} \cdot g \, d\nu \\
&= \ \lim_{n \to \infty} \int_X f \cdot \chi_{X_n} \cdot g \, d\nu \\
&= \ \lim_{n \to \infty} \varphi_n(f) \ = \ \lim_{n \to \infty} \varphi(f \cdot \chi_{X_n}).
\end{aligned}$$

Por otro lado, el hecho de que $f \in \mathrm{L}_p(X, \nu)$ combinado con el Teorema de la Convergencia Dominada nos conduce a

$$\left\| f - f \cdot \chi_{X_n} \right\|_p^p = \left\| f \cdot \chi_{X_n^c} \right\|_p^p = \int_X \chi_{X_n^c} |f|^p \, d\nu \to \int_X 0 \, d\nu = 0,$$

en otras palabras, $f \cdot \chi_{X_n} \to f$. Puesto que φ es una aplicación continua, se tiene que

$$\varphi(f) = \lim_{n\to\infty} \varphi(f \cdot \chi_{X_n}) = \int_X fg \, d\nu.$$

Esto termina la prueba. ■

Nota Adicional 12.2.4 El Teorema de Representación de Riesz dice, en particular, que el dual de $\mathrm{L}_1(X, \nu)$ es $\mathrm{L}_\infty(X, \nu)$, es decir,

$$\mathrm{L}_1(X)^* = \mathrm{L}_\infty(X),$$

siempre que $(X, \mathcal{M}, \nu)$ sea un espacio de medida σ-finita. Ello no es cierto, en general, si $(X, \mathcal{M}, \nu)$ es un espacio de medida arbitrario no σ-finito, véase el Problema 5, página 830, al final de éste capítulo. Por otro lado, es importante advertir que la aplicación T definida en la prueba del Teorema de Representación de Riesz no se aplica al caso cuando $p = +\infty$ y $q = 1$; pues, en esta circunstancia, T no sería sobreyectiva; en otras palabras, el dual de $\mathrm{L}_\infty(X)$ no es $\mathrm{L}_1(X)$. Por ejemplo, para verificar dicha afirmación en el caso en que $X = \mathbb{R}$, recordemos del Análisis Funcional que *si X es un espacio de Banach cuyo dual X^* es **separable**, entonces X es **separable***. Disponiendo de esta información es fácil concluir que el dual de $\mathrm{L}_\infty(X)$ no puede ser $\mathrm{L}_1(X)$. En efecto, si ello fuese verdad, entonces de la igualdad $\mathrm{L}_\infty(\mathbb{R}, \mu)^* = \mathrm{L}_1(\mathbb{R}, \mu)$ se concluiría que $\mathrm{L}_\infty(\mathbb{R}, \mu)$ sería, por la observación anterior, separable ya que $\mathrm{L}_1(\mathbb{R}, \mu)$ lo es; sin embargo, como ya demostramos, $\mathrm{L}_\infty(\mathbb{R}, \mu)$ no es separable. De hecho, $\mathrm{L}_\infty(\mathbb{R}, \mu)^*$ es inmensamente más grande que $\mathrm{L}_1(\mathbb{R}, \mu)$.

Por otro lado, si p y q son exponentes conjugados con $1 < p, q < +\infty$, el Teorema de Representación de Riesz nos dice que

$$\mathrm{L}_p(X)^* = \mathrm{L}_q(X) \qquad \text{y} \qquad \mathrm{L}_q(X)^* = \mathrm{L}_p(X) \tag{4}$$

de donde podemos concluir que $\mathrm{L}_p(X)^{**} = \mathrm{L}_p(X)$, donde A^{**} es, por definición, $(A^*)^*$ y se denomina el **segundo dual** de A. En general, un espacio de Banach X se dice que es **reflexivo** si la aplicación $J : X \to X^{**}$ definida por

$$J(x)(x^*) = x^*(x) \quad \text{para todo } x \in X \text{ y todo } x^* \in X^*$$

es una isometría lineal de X sobre X^{**}. Aunque los espacios $\mathrm{L}_p(X)$, con $1 < p < +\infty$, son reflexivos, la conclusión $\mathrm{L}_p(X)^{**} = \mathrm{L}_p(X)$ derivada de las dos igualdades dadas en (4) no constituye una demostración de ese hecho.

12.2.4. Existencia de la Esperanza Condicional

Uno de los ingredientes fundamentales en el estudio de las martingalas en la Teoría de Probabilidad es la noción de esperanza condicional que describiremos ahora. Fijemos un espacio de medida $(X, \mathcal{M}, \nu)$ tal que $\nu(X) = 1$. Los probabilistas gustan en llamar a un tal espacio de medida un **espacio de probabilidad**. Similarmente, a cualquier función $\mathcal{M}$-medible $f : X \to \mathbb{R}^*$ ellos la denominan una **variable aleatoria**. Si $\mathcal{M}_0$ es una σ-álgebra tal que $\mathcal{M}_0 \subseteq \mathcal{M}$, entonces diremos que $\mathcal{M}_0$ es una **sub-σ-álgebra** de $\mathcal{M}$.

Definición 12.2.4. *Sean $(X, \mathcal{M}, \nu)$ un **espacio de probabilidad**, $f \in L_1(X, \mathcal{M}, \nu)$ y $\mathcal{M}_0$ una sub-σ-álgebra de $\mathcal{M}$. La **esperanza condicional de f dado** $\mathcal{M}_0$ es cualquier función $g : X \to \mathbb{R}$ que satisfaga las siguientes propiedades:*

(a) g es $\mathcal{M}_0$-medible, y

(b) para cualquier conjunto $E \in \mathcal{M}_0$,

$$\int_E g\,d\nu \;=\; \int_E f\,d\nu. \tag{1}$$

Observe que (b), en la definición anterior, implica que g es ν-integrable respecto a $\mathcal{M}_0$; es decir, $g \in L_1(X, \mathcal{M}_0, \nu)$. En efecto, pongamos $E = \{x \in X : g(x) > 0\} \in \mathcal{M}_0$ y usemos dos veces la condición (b) para obtener:

$$\int_E g\,d\nu \;=\; \int_E f\,d\nu \;\leq\; \int_E |f|\,d\nu$$

$$\int_{E^c} (-g)\,d\nu \;=\; \int_{E^c} (-f)\,d\nu \;\leq\; \int_{E^c} |f|\,d\nu$$

De esto se sigue que

$$\int_X |g|\,d\nu \;\leq\; \int_X |f|\,d\nu \;<\; +\infty.$$

En lo que sigue, denotaremos a g por $\mathbb{E}^{\mathcal{M}_0}(f)$. Con esta notación, la ecuación (1) se expresa en la forma:

$$\int_E \mathbb{E}^{\mathcal{M}_0}(f)\,d\nu \;=\; \int_E f\,d\nu, \quad \text{para todo } E \in \mathcal{M}_0.$$

Teorema 12.2.5 (Esperanza Condicional). *Sean $(X, \mathcal{M}, \nu)$ un **espacio de probabilidad**, $\mathcal{M}_0$ una **sub-σ-álgebra** de $\mathcal{M}$ y $f \in L_1(X, \mathcal{M}, \nu)$. Entonces $\mathbb{E}^{\mathcal{M}_0}(f)$ existe y es única.*

Prueba. Suponga, en primer lugar, que $f \geq 0$ y defina $\lambda : \mathcal{M}_0 \to [0, +\infty)$ por

$$\lambda(E) \;=\; \int_E f\,d\nu \quad \text{para todo } E \in \mathcal{M}_0.$$

Claramente λ es una medida positiva y finita sobre $\mathcal{M}_0$ y se cumple que $\lambda \ll \nu$. Por el Teorema de Radon-Nikodým, existe una única función no-negativa $f_0 : X \to [0, +\infty)$ la cual es $\mathcal{M}_0$-medible, pertenece a $L_1(X, \mathcal{M}_0, \nu)$ y cumple que

$$\int_E f_0\,d\nu \;=\; \int_E f\,d\nu$$

para todo $E \in \mathcal{M}_0$. Para el caso general, escriba a f como $f = f^+ - f^-$. Por el argumento anterior, existen funciones no-negativas $f_1, f_2 : X \to \mathbb{R}$ la cuales son $\mathcal{M}_0$-medibles, ambas pertenecen a $L_1(X, \mathcal{M}_0, \nu)$ y satisfacen

$$\int_E f_1\,d\nu \;=\; \int_E f^+\,d\nu \qquad \text{y} \qquad \int_E f_2\,d\nu \;=\; \int_E f^-\,d\nu$$

para todo $E \in \mathcal{M}_0$. Si ahora definimos $g = f_1 - f_2$, resulta que g es $\mathcal{M}_0$-medible, pertenece al espacio $\mathrm{L}_1(X, \mathcal{M}_0, \nu)$ y, además,

$$\int_E g\,d\nu \;=\; \int_E f\,d\nu$$

para todo $E \in \mathcal{M}_0$. La prueba es completa. ■

Algunas de las buenas propiedades que posee la esperanza condicional se muestra a continuación. Siguiendo el lenguaje probabilístico, escribiremos, en lo que resta de esta sección:

$$\mathbb{E}(f) \;=\; \int_X f\,d\nu,$$

para cualquier función $f \in \mathrm{L}_1(X, \mathcal{M}, \nu)$. En este caso,

$$|\mathbb{E}(f)| \;\le\; \mathbb{E}(|f|).$$

Teorema 12.2.6. *Sean $(X, \mathcal{M}, \nu)$ un **espacio de probabilidad**, $\mathcal{M}_0$ una sub-σ-álgebra de $\mathcal{M}$ y suponga que $f, g \in \mathrm{L}_1(X, \mathcal{M}, \nu)$. Entonces:*

(a) $\mathbb{E}\big(\mathbb{E}^{\mathcal{M}_0}(f)\big) = \mathbb{E}(f)$.

(b) $\mathbb{E}(af + bg;\, \mathcal{M}_0) \;=\; a \cdot \mathbb{E}^{\mathcal{M}_0}(f) \,+\, b \cdot \mathbb{E}^{\mathcal{M}_0}(g)$ ν-casi-siempre, para cualesquiera $a, b \in \mathbb{R}$.

(c) Si $f \in \mathrm{L}_1(X, \mathcal{M}_0, \nu)$, entonces $\mathbb{E}^{\mathcal{M}_0}(f) = f$ ν-casi-siempre.

(d) Si $f \ge 0$, entonces $\mathbb{E}^{\mathcal{M}_0}(f) \ge 0$.

Prueba. Se deja a cargo del lector. ■

La esperanza condicional posee, además de las propiedades mencionadas en el resultado anterior, otra que es bastante interesante.

Teorema 12.2.7. *Sea $(X, \mathcal{M}, \nu)$ un **espacio de probabilidad** y suponga que $\mathfrak{G}$ es una familia arbitraria de **sub-σ-álgebras** de $\mathcal{M}$. Para cada $f \in \mathrm{L}_1(X, \mathcal{M}, \nu)$, la familia*

$$\mathcal{U} \;=\; \big\{\mathbb{E}^{\mathcal{B}}(f) \,:\, \mathcal{B} \in \mathfrak{G}\big\}$$

*es **uniformemente integrable**.*

Prueba. Puesto que $|\mathbb{E}^{\mathcal{B}}(f)| \le \mathbb{E}^{\mathcal{B}}(|f|)$ para cualquier $\mathcal{B} \in \mathfrak{G}$, resulta entonces que

$$\int_{[\mathbb{E}^{\mathcal{B}}(f) > c]} |\mathbb{E}^{\mathcal{B}}(f)|\,d\nu \;\le\; \int_{[\mathbb{E}^{\mathcal{B}}(|f|) > c]} \mathbb{E}^{\mathcal{B}}(|f|)\,d\nu \;=\; \int_{[\mathbb{E}^{\mathcal{B}}(|f|) > c]} |f|\,d\nu$$

para todo $c > 0$. Fijemos un número real $b > 0$ y observe que como $X = [|f| > b] \cup [|f| \le b]$, entonces

$$\int_{[\mathbb{E}^{\mathcal{B}}(|f|) > c]} |f|\,d\nu \;\le\; b \cdot \nu\big([\,\mathbb{E}^{\mathcal{B}}(|f|) > c\,]\big) \,+\, \int_{[\,|f| > b\,]} |f|\,d\nu.$$

Por otro lado, por la Desigualdad de Chevyshev,

$$\nu\big([\,\mathbb{E}^{\mathcal{B}}(|f|) > c\,]\big) \;\le\; c^{-1} \cdot \mathbb{E}(\mathbb{E}^{\mathcal{B}}(|f|)) \;=\; c^{-1} \cdot \mathbb{E}(|f|)$$

de donde se obtiene

$$\sup_{\mathcal{B}\in\mathfrak{A}} \int_{[\mathbb{E}^{\mathcal{B}}(|f|)>c]} |f|\,d\nu \;\leq\; \frac{b}{c}\cdot\mathbb{E}(|f|) \;+\; \int_{[|f|>b]} |f|\,d\nu.$$

Si ahora elegimos $b = \sqrt{c}$ y hacemos que $c \to \infty$ se obtiene el resultado. ■

Intuitivamente, podríamos pensar la esperanza condicional $\mathbb{E}^{\mathcal{M}_0}(f)$ como la "predicción óptima de f, dada la información $\mathcal{M}_0$". Esta interpretación de la esperanza condicional permite, de modo natural, la idea de martingala. Pero, ¿qué es una martingala? En términos sencillos, una martingala no es otra cosa que *un modelo matemático para una apuesta justa*, es decir, que las ganancias esperadas deberían ser cero.

Definición 12.2.8. *Sea $(X, \mathcal{M}, \nu)$ un espacio de probabilidad. Una **filtración** sobre $(X, \mathcal{M}, \nu)$ es una sucesión creciente $(\mathcal{M}_n)_{n=1}^{\infty}$ de sub-σ-algebras de $\mathcal{M}$, es decir,*

$$\mathcal{M}_1 \;\subseteq\; \mathcal{M}_2 \;\subseteq\; \cdots \;\subseteq \mathcal{M}_n \;\subseteq \cdots$$

*Una sucesión $(f_n)_{n=1}^{\infty}$ de funciones $\mathcal{M}$-medibles se dice **adaptada** a la filtración $(\mathcal{M}_n)_{n=1}^{\infty}$ si, para todo $n \in \mathbb{N}$, f_n es $\mathcal{M}_n$-medible.*

Definición 12.2.9. *Sean $(X, \mathcal{M}, \nu)$ un espacio de probabilidad, $(\mathcal{M}_n)_{n=1}^{\infty}$ una filtración sobre $(X, \mathcal{M}, \nu)$ y $(f_n)_{n=1}^{\infty}$ una sucesión adaptada a $(\mathcal{M}_n)_{n=1}^{\infty}$. Diremos que $(f_n, \mathcal{M}_n)_{n=1}^{\infty}$ es una **martingala** si, para cada $n \geq 1$:*

(a) $f_n \in L_1(X, \mathcal{M}_n, \nu)$ *y*

(b) $\mathbb{E}^{\mathcal{M}_n}(f_{n+1}) = f_n$.

Nótese que si $(f_n, \mathcal{M}_n)_{n=1}^{\infty}$ es una martingala, entonces para todo $m, n \in \mathbb{N}$ con $m < n$, se cumple que

$$\mathbb{E}(f_n) \;=\; \mathbb{E}\big(\mathbb{E}^{\mathcal{M}_m}(f_n)\big) \;=\; \mathbb{E}(f_m).$$

Esto significa que una martingala tiene esperanza constante. Uno de los resultados fundamentales sobre martingalas es el siguiente.

Teorema 12.2.10. *Sea $(X, \mathcal{M}, \nu)$ un espacio de probabilidad. Si $(f_n, \mathcal{M}_n)_{n=1}^{\infty}$ es una **martingala uniformemente integrable**, entonces ella **converge** ν-**casi-siempre** y en la **norma** de $L_1(X, \mathcal{M}, \nu)$.*

12.2.5. Unicidad de la Medida de Lebesgue

En el Teorema 11.3.14, página 743 dimos una demostración de la unicidad de la medida de Lebesgue. En esta sección veremos una nueva demostración usando el Teorema de Radon-Nikodym.

Teorema 12.2.11 (Unicidad de la Medida de Lebesgue). *Sea $\lambda : \mathfrak{Bo}(\mathbb{R}^n) \to [0, +\infty]$ una medida σ-**finita**, **invariante por traslación** tal que $\lambda(K) < +\infty$ para cualquier compacto $K \subseteq \mathbb{R}^n$. Entonces existe una constante $c \geq 0$ tal que*

$$\lambda(E) \;=\; c\cdot\mu_n(E), \tag{3}$$

para cualquier $E \in \mathfrak{Bo}(\mathbb{R}^n)$.

Prueba. Nuestra primera tarea es verificar que $\lambda \ll \mu_n$. Sea $E \in \mathfrak{B}o(\mathbb{R}^n)$ con $\mu_n(E) = 0$. Puesto que la aplicación $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ dada por $f(x,y) = x + y$ es continua, en particular, medible Borel, resulta que el conjunto

$$\widetilde{E} = f^{-1}(E) = \big\{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n : x + y \in E\big\}$$

es Borel medible. Observe que la x-sección de $\widetilde{E}$ es el conjunto

$$\widetilde{E}_x = \big\{y \in \mathbb{R}^n : (x,y) \in \widetilde{E}\big\} = E - x.$$

Similarmente, su y-sección es $\widetilde{E}^y = E - y$. Teniendo en cuenta que λ y μ_n son ambas σ-finitas, podemos aplicar el Teorema de Fubini a $(\mu_n \times \lambda)(\widetilde{E})$. En efecto, como $\mu_n(E) = 0$, entonces $(\mu_n \times \lambda)(\widetilde{E}) = 0$. De aquí se sigue que $\lambda(E_x) = \lambda(E - x) = 0$ para μ_n-casi todo x. El hecho de λ es invariante por traslación conduce a que $\lambda(E) = 0$ y, por lo tanto, $\lambda \ll \mu_n$. Un llamado al Teorema de Radon-Nikodym nos garantiza la existencia de una única función μ_n-integrable $f : \mathbb{R}^n \to [0, +\infty]$ tal que

$$\lambda(E) = \int_E f \, d\mu_n \quad \text{para cualquier } E \in \mathfrak{B}o(\mathbb{R}^n).$$

En particular, como la integral de Lebesgue es invariante por traslación

$$\lambda(E) = \lambda(E + t) = \int_{E+t} f(x) \, d\mu_n(x) = \int_E f(x - t) \, d\mu_n(x)$$

para cualquier t y cualquier boreliano $E \subseteq \mathbb{R}^n$. La unicidad de f implica entonces que para todo t y μ_n-casi todo x, la ecuación

$$f(x - t) = f(x) \tag{1}$$

se cumple. Para completar la prueba, sólo debemos verificar la existencia de una constante $c \geq 0$ tal que $f(x) = c$ para μ_n-casi todo x. Suponga, para generar una contradicción, que existen números reales $r_1 < r_2$ tales que los conjuntos borelianos

$$E_1 = \big\{x \in \mathbb{R} : f(x) < r_1\big\} \qquad \text{y} \qquad E_2 = \big\{x \in \mathbb{R} : f(x) > r_2\big\}$$

son de medida de Lebesgue positiva. Sean d_1 y d_2 dos puntos de densidad de Lebesgue de E_1 y E_2 respectivamente y sea $t = d_2 - d_1$. Puesto que d_2 también es un punto de densidad del conjunto $\big\{x \in \mathbb{R} : f(x - t) < r_1\big\}$, resulta que

$$\mu_n\big(\big\{x \in \mathbb{R} : f(x - t) < r_1\big\} \cap E_2\big) > 0$$

lo que, evidentemente, contradice la ecuación (1). ■

12.3. Diferenciación de Medidas

El Teorema Fundamental del Cálculo para la integral de Riemann posee dos partes: una de ellas establece que si $f : [a, b] \to \mathbb{R}$ es integrable según Riemann, entonces

$$f(x) = \frac{d}{dx} \int_a^x f(t) \, dt. \tag{1}$$

La otra parte dice que si f es derivable con f' acotada y Riemann-integrable, entonces podemos recobrar la función original integrando su derivada; esto es,

$$f(x) \;=\; f(a) + \int_a^x f'(t)\,dt. \tag{2}$$

El objetivo de esta sección es obtener una generalización de (1) en $\mathbb{R}^n$ para cierto tipo de funciones. Para motivar nuestro resultado, observe que si definimos

$$F(x) \;=\; \int_a^x f(t)\,dt,$$

entonces la igualdad dada en (1) se puede reescribir en la forma

$$f(x) \;=\; \lim_{h\to +} \frac{F(x+h)-F(x)}{h} \;=\; \lim_{h\to 0+} \frac{1}{h}\int_x^{x+h} f(t)\,dt.$$

Más aun, del Teorema 10.2.55, página 629, sabemos que si $f \in L_1([a,b])$, entonces existe un conjunto $N \subseteq [a,b]$ con $\mu(N)=0$ tal que

$$\lim_{h\to 0}\frac{1}{h}\int_x^{x+h} |f(t)-f(x)|\,d\mu \;=\; 0 \tag{3}$$

para todo $x \in [a,b]\setminus N$.

Una versión n-dimensional de la igualdad (3) debería ser expresada en la forma:

$$f(x) \;=\; \lim_{r\to 0+} \frac{1}{\mu_n(B(x,r))}\int_{B(x,r)} f(t)\,d\mu_n(t),$$

donde μ_n es la medida de Lebesgue sobre $\mathfrak{Bo}(\mathbb{R}^n)$, $B(x,r)$ es la bola abierta con centro en $x \in \mathbb{R}^n$ y radio $r>0$ y $f : \mathbb{R}^n \to \mathbb{R}$ es una función localmente integrable. ¿Qué significa esto último?

Definición 12.3.1. *Una función $f : \mathbb{R}^n \to \mathbb{R}$ se dice que es* ***localmente integrable*** *si ella es medible según Borel y para cada $x \in \mathbb{R}^n$, existe un entorno abierto U_x de x tal que*

$$\int_{U_x} |f|\,d\mu_n \;<\; +\infty.$$

El espacio de todas las funciones $f : \mathbb{R}^n \to \mathbb{R}$ que son localmente integrables será denotado por $L^1_{\mathrm{loc}}(\mathbb{R}^n)$.

Teorema 12.3.2. *Sea $f : \mathbb{R}^n \to \mathbb{R}$ una* ***función medible Borel****. Las siguientes condiciones son equivalentes:*

(1) $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$.

(2) $\int_K |f|\,d\mu_n \;<\; +\infty$ *para cada* ***conjunto compacto*** $K \subseteq \mathbb{R}^n$.

Prueba. (1) $\Rightarrow$ (2). Suponga que f es localmente acotada y sea $K \subseteq \mathbb{R}^n$ compacto. Por hipótesis, para cada $x \in K$, existe un entorno abierto U_x de x tal que $\int_{U_x} |f|\, d\mu_n < +\infty$. Puesto que la colección $\{U_x : x \in K\}$ es un cubrimiento abierto de K, resulta, por compacidad, que existen $x_1, \ldots, x_k$ en K tal que $K \subseteq U_{x_1} \cup \ldots \cup U_{x_k}$. Esto implica que

$$\int_K |f|\, d\mu_n \;\leq\; \int_{U_{x_1}} |f|\, d\mu_n + \cdots + \int_{U_{x_k}} |f|\, d\mu_n \;<\; +\infty.$$

La implicación (2) $\Rightarrow$ (1) es inmediata. En efecto, sea $x \in \mathbb{R}^n$. Entonces $U_x = B(x,1)$ es un entorno abierto de x cuya clausura, $\overline{U}_x$, es compacta y así,

$$\int_{U_x} |f|\, d\mu_n \;\leq\; \int_{\overline{U}_x} |f|\, d\mu_n \;<\; +\infty.$$

La prueba es completa. ■

Nuestra generalización de (3) se expresa del modo siguiente:

Teorema 12.3.3 (Diferenciación de Lebesgue). *Si $f \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^n)$, entonces existe un conjunto medible $N \subseteq \mathbb{R}^n$ con $\mu_n(N) = 0$ tal que*

$$\lim_{r\to 0} \frac{1}{\mu_n(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\, d\mu_n(y) \;=\; 0 \tag{4}$$

para todo $x \notin N$.

Observe que si f es continua en $a \in \mathbb{R}^n$, entonces la igualdad (4) es válida. En efecto, dado $\varepsilon > 0$, escoja un $\delta > 0$ de modo tal que $|f(x) - f(a)| < \varepsilon$ siempre $\|x - a\| < \delta$. Si tomamos $B = B(a, \delta/2)$, resulta que

$$\begin{aligned} \left| \frac{1}{\mu_n(B)} \int_B f(x)\, d\mu_n(x) - f(a) \right| \;&=\; \frac{1}{\mu_n(B)} \left| \int_B (f(x) - f(a))\, d\mu_n \right| \\ &\leq\; \frac{1}{\mu_n(B)} \int_B |f(x) - f(a)|\, d\mu_n \;<\; \varepsilon. \end{aligned} \tag{5}$$

El lector podrá observar que la integral puede ser evaluada sobre bolas abiertas o cerradas indistintamente, ya que la esfera tiene medida cero. En efecto, como $\mu_n(B(0,1)) < +\infty$, no se pierde generalidad en asumir que $\mu_n(B(0,1)) = 1$. Puesto que $B(x,r) = x + rB(0,1)$, el Teorema 11.3.14, página 743, nos dice entonces

$$\mu_n(B(x,r)) \;=\; \mu_n(rB(0,1)) \;=\; r^n$$

y como $S(x,r) \subseteq B(x, r+\varepsilon) \setminus B(x, r-\varepsilon)$ para cada $\varepsilon > 0$, resulta que

$$\mu_n(S(x,r)) \;\leq\; (r+\varepsilon)^n - (r-\varepsilon)^n.$$

Haciendo que $\varepsilon \to 0$ se obtiene que $\mu_n(S(x,r)) = 0$. De esto se sigue que $\mu_n(B(x,r)) = \mu_n(\overline{B(x,r)})$.

12.3.1. La Función Maximal de Hardy-Littlewood

Para probar el Teorema de Diferenciación de Lebesgue, introduciremos un función especial, llamada la función maximal de Hardy-Littlewood, que constituye una pieza clave en su demostración.

Definición 12.3.4. *Si $f \in L^1_{\text{loc}}(\mathbb{R}^n)$, entonces la* ***función maximal de Hardy-Littlewood*** *de f, denotada por $\mathrm{M}f$, se define por*

$$\mathrm{M}f(x) \;=\; \sup_{r>0} \frac{1}{\mu_n(B(x,r))} \int_{B(x,r)} |f(y)|\, d\mu_n(y).$$

Las bolas abiertas $B(x,r)$ con centro en x usadas para definir $\mathrm{M}f$ se eligen sólo por conveniencia, es decir, podríamos haber tomado cualquier bola abierta con centro distinto de x pero conteniendo a x o bolas cerradas, o cualquier otro conjunto tal como cubos, etc. Sin embargo, algún tipo de restricción hay que imponerle a tales conjuntos: por ejemplo, no se pueden usar rectángulos arbitrarios ya que los promedios sobre rectángulos con "centro" en un punto x y cuyos lados se hacen cada vez más largos pero su volumen se contrae a cero no convergen a $f(x)$, aun en el caso en que f es continua.

Lema 12.3.5. *Si $f \in L^1_{\text{loc}}(\mathbb{R}^n)$, entonces $\mathrm{M}f$ es* ***inferiormente semicontinua****. En particular, f es* ***medible Borel****.*

Prueba. Observe que $\mathrm{M}f \geq 0$. Para demostrar que $\mathrm{M}f$ es inferiormente semicontinua tenemos que verificar que el conjunto

$$E_t \;=\; \{x \in \mathbb{R}^n : \mathrm{M}f(x) > t\}$$

es abierto en $\mathbb{R}^n$ para cualquier $0 < t < +\infty$. Fijemos $t > 0$ y sea $x \in E_t$. Entonces $\mathrm{M}f(x) > t$ y, por lo tanto, existe un $r > 0$ tal que

$$\frac{1}{\mu_n(B(x,r))} \int_{B(x,r)} |f(y)|\, d\mu_n(y) \;>\; t.$$

Si ahora tomamos cualquier $x' \in B(x,r)$ resulta que

$$\mathrm{M}f(x') \;\geq\; \frac{1}{\mu_n(B(x,r))} \int_{B(x,r)} |f(y)|\, d\mu_n(y) \;>\; t.$$

Esto prueba que $x' \in E_t$ y, por lo tanto, $B(x,r) \subseteq E_t$, es decir, E_t es abierto. ∎

Es importante destacar que si $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ con $f \neq 0$, entonces no es necesariamente cierto que $\mathrm{M}f \in L_1(\mathbb{R}^n)$. En efecto, sea $a > 0$ y suponga que $\|x\| \geq a$. Si ahora consideramos el promedio de $|f|$ en x sobre una bola abierta de radio $r = 2\|x\|$ y usando el hecho de que $B(0,a) \subseteq B(x, 2\|x\|)$, resulta que

$$\begin{aligned}
\mathrm{M}f(x) \;&\geq\; \frac{1}{\mu_n(B(x,2\|x\|))} \int_{B(x,2\|x\|)} |f(y)|\, d\mu_n(y) \\
&\geq\; \frac{C}{\|x\|^n} \int_{B(0,a)} |f(y)|\, d\mu_n(y).
\end{aligned}$$

para alguna constante $C > 0$. Puesto que la función $1/\parallel x \parallel^n$ no es integrable sobre $\mathbb{R}^n \setminus B(0,a)$, resulta que $\mathrm{M}f$ no puede ser integrable sobre $\mathbb{R}^n$. En efecto, si $\mathrm{M}f$ fuese μ_n-integrable, entonces

$$\int_{B(0,a)} |f(y)|\, d\mu_n(y) \;=\; 0$$

y como $a > 0$ fue elegido de modo arbitrario, tendríamos que $f = 0$ casi-siempre en $\mathbb{R}^n$, lo que negaría nuestra suposición.

12.3.2. El Teorema de Hardy-Littlewood

Aunque la función maximal de una función localmente integrable no tiene necesariamente que ser integrable, ni aun localmente integrable, existen ciertos tipos de funciones que son adecuadas para nuestro propósito.

Definición 12.3.6. *Una función medible $f : \mathbb{R}^n \to \mathbb{R}$ se dice que es **integrable en el sentido débil**, lo que escribiremos como $f \in \mathrm{L}_1^{\omega}(\mathbb{R}^n)$, si existe una constante $C > 0$, dependiendo sólo de n, tal que*

$$\mu_n\big(\{x \in \mathbb{R}^n \,:\, |f(x)| > t\}\big) \;\leq\; \frac{C}{t}$$

para todo $t \in (0, +\infty)$.

Observe que si $f \in \mathrm{L}_1^{\omega}(\mathbb{R}^n)$, entonces f es **finita casi-siempre**. En efecto, como

$$E_f^{\infty} \;=\; \big\{x \in \mathbb{R}^n \,:\, |f(x)| = +\infty\big\} \;=\; \bigcap_{k=1}^{\infty} \big\{x \in \mathbb{R}^n \,:\, |f(x)| > k\big\}$$

resulta que

$$\mu_n(E_f^{\infty}) \;\leq\; \mu_n\big(\{x \in \mathbb{R}^n \,:\, |f(x)| > k\}\big) \;\leq\; \frac{C}{k}$$

para todo $k \in \mathbb{N}$, de donde se sigue $\mu_n(E_f^{\infty}) = 0$. También es claro, usando la Desigualdad de Chebyshev, Teorema 11.1.39, página 714, que si $f \in \mathrm{L}_1(\mathbb{R}^n)$, entonces $f \in \mathrm{L}_1^{\omega}(\mathbb{R}^n)$, es decir, $\mathrm{L}_1(\mathbb{R}^n) \subseteq \mathrm{L}_1^{\omega}(\mathbb{R}^n)$.

El siguiente resultado, conocido como la desigualdad de Hardy-Littlewood, establece que la función maximal de Hardy-Littlewood de cualquier función μ_n-integrable es integrable en el sentido débil. La demostración de ésta desigualdad requiere del siguiente resultado sobre cubrimientos de Wiener que es una versión no tan sofisticada como lo son los cubrimientos de Vitali.

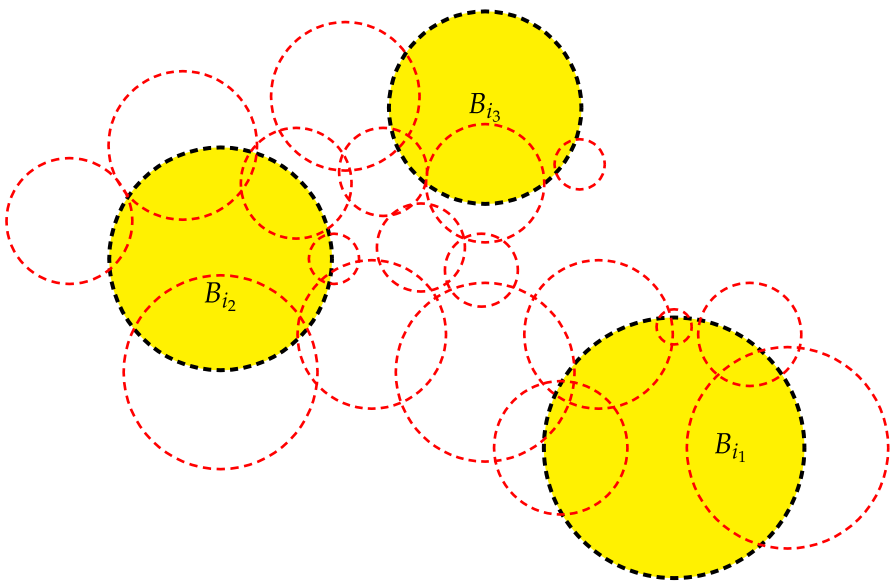
Lema 12.3.7 (Wiener). *Sea $\mathcal{C} = \{B_1, \ldots, B_N\}$ una **colección finita de bolas abiertas** en $\mathbb{R}^n$. Entonces existe una **subcolección disjunta** $\{B_{i_1}, \ldots, B_{i_m}\}$ de $\mathcal{C}$ tal que*

$$\mu_n\Big(\bigcup_{i=1}^{N} B_i\Big) \;\leq\; 3^n \sum_{j=1}^{m} \mu_n(B_{i_j}).$$

Prueba. Pongamos $B_i = B(x_i, r_i)$ para $i = 1, \ldots, N$ y sea $B_i^* = 3B_i = B(x_i, 3r_i)$. Entonces

$$\mu_n(B_i^*) \;=\; 3^n \mu_n(B_i)$$

para cada $i = 1, \ldots, N$. De la colección $\mathcal{C} = \{B_1, \ldots, B_N\}$ escojamos una bola, digamos B_{i_1}, que posee el mayor radio (por supuesto, puede suceder que exista más de una bola con el mismo radio maximal y entonces seleccione una de ellas).



Miremos ahora todas la bolas en $\mathcal{C}$, incluyendo a B_{i_1}, que intersectan a B_{i_1}, esto es, sea

$$\mathcal{D}_1 = \{B_i \in \mathcal{C} \,:\, B_i \cap B_{i_1} \neq \varnothing\}.$$

Puesto que el radio de cada una de las bolas en $\mathcal{D}_1$ es menor o igual que el radio de B_{i_1}, resulta que $\bigcup \mathcal{D}_1 \subseteq B_{i_1}^*$ y, por lo tanto,

$$\mu_n\Big(\bigcup \mathcal{D}_1\Big) \;\leq\; \mu_n(B_{i_1}^*) \;=\; 3^n \mu_n(B_{i_1}).$$

Las bolas de $\mathcal{C}$ que no están en $\mathcal{D}_1$ tienen intersección vacía con B_{i_1}. Denote por $\mathcal{C}_1$ la colección de tales bolas y de ellas escoja una bola B_{i_2} que posea el radio más grande. Como antes, sea

$$\mathcal{D}_2 = \{B_i \in \mathcal{C}_1 \,:\, B_i \cap B_{i_2} \neq \varnothing\}.$$

Nótese que $B_{i_1} \cap B_{i_2} = \varnothing$ y

$$\mu_n\Big(\bigcup \mathcal{D}_2\Big) \;\leq\; \mu_n(B_{i_2}^*) \;=\; 3^n \mu_n(B_{i_2}).$$

Continúe. Como la colección $\mathcal{C}$ es finita, el proceso anterior debe culminar en algún paso $m \leq N$. Por construcción, las bolas $\{B_{i_1}, \ldots, B_{i_m}\}$ son disjuntas y

$$\bigcup_{i=1}^{N} B_i \;\subseteq\; \bigcup_{j=1}^{m} B_{i_j}^* \;=\; \bigcup_{j=1}^{m} 3B_{i_j}.$$

De aquí se sigue que

$$\mu_n\Big(\bigcup_{i=1}^{N} B_i\Big) \;\leq\; 3^n \sum_{j=1}^{m} \mu_n(B_{i_j})$$

y termina la prueba. ■

Teorema 12.3.8 (Desigualdad de Hardy-Littlewood). *Si $f \in \mathrm{L}_1(\mathbb{R}^n)$, entonces $\mathrm{M}f \in \mathrm{L}_1^{\omega}(\mathbb{R}^n)$. Más aun, para todo $t > 0$ se cumple que*

$$\mu_n\big(\{x \in \mathbb{R}^n \,:\, \mathrm{M}f(x) > t\}\big) \;\leq\; \frac{3^n}{t}\, \| f \|_1 \,.$$

Prueba. Fijemos $t > 0$ y sea $E_t = \{x \in \mathbb{R}^n : \mathrm{M}f(x) > t\}$. Puesto que E_t es medible, la regularidad de μ_n nos dice que

$$\mu_n(E_t) \;=\; \sup\big\{\mu_n(K) \,:\, K \subseteq E_t,\ K \text{ compacto}\big\},$$

de modo que será suficiente demostrar que

$$\mu_n(K) \;\leq\; \frac{3^n}{t}\, \| f \|_1$$

para cualquier compacto $K \subseteq E_t$. Suponga entonces que $K \subseteq E_t$ es compacto y sea $x \in K$. Puesto que $x \in E_t$, existe un $r_x > 0$ tal que

$$\frac{1}{\mu_n(B(x, r_x))} \int_{B(x,r_x)} |f(y)|\, d\mu_n(y) \;>\; t.$$

Si hacemos $B_x = B(x, r_x)$ para cada $x \in K$, resulta que la colección $\{B_x : x \in K\}$ es un cubrimiento abierto de K, el cual, por la compacidad de K, se reduce a un subcubrimiento finito, digamos $K \subseteq B_1 \cup \ldots \cup B_N$. Por el Lema 12.3.7 existe una subfamilia disjunta $\{B_{i_1}, \ldots, B_{i_m}\}$ de $\{B_1, \ldots, B_N\}$ tal que

$$\begin{aligned}
\mu(K) \;&\leq\; \sum_{i=1}^{N} \mu_n(B_i) \;\leq\; 3^n \sum_{j=1}^{m} \mu_n(B_{i_j}) \\
&\leq\; \frac{3^n}{t} \sum_{j=1}^{m} \int_{B_{i_j}} |f|\, d\mu_n \;\leq\; \frac{3^n}{t} \int_{\mathbb{R}^n} |f|\, d\mu_n \\
&=\; \frac{3^n}{t}\, \| f \|_1
\end{aligned}$$

lo cual prueba el resultado. ■

12.3.3. El Teorema de Diferenciabilidad de Lebesgue

La función maximal de Hardy-Littlewood es el elemento fundamental para obtener el siguiente resultado.

Teorema 12.3.9 (Diferenciación de Lebesgue). *Si $f \in L^1_{\text{loc}}(\mathbb{R}^n)$, entonces existe un conjunto medible $N \subseteq \mathbb{R}^n$ con $\mu_n(N) = 0$ tal que*

$$\lim_{r\to 0} \frac{1}{\mu_n(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\, d\mu_n(y) \;=\; 0 \tag{4}$$

para todo $x \notin N$. En particular,

$$\lim_{r\to 0} \frac{1}{\mu_n(B(x,r))} \int_{B(x,r)} f(y)\, d\mu_n(y) \;=\; f(x) \tag{4}$$

para todo $x \notin N$.

Prueba. Puesto que

$$\begin{aligned} \left| \frac{1}{\mu_n(B(x,r))} \int_{B(x,r)} f(y)\, d\mu_n(y) - f(x) \right| \;&=\; \frac{1}{\mu_n(B(x,r))} \left| \int_{B(x,r)} (f(y) - f(x))\, d\mu_n(y) \right| \\ &\leq\; \frac{1}{\mu_n(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\, d\mu_n(y), \end{aligned}$$

todo lo que tenemos que demostrar es la otra desigualdad. Para este fin, defina la función $f^* : \mathbb{R}^n \to \mathbb{R}$ por

$$f^*(x) \;=\; \limsup_{r\to 0+} \left[\frac{1}{\mu_n(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\, d\mu_n(y) \right]$$

y veamos que $f^* = 0$ casi-siempre. Esto, por supuesto, es equivalente a demostrar que

$$\mu_n\big(\{x \in \mathbb{R}^n \,:\, f^*(x) > 0\}\big) \;=\; 0.$$

Observación 1. *Si f es continua sobre $\mathbb{R}^n$, entonces la desigualdad obtenida en* (5) *nos muestra que* $f^* = 0$.

Suponga ahora que $f \in L^1_{\text{loc}}(\mathbb{R}^n)$. Nótse que si el resultado se cumple para $f \cdot \chi_{B(0,k)} \in L_1(\mathbb{R}^n)$ para cada $k \in \mathbb{N}$, excepto sobre un conjunto medible E_k de medida cero, entonces éste se cumple para $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ excepto sobre $\bigcup_{k=1}^{\infty} E_k$, el cual es medible y tiene medida cero. Esta observación nos permite suponer, sin perder generalidad, que $f \in L_1(\mathbb{R}^n)$.

Sea $\varepsilon > 0$. Como $f \in L_1(\mathbb{R}^n)$, el Teorema de Aproximación nos garantiza la existencia de una función $g \in C_c(\mathbb{R}^n)$ tal que $\| f - g \|_1 < \varepsilon$. Afirmamos que existe una constante $C > 0$, dependiendo sólo de n, tal que para cualquier $t > 0$,

$$\mu_n\big(\{x \in \mathbb{R}^n \,:\, f^*(x) > t\}\big) \;\leq\; \frac{C}{t} \| f \|_1 . \tag{6}$$

Para demostrar nuestra afirmación, observe que

$$\begin{aligned} f^*(x) \;&\leq\; \sup_{r>0} \left[\frac{1}{\mu_n(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\, d\mu_n(y) \right] \\ &\leq\; \sup_{r>0} \left[\frac{1}{\mu_n(B(x,r))} \int_{B(x,r)} |f(y)|\, d\mu_n(y) \right] \,+\, |f(x)| \\ &\leq\; \mathrm{M}f(x) \,+\, |f(x)|, \end{aligned}$$

de donde se sigue que

$$\begin{aligned}\{x \in \mathbb{R}^n : f^*(x) > t\} &\subseteq \{x \in \mathbb{R}^n : \mathrm{M}f(x) + |f(x)| > t\} \\ &\subseteq \{x \in \mathbb{R}^n : \mathrm{M}f(x) > t/2\} \cup \{x \in \mathbb{R}^n : |f(x)| > t/2\}.\end{aligned}$$

Ahora bien, por la Desigualdad de Hardy-Littlewood

$$\mu_n\big(\{x \in \mathbb{R}^n : \mathrm{M}f(x) > t/2\}\big) \le \frac{2 \cdot 3^n}{t} \| f \|_1 ,$$

mientras que la Desigualdad de Chebyshev nos dice que

$$\mu_n\big(\{x \in \mathbb{R}^n : |f(x)| > t/2\}\big) \le \frac{2}{t} \| f \|_1 .$$

La combinación de estos hechos nos muestra que (6) se cumple tomando $C = 2(1 + 3^n)$. Finalmente, el hecho de que $(f - g)^* = f^*$ combinado con (6) produce la siguiente desigualdad:

$$\begin{aligned}\mu_n\big(\{x \in \mathbb{R}^n : f^*(x) > t\}\big) &= \mu_n\big(\{x \in \mathbb{R}^n : (f - g)^*(x) > t\}\big) \\ &\le \frac{C}{t} \| f - g \|_1 \\ &< \frac{C}{t}\varepsilon.\end{aligned}$$

Puesto que $\varepsilon > 0$ es arbitrario, se tiene que

$$\mu_n\big(\{x \in \mathbb{R}^n : f^*(x) > t\}\big) = 0$$

y como

$$\{x \in \mathbb{R}^n : f^*(x) > 0\} = \bigcup_{k=1}^{\infty} \{x \in \mathbb{R}^n : f^*(x) > 1/k\}$$

se concluye que

$$\mu_n\big(\{x \in \mathbb{R}^n : f^*(x) > 0\}\big) = 0.$$

Esto termina la prueba. ■

Nota Adicional 12.3.5 La afirmación de que $(f - g)^* = f^*$ utilizada en la demostración del resultado anterior es realmente fácil de obtener. En efecto, como $\limsup(A + B) \le \limsup A + \limsup B$, resulta que

$$(f + g)^* \le f^* + g^*$$

y puesto que $g^* = 0$, se tiene que $(f - g)^* \le f^* + g^* = f^*$. Finalmente, como

$$f^* = (f - g + g)^* \le (f - g)^* + g^* = (f - g)^*$$

concluimos que $(f - g)^* = f^*$.

Un caso particularmente interesante del Teorema de Diferenciación de Lebesgue ocurre cuando ν es una medida real de Borel sobre $\mathbb{R}^n$ tal que $\nu \ll \mu_n$. En efecto, en este caso el Teorema de Radon-Nikodým nos garantiza la existencia de una función $f \in \mathrm{L}_1(\mathbb{R}^n)$ tal que

$$\nu(E) = \int_E f \, d\mu_n, \quad E \in \mathfrak{Bo}(\mathbb{R}^n),$$

de donde se sigue, por el Teorema de Diferenciación de Lebesgue, que para cada casi-todo $a \in \mathbb{R}^n$

$$\begin{aligned} \lim_{r\to 0}\left|\frac{\nu(B)}{\mu_n(B)} - f(a)\right| &= \lim_{r\to 0}\frac{1}{\mu_n(B)}\left|\int_B (f(x)-f(a))\,d\mu_n\right| \\ &\leq \lim_{r\to 0}\frac{1}{\mu_n(B)}\int_B |f(x)-f(a)|\,d\mu_n = 0, \end{aligned}$$

donde hemos escrito $B = B(a,r)$. Esto sugiere la siguiente fórmula para computar la derivada de Radon-Nikodým:

$$f(a) = \lim_{B\to a}\frac{\nu(B)}{\mu_n(B)}$$

donde $B \to a$ significa que $a \in B$ y el volumen de B se contrae a 0.

12.4. Problemas

(1) Sea μ la medida de Lebesgue sobre $\mathbb{R}$ y sea $\nu : \mathfrak{M}_\mu(\mathbb{R}) \to [0,+\infty]$ la medida de conteo. Demuestre que:

(a) $\mu \ll \nu$ pero $\frac{d\mu}{d\nu}$ no existe.

(b) ν no tiene descomposición de Lebesgue con respecto a μ.

(2) Sean λ y ν dos medidas positivas sobre un espacio medible $(X,\mathcal{A})$. Suponga que $\frac{d\nu}{d\lambda}$ existe tal que $\nu \ll \lambda$.

(a) Demuestre que si $\frac{d\nu}{d\lambda} > 0$ λ-casi-siempre sobre X, entonces $\lambda \ll \nu$ y así, $\lambda \sim \nu$.

(b) Demuestre que si $\frac{d\nu}{d\lambda} > 0$ λ-casi-siempre sobre X y si λ y ν son σ-finitas, entonces $\frac{d\lambda}{d\nu}$ existe y

$$\frac{d\lambda}{d\nu} = \left(\frac{d\nu}{d\lambda}\right)^{-1}, \quad \lambda - \text{casi-siempre} \;\text{ y }\; \nu - \text{casi-siempre} \;\text{ sobre } X.$$

(3) Sea $(X,\mathcal{A},\lambda)$ un espacio de medida. Asuma que existe una función medible $f : X \to (0,+\infty)$ satisfaciendo la condición $\lambda(\{x \in X : f(x) \leq n\}) < +\infty$ para cada $n \in \mathbb{N}$.

(a) Pruebe que λ es σ-finita.

(b) Defina $\nu : \mathcal{A} \to [0,+\infty]$ por

$$\nu(E) = \int_E f\,d\lambda \quad \text{para todo } E \in \mathcal{A}.$$

Demuestre que ν es σ-finita.

(c) Demuestre que $\frac{d\lambda}{d\nu}$ existe y

$$\frac{d\lambda}{d\nu} = \frac{1}{f}, \quad \lambda - \text{casi-siempre} \;\text{ y }\; \nu - \text{casi-siempre} \;\text{ sobre } X.$$

(4) Sean λ y ν medidas finitas definidas sobre $(X, \mathcal{A})$. Sea $(\nu_n)_{n=1}^{\infty}$ una sucesión de medidas finitas definidas sobre $\mathcal{A}$ tal que $\nu = \sum_{n=1}^{\infty} \nu_n$. Suponga que, para cada $n \in \mathbb{N}$ y $E \in \mathcal{A}$,

$$\nu_n(E) \ = \ \int_E f_n \, d\lambda \ + \ \nu_n'(E),$$

donde cada $f_n : X \to [0, +\infty)$ es una función medible sobre X y $\nu_n' \perp \lambda$. Finalmente, asuma que para cada $E \in \mathcal{A}$ se cumple que

$$\nu(E) \ = \ \int_E f \, d\lambda \ + \ \nu'(E),$$

donde f es no-negativa sobre X y $\nu' \perp \lambda$. Demuestre que

(a) $f = \sum_{n=1}^{\infty} f_n$ existe λ-casi-siempre sobre X.

(b) $\nu = \sum_{n=1}^{\infty} \nu_n'$.

(c) $\nu \ll \lambda$ si, y sólo si, $\nu_n \ll \lambda$ para cualquier $n \in \mathbb{N}$.

(5) En este problema se establece que el Teorema de Representación de Riesz no es válido cuando $(X, \mathcal{M}, \nu)$ es un espacio de medida arbitrario no σ-finito y $p = 1$.

Sea $X = [0,1] \times [0,1]$ y sea $\mathcal{M}$ la σ-álgebra de Borel de X. Por cada $E \in \mathcal{M}$ defina

$$\nu(E) \ = \sum_{x \in [0,1]} \mu(E_x) \qquad \text{y} \qquad \lambda(E) \ = \ \nu(E) \ + \sum_{y \in [0,1]} \mu(E^y)$$

donde $E_x = \{y \in [0,1] : (x,y) \in E\}$ y $E^y = \{x \in [0,1] : (x,y) \in E\}$. Observe que ambas medidas existen, véase la Definición 2.1.52, página 114 y, además, se cumple que $\nu \leq \lambda$.

(a) Verifique que $\mathrm{L}_1(X, \lambda) \subseteq \mathrm{L}_1(X, \nu)$. Esto permite definir, sin ambigüedad, la aplicación $\varphi : \mathrm{L}_1(X, \lambda) \to \mathbb{R}$ por

$$\varphi(f) \ = \ \int_X f \, d\nu \quad \text{para cualquier } f \in \mathrm{L}_1(X, \lambda)$$

y comprobar que ella es un funcional lineal continuo sobre $\mathrm{L}_1(X, \lambda)$ con $\| \varphi \| \leq 1$.

(b) Suponga ahora que la conclusión del Teorema de Representación de Riesz se cumple; es decir, que existe una función $g \in \mathrm{L}_\infty(X, \lambda)$ tal que $\varphi = \varphi_g$ para esta g. Esto es, asuma que $\varphi(f) = \int_X fg \, d\lambda$ para cualquier $f \in \mathrm{L}_1(X, \lambda)$. Demuestre que existe un $x_0 \in [a, b]$ tal que

$$\int_{[0,1]} |g(x_0, y)| d\mu(y) \ = \ 0.$$

Sean $V = \{x_0\} \times [0,1]$ y $f = \chi_V$. Verifique $\nu(V) = 1$, $f \in \mathrm{L}_1(X, \lambda)$ y

$$1 \ = \ \nu(V) \ = \ \int_X f \, d\nu \ = \ \int_X fg \, d\lambda \ = \ \int_V g \, d\lambda \ \leq \ \int_{[0,1]} g(x_0, y) \, d\mu(y) \ = \ 0.$$

Esta contradicción se produce por el hecho de que λ no es σ-finita.

(6) El siguiente problema describe cómo es el dual de $L_\infty(X, \mathcal{M}, \lambda)$. Sea $(X, \mathcal{M}, \lambda)$ un espacio de medida σ-finito y denote por $\text{ba}(X, \mathcal{M}, \lambda)$ *la colección de todas las funciones funciones de conjuntos* $\nu : \mathcal{M} \to \mathbb{R}$ *que satisfacen las siguientes condiciones*:

(a) $\nu(A \cup B) = \nu(A) + \nu(B)$ siempre que $A, B \in \mathcal{M}$ sean disjuntos,

(b) $\sup\{|\nu(A)| : A \in \mathcal{M}\} < +\infty$, y

(c) $\nu(A) = 0$ siempre que $A \in \mathcal{M}$ y $\lambda(A) = 0$.

Dado $\nu \in \text{ba}(X, \mathcal{M}, \lambda)$, defina $|\nu|$ sobre $\mathcal{M}$ escribiendo

$$|\nu|(A) \;=\; \sup\left\{ \sum_{i=1}^{n} |\nu(A_i)| \;:\; A_1, \ldots, A_n \text{ es una partición medible de } A \right\}$$

para cualquier $A \in \mathcal{M}$. Recuerde que una colección finita $\{A_1, \ldots, A_n\}$ es una partición medible de $A \in \mathcal{M}$ si ellos son disjuntos, pertenecen a $\mathcal{M}$ y su unión es A. Defina

$$\|\nu\| \;=\; |\nu|(X) \quad \text{para cada } \nu \in \text{ba}(X, \mathcal{M}, \lambda).$$

Demuestre que:

(1) $|\nu| \in \text{ba}(X, \mathcal{M}, \lambda)$.

(2) $(\text{ba}(X, \mathcal{M}, \lambda), \|\cdot\|)$ es un espacio de Banach.

Si $\{A_1, \ldots, A_n\}$ es una partición medible de X y $s = \sum_{i=1}^{n} a_i \chi_{A_i}$ es una función simple, entonces para cada $\nu \in \text{ba}(X, \mathcal{M}, \lambda)$ se tiene que

$$\left| \sum_{i=1}^{n} a_i \, \nu(A_i) \right| \;\leq\; \sum_{i=1}^{n} |a_i| \big|\nu(A_i)\big| \;\leq\; \|s\|_\infty \, \|\nu\|$$

Defina entonces

$$\int_X s \, d\nu \;=\; \sum_{i=1}^{n} a_i \, \nu(A_i).$$

Si $f \in L_\infty(X, \mathcal{M}, \lambda)$, escoja una sucesión $(s_n)_{n=1}^\infty$ de funciones simples definidas sobre X tal que $\|s_n - f\|_\infty \to 0$.

(3) Demuestre que la sucesión $\left(\int_X s_n \, d\nu\right)_{n=1}^\infty$ es de Cauchy en $\mathbb{R}$.

Defina $\displaystyle\int_X f \, d\nu = \lim_{n\to\infty} \int_X s_n \, d\nu$.

(4) Demuestre que la integral de f está bien definida; en otras palabras, no depende de la elección particular de funciones simples que converge a f. Además, pruebe que

(a) $\displaystyle\int_X (f + g) \, d\nu = \int_X f \, d\nu + \int_X g \, d\nu$.

(b) $\displaystyle\int_X a f \, d\nu = a \int_X f \, d\nu$.

(c) $\displaystyle\left| \int_X f \, d\nu \right| \leq \int_X |f| \, d|\nu|$.

(d) $\displaystyle\int_X f \, d|\nu| \leq \int_X g \, d|\nu|$ siempre que $0 \leq f \leq g$.

(4) Finalmente, demuestre que $\big(L_\infty(X, \mathcal{M}, \lambda)\big)^* = \text{ba}(X, \mathcal{M}, \lambda)$. (véase, por ejemplo, [133], p. 416-418 para ciertos detalles).

CAPÍTULO 13

Convergencia en $\text{ca}(\mathcal{M}, \mathbb{R})$

Si $(\nu_n)_{n=1}^{\infty}$ en una sucesión en $\text{ca}(\mathcal{M}, \mathbb{R})$ que converge en la norma de la variación a una medida $\nu \in \text{ca}(\mathcal{M}, \mathbb{R})$, entonces de la desigualdad $|\nu(E)| \leq |\nu|(X)$, la cual se cumple para todo $E \in \mathcal{M}$, se sigue que

$$|\nu_n(E) - \nu(E)| \ \leq \ \| \nu_n - \nu \| \ \to \ 0,$$

es decir,

$$\nu(E) \ = \ \lim_{n\to\infty} \nu_n(E) \quad \text{para todo } E \in \mathcal{M}.$$

En este caso se dice que la sucesión $(\nu_n)_{n=1}^{\infty}$ **converge puntualmente** sobre $\mathcal{M}$ a ν. Si bien es cierto que la convergencia puntual de una sucesión en $\text{ca}(\mathcal{M}, \mathbb{R})$ no garantiza la convergencia en la norma de la variación, sin embargo, se puede obtener cierta información de su límite puntual.

El objetivo de esta sección es presentar tres resultados fundamentales acerca de la convergencia puntual de medidas en $\text{ca}(\mathcal{M}, \mathbb{R})$. El primero de ellos se conoce con el nombre de **Teorema de la Convergencia de Nikodým** el cual establece que si $(\nu_n)_{n=1}^{\infty}$ en una sucesión de medidas en $\text{ca}(\mathcal{M}, \mathbb{R})$ convergiendo puntualmente a ν, entonces $\nu \in \text{ca}(\mathcal{M}, \mathbb{R})$. El segundo es el **Teorema de Acotación de Nikodým** que afirma que si $\mathcal{K}$ es un subconjunto de $\text{ca}(\mathcal{M}, \mathbb{R})$ el cual es puntualmente acotado, esto es, para cada $E \in \mathcal{M}$, existe una constante $M_E > 0$ tal que

$$|\nu(E)| \ \leq \ M_E \quad \text{para cada } \nu \in \mathcal{K},$$

entonces $\mathcal{K}$ es uniformemente acotada en el siguiente sentido: existe una constante $M > 0$ tal que

$$\sup_{\nu \in \mathcal{K}} \| \nu \| \ \leq \ M.$$

Finalmente, el **Teorema de Vitali-Hahn-Saks** que nos asegura que si $(\nu_n)_{n=1}^{\infty}$ en una sucesión de medidas en $\text{ca}(\mathcal{M}, \mathbb{R})$ convergiendo puntualmente a ν y tal que $\nu_n \ll \lambda$ para todo $n \geq 1$, entonces $\nu \ll \lambda$ y la sucesión $(\nu_n)_{n=1}^{\infty}$ es uniformemente σ-aditiva con respecto a λ. Existen, por supuestos, varias extensiones y demostraciones de estos resultados. Nuestro programa para alcanzar estos objetivos comienza con un criterio general de convergencia creado por P. Antonsik y J. Mikusiński sobre una matriz infinita verificando ciertas propiedades que facilita la demostración del Teorema de la Convergencia de Nikodým.

13.1. Convergencia según Antosik-Mikusiński

En esta sección se establecen dos resultados básicos sobre matrices infinitas que será usados posteriormente para demostrar el Teorema de la Convergencia de Nikodým. Comenzaremos estableciendo ciertas notaciones. Sea $A = [a_{ij}]_{i,j=1}^{\infty}$ una matriz infinita de números reales, es decir,

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & \cdots \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} & \cdots \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \end{pmatrix},$$

donde los $a_{ij} \in \mathbb{R}$ para todo $i, j \in \mathbb{N}$. Diremos que la matriz A es no-negativa si $a_{ij} \geq 0$ para todo $i, j \in \mathbb{N}$.

Lema 13.1.1. *Sea $A = [a_{ij}]_{i,j=1}^{\infty}$ una **matriz infinita no-negativa** y sea $(\varepsilon_{ij})_{i,j\in\mathbb{N}}$ una sucesión doble de números reales con $\varepsilon_{ij} > 0$ para todo $i, j \in \mathbb{N}$. Si*

(a) $\lim_{i\to\infty} a_{ij} = 0$ *para cada* $j \in \mathbb{N}$, *y*

(b) $\lim_{j\to\infty} a_{ij} = 0$ *para cada* $i \in \mathbb{N}$,

entonces existe una subsucesión $(m_i)_{i=1}^{\infty}$ de $\mathbb{N}$ tal que

$$a_{m_i m_j} < \varepsilon_{ij} \quad \text{para todo } i, j \in \mathbb{N} \text{ con } i \neq j.$$

Prueba. Comencemos definiendo $m_1 = 1$. Usemos ahora el hecho de que

$$\lim_{j\to\infty} a_{m_1 j} = 0 \qquad \text{y} \qquad \lim_{i\to\infty} a_{i m_1} = 0,$$

para hallar un $m_2 > m_1$ de modo que para todo $m \geq m_2$,

$$a_{m_1 m} < \varepsilon_{12} \qquad \text{y} \qquad a_{m m_1} < \varepsilon_{21}.$$

En particular, tomando $m = m_2$ encontramos que:

$$a_{m_1 m_2} < \varepsilon_{12} \qquad \text{y} \qquad a_{m_2 m_1} < \varepsilon_{21}.$$

De nuevo, teniendo en cuenta que

$$\lim_{j\to\infty} a_{m_1 j} = 0, \quad \lim_{j\to\infty} a_{m_2 j} = 0, \quad \lim_{i\to\infty} a_{i m_1} = 0 \quad \text{y} \quad \lim_{i\to\infty} a_{i m_2} = 0,$$

existe un $m_3 > m_2$ tal que las desigualdades

$$a_{m_1 m} < \varepsilon_{13}, \quad a_{m_2 m} < \varepsilon_{23}, \quad a_{m m_1} < \varepsilon_{31} \quad \text{y} \quad a_{m m_2} < \varepsilon_{32}.$$

se cumplen para todo $m \geq m_3$. Eligiendo el valor $m = m_3$, se obtiene

$$a_{m_1 m_3} < \varepsilon_{13}, \qquad a_{m_2 m_3} < \varepsilon_{23}, \qquad a_{m_3 m_1} < \varepsilon_{31} \qquad \text{y} \qquad a_{m_3 m_2} < \varepsilon_{32}.$$

Continuando inductivamente con éste procedimiento se obtiene la conclusión deseada. ■

El Lema 13.1.1 se puede aplicar en muchas situaciones. Una de ellas, y que es importante en esta sección, es el siguiente resultado.

Teorema 13.1.2 (Antosik-Mikusińki). *Sea $A = [a_{ij}]_{i,j=1}^{\infty}$ una **matriz infinita de números reales**. Suponga que*

(1) $\lim\limits_{i\to\infty} a_{ij} = a_j$ *existe para cada* $j \in \mathbb{N}$, *y*

(2) *para cada subsucesión* $(m_j)_{j=1}^{\infty}$ *de* $\mathbb{N}$, *existe una subsucesión* $(n_{k_j})_{j=1}^{\infty}$ *de* $(m_j)_{j=1}^{\infty}$ *tal que*

$$\lim_{i\to\infty} \sum_{j=1}^{\infty} a_{i\, n_{k_j}}$$

existe. Entonces $\lim_{i\to\infty} a_{ij} = a_j$ ***uniformemente para*** $j \in \mathbb{N}$. *Más aun,* $\lim\limits_{j\to\infty} a_{ij} = 0$ ***uniformemente para*** $i \in \mathbb{N}$ *y* $\lim\limits_{i\to\infty} a_{ii} = 0$.

Prueba. Suponga que la conclusión es falsa. Entonces existe un $\varepsilon > 0$ y una subsucesión $(k_i)_{i=1}^{\infty}$ de $\mathbb{N}$ tal que

$$\sup_{j\in\mathbb{N}} \left|a_{k_i j} - a_j\right| > \varepsilon \quad \text{para } i = 1, 2, \ldots \tag{α}$$

Para evitar el uso engorroso de los subíndices, asumiremos que $k_i = i$ para todo $i \in \mathbb{N}$. Pongamos $i_1 = 1$ y usemos (α) para elegir un $j_1 \in \mathbb{N}$ de modo que

$$\left|a_{i_1 j_1} - a_{j_1}\right| > \varepsilon.$$

Por otro lado, la condición (1) nos garantiza la existencia de un $i_2 > i_1$ tal que

$$\left|a_{ij} - a_j\right| < \varepsilon \quad \text{para todo } i \geq i_2 \quad \text{y} \quad j \in \{1, \ldots, j_1\}$$

y

$$\left|a_{i_1 j_1} - a_{i_2 j_1}\right| > \varepsilon.$$

Usemos de nuevo la desigualdad (α) para seleccionar un $j_2 \in \mathbb{N}$ tal que

$$\left|a_{i_2 j_2} - a_{j_2}\right| > \varepsilon.$$

Observe que $j_2 > j_1$. Aplicando una vez más la condición (1) nos produce un $i_3 > i_2$ tal que

$$\left|a_{ij} - a_j\right| < \varepsilon \quad \text{para todo } i \geq i_3 \quad \text{y} \quad j \in \{j_1 + 1, \ldots, j_2\}$$

y

$$\left|a_{i_2 j_2} - a_{i_3 j_2}\right| > \varepsilon.$$

Continuando inductivamente con el procedimiento anterior se logra obtener un par de sucesiones estrictamente crecientes $(i_k)_{k=1}^\infty$ y $(j_k)_{k=1}^\infty$ tales que

$$\left|a_{i_kj_k} - a_{i_{k+1}j_k}\right| > \varepsilon \quad \text{para todo } k \in \mathbb{N}.$$

Pongamos $z_{kl} = a_{i_kj_l} - a_{i_{k+1}j_l}$ para todo $k, l \in \mathbb{N}$ y observe que

$$\left|z_{kk}\right| > \varepsilon. \tag{β}$$

Estamos listo para aplicar el lema anterior. En efecto, considere la matriz infinita positiva $A = [|z_{kl}|]_{k,l=1}^\infty$. Por (1) las columnas de ésta matriz convergen a 0. Por otro lado, por (2), las filas de la matriz $B = [a_{ij}]_{i,j=1}^\infty$ convergen a 0 de modo que, por construcción, las filas de A también convergen a 0. Fijemos una sucesión $(\varepsilon_{ij})_{i,j=1}^\infty$ de números positivos tal que $\sum_{i,j=1}^\infty \varepsilon_{ij} < +\infty$. Por ejemplo, tome $\varepsilon_{kl} = 2^{-(k+l)}$. Por el Lema 13.1.1, existe una subsucesión $(m_k)_{k=1}^\infty$ de $\mathbb{N}$ tal que

$$\left|z_{m_km_l}\right| < \frac{1}{2^{k+l}} \quad \text{para todo } k, l \in \mathbb{N} \text{ con } k \neq l.$$

Para la subsucesión $(m_k)_{k=1}^\infty$ que acabamos de hallar, existe, por una aplicación de (2), una subsucesión $(n_k)_{k=1}^\infty$ tal que

$$\lim_{k\to\infty} \sum_{l=1}^\infty z_{n_kn_l} = 0. \tag{γ}$$

Ahora bien, como

$$z_{n_kn_k} = \sum_{l=1}^\infty z_{n_kn_l} - \sum_{\substack{l=1\\ l\neq k}}^\infty z_{n_kn_l}$$

resulta que

$$\begin{aligned}\left|z_{n_kn_k}\right| &\leq \left|\sum_{l=1}^\infty z_{n_kn_l}\right| + \left|\sum_{\substack{l=1\\ l\neq k}}^\infty z_{n_kn_l}\right| \leq \left|\sum_{l=1}^\infty z_{n_kn_l}\right| + \sum_{\substack{l=1\\ l\neq k}}^\infty \left|z_{n_kn_l}\right| \\ &\leq \left|\sum_{l=1}^\infty z_{n_kn_l}\right| + \sum_{l=1}^\infty \frac{1}{2^{k+l}} = \left|\sum_{l=1}^\infty z_{n_kn_l}\right| + \frac{1}{2^k}\end{aligned}$$

y se sigue entonces de (γ) que

$$\lim_{k\to\infty} \left|z_{n_kn_k}\right| = 0$$

lo que, por supuesto, contradice a (β). Por esto, $\lim_{i\to\infty} a_{ij} = a_j$ uniformemente para $j \in \mathbb{N}$.

Para demostrar la última parte, observe que la convergencia uniforme de $\lim_{i\to\infty} a_{ij} = a_j$, combinado con el hecho de que $\lim_{j\to\infty} a_{ij} = 0$ para cada $i \in \mathbb{N}$, implica que

$$0 = \lim_{i\to\infty}\lim_{j\to\infty} a_{ij} = \lim_{j\to\infty}\lim_{i\to\infty} a_{ij} = \lim_{i,j\to\infty} a_{ij}.$$

Esto prueba que $\lim_{j\to\infty} a_{ij} = 0$ uniformemente para $i \in \mathbb{N}$. En particular, $\lim_{i\to\infty} a_{ii} = 0$. Fin de la prueba. ■

13.2. Medidas Uniformemente σ-aditivas

La presentación de esta sección comienza definiendo lo que significa que una familia de medidas reales sea uniformemente σ-aditiva y luego, a partir de allí, obtener algunas de sus equivalencias. Esto resultados serán fundamentales para la obtención de los dos Teoremas de Nikodým mencionados con anterioridad.

Recordemos que si $\nu \in \text{ca}(\mathcal{M}, \mathbb{R})$, entonces la serie $\sum_{j=1}^{\infty} \nu(E_j)$ converge (absolutamente) para cualquier sucesión disjunta de conjuntos $(E_j)_{j=1}^{\infty}$ en $\mathcal{M}$. Si hacemos $s_n = \sum_{j=1}^{n} \nu(E_j)$ y $s = \sum_{j=1}^{\infty} \nu(E_j)$, entonces

$$0 \;=\; \lim_{n\to\infty} (s_n - s) \;=\; \lim_{n\to\infty} \sum_{j=n+1}^{\infty} \nu(E_j).$$

Por supuesto, si $\nu_1, \ldots, \nu_k$ es un conjunto finito de medidas en $\text{ca}(\mathcal{M}, \mathbb{R})$, entonces la conclusión es la misma:

$$\sup\left\{ \sum_{j=1}^{\infty} \nu_n(E_j) \;:\; n = 1, \ldots, k \right\}$$

converge para cualquier sucesión disjunta de conjuntos $(E_j)_{j=1}^{\infty}$ en $\mathcal{M}$.

Definición 13.2.1. *Una familia de medidas $\mathcal{K} \subseteq \text{ca}(\mathcal{M}, \mathbb{R})$ se dice que es **uniformemente σ-aditiva** si para cualquier sucesión disjunta de conjuntos $(E_j)_{j=1}^{\infty}$ en $\mathcal{M}$, la serie $\sum_{j=1}^{\infty} \nu(E_j)$ converge uniformemente con respecto a $\nu \in \mathcal{K}$, o equivalentemente,*

$$\lim_{n\to\infty} \sum_{j=n}^{\infty} \nu(E_j) \;=\; 0$$

uniformemente con respecto a $\nu \in \mathcal{K}$.

Fijemos una sucesión disjunta $(E_j)_{j=1}^{\infty}$ en $\mathcal{M}$. Que la serie $\sum_{j=1}^{\infty} \nu(E_j)$ converge uniformemente con respecto a $\nu \in \mathcal{K}$ significa que (usando el Criterio de Cauchy para Series), para cada $\varepsilon > 0$, siempre se puede determinar un $N \in \mathbb{N}$ de modo que cualesquiera sean $m > n > N$ se cumple que

$$\left| \sum_{j=n+1}^{m} \nu(E_j) \right| \;<\; \varepsilon \quad \text{uniformemente respecto a } \nu \in \mathcal{K}.$$

Por otro lado, como la serie $\sum_{j=1}^{\infty} \nu(E_j)$ converge (uniformemente con respecto a $\nu \in \mathcal{K}$,) resulta que $\lim\limits_{j\to\infty} \nu(E_j) = 0$ uniformemente en $\nu \in \mathcal{K}$, o de forma equivalente,

$$\lim_{j\to\infty} \big(\sup\{\nu(E_j) \;:\; \nu \in \mathcal{K}\} \big) \;=\; 0.$$

Un peldaño, aun por debajo, pero necesario para alcanzar la prueba del Teorema de la Convergencia de Nikodým es el siguiente resultado.

Lema 13.2.2. *Sea $\mathcal{K} \subseteq \text{ca}(\mathcal{M}, \mathbb{R})$. Las siguientes condiciones son equivalentes:*

(1) *$\mathcal{K}$ es **uniformemente σ-aditiva**.*

(2) $\mathcal{K}$ *es* ***uniformemente exhaustiva****, esto es, para cada sucesión disjunta de conjuntos* $(E_j)_{j=1}^{\infty}$ *en* $\mathcal{M}$*,*

$$\lim_{j\to\infty} \nu(E_j) = 0 \quad \textbf{\textit{uniformemente en}} \ \nu \in \mathcal{K}.$$

(3) *Para cada* ***sucesión decreciente*** $(E_j)_{j=1}^{\infty}$ *en* $\mathcal{M}$ *con* $\bigcap_{j=1}^{\infty} E_j = \varnothing$*,*

$$\lim_{j\to\infty} \nu(E_j) = 0 \quad \textbf{\textit{uniformemente}} \textit{ en } \nu \in \mathcal{K}.$$

Prueba. (1) $\Rightarrow$ (2). Ya vimos que si $\mathcal{K}$ es uniformemente σ-aditiva, entonces $\lim_{j\to\infty} \nu(E_j) = 0$ uniformemente para $\nu \in \mathcal{K}$ para cualquier sucesión disjunta de conjuntos $(E_j)_{j=1}^{\infty}$ incluida en $\mathcal{M}$.

(2) $\Rightarrow$ (1). Suponga que (2) se cumple pero no así (1). Esto, por supuesto, significa que existe una sucesión disjunta de conjuntos $(E_j)_{j=1}^{\infty}$ en $\mathcal{M}$ y un $\varepsilon > 0$ con la siguiente propiedad: para cada $k \in \mathbb{N}$, existen enteros positivos $m_k > n_k > k$ tal que

$$\left| \sum_{j=n_k+1}^{m_k} \nu_k(E_j) \right| \geq \varepsilon.$$

Por inducción, podemos elegir sucesiones $(m_k)_{k=1}^{\infty}$ y $(n_k)_{k=1}^{\infty}$ de modo que $n_k < m_k < n_{k+1}$ para todo $k \geq 1$. Considere, para cada $k \in \mathbb{N}$ el conjunto

$$F_k = \bigcup_{j=n_k+1}^{m_k} E_j.$$

Resulta que los conjuntos $(F_k)_{k=1}^{\infty}$ son disjuntos dos a dos y se tiene que

$$|\nu_k(F_k)| = \left| \sum_{j=n_k+1}^{m_k} \nu_k(E_j) \right| \geq \varepsilon,$$

para cada $k \in \mathbb{N}$. Esto, evidentemente, constituye una violación a la hipótesis de que $\nu(F_k) \to \infty$ uniformemente para $\nu \in \mathcal{K}$.

(1) $\Leftrightarrow$ (3). Esta equivalencia es consecuencia del Teorema 12.1.5. La prueba es completa. ■

Si $\mathcal{K}$ es una familia de medidas incluidas en $\mathrm{ca}(\mathcal{M}, \mathbb{R})$, entonces escribiremos

$$|\mathcal{K}| = \big\{ |\nu| : \nu \in \mathcal{K} \big\}.$$

Lema 13.2.3. *Sea* $\mathcal{K} \subseteq \mathrm{ca}(\mathcal{M}, \mathbb{R})$*. Las siguientes condiciones son equivalentes:*

(1) $\mathcal{K}$ *es* ***uniformemente*** $\boldsymbol{\sigma}$***-aditiva****.*

(2) $|\mathcal{K}|$ *es* ***uniformemente*** $\boldsymbol{\sigma}$***-aditiva****.*

Prueba. (1) $\Rightarrow$ (2). Suponga que (1) se cumple pero que (2) es falsa. Entonces existe una sucesión $\big(E_n\big)_{n=1}^{\infty}$ de conjuntos disjuntos en $\mathcal{M}$, un $\varepsilon > 0$ y una sucesión $(\nu_n)_{n=1}^{\infty}$ en $\mathcal{K}$ tal que

$$\big|\nu_n\big|(E_n) > 2\varepsilon \quad \text{para cada } n \in \mathbb{N}.$$

Por el Teorema 12.1.16, página 793, por cada $n \in \mathbb{N}$, existe un conjunto $F_n \in \mathcal{M}$ con $F_n \subseteq E_n$ tal que

$$|\nu_n|(E_n) \ \leq \ 2|\nu_n(F_n)|.$$

De esto de sigue que

$$|\nu_n(F_n)| \ > \ \varepsilon \quad \text{para cada } n \in \mathbb{N}. \tag{$*$}$$

Puesto que la sucesión $(F_n)_{n=1}^{\infty}$ es disjunta y $\mathcal{K}$ es uniformemente σ-aditiva, se sigue del Lema 13.2.2 que

$$\lim_{n\to\infty} \nu(F_n) \ = \ 0 \quad \text{uniformemente para } \nu \in \mathcal{K},$$

lo que, por supuesto, contradice la desigualdad $(*)$.

$(2) \Rightarrow (1)$. Si $|\mathcal{K}|$ es uniformemente σ-aditiva, entonces claramente $\mathcal{K}$ es uniformemente σ-aditiva y termina la prueba. ■

El siguiente lema será usado para demostrar la existencia de una medida positiva λ tal que $\nu \ll \lambda$ uniformemente respecto a $\nu \in \mathcal{K}$.

Lema 13.2.4. *Sea $\mathcal{K} \subseteq \mathrm{ca}(\mathcal{M}, \mathbb{R})$. Si $\mathcal{K}$ es **uniformemente σ-aditiva**, entonces, para cada $\varepsilon > 0$, existe un $\delta > 0$ y un **conjunto finito** $J \subseteq \mathcal{K}$ con la siguiente propiedad: para cualquier $E \in \mathcal{M}$*

$$\sup\{|v|(E) : v \in J\} \ < \ \delta \qquad \Rightarrow \qquad \sup_{\nu \in \mathcal{K}} |\nu(E)| \ < \ \varepsilon.$$

Prueba. Suponga que la conclusión es falsa. Entonces existe un $\varepsilon > 0$ tal que: *para cada $\delta > 0$ y cualquier conjunto finito $J \subseteq \mathcal{K}$, existe un conjunto $E_J \in \mathcal{M}$ y una medida $\nu_J \in \mathcal{K}$ tal que*

$$|v|(E_J) \ < \ \varepsilon/2 \quad \text{para todo } v \in J \qquad \text{y} \qquad |\nu_J(E_J)| \ > \ \varepsilon.$$

Vamos a repetir indefinidamente, pero adecuadamente, el paso anterior para construir una sucesión de conjuntos medibles que decrece al vacío y con ella producir una contradicción. Veamos esto. Comencemos tomando $\delta = \varepsilon/2$ y una medida arbitraria $\nu_1 \in \mathcal{K}$. Si hacemos $J = \{\nu_1\}$, entonces existe un conjunto $E_1 \in \mathcal{M}$ y una medida $\nu_2 \in \mathcal{K}$ tal que

$$|\nu_1|(E_1) \ < \ \varepsilon/2 \qquad \text{y} \qquad |\nu_2(E_1)| \ > \ \varepsilon.$$

Tome ahora $\delta = \varepsilon/2^2$ y $J = \{\nu_1, \nu_2\}$. Como antes, existe un conjunto $E_2 \in \mathcal{M}$ y una medida $\nu_3 \in \mathcal{K}$ tal que

$$|\nu_1|(E_2) \ < \ \varepsilon/2^2, \quad |\nu_2|(E_2) \ < \ \varepsilon/2^2 \qquad \text{y} \qquad |\nu_3(E_2)| \ > \ \varepsilon.$$

Continuando inductivamente con este procedimiento, se logra obtener una sucesión de conjuntos $(E_n)_{n=1}^{\infty}$ en $\mathcal{M}$ y una sucesión de medidas $(\nu_n)_{n=1}^{\infty}$ en $\mathcal{K}$ verificando lo siguiente: para cada $n \in \mathbb{N}$,

$$|\nu_1|(E_n) \ < \ \varepsilon/2^n, \quad |\nu_2|(E_n) \ < \ \varepsilon/2^n, \ \ldots, \quad |\nu_n|(E_n) \ < \ \varepsilon/2^n \qquad \text{y} \qquad |\nu_{n+1}(E_n)| \ > \ \varepsilon.$$

Sea $F_n = \bigcup_{i=n}^{\infty} E_j$. Entonces, para $j \leq n \leq i$ se tiene que

$$|\nu_j|(F_n) \ \leq \ \sum_{i=n}^{\infty} |\nu_j|(E_i) \ \leq \ \sum_{i=n}^{\infty} \frac{\varepsilon}{2^i} \ = \ \frac{\varepsilon}{2^{n-1}}$$

y

$$\begin{aligned} \left|\nu_{n+1}(F_n)\right| &= \left|\nu_{n+1}(E_n) - \nu_{n+1}(F_n \setminus E_n)\right| \\ &\geq \left|\nu_{n+1}(E_n)\right| - \left|\nu_{n+1}(F_n \setminus E_n)\right| \\ &\geq \varepsilon - \left|\nu_{n+1}(F_n \setminus E_n)\right| \geq \varepsilon - \left|\nu_{n+1}\right|(F_{n+1}) \\ &\geq \varepsilon - \varepsilon/2^n > \varepsilon/2. \end{aligned}$$

Hagamos $F = \bigcap_{n=1}^{\infty} F_n$. Entonces, para $n > j$ tenemos que

$$|\nu_j|(F) \leq |\nu_j|(F_n) \leq \varepsilon/2^{n-1} \to 0 \quad \text{cuando } n \to \infty$$

y, por consiguiente,

$$|\nu_j|(F) = 0 \quad \text{para cada } j \in \mathbb{N}.$$

De la desigualdad $|\nu_j(A)| \leq |\nu_j|(A)$ se sigue que $\nu_j(F) = 0$ para cada $j \in \mathbb{N}$. Por esto,

$$\left|\nu_{n+1}(F_n \setminus F)\right| = \left|\nu_{n+1}(F_n) - \nu_{n+1}(F)\right| = \left|\nu_{n+1}(F_n)\right| > \varepsilon/2. \tag{$**$}$$

Puesto que la sucesión $(F_n \setminus F)_{n=1}^{\infty}$ decrece a $\varnothing$ y $\mathcal{K}$ es uniformemente σ-aditiva, se sigue del Lema 13.2.2 que $\lim_{n\to\infty} \nu(F_n \setminus F) = 0$ uniformemente en $\nu \in \mathcal{K}$. De esto último se sigue que, para el valor $\varepsilon/2$, existe un $N \in \mathbb{N}$ tal que

$$|\nu(F_n \setminus F)| < \varepsilon/2 \quad \text{para todo } n \geq N \text{ y todo } \nu \in \mathcal{K}.$$

Esto, por supuesto, contradice la desigualdad $(**)$ y termina la prueba. ■

Definición 13.2.5. *Sea $\mathcal{K} \subseteq \mathrm{ca}(\mathcal{M},\mathbb{R})$ y sea $\lambda : \mathcal{M} \to [0, +\infty]$ una medida positiva.*

*(a) Diremos que $\mathcal{K}$ es **absolutamente continua respecto a** λ, en notación $\mathcal{K} \ll \lambda$, si $\nu \ll \lambda$ para cada medida $\nu \in \mathcal{K}$.*

*(b) Diremos que $\mathcal{K} \ll \lambda$ **uniformemente** si, para cada $\varepsilon > 0$, existe un $\delta > 0$ tal que para cualquier $E \in \mathcal{M}$ con $\lambda(E) < \delta$ se cumple que*

$$\sup_{\nu \in \mathcal{K}} |\nu(E)| < \varepsilon.$$

Observe que la afirmación: $\mathcal{K} \ll \lambda$ *uniformemente*, significa que

$$\lim_{\lambda(E)\to 0} \nu(E) = 0 \quad \text{uniformemente en } \nu \in \mathcal{K}.$$

La expresión $\mathcal{K}$ es **uniformemente λ-continua** o **uniformemente λ-absolutamente continua** se usa con frecuencia en lugar de $\mathcal{K} \ll \lambda$ uniformemente.

En el siguiente resultado se establece condiciones bajo la cual una sucesión $\mathcal{K}$ de medidas reales satisfaciendo $\mathcal{K} \ll \lambda$ para alguna medida positiva λ, es uniforme.

Lema 13.2.6. *Sea $\mathcal{K} = \{\nu_n : n \in \mathbb{N}\}$ una colección numerable de medidas incluidas en $\mathrm{ca}(\mathcal{M},\mathbb{R})$ y sea λ una **medida positiva** tal que $\mathcal{K} \ll \lambda$. Si $\mathcal{K}$ es **uniformemente σ-aditiva**, entonces $\mathcal{K} \ll \lambda$ **uniformemente**.*

Prueba. Suponga que la conclusión es falsa. Entonces *existe un* $\varepsilon > 0$ *con la siguiente propiedad: para cualquier* $\delta > 0$, *siempre es posible determinar un* $k \in \mathbb{N}$ *y un conjunto* $E \in \mathcal{M}$ *tal que*

$$|\nu_k(E)| \geq \varepsilon \qquad \text{y} \qquad \lambda(E) < \delta.$$

Comenzando con $\delta = 1$, seleccione un entero $n_1 \in \mathbb{N}$ y un conjunto medible E_1 tal que

$$|\nu_{n_1}(E_1)| \geq \varepsilon \qquad \text{y} \qquad \lambda(E_1) < 1.$$

Por otro lado, como $\nu_{n_1} \ll \lambda$, existe un $\delta_1 > 0$ tal que

$$|\nu_{n_1}(E)| < \varepsilon/2 \quad \text{siempre que} \quad \lambda(E) < \delta_1$$

Ahora, con $\delta = \delta_1$, escoja un entero $n_2 > n_1$ y un conjunto medible E_2 tal que

$$|\nu_{n_2}(E_2)| \geq \varepsilon \qquad \text{y} \qquad \lambda(E_2) < \delta_1/2.$$

De nuevo, como $\nu_{n_2} \ll \lambda$, existe un $\delta_2 > 0$ tal que

$$|\nu_{n_2}(E)| < \varepsilon/2 \quad \text{siempre que} \quad \lambda(E) < \delta_2.$$

Continuando inductivamente con esta construcción se logra producir una sucesión de conjuntos $(E_k)_{k=1}^{\infty}$ en $\mathcal{M}$, una sucesión de números positivos $(\delta_k)_{k=1}^{\infty}$ y una sucesión estrictamente creciente de enteros positivos $(n_k)_{k=1}^{\infty}$ tales que

$$|\nu_{n_k}(E_k)| \geq \varepsilon, \quad \lambda(E_k) < \delta_{k-1}/2 \qquad \text{y} \qquad |\nu_{n_k}(E)| < \varepsilon/2 \quad \text{siempre que} \quad \lambda(E) < \delta_k$$

para todo $k \geq 2$. De esto y la subaditividad de λ se sigue que

$$\begin{aligned}\lambda\left(\bigcup_{j=k+1}^{\infty} E_j\right) &\leq \sum_{j=k+1}^{\infty} \lambda(E_j) < \frac{\delta_k}{2} + \frac{\delta_{k+1}}{2} + \cdots \\ &< \frac{\delta_k}{2} + \frac{\delta_k}{2^2} + \cdots = \delta_k\end{aligned}$$

y, entonces, como $\mathcal{K} \ll \lambda$, resulta que

$$\left|\nu_{n_k}\left(E_k \cap \bigcup_{j=k+1}^{\infty} E_j\right)\right| < \frac{\varepsilon}{2}.$$

Tomando $A_k = E_k \setminus \bigcup_{j=k+1}^{\infty} E_j$ para cada $k \in \mathbb{N}$, se tiene que la sucesión $(A_k)_{k=1}^{\infty}$ es disjunta y

$$|\nu_{n_k}(A_k)| \geq |\nu_{n_k}(E_k)| - \left|\nu_{n_k}\left(E_k \cap \bigcup_{j=k+1}^{\infty} E_j\right)\right| \geq \varepsilon - \frac{\varepsilon}{2} = \frac{\varepsilon}{2}.$$

Sin embrago, como $\mathcal{K}$ es uniformemente σ-aditiva, el Teorema 13.2.2 nos asegura que

$$\lim_{k\to\infty} \nu_n(A_k) = 0 \quad \text{uniformemente para } n \in \mathbb{N}$$

lo que constituye una contradicción a la desigualdad anterior. La prueba es completa. ■

La existencia de una medida de control λ para familias de medidas uniformemente σ-aditiva tal que $\mathcal{K} \ll \lambda$ uniformemente es demostrada en el siguiente resultado.

Teorema 13.2.7. *Sea $\mathcal{K} \subseteq \mathrm{ca}(\mathcal{M},\mathbb{R})$. Las siguientes condiciones son equivalentes:*

(1) *$\mathcal{K}$ es **uniformemente σ-aditiva**.*

(2) *Existe una **medida positiva y finita** $\lambda : \mathcal{M} \to [0,+\infty)$ tal que $\mathcal{K} \ll \lambda$ **uniformemente**. Además,*

$$\lambda(E) \;\leq\; \sup_{\nu\in\mathcal{K}}\sup_{F\subseteq E}|\nu(F)| \quad \text{para todo } E \in \mathcal{M}.$$

Prueba. (1) $\Rightarrow$ (2). Suponga que $\mathcal{K}$ es uniformemente σ-aditiva. Para cada $n \in \mathbb{N}$, considere el número positivo $\varepsilon_n = 1/2^n$. Usemos ahora el Lema 13.2.4 para hallar un $\delta_n > 0$ y un conjunto finito de medidas $J_n = \{\nu_1,\ldots,\nu_{k(n)}\}$ en $\mathcal{K}$ tal que: para cualquier conjunto $E \in \mathcal{M}$ con $|\nu_i|(E) < \delta_n$ para $1 \leq i \leq k(n)$ se cumpla que

$$|\nu(E)| \;<\; 1/2^n \quad \text{para todo } \nu \in \mathcal{K}. \tag*{$(*)^3$}$$

Definamos, para cada $n \in \mathbb{N}$, la medida positiva

$$\lambda_n \;=\; \sum_{i=1}^{k(n)} \frac{|\nu_i|}{2^i}.$$

Observe que cada λ_n es finita. La medida λ que nos servirá puede ser definida del modo siguiente:

$$\lambda \;=\; \frac{1}{2}\cdot\sum_{n=1}^{\infty}\frac{\lambda_n}{2^n\big(1+\lambda_n(X)\big)}.$$

Observe que λ es una medida positiva y, por supuesto, finita. Veamos que $\mathcal{K} \ll \lambda$ uniformemente. En efecto, sea $\varepsilon > 0$ y elija un $n \in \mathbb{N}$ adecuadamente de modo que $1/2^n < \varepsilon$. Si tomamos

$$\delta \;=\; \frac{1}{2}\cdot\delta_n\cdot\frac{1}{2^{n+k(n)}}\cdot\frac{1}{1+\lambda_n(X)},$$

resulta que para cualquier conjunto $E \in \mathcal{M}$ para el cual $\lambda(E) < \delta$, se tiene que

$$\frac{1}{2}\cdot\frac{\lambda_n(E)}{2^n\big(1+\lambda_n(X)\big)} \;\leq\; \lambda(E) \;<\; \delta \;=\; \frac{1}{2}\cdot\delta_n\cdot\frac{1}{2^{n+k(n)}}\cdot\frac{1}{1+\lambda_n(X)},$$

de donde se sigue que

$$\lambda_n(E) \;<\; \delta_n\cdot\frac{1}{2^{k(n)}}$$

y, entonces, por la definición de λ_n se obtiene que

$$\frac{|\nu_i|(E)}{2^i} \;<\; \delta_n\cdot\frac{1}{2^{k(n)}} \quad \text{para todo } i \in \{1,2,\ldots,k(n)\}.$$

De esto último se concluye que

$$|\nu_i|(E) \;<\; \delta_n \quad \text{para todo } i \in \{1,2,\ldots,k(n)\}.$$

Un llamado a $(*)^3$ nos revela que $|\nu(E)| < \varepsilon$ para toda $\nu \in \mathcal{K}$. Esto prueba que $\mathcal{K} \ll \lambda$ uniformemente.

Para demostrar la última parte, observe que $|\mathcal{K}| \ll \lambda$ uniformemente gracias al Lema 13.2.3. Fijemos cualquier conjunto $E \in \mathcal{M}$ e invoquemos el Teorema 12.1.16 para obtener

$$\begin{aligned} \lambda_n(E) \ &\leq \ \sup_{\nu \in \mathcal{K}} |\nu|(E) \cdot \sum_{i=1}^{k(n)} \frac{1}{2^i} \\ &\leq \ \sup_{\nu \in \mathcal{K}} |\nu|(E) \ \leq \ 2 \sup_{\nu \in \mathcal{K}} \sup_{F \subseteq E} |\nu(F)|. \end{aligned}$$

Puesto que $\lambda_n/(1+\lambda_n(X)) \leq \lambda_n \leq \sup \lambda_n$, resulta que

$$\begin{aligned} \lambda(E) \ &= \ \frac{1}{2} \cdot \sum_{n=1}^{\infty} \frac{\lambda_n(E)}{2^n\big(1+\lambda_n(X)\big)} \\ &\leq \ \frac{1}{2} \cdot \sup_{n \in \mathbb{N}} \lambda_n(E) \sum_{n=1}^{\infty} \frac{1}{2^n} \\ &\leq \ \sup_{\nu \in \mathcal{K}} \sup_{F \subseteq E} |\nu(F)|. \end{aligned}$$

(2) $\Rightarrow$ (1). Suponga que (2) se cumple y sea $(E_n)_{n=1}^{\infty}$ una sucesión disjunta de conjuntos en $\mathcal{M}$. Entonces

$$\lim_{j \to \infty} \lambda\left(\bigcup_{n=j}^{\infty} \lambda(E_n) \right) \ = \ 0,$$

y como $\mathcal{K} \ll \lambda$ uniformemente, se sigue que

$$0 \ = \ \lim_{j \to \infty} \sup_{\nu \in \mathcal{K}} \left| \nu\left(\bigcup_{n=j}^{\infty} E_n \right) \right| \ = \ \lim_{j \to \infty} \sup_{\nu \in \mathcal{K}} \left| \sum_{n=j}^{\infty} \nu(E_n) \right|$$

Esto nos dice que $\mathcal{K}$ es uniformemente σ-aditiva y termina la prueba. ■

Si $\mathcal{K}$ es una familia uniformemente σ-aditiva de medidas en $\mathrm{ca}(\mathcal{M}, \mathbb{R})$, entonces la medida λ obtenida en resultado anterior se llama una **medida de control** para $\mathcal{K}$. Observe que λ se puede escribir en la forma

$$\lambda \ = \ \sum_{n=1}^{\infty} a_n |\nu_n|,$$

donde $a_n \geq 0$, $\sum_{n=1}^{\infty} a_n \leq 1$ y $(\nu_n)_{n=1}^{\infty}$ está incluida en $\mathcal{K}$. En el caso especial cuando $\mathcal{K}$ consta de una única medida se obtiene un resultado que es fundamental en la Teoría de la Medida.

Corolario 13.2.8 (Bartle-Dunford-Schwartz). *Sea $\nu \in \mathrm{ca}(\mathcal{M}, \mathbb{R})$. Entonces existe una* ***medida positiva y finita*** *$\lambda : \mathcal{M} \to [0, +\infty)$ tal que $\nu \ll \lambda$. Más aun,*

$$\lambda(E) \ \leq \ \sup_{F \subseteq E} |\nu(F)| \quad \textit{para todo } E \in \mathcal{M}$$

13.3. Los Teoremas de Nikodým y el de Vitali-Hahn-Saks

Con los resultados establecidos en las dos secciones anteriores podemos ahora formular y probar tres joyas de la corona: los tres teoremas anunciados en la introducción. El primero es el Teorema de Convergencia de Nikodým el cual establece que *el limite de sucesiones de medidas reales es una medida real*. Además, en este caso, la sucesión es *uniformemente σ-aditiva*.

Teorema 13.3.1 (Convergencia de Nikodým). *Sea $(\nu_n)_{n=1}^{\infty}$ una **sucesión de medidas** en $\operatorname{ca}(\mathcal{M}, \mathbb{R})$. Si*

$$\lim_{n\to\infty} \nu_n(E) \;=\; \nu(E)$$

existe para cada $E \in \mathcal{M}$, entonces:

*(a) la familia $\mathcal{K} = \{\nu_n : n \in \mathbb{N}\}$ es **uniformemente σ-aditiva**, y*

(b) $\nu \in \operatorname{ca}(\mathcal{M}, \mathbb{R})$.

Prueba. (a). Sea $(E_j)_{j=1}^{\infty}$ una sucesión disjunta de conjuntos en $\mathcal{M}$ y considere la matriz infinita de números reales

$$A \;=\; \begin{pmatrix} \nu_1(E_1) & \nu_1(E_2) & \nu_1(E_3) & \cdots & \nu_1(E_j) & \cdots \\ \nu_2(E_1) & \nu_2(E_2) & \nu_2(E_3) & \cdots & \nu_2(E_j) & \cdots \\ \nu_3(E_1) & \nu_3(E_2) & \nu_3(E_3) & \cdots & \nu_3(E_j) & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ \nu_i(E_1) & \nu_i(E_2) & \nu_i(E_3) & \cdots & \nu_i(E_j) & \cdots \\ \vdots & \vdots & \vdots & & \vdots & \ddots \end{pmatrix}.$$

Por hipótesis, las columnas de A convergen y si $(m_j)_{j=1}^{\infty}$ es cualquier sucesión creciente de $\mathbb{N}$, entonces la numerabilidad aditiva de las ν_n y nuestra hipótesis nos revelan que

$$\lim_{n\to\infty} \sum_{j=1}^{\infty} \nu_n(E_{m_j}) \;=\; \lim_{n\to\infty} \nu_n\Big(\bigcup_{j=1}^{\infty} E_{m_j}\Big) \;=\; \nu\Big(\bigcup_{j=1}^{\infty} E_{m_j}\Big)$$

existe. La conclusión es simple: la matriz A satisface las condiciones (1) y (2) del Teorema de Antosik-Mikusiński, página 835 y, en consecuencia, $\lim_{j\to\infty} \nu_n(E_j) = 0$ converge uniformemente para $n \in \mathbb{N}$. Un llamado al Lema 13.2.2 nos muestra que la familia $\mathcal{K} = \{\nu_n : n \in \mathbb{N}\}$ es uniformemente σ-aditiva.

(b) Es claro que ν es finitamente aditiva. Sea $(E_n)_{n=1}^{\infty}$ una sucesión decreciente de conjuntos en $\mathcal{M}$ tal que $\bigcap_{n=1}^{\infty} E_n = \varnothing$. Por la parte (a) sabemos que la familia $\mathcal{K} = \{\nu_n : n \in \mathbb{N}\}$ es uniformemente σ-aditiva y, así, por una nueva aplicación del Lema 13.2.2 se tiene que $\lim_{j\to\infty} \nu_n(E_j) = 0$ converge uniformemente para $n \in \mathbb{N}$. De esto se sigue que

$$\lim_{j\to\infty} \nu(E_j) \;=\; \lim_{j\to\infty} \lim_{n\to\infty} \nu_n(E_j) \;=\; \lim_{n\to\infty} \lim_{j\to\infty} \nu_n(E_j) \;=\; 0$$

y, por lo tanto, gracias al Teorema 12.1.5 se concluye que $\nu \in \operatorname{ca}(\mathcal{M}, \mathbb{R})$. Fin de la prueba. ■

Es importante estar en conocimiento de que el Teorema de Convergencia de Nikodým no se cumple, en general, para álgebras. Un ejemplo de esta situación se puede ver en el artículo de P. de Lucia y E. Pap titulado "*Convergence Theorems for Set Functions*" en [106], pág. 125-178.

Nuestro próximo resultado, el Teorema de Acotación de Nikodým, requiere de la siguiente definición.

Definición 13.3.2. *Una familia de medidas $\mathcal{K} \subseteq \mathrm{ca}(\mathcal{M}, \mathbb{R})$ se dice que está **puntualmente acotada** si, para cada $E \in \mathcal{M}$, existe una constante $M_E > 0$ tal que*

$$\sup_{\nu \in \mathcal{K}} |\nu(E)| \leq M_E.$$

*La familia $\mathcal{K}$ se dice que está **acotada en la norma** de $\mathrm{ca}(\mathcal{M}, \mathbb{R})$, o que es **uniformemente acotada**, si*

$$\sup_{\nu \in \mathcal{K}} \| \nu \| = \sup_{\nu \in \mathcal{K}} |\nu|(X) < +\infty.$$

Teorema 13.3.3 (Acotamiento de Nikodým). *Sea $\mathcal{K}$ una familia de medidas en $\mathrm{ca}(\mathcal{M}, \mathbb{R})$ **puntualmente acotada**, es decir,*

$$\sup_{\nu \in \mathcal{K}} |\nu(E)| < +\infty \tag{α_1}$$

*para cada $E \in \mathcal{M}$. Entonces $\mathcal{K}$ es **uniformemente acotada**.*

Prueba. Para generar una contradicción, suponga que $\mathcal{K}$ no es uniformemente acotada, es decir,

$$\sup_{\nu \in \mathcal{K}} |\nu|(X) = +\infty.$$

Se sigue del Teorema 12.1.16 que

$$\sup \big\{ |\nu(E)| \, : \, \nu \in \mathcal{K}, E \in \mathcal{M} \big\} = +\infty$$

y entonces seleccione una sucesión $(\nu_n)_{n=1}^{\infty}$ en $\mathcal{K}$ tal que

$$\sup \big\{ |\nu_n(E)| \, : \, n \in \mathbb{N}, E \in \mathcal{M} \big\} = +\infty. \tag{α_2}$$

Afirmamos que: *para cada $\rho > 0$, se puede elegir un $n \in \mathbb{N}$ y una partición medible $\{E, F\}$ de X tal que*

$$\min\{|\nu_n(E)|, |\nu_n(F)|\} > \rho.$$

En efecto, por (α_1) sabemos que $\sup_{k \in \mathbb{N}} |\nu_k(X)| < +\infty$. Usemos ahora (α_2) para elegir un $n \in \mathbb{N}$ y un $E \in \mathcal{M}$ de modo tal que

$$\nu_n(E) > \sup_{k \in \mathbb{N}} |\nu_k(X)| + \rho.$$

Tomando $F = X \setminus E$, se verifica que

$$\begin{aligned} |\nu_n(F)| = |\nu_n(X \setminus E)| &= |\nu_n(X) - \nu_n(E)| \\ &\geq |\nu_n(X)| - |\nu_n(E)| > \rho. \end{aligned}$$

Esto termina la prueba de la afirmación. Vamos a aplicar reiteradamente nuestra afirmación para obtener una sucesión disjunta $(G_k)_{k=1}^{\infty}$ de miembros de $\mathcal{M}$ y una sucesión estrictamente creciente $(n_k)_{k=1}^{\infty}$ de enteros positivos tal que para cada $k > 1$

$$\big|\nu_{n_k}(G_k)\big| > \sum_{j=1}^{k-1} \big|\nu_{n_k}(G_j)\big| + k + 1. \tag{α_2}$$

En efecto, comenzando con $\rho = 1$, elija el primer entero positivo n_1 para el cual existe una partición medible $\{E, F\}$ de X tal que

$$\text{mín}\{|\nu_n(E)|, |\nu_n(F)|\} > 1 + 1.$$

Observe que una de las cantidades

$$\sup\{|\nu_n(E \cap B)| : n \in \mathbb{N}, B \in \mathcal{M}\} \qquad \text{o} \qquad \sup\{|\nu_n(F \cap B)| : n \in \mathbb{N}, B \in \mathcal{M}\}$$

es infinita. Si es la primera, hagamos $B_1 = E$ y $G_1 = F$; en otro caso, sea $B_1 = F$ y $G_1 = E$. Trabajemos ahora con B_1 y con $\rho = 2$, escoja el entero más pequeño $n_2 > n_1$ para el cual existe una partición medible $\{E, F\}$ de B_1 tal que

$$\text{mín}\{|\nu_{n_2}(E)|, |\nu_{n_2}(F)|\} > |\nu_{n_2}(G_1)| + 3.$$

De nuevo, una de las dos cantidades

$$\sup\{|\nu_n(E \cap B)| : n \in \mathbb{N}, B \in \mathcal{M}\} \qquad \text{o} \qquad \sup\{|\nu_n(F \cap B)| : n \in \mathbb{N}, B \in \mathcal{M}\}$$

es infinita. Si es la primera, hacemos $B_2 = E$ y $G_2 = F$; en otro caso, sea $B_2 = F$ y $G_2 = E$. Continuando indefinidamente con éste procedimiento, se obtienen las sucesiones disjunta $(G_k)_{k=1}^{\infty}$ y $(n_k)_{k=1}^{\infty}$ que satisfacen (α_3).

Para evitar el uso engorroso de los subíndices, pondremos $\nu_{n_k} = \nu_k$ para todo $k \in \mathbb{N}$ lo que implica que (α_2) se puede escribir en la forma:

$$\big|\nu_k(G_k)\big| > \sum_{j=1}^{k-1} \big|\nu_k(G_j)\big| + k + 1 \quad \text{para todo } k \geq 2.$$

Nuestro siguiente paso consistirá en construir una subsucesión de $(G_k)_{k=1}^{\infty}$ que nos conduzca a la contradicción que buscamos. Sea $n_1 = 1$. Invoquemos el Teorema 1.2.12, página 17, para obtener una partición infinita $(N_k)_{k=1}^{\infty}$ de $\mathbb{N} \setminus \{n_1\}$ donde cada N_k es infinito. La numerabilidad aditiva de $|\nu_{n_1}|$ nos dice que:

$$\sum_{k=1}^{\infty} |\nu_{n_1}|\Big(\bigcup_{j \in N_k} G_j\Big) = |\nu_{n_1}|\Big(\bigcup_{i=1}^{\infty} \bigcup_{j \in N_i} G_j\Big) \leq |\nu_{n_1}|(X). \tag{α_4}$$

Recordemos ahora que si una serie $\sum_{n=1}^{\infty} a_n$ converge, entonces $\text{lím}_{n\to\infty} a_n = 0$ y, por consiguiente, existe al menos un $n_0 \in \mathbb{N}$ tal que $|a_n| < 1$. Pues bien, usemos éste hecho aplicándolo a la serie (α_4) para obtener algún N_k, el cual denotaremos por $\widehat{N}_1$, tal que

$$|\nu_{n_1}|\Big(\bigcup_{j \in \widehat{N}_1} G_j\Big) < 1.$$

Pongamos $n_2 = \text{mín}\, \widehat{N}_1$ y observe que $n_2 > n_1$. Repitamos el argumento anterior con $\mathbb{N} \setminus \{n_1\}$ reemplazado por $\widehat{N}_1 \setminus \{n_2\}$ y $|\nu_{n_1}|$ por $|\nu_{n_2}|$ para obtener un conjunto infinito $\widehat{N}_2 \subseteq \widehat{N}_1 \setminus \{n_2\}$ tal que

$$|\nu_{n_2}|\Big(\bigcup_{j \in \widehat{N}_2} G_j\Big) < 1.$$

Defina $n_3 = \min \widehat{N}_2$. Continuando con éste argumento indefinidamente se obtiene una sucesión estrictamente creciente $(n_k)_{k=1}^{\infty}$ de enteros positivos y una sucesión decreciente $(\widehat{N}_k)_{k=1}^{\infty}$ de subconjuntos infinitos de $\mathbb{N}$ tal que

$$|\nu_{n_k}|\Big(\bigcup_{j\in\widehat{N}_k} G_j\Big) \ < \ 1.$$

para todo $k \geq 1$. De nuevo, renombre G_{n_k} por G_k y ν_{n_k} por ν_k para todo $k \in \mathbb{N}$ y defina

$$E \ = \ \bigcup_{k=1}^{\infty} G_k.$$

Este conjunto E es el adecuado para producir la contradicción buscada. En efecto, para cada $k \geq 2$

$$\begin{aligned}
\big|\nu_k(E)\big| \ &\geq \ \big|\nu_k(G_k)\big| \ - \ \sum_{j=1}^{k-1}\big|\nu_k(G_j)\big| \ - \ \Big|\nu_k\Big(\bigcup_{j=k+1}^{\infty} G_j\Big)\Big| \\
&\geq \ k+1 \ - \ |\nu_k|\Big(\bigcup_{j=k+1}^{\infty} G_j\Big) \ \geq \ k+1-1 \ = \ k
\end{aligned}$$

y entonces,

$$\sup_{k\geq 2}\big|\nu_k(E)\big| \ = \ +\infty$$

lo que, por supuesto, contradice la hipótesis de que la colección $\{\nu_k : k \geq 2\}$ está puntualmente acotada. La prueba es completa. ■

Finalmente presentamos el Teorema de Vitali-Hahn-Saks cuya demostración, en estas notas, se basa en el Teorema de Convergencia de Nikodým y el Lema 13.2.6.

Teorema 13.3.4 (Vitali-Hahn-Saks). *Sea $(\nu_n)_{n=1}^{\infty}$ una* ***sucesión de medidas*** *en $\mathrm{ca}(\mathcal{M},\mathbb{R})$ y sea λ una* ***medida positiva*** *tal que $\nu_n \ll \lambda$ para cada $n \in \mathbb{N}$. Si*

$$\lim_{n\to\infty} \nu_n(E) \ = \ \nu(E)$$

para cada $E \in \mathcal{M}$, entonces:

(a) la familia $\mathcal{K} = \{\nu_n : n \in \mathbb{N}\} \ll \lambda$ ***uniformemente****, y*

(b) $\nu \ll \lambda$.

Prueba. (a) Por el Teorema de Convergencia de Nikodým se tiene que $\mathcal{K} = \{\nu_n : n \in \mathbb{N}\}$ es uniformemente σ-aditiva y como por hipótesis $\mathcal{K} \ll \lambda$, entonces el Lema 13.2.6 nos revela que $\mathcal{K} \ll \lambda$ uniformemente.

Para verificar (b) observe que el Teorema de Convergencia de Nikodým es, de nuevo, el responsable de garantizarnos que $\nu \in \mathrm{ca}(\mathcal{M},\mathbb{R})$. Sea $E \in \mathcal{M}$ tal que $\lambda(E) = 0$. Se sigue de (a) que $\nu_n(E) = 0$ para todo $n \geq 1$ y, por lo tanto, $\nu(E) = \lim_{n\to\infty} \nu_n(E) = 0$. Fin de la prueba. ■

Nota Adicional 13.3.1 El Teorema de Radon-Nikodým, los dos teorema de Nikodým más el Teorema de Vitali-Hahn-Saks constituyen algunas de las joyas más preciadas de la Teoría de la Medida e Integración. La evolución de esos cuatro resultados refleja, en gran medida, el desarrollo actual de dicha teoría.

En 1907 Vitali demostró si $(f_n)_{n=1}^{\infty}$ es una sucesión de funciones en $\mathrm{L}_1(\mu)$ tal que $f_n \to f$ puntualmente, entonces

$$\lim_{n\to\infty} \int_E f_n \, d\mu \;=\; \int_E f \, d\mu \quad \text{para cualquier } E \in \mathfrak{M}_\mu(\mathbb{R})$$

si, y sólo si, la familia de medidas $\{\lambda_n : n \in \mathbb{N}\}$ es uniformemente μ-continua, donde $\lambda_n(E) = \int_E f_n \, d\mu$ para todo $n \in \mathbb{N}$. En 1909 Lebesgue debilita la hipótesis de Vitali demostrando que si $(f_n)_{n=1}^{\infty}$ es una sucesión en $\mathrm{L}_1(\mu)$ tal que

$$\lim_{n\to\infty} \int_E f_n \, d\mu \;=\; 0 \quad \text{para cualquier } E \in \mathfrak{M}_\mu(\mathbb{R})$$

entonces $\{\lambda_n : n \in \mathbb{N}\} \ll \mu$ uniformemente. Posteriormente, Hahn mejora el resultado de Lebesgue exigiendo sólo que

$$\lim_{n\to\infty} \int_E f_n \, d\mu \quad \text{exista para cualquier } E \in \mathfrak{M}_\mu(\mathbb{R}).$$

En 1931, Nikodým extiende estos resultados en un contexto abstracto. Es importante señalar que el Teorema de Vitali-Hahn-Saks fue demostrado en 1933 por Saks usando el Método de Categoría de Baire, lo que provocó una profunda revisión de dicho método. Con el correr de los años, el Teorema de Categoría de Baire se convirtió en una herramienta fundamental y de primer orden para demostrar, de modo simple, muchísimos resultados importantes y profundos en matemáticas [23].

CAPÍTULO 14

La Integral de Henstock-Kurzweil

En el Capítulo 8, página 405, introducimos brevemente la integral de Riemann y vimos que ella era capaz de integrar cualquier función continua $f : [a, b] \to \mathbb{R}$, así como también cualquier función discontinua $f : [a, b] \to \mathbb{R}$ cuya discontinuidades formaban un conjunto de medida cero (Teorema de Vitali-Lebesgue). Sin embrago, a pesar de que dicha integral posee una definición sencilla y elegante y ser, además, la más popular en los cursos de Cálculo, ella posee, como pudimos observar, algunas severas deficiencias teóricas lo que impide que los matemáticos la puedan consideren como su integral predilecta. La integral de Lebesgue, en cambio, ha sido considerada desde sus inicios como la integral preferida por los matemáticos a pesar de que su construcción es altamente sofisticada: hay que desarrollar gran parte de la Teoría de la Medida y de las Funciones Medibles para llegar a ella. Esto, por supuesto, requiere que el estudiante posea algunos sólidos conocimientos del Análisis para entenderla. Sin embargo, al igual que la integral de Riemann, ella no es capaz de integrar a todas las derivadas. De hecho, en su tesis de 1902, Lebesgue hace notar, un resultado que ya hemos verificado, que su integral no puede integrar todas las derivadas no-acotadas y deja abierto el problema de definir una integral más general que resuelva el problema de las primitivas, es decir, la validez de la fórmula de Newton-Leibniz para cualquier derivada. Afortunadamente, dicho problema fue resuelto una década después. En efecto, en 1912, el matemático francés, Arnaud Denjoy (1884-1974) construyó una nueva integral donde el Teorema Fundamental de Cálculo se cumple en toda su generalidad, es decir, todas las derivadas son integrables. Posteriormente, en 1914, el matemático alemán, Oskar Perron (1880-1975) obtiene otra teoría de integración donde, al igual que la integral de Denjoy, el Teorema Fundamental de Cálculo se cumple sin restricciones. Aunque la definición de Perron es muy diferente a la de Denjoy, resulta que ambas integrales son equivalentes como fue demostrado, casi diez años más tarde, por Alexandrov y Looman (véase, por ejemplo, [66], Capítulo 11, o [109], Sección 9.3). Un extraordinario avance en la dirección establecida por Denjoy-Perron se da entre 1957-1961 cuando los matemáticos Jaroslav Kurzweil (nacido el 7 de Mayo en Praga en al año 1926) y el inglés, Ralph Henstock (1923-2007) independientemente, construyen una nueva integral tipo-Riemann que también posee el privilegio de la validez del Teorema Fundamental de Cálculo en toda su generalidad y que, además, resulta ser equivalente a las dos anteriores.

Esta integral es la que llamaremos, en estas notas, la integral de Henstock-Kurzweil y que será desarrollada brevemente en este capítulo. La integral de Henstock-Kurzweil tiene la virtud de ser, en muchos aspectos, tanto o más poderosa que la integral de Lebesgue en varios aspectos. De hecho, el conjunto $\mathcal{HK}([a,b])$ de las funciones $f : [a,b] \to \mathbb{R}$ que son integrables en el sentido de Henstock-Kurzweil incluye a todas las funciones que son integrables según Riemann, según Newton y también según Lebesgue. Sin embargo, la integral de Henstock-Kurzweil no es una integral absoluta, vale decir, la integrabilidad de una función f en el sentido de Henstock-Kurzweil no implica la integrabilidad de $|f|$. Este hecho, por supuesto, no permite desarrollar, en toda su generalidad, una teoría $\mathcal{HK}_p$ similar a la de los espacios L_p y, por supuesto, no tampoco permite la construcción de una norma completa sobre $\mathcal{HK}([a,b])$ similar a la norma $\|\cdot\|_1$.

El siguiente aspecto es, tal vez, el más importante en relación a la integral de Henstock-Kurzweil. En ciertas ramas del Análisis, como por ejemplo, el Análisis Funcional, la Teoría de Ecuaciones Diferenciales, etc. es de mayor interés tener a la mano un concepto de integral que no se desarrolle a partir de una Teoría de la Medida previamente desarrollada. Por ejemplo, la integral de Daniell, creada por P. J. Daniell en el año 1916 y posteriormente profundizada por M. H. Stone en los años siguientes, es una integral equivalente a la de Lebesgue que no requiere, en lo absoluto, ninguna teoría de la medida para su construcción. En esta sección veremos que la integral de Henstock-Kurzweil tampoco necesita de la noción de medida para su existencia, aunque la Teoría de la Medida de Lebesgue puede ser de gran ayuda en ciertos aspectos de su desarrollo. Otra integral que también comparte este enfoque es la integral de McShane. Estas ideas serán expuestas en lo que sigue.

14.1. Construcción de la Integral de Henstock-Kurzweil

Sea $[a,b]$ un intervalo cerrado y acotado de $\mathbb{R}$. Recordemos que una partición de $[a,b]$ es una colección finita de puntos $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ de $[a,b]$ tal que

$$a = x_0 < x_1 < x_2 < \cdots < x_n = b.$$

Los intervalos asociados a la partición $\mathcal{P}$ serán denotados por $I_1 = [x_0, x_1], \ldots, I_n = [x_{n-1}, x_n]$. Un conjunto de **etiquetas** asociadas a los intervalos $I_1, \ldots, I_n$ de la partición $\mathcal{P}$ es cualquier colección finita de puntos de $[a,b]$, digamos $\{t_1, t_2, \ldots, t_n\}$ tal que $t_i \in [x_{i-1}, x_i]$ para $i = 1, \ldots, n$. Cada t_i es llamada la **etiqueta**, o **marca**, asociada con I_i. Cualquier conjunto de pares ordenados

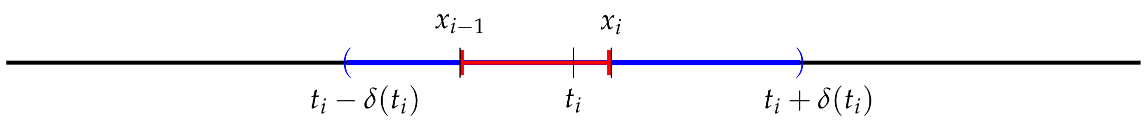
$$\mathcal{P}_e = \big\{\big(t_i, [x_{i-1}, x_i]\big) : i = 1, \ldots, n\big\},$$

donde $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ es una partición de $[a,b]$ y $e = \{t_1, t_2, \ldots, t_n\}$ es un conjunto de etiquetas asociadas a los intervalos $I_1, \ldots, I_n$ de la partición $\mathcal{P}$, lo llamaremos una **partición etiquetada** de $[a,b]$.

También recordemos, Lema de Cousin, Lema 8.3.11, página 419, que si $\delta : [a,b] \to \mathbb{R}$ es un **calibrador** sobre $[a,b]$, es decir, una función **estrictamente positiva**, entonces existe una partición etiquetada en $[a,b]$, digamos $\mathcal{P} = (t_i, [x_{i-1}, x_i])_{i=1}^n$, que es δ-**fina**, o **subordinada** a δ, es decir, tal que que

$$t_i \in [x_{i-1}, x_i] \subseteq \big(t_i - \delta(t_i), t_i + \delta(t_i)\big) = \gamma(t_i),$$

para cada índice $i \in \{1, 2, \ldots, n\}$, donde γ es la función que controla a $\mathcal{P}_e$.



Denotaremos por $\mathfrak{Pe}([a,b],\delta)$ *el conjunto de todas las particiones etiquetadas de* $[a,b]$ *que están subordinadas a* δ. Como antes, si $f:[a,b]\to\mathbb{R}$ es una función arbitraria y $\mathcal{P}=(t_i,[x_{i-1},x_i])_{i=1}^n$ es una partición etiquetada de $[a,b]$, entonces la suma de Riemann asociada a $\mathcal{P}$ la escribiremos como:

$$S(f,\mathcal{P}) \;=\; \sum_{i=1}^{n} f(t_i)(x_i - x_{i-1}).$$

Si la partición etiquetada $\mathcal{P}$ no está descrita explícitamente, entonces escribiremos

$$S(f,\mathcal{P}) \;=\; \sum_{(t,[x,y])\,\in\,\mathcal{P}} f(t)(y-x).$$

Las siguientes observaciones serán importantes en lo que sigue:

Observación 1. *Sea* δ *un calibrador sobre* $[a,b]$ *y fijemos un número* $a<c<b$. *Si* $\mathcal{P}$ *es una partición etiquetada de* $[a,c]$ *que es* δ*-fina y* $\mathcal{Q}$ *es otra partición etiquetada de* $[c,b]$ *que también es* δ*-fina, entonces* $\mathcal{P}\cup\mathcal{Q}$ *es una partición etiquetada de* $[a,b]$ *que es* δ*-fina.*

Observación 2. *Sean* δ_1 *y* δ_2 *calibradores sobre* $[a,b]$ *y suponga que* $\delta_1(t)\le\delta_2(t)$ *para cada* $t\in[a,b]$. *Si* $\mathcal{P}=((t_i,[x_{i-1},x_i]))_{i=1}^n$ *es una partición etiquetada de* $[a,b]$ *que es* δ_1*-fina, entonces también será* δ_2*-fina.* En efecto, basta observar que como

$$t_i \,-\, \delta_2(t_i) \;<\; t_i \,-\, \delta_1(t_i) \qquad\quad y \qquad\quad t_i \,+\, \delta_1(t_i) \;<\; t_i \,+\, \delta_2(t_i),$$

entonces

$$[x_{i-1},x_i] \;\subseteq\; (t_i-\delta_1(t_i),\, t_i+\delta_1(t_i)) \;\subseteq\; (t_i-\delta_2(t_i),\, t_i+\delta_2(t_i)).$$

Por consiguiente, si δ_1 *y* δ_2 *son calibradores sobre* $[a,b]$ *y si se define*

$$\delta(t) \;=\; \text{mín}\{\delta_1(t),\, \delta_2(t)\} \quad \textit{para todo } t\in[a,b],$$

entonces cualquier partición etiquetada de $[a,b]$ *que es* δ*-fina es a la vez* δ_1*-fina y* δ_2*-fina.*

Observación 3. Un procedimiento que es utilizado muchas veces cuando se trabaja con particiones etiquetadas δ-finas es el siguiente: *siempre podemos, cuando sea conveniente, asumir que cada etiqueta, en cualquier partición etiquetada de* $[a,b]$ *subordinada a* δ*, es un* ***extremo*** *de cada uno de los intervalos asociados a dicha partición*. Veamos por qué esto es así:

Lema 14.1.1. *Si* $\delta:[a,b]\to\mathbb{R}$ *es un* ***calibrador*** *sobre* $[a,b]$ *y* $\mathcal{P}=((t_i,[x_{i-1},x_i]))_{i=1}^n$ *es una partición etiquetada de* $[a,b]$ *subordinada a* δ*, entonces siempre se puede construir una nueva partición etiquetada de* $[a,b]$ *subordinada a* δ*, digamos* $\mathcal{Q}=((s_j,[u_{j-1},u_j]))_{j=1}^m$*, de modo tal que cada* $t_i=s_j$ *para algún* $j\in\{1,\dots,m\}$ *y los* $s_j\in\{u_{j-1},u_j\}$ *para todo* $j=1,\dots,m$.

Prueba. Sea $t \in (a,b)$ y asuma que $(t,[a,b])$ es una partición etiquetada de $[a,b]$ subordinada a δ. Entonces $[a,b] \subseteq (t-\delta(t), t+\delta(t))$ y puesto que $[a,t]$ y $[t,b]$ están incluidos en $[a,b]$, resulta que $\{(t,[a,t]),\,(t,[t,b])\}$ es una partición etiquetada de $[a,b]$ subordinada a δ. Para demostrar el caso general, sea $\mathcal{P} = ((t_i,[x_{i-1},x_i]))_{i=1}^n$ una partición etiquetada de $[a,b]$ subordinada a δ y defina

$$E_1 \;=\; \{i \in \{1,\ldots,n\} \,:\, t_i \in (x_{i-1},x_i)\} \qquad \text{y} \qquad E_2 \;=\; \{1,\ldots,n\} \setminus E_1.$$

Observe que si $j \in E_2$, entonces t_j es un extremo del intervalo correspondiente $[x_{j-1},x_j]$. De esto se sigue que

$$\mathcal{Q} \;=\; \Big\{(t_i,[x_{i-1},t_i]),\,(t_i,[t_i,x_i]) \,:\, i \in E_1\Big\} \cup \Big\{(t_j,[x_{j-1},x_j]) \,:\, j \in E_2\Big\}$$

es partición etiquetada de $[a,b]$ subordinada a δ, donde todas las etiquetas de $\mathcal{P}$ son extremos de la partición $\mathcal{Q}$. ■

Un simple, pero profundo refinamiento de la integral de Riemann, lo constituye la siguiente definición.

Definición 14.1.2. *Una función $f : [a,b] \to \mathbb{R}$ es* ***Henstock-Kurzweil integrable*** *sobre $[a,b]$, si existe un número $A \in \mathbb{R}$ con la siguiente propiedad: para cada $\varepsilon > 0$, existe un calibrador $\delta : [a,b] \to (0,+\infty)$ tal que si $\mathcal{P}_e = (t_i,[x_{i-1},x_i])_{i=1}^n$ es cualquier partición δ-fina de $[a,b]$, entonces*

$$\big|S(f,\mathcal{P}_e) - A\big| \;<\; \varepsilon.$$

El número A de la definición anterior, si existe, es único y se llama la **integral de Henstock-Kurzweil** de f el cual será representado por el símbolo

$$A \;=\; (\mathrm{HK})\int_a^b f(t)\,dt.$$

Denotaremos por $\mathcal{HK}([a,b])$ *el conjunto de todas las funciones $f : [a,b] \to \mathbb{R}$ que son* ***Henstock-Kurzweil integrables*** *sobre* $[a,b]$. En general, si $E \subseteq [a,b]$ es un conjunto medible, diremos que f es **Henstock-Kurzweil integrable sobre** E si $f \cdot \chi_E$ es Henstock-Kurzweil integrable sobre $[a,b]$ y escribiremos

$$(\mathrm{HK})\int_E f(t)\,dt \;=\; (\mathrm{HK})\int_a^b f \cdot \chi_E(t)\,dt.$$

La integral de Henstock-Kurzweil también es conocida como la integral Generalizada de Riemann, integral Completa de Riemann, integral de Henstock, integral de Kurzweil-Henstock, integral Gauge, integral Norma, y también como integral de Denjoy-Perron-Kurzweil-Henstock. Percibir la profundidad de la integral formulada por Riemann con respecto a la integral de Cauchy no fue, en principio, una tarea inmediata. Tardó un poco. Modificar la elección de las etiquetas en cada partición del intervalo $[a,b]$ eligiéndolas de modo arbitrario tal como lo hizo Riemann no parecía ser gran cosa, pero en realidad, sí lo era. Mientras que la integral de Cauchy se limitaba sólo a integrar funciones continuas, la integral de Riemann, con ese pequeño cambió, fue capaz de integrar tanto a las funciones continuas, así como a una gran cantidad de funciones discontinuas, incluso con un conjunto denso de discontinuidades. Un sentimiento similar ocurrió con la integral de Henstock-Kurzweil. Es difícil imaginar cómo, con sólo modificar particiones etiquetadas subordinadas a una constante por particiones etiquetadas subordinadas a una variable, se puede generar una integral tan poderosa que, incluso, es capaz de integrar cualquier función que es Lebesgue integrable y, además, satisfacer la idílica fórmula de Newton-Leibniz. Tal vez un poco de practica con las funciones gauges ayude al lector a familiarizarse con esta nueva integral.

Nota Adicional 14.1.1 Revelamos dos hechos importantes referentes a la definición de la integral de Henstock-Kurzweil. El primero: contrario a lo que ocurre con la integral de Riemann, la integral de Henstock-Kurzweil se puede extender a cualquier subintervalo cerrado de longitud infinita incluido en $\mathbb{R}$. De hecho, se puede extender a cualquier rectángulo cerrado en $\mathbb{R}^n$, (véase, por ejemplo, [90]).

Segundo: aunque no se demanda ninguna condición adicional al calibrador en la definición de la integral de Henstock-Kurzweil, se puede demostrar, sin embargo, que:

Si $f \in \mathcal{HK}([a,b])$, entonces para cada $\varepsilon > 0$, existe un ***calibrador medible*** *δ sobre $[a,b]$ tal que*

$$\left| S(f,\mathcal{P}) - (\mathrm{HK})\int_a^b f(t)\,dt \right| < \varepsilon.$$

para cualquier partición etiquetada $\mathcal{P}$ de $[a,b]$ subordinada a δ (véase [66], Theorem 9.24, p.153).

Puede ser de interés al lector estar al corriente que la integral de Lebesgue también se puede aproximar por sumas de Riemann. En efecto, el siguiente resultado fue probado por H. Lebesgue en [89]:

Teorema 14.1.3 (Lebesgue). *Sea $f \in \mathrm{L}_1([a,b])$. Entonces existe una red consistente de una sucesión de particiones etiquetadas $(e_n, \mathcal{P}_n)_{n=1}^{\infty}$, donde cada $e_n = (t_i^n)_{i=1}^{i_n}$ son las etiquetas asociadas a la partición $\mathcal{P}_n = \{x_0, \ldots, x_{i_n}\}$ con $\|\mathcal{P}_n\| \to 0$ tal que*

$$\lim_{n\to\infty} \left| S(f,\mathcal{P}_n) - (\mathrm{L})\int_a^b f\,d\mu \right| = 0$$

De hecho, Lebesgue demostró en ese artículo una afirmación más fuerte: *existe una red como en el resultado anterior que trabaja para cualquier colección finita o infinita numerable de funciones en* $\mathrm{L}_1([a,b])$. Debemos advertir, sin embargo, que el Teorema 14.1.3 no puede ser considerado como una nueva definición descriptiva de la integral de Lebesgue debido al hecho de que, para construir la red de particiones etiquetadas, se requiere la integral de Lebesgue (véase, [109], p. 103-105).

Nuestro primer resultado consistirá en verificar que, efectivamente, cada función f que es Henstock-Kurzweil integrable posee una única integral.

Teorema 14.1.4. *Si $f \in \mathcal{HK}([a,b])$, entonces $(\mathrm{HK})\int_a^b f(t)\,dt$ es única.*

Prueba. Suponga que A_1 y A_2 son números reales que satisfacen la Definición 14.1.2. Entonces, para cada $\varepsilon > 0$, existe un calibrador $\delta_1 : [a,b] \to \mathbb{R}$ tal que $|S(f,\mathcal{P}') - A_1| < \varepsilon/2$ para cualquier partición etiquetada $\mathcal{P}'$ de $[a,b]$ subordinada a δ_1. Similarmente, existe un calibrador $\delta_2 : [a,b] \to \mathbb{R}$ tal que $|S(f,\mathcal{P}'') - A_2| < \varepsilon/2$ para cualquier partición etiquetada $\mathcal{P}''$ de $[a,b]$ subordinada a δ_2. Sea $\delta = \text{mín}\{\delta_1, \delta_2\}$ y sea $\mathcal{P}$ una partición etiquetada de $[a,b]$ subordinada a δ. Entonces $\mathcal{P}$ está subordinada tanto a δ_1 así como a δ_2 y, en consecuencia,

$$|A_1 - A_2| \;\leq\; |S(f,\mathcal{P}) - A_1| + |S(f,\mathcal{P}) - A_2| \;<\; \varepsilon.$$

Puesto que $\varepsilon > 0$ es arbitrario, se obtiene que $A_1 = A_2$. ■

Similar a como ocurre con la integral de Riemann, la integral de Henstock-Kurzweil también posee un criterio á la Cauchy; en otras palabras:

Teorema 14.1.5 (Criterio de Cauchy). *Sea $f : [a,b] \to \mathbb{R}$ una función. Las siguientes condiciones son equivalentes:*

(1) $f \in \mathcal{HK}([a,b])$.

(2) *Para cada $\varepsilon > 0$, existe un **calibrador** $\delta : [a,b] \to \mathbb{R}$ tal que*

$$\mathcal{P},\, \mathcal{Q} \in \mathfrak{Pe}([a,b],\delta) \qquad \Rightarrow \qquad \big|S(f,\mathcal{P}) \,-\, S(f,\mathcal{Q})\big| \;<\; \varepsilon \tag{a}$$

Prueba. Claramente (1) implica (2). Suponga que (2) se cumple. Por cada $n \geq 1$, seleccione (tomando $\varepsilon = 1/n$) un calibrador $\delta_n : [a,b] \to \mathbb{R}$ tal que (a). Reemplazando δ_n por $\text{mín}\{\delta_1, \ldots, \delta_n\}$ si fuese necesario, asumiremos que $\delta_1 > \delta_2 > \cdots$. Usemos ahora el Lema de Cousin para obtener, por cada entero $n \geq 1$, una partición etiquetada $\mathcal{D}_n$ de $[a,b]$ subordinada a δ_n. Observe que para cada $m > n$, puesto que $\delta_n > \delta_m$, resulta que $\mathcal{D}_m$ es una partición etiquetada de $[a,b]$ subordinada a δ_n. De esto se sigue que

$$\big|S(f,\mathcal{D}_m) \,-\, S(f,\mathcal{D}_n)\big| \;<\; \frac{1}{m},$$

lo cual implica que $(S(f,\mathcal{D}_n))_{n=1}^{\infty}$ es una sucesión de Cauchy en $\mathbb{R}$ y, en consecuencia, converge a algún $A \in \mathbb{R}$. Se sigue de la desigualdad anterior que

$$\big|S(f,\mathcal{D}_m) \,-\, A\big| \;<\; \frac{1}{m},$$

Fijemos $\varepsilon > 0$ y escoja un entero $m > 2/\varepsilon$. Si $\mathcal{D}$ es cualquier partición etiquetada de $[a,b]$ subordinada a $\delta = \delta_m$, entonces

$$\big|S(f,\mathcal{D}) \,-\, A\big| \;\leq\; \big|S(f,\mathcal{D}) \,-\, S(f,\mathcal{D}_m)\big| \;+\; \big|S(f,\mathcal{D}_m) \,-\, A\big| \;<\; \frac{1}{m} + \frac{1}{m} \;<\; \varepsilon.$$

Esto nos indica que $f \in \mathcal{HK}([a,b])$ y termina la prueba. ■

El Criterio de Cauchy para la integral de Henstock-Kurzweil es un elemento fundamental para establecer la siguiente caracterización de las funciones que son Henstock-Kurzweil integrables.

Teorema 14.1.6. *Sea $f : [a,b] \to \mathbb{R}$ una función. Las siguientes condiciones son equivalentes:*

(1) $f \in \mathcal{HK}([a,b])$.

(2) *Para cada* $\varepsilon > 0$, *existen funciones* $\phi, \Phi \in \mathcal{HK}([a,b])$ *tales que*

$$\phi \leq f \leq \Phi \qquad y \qquad (\text{HK})\int_a^b (\Phi - \phi)\, dt \leq \varepsilon.$$

Prueba. (1) $\Rightarrow$ (2) Si $f \in \mathcal{HK}([a,b])$ y $\varepsilon > 0$, entonces podemos tomar $\phi = \Phi = f$.
(2) $\Rightarrow$ (1) Suponga que (2) se cumple y fijemos un $\varepsilon > 0$. Como $\phi \in \mathcal{HK}([a,b])$, existe un calibrador δ_1 sobre $[a,b]$ tal que

$$\left| S(\phi, \mathcal{R}) - (\text{HK})\int_a^b \phi\, dt \right| < \varepsilon/2$$

para cualquier partición etiquetada $\mathcal{R}$ de $[a,b]$ subordinada a δ_1, de donde se tiene que

$$(\text{HK})\int_a^b \phi\, dt - \varepsilon/2 \leq S(\phi, \mathcal{R})$$

Con un argumento similar, siendo $\Phi \in \mathcal{HK}([a,b])$, existe un calibrador δ_2 sobre $[a,b]$ tal que si $\mathcal{S}$ es una partición etiquetada de $[a,b]$ subordinada a δ_2, entonces

$$S(\Phi, \mathcal{S}) \leq (\text{HK})\int_a^b \Phi\, dt + \varepsilon/2$$

Sea $\delta = \text{mín}\{\delta_1, \delta_2\}$. Si $\mathcal{P}$ es una partición etiquetada de $[a,b]$ subordinada a δ, entonces ella está subordinada tanto a δ_1 así como a δ_2 y entonces, de las dos desigualdades anteriores se tiene que

$$(\text{HK})\int_a^b \phi\, dt - \varepsilon/2 \leq S(f, \mathcal{P}) \leq (\text{HK})\int_a^b \Phi\, dt + \varepsilon/2.$$

De igual forma, pero ahora trabajando con $-f$, si $\mathcal{Q}$ es una partición etiquetada de $[a,b]$ subordinada a δ entonces

$$-(\text{HK})\int_a^b \Phi\, dt - \varepsilon/2 \leq -S(f, \mathcal{Q}) \leq -(\text{HK})\int_a^b \phi\, dt + \varepsilon/2$$

Sumando estas dos desigualdades, obtenemos

$$-(\text{HK})\int_a^b (\Phi - \phi)\, dt - \varepsilon \leq S(f, \mathcal{P}) - S(f, \mathcal{Q}) \leq (\text{HK})\int_a^b (\Phi - \phi)\, dt + \varepsilon,$$

de donde se concluye que

$$\left| S(f, \mathcal{P}) - S(f, \mathcal{Q}) \right| \leq (\text{HK})\int_a^b (\Phi - \phi)\, dt + \varepsilon \leq 2\varepsilon.$$

Siendo $\varepsilon > 0$ arbitrario, resulta que f satisface el Criterio de Cauchy y, en consecuencia, es Henstock-Kurzweil integrable sobre $[a,b]$. ∎

Ahora mostraremos que cualquier función Riemann integrable es también Henstock-Kurzweil integrable y que ambas integrales coinciden.

Teorema 14.1.7. *Si $f \in \mathcal{R}([a,b])$, entonces $f \in \mathcal{HK}([a,b])$ y*

$$(\mathrm{R})\int_a^b f(t)\,dt \;=\; (\mathrm{HK})\int_a^b f(t)\,dt. \tag{1}$$

Prueba. Sea $\varepsilon > 0$. Puesto que $f \in \mathcal{R}([a,b])$, existe un $\delta_\varepsilon > 0$ tal que

$$\left| S(f,\mathcal{P}) - (\mathrm{R})\int_a^b f(t)\,dt \right| \;<\; \varepsilon$$

para cualquier partición etiquetada $\mathcal{P}$ de $[a,b]$ satisfaciendo $\|\,\mathcal{P}\,\| < \delta_\varepsilon$. Defina $\delta : [a,b] \to \mathbb{R}$ por $\delta(t) = \frac{1}{4}\delta_\varepsilon$ para todo $t \in [a,b]$. Entonces δ es un calibrador sobre $[a,b]$. Si $\mathcal{P}_e = \{(t_i, [x_{i-1}, x_i]) : i = 1,\ldots,n\}$ es una partición etiquetada de $[a,b]$ subordinada a δ, entonces, puesto que

$$[x_{i-1}, x_i] \;\subseteq\; (t_i - \delta(t_i),\, t_i + \delta(t_i)) \;=\; \left(t_i - \frac{1}{4}\delta_\varepsilon,\, t_i + \frac{1}{4}\delta_\varepsilon\right)$$

resulta que $0 < x_i - x_{i-1} < \frac{1}{2}\delta_\varepsilon \;<\; \delta_\varepsilon$ para cada $i \in \{1,\ldots,n\}$. Por consiguiente, la partición $\mathcal{P}_e$ también satisface $\|\,\mathcal{P}_e\,\| < \delta_\varepsilon$ y, por lo tanto,

$$\left| S(f,\mathcal{P}_e) - (\mathrm{R})\int_a^b f(t)\,dt \right| \;<\; \varepsilon. \tag{2}$$

Hemos demostrado que cualquier partición etiquetada $\mathcal{P}_e$ de $[a,b]$ que esté subordinada a δ cumple con la desigualdad (2). Esto nos indica que $f \in \mathcal{HK}([a,b])$ y, por la unicidad de la integral, (1) se cumple. ■

El resultado anterior combinado con el conocimiento que toda función continua $f : [a,b] \to \mathbb{R}$ es Riemann integrable, permite concluir que:

$$C([a,b]) \;\subseteq\; \mathcal{R}([a,b]) \;\subseteq\; \mathcal{HK}([a,b]).$$

En particular, toda función $f : [a,b] \to \mathbb{R}$ que sea monótona, en escalera o regulada es Henstock-Kurzweil integrable. Por supuesto, no tendría ningún interés construir las funciones Henstock-Kurzweil integrables salvo, tal vez, por una perspicaz curiosidad, si cualquier función Henstock-Kurzweil integrable fuese Riemann integrable. Los dos ejemplos que ahora mostraremos confirman que, efectivamente,

$$\mathcal{R}([a,b]) \;\subsetneqq\; \mathcal{HK}([a,b]).$$

Ejemplo 14.1.1. *La función de Dirichlet $f = \chi_{\mathbb{Q}\cap[0,1]}$ es **acotada**, pertenece a $\mathcal{HK}([0,1])$ pero no a $\mathcal{R}([0,1])$.*

Prueba. Sea $\varepsilon > 0$ y suponga que la sucesión $(r_k)_{k=1}^{\infty}$ es una enumeración de los racionales en $[0,1]$. Defina $\delta : [0,1] \to \mathbb{R}$ por

$$\delta(x) \;=\; \begin{cases} \dfrac{\varepsilon}{2^{k+1}} & \text{si } x = r_k, \quad k = 1,2,\ldots \\ 1 & \text{si } x \in [0,1] \text{ es irracional.} \end{cases}$$

Claramente, δ es un gauge. Fijemos una partición etiquetada de $[0,1]$ subordinada a δ, digamos $\mathcal{P}_e = (t_i, [x_{i-1}, x_i])_{i=1}^n$. En este caso, $x_i - x_{i-1} < 2\delta(t_i)$ para cada $i = 1, \ldots, n$. Observe que si la etiqueta t_i es irracional, entonces $f(t_i) = 0$, de modo que en la suma de Riemann $S(f, \mathcal{P}_e) = f(t_1)(x_1 - x_0) + \cdots + f(t_n)(x_n - x_{n_1})$ sólo interesan las etiquetas que son racionales. En este caso, si $t_i = r_k$ para algún $k \in \mathbb{N}$, entonces

$$0 \;<\; f(r_k)(x_i - x_{i-1}) \;=\; 1 \cdot (x_i - x_{i-1}) \;\leq\; 2\frac{\varepsilon}{2^{k+1}} \;=\; \frac{\varepsilon}{2^k}$$

y, por lo tanto,

$$\left|S(f, \mathcal{P}_e)\right| \;\leq\; \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} \;=\; \varepsilon.$$

Puesto que $\varepsilon > 0$ fue elegido de modo arbitrario, se tiene que $f \in \mathcal{HK}([a,b])$ y $\int_0^1 f\,dt = 0$. ■

Ejemplo 14.1.2. *La función de* $f : [0,1] \to \mathbb{R}$ *definida por*

$$f(x) \;=\; \begin{cases} k & si \;\; x = 1/k, \quad k = 1, 2, \ldots \\ 0 & si \;\; x \in [0,1] \setminus \{1/k \,:\, k = 1, 2, \ldots\}. \end{cases}$$

es claramente ***no acotada*** *sobre* $[0,1]$ *y, por consiguiente, no pertenece a* $\mathcal{R}([0,1])$ *pero sí a* $\mathcal{HK}([0,1])$.

Prueba. Fijemos $\varepsilon > 0$ y defina $\delta : [0,1] \to (0, +\infty)$ por

$$\delta(x) \;=\; \begin{cases} \dfrac{\varepsilon}{k\,2^{k+1}} & \text{si } \; x = 1/k, \quad k = 1, 2, \ldots \\ 1 & \text{si } \; x \in [0,1] \setminus \{1/k \,:\, k = 1, 2, \ldots\}. \end{cases}$$

Si $\mathcal{P}_e = (t_i, [x_{i-1}, x_i])_{i=1}^n$ es una partición etiquetada de $[0,1]$ subordinada a δ, entonces debemos tener que $x_i - x_{i-1} < 2\delta(t_i)$ para $i = 1, 2, \ldots, n$. Similar al ejemplo anterior, las únicas etiquetas que realmente importan son aquellas que tienen la forma $t_i = 1/k$ para algún $k \in \mathbb{N}$ y, por consiguiente,

$$0 \;<\; f(1/k)(x_i - x_{i-1}) \;=\; k \cdot (x_i - x_{i-1}) \;\leq\; 2k\frac{\varepsilon}{k\,2^{k+1}} \;=\; \frac{\varepsilon}{2^k}.$$

De esto se sigue que

$$\left|S(f, \mathcal{P}_e)\right| \;\leq\; \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} \;=\; \varepsilon.$$

Puesto que $\varepsilon > 0$ es arbitrario, se tiene que $f \in \mathcal{HK}([a,b])$ y (HK) $\int_0^1 f\,dt = 0$. ■

El hecho de f sea Henstock-Kurzweil integrable en los dos ejemplos anteriores son casos particulares de un resultado más general.

Teorema 14.1.8. *Sea* $f : [a,b] \to \mathbb{R}$ *una función. Si* $f = 0$ ***casi-siempre***, *entonces* $f \in \mathcal{HK}([a,b])$ *y*

$$(\text{HK})\int_0^1 f\,dt \;=\; 0.$$

Prueba. Sea $E = \{x \in [a,b] : f(x) \neq 0\}$. Por hipótesis, $\mu(E) = 0$. Para cada $n \in \mathbb{N}$, considere el conjunto

$$E_n = \{x \in E : n-1 \leq |f(x)| < n\}.$$

Por supuesto, cada uno de los conjuntos E_n tiene medida cero, son disjuntos dos a dos y $E = \bigcup_{n=1}^{\infty} E_n$. Fijemos un $\varepsilon > 0$ y, por cada $n \in \mathbb{N}$, seleccione un conjunto abierto O_n tal que

$$E_n \subseteq O_n \qquad \text{y} \qquad \mu(O_n) < \frac{\varepsilon}{n2^n}.$$

Esta información es suficiente para obtener un calibrador apropiado sobre $[a,b]$. En efecto, basta con definir la función $\delta : [a,b] \to (0,+\infty)$ por

$$\delta(x) = \begin{cases} \operatorname{dist}(x, O_n^c) & \text{si } x \in E_n, \\ 1 & \text{si } x \in [0,1] \setminus E. \end{cases}$$

Suponga ahora que $\mathcal{P}_e = (t_i, [x_{i-1}, x_i])_{i=1}^n$ es una partición etiquetada de $[a,b]$ subordinada a δ. Para cada $n \in \mathbb{N}$, sea $\mathcal{P}_n$ el subconjunto de $\mathcal{P}_e$ cuyas etiquetas están en E_n. Nótese que si J es un intervalo de $\mathcal{P}_n$, entonces $J \subseteq O_n$ y, por lo tanto,

$$\left|S(f, \mathcal{P}_n)\right| = \left|\sum_{(t_i,J_i)\in\mathcal{P}_n} f(t_i)\ell(J_i)\right| < n \sum_{(t_i,J_i)\in\mathcal{P}_n} \mu(O_n) < \frac{\varepsilon}{2^n}.$$

De esta última desigualdad se tiene entonces que

$$\left|S(f, \mathcal{P}_e)\right| \leq \sum_{n=1}^{\infty} \left|S(f, \mathcal{P}_n)\right| < \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon.$$

Esto nos revela que $f \in \mathcal{HK}([a,b])$ y (HK)$\int_0^1 f\,dt = 0$. ■

Como consecuencia del resultado anterior se obtiene lo siguiente: *Si $f : \mathbb{R} \to \mathbb{R}$ es una función **nula casi-siempre**, entonces*

$$(\text{HK})\int_I f\,dt = 0 \quad \text{para cualquier intervalo } I \subseteq \mathbb{R}.$$

*En particular, si $E \subseteq \mathbb{R}$ es un **conjunto nulo**, entonces χ_E es Henstock-Kurzweil integrable con*

$$(\text{HK})\int_I \chi_E\,dt = 0 \quad \text{para cualquier intervalo } I \subseteq \mathbb{R}.$$

14.1.1. El Teorema Fundamental del Cálculo

Recordemos que el Teorema Fundamental del Cálculo para la integral de Riemann establece que:

TFC - Riemann: *Si $F : [a,b] \to \mathbb{R}$ es **diferenciable** en cada punto $t \in [a,b]$ y F' es **Riemann integrable** sobre $[a,b]$, entonces*

$$(\text{R})\int_a^b F'\,dt = F(b) - F(a).$$

Para la integral de Lebesgue, dicho teorema se expresa de la siguiente forma:

TFC - Lebesgue: *Si $F : [a,b] \to \mathbb{R}$ es **diferenciable** en cada punto $t \in [a,b]$ y F' es **acotada** sobre $[a,b]$, entonces $F' \in \mathrm{L}_1([a,b])$ y se cumple que*

$$(\mathrm{L})\int_a^b F'\,dt \;=\; F(b) - F(a).$$

Sin embrago, como ya habíamos informado, para la integral de Henstock-Kurzweil, no se requiere ninguna hipótesis adicional sobre la función derivada F', en otras palabras:

TFC - Henstock-Kurzweil: *Si $F : [a,b] \to \mathbb{R}$ es **diferenciable** en cada punto $t \in [a,b]$, entonces $F' \in \mathcal{HK}([a,b])$ y se tiene que*

$$(\mathrm{HK})\int_a^b F'\,dt \;=\; F(b) - F(a).$$

Por consiguiente, si una función $f : [a,b] \to \mathbb{R}$ posee una primitiva, la integrabilidad de f según Henstock-Kurzweil se convierte en una conclusión en lugar de una hipótesis como ocurre con las integrales de Riemann y Lebesgue. La demostración del Teorema Fundamental de Henstock-Kurzweil se obtendrá casi inmediatamente si se establece el siguiente resultado, el cual es un modo alternativo de expresar la derivada de una función en un punto.

Lema 14.1.9 (Definición alternativa de Derivada). *Sea $F : [a,b] \to \mathbb{R}$ una función **diferenciable** en un punto $t \in [a,b]$ y sea $\varepsilon > 0$. Entonces existe un $\delta(t) > 0$ tal que cualesquiera sean $u, v \in [a,b]$ satisfaciendo*

$$t - \delta(t) \;<\; u \;\le\; t \;\le\; v \;<\; t + \delta(t), \tag{1}$$

se cumple que

$$\big|F(v) \;-\; F(u) \;-\; F'(t)(v-u)\big| \;\le\; \varepsilon(v-u). \tag{2}$$

Prueba. Puesto que, por definición,

$$F'(t) \;=\; \lim_{x\to t} \frac{F(x) \;-\; F(t)}{x \;-\; t}$$

entonces, para el $\varepsilon > 0$, existe un $\delta(t) > 0$ tal que, para cada $x \in [a,b]$,

$$0 \;<\; |x-t| \;<\; \delta(t) \qquad \Rightarrow \qquad \left|\frac{F(x) - F(t)}{x - t} \;-\; f(t)\right| \;<\; \varepsilon,$$

o de forma equivalente,

$$|x-t| \;<\; \delta(t) \qquad \Rightarrow \qquad |F(x) - F(t) - f(t)(x-t)| \;\le\; \varepsilon|x-t| \tag{3}$$

Sean ahora u, v verificando la condición (1). Entonces $|u-t| = t-u$, $|v-t| = v-t$ y por (3),

$$\begin{aligned}
\big|F(v) \;-\; F(u) \;-\; F'(t)(v-u)\big| \;&=\; \big|F(v) \;-\; F(t) \;+\; F(t) \;-\; F(u) \;-\; F'(t)(v-t+t-u)\big| \\
&\le\; \big|F(v) \;-\; F(t) \;-\; F'(t)(v-t)\big| \;+\; \big|F(u) \;-\; F(t) \;-\; F'(t)(u-t)\big| \\
&\le\; \varepsilon|v-t| \;+\; \varepsilon|u-t| \;=\; \varepsilon(v-t) \;+\; \varepsilon(t-u) \;=\; \varepsilon(v-u)
\end{aligned}$$

y termina la prueba. ■

Teorema 14.1.10 (Fundamental del Cálculo (HK). *Sea $F : [a,b] \to \mathbb{R}$ una función* ***diferenciable*** *en $[a,b]$. Entonces $F' \in \mathcal{HK}([a,b])$ y vale la igualdad*

$$(\mathrm{HK})\int_a^b F'\,dt \;=\; F(b) - F(a). \tag{1}$$

Prueba. Fijemos un $\varepsilon > 0$ arbitrario y pongamos $\varepsilon' = \varepsilon/(b-a)$. El Lema 14.1.9 aplicado a F y a ε' nos dice que para cada $t \in [a,b]$, existe un $\delta(t) > 0$ con la siguiente propiedad: cualesquiera sean $u, v \in [a,b]$ tal que $t - \delta(t) < u \le t \le v < t + \delta(t)$ se tiene que

$$\big|F(v) \,-\, F(u) \,-\, F'(t)(v-u)\big| \;<\; \varepsilon'(v-u). \tag{2}$$

Por supuesto, la función $\delta : [a,b] \to \mathbb{R}$ es un calibrador. Considere ahora cualquier partición etiquetada $\mathcal{P}_e = (t_i, [x_{i-1}, x_i])_{i=1}^n$ de $[a,b]$ subordinada a δ. Entonces, para cada $i = 1, 2, \ldots, n$,

$$t_i \,-\, \delta(t_i) \;<\; x_{i-1} \;\le\; t_i \;\le\; x_i \;<\; t_i \,+\, \delta(t_i).$$

Aplicando la desigualdad (2) con $u = x_{i-1}$, $v = x_i$ y $t = t_i$ se obtiene

$$\big|F(x_i) \,-\, F(x_{i-1}) \,-\, F'(t_i)(x_i - x_{i-1})\big| \;<\; \varepsilon'(x_i - x_{i-1})$$

para cada $i = 1, 2, \ldots, n$. Por otro lado, puesto que $F(b) - F(a) = \sum_{i=1}^n F(x_i) \,-\, F(x_{i-1})$ resulta entonces que

$$\begin{aligned}
\big|F(b) - F(a) \,-\, S(f, \mathcal{P}_e)\big| \;&=\; \left|\sum_{i=1}^n \big[F(x_i) \,-\, F(x_{i-1})\big] \,-\, \sum_{i=1}^n F'(t_i)(x_i - x_{i-1})\right| \\
&=\; \left|\sum_{i=1}^n \Big[F(x_i) \,-\, F(x_{i-1}) \,-\, F'(t_i)(x_i - x_{i-1})\Big]\right| \\
&\le\; \sum_{i=1}^n \big|F(x_i) \,-\, F(x_{i-1}) \,-\, F'(t_i)(x_i - x_{i-1})\big| \\
&<\; \sum_{i=1}^n \varepsilon'(x_i - x_{i-1}) \;=\; \frac{\varepsilon}{b-a}(b-a) \;=\; \varepsilon.
\end{aligned}$$

Esto nos revela que F' es Henstock-Kurzweil integrable y (1) se cumple. ■

Del Teorema Fundamental del Cálculo para la integral de Henstock-Kurzweil y los resultados vistos previamente se obtiene que

$$\begin{array}{ccccc}
 & & \mathcal{N}([a,b]) & & \\
 & \rotatebox{45}{$\subseteq$} & & \rotatebox{-45}{$\subseteq$} & \\
C([a,b]) & & & & \mathcal{HK}([a,b]) \\
 & \rotatebox{-45}{$\subseteq$} & & \rotatebox{45}{$\subseteq$} & \\
 & & \mathcal{R}([a,b]) & &
\end{array}$$

Más aun, del Teorema Fundamental del Cálculo para la integral de Henstock-Kurzweil se obtiene, similar al caso de la integral de Riemann, la regla de integración por sustitución o del cambio de variable. Como siempre, si $f \in \mathcal{HK}([a,b])$ definimos

$$(\text{HK})\int_a^b f\,dt = -(\text{HK})\int_b^a f\,dt \qquad \text{y} \qquad (\text{HK})\int_a^a f\,dt = 0.$$

Teorema 14.1.11 (Integración por Sustitución). *Sean $f : [a,b] \to \mathbb{R}$ y $\varphi : [\alpha,\beta] \to [a,b]$* ***funciones diferenciables****. Entonces*

$$(\text{HK})\int_{\varphi(\alpha)}^{\varphi(\beta)} f'\,dt = (\text{HK})\int_a^b (f' \circ \varphi)\varphi'\,dt.$$

Prueba. Por la Regla de la Cadena, $(f \circ \varphi)' = (f' \circ \varphi)\varphi'$. El hecho de que

$$(\text{HK})\int_\alpha^\beta (f' \circ \varphi)\varphi'\,dt = f(\varphi(\beta)) - f(\varphi(\alpha)) = (\text{HK})\int_{\varphi(\alpha)}^{\varphi(\beta)} f'\,dt$$

sigue del Teorema 14.1.10. ■

El siguiente ejemplo nos muestra una derivada que, por supuesto, es Henstock-Kurzweil integrable pero que no es integrable ni en el sentido de Riemann, ni el sentido de Lebesgue.

Ejemplo 14.1.3. *La función*

$$F(x) = \begin{cases} 2x\cos\left(\dfrac{\pi}{x^2}\right) & \text{si } 0 < x \le 1, \\ 0 & \text{si } x = 0. \end{cases}$$

posee, hecho que ya hemos demostrado, una ***derivada no acotada***

$$F'(x) = \begin{cases} 2x\cos\left(\dfrac{\pi}{x^2}\right) + \dfrac{2\pi}{x^2}\,\text{sen}\left(\dfrac{\pi}{x^2}\right) & \text{si } 0 < x \le 1, \\ 0 & \text{si } x = 0. \end{cases}$$

la cual, por el Teorema Fundamental del Cálculo, es Henstock-Kurzweil integrable sobre $[0,1]$ y

$$(\text{HK})\int_0^1 F'(t)\,dt = F(1) - F(0) = -1.$$

Sin embargo, como ya hemos visto, F' no es ***ni Riemann integrable****,* ***ni Lebesgue integrable*** *sobre $[0,1]$.*

Una mejora substancial del Teorema Fundamental del Cálculo para la integral de Henstock-Kurzweil se puede obtener si recordamos lo siguiente: *Sea $f : [a,b] \to \mathbb{R}$ una función. Una función $F : [a,b] \to \mathbb{R}$ es una $\aleph_0$-**primitiva** de f sobre $[a,b]$, si F es continua sobre $[a,b]$ y existe un conjunto a lo más numerable $E \subseteq [a,b]$ tal $F'(t) = f(t)$ para cada $t \in [a,b] \setminus E$ y $F'(t)$ no existe, o si existe, no es igual a $f(t)$ para cada $t \in E$.* El conjunto E es llamado un **conjunto excepcional** de f.

Observe que si $E = \{c_k : k = 1,2,\ldots\}$ es el conjunto de puntos $t \in [a,b]$ para los cuales $F'(t)$ no existe, entonces, puesto que $\mu(E) = 0$, siempre se puede redefinir F' en tales puntos de modo arbitrario y, en consecuencia, convenir que $F'(c_k) = 0$ para todo $k \in \mathbb{N}$.

Teorema 14.1.12 (Fundamental del Cálculo $\aleph_0$-HK). *Si una función $f : [a,b] \to \mathbb{R}$ posee una $\aleph_0$-**primitiva** $F : [a,b] \to \mathbb{R}$, entonces $f \in \mathcal{HK}([a,b])$ y se cumple que*

$$(\mathrm{HK})\int_a^b f\,dt \;=\; F(b) - F(a). \tag{1}$$

Prueba. Sea $E = \{c_k : k = 1,2,\ldots\}$ el conjunto excepcional para la $\aleph_0$-primitiva F. Nuestro propósito inmediato es construir un calibrador $\delta : [a,b] \to \mathbb{R}$. Veamos cómo se hace esto. Fijemos un $\varepsilon > 0$ y sea $t \in [a,b] \setminus E$. Como F es diferenciable en t, seleccione, aplicando el Lema 14.1.9, un $\delta(t) > 0$ tal que si u, v satisfacen

$$t - \delta(t) \;<\; u \;\leq\; t \;\leq\; v \;<\; t + \delta(t), \tag{2}$$

entonces

$$\big|F(v) \;-\; F(u) \;-\; F'(t)(v-u)\big| \;\leq\; \varepsilon(v-u). \tag{3}$$

Por otro lado, si $t = c_k$ para algún k, entonces use la continuidad de F para escoger un $\delta(c_k) > 0$ de modo que, para todo $x \in [a,b]$ tal que $0 < |x - c_k| < \delta(c_k)$, se cumpla que

$$|F(z) - F(c_k)| \;\leq\; \varepsilon/2^{k+3} \qquad \text{y} \qquad |f(c_k)||x - c_k| \;<\; \varepsilon/2^{k+3}.$$

De esta forma queda determinado el calibrador deseado $\delta : [a,b] \to (0,+\infty)$. Considere ahora cualquier partición etiquetada $\mathcal{P}_e = (t_i, [x_{i-1}, x_i])_{i=1}^n$ de $[a,b]$ subordinada a δ. Si ninguna de las etiquetas t_i pertenecen a E, entonces la prueba dada en el Teorema 14.1.10 se aplica sin cambio para obtener

$$\big|F(b) - F(a) \;-\; S(f, \mathcal{P}_e)\big| \;<\; \varepsilon.$$

Por otro lado, si t_i es una etiqueta perteneciente a E, digamos $t_i = c_k$ para algún k, entonces

$$\begin{aligned}\big|F(x_i) - F(x_{i-1}) - f(t_i)(x_i - x_{i-1})\big| &\leq \big|F(x_i) - F(c_k)\big| + \big|F(c_k) - F(x_{i-1})\big| + \big|f(c_k)(x_i - x_{i-1})\big| \\ &\leq \varepsilon/2^{k+3} \;+\; \varepsilon/2^{k+3} \;+\; \varepsilon/2^{k+3} \;<\; \varepsilon/2^{k+1}.\end{aligned}$$

Ahora bien, como cada punto de E puede ser la etiqueta de a lo sumo dos subintervalos de $\mathcal{P}_e$, resulta que la suma de Riemann de los términos con $t_i \in E$ debe satisfacer

$$\sum_{t_i \in E} \big|F(x_i) \;-\; F(x_{i-1}) \;-\; f(t_i)(x_i - x_{i-1})\big| \;\leq\; 2\sum_{k=1}^{\infty} \varepsilon/2^{k+1} \;=\; \varepsilon.$$

También, la suma Riemann de los términos $t_i \notin E$ satisface, por (2) y (3) la desigualdad

$$\begin{aligned}\sum_{t_i \notin E} \big|F(x_i) \;-\; F(x_{i-1}) \;-\; f(t_i)(x_i - x_{i-1})\big| \;&=\; \sum_{t_i \notin E} \big|F(x_i) \;-\; F(x_{i-1}) \;-\; F'(t_i)(x_i - x_{i-1})\big| \\ &\leq\; \varepsilon \sum_{t_i \notin E} (x_i - x_{i-1}) \;\leq\; \varepsilon(b-a).\end{aligned}$$

Combinando las dos desigualdades anteriores se tiene que si $\mathcal{P}_e = (t_i, [x_{i-1}, x_i])_{i=1}^n$ es cualquier partición etiquetada de $[a,b]$ subordinada a δ, entonces

$$\big|F(b) - F(a) \;-\; S(f, \mathcal{P}_e)\big| \;<\; \varepsilon(1 + b - a).$$

Siendo $\varepsilon > 0$ arbitrario se concluye que $f \in \mathcal{HK}([a,b])$ y (1) se cumple. ■

Puesto que todo subconjunto numerable de $[a,b]$ es de medida cero, es natural formularse, en virtud del Teorema 14.1.12, la siguiente pregunta:

Pregunta: *Sea $f : [a,b] \to \mathbb{R}$ una función. Si F es una función **continua** sobre $[a,b]$ y existe un conjunto nulo $E \subseteq [a,b]$ tal que $F'(t) = f(t)$ para todo $t \in [a,b] \setminus E$ ¿será cierto que $f \in \mathcal{HK}([a,b])$ y*

$$(\mathrm{HK})\int_a^b f\,dt \;=\; F(b) - F(a)?$$

Desafortunadamente, la respuesta es: **no**. En efecto, la función de Cantor $\varphi_\Gamma : [0,1] \to \mathbb{R}$ es continua sobre $[0,1]$ y diferenciable casi-siempre ya que $\varphi'_\Gamma(t) = 0$ para todo $t \in [0,1] \setminus \Gamma$ y $\mu(\Gamma) = 0$. Se sigue del Teorema 14.1.8 que $\varphi'_\Gamma \in \mathcal{HK}([a,b])$ y

$$(\mathrm{HK})\int_0^1 \varphi'_\Gamma(t)\,dt \;=\; 0.$$

Sin embargo, la igualdad $(\mathrm{HK})\int_0^1 \varphi'_\Gamma(t)\,dt = \varphi_\Gamma(1) - \varphi_\Gamma(0)$ no puede ocurrir pues ello nos conduciría al siguiente disparate:

$$0 \;=\; (\mathrm{HK})\int_0^1 \varphi'_\Gamma(t)\,dt \;=\; \varphi_\Gamma(1) - \varphi_\Gamma(0) \;=\; 1.$$

Ejemplo 14.1.4. *Sea $f : [0,1] \to \mathbb{R}$ definida por*

$$f(x) \;=\; \begin{cases} \dfrac{1}{\sqrt{x}} & si\;\; x \in (0,1] \\ 0 & si\;\; x = 0. \end{cases}$$

La función $F : [0,1] \to \mathbb{R}$ dada por $F(x) = 2\sqrt{x}$ es claramente una $\aleph_0$-primitiva de f y, entonces, por el Teorema 14.1.12, $f \in \mathcal{HK}([0,1])$ y

$$(\mathrm{HK})\int_0^1 f(t)\,dt \;=\; F(1) - F(0) \;=\; 2.$$

Observe que f no es Riemann integrable ya que ella no es acotada sobre $[0,1]$, aunque $(\mathrm{HK})\int_0^1 f(t)\,dt$ coincide con la integral de Cauchy-Riemann, es decir, la *integral impropia* en el sentido de Riemann. Otra de las virtudes de la integral de Henstock-Kurzweil es que cualquier integral impropia, en el sentido de Riemann, es una integral de Henstock-Kurzweil.

14.1.2. Propiedades Básicas

Como ocurre con las integrales de Riemann y Lebesgue, veremos que el conjunto $\mathcal{HK}([a,b])$ es un espacio vectorial sobre $\mathbb{R}$ y que la integral de Henstock-Kurzweil es un funcional lineal sobre $\mathcal{HK}([a,b])$. Además, muchas otras de las operaciones algebraicas que se cumplen para las integrales anteriores también son válidas en $\mathcal{HK}([a,b])$.

Teorema 14.1.13. *Sean* $f, g \in \mathcal{HK}([a,b])$ *y* $\lambda \in \mathbb{R}$. *Entonces:*

(a) $f + g \in \mathcal{HK}([a,b])$ *y* $$(\mathrm{HK})\int_a^b (f+g)\,dt \;=\; (\mathrm{HK})\int_a^b f\,dt \;+\; (\mathrm{HK})\int_a^b g\,dt.$$

(b) $\lambda f \in \mathcal{HK}([a,b])$ *y* $$(\mathrm{HK})\int_a^b \lambda f\,dt \;=\; \lambda \cdot (\mathrm{HK})\int_a^b f\,dt.$$

Prueba. (a) Sean $f, g \in \mathcal{HK}([a,b])$ y fijemos $\varepsilon > 0$ tomado de modo arbitrario. Por definición, existen funciones gauges $\delta_1, \delta_2 : [a,b] \to (0, +\infty)$ con las siguientes propiedades: si $\mathcal{Q}_e$ y $\mathcal{R}_e$ son particiones etiquetadas de $[a,b]$ subordinadas a δ_1 y δ_2 respectivamente, entonces

$$\left| S(f, \mathcal{Q}_e) - (\mathrm{HK})\int_a^b f\,dt \right| \;<\; \frac{\varepsilon}{2} \qquad \text{y} \qquad \left| S(f, \mathcal{R}_e) - (\mathrm{HK})\int_a^b g\,dt \right| \;<\; \frac{\varepsilon}{2}$$

Defina $\delta(t) = \min\{\delta_1(t), \delta_2(t)\}$ para todo $t \in [a,b]$. Resulta que δ es un gauge. Si $\mathcal{P}_e$ es una partición etiquetada de $[a,b]$ subordinada a δ, entonces ella también está subordinada tanto a δ_1 así como a δ_2 y, por lo tanto,

$$\begin{aligned}
\left| S(f+g, \mathcal{P}_e) - (\mathrm{HK})\int_a^b (f+g)\,dt \right| \;&=\; \left| S(f, \mathcal{P}_e) + S(f, \mathcal{P}_e) - (\mathrm{HK})\int_a^b f\,dt - (\mathrm{HK})\int_a^b g\,dt \right| \\
&\leq\; \left| S(f, \mathcal{P}_e) - (\mathrm{HK})\int_a^b f\,dt \right| + \left| S(g, \mathcal{P}_e) - (\mathrm{HK})\int_a^b g\,dt \right| \\
&<\; \varepsilon.
\end{aligned}$$

La demostración de (b) es similar y se deja a cargo del lector. ■

El siguiente resultado, el cual es válido para la integral de Lebesgue (pero no para la integral de Riemann), también se cumple para la integral de Henstock-Kurzweil.

Corolario 14.1.14. *Sea* $f \in \mathcal{HK}([a,b])$. *Si* $g : [a,b] \to \mathbb{R}$ *es cualquier función tal que* $f = g$ ***casi-siempre***, *entonces* $g \in \mathcal{HK}([a,b])$.

Prueba. Como $g - f = 0$ casi-siempre, el Teorema 14.1.8 nos muestra que $g - f \in \mathcal{HK}([a,b])$ y entonces, por ser $\mathcal{HK}([a,b])$ un espacio vectorial, se tiene que $g = f + (g - f) \in \mathcal{HK}([a,b])$. ■

En lo que sigue, adoptaremos la siguiente convención:

Si $f : [a,b] \to \mathbb{R}$ *es una función y* $[c,d] \subseteq [a,b]$, *entonces la expresión* $f \in \mathcal{HK}([c,d])$ *significa que la función* $f \cdot \chi_{[c,d]} \in \mathcal{HK}([a,b])$, *en otras palabras,*

$$(\mathrm{HK})\int_c^d f\,dt \;=\; (\mathrm{HK})\int_a^b f \cdot \chi_{[c,d]}\,dt.$$

La demostración del próximo resultado requiere forzar a un punto particular del intervalo $[a,b]$ para que sea una etiqueta del cualquier partición etiquetada δ-fina.

Teorema 14.1.15. *Sea* $f : [a,b] \to \mathbb{R}$ *una función y sea* $c \in (a,b)$.

(*a*) *Si* $f \in \mathcal{HK}([a,c]) \cap \mathcal{HK}([c,b])$, *entonces* $f \in \mathcal{HK}([a,b])$ *y*

$$(\mathrm{HK})\int_a^b f\,dt \;=\; (\mathrm{HK})\int_a^c f\,dt \;+\; (\mathrm{HK})\int_c^b f\,dt. \tag{1}$$

(*b*) *Si* $f \in \mathcal{HK}([a,b])$, *entonces* $f \in \mathcal{HK}([c,d])$ *para cualquier subintervalo* $[c,d] \subseteq [a,b]$.

Prueba. (*a*) Sea $\varepsilon > 0$. Puesto que $f \in \mathcal{HK}([a,c])$ existe un calibrador $\delta_1 : [a,c] \to (0,+\infty)$ tal que

$$\left| S(f,\mathcal{P}_1) \;-\; (\mathrm{HK})\int_a^c f\,dt \right| \;<\; \varepsilon/2$$

para cualquier partición etiquetada $\mathcal{P}_1$ de $[a,c]$ subordinada a δ_1. Similarmente, como $f \in \mathcal{HK}([c,b])$ existe un calibrador $\delta_2 : [c,b] \to (0,+\infty)$ tal que

$$\left| S(f,\mathcal{P}_2) \;-\; (\mathrm{HK})\int_c^b f\,dt \right| \;<\; \varepsilon/2$$

para cualquier partición etiquetada $\mathcal{P}_2$ de $[c,b]$ subordinada a δ_2. Defina $\delta : [a,b] \to (0,+\infty)$ por

$$\delta(x) \;=\; \begin{cases} \text{mín}\{\delta_1(x),\, c-x\} & \text{si } x \in [a,c) \\ \text{mín}\{\delta_1(x),\, \delta_2(x)\} & \text{si } x = c \\ \text{mín}\{\delta_1(x),\, x-c\} & \text{si } x \in (x,b]. \end{cases}$$

Sea $\mathcal{P} = (t_i, [x_{i-1}, x_i])_{i=1}^n$ una partición etiquetada de $[a,b]$ subordinada a δ. Por el Lema 14.1.1 podemos asumir que c es una etiqueta de al menos un subintervalo de $\mathcal{P}$ y, por consiguiente, $\mathcal{P}$ debe ser de la forma $\mathcal{P}_a \cup (c,[u,v]) \cup \mathcal{P}_b$, donde las etiquetas de $\mathcal{P}_a$ son menores que c y las etiquetas de $\mathcal{P}_b$ son mayores que c. Sean

$$\mathcal{P}_1 \;=\; \mathcal{P}_a \cup (c,[u,c]) \qquad\qquad \text{y} \qquad\qquad \mathcal{P}_2 \;=\; \mathcal{P}_b \cup (c,[c,v]).$$

Entonces $\mathcal{P}_1$ es una partición etiquetada de $[a,c]$ subordinada a δ_1 y $\mathcal{P}_2$ es una partición etiquetada de $[c,b]$ subordinada a δ_2. Puesto que $S(f,\mathcal{P}) = S(f,\mathcal{P}_1) + S(f,\mathcal{P}_2)$, resulta que

$$\begin{aligned} \left| S(f,\mathcal{P}) - (\mathrm{HK})\int_a^c f\,dt - (\mathrm{HK})\int_c^b f\,dt \right| &\le \left| S(f,\mathcal{P}_1) - (\mathrm{HK})\int_a^c f\,dt \right| + \left| S(f,\mathcal{P}_2) - (\mathrm{HK})\int_c^b f\,dt \right| \\ &< \varepsilon. \end{aligned}$$

Esto nos confirma que $f \in \mathcal{HK}([a,b])$ y (1) se cumple.
(*b*) Suponga que $[c,d]$ es un subintervalo propio de $[a,b]$ y sea $\varepsilon > 0$. Como $f \in \mathcal{HK}([a,b])$, existe, por el Criterio de Cauchy, un calibrador $\delta : [a,b] \to (0,+\infty)$ tal que

$$\left| S(f,\mathcal{P}) \;-\; S(f,\mathcal{Q}) \right| \;<\; \varepsilon$$

para cualquier par de particiones etiquetadas $\mathcal{P}, \mathcal{Q}$ de $[a,b]$ subordinadas a δ. Siendo $[c,d]$ un subintervalo propio de $[a,b]$, existe una colección finita $\{[u_1,v_1],\ldots,u_n,v_n\}$ de subintervalos no-superpuestos de $[a,b]$ tal que

$$[c,d] \notin \{[u_1,v_1],\ldots,u_n,v_n\} \qquad \text{y} \qquad [a,b] \;=\; [c,d] \cup \bigcup_{k=1}^{n}[u_k,v_k].$$

Usemos el Lema de Cousin para seleccionar, por cada $k \in \{1,\ldots,n\}$, una partición etiquetada $\mathcal{P}_k$ de $[u_k,v_k]$ subordinada a δ. Si $\mathcal{P}_c$ y $\mathcal{P}_d$ son particiones etiquetadas de $[c,d]$ subordinadas a δ, entonces $\mathcal{P}_c \cup \bigcup_{k=1}^{n} \mathcal{P}_k$ y $\mathcal{P}_d \cup \bigcup_{k=1}^{n} \mathcal{P}_k$ son particiones etiquetadas de $[a,b]$ subordinadas a δ y, por lo tanto,

$$\begin{aligned}
\Big|S(f,\mathcal{P}_c) \;-\; S(f,\mathcal{P}_d)\Big| \;&=\; \left|S(f,\mathcal{P}_c) + \sum_{k=1}^{n} S(f,\mathcal{P}_k) - S(f,\mathcal{P}_d) - \sum_{k=1}^{n} S(f,\mathcal{P}_k)\right| \\
&=\; \left|S\left(f,\, \mathcal{P}_c \cup \bigcup_{k=1}^{n} \mathcal{P}_k\right) - S\left(f,\, \mathcal{P}_d \cup \bigcup_{k=1}^{n} \mathcal{P}_k\right)\right| \\
&<\; \varepsilon
\end{aligned}$$

Un nuevo llamado al Criterio de Cauchy nos revela que $f \in \mathcal{HK}([c,d])$. ■

Sea $f \in \mathcal{HK}([a,b])$. Recordemos que la integral indefinida de f se define por

$$F(x) \;=\; (\mathrm{HK})\!\int_a^x f\,dt \quad \text{para todo } x \in [a,b].$$

Del resultado anterior se tiene que si $a \le c < d \le b$, entonces

$$(\mathrm{HK})\!\int_a^d f\,dt \;=\; (\mathrm{HK})\!\int_a^c f\,dt \;+\; (\mathrm{HK})\!\int_c^d f\,dt,$$

es decir,

$$F(d) - F(c) \;=\; (\mathrm{HK})\!\int_c^d f\,dt.$$

Este hecho permite pensar a F como una función de conjuntos, es decir, si $\mathcal{I}c([a,b])$ denota a *la familia de todos los subintervalos cerrados* $[c,d]$ *incluidos en* $[a,b]$, entonces podemos reescribir a F como la función $F : \mathcal{I}c([a,b]) \to \mathbb{R}$ dada por:

$$F([c,d]) \;=\; F(d) - F(c) \quad \text{para todo } [c,d] \in \mathcal{I}c([a,b])$$

Por consiguiente, si $\mathcal{P} = \{a = x_0, x_1, \ldots, x_n = b\}$ es una partición de $[a,b]$, entonces

$$(\mathrm{HK})\!\int_a^b f\,dt \;=\; F(b) - F(a) \;=\; \sum_{i=1}^{n} F(x_i) - F(x_{i-1}) \;=\; \sum_{i=1}^{n} F([x_{i-1},x_i]).$$

Teorema 14.1.16. *Sea $f \in \mathcal{HK}([a,b])$. Si $f \geq 0$ sobre $[a,b]$, entonces*

$$(\mathrm{HK})\int_a^b f\,dt \;\geq\; 0.$$

En particular, si $f, g \in \mathcal{HK}([a,b])$ y $f \leq g$ sobre $[a,b]$, entonces

$$(\mathrm{HK})\int_a^b f\,dt \;\leq\; (\mathrm{HK})\int_a^b g\,dt.$$

Prueba. Suponga que $f \in \mathcal{HK}([a,b])$ es no-negativa y sea $\varepsilon > 0$. Por definición, existe un calibrador $\delta : [a,b] \to (0,+\infty)$ tal que

$$\left| S(f,\mathcal{P}_e) - (\mathrm{HK})\int_a^b f\,dt \right| \;<\; \varepsilon,$$

para cualquier partición etiquetada $\mathcal{P}_e$ de $[a,b]$ subordinada a δ. Puesto que $f \geq 0$, entonces

$$0 \;\leq\; S(f,\mathcal{P}_e) \;<\; (\mathrm{HK})\int_a^b f\,dt + \varepsilon$$

y como $\varepsilon > 0$ es arbitrario, se sigue que $(\mathrm{HK})\int_a^b f\,dt \geq 0$. ■

Observe que cualquier función $g : [a,b] \to \mathbb{R}$ que sea constante sobre $[a,b]$ es Henstock-Kurzweil integral y, por consiguiente, del resultado anterior se sigue que si $f \in \mathcal{HK}([a,b])$ es tal que $|f| \leq M$ para alguna constante $M \geq 0$, entonces

$$\left| (\mathrm{HK})\int_a^b f\,dt \right| \;\leq\; M(b-a).$$

Corolario 14.1.17. *Si $f \in \mathcal{HK}([a,b])$ es **no-negativa** y F es la **integral indefinida** de f, entonces F es **no-decreciente**.*

Nota Adicional 14.1.2 Podemos usar el resultado anterior para mostrar una función que no es Henstock-Kurzweil integrable. En efecto, la función $u : [0,1] \to \mathbb{R}$ dada por

$$u(x) \;=\; \begin{cases} 1/x & \text{si } x \in (0,1] \\ 0 & \text{si } x = 0 \end{cases}$$

no es Henstock-Kurzweil integrable. Suponga, para generar una contradicción, que u es Henstock-Kurzweil integrable. Por cada $n \in \mathbb{N}$, considere la función $u_n : [0,1] \to \mathbb{R}$, definida por

$$u_n(x) \;=\; \sum_{k=1}^{n-1} k \cdot \chi_{\left(\frac{1}{k+1},\frac{1}{k}\right]}(x).$$

Puesto que $u_n \leq u$ y $u_n \in \mathcal{R}([0,1]) \subseteq \mathcal{HK}([0,1])$, entonces

$$\sum_{i=1}^{n-1} \frac{1}{k+1} = (\mathrm{R})\int_0^1 u_n\,dt = (\mathrm{HK})\int_0^1 u_n\,dt \leq (\mathrm{HK})\int_0^1 u\,dt$$

de donde se tiene que la serie $\sum_{i=1}^{\infty} \frac{1}{k}$ converge, lo cual es absurdo.

Si f y g son funciones Riemann integrables, o Lebesgue integrables, entonces $f \cdot g$ es integrable; sin embargo, esto no ocurre con la integral de Henstock-Kurzweil. Para ver esto, usaremos el ejemplo anterior. En efecto, la función $v : [0,1] \to \mathbb{R}$ dada por

$$v(x) = \begin{cases} 1/\sqrt{x} & \text{si } x \in (0,1] \\ 0 & \text{si } x = 0. \end{cases}$$

es Henstock-Kurzweil integrable, pero $v \cdot v = u \notin \mathcal{HK}([a,b])$. En general, vale el siguiente resultado.

Teorema. *Sean $f, g \in \mathcal{HK}([a,b])$ con g de **variación acotada**. Entonces $f \cdot g$ es Henstock-Kurzweil integrable.*

14.1.3. El Lema de Saks-Henstock

Como ocurre en toda buena teoría, siempre existen resultados que brillan con luz propia y, por lo tanto, su importancia se aprecia por el amplio espectro de resultados importantes que de él se derivan. En esta sección discutiremos uno de esos resultados conocido como el Lema de Saks-Henstock o Lema de Henstock. Como afirma Gordon en su libro [66], pág. 144: casi cualquier resultado importante en la teoría de la integral de Henstock-Kurzweil usa dicho lema en algún punto de la prueba. Por ejemplo, el lema se usa para demostrar la validez de los teoremas de convergencia de la forma

$$\lim_{n\to\infty} (\mathrm{HK})\int_a^b f_n\,dt = (\mathrm{HK})\int_a^b f\,dt.$$

El lema en cuestión establece que *las sumas de Riemann no sólo aproximan la integral sobre el intervalo $[a,b]$, sino también sobre uniones de subintervalos.*

Definición 14.1.18. *Una **subpartición etiquetada** de $[a,b]$ consiste de una colección finita de la forma $\mathcal{S}_e = \{(s_1, J_1), \ldots, (s_k, J_k)\}$ donde cada etiqueta $s_i \in J_i \subseteq [a,b]$ para $i \in \{1,\ldots,k\}$ y los intervalos $J_1, \ldots, J_k$ son cerrados, no-degenerados y no-superpuestos.*

Observe que cada subpartición etiquetada $\mathcal{S}_e = (s_i, J_i)_{i=1}^k$ de $[a,b]$ satisface

$$\bigcup_{i=1}^k J_i \subseteq [a,b], \quad \text{pero puede ocurrir que} \quad \bigcup_{i=1}^k J_i \neq [a,b].$$

Más aun, si $f \in \mathcal{HK}([a,b])$ y $J = \bigcup_{i=1}^k J_i$, entonces

$$(\mathrm{HK})\int_J f\,dt = \sum_{i=1}^k (\mathrm{HK})\int_{J_i} f\,dt.$$

Fijemos ahora un calibrador $\delta : [a,b] \to (0,+\infty)$. Una subpartición etiquetada $\mathcal{S}_e = (s_i, J_i)_{i=1}^k$ de $[a,b]$ se dice δ-fina, o subordinada a δ, si

$$J_i \subseteq (s_i - \delta(s_i),\, s_i + \delta(s_i)) \quad \text{para } i = 1, \ldots, k.$$

Lema 14.1.19 (Saks-Henstock). *Sea $f \in \mathcal{HK}([a,b])$ y sea $\varepsilon > 0$. Si δ es un **calibrador** sobre $[a,b]$ tal que*

$$\left| S(f,\mathcal{P}) \;-\; (\text{HK})\int_a^b f\,dt \right| \;<\; \varepsilon$$

*para cualquier **partición etiquetada** $\mathcal{P}$ de $[a,b]$ subordinada a δ, entonces*

$$\left| \sum_{i=1}^{k} \left(f(s_i)\ell(J_i) \;-\; (\text{HK})\int_{J_i} f\,dt \right) \right| \;=\; \left| S(f,\mathcal{S}) \;-\; (\text{HK})\int_{\bigcup_{i=1}^{k} J_i} f\,dt \right| \;\le\; \varepsilon \tag{1}$$

*para cualquier **subpartición etiquetada** $\mathcal{S} = (s_i, J_i)_{i=1}^k$ de $[a,b]$ subordinada a δ.*

Prueba. Sea $\mathcal{S} = (s_i, [u_i, v_i])_{i=1}^k$ una subpartición etiquetada de $[a,b]$ subordinada a δ. Si $\bigcup_{i=1}^k [u_i, v_i] = [a,b]$, entonces el resultado sigue del Teorema 14.1.15. Suponga entonces que $\bigcup_{i=1}^k [u_i, v_i] \subsetneqq [a,b]$. En este caso, existen intervalos no rampantes $[c_1, d_1], \ldots, [c_m, d_m]$ incluidos en $[a,b]$ tales que

$$[a,b] \setminus \bigcup_{i=1}^{k} (u_i, v_i) \;=\; \bigcup_{j=1}^{m} [c_j, d_j].$$

Una nueva aplicación del Teorema 14.1.15 nos garantiza que $f \in \mathcal{HK}([c_j, d_j])$ para cada $j \in \{1, \ldots, m\}$. Fijemos un $\tau > 0$ y por cada $j \in \{1, \ldots, m\}$, escoja, usando el Lema de Cousin, un calibrador δ_j sobre $[c_j, d_j]$ de modo que

$$\left| S(f,\mathcal{P}_j) \;-\; (\text{HK})\int_{c_j}^{d_j} f \right| \;<\; \frac{\tau}{m}$$

para cada partición etiquetada $\mathcal{P}_j$ de $[c_j, d_j]$ subordinada a δ_j. Puesto que δ y δ_j son calibradores sobre $[c_j, d_j]$, podemos aplicar el Lema de Cousin al calibrador $\min\{\delta, \delta_j\}$ para hallar una partición etiquetada $\mathcal{Q}_j$ de $[c_j, d_j]$ subordinada a $\min\{\delta, \delta_j\}$. Claramente $\mathcal{Q} = \mathcal{S} \cup \bigcup_{j=1}^m \mathcal{Q}_j$ es partición etiquetada de $[a,b]$ subordinada a δ tal que

$$S(f,\mathcal{Q}) \;=\; S(f,\mathcal{S}) \;+\; \sum_{j=1}^{m} S(f,\mathcal{Q}_j)$$

y

$$(\text{HK})\int_a^b f\,dt \;=\; \sum_{i=1}^{k} (\text{HK})\int_{u_i}^{v_i} f\,dt \;+\; \sum_{j=1}^{m} (\text{HK})\int_{c_j}^{d_j} f\,dt.$$

Por consiguiente,

$$\begin{aligned}
&\left|\sum_{i=1}^{k}\left(f(s_i)(v_i-u_i) \;-\; \text{(HK)}\int_{u_i}^{v_i} f\,dt\right)\right| \\
&= \left|\left(S(f,\mathcal{Q})-\sum_{j=1}^{m} S(f,\mathcal{Q}_j)\right) \;-\; \left(\text{(HK)}\int_a^b f\,dt-\sum_{j=1}^{m}\text{(HK)}\int_{c_j}^{d_j} f\,dt\right)\right| \\
&\leq \left|S(f,\mathcal{Q})-\text{(HK)}\int_a^b f\,dt\right| \;+\; \sum_{j=1}^{m}\left|S(f,\mathcal{Q}_j)-\text{(HK)}\int_{c_j}^{d_j} f\,dt\right| \\
&< \;\varepsilon+\tau
\end{aligned}$$

Puesto que $\tau > 0$ es arbitrario, la desigualdad en (1) se cumple. ■

El siguiente corolario también recibe el nombre de Lema de Saks-Henstock.

Corolario 14.1.20 (Saks-Henstock). *Sea $f \in \mathcal{HK}([a,b])$. Entonces, para cada $\varepsilon > 0$, existe un* ***calibrador*** *δ sobre $[a,b]$ tal que*

$$\sum_{i=1}^{k}\left|f(s_i)\ell(J_i) \;-\; \text{(HK)}\int_{J_i} f\,dt\right| \;\leq\; 2\varepsilon$$

para cualquier ***subpartición etiquetada*** *$\mathcal{S} = (s_i, J_i)_{i=1}^{k}$ de $[a,b]$ subordinada a δ.*

Prueba. Puesto que $f \in \mathcal{HK}([a,b])$, el Lema 14.1.19 nos garantiza la existencia de un calibrador δ sobre $[a,b]$ tal que

$$\left|\sum_{i=1}^{m}\left(f(s_i)\ell(J_i) \;-\; \text{(HK)}\int_{J_i} f\,dt\right)\right| \;=\; \left|S(f,\mathcal{Q}) \;-\; \text{(HK)}\int_{\bigcup\limits_{i=1}^{m} J_i} f\,dt\right| \;\leq\; \varepsilon$$

para cada subpartición etiquetada $\mathcal{S} = (s_i, J_i)_{i=1}^{m}$ de $[a,b]$ subordinada a δ. Fijemos una subpartición etiquetada $\mathcal{S} = (s_i, J_i)_{i=1}^{m}$ de $[a,b]$ subordinada a δ y considere el conjunto

$$\mathcal{S}_1 \;=\; \left\{(s_i, J_i) \in \mathcal{S} \;:\; f(s_i)\ell(J_i) - \text{(HK)}\int_{J_i} f\,dt \;\geq\; 0\right\}.$$

Tomando $\mathcal{S}_2 = \mathcal{S} \setminus \mathcal{S}_1$, resulta que

$$\begin{aligned}
&\sum_{i=1}^{m}\left|f(s_i)\ell(J_i) \;-\; \text{(HK)}\int_{J_i} f\,dt\right| \\
&= \sum_{(s_i,J_i)\in\mathcal{S}_1}\left(f(s_i)\ell(J_i) \;-\; \text{(HK)}\int_{J_i} f\,dt\right) \;-\; \sum_{(s_i,J_i)\in\mathcal{S}_2}\left(f(s_i)\ell(J_i) \;-\; \text{(HK)}\int_{J_i} f\,dt\right) \\
&\leq\; 2\varepsilon.
\end{aligned}$$

La prueba es completa. ■

Corolario 14.1.21. *Sea $f \in \mathcal{HK}([a,b])$. Entonces, para cada $\varepsilon > 0$, existe un **calibrador** δ sobre $[a,b]$ tal que*

$$\left| \sum_{i=1}^{k} |f(s_i)|\ell(J_i) \;-\; \sum_{i=1}^{k} \left|(\mathrm{HK})\int_{J_i} f\,dt\right| \right| \;\leq\; 2\varepsilon$$

*para cualquier **subpartición etiquetada** $\mathcal{S} = (s_i, J_i)_{i=1}^{k}$ de $[a,b]$ subordinada a δ.*

Prueba. Fijemos $\varepsilon > 0$. Por el Corolario 14.1.20, existe un calibrador δ sobre $[a,b]$ tal que

$$\sum_{i=1}^{k} \left| f(s_i)\ell(J_i) \;-\; (\mathrm{HK})\int_{J_i} f\,dt \right| \;\leq\; 2\varepsilon$$

para cualquier subpartición etiquetada $\mathcal{S} = (s_i, J_i)_{i=1}^{k}$ de $[a,b]$ subordinada a δ. Fijemos una tal subpartición $\mathcal{S} = (s_i, J_i)_{i=1}^{k}$ de $[a,b]$ subordinada a δ y observe que de la desigualdad $\big||a| - |b|\big| \leq |a-b|$, la cual es válida para cualesquiera sean los números $a, b \in \mathbb{R}$ se tiene, tomando $a_i = f(s_i)\ell(J_i)$ y $b_i = \int_{J_i} f\,dt$ y luego sumando, que

$$\begin{aligned}\left| \sum_{i=1}^{k} |f(s_i)|\ell(J_i) \;-\; \sum_{i=1}^{k} \left|(\mathrm{HK})\int_{J_i} f\,dt\right| \right| \;&=\; \left| \sum_{i=1}^{k} \left(|f(s_i)\ell(J_i)| \;-\; \left|(\mathrm{HK})\int_{J_i} f\,dt\right| \right) \right| \\ &\leq\; \sum_{i=1}^{k} \left| f(s_i)\ell(J_i) \;-\; (\mathrm{HK})\int_{J_i} f\,dt \right| \;\leq\; 2\varepsilon\end{aligned}$$

Fin de la prueba. ■

Una pieza fundamental para demostrar que en $\mathcal{HK}([a,b])$ las integrales indefinidas permanecen en dicho conjunto la proporciona el siguiente resultado.

Teorema 14.1.22. *Sea $f \in \mathcal{HK}([a,b])$ y sea $F : [a,b] \to \mathbb{R}$ la **integral indefinida** de f, es decir,*

$$F(x) \;=\; (\mathrm{HK})\int_{a}^{x} f\,dt \quad \text{para todo } x \in [a,b].$$

*Entonces F es **continua** sobre $[a,b]$.*

Prueba. Sea $\varepsilon > 0$. Para demostrar que F es continua sobre $[a,b]$, tomemos un número arbitrario $x \in [a,b]$. Puesto que $f \in \mathcal{HK}([a,b])$ existe, por definición, un calibrador δ sobre $[a,b]$ tal que

$$\left| S(f, \mathcal{P}) \;-\; (\mathrm{HK})\int_{a}^{x} f\,dt \right| \;<\; \varepsilon.$$

para cada partición etiquetada $\mathcal{P}$ de $[a,b]$ subordinada a δ. Fijemos una tal partición etiqueta $\mathcal{P} = (t_i, I_i)_{i=1}^{n}$ de $[a,b]$ subordinada a δ y sea I_{i_0} un intervalo de la partición conteniendo a x. Puesto que $x \in I_{i_0} \subseteq (t_{i_0} - \delta(t_{i_0}), t_{i_0} + \delta(t_{i_0}))$, podemos considerar el número γ definido por

$$\gamma \;=\; \min\Big\{ x - t_{i_0} - \delta(t_{i_0}),\; t_{i_0} + \delta(t_{i_0}) - x,\; \varepsilon/(1 + |f(x)|) \Big\}.$$

Tomemos ahora que cualquier $y \in [a,b]$ que satisfaga $|y-x| < \gamma$. Entonces, aplicando el Lema de Saks-Henstock a la subpartición etiquetada $\{(x,[y,x])\}$ o $\{(x,[x,y])\}$, se tiene que

$$\left| f(x)(y-x) - (\mathrm{HK})\int_x^y f\,dt \right| \leq \varepsilon \qquad \text{o} \qquad \left| f(x)(x-y) - (\mathrm{HK})\int_y^x f\,dt \right| \leq \varepsilon.$$

En cualquier caso

$$\begin{aligned} \big|F(y)-F(x)\big| &\leq \big|f(x)(y-x) - \big(F(y)-F(x)\big)\big| + \big|f(x)\big|\big|(y-x)\big| \\ &\leq \varepsilon + \varepsilon = 2\varepsilon. \end{aligned}$$

Esto nos indica que F es continua en x y termina la prueba. ■

Como acabamos de ver, la integral indefinida F de cualquier función $f \in \mathcal{HK}([a,b])$, es siempre continua sobre $[a,b]$, pero en general no será absolutamente continua salvo que f sea Lebesgue integrable sobre $[a,b]$.

Hemos visto que si $f : [a,b] \to \mathbb{R}$ es una función tal que f es Riemann integrable sobre cada subintervalo $[a,c] \subseteq [a,b]$ donde $c \in (a,b)$, entonces puede ocurrir que la función $f \notin \mathcal{R}([a,b])$. Esto sucede, por ejemplo, si f posee una singularidad en algún punto $x_0 \in [a,b]$, lo cual significa que $\lim_{x \to x_0} |f(x)| = +\infty$. Por ejemplo, $x_0 = 0$ es una punto singular de la función $f(x) = 1/\sqrt{x}$ si $x \in (0,1]$ y $f(0) = 0$. Suponga ahora que $x_0 = b$ es una singularidad de una función $f : [a,b] \to \mathbb{R}$ y que f es Riemann integrable sobre cada subintervalo $[a,c] \subseteq [a,b]$ donde $c \in (a,b)$. Se define la **integral de Cauchy-Riemann** o **integral impropia de Riemann** como

$$(\mathrm{CR})\int_a^b f\,dt = \lim_{c \to b^-} (\mathrm{R})\int_a^c f\,dt$$

siempre que dicho límite exista. Consideraciones similares se aplican en los otros casos. Observe que $(\mathrm{CR})\int_a^b f\,dt$ no es la integral de Riemann. En este sentido, la integral impropia de Riemann generaliza la integral de Riemann. Más aun, también puede ocurrir que f sea Lebesgue integrable sobre cada subintervalo $[a,c] \subseteq [a,b]$ donde $c \in (a,b)$, $(\mathrm{CR})\int_a^b f\,dt$ exista y tampoco f sea Lebesgue integrable. En el siguiente resultado veremos que para la integral de Henstock-Kurzweil no existen integrales impropias.

Teorema 14.1.23 (Hake). *Sea $f : [a,b] \to \mathbb{R}$ una función. Las siguientes condiciones son equivalentes:*

(1) $f \in \mathcal{HK}([a,b])$.

(2) $f \in \mathcal{HK}([a,c])$ para cada $c \in (a,b)$ y $\lim_{c \to b^-} (\mathrm{HK})\int_a^c f\,dt$ existe.

En este caso se tiene que

$$(\mathrm{HK})\int_a^b f\,dt = \lim_{c \to b^-} (\mathrm{HK})\int_a^c f\,dt.$$

Prueba. (1) $\Rightarrow$ (2) Por el Teorema 14.1.15 sabemos que $f \in \mathcal{HK}([a,c])$ para cada $c \in (a,b)$. Además, por el Teorema 14.1.22, la integral indefinida F de f es continua sobre $[a,b]$. En particular, continua en $x = b$ y, por lo tanto, se verifica la igualdad:

$$(\mathrm{HK})\int_a^b f\,dt = F(b) = \lim_{c \to b^-} F(c) = \lim_{c \to b^-} (\mathrm{HK})\int_a^c f\,dt.$$

(2) $\Rightarrow$ (1) Fijemos $\varepsilon > 0$ y seleccione una sucesión $(c_n)_{n=0}^{\infty}$ en $[a,b]$ que sea estrictamente creciente tal que $c_0 = a$ y $\lim_{n\to\infty} c_n = b$. Nuestra hipótesis nos garantiza la existencia de $A \in \mathbb{R}$ tal que

$$A \;=\; \lim_{c\to b^-}\,(\mathrm{HK})\!\int_a^c f\,dt.$$

Ahora bien, como $\lim_{n\to\infty} c_n = b$ y $\lim_{c_n\to b}(\mathrm{HK})\!\int_a^{c_n} f\,dt = A$, podemos elegir un $N \in \mathbb{N}$ de modo tal que $b - c_N \leq \varepsilon/4(1+|f(b)|)$ y también que

$$\left|(\mathrm{HK})\!\int_a^x f\,dt \;-\; A\right| \;<\; \frac{\varepsilon}{4}$$

para cualquier $x \in [c_N, b)$.

Por otro lado, puesto que $\bigcup_{n=1}^{\infty}[c_{n-1}, c_n] = [a,b)$, una aplicación del Corolario 14.1.20 nos conduce a obtener, por cada $n \in \mathbb{N}$, un calibrador δ_n sobre $J_n = [c_{n-1}, c_n]$ tal que la desigualdad

$$\sum_{(s,[u,v])\in\mathcal{P}_n}\left|f(s)\ell(v-u) \;-\; (\mathrm{HK})\!\int_u^v f\,dt\right| \;<\; \frac{\varepsilon}{4(2^n)}$$

se cumpla para cada subpartición etiquetada $\mathcal{P}_n$ de $[c_{n-1}, c_n]$ subordinada a δ_n. Defina un calibrador δ sobre $[a,b]$ del modo siguiente:

$$\delta(t) \;=\; \begin{cases} \frac{1}{2}(c_1 - c_0) & \text{si } t = c_0 \\ \text{mín}\left\{\delta_n(c_n),\, \delta_{n+1}(c_n),\, \frac{1}{2}(c_n - c_{n-1}),\, \frac{1}{2}(c_{n+1} - c_n)\right\} & \text{si } t = c_n \text{ para algún } n \geq 1 \\ \text{mín}\left\{\delta_n(t),\, \frac{1}{2}(t - c_{n-1}),\, \frac{1}{2}(c_n - t)\right\} & \text{si } t \in (c_{n-1}, c_n) \text{ para algún } n \geq 1 \\ b - c_N & \text{si } t = b \end{cases}$$

y sea $\mathcal{P} = ((t_i, [x_{i-1}, x_i]))_{i=1}^{p}$ una partición etiquetada de $[a,b]$ subordinada a δ. Puesto que $b \notin \bigcup_{n=1}^{\infty}[c_{n-1}, c_n] = [a,b)$, nuestra elección de δ implica que $t_p = b$ y $x_{p-1} \in (c_m, c_{m+1}]$ para algún único entero $m \geq 1$. También, observe que si $k \in \{1,\ldots,m\}$, nuestra elección de δ implica que

$$\big\{(t,[x,y]) \in \mathcal{P} \,:\, [x,\,y] \subseteq [c_{k-1},\, c_k]\big\}$$

es una partición etiquetada de $[c_{k-1}, c_k]$ subordinada a δ_k. Por esto,

$$\begin{aligned} &\big|S(f,\mathcal{P}) \;-\; A\big| \\ \leq\;& \left|\sum_{k=1}^{m}\left(\sum_{\substack{(t,[x,y])\in\mathcal{P} \\ [x,y]\subseteq[c_{k-1},c_k]}} f(t)(y-x) \;-\; (\mathrm{HK})\!\int_{c_{k-1}}^{c_k} f\,dt\right)\right| \\ +\;& \left|\sum_{\substack{(t,[x,y])\in\mathcal{P} \\ [x,y]\subseteq[c_m,x_{p-1}]}} f(t)(y-x) \;-\; (\mathrm{HK})\!\int_{c_m}^{x_{p-1}} f\,dt\right| \;+\; |f(b)|(b - x_{p-1}) \;+\; \left|(\mathrm{HK})\!\int_a^{x_{p-1}} f\,dt - A\right| \\ <\;& \varepsilon. \end{aligned}$$

Esto finaliza la prueba. ■

Nota Adicional 14.1.3 Otro caso similar al Teorema de Hake viene dado:

(1) *Sea* $f : [a,b] \to \mathbb{R}$ *una función. Las siguientes condiciones son equivalentes:*

(α) $f \in \mathcal{HK}([a,b])$.

(β) $f \in \mathcal{HK}([c,b])$ *para cada* $c \in (a,b)$ *y* $\lim\limits_{c \to a^+} (\mathrm{HK})\int_c^b f\,dt$ *existe*.

En este caso se tiene que

$$(\mathrm{HK})\int_a^b f\,dt = \lim_{c \to a^+} (\mathrm{HK})\int_c^b f\,dt.$$

En general, si para cada intervalo $[c,d] \subseteq (a,b)$, el límite

$$\lim_{\substack{c \to a^+ \\ d \to b^-}} (\mathrm{HK})\int_c^d f\,dt$$

existe, entonces $f \in \mathcal{HK}([a,b])$ y

$$(\mathrm{HK})\int_a^b f\,dt = \lim_{\substack{c \to a^+ \\ d \to b^-}} (\mathrm{HK})\int_c^d f\,dt.$$

Otra consecuencia importante del Teorema de Hake es el siguiente:

(2) *Sea* $f : [a,b] \to \mathbb{R}$ *una función que es* ***Riemann integrable*** *sobre cada subintervalo* $[a,c] \subseteq [a,b]$ *con* $c \in (a,b)$. *Si*

$$\lim_{c \to b^-} (\mathrm{R})\int_a^c f\,dt$$

existe, entonces $f \in \mathcal{HK}([a,b])$ *y*

$$(\mathrm{CR})\int_a^b f\,dt = \lim_{c \to b^-} (\mathrm{R})\int_a^c f\,dt = (\mathrm{HK})\int_a^c f\,dt.$$

Como sucede con la integral de Lebesgue, no cualquier función $f : [a,b] \to \mathbb{R}$ es Henstock-Kurzweil integrable. En efecto, la función $f : [0,1] \to \mathbb{R}$ definida por

$$f(x) = \begin{cases} 1/x & \text{si } x \in (0,1] \\ 0 & \text{si } x = 0 \end{cases}$$

no es integrable en el sentido de Henstock-Kurzweil. Para ver esto, observe que para cualquier $c \in (0,1)$,

$$(\mathrm{R})\int_c^1 \frac{1}{x}\,dt = -\ln(c)$$

existe, pero $\lim\limits_{c \to 0^-} (\mathrm{R})\int_c^1 f\,d = +\infty$. Por esto, $f \notin \mathcal{HK}([0,1])$.

14.1.4. El Segundo Teorema Fundamental del Cálculo

El segundo Teorema Fundamental del Cálculo para la integral de Lebesgue establece que si $f \in L_1([a,b])$, entonces su integral indefinida es absolutamente continua (en particular, diferenciable) y F' es igual a f casi-siempre. En esta sección demostraremos un resultado similar para la integral de Henstock-Kurzweil, es decir, si $f \in \mathcal{HK}([a,b])$, entonces su integral indefinida es diferenciable y su derivada es igual a f casi-siempre. Es en este resultado donde el Lema de Saks-Henstock encuentra su más importante aplicación. Además del Lema de Saks-Henstock, usaremos las derivadas de Dini y el Teorema del Cubrimiento de Vitali para su demostración.

Teorema 14.1.24 (Fundamental del Cálculo 2). *Sea $f \in \mathcal{HK}([a,b])$. Si*

$$F(x) \;=\; (\mathrm{HK})\int_a^x f\,dt \quad \textit{para cada } x \in [a,b],$$

*entonces F es **diferenciable casi-siempre** sobre $[a,b]$ y $F'(x) = f(x)$ para casi-todo $x \in [a,b]$. En particular, f es **Lebesgue medible**.*

Prueba. Demostraremos que $D^+F(x) = f(x)$ para casi-todo $x \in [a,b]$. Los otros tres casos poseen pruebas similares y, por lo tanto, las omitiremos. Considere el conjunto

$$E \;=\; \big\{x \in [a,b] \,:\, D^+F(x) \neq f(x)\big\}$$

y veamos que $\mu^*(E) = 0$. Esta afirmación requiere un poquito de trabajo. Observe que si $x \in E$, entonces existe un $\eta_x > 0$ con la siguiente propiedad: para cada $h > 0$, existe un punto $y_h^x \in (x, x+h) \cap [a,b]$ tal que

$$\big|F(y_h^x) \,-\, F(x) \,-\, f(x)(y_h^x - x)\big| \;\geq\; \eta_x\,(y_h^x - x).$$

Por cada entero $n \geq 1$, sea $E_n = \{x \in E : \eta_x \geq 1/n\}$. Puesto que $E = \bigcup_{n=1}^{\infty} E_n$, es suficiente demostrar que $\mu^*(E_n) = 0$ para cada $n \geq 1$.

Fijemos un $n \in \mathbb{N}$ y sea $\varepsilon > 0$. Puesto que $f \in \mathcal{HK}([a,b])$, existe una calibrador δ sobre $[a,b]$ tal que

$$\Bigg|\sum_{(t,[x,y])\,\in\,\mathcal{P}} f(t)(y-x) \;-\; (F(b) - F(a))\Bigg| \;=\; \Bigg|S(f,\mathcal{P}) \;-\; (\mathrm{HK})\int_a^b f\,dt\Bigg| \;<\; \varepsilon/4n$$

para cualquier partición etiquetada $\mathcal{P}$ de $[a,b]$ subordinada a δ. Observe ahora que la elección de δ nos permite concluir que la siguiente colección de intervalos cerrados

$$\mathcal{V} \;=\; \Big\{[x, y_h^x] \;:\; x \in E_n,\;\; 0 < h < \delta(x)\Big\}$$

forma un cubrimiento en el sentido de Vitali de E_n y, entonces, por el Teorema del Cubrimiento de Vitali, existe una colección finita $\{[c_i, d_i] : 1 \leq i \leq m\}$ de intervalos disjuntos en $\mathcal{V}$ tal que

$$\mu^*(E_n) \;<\; \sum_{i=1}^{m}(d_i - c_i) \,+\, \varepsilon/2. \tag{1}$$

Nótese que cada intervalo etiquetado $(c_i, [c_i, d_i])$ está subordinado a δ y que

$$\eta_{c_i}(d_i - c_i) \;\leq\; \big|F(d_i) \,-\, F(c_i) \,-\, f(c_i)(d_i - c_i)\big|.$$

Estamos ahora en posición de invocar el Lema de Saks-Henstock. En efecto, puesto que la colección $\{(c_i, [c_i, d_i]) : 1 \leq i \leq m\}$ es una subpartición etiquetada de $[a, b]$ subordinada a δ, entonces el Lema de Saks-Henstock, Corolario 14.1.21, nos asegura que

$$\begin{aligned}\sum_{i=1}^{m} \varepsilon_{c_i}(d_i - c_i) &\leq \sum_{i=1}^{m} \frac{1}{\eta_{c_i}} \big|F(d_i) - F(c_i) - f(c_i)(d_i - c_i)\big| \\ &\leq n \sum_{i=1}^{m} \big|f(c_i)(d_i - c_i) - (F(d_i) - F(c_i))\big| \\ &\leq n\frac{2\varepsilon}{4n} = \frac{\varepsilon}{2}\end{aligned}$$

lo que combinado con (1) nos dice que $\mu^*(E_n) < \varepsilon$. Siendo $\varepsilon > 0$ arbitrario, concluimos que $\mu^*(E) = 0$. De aquí se sigue que las cuatro derivadas de Dini de F son iguales a f casi-siempre en $[a, b]$ y, entonces, $F' = f$ casi-siempre en $[a, b]$.

Finalmente, el hecho de $f = F'$ casi-siempre en $[a, b]$ nos garantiza, por una aplicación del Teorema 7.2.6, página 380, que f es medible. ■

14.1.5. Integrabilidad Absoluta

Hay que advertir que, a diferencia de lo que ocurre con las integrales de Riemann y de Lebesgue, la siguiente desigualdad **no es válida**, en general, para la integral de Henstock-Kurzweil:

$$\left|(\mathrm{HK})\int_a^b f\,dt\right| \leq (\mathrm{HK})\int_a^b |f|\,dt.$$

La razón de por qué ello es así se encuentra en el hecho de que *la integral de Henstock-Kurzweil no es una integral absoluta*, es decir, $f \in \mathcal{HK}([a, b])$ no implica que $|f| \in \mathcal{HK}([a, b])$. He aquí un ejemplo.

Ejemplo 14.1.5. *La función* $f : [a, b] \to \mathbb{R}$ *definida por*

$$f(x) = \begin{cases} \cos(\pi/x) + (\pi/x)\operatorname{sen}(\pi/x) & \text{si } x \in (0, 1] \\ 0 & \text{si } x = 0 \end{cases}$$

satisface $f \in \mathcal{HK}([a, b])$, *pero* $|f| \notin \mathcal{HK}([a, b])$.

Prueba. Claramente f es continua en el intervalo $(0, 1]$ y, por consiguiente, gracias al Teorema de Vitali-Lebesgue, ella es Riemann integrable sobre $[0, 1]$. El Teorema 14.1.7 nos indica entonces que $f \in \mathcal{HK}([0, 1])$. La función $F : [0, 1] \to \mathbb{R}$ definida por

$$F(x) = \begin{cases} x\cos(\pi/x) & \text{si } x \in (0, 1] \\ 0 & \text{si } x = 0 \end{cases}$$

es una primitiva de f. Se sigue del Teorema Fundamental del Cálculo para la integral de Henstock-Kurzweil que

$$\int_{a_k}^{b_k} f(t)\,dt = F(b_k) - F(a_k) = \frac{(-1)^k}{k}$$

donde $a_k = 2/(2k+1)$ y $b_k = 1/k$ para cada $k \in \mathbb{N}$. Suponga ahora que $|f| \in \mathcal{HK}([0,1])$. Entonces, de la desigualdad

$$\sum_{k=1}^{n} \frac{1}{k} \;\le\; \sum_{k=1}^{n} (\mathrm{HK})\int_{a_k}^{b_k} |f|\,dt \;\le\; (\mathrm{HK})\int_{0}^{1} |f|\,dt,$$

la cual es válida para cualquier $n \in \mathbb{N}$, se deduce que la serie $\sum_{k=1}^{\infty} 1/k$ converge. Esta contradicción establece que $|f| \notin \mathcal{HK}([a,b])$. ■

Por supuesto, si $f, |f| \in \mathcal{HK}([a,b])$, entonces de las desigualdades $-|f| \le f \le |f|$ y el Teorema 14.1.16, se sigue que $\int_a^b |f|\,dt \le \int_a^b f\,dt \le \int_a^b |f|\,dt$ y, en consecuencia:

Corolario 14.1.25. *Si $f, |f| \in \mathcal{HK}([a,b])$, entonces*

$$\left|(\mathrm{HK})\int_a^b f\,dt\right| \;\le\; (\mathrm{HK})\int_a^b |f|\,dt.$$

Nuestro objetivo principal en esta sección es indagar cuándo, dada $f \in \mathcal{HK}([a,b])$, ocurre que $|f| \in \mathcal{HK}([a,b])$. Una respuesta viene dada por el siguiente teorema. Antes, recordemos (véase, la Definición 9.1.12, página 462) que una función $f : [a,b] \to \mathbb{R}$ es de variación acotada en $[a,b]$ si $V(f,[a,b]) < +\infty$, donde

$$V(f,[a,b]) \;=\; \sup\left\{ \sum_{i=1}^{n} \big|f(x_i) - f(x_{i-1})\big| \;:\; \mathcal{P} \in \mathbb{P}(a,b) \right\}.$$

También, si $f \in \mathcal{HK}([a,b])$, entonces su integral indefinida es la función $F : [a,b] \to \mathbb{R}$ dada por

$$F(x) \;=\; (\mathrm{HK})\int_a^x f\,dt \quad \text{para cada } x \in [a,b].$$

Teorema 14.1.26 (Caracterización de Integrabilidad Absoluta). *Sea $f \in \mathcal{HK}([a,b])$. Las siguientes condiciones son equivalentes:*

(1) $|f| \in \mathcal{HK}([a,b])$.

(2) *F, la* ***integral indefinida*** *de f, es de* ***variación acotada*** *en $[a,b]$. En este caso*

$$(\mathrm{HK})\int_a^b |f|\,dt \;=\; V(F,[a,b]). \qquad (*)$$

Prueba. (1) $\Rightarrow$ (2) Suponga que $|f| \in \mathcal{HK}([a,b])$ y sea $\mathcal{P} = \{a = x_0, x_1, \ldots, x_n = b\}$ una partición de $[a,b]$. Las propiedades de la integral y el corolario anterior nos muestran que

$$\sum_{i=1}^{n} \big|F(x_i) - F(x_{i-1})\big| \;=\; \sum_{i=1}^{n} \left|(\mathrm{HK})\int_{x_{i-1}}^{x_i} f\,dt\right| \;\le\; \sum_{i=1}^{n} (\mathrm{HK})\int_{x_{i-1}}^{x_i} |f|\,dt \;=\; (\mathrm{HK})\int_a^b |f|\,dt.$$

Esto prueba que F es de variación acotada y

$$V(F,[a,b]) \;\le\; (\mathrm{HK})\int_a^b |f|\,dt.$$

(2) $\Rightarrow$ (1) Fijemos $\varepsilon > 0$. Como $V(F, [a,b]) < +\infty$, existe una partición de $[a,b]$, la cual escribiremos como $\mathcal{P}_0 = \{a = x_0, x_1, \ldots, x_n = b\}$, tal que

$$V(F,[a,b]) - \varepsilon < \sum_{i=1}^{n} \big|F(x_i) - F(x_{i-1})\big| \leq V(F,[a,b]).$$

Observe que si $\mathcal{Q} = \{a = z_0, z_1, \ldots, z_m\}$ es cualquier partición más fina que $\mathcal{P}_0$, entonces

$$V(F,[a,b]) - \varepsilon < \sum_{i=1}^{n} \big|F(x_i) - F(x_{i-1})\big| \leq \sum_{i=1}^{m} \big|F(z_i) - F(z_{i-1})\big| \leq V(F,[a,b]). \qquad (\alpha)$$

La idea ahora es trabajar con particiones etiquetadas de modo que todo punto de la partición $\mathcal{P}_0$ sea una etiqueta de tales particiones. Esto, por supuesto, siempre es posible gracias al Lema 14.1.1 y, en consecuencia, dichas particiones serán más finas que $\mathcal{P}_0$, lo que garantizará la validez de la desigualdad (α).

Ahora bien, puesto que $f \in \mathcal{HK}([a,b])$, existe un calibrador δ sobre $[a,b]$ tal que

$$\left| S(f, \mathcal{P}_e) - (\mathrm{HK})\int_a^b f\,dt \right| \leq \varepsilon$$

para cualquier partición etiquetada $\mathcal{P}_e$ de $[a,b]$ subordinada a δ. Una aplicación del Corolario 14.1.21 nos conduce a la desigualdad

$$\left| \sum_{i=1}^{k} |f(s_i)|(v_i - u_i) - \sum_{i=1}^{k} \left|(\mathrm{HK})\int_{u_i}^{v_i} f\,dt\right| \right| \leq 2\varepsilon. \qquad (\beta)$$

para cualquier subpartición etiquetada $\mathcal{S} = ((s_i, [u_i, v_i]))_{i=1}^{k}$ de $[a,b]$ subordinada a δ.

Sea E el conjunto de los puntos de la partición $\mathcal{P}_0$, es decir, $E = \{x_0, \ldots, x_n\}$ y defina un calibrador δ_1 sobre $[a,b]$ del modo siguiente:

$$\delta_1(x) = \min\left\{\delta(x), \frac{1}{2}\mathrm{dist}\big(x, E \setminus \{x\}\big)\right\} \quad \text{para cada } x \in [a,b].$$

Tomemos cualquier partición etiquetada $\mathcal{S} = ((s_i, [u_i, v_i]))_{i=1}^{k}$ de $[a,b]$ subordinada a δ_1 la cual, como ya acordamos, supondremos que contiene a todos los puntos de $\mathcal{P}_0$ como etiquetas. Puesto que $\mathcal{S}$ está también subordinada a δ, resulta que la desigualdad dada en (β) sigue siendo válida para tal subpartición y, en consecuencia, por (α) se tiene que

$$V(F,[a,b]) - \varepsilon < \sum_{i=1}^{k} \big|F(v_i) - F(u_{i-1})\big| = \sum_{i=1}^{k} \left|(\mathrm{HK})\int_{u_i}^{v_i} f\,dt\right| \leq V(F,[a,b]).$$

De esto se sigue que

$$\begin{aligned}\left|S(|f|,\mathcal{S}) - V(F,[a,b])\right| &= \left|\sum_{i=1}^{k}|f(s_i)|(v_i-u_i) - V(F,[a,b])\right| \\ &= \left|\sum_{i=1}^{k}|f(s_i)|(v_i-u_i) - \sum_{i=1}^{k}\left|(\text{HK})\int_{u_i}^{v_i} f\,dt\right|\right| \\ &+ \left|\sum_{i=1}^{k}\left|(\text{HK})\int_{u_i}^{v_i} f\,dt\right| - V(F,[a,b])\right| \\ &\leq 2\varepsilon+\varepsilon = 3\varepsilon.\end{aligned}$$

Siendo $\varepsilon > 0$ arbitrario, se concluye que $|f| \in \mathcal{HK}([a,b])$ y $(*)$ se cumple. ■

El siguiente criterio es una herramienta muy útil a la hora de establecer la integrabilidad absoluta de una función $f \in \mathcal{HK}([a,b])$.

Teorema 14.1.27 (Criterio de Integrabilidad Absoluta). *Sean $f,g \in \mathcal{HK}([a,b])$. Si $|f| \leq g$ sobre $[a,b]$, entonces $|f| \in \mathcal{HK}([a,b])$ y*

$$\left|(\text{HK})\int_a^b f\,dt\right| \leq (\text{HK})\int_a^b |f|\,dt \leq (\text{HK})\int_a^b g\,dt. \tag{1}$$

Prueba. Sea $\mathcal{P} = \{a = x_0, x_1, \ldots, x_n\}$ una partición de $[a,b]$. Si F es la integral indefinida de f, entonces

$$\sum_{i=1}^{n}\left|F(x_i)-F(x_{i-1})\right| = \sum_{i=1}^{n}\left|(\text{HK})\int_{x_{i-1}}^{x_i} f\,dt\right| \leq \sum_{i=1}^{n}(\text{HK})\int_{x_{i-1}}^{x_i} g\,dt = (\text{HK})\int_a^b g\,dt$$

de donde se obtiene que $V(F,[a,b]) < +\infty$. Del Teorema 14.1.26 se concluye que $|f| \in \mathcal{HK}([a,b])$ y (1) se cumple. ■

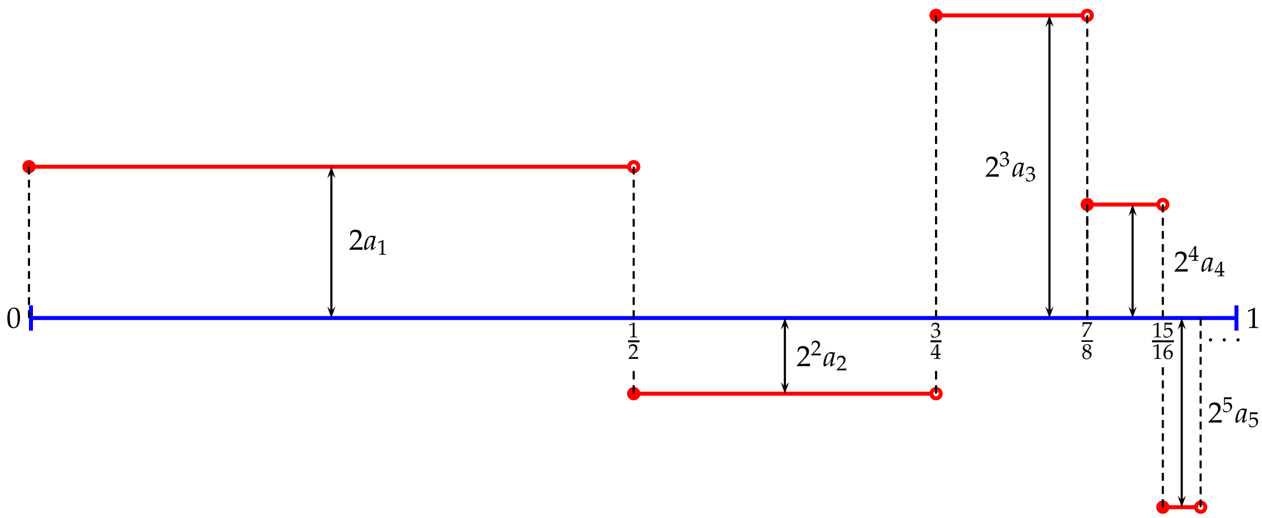
En el siguiente resultado se muestra un método sencillo de construir funciones Henstock-Kurzweil integrables cuyo valor absoluto puede ser o no Henstock-Kurzweil integrable.

Teorema 14.1.28. *Sea $\sum_{n=1}^{\infty} a_n$ una serie de números reales que **converge** en $\mathbb{R}$. Entonces, existe una función $f \in \mathcal{HK}([0,1])$ tal que*

$$(\text{HK})\int_0^1 f\,dt = \sum_{n=1}^{\infty} a_n.$$

Prueba. Por cada $n = 0,1,2,\ldots,$ considere el número $c_n = 1 - 1/2^n$. Observe que la sucesión $(c_n)_{n=0}^{\infty}$ es estrictamente creciente. Defina ahora la función $f : [0,1] \to \mathbb{R}$ por

$$f(x) = \begin{cases} 2^n a_n & \text{si } x \in [c_{n-1}, c_n) \text{ para } n \in \mathbb{N} \\ 0 & \text{si } x = 0. \end{cases}$$



Nótese que la función f toma los valores $2a_1, 2^2a_2, 2^3a_3, \ldots$ sobre los intervalos

$$[0,1/2),\ [1/2,3/4),\ [3/4,7/8),\ \ldots$$

Para cada $c \in (0,1)$, la restricción de f al intervalo $[0,c]$ es una función en escalera y, por consiguiente, Riemann integrable. Más aun, como $\mathcal{R}([0,1]) \subseteq \mathcal{HK}([0,1])$, Teorema 14.1.7, resulta que $f \in \mathcal{HK}([0,1])$. Como los intervalos $[c_0,c_1),[c_1,c_2)\ldots$ son disjuntos y $\bigcup_{n=1}^{\infty}[c_{n-1},c_n) = [0,1)$, entonces existe un único $n \in \mathbb{N}$ tal que $c \in [c_{n-1},c_n)$ y

$$\begin{aligned}
(\mathrm{HK})\int_0^c f\,dt &= (\mathrm{R})\int_0^c f\,dt \\
&= 2a_1(c_1-c_0) + \cdots + 2^{n-1}a_{n-1}(c_{n-1}-c_{n-2}) + 2^n a_n(c-c_{n-1}) \\
&= 2a_1\frac{1}{2} + \cdots + 2^{n-1}a_{n-1}\frac{1}{2^n-1} + 2^n a_n(c-c_{n-1}) \\
&= \sum_{i=1}^{n-1} a_i + 2^n a_n(c-c_{n-1}).
\end{aligned}$$

Sin embrago, como $(c-c_{n-1}) \le 1/2^n$, debemos tener que

$$\left|2^n a_n(c-c_{n-1})\right| \le |a_n|.$$

Teniendo en cuenta que nuestra serie $\sum_{n=1}^{\infty} a_n$ converge, resulta que $\lim_{n\to\infty} a_n = 0$ y, en consecuencia, por la desigualdad anterior, se tiene que $\lim_{n\to\infty} 2^n a_n(c-c_{n-1}) = 0$. Esto implica que

$$\lim_{c\to 1^-} (\mathrm{HK})\int_0^c f\,dt = \lim_{n\to\infty}\sum_{i=1}^{n-1} a_i = \sum_{i=1}^{\infty} a_i,$$

y, entonces, por el Teorema de Hake, $f \in \mathcal{HK}([0,1])$ con integral $(\mathrm{HK})\int_0^1 f\,dt = \sum_{i=1}^{\infty} a_i.$ ■

Nota Adicional 14.1.4 Observe que si la serie $\sum_{n=1}^{\infty} a_n$ es absolutamente convergente, entonces la función f definida en la prueba del Teorema 14.1.28 satisface que $|f| \in \mathcal{HK}([0,1])$ y

$$(\mathrm{HK})\int_0^1 |f|\,dt \;=\; \sum_{n=1}^{\infty} |a_n|.$$

Sin embargo, si la serie $\sum_{n=1}^{\infty} a_n$ es **condicionalmente convergente**, entonces $f \in \mathcal{HK}([0,1])$ pero $|f| \notin \mathcal{HK}([0,1])$. Por ejemplo, la serie $\sum_{n=1}^{\infty}(-1)^{n+1}/n$ es condicionalmente convergente y, por consiguiente, la función

$$f(x) \;=\; \begin{cases} (-1)^{n+1}2^n/n & \text{si } x \in [c_{n-1},\, c_n) \text{ para } n \in \mathbb{N} \\ 0 & \text{si } x = 0, \end{cases}$$

donde $c_n = 1 - 1/2^n$ para $n = 1, 2, \ldots$ es Henstock-Kurzweil integrable pero no así su valor absoluto $|f|$.

Otro aspecto que se deduce del resultado anterior es que toda serie condicionalmente convergente conduce a la existencia de una función que es integrable en el sentido impropio de Riemann pero que no es absolutamente integrable.

El hecho de que la integral de Henstock-Kurzweil no es una integral absoluta, no debe ser pensada como una fatalidad, más bien, esa afortunada circunstancia es lo que permitió la creación de integrales tales como la de Denjoy, Perron, Henstock-Kurzweil y algunas otras sean extensiones propias de las integrales de Lebesgue y de Newton. De hecho, el motivo por el cual la integral de Lebesgue no puede integrar a todas la derivadas reside precisamente en el hecho de ella es absolutamente continua.

14.1.6. La clase $\mathcal{L}_{\mathrm{HK}}([a,b])$

El Criterio de Integrabilidad Absoluta dado en el Teorema 14.1.27, permite introducir una subclase muy importante de funciones en el espacio vectorial $\mathcal{HK}([a,b])$ que remueve algunas propiedades "indeseables". Lo interesante de esta subclase es que ella puede ser identificada con el espacio $\mathrm{L}_1([a,b])$ de las funciones medibles $f : [a,b] \to \mathbb{R}$ que son Lebesgue integrables.

Definición 14.1.29. *Una función $f : [a,b] \to \mathbb{R}$ se dice que es **absolutamente HK-integrable**, o simplemente, **absolutamente integrable** sobre $[a,b]$, si $f, |f| \in \mathcal{HK}([a,b])$.*

En lo que sigue, usaremos el símbolo $\mathcal{L}_{\mathrm{HK}}([a,b])$ para designar el conjunto de *todas las funciones $f : [a,b] \to \mathbb{R}$ que son **absolutamente integrables** sobre $[a,b]$.*

Observe que si $f \in \mathcal{HK}([a,b])$ y $f \geq 0$, entonces $f \in \mathcal{L}_{\mathrm{HK}}([a,b])$. Además, todas las funciones $f : [a,b] \to \mathbb{R}$ que son constantes sobre $[a,b]$ pertenecen a $\mathcal{L}_{\mathrm{HK}}([a,b])$.

Teorema 14.1.30. *$\mathcal{L}_{\mathrm{HK}}([a,b])$ es un **espacio vectorial** sobre $\mathbb{R}$.*

Prueba. Sean $f, g \in \mathcal{L}_{\mathrm{HK}}([a,b])$ y sean $c_1, c_2 \in \mathbb{R}$. Como $\mathcal{HK}([a,b])$ es un espacio vectorial, resulta que $c_1 f + c_2 g \in \mathcal{HK}([a,b])$. Por otro lado, gracia a la desigualdad triangular se tiene que

$$|c_1 f + c_2 g| \;\leq\; |c_1||f| + |c_2||g|$$

y, en consecuencia $|c_1 f + c_2 g| \in \mathcal{HK}([a,b])$ ya que $|f|, |g| \in \mathcal{HK}([a,b])$. En conclusión, hemos demostrado que $c_1 f + c_2 g \in \mathcal{L}_{\mathrm{HK}}([a,b])$ y termina la prueba. ■

Una de las propiedades que puede ser pensada como"indeseable" en el espacio $\mathcal{HK}([a,b])$ es la siguiente: si $f \in \mathcal{HK}([a,b])$, entonces no es cierto, en general, que f^+ y f^- sean Henstock-Kurzweil integrables. Sin embargo, si $f \in \mathcal{L}_{\mathrm{HK}}([a,b])$, entonces $f^+, f^- \in \mathcal{HK}([a,b])$ ya que

$$f^+ = \frac{|f| + f}{2} \qquad \text{y} \qquad f^- = \frac{|f| - f}{2} \tag{1a}$$

y entonces usar el hecho de que $\mathcal{HK}([a,b])$ es un espacio vectorial. De modo más general vale el siguiente resultado.

Teorema 14.1.31. *Sea $f \in \mathcal{HK}([a,b])$. Las siguientes condiciones son equivalentes:*

(1) $f \in \mathcal{L}_{\mathrm{HK}}([a,b])$.

(2) *Existe una función $\beta \in \mathcal{L}_{\mathrm{HK}}([a,b])$ tal que $f(x) \leq \beta(x)$ para cada $x \in [a,b]$.*

(3) *Existe una función $\alpha \in \mathcal{L}_{\mathrm{HK}}([a,b])$ tal que $\alpha(x) \leq f(x)$ para cada $x \in [a,b]$.*

(4) $f^+, f^- \in \mathcal{HK}([a,b])$.

Prueba. $(1) \Rightarrow (2)$ Basta tomar $\beta = f$.
$(2) \Rightarrow (1)$ Suponga que (2) se cumple. Como $\beta - f \in \mathcal{HK}([a,b])$ y $\beta - f \geq 0$, resulta que $\beta - f \in \mathcal{L}_{\mathrm{HK}}([a,b])$. Finalmente, puesto que $f = \beta - (\beta - f)$ y $\mathcal{L}_{\mathrm{HK}}([a,b])$ es un espacio vectorial, se tiene que $f \in \mathcal{L}_{\mathrm{HK}}([a,b])$.
$(2) \Leftrightarrow (3)$ Es inmediata.
$(1) \Rightarrow (4)$ Sigue de las igualdades dadas en $(1a)$.
$(4) \Rightarrow (1)$ Sigue de la igualdad $f = f^+ - f^-$ y de que $\mathcal{L}_{\mathrm{HK}}([a,b])$ es un espacio vectorial. ■

Del hecho de que toda función constante definida sobre $[a,b]$ pertenece a $\mathcal{L}_{\mathrm{HK}}([a,b])$ y del resultado anterior se sigue que:

Corolario 14.1.32. *Sea $f : [a,b] \to \mathbb{R}$ una **función acotada** sobre $[a,b]$. Las siguientes condiciones son equivalentes:*

(1) $f \in \mathcal{HK}([a,b])$.

(2) $f \in \mathcal{L}_{\mathrm{HK}}([a,b])$.

Corolario 14.1.33. *Si $f \in \mathcal{L}_{\mathrm{HK}}([a,b])$ y $g : [a,b] \to \mathbb{R}$ es una **función acotada** sobre $[a,b]$, entonces $fg \in \mathcal{HK}([a,b])$.*

Prueba. Sea $M > 0$ tal que $|g| \leq M$ sobre $[a,b]$ y considere las funciones $\alpha(x) = -M|f(x)|$ y $\beta(x) = M|f(x)|$ para todo $x \in [a,b]$. Entonces $\alpha(x) \leq f(x)g(x) \leq \beta(x)$ para todo $x \in [a,b]$ y aplicando el Teorema 14.1.31 se tiene que $fg \in \mathcal{HK}([a,b])$. ■

Corolario 14.1.34. *Sean $f, g \in \mathcal{HK}([a,b])$. Las siguientes condiciones son equivalentes*

(1) $f, g \in \mathcal{L}_{\mathrm{HK}}([a,b])$.

(2) $f \vee g = \text{máx}\{f, g\} \in \mathcal{L}_{\mathrm{HK}}([a,b])$.

(3) $f \wedge g = \text{mín}\{f, g\} \in \mathcal{L}_{\mathrm{HK}}([a,b])$.

Prueba. (1) $\Rightarrow$ (2) De la igualdad

$$f \vee g \;=\; \text{máx}\{f, g\} \;=\; \frac{1}{2}\big[(f+g) + |f-g|\big]$$

y del hecho de que $(f+g)$ y $|f-g|$ son absolutamente integrables y $\mathcal{L}_{\mathrm{HK}}([a,b])$ es un espacio vectorial, se tiene que $f \vee g \in \mathcal{L}_{\mathrm{HK}}([a,b])$.

(2) $\Rightarrow$ (1) Puesto que $f, g \le f \vee g$, el resultado sigue del Teorema 14.1.31.

(2) $\Leftrightarrow$ (3) sigue de la igualdad $f \wedge g = -\,\text{máx}\{-f, -g\}$. ■

El siguiente resultado se aplica a funciones que están en $\mathcal{HK}([a,b])$ y no únicamente en $\mathcal{L}_{\mathrm{HK}}([a,b]$.

Teorema 14.1.35. *Sean $f, g, \alpha, \beta \in \mathcal{HK}([a,b])$.*

(a) Si $f \le \beta$ y $g \le \beta$, entonces $f \vee g$ y $f \wedge g$ están en $\mathcal{HK}([a,b])$.

(b) Si $\alpha \le f$ y $\alpha \le g$, entonces $f \vee g$ y $f \wedge g$ están en $\mathcal{HK}([a,b])$.

Prueba. (a) La hipótesis implica que $f \vee g \le \beta$ y de la igualdad

$$f \vee g \;=\; \frac{1}{2}\big[(f+g) + |f-g|\big]$$

se sigue que $0 \le |f-g| = 2(f \vee g) - f - g \le 2\beta - f - g$. Por el Criterio de Integrabilidad Absoluta, Teorema 14.1.27, se tiene que $|f-g|$ es Henstock-Kurzweil integrable y como $\mathcal{HK}([a,b])$ es un espacio vectorial, se concluye que $f \vee g$ y $f \wedge g$ están en $\mathcal{HK}([a,b])$. La parte (b) es similar y se deja a cargo del lector. ■

14.1.7. Comparando Integrabilidad: Lebesgue y Henstock-Kurzweil

En esta sección veremos que toda función $f \in \mathrm{L}_1([a,b])$ es Henstock-Kurzweil integrable y, además, se cumple que

$$(\mathrm{L})\int_a^b f\,d\mu \;=\; (\mathrm{HK})\int_a^b f\,dt.$$

Más aun, mostraremos que $\mathrm{L}_1([a,b]) = \mathcal{L}_{\mathrm{HK}}([a,b])$.

Teorema 14.1.36. *Si $f \in \mathrm{L}_1([a,b])$, entonces $f \in \mathcal{HK}([a,b])$ y se cumple que*

$$(\mathrm{L})\int_a^b f\,d\mu \;=\; (\mathrm{HK})\int_a^b f\,dt. \tag{1}$$

En particular,

$$\mathrm{L}_1([a,b]) \;\subseteq\; \mathcal{L}_{\mathrm{HK}}([a,b]).$$

Prueba. Sea $f \in \mathrm{L}_1([a,b])$ y fijemos un $\varepsilon > 0$. El Teorema de Vitali-Carathéodory, Teorema 10.2.46, página 619, permite obtener una función inferiormente semicontinua $u > f$ y una función superiormente semicontinua $v < f$ tal que $u, v \in \mathrm{L}_1([a,b])$ y

$$(\mathrm{L})\int_a^b (u - v)\,d\mu \;<\; \varepsilon.$$

Para cada $x \in [a,b]$, encuentre un $\delta(x) > 0$ de modo tal que $v(t) < f(x) < u(t)$ para todo $t \in (x - \delta(x), x + \delta(x)) \cap [a,b]$. Claramente δ es un calibrador sobre $[a,b]$. Sea ahora $\mathcal{P} = ((t_i, [x_{i-1}, x_i]))_{i=1}^n$ una partición etiquetada de $[a,b]$ subordinada a δ y observe que para cada $i \in \{1, \ldots, n\}$,

$$\text{(L)}\int_{x_{i-1}}^{x_i} v\, d\mu \;\leq\; f(t_i)(x_i - x_{i-1}) \;\leq\; \text{(L)}\int_{x_{i-1}}^{x_i} u\, d\mu.$$

Sumando se tiene que

$$\text{(L)}\int_a^b v\, d\mu \;\leq\; \sum_{i=1}^n f(t_i)(x_i - x_{i-1}) \;\leq\; \text{(L)}\int_a^b u\, d\mu.$$

Por otro lado, como $v < f < u$, entonces

$$\text{(L)}\int_a^b v\, d\mu \;\leq\; \text{(L)}\int_a^b f\, d\mu \;\leq\; \text{(L)}\int_a^b u\, d\mu.$$

y, en consecuencia,

$$\left| S(f, \mathcal{P}) \;-\; \text{(L)}\int_a^b f\, d\mu \right| \;<\; \varepsilon.$$

Esto nos dice que $f \in \mathcal{HK}([a,b])$ y (1) se cumple.

Finalmente, observe que si $f \in \mathrm{L}_1([a,b])$, entonces $|f| \in \mathrm{L}_1([a,b])$ y se sigue de la primera parte que $|f| \in \mathcal{HK}([a,b])$. Por esto,

$$\mathrm{L}_1([a,b]) \;\subseteq\; \mathcal{L}_{\mathrm{HK}}([a,b])$$

y concluye la prueba. ■

Con el resultado que acabamos de demostrar se constata la importancia de la clase $\mathcal{HK}([a,b])$: la integral de Henstock-Kurzweil es, realmente, una *extensión propia de la integral de Lebesgue*.

$$\begin{tikzcd}
C([a,b]) \arrow[r, phantom, "\subseteq"] \arrow[dd, phantom, sloped, "\subseteq"] & \mathcal{N}([a,b]) \arrow[dr, phantom, sloped, "\subseteq"] & \\
& & \mathcal{HK}([a,b]) \\
\mathcal{R}([a,b]) \arrow[r, phantom, "\subseteq"] & \mathrm{L}_1([a,b]) \arrow[ur, phantom, sloped, "\subseteq"] &
\end{tikzcd}$$

El siguiente resultado es fundamental para establecer que, efectivamente, las clases $\mathrm{L}_1([a,b])$ y $\mathcal{L}_{\mathrm{HK}}([a,b])$ son idénticas.

Teorema 14.1.37. *Sea $f \in \mathcal{HK}([a,b])$.*

*(a) Si f es **acotada** sobre $[a,b]$, entonces $f \in \mathrm{L}_1([a,b])$.*

*(b) Si f es **no-negativa** sobre $[a,b]$, entonces $f \in \mathrm{L}_1([a,b])$.*

Prueba. (a) Suponga que f es acotada. Puesto que también f es medible por el Teorema 14.1.24, entonces f es Lebesgue integrable sobre $[a,b]$. Para demostrar (b), suponga que f es no-negativa sobre $[a,b]$. Por supuesto, si f es acotada, entonces f es Lebesgue integral sobre $[a,b]$. Suponga que f no es acotada y sea $F : [a,b] \to \mathbb{R}$ la integral indefinida de f, es decir,

$$F(x) \;=\; \text{(HK)}\int_a^b f\, dt \quad \text{para cada } x \in [a,b]$$

Por el Corolario 14.1.17 sabemos que F es creciente y así, por una aplicación del Lema 10.2.35, página 608, se tiene que $F' \in \mathrm{L}_1([a,b])$. Finalmente, gracias al Segundo Teorema Fundamental del Cálculo para la integral de Henstock-Kurzweil, Teorema 14.1.24, página 875, tenemos que $F' = f$ casi-siempre en $[a,b]$ y, por consiguiente, $f \in \mathrm{L}_1([a,b])$. ■

Teorema 14.1.38. *Si $f \in \mathcal{L}_{\mathrm{HK}}([a,b])$, entonces f es* ***Lebesgue integrable*** *sobre $[a,b]$. Por consiguiente,*

$$\mathcal{L}_{\mathrm{HK}}([a,b]) \;=\; \mathrm{L}_1([a,b]) \qquad y \qquad (\mathrm{HK})\int_a^b f\,dt \;=\; (\mathrm{L})\int_a^b f\,d\mu$$

para cualquier $f \in \mathcal{L}_{\mathrm{HK}}([a,b]$.

Prueba. Sea $f \in \mathcal{L}_{\mathrm{HK}}([a,b])$. Puesto que $|f| \in \mathcal{HK}([a,b])$ y $|f| \geq 0$, el Teorema 14.1.37 nos garantiza que $|f| \in \mathrm{L}_1([a,b])$. De nuevo, como $|f| - f \in \mathcal{HK}([a,b])$ y $|f| - f \geq 0$, por una nueva aplicación del Teorema 14.1.37 se tiene que $|f| - f \in \mathrm{L}_1([a,b])$ y, en consecuencia, $f = |f| - (|f| - f) \in \mathrm{L}_1([a,b])$. Esto prueba que

$$\mathcal{L}_{\mathrm{HK}}([a,b]) \subseteq \mathrm{L}_1([a,b]).$$

Un llamado al Teorema 14.1.36 termina la prueba. ■

Del Corolario 14.1.32 y el resultado anterior se tiene que

$$\mathrm{L}_1([a,b]) \;=\; \mathcal{L}_{\mathrm{HK}}([a,b]) \;=\; \mathcal{HK}([a,b]) \cap \mathcal{B}_\infty([a,b]),$$

donde $\mathcal{B}_\infty([a,b])$ es el espacio vectorial de todas las funciones $f : [a,b] \to \mathbb{R}$ que son acotadas sobre $[a,b]$.

El Teorema 14.1.38 más el hecho de que $\mathrm{L}_1([a,b]) \subsetneqq \mathcal{HK}([a,b])$ permiten visualizar que, efectivamente, $\mathcal{HK}([a,b])$ califica como un espacio super interesante. Falta por indagar cómo se comportan sus sucesiones en el intercambio del límite, es decir, si en dicho espacio se verifican teoremas tales como el de la Convergencia Monótona, la Convergencia Dominada, etc. Este será nuestro análisis en la próxima sección.

Otra manera de caracterizar a las funciones absolutamente integrables es por medio del siguiente resultado.

Teorema 14.1.39. *Sea $f \in \mathcal{HK}([a,b])$ y sea F su* ***integral indefinida****. La siguientes condiciones son equivalentes:*

(1) $f \in \mathcal{L}_{\mathrm{HK}}([a,b])$.

(2) F es ***absolutamente continua*** *sobre $[a,b]$.*

(3) F es de ***variación acotada*** *sobre $[a,b]$.*

Prueba. $(1) \Rightarrow (2)$ Suponga que $f \in \mathcal{L}_{\mathrm{HK}}([a,b])$. Por el Teorema 14.1.38, $f \in \mathrm{L}_1([a,b])$ y

$$F(x) \;=\; (\mathrm{HK})\int_a^x f\,dt \;=\; (\mathrm{L})\int_a^x f\,dt \quad \text{para cada } x \in [a,b]$$

Se sigue del Segundo Teorema Fundamental del Cálculo para la integral de Lebesgue, Teorema 10.2.39, página 612, que F es absolutamente continua.

$(2) \Rightarrow (3)$ es el Teorema 9.1.49, página 501.

$(3) \Rightarrow (1)$ Por el Teorema 14.1.26 sabemos que $|f| \in \mathcal{HK}([a,b])$ y como $f \in \mathcal{HK}([a,b])$, entonces $f \in \mathcal{L}_{\mathrm{HK}}([a,b])$. ∎

Ya vimos que si $f \in \mathrm{L}_1([a,b]) = \mathcal{L}_{\mathrm{HK}}([a,b])$, entonces $f \cdot \chi_E \in \mathrm{L}_1([a,b])$ para cualquier conjunto medible $E \subseteq [a,b]$. Sin embargo, la misma conclusión no es válida si reemplazamos el espacio $\mathrm{L}_1([a,b])$ por $\mathcal{HK}([a,b])$. En este sentido, lo mejor que podemos afirmar es lo siguiente.

Corolario 14.1.40. *Sea $f : [a,b] \to \mathbb{R}$ una función. Si f es* ***Henstock-Kurzweil integrable*** *sobre cualquier conjunto medible $E \subseteq [a,b]$, entonces $f \in \mathrm{L}_1([a,b])$.*

Prueba. Suponga que f es Henstock-Kurzweil integrable sobre cualquier conjunto medible $E \subseteq [a,b]$. Puesto que el conjunto $E = \{x \in [a,b] : f(x) \geq 0\}$ es medible y $f \cdot \chi_E = f^+$, entonces f^+ es Henstock-Kurzweil integrable y el Teorema 14.1.37 nos indica que f^+ es Lebesgue integrable sobre $[a,b]$. Un argumento similar nos dice que también f^- es Lebesgue integrable sobre $[a,b]$ y, en consecuencia, $f = f^+ - f^- \in \mathrm{L}_1([a,b])$. ∎

Otra interesante caracterización de las funciones $f : [a,b] \to \mathbb{R}$ que son Lebesgue integrables en función de las integrales de Henstock-Kurzweil viene dada por el siguiente teorema.

Teorema 14.1.41. *Sea $f \in \mathcal{HK}([a,b])$. La siguientes condiciones son equivalentes:*

(1) $f \in \mathrm{L}_1([a,b])$.

(2) $\sup\left\{ \sum_{I \in \mathcal{P}} \left| (\mathrm{HK})\int_I f\,dt \right| \,:\, \mathcal{P} \in \mathbb{P}[a,b] \right\} < +\infty$.

En este caso,

$$(\mathrm{L})\int_{[a,b]} |f|\,d\mu \;=\; \sup\left\{ \sum_{I \in \mathcal{P}} \left| (\mathrm{HK})\int_I f\,dt \right| \,:\, \mathcal{P} \in \mathbb{P}[a,b] \right\}.$$

Prueba. $(1) \Rightarrow (2)$ Suponga que $f \in \mathrm{L}_1([a,b])$ y sea $\mathcal{P}$ una partición de $[a,b]$. Entonces

$$\begin{aligned} \sum_{I \in \mathcal{P}} \left| (\mathrm{HK})\int_I f\,dt \right| &= \sum_{I \in \mathcal{P}} \left| (\mathrm{L})\int_I f\,d\mu \right| \\ &\leq \sum_{I \in \mathcal{P}} (\mathrm{L})\int_I |f|\,d\mu \;=\; (\mathrm{L})\int_{[a,b]} |f|\,d\mu \;<\; +\infty \end{aligned}$$

Puesto que la partición $\mathcal{P}$ es arbitraria se sigue que (2) se cumple y

$$\sup\left\{ \sum_{I \in \mathcal{P}} \left| (\mathrm{HK})\int_I f\,dt \right| \,:\, \mathcal{P} \in \mathbb{P}[a,b] \right\} \;\leq\; (\mathrm{L})\int_{[a,b]} |f|\,d\mu.$$

Para demostrar que $(2) \Rightarrow (1)$ es suficiente verificar, en virtud del Teorema 14.1.38, que la función $|f| \in \mathcal{HK}([a,b])$ y que

$$(\mathrm{HK})\int_{[a,b]} |f|\,d\mu \;=\; \sup\left\{ \sum_{I \in \mathcal{P}} \left| (\mathrm{HK})\int_I f\,dt \right| \,:\, \mathcal{P} \in \mathbb{P}[a,b] \right\}. \tag{2}$$

Denote por A el lado derecho de (2) el cual, por hipótesis, es finito y sea $\varepsilon > 0$. Seleccione una partición $\mathcal{P}_0 = \{J_1, \ldots, J_m\} \in \mathbb{P}[a,b]$ tal que

$$A \;<\; \sum_{J \in \mathcal{P}_0} \left|(\mathrm{HK})\int_J f\,dt\right| \;+\; \frac{\varepsilon}{2}.$$

Por el Lema de Saks-Henstock, existe un calibrador δ sobre $[a,b]$ tal que

$$\sum_{(s,I') \in \mathcal{Q}} \left|f(s)\ell(I') \;-\; (\mathrm{HK})\int_{I'} f\,dt\right| \;+\; \frac{\varepsilon}{2}$$

para cada subpartición etiquetada $\mathcal{Q}$ de $[a,b]$ subordinada a δ. Fijemos una partición etiquetada $\mathcal{P}$ de $[a,b]$ subordinada a δ. Por el Lema 14.1.1 podemos asumir que $\mathcal{Q}$ es más fina que $\mathcal{P}_0$ y que todos los extremos de los intervalos de la partición $\mathcal{P}_0$ son etiquetas de $\mathcal{Q}$. De esto se sigue que

$$(\mathrm{HK})\int_J f\,dt \;=\; \sum_{\substack{(s,I') \in \mathcal{Q} \\ I' \subseteq J}} (\mathrm{HK})\int_{I'} f\,dt \quad \text{para todo } J \in \mathcal{P}_0$$

y, por consiguiente,

$$\begin{aligned}
\left|\sum_{(t,I) \in \mathcal{P}} |f(t)|\ell(I) \;-\; A\right| \;&\leq\; \sum_{(t,I) \in \mathcal{P}} \left||f(t)|\ell(I) \;-\; \left|(\mathrm{HK})\int_I f\,dt\right|\right| \;+\; A \;-\; \sum_{(t,I) \in \mathcal{P}} \left|(\mathrm{HK})\int_I f\,dt\right| \\
&\leq\; \sum_{(t,I) \in \mathcal{P}} \left|f(t)\ell(I) \;-\; (\mathrm{HK})\int_I f\,dt\right| \;+\; A \;-\; \sum_{J \in \mathcal{P}_0} \left|(\mathrm{HK})\int_J f\,dt\right| \\
&<\; \varepsilon.
\end{aligned}$$

Siendo $\varepsilon > 0$ arbitrario, se concluye que $|f| \in \mathcal{HK}([a,b])$ y termina la prueba. ■

Sea $f \in \mathcal{HK}([a,b])$. Si bien es cierto que $f \cdot \chi_E \notin \mathcal{HK}([a,b])$ para todo conjunto medible $E \subseteq [a,b]$ vale, sin embrago, lo siguiente:

Teorema 14.1.42 (Buczolich). *Si $f \in \mathcal{HK}([a,b])$, entonces existe un* ***intervalo cerrado*** *no-degenerado $[c,d] \subseteq [a,b]$ tal que $f \in \mathrm{L}_1([c,d])$.*

Prueba. Con $\varepsilon = 1$, apliquemos el Lema de Saks-Henstock para seleccionar un calibrador δ sobre $[a,b]$ tal que

$$\sum_{(t,I) \in \mathcal{P}} \left|f(t)\ell(I) \;-\; (\mathrm{HK})\int_I f\,dt\right| \;<\; 1 \tag{2}$$

para cualquier subpartición etiquetada $\mathcal{P}$ de $[a,b]$ subordinada a δ. Por cada $n \in \mathbb{N}$, considere el conjunto

$$E_n \;=\; \left\{x \in [a,b] \;:\; |f(x)| < n \quad \text{y} \quad \delta(x) > \frac{1}{n}\right\}$$

y sea $F_n = \overline{E}_n$. Puesto que $(F_n)_{n=1}^{\infty}$ es una sucesión creciente de conjuntos cerrados con $\bigcup_{n=1}^{\infty} F_n = [a,b]$, el Teorema de Categoría de Baire nos revela la existencia de un $n_0 \in \mathbb{N}$ tal que $\mathrm{int}(F_{n_0}) \neq \varnothing$ y, por lo tanto, existe un intervalo cerrado no-degenerado $[c,d] \subseteq [a,b]$ tal que $[c,d] \subseteq F_{n_0}$. Sin perder generalidad, asumiremos que $d - c < 1/n_0$.

Afirmamos que $f \in \mathrm{L}_1([c,d])$. Para verificar esto último, sea $\{[u_1,v_1],\ldots,[u_m,v_m]\}$ una partición de $[c,d]$. Por cada $i = 1,\ldots,m$, usemos el hecho de que el intervalo $[u_i,v_i] \subseteq F_{n_0}$, para elegir un $t_i \in E_{n_0} \cap (u_i,v_i)$. Puesto que $\ell([c,d]) < 1/n_0$, se concluye que la colección $\{(t_1,[u_1,v_1]),\ldots,(t_m,[u_m,v_m])\}$ es una partición etiquetada de $[c,d]$ subordinada a $\delta_0 = 1/n_0$ y, en consecuencia, subordinada a δ. Por (2) tenemos que

$$\sum_{i=1}^{m} \left| f(t_i)(v_i - u_i) - (\mathrm{HK})\int_{u_i}^{v_i} f\,dt \right| < 1 \tag{2}$$

y, así,

$$\sum_{i=1}^{m} \left| (\mathrm{HK})\int_{u_i}^{v_i} f\,dt \right| < 1 + \sum_{i=1}^{m} |f(t_i)|(v_i - u_i) < 1 + n_0(v-u).$$

Puesto que $\{[u_1,v_1],\ldots,[u_m,v_m]\}$ es una partición arbitraria de $[c,d]$, el Teorema 14.1.41 nos revela que $f \in \mathrm{L}_1([c,d])$ y termina la prueba. ■

Nota Adicional 14.1.5 Un resultado más fuerte que el anterior se puede obtener. En efecto, Kurzweil y Jarník obtuvieron lo siguiente:

Teorema 14.1.43 (Kurzweil-Jarník). *Si $f \in \mathcal{HK}([a,b])$, entonces existe una sucesión $(F_n)_{n=1}^{\infty}$ de* ***conjuntos cerrados*** *tal que*

$$\bigcup_{n=1}^{\infty} F_n = [a,b] \qquad \text{y} \qquad f \cdot \chi_{F_n} \in \mathrm{L}_1([a,b]) \quad \text{para todo } n \in \mathbb{N}.$$

Una demostración de este teorema se puede ver, por ejemplo, en [90], Theorem 4.2.5, p.111-112.

14.1.8. Los Teoremas de Convergencia en $\mathcal{HK}([a,b])$

Una de las buenas razones por la cual la integral de Lebesgue ha sido favorecida por encima de la integral de Riemann tiene que ver, fundamentalmente, con la validez de los teoremas de convergencia: los poderosos Teoremas de la Convergencia Monótona, de la Convergencia Acotada, de la Convergencia Dominada y algunos otros. En esta sección estudiaremos las versiones de esos teoremas de convergencia que son válidos en $\mathcal{HK}([a,b])$. Aprovechando la maquinaria ya disponible de la integral de Lebesgue, usaremos algunas de ellas para demostrar tales teoremas de convergencia. Existen, por supuestos, textos dedicados única y exclusivamente a la construcción y desarrollo de la integral de Henstock-Kurzweil y donde dichos teoremas son demostrados con sus propias herramientas sin apelar a la integral de Lebesgue.

Teorema 14.1.44 (Convergencia Uniforme). *Sea $(f_n)_{n=1}^{\infty}$ una sucesión en $\mathcal{HK}([a,b])$ que* ***converge uniformemente*** *sobre $[a,b]$ a una función f. Entonces $f \in \mathcal{HK}([a,b])$ y*

$$(\mathrm{HK})\int_a^b f\,dt = \lim_{n\to\infty} (\mathrm{HK})\int_a^b f\,dt. \tag{1}$$

Prueba. Sea $\varepsilon > 0$. La convergencia uniforme garantiza la existencia de entero $N_\varepsilon > 0$ tal que

$$\left|f_n(x) - f(x)\right| < \varepsilon \quad \text{para todo } x \in [a,b] \ \text{ y } \ \text{todo } n \geq N_\varepsilon.$$

De esto se deduce que si $m, n \geq N_\varepsilon$, entonces

$$-2\varepsilon < f_m(x) - f_n(x) < 2\varepsilon \quad \text{para todo } x \in [a,b].$$

El Teorema 14.1.16 en combinación con la linealidad de la integral producen la siguiente desigualdad:

$$-2\varepsilon(b-a) \leq (\mathrm{HK})\int_a^b f_m\,dt - (\mathrm{HK})\int_a^b f_n\,dt \leq 2\varepsilon(b-a),$$

en otras palabras,

$$\left|(\mathrm{HK})\int_a^b f_m\,dt - (\mathrm{HK})\int_a^b f_n\,dt\right| \leq 2\varepsilon(b-a) \quad \text{para todo } m,n \geq N_\varepsilon.$$

Como $\varepsilon > 0$, resulta que la sucesión $\left((\mathrm{HK})\int_a^b f_n\,dt\right)_{n=1}^{\infty}$ es de Cauchy en $\mathbb{R}$ y, en consecuencia, converge a algún $A \in \mathbb{R}$. Mostraremos ahora que $f \in \mathcal{HK}([a,b])$ y

$$(\mathrm{HK})\int_a^b f\,dt = A = \lim_{n\to\infty} (\mathrm{HK})\int_a^b f_n\,dt.$$

Para ver esto último, sea $\mathcal{Q} = ((t_i, [x_{i-1}, x_i]))_{i=1}^k$ una partición etiquetada de $[a,b]$. Entonces, si $n \geq N_\varepsilon$ se tiene que

$$\begin{aligned}\left|S(f_n,\mathcal{Q}) - S(f,\mathcal{Q})\right| &= \left|\sum_{i=1}^k [f_n(t_i) - f(t_i)](x_i - x_{i-1})\right| \\ &\leq \sum_{i=1}^k \left|f_n(t_i) - f(t_i)\right|(x_i - x_{i-1}) \\ &< \varepsilon \sum_{i=1}^k (x_i - x_{i-1}) = \varepsilon(b-a).\end{aligned}$$

Puesto que $A = \lim_{n\to\infty} (\mathrm{HK})\int_a^b f\,dt$, podemos usar el mismo entero $N_\varepsilon > 0$ de modo que

$$\left|(\mathrm{HK})\int_a^b f_n\,dt - A\right| < \varepsilon$$

para todo $n \geq N_\varepsilon$. Fijemos un $n_0 \geq N_\varepsilon$. Usemos ahora el hecho de que $f_{n_0} \in \mathcal{HK}([a,b])$ para hallar un calibrador δ_{n_0} sobre $[a,b]$ tal que

$$\left|(\mathrm{HK})\int_a^b f_{n_0}\,dt - S(f_{n_0},\mathcal{P})\right| < \varepsilon$$

para cualquier partición etiquetada $\mathcal{P}$ de $[a,b]$ subordinada a δ_{n_0}. Fijemos una tal partición etiquetada $\mathcal{P}$ de $[a,b]$ subordinada a δ_{n_0}. Entonces

$$\begin{aligned}\left|S(f,\mathcal{P}) - A\right| &\leq \left|S(f,\mathcal{P}) - S(f_{n_0},\mathcal{P})\right| + \left|S(f_{n_0},\mathcal{P}) - (\mathrm{HK})\int_a^b f_{n_0}\,dt\right| + \left|(\mathrm{HK})\int_a^b f_{n_0}\,dt - A\right| \\ &< \varepsilon(b-a) + \varepsilon + \varepsilon = \varepsilon(b-a+2).\end{aligned}$$

Siendo $\varepsilon > 0$ se concluye que $f \in \mathcal{HK}([a,b])$ y (1) se cumple. ■

Aunque el resultado anterior es elegante y simple, la exigencia de la convergencia uniforme es muy fuerte y, por consiguiente, restringe su utilidad. Por tal motivo J. Kurzweil ideó otra condición de uniformidad que permite el intercambio del límite. De modo más preciso, Kurzweil sustituyó convergencia uniforme por convergencia simple pero añadió una nueva condición de uniformidad a la sucesión. La idea detrás de este nuevo concepto es que debe existir un sólo calibrador que funcione para todas las funciones de la sucesión.

Definición 14.1.45. *Una sucesión $(f_n)_{n=1}^{\infty}$ en $\mathcal{HK}([a,b])$ se dice **equi-integrable** sobre $[a,b]$ si, para cada $\varepsilon > 0$, existe un calibrador δ_ε sobre $[a,b]$ tal que*

$$\left| S(f_n, \mathcal{P}) - (\mathrm{HK})\int_a^b f_n\, dt \right| < \varepsilon$$

para todo $n \in \mathbb{N}$ y cualquier partición etiquetada $\mathcal{P}$ de $[a,b]$ subordinada a δ_ε.

El término **sucesión uniformemente HK-integrable** es otra forma muy común de nombrar a una sucesión equi-integrable. El siguiente resultado fue demostrado por R. A. Gordon.

Teorema 14.1.46 (Gordon). *Sea $(f_n)_{n=1}^{\infty}$ una sucesión en $\mathcal{HK}([a,b])$ tal que $\lim\limits_{n\to\infty} f_n(x) = f(x)$ para cada $x \in [a,b]$. Entonces*

$$f \in \mathcal{HK}([a,b]) \qquad y \qquad (\mathrm{HK})\int_a^b f\, dt = \lim_{n\to\infty} (\mathrm{HK})\int_a^b f_n\, dt. \tag{1}$$

*si, y sólo si, para cada $\varepsilon > 0$, existe un **calibrador** δ sobre $[a,b]$ con la siguiente propiedad: si $\mathcal{P}$ es cualquier partición etiquetada de $[a,b]$ subordinada a δ, entonces existe un $N_{\mathcal{P}} \in \mathbb{N}$ tal que*

$$\left| S(f_n, \mathcal{P}) - (\mathrm{HK})\int_a^b f_n\, dt \right| \le \varepsilon \quad \textit{para todo } n \ge N_{\mathcal{P}}. \tag{2}$$

Prueba. Suponga que (1) se cumple y sea $\varepsilon > 0$. Puesto que

$$(\mathrm{HK})\int_a^b f\, dt = \lim_{n\to\infty} (\mathrm{HK})\int_a^b f_n\, dt$$

existe un $N_0 \in \mathbb{N}$ tal que si $n \ge N_0$, entonces

$$\left| (\mathrm{HK})\int_a^b f_n\, dt - (\mathrm{HK})\int_a^b f\, dt \right| < \varepsilon/3.$$

También, como $f \in \mathcal{HK}([a,b])$, existe un calibrador δ sobre $[a,b]$ tal que

$$\left| (\mathrm{HK})\int_a^b f\, dt - S(f, \mathcal{P}) \right| < \varepsilon/3$$

para cualquier partición etiquetada $\mathcal{P}$ de $[a,b]$ subordinada a δ_ε. Fijemos una partición etiquetada $\mathcal{P} = \{(t_i, [x_{i-1}, x_i]) : i = 1, 2, \ldots, k\}$ de $[a,b]$ subordinada a δ_ε y use el hecho de que $f_n \to f$ puntualmente, para elegir un $N_{\mathcal{P}} \geq N_0$ tal que si $n \geq N_{\mathcal{P}}$, entonces

$$\left|f_n(t_i) - f(t_i)\right| < \frac{\varepsilon}{3(b-a)} \quad \text{para todo } i \in \{1, 2, \ldots, k\}.$$

Dde lo anterior se tiene que

$$\left|S(f_n, \mathcal{P}) - S(f_m, \mathcal{P})\right| \leq \sum_{i=1}^{k} |f_n(t_i) - f_m(t_i)|(x_i - x_{i-1}) < \varepsilon/3.$$

y, por lo tanto, si $n \geq N_{\mathcal{P}} \geq N_0$, entonces

$$\begin{aligned}
\left|S(f_n, \mathcal{P}) - (\mathrm{HK})\int_a^b f_n\,dt\right| &\leq \left|S(f_n, \mathcal{P}) - S(f, \mathcal{P})\right| + \left|S(f, \mathcal{P}) - (\mathrm{HK})\int_a^b f\,dt\right| \\
&\quad + \left|(\mathrm{HK})\int_a^b f\,dt - (\mathrm{HK})\int_a^b f_n\,dt\right| \\
&< \varepsilon.
\end{aligned}$$

Siendo $\varepsilon > 0$, se tiene que (2) se cumple para todo $n \geq N_{\mathcal{P}}$.

Sea $\varepsilon > 0$ y suponga que existe un calibrador δ sobre $[a,b]$ tal que por cada partición etiquetada $\mathcal{P}$ de $[a,b]$ subordinada a δ, existe un $N_{\mathcal{P}} \in \mathbb{N}$ para el cual (2) se cumple. En primer lugar, vamos a demostrar que la sucesión $\left((\mathrm{HK})\int_a^b f_n\,dt\right)_{n=1}^{\infty}$ es de Cauchy en $\mathbb{R}$. Para ello, fijemos una partición etiquetada $\mathcal{P} = \{(t_i, [x_{i-1}, x_i]) : i = 1, 2, \ldots, k\}$ de $[a,b]$ subordinada a δ. Puesto que $f_n \to f$ puntualmente, elija un entero $N_0 \geq N_{\mathcal{P}}$ tal que si $m, n \geq N_0$, entonces $|f_n(t_i) - f_m(t_i)| < \varepsilon/(b-a)$ para cada $i \in \{1, 2, \ldots .k\}$. Por esto,

$$\left|S(f_n, \mathcal{P}) - S(f_m, \mathcal{P})\right| \leq \sum_{i=1}^{k} |f_n(t_i) - f_m(t_i)|(x_i - x_{i-1}) < \varepsilon.$$

En particular,

$$\left|S(f_n, \mathcal{P}) - S(f, \mathcal{P})\right| = \lim_{m\to\infty} \left|S(f_n, \mathcal{P}) - S(f_m, \mathcal{P})\right| \leq \varepsilon \quad \text{para todo } n \geq N_0.$$

Por último, si $m, n \geq N_0$, entonces

$$\begin{aligned}
\left|(\mathrm{HK})\int_a^b f_n\,dt - (\mathrm{HK})\int_a^b f_m\,dt\right| &\leq \left|(\mathrm{HK})\int_a^b f_n\,dt - S(f_n, \mathcal{P})\right| + \left|S(f_n, \mathcal{P}) - S(f_m, \mathcal{P})\right| \\
&\quad + \left|S(f_m, \mathcal{P}) - (\mathrm{HK})\int_a^b f_m\,dt\right| \\
&\leq \varepsilon + \varepsilon + \varepsilon = 3\varepsilon
\end{aligned}$$

Como $\varepsilon > 0$ es arbitrario, concluimos que $\left((\text{HK})\int_a^b f_n\,dt\right)_{n=1}^{\infty}$ es de Cauchy en $\mathbb{R}$ y, por lo tanto, converge a algún $A \in \mathbb{R}$. Haciendo que $m \to \infty$ en la última desigualdad se obtiene que

$$\left|(\text{HK})\int_a^b f_n\,dt - A\right| \leq 3\varepsilon \quad \text{para todo } n \geq N_0.$$

Veamos ahora que $f \in \mathcal{HK}([a,b])$. En efecto, sea $\mathcal{P}$ una partición etiquetada de $[a,b]$ subordinada a δ y tome cualquier $n \geq N_0 \geq N_{\mathcal{P}}$. Entonces

$$\begin{aligned}\left|S(f,\mathcal{P}) - A\right| &\leq \left|S(f,\mathcal{P}) - S(f_n,\mathcal{P})\right| + \left|S(f_n,\mathcal{P}) - (\text{HK})\int_a^b f_n\,dt\right| + \left|(\text{HK})\int_a^b f_n\,dt - A\right| \\ &\leq \varepsilon + \varepsilon + 3\varepsilon = 5\varepsilon.\end{aligned}$$

Puesto que $\varepsilon > 0$ es arbitrario, resulta que $f \in \mathcal{HK}([a,b])$ con integral A. ■

Si se toma $N_{\mathcal{P}} = 1$ en la condición (2) del Teorema de Gordon se obtiene el siguiente resultado.

Corolario 14.1.47 (Equi-integrabilidad). *Sea $(f_n)_{n=1}^{\infty}$ una sucesión en $\mathcal{HK}([a,b])$* ***equi-integrable*** *sobre $[a,b]$. Si $\lim_{n\to\infty} f_n(x) = f(x)$ para cada $x \in [a,b]$, entonces $f \in \mathcal{HK}([a,b])$ y*

$$(\text{HK})\int_a^b f\,dt = \lim_{n\to\infty} (\text{HK})\int_a^b f_n\,dt.$$

Similar a la convergencia uniforme, la equi-integrabilidad posee sus encantos y, aún así, es difícil de aplicar debido a que no es fácil construir un calibrador δ que funcione para todas los elementos de la sucesión. Sin embargo, como en el caso del espacio $\mathrm{L}_1([a,b])$, se tiene la siguiente extensión del Teorema de la Convergencia Dominada de Lebesgue.

Teorema 14.1.48 (Convergencia Dominada). *Sea $(f_n)_{n=1}^{\infty}$ una sucesión en $\mathcal{HK}([a,b])$ que* ***converge puntualmente*** *sobre $[a,b]$ a una función f. Si existen funciones $\alpha, \beta \in \mathcal{HK}([a,b])$ tales que*

$$\alpha(x) \leq f_n(x) \leq \beta(x) \quad \text{para cada } x \in [a,b] \quad \text{y} \quad \text{todo } n \in \mathbb{N},$$

entonces $(f_n)_{n=1}^{\infty}$ es ***equi-integrable*** *sobre $[a,b]$. Por consiguiente, $f \in \mathcal{HK}([a,b])$ y*

$$(\text{HK})\int_a^b f\,dt = \lim_{n\to\infty} (\text{HK})\int_a^b f_n\,dt.$$

Prueba. Sea $\varphi = \beta - \alpha$ y observe que

$$\left|f_n - f_m\right| \leq \varphi \quad \text{para todo } m,n \in \mathbb{N}.$$

Puesto que $\varphi \in \mathcal{HK}([a,b])$ y es no-negativa, el Teorema 14.1.37 nos muestra que $\varphi \in \mathrm{L}_1([a,b])$. Más aun, como la sucesión $(f_n - f_1)_{n=1}^{\infty}$ es medible y $|f_n - f_1| \leq \varphi$ para todo $n \in \mathbb{N}$, el Teorema de la Convergencia Dominada de Lebesgue, Teorema 10.1.40, página 561, nos garantiza que cada una de las funciones $f_n - f_1$, así como $f - f_1$, pertenecen a $\mathrm{L}_1([a,b])$ y

$$(\text{L})\int_a^b (f - f_1)\,d\mu = \lim_{n\to\infty} (\text{L})\int_a^b (f_n - f_1)\,d\mu.$$

Usando ahora el hecho de que $L_1([a,b]) \subseteq \mathcal{HK}([a,b])$ se tiene que $f - f_1 \in \mathcal{HK}([a,b])$ y, por lo tanto, $f = (f - f_1) + f_1 \in \mathcal{HK}([a,b])$. Observe que no podemos afirmar que $f \in L_1([a,b])$ pues no tenemos la certeza de que $f_1 \in L_1([a,b])$. Lo que si podemos garantizar es que la sucesión $\big((\mathrm{HK})\int_a^b f_n\,dt\big)_{n=1}^{\infty}$ converge. En efecto, como $f_n = f_1 + (f_n - f_1)$, entonces

$$\lim_{n\to\infty} (\mathrm{HK})\int_a^b f_n\,dt \;=\; (\mathrm{HK})\int_a^b f_1\,dt \;+\; \lim_{n\to\infty} (\mathrm{HK})\int_a^b (f_n - f_1)\,dt.$$

lo cual prueba nuestra afrimación. Sea $\varepsilon > 0$. Para demostrar que la sucesión $(f_n)_{n=1}^{\infty}$ es, efectivamente, equi-integrable, empecemos por considerar la integral indefinida (según Lebesgue) de la función φ, es decir,

$$F(x) \;=\; (\mathrm{L})\int_a^x \varphi\,d\mu \quad \text{para todo } x \in [a,b].$$

Por el Segundo Teorema Fundamental del Cálculo para la integral de Lebesgue, Teorema 10.2.39, página 612, sabemos que F es absolutamente continua sobre $[a,b]$ y, por consiguiente, existe $\eta > 0$ tal que

$$\left|\sum_{i=1}^{k} F(d_i) - F(c_i)\right| \;<\; \varepsilon$$

para cualquier colección finita $\{[c_i,d_i] : i = 1,\dots,k\}$ de subintervalos cerrados no-superpuestos incluidos en $[a,b]$ que satisfaga $\sum_{i=1}^{k}(d_i - c_i) < \eta$. Por el Teorema de Severini-Egoroff, Teorema 7.2.12, página 386, existe un conjunto abierto $O \subseteq [a,b]$ con $\mu(O) < \eta$ tal que $f_n \to f$ uniformemente sobre $[a,b] \setminus O$. Fijemos un $x \in [a,b] \setminus O$. Puesto que las sucesiones $\big((\mathrm{HK})\int_a^b f_n\,dt\big)_{n=1}^{\infty}$ y $(f_n(x))_{n=1}^{\infty}$ son ambas convergentes, en particular, de Cauchy, podemos elegir un entero $N \geq 1$ de modo que

$$\left|(\mathrm{HK})\int_a^b f_n\,d\mu \;-\; (\mathrm{HK})\int_a^b f_m\,d\mu\right| \;<\; \varepsilon \qquad \text{y} \qquad \big|f_n(x) \;-\; f_m(x)\big| \;<\; \varepsilon$$

para todo $m,n \geq N$. Por otro lado, como $\{f_1,\dots,f_N,\varphi\} \subseteq \mathcal{HK}([a,b])$, exsite un calibrador δ_φ sobre $[a,b]$ tal que para cada $n \in \{1,\dots,N\}$,

$$\left|S(f_n,\mathcal{P}) \;-\; (\mathrm{HK})\int_a^b f_n\,dt\right| \;<\; \varepsilon \qquad \text{y} \qquad \left|S(\varphi,\mathcal{P}) \;-\; (\mathrm{HK})\int_a^b \varphi\,dt\right| \;<\; \varepsilon$$

cualquiera sea $\mathcal{P}$ una partición etiquetada de $[a,b]$ subordinada a δ_φ. Lo anterior nos permite definir el siguiente calibrador δ sobre $[a,b]$:

$$\delta(x) \;=\; \begin{cases} \delta_\varphi(x) & \text{si } x \in [a,b] \setminus O, \\[1ex] \text{mín}\,\big\{\delta_\varphi(x),\, \mathrm{dist}(x,\, [a,b] \setminus O)\big\} & \text{si } x \in O. \end{cases}$$

Tomemos ahora cualquier partición etiquetada de $[a,b]$, digamos $\mathcal{P}$, subordinada a δ y fijemos un $n > N$. Sea $\mathcal{P}_1$ el subconjunto de $\mathcal{P}$ cuyas etiquetas viven en $[a,b] \setminus O$ y sea $\mathcal{P}_2 = \mathcal{P} \setminus \mathcal{P}_1$.

Usando el Lema de Saks-Henstock y el hecho de que $\mu(\mathcal{P}_2) < \eta$ se obtiene que

$$\begin{aligned}
\left|S(f_n, \mathcal{P}) - S(f_N, \mathcal{P})\right| &\leq \left|S(f_n, \mathcal{P}_1) - S(f_N, \mathcal{P}_1)\right| + \left|S(f_n, \mathcal{P}_2) - S(f_N, \mathcal{P}_2)\right| \\
&\leq \varepsilon\mu(\mathcal{P}_1) + S(\varphi, \mathcal{P}_2) \\
&\leq \varepsilon(b-a) + \left|S(\varphi, \mathcal{P}_2) - S(F, \mathcal{P}_2)\right| + \left|S(F, \mathcal{P}_2)\right| \\
&< \varepsilon(b-a) + \varepsilon + \varepsilon.
\end{aligned}$$

De esto último se deduce que

$$\begin{aligned}
\left|S(f_n, \mathcal{P}) - (\mathrm{HK})\int_a^b f_n\, dt\right| &\leq \left|S(f_n, \mathcal{P}) - S(f_N, \mathcal{P})\right| + \left|S(f_N, \mathcal{P}) - (\mathrm{HK})\int_a^b f_N\, dt\right| \\
&\quad + \left|(\mathrm{HK})\int_a^b f_N\, dt - (\mathrm{HK})\int_a^b f_n\, dt\right| \\
&< \varepsilon(b - a + 2) + \varepsilon + \varepsilon.
\end{aligned}$$

Esto termina la prueba. ■

Observe que si en el teorema anterior $\alpha, \beta \in \mathcal{L}_{\mathrm{HK}}([a,b])$ y si definimos $\varphi = \max\{|\alpha|, |\beta|\}$, entonces $\varphi \in \mathcal{L}_{\mathrm{HK}}([a,b])$ y, en consecuencia, podemos reemplazar la condición

$$\alpha \leq f_n \leq \beta \quad \text{para todo } n \geq 1,$$

por esta otra

$$|f_n| \leq \varphi \quad \text{para todo } n \geq 1.$$

Teorema 14.1.49 (Convergencia Monótona). *Sea $(f_n)_{n=1}^{\infty}$ una **sucesión monótona** en $\mathcal{HK}([a,b])$ tal que $f(x) = \lim\limits_{n\to\infty} f_n(x)$ para cada $x \in [a,b]$. Si*

$$\lim_{n\to\infty} (\mathrm{HK})\int_a^b f_n\, dt < +\infty,$$

*entonces la sucesión $(f_n)_{n=1}^{\infty}$ es **equi-integrable** sobre $[a,b]$. Por consiguiente, $f \in \mathcal{HK}([a,b])$ y*

$$(\mathrm{HK})\int_a^b f\, dt = \lim_{n\to\infty} (\mathrm{HK})\int_a^b f_n\, dt.$$

Prueba. Es suficiente, para demostrar el corolario, suponer que la sucesión $(f_n)_{n=1}^{\infty}$ es creciente. En este caso, resulta que para cada $n \in \mathbb{N}$, la función $f_n - f_1$ es no-negativa y Henstock-Kurzweil integrable sobre $[a,b]$. Se sigue entonces del Teorema 14.1.37 que $f_n - f_1 \in \mathrm{L}_1([a,b])$ y así, aplicando el Teorema de la Convergencia Monótona de Lebesgue, Teorema 10.1.25, página 548, se tiene que $f - f_1 \in \mathrm{L}_1([a,b])$. Usando de nuevo el hecho de que $\mathrm{L}_1([a,b]) \subseteq \mathcal{HK}([a,b])$, se tiene que $f = f_1 + (f - f_1) \in \mathcal{HK}([a,b])$. Finalmente, puesto que $f_1 \leq f_n \leq f$ para todo $n \in \mathbb{N}$, entonces el Teorema 14.1.48 nos revela que la sucesión $(f_n)_{n=1}^{\infty}$ es equi-integrable sobre $[a,b]$ y termina la prueba. ■

Recordemos que si $(f_n)_{n=1}^{\infty}$ es una sucesión de funciones, entonces para cada $n \in \mathbb{N}$, la n-ésima suma parcial de la sucesión viene dada por

$$s_n = f_1 + \cdots + f_n.$$

A $(s_n)_{n=1}^{\infty}$ se le llama la sucesión de las sumas parciales asociada a $(f_n)_{n=1}^{\infty}$.

Corolario 14.1.50 (Beppo-Levi). *Sea $(f_n)_{n=1}^{\infty}$ una sucesión en $\mathcal{HK}([a,b])$ tal que $f_n \geq 0$ para todo entero $n \geq 1$. Suponga que la sucesión $(s_n)_{n=1}^{\infty}$ de las sumas parciales asociada a $(f_n)_{n=1}^{\infty}$ **converge puntualmente** a una función $f : [a,b] \to \mathbb{R}$. Si*

$$\sum_{n=1}^{\infty} (\mathrm{HK})\int_a^b f_n \, dt < +\infty$$

entonces:

*(a) $(s_n)_{n=1}^{\infty}$ es **equi-integrable** sobre $[a,b]$.*

(b) $f \in \mathcal{HK}([a,b])$ y

$$\sum_{n=1}^{\infty} (\mathrm{HK})\int_a^b f_n \, dt = (\mathrm{HK})\int_a^b f \, dt = (\mathrm{HK})\int_a^b \sum_{n=1}^{\infty} f_n \, dt.$$

Prueba. Puesto que la sucesión $(s_n)_{n=1}^{\infty}$ es monótona creciente y

$$(\mathrm{HK})\int_a^b s_n \, dt = \sum_{k=1}^{n} (\mathrm{HK})\int_a^b f_k \, dt \quad \text{para todo } n \geq 1,$$

entonces el resultado sigue del Teorema de la Convergencia Monótona. ■

Lema 14.1.51 (Fatou). *Sea $(f_n)_{n=1}^{\infty}$ una sucesión en $\mathcal{HK}([a,b])$ y suponga que para alguna función $\alpha \in \mathcal{HK}([a,b])$ se cumple que $\alpha \leq f_n$ todo $n \in \mathbb{N}$. Si*

$$\liminf_{n\to\infty} (\mathrm{HK})\int_a^b f_n \, dt < +\infty,$$

entonces $\liminf\limits_{n\to\infty} f_n \in \mathcal{HK}([a,b])$ y

$$(\mathrm{HK})\int_a^b \liminf_{n\to\infty} f_n \, dt \leq \liminf_{n\to\infty} (\mathrm{HK})\int_a^b f_n \, dt.$$

Prueba. Sea $n \in \mathbb{N}$. Por cada entero $k \geq n$, defina la función $g_{n,k} : [a,b] \to \mathbb{R}$ por

$$g_{n,k} = \min\{f_n, f_{n+1}, \ldots, f_k\}.$$

y observe que, gracias al Teorema 14.1.35, $g_{n,k} \in \mathcal{HK}([a,b])$. Como la sucesión $(g_{n,k})_{k\geq n}$ es creciente,

$$(\mathrm{HK})\int_a^b g_{n,k} \, dt \leq (\mathrm{HK})\int_a^b f_n \, dt \quad \text{para todo } k \geq n, \tag{1}$$

y $\lim_{n\to\infty} g_{n,k} = \liminf_{n\to\infty} f_n$ puntualmente, se sigue del Teorema de la Convergencia Monótona que $\liminf_{n\to\infty} f_n \in \mathcal{HK}([a,b])$ y

$$\begin{aligned}(\mathrm{HK})\int_a^b \liminf_{n\to\infty} f_n\,dt &= \lim_{n\to\infty}(\mathrm{HK})\int_a^b g_{n,k}\,dt\\ &\le \liminf_{n\to\infty}(\mathrm{HK})\int_a^b f_n\,dt,\end{aligned}$$

donde la última desigualdad proviene de (1). ■

14.1.9. La Norma de Alexiewicz

Si $f \in \mathrm{L}_1([a,b])$, entonces la L_1-norma de f se define como

$$\|f\|_1 = (\mathrm{L})\int_a^b |f|\,d\mu.$$

Aunque $f \in \mathcal{HK}([a,b])$, resulta que

$$(\mathrm{HK})\int_a^b |f|\,dt.$$

carece de sentido si $|f| \notin \mathcal{HK}([a,b])$. A pesar de este inconveniente, el espacio $\mathcal{HK}([a,b])$ posee una semi-norma natural. En efecto, si $f \in \mathcal{HK}([a,b])$, entonces la **semi-norma de Alexiewicz** de f se define como

$$\|f\|_{\mathrm{A}} = \sup\left\{\left|(\mathrm{HK})\int_a^x f\,dt\right| \,:\, a \le x \le b\right\}.$$

De las propiedades del espacio $\mathcal{HK}([a,b])$ se comprueba sin dificultad que $\|\cdot\|_{\mathrm{A}}$ es una semi-norma sobre $\mathcal{HK}([a,b])$. Si se identifican las funciones en $\mathcal{HK}([a,b])$ que son iguales casi-siempre, resulta entonces que $\|\cdot\|_{\mathrm{A}}$ es una norma sobre $\mathcal{HK}([a,b])$.

Por el Teorema de Riesz-Fischer, Teorema 10.3.7, página 634, sabemos que $\mathrm{L}_1([a,b])$ es completo bajo la norma $\|\cdot\|_1$. Ahora mostraremos que $\mathcal{HK}([a,b])$ no es completo bajo la norma de Alexiewicz.

Ejemplo 14.1.6. *$\mathcal{HK}([a,b])$ **no es completo** bajo la norma de Alexiewicz.*

Prueba. Sin perder generalidad en el razonamiento, supondremos que $[a,b] = [0,1]$. Considere una función $F : [0,1] \to \mathbb{R}$ que sea **continua** y **nunca-diferenciable** sobre $[0,1]$ con $F(0) = 0$. Seleccionemos, usando el Teorema de Aproximación de Weierstrass, página 191, una sucesión de polinomios $(p_n)_{n=1}^{\infty}$ tal que $p_n \to F$ uniformemente sobre $[0,1]$ y $p_n(0) = 0$ para todo $n \in \mathbb{N}$. Por el Teorema Fundamental del Cálculo para la integral de Henstock-Kurzweil,

$$p_n(x) = (\mathrm{HK})\int_0^x p_n'\,dt \quad \text{para cualquier } x \in [0,1].$$

Por esto,

$$\begin{aligned}\|p_m' - p_n'\|_{\mathrm{A}} &= \sup\left\{\left|(\mathrm{HK})\int_a^x (p_m' - p_n')\,dt\right| \,:\, a \le x \le b\right\}\\ &= \sup\left\{|(p_m - p_n)(x)| \,:\, a \le x \le b\right\}\end{aligned}$$

Puesto que $(p_n)_{n=1}^{\infty}$ converge uniformemente a F, se sigue que $(p_n')_{n=1}^{\infty}$ es una sucesión de Cauchy en $\mathcal{HK}([0,1])$ con respecto a la norma de Alexiewicz. Suponga, para generar una contradicción, que existe una función $g \in \mathcal{HK}([0,1])$ tal que $\| p_n' - g \|_{\mathrm{A}} \to 0$ cuando $n \to \infty$. Entonces

$$p_n(x) \;=\; (\mathrm{HK})\!\int_0^x p_n'\,dt \;\to\; (\mathrm{HK})\!\int_0^x g\,dt \quad \text{uniformemente en } x \in [0,1]$$

de modo que

$$F(x) = (\mathrm{HK})\!\int_0^x g\,dt \quad \text{para todo } x \in [0,1].$$

Una aplicación del Segundo Teorema Fundamental del Cálculo, Teorema 14.1.24, nos indica que F es diferenciable casi-siempre lo cual es absurdo ya que habíamos supuesto que F no era derivable en ningún punto de $[0,1]$. De allí que $(\mathcal{HK}([0,1]), \|\cdot\|_{\mathrm{A}})$ no es completo. ■

Aunque $(\mathcal{HK}([a,b]), \|\cdot\|_{\mathrm{A}})$ no es un espacio de Banach, se puede demostrar (véase, por ejemplo, [125], Chap. 7) que dicho espacio posee algunas otras propiedades interesantes. Por ejemplo:

(a) El dual de $(\mathcal{HK}([a,b]), \|\cdot\|_{\mathrm{A}})$ es isométricamente isomorfo a $\mathrm{BV}([a,b])$, el espacio de las funciones $f : [a,b] \to \mathbb{R}$ que son de variación acotada sobre $[a,b]$.

(b) El espacio de las funciones en escaleras $\mathrm{Esc}([a,b])$ es denso en $(\mathcal{HK}([a,b]), \|\cdot\|_{\mathrm{A}})$

(c) Como espacio topológico, $(\mathcal{HK}([a,b]), \|\cdot\|_{\mathrm{A}})$ es de primera categoría.

(d) $(\mathcal{HK}([a,b]), \|\cdot\|_{\mathrm{A}})$ es un β-espacio o espacio de Sargent, etc.

14.2. La Integral de McShane

La integral de Henstock-Kurzweil, como acabamos de ver, se construyó de modo muy similar a la integral de Riemann y, por supuesto, no se utilizó ninguna herramienta de la Teoría de la Medida de Lebesgue para su construcción. Existe otra modo de considerar sumas à la Riemann, muy similar a la definición de la integral de Henstock-Kurzweil, que da origen a una nueva integral que, a la postre, resulta ser equivalente a la integral de Lebesgue. La diferencia entre ambas radica, fundamentalmente, en la clase de particiones usadas para formar las sumas de Riemann, o de modo más específico, en la localización de las etiquetas en cada partición. Este enfoque fue ideado por Edward J. McShane (1904-1989) en los años 60's del siglo XX quien la pensó de la siguiente manera: imaginemos que en cualquier partición etiquetada $\mathcal{P} = \{(t_1, [x_0, x_1], (t_2, [x_1, x_2]), \ldots, (t_n, [x_{n-1}, x_n]))\}$ de $[a,b]$ subordinada a algún calibrador δ sobre $[a,b]$ se elimina el requerimiento de que cada etiqueta $t_i \in [x_{i-1}, x_i]$ para cada i, es decir, se permite que t_i esté en otro intervalo distinto a $[x_{i-1}, x_i]$. ¿Qué tipo de integral se obtiene? La respuesta es: se obtiene una integral, conocida como la integral de McShane, equivalente a la integral de Lebesgue. Lo interesante de esta versión modificada de la integral de Henstock-Kurzweil es que la integral de Lebesgue puede ser pensada como límite de sumas à la Riemann.

Definición 14.2.1. *Sea δ un calibrador sobre $[a,b]$. Un **intervalo libremente etiquetado** $(t, [c,d])$ consiste de un intervalo $[c,d] \subseteq [a,b]$ y un punto $t \in [a,b]$. Un intervalo libremente etiquetado $(t, [c,d])$ está subordinado a δ si*

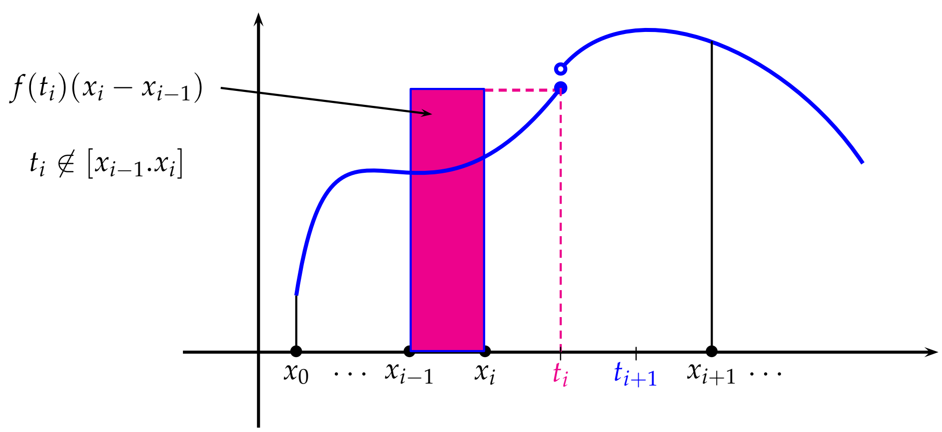
$$[c,\, d] \;\subseteq\; (t - \delta(t),\, t + \delta(t)). \tag{1}$$

*Por una **partición libremente etiquetada** $\mathcal{P}$ de $[a,b]$ subordinada a δ, se entiende una colección finita de intervalos libremente etiquetados subordinados a δ, digamos $\mathcal{P} = \{(t_1, [x_0, x_1]), \ldots, (t_n, [x_{n-1}, x_i])\}$, tal que $[a,b] = \bigcup_{i=1}^{n} [x_{n-1}, x_i]$.*

Aunque no se insiste que la etiqueta t en un intervalo libremente etiquetado $(t, [c,d])$ subordinado a δ pertenezca al intervalo $[c,d]$, la función δ, gracias a la inclusión dada (1), mantiene a tal etiqueta muy cerca del intervalo $[c,d]$. En lo que sigue, y sólo por comodidad, a cualquier partición libremente etiquetada la llamaremos una **partición de McShane**. Observe que *cualquier partición etiquetada de $[a,b]$ subordinada a δ es una partición libremente etiquetada subordinada a δ*, pero lo contrario no es, en general, cierto.

Como antes, si $f : [a,b] \to \mathbb{R}$ es una función y $\mathcal{P} = \{(t_1, [x_0, x_1]), \ldots, (t_n, [x_{n-1}, x_i])\}$ es una partición de McShane, la suma de Riemann correspondiente a $\mathcal{P}$ se define como

$$S(f, \mathcal{P}) \;=\; \sum_{i=1}^{n} f(t_i)(x_i - x_{i-1})$$



Definición 14.2.2. *Una función $f : [a,b] \to \mathbb{R}$ es llamada **McShane integrable** sobre $[a,b]$ si existe $A \in \mathbb{R}$ con la siguiente propiedad: para cada $\varepsilon > 0$, existe un calibrador δ sobre $[a,b]$ tal que*

$$\left| S(f, \mathcal{P}) \;-\; A \right| \;<\; \varepsilon$$

para cualquier partición de McShane $\mathcal{P} = (t_i, I_i)_{i=1}^{n}$ de $[a,b]$ subordinada a δ.

El número A de la definición anterior, si existe, se llama la **integral de McShane** de f y denotado por el símbolo

$$A \;=\; (\mathrm{M})\int_a^b f\,dt.$$

Designaremos por $\mathcal{M}c([a,b])$ *el conjunto de todas las funciones $f : [a,b] \to \mathbb{R}$ que son **McShane integrables** sobre $[a,b]$*. En general, si $E \subseteq [a,b]$ es un conjunto medible, diremos que f es **McShane integrable sobre** E si $f \cdot \chi_E$ es McShane integrable sobre $[a,b]$ y escribiremos

$$(\mathrm{M})\int_E f\,dt \;=\; (\mathrm{M})\int_a^b f \cdot \chi_E\,dt.$$

Los argumentos que se utilizaron para demostrar que la integral de Henstock-Kurzweil está bien definida se pueden usar para mostrar que la integral de McShane también está bien definida. Es claro que cualquier función McShane integrable es Henstock-Kurzweil integrable y sus integrales coinciden:

$$\mathcal{M}c([a,b]) \subseteq \mathcal{HK}([a,b])$$

y

$$(\mathrm{M})\int_a^b f\,dt = (\mathrm{HK})\int_a^b f\,dt$$

para cualquier $f \in \mathcal{M}c([a,b])$. En particular, toda función McShane integrable es medible. Similarmente, cualquier función Riemann integrable es McShane integrable y sus integrales son iguales. Por supuesto, las integrales de McShane y Henstock-Kurzweil comparten muchas propiedades y, en muchos casos, las pruebas son casi idénticas. Sin embargo, en algunos casos, hay que tener cierta precaución. Por ejemplo, en ciertas demostraciones sobre teoremas de la teoría de integración de Henstock-Kurzweil asumíamos que todas las etiquetas debían aparecer una, y sóla una vez en cada intervalo de la partición, o que todas las etiquetas tenían que ser extremos de dichos intervalos. Tales suposiciones no se pueden asumir en la integral de McShane. Esto, por supuesto, es una diferencia crucial entre los dos tipos de particiones etiquetadas.

Lo que haremos en esta sección es demostrar que las integrales de McShane y Lebesgue son, en realidad, equivalentes con integrales idénticas. Para ello, vamos a mencionar algunas de la propiedades de la integral de McShane que necesitaremos. La demostraciones, por ser muy similares a las de la integral de Henstock-Kurzweil, serán omitidas.

Teorema 14.2.3. *Sea $f : [a,b] \to \mathbb{R}$ una función.*

(1) *Las siguientes condiciones son equivalentes:*

   (a) $f \in \mathcal{M}c([a,b])$.

   (b) (**Criterio de Cauchy**) *Para cada $\varepsilon > 0$, existe un calibrador δ sobre $[a,b]$ tal que*

$$\left|S(f,\mathcal{P}) - S(f,\mathcal{Q})\right| < \varepsilon$$

   siempre que $\mathcal{P}$ y $\mathcal{Q}$ sean particiones de McShane subordinadas a δ.

   (c) *f es McShane integrable sobre $[a,c]$ y $[c,b]$ para cualquier $c \in (a,b)$. En este caso,*

$$(\mathrm{M})\int_a^b f\,dt = (\mathrm{M})\int_a^c f\,dt + (\mathrm{M})\int_c^b f\,dt.$$

(2) *$\mathcal{M}c([a,b])$ es un **espacio vectorial** sobre $\mathbb{R}$. Además, si $f,g \in \mathcal{M}c([a,b])$ con $f \leq g$, entonces*

$$(\mathrm{M})\int_a^b f\,dt \leq (\mathrm{M})\int_a^b g\,dt.$$

Por supuesto, si definimos *subpartición libremente etiquetada* o *subpartición de McShane* similar a la definición de subpartición etiquetada, el Lema de Saks-Henstock en $\mathcal{M}c([a,b])$ se expresa como antes:

Lema 14.2.4 (Saks-Henstock). *Sea* $f \in \mathcal{M}c([a,b])$. *Dado* $\varepsilon > 0$, *sea* δ *un calibrador sobre* $[a,b]$ *tal que*

$$\left| S(f,\mathcal{P}) - (\mathrm{M})\int_a^b f\,dt \right| < \varepsilon$$

para cualquier partición $\mathcal{P}$ *de McShane subordinada a* δ. *Entonces*

$$\left| \sum_{i=1}^{k} \left(f(s_i)\ell(J_i) - (\mathrm{M})\int_{J_i} f\,dt \right) \right| \leq \varepsilon \qquad y \qquad \sum_{i=1}^{k} \left| f(s_i)\ell(J_i) - (\mathrm{M})\int_{J_i} f\,dt \right| \leq 2\varepsilon$$

para cualquier ***subpartición de McShane*** $\mathcal{S} = (s_i, J_i)_{i=1}^{k}$ *de* $[a,b]$ *subordinada a* δ.

Ahora demostraremos que si $f \in \mathcal{M}c([a,b])$, entonces $|f| \in \mathcal{M}c([a,b])$. Como ya vimos, esta propiedad de la integral de McShane no es válida para la integral de Henstock-Kurzweil. El siguiente resultado es clave para lograr nuestro objetivo.

Lema 14.2.5. *Sea* $f \in \mathcal{M}c([a,b])$. *Dado* $\varepsilon > 0$, *sea* δ *un calibrador sobre* $[a,b]$ *tal que*

$$\left| S(f,\mathcal{P}) - (\mathrm{M})\int_a^b f\,dt \right| < \varepsilon$$

para cualquier partición de McShane $\mathcal{P}$ *de* $[a,b]$ *subordinada a* δ. *Si*

$$\mathcal{R} = \{(x_i, I_i) : 1 \leq i \leq N\} \qquad y \qquad \mathcal{S} = \{(y_j, J_j) : 1 \leq j \leq M\}$$

son ***subparticiones de McShane*** *subordinadas a* δ, *entonces*

$$\sum_{i=1}^{N}\sum_{j=1}^{M} |f(x_i) - f(y_i)|\, \mu(I_i \cap J_j) < 2\varepsilon.$$

Prueba. Los intervalos no-degenerados de la colección

$$\mathcal{J} = \{ I_i \cap J_j \,:\, 1 \leq i \leq N,\ 1 \leq j \leq M \}$$

forman una partición de $[a,b]$. Usando esos intervalos, formemos dos particiones de McShane $\mathcal{P}$ y $\mathcal{Q}$ como sigue:

$$\text{si } f(x_i) \geq f(y_j), \text{ entonces ponga } (x_i, I_i \cap J_j) \in \mathcal{P} \text{ y } (y_j, I_i \cap J_j) \in \mathcal{Q};$$
$$\text{si } f(x_i) < f(y_j), \text{ entonces ponga } (y_j, I_i \cap J_j) \in \mathcal{P} \text{ y } (x_i, I_i \cap J_j) \in \mathcal{Q}.$$

Observe que

$$S(f,\mathcal{P}) - S(f,\mathcal{Q}) = \sum_{i=1}^{N}\sum_{j=1}^{M} |f(x_i) - f(y_i)|\, \mu(I_i \cap J_j).$$

Puesto que $\mathcal{P}$ y $\mathcal{Q}$ son particiones de McShane de $[a,b]$ subordinadas a δ, se tiene que

$$S(f,\mathcal{P}) - S(f,\mathcal{Q}) \leq \left| S(f,\mathcal{P}) - (\mathrm{M})\int_a^b f\,dt \right| + \left| (\mathrm{M})\int_a^b f\,dt - S(f,\mathcal{Q}) \right| < 2\varepsilon.$$

La prueba es completa. ∎

Teorema 14.2.6. *Si $f \in \mathcal{M}c([a,b])$, entonces $|f| \in \mathcal{M}c([a,b])$.*

Prueba. Sea $\varepsilon > 0$. Puesto que $f \in \mathcal{M}c([a,b])$, existe un calibrador δ sobre $[a,b]$ tal que

$$\left| S(f,\mathcal{P}) - (\mathrm{M})\int_a^b f\,dt \right| < \varepsilon/2$$

para cualquier partición de McShane $\mathcal{P}$ de $[a,b]$ subordinada a δ. Mostraremos ahora que $|f|$ satisface el Criterio de Cauchy para la integral de McShane. Para ver esto, sean

$$\mathcal{P} = \{(x_i, I_i) : 1 \leq i \leq N\} \qquad \text{y} \qquad \mathcal{Q} = \{(y_j, J_j) : 1 \leq j \leq M\}$$

particiones de McShane sobre $[a,b]$ subordinadas a δ. Use los intervalos no-degenerados de la colección $\mathcal{J} = \{I_i \cap J_j : 1 \leq i \leq N,\ 1 \leq j \leq M\}$ para formar dos particiones etiquetadas de $[a,b]$:

$$\mathcal{P}_1 = \{(x_i, I_i \cap J_j) : i \in N_1\} \qquad \text{y} \qquad \mathcal{Q}_1 = \{(y_j, I_i \cap J_j) : j \in N_2\}.$$

Observe que, en realidad, $\mathcal{P}_1$ y $\mathcal{Q}_1$ son particiones de McShane de $[a,b]$ subordinadas a δ y que

$$S(|f|,\mathcal{P}_1) = \sum_{i=1}^{N}\sum_{j=1}^{M} |f(x_i)|\,\mu(I_i \cap J_j) = \sum_{i=1}^{N} |f(x_i)|\,\mu(I_i) = S(|f|,\mathcal{P}),$$

$$S(|f|,\mathcal{Q}_1) = \sum_{i=1}^{M}\sum_{j=1}^{N} |f(y_j)|\,\mu(I_i \cap J_j) = \sum_{i=1}^{M} |f(y_j)|\,\mu(J_j) = S(|f|,\mathcal{Q}).$$

Se sigue del Lema 14.2.5 que

$$\begin{aligned}
\Big|S(|f|,\mathcal{P}) - S(|f|,\mathcal{Q})\Big| &= \Big|S(|f|,\mathcal{P}_1) - S(|f|,\mathcal{Q}_1)\Big| \\
&\leq \sum_{i=1}^{N}\sum_{j=1}^{M} \Big||f(x_i)| - |f(y_j)|\Big|\,\mu(I_i \cap J_j) \\
&\leq \sum_{i=1}^{N}\sum_{j=1}^{M} |f(x_i) - f(y_j)|\,\mu(I_i \cap J_j) \\
&< 2\frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

Esto nos indica que $|f|$ es McShane integrable y termina la prueba. ■

Ya poseemos los argumentos necesarios para demostrar que las clases $\mathcal{M}c([a,b])$ y $\mathrm{L}_1([a,b])$ son iguales con idénticas integrales.

Teorema 14.2.7 (McShane-Lebesgue). *Sea $f : [a,b] \to \mathbb{R}$ una función. Las siguientes condiciones son equivalentes:*

(1) $f \in \mathcal{M}c([a,b])$.

(2) $f \in \mathrm{L}_1([a,b])$.

Además, para toda función $f \in \mathcal{M}c([a,b]) = \mathrm{L}_1([a,b])$,

$$(\mathrm{M})\int_a^b f\,dt = (\mathrm{L})\int_a^b f\,d\mu.$$

Prueba. (1) $\Rightarrow$ (2) Suponga que $f \in \mathcal{M}c([a,b])$. Por el Teorema 14.2.6, $|f| \in \mathcal{M}c([a,b])$ y como $\mathcal{M}c([a,b]) \subseteq \mathcal{HK}([a,b])$, resulta que $f, |f| \in \mathcal{HK}([a,b])$. Esto, por supuesto, significa que $f \in \mathcal{L}_{\mathrm{HK}}([a,b]) = \mathrm{L}_1([a,b])$.

(2) $\Rightarrow$ (1) Suponga que $f \in \mathrm{L}_1([a,b])$ y sea $\varepsilon > 0$. Puesto que $f = f^+ - f^-$, podemos asumir, y así lo haremos, que $f \geq 0$. Por el Teorema 10.2.7, página 581, existe un $\eta > 0$ tal que

$$(\mathrm{L})\int_B |f|\, d\mu \;<\; \varepsilon \tag{1}$$

para todo conjunto medible $B \subseteq [a,b]$ que satisfaga $\mu(B) < \eta$. Pongamos $\gamma = \mathrm{mín}\{\varepsilon, \eta\}$.

Sea $\beta = \mathrm{mín}\{1, \varepsilon/(\eta + b - a)\}$ y observe que $\gamma\beta \leq \varepsilon$. Por cada entero positivo k, defina

$$E_k \;=\; \big\{t \in [a,b] \,:\, (k-1)\beta \;<\; f(t) \;\leq\; k\beta\big\}.$$

Los conjuntos E_k son medibles, disjuntos y, además, $[a,b] = \bigcup_{k=1}^{\infty} E_k$. Usemos ahora la definición de medida exterior para seleccionar, por cada $k \in \mathbb{N}$, un conjunto abierto G_k tal que

$$E_k \subseteq G_k \qquad \text{y} \qquad \mu(G_k \setminus E_k) \;<\; \frac{\gamma}{k\,2^k}.$$

Con esta información podemos definir un calibrador δ sobre $[a,b]$ del modo siguiente: puesto que cada $t \in [a,b]$ pertenece a un único E_k, defina

$$\delta(t) \;=\; \mathrm{dist}(t, G_k^c) \quad \text{si } t \in E_k \subseteq [a,b].$$

La prueba estará completa una vez que verifiquemos que

$$\left| S(f,\mathcal{P}) \;-\; (\mathrm{L})\int_a^b f\, d\mu \right| \;<\; \varepsilon$$

para cualquier partición de McShane $\mathcal{P}$ de $[a,b]$ subordinada a δ. Sea entonces $\mathcal{P} = ((t_i, I_i))_{i=1}^n$ una partición de McShane de $[a,b]$ subordinada a δ. Para cada $i \in \{1, \ldots, n\}$, sea k_i el único entero positivo tal que $t_i \in E_{k_i}$ y sean

$$A_i \;=\; I_i \cap E_{k_i} \qquad \text{y} \qquad B_i \;=\; I_i \setminus E_{k_i}.$$

Puesto que $(\mathrm{L})\displaystyle\int_a^b f\, d\mu = \sum_{i=1}^n (\mathrm{L})\int_{I_i} f\, d\mu$, $A_i \cup B_i = I_i$ y estamos asumiendo que $f \geq 0$, resulta que

$$\begin{aligned}
&\left| S(f,\mathcal{P}) \;-\; (\mathrm{L})\int_a^b f\, d\mu \right| \\
&= \left| \sum_{i=1}^n (\mathrm{L})\int_{I_i} \big[f(t_i) \;-\; f(t)\big]\, d\mu \right| \;\leq\; \sum_{i=1}^n (\mathrm{L})\int_{I_i} \big|f(t_i) \;-\; f(t)\big|\, d\mu \\
&\leq\; \sum_{i=1}^n (\mathrm{L})\int_{A_i} \big|f(t_i) \;-\; f(t)\big|\, d\mu \;+\; \sum_{i=1}^n (\mathrm{L})\int_{B_i} f(t_i)\, d\mu \;+\; \sum_{i=1}^n (\mathrm{L})\int_{B_i} f(t)\, d\mu
\end{aligned}$$

Veremos ahora que cada uno de los términos de la derecha de la última desigualdad es menor que ε. En efecto, para el primero, observe que si $t, t_i \in A_i$, entonces como $A_i \subseteq E_{k_i}$ se tiene que $f(t), f(t_i) \in (k_i - 1)\beta, k_i\beta]$ y, por consiguiente, $|f(t_i) - f(t)| < \beta$. Por esto,

$$\sum_{i=1}^{n} (\mathrm{L})\int_{A_i} \big|f(t_i) - f(t)\big|\, d\mu \;<\; \beta \cdot \mu\Big(\bigcup_{i=1}^{n} A_i\Big) \;\le\; \beta(b-a) \;\le\; \varepsilon.$$

Para estimar $\sum_{i=1}^{n} (\mathrm{L})\int_{B_i} \big|f(t_i)\big|\, d\mu$, nótese que si $N_k = \{i \in \{1,\ldots,n\} : k_i = k\}$ para cada $k \in \mathbb{N}$, entonces $\bigcup_{k=1}^{\infty} N_k = \{1,\ldots,n\}$. Como $I_i \subseteq (t_i - \delta(t_i), t_i + \delta(t_i)) \subseteq G_{k_i}$ y $f(t_i) \le k_i\beta$, se tiene que

$$\begin{aligned}
\sum_{i=1}^{n} (\mathrm{L})\int_{B_i} f(t_i)\, d\mu \;&\le\; \sum_{k=1}^{n} k_i\beta \cdot \mu(B_i) \;=\; \sum_{k=1}^{\infty} \sum_{i\in N_k} k\beta \cdot \mu(I_i \setminus E_{k_i}) \\
&\le\; \sum_{k=1}^{\infty} k\beta \cdot \mu(G_k \setminus E_k) \;\le\; \sum_{k=1}^{\infty} k\beta \frac{\gamma}{k\, 2^k} \\
&\le\; \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} \;=\; \varepsilon.
\end{aligned}$$

Finalmente, tomando $B = \bigcup_{i=1}^{n} B_i$ y teniendo en cuenta que $I_i \subseteq G_{k_i}$, $i = 1, \ldots, n$, resulta que

$$\begin{aligned}
\mu(B) \;&=\; \sum_{i=1}^{n} \mu(I_i \setminus E_{k_i}) \;=\; \sum_{k=1}^{\infty} \sum_{i\in N_k} \mu(I_i \setminus E_{k_i}) \\
&\le\; \sum_{k=1}^{\infty} \mu(G_k \setminus E_k) \;\le\; \sum_{k=1}^{\infty} \frac{\gamma}{k\, 2^k} \;<\; \sum_{k=1}^{\infty} \frac{\eta}{2^k} \;=\; \eta.
\end{aligned}$$

Se sigue entonces de (1) que

$$\sum_{i=1}^{n} (\mathrm{L})\int_{B_i} f(t)\, d\mu \;=\; (\mathrm{L})\int_{B} \big|f(t)\big|\, d\mu \;<\; \varepsilon.$$

Reuniendo las tres estimaciones anteriores se concluye que

$$\left| S(f, \mathcal{P}) \;-\; (\mathrm{L})\int_a^b f\, d\mu \right| \;<\; 3\varepsilon$$

probándose, de este modo, que f es McShane integrable y que

$$(\mathrm{M})\int_a^b f\, dt \;=\; (\mathrm{L})\int_a^b f\, d\mu.$$

Fin de la prueba. ■

Nota Adicional 14.2.6 Recordemos que una función acotada $f : [a,b] \to \mathbb{R}$ es Riemann integrable si, para cada $\varepsilon > 0$ se puede determinar un $\delta > 0$ de modo que

$$\left| S(f, \mathcal{P}) \;-\; (\mathrm{R})\int_a^b f\, d\mu \right| \;<\; \varepsilon$$

para cualquier partición etiquetada $\mathcal{P} = ((t_i, I_i))_{i=1}^n$ de $[a,b]$ subordinada a δ. Puesto que $\mathcal{R}([a,b]) \subseteq L_1([a,b]) = \mathcal{M}c([a,b])$, resulta que toda función Riemann integrable es McShane integrable y, en consecuencia, la integral de Riemann no cambia si en lugar de utilizar particiones etiquetadas subordinadas a δ, se usan particiones **libremente** etiquetadas subordinadas a δ. En efecto, tal como fue demostrado por R. Gordon en [65], Theorem 2, la definición de la integral de Riemann no se afecta en lo absoluto si las etiquetas t_i pertenecen o no al intervalo I_i. Sin embargo, como vimos con las integrales de Henstock-Kurzweil y de McShane, cambiar una constante por una variable δ y demandar que las etiquetas permanezcan en cada intervalo correspondiente genera la integral de Henstock-Kurzweil, mientras que si se permite que las etiquetas se muevan libremente en dichos intervalos se obtiene la integral de McShane o, equivalentemente, la integral de Lebesgue. Esa diferencia es la que impide que las funciones Henstock-Kurzweil integrables sean, en general, absolutamente integrables.

14.3. La C-integral

.

Ya hemos visto que $\mathcal{HK}([a,b])$ contiene propiamente tanto a $L_1([a,b])$, así como a $\mathcal{N}([a,b])$. Surge entonces la pregunta: ¿es posible obtener una colección minimal de funciones, "integrables" en algún sentido, que contenga a las integrales de Lebesgue y Newton y esté contenida propiamente en $\mathcal{HK}([a,b])$? La respuesta, aunque tardó un poco, es positiva. En efecto, en el año 1986, Bruckner, Fleissner y Foran [26] obtuvieron una definición descriptiva de una extensión minimal de la integral de Lebesgue la cual integra la derivada de cualquier función diferenciable sobre $[a,b]$. Específicamente, Bruckner, Fleissner y Foran demostraron el siguiente resultado:

Teorema 14.3.1 (Bruckner, Fleissner, Foran). *Una función $f : [a,b] \to \mathbb{R}$ es **C-integrable** sobre $[a,b]$ si, y sólo si, existe una **función derivada** g sobre $[a,b]$ tal que $f - g$ es **Lebesgue integrable** sobre $[a,b]$.*

Un poco más tarde, en el año 2000, Bongiorno, Di Piazza y Preiss [18] demostraron que esa integral minimal se puede obtener a partir de la definición de la integral de McShane imponiendo una pequeña condición de regularidad en las particiones de McShane.

Definición 14.3.2. *Una función $f : [a,b] \to \mathbb{R}$ se dice que es **C-integrable** sobre $[a,b]$, si existe un número real A satisfaciendo la siguiente propiedad: para cada $\varepsilon > 0$, existe un calibrador δ sobre $[a,b]$ tal que*

$$\big|S(f,\mathcal{P}) - A\big| < \varepsilon$$

para cualquier partición de McShane $\mathcal{P} = ((t_i, I_i))_{i=1}^n$ de $[a,b]$ subordinada a δ y $1/\varepsilon$-controlada, es decir, que cumpla la condición

$$\sum_{i=1}^n \operatorname{dist}(t_i, I_i) < 1/\varepsilon.$$

El número A en la definición anterior se llama la C-**integral** de f sobre $[a,b]$ y se escribirá como

$$A = \mathrm{C}\!\int_a^b f.$$

Denote por $\mathcal{C}-\mathrm{int}([a,b])$ *el conjunto de todas las funciones* $f:[a,b]\to\mathbb{R}$ *que son* ***C-integrables*** *sobre* $[a,b]$.

Propiedades de la C-integral.

(1) $\mathcal{C}-\mathrm{int}([a,b])$ es un espacio vectorial sobre $\mathbb{R}$.

(2) Es claro que $\mathcal{C}-\mathrm{int}([a,b])$ incluye a cualquier función McShane integrable sobre $[a,b]$ y, en consecuencia, como las integrales de McShane y Lebesgue coinciden sobre $[a,b]$, entonces $\mathcal{C}-\mathrm{int}([a,b])$ incluye a cualquier $f\in \mathrm{L}_1([a,b])$. En particular, si f es C-integrable sobre $[a,b]$, entonces f es Henstock-Kurzweil integrable sobre $[a,b]$ con la misma integral.

(3) Si f es C-integrable sobre $[a,b]$, entonces f es C-integrable sobre $[a,x]$ para cualquier $x\in(a,b)$. Además, la función

$$F(x) \;=\; \mathrm{C}\int_a^x f$$

es continua sobre $[a,b]$.

Teorema 14.3.3. $\mathcal{N}([a,b])\subseteq \mathcal{C}-\mathrm{int}([a,b])$.

Prueba. Sea $f\in\mathcal{N}([a,b])$. Escoja una función siempre diferenciable $F:[a,b]\to\mathbb{R}$ tal que $F'(x)=f(x)$ para todo $x\in[a,b]$. Sea $\varepsilon\in(0,1/(b-a))$ y sea $x\in[a,b]$. Usemos la definición de derivada para seleccionar un $\delta(x)>0$ de modo que

$$\left|\frac{F(y)-F(x)}{y-x}-f(x)\right| \;<\; \frac{\varepsilon^2}{4} \tag{2}$$

para cada $y\in(x-\delta(x),x+\delta(x))\cap[a,b]$. Observe que si $I=[c,d]$ es un intervalo incluido en $(x-\delta(x),x+\delta(x))\cap[a,b]$, entonces de (2) se tiene que

$$\begin{aligned}\big|F(d)-F(c)-f(x)(d-c)\big| \;&\le\; \big|F(d)-F(x)-f(x)(d-x)\big| \;+\; \big|F(c)-F(x)-f(x)(c-x)\big|\\ &<\; \frac{\varepsilon^2}{4}|d-x| \;+\; \frac{\varepsilon^2}{4}|c-x|\\ &\le\; \frac{\varepsilon^2}{2}\big(\mathrm{dist}(x,I)+\ell(I)\big).\end{aligned}$$

De allí que, si $\mathcal{P}=\{(x_1,I_1),\ldots,(x_m,I_m)\}$ es una partición de McShane de $[a,b]$ subordinada a δ satisfaciendo la condición $\sum_{i=1}^m \mathrm{dist}(t_i,I_i)<1/\varepsilon$, entonces, escribiendo $F(I_i)=F(d_i)-F(c_i)$, resulta de las desigualdades anteriores que

$$\begin{aligned}\left|\sum_{i=1}^m f(x_i)\ell(I_i) \;-\; \big[F(b)\;-\;F(a)\big]\right| \;&\le\; \sum_{i=1}^m \big|f(x_i)\ell(I_i)\;-\;F(I_i)\big|\\ &<\; \frac{\varepsilon^2}{2}\sum_{i=1}^m\big(\mathrm{dist}(x,I_i)+\ell(I_i)\big)\\ &<\; \frac{\varepsilon}{2}+\frac{\varepsilon}{2}\;=\;\varepsilon.\end{aligned}$$

Esto prueba que f es C-integrable sobre $[a,b]$. ■

Combinando los resultados anteriores se tienen las siguientes inclusiones:

$$\mathcal{M}c([a,b]) \subseteq \mathcal{C}-\mathrm{int}([a,b]) \subseteq \mathcal{HK}([a,b]).$$

Veamos ahora que tales inclusiones son todas propias. En efecto, la función $f : [0,1] \to \mathbb{R}$ definida por

$$f(x) = 2x \operatorname{sen}\Big(\frac{1}{x^2}\Big) - \frac{2}{x}\cos\Big(\frac{1}{x^2}\Big) \quad \text{con } f(0) = 0$$

pertenece a $\mathcal{C}-\mathrm{int}([a,b])$ ya que es una derivada, pero no a $\mathrm{L}_1([a,b]) = \mathcal{M}c([a,b])$, mientras que la función $g : [0,1] \to \mathbb{R}$ definida por

$$g(x) = \begin{cases} x \operatorname{sen}\Big(\dfrac{1}{x^2}\Big) & \text{si } x \in (0,1], \\ 0 & \text{si } x = 0 \end{cases}$$

es Henstock-Kurzweil integrable sobre $[0,1]$ pero no es C-integrable sobre $[0,1]$.

Para la C-integral valen los Teoremas de la Convergencia Monótona y la Convergencia Dominada.

Teorema 14.3.4 (Convergencia Monótona). *Sea $(f_n)_{n=1}^{\infty}$ una sucesión **monótona creciente** de funciones C-integrables sobre $[ab]$ tal que*

$$\lim_{n\to} \mathrm{C}\int_a^b f_n < +\infty.$$

Si $f(x) = \lim_{n\to\infty} f_n(x)$ para cada $x \in [a,b]$, entonces $f \in \mathcal{C}-\mathrm{int}([a,b])$ y se cumple que

$$\mathrm{C}\int_a^b f = \lim_{n\to\infty} \mathrm{C}\int_a^b f_n.$$

Prueba. Puesto que cada función C-integrable es Henstock-Kurzweil integrable, se sigue del Teorema 14.1.37, página 884, que cada función no-negativa C-integrable es Lebesgue integrable. De esto se sigue que podemos aplicar el Teorema de la Convergencia Monótona de Lebesgue a la sucesión no-negativa y creciente

$$0 \le f_2 - f_1 \le f_3 - f_1 \le \cdots f_n - f_1 \le \cdots$$

para obtener

$$(\mathrm{L})\int_a^b (f - f_1)\, d\mu = \lim_{n\to\infty} (\mathrm{L})\int_a^b (f_n - f_1)\, d\mu.$$

Por otro lado, como

$$\begin{aligned}(\mathrm{L})\int_a^b (f_n - f_1)\, d\mu &= \mathrm{C}\int_a^b (f_n - f_1)\, d\mu \\ &= \mathrm{C}\int_a^b f_n\, d\mu - \mathrm{C}\int_a^b f_1\, d\mu\end{aligned}$$

y

$$\lim_{n\to\infty} \mathrm{C}\int_a^b f_n \; < \; +\infty$$

resulta que $f - f_1$ es Lebesgue integrable sobre $[a,b]$ y, por lo tanto, $f = (f - f_1) + f_1$ también lo es. En particular, $f \in \mathcal{C} - \mathrm{int}([a,b])$ y

$$\begin{aligned}
\mathrm{C}\int_a^b f \; &= \; \mathrm{C}\int_a^b (f - f_1) \; + \; \mathrm{C}\int_a^b f_1 \\
&= \; \lim_{n\to\infty} (\mathrm{L})\int_a^b f_n \; - \; \mathrm{C}\int_a^b f_1 \; + \; \mathrm{C}\int_a^b f_1 \\
&= \; \lim_{n\to\infty} (\mathrm{L})\int_a^b f_n
\end{aligned}$$

y termina la prueba. ■

Teorema 14.3.5 (Convergencia Dominada). *Sea $(f_n)_{n=1}^{\infty}$ una sucesión de **funciones medibles** sobre $[a,b]$ tal que:*

(a) $f(x) = \lim_{n\to\infty} f_n(x)$ para casi-todo $x \in [a,b]$, y

(b) $g \le f_n \le h$ casi-siempre en $[a,b]$ con $g, h \in \mathcal{C} - \mathrm{int}([a,b])$.

Entonces $f \in \mathcal{C} - \mathrm{int}([a,b])$ y se cumple que

$$\mathrm{C}\int_a^b f \; = \; \lim_{n\to\infty} \mathrm{C}\int_a^b f_n.$$

Prueba. Por (b) se tiene que $0 \le f_n - g \le h - g$ casi-siempre en $[a,b]$. Más aun, como $h - g$ es no-negativa y C-integrable, el Teorema 14.1.37 nos garantiza que $h - g \in L_1([a,b])$. Por el Teorema de la Convergencia Dominada de Lebesgue se tiene que $f - g \in L_1([a,b])$ y

$$(\mathrm{L})\int_a^b (f - g)\, d\mu \; = \; \lim_{n\to\infty} (\mathrm{L})\int_a^b (f_n - g)\, d\mu.$$

Por esto, como $f = (f - g) + g$ se sigue que f es C-integrable sobre $[a,b]$ y

$$\begin{aligned}
\mathrm{C}\int_a^b f \; &= \; \mathrm{C}\int_a^b (f - g) \; + \; \mathrm{C}\int_a^b g \\
&= \; (\mathrm{L})\int_a^b (f - g) \; + \; \mathrm{C}\int_a^b g \\
&= \; \lim_{n\to\infty} \mathrm{C}\int_a^b f_n \; - \; \mathrm{C}\int_a^b g \; + \; \mathrm{C}\int_a^b g \\
&= \; \lim_{n\to\infty} \mathrm{C}\int_a^b f_n.
\end{aligned}$$

Fin de la prueba. ■

14.4. Las Integrales de Denjoy, Perron y Distribucional

Esta corta sección trata, exclusivamente, sobre definiciones descriptivas de las integrales de Denjoy, Perron y Distribucional de Denjoy. Las propiedades de las dos primeras de estas integrales pueden ser estudiadas, por ejemplo, en el libro [66]. Un buen artículo para estudiar la integral distribucional de Denjoy es [128]. Como ya hemos informado, en 1912 A. Denjoy desarrolló una integral, la cual es conocida hoy en día como la integral de Denjoy y que tiene la propiedad de que el Teorema Fundamental del Cálculo es válido en todo su generalidad. El proceso de construcción ideado por Denjoy era muy complicado ya que utilizaba los números transfinitos. Por suerte, N. Lusin, unos meses después del trabajo de Denjoy, tendió un puente entre la integral de Denjoy y la noción de funciones absolutamente continuas generalizadas. Ese enfoque es el que, por lo general, se usa para definir descriptivamente la integral de Denjoy.

14.4.1. La Integral de Denjoy

.

En cierto sentido, las funciones absolutamente continuas generalizadas son extensiones numerables del concepto de función absolutamente continua.

Definición 14.4.1. *Sean $F : [a,b] \to \mathbb{R}$ y $X \subseteq [a,b]$. La función F es* ***absolutamente continua en el sentido restringido*** *sobre X, en notación $F \in \mathrm{AC}_*(X)$, si F es acotada sobre un intervalo cerrado $I \supseteq X$ y para cada $\varepsilon > 0$, existe un $\delta > 0$ tal que*

$$\sum_{i=1}^{n} \mathrm{osc}(F, [c_i, d_i]) \ < \ \varepsilon$$

para cualquier colección finita de intervalos no-superpuestos $([c_i, d_i])_1^n$ con $\{c_i, d_i : i = 1, \dots n\} \subseteq X$ que satisfaga $\sum_{i=1}^n (d_i - c_i) < \delta$. La función F es ***absolutamente continua generalizada en el sentido restringido*** *sobre X si F es continua sobre X y existe una sucesión disjunta de conjuntos $\big(X_n\big)_{n=1}^{\infty}$ tal que*

$$[a,b] \ = \ \bigcup_{n=1}^{\infty} X_n \qquad\qquad y \qquad\qquad f \in \mathrm{AC}_*(X_n)$$

para cada $n \in \mathbb{N}$.

La notación $F \in \mathrm{ACG}_*$ significa que F es absolutamente continua generalizada en el sentido restringido sobre $[a,b]$. Es interesante observar que $C^1([a,b]) \subseteq \mathrm{AC}([a,b]) \subseteq \mathrm{ACG}_*([a,b]) \subseteq C([a,b])$ y, por lo tanto, se tiene un espacio más grande en el cual las funciones tienen derivadas casi-siempre. Más aun, $\mathrm{ACG}_*([a,b])$ es denso en $C([a,b])$ ya que $\mathrm{AC}([a,b])$ lo es.

El Corolario 10.2.44, página 618, ofrece una manera sencilla de presentar la integral de Lebesgue sobre un intervalo $[a,b]$ a partir de una definición descriptiva de la misma.

Definición 14.4.2. *Una función $f : [a,b] \to \mathbb{R}$ es* ***Lebesgue integrable*** *sobre $[a,b]$ si existe una función $F \in \mathrm{AC}([a,b])$ tal que $F' = f$ casi-siempre sobre $[a,b]$.*

N. Lusin, en lugar de utilizar funciones absolutamente continuas, propuso, usando funciones absolutamente continuas generalizadas en el sentido restringido sobre $[a,b]$, la siguiente definición, también descriptiva, de la integral de Denjoy.

Definición 14.4.3. *Una función $f : [a, b] \to \mathbb{R}$ es* ***Denjoy integrable*** *sobre $[a, b]$ si existe una función $F \in \mathrm{ACG}_*$ tal que $F' = f$ casi-siempre sobre $[a, b]$.*

Se puede demostrar, véase [66], Corollary 6.23, p. 103, que *si $F : [a, b] \to \mathbb{R}$ es* ***continua*** *sobre $[a, b]$ y* ***diferenciable*** *excepto sobre un conjunto a lo más numerable de $[a, b]$, entonces $F \in \mathrm{ACG}_*$.* En virtud de este resultado, la integral de Denjoy de una función está unívocamente determinada salvo por una constante. Si se añade la condición $F(a) = 0$, entonces la función F es única y se llama la **integral de Denjoy**, o **integral de Denjoy-Khintchine** de f sobre $[a, b]$, la cual denotaremos por

$$(\mathrm{De})\int_a^b f\,dt \;=\; F(b) - F(a).$$

El símbolo $\mathcal{DK}([a, b])$ será usado para denotar *a todas funciones $f : [a, b] \to \mathbb{R}$ que son Denjoy integrables sobre $[a, b]$.* Todas las propiedades usuales que, en general, una buena integral posee son válidas para la integral de Denjoy. La diferencia estructural entre las integrales de Lebesgue y Denjoy es que la integral indefinida en la integral de Denjoy es una función absolutamente continua generalizada en el sentido restringido en lugar de una función absolutamente continua como es el caso en la integral de Lebesgue. Por otro lado, la integral de Denjoy no es absoluta.

Recordemos que un punto $x \in (a, b)$ es punto de densidad de Lebesgue de un conjunto $E \subseteq [a, b]$ si

$$\lim_{h\to 0^+} \frac{\mu^*(E \cap [x-h,\, x+h])}{2h} \;=\; 1.$$

Una función $F : [a, b] \to \mathbb{R}$ se dice que es **aproximadamente derivable** en un punto $x \in (a, b)$, con **derivada aproximada** $F'_{\mathrm{ap}} \in \mathbb{R}$, si existe un conjunto medible $E \subseteq [a, b]$ teniendo a x como un punto de densidad tal que

$$\lim_{\substack{s\to x\\ s\in E}} \frac{F(s) \,-\, F(x)}{s \,-\, x} \;=\; F'_{\mathrm{ap}}(x).$$

Si en la Definición 14.4.3 reemplazamos el espacio ACG_* por ACG y F' por F'_{ap} obtenemos la **integral de Khinchin** o **Denjoy-Khinchin**.

Un resultado que expresa una íntima relación entre la integral de Lebesgue y la integral de Denjoy es el siguiente:

Teorema 14.4.4. *Si $f : [a, b] \to \mathbb{R}$ es* ***Denjoy integrable****, entonces para cada $\varepsilon > 0$, existe un conjunto medible $E \subseteq [a, b]$ tal que $\mu([a, b] \setminus E) < \varepsilon$, f es* ***Lebesgue integrable*** *sobre E y*

$$(\mathrm{L})\int_E f\,d\mu \;=\; (\mathrm{De})\int_a^b f\,dt.$$

La definición descriptiva de la integral de Denjoy (= Denjoy-Khinchin) que acabamos de dar constituyen dos de las tres soluciones diferentes al problema de integrar derivadas que Denjoy produjo. La otra solución consistió en utilizar su proceso de totalización a través de inducción transfinita.

14.4.2. La Integral de Perron

.

En 1914, O. Perron desarrolló otra integral por un método completamente distinto al seguido por Denjoy. De hecho, dichos enfoques eran tan diferentes que se tardó casi una década en demostrarse que ambas integrales eran equivalentes. Las nociones de funciones menor y mayor introducidas y estudiadas por primera vez por Ch. J. de la Vallée-Poussin sólo para funciones absolutamente continuas son las que serán utilizadas para definir la integral de Perron. Ellas se definen a partir de las derivadas superior e inferior de Dini (véase la página 485 para recordar sus definiciones). Eliminando el requerimiento de que las funciones mayor y menor sean absolutamente continuas, Perron define su integral.

Definición 14.4.5. *Sea $f : [a,b] \to \overline{\mathbb{R}}$ una función.*

*(a) Una función $U : [a,b] \to \mathbb{R}$ es una **función mayor** de f sobre $[a,b]$ si*

$$-\infty < \underline{D}U(x) \qquad y \qquad f(x) \leq \underline{D}U(x) \quad \text{para todo } x \in [a,b].$$

*(b) Una función $V : [a,b] \to \mathbb{R}$ es una **función menor** de f sobre $[a,b]$ si*

$$\overline{D}\,V(x) < +\infty \qquad y \qquad \overline{D}\,V(x) \leq f(x) \quad \text{para todo } x \in [a,b].$$

De la Vallée-Poussin usó sus nociones de funciones absolutamente continuas mayor y menor para dar otra caracterización de las funciones Lebesgue integrables sobre $[a,b]$.

Teorema 14.4.6 (De la Vallée-Poussin). *Una función medible $f : [a,b] \to \overline{\mathbb{R}}$ es **Lebesgue integrable** sobre $[a,b]$ si, y sólo si, para cada $\varepsilon > 0$ existen funciones **absolutamente continuas mayor** y **menor** U y V de f sobre $[a,b]$ tal que $U_a^b - V_a^b < \varepsilon$.*

Prueba. Suponga que f es Lebesgue integrable sobre $[a,b]$ y sea $\varepsilon > 0$. Invoquemos el Teorema de Vitali-Carathéodory, Teorema 10.2.46, página 619, para hallar funciones u, v con u semicontinua inferiormente y v semicontinua superiormente sobre $[a,b]$ tales que

$$u < f < v \qquad \text{y} \qquad \int_{[a,b]} u\,d\mu - \frac{\varepsilon}{2} < \int_{[a,b]} f\,d\mu < \int_{[a,b]} v\,d\mu + \frac{\varepsilon}{2}.$$

Como u y v son Lebesgue integrables, resulta que las funciones

$$U(x) = \int_{[a,x]} u\,d\mu \qquad \text{y} \qquad V(x) = \int_{[a,x]} v\,d\mu$$

son absolutamente continuas sobre $[a,b]$ gracias al Segundo Teorema Fundamental del Cálculo, Teorema 10.2.39, página 612. Además, de la observación después de Teorema de Vitali-Carathéodory, página 622, se sigue que

$$\underline{D}U(x) \geq u(x) \qquad \text{y} \qquad \overline{D}V(x) \leq v(x)$$

para todo $x \in [a,b]$. Teniendo en cuenta que $-\infty < u(x)$ y $v(x) < +\infty$, se concluye que

$$-\infty < \underline{D}U(x), \quad f(x) \leq \underline{D}U(x) \qquad \text{y} \qquad \overline{D}\,V(x) < +\infty, \quad \overline{D}\,V(x) \leq f(x)$$

para todo $x \in [a,b]$. Esto nos indica que U es una función mayor para f y V es una función menor para f sobre $[a,b]$. Finalmente,

$$U_a^b - V_a^b = \int_{[a,b]} u\, d\mu - \int_{[a,b]} f\, d\mu + \int_{[a,b]} f\, d\mu - \int_{[a,b]} v\, d\mu < \frac{\varepsilon}{2} + \frac{\varepsilon}{2}.$$

Para demostrar la otra implicación, suponga que para cada $\varepsilon > 0$, existen funciones absolutamente continuas mayor y menor U y V de f sobre $[a,b]$ tal que $U_a^b - V_a^b < \varepsilon$. Por el Primer Teorema Fundamental del Cálculo, Teorema 10.2.41, página 615, las funciones U' y V' existen casi-siempre y son Lebesgue integrables. Ahora bien, como $\underline{D}U = U'$ y $\overline{D}V = V'$ casi-siempre sobre $[a,b]$, resulta que tales funciones son también Lebesgue integrables sobre $[a,b]$. Finalmente, puesto $\overline{D}V \leq f \leq \underline{D}U$, entonces la función f es Lebesgue integrable sobre $[a,b]$. La prueba es completa. ∎

Si $F : [a,b] \to \mathbb{R}$ es una función arbitraria, escribiremos $F_a^b = F(b) - F(a)$. Suponga que U es una función mayor para f y V es una función menor para f sobre $[a,b]$. Puesto que $\underline{D}(U - V) \geq \underline{D}U - \overline{D}V \geq 0$, se sigue del Teorema 9.1.43, página 491, que $U - V$ es creciente sobre $[a,b]$. Por lo tanto,

$$V_a^b \leq U_a^b \qquad \text{y} \qquad 0 \leq U_c^d - V_c^d \leq U_a^b - V_a^b$$

para cualesquiera $a < c < d < b$. De allí que

$$\begin{aligned} -\infty &< \sup\left\{V_a^b : V \text{ es una función menor para } f\right\} \\ &\leq \inf\left\{U_a^b : U \text{ es una función mayor para } f\right\} < +\infty. \end{aligned}$$

La integral de Perron, la cual es una generalización del resultado anterior, Teorema 14.4.6, se obtiene eliminando el requerimiento de que las funciones mayor y menor sean absolutamente continuas.

Definición 14.4.7. *Una función $f : [a,b] \to \overline{\mathbb{R}}$ es **Perron integrable** sobre $[a,b]$ si f tiene al menos una función mayor V y al menos una función menor sobre $[a,b]$ y son iguales los números*

$$\sup\left\{V_a^b : V \text{ es una función menor para } f\right\},$$
$$\inf\left\{U_a^b : U \text{ es una función mayor para } f\right\}.$$

Si f es Perron integrable sobre $[a,b]$, entonces el valor común en la definición anterior se llama la **integral de Perron** de f sobre $[a,b]$ y es denotada por

$$(\mathrm{P})\int_a^b f\, dt.$$

El conjunto de todas las funciones $f : [a,b] \to \overline{\mathbb{R}}$ que son Perron integrables sobre $[a,b]$ se denotará por $\mathcal{P}\mathrm{e}([a,b])$.

El siguiente resultado es una consecuencia inmediata de la definición anterior.

Teorema 14.4.8. *Una función medible $f : [a,b] \to \overline{\mathbb{R}}$ es **Perron integrable** sobre $[a,b]$ si, y sólo si, para cada $\varepsilon > 0$ existe una **función mayor** U y una **función menor** V de f sobre $[a,b]$ tal que $U_a^b - V_a^b < \varepsilon$.*

El Teorema Fundamental del Cálculo para la integral de Perron es totalmente trivial. En efecto, si $F : [a,b] \to \mathbb{R}$ es continua sobre $[a,b]$ y es diferenciable sobre $[a,b]$, entonces F es claramente una función mayor y menor para F' y, por lo tanto, $F' \in \mathcal{P}\mathrm{e}([a,b])$. Así,

$$(\mathrm{P})\int_a^b F'\,dt \;=\; F(b) - F(a).$$

Teorema 14.4.9 (Denjoy-Perron-Henstock-Kurzweil). *Sea $f : [a,b] \to \mathbb{R}$ una función. Las siguientes condiciones son equivalentes:*

(1) *f es Denjoy integrable.*

(2) *f es Perron integrable.*

(3) *f es Henstock-Kurzweil integrable.*

Una demostración de que las integrales de Denjoy, Perron y Henstock-Kurzweil son todas equivalentes se puede consultar, por ejemplo, en [66], Chapter 11.

14.4.3. La Integral de Denjoy Distribucional

.

Ahora presentaremos otra teoría de integración basada en el método descriptivo de Denjoy. Esta nueva integral, al igual que las dos anteriores, no necesita ninguna Teoría de la Medida para su construcción, así como tampoco particiones. Más aún, dicha integral incluye a todas las integrales estudiadas anteriormente en estas notas. Lo único que se requiere para definirla es la noción de derivada distribucional y la integración de Riemann de funciones continuas.

Las funciones generalizadas, un objeto matemático que generaliza la noción de función y la de medida, fueron introducidas por Serguéi L. Sóbolev (1908-1989) en el año 1935 y formalizadas por Laurent Schwartz (1915-2002) a finales de la década de 1940 en una nueva teoría llamada Teoría de Distribuciones.

Sea $C^0(\mathbb{R}) = C(\mathbb{R})$ y para cada entero $m \geq 1$ denote por $C^m(\mathbb{R})$ la clase de todas las funciones $f : \mathbb{R} \to \mathbb{R}$ que poseen *derivadas continuas de orden menor o igual a m*. Por $C^\infty(\mathbb{R})$ entenderemos el espacio vectorial de todas las funciones $\phi : \mathbb{R} \to \mathbb{R}$ que tienen *derivadas continuas de todos los ordenes*. Observe que

$$C^\infty(\mathbb{R}) \;=\; \bigcup_{m=0}^{\infty} C^m(\mathbb{R}).$$

Defina ahora

$$C_c^\infty(\mathbb{R}) \;=\; \big\{\phi \in C^\infty(\mathbb{R}) \,:\, \phi \text{ tiene soporte compacto}\big\}.$$

Si $K \subseteq \mathbb{R}$ es compacto, entonces escribiremos

$$C_c^\infty(K) \;=\; \big\{\phi \in C_c^\infty(\mathbb{R}) \,:\, \mathrm{sop}(\phi) \subseteq K\big\}.$$

Observe que si $\phi \in C_c^\infty(K)$, entonces $\phi^{(m)}$ es uniformemente continua sobre K para cada entero $m \geq 0$, por lo que podemos considerar

$$\|\phi^{(m)}\|_{\infty,K} = \sup_{x\in K} |\phi^{(m)}(x)|.$$

Cualquier función $\phi \in C_c^\infty(\mathbb{R})$ será llamada una **función de prueba** o **función test**. Por ejemplo, la función

$$\phi_0(t) = \begin{cases} e^{1/t} & \text{si } t < 0, \\ 0 & \text{si } t \geq 0 \end{cases}$$

es un miembro de $C^\infty(\mathbb{R})$, mientras

$$\phi(x) = \phi_0(|x|-1) \quad \text{para todo } x \in \mathbb{R} \tag{1}$$

pertenece a $C_c^\infty(\mathbb{R})$. Observe que el soporte de ϕ es igual a $[-1,1]$.

Un puede introducir una topología sobre $C_c^\infty(\mathbb{R})$ declarando que:

Definición 14.4.10. *Una sucesión $(\varphi_n)_{n=1}^\infty$ en $C_c^\infty(\mathbb{R})$ **converge** a $\phi \in C_c^\infty(\mathbb{R})$ si existe un conjunto compacto $K \subseteq \mathbb{R}$ tal que $\operatorname{sop}(\phi_n) \subseteq K$ para todo $n \in \mathbb{N}$ y para cada entero $m \geq 0$, la sucesión de derivadas $\phi_n^{(m)}$ converge a $\phi^{(m)}$ uniformemente sobre K.*

El espacio topológico $C_c^\infty(\mathbb{R})$ dotado de la topología anterior será denotado por $\mathcal{D}$. El dual topológico de $\mathcal{D}$, como es costumbre, lo escribiremos como $\mathcal{D}'$. Cada elemento $f \in \mathcal{D}'$ será llamado una **distribución** sobre $\mathbb{R}$ y a $\mathcal{D}$ lo nombraremos como el **espacio de todas las distribuciones** sobre $\mathbb{R}$. Si $f \in \mathcal{D}'$ escribiremos $\langle f, \phi\rangle$ para denotar la imagen de cualquier $\phi \in \mathcal{D}$ por medio de f.

Una manera de caracterizar a los elementos de $\mathcal{D}'$ es por medio del siguiente resultado.

Teorema 14.4.11. *Un funcional $f \in \mathcal{D}'$ si, y sólo, si para cada compacto $K \subseteq \mathbb{R}$, existe una constante $C \geq 0$ y un entero $N \geq 0$ tal que*

$$|\langle f, \phi\rangle| \leq C \sum_{m=0}^{N} \|\phi^{(m)}\|_{\infty,K}$$

para cualquier $\phi \in C_c^\infty(\mathbb{R})$ con $\operatorname{sop}(\phi) \subseteq K$.

Ejemplo 14.4.1. *Si $f \in L^1_{\text{loc}}(\mathbb{R})$, entonces $T_f \in \mathcal{D}'$, donde*

$$\langle T_f, \varphi\rangle = (\mathrm{L})\int_{\mathbb{R}} f\varphi\, d\mu \quad \textit{para toda } \varphi \in C_c^\infty(\mathbb{R}).$$

En efecto, como f es localmente integrable, se sigue del Teorema 12.3.2, página 821, que

$$(\mathrm{L})\int_K |f|\, d\mu < +\infty \quad \text{para cualquier compacto } K \subseteq \mathbb{R}.$$

Suponga ahora que $\phi \in C_c^\infty(\mathbb{R})$ y sea K un subconjunto compacto tal que $\operatorname{sop}(\phi) \subseteq K$. Entonces

$$\left|(\mathrm{L})\int_{\mathbb{R}} f\phi\, d\mu\right| \;\leq\; \|\,\phi\,\|_{\infty,K}\,(\mathrm{L})\int_K |f|\, d\mu \;<\; +\infty$$

de donde se sigue que T_f es continua y, por supuesto, lineal.

Observe que si $f \in C^1(\mathbb{R})$ y $\phi \in C_c^\infty(\mathbb{R})$, entonces integrando por partes resulta que

$$(\mathrm{L})\int_{\mathbb{R}} f'\phi\, d\mu \;=\; -\,(\mathrm{L})\int_{\mathbb{R}} f\phi'\, d\mu$$

Imitando este comportamiento podemos convenir que:

Definición 14.4.12. *Sea $f \in \mathcal{D}'$. Definimos la* ***derivada distribucional*** $\mathrm{D_d} f \in \mathcal{D}'$ *por*

$$\langle \mathrm{D_d} f, \phi\rangle \;=\; -\langle f, \phi'\rangle \quad \textit{para cualquier } \phi \in C_c^\infty(\mathbb{R}).$$

No es difícil establecer que la aplicación $f \mapsto \mathrm{D_d} f$ es lineal y continua de $\mathcal{D}'$ en $\mathcal{D}'$. De allí que la derivada distribucional posee propiedades más agradables que la derivada clásica ya que, por ejemplo, límite y diferenciación se pueden intercambiar indistintamente; en otras palabras, si $f_n \to f$ en $\mathcal{D}'$, entonces $\mathrm{D_d} f_n \to \mathrm{D_d} f$ en $\mathcal{D}'$. Por otro lado, la derivada clásica de la función de Cantor φ_Γ es 0 casi-siempre pero su derivada distribucional no lo es. En efecto, si $\phi \in C_c^\infty((0,1))$, se tiene que

$$\langle \mathrm{D_d}\varphi_\Gamma, \phi\rangle \;=\; -\langle \varphi_\Gamma, \phi'\rangle \;=\; -\int_0^1 \varphi_\Gamma \phi'\, d\mu,$$

lo cual muestra que $\mathrm{D_d}\varphi_\Gamma \neq 0$.

Denotemos por $C_{-\infty}^{+\infty}(\mathbb{R})$ *el espacio vectorial de todas las* ***funciones continuas*** $F : \mathbb{R} \to \mathbb{R}$ *tales que los límites*

$$\lim_{x\to-\infty} F(x) \qquad \text{y} \qquad \lim_{x\to+\infty} F(x)$$

ambos existen en $\mathbb{R}$. Si $F \in C_{-\infty}^{+\infty}(\mathbb{R})$, definimos

$$F(-\infty) \;=\; \lim_{x\to-\infty} F(x) \qquad \text{y} \qquad F(+\infty) \;=\; \lim_{x\to+\infty} F(x).$$

Por ejemplo, si $F(x) = x$ para todo $x \in \mathbb{R}$, resulta que F es continua en $\mathbb{R}$ pero no pertenece a $C_{-\infty}^{+\infty}(\mathbb{R})$. Por otro lado, la función $F : \mathbb{R} \to \mathbb{R}$ dada por $F(x) = \operatorname{arctang}(x)$ pertenece a $C_{-\infty}^{+\infty}(\mathbb{R})$ ya que $F(-\infty) = -\pi/2$ y $F(+\infty) = \pi/2$.

Defina

$$\mathcal{B}_C \;=\; \Big\{F \in C_{-\infty}^{+\infty}(\mathbb{R}) \,:\, F(-\infty) = 0\Big\}.$$

Observe que la función $F(x) = \operatorname{arctang}(x) + \pi/2 \in \mathcal{B}_C$. Similarmente, la función dada en (1) también pertenece a $\mathcal{B}_C$ por tener soporte compacto.

Nótese que $\mathcal{B}_C \subseteq C_b(\mathbb{R})$. Para ver esto, sea $F \in \mathcal{B}_C$ y sea $\varepsilon = 1$. Por definición, existen $a, b \in \mathbb{R}$ con $a < b$ tales que

$$|F(x)| \;=|F(x) - F(-\infty)| \;<\; 1 \quad \text{si } x < a \qquad \text{y} \qquad |F(x) - F(+\infty)| \;<\; 1 \quad \text{si } x > b.$$

En particular, si $x > b$, entonces $|F(x)| \leq |F(x) - F(+\infty)| + |F(+\infty)| < 1 + |F(+\infty)|$. Puesto que F es acotada sobre el intervalo $[a, b]$, digamos por $C > 0$, resulta que

$$\| F \|_\infty \leq \text{máx}\{C, 1 + |F(+\infty)|\}$$

y, por lo tanto, $F \in C_b(\mathbb{R})$. Es un ejercicio sencillo establecer $\mathcal{B}_C$ es un espacio de Banach bajo la norma $\| F \|_\infty = \sup\{|F(x)| : x \in \mathbb{R}\}$. Ahora definiremos la integral distribucional de Denjoy.

Definición 14.4.13. *Una función $f \in \mathcal{D}'$ se dice que es* ***distribucionalmente Denjoy integrable*** *sobre $\mathbb{R}$ si existe una función $F \in \mathcal{B}_C$ tal que $\mathrm{D_d}F = f$.*

Por brevedad escribiremos f es **DD-integrable** en lugar de "distribucionalmente Denjoy integrable" y denotaremos por $\mathcal{D}_{\text{dist}}(\mathbb{R})$ el espacio de todas las funciones DD-integrables sobre $\mathbb{R}$, es decir,

$$\mathcal{D}_{\text{dist}}(\mathbb{R}) = \{f \in \mathcal{D}' : f = \mathrm{D_d}F \text{ para alguna } F \in \mathcal{B}_C\}.$$

Cualquier función $F \in \mathcal{B}_C$ tal que $\mathrm{D_d}F = f$ se llamará una **DD-primitiva** o **primitiva distribucional** de f. Observe que si $f \in \mathcal{D}_{\text{dist}}(\mathbb{R})$, entonces para toda $\phi \in C_c^\infty(\mathbb{R})$ se tiene que

$$\langle f, \phi\rangle = \langle \mathrm{D_d}F, \phi\rangle = -\langle F, \phi'\rangle = -\int_{-\infty}^{+\infty} F\phi' \, d\mu.$$

Puesto que F y ϕ' son ambas continuas sobre $\mathbb{R}$ y ϕ' tiene soporte compacto, resulta que la integral

$$\int_{-\infty}^{+\infty} F\phi' \, d\mu$$

existe como una integral de Riemann. En este caso, definimos la **integral distribucional de Denjoy** de f como

$$(\mathrm{DD})\int_{-\infty}^{+\infty} f = F(+\infty). \tag{2}$$

Si $-\infty \leq a < b \leq +\infty$, la **integral distribucional de Denjoy** de f sobre $[a, b]$ se define como

$$(\mathrm{DD})\int_a^b f = F(b) - F(a).$$

Observaciones.

(1) Nótese que si $f \in \mathcal{D}_{\text{dist}}(\mathbb{R})$, entonces f posee muchas DD-primitivas, todas diferentes por una constante, pero f tiene exactamente una DD-primitiva en $\mathcal{B}_C$. En efecto, si $F_1, F_2 \in \mathcal{B}_C$ y $\mathrm{D_d}F_1 = f$ y $\mathrm{D_d}F_2 = f$, entonces la linealidad de la derivada nos muestra que $\mathrm{D_d}(F_1 - F_2) = \mathrm{D_d}F_1 - \mathrm{D_d}F_2 = 0$ y como $F_1(-\infty) = F_2(-\infty) = 0$, resulta que $F_1 = F_2$. Esto prueba que **la integral distribucional de Denjoy es única**.

Teorema 14.4.14 (Fundamental del Cálculo). *Sea $f \in \mathcal{D}_{\text{dist}}(\mathbb{R})$ y defina*

$$F(x) = (\mathrm{DD})\int_{-\infty}^{x} f \quad \textit{para todo } x \in \mathbb{R}.$$

Entonces $F \in \mathcal{B}_C$ y $\mathrm{D_d}F = f$.

Prueba. Puesto que $f \in \mathcal{D}_{\text{dist}}(\mathbb{R})$, existe un único $G \in \mathcal{B}_C$ tal que $D_d G = f$ y, por lo tanto,

$$G(x) \;=\; (\text{DD})\int_{-\infty}^{x} f \;=\; F(x) \quad \text{para todo } x \in \mathbb{R}.$$

Por esto, $F = G \in \mathcal{B}_C$ y $D_d F = D_d G = f$. ■

(2) Si f es Riemann integrable sobre $[a, b]$, entonces su integral indefinida $F(x) = (\text{R})\int_a^b f\,dx$ es una función Lipschitz y $F'(x) = f(x)$ para todos los puntos $x \in [a, b]$ donde f es continua. Por el Teorema de Vitali-Lebesgue, sabemos que f es continua casi-siempre. Sea $E \subseteq [a, b]$ el conjunto nulo donde f no es continua. Entonces

$$\langle F', \phi \rangle \;=\; \int_{-\infty}^{+\infty} F'\phi \;=\; \int_{-\infty}^{+\infty} f\phi$$

ya que la última integral no cambia sobre E. Esto nos indica que $f \in \mathcal{D}_{\text{dist}}(\mathbb{R})$ y, por lo tanto, **la integral distribucional de Denjoy contiene a la integral de Riemann**.

(3) Si $f \in L_1(\mathbb{R}, \mu)$, entonces $F(x) = (\text{L})\int_{-\infty}^{x} f\,d\mu$ define una función perteneciente al conjunto $AC(\mathbb{R}) \cap C(\mathbb{R}) \subsetneqq \mathcal{B}_C$. Puesto que $F' = f$ casi-siempre, **la integral distribucional de Denjoy contiene a la integral de Lebesgue**. Más aun, si φ_Γ es la función de Cantor, entonces $\varphi'_\Gamma(x) = 0$ para casi todo $x \in [0, 1]$ y como φ_Γ es de variación acotada, su derivada es Lebesgue integrable sobre $[0, 1]$ y

$$(\text{L})\int_0^x \varphi'_\Gamma(x)\,d\mu \;=\; 0 \quad \text{para todo } x \in [0, 1].$$

Sin embargo, como $\varphi'_\Gamma \in \mathcal{D}_{\text{dist}}(\mathbb{R})$ resulta que

$$(\text{DD})\int_0^x \varphi'_\Gamma \;=\; \varphi_\Gamma(x) \quad \text{para todo } x \in [0, 1].$$

(4) Si f es Denjoy integrable, entonces su primitiva es una función ACG_* y por un razonamiento similar al caso anterior, **la integral distribucional de Denjoy contiene a la integral de Denjoy**. En particular, como las integrales de Denjoy, Perron y Henstock-Kurzweil son todas equivalentes y sus respectivas integrales son iguales, resulta que **la integral distribucional de Denjoy las contiene a todas ellas**.

(5) $\mathcal{D}_{\text{dist}}(\mathbb{R})$ es un **espacio de Banach** bajo la norma $\| f \| = \| F \|_\infty$, donde F es la única función en $\mathcal{B}_C$ tal que (2) se cumple. Además, $\mathcal{D}_{\text{dist}}(\mathbb{R})$ es **separable** y $L_1(\mathbb{R}, \mu)$ es **denso** en $\mathcal{D}_{\text{dist}}(\mathbb{R})$. Observe que los espacios de Banach $\mathcal{D}_{\text{dist}}(\mathbb{R})$ y $\mathcal{B}_C$ son isométricamente isomorfos.

Una Pequeña Reflexión

No queremos finalizar estas notas sin antes hacer la siguiente reflexión: R. A. Gordon, en su libro [66], p. 154, se plantea la siguiente pregunta: “Suponga que una persona ya ha estudiado la integral de Riemann. ¿Cuál es la siguiente integral que esa persona debería estudiar?” La respuesta casi universal a esta interrogante, responde Gordon, es, sin lugar a dudas, la *integral de Lebesgue*. Por supuesto, como se ha evidenciado en estas notas, la integral de Lebesgue es, desde todo punto de vista, una “mejor integral” que la integral de Riemann lo que se logra con una gran dosis de conocimiento en teoría de la medida, funciones medibles y, por supuesto, del análisis real. Los grandes beneficios que se logran con las aplicaciones de la integral de Lebesgue bien vale tal esfuerzo. Por otro lado, algunos matemáticos soportan la idea de que la integral de Henstock-Kurzweil debería ser la siguiente integral a estudiar después de la de Riemann. Un buen argumento a su favor es que la gran mayoría de los resultados obtenidos referentes a la integral de Henstock-Kurzweil se pueden demostrar con muy poco material de la teoría de la medida y, además, como ya vimos, dicha integral posee todo el poder de la integral de Lebesgue y satisface la fórmula de Newton-Leibniz. Es decir, con menos trabajo, se obtiene una mejor integral. A pesar de todos esos beneficios que uno encuentra en la integral de Henstock-Kurzweil, por qué, se pregunta de nuevo Gordon, no ha habido una revolución en los cursos de Análisis Real para enseñar dicha integral a estudiantes graduados? Gordon se pasea por un conjunto de posibles respuestas a esa interrogante. Una de la más acertadas, en mi humilde opinión, es que, si bien la definición de la integral de Henstock-Kurzweil es simple, las demostraciones de las propiedades interesantes y profundas son sutiles y difíciles. Pero además, no es fácil extender la integral de Henstock-Kurzweil a cualquier conjunto medible en un contexto abstracto. Tampoco se cumple el Teorema de Riemann-Lebesgue lo que dificulta una rica Teoría de las Series de Fourier como en el caso de $L_1([a,b])$, etc. Mi reflexión es la siguiente: ¿cómo podemos saber realmente que la integral de Lebesgue es tan poderosa si ella no se estudia en profundidad? No basta, a mi parecer, demostrar que $L_1([a,b]) = \mathcal{L}_{\mathrm{HK}}([a,b]) = \mathcal{M}c([a,b]) \subsetneqq \mathcal{HK}([a,b])$ si no se conocen previamente los exquisitos y profundos resultados que se pueden lograr con la integral de Lebesgue y sus inimaginables aplicaciones en muchas ramas del quehacer matemático. De modo que si se desea comparar la integral de Henstock-Kurzweil con la de Lebesgue o, en su defecto, elegir cuál de ellas se debe enseñar después de conocer la integral de Riemann, me inclino por la enseñanza de la integral de Lebesgue para luego ir a la de Henstock-Kurzweil. Esto, en definitiva, es el objetivo fundamental de estas notas: presentar las integrales de Riemann, Lebesgue y de Henstock-Kurzweil en ese orden, mostrando algunas de las relaciones más profundas que existen entre ellas. Tal vez la poca bibliografía existente en español sobre el tema y lo maravilloso que son tales integrales me animaron a emprender tan agradable tarea.

14.5. Problemas

(1) Sea $f : [0,1] \to \mathbb{R}$ definida por

$$f(x) = \begin{cases} q & \text{si } x = p/q \in [0,1] \text{ con } q, q \text{ primos relativos,} \\ 0 & \text{si } x \in [0,1],\ x \text{ irracional.} \end{cases}$$

Pruebe que $f \in \mathcal{HK}([a,b])$ pero no es Riemann integrable.

(2) Un conjunto $E \subseteq [a,b]$ se llama **integrable** si $\chi_E \in \mathcal{HK}([a,b])$. Denote por $\mathcal{M}_{\mathrm{HK}}([a,b])$ la colección de todos conjuntos integrables en $[a,b]$. Demuestre que $\mathcal{M}_{\mathrm{HK}}([a,b])$ es una σ-álgebra.

(3) Sea $f \in \mathcal{HK}([a,b])$. Demuestre que $f \in L_1([a,b])$ si, y sólo si, para cada $\varepsilon > 0$, existe un calibrador δ sobre $[a,b]$ y un número positivo η tal que $|S(f,\mathcal{P})| < \varepsilon$ para cualquier subpartición $\mathcal{P} = (t_i, I_i)_{i=1}^n$ subordinada a δ y $\sum_{i=1}^n \mu(I_i) < \eta$.

(4) Sea $f \in \mathcal{HK}([a,b])$. Demuestre que el conjunto $\mathcal{F} = \{f \cdot \chi_{[a,c]} : c \in (a,b)\}$ es equi-integrable sobre $[a,b]$.

(5) Sea $\mathcal{F} \subseteq \mathcal{HK}([a,b])$ equi-integrable. Si $\mathcal{F}^*$ consiste de todos los límite puntuales de sucesiones de funciones en $\mathcal{F}$, entonces pruebe que $\mathcal{F}^*$ también es equi-integrable.

(6) Una colección de funciones $\mathcal{F}$ incluidas en $\mathbb{R}^{[a,b]}$ se dice que es **equi-continua** en $c \in [a,b]$ si para cada $\varepsilon > 0$, existe un $\delta_\varepsilon(c) > 0$ tal que, para cualquier $x \in [a,b]$ satisfaciendo $|x - c| < \delta_\varepsilon(c)$, se cumple que

$$|F(x) - F(c)| \;<\; \varepsilon \quad \text{para todo } F \in \mathcal{F}.$$

(a) Demuestre que si $\mathcal{F} \subseteq \mathcal{HK}([a,b])$ es equi-integrable sobre $[a,b]$ y si el conjunto $\mathcal{F}(c) = \{F(c) : F \in \mathcal{F}\}$ es acotado, entonces la colección

$$\mathcal{F}\mathrm{ind} \;=\; \left\{F \,:\, F(x) = (\mathrm{HK})\!\int_a^x f\,dt,\; f \in \mathcal{F}\right\}$$

es equi-continua en c.

(b) Si $\mathcal{F}$ es equi-continua en cualquier punto de $[a,b]$, entonces $\mathcal{F}$ es **uniformemente equi-continua** sobre $[a,b]$, lo cual quiere decir que: para cada $\varepsilon > 0$, existe un $\delta_\varepsilon > 0$ tal que, para cualquiera $x, y \in [a,b]$ satisfaciendo $|x - y| < \delta_\varepsilon$, se cumple que

$$|F(x) - F(c)| \;<\; \varepsilon \quad \text{para todo } F \in \mathcal{F}.$$

(7) Sea $(f_n)_{n=1}^{\infty}$ una **sucesión creciente** en $\mathcal{HK}([a,b])$ tal que

$$\sup\left\{(\mathrm{HK})\!\int_a^b f_n\,dt \,:\, n \in \mathbb{N}\right\} \;<\; +\infty.$$

(a) Pruebe que el conjunto $E = \{x \in [a,b] : \lim_{n\to\infty} f_n(x) = \infty\}$ es de medida cero.

(b) Si se define

$$f(x) \;=\; \begin{cases} \lim\limits_{n\to\infty} f_n(x) & \text{si } x \in [a,b] \setminus E, \\ 0 & \text{si } x \in E, \end{cases}$$

demuestre entonces que $f \in \mathcal{HK}([a,b])$ y

$$(\mathrm{HK})\!\int_a^b f\,dt \;=\; \lim_{n\to\infty} (\mathrm{HK})\!\int_a^b f_n\,dt.$$

(8) Demuestre que $\mathrm{AC}([a,b])$ es denso en $C([a,b])$.

Bibliografía

[1] S. Abbott, **Understending Analysis**, Springer, 2010.

[2] M. Aigner, G. M. Ziegler, **Proofs from THE BOOK**, Fourth Edition, Springer-Verlag Berlin Heidelberg, 2010.

[3] C. D. Aliprantis and K. C. Border, **Infinity Dimensional Analysis: A Hitchhiker's Guide** (3rd ed.), Springer, 2005.

[4] P. Antosik and C. Swartz, **Matrix Methods in Analysis**, Lecture Notes in Math. 1113, Springer-Verlag, 1985.

[5] T. M. Apostol, **Análisis Matemático**, 2a. Edición, Editorial Reverté, España, 1976.

[6] I. Arandelović, *An inequality for the Lebesgue measure*, Univ. Beog. Publ. Elek. Fak. ser. Math. 15 (2004) 85-86.

[7] S. Banach, *Sur le problème de la mesure*, Fundamenta Mathematicae 4(1923), 7-33.

[8] S. Banach, C. Kuratowski, *Sur une généralisation du probème de la mesure*, Fundamenta Mathematicae 16(1929), 127-131.

[9] S. Banach, A. Tarski, *Sur la décomposition des ensembles de points en parties respectivement congruentes*, Fundamenta Mathematicae VI, (1924) 244-277.

[10] R. G. Bartle **A Modern Theory of Integration**, The Amer. Math. Soc., 2001

[11] B. Levi, *Sopra l'integrazione delle serie*, Rendiconti Istituto Lombardo Scienze 2 (1906), no. 39, 775–780.

[12] J. J. Benedetto and W. Czaja, **Integration and Modern Analysis**, Birhaüser Boston, 2009.

[13] R. P. Boas, **A Primer of Real Functions**, The Carns Mathematical Monographs no. 13, 4th, ed. The Mathematical Association of America, Washington, DC, 1996.

[14] V. I. Bogachev, **Measure Theory, Vol. I**, Springer-Verlag Berlin Heidelberg, 2007.

[15] B. Bolzano, **Paradoxes of the Infinite**, Routledge and Kegan Paul, London 1950.

[16] B. Bongiorno, *On the C-integral*, Conferencia en Toronto, ON Canada, Sep. 23-24, 2000.

[17] B. Bongiorno, *Un nuovo integrale per il problema delle primitive*, Le Matematiche, Vol. LI(1996), 299-313.

[18] B. Bongiorno, L. Di Piazza, D. Preiss, *A constructive minimal integral which includes Lebesgue integral functions and derivatives*, J. London Math. Soc. (2) 62 (2000), 117-126.

[19] M. W. Botsko, *An Elementary Proof That a Bounded a.e. Continuous Functions is Riemann Integrable*, The Amer. Math. Montly, 95 (1988), 249-252.

[20] M. W. Botsko, *An elementary proof of Lebesgue's differentiation theorem*, The Amer. Math. Monthly, 110 (2003), 834–838.

[21] M. W. Botsko, *A Fundamental Theorem of Calculus that Applies to All Riemann Integrable Functions*, Math. Magazine, 64 (1991), 347-348.

[22] N. Bourbaki, **Topologie générale**, Chap. 9, Paris 1958.

[23] W. Brito, **El Teorema de Categoría de Baire. Sus Aplicaciones**. Editorial Académica Española, Alemania, 2011.

[24] A. M. Bruckner, **Differentiation of Real Functions**, Lecture Notes in Math. Vol. 659, Spriger Verlag, Berlin, 1978.

[25] A. M. Bruckner, J. B. Bruckner and B. S. Thomson, **Real Anlysis**, Prentice-Hall, 1994.

[26] A. M. Bruckner, R. J. Fleissner and J. Foran, *The minimal integral which includes Lebesgue integrable functions and derivatives*, Colloq. Math. 50 (1986), 289-293.

[27] P. S. Bullen, **Handbook of Means and their inequalities**, Kluwer, ordrecht, 2003.

[28] F. E. Burk, **A Garden of Integrals**, The Mathematical Asoc. of America, 2007.

[29] J. F. Burnol, *Le Théorème de la convergence dominée pour les fonctions intégrables au sens de Riemann*, Por publicarse.

[30] T. Carlson, *Extending Lebesgue by infinitely many sets*, Pacific J. Math. 15(1984), 33-45.

[31] N. L. Carothers, **Real Analysis**, Cambridge University Press, 2000.

[32] A. L. Cauchy, **Course d'Analyse**, (1821).

[33] T. K. Chandra, **The Borel-Cantelli Lemma**, Springer, 2012.

[34] K. Cielsielski, *How Good is Lebesgue Measure?*, The Math. Intelligencer, 11(1989), 54-58.

[35] G. Choquet, **Lectures on Analysis, Vol. I**, W. A. Benjamin, Inc., New York-Amstrdam, 1969.

[36] R. G. Cooke, **Infinite Matrices and Sequence Spaces**, MacMillan, London, 1950.

[37] G. de Barra, **Measure Theory and Integration**, 2nd Edition, Woodhead Publishing, 2003.

[38] R. G. Bartle, **A Modern Theory of Integration**, The Amer. Math. Soc., 2001.

[39] P. J. Daniell, *A General form of Integral*, Annals of Mathematics, 10(1918), 279-294.

[40] A. G. Das, **The Riemann, Lebesgue and Generalized Riemann Integrals**, Narcosa Publishing House, New Delhi, 2008.

[41] A. Dasgupta, **Set Theory** - With an Introduction to Real Point Sets, Springer, N. Y., 2014.

[42] R. de Silva, *A concise elementary proof of Arzelà's bounded convergence theorem*. Por publicarse.

[43] J. Diestel, **Sequences abd Series in Banach Spaces**, Springer-Verlag, 1984.

[44] J. Diestel, *Uniform Integrability: An Introduction*, Rendiconti dellínstituto di Matematica dellÚniversitá di Trieste, 1991, 41-80.

[45] U. Dini, **Fondamenti per la teorica della funzioni di variabili reali**, T. Nistri e C., Pisa, 1878.

[46] J. L. Doob, *Regularity properties of certain chance variables*, Trans. Amer. Math. Soc. 47 (1940), 455–486.

[47] F. del Castillo, **Análisis Matemático II**, Alhambra Universidad, 1980.

[48] J. Diestel and J. J. Uhl Jr., **Vector Measures**, AMS, 1979.

[49] J. Dieudonné, **Fundamentos de Análisis Moderno**, Editorial Reverté, S. A., 1960.

[50] P. Dugac, **Histoire de l'Analyse - Autour de la Notion de Limite et de ses Voisinages**, Vuibert, Paris, 2003.

[51] J. Dugundji, **Topology**, Allyn and Bacon, Inc., 1966.

[52] N. Dunford and J. T. Schwartz, **Linear Operators, Part I**, Interscience Publishers, Inc., New York, 1957.

[53] W. Dunham, *A Historical Gem from Vito Volterra*, Math. Magazine, 63 (1990), 234-237.

[54] W. Dunham, **The Calculus Gallery: Masterpieces from Newton to Lebesgue**, Pnnceton University Press, 2005.

[55] M. Eisenberg, **Axiomatic Theory of Seta and Classes**, Holt, Rinehart and Winston, Inc., 1971.

[56] V. Ene, *An Elementary Proof of the Banach-Zarecki Theorem*, Real Analysis Exchange, 23 (1998-1999), 295-301.

[57] R. Engelking, **General Topology**, Polish Scientific Publishers, Warszawa, 1977. Amer. J. Math., 74 (1952), 168-186.

[58] C. A. Faure, *A Short Proof of Lebesgue's Density Theorem*, The Amer. Math. Monthly, 109(2002), 194-196.

[59] J. Foran, **Fundamentals of Real Analysis**, Marcel Dekker, Inc., 1991.

[60] D. Fremlin **Measure Theory Vol. I-V**

[61] B. R. Gelbaum and J. M. H. Olmsted, **Counterexamples in Analysis**, Holden-Day 3rd printing (1966).

[62] L. Gillman *Two Classical Surprises Concerning the Axiom of Choice and the Continuum Hypothesis*, Amer. Math. Soc. 109 (2002), 544-553.

[63] D. Goldrei, **Classic Set Theory - A guided independent study**, Chapman & Hall, 1996.

[64] R. A. Gordon, *A Convergence Theorem for the Riemann Integral*, Math. Magazine, 73 (2000), 141-147.

[65] R. A. Gordon, *Some Comments on the McShane and Henstock Integrals*, Real Anal. Exchange, 23(1),(1998-99), 329-342.

[66] R. A. Gordon, **The Integrals of Lebesgue, Denjoy, Perron, and Henstock**, Graduate Studies in Math. 4, 1994, Amer. Math. Soc.

[67] R. A. Gordon, *The Use of Tagged Partions in Elementary Analysis*, The Amer. Math. Montly, 105 (1998), 107-117.

[68] S. B. Guthery, **A Motif of Mathematics: History and Applications of the Mediant and Farey Sequences**, Docent Press, Boston, 2011.

[69] J. W. Hagood, *The Lebesgue Differentiation Theorem via nonoverlapping interval covers*, Real Analysis Exchange, 29(2003-2004), 953-956.

[70] F. Hernández Hernández, **Teoría de Conjuntos**, Sociedad Matemática Mexicana (2da. Edition), 1998.

[71] H. Herrlich, **Axiom of Choice**, Springer-Verlag Berlin Heidelberg, 2006.

[72] E. Hewitt and K. Stromberg, **Real and Abstract Analysis**, Springer-Verlag, 1975.

[73] T. H. Hildebrandt, *Integration in Abstract Spaces*, Bull. Amer. Mmath. Soc., 59(1953), 111-139.

[74] T. Jech, **Set Theory** - The Third Millennium Edition, revised and expanded. Springer, 2002.

[75] A. B. Kharazishvili, **Set Theoretical Aspect of Real Analysis**, CRC Press Taylor and Francis Group, 2015.

[76] A. B. Kharazishvili, **Strange Functions in Real Analysis**, 2nd Edition, Chapman and Hill/CRC, Taylor and Francis Group, LLC, N.Y., 2006.

[77] A. B. Kharazishvili, **Nonmeasurable Sets and Functions**, Elsevier B. V., 2004.

[78] I. Kluvánek, *What is wrong with the teaching of calculus?*, Mathematické obzory 36 (1991), 23-49.

[79] K. Knopp, **Infinite Sequences and Series**, Dover Publications, 1956.

[80] R. I. Kraft, *What's the Difference Between Cantor Sets?*, The Amer. Math. Monthly, 102, No. 7(1993), 640-650.

[81] R. I. Kraft, *A Golden Cantor Sets?*, The Amer. Math. Monthly, 105, No. 8(1998), 718-725.

[82] M. Kuczma, **An Introduction to the Theory of Functional Equations and Inequalities**, 2nd Ed., Birkhäuser Verlag AG, 2009.

[83] K. Kuratowski, **Topology, Vol I**, Academic Press, New York, 1966.

[84] D. S. Kurtz, C. W. Swartz, **Theories of Integration - The Integrals of Riemann, Lebesgue, Henstock-Kurzweil, and McShane**, 2nd Edition, World Scientific, 2012.

[85] E. Lages Lima, **Espaços Métricos**, Projeto Euclides, São Paulo, SP, Brasil, 1977.

[86] S. Lang, **Real Analysis**, 2nd Edition, Addison-Wesley Publishing Company, Inc., 1983.

[87] J. A. LaVita, *A Necessary and Sufficient Condition for Riemann Integration*, The Amer. Math. Montly, 71 (1964), 193-196.

[88] H. Lebesgue, **Leçons sur l'Integration et la Recherche des Fonctions Primitives**, Gauthier Villars, Paris, 1904.

[89] H. Lebesgue, *Sur les intégrales singulières*, Ann. Fac. Sci. Toulouse 1 (1909), 4-117.

[90] T. Y. Lee, **Henstock-Kurzweil Integration on Euclidean Spaces**, World Scientific, 2011.

[91] G. Letta, *Une démonstration élémentaire du théorème de Lebesgue sur la derivation des fonctions croissantes*, L'Enseignement Math. 16 (1970), 177–184.

[92] L. M. Levine, *On a Necessary and Sufficient Condition for Riemann Integrability*, The Amer. Math. Montly, 84 (1977), 205.

[93] J. W. Lewin, *A truly elementary approach to the Bounded Convergence Theorem*, The Amer. Math. Monthly, 93 (1986), 395-397.

[94] P. Lévy, *An Extension of The Lebesgue Measure of Linear Sets*, Proceedings of the Fourth Berkeley Symposium in Mathematical Statistics and Probability, University of California Press, Berkeley, 1961, 275-287.

[95] S. Lojasiewicz, **An Introduction to the Theory of Real Functions**, John Wiley and Sons Ltd., 1988.

[96] J. Łoś and E. Marczewski, *Extension Measure*, Fundamenta Mathematicae 36(1949), 267-276.

[97] W. A. J. Luxemburg, *Arzelà's dominated convergence theorem for the Riemainn integral*, The Amer. Math. Monthly, 78 (1971), 970-979.

[98] F. A. Medvedev, **Scenes from the History of Real Functions**, Springer Basel AG, 1991.

[99] P. Mikusiński, M. D. Taylor, **An Introduction to Multivariable Analysis from Vector Manifold**, Springer Science+Business Media, LLC, 2002.

[100] G. H. Moore, *Lebesgue's Measure Pproblem and Zermelo's Axiom of Choice: The Mathematical Effects of a Philosophical Dispute*, Annals of the New York Academy of Sciences, 412(1983), 129-154.

[101] G. H. Moore, **Zermelo's Axiom of Choice: Its Origins, Development and Influence**, Springer-Verlag, 1982.

[102] R. L. Moore, **Foundations of Point Set Theory**, Revised ed., American Mathematical Society, Providence, RI 1962.

[103] J. R. Munkres, **Topología**, 2^a edición, Prentice Hall Inc. 2002.

[104] I. P. Natanson, **Theory of Functions of a Real Variable**, Vol. 1, Revised ed., Frederick Ungar Publishing Co., Inc., 1964.

[105] J. C. Oxtoby, **Measure and Category**, 2nd Edition, Springer-Verlag, New York, 1971.

[106] E. Pap, **Handbook of Measure Theory**, Vol. I, Elsevier Science B. V., 2002.

[107] K. R. Parthasarathy, **Probability Measures on Metric Spaces**, Academic Press, New York and London. 1967.

[108] M. de J. Perez-Jimenez, **Teoría de Clases y Conjuntos**, Edunsa-Viladomat, Barcelona, 1988.

[109] I. N. Pesin, **Classical and Modern Theories to Integration**, Cambridge University Press, Cambridge, 1993.

[110] J. F. Randolph, **Basic Real and Abstract Analysis**, Academic Press, New York and London, 1968.

[111] M. M. Rao, **Measure Theory and Integration**, 2nd Edition, Marcel Dekker, New York, 2006.

[112] F. Riesz, *Sur l'existence de la dérivée des fonctions monotones et sur quelques problemes qui s'y rattachent*, Acta Sci. Math. 5(1930-1932), 208-221.

[113] A, C. M. van Rooij and W. H. Schikhof, **A Second Course on Real Functions**, Cambridge University Press, 1982.

[114] H. Royden, **Real Analysis**, 2nd ed., Macmillan, New York, 1968.

[115] W. Rudin, **Real and Complex Analysis**, Tata McGraw-Hill Book Company, 1974.

[116] W. Rudin, **Principles of Mathematical Analysis**, International Student Analysis, McGraw-Hill, Inc., 3rd Edition, 1976.

[117] S. Saks, **Theory of the Integral**, Dover, New York, 1964.

[118] E. Schechter, **Handbook of Analysis and Its Foundations**, Academic Press, INC., USA, 1997.

[119] S. Schwabik, Ye Guoju, **Topics in Banach Space Integration**, World Scientific, 2005.

[120] S. Shelah, *Can you take Solovay's inaccesible away?*, Israel J. Math. 48(1984), 1-47.

[121] G. E. Shilov, B. L. Gurevich, **Integral, Measure and Derivative - A Unified Approach**, Dover Publications, Inc., New York, 1977.

[122] R. M. Solovay, *A model of set-theory in which every set of reals is Lebesgue measurable*, Annals of Math. Second Series 92(1970), 1-56.

[123] L. A. Steen and J. A. Seebach, Jr., **Counterexamples in Topology**, Dover Publications, Inc., New York, 1995.

[124] J. Stern, *Le Problème de la Mesure*, Seminaire Bourbaki, 36e (1983-84), 325-346.

[125] C. Swartz, **Introduction to Gauge Integrals**, World Scientific, 2001.

[126] T. Takagi, *A simple example of the continuous function without derivative*, Proc. Phys. Math. Soc. Japan, 1 (1903), 176-177.

[127] G. Takeuti, W. M. Zaring, **Introduction to Axiomatic Set Theory**, Springer-Verlag, 1970.

[128] E. Talvila, *The Distributional Denjoy Integral*, Real Anal. Exchange, vol. 33(1), 2007/2008, 51-82.

[129] A. E, Taylor, **General Theory of Functions and Integration**, Blaisdell Publishing Co., London, 1965.

[130] J. Thim, *Continuous Nowhere Differentiable Functions*, Master Thesis, Luleâ University of Technology, (2003), 1-94.

[131] S. Todorcevic, **Topics in Topology**, Lecture Notes in Math., Vol. 1652, Springer Verlag, 1997.

[132] C. J. de la Vallée Poussin, *Sur l'intégrale de Lebesgue*, Trans. Amer. Math. Soc. 16 (1915), 435–501.

[133] E. M. Vestrup, **The Theory of Measures and Integration**, John Wiley & Sons, 2003.

[134] G. Vitali, *Sul problema della misura dei gruppi di punti di una retta*, Bologna, 1905.

[135] G. Vitali, *Sull' integrazione per serie*, Palermo Rend. 23 (1907), 137–155.

[136] E. B. van Vleck, *On non-measurable sets of points, with an example*, Trans. Amer. Math. Soc. 9(1908), 237–244.

[137] S. Wagon, **The Banach-Tarski Paradox**, Cambridge University Press, 1985.

[138] A. J. Weir, **General Integration and Measure**, Cambridge University Press, 1974.

[139] S. Willard, **General Topology**, Addison-Wesley Publishing Company, Inc., 1970.

[140] G. L. Wise and E. B. Hall, **Counterexamples in Probability and Real Analysis**, Oxford University Press, Inc., 1993.

[141] J. Jeh, **Real Analysis - Theory of Measure and Integration**, 2nd Edition, World Scientific Pub., 2006.

[142] W. H. Young, *On the general theory of integration.*, Lond. Phil. Trans. (A) 204 (1905), 221–252.

[143] L. Zajiček, *An Elementary Proof of the one-dimensional Density Theorem*, The Amer. Math. Monthly 86(1979), 297-298.

Índice alfabético

2^{o} numerable, 119
M-Test de Weierstrass, 179
$\aleph$, 69
$\aleph_0$-primitiva, 436
δ-cubrimiento, 365
λ-sistema, 699
μ es completa, 280
μ es continua, 286, 287
μ es invariante por traslación, 285
μ es numerablemente aditiva, 280
μ-c.s., 375
μ-casi-siempre, 375
μ_n es invariante
  por rotación, 750
  por traslación, 742
π-sistema, 699
ρ-cubo, 735
σ-álgebra
  producto de Borel, 770
  de Borel, 266, 681
  de Borel en $\mathbb{R}^n$, 740
  de conjuntos, 258
  de Lebesgue, 261
  de Lebesgue en $\mathbb{R}^n$, 739
  de Lebesgue inducida, 262
  generada, 263
  numerablemente generada, 267
  producto, 720, 768
σ-álgebra de Baire, 372
σ-aditiva, 282
σ-aditividad de la integral, 532
C-integral, 904
$\mathrm{L}_1(X, \mathcal{M}, \lambda)$, 714
$\mathrm{L}_p(\mathbb{R}, \mathfrak{M}_\mu(\mathbb{R}), \mu)$, $1 < p < +\infty$, 630
$\mathrm{L}_\infty(X, \mu)$, 647
x-sección de f, 723
x-sección de un conjunto, 720
y-sección de f, 723
álgebra de conjuntos, 259, 716
álgebra de los cilindros, 769
árbol
  de Calkin-Wilf, 24
  de Stern-Brocot, 35
árbol de Calkin-Wilf, 24
átomo, 686
ínfimo de un conjunto, 86

absoluta continuidad de la integral, 581
anti-derivada de una función, 430
antiderivada, 409
aplicación abierta, 758
Axioma
  de Completitud de $\mathbb{R}$, 90
  de Determinación, 353
  de Elección, 47
  del Infimo, 90
  del Supremo, 90

base
  de Burstin, 368
  de Hamel, 51, 332, 333
  de una topología, 119
  en un espacio métrico, 127
  numerable, 127, 267
buen orden en un conjunto, 53

cápsula convexa, 193
círculos de Ford, 34
cadena, 50
cambio de variable, 759

caracterización de conjuntos abiertos, 122, 124
cardinal
  inaccesible, 73
  límite fuerte, 72
  regular, 73
  singular, 73
clase
  de equivalencia, 9
  monótona, 716
  monótona generada por conjuntos, 716
colección
  de conjuntos, 3
  independiente, 289
combinación convexa, 193
complemento de un conjunto, 5
completación
  de $\mathcal{R}([a,b])$, 593
  de $C([a,b])$, 593
  de un espacio métrico, 592
completación de una medida, 689
componente conexa, 123
condición (N) de Lusin, 455
condición de Banach, 357
conjunto
  F_σ, 120
  G_δ, 121, 168
  λ^*-medible, 693
  n-ario de Cantor, 237
  ínfimo de un, 86
  SVC, 230
  a lo más numerable, 11
  abierto, 119
  abierto en $\mathbb{R}$, 120
  abundante, 140
  acotado, 85, 128
  acotado inferiormente, 85
  acotado superiormente, 85
  ambiguo, 121
  analítico, 273
  básico, 148
  bien ordenado, 53
  cerrado en $\mathbb{R}$, 120
  cilindro, 768
  clausura, 121
  co-finito, 778
  convexo, 193
  de Bernstein, 78, 323
  de contenido cero, 253
  de etiquetas, 410
  de los números complejos, $\mathbb{C}$, 83
  de los números enteros, $\mathbb{Z}$, 83
  de los números irracionales, $\mathbb{I}$, 83
  de los números naturales, $\mathbb{N}$, 83
  de los números racionales, $\mathbb{Q}$, 83
  de los números reales, $\mathbb{R}$, 83
  de Lusin, 142
  de medida cero según Jordan, 253
  de medida fuertemente cero, 305
  de nivel, 369
  de primera categoría, 137
  de segunda categoría, 137
  de Sierpiński, 324
  de Vitali, 312
  denso, 125
  denso en alguna parte, 136
  diámetro de un, 128
  dirigido, 9
  elemental, 241
  finito, 11
  fuertemente de medida cero, 367
  gordo de Cantor, 230
  integrable, 917
  interior, 122
  linealmente dependiente, 51
  linealmente independiente, 51
  localmente compacto, 158
  medible, 256, 680, 782
  medible Lebesgue, 738
  negativo, 786
  no-medible, 257
  no-negativo, 88
  no-numerable, 11
  norma-acotado, 203
  nulo, 247, 786
  numerable, 11
  nunca-denso, 136
  parcialmente ordenado, 50
  perfecto, 130
  positivo, 786
  potencia, 4
  que genera a otro, 51
  raro, 136
  saturado no-medible, 320
  simétrico, 89

supremo de un, 86
tenso, 606
ternario de Cantor, 216
tipo-Cantor de medida cero, 227
tipo-Cantor de medida positiva, 230
totalmente disconexo, 123
totalmente ordenado, 50
uniformemente absolutamente continuo, 594
uniformemente integrable, 593
uniformemente Vitali-pequeño, 606
universal para una clase $\mathcal{C}$, 277
vacío, 4

conjuntos
abiertos, 119
biyectables, 10
casi-disjuntos, 82
de Borel, 266
disjuntos o ajenos, 4
equipotentes, 10
estrictamente separados, 250
orden-isomorfos, 56
ordenadamente separados, 87

constante de Lipschitz, 454

continuidad
de la integral, 569
de una transformación lineal, 202

contradominio de una función, 7

convergencia
$\mu -$ c.s., 376
casi-uniforme, 387
débil en $L_p(X, \mu)$, 637
de una sucesión, 91
en la p-norma, 635
en la norma, 203
en la norma de $L_1(\mu)$, 580
en medida, 394
en probabilidad, 394
fuerte en L_p, 635
puntual, 176
puntual de medidas, 833
uniforme, 176

Convergencia Acotada, 560

convolución de dos funciones, 660

cota
inferior de un conjunto, 85
superior, 50
superior de un conjunto, 85

Criterio
de Carathéodory, 295
de Cauchy para familias sumables, 116
de Cauchy para Series, 108
de la Vallée-Poussin, 595
Uniforme de Cauchy, 178

cubo diádico, 736

cubrimiento
abierto, 131
de un conjunto, 6
de Vitali, 472

DD-primitiva, 915

densidad
de las funciones simples, 383
en $L_p(X, \mu)$, 645

derivada
aproximada, 909
de Radon-Nikodým, 805
distribucional, 914
inferior, 484, 485
superior, 484, 485

derivadas de Dini, 485

descomposición
de Hahn, 788, 789
de Jordan, 789

Desigualdad
AM-GM, 196
de Cauchy-Schwarz, 200, 633
de Chebyshev, 538, 551, 714
de Hölder, 201
de Hölder con $p = 1$ y $q = \infty$, 655
de Hölder en L_p, 631
de Hardy-Littlewood, 826
de Jensen, 600
de Jensen finita, 195
de la Media Generalizada, 198
de Minkowski, 200
de Minkowski en L_p, 633
de Sard, 759
de Young, 199
de Young para convoluciones, 662
Triangular, 144

determinante Jacobiano, 755

diámetro de un conjunto, 128

difeomorfismo, 759

diferencia

de dos conjuntos, 5
simétrica de conjuntos, 5
diferenciación bajo el signo integral, 570
diferencial de una función, 755
discontinuidad
de la primera especie, 171
de la segunda especie, 171
de salto, 171
removible, 171
distancia
de un punto a un conjunto, 128
entre dos conjuntos, 128, 250
dominio de una función, 7
dual
de un espacio de Banach, 145

el continuo, 73
El Problema
de la Medida de Hausdorff, 350
de la Medida de Lebesgue, 241
de las Primitivas, 408
elemento
mínimo, 53
maximal, 50
entorno abierto en $\mathbb{R}$, 120
equivalencia de las p-normas en $\mathbb{R}^n$, 201
escalera del Diablo, 234
espacio
$L_1(X, \mu)$, 576
$L_\infty(X, \mu)$, 648
cociente, 9
compacto, 131
conexo, 123
de Baire, 150
de Banach, 203
de Cantor, 771
de Hausdorff, 119
de las distribuciones, 913
de medida, 680
de medida completa, 690
de medida finita, 681
de probabilidad, 681, 816
de Stone, 571
métrico, 126
métrico completo, 127
métrico separable, 127
medible, 680
medible producto, 768
normado, 144
Polaco, 127
reflexivo, 816
topológico, 119
totalmente disconexo, 151
espacios
ℓ_p^n, 201
$L_p(\mu)$, $1 < p < +\infty$, 630
homeomorfos, 163
isométricos, 163
estrategia, 353
ganadora, 353
etiquetas
de una partición, 850
exponentes conjugados, 631
extensión
continua de una medida, 345
de una medida, 345
universal, 350
universal de una medida, 345

Fórmula de Cavalieri, 567
fórmula de Newton-Leibniz, 408
familia
absolutamente continua, 840
absolutamente continua uniformemente, 840
de conjuntos, 3, 9
de conjuntos disjuntos dos a dos, 6
de medidas puntualmente acotada, 845
de medidas uniformemente acotada, 845
disjunta de conjuntos, 6
equi-continua, 191, 918
puntualmente acotada, 191
sumable, 115
uniformemente acotada, 191, 440
uniformemente equi-continua, 918
filtración, 819
filtro, 326
co-finito, 326
de Fréchet, 326
generado, 327
maximal, 327
principal, 326
forma canónica de funciones simples, 382
fracciones de Farey, 30
función, 7

M-Lipschitz, 156
$\aleph_0$-simple, 398
λ-integrable, 714
$\mathbb{Q}$-lineal, 337
p-integrable, 630
C-integrable, 904
absolutamente continua, 498
absolutamente continua generalizada, 908
absolutamente continua sobre $\mathbb{R}$, 709
absolutamente integrable, 881
acotada, 85
acotada casi-siempre, 376
acotada por arriba, 187
acotada por debajo, 187
aproximadamente continua, 483
biyectiva, 7
calibrador, 419
característica, 8
casi-continua, 391
cociente, 9
compuesta, 8
continua, 155
continua casi-siempre, 376
continua con soporte compacto, 159
continua por la derecha, 174
controladora, 419
convexa, 193
Daniell integrable, 574
DD-integrable, 915
de Cantor, 234, 354
de Cantor-Lebesgue, 234
de Cauchy, 337
de clase C^1, 758
de Darboux, 163
de Dirichlet, 8
de distribución, 701
de elección, 10
de la primera clase de Baire, 143
de prueba, 913
de Takagi, 182
de Thomae, 165
de variación acotada, 462
de variación acotada sobre $\mathbb{R}$, 709
derivada, 408, 430
diferenciable (según Fréchet), 755
distribución, 913
distribución de Schwartz, 913
distribucionalmente Denjoy integrable, 915
en escalera, 172, 383, 417
escalonada, 172
esencialmente acotada, 648
extensión de una, 8
finita casi-siempre, 376
Fréchet diferenciable, 755
gauge, 419
identidad, 8
inclusión, 8
indicatriz de Banach, 672
integrable en el sentido débil, 824
inversa, 8
inyectiva, 7
Lebesgue integrable, 530, 547, 557
Lipschitz, 454
localmente de variación acotada, 709
localmente integrable, 821
maximal de Hardy-Littlewood, 823
mayor, 910
McShane integrable, 898
medible, 371
medible Borel, 381
medible Lebesgue, 371
medible según Borel, 273
menor, 910
nula casi-siempre, 376
nunca diferenciable, 182
Pfeffer integrable, 904
proyección, 8
que se anula en el infinito, 160
regulada, 175
restricción, 8
Riemann integrable, 411, 413
semicontinua inferiormente, 183
semicontinua superiormente, 183
siempre-sobreyectiva, 340
signo, 385
simple, 382
singular, 458
sobreyectiva, 7
test, 913
traslación, 656
truncada, 549
uniformemente continua, 156
uniformemente diferenciable, 757
valor absoluto, 8

funcional lineal, 145

gráfico de una función, 7
grupo lineal, 747

Hipótesis del Continuo, 77
  Débil, 77
  Generalizada, 77
homeomorfismo, 163

igualdad
  casi-siempre, 376
  de conjuntos, 4
  de funciones, 7
integración
  por sustitución, 861
integración por parte, 624
integral
  de Cauchy, 409
  de Daniell, 574
  de Darboux, 413
  de Denjoy, 909
  de Denjoy-Khintchine, 909
  de Gauss, 763
  de Henstock-Kurzweil, 852
  de Lebesgue, 530, 556
  de Lebesgue-Stieltjes, 707
  de McShane, 898
  de Mikusiński, 575
  de Newton, 408
  de Perron, 911
  de Pfeffer, 904
  de Riemann, 411
  distribucional de Denjoy, 915
  elemental de Daniell, 572
  impropia de Riemann, 566
  inferior de Darboux, 413
  inferior de Lebesgue, 530
  superior de Darboux, 413
  superior de Lebesgue, 530
intercambio entre el límite y la integral, 540
intersección de una familia de conjuntos, 6
intervalo
  degenerado, 85
  no-degenerado, 85
intervalos
  adyacentes, 124
  casi-disjuntos, 124
  contiguos, 124
  degenerados, 240
  diádicos, 124
  en $\mathbb{R}$, 85
  no-degenerados, 240
  no-superpuestos, 124

jerarquía de Borel, 271
juego
  $G(X, A)$, 353
  determinado, 353

límite
  de una sucesión, 91, 106
  inferior de una función, 185
  inferior de una sucesión, 99, 105
  superior de una función, 186
  superior de una sucesión, 99, 105
la recta real, 73
Lema
  de Borel-Cantelli (primer), 288
  de Borel-Cantelli (segundo), 289
  de Cousin, 419
  de Fatou, 542, 554, 713
  de Fatou en $\mathcal{HK}([a,b])$, 895
  de Riemann-Lebesgue, 590
  de Saks-Henstock, 869
  de Urysohn, 159
  de Zorn, 51
  del Sol Naciente, 497
Ley de Tricotomía
  de los Cardinales, 69
Leyes de Morgan, 6
Littlewood,
  Primer Principio de, 298
longitud
  de un intervalo, 85, 240

máximo de un conjunto, 86
métrica sobre un conjunto, 126
mínimo de un conjunto, 87
marca de una partición, 850
martingala, 819
matriz Jacobiana, 755
mediana de dos fracciones, 31
medida, 680
  κ-aditiva, 352
  σ-aditiva, 241

σ-finita, 282, 681
acotada, 794
atómica, 686
completa, 258
con signo, 783
con signo finita, 793
de Borel, 681
de Borel en $\mathbb{R}$, 280
de Borel en $\mathbb{R}^n$, 740
de conteo, 681
de control, 843
de Dirac, 682, 708
de Lebesgue, 280
de Lebesgue en $\mathbb{R}^n$, 739
de Lebesgue-Stieltjes, 707
de probabilidad, 681
de Radon, 745
de una transformación, 749
descomposable, 809
difusa, 686
exterior, 692
exterior de Lebesgue, 240, 242
exterior en $\mathbb{R}^n$, 738
finitamente aditiva, 282, 680
finitamente subaditiva, 247
imagen, 682
interior de Lebesgue, 300
invariante por traslación, 241, 245
no-atómica, 686
numerablemente aditiva, 242, 680
numerablemente subaditiva, 246
positiva, 783
producto, 727
puramente irregular, 780
real, 793
sobre una σ-álgebra, 280
super σ-finita, 809
medidas
absolutamente continuas, 582
mutuamente singulares, 789
uniformemente λ-continuas, 840
uniformemente σ-aditivas, 837
uniformemente μ-continuas, 594
uniformemente exhaustivas, 838
Monstruo de Bernstein, 78

número
algrbraico, 39
cardinal, 67
de oro, 39, 229
ordinal, 58
trascendente, 40
números de Fibonacci, 38
Números Reales Extendidos, 95
norma
de la variación, 794
de una partición, 410
dual, 145
euclídea, 200
uniforme, 158
normas equivalentes, 145

orden
canónico, 62
del diccionario, 53
lexicográfico, 53
parcial, 50
total o lineal, 50
ordinal
inicial, 67
límite, 64
numerable, 64
predecesor, 64
sucesor, 64
supremo, 64
oscilación de una función, 169

Paradoja
de Banach-Tarski, 351
de Russell, 4
parte
entera de un número, 83
negativa de f, 8
positiva de f, 8
partición
δ-fina, 410, 419
de McShane, 898
de un intervalo, 410
etiqueta de una, 850
etiquetada, 410, 850
libremente etiquetada, 897
más fina, 412
marca de una, 850
medible, 382
subordinada a δ, 419

subordinada a δ, 410
partición de $\mathbb{N}$, 17
partición de un conjunto, 6
paso al límite, 540
pre-medida, 696
Primer
Lema de Borel-Cantelli, 288
Principio de Littlewood, 298
primer
elemento de un conjunto, 53
ordinal no-numerable, 68
primer ordinal no-numerable, 65
primitiva, 409
distribucional, 915
primitiva de una función, 430
Principio
de Arquímedes, 84
de Cavalieri, 726
de Inducción Transfinita, 56
del Buen-Orden, 12, 54
del Dirichlet, 135
del Palomar, 135
del Palomar Infinito, 16
Maximal de Hausdorff, 51
producto
cartesiano de dos conjuntos, 6
interno sobre $\mathbb{R}^n$, 199
producto algebraico, 88
propiedad
casi-siempre, 375
de intersección finita, 134
de monotonía, 716
monótona de μ^*, 243
Propiedad del Valor Intermedio, 163
proyección sobre un producto, 10, 148
punto
aislado, 126
de acumulación, 126
de clausura, 121
de condensación, 129
de constancia, 356
de densidad de Lebesgue, 480
de dispersión de Lebesgue, 481
de Lebesgue de f, 628
de Lebesgue de un conjunto, 360
de sombra, 497
límite, 126
racionales diádicos, 125
radio de oro, 229
rango
de una medida, 689
rectángulo
en $\mathbb{R}^n$, 735
medible, 719
recta extendida, 95
red de conjuntos, 9
regularidad de la medida de Lebesgue, 301
relación binaria, 6
relación de equivalencia, 9
reordenamiento de una serie, 110
representación
binaria, 213
binaria infinita, 214
canónica de funciones simples, 382
ternaria, 210
retículo vectorial, 159
Riemann integrable es Lebesgue integrable, 539
rotación en $\mathbb{R}^n$, 750

Segundo Principio de Littlewood, 386
semi-norma
de Alexiewicz, 896
serie
absolutamente convergente, 109, 203
convergente, 107, 203
de funciones, 179
incondicionalmente convergente, 110
sistema de Dynkin, 698
soporte
de una función, 158
de una función medible, 660
sub-σ-álgebra, 816
subconjunto
de un conjunto, 3
propio, 4
subcubrimiento finito, 131
subespacio, 120
subpartición
etiquetada, 868
subsucesión de una sucesión, 93
sucesión
adaptada a σ-álgebras, 819
convergente, 127
creciente de conjuntos, 9

de Cauchy, 94, 127
de conjuntos, 9
de Farey, 30
decreciente de conjuntos, 9
divergente, 91
en un conjunto, 9
equi-integrable según Kurzweil, 890
estrictamente creciente, 9
estrictamente decreciente, 9
monótona creciente, 92
monótona decreciente, 92
puntualmente acotada, 404
transfinita, 62
transfinita no creciente, 62
transfinita no decreciente, 62
uniformemente HK-integrable, 890
uniformemente absolutamente continuo, 594
uniformemente integrable, 594

suma
de Riemann, 411
de una serie, 179
inferior de Darboux, 412
inferior de Lebesgue, 529
superior de Darboux, 412
superior de Lebesgue, 529

suma algebraica, 88
super-descomposición, 809
supremo
de un conjunto, 50, 86
esencial de una función, 649

Teoema
de Saks, 687

Teorema
de Acotación Uniforme, 145
de Acotamiento de Nikodým, 845
de Antosik-Mikusińki, 835
de Aproximación de Weierstrass, 191
de Aproximación en $L_1(\mathbb{R}^n)$, 746
de Arandelović, 290
de Baire, 142
de Banach-Kuratowski, 347
de Banach-Vitali, 672
de Banach-Zarecki, 517, 611
de Bartle-Dunford-Schwartz, 843
de Beppo Levi, 544, 553, 712
de Beppo-Levi en $\mathcal{HK}([a,b])$, 895
de Bernstein, 78
de Bolzano-Weierstrass, 94
de Buczolich, 887
de Cantor, 41, 42
de Cantor-Bendixson, 130
de Cantor-Bernstein-Schröeder, 43
de Cantor-Lebesgue, 591
de Categoría de Baire, 137
de Convergencia de Nikodým, 844
de Convergencia de Vitali, 602, 607
de Convergencia de Vitali en L_p, 644
de Convexidad de Lyapunov, 309
de Darboux, 164
de Densidad de Lebesgue, 480
de Descomposición de Hahn, 788
de Descomposición de Lebesgue, 806
de Diferenciabilidad de Fubini, 492
de Diferenciación de Lebesgue, 629, 822, 827
de Dini, 389
de Drewnowski, 796
de Dunford-Pettis, 601
de Dynkin π-λ, 699
de Encaje de Cantor, 128
de Erdös-Sierpiński, 326
de Extensión Continua, 161
de Extensión de Carathéodory, 696
de Extensión de Tietze, 391
de Fichtenholz-Sierpińki, 687
de Fubini-Tonelli, 728
de Gordon, 890
de Hake, 872
de Halmos, 717
de Heine, 157
de Jordan, 467
de Kronecker, 135
de Lévy, 334
de la Aplicación Abierta, 758
de la Clase Monótona, 717
de la Convergencia Acotada, 441, 541
de la Convergencia Dominada, 544, 714
de la Convergencia Dominada de Lebesgue, 561
de la Convergencia Dominada en $\mathcal{HK}([a,b])$, 892
de la Convergencia Dominada Generalizada, 568
de la Convergencia Monótona, 438, 542, 559,

712
de la Convergencia Monótona de Lebesgue, 548
de la Convergencia Monótona en $\mathcal{HK}([a,b])$, 894
de la Convergencia Uniforme, 540
de la Convergencia Uniforme de Riemann, 438
de la Convergencia Uniforme en $\mathcal{HK}([a,b])$, 888
de la equi-integrabilidad, 892
de la Existencia de la Esperanza Condicional, 817
de la Igualdad Riemann-Darboux, 416
de la Integrabilidad Absoluta, 877
de la Regla de la Cadena, 624
de la Serie Geométrica, 210
de la Vallée-Poussin, 910
de Lebesgue, 395
de Lebesgue-Stieltjes, 705
de Lindelöf, 131
de Lusin, 141, 390
de McShane-Lebesgue, 901
de Rademacher, 315, 455, 491
de Radon-Nikodým - Caso σ-finito, 806
de Radon-Nikodým - Caso finito, 804
de Radon-Nikodým - Caso general, 808
de Representación de Riesz, 812
de Riemann-Darboux, 415
de Riesz, 396, 584
de Riesz-Fischer, 634
de Sard, 765
de Severini-Egoroff, 386, 711
de Steinhaus, 292
de Tarski, 329
de Tychonoff, 149
de Ulam, 349
de Vitali-Carathéodory, 619
de Vitali-Hahn-Saks, 847
de Vitali-Lebesgue, 420
de Vitali-Scheffé, 605
de Volterra, 444
de Weierstrass, 133
de Weierstrass-Riemann, 110
de Wiener, 824
del Cambio de Variable lineal, 752
del Cambio de Variable no-lineal, 762
del Ultrafiltro, 329
del Valor Intermedio, 163
del Valor Medio en $\mathbb{R}^n$, 756
del Valor Medio para Derivadas, 429
del Valor Medio para Integrales, 428
Diferenciabilidad de Lebesgue, 488
Fundamental del Cálculo, 431, 461, 612
Fundamental del Cálculo 2 para HK, 875
Fundamental del Cálculo de Lebesgue, 545, 615
Fundamental del Cálculo en $\mathcal{HK}([a,b])$, 860, 862
que preserva volumen, 763
Tercer Principio de Littlewood, 386
topología, 119
débil, 601
densidad, 482
discreta, 120
generada por una colección, 147
inducida, 120
inicial, 147
producto, 146
trivial, 120
uniforme, 158
transformación
dilatación, 752
lineal, 202
lineal invertible, 747
lineal(continuidad), 202
ortogonal, 750
traslación invariante de $L_1(\mu)$, 656
triángulo de Pascal, 38
tricotomía de ordinales, 62

ultrafiltro, 327
existencia, 329
libre, 327
no-medible, 326
no-principal, 327
unión de una familia de conjuntos, 6
unión disjunta, 6
unicidad
de μ_n, 743
de la medida de Borel, 742

variable aleatoria, 816
variación
negativa, 789

positiva, 789
total, 789
volumen de cubo en $\mathbb{R}^n$, 735

Weierstrass K., 180