Introducción a la Teoría de la Medida

Daniel Azagra Rueda

DEPARTAMENTO DE ANÁLISIS MATEMÁTICO Y MATEMÁTICA APLICADA

FACULTAD DE CIENCIAS MATEMÁTICAS

UNIVERSIDAD COMPLUTENSE DE MADRID

Diciembre de 2023

Revisado por última vez en diciembre de 2024.

Índice general

Introducción

Uno de los problemas básicos de la teoría de la medida es determinar sobre qué clase $\mathcal{A}$ de subconjuntos de $\mathbb{R}^n$ (estables por uniones finitas, complementación, traslaciones y rotaciones) puede definirse una función $\mu : \mathcal{A} \to [0, \infty]$ no trivial que sea finitamente aditiva e invariante por traslaciones y rotaciones, es decir, que si $A, A_1, ..., A_k \in \mathcal{A}$ son disjuntos entonces $\mu(\bigcup_{j=1}^k A_j) = \sum_{j=1}^k \mu(A_j)$, y $\mu(T(A)) = \mu(A)$ para toda T que sea traslación o rotación. El teorema de Banach-Tarski nos dice que $\mathcal{A}$ no puede coincidir con el conjunto de partes de $\mathbb{R}^n$ al menos cuando $n \geq 3$ (dados dos conjuntos cualesquiera acotados y con interior no vacío en $\mathbb{R}^3$, puede descomponerse uno de ellos en un número finito de piezas que, mediante traslaciones y rotaciones se ensamblan para recomponer el otro). Por otro lado, para que las operaciones de paso al límite que a los analistas nos gusta hacer puedan tener éxito en condiciones razonables, es necesario pedir también que la clase $\mathcal{A}$ sea estable por uniones e intersecciones numerables y que contenga al menos a todos los conjuntos abiertos.

Esto nos lleva al problema de averiguar cuál es la mayor clase $\mathcal{A}$ de subconjuntos de $\mathbb{R}^n$, que sea estable por uniones numerables, por complementación, por traslaciones y rotaciones, y que admita una función $\mu : \mathcal{A} \to [0, \infty]$ de manera que:

1. $\mu(\emptyset) = 0$;
2. $\mu\left(\bigcup_{j=1}^\infty A_j\right) = \sum_{j=1}^\infty \mu(A_j)$ siempre que $\{A_j\}_{j \in \mathbb{N}} \subset \mathcal{A}$ y $A_j \cap A_k = \emptyset$ para $j \neq k$;
3. $\mu(T(A)) = \mu(A)$ si $A \in \mathcal{A}$ y $T : \mathbb{R}^n \to \mathbb{R}^n$ es una isometría;
4. $\mu([0, 1]^n) = 1$.

En la primera parte de esta asignatura resolveremos esta cuestión fundacional: $\mathcal{A}$ será la clase de los conjuntos *medibles Lebesgue*, y μ la medida de Lebesgue en $\mathbb{R}^n$.

El concepto de medida está ligado al de dimensión: incluso dentro de un mismo espacio euclídeo $\mathbb{R}^n$ nos interesa poder medir longitudes, áreas (bidimensionales), volúmenes (tridimensionales), o en general volúmenes de objetos de dimensión k. Para ello debemos considerar diferentes tipos de medidas, porque cualquier medida que nos permita medir áreas de forma adecuada asignará medida cero a los objetos que intuitivamente pueden considerarse unidimensionales (por ejemplo las curvas diferenciables), y la medidas que nos ayudan a medir longitudes de curvas necesariamente asignarán medida infinita a los objetos bidimensionales. Lo mismo sucede con medidas o conjuntos de dimensiones k, n con $k \neq n$. Todo esto suena muy impreciso, y lo es: precisamente, uno de los objetivos de esta asignatura es dotar de precisión y rigor a estas ideas. Veremos cómo asignar medidas razonables a objetos de $\mathbb{R}^n$ que son de dimensión t (con $0 \leq t \leq n$), en algún sentido más o menos intuitivo cuando t es entero (y quizá muy poco intuitivo cuando t no es entero): aparecerán así las medidas de Hausdorff en $\mathbb{R}^n$.

En general, de una clase $\mathcal{A}$ de subconjuntos de un conjunto dado X que sea estable por uniones numerables, por complementación, y que contenga al propio X, se dice que es una σ-álgebra.[1] Y de una función μ con las dos primeras propiedades de las cuatro señaladas antes se dice que es una *medida*.

Vemos así que incluso limitándonos a los objetos de $\mathbb{R}^n$ es necesario ya considerar una infinidad de medidas diferentes (y clases diferentes de objetos medibles) con interés propio. Fuera de $\mathbb{R}^n$ hay también

[1] Si cambiamos estabilidad por uniones numerables por estabilidad por uniones finitas, se tiene lo que se llama un álgebra de conjuntos, pero este concepto no resulta tan útil en Análisis.

espacios métricos más generales en los que pueden definirse medidas con diversas utilidades. Sin hablar de los espacios probabilísticos, que son conjuntos X (no necesariamente espacios métricos) con una medida que asigna el valor 1 al conjunto X.

Todo esto hace que resulte muy útil estudiar de manera unificada las propiedades generales de las σ-álgebras y las medidas abstractas, y cómo a partir de una medida abstracta puede definirse la integral de Lebesgue asociada a esa medida.

La medida y la integral son conceptos ligados muy estrechamente (la integral nos da la medida del conjunto bajo la gráfica de una función positiva, y por otro lado la integral de Lebesgue se sustenta en una medida en el conjunto donde está definida la función que se integra). Y sobre todo, para poder calcular medidas de forma eficiente, necesitamos la maquinaria de la integral.

Así como no todo conjunto puede ser medido en general, tampoco toda función puede ser integrada. Si $\mathcal{A}$ es una σ-álgebra en un conjunto X y μ es una medida en $\mathcal{A}$, diremos que una función $f : X \to \mathbb{R}$ es $\mathcal{A}$-medible si $f^{-1}(I) \in \mathcal{A}$ para todo I intervalo de $\mathbb{R}$. Las funciones medibles positivas $f : X \to [0, \infty]$ sí las podremos integrar: definiremos su integral de Lebesgue $\int_X f d\mu$ como el supremo de $\sum_{i=1}^n \alpha_i \mu(A_i)$, donde el supremo se toma sobre los $n \in \mathbb{N}$, $A_1, ..., A_n \in \mathcal{A}$ disjuntos, $\alpha_i \in [0, \infty)$, $i = 1, ..., n$, y tales que $\sum_{i=1}^n \alpha_i \chi_{A_i} \leq f$, siendo χ_{A_i} la función característica del conjunto A_i (o sea la función que vale 1 en A_i y 0 fuera de A_i). Esto generaliza la idea intuitiva de aproximar el *área* del conjunto $\{(x, t) : x \in X, 0 \leq t \leq f(x)\}$ por medidas de uniones finitas de conjuntos más o menos razonables, de tipo *rectangular en* $X \times \mathbb{R}$ (o sea los conjuntos $A_i \times [0, \alpha_i]$), que quedan por debajo de la gráfica de la función f.

La teoría básica de la medida y la integral de Lebesgue la desarrollaremos en los capítulos 1-8, y 10, aunque sin perder de vista los dos ejemplos motivadores y fundamentales de las medidas de Lebesgue y de Hausdorff en $\mathbb{R}^n$.

En los capítulos 9, 11, 12 y 13 investigaremos algunas de las propiedades más particulares y sutiles de las medidas de Lebesgue y de Hausdorff en $\mathbb{R}^n$, y veremos cómo las técnicas desarrolladas de forma general pueden utilizarse también para resolver cuestiones importantes del cálculo integral en $\mathbb{R}$ (por ejemplo, para caracterizar la clase de funciones para las que el teorema fundamental del cálculo es válido).

La teoría de la medida es una asignatura de gran interés intrínseco que además resulta fundamental para entender otras más avanzadas (por supuesto todas las de Análisis Matemático, pero no sólo estas: también es esencial en algunos problemas de Geometría Diferencial, en EDPs, y por supuesto en Probabilidad). Por eso te animo, querido lector, a estudiarla con el mayor cuidado, y espero que este manual te ayude en esa tarea.

Capítulo 1

Medidas y σ-álgebras

Como ya hemos apuntado en la introducción, en general no puede definirse una medida no trivial con propiedades adecuadas (invariancia por traslaciones y rotaciones) para todos los subconjuntos de $\mathbb{R}^n$, y lo mismo pasa en otras situaciones interesantes. Por este motivo es necesario estudiar las propiedades de las clases de conjuntos para las que sí pueden definirse medidas no triviales. Para ser útiles en Análisis, dichas clases deben ser estables por uniones numerables, intersecciones numerables y complementación. Esto da pie a la definición de σ-álgebra.

1.1. σ-álgebras.

Definición 1.1. Sea X un conjunto. Una colección $\mathcal{A}$ de subconjuntos de X se dice que es una σ-álgebra en X si $\mathcal{A}$ tiene las siguientes propiedades:

1. $X \in \mathcal{A}$;
2. $A \in \mathcal{A} \implies A^c := X \setminus A \in \mathcal{A}$;
3. Si $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{A}$ entonces $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

Del par $(X, \mathcal{A})$ se dice que es un *espacio medible*, y a los elementos de $\mathcal{A}$ se les llama conjuntos medibles (o, para ser más precisos, $\mathcal{A}$-medibles).

Ejemplos 1.2. Sea X un conjunto.

1. El conjunto 2^X de las partes de X es una σ-álgebra en X (esta es la mayor σ-algebra que puede existir en X).
2. $\{\emptyset, X\}$ es una σ-algebra en X (esta es la menor σ-algebra que puede existir en X).
3. Si $E \subset X$, entonces $\mathcal{A} = \{\emptyset, X, E, X \setminus E\}$ es una σ-álgebra en X (y $\mathcal{A}$ es la menor σ-algebra en X tal que $E \in \mathcal{A}$).

Las demostraciones de las dos proposiciones siguientes son muy fáciles y se dejan al cuidado del lector.

Proposición 1.1. *Sea $(X, \mathcal{A})$ un espacio medible. Entonces:*

1. *$\emptyset \in \mathcal{A}$;*
2. *$A, B \in \mathcal{A} \implies A \setminus B \in \mathcal{A}$;*
3. *$\mathcal{A}$ es cerrado por uniones finitas o numerables, y por intersecciones finitas o numerables.*

Proposición 1.2. *Si $\{\mathcal{A}_i\}_{i\in I}$ es una familia de σ-álgebras en un conjunto X entonces $\mathcal{A} := \bigcap_{i\in I} \mathcal{A}_i$ es una σ-álgebra en X.*

Definición 1.3. Sea $\mathcal{R}$ una familia cualquiera de subconjuntos de un conjunto X. La proposición anterior muestra que la intersección de todas las σ-álgebras en X que contienen a $\mathcal{R}$ es una σ-álgebra. Dicha intersección es no vacía ya que hay al menos una σ-álgebra en X que contiene a $\mathcal{R}$ (a saber, el conjunto 2^X de las partes de X), y todas las σ-álgebras en X contienen al menos a $\{\emptyset, X\}$. A esta intersección la llamaremos *la σ-álgebra generada por* $\mathcal{R}$, y la denotaremos $\sigma(\mathcal{R})$.

Así, $\sigma(\mathcal{R})$ queda caracterizada por las siguientes dos propiedades:

1. $\sigma(\mathcal{R})$ es una σ-álgebra en X que contiene a $\mathcal{R}$.

2. $\sigma(\mathcal{R})$ es la menor σ-álgebra que contiene a $\mathcal{R}$, en el sentido de que si $\mathcal{A}$ es una σ-álgebra en X que contiene a $\mathcal{R}$ entonces $\sigma(\mathcal{R}) \subset \mathcal{A}$.

Ejemplo 1.1. Si $\mathcal{R} = \{E\}$, donde $E \subset X$, entonces $\sigma(\mathcal{R}) = \{\emptyset, X, E, X \setminus E\}$.

Proposición 1.3. *Un conjunto $\mathcal{A} \subset 2^X$ es una σ-álgebra en X si y sólo si satisface las siguientes tres propiedades:*

1. *$X \in \mathcal{A}$;*

2. *Si $A, B \in \mathcal{A}$ entonces $A \setminus B \in \mathcal{A}$;*

3. *Si $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{A}$ y $A_i \cap A_j = \emptyset$ para $i \neq j$, entonces $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.*

Demostración. La necesidad de estas tres propiedades es obvia. Veamos que son suficientes para que $\mathcal{A}$ sea una σ-álgebra en X. Lo único que hay que demostrar es que la condición (3) se cumple incluso si los conjuntos A_i no son disjuntos dos a dos. Sea pues una familia numerable cualquiera $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{A}$, y definamos $B_1 = A_1$, $B_2 = A_2 \setminus B_1$, y en general, por inducción, supuesto definido B_n, pongamos

$$B_{n+1} := A_{n+1} \setminus (B_1 \cup ... \cup B_n) = A_{n+1} \setminus (A_1 \cup ... \cup A_n).$$

Es obvio que $B_1 \in \mathcal{A}$, y $B_2 \in \mathcal{A}$ por la propiedad (2). Obsérvese también que los conjuntos B_n son disjuntos dos a dos. Suponiendo que $B_1, ..., B_n \in \mathcal{A}$, es entonces también claro, usando (2) y (3), que $B_{n+1} \in \mathcal{A}$. Por tamto, por inducción, $\{B_n\}_{n\in\mathbb{N}} \subset \mathcal{A}$. Entonces, usando (3) obtenemos que $\bigcup_{n=1}^{\infty} B_n \in \mathcal{A}$. Y así

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n \in \mathcal{A},$$

lo que completa la prueba. $\square$

1.2. Conjuntos borelianos

Sea X un espacio métrico.[1] Denotaremos por $\mathcal{B}(X)$ la σ-álgebra en X generada por la familia de todos los conjuntos abiertos de X. A los elementos de $\mathcal{B}(X)$ los llamaremos *conjuntos borelianos* (o de Borel), y de $\mathcal{B}(X)$ diremos que es la σ-álgebra de los borelianos en X.

Se sigue inmediatamente de las propiedades de las σ-álgebras que:

1. Todos los conjuntos cerrados son borelianos.

2. $\mathcal{B}(X)$ es también la σ-álgebra generada por los subconjuntos cerrados de X.

3. Si $\{U_n\}_{n\in\mathbb{N}}$ es una sucesión de abiertos de X, entonces $\bigcap_{n=1}^{\infty} U_n$ es boreliano. De este tipo de borelianos se dice que son *conjuntos G_δ*.

4. Si $\{C_n\}_{n\in\mathbb{N}}$ es una sucesión de cerrados de X, entonces $\bigcup_{n=1}^{\infty} C_n$ es boreliano. De este tipo de borelianos se dice que son *conjuntos F_σ*.

En $\mathbb{R}^d$ existen conjuntos borelianos que no son ni F_σ ni G_δ. Véase la sección de problemas.

[1] X también puede ser un espacio topológico, pero los espacios métricos separables bastarán para todos los propósitos de esta asignatura.

1.3. Medidas

Definición 1.4. Sea $(X, \mathcal{A})$ un espacio de medida. Una *medida* en $(X, \mathcal{A})$ es una función $\mu : \mathcal{A} \to [0, \infty]$ tal que:

1. $\mu(\emptyset) = 0$;

2. μ es numerablemente aditiva, esto es, si $\{A_j\}_{j\in\mathbb{N}} \subset \mathcal{A}$ son tales que $A_j \cap A_k = \emptyset$ para $k \neq j$, entonces
$$\mu\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mu(A_j). \tag{1.1}$$

Si μ es una medida en $(X, \mathcal{A})$, se dice que $(X, \mathcal{A}, \mu)$ es un *espacio de medida*.

Si $\mu(X) < \infty$, diremos que μ es una medida finita.
Cuando $\mu(X) = 1$, también se dice que μ es una medida de probabilidad.
Cuando existen $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{A}$ tales que $X = \bigcup_{n=1}^{\infty} A_n$ y $\mu(A_n) < \infty$ para todo n, se dice que μ es una medida σ-finita.

Ejemplo 1.2. Sea X un conjunto cualquiera, y definamos $\mu : 2^X \to [0, \infty]$ como $\mu(E)$ igual al cardinal de E si E es finito, y $\mu(E) = \infty$ en otro caso. Entonces μ es una medida, a la que se llama *medida de contar*. En esta asignatura denotaremos esta medida por $\sharp$. Es fácil ver que la medida $\sharp$ es finita si y sólo si X es finito, y también que $\sharp$ es σ-finita si y sólo si X es a lo sumo numerable. En particular, cuando $X = \mathbb{R}^d$, la medida $\sharp$ no es σ-finita.

Ejemplo 1.3. Sean X un conjunto, y $a \in X$. Definamos $\delta_a : 2^X \to [0, \infty]$ por
$$\delta_a(E) = \begin{cases} 1 & \text{si } a \in E \\ 0 & \text{si } a \notin E. \end{cases}$$

Es fácil ver que δ_a es una medida, a la que se llama medida de Dirac en a. Esta medida es finita cualquiera que sea X.

Las medidas de los dos ejemplos anteriores, a pesar de su simplicidad casi trivial, tienen cierta utilidad general. Para construir medidas menos triviales y que sean útiles en $\mathbb{R}^n$ (por ejemplo para obtener una medida que nos permita medir las áreas de los conjuntos abiertos y cerrados del plano), necesitaremos algunas herramientas más. Dejamos aparcada esta cuestión por un momento, y en lo que queda de capítulo nos centraremos en estudiar las propiedades más elementales comunes a todas las medidas.

Teorema 1.4. *[Propiedades elementales de las medidas]. Sea $(X, \mathcal{A}, \mu)$ un espacio de medida. Entonces:*

(a) Para toda colección finita y disjunta $A_1, ..., A_n \in \mathcal{A}$ se tiene $\mu(A_1 \cup ... \cup A_n) = \mu(A_1) + ... + \mu(A_n)$.

(b) Si $A, B \in \mathcal{A}$, $A \subset B$ y $\mu(B) < \infty$, entonces $\mu(B \setminus A) = \mu(B) - \mu(A)$.

(c) Si $A, B \in \mathcal{A}$ y $A \subset B$ entonces $\mu(A) \leq \mu(B)$.

(d) Si $\{A_j\}_{j\in\mathbb{N}} \subset \mathcal{A}$ entonces
$$\mu\left(\bigcup_{j=1}^{\infty} A_j\right) \leq \sum_{j=1}^{\infty} \mu(A_j).$$

(e) Si $\{A_j\}_{j\in\mathbb{N}} \subset \mathcal{A}$ y $\mu(A_j) = 0$ para todo j, entonces $\mu\left(\bigcup_{j=1}^{\infty} A_j\right) = 0$.

(f) Si $\{A_j\}_{j\in\mathbb{N}} \subset \mathcal{A}$ y $A_j \subset A_{j+1}$ para todo j, entonces

$$\mu\left(\bigcup_{j=1}^{\infty} A_j\right) = \lim_{j\to\infty} \mu(A_j).$$

(g) Si $\{A_j\}_{j\in\mathbb{N}} \subset \mathcal{A}$, $A_{j+1} \subset A_j$ para todo j, y $\mu(A_{j_0}) < \infty$ para algún j_0, entonces

$$\mu\left(\bigcap_{j=1}^{\infty} A_j\right) = \lim_{j\to\infty} \mu(A_j).$$

Demostración. (a): es obvio, puesto que en (1.1) puede tomarse $A_j = \emptyset$ para $j > n$.

(b) y (c): si $A \subset B$ se tiene $B = A \cup (B \setminus A)$, unión disjunta, luego

$$\mu(B) = \mu(A) + \mu(B \setminus A) \geq \max\{\mu(A), \mu(B \setminus A)\},$$

lo que prueba (c), y si además $\mu(B) < \infty$ se deduce que también $\mu(A)$ y $\mu(B \setminus A)$ son finitos y que $\mu(B \setminus A) = \mu(B) - \mu(A)$.

(d): Definamos $B_1 = A_1$ y, para $n \geq 2$, $B_n = A_n \setminus (A_1 \cup ... \cup A_{n-1})$, como en la demostración de la Proposición 1.3. Se tiene que

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n,$$

los conjuntos B_n son disjuntos dos a dos, y $B_n \subset A_n$ para todo n. Luego, usando (1.1) y (c) deducimos que

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} \mu(B_n) \leq \sum_{n=1}^{\infty} \mu(A_n).$$

(e): se deduce inmediatamente de (d).

(f): definiendo los conjuntos B_n como en la demostración de (d), puesto que $A_j \subset A_{j+1}$ ahora se tiene $B_n = A_n \setminus A_{n-1}$ para $n \geq 2$. Como los conjuntos B_n son disjuntos dos a dos, y $(B_1 \cup ... \cup B_n) = A_n$, concluimos que

$$\mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} \mu(B_n) = \lim_{n\to\infty} \sum_{j=1}^{n} \mu(B_j) = \lim_{n\to\infty} \mu\left(B_1 \cup ... \cup B_n\right) = \lim_{n\to\infty} \mu(A_n).$$

(g): se deduce de (f) aplicándolo a los conjuntos $A_{j_0} \setminus A_j$ para $j \geq j_0$. En efecto, como $A_{j+1} \subset A_j$, entonces $A_{j_0} \setminus A_j \subset A_{j_0} \setminus A_{j+1}$ para todo $j \geq j_0$. Aplicando (f) obtenemos

$$\mu\left(A_{j_0} \setminus \bigcap_{j=j_0}^{\infty} A_j\right) = \mu\left(\bigcup_{j=j_0}^{\infty} (A_{j_0} \setminus A_j)\right) = \lim_{j\to\infty} \mu\left(A_{j_0} \setminus A_j\right),$$

y usando ahora (b), deducimos que

$$\mu(A_{j_0}) - \mu\left(\bigcap_{j=j_0}^{\infty} A_j\right) = \lim_{j\to\infty} \left(\mu(A_{j_0}) - \mu(A_j)\right) = \mu(A_{j_0}) - \lim_{j\to\infty} \mu(A_j),$$

de donde

$$\mu\left(\bigcap_{j=1}^{\infty} A_j\right) = \mu\left(\bigcap_{j=j_0}^{\infty} A_j\right) = \lim_{j\to\infty} \mu(A_j),$$

lo que termina la prueba. □

Observación 1.5. En la propiedad (g) del Teorema anterior, la condición de que alguno de los conjuntos tenga medida finita es esencial. Por ejemplo (g) no se cumple en $(\mathbb{N}, 2^{\mathbb{N}}, \sharp)$ con $A_n = \{k \in \mathbb{N} : k \geq n\}$, $n \in \mathbb{N}$.

1.4. Problemas

Problema 1.1. Dar un ejemplo que muestre que la unión de una familia de σ-álgebras no necesariamente es una σ-álgebra.

Problema 1.2. Demostrar que la σ-álgebra de los borelianos de $\mathbb{R}^n$ está generada por los compactos de $\mathbb{R}^n$

Problema 1.3. Demostrar que $\mathbb{Q}$ es un F_σ, pero no un G_δ. [Para la segunda parte, usar el teorema de categoría de Baire.]

Problema 1.4. Encontrar un subconjunto de $\mathbb{R}$ que no sea ni un F_σ ni un G_δ.

Problema 1.5. Sean X, Y conjuntos, y $f : X \to Y$ una función. Si $\mathcal{S}$ es una σ-álgebra en X, demostrar que la familia $\mathcal{S}_Y := \{B \subset Y : f^{-1}(B) \in \mathcal{S}\}$ es una σ-álgebra en Y.

Problema 1.6. Sean X e Y dos conjuntos, y $f : X \to Y$ una función. Si $\mathcal{S}$ es una σ-álgebra en Y, demostrar que:

1. $\Sigma_X := \{f^{-1}(B) : B \in \mathcal{S}\}$ es una σ-álgebra en X.
2. Si $\mathcal{S}$ está generada por un conjunto $\mathcal{E} \subset 2^Y$ entonces Σ_X es la σ-álgebra generada por $\mathcal{E}_X := \{f^{-1}(B) : B \in \mathcal{E}\}$.

Problema 1.7. Dar siete ejemplos de subconjuntos de $2^{\mathbb{R}}$ cada uno de los cuales genere la σ-algebra de los borelianos de $\mathbb{R}$.

Problema 1.8. Sean X un conjunto infinito, y $\mathcal{E} \subset 2^X$. Demostrar que para cada $A \in \sigma(\mathcal{E})$ existe $\mathcal{E}_0 \subset \mathcal{E}$ finito o numerable tal que $A \in \sigma(\mathcal{E}_0)$.
Indicación: Sea $\mathcal{A}$ la unión de las σ-álgebras $\sigma(\mathcal{C})$, donde $\mathcal{C}$ recorre las subfamilias contables de $\mathcal{E}$. Probar que $\mathcal{A}$ es una σ-álgebra tal que $\mathcal{E} \subseteq \mathcal{A} \subseteq \sigma(\mathcal{E})$ y por tanto $\mathcal{A} = \sigma(\mathcal{E})$.

Problema 1.9. Sea $X = \mathbb{R}^n$, y sea $\mathcal{E} = \{\{x\} : x \in \mathbb{R}^n\}$. Demostrar que la σ-álgebra generada por $\mathcal{E}$ está estrictamente contenida en $\mathcal{B}(\mathbb{R}^n)$.

Problema 1.10. Demostrar que, para todo conjunto no vacío X, la medida de contar $\mu : 2^X \to [0, \infty]$, definida por $\mu(E) = m$ si E es finito con m elementos, y $\mu(E) = \infty$ en otro caso, es en efecto una medida. Demostrar también el resto de afirmaciones hechas en el Ejemplo 1.2.

Problema 1.11. Si en el problema anterior ponemos $\mu(E) = 0$ cuando E es finito, ¿es μ una medida?

Problema 1.12. Sean X un conjunto no numerable, y $\mathcal{A}$ la colección de todos los subconjuntos $A \subset X$ tales que A o $X \setminus A$ es numerable, y definamos $\nu(A) = 0$ si A es numerable y $\nu(A) = 1$ en caso contrario. Comprobar que $(X, \mathcal{A}, \mu)$ es un espacio de medida.

Problema 1.13. Sea $\mathcal{A}$ una σ-álgebra de subconjuntos de X, y $(\mu_n)_{n\in\mathbb{N}}$ una sucesión de medidas en $(X, \mathcal{A})$ que es creciente (en el sentido de que $\mu_n(A) \leq \mu_{n+1}(A)$ para todos $n \in \mathbb{N}$, $A \in \mathcal{A}$). Demostrar que $\mu(A) := \lim_{n\to\infty} \mu_n(A)$ define una medida en $(X, \mathcal{A})$.

Problema 1.14. Dar un ejemplo de una sucesión de medidas (μ_n) tales que $\mu = \lim_n \mu_n$ no sea una medida.

Problema 1.15. Si $(\mu_n)_{n\in\mathbb{N}}$ es una sucesión arbitraria de medidas en $(X, \mathcal{A})$, probar que $\mu(A) := \sum_{n\in\mathbb{N}} \mu_n(A)$ define una medida en $(X, \mathcal{A})$.

Problema 1.16. Sea $(x_n)_{n\in N}$ una sucesión de $\mathbb{R}^d$, y definamos una medida en $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ por $\mu = \sum_{n\in\mathbb{N}} \delta_{x_n}$. ¿Para qué sucesiones (x_n) se tiene que $\mu(K) < \infty$ para todo compacto $K \subset \mathbb{R}^d$? ¿Y para qué sucesiones (x_n) es μ σ-finita?

Problema 1.17. Si μ es una medida en $(X, \mathcal{A})$ y fijamos $A \in \mathcal{A}$, probar que $\nu(B) := \mu(A \cap B)$ define una medida en $(X, \mathcal{A})$.

Problema 1.18. Sean $(X, \mathcal{A}, \mu)$ un espacio de medida, y $A_1, ..., A_n \in \mathcal{A}$ de medida finita. Si $\mathcal{I}_k^n = \{I \subset \{1, ..., n\} : \sharp(I) = k\}$, demostrar el principio de inclusión-exclusión:

$$\mu(A_1 \cup ... \cup A_n) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{I \in \mathcal{I}_k^n} \mu\left(\bigcap_{i \in I} A_i\right).$$

Como caso particular, deducir la fórmula de Poincaré para el cardinal de una unión de conjuntos finitos:

$$\sharp\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{I \in \mathcal{I}_k^n} \sharp\left(\bigcap_{i \in I} A_i\right).$$

Problema 1.19. Sea $(X, \mathcal{A}, \mu)$ un espacio de medida. Se dice que una sucesión $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A}$ es casi disjunta si $\mu(A_n \cap A_k) = 0$ siempre que $k \neq n$. Demostrar:

1. Si (A_n) es casi disjunta, $\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} \mu(A_n)$.
2. Si (A_n) verifica que $\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} \mu(A_n) < \infty$, entonces (A_n) es casi disjunta.
3. Si $\sum_{n \in \mathbb{N}} \mu(A_n) = \infty$ entonces la conclusión del apartado anterior puede ser falsa.

Problema 1.20 (Lema de Borel-Cantelli)**.** Si $\sum_{n \in \mathbb{N}} \mu(A_n) < \infty$, demostrar que

$$\mu\left(\{x : x \in A_k \text{ para infinitos valores de } k\}\right) = 0.$$

Indicación: considerar $\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$.

Capítulo 2

Medidas exteriores. Construcción de Carathéodory. Más propiedades de las medidas.

2.1. Medidas exteriores. Construcción de Carathéodory

En general construir medidas con propiedades deseables en $\mathbb{R}^n$ es difícil. El proceso se facilita si se construye primero lo que se llama una *medida exterior*, más sencilla de definir, y luego a partir de ella, por un proceso debido a Carathéodory (que se puede aplicar a todo tipo de medidas exteriores) se obtiene una medida. De esta forma se ahorra también trabajo si se quiere manejar varias medidas, porque el proceso de pasar de medida exterior a medida es el mismo siempre y puede establecerse una sóla vez de forma general. Utilizaremos esta estrategia en el siguiente capítulo para definir una infinidad de medidas muy útiles en $\mathbb{R}^n$: la medida de Lebesgue $\mathcal{L}_n$, y las medidas de Hausdorff $\mathcal{H}^t$ en $\mathbb{R}^n$, con $0 \leq t \leq n$.

Definición 2.1. Sea X un conjunto. Una aplicación $\mu^* : 2^X \to [0, \infty]$ se dice que es una *medida exterior* en X si:

1. $\mu^*(\emptyset) = 0$;
2. $A \subset B \implies \mu^*(A) \leq \mu^*(B)$;
3. Para toda familia numerable $\{A_j\}_{j\in\mathbb{N}} \subset 2^X$ se tiene
$$\mu^*\left(\bigcup_{j=1}^{\infty} A_j\right) \leq \sum_{j=1}^{\infty} \mu^*(A_j).$$

Nos referiremos a la segunda propiedad diciendo que la medida exterior es *monótona*, y a la tercera hablando de *subaditividad* numerable de la medida exterior. Cualquier medida definida en la σ-álgebra 2^X es obviamente una medida exterior en X. Como ejemplos fundamentales y no triviales de medidas exteriores, nos remitimos al siguiente capítulo, donde se definen las medidas exteriores de Lebesgue y de Hausdorff en $\mathbb{R}^n$.

Definición 2.2. Sea μ^* una medida exterior en X. Diremos que un subconjunto $E \subset X$ es μ^*-medible si
$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c) \quad \text{para todo} \quad A \subset X. \tag{2.1}$$

Aquí E^c denota $X \setminus E$, el conjunto complementario de E respecto de X.

A la condición (2.1) se le llama *condición de Carathéodory*. Obsérvese que, puesto que $(E^c)^c = E$, esta condición es simétrica, y resulta que E es medible si y sólo si E^c lo es. También es obvio que X y $\emptyset$ son siempre conjuntos μ^*-medibles. Por otra parte, puesto que por la propiedad (3) de la Definición 2.1, siempre es verdad que

$$\mu^*(A) \leq \mu^*(A \cap E) + \mu^*(A \cap E^c),$$

para comprobar que un conjunto E es medible basta demostrar que

$$\mu^*(A) \geq \mu^*(A \cap E) + \mu^*(A \cap E^c)$$

para todo $A \subset X$.

Proposición 2.1. *Si $\mu^*(E) = 0$ entonces E es μ^*-medible.*

Demostración. Si $\mu^*(E) = 0$ entonces también $\mu^*(A \cap E) = 0$, ya que $0 \leq \mu^*(A \cap E) \leq \mu^*(E)$. Por tanto

$$\mu^*(A) \geq \mu^*(A \cap E^c) = \mu^*(A \cap E^c) + \mu^*(A \cap E).$$

□

Teorema 2.2 (Carathéodory)**.** *Sea μ^* una medida exterior en un conjunto X, y definamos*

$$\mathcal{M}^* = \{E \subset X : E \text{ es } \mu^*\text{-medible}\}.$$

Entonces $\mathcal{M}^*$ es una σ-álgebra, y la restricción de μ^* a $\mathcal{M}^*$ es una medida.

Demostración. Dividiremos la prueba en varios pasos.

Paso 1. $X \in \mathcal{M}^*$. Esto es obvio, como ya hemos observado: $\emptyset$ y X siempre son medibles.

Paso 2. Si $E, F \in \mathcal{M}^*$, entonces $E \cup F \in \mathcal{M}^*$. En efecto, para cada $A \subset X$ tenemos

$$A \cap (E \cup F) = (A \cap E) \cup (A \cap F \cap E^c),$$

luego

$$\begin{aligned}
&\mu^*(A \cap (E \cup F)) + \mu^*(A \cap (E \cup F)^c) \\
&\leq \mu^*(A \cap E) + \mu^*(A \cap F \cap E^c) + \mu^*(A \cap E^c \cap F^c) \\
&= \mu^*(A \cap E) + \mu^*(A \cap E^c) = \mu^*(A),
\end{aligned}$$

donde en la primera desigualdad hemos usado la subaditividad de la medida exterior, en la penúltima igualdad hemos usado que F es medible, y en la última igualdad que E también lo es.

Paso 3. Si $E, F \in \mathcal{M}^*$ entonces $E \cap F, E \setminus F \in \mathcal{M}^*$. En efecto, sabemos que un conjunto es medible si y sólo si lo es su complementario. Puesto que ya sabemos que la unión de dos medibles es medible y se tiene

$$E \cap F = (E^c \cup F^c)^c,$$

se deduce de esa observación que $E \cap F$ es medible si E y F lo son. Además

$$E \setminus F = E \cap F^c$$

también es medible, volviendo a aplicar lo ya demostrado y observado.

Paso 4. Por inducción, se sigue de los pasos anteriores que $\mathcal{M}^*$ es cerrado por complementaciones, uniones e intersecciones finitas.

Paso 5. Si $\{E_j\}_{j \in \mathbb{N}} \subset \mathcal{M}^*$ son disjuntos dos a dos, entonces para cada $A \subset X$, se tiene que

$$\mu^* \left(A \cap \bigcup_{j=1}^{\infty} E_j \right) = \sum_{j=1}^{\infty} \mu^*(A \cap E_j).$$

En efecto: para una familia de a lo sumo dos elementos no vacíos esto es verdad porque, si $E, F \in \mathcal{M}^*$ son disjuntos entonces para cada $A \subset X$ se tiene (usando que E es medible)

$$\mu^*(A \cap (E \cup F)) = \mu^*(A \cap (E \cup F) \cap E) + \mu^*(A \cap (E \cup F) \cap E^c) = \mu^*(A \cap E) + \mu^*(A \cap F).$$

Luego, por inducción también se tiene

$$\mu^* \left(A \cap \bigcup_{j=1}^{n} E_j \right) = \sum_{j=1}^{n} \mu^*(A \cap E_j)$$

para todo $n \in \mathbb{N}$. Por tanto, usando la monotonía de μ^*,

$$\sum_{j=1}^{n} \mu^*(A \cap E_j) = \mu^* \left(A \cap \bigcup_{j=1}^{n} E_j \right) \leq \mu^* \left(A \cap \bigcup_{j=1}^{\infty} E_j \right),$$

de donde, haciendo $n \to \infty$ obtenemos

$$\sum_{j=1}^{\infty} \mu^*(A \cap E_j) \leq \mu^* \left(A \cap \bigcup_{j=1}^{\infty} E_j \right),$$

y como la otra desigualdad siempre se tiene por subaditividad, tenemos lo que queremos.

Paso 6. Si $\{E_j\}_{j \in \mathbb{N}} \subset \mathcal{M}^*$ son disjuntos dos a dos, entonces $\bigcup_{j=1}^{\infty} E_j \in \mathcal{M}^*$. En efecto, para todo $A \subset X$, puesto que $\bigcup_{j=1}^{n} E_j \in \mathcal{M}^*$ para cada $n \in \mathbb{N}$,

$$\mu^*(A) = \mu^* \left(A \cap \bigcup_{j=1}^{n} E_j \right) + \mu^* \left(A \setminus \bigcup_{j=1}^{n} E_j \right) \geq \sum_{j=1}^{n} \mu^* (A \cap E_j) + \mu^* \left(A \setminus \bigcup_{j=1}^{\infty} E_j \right),$$

y tomando límites

$$\mu^*(A) \geq \sum_{j=1}^{\infty} \mu^* (A \cap E_j) + \mu^* \left(A \setminus \bigcup_{j=1}^{\infty} E_j \right) = \mu^* \left(A \cap \bigcup_{j=1}^{\infty} E_j \right) + \mu^* \left(A \setminus \bigcup_{j=1}^{\infty} E_j \right),$$

donde en la igualdad hemos usado el Paso 5. Por tanto $\bigcup_{j=1}^{\infty} E_j$ es medible.

Paso 7. Ya sabemos que $\mathcal{M}^*$ es cerrado por uniones numerables disjuntas y por complementación, y que contiene a X. Por tanto, aplicando la Proposición 1.3 tenemos que $\mathcal{M}^*$ es una σ-álgebra. Además, por el Paso 5 con $A = X$, se tiene que la restricción de μ^* a $\mathcal{M}^*$ es una medida. □

Definición 2.3. Se dice que una medida $\mu : \mathcal{A} \to [0, \infty]$ es *completa* si todo subconjunto de un conjunto medible de medida cero es también medible. Por la Proposición 2.1 y la monotonía de μ^*, la medida descrita por el teorema de Carathéodory es completa.

Sea (X, d) un espacio métrico. Recordemos que, si E, F son subconjuntos no vacíos de X, se definen

$$d(x, E) = \inf_{y \in E} d(x, y),$$

$$d(E, F) = \inf_{x \in E, y \in F} d(x, y),$$

$$\operatorname{diam}(E) = \sup_{x, y \in E} d(x, y),$$

que la aplicación $X \ni x \mapsto d(x, E)$ es 1-Lipschitz (esto es, $|d(x, E) - d(y, E)| \leq d(x, y)$ para todo $x, y \in X$), y que se tiene $E = d(\cdot, E)^{-1}(\{0\})$ si y sólo si E es cerrado.

Definición 2.4. Una medida exterior μ^* en un espacio métrico X se dice que es una *medida exterior métrica* si

$$\mu^*(E \cup F) = \mu^*(E) + \mu^*(F) \quad \text{siempre que} \quad d(E, F) > 0.$$

Las medidas exteriores métricas tienen la importante propiedad de que los borelianos son conjuntos medibles para ellas, como vemos a continuación.

Teorema 2.3. *Si μ^* es una medida exterior métrica en X, entonces todos los subconjuntos borelianos de X son μ^*-medibles; es decir, con la notación del teorema de Carathéodory, $\mathcal{B}(X) \subset \mathcal{M}^*$.*

Demostración. Puesto que $\mathcal{M}^*$ es una σ-álgebra y $\mathcal{B}(X)$ es la menor σ-álgebra que contiene a los abiertos de X, basta probar que todos los abiertos son μ^*-medibles: dados G abierto de X y $A \subset X$, tenemos que ver que $\mu^*(A \cap G) + \mu^*(A \setminus G) \leq \mu^*(A)$. Podemos suponer que $\mu^*(A) < \infty$ ya que en otro caso la desigualdad es obvia. Para cada $n \in \mathbb{N}$, definamos

$$G_n = \left\{ x \in G : d(x, X \setminus G) > \frac{1}{n} \right\},$$

que es un abierto con la propiedad de que

$$d(G_n, X \setminus G) \geq \frac{1}{n} > 0.$$

Sea

$$D_n = G_{n+1} \setminus G_n = \left\{ x \in G : \frac{1}{n+1} < d(x, X \setminus G) \leq \frac{1}{n} \right\}.$$

Obviamente

$$G \setminus G_n = \bigcup_{j=n}^{\infty} D_j,$$

y

$$d(D_i, D_j) \geq \frac{1}{i+1} - \frac{1}{j} > 0 \quad \text{si} \quad i+2 \leq j \quad \text{y} \quad D_i \neq \emptyset \neq D_j$$

Puesto que las distancias entre dos cualesquiera de los conjuntos $D_1, D_3, ..., D_{2n-1}$ son estrictamente positivas cuando ambos son no vacios y μ^* es medida exterior métrica tenemos que

$$\mu^*(A \cap D_1) + \mu^*(A \cap D_3) + ... + \mu^*(A \cap D_{2n-1}) = \mu^* \left(A \cap (D_1 \cup D_3 \cup ...D_{2n-1})\right) \leq \mu^*(A),$$

ya que los D_{2k-1} que sean vacíos, si los hubiera, se pueden eliminar de la colección sin consecuencias para probar la desigualdad anterior. De forma análoga

$$\mu^*(A \cap D_2) + \mu^*(A \cap D_4) + ... + \mu^*(A \cap D_{2n}) \leq \mu^*(A).$$

Luego

$$\sum_{j=1}^{\infty} \mu^*(A \cap D_i) \leq 2\mu^*(A) < \infty.$$

Como $G \setminus G_n = \bigcup_{j=n}^{\infty} D_j$, se tiene (por la subaditividad de μ^*) que

$$\mu^*(A \cap (G \setminus G_n)) \leq \sum_{j=n}^{\infty} \mu^*(A \cap D_j),$$

de donde haciendo $n \to \infty$ se deduce que

$$\lim_{n \to \infty} \mu^*(A \cap (G \setminus G_n)) = 0.$$

Puesto que μ^* es medida exterior métrica y $d(G_n, X \setminus G) > 0$ tenemos

$$\mu^*(A \cap G_n) + \mu^*(A \setminus G) = \mu^*((A \cap G_n) \cup A \setminus G)) \leq \mu^*(A),$$

y así

$$\begin{aligned} \mu^*(A \cap G) + \mu^*(A \setminus G) &\leq \mu^*(A \cap (G \setminus G_n)) + \mu^*(A \cap G_n) + \mu^*(A \setminus G) \\ &\leq \mu^*(A \cap (G \setminus G_n)) + \mu^*(A), \end{aligned}$$

con lo que haciendo $n \to \infty$ concluimos

$$\mu^*(A \cap G) + \mu^*(A \setminus G) \leq \mu^*(A).$$

□

2.2. Más propiedades de las medidas

Usaremos medidas exteriores métricas para demostrar un resultado que nos será de gran utilidad más adelante: nos permite aproximar la medida de cualquier boreliano por las medidas de abiertos que lo contienen, y también por las de los cerrados que están contenidos en él.

Teorema 2.4. *Sean X un espacio métrico, y μ una medida de Borel en X, es decir, una medida $\mu : \mathcal{B}(X) \to [0, \infty]$. Supongamos que X es unión numerable de conjuntos abiertos de medida finita. Entonces*

$$\mu(E) = \text{ínf}\{\mu(U) : E \subset U, U \textit{ abierto}\} = \sup\{\mu(C) : C \subset E, C \textit{ cerrado}\} \tag{2.2}$$

para cualquier $E \in \mathcal{B}(X)$.

Demostración. Podemos suponer sin pérdida de generalidad que

$$X = \bigcup_{n=1}^{\infty} V_n,$$

donde los V_n son abiertos, $V_n \subset V_{n+1}$ y $\mu(V_n) < \infty$ para todo n. Definamos, para cada $E \subset X$,

$$\mu^*(E) = \inf_{U \supset E, U \text{ abierto}} \mu(U).$$

Es fácil ver (y queda como ejercicio para el lector) que μ^* es una medida exterior métrica. Por tanto, al restringir μ^* a los conjuntos borelianos se tiene una medida de Borel. Es claro que

$$\mu^*(U) = \mu(U) \quad \text{para todo } U \text{ abierto}, \tag{2.3}$$

y que

$$\mu(E) \leq \mu^*(E) \quad \text{para todo } E \in \mathcal{B}(X). \tag{2.4}$$

Por la desigualdad anterior tenemos, para cada $E \in \mathcal{B}(X)$,

$$\mu(V_n \setminus E) \leq \mu^*(V_n \setminus E), \quad \text{y} \quad \mu(V_n \cap E) \leq \mu^*(V_n \cap E).$$

De hecho las desigualdades anteriores son igualdades, porque en caso contrario tendríamos

$$\mu(V_n) = \mu(V_n \setminus E) + \mu(V_n \cap E) < \mu^*(V_n \setminus E) + \mu^*(V_n \cap E) = \mu^*(V_n),$$

lo cual contradiría (2.3). Por tanto

$$\mu(V_n \cap E) = \mu^*(V_n \cap E),$$

y tomando límites obtenemos, gracias al Teorema 1.4(f), que

$$\mu(E) = \lim_{n\to\infty} \mu(V_n \cap E) = \lim_{n\to\infty} \mu^*(V_n \cap E) = \mu^*(E).$$

Esto implica que

$$\mu(E) = \inf_{U\supset E,\, U \text{ abierto}} \mu(U). \tag{2.5}$$

Veamos ahora que $\mu(E)$ también puede aproximarse por las medidas de cerrados contenidos en E. Obsérvese que $\mu(V_n \setminus E) < \infty$, y por tanto se sigue de (2.5) que existe un abierto G_n tal que

$$V_n \setminus E \subset G_n, \;\; \mu(G_n \setminus (V_n \setminus E)) < \frac{\varepsilon}{2^n}.$$

Entonces $G = \bigcup_{n=1}^{\infty} G_n$ es abierto, y $C := X \setminus G$ es cerrado y está contenido en E. Puesto que

$$E \setminus C = E \cap \bigcup_{n=1}^{\infty} G_n \subset \bigcup_{n=1}^{\infty} G_n \setminus (V_n \setminus E))$$

se deduce que

$$\mu(E \setminus C) \leq \sum_{n=1}^{\infty} \mu(G_n \setminus (V_n \setminus E)) < \varepsilon,$$

lo que completa la prueba. □

Corolario 2.5. *Si μ es como en el Teorema 2.4 y $\nu : \mathcal{B}(X) \to [0,\infty]$ es otra medida que cumple*

$$\nu(U) = \mu(U) \;\; \textit{para todo } U \textit{ abierto},$$

entonces

$$\mu(E) = \nu(E) \;\; \textit{para todo } E \in \mathcal{B}(X).$$

Demostración. Como μ y ν coinciden en los abiertos, ν también satisface las hipótesis del Teorema 2.4. Entonces se sigue inmediatamente de dicho teorema que $\mu = \nu$. □

Definición 2.5. Sea X un espacio métrico, y μ una medida de Borel en X. Diremos que μ es una *medida de Radon* si $\mu(K) < \infty$ para todo compacto $K \subset X$ y

$$\mu(E) = \inf\{\mu(U) : E \subset U,\, U \text{ abierto}\} = \sup\{\mu(K) : K \subset E,\, K \text{ compacto}\}$$

para todo $E \in \mathcal{B}(X)$.

Corolario 2.6. *Sean X un espacio métrico separable y localmente compacto, y μ una medida de Borel tal que $\mu(K) < \infty$ para cada compacto $K \subset X$. Entonces X es unión de una familia numerable de conjuntos abiertos de medida finita, y μ es una medida de Radon.*

Demostración. Como X es localmente compacto, es unión de una familia de abiertos con adherencias compactas, y al ser también separable podemos extraer un subrecubrimiento numerable; es decir, podemos escribir

$$X = \bigcup_{j=1}^{\infty} U_i,$$

donde $\overline{U_j}$ es compacto para cada $j \in \mathbb{N}$. Puesto que $\mu(U_j) \leq \mu(\overline{U_j}) < \infty$, esto prueba la primera parte del enunciado. Para probar la segunda parte, observemos que para cada cerrado $C \subset X$,

$$C = \bigcup_{n=1}^{\infty} \left(C \cap \bigcup_{j=1}^{n} \overline{U_j} : \right) := \bigcup_{n=1}^{\infty} K_n,$$

unión creciente de compactos, por lo que

$$\mu(C) = \lim_{j\to\infty} \mu(K_j) \leq \sup_{K \subset C,\, K \text{ compacto}} \mu(K) \leq \mu(C),$$

y así

$$\mu(C) = \sup_{K \subset C,\, K \text{ compacto}} \mu(K),$$

lo que en combinación con (2.2) implica que

$$\mu(E) = \inf\{\mu(U) : E \subset U,\, U \text{ abierto}\} = \sup\{\mu(K) : K \subset E,\, K \text{ compacto}\}$$

para todo $E \in \mathcal{B}(X)$. □

2.3. Problemas

Problema 2.1. Sean X un espacio métrico y μ una medida de Borel en X. Para cualquier $E \subset X$ definamos $\mu^*(E) = \inf\{\mu(U) : E \subset U,\, U \text{ abierto}\}$. Demostrar que μ^* es una medida exterior métrica.

Problema 2.2. Si $f : X \to Y$ es un homeomorfismo entre dos espacios métricos, probar que $f(A)$ es boreliano si y sólo si A es boreliano.

Problema 2.3. Demostrar que si μ^* es una medida exterior en un conjunto X que verifica $\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$ para todos $A, B \subseteq X$ disjuntos, entonces μ^* es una medida en $(X, 2^X)$.

Problema 2.4 (Lema de Borel-Cantelli)**.** Sea μ^* una medida exterior en un conjunto X, y sea $(A_n)_{n\in\mathbb{N}} \subset 2^X$ tal que $\sum_{n\in\mathbb{N}} \mu^*(A_n) < \infty$. Demostrar que $\mu^*\left(\{x \in X : x \in A_n \text{ para infinitos } n\}\right) = 0$.

Problema 2.5. Sea $(X, \mathcal{A}, \mu)$ un espacio de medida. Probar que $\mu^*(A) := \inf\{\mu(E) : E \in \mathcal{A}, A \subset E\}$ define una medida exterior en X. [Se le llama *medida exterior generada por* μ.]

Capítulo 3

Dos ejemplos fundamentales: las medidas de Lebesgue y de Hausdorff

En este capítulo, usando el teorema de Carathéodory, a partir de medidas exteriores construiremos las medidas de Lebesgue (que nos permitirán medir volúmenes n-dimensionales de subconjuntos de $\mathbb{R}^n$), y las de Hausdorff (que para cada $t \in [0, n]$ nos permitirán medir subconjuntos t-dimensionales de $\mathbb{R}^n$, o incluso de un espacio métrico).

3.1. Medidas de Hausdorff

Para cada $s \in [0, \infty)$, sea
$$\omega_s := \frac{\pi^{s/2}}{\Gamma\left(1 + \frac{s}{2}\right)},$$
donde recordemos que la función Γ está definida por
$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt.$$
Cuando $s = n$ es un número natural, ω_n es justamente el volumen de la bola unidad de $\mathbb{R}^n$.[1]

Definición 3.1. Sea X un espacio métrico. Para cada $\varepsilon > 0$ se define
$$\mathcal{H}^s_\varepsilon(E) = \text{ínf}\left\{ \frac{\omega_s}{2^s} \sum_{j=1}^\infty \text{diam}(A_j)^s : E \subset \bigcup_{j=1}^\infty A_i,\ \text{diam}(A_j) < \varepsilon \right\},$$
conviniendo que si alguno de los conjuntos A_j no es vacío pero tiene diámetro cero (lo que ocurre si y sólo si A_j consta de un único punto) entonces $\text{diam}(A_j)^s = 0$ para $s > 0$, pero $\text{diam}(A_j)^s = 1$ cuando $s = 0$. En cambio, si A_j es vacío, entenderemos que $\text{diam}(A_j)^s = 0$ en todos los casos.

Puesto que la función $(0, \infty) \ni \varepsilon \to \mathcal{H}^s_\varepsilon(E)$ es decreciente (lo cual es fácil demostrar y se deja como ejercicio para el lector), el límite
$$\mathcal{H}^s(E) = \lim_{\varepsilon \to 0^+} \mathcal{H}^s_\varepsilon(E)$$
existe para cada $E \subset X$.

A la función $2^X \ni E \mapsto \mathcal{H}^s(E)$ se le llama *medida de Hausdorff* (aunque en general es una medida exterior y no una medida, es muy usual denominarla así).

[1] Algunos autores prescinden de esta constante en la definición de las medidas de Hausdorff. Se obtiene en ese caso una medida proporcional a la medida de Hausdorff aquí considerada. La simplicidad de notación que supondría prescindir de la constante ω_s (o incluso de $\omega_s/2^s$) exigiría pagar el precio de obtener, en el caso $s = n$, una medida $\mathcal{H}^n$ que es proporcional a la de Lebesgue $\mathcal{L}_n$, pero no igual a ella. Dicho de otro modo, si queremos que $\mathcal{H}^n$ asigne medida 1 al cubo unidad $[0, 1]^n$ (o que $\mathcal{H}^k$ coincida, en variedades de dimensión k, con la noción habitual de volumen k-dimensional) es necesario introducir la constante $\omega_s/2^s$ en la definición de la medida $\mathcal{H}^s$.

Es fácil ver que cuando $s = 0$, $\mathcal{H}^0$ es la medida de contar en X.

Teorema 3.1. *$\mathcal{H}^s$ es una medida exterior métrica en X.*

Demostración. Es obvio que $\mathcal{H}^s(\emptyset) = 0$, y muy fácil de ver que si $A \subset B$ entonces $\mathcal{H}^s_\varepsilon(A) \leq \mathcal{H}^s_\varepsilon(B)$, y por tanto también $\mathcal{H}^s(A) \leq \mathcal{H}^s(B)$. Veamos que $\mathcal{H}^s$ es numerablemente subaditiva. Dada $\{E_j\}_{j\in\mathbb{N}}$, si $\sum_{j=1}^\infty \mathcal{H}^s(E_j) = \infty$ no hay nada que probar, así que podemos suponer que esta suma es finita. En este caso también se tiene

$$\sum_{j=1}^\infty \mathcal{H}^s_\varepsilon(E_j) < \infty$$

para cada $\varepsilon > 0$. Dado $\delta > 0$, para cada $j \in \mathbb{N}$ podemos encontrar $\{A_{j,k}\}_{k\in\mathbb{N}}$ tal que

$$E_j \subset \bigcup_{k=1}^\infty A_{j,k}, \quad \text{con} \quad \operatorname{diam}(A_{j,k}) < \varepsilon, \quad \text{y} \quad \mathcal{H}^s_\varepsilon(E_j) \geq \frac{\omega_s}{2^s} \sum_{k=1}^\infty \operatorname{diam}(A_{j,k})^s - \frac{\delta}{2^j}.$$

Por tanto, puesto que $\{A_{j,k}\}_{j,k\in\mathbb{N}}$ es un recubrimiento de $\bigcup_{j=1}^\infty E_j$, tenemos

$$\sum_{j=1}^\infty \mathcal{H}^s_\varepsilon(E_j) \geq \frac{\omega_s}{2^s} \sum_{j,k\in\mathbb{N}} \operatorname{diam}(A_{j,k})^s - \delta \geq \mathcal{H}^s_\varepsilon\left(\bigcup_{j=1}^\infty E_j\right) - \delta.$$

Haciendo $\delta \to 0^+$ deducimos que

$$\sum_{j=1}^\infty \mathcal{H}^s(E_j) \geq \sum_{j=1}^\infty \mathcal{H}^s_\varepsilon(E_j) \geq \mathcal{H}^s_\varepsilon\left(\bigcup_{j=1}^\infty E_j\right).$$

Luego, pasando al límite cuando $\varepsilon \to 0^+$, obtenemos que

$$\mathcal{H}^s\left(\bigcup_{j=1}^\infty E_j\right) \leq \sum_{j=1}^\infty \mathcal{H}^s(E_j).$$

Esto prueba que $\mathcal{H}^s$ es una medida exterior (y de hecho que $\mathcal{H}^s_\varepsilon$ también lo es, para cualquier $\varepsilon > 0$). Veamos por último que $\mathcal{H}^s(A \cup B) = \mathcal{H}^s(A) + \mathcal{H}^s(B)$ siempre que $d(E, F) > 0$. Para ello basta ver que, si $0 < \varepsilon < d(A, B)$ entonces $\mathcal{H}^s_\varepsilon(A \cup B) = \mathcal{H}^s_\varepsilon(A) + \mathcal{H}^s_\varepsilon(B)$. Sea $\{C_j\}_{j\in\mathbb{N}}$ una familia de conjuntos tal que

$$A \cup B \subset \bigcup_{j=1}^\infty C_j \quad \text{y} \quad \operatorname{diam}(C_j) < \varepsilon.$$

Podemos suponer que $C_j \cap (A \cup B) \neq \emptyset$ para todo j, porque en otro caso podemos eliminar C_j del recubrimiento sin ninguna consecuencia. Para cada $j \in \mathbb{N}$, como $\operatorname{diam}(C_j) < d(A, B)$, el conjunto C_j corta a A o a B, pero no a ambos. Por tanto podemos escribir $\mathbb{N} = J_A \cup J_B$, con

$$J_A = \{j \in \mathbb{N} : C_j \cap A \neq \emptyset, C_j \cap B = \emptyset\}, \quad J_B = \{j \in \mathbb{N} : C_j \cap B \neq \emptyset, C_j \cap A = \emptyset\}.$$

Entonces

$$\mathcal{H}^s_\varepsilon(A) + \mathcal{H}^s_\varepsilon(B) \leq \sum_{j\in J_A} \frac{\omega_s}{2^s} \operatorname{diam}(C_j)^s + \sum_{j\in J_A} \frac{\omega_s}{2^s} \operatorname{diam}(C_j)^s = \sum_{j\in\mathbb{N}} \frac{\omega_s}{2^s} \operatorname{diam}(C_j)^s.$$

Tomando ínfimos sobre todos los recubrimientos C_j de este tipo obtenemos que

$$\mathcal{H}^s_\varepsilon(A) + \mathcal{H}^s_\varepsilon(B) \leq \mathcal{H}^s_\varepsilon(A \cup B),$$

y como hemos visto antes que la desigualdad opuesta es siempre verdad, concluimos que $\mathcal{H}^s_\varepsilon(A \cup B) = \mathcal{H}^s_\varepsilon(A) + \mathcal{H}^s_\varepsilon(B)$. $\square$

Observación 3.2. La demostración anterior también prueba que $\mathcal{H}^s_\varepsilon$ es una medida exterior en X, para cada $\varepsilon > 0$, $s \in [0, \infty)$. Sin embargo, $\mathcal{H}^s_\varepsilon$ no necesariamente es una medida exterior métrica: en general sólo se cumple que $d(A, B) > 0 \implies \mathcal{H}^s_\varepsilon(A) + \mathcal{H}^s_\varepsilon(B) = \mathcal{H}^s_\varepsilon(A \cup B)$ cuando $0 < \varepsilon < d(A, B)$.

Proposición 3.2. *Sean X un espacio métrico, $s \in [0, \infty)$. Si $T : X \to X$ es una isometría sobreyectiva (es decir, T es sobreyectiva y $d(T(x), T(y)) = d(x, y)$ para todo $x, y \in X$), entonces*

$$\mathcal{H}^s(T(A)) = \mathcal{H}^s(A)$$

para todo $A \subset X$.

Demostración. Es fácil y se deja como ejercicio. □

Corolario 3.3. *Si $L : \mathbb{R}^n \to \mathbb{R}^n$ es una isometría afín, entonces para todo $A \subseteq \mathbb{R}^n$ se tiene $\mathcal{H}^s(L(A)) = \mathcal{H}^s(A)$. En particular $\mathcal{H}^s$ es invariante por rotaciones.*

Proposición 3.4. *Para todos $A \subseteq \mathbb{R}^n$, $\lambda > 0$, se tiene que $\mathcal{H}^s(\lambda A) = \lambda^s \mathcal{H}^s(A)$.*

Demostración. Se deja como ejercicio. □

Proposición 3.5. *Sea E un subconjunto de un espacio métrico X. Se tiene que:*

1. *Si $\mathcal{H}^s(E) < \infty$ entonces $\mathcal{H}^t(E) = 0$ para todo $t > s$.*
2. *Si $\mathcal{H}^s(E) > 0$ entonces $\mathcal{H}^t(E) = \infty$ para todo $t \in (0, s)$.*

Demostración. Se deja como ejercicio. □

Definición 3.3. Se deduce de la proposición anterior que para cada $E \subset \mathbb{R}^n$ existe un único $s = s(E) \in [0, \infty]$ tal que $\mathcal{H}^t(E) = 0$ para $t > s$ y $\mathcal{H}^t(E) = \infty$ para $0 < t < s$. A este s se le llama *dimensión de Hausdorff de E*. Más precisamente, la dimensión de E se define como sigue. Si $\mathcal{H}^s(E) > 0$ para todo $s \geq 0$ entonces $\dim_H(E) = \infty$. En cualquier otro caso, se define

$$\dim_H(E) = \operatorname{ínf}\{s \in [0, \infty) : \mathcal{H}^s(E) = 0\}.$$

Proposición 3.6. *Sean X un espacio métrico, y $\{A_k\}_{k \in \mathbb{N}}$ una familia numerable de subconjuntos de X. Entonces*

$$dim_H \left(\bigcup_{k=1}^{\infty} A_k \right) = \sup_{k \in \mathbb{N}} dim_H (A_k).$$

En particular, si A es numerable entonces $dim_H (A) = 0$.

Demostración. Es fácil y se propone como ejercicio. □

Recordemos que una función $f : X \to Y$ entre dos espacios métricos es de Lipschitz si existe una constante $C \geq 0$ tal que $d(f(x), f(y)) \leq d(x, y)$ para todos $x, y \in X$. Al ínfimo de tales C se le llama la constante de Lipschitz de f, y se denota $\mathrm{Lip}(f)$.

Proposición 3.7. *Sean X, Y espacios métricos, y $f : X \to Y$ de Lipschitz. Entonces, para todo $A \subseteq X$ se tiene $\mathcal{H}^s(f(A)) \leq (Lip(f))^s \mathcal{H}^s(A)$.*

Demostración. Sea $C > \mathrm{Lip}(f)$. Es fácil ver que $\mathrm{diam}(f(B)) \leq C\mathrm{diam}(B)$ para cualquier $B \subset X$. Entonces, si $\{A_j\}_{j \in \mathbb{N}}$ es tal que $A \subset \bigcup_{j=1}^{\infty} A_j$ y $\mathrm{diam}(A_j) < \delta$ para todo j, se tiene que $f(A) \subset \bigcup_{j=1}^{\infty} f(A_j)$ y $\mathrm{diam}(f(A_j)) \leq C\mathrm{diam}(A_j) < C\delta$, luego

$$\mathcal{H}^s_{C\delta}(f(A)) \leq \frac{\omega_s}{2^s} \sum_{j=1}^{\infty} \mathrm{diam}(f(A_j))^s \leq C^s \frac{\omega_s}{2^s} \sum_{j=1}^{\infty} \mathrm{diam}(A_j)^s.$$

Tomando ínfimos sobre tales $\{A_j\}$, deducimos que

$$\mathcal{H}^s_{C\delta}(f(A)) \leq C^s \mathcal{H}^s_\delta(A),$$

y haciendo $\delta \to 0$ concluimos que $\mathcal{H}^s(f(A)) \leq C^s\mathcal{H}^s(A)$. Como $C > \mathrm{Lip}(f)$ es arbitrario, esta desigualdad también vale con $C = \mathrm{Lip}(f)$. □

Corolario 3.8. *Si $n > k$ y $P : \mathbb{R}^n \to \mathbb{R}^k$ es la proyección lineal canónica, entonces para todos $A \subseteq \mathbb{R}^n$, $s \in [0, \infty)$ se tiene $\mathcal{H}^s(P(A)) \leq \mathcal{H}^s(A)$.*

3.2. La medida de Lebesgue en $\mathbb{R}^n$

Dado cualquier rectángulo

$$R = [a_1, b_1] \times ... \times [a_n, b_n] \subset \mathbb{R}^n,$$

definimos su volumen por

$$|R| = \prod_{j=1}^{n}(b_j - a_j),$$

y, para cada $A \subset \mathbb{R}^n$, la medida exterior de Lebesgue de A como

$$\mathcal{L}^*_n(A) = \inf \sum_{j=1}^{\infty} |R_j|,$$

donde el ínfimo se toma sobre todas las colecciones a lo sumo numerables de rectángulos cerrados $\{R_j\}_{j\in\mathbb{N}}$ tales que $A \subset \bigcup_{j=1}^{\infty} R_j$ (no excluimos la posibilidad de que en algunos casos sean $R_j = \emptyset$ para todo $j \geq k$).

Teorema 3.9. *$\mathcal{L}^*_n$ es una medida exterior métrica.*

Demostración. Es muy fácil probar, y se deja como ejercicio, que $\mathcal{L}^*_n(\emptyset) = 0$ y que $A \subset B \implies \mathcal{L}^*_n(A) \leq \mathcal{L}^*_n(B)$. Veamos que $\mathcal{L}^*_n$ es numerablemente subaditiva. Sea $\{A_j\}_{j\in\mathbb{N}} \subset \mathbb{R}^n$. Queremos ver que

$$\mathcal{L}^*_n\left(\bigcup_{j=1}^{\infty} A_j\right) \leq \sum_{j=1}^{\infty} \mathcal{L}^*_n(A_j).$$

Si la suma de la derecha es infinita no hay nada que probar. Si es finita, también es finita $\mathcal{L}^*_n(A_j)$ para cada j, y dado cualquier $\varepsilon > 0$, por su definición podemos encontrar una colección de rectángulos $\{R_{j,k}\}_{k\in\mathbb{N}}$ tales que

$$A_j \subset \bigcup_{k=1}^{\infty} R_{j,k}, \quad \text{y} \quad \sum_{k=1}^{\infty} |R_{j,k}| < \mathcal{L}^*_n(A_j) + \frac{\varepsilon}{2^j}.$$

Entonces, puesto que $\bigcup_{j,k\in\mathbb{N}} R_{j,k} \supset \bigcup_{j=1}^{\infty} A_j$, se tiene que

$$\mathcal{L}^*_n\left(\bigcup_{j=1}^{\infty} A_j\right) \leq \sum_{j,k\in\mathbb{N}} |R_{j,k}| \leq \sum_{j=1}^{\infty} \mathcal{L}^*_n(A_j) + \varepsilon.$$

Como $\varepsilon > 0$ es arbitario, se deduce que

$$\mathcal{L}^*_n\left(\bigcup_{j=1}^{\infty} A_j\right) \leq \sum_{j=1}^{\infty} \mathcal{L}^*_n(A_j),$$

como queríamos probar.

Finalmente, veamos que si $d(A,B)>0$ entonces $\mathcal{L}_n^*(A\cup B)=\mathcal{L}_n^*(A)+\mathcal{L}_n^*(B)$. Basta probar que $\mathcal{L}_n^*(A\cup B)\geq\mathcal{L}_n^*(A)+\mathcal{L}_n^*(B)$, y podemos suponer que $\mathcal{L}_n^*(A\cup B)<\infty$. Dado $\varepsilon>0$ existe $\{R_j\}_{j\in\mathbb{N}}$ familia de rectángulos tal que

$$A\cup B\subset\bigcup_{j=1}^{\infty}R_j \quad \text{y} \quad \mathcal{L}_n^*(A\cup B)+\varepsilon\geq\sum_{j=1}^{\infty}|R_j|.$$

Podemos suponer que $R_j\cap(A\cup B)\neq\emptyset$ para todo j, porque en otro caso podemos eliminar R_j del recubrimiento. Subdividiendo cada rectángulo R_j en una cantidad finita de rectángulos suficientemente pequeños con interiores disjuntos dos a dos, digamos $R_j=\bigcup_{k=1}^{m_j}Q_{j,k}$, observando que $|R_j|=\sum_{k=1}^{m_j}|Q_{j,k}|$, y reindexando la familia numerable $Q_{j,k}$, podemos suponer sin pérdida de generalidad que el diámetro de cada rectángulo R_j es menor que $\frac{1}{2}d(A,B)$. Entonces cada rectángulo R_j corta a A o a B, pero no a ambos. Por tanto podemos escribir $\mathbb{N}=J_A\cup J_B$, con

$$J_A=\{j\in\mathbb{N}:R_j\cap A\neq\emptyset,R_j\cap B=\emptyset\},\ \ J_B=\{j\in\mathbb{N}:R_j\cap B\neq\emptyset,R_j\cap A=\emptyset\}.$$

Entonces

$$\mathcal{L}_n^*(A)+\mathcal{L}_n^*(B)\leq\sum_{j\in J_A}|R_j|+\sum_{j\in J_B}|R_j|=\sum_{j\in\mathbb{N}}|R_j|\leq\mathcal{L}_n^*(A\cup B)+\varepsilon,$$

es decir,

$$\mathcal{L}_n^*(A)+\mathcal{L}_n^*(B)\leq\mathcal{L}_n^*(A\cup B)+\varepsilon,$$

y como $\varepsilon>0$ es arbitrario se deduce que $\mathcal{L}_n^*(A)+\mathcal{L}_n^*(B)\leq\mathcal{L}_n^*(A\cup B)$. □

Definición 3.4. Así pues, por el teorema de Carathéodory, la familia de los subconjuntos E de $\mathbb{R}^n$ tales que $\mathcal{L}_n^*(A\cap E)+\mathcal{L}_n^*(A\cap E^c)=\mathcal{L}_n^*(A)$ es una σ-álgebra que contiene a los borelianos $\mathcal{B}(\mathbb{R}^n)$. A tales subconjuntos E se les llama *medibles Lebesgue*.

A la restricción de $\mathcal{L}_n^*$ a la σ-álgebra de los subconjuntos medibles Lebesgue se le llama *medida de Lebesgue* en $\mathbb{R}^n$, y se denotará $\mathcal{L}_n$.

Observación 3.5. Todos los conjuntos borelianos son medibles Lebesgue. Y todos los subconjuntos $E\subset\mathbb{R}^n$ tales que $\mathcal{L}_n^*(E)=0$ son medibles Lebesgue (aunque no necesariamente son borelianos; ver el problema 4.10), y verifican $\mathcal{L}_n(E)=0$. Dicho de otro modo, la medida de Lebesgue es completa en la σ-álgebra de los conjuntos medibles Lebesgue (pero no lo es al restringirla a la σ-álgebra de los borelianos).

Nótese también que un conjunto $E\subset\mathbb{R}^n$ tiene medida de Lebesgue cero si y sólo si puede recubrirse por una familia numerable de rectángulos cuya suma de volúmenes puede hacerse tan pequeña como se quiera.

En lo sucesivo, cuando hablemos de subconjuntos medibles de $\mathbb{R}^n$, si no hay mención expresa en otro sentido, entenderemos que se trata de subconjuntos medibles Lebesgue.

Proposición 3.10. *Se tiene que*

$$\mathcal{L}_n^*(A)=\mathcal{L}_n^*(x+A)$$

para todo $x\in\mathbb{R}^n$, $A\subseteq\mathbb{R}^n$, *y que*

$$\mathcal{L}_n^*(rA)=r^n\mathcal{L}_n^*(A)$$

para todo $r>0$.

Demostración. Es fácil y se deja como ejercicio. □

Corolario 3.11. *Si* $E\subset\mathbb{R}^n$ *es medible entonces también lo son* $a+E$ *y* rE*, para cualesquiera* $a\in\mathbb{R}^n, r>0$.

Demostración. Se deja al cuidado del lector. □

Proposición 3.12. *Si R es un rectángulo cerrado de $\mathbb{R}^n$, entonces $\mathcal{L}_n(R) = |R|$.*

Demostración. Es obvio que $\mathcal{L}_n(R) \leq |R|$. Para probar la otra desigualdad tenemos que demostrar que si $\{R_j\}_{j\in\mathbb{N}}$ es una familia de rectángulos tales que $R \subset \bigcup_{j=1}^{\infty} R_j$ entonces $|R| \leq \sum_{j=1}^{\infty} |R_j|$. Dado $\varepsilon > 0$, observemos que cada rectángulo R_j está contenido en un rectángulo abierto R_j^ε ligeramente mayor tal que

$$|\overline{R}_j^\varepsilon| \leq |R_j| + \frac{\varepsilon}{2^j}.$$

Como $R \subset \bigcup_{j=1}^{\infty} R_j^\varepsilon$ y R es compacto, existe $k \in \mathbb{N}$ tal que

$$R \subset \bigcup_{j=1}^{k} R_j^\varepsilon \subset \bigcup_{j=1}^{k} \overline{R}_j^\varepsilon$$

Proyectando los lados de los rectánculos $\overline{R}_j^\varepsilon$ en las direcciones coordenadas y haciendo todos los productos cartesianos de los intervalos así obtenidos (un sólo intervalo por cada dirección para cada producto), se obtiene una familia finita de rectánculos con interiores disjuntos dos a dos cuya unión recubre a R. Al intersecar cada miembro de esta familia con el rectángulo R se obtiene otra familia finita de rectángulos Q_i con interiores disjuntos dos a dos cuya unión es exactamente R, y con la propiedad de que cada rectángulo Q_i está contenido en alguno de los rectángulos $\overline{R}_j^\varepsilon$. Entonces se tiene

$$|R| = \sum_j |Q_j| \leq \sum_{j=1}^{k} |\overline{R}_j^\varepsilon| \leq \sum_{j=1}^{\infty} |R_j| + \varepsilon.$$

Como $\varepsilon > 0$ es arbitrario, se deduce que $|R| \leq \sum_{j=1}^{\infty} |R_j|$, como queríamos probar. □

Notación 3.6. En lo que sigue, cuando no haya ambigüedad respecto de la dimensión n, denotaremos

$$\mathcal{L}_n(A) = |A|$$

para cada $A \subset \mathbb{R}^n$ medible Lebesgue.

Proposición 3.13. *Para cada rectángulo abierto y acotado $R = (a_1, b_1) \times ... \times (a_n, b_n)$ en $\mathbb{R}^n$, se tiene que*

$$|R| = \prod_{j=1}^{n} (b_j - a_j),$$

y

$$|\partial R| = 0.$$

Por otro lado, si R es un producto cartesiano de n intervalos, alguno de los cuales es no acotado, y con interior no vacío, entonces $|R| = \infty$.

Demostración. Para cada $\varepsilon > 0$, los rectángulos

$$Q^\varepsilon = [a_1 - \varepsilon, b_1 + \varepsilon] \times ... \times [a_n - \varepsilon, b_n + \varepsilon] \quad \text{y} \quad Q_\varepsilon = [a_1 + \varepsilon, b_1 - \varepsilon] \times ... \times [a_n + \varepsilon, b_n - \varepsilon]$$

cumplen que

$$Q_\varepsilon \subset R \subset Q^\varepsilon.$$

Por tanto

$$|Q_\varepsilon| \leq |R| \leq |Q^\varepsilon|$$

para todo ε, y en el límite

$$\prod_{j=1}^{n} (b_j - a_j) = \lim_{\varepsilon\to 0^+} |Q_\varepsilon| \leq |R| \leq \lim_{\varepsilon\to 0^+} |Q^\varepsilon| = \prod_{j=1}^{n} (b_j - a_j).$$

La segunda afirmación se sigue de los hechos de que $|\overline{R}| = |R| + |\partial R|$ y que $|R| = |\overline{R}|$. La tercera afirmación es obvia porque R contiene rectángulos acotados de medida arbitrariamente grande. □

Corolario 3.14. *Para cualquier $a \in \mathbb{R}^{n-k}$, se tiene $\mathcal{L}_n(\mathbb{R}^k \times \{a\}) = 0$; en particular $\mathcal{L}_n(\mathbb{R}^k \times \{0\}) = 0$.*

Demostración. $\mathbb{R}^k \times \{a\}$ puede recubrirse por una familia numerable de fronteras de rectángulos acotados, y por la proposición anterior cada una de estas fronteras tiene medida cero. □

El siguiente resultado nos dice que se puede calcular la medida de cualquier abierto representándolo como una unión numerable de cubos con interiores disjuntos dos a dos.[2] Diremos que un rectángulo es un *cubo diádico* si sus lados tienen todos la misma longitud, igual a 2^{-k} para algún $k \in \mathbb{N} \cup \{0\}$.

Teorema 3.15. *Todo abierto U de $\mathbb{R}^n$ es unión numerable de cubos diádicos con interiores disjuntos dos a dos. Por tanto $|U|$ es igual a la suma de las medidas de estos cubos.*

Demostración. Denotemos por $\mathcal{F}$ la familia de todos los cubos cerrados de la forma

$$\left[\frac{k_1}{2^m}, \frac{k_1+1}{2^m}\right] \times \ldots \times \left[\frac{k_n}{2^m}, \frac{k_n+1}{2^m}\right],$$

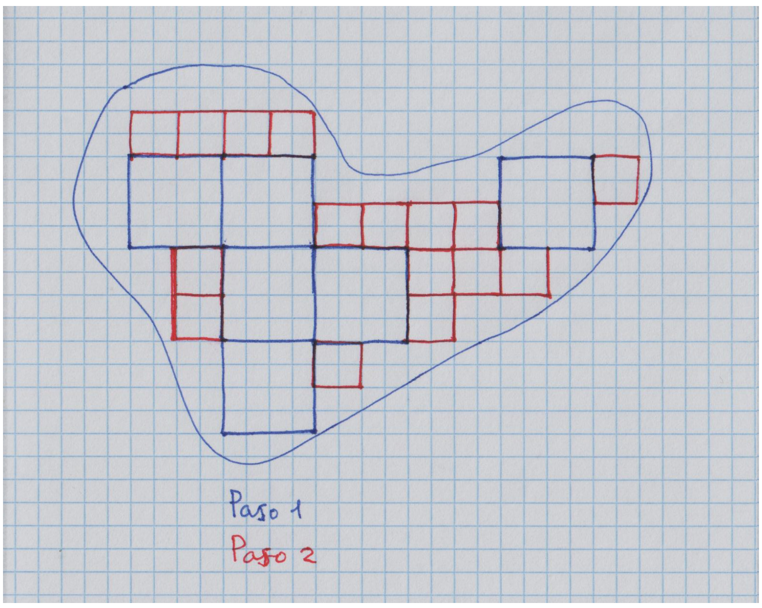
con $k_i \in \mathbb{Z}$ y $m \in \mathbb{N}$. Sea $\mathcal{Q}_1$ la familia de todos los cubos cerrados Q de la forma $[k_1, k_1+1] \times \ldots \times [k_n, k_n+1]$, donde los $k_i \in \mathbb{Z}$, y tales que $Q \subset U$. Supuesto definida $\mathcal{Q}_m$, sea $\mathcal{Q}_{m+1}$ la familia de todos los cubos Q de la forma $[\frac{k_1}{2^m}, \frac{k_1+1}{2^m}] \times \ldots \times [\frac{k_n}{2^m}, \frac{k_n+1}{2^m}]$, donde $k_i \in \mathbb{Z}$, tales que no están contenidos en ningún cubo $Q' \in \mathcal{Q}_j$ para $j \leq m$, y tales que $Q \subset U$. Por inducción queda definida $\mathcal{Q}_m$ para todo $m \in \mathbb{N}$, y ponemos

$$\mathcal{Q} = \bigcup_{m=1}^{\infty} \mathcal{Q}_m.$$

Es obvio por construcción que si $Q, Q' \in \mathcal{Q}$ y $Q \neq Q'$, entonces Q y Q' tienen interiores disjuntos. También es claro que que $\bigcup_{Q \in \mathcal{Q}} Q \subset U$. Veamos que de hecho

$$U = \bigcup_{Q \in \mathcal{Q}} Q.$$

Dado $x \in U$, usando que U es abierto y que el conjunto $\{k/2^m : k \in \mathbb{Z}, m \in \mathbb{N} \cup 0\}$ es denso en $\mathbb{R}$, es fácil ver que existe algún cubo $Q_x \in \mathcal{F}$ tal que $x \in Q$ y $Q \subset U$. El lado de Q_x mide 2^{-m_x} para algún $m_x \in \mathbb{N} \cup \{0\}$. Si $Q_x \in \mathcal{Q}_{m_x}$ ya hemos terminado. En otro caso, por definición de $\mathcal{Q}_{m_x}$, existe algún $j < m_x$ tal que Q_x está contenido en algún cubo $Q'_x \in \mathcal{Q}_j$, y por tanto x pertenece a este cubo. En cualquier caso se ve que $x \in \bigcup_{i=1}^{\infty} \mathcal{Q}_i$. □



[2] Sin embargo las aplicaciones de este teorema serán más teóricas que prácticas.

Teorema 3.16. *Para cualquier $E \subset \mathbb{R}^n$, se tiene*

$$\mathcal{L}_n^*(E) = \inf_{E \subset U,\, U \text{ abierto}} \mathcal{L}_n(U).$$

Demostración. Es obvio que $\mathcal{L}_n^*(E) \leq \inf_{E\subset U,\, U \text{ abierto}} \mathcal{L}_n(U)$. Para ver la otra desigualdad, podemos suponer que $\mathcal{L}_n^*(E) < \infty$. Entonces, dado cualquier $\varepsilon > 0$, existe $\{R\}_{j\in\mathbb{N}}$ familia de rectángulos cerrados tales que $E \subset \bigcup_{j=1}^{\infty} R_j$ y $\sum_{j=1}^{\infty} |R_j| \leq \mathcal{L}_n^*(E) + \varepsilon/2$. Cada R_j está contenido en un rectángulo abierto R_j^ε de tamaño ligeramente mayor, digamos $|R_j^\varepsilon| < |R_j| + \varepsilon/2^{j+1}$. Entonces $V_\varepsilon = \bigcup_{j=1}^{\infty} R_j^\varepsilon$ es abierto, contiene a E, y tiene medida $|V_\varepsilon| \leq \sum_{j=1}^{\infty} |R_j^\varepsilon| \leq \sum_{j=1}^{\infty} |R_j| + \varepsilon/2 \leq \mathcal{L}_n^*(E) + \varepsilon$. Por tanto $\inf_{E\subset U,\, U \text{ abierto}} \mathcal{L}_n(U) \leq \mathcal{L}_n^*(E)$. □

El siguiente teorema de estructura de los conjuntos medibles Lebesgue es fundamental. Nos dice, entre otras cosas, que cualquier conjunto medible Lebesgue, aunque pueda no ser boreliano, sólo se diferencia de algún boreliano en un conjunto de medida cero.

Teorema 3.17. *Sea $A \subset \mathbb{R}^n$. Las siguientes afirmaciones son equivalentes:*

1. *A es medible Lebesgue;*
2. *Para cada $\varepsilon > 0$ existe un abierto G tal que $A \subset G$ y $\mathcal{L}_n^*(G \setminus A) < \varepsilon$;*
3. *Existe H, conjunto G_δ, tal que $A \subset H$ y $\mathcal{L}_n^*(H \setminus A) = 0$,*
4. *Para cada $\varepsilon > 0$ existe F cerrado tal que $F \subset A$ y $\mathcal{L}_n^*(A \setminus F) < \varepsilon$;*
5. *Existe M, conjunto F_σ, tal que $M \subset A$ y $\mathcal{L}_n^*(A \setminus M) = 0$;*
6. *Para cada $\varepsilon > 0$ existen G abierto y F cerrado tales que $F \subset A \subset G$ y $\mathcal{L}_n(G \setminus F) < \varepsilon$.*

Demostración. $(1) \implies (2)$: todo conjunto medible A puede escribirse $A = \bigcup_{j=1}^{\infty} A_j$, con $|A_j| < \infty$ (por ejemplo puede tomarse $A_j = A \cap B(0, j)$ para cada j). Por el teorema anterior, para cada j existe G_j abierto tal que $A_j \subset G_j$ y $|G_j \setminus A_j| = |G_j| - |A_j| < \varepsilon/2^j$. Entonces $A \subset G := \bigcup_{j=1}^{\infty} G_j$, y $|G \setminus A| < \varepsilon$.

$(2) \implies (3)$: Basta definir $H = \bigcap_{j=1}^{\infty} U_j$, donde los U_j son abiertos tales que $\mathcal{L}_n^*(U_j \setminus A) < 1/j$.

$(3) \implies (1)$ es obvio.

$(1) \iff (4) \iff (5)$ se sigue fácilmente de las equivalencias $(1) \iff (2) \iff (3)$ aplicadas al conjunto $\mathbb{R}^n \setminus A$ (téngase en cuenta que los complementarios de los G_δ son los F_σ).

$(1) \implies (6)$: esto se sigue inmediatamente de que $(1) \implies (2)$ y $(1) \implies (4)$.

$(6) \implies (2)$ es inmediato. □

El siguiente resultado muestra que existen subconjuntos de $\mathbb{R}^n$ que no son medibles Lebesgue. Esto no es un defecto particular de la medida de Lebesgue: la misma prueba muestra que si μ es una medida en $\mathbb{R}^n$ invariante por traslaciones y que asigna un número positivo al cubo unidad, entonces hay conjuntos que μ no puede medir.

Teorema 3.18 (Vitali)**.** *Existe un conjunto $E \subset \mathbb{R}^n$ que no es medible Lebesgue.*

Demostración. Por simplicidad de notación haremos la prueba para $n = 1$, pero el mismo argumento vale para cualquier n con cambios obvios. La idea básica es encontrar una familia numerable de conjuntos disjuntos dos a dos que son traslaciones unos de otros, cuya unión E contiene al intervalo $(0, 1)$ y está contenida en $(-1, 2)$. Si estos conjuntos disjuntos son distintas traslaciones de un mismo conjunto V, entonces V no puede ser medible, porque si lo fuera entonces su medida, un número entre 1 y 3, sería igual a una suma infinita de números iguales.

Para cada $x, y \in (0, 1)$, definimos $x\mathcal{R}y$ si y sólo si $x - y \in \mathbb{Q}$. Es claro que $\mathcal{R}$ es una relación de equivalencia. Por tanto, denotando $\mathcal{F}$ el conjunto cociente $(0, 1)/\mathcal{R}$, tenemos que

$$(0, 1) = \bigcup_{F\in\mathcal{F}} F,$$

unión disjunta. Por el axioma de elección, existe un conjunto $V \subset (0,1)$ que contiene exactamente un elemento de cada conjunto $F \in \mathcal{F}$; es decir, V contiene un y sólo un representante $x_F \in (0,1)$ de cada clase de equivalencia $F = [x_F] = (x_F + \mathbb{Q}) \cap (0,1)$. Sea

$$E = \bigcup_{a \in \mathbb{Q} \cap (-1,1)} V_a, \quad \text{donde} \quad V_a := a + V = \{a + x : x \in V\}.$$

Es fácil ver que

$$V_a \cap V_b = \emptyset$$

si $a, b \in \mathbb{Q}$ con $a \neq b$, y que

$$(0,1) \subset E \subset (-1,2).$$

Supongamos que V fuera medible. Entonces V_a es medible para cualquier $a \in \mathbb{Q}$, con $\mathcal{L}_1(V) = \mathcal{L}_1(V_a)$, y E es medible por ser unión numerable de medibles. Si $\mathcal{L}_1(V) > 0$ entonces $\mathcal{L}_1(E) = \sum_{a \in \mathbb{Q} \cap (-1,1)} \mathcal{L}_1(V_a) = \sum_{a \in \mathbb{Q} \cap (-1,1)} \mathcal{L}_1(V) = \infty$, lo cual es imposible porque $E \subset (-1,2)$. Y si $\mathcal{L}_1(V) = 0$ entonces $\mathcal{L}_1(E) = 0$, lo cual también es imposible ya que $(0,1) \subset E$. Por tanto V no puede ser medible. $\square$

Sabemos que la medida de Lebesgue es invariante por traslaciones y asigna el número 1 al cubo unidad $[0,1]^n$. Vamos a ver que cualquier otra medida que tenga estas dos propiedades coincide con la de Lebesgue. Por tanto la medida de Lebesgue nos proporciona la única manera natural de medir volúmenes n-dimensionales en $\mathbb{R}^n$.

Teorema 3.19. *Si $\mu : \mathcal{B}(\mathbb{R}^n) \to [0,\infty]$ es una medida tal que $\mu(a+E) = \mu(E)$ para todo $a \in \mathbb{R}^n$ y todo $E \in \mathcal{B}(\mathbb{R}^n)$ y $\mu([0,1]^n) = 1$, entonces $\mu(E) = \mathcal{L}_n(E)$ para todo $E \in \mathcal{B}(\mathbb{R}^n)$.*

Demostración. Por el Corolario 2.5, basta probar que $\mu(U) = \mathcal{L}_n(U)$ para cada abierto U de $\mathbb{R}^n$. A su vez, puesto que cada abierto es unión numerable de cubos diádicos con interiores disjuntos dos a dos, basta probar que μ y $\mathcal{L}_n$ coinciden en cada cubo diádico, y que ambas asignan medida cero a la frontera de cada cubo (esto último ya lo sabemos de la medida de Lebesgue, pero habrá que probarlo para μ).

Consideremos el cubo abierto $Q_k = (0, 2^{-k})^n$ de lado $1/2^k$. Hay 2^{kn} cubos abiertos disjuntos dos a dos contenidos en el cubo $[0,1]^n$, cada uno de los cuales es una traslación $a_j + Q_k$ de Q_k. Como μ es invariante por traslaciones, se tiene que

$$2^{kn}\mu(Q_k) = \sum_{j=1}^{2^{kn}} \mu(a_j + Q_k) = \mu\left(\bigcup_{j=1}^{2^{kn}} (a_j + Q_k)\right) \leq \mu\left([0,1]^n\right) = 1,$$

de donde

$$\mu(Q_k) \leq 2^{-kn}. \tag{3.1}$$

Supongamos por un momento que ya sabemos que $\mu(\partial Q) = 0$ para cada cubo diádico Q. Entonces, como $[0,1]^n$ se puede recubrir por 2^{kn} cubos cada uno de los cuales es una traslación $a_j + \overline{Q_k}$ de $\overline{Q_k}$ y μ es invariante por traslaciones, se tiene que

$$1 = \mu([0,1]^n) \leq \mu\left(\bigcup_{j=1}^{2^{kn}} (a_j + \overline{Q_k})\right) \leq \sum_{j=1}^{2^{kn}} \mu(a_j + \overline{Q_k}) = 2^{kn}\mu(\overline{Q_k}),$$

luego

$$\mu(\overline{Q_k}) \geq 2^{-kn}.$$

Por tanto

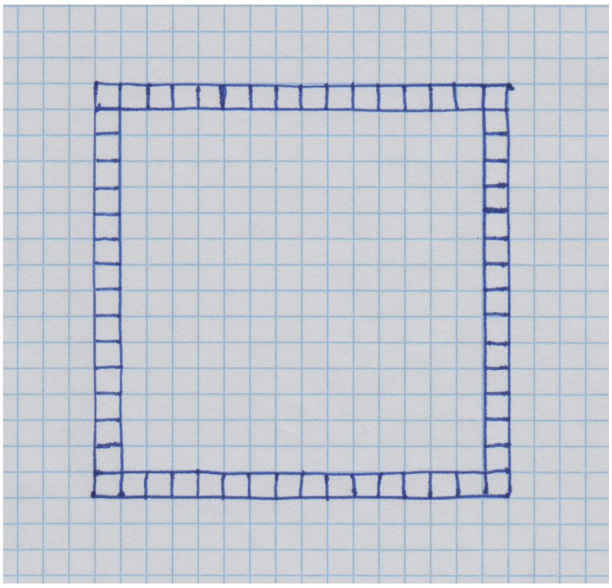
$$2^{-kn} \leq \mu(\overline{Q_k}) = \mu(Q_k) \leq 2^{-kn},$$

y así

$$\mu(Q_k) = \mu(\overline{Q_k}) = 2^{-kn} = \mathcal{L}_n(\overline{Q_k}),$$

que es lo que queremos (nótese que cualquier otro cubo diádico es traslación de un Q_k y tiene la misma medida que este).

Por tanto sólo queda ver que $\mu(\partial Q) = 0$ para cada cubo diádico Q en $\mathbb{R}^n$. En dimensión n, las $2n$ caras de la frontera de un cubo diádico de lado 2^{-i} pueden recubrirse por $2n2^{k(n-1)}$ cubos diádicos de lado 2^{-i-k}, la suma de cuyas μ-medidas, en vista de (3.1), es menor o igual que $2n2^{k(n-1)}2^{-in-kn} = n2^{-k-in}$. Como n está fijo y k es arbitrario y $\lim_{k\to\infty} n2^{-k-in} = 0$, esto muestra que $\mu(\partial Q) = 0$. $\square$



Demostraremos más adelante (ver el Teorema 12.5) que en $\mathbb{R}^n$ las medidas $\mathcal{H}^n$ y $\mathcal{L}_n^*$ coinciden. De momento vamos a probar un resultado más elemental que nos dice que estas dos medidas son proporcionales; esto será suficiente para poder deducir algunas propiedades para una de las medidas del hecho de que la otra las tiene. Por ejemplo, es muy sencillo ver que $\mathcal{H}^n$ es invariante por rotaciones, pero no lo es el ver directamente que $\mathcal{L}_n^*$ también tiene esta propiedad.

Proposición 3.20. *Si A es un abierto acotado de $\mathbb{R}^n$ entonces $0 < \mathcal{H}^n(A) < \infty$. Por tanto la dimensión de Hausdorff de cualquier abierto de $\mathbb{R}^n$ es n.*

Demostración. Observemos que si $A_j \subset \mathbb{R}^n$ tiene diámetro δ_j entonces, dado cualquier $x_j \in A_j$, $Q_J := B_\infty(x_j, \delta_j)$ es un cubo de lado $2\delta_j$ que contiene a A_j. Por tanto, si $A \subset \bigcup_{j=1}^\infty A_j$ y $\operatorname{diam}(A_j) = \delta_j < \delta$, también se tiene

$$A \subset \bigcup_{j=1}^\infty Q_j,$$

y

$$\frac{\omega_n}{2^n}\sum_{j=1}^\infty \operatorname{diam}(A_j)^n = \frac{\omega_n}{4^n}\sum_{j=1}^\infty |Q_j| \geq \frac{\omega_n}{4^n}\mathcal{L}_n^*(A)$$

En particular, si A es abierto se tiene, tomando ínfimos sobre tales $\{A_j\}$, que

$$\mathcal{H}^n_\delta(A) \geq \frac{\omega_n}{4^n}\mathcal{L}_n(A) > 0.$$

Haciendo $\delta \to 0$ obtenemos que

$$\mathcal{H}^n(A) \geq \frac{\omega_n}{4^n}\mathcal{L}_n(A) > 0.$$

Por otro lado, si $R = [0, 1/2^i]^n$, para todo $k \in \mathbb{N}$, este cubo diádico R puede descomponerse como la unión de 2^{kn} subcubos diádicos R_j de lado $\delta_k := 1/2^{k+i}$. Entonces

$$\mathcal{H}^n_{\delta_k}(R) \leq \frac{\omega_n}{2^n}\sum_{j=1}^{2^{kn}} \operatorname{diam}(R_j)^n = \frac{\omega_n}{2^n}\sum_{j=1}^{2^{kn}} \delta_k^n n^{n/2} = \frac{\omega_n}{2^n} n^{n/2}\sum_{j=1}^{2^{kn}} |R_j| = \frac{\omega_n}{2^n} n^{n/2}|R|.$$

Haciendo $k \to \infty$ deducimos que

$$\mathcal{H}^n(R) \le \frac{\omega_n}{2^n} n^{n/2} |R|.$$

Como $\mathcal{H}^n$ es invariante por traslaciones y todo abierto A es unión numerable de cubos diádicos con interiores disjuntos dos a dos, esto implica que, para todo abierto A,

$$\mathcal{H}^n(A) \le \frac{\omega_n}{2^n} n^{n/2} \mathcal{L}_n(A).$$

En particular, si A es un abierto acotado, se ve que

$$0 < \frac{\omega_n}{4^n} \mathcal{L}_n(A) \le \mathcal{H}^n(A) \le \frac{\omega_n}{2^n} n^{n/2} \mathcal{L}_n(A) < \infty.$$

□

Corolario 3.21. *Existe una constante $\lambda_n > 0$ tal que $\mathcal{H}^n(E) = \lambda_n \mathcal{L}_n(E)$ para todo $E \in \mathcal{B}(\mathbb{R}^n)$.*[3]

Demostración. Puesto que la medida de Hausdorff $\mathcal{H}^n$ es invariante por traslaciones, se deduce que la medida

$$\mu(E) := \frac{1}{\mathcal{H}^n\left([0,1]^n\right)} \mathcal{H}^n(E), \quad E \in \mathcal{B}(\mathbb{R}^n)$$

también lo es (esta medida está bien definida porque, como consecuencia de la proposición anterior, $0 < \mathcal{H}^n([0,1]^n) < \infty$). Como además $\mu\left([0,1]^n\right) = 1$, el Teorema 3.19 nos dice que $\mu = \mathcal{L}_n$ en $\mathcal{B}(\mathbb{R}^n)$. □

Corolario 3.22. *Para toda isometría afín sobreyectiva $T : \mathbb{R}^n \to \mathbb{R}^n$ y todo $A \subset \mathbb{R}^n$ se tiene que*

$$\mathcal{L}_n^*(T(A)) = \mathcal{L}_n^*(A).$$

En particular $\mathcal{L}_n^*$ es invariante por rotaciones.

Demostración. Supongamos primero que A es abierto. Entonces, como $\mathcal{H}^n$ es invariante por isometrías sobreyectivas,

$$\mathcal{L}_n(T(A)) = \frac{1}{\lambda_n} \mathcal{H}^n(T(A)) = \frac{1}{\lambda_n} \mathcal{H}^n(A) = \mathcal{L}_n(A).$$

Una vez probado para abiertos, el resultado se extiende inmediatamente a cualquier subconjunto de $\mathbb{R}^n$ usando el Teorema 3.16. □

3.3. Problemas

Problema 3.1. Probar que $\mathcal{H}^0$ coincide con la medida de contar.

Problema 3.2. Para todos $A \subseteq \mathbb{R}^n$, $\lambda > 0$, probar que $\mathcal{H}^s(\lambda A) = \lambda^s \mathcal{H}^s(A)$.

Problema 3.3. Si $L : \mathbb{R}^n \to \mathbb{R}^n$ es una isometría, probar que $\mathcal{H}^s(L(A)) = \mathcal{H}^s(A)$ para todo $A \subseteq \mathbb{R}^n$.

Problema 3.4. Sea E un subconjunto de un espacio métrico. Probar que:

1. Si $\mathcal{H}^s(E) < \infty$ entonces $\mathcal{H}^t(E) = 0$ para todo $t > s$.
2. Si $\mathcal{H}^s(E) > 0$ entonces $\mathcal{H}^t(E) = \infty$ para todo $t \in (0, s)$.

Deducir que para cada $E \subset \mathbb{R}^n$ existe un único $s = s(E) \in [0, \infty]$ tal que $\mathcal{H}^t(E) = 0$ para $t > s$ y $\mathcal{H}^t(E) = \infty$ para $0 < t < s$.

[3] Como hemos dicho antes, probaremos más adelante que $\lambda_n = 1$.

Problema 3.5. Demostrar que $\dim_H \left(\bigcup_{k=1}^{\infty} A_k\right) = \sup_{k\in\mathbb{N}} \dim_H (A_k)$. Deducir que si A es numerable entonces $\dim_H (A) = 0$.

Problema 3.6. Probar que $\mathcal{H}^n(A) = \inf\{\mathcal{H}^n(U) : U$ abierto, $A \subset U\}$ para todo $A \subset \mathbb{R}^n$ (es decir, la medida de Hausdorff es exteriormente regular). Suponiendo que $\mathcal{H}^n = \mathcal{L}_n$ en $\mathcal{B}(\mathbb{R}^n)$, deducir que $\mathcal{H}^n = \mathcal{L}_n^*$ en $\mathcal{P}(\mathbb{R}^n)$.

Problema 3.7. Probar que $\mathcal{L}_n^*(A) = \mathcal{L}_n^*(x + A)$ para todo $x \in \mathbb{R}^n$, $A \subseteq \mathbb{R}^n$, y que $\mathcal{L}_n^*(rA) = r^n \mathcal{L}_n^*(A)$ si $r > 0$.

Problema 3.8. Probar que si $n \leq m$, $f : \mathbb{R}^n \to \mathbb{R}^m$ es de Lipschitz, y $\mathcal{L}_n(A) = 0$, entonces $\mathcal{L}_m(f(A)) = 0$.

Problema 3.9. Demostrar que si U es un abierto de $\mathbb{R}^n$, $n \leq m$, y $f \in C^1(U, \mathbb{R}^m)$, entonces $f(A)$ tiene medida cero en $\mathbb{R}^m$ para todo $A \subset U$ de medida cero en $\mathbb{R}^n$. Probar también que si $n < m$ entonces $\mathcal{L}_m(f(U)) = 0$.

Problema 3.10. ¿Son ciertos los enunciados de los dos problemas anteriores si se supone $n > m$?

Problema 3.11. Si $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ es de Lipschitz y $\mathcal{H}^n(A) > 0$, probar que la dimensión de Hausdorff de la gráfica de f en $\mathbb{R}^n \times \mathbb{R}^m \equiv \mathbb{R}^{n+m}$ es n.
Indicación: $G_f = g(A)$, donde $g(x) = (x, f(x))$ es Lipschitz.

Problema 3.12. Sean $s \in (0, n]$ y $M \subset \mathbb{R}^n$ un conjunto cerrado tal que $\mathcal{H}^s(K) < \infty$ para todo compacto $K \subset M$. Demostrar que $\mathcal{H}^s(A) = \sup\{\mathcal{H}^s(C) : C \subset A,\, C$ cerrado$\}$ para todo boreliano $A \subset M$.

Problema 3.13. Demostrar que el resultado anterior se aplica con $s = k$ a cualquier subvariedad diferenciable M de dimensión k de $\mathbb{R}^n$. Más en general, también es cierto si $M = \bigcup_{j=1}^{\infty} f_j(C_j)$, donde $f_j : C_j \subset \mathbb{R}^k \to \mathbb{R}^n$ es de Lipschitz, C_j es abierto acotado de $\mathbb{R}^k$, y $f(C_j)$ es abierto relativo a M para todo j.

Problema 3.14. Sea $V \subset \mathbb{R}$ el conjunto no medible de Vitali. Demostrar que, para todo $A \subset V$, si A es medible entonces $\mathcal{L}_1(A) = 0$.
Indicación: Sea $\{r_n\}_{n\in\mathbb{N}} = (0, 1) \cap \mathbb{Q}$, y considerar los conjuntos disjuntos $B_n = r_n + V$, $n \in \mathbb{N}$.

Problema 3.15. Demostrar que si A es un subconjunto de $\mathbb{R}$ con $\mathcal{L}_1^*(A) > 0$, entonces A contiene un subconjunto no medible.
Indicación: Suponer primero que $A \subseteq (0, 1)$. Considerar $A_r = A \cap V_r$, donde $\{V_r\}$ es la familia de conjuntos considerados en la demostración de la existencia de un conjunto no medible $V \subset (0, 1)$. Si cada A_r es medible, por el ejercicio anterior se tiene $|A_r| = 0$. Pero por otra parte $0 < \mathcal{L}_1^*(A) \leq \sum_r \mathcal{L}_1^*(A_r)$.

Problema 3.16. Usar las propiedades de la medida exterior de Lebesgue para demostrar que, para cualesquiera $a, b \in \mathbb{R}$ con $a < b$, el intervalo $[a, b]$ es un conjunto no numerable.

Problema 3.17. Demostrar que si $A \subset \mathbb{R}^n$ es medible Lebesgue y $\mathcal{L}_n(A) > 0$ entonces existe $K \subset A$ compacto tal que $\mathcal{L}_n(K) > 0$.

Problema 3.18. Sea $E \subset \mathbb{R}$ medible Lebesgue y con medida positiva. Demostrar que para todo $\alpha \in (0, 1)$ existe un intervalo abierto I tal que $|E \cap I| > \alpha |I|$.
Indicación: aproximar E por un abierto adecuado, y escribir dicho abierto como unión numerable y disjunta de intervalos abiertos.

Problema 3.19. Demostrar que si $\mathcal{L}_1(E) > 0$ entonces $E - E := \{x - y : x, y \in E\}$ contiene un intervalo abierto centrado en 0.
Indicación: usar el ejercicio anterior.

Problema 3.20. Sea $\{r_n\}_{n\in\mathbb{N}}$ una enumeración de $\mathbb{Q}$. Demostrar que existen infinitos números irracionales x que no cumplen $|x - r_n| < 1/n^2$ para ningún n.

Problema 3.21. Demostrar que el conjunto $\{x \in [0, 1] : x$ no tiene el dígito 5 en su desarrollo decimal$\}$ tiene medida cero.

Problema 3.22. Dar un ejemplo de conjunto $A \subset \mathbb{R}$ que no sea medible en $\mathbb{R}$ y que $A\times\{y\}$ sea medible en $\mathbb{R}^2$ para todo $y \in \mathbb{R}$.

Problema 3.23. Demostrar que si $A \subset \mathbb{R}$ es medible Lebesgue entonces $\{(x, y) \in \mathbb{R}^2 : x \in A\}$ es medible Lebesgue en $\mathbb{R}^2$.

Problema 3.24. Sea $(X, \mathcal{A}, \mu)$ un espacio de medida, y sea $\mathcal{A}^*$ la colección de todos los $E \subset X$ para los cuales existen $A, B \in \mathcal{A}$ tales que $A \subset E \subset B$ y $\mu(B \setminus A) = 0$. Abusando de notación, definamos $\mu(E) = \mu(A)$ en tal situación. Probar que μ está bien definida y $\mathcal{A}^*$ es una σ-álgebra, sobre la cual esta extensión de μ es una medida completa.

Problema 3.25. Demostrar que la medida de Lebesgue $\mathcal{L}_n$ restringida a $\mathcal{B}(\mathbb{R}^n)$ no es completa, y que su compleción es igual $\mathcal{L}_n$ definida en la σ-álgebra de los conjuntos medibles Lebesgue.

Problema 3.26 (Construcción de conjuntos de Cantor)**.** Sea $(m_n)_{n\in\mathbb{N}}$ una sucesión estrictamente decreciente de números reales positivos tal que $m_0 = 1$. Para cada n se denota $\ell_n = 2^{-n}m_n$. Definamos $C_0 = [0, 1]$, y por inducción, supuesto definido C_n como una unión de 2^n intervalos cerrados disjuntos de longitud ℓ_n, se define C_{n+1} como la unión de los 2^{n+1} intervalos cerrados disjuntos de longitud ℓ_{n+1} que quedan al eliminar, de cada uno de los intervalos de C_n, un invervalo abierto con el mismo centro que dicho intervalo y longitud $\ell_n - 2\ell_{n+1} = (m_n - m_{n+1})2^{-n}$. Definamos entonces $C = \bigcap_{n=0}^{\infty} C_n$. A C se le llama conjunto (generalizado) de Cantor. El caso particular en que $m_n = (2/3)^n$ se le llama conjunto ternario (o clásico) de Cantor. Demostrar:

1. $C_n \subset C_{n-1}$ para todo $n \in \mathbb{N}$.
2. C es un compacto no vacío, tiene interior vacío, y no tiene puntos aislados.
3. $\mathcal{L}_1(C) = \lim_{n\to\infty} m_n$.

En los casos en que $\lim_{n\to\infty} m_n > 0$, diremos que C es un conjunto de Cantor gordo.

Problema 3.27. Observar que el conjunto ternario de Cantor es el conjunto C de los números reales del intervalo $[0, 1]$ cuyo desarrollo en base 3 sólo tiene ceros y doses. Concluir que C es no numerable.

Problema 3.28. Para cada conjunto de Cantor C del ejercicio 3.26, construir una función estrictamente creciente, continua y biyectiva $f : [0, 1] \to [0, 1]$ tal que $f(C) = K$, donde K es el conjunto ternario de Cantor. Deducir que todos los conjuntos de Cantor son no numerables y homeomorfos entre sí.

Problema 3.29. Sea C el conjunto de Cantor asociado a la sucesión $m_n = (2\lambda)^n$, donde $0 < \lambda < 1/2$. Demostrar que $\dim_H(C) = -\log 2/\log\lambda$. [Por tanto hay conjuntos con dimensión de Hausdorff igual a cualquier número real del intervalo $(0, 1)$.]
Indicación: De $C \subset C_n$ deducir que $\mathcal{H}^p_{\lambda^n} \leq (2\lambda^p)^n$; eligiendo p de forma que $2\lambda^p = 1$ se obtiene $\mathcal{H}^p(C) \leq 1$. Por otro lado, si $(I_n)_{n\in\mathbb{N}}$ es un recubrimiento de C por intervalos abiertos, por compacidad puede extraerse un recubrimiento finito: $C \subset I_1 \cup ... \cup I_m$. Para cada $k \in \{1, ..., m\}$, sea $n_k \in \mathbb{N}$ tal que $\lambda^{n_k} \leq (1 - 2\lambda)^{-1}\mathcal{L}^1(I_k) < \lambda^{n_k-1}$; puesto que los intervalos de C_n distan entre sí al menos $(1 - 2\lambda)\lambda^{n-1}$, $\overline{I_k}$ corta a lo sumo a un intervalo de C_{n_k}. Luego, si $j \geq n_k$, $\overline{I_k}$ corta a lo sumo 2^{j-n_k} intervalos de C_j. Tomando j suficientemente grande para que $(1 - 2\lambda)\lambda^j \leq \mathcal{L}^1(I_k)$ para todo $k \in \{1, ..., m\}$, deducir que

$$2^j \leq \sum_{k=1}^{m} 2^{j-n_k} = \sum_{k=1}^{m} 2^j\lambda^{pn_k} \leq 2^j(1 - 2\lambda)^{-p}\sum_{k=1}^{m} \mathcal{L}^1(I_k)^p,$$

y concluir que $0 < (1 - 2\lambda)^p \leq \mathcal{H}^p(C)$.

Capítulo 4

Funciones medibles

Así como en general no todo conjunto puede medirse, tampoco toda función puede integrarse. En este capítulo definiremos la clase natural de funciones que luego podremos integrar, y estudiaremos sus propiedades más relevantes.

4.1. Funciones medibles

Definición 4.1. Sean $(X, \mathcal{A})$ un espacio medible, e Y un espacio métrico. Diremos que una función $f : X \to Y$ es *medible* si

$$f^{-1}(U) \in \mathcal{A} \quad \text{para todo abierto } U \subset Y.$$

Puesto que $f^{-1}(Y \setminus U) = X \setminus f^{-1}(U)$ y los complementarios de los abiertos de Y son los cerrados de Y, es obvio que una función es medible si y sólo si la imagen inversa de cualquier cerrado es medible.

Definición 4.2. Si $(X, \mathcal{A})$ es un espacio medible, Y es un espacio métrico y $E \in \mathcal{A}$, se dice que una función $f : E \to Y$ es medible si $f^{-1}(U) \in \mathcal{A}$ para cada abierto $U \subset Y$.

Ejemplo 4.1. Si $f : \mathbb{R}^n \to \mathbb{R}^m$ es continua, entonces f es medible (tanto respecto de la σ-álgebra de los borelianos de $\mathbb{R}^n$ como con respecto a la σ-álgebra de los conjuntos medibles Lebesgue). En efecto, para todo $U \subset \mathbb{R}^m$, por ser f continua, el conjunto $f^{-1}(U)$ es abierto, y en particular medible en $\mathbb{R}^n$ (con respecto a estas σ-álgebras).

Ejemplo 4.2. Recordemos que la función característica de un conjunto $A \subset X$ es la función $\chi_A : X \to [0, 1]$ definida por

$$\chi_A(x) := \begin{cases} 1 & \text{si } x \in A, \\ 0 & \text{si } x \notin A. \end{cases}$$

Si $(X, \mathcal{A})$ es un espacio medible, es fácil ver que $\chi_A : X \to \mathbb{R}$ es medible si y sólo si $A \in \mathcal{A}$.

Proposición 4.1. *Sean $(X, \mathcal{A})$ un espacio medible, Y un espacio métrico, $E \in \mathcal{A}$ no vacío, y $f : E \to Y$ una función. Las siguientes afirmaciones son equivalentes:*

1. *$f : E \to Y$ es medible;*
2. *Existe $\widetilde{f} : X \to Y$ medible tal que $\widetilde{f} = f$ en E;*
3. *$f : E \to Y$ es medible respecto de la σ-álgebra $\mathcal{A}_E := \{A \subset E : A \in \mathcal{A}\} = \{B \cap E : B \in \mathcal{A}\}$.*

Demostración. $(1) \implies (2)$: dado $y_0 \in Y$, definamos

$$\widetilde{f}(x) = \begin{cases} f(x) & \text{si } x \in E \\ y_0 & \text{si } x \in X \setminus E. \end{cases}$$

Entonces, para todo $U \subset Y$ abierto se tiene que

$$\widetilde{f}^{-1}(U) = \{x \in E : f(x) \in U\} \cup \{x \in X \setminus E : y_0 \in U\} = \begin{cases} f^{-1}(U) & \text{si } y_0 \notin U \\ f^{-1}(U) \cup (X \setminus E) & \text{si } y_0 \in U, \end{cases}$$

y como $f^{-1}(U) \in \mathcal{A}$ por ser f medible, y $X \setminus E \in \mathcal{A}$, es obvio que $\widetilde{f}^{-1}(U) \in \mathcal{A}$.

(2) $\implies$ (3): es fácil comprobar que $\mathcal{A}_E$ es una σ-álgebra. Si $\widetilde{f} = f$ en E y $\widetilde{f} : X \to Y$ es medible, entonces, para todo U abierto es $\widetilde{f}^{-1}(U) \in \mathcal{A}$. Luego

$$f^{-1}(U) = \{x \in E : f(x) \in U\} = \{x \in E : \widetilde{f}(x) \in U\} = \widetilde{f}^{-1}(U) \cap E \in \mathcal{A}_E,$$

y esto prueba que f es $\mathcal{A}_E$-medible.

(3) $\implies$ (1): si $U \subset Y$ es abierto entonces por hipótesis $f^{-1}(U) \in \mathcal{A}_E$, y como $\mathcal{A}_E \subset \mathcal{A}$ (por ser $E \in \mathcal{A}$), se tiene $f^{-1}(U) \in \mathcal{A}$. □

Proposición 4.2. *Sean $(X, \mathcal{A})$ un espacio medible, $f : X \to Y$ una función medible, y $g : Y \to Z$ una función continua. Entonces la función $g \circ f : X \to Z$ es medible.*[1]

Demostración. Para cada U abierto de Z, $g^{-1}(U)$ es abierto de Y (por ser g continua), luego, al ser f medible $f^{-1}(g^{-1}(U)) \in \mathcal{A}$. Por tanto

$$(g \circ f)^{-1}(U) = f^{-1}(g^{-1}(U)) \in \mathcal{A},$$

y así $g \circ f$ es medible. □

Proposición 4.3. *Sean $(X, \mathcal{A})$ un espacio medible, e Y un espacio métrico. Si $u, v : X \to \mathbb{R}$ son funciones medibles y $\Phi : \mathbb{R}^2 \to Y$ es continua, entonces la función $h : X \to Y$ definida por*

$$h(x) = \Phi(u(x), v(x))$$

es medible.

Demostración. Definamos $f : X \to \mathbb{R}^2$ por $f(x) = (u(x), v(x))$. Bastará probar que esta función es medible, porque entonces por la Proposición 4.2 tendremos que $h = \Phi \circ f$ será medible. Sea $R = (a, b) \times (c, d)$ un rectángulo abierto en $\mathbb{R}^2$. Se tiene

$$f^{-1}(R) = u^{-1}((a, b)) \cap v^{-1}((c, d)) \in \mathcal{A}.$$

Entonces, si U es un abierto de $\mathbb{R}^2$, U es unión (no necesariamente disjunta) de rectángulos abiertos, digamos $U = \bigcup_{j=1}^{\infty} R_j$, y por tanto

$$f^{-1}(U) = \bigcup_{j=1}^{\infty} f^{-1}(R_j) \in \mathcal{A}.$$

Esto completa la prueba. □

La proposición anterior tiene bastantes aplicaciones: por ejemplo, deduciremos fácilmente que las combinaciones algebraicas más comunes (suma, producto, etc) de funciones medibles dan lugar a funciones medibles.

Proposición 4.4.

(a) Si $f = u + iv : X \to \mathbb{C}$, donde $u, v : X \to \mathbb{R}$ son medibles, entonces f es medible.

[1] Aquí es muy importante el orden: la función que actúa en segundo lugar debe ser la continua. La composición de una función continua seguida de una medible no tiene por qué ser medible; ver el problema 4.10.

(b) Si $f = u + iv : X \to \mathbb{C}$ es medible, entonces las funciones $u, v, |f| : X \to \mathbb{R}$ son medibles.

(c) Si $f, g : X \to \mathbb{R}$ (o con valores complejos) son medibles, entonces las funciones $f + g$ y fg son medibles.

(d) Un conjunto $E \subset X$ es medible si y sólo si su función característica,

$$\chi_E(x) = \begin{cases} 1 & \text{si } x \in E \\ 0 & \text{si } x \notin E \end{cases},$$

es medible.

(e) Si $f : X \to \mathbb{C}$ es medible entonces existe $\alpha : X \to X$ tal que $|\alpha| = 1$ y $f = \alpha|f|$.

Demostración. (a): basta tomar $\Phi(u, v) = u + iv$ y aplicar la proposición anterior.

(b): puede tomarse $\Phi_1(u+iv) = u$, $\Phi_2(u+iv) = v$, $\Phi_3(u+iv) = |u+iv|$ y aplicar la proposición anterior.

(c): puede tomarse $\Phi_1(z_1, z_2) = z_1 + z_2$ y $\Phi_2(z_1, z_2) = z_1 z_2$ y aplicar la proposición anterior.

(d): esto se sigue del hecho de que

$$(\chi_E)^{-1}(U) = \begin{cases} \emptyset & \text{si } 0 \notin U, 1 \notin U \\ E & \text{si } 0 \notin U, 1 \in U \\ X \setminus E & \text{si } 0 \in U, 1 \notin U \\ X & \text{si } 0 \in U, 1 \in U \end{cases}$$

(e): Sea $E = f^{-1}(\{0\})$, que es medible por serlo f. Entonces la función $h(x) := f(x) + \chi_E(x)$ es medible, y $h(x) \neq 0$ para todo $x \in X$. Luego $h : \mathbb{C} \to \mathbb{C} \setminus \{0\}$ es medible, y como $\varphi : \mathbb{C} \setminus \{0\} \to \mathbb{C}$, $\varphi(z) = z/|z|$ es continua, se sigue de la Proposición 4.2 que

$$\alpha(x) := \varphi\left(f(x) + \chi_E(x)\right)$$

es medible. Para $x \in E$ es $\alpha(x) = 1$, y si $x \notin E$ entonces $\alpha(x) = f(x)/|f(x)|$. Por tanto $f = \alpha|f|$. $\square$

Definición 4.3. Sean X, Y espacios métricos. Se dice que $f : X \to Y$ es de Borel si es medible con respecto a la σ-álgebra de los borelianos de X; es decir, si $f^{-1}(U) \in \mathcal{B}(X)$ para todo U abierto de Y.

Es obvio que toda función continua es de Borel.

Proposición 4.5. *Sean $(X, \mathcal{A})$ un espacio medible, Y un espacio métrico y $f : X \to Y$ una función medible.*

(a) Si $E \subset Y$ es boreliano, entonces $f^{-1}(E) \in \mathcal{A}$.

(b) Si Z es otro espacio métrico y $g : Y \to Z$ es de Borel, entonces $g \circ f : X \to Z$ es medible.

Demostración. (a): Sea $\mathcal{M} = \{A \subset Y : f^{-1}(A) \in \mathcal{A}\}$. Por definición de función medible, $\mathcal{M}$ contiene a todos los abiertos de Y. Bastará probar que $\mathcal{M}$ es una σ-álgebra, porque entonces se tendrá, por definición de $\mathcal{B}(Y)$, que $\mathcal{B}(Y) \subset \mathcal{M}$, que es justo lo que queremos ver. Y en efecto se tiene que: 1) $Y \in \mathcal{M}$, ya que $f^{-1}(Y) = X \in \mathcal{A}$; 2) si $A \in \mathcal{M}$ entonces $f^{-1}(A) \in \mathcal{A}$, luego $X \setminus f^{-1}(A) \in \mathcal{A}$, y como $f^{-1}(Y \setminus A) = X \setminus f^{-1}(A)$, se obtiene que $Y \setminus A \in \mathcal{M}$; y 3) si $\{A_j\}_{j\in\mathbb{N}} \subset \mathcal{M}$ entonces $f^{-1}\left(\bigcup_{j=1}^{\infty} A_j\right) = \bigcup_{j=1}^{\infty} f^{-1}(A_j) \in \mathcal{A}$, por ser $\mathcal{A}$ una σ-álgebra, y así $\bigcup_{j=1}^{\infty} A_j \in \mathcal{M}$.

(b): si U es un abierto de Z entonces por ser g de Borel $g^{-1}(U) \in \mathcal{B}(Y)$, luego por (a) se tiene $(g \circ f)^{-1}(U) = f^{-1}(g^{-1}(U)) \in \mathcal{A}$. $\square$

Definición 4.4. Sea $(X, \mathcal{A})$ un espacio medible. Diremos que una función $f : X \to [-\infty, \infty]$ es medible si $f^{-1}(U)$ es medible para cada $U \subset \mathbb{R}$ abierto y además los conjuntos $f^{-1}(\{+\infty\})$ y $f^{-1}(\{-\infty\})$ son medibles.

Proposición 4.6. *Sea $(X, \mathcal{A})$ un espacio medible. Una función $f : X \to [-\infty, \infty]$ es medible si y sólo si $f^{-1}((t, \infty]) \in \mathcal{A}$ para todo $t \in \mathbb{R}$.*

Demostración. Es fácil y se propone como ejercicio. □

4.2. Límites de funciones medibles

Sea $(a_n)_{n\in\mathbb{N}} \subset [-\infty, \infty]$. Recordemos que

$$\limsup_{n\to\infty} a_n = \lim_{k\to\infty} \sup\{a_n : n \geq k\}.$$

Obsérvese que $b_k := \sup\{a_n : n \geq k\}$ define una sucesión decreciente, de forma que

$$\limsup_{n\to\infty} a_n = \lim_{k\to\infty} b_k = \inf_{k\in\mathbb{N}}\{b_k\}.$$

Este valor es el supremo de todos los posibles valores de límites de subsucesiones de (a_n), y es igual al límite de alguna de ellas. Análogamente se define

$$\liminf_{n\to\infty} a_n = \lim_{k\to\infty} \inf\{a_n : n \geq k\} = \sup_{k\in\mathbb{N}} \inf_{n\geq k} a_n,$$

valor que es el ínfimo de todos los posibles límites de subsucesiones de (a_n), y que es igual al límite de alguna de ellas. Es fácil ver que

$$\liminf_{n\to\infty} a_n = -\limsup_{n\to\infty}(-a_n).$$

Recordemos también que el límite $\lim_{n\to\infty} a_n$ existe si y sólo si existen $\limsup_{n\to\infty} a_n$ y $\liminf_{n\to\infty} a_n$ y son iguales (y en ese caso son iguales también a $\lim_{n\to\infty} a_n$).

Si $f_n : X \to [-\infty, \infty]$ es una sucesión de funciones, se definen las funciones $\sup_{n\in\mathbb{N}} f_n$, $\inf_{n\in\mathbb{N}} f_n$, $\limsup_{n\to\infty} f_n$, $\liminf_{n\to\infty} f_n : X \to [-\infty, \infty]$ por

$$\left(\sup_{n\in\mathbb{N}} f_n\right)(x) = \sup_{n\in\mathbb{N}} f_n(x), \quad \left(\inf_{n\in\mathbb{N}} f_n\right)(x) = \inf_{n\in\mathbb{N}} f_n(x),$$

$$\left(\limsup_{n\to\infty} f_n\right)(x) = \limsup_{n\to\infty} f_n(x), \quad \left(\liminf_{n\to\infty} f_n\right)(x) = \liminf_{n\to\infty} f_n(x)$$

para cada $x \in X$.

Si $\lim_{n\to\infty} f_n(x) := f(x)$ existe para cada $x \in X$, diremos que la sucesión f_n converge puntualmente a f.

Teorema 4.7. *Sea $f_n : X \to [-\infty, \infty]$ una sucesión de funciones medibles. Entonces las funciones $\sup_{n\in\mathbb{N}} f_n$ $\inf_{n\in\mathbb{N}} f_n$, $\limsup_{n\to\infty} f_n$ y $\liminf_{n\to\infty} f_n : X \to [-\infty, \infty]$ son medibles.*

Demostración. Para todo $t \in \mathbb{R}$ se tiene

$$\left(\sup_{n\in\mathbb{N}} f_n\right)^{-1}((t, \infty]) = \bigcup_{n=1}^{\infty} f_n^{-1}((t, \infty]) \in \mathcal{A},$$

luego por la proposición anterior $\sup_{n\in\mathbb{N}} f_n$ es medible. Entonces

$$\inf_{n\in\mathbb{N}} f_n = -\sup_{n\in\mathbb{N}}(-f_n)$$

es también medible. Y por tanto, combinando estos dos hechos,

$$\limsup_{n\to\infty} f_n = \inf_{k\in\mathbb{N}}\{\sup_{j\geq k} f_j\}$$

es medible. Se deduce entonces también que

$$\liminf_{n\to\infty} f_n = -\limsup_{n\to\infty}(-f_n)$$

es medible. $\square$

Corolario 4.8. *Si $f_n : X \to [-\infty,\infty]$ es una sucesión de funciones medibles que converge puntualmente a f, entonces $f : X \to [-\infty,\infty]$ es medible.*

En particular, si $g_n : X \to [0,\infty]$ es medible para todo n entonces $\sum_{n=1}^{\infty} g_n : X \to [0,\infty]$ es medible

Corolario 4.9. *Si $f, g : X \to [-\infty,\infty]$ son medibles entonces también lo son las funciones*

$$\text{máx}\{f,g\},\ \ \text{mín}\{f,g\},$$

$$f^+ = \text{máx}\{f,0\} \ \ y \ \ f^- = -\text{mín}\{f,0\}.$$

La función f^+ se llama la *parte positiva* de f, y f^- es la *parte negativa* de f. Nótese que $f^+, f^- \geq 0$, $f = f^+ - f^-$, y $|f| = f^+ + f^-$.

El Corolario 4.8 se puede generalizar al caso de funciones con valores en espacios métricos como sigue.

Teorema 4.10. *Sean $(X,\mathcal{A})$ un espacio medible, Y un espacio métrico, y $f_n : X \to Y$ una sucesión de funciones medibles. Supongamos que $f_n(x)$ converge a $f(x)$ para cada $x \in X$. Entonces la función $f : X \to Y$ es medible.*

Demostración. Sea U un abierto de Y. Tenemos que probar que $f^{-1}(U) \in \mathcal{A}$. Para cada $m \in \mathbb{N}$, definamos

$$F_m = \{y \in U : d(y, X\setminus U) \geq 1/m\}.$$

Es fácil ver que F_m es cerrado, y

$$U = \bigcup_{m=1}^{\infty} F_m.$$

Por consiguiente, si $f(x) \in U$ entonces existe m tal que $f(x) \in F_m$, y como $f_n(x)$ converge a $f(x)$ existe $k \in \mathbb{N}$ tal que $d(f_n(x), f(x)) < 1/2m$ para todo $n \geq k$, lo que implica que $f_n(x) \in F_{2m}$ si $n \geq k$. Por otro lado, si existen ℓ y j tal que $f_n(x) \in F_\ell$ para todo $n \geq j$, entonces por ser F_ℓ cerrado se tiene $f(x) \in F_\ell \subset U$. Esto prueba que

$$f^{-1}(U) = \bigcup_{m\in\mathbb{N}} \bigcup_{k\in\mathbb{N}} \bigcap_{n\geq k} f_n^{-1}(F_m),$$

y como el conjunto a la derecha de la igualdad está en $\mathcal{A}$ (por ser cada f_n medible y F_m cerrado), tenemos lo que necesitamos. $\square$

Definición 4.5. Sea $(X,\mathcal{A})$ un espacio medible. Una función $s : X \to \mathbb{R}$ (o con valores en $\mathbb{C}$) se dice que es simple medible si es medible y toma un número finito de valores. Si dichos valores son $\alpha_1, ..., \alpha_n$ entonces

$$s = \sum_{j=1}^{n} \alpha_j \chi_{A_j},$$

donde los conjuntos $A_j = s^{-1}(\{\alpha_j\})$ son disjuntos dos a dos y medibles. Recíprocamente, si s puede representarse de esta forma (con los conjuntos A_j disjuntos dos a dos y medibles), entonces s es simple medible.

Aquí, de forma explícita o implícita, uno de los α_j puede ser igual a cero. Más precisamente, si $s = 0$ entonces sólo hay un $\alpha_j = \alpha_1$ que es cero, con $A_1 = s^{-1}(0) = X$. Si $s \neq 0$ pero $s(x) = 0$ para algún x, entonces s toma al menos dos valores, uno de los cuales es cero, y podremos escribir $s = \sum_{j=1}^{n} \alpha_j \chi_{A_j} = \sum_{j=0}^{n} \alpha_j \chi_{A_j}$, con $\alpha_0 = 0$ y $A_0 = s^{-1}(\{0\})$. Es decir, el valor cero y su imagen inversa pueden aparecer (o no) explícitamente en la suma que define a s, según nos convenga (o no), cuando la función toma al menos dos valores y uno es nulo. Obviamente el cero es el único valor para el que tenemos esta libertad. El resto de valores de s y sus imágenes inversas siempre deben aparecer explícitamente en la suma que define a s.

Teorema 4.11. *Sea $(X, \mathcal{A})$ un espacio medible. Si $f : X \to [0, \infty]$ es medible, entonces existe una sucesión (s_n) de funciones simples medibles en X tales que.*

1. $0 \leq s_n \leq s_{n+1}$ *para todo* $n \in \mathbb{N}$, *y*
2. $\lim_{n\to\infty} s_n(x) = f(x)$ *para todo* $x \in X$.

Además s_n converge uniformemente a f en cualquier subconjunto de X donde f esté acotada. En particular, si f es acotada, $\lim_{n\to\infty} s_n = f$ uniformemente en X.

Demostración. Consideremos primero el caso en que $f = \varphi : [0, \infty] \to [0, \infty]$ es la identidad, $\varphi(t) = t$ para todo $t \geq 0$. Entonces es fácil ver que la sucesión

$$\varphi_n(t) = \begin{cases} n & \text{si } t \geq n, \\ \frac{k}{2^n} & \text{si } \frac{k}{2^n} \leq t < \frac{k+1}{2^n} \leq n \end{cases}$$

tiene las propiedades del enunciado. En el caso general, sea $s_n = \varphi_n \circ f$, es decir,

$$s_n(x) = \begin{cases} n & \text{si } f(x) \geq n, \\ \frac{k}{2^n} & \text{si } \frac{k}{2^n} \leq f(x) < \frac{k+1}{2^n} \leq n. \end{cases}$$

Es claro que $s_n = \sum_{j=1}^{n2^n} \alpha_j \chi_{A_j}$, con $\{\alpha_1, ..., \alpha_n\} = \{1/2^n, ..., n\}$ donde los conjuntos $A_1, ..., A_{n2^n} = f^{-1}([1/2^n, 2/2^n), ..., f^{-1}([n, \infty])$ son medibles y disjuntos dos a dos; por tanto s_n es simple medible. Además, puesto que φ_n converge a φ puntualmente, se obtiene inmediatamente que $s_n = \varphi_n \circ f$ converge puntualmente a $\varphi \circ f = f$. De hecho se tiene $0 \leq f(x) - s_n(x) \leq 1/2^n$ si $f(x) \leq n$, y por tanto la convergencia es uniforme en cualquier conjunto donde f esté acotada. $\square$

4.3. Problemas

Problema 4.1. Sean $f, g : X \to Y$ medibles, $A \subset X$ medible, y definamos $h : X \to Y$ por $h(x) = f(x)$ si $x \in A$, y $h(x) = g(x)$ si $x \in X \setminus A$. Comprobar que h es medible.

Problema 4.2. Demostrar que una función $f = (f_1, ..., f_m) : X \to \mathbb{R}^m$ es medible si y sólo si $f_j : X \to \mathbb{R}$ es medible para cada $j = 1, ..., n$.

Problema 4.3. Demostrar que $f : X \to [-\infty, \infty]$ es medible (en el sentido de que $f_{|_{f^{-1}(\mathbb{R})}}$ es una función medible y $f^{-1}(\{\infty\})$, $f^{-1}(\{-\infty\})$ son conjuntos medibles) si y sólo si $f^{-1}((t, +\infty])$ es medible para cada $t \in \mathbb{R}$.

Problema 4.4. Sean X un espacio medible, y (f_n) una sucesión de funciones medibles de X en $[-\infty, +\infty]$. Demostrar que el conjunto $\{x \in X : \exists \lim_{n\to\infty} f_n(x)\}$ es medible.

Problema 4.5. Demostrar que toda función monótona $f : \mathbb{R} \to \mathbb{R}$ es medible Borel.

Problema 4.6. Demostrar que si $f : \mathbb{R}^n \to \mathbb{R}$ es continua en casi todo punto entonces f es medible Lebesgue.

Problema 4.7. Sea $f : \mathbb{R} \to \mathbb{R}$ derivable. Demostrar que f' es medible Borel.

Problema 4.8 (Construcción de funciones de Cantor)**.** Con la notación del ejercicio 3.26, para cada n definamos $f_n : [0,1] \to [0,1]$ por $f_n(x) = \mathcal{L}_1 \left(C_n \cap [0,x]\right)/m_n$. Representar gráficamente f_1 y f_2 en el caso $m_1 = 2/3$ y $m_2 = (2/3)^2$. Probar lo siguiente:

1. Si $0 \leq x < y \leq 1$ entonces $0 \leq f_n(y) - f_n(x) \leq (y-x)/m_n$, y en particular f_n es creciente y $1/m_n$-Lipschitz para cada n.
2. En cada uno de los intervalos abiertos disjuntos que forman $[0,1] \setminus C_n$, f_n es constante; en particular $f_n'(x) = 0$ para todo $x \in [0,1] \setminus C_n$.
3. Si (a,b) es el k-ésimo intervalo abierto de $[0,1] \setminus C_n$, entonces, para todo $x \in (a,b)$, $f_n(x) = f_n(a) = k/2^n$.
4. $f_{n+1}(x) = f_n(x)$ si $x \notin C_n$.
5. $|f_n(x) - f_{n+1}(x)| \leq 2^{-n}$ si $x \in C_n$.
6. Para todo $x \in [0,1]$, $|f_n(x) - f_{n+p}(x)| \leq 1/2^{n-1}$.
7. (f_n) converge uniformemente a una función creciente $f : [0,1] \to [0,1]$ tal que $f(0) = 0$, $f(1) = 1$, y $f'(x) = 0$ para todo $x \in [0,1] \setminus C$.

A f se le llama función de Cantor.

Problema 4.9 (Existencia de conjuntos medibles Lebesgue que no son borelianos)**.** Sean C un conjunto de Cantor gordo, y $f : [0,1] \to [0,1]$ la función del problema 3.28 tal que $f(C) = K$. Por el Problema 3.15, existe $A \subset C$ no medible. Probar que $B = f(A)$ es medible, pero no boreliano, y $f^{-1} : [0,1] \to [0,1]$ es un homeomorfismo que lleva un conjunto medible en otro no medible.

Problema 4.10. Sea $f : [0,1] \to [0,1]$ la función de Cantor correspondiente al conjunto ternario, y sea $g : [0,1] \to [0,2]$ definida por $g(x) = x + f(x)$. Demostrar: g es un homeomorfismo, $\mathcal{L}^1(g(C)) = 1$, existe $A \subset g(C)$ no medible Lebesgue, y $B = g^{-1}(A)$ es medible Lebesgue pero no boreliano. Luego $\chi_A = \chi_B \circ g$ no es medible Lebesgue, pese a que χ_B sí lo es y g es continua.

Problema 4.11. Si $f, g : X \to \mathbb{R}$ son medibles, demostrar que los conjuntos $\{x \in X : f(x) < g(x)\}$, $\{x \in X : f(x) \leq g(x)\}$ y $\{x \in X : f(x) = g(x)\}$ son medibles.

Problema 4.12. Dar un ejemplo de una función $f : \mathbb{R} \to \mathbb{R}$ continua en casi todo punto pero que no coincida en casi todo punto con ninguna función continua. Dar otro ejemplo de una función $g : \mathbb{R} \to \mathbb{R}$ que coincida en casi todo punto con una función continua pero que no sea continua en casi todo punto.

Problema 4.13. Dar un ejemplo de una función $f : [0,1] \to \mathbb{R}$ no medible tal que $f^{-1}(\{t\})$ sea medible para todo $t \in \mathbb{R}$.

Problema 4.14. Sea $(X, \mathcal{A}, \mu)$ un espacio de medida, y supongamos que μ es completa. Demostrar que si $f = g$ en casi todo punto y f es medible, entonces g también lo es. Si (φ_n) es una sucesión de funciones medibles que converge a una función φ en casi todo punto, demostrar que φ es medible. Dar ejemplos que prueben que estas afirmaciones son falsas si μ no es completa.

Problema 4.15. Sean $(X, \mathcal{A}, \mu)$ espacio de medida σ-finito, y $f : X \to \mathbb{R}$ medible. Demostrar que $\{t \in \mathbb{R} : \mu(f^{-1}(\{t\})) > 0\}$ es a lo sumo numerable.

Problema 4.16. Construir un conjunto boreliano $E \subset [0,1]$ tal que para todo subintervalo abierto no vacío $I \subset [0,1]$ se tiene $|E \cap I| > 0$ y $|E^c \cap I| > 0$ (o, dicho de otra forma, $0 < |E \cap I| < |I|$). *Indicación:* considerar un conjunto de Cantor gordo en $[0,1]$, y añadir, en cada uno de los intervalos que se quitan en el primer paso de su construcción, otro conjunto de Cantor gordo. Continuar este proceso indefinidamente.

Problema 4.17. Construir una función medible Borel $f : [0,1] \to \mathbb{R}$ tal que, para toda $g : [0,1] \to \mathbb{R}$, si $\mathcal{L}_1^*(\{x \in [0,1] : g(x) \neq f(x)\}) = 0$, entonces g es discontinua en todo punto de $[0,1]$. *Indicación:* considerar $f = \chi_E$, donde E es el conjunto del problema anterior.

Problema 4.18. Probar que no existe ninguna función continua $f : \mathbb{R} \to \mathbb{R}$ tal que $f = \chi_{[0,1]}$ en casi todo punto.

Capítulo 5

Integración de funciones. Teoremas de convergencia

En todo lo que sigue, cuando aparezcan valores de funciones o medidas de conjuntos posiblemente infinitos o nulos, convendremos que $0 \cdot \infty = 0$.

En todo este capítulo, si no se dice nada en otro sentido, $(X, \mathcal{A}, \mu)$ será un espacio de medida.

5.1. Integración de funciones positivas

Definición 5.1. Sea $s : X \to [0, \infty)$ una función simple medible, escrita en forma canónica como

$$s = \sum_{j=1}^{n} \alpha_j \chi_{A_j},$$

donde los $\alpha_i \neq \alpha_j$ y $A_i \cap A_j = \emptyset$ para todo $i \neq j$. Para cada $E \in \mathcal{A}$ definiremos

$$\int_E s\, d\mu := \sum_{j=1}^{n} \alpha_j \mu(A_j \cap E). \tag{5.1}$$

Si $f : X \to [0, \infty]$ es medible y $E \in \mathcal{A}$, definimos la integral de Lebesgue de f en E como

$$\int_E f d\mu = \sup \left\{ \int_E s d\mu : 0 \leq s \leq f,\ s \text{ es simple medible} \right\}. \tag{5.2}$$

Ahora mismo, si f es una función simple medible, tenemos dos definiciones para $\int_E f d\mu$, a saber la dada por (5.1) y la dada por (5.2) y lo primero que debemos hacer es comprobar que las dos definiciones coinciden.

Proposición 5.1. *Si $f : X \to [0, \infty)$ es una función simple medible, entonces las definiciones* (5.1) *y* (5.2) *coinciden para f.*

Demostración. En esta demostración $\int_E$ se interpretará exclusivamente en el sentido de (5.1). Sean $f = \sum_{j=1}^{n} \alpha_j \chi_{A_j}$, donde los A_j son disjuntos dos a dos, y $s = \sum_{j=1}^{m} \beta_j \chi_{B_j}$, donde los B_j son disjuntos dos a dos. Supongamos que $0 \neq s \leq f$. Nótese que si $x \in A_i \cap B_j \cap E$ entonces de $s(x) \leq f(x)$ se sigue que $\beta_j \leq \alpha_i$. Es decir, $A_i \cap B_j \cap E \neq \emptyset \implies \beta_j \leq \alpha_i$. Entonces, puesto que

$$X = \bigcup_{i=1}^{n} A_i = \bigcup_{j=1}^{m} B_j = \bigcup_{i,j} A_i \cap B_j,$$

donde todas las uniones son disjuntas, se tiene, para todo $E \in \mathcal{A}$,

$$\int_E s d\mu = \sum_{j=1}^m \beta_j \mu(B_j \cap E) = \sum_{j=1}^m \beta_j \sum_{i=1}^n \mu(B_j \cap A_i \cap E) = \sum_{i,j} \beta_j \mu(A_i \cap B_j \cap E)$$
$$\leq \sum_{i,j} \alpha_i \mu(A_i \cap B_j \cap E) = \sum_{i=1}^n \alpha_i \sum_{j=1}^m \mu(A_i \cap B_j \cap E) = \sum_{i=1}^n \alpha_i \mu(A_i \cap E) = \int_E f d\mu.$$

Es decir, para toda función simple medible s con $0 \leq s \leq f$, es $\int_E s \leq \int_E f$. Por tanto

$$\sup \left\{ \int_E s d\mu : s \text{ simple medible, } 0 \leq s \leq f \right\} \leq \int_E f \leq \sup \left\{ \int_E s d\mu : s \text{ simple medible, } 0 \leq s \leq f \right\},$$

de donde se obtiene la igualdad entre las dos definiciones de integral de f en E. □

En lo sucesivo, si la medida μ se sobrentiende, con frecuencia escribiremos $\int_E f$ en lugar de $\int_E f d\mu$. También, si el espacio X y la medida μ se sobrentienden, escribiremos $\int f$ en lugar de $\int_X f$.

Proposición 5.2. *Sean $f, g : X \to [0, \infty]$ funciones medibles, y $A, B, E \in \mathcal{A}$. Se cumplen:*

1. *Si $0 \leq f \leq g$ entonces $\int_E f d\mu \leq \int_E g d\mu$.*
2. *Si $A \subset B$ y $f \geq 0$ entonces $\int_A f d\mu \leq \int_B f d\mu$.*
3. *Si $f(x) = 0$ para todo $x \in E$ entonces $\int_E f d\mu = 0$ (incluso si $\mu(E) = \infty$).*
4. *Si $c \in [0, \infty)$ entonces $\int_E (cf) d\mu = c \int_E f d\mu$.*
5. *Si $\mu(E) = 0$ entonces $\int_E f d\mu = 0$ (incluso si $f(x) = \infty$ para todo $x \in E$).*
6. *$\int_E f d\mu = \int_X f \chi_E d\mu$.*

Demostración. (1): Si s es una función simple medible y $0 \leq s \leq f$, entonces, como $f \leq g$, también se tiene $0 \leq s \leq g$. Esto implica que el conjunto cuyo supremo define a $\int_E f$ según (5.2) está contenido en el conjunto cuyo supremo define a $\int_E g$, y por tanto el primero de estos supremos es menor o igual que el segundo.

(2): es un argumento parecido y se deja como ejercicio.

(3): Sea $s = \sum_{j=1}^m \beta_j \chi_{B_j}$ (donde los B_j son disjuntos) una función simple medible tal que $0 \leq s \leq f$. Si $x \in E \cap B_j$ entonces $0 \leq s(x) = \beta_j \leq f(x) = 0$, luego $\beta_j = 0$. Es decir, $\beta_j = 0$ siempre que $E \cap B_j \neq \emptyset$. Por tanto $\sum_{j=1}^m \beta_j \mu(B_j \cap E) = 0$. En otras palabras $\int_E s = 0$, y así $\int_E f = \sup_{s \leq f} \int_E s = 0$.

(4): Si $c = 0$ esto es consecuencia de la propiedad anterior. Si $c > 0$ se tiene que

$$c \int_E f = c \sup_{s \leq f} \int_E s = \sup_{s \leq f} c \int_E s = \sup_{s \leq f} \int_E cs = \sup_{\varphi \leq cf} \int_E \varphi = \int_E cf,$$

donde s, φ denotan funciones simples medibles positivas.

(5): es fácil y se deja como ejercicio.

(6): comprobémoslo primero en el caso en que $f = s = \sum_{i=1}^n \alpha_i \chi_{A_i}$ es una función simple. Observemos que entonces $s\chi_E$ también es una función simple. Puesto que, para cualesquiera conjuntos A, B se cumple $\chi_{A \cap B} = \chi_A \chi_B$, tenemos:

$$\int_E s = \sum_{i=1}^n \alpha_i \mu(A_i \cap E) = \int_X \sum_{i=1}^n \alpha_i \chi_{A_i \cap E} = \int_X \sum_{i=1}^n \alpha_i \chi_{A_i} \chi_E = \int_X s\chi_E.$$

Por tanto, en el caso general de una función medible positiva f, se tiene

$$\int_E f = \sup_{s\leq f}\int_E s = \sup_{s\leq f}\int_X s\chi_E = \sup_{\varphi\leq f\chi_E}\int_X \varphi = \int_X f\chi_E$$

donde s, φ denotan funciones simples medibles positivas. La tercera igualdad se justifica observando que si $0 \leq s \leq f$ es simple medible entonces $\varphi = s\chi_E$ también es simple medible y cumple $0 \leq \varphi \leq f\chi_E$, lo que da media igualdad ($\leq$), y por otro lado si $0 \leq \varphi \leq f\chi_E$ es simple medible entonces $0 \leq \varphi \leq f$ y $\varphi = 0$ fuera de E, luego $\varphi = \varphi\chi_E$ y $\int_X \varphi = \int_X \varphi\chi_E$, lo cual da la otra media igualdad ($\geq$). □

Para demostrar los dos teoremas siguientes usaremos el siguiente lema.

Lema 5.3. *Sean $s, t : X \to [0, \infty)$ funciones simples medibles. Entonces la función*

$$\mathcal{A} \ni E \mapsto \varphi_s(E) := \int_E s d\mu$$

define una medida en $\mathcal{A}$. Además, $\varphi_{s+t} = \varphi_s + \varphi_t$; es decir,

$$\int_E (s+t)d\mu = \int_E s d\mu + \int_E t d\mu$$

para todo $E \in \mathcal{A}$.

Demostración. Es obvio que $\varphi_s(\emptyset) = 0$. Veamos que φ_s es numerablemente aditiva. Escribamos $s = \sum_{j=1}^n \alpha_j\chi_{A_j}$. Si $\{E_k\}_{k\in\mathbb{N}} \subset \mathcal{A}$ es una familia numerable de conjuntos disjuntos entonces

$$\begin{aligned}\varphi\left(\bigcup_{k=1}^\infty E_k\right) &= \sum_{i=1}^n \mu\left(\bigcup_{k=1}^\infty E_k \cap A_i\right) = \sum_{i=1}^n\sum_{k=1}^\infty \mu\left(E_k\cap A_i\right)\\ &= \sum_{k=1}^\infty\sum_{i=1}^n \alpha_i\mu\left(E_k\cap A_i\right) = \sum_{k=1}^\infty \varphi(E_k).\end{aligned}$$

Esto prueba que φ_s es una medida en $\mathcal{A}$. Para ver la segunda parte del enunciado, escribamos $t = \sum_{j=1}^m \beta_j\chi_{B_j}$, con los B_j disjuntos. Obsérvese que los conjuntos $E_{ij} := A_i\cap B_j$, $1 \leq i \leq n, 1 \leq j \leq m$, son disjuntos, y en E_{ij} el único valor que toma $s+t$ es $\alpha_i + \beta_j$. Entonces, para todo $E \in \mathcal{A}$ se tiene

$$\int_{E_{ij}\cap E}(s+t)d\mu = (\alpha_i+\beta_j)\mu(E_{ij}\cap E) = \alpha_i\mu(E_{ij}\cap E)+\beta_j\mu(E_{ij}\cap E) = \int_{E_{ij}\cap E} s d\mu + \int_{E_{ij}\cap E} t d\mu.$$

Luego, como φ_s, φ_t y φ_{s+t} son medidas y los $E_{ij}\cap E$ son disjuntos y su unión es E, se tiene

$$\begin{aligned}\varphi_{s+t}(E) &= \varphi_{s+t}\left(\bigcup_{i,j} E_{ij}\cap E\right) = \sum_{i,j}\varphi_{s+t}\left(E_{ij}\cap E\right) = \sum_{i,j}\left[\varphi_s\left(E_{ij}\cap E\right) + \varphi_t\left(E_{ij}\cap E\right)\right]\\ &= \sum_{i,j}\varphi_s\left(E_{ij}\cap E\right) + \sum_{i,j}\varphi_t\left(E_{ij}\cap E\right) = \varphi_s(E)+\varphi_t(E).\end{aligned}$$

Esto prueba la aditividad de la integral para funciones simples medibles. □

El siguiente teorema tiene una importancia fundamental en la teoría de integración.

Teorema 5.4 (de la convergencia monótona de Lebesgue)**.** *Sea (f_n) una sucesión de funciones medibles en X tales que $0 \leq f_n(x) \leq f_{n+1}(x)$ para todo $x \in X$, y sea $f(x) = \lim_{n\to\infty} f_n(x)$, definida para cada $x \in X$. Entonces $f : X \to [0,\infty]$ es medible, y*

$$\lim_{n\to\infty}\int_X f_n d\mu = \int_X f d\mu.$$

Demostración. Ya sabemos que f es medible por el Corolario 4.8. Como $f_n \leq f_{n+1} \leq f$ se tiene $\int_X f_n d\mu \leq \int_X f_{n+1} d\mu \leq \int_X f d\mu$ para todo n, y por tanto existe

$$\alpha := \lim_{n\to\infty} \int_X f_n d\mu \leq \int_X f d\mu.$$

Por tanto el trabajo se reduce a probar que $\int_X f d\mu \leq \alpha$. Sea $s \leq f$ una función simple medible positiva, y fijemos una constante $c \in (0,1)$. Definamos

$$E_n = \{x \in X : f_n(x) \geq cs(x)\}$$

para cada $n \in \mathbb{N}$. Como $f_n(x) \nearrow f(x)$ cuando $n \to \infty$, se tiene que $E_n \subset E_{n+1}$ para todo $n \in \mathbb{N}$, y

$$X = \bigcup_{n=1}^{\infty} E_n.$$

En efecto, dado $x \in X$, si $f(x) = 0$ entonces $f_n(x) = s(x) = 0$ y por tanto $x \in E_n$ para todo n; por otra parte, si $f(x) > 0$ entonces $f(x) > cf(x) \geq cs(x)$, y como $f_n(x) \to f(x)$ existe n_x tal que $cs(x) \leq f_n(x) \leq f(x)$ para todo $n \geq n_x$, y por tanto $x \in E_n$ si $n \geq n_x$.

Entonces, por definición de E_n,

$$\int_X f_n d\mu \geq \int_{E_n} f_n d\mu \geq c \int_{E_n} s d\mu. \tag{5.3}$$

Puesto que por el lema anterior $\varphi(E) := \int_E s d\mu$ define una medida en $\mathcal{A}$ y la sucesión de conjuntos $\{E_n\}$ es creciente con unión X, tenemos que

$$\lim_{n\to\infty} \int_{E_n} s d\mu = \lim_{n\to\infty} \varphi(E_n) = \varphi(X) = \int_X s d\mu.$$

Por tanto, tomando límite cuando $n \to \infty$ en (5.3) obtenemos

$$\alpha \geq c \int_X s d\mu.$$

Y al ser esto verdad para cualquier $c \in (0,1)$, deducimos que

$$\alpha \geq \int_X s d\mu,$$

como queríamos ver. □

Teorema 5.5. *Si $(f_n)_{n\in\mathbb{N}}$ es una sucesión de funciones medibles $f_n : X \to [0,\infty]$ y*

$$f(x) := \sum_{n=1}^{\infty} f_n(x), \quad x \in X,$$

entonces f es medible y

$$\int_X f d\mu = \sum_{n=1}^{\infty} \int_X f_n d\mu.$$

Demostración. Como $f = \lim_{N\to\infty} \sum_{n=1}^{N} f_n$ es límite de una sucesión de funciones medibles, es medible. Sean (s_k) y (t_k) sucesiones crecientes de funciones simples medibles tales que $f_1 = \lim_{k\to\infty} s_k$ y $f_2 = \lim_{k\to\infty} t_k$. Entonces $(s_k + t_k)$ converge de forma creciente a $f_1 + f_2$, y por el Lema 5.3 y el Teorema de la Convergencia Monótona obtenemos que

$$\begin{aligned}\int_X (f_1 + f_2) d\mu &= \lim_{k\to\infty} \int_X (s_k + t_k) d\mu = \lim_{k\to\infty} \left(\int_X s_k d\mu + \int_X t_k d\mu \right) \\ &= \lim_{k\to\infty} \int_X s_k d\mu + \lim_{k\to\infty} \int_X t_k d\mu = \int_X f_1 d\mu + \int_X f_2 d\mu.\end{aligned}$$

Entonces por inducción se obtiene que

$$\int_X \left(\sum_{n=1}^{N} f_n \right) d\mu = \sum_{n=1}^{N} \int_X f_n d\mu$$

para cualquier $N \in \mathbb{N}$, y tomando límites y aplicando el Teorema de la Convergencia Monótona otra vez concluimos que $\int_X f d\mu = \sum_{n=1}^{\infty} \int_X f_n d\mu$. $\square$

Corolario 5.6. *Si $f_1, f_2 : X \to [0, \infty]$ son medibles entonces*

$$\int_X (f_1 + f_2) = \int_X f_1 + \int_X f_2.$$

Teorema 5.7 (Lema de Fatou)**.** *si $f_n : X \to [0, \infty]$ es medible para cada $n \in \mathbb{N}$ entonces*

$$\int_X \liminf_{n\to\infty} f_n \leq \liminf_{n\to\infty} \int_X f_n.$$

Demostración. Definamos $g_n(x) = \inf_{k\geq n} f_k(x)$; entonces $g_n \leq f_n$ para todo n, y (g_n) es una sucesión monótona creciente de funciones medibles cuyo límite es precisamente $\liminf_{n\to\infty} f_n$. Luego podemos aplicar el Teorema de la convergencia monótona y obtener

$$\int_X \liminf f_n = \int_X \lim g_n = \lim \int_X g_n = \liminf \int_X g_n \leq \liminf \int_X f_n.$$

$\square$

Teorema 5.8. *Sea $f : X \to [0, \infty]$ medible. Entonces*

$$\varphi(E) := \int_E f d\mu$$

define una medida en $\mathcal{A}$. Además,

$$\int_X g d\varphi = \int_X g f d\mu \tag{5.4}$$

para toda función medible $g : X \to [0, \infty]$.

Demostración. Es obvio que $\varphi(\emptyset) = 0$. Veamos que φ es numerablemente aditiva. Sea $\{E_n\}_{n\in\mathbb{N}} \subset \mathcal{A}$, disjuntos dos a dos, y pongamos $E = \bigcup_{n=1}^{\infty} E_n$. Entonces

$$f\chi_E = \sum_{n=1}^{\infty} f\chi_{E_n},$$

luego

$$\varphi(E) = \int_X f\chi_E d\mu = \int_X \sum_{n=1}^{\infty} f\chi_{E_n} d\mu = \sum_{n=1}^{\infty} \int_X f\chi_{E_n} d\mu = \sum_{n=1}^{\infty} \varphi(E_n).$$

Esto prueba que φ es una medida. Para ver la segunda parte, observemos primero que (5.4) es verdad cuando $g = \chi_E$ es la función característica de un conjunto medible, ya que

$$\int_X \chi_E d\varphi = \varphi(E) = \int_E f d\mu = \int_X \chi_E f d\mu.$$

Entonces también es cierto si $g = \sum_{j=1}^{n} \alpha_j \chi_{A_j}$ es una función simple medible positiva:

$$\int_X g d\varphi = \sum_{j=1}^{n} \alpha_j \int_X \chi_{A_j} d\varphi = \sum_{j=1}^{n} \alpha_j \int_X \chi_{A_j} f d\mu = \int_X g f d\mu.$$

En el caso general, si $g : X \to [0, \infty]$ es una función medible, existe una sucesión de funciones medibles positivas $s_n \nearrow g$, luego también $s_n f \nearrow gf$, y usando lo anterior y el Teorema de la Convergencia Monótona deducimos que

$$\int_X g d\varphi = \lim_{n\to\infty} \int_X s_n d\varphi = \lim_{n\to\infty} \int_X s_n f d\mu = \lim_{n\to\infty} \int_X g f d\mu.$$

□

Corolario 5.9. *Sea $f : X \to [0, \infty]$ medible. Si $A, B \in \mathcal{A}$ y $\mu(A \cap B) = 0$, entonces*

$$\int_{A\cup B} f d\mu = \int_A f d\mu + \int_B f d\mu.$$

Demostración. Con la notación del teorema anterior,

$$\varphi(A \cap B) = \int_{A\cap B} f d\mu = 0,$$

ya que $\mu(A \cap B) = 0$. Luego, por ser φ una medida,

$$\varphi(B \setminus (A \cap B)) = \varphi(B \setminus (A \cap B)) + \varphi(A \cap B) = \varphi(B)$$

y así

$$\varphi(A \cup B) = \varphi(A) + \varphi(B \setminus (A \cap B)) = \varphi(A) + \varphi(B),$$

lo que equivale a la igualdad del enunciado. □

5.2. Integración de funciones con valores reales o complejos

Definición 5.2. Sean $(X, \mathcal{A}, \mu)$ un espacio de medida y $f : X \to \mathbb{R}$ una función. Diremos que f es integrable si f es medible y

$$\int_X |f| d\mu < \infty.$$

En este caso, definiremos la integral de $f = f^+ - f^-$ en X por

$$\int_X f d\mu = \int_X f^+ d\mu - \int_X f^- d\mu. \tag{5.5}$$

Obsérvese que en el caso de que $f : X \to [0, \infty]$ se tiene $f^- = 0$, $|f| = f = f^+$, y por tanto, suponiendo $\int_X f d\mu < \infty$, la integral de f conincide con la ya conocida para funciones positivas.[1] Nótese también que (5.5) está bien definida ya que

$$\int_X f^\pm d\mu \leq \int_X |f| d\mu < \infty.$$

Si $f = u + iv : X \to \mathbb{C}$ toma valores complejos, diremos que f es integrable si lo son sus partes real e imaginaria, las funciones u y v. Puesto que $|u|, |v| \leq |f|$, esto equivale a decir que f es medible y

$$\int_X |f| d\mu < \infty.$$

[1]Como hemos visto en la sección anterior, toda función medible positiva tiene asociada una integral, pero eso no quiere decir que sea integrable (y nos hemos cuidado de no usar esta terminología). Decimos que es integrable sólo cuando su integral es finita.

En ese caso, definiremos la integral de f por

$$\int_X f d\mu = \int_X u d\mu + i \int_X v d\mu.$$

Definiremos el espacio $L^1(\mu)$ como el conjunto de todas las funciones integrables $f : X \to \mathbb{C}$. También usaremos la misma notación para el espacio de todas las funciones integrables $f : X \to \mathbb{R}$. Por el contexto quedará claro si nos referimos a uno u otro espacio. Asimismo $L^1(X)$ denotará cualquiera de estos dos espacios. Usaremos $L^1(X)$ cuando queramos enfatizar el espacio en el que trabajamos, o $L^1(\mu)$ cuando queramos hacer lo propio con la medida. En los casos en que consideremos varias medidas sobre un mismo espacio fijo, deberemos usar la notación $L^1(\mu)$ para evitar confusiones. Si en algún momento necesitamos la máxima precisión notacional, usaremos $L^1(X, \mu, \mathbb{C})$ o $L^1(X, \mu, \mathbb{R})$ para referirnos a los conjuntos de funciones integrables sobre X, respecto de la medida μ, que toman valores complejos o reales.

Proposición 5.10.

(a) $L^1(X, \mu, \mathbb{R})$ *es un espacio vectorial real, y la integral es un funcional* $\mathbb{R}$*-lineal en este espacio; es decir,*

$$\int_X (\alpha f + \beta g) d\mu = \alpha \int_X f d\mu + \beta \int_X g d\mu$$

para todas $f, g \in L^1(X, \mu, \mathbb{R})$ *y todos* $\alpha, \beta \in \mathbb{R}$.

(b) $L^1(X, \mu, \mathbb{C})$ *es un espacio vectorial complejo, y la integral es un funcional* $\mathbb{C}$*-lineal en este espacio; es decir,*

$$\int_X (\alpha f + \beta g) d\mu = \alpha \int_X f d\mu + \beta \int_X g d\mu$$

para todas $f, g \in L^1(X, \mu, \mathbb{C})$ *y todos* $\alpha, \beta \in \mathbb{C}$.

Demostración. Haremos la demostración de (a), dejando la (b) como ejercicio. Si $f, g \in L^1(X)$ entonces

$$\int_X |\alpha f + \beta g| \leq \int_X (|\alpha||f| + |\beta||g|) = |\alpha| \int_X |f| + |\beta| \int_X |g| < \infty$$

y por tanto $\alpha f + \beta g \in L^1(X)$. Veamos que la integral es lineal. Comenzamos probando que es aditiva: puesto que

$$f^+ - f^- + g^+ - g^- = f + g = (f+g)^+ - (f+g)^-,$$

tenemos también

$$f^+ + g^+ + (f+g)^- = (f+g)^+ + f^- + g^-,$$

donde todas las funciones son positivas. Integrando esta igualdad y usando la aditividad ya conocida de la integral para funciones positivas, obtenemos

$$\int_X f^+ + \int_X g^+ + \int_X (f+g)^- = \int_X (f+g)^+ + \int_X f^- + \int_X g^-,$$

de donde

$$\int_X (f+g)^+ - \int_X (f+g)^+ = \int_X f^+ - \int_X f^- + \int_X g^+ - \int_X g^-,$$

es decir, $\int_X (f+g) = \int_X f + \int_X g$. Veamos ahora que $\int_X \alpha f = \alpha \int_X f$ para toda $f \in L^1(X)$ y todo $\alpha \in \mathbb{R}$; combinando esto con la aditividad ya probada, la linealidad queda clara. Para $\alpha = 0$ esto es obvio. Para $\alpha > 0$, como $f^{\pm}$ son positivas, ya sabemos que

$$\alpha \int_X f^{\pm} = \int (\alpha f^{\pm}),$$

luego

$$\alpha \int_X f = \alpha \int_X f^+ - \alpha \int_X f^- = \int_X \alpha f^+ - \int_X \alpha f^- = \int_X (\alpha f)^+ - \int_X (\alpha f)^- = \int_X \alpha f.$$

Para $\alpha < 0$ tenemos, usando lo ya probado para $\alpha > 0$ y toda f, y el hecho de que $(-f)^\pm = f^\mp$,

$$\int_X \alpha f = \int_X (-\alpha)(-f) = -\alpha \int_X (-f) = -\alpha \left(\int_X f^- - \int_X f^+ \right) = \alpha \int_X f.$$

En todo caso se cumple lo que queremos. □

Definición 5.3. Definimos $L^1(X, \mu, \mathbb{R}^m)$ como el espacio de las funciones medibles $f : X \to \mathbb{R}^m$ tales que $\int_X |f| d\mu < \infty$. En este caso definimos también la integral de $f = (f_1, ..., f_m)$ por

$$\int_X f d\mu = \left(\int_X f_1 d\mu, ..., \int_X f_m d\mu \right).$$

Es fácil comprobar que $L^1(X, \mu, \mathbb{R}^m)$ es un espacio vectorial real y que la integral así definida es una función lineal de este espacio en $\mathbb{R}^m$. Además se tiene lo siguiente.

Proposición 5.11. *Sea L^1 cualquiera de los espacios $L^1(X, \mu, \mathbb{R})$, $L^1(X, \mu, \mathbb{C})$ o $L^1(X, \mu, \mathbb{R}^m)$. Para toda $f \in L^1$ se cumple que*[2]

$$\left| \int_X f d\mu \right| \le \int_X |f| d\mu.$$

Demostración. Puesto que $|u + iv|$ coincide con la norma euclídea de (u, v), el caso de $L^1(X, \mu, \mathbb{R}^m)$ engloba a todos los demás, y podemos limitarnos a considerar sólo este. Denotemos

$$\alpha = (\alpha_1, ..., \alpha_n) = \left(\int_X f_1, ..., \int_X f_m \right) = \int_X f.$$

Podemos suponer $\alpha \neq 0$. Tenemos entonces, usando la linealidad de la integral de funciones con valores en $\mathbb{R}$ y la desigualdad de Cauchy-Schwarz en $\mathbb{R}^m$, que

$$|\alpha|^2 = \langle \alpha, \alpha \rangle = \sum_{j=1}^m \alpha_j \int_X f_j d\mu = \int_X \sum_{j=1}^m \alpha_j f_j d\mu \le \int_X |\alpha|\, |f| d\mu = |\alpha| \int_X |f| d\mu,$$

de donde deducimos la desigualdad deseada al dividir por $|\alpha|$. □

Definición 5.4. En el espacio vectorial $L^1(\mu)$ se define

$$\|f\|_1 := \int_X |f| d\mu$$

para cada $\in L^1(X)$. Es inmediato comprobar que

$$\|f + g\|_1 \le \|f\|_1 + \|g\|_1$$

para todas $f, g \in L^1(\mu)$. También es obvio que $\|0\|_1 = 0$, y que $\|\lambda f\|_1 = |\lambda| \|f\|_1$ si $f \in L^1(\mu)$, y $\lambda \in \mathbb{R}$ (o $\mathbb{C}$, dependiendo del caso). Por tanto $\| \cdot \|_1$ define una seminorma en $L^1(\mu)$.

Sin embargo $\| \cdot \|_1$ no es una norma, puesto que si f es igual a cero salvo en un conjunto no vacío de medida cero entonces $\|f\|_1 = 0$ aunque $f \neq 0$.[3]

[2] En el caso $L^1(X, \mu, \mathbb{R}^m)$ la norma $| \cdot |$ considerada en $\mathbb{R}^m$ es la euclídea. El resultado es cierto para cualquier norma, pero una demostración de este hecho basada en la misma idea requeriría el uso del teorema de Hahn-Banach.

[3] En la próxima sección veremos que si identificamos las funciones que son iguales salvo en subconjuntos de medida cero entonces $\| \cdot \|_1$ sí es una norma en el espacio cociente respecto de esta relación de equivalencia.

El siguiente teorema es otro de los resultados fundamentales de la teoría de integración.

Teorema 5.12 (de la Convergencia Dominada de Lebesgue)**.** *Sea $(f_n)_{n\in\mathbb{N}}$ una sucesión de funciones medibles en X (con valores reales, complejos o vectoriales) tales que existe* $\lim_{n\to\infty} f_n(x) := f(x)$ *para todo $x \in X$. Supongamos que existe $g \in L^1(X)$ tal que*

$$|f_n(x)| \leq g(x) \quad \textit{para todo } x \in X,\ n \in \mathbb{N}.$$

Entonces: $f \in L^1(X)$,

$$\lim_{n\to\infty} \int_X |f_n - f| d\mu = 0,$$

y

$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X f d\mu.$$

Demostración. Es claro que f es medible, y $|f| \leq g \in L^1(X)$. Por tanto $f \in L^1(X)$. Es fácil ver que $2g - |f_n - f| \geq 0$, y esta sucesión tiene límite $2g$. Aplicando el lema de Fatou tenemos entonces

$$\begin{aligned} \textstyle\int_X 2g \leq \quad & \liminf_{n\to\infty} \int_X (2g - |f_n - f|) \\ &= \int_X 2g + \liminf_{n\to\infty} \left(- \int_X |f_n - f| \right) \\ &= \int_X 2g - \limsup_{n\to\infty} \int_X |f_n - f|. \end{aligned}$$

Puesto que $0 \leq \int_X 2g < \infty$, deducimos que

$$\limsup_{n\to\infty} \int_X |f_n - f| \leq 0,$$

lo que implica que $\lim_{n\to\infty} \int_X |f_n - f| = 0$. Por otra parte,

$$\left| \int_X f_n - \int_X f \right| = \left| \int_X (f_n - f) \right| \leq \int_X |f_n - f| \to 0$$

cuando $n \to \infty$, y esto prueba la última afirmación. $\square$

Corolario 5.13. *Si $\mu(X) < \infty$ y $(f_n)_{n\in\mathbb{N}} \subset L^1(\mu)$ converge uniformemente a f, entonces $f \in L^1(\mu)$,*

$$\lim_{n\to\infty} \int_X |f_n - f| d\mu = 0,$$

y

$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X |f| d\mu.$$

Demostración. Para todo $n \geq n_0$, con algún n_0 suficientemente grande, se tiene que $|f_n - f_{n_0}| \leq 1$, luego $|f_n| \leq |f_{n_0}| + 1 := g \in L^1(\mu)$ (ya que $\mu(X) < \infty$), y puede aplicarse el teorema anterior. $\square$

Teorema 5.14 (Continuidad absoluta de la integral)**.** *Sea $f \in L^1(\mu)$. Para todo $\varepsilon > 0$ existe $\delta > 0$ tal que si $\mu(E) < \delta$ entonces $\int_E |f| < \varepsilon$.*

Demostración. Por definición de integral, para cada $\varepsilon > 0$ existe φ_ε simple medible positiva menor o igual que $|f|$ y tal que

$$\int_X \varphi_\varepsilon \geq \int_X |f| - \varepsilon/2,$$

de donde, para cada $E \in \mathcal{A}$ se tiene

$$\int_E (|f| - \varphi_\varepsilon) \leq \int_X (|f| - \varphi_\varepsilon) \leq \varepsilon/2;$$

luego

$$\int_E |f| \leq \int_E \varphi_\varepsilon + \varepsilon/2 \leq (\text{máx}\, \varphi_\varepsilon)\, \mu(E) + \varepsilon/2 \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$$

siempre que

$$\mu(E) < \delta := \frac{\varepsilon}{2(1 + \text{máx}\, \varphi_\varepsilon)}.$$

□

5.3. Propiedades que ocurren en casi todo punto

Definición 5.5. Sea $(X, \mathcal{A}, \mu)$ un espacio de medida. Si existe un conjunto $\mathcal{N} \in \mathcal{A}$ con $\mu(\mathcal{N}) = 0$ y tal que una propiedad $P(x)$ es verdad para todo $x \in X \setminus \mathcal{N}$, entonces decimos que $P(x)$ se cumple *para casi todo* $x \in X$.

Cuando la medida μ es completa, esto equivale a decir que el conjunto $\{x \in X : \neg P(x)\}$ tiene medida cero. Pero si la medida no es completa, puede suceder que este conjunto no sea medible; entonces decir que $P(x)$ es verdad para casi todo $x \in X$ equivale a decir que existe un conjunto medible de medida cero que contiene a $\{x \in X : \neg P(x)\}$.

Observación 5.6. No obstante, recordemos que por el Problema 3.24, incluso si μ no es completa, siempre existe una extensión $\widetilde{\mathcal{A}}$ de la σ-álgebra $\mathcal{A}$ que incluye a todos los sunbconjuntos de conjuntos con medida cero, y puede definirse en esta nueva σ-álgebra una extensión $\widetilde{\mu}$ de la medida μ que es completa.[4] También es útil observar que si una función es $\mathcal{A}$-medible entonces en particular es $\widetilde{\mathcal{A}}$-medible, y en caso de ser positiva (o integrable) su integral respecto de μ en cualquier $E \in \mathcal{A}$ es igual a su integral respecto de la compleción $\widetilde{\mu}$ de μ. En efecto, dada $f : X \to [0, \infty]$, si f es $\mathcal{A}$-medible, existe una sucesión (s_n) de funciones simples $\mathcal{A}$-medibles (luego también $\widetilde{\mathcal{A}}$-medibles) tales que $s_n \nearrow f$; entonces, puesto que $\mu = \widetilde{\mu}$ en $\mathcal{A}$, es obvio que

$$\int_X s_n d\mu = \int_X s_n d\widetilde{\mu}$$

para todo n, y tomando límites y aplicando el Teorema de la Convergencia Monótona se deduce que

$$\int_X f d\mu = \int_X f d\widetilde{\mu}.$$

Por tanto, si en algún momento lo necesitamos para demostrar algún resultado de integración, podremos trabajar con la compleción (o incluso, en determinadas situaciones, y con las debidas precauciones, suponer sin pérdida de generalidad que la medida es completa).

Ejemplo 5.1. Una sucesión de funciones (f_n) está definida en casi todo punto de X si y sólo si para cada n existen $D_n, E_n \subset X$ con $f : D_n \to [-\infty, \infty]$, $E_n \in \mathcal{A}$, $\mu(X \setminus E_n) = 0$, $E_n \subset D_n$. Además, $f_n(x)$ converge en casi todo $x \in X$ a $f(x)$ si y sólo si existe $\mathcal{N} \in \mathcal{A}$ con $\mu(\mathcal{N}) = 0$ de modo que $X \setminus \mathcal{N} \subset \bigcap_{n=1}^\infty D_n$ y $\text{lím}_{n\to\infty} f_n(x) = f(x)$ para todo $x \in X \setminus \mathcal{N}$.

Ejemplo 5.2. Sean $(X, \mathcal{A}, \mu)$ un espacio de medida, e Y un conjunto cualquiera. En el conjunto de las funciones $f : X \to Y$ definimos $f\mathcal{R}g$ si $f(x) = g(x)$ para casi todo $x \in X$. Es fácil comprobar que $\mathcal{R}$ define una relación de equivalencia en este conjunto (la propiedad transitiva se sigue del hecho de que la unión de dos conjuntos de medida cero tiene medida cero). Cuando $Y = \mathbb{R}$ o $\mathbb{C}$ y restringimos esta relación al espacio $L^1(\mu)$ obtenemos un conjunto cociente $L^1(\mu)/\mathcal{R}$ cuyos elementos son las distintas clases

[4] Esto será particularmente útil cuando estudiemos las medidas producto, que rara vez son completas.

de equivalencia $[f]$ así generadas. Por ejemplo, si $X = \mathbb{R}$, $[\chi_{\mathbb{Q}}] = [0]$, ya que la función característica de los racionales es igual a cero en casi todo punto.

Por abuso de notación, salvo en casos muy especiales en los que para la claridad de la discusión sea imprescindible hacerlo de otra forma, seguiremos denotando este conjunto cociente $L^1(\mu)/\mathcal{R}$ (que es también un espacio vectorial cociente) como $L^1(\mu)$, y la clase $[f]$ de f por f. Por el contexto quedará claro si identificamos o no funciones que coinciden en casi todo punto, y si estamos hablando de clases de equivalencia de funciones o de funciones propiamente dichas. En muchos de los casos dará lo mismo si lo hacemos o no, ya que la integral de dos funciones que coinciden en casi todo punto es siempre la misma: para todas $f, g \in L^1(\mu)$,

$$\text{si } f = g \text{ en casi todo punto, entonces } \int_E f d\mu = \int_E g d\mu \text{ para todo } E \in \mathcal{A}. \tag{5.6}$$

En efecto, si $f = g$ en $X \setminus N$, donde $\mu(N) = 0$, entonces, dado $E \in \mathcal{A}$, es

$$\int_E f d\mu = \int_{E\setminus N} f d\mu + \int_N f d\mu = \int_{E\setminus N} f d\mu = \int_{E\setminus N} g d\mu = \int_{E\setminus N} g d\mu + \int_N g d\mu = \int_E g d\mu.$$

Por tanto puede definirse la integral de una clase de equivalencia como la integral de uno cualquiera de sus representantes, conservando las propiedades ya conocidas (linealidad de la integral, etc). También podemos decir que una sucesión $\{[f_n]\}_{n\in\mathbb{N}}$ de clases de equivalencia converge a otra clase de equivalencia $[f]$ si f_n converge a f en casi todo punto. Es fácil comprobar que si $g_n \in [f_n]$ para cada n y $g \in [f]$ entonces f_n converge a f en casi todo punto si y sólo si g_n converge a g en casi todo punto, y por eso esta definición tiene sentido.

El siguiente resultado nos da un recíproco de (5.6).

Teorema 5.15. *Sea $(X, \mathcal{A}, \mu)$ un espacio de medida.*

1. *Si $f : X \to [0, \infty]$ es medible y $\int_E f d\mu = 0$, entonces $f = 0$ en casi todo punto $x \in E$.*
2. *Si $f \in L^1(\mu)$ y $\int_E f d\mu = 0$ para todo $E \in \mathcal{A}$, entonces $f = 0$ en casi todo punto.*
3. *Si $f, g \in L^1(\mu)$ y $\int_E f d\mu = \int_E g d\mu$ para todo $E \in \mathcal{A}$, entonces $f = g$ en casi todo punto.*

Demostración. (1). Definamos $A_n := \{x \in E : f(x) \geq 1/n\}$ para cada $n \in \mathbb{N}$. Entonces

$$\frac{1}{n}\mu(A_n) \leq \int_{A_n} f d\mu \leq \int_E f d\mu = 0,$$

de donde $\mu(A_n) = 0$ para todo n, y por tanto $\{x \in E : f(x) > 0\} = \bigcup_{n=1}^{\infty} A_n$ tiene medida cero.

(2). En el caso de valores reales, sea $E = \{x \in X : f(x) \geq 0\}$. Entonces

$$\int_E f^+ d\mu = \int_E f d\mu = 0,$$

y por tanto por (1) se deduce $f^+ = 0$ en casi todo punto de E; por otro lado por definición de parte positiva es $f^+ = 0$ fuera de E. Luego $f^+ = 0$ en casi todo punto de X. Análogamente se ve que $f^- = 0$ en casi todo punto de X. Por consiguiente $f(x) = 0$ para casi todo $x \in X$. El caso complejo se obtiene aplicando el mismo tipo de argumento a las partes reales e imaginarias de f.

(3) es consecuencia inmediata de (2) aplicado a $f - g$. □

A propósito de la demostración anterior, al tipo de desigualdad usada para probar (1) se le suele llamar desigualdad de Tchebychev. Aunque su demostración es trivial, enunciamos explícitamente este resultado porque lo usaremos más adelante en diferentes contextos.

Proposición 5.16 (desigualdad de Tchebychev). *Si $f \in L^1(\mu)$ entonces para cualquier $\lambda > 0$ se tiene*

$$\mu\left(\{x \in X : f(x) \geq \lambda\}\right) \leq \frac{1}{\lambda}\|f\|_1.$$

Antes hemos visto que $\|f\|_1 := \int_X |f| d\mu$ define una seminorma en el espacio $L^1(\mu)$ (sin identificar funciones que coinciden en casi todo punto). Al identificar funciones que coinciden en casi todo punto y poner

$$\|[f]\|_1 = \int_X |f| d\mu,$$

se ve fácilmente que $\|\cdot\|_1$ está bien definida en $L^1(\mu)/\mathcal{R}$, y como consecuencia del teorema anterior,

$$\|[f]\|_1 = 0 \implies [f] = 0.$$

Así pues, identificando funciones que coinciden en casi todo punto, $(L^1(\mu), \|\cdot\|_1)$ es un espacio vectorial normado. En el Capítulo 9 veremos que este espacio normado es completo (es decir, toda sucesión de Cauchy en este espacio es convergente).

Observación 5.7. Supongamos que μ es una medida completa en $(X, \mathcal{A})$. Si una función f está definida sólo en casi todo punto de X, es decir, $f : X \setminus N \to Y$, donde $\mu(N) = 0$, entonces por la Proposición 4.1 f es medible si y sólo si tiene alguna extensión medible definida en todo X. Equivalentemente, dado cualquier $y_0 \in Y$, la función

$$\widetilde{f}(x) = \begin{cases} f(x) & \text{si } x \in X \setminus N, \\ y_0 & \text{si } x \in N \end{cases}$$

es medible en X. En este tipo de situación, abusando de notación escribimos $f : X \to Y$ definida en casi todo punto. La siguiente proposición nos dice que, en cuanto a la medibilidad de sus extensiones a X, da exactamente igual cómo las definamos en N.

En particular los representantes de las funciones de $L^1(\mu)$ sólo necesitan definirse en casi todo punto para poder definir una clase de equivalencia de este espacio.

Proposición 5.17. *Sean $(X, \mathcal{A}, \mu)$ un espacio de medida completa, y f una función definida en casi todo punto de X, es decir, $f : X \setminus N \to Y$, donde $\mu(N) = 0$ e Y es un espacio métrico. Las siguientes afirmaciones son equivalentes:*

1. *$f : X \setminus N \to Y$ es medible;*
2. *Toda extensión de f a X es medible.*

Demostración. (1) $\implies$ (2): Sea $F : X \to Y$ tal que $F = f$ en $X \setminus N$. Entonces, para todo $U \subset Y$ abierto,

$$F^{-1}(U) = \left(F^{-1}(U) \cap (X \setminus N)\right) \cup \left(F^{-1}(U) \cap N\right) = f^{-1}(U) \cup \left(F^{-1}(U) \cap N\right).$$

Como μ es completa, $\mu(N) = 0$ y $F^{-1}(U) \cap N \subset N$, se tiene que $F^{-1}(U) \cap N \in \mathcal{A}$. Como también es $f^{-1}(U) \in \mathcal{A}$, por ser f medible, resulta que $F^{-1}(U) \in \mathcal{A}$.

(2) $\implies$ (1) es trivial usando la Proposición 4.1. □

Es importante observar que en todos los teoremas de convergencia estudiados en este capítulo, las hipótesis sobre convergencia puntual pueden reemplazarse por las de convergencia en casi todo punto. Por ejemplo, el Teorema de la Convergencia Dominada puede formularse como sigue.

Teorema 5.18 (Teorema de la Convergencia Dominada). *Sea $(f_n)_{n\in\mathbb{N}}$ una sucesión de funciones medibles definidas en casi todo punto de X (con valores reales, complejos o vectoriales) tales que existe $\lim_{n\to\infty} f_n(x) := f(x)$ para casi todo $x \in X$. Supongamos que existe $g \in L^1(X)$ tal que*

$$|f_n(x)| \leq g(x) \quad \textit{para casi todo } x \in X,\ n \in \mathbb{N}.$$

Entonces: f es igual en casi todo punto a una función (denotada igual) $f \in L^1(X)$,

$$\lim_{n\to\infty} \int_X |f_n - f| d\mu = 0,$$

y

$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X f d\mu.$$

Demostración. Sean D_n el conjunto donde está definida f_n, C el conjunto donde f_n converge a f, y G_n el conjunto donde $|f_n| \leq g$. Por hipótesis existen $A_n, B_n, E \in \mathcal{A}$ de medida cero tales que $X \setminus D_n \subset A_n$, $X \setminus G_n \subset B_n$, $X \setminus C \subset E$. Entonces

$$\mathcal{N} = E \cup \bigcup_{n=1}^{\infty} A_n \cup B_n$$

es medible, tiene medida cero, y

$$X \setminus \mathcal{N} \subset C \cup \bigcup_{n=1}^{\infty} D_n \cup G_n;$$

es decir, en cada punto de $X \setminus \mathcal{N}$ se cumplen las hipótesis del teorema. Así podemos redefinir todas las funciones que intervienen en el enunciado de forma que valgan 0 en $\mathcal{N}$ y coincidan con las antiguas en $X \setminus \mathcal{N}$. Entonces las nuevas funciones así definidas, que denotaremos $\widetilde{f_n}$, $\widetilde{f}$, etc, satisfacen las hipótesis del Teorema 5.12 (es decir, las cumplen en todo punto $x \in X$). Por dicho teorema, se tiene que $\widetilde{f} \in L^1(\mu)$, $\lim_{n\to\infty} \int_X |\widetilde{f_n} - \widetilde{f}| d\mu = 0$, y $\lim_{n\to\infty} \int_X \widetilde{f_n} d\mu = \int_X \widetilde{f} d\mu$. Como la integral no distingue entre funciones que difieren en conjuntos de medida cero, estas mismas conclusiones valen con f_n en lugar de $\widetilde{f_n}$, y f en lugar de $\widetilde{f}$. □

El mismo tipo de argumento sirve para demostrar la versión del Teorema de la Convergencia Monótona que enunciamos a continuación.

Teorema 5.19. *Sea (f_n) una sucesión de funciones medibles definidas en casi todo punto de X tales que $0 \leq f_n(x) \leq f_{n+1}(x)$ para casi todo $x \in X$, y sea $f(x) = \lim_{n\to\infty} f_n(x)$ (definida para cada casi todo $x \in X$). Entonces f coincide en casi todo punto con una función medible (que denotamos también f), y*

$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X f d\mu.$$

Igualmente puede formularse una versión del lema de Fatou para funciones definidas en ctp.

Sabemos (ver el Problema 4.10) que una aplicación medible Lebesgue no tiene por qué ser medible Borel. Sin embargo el siguiente resultado muestra que al menos coincide en casi todo punto con una función medible Borel.[5]

Teorema 5.20. *Sea Y un espacio métrico separable, y $f : \mathbb{R}^n \to Y$ medible Lebesgue. Entonces existe $g : \mathbb{R}^n \to Y$ medible Borel tal que $f = g$ en casi todo punto.*

Demostración. Sea $\{U_j\}_{j\in\mathbb{N}}$ una familia numerable de abiertos de Y tales que todo abierto de Y es unión de alguna subfamilia suya (por ejemplo las bolas abiertas de radio racional con centro en un subconjunto denso numerable de Y). Para cada $j \in \mathbb{N}$ sea M_j un F_σ tal que

$$M_j \subset f^{-1}(U_j) \quad \text{y} \quad \mathcal{L}_n\left(f^{-1}(U_j) \setminus M_j\right) = 0.$$

[5] Así pues, todos los fenómenos relativamente extraños que suceden con las funciones que son medibles Lebesgue pero no de Borel, ocurren en subconjuntos de medida cero del espacio donde están definidas; sin embargo, la repercusión de estos efectos en los conjuntos imágenes de esas aplicaciones puede ser brutal.

El conjunto

$$E := \bigcup_{j=1}^{\infty} \left(f^{-1}(U_j) \setminus M_j\right)$$

tiene medida cero, y por tanto existe H conjunto G_δ tal que $E \subset H$ y $\mathcal{L}_n(H) = 0$. Fijemos $y_0 \in Y$ y definamos

$$g(x) = \begin{cases} f(x) & \text{si } x \notin H, \\ y_0 & \text{si } x \in H. \end{cases}$$

Es obvio que $f = g$ en casi todo punto. Sólo queda por ver que g es medible Borel. No es difícil ver que

$$g^{-1}(U_j) = \begin{cases} M_j \cup H & \text{si } y_0 \in U_j, \\ M_j \setminus H & \text{si } y_o \notin U_j, \end{cases}$$

y en cualquiera de los casos resulta $g^{-1}(U_j) \in \mathcal{B}(\mathbb{R}^n)$. Entonces, dado cualquier abierto $U \subset Y$, podemos escribir $U = \bigcup_{k=1}^{\infty} U_{j_k}$ para alguna sucesión $(j_k) \subset \mathbb{N}$, y por tanto

$$g^{-1}(U) = \bigcup_{k=1}^{\infty} g^{-1}(U_{j_k})$$

es boreliano. □

5.4. Observaciones sobre la integración de funciones en $\mathbb{R}^n$

Hasta aquí hemos expuesto la parte más básica de la teoría de la integral de Lebesgue. Obviamente en general no resulta muy práctico calcular integrales a partir de la definición. Para ello, en el caso principal de $X = \mathbb{R}^n$ con $\mu = \mathcal{L}_n$, las herramientas esenciales de cálculo son el Teorema Fundamental del Cálculo y el Teorema de Fubini (y también, en menor medida, el teorema del cambio de variables). La versión general del teorema de Fubini (para producto de medidas cualesquiera) la estudiaremos en el Capítulo 8.

A partir de este momento, si no especificamos nada en sentido contrario, la medida que consideraremos en $\mathbb{R}^n$ y sus subconjuntos más usuales (abiertos, intervalos, etc) será siempre $\mathcal{L}_n$, y la integral en $\mathbb{R}^n$ será la de Lebesgue respecto de la medida $\mathcal{L}_n$. Denotaremos en este caso

$$\int_E f d\mathcal{L}_n = \int_E f(x)dx.$$

El Teorema Fundamental del Cálculo que se estudia en primero es válido, con la misma demostración, cambiando la integral de Riemann por la de Lebesgue en $\mathbb{R}$. Lo enunciamos a continuación y dejamos a cargo del lector la verificación de esta afirmación.

Teorema 5.21. *Si $f : [a,b] \to \mathbb{R}$ es continua y $F(x) := \int_a^x f(t)dt$, entonces F es derivable, con $F'(x) = f(x)$.*

La regla de Barrow y el método de integración por partes, que son consecuencias inmediatas de este teorema, valen por tanto igualmente para la integral de Lebesgue. Todas las herramientas para el cálculo de primitivas son válidas en este contexto y pueden usarse sin problema a partir de ahora.

El teorema del cambio de variables para la integral de Lebesgue en $\mathbb{R}^n$ también es cierto, pero no lo estudiaremos tal cual en este curso, sino que veremos resultados mucho más potentes que lo generalizan, a saber las fórmulas del área y de la coárea, y que sirven además para calcular medidas de Hausdorff de dimensión k de subconjuntos de $\mathbb{R}^n$, con k entero menor o igual que n; estos resultados más avanzados los expondremos en el capítulo 13.

Por tanto en este curso podemos prescindir totalmente de la integral de Riemann. No obstante, en lo que resta de este capítulo vamos a hacer un breve estudio de las relaciones fundamentales entre la integral de Riemann y la de Lebesgue en $\mathbb{R}^n$. En particular veremos que toda función integrable Riemann es integrable Lebesgue, con la misma integral; sin embargo el recíproco no es cierto. Vamos a suponer $n = 1$ por simplicidad, aunque todo lo que vamos a decir a continuación es cierto para dimensiones mayores cambiando intervalos por rectángulos.

En la teoría de integración de Riemann se demuestra que una función acotada $f : [a, b] \to \mathbb{R}$ es integrable Riemann si y sólo si el conjunto $D := \{x \in [a, b] : f \text{ es discontinua en } x\}$ tiene medida cero. Entonces, si f es integrable Riemann, f es medible, porque D y $A := [a, b] \setminus D$ lo son, y para cada abierto U de $\mathbb{R}$, al ser $f|_A$ continua, $(f|_A)^{-1}(U)$ es un abierto relativo a A, es decir, un abierto de $\mathbb{R}$ intersecado con A, y en particular un conjunto medible, y así resulta que

$$f^{-1}(U) = (f|_A)^{-1}(U) \cup \big(f^{-1}(U) \cap D\big)$$

es medible (pues el primer conjunto de esta unión es medible y el segundo también, al ser un subconjunto de un conjunto de medida cero). Como toda función medible y acotada en un espacio de medida finita es obviamente integrable, deducimos que $f \in L^1([a, b])$.

Esto prueba que toda función integrable Riemann es integrable Lebesgue. El recíproco no es cierto: por ejemplo, $f = \chi_{\mathbb{Q}\cap[0,1]}$ es discontinua en todo punto, y por tanto no es integrable Riemann; sin embargo sí es integrable Lebesgue, con integral cero, ya que es igual a cero en casi todo punto.

Veamos ahora que si $f : [a, b] \to \mathbb{R}$ es integrable Riemann entonces su integral de Riemann coincide con la de Lebesgue. Dada una partición P_n de $[a, b]$ en subintervalos I_j^n de longitud $(b - a)/n$ y con extremos $t_{j-1}^n < t_j^n$, definamos

$$M_j^n = \sup_{x\in I_j^n} f(x),\ \ m_j^n = \inf_{x\in I_j^n} f(x),\ \ L(f, P_n) = \sum_{j=1}^{n} m_j^n(t_j^n - t_{j-1}^n),\ \ U(f, P_n) = \sum_{j=1}^{n} M_j^n(t_j^n - t_{j-1}^n).$$

Consideremos la función

$$g_n = \sum_{j=1}^{n} s_j^n \chi_{(t_{j-1}^n, t_j^n)},$$

donde, para cada $j = 1, ..., n$, s_j^n es un número cualquiera del intervalo $[m_j^n, M_j^n]$.

Veamos que $\lim_{n\to\infty} g_n(x) = f(x)$ para todo $x \in E := [a, b] \setminus \Big(D \cup \{t_j^n : n \in \mathbb{N}, j = 1, ..., n\}\Big)$. En efecto, si $x \in E$ entonces f es continua en x, luego dado $\varepsilon > 0$ existe $\delta > 0$ tal que si $|y - x| \leq \delta$ entonces $|f(y) - f(x)| \leq \varepsilon/2$. Sea $n_0 \in \mathbb{N}$ tal que $(b - a)/n_0 < \delta$. Entonces para todo $n \geq n_0$, como $x \in E$, existe un único $k = k_x^n \in \{1, ..., n\}$ tal que $x \in (t_{k-1}^n, t_k^n)$, con $(t_k^n - t_{k-1}^n) = (b - a)/n < \delta$, y por tanto

$$|g_n(x) - f(x)| = |s_k^n - f(x)| \leq M_k^n - m_k^n \leq M_k - f(x) + f(x) - m_k \leq \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Así g_n converge a f en casi todo punto.

Como además g_n está acotada por $M = \sup_{x\in[a,b]} |f(x)|$, que es independiente de n, y $[a, b]$ tiene medida finita, usando el Teorema de la Convergencia Dominada se ve que

$$\lim_{n\to\infty} \int_a^b g_n(t)dt = \int_a^b f(t)dt.$$

Por otro lado, es claro que

$$L(f, P_n) \leq \int_a^b g_n(t)dt \leq U(f, P_n)$$

para todo $n \in \mathbb{N}$, y por ser f integrable Riemann se tiene que

$$\lim_{n\to\infty} U(f, P_n) = \lim_{n\to\infty} L(f, P_n) = \text{(integral de Riemann)} \int_a^b f.$$

Esto, junto con la desigualdad anterior, implica que

$$\int_a^b f(t)dt = \text{(integral de Riemann)} \int_a^b f.$$

Finalmente, permitámonos una breve observación sobre las integrales de Riemann impropias: no se trata de verdaderas integrales, sino de límites de integrales. Por ejemplo,

$$\text{(integral de Riemann impropia)} \int_{-\infty}^{\infty} f(t)dt = \lim_{a\to\infty} \text{(integral de Riemann)} \int_{-a}^{a} f(t)dt.$$

Estos límites pueden existir o no, y si existen para la integral de Riemann, entonces de forma obvia (puesto que la integral de Riemann de una función acotada en un intervalo compacto coincide con la de Lebesgue) también existirán para la integral de Lebesgue. Sin embargo de la existencia de estos límites no se puede concluir que la función sea integrable Lebesgue (por ejemplo en $\mathbb{R}$, en el caso del límite anteriormente considerado). Conviene insitir en este hecho: las integrales impropias de Riemann son límites de integrales, no son integrales.

5.5. Problemas

Problema 5.1. Sea X un espacio métrico, μ una medida en $(X, \mathcal{A})$, y $\mathcal{G}$ un grupo de isometrías sobreyectivas de X. Supongamos que $T(A) \in \mathcal{A}$ y $\mu(T(A)) = \mu(A)$ para toda $A \in \mathcal{A}$ y todo $T \in \mathcal{G}$. Demostrar que entonces $\int_X f d\mu = \int_X f \circ T d\mu$ para toda $T \in \mathcal{G}$ y toda $f \in L^1(\mu)$. *Indicación:* considerar primero el caso en que $f = \chi_E$ para $E \in \mathcal{A}$.

Problema 5.2. Sea X un espacio métrico, $\mathcal{A}$ la σ-álgebra de los conjuntos $\mathcal{H}^s$-medibles, y $T : X \to X$ una isometría sobreyectiva (es decir, una aplicación sobreyectiva tal que $d(T(x), T(y)) = d(x, y)$ para todos $x, y \in X$). Demostrar que entonces, para todo $s > 0$ y toda toda $f \in L^1(\mathcal{H}^s)$, se cumple $\int_X f d\mathcal{H}^s = \int_X (f \circ T) d\mathcal{H}^s$.

Problema 5.3. Deducir del problema anterior que para toda $f : \mathbb{R}^n \to \mathbb{R}$ integrable se tiene que $\int_{\mathbb{R}^n} f(x + h)dx = \int_{\mathbb{R}^n} f(x)dx$ para todo $h \in \mathbb{R}^n$, y que $\int_{\mathbb{R}^n} f(R(x))dx = \int_{\mathbb{R}^n} f(x)dx$ para toda rotación R de $\mathbb{R}^n$.

Problema 5.4. Si $f : X \to [0, \infty]$ es medible y $f_n = \min\{f, n\}$, probar que $\int_X f = \lim_{n\to\infty} \int_X f_n$.

Problema 5.5. Dar un ejemplo de $f, g \in L^1(\mathbb{R})$ tales que $fg \notin L^1(\mathbb{R})$.

Problema 5.6. Construir una sucesión $(f_n)_{n\in\mathbb{N}}$ de funciones medibles $f_n : \mathbb{R} \to [0, \infty)$ tal que la desigualdad del Lema de Fatou es estricta.

Problema 5.7. Probar que $\lim_{n\to\infty} \int_0^\pi e^{-n|\operatorname{sen} t|} dt = 0$.

Problema 5.8. Sea $(X, \mathcal{A}, \mu)$ un espacio de medida. Supongamos que $\mathcal{A} \supset \mathcal{B}(X)$, y que μ es finita en cada bola de X. Demostrar que, para toda función medible $f : X \to \mathbb{R}$ estrictamente positiva en casi todo punto, para todo $x_0 \in X$ y para todo $R > 0$ se tiene que

$$\lim_{t\to\infty} \int_{B(x_0,R)} e^{-tf(x)} d\mu(x) = 0.$$

Problema 5.9. Si $\sum_{n\in\mathbb{N}} \int_X |f_n| d\mu < \infty$, demostrar que $\sum_{n=1}^\infty f_n(x)$ converge para casi todo $x \in X$ a una función f integrable, para la cual se tiene $\int_X f d\mu = \sum_{n\in\mathbb{N}} \int_X f_n d\mu$. Deducir que si (g_n) converge a g en $L^1(\mu)$ entonces (g_n) tiene una subsucesión que converge a g en μ-casi todo punto.

Problema 5.10. En el espacio de medida $(\mathbb{N}, 2^{\mathbb{N}}, \sharp)$, donde $\sharp$ denota la medida de contar, demostrar que una función $a : \mathbb{N} \to \mathbb{R}$ es integrable si y sólo si $\sum_{n\in\mathbb{N}} a(n)$ es absolutamente convergente, en cuyo caso $\int_{\mathbb{N}} a\, d\sharp = \sum_{n=1}^\infty a(n)$.

Problema 5.11. Dar ejemplos de sucesiones de funciones (f_n), (g_n), (h_n) en $L^1(\mathbb{R})$ que converjan a cero uniformemente y tales que: $\lim_{n\to\infty}\int_{\mathbb{R}} f_n = \infty$; $\int_{\mathbb{R}} g_n = 1$ para todo n; $\limsup_{n\to\infty}\int_{\mathbb{R}} h_n = 1$, $\liminf_{n\to\infty}\int_{\mathbb{R}} h_n = -1$.

Problema 5.12. Demostrar que si $f \in L^1(\mathbb{R}^n)$ entonces $\liminf_{|x|\to\infty} |f(x)| = 0$.

Problema 5.13. Dar un ejemplo de $f \in C(\mathbb{R}) \cap L^1(\mathbb{R})$ tal que $\limsup_{|x|\to\infty} f(x) = \infty$.

Problema 5.14. Probar que si $f \in L^1(\mathbb{R}^n)$ es uniformemente continua, entonces $\lim_{|x|\to\infty} f(x) = 0$.

Problema 5.15. Si $f \in L^1(X)$ con $\mu(X) < \infty$, probar que $|f|^{1/n}$ es integrable para todo $n \in \mathbb{N}$, y calcular $\lim_{n\to\infty}\int_X |f|^{1/n} d\mu$.

Problema 5.16. Si X es un conjunto no vacío y δ_a es la medida de Dirac en a, demostrar que toda función $f : X \to \mathbb{R}$ es integrable, con $\int_X f d\delta_a = f(a)$.

Problema 5.17. Demostrar que no existe ninguna serie trigonométrica del tipo $\sum_{n=1}^{\infty} a_n \operatorname{sen}(nx) + b_n \cos(nx)$ que converja a 1 en casi todo x y tal que $\sum_{n\in\mathbb{N}} |a_n| + |b_n| < \infty$.

Problema 5.18. Sean $(X, \mathcal{A}, \mu)$ espacio de medida finita, $f : X \to \mathbb{R}$ medible, y definamos $A_n = \{x \in X : |f(x)| \geq n\}$, $B_n = \{x \in X : n-1 \leq |f(x)| < n\}$, $n \in \mathbb{N}$. Demostrar que son equivalentes:

1. $f \in L^1(X)$;
2. $\sum_{n\in\mathbb{N}} n\,\mu(B_n) < \infty$;
3. $\sum_{n\in\mathbb{N}} \mu(A_n) < \infty$.

Problema 5.19. Calcular los siguientes límites, justificando cada paso:
a) $\lim_{n\to\infty}\int_0^n \left(1+\frac{x}{n}\right)^n e^{\alpha x}dx$; b) $\lim_{n\to\infty}\int_0^n \left(1-\frac{x}{n}\right)^n e^{-x}dx$; c) $\lim_{n\to\infty}\int_0^\infty \frac{n\operatorname{sen}(x/n)}{x(1+x^2)}dx$.

Problema 5.20. Dar un ejemplo que muestre que el Corolario 5.13 es falso si X tiene medida infinita.

Problema 5.21. Sea $f : \mathbb{R} \to [0,\infty)$ medible. Si para todo $n \in \mathbb{N}$ es $\int_{\mathbb{R}} \frac{n^2}{n^2+x^2} f(x)dx \leq 1$, deducir que f es integrable y $\int_{\mathbb{R}} f \leq 1$.

Problema 5.22. Sea (f_n) una sucesión de funciones medibles $f_n : X \to [0,\infty]$ que converge puntualmente a una función f tal que $f_n \leq f$ para todo n. Probar que $\int_X f = \lim_{n\to\infty}\int_X f_n$.

Problema 5.23. Sea (f_n) una sucesión decreciente de funciones medibles positivas. Si $\int_X f_{n_0} < \infty$ para algún n_0, demostrar que $\int_X \lim_{n\to\infty} f_n = \lim_{n\to\infty}\int_X f_n$.

Problema 5.24. Sea (f_n) una sucesión de funciones positivas e integrables que converge puntualmente a una función integrable f, y supongamos que $\int_X f = \lim_{n\to\infty}\int_X f_n$. Probar que $\lim_{n\to\infty}\int_X |f_n - f| = 0$. Deducir que para todo $E \subset X$ medible se tiene $\int_E f = \lim_{n\to\infty}\int_E f_n$. Mostrar que este enunciado es falso en general si $f_n \geq 0$ para todo n pero $\int_X f = \infty$.

Problema 5.25. Si $f \in C[0,1]$, calcular $\lim_{n\to\infty} n\int_0^1 x^n f(x)dx$.

Problema 5.26. Si $f : [0,1] \to [0,\infty]$ medible tiene la propiedad de que $\int_0^1 f(x)^n dx$ es finito e independiente de n, probar que existe E medible tal que $f = \chi_E$ en casi todo punto.

Problema 5.27. Transcribir el enunciado del teorema de la convergencia dominada en el caso particular $X = \mathbb{N}$, $\mathcal{S} = 2^{\mathbb{N}}$, y $\mu = \sharp$.

Problema 5.28. Demostrar que $\lim_{n\to\infty}\int_a^n \frac{n^2xe^{-n^2x^2}}{1+x^2}dx = 0$ para todo $a > 0$, pero que esto no es cierto para $a = 0$.

Problema 5.29. Si $f \in L^1(\mathbb{R})$, probar que $g(x) = \int_{-\infty}^x f$ es uniformemente continua en $\mathbb{R}$.

Problema 5.30. Sean $(X, \mathcal{A}, \mu)$ un espacio de medida completa, $f : X \to [0, \infty]$ una función arbitraria, y $(f_n)_{n\in\mathbb{N}}$, $(g_n)_{n\in\mathbb{N}}$ sucesiones de funciones medibles en X tales que

$$0 \leq f_n \leq f_{n+1} \leq f \leq g_{n+1} \leq g_n$$

para todo $n \in \mathbb{N}$. Supongamos que $\lim_{n\to\infty} \int_X (g_n - f_n) d\mu = 0$. Probar que f es medible.

Problema 5.31. Probar que si $f, f_k \in L^1(\mu)$, $\lim_{k\to\infty} \int_X |f_k| d\mu = \int_X |f| d\mu$, y $\lim_{k\to\infty} f_k = f$ en casi todo punto, entonces

$$\lim_{k\to\infty} \int |f_k - f| d\mu = 0.$$

Problema 5.32. Si $E \subset \mathbb{R}^n$ es medible, demostrar que existe $A \subset E$ medible tal que $|A| = \frac{1}{\pi}|E|$.

Capítulo 6

Los tres principios de Littlewood. Teoremas de Lusin y Egoroff.

> "There are three principles, roughly expressible in the following terms: Every (measurable) set is nearly a finite sum of intervals; every function (of class L^λ) is nearly continuous; every convergent sequence of functions is nearly uniformly convergent."
> — J. E. Littlewood, Lectures on the Theory of Functions (1944)

La cita de Littlewood alude a los tres principios básicos que él recomendaba seguir, como guía intuitiva, a los principiantes en la teoría de la medida en $\mathbb{R}^n$. El primero de ellos ya lo hemos visto en capítulos anteriores: todo conjunto de medida finita se aproxima interiormente en medida por compactos, y los compactos se aproximan exteriormente en medida por uniones finitas de rectángulos; por tanto podría decirse que todo conjunto medible Lebesgue y de medida finita en $\mathbb{R}^n$ es, aproximadamente, una unión finita de rectángulos.

El segundo principio alude al teorema de Lusin, que estudiaremos en este capítulo: toda función medible definida en un conjunto medible con medida finita de $\mathbb{R}^n$ es, salvo un conjunto de medida arbitrariamente pequeña, la restricción de alguna función continua.

Y el tercer principio parafrasea el teorema de Egoroff, que probaremos también en este capítulo: si una sucesión de funciones, definidas en un conjunto medible de medida finita, converge puntualmente a una función, entonces también lo hace uniformemente salvo en un conjunto de medida arbitrariamente pequeña.

6.1. El teorema de Lusin

Daremos dos versiones del teorema de Lusin: una para funciones medibles Borel definidas en espacios métricos (y con valores en espacios métricos separables), y otra para funciones medibles Lebesgue definidas en subconjuntos de $\mathbb{R}^n$.

Teorema 6.1 (Lusin). *Sea X un espacio métrico, y μ una medida en $\mathcal{B}(X)$ tal que X es unión numerable de conjuntos abiertos de medida finita. Si Y es un espacio métrico separable y $f : X \to Y$ es medible Borel, entonces para todo $\varepsilon > 0$ existe un cerrado $F \subset X$ tal que $\mu(X \setminus F) < \varepsilon$ y $f|_F$ es continua.*

Demostración. Sea $\{U_i\}_{i\in\mathbb{N}}$ una familia numerable de abiertos de Y tal que cualquier abierto de Y es unión de una subfamilia de $\{U_i\}_{i\in\mathbb{N}}$. Por el Teorema 2.4, para cada $i \in \mathbb{N}$ existe un abierto $V_i \subset X$ tal que

$$f^{-1}(U_i) \subset V_i, \;\; \mu(V_i \setminus f^{-1}(U_i)) < \frac{\varepsilon}{2^{i+1}}.$$

Sea
$$E = \bigcup_{i=1}^{\infty} V_i \setminus f^{-1}(U_i).$$
Es claro que $\mu(E) < \varepsilon/2$. Vamos a probar que $g = f|_{X\setminus E}$ es continua. Observemos que, para cada $i \in \mathbb{N}$,
$$g^{-1}(U_i) = V_i \cap (X \setminus E). \tag{6.1}$$
En efecto, $g^{-1}(U_i) \subset V_i \cap (X \setminus E)$, y por otro lado
$$V_i \cap (X \setminus E) \subset V_i \cap (X \setminus (V_i \setminus f^{-1}(U_i))) = f^{-1}(U_i),$$
luego
$$V_i \cap (X \setminus E) \subset f^{-1}(U_i) \cap (X \setminus E) = g^{-1}(U_i).$$
Para ver que g es continua bastará probar que para cada abierto $U \subset Y$ existe un abierto $V \subset X$ tal que $g^{-1}(U) = V \cap (X \setminus E)$. Escribamos $U = \bigcup_{k=1}^{\infty} U_{i_k}$ para una subsucesión $(i_k) \subset \mathbb{N}$ adecuada. Entonces por (6.1) se tiene $g^{-1}(U) = (\bigcup_{k=1}^{\infty} V_{i_k}) \cap (X \setminus E)$, y tenemos lo que queremos.

El conjunto $X \setminus E$ no es necesariamente cerrado, pero por el Teorema 2.4 existe $G \subset X$ abierto tal que $E \subset G$ y $\mu(G \setminus E) < \varepsilon/2$. Poniendo $F = X \setminus G$, resulta que $f|_F$ es continua, F es cerrado y $\mu(X \setminus F) = \mu(G) = \mu(E) + \mu(G \setminus E) < \varepsilon$. □

Observación 6.1. El teorema no dice que f sea continua en F: esto puede ser falso. Por ejemplo, si $f = \chi_{\mathbb{Q}}$, entonces f es discontinua en todos los puntos. Sin embargo, incluso para una función discontinua en todo punto como esta (o, de forma aún más radical, incluso para la función del Problema 4.17 para la cual toda función que coincida con ella es discontinua en todo punto), sigue siendo verdad que para todo $\varepsilon > 0$ existe un cerrado $F \subset \mathbb{R}$ tal que la restricción de f a F es continua y $|\mathbb{R} \setminus F| < \varepsilon$.

Teorema 6.2 (Lusin). *Sean $E \subset \mathbb{R}^n$ medible Lebesgue, y $f : E \to \mathbb{R}$. Las siguientes afirmaciones son equivalentes:*

1. *f es medible Lebesgue;*
2. *Para todo $\varepsilon > 0$ existe F cerrado de $\mathbb{R}^n$ tal que $F \subset E$, $|E \setminus F| < \varepsilon$, y $f|_F$ es continua;*
3. *Para cada $\varepsilon > 0$ existe $g : \mathbb{R}^n \to \mathbb{R}$ continua tal que $\mathcal{L}_n^* (\{x \in E : f(x) \neq g(x)\}) < \varepsilon$.*

Demostración. (1) $\implies$ (2): Si $f : E \to \mathbb{R}$ es medible Lebesgue entonces $\widetilde{f} : \mathbb{R}^n \to \mathbb{R}$ definida por $\widetilde{f} = f$ en E y $\widetilde{f} = 0$ fuera de E también lo es, luego por el Teorema 5.20 existe $\widetilde{h} : \mathbb{R}^n \to \mathbb{R}$ medible Borel tal que $\widetilde{f} = \widetilde{h}$ en $\mathbb{R}^n \setminus Z$, donde Z es un boreliano de medida cero. Como $E \setminus Z$ es medible Lebesgue, dado $\varepsilon > 0$ existe F_1 cerrado de $\mathbb{R}^n$ tal que $F_1 \subset E \setminus Z$ y $|(E \setminus Z) \setminus F_1| < \varepsilon/2$, luego también $|E \setminus F_1| < \varepsilon/2$. La función $\widetilde{h}_{|F_1}$ es medible Borel, y coincide con $f_{|F_1}$ (ya que $F_1 \subset E$). Por el teorema anterior, existe $F \subset F_1$ cerrado de F_1 (y por tanto también cerrado de $\mathbb{R}^n$) tal que $|F \setminus F_1| < \varepsilon/2$ y la función
$$\left(\widetilde{h}_{|F_1}\right)_{|F} = \widetilde{h}_{|F} = f_{|F}$$
es continua. Además $|E \setminus F| \leq |E \setminus F_1| + |F_1 \setminus F| < \varepsilon$.

(2) $\implies$ (3): Siguiendo con el argumento anterior, como F es cerrado de $\mathbb{R}^n$ y $f|_F$ es continua, por el teorema de extensión de Tietze existe $g \in C(\mathbb{R}^n)$ tal que $g|_F = f|_F$. Y como $|E \setminus F| < \varepsilon$ y f coincide con g en F, se tiene también que
$$\mathcal{L}_n^* (\{x \in E : g(x) \neq f(x)\}) \leq |E \setminus F| < \varepsilon.$$

(3) $\implies$ (1): Para cada $k \in \mathbb{N}$, sea $g_k \in C(\mathbb{R}^n)$ tal que
$$\mathcal{L}_n^* (\{x \in E : f(x) \neq g_k(x)\}) < \frac{1}{2^k},$$

y denotemos $A_k = \{x \in E : f(x) \neq g_k(x)\}$. Como $\sum_{k=1}^{\infty} \mathcal{L}_n^*(A_k) < \infty$, por el lema de Borel-Cantelli (Problema 2.4), se tiene

$$\mathcal{L}_n^* \left(\{x \in E : x \in A_k \text{ para infinitos } k\} \right) = 0;$$

es decir, para casi todo punto $x \in E$ se tiene que x pertenece a lo sumo a una cantidad finita de los conjuntos A_k, lo que equivale a decir que existe $k_x \in \mathbb{N}$ tal que si $k \geq k_x$ entonces $g_k(x) = f(x)$. Esto implica que $\lim_{k\to\infty} g_k(x) = f(x)$ para casi todo $x \in E$. O sea, que salvo en un conjunto de medida cero, f es el límite puntual de una sucesión de funciones continuas. Como el límite puntual de una sucesión de funciones medibles es medible y la medida $\mathcal{L}_n$ es completa, de esto se deduce que f es medible Lebesgue. □

Corolario 6.3. *Toda función medible $f : E \subset \mathbb{R}^n \to \mathbb{R}$ es límite en casi todo punto de una sucesión de funciones continuas.*[1]

Demostración. Se ha visto en la prueba del anterior teorema. □

6.2. El teorema de Egoroff

En el teorema de Egoroff X puede ser cualquier espacio de medida (no necesariamente un espacio métrico), pero es crucial que tenga medida finita.

Teorema 6.4 (Egoroff)**.** *Sean $(X, \mathcal{A}, \mu)$ un espacio de medida con $\mu(X) < \infty$, y $(f_n)_{n\in\mathbb{N}}$ una sucesión de funciones medibles $f_n : X \to Y$, donde Y es un espacio métrico separable. Si $f_n \to f$ en casi todo punto, entonces para cada $\varepsilon > 0$ existe $E \subset X$ medible tal que $\mu(X \setminus E) < \varepsilon$ y f_n converge a f uniformemente en E.*

Demostración. Quitando a X un subconjunto de medida cero, lo que no afecta a la conclusión, podemos suponer que f es medible.[2] Entonces

$$A_{m,n} := \{x : d(f_k(x), f(x)) \leq 1/m \text{ para todo } k \geq n\}$$

es medible para cada $n, m \in \mathbb{N}$. En efecto, observemos que la función $F_k = (f_k, f) : X \to Y \times Y$ es medible, ya que si $\{U_i\}_{i\in\mathbb{N}}$ es una colección de abiertos de Y tal que cualquier abierto de Y es unión de una subfamilia suya, entonces todo abierto de $Y \times Y$ es unión de una subfamilia de $\{U_i \times U_j\}_{i,j\in\mathbb{N}}$, y puesto que

$$F_k^{-1}(U_i \times U_j) = f_k^{-1}(U_i) \cap f^{-1}(U_j)$$

y f_k, f son medibles, se ve que F es medible. Entonces, como $A_{m,n} = \bigcap_{k\geq n} F_k^{-1}\left(d^{-1}\left([0, 1/m]\right)\right)$, resulta que $A_{m,n}$ es medible.

Para cada $m \in \mathbb{N}$, $(A_{m,n})_{n\in\mathbb{N}}$ forma una sucesión creciente de conjuntos, y por tanto sus complementarios una sucesión decreciente; como $\mu(X) < \infty$ se obtiene que

$$\lim_{n\to\infty} \mu(X \setminus A_{m,n}) = \mu\left(\bigcap_{n=1}^{\infty} (X \setminus A_{m,n})\right) = \mu\left(X \setminus \bigcup_{n=1}^{\infty} A_{m,n}\right) = 0,$$

[1] No es cierto en general que f sea límite en *todo punto* de una sucesión de funciones continuas. Por ejemplo, si $f = \chi_{\mathbb{Q}}$ esto es falso. Las funciones que son límite en todo punto de sucesiones de funciones continuas se llaman de clase Baire-1, y puede probarse que tales funciones tienen la propiedad de que su restricción a cada cerrado tiene al menos un punto de continuidad. Sin embargo, $\chi_{\mathbb{Q}}$ sí es límite puntual de una sucesión de funciones de clase Baire-1 (a las funciones con esta propiedad se les llama de clase Baire-2): se tiene $\chi_{\mathbb{Q}}(x) = \lim_{k\to\infty}\left(\lim_{j\to\infty} \cos(k!\pi x)^{2j}\right)$. Por inducción puede definirse la clase de funciones Baire-k para cualquier k (como límites puntuales de funciones de la clase Baire-$(k-1)$), y tales funciones son siempre medibles Borel. Por tanto, si f es medible Lebesgue pero no medible Borel, entonces f no es de clase Baire-k para ningún k.

[2] Sabemos que el límite puntual de funciones medibles es medible, pero si la medida μ no es completa, un límite en casi todo punto de una sucesión de funciones medibles puede no ser medible.

ya que como f_k converge puntualmente a f en casi todo punto, casi todo $x \in X$ está en $\bigcup_{n=1}^{\infty} A_{m,n}$. Por tanto para cada $m \in \mathbb{N}$ podemos elegir $n(m) \in \mathbb{N}$ tal que

$$\mu\left(X \setminus A_{m,n(m)}\right) < \frac{\varepsilon}{2^m}.$$

Definiendo $E = \bigcap_{m=1}^{\infty} A_{m,n(m)}$, se obtiene que $\mu(X \setminus E) < \varepsilon$, y f_n converge uniformemente a f en E: para todo $m \in \mathbb{N}$, si $k \geq n(m)$ entonces $d(f_k(x), f(x)) \leq 1/m$ para todo $x \in E$. $\square$

Como hemos avanzado antes, la conclusión del teorema de Egoroff es falsa en general si $\mu(X) = \infty$. Por ejemplo, la sucesión de funciones $f_n(x) = x^2/n$ converge a 0 cuando $n \to \infty$ para todo $x \in \mathbb{R}$, y no puede existir ningún conjunto $E \subset \mathbb{R}$ tal que $|\mathbb{R} \setminus E| < \infty$ y (f_n) converja a 0 uniformemente en E. En efecto, tal conjunto E no podría estar acotado, con lo cual existiría $(x_n)_{n\in\mathbb{N}} \subset E$ con $|x_n| \geq n$, y se tendría $f_n(x_n) \geq n \to \infty$ cuando $n \to \infty$.

6.3. Problemas

Problema 6.1. Si (f_n) converge a f en ctp de $\mathbb{R}^d$, demostrar que para cada $\varepsilon > 0$ y cada compacto $K \subset \mathbb{R}^d$ existe otro compacto $C \subset K$ tal que $|K \setminus C| < \varepsilon$ y (f_n) converge uniformemente a f en C.

Problema 6.2. Dar una versión del teorema de Egoroff para sucesiones de funciones medibles que convergen a ∞ en casi todo punto.

Problema 6.3. Sea $K \subset \mathbb{R}^n$ compacto. Si $f_j : K \to \mathbb{R}$ es medible para cada $j \in \mathbb{N}$, probar que existe una sucesión $(c_j) \subset \mathbb{R}$ tal que $\lim_{j\to\infty} f_j(x)/c_j = 0$ para casi todo $x \in \mathbb{R}^n$.
Indicación: primero demostrar que existe (c_j) tal que $|\{x : |f_j(x)/c_j| > 1/j\}| < 1/2^j$; luego aplicar el lema de Borel-Cantelli.

Problema 6.4. Dar un ejemplo de una función $f : \mathbb{R} \to \mathbb{R}$ medible Borel tal que no existe ningún conjunto $B \subset \mathbb{R}$ con $|\mathbb{R} \setminus B| = 0$ y $f|_B$ continua.

Problema 6.5. Sean X un espacio métrico y μ una medida de Borel en X tal que X es unión numerable de conjuntos abiertos de medida finita. Demostrar que para cada $f : X \to \mathbb{R}$ medible existe una sucesión $(f_k) \subset C(X)$ tal que para casi todo $x \in X$ existe $k_x \in \mathbb{N}$ tal que $f_k(x) = f(x)$ para todo $k \geq k_x$. En particular $\lim_{k\to\infty} f_k(x) = f(x)$ para casi todo $x \in X$.

Si además suponemos que X es separable y localmente compacto, usar lo anterior para demostrar lo siguiente: dada $f \in L^1_{loc}(\mu)$, si $\int_X f\varphi d\mu = 0$ para toda $\varphi \in C_c(X)$, entonces $f = 0$ en casi todo punto.

Capítulo 7

Diversas nociones de convergencia

7.1. Convergencia en media y convergencia en casi todo punto

El el capítulo 5 hemos definido el espacio vectorial normado $L^1(\mu)$. Como veremos en el capítulo 9, este espacio es completo. El lector puede estudiar ahora la demostración del Teorema 9.6; no incluimos la demostración de la completitud de $L^1(\mu)$ en este capítulo simplemente porque es la misma que la que se hace para $L^p(\mu)$, para cualquier $p \in [1, \infty)$.

Cuando una sucesión de funciones $(f_n) \subset L^1(\mu)$ converge en el espacio $(L^1(\mu), \|\cdot\|_1)$ a una función f, a veces se dice que f_n *converge a* f *en media*.[1]

La convergencia en media no implica convergencia en casi todo punto, ni la convergencia en casi todo punto implica la convergencia en media. El lector queda invitado a encontrar ejemplos de estas situaciones. Sin embargo (como consecuencia del Problema 5.9, o también como corolario del Teorema de Riesz que demostraremos en la siguiente sección), toda sucesión que converja en $L^1(\mu)$ tiene una subsucesión convergente en casi todo punto (a la misma función límite). Por otro lado, el Teorema de la Convergencia Dominada nos proporciona condiciones que añadidas a la convergencia en casi todo punto implican la convergencia en media.

Otra noción útil, que se usa por ejemplo en probabilidad, es la de convergencia en medida.

7.2. Convergencia en medida

Se dice que una sucesión de funciones medibles $f_n : X \to \mathbb{R}$ converge en medida a una función medible $f : X \to \mathbb{R}$ si para todo $\varepsilon > 0$ es

$$\lim_{n\to\infty} \mu\left(\{x \in X : |f_n(x) - f(x)| \geq \varepsilon\}\right) = 0.$$

Se suele denotar, en este caso, $f_n \xrightarrow{\mu} f$.

Proposición 7.1. *Si* (f_n) *converge a* f *en* $L^1(\mu)$ *entonces* $f_n \xrightarrow{\mu} f$.

Demostración. Por la desigualdad de Tchebychev y la convergencia en media,

$$\mu\left(\{x \in X : |f_n(x) - f(x)| \geq \varepsilon\}\right) \leq \frac{1}{\varepsilon}\|f_n - f\|_1 \to 0.$$

□

El recíproco de la proposición anterior es falso (queda propuesto como ejercicio justificar esta afirmación).

[1] Similarmente, si (f_n) converge a f en el espacio $L^2(\mu)$, a veces también se dice que converge *en media cuadrática*.

Teorema 7.2 (Lebesgue)**.** *Si $\mu(X) < \infty$ y una sucesión de funciones medibles $f_n : X \to \mathbb{R}$ converge a una función medible f en casi todo punto, entonces $f_n \xrightarrow{\mu} f$.*

Demostración. Dado $\varepsilon > 0$, definamos

$$E_j = \{x \in X : |f_j(x) - f(x)| \geq \varepsilon\}.$$

La sucesión de conjuntos $\bigcup_{j=n}^{\infty} E_j$ es decreciente, y son conjuntos de medida finita, luego se tiene

$$\lim_{n\to\infty} \mu\left(\bigcup_{j=n}^{\infty} E_j\right) = \mu\left(\bigcap_{n=1}^{\infty}\bigcup_{j=n}^{\infty} E_j\right) = 0.$$

La última igualdad es porque si $x \in \bigcap_{n=1}^{\infty}\bigcup_{j=n}^{\infty} E_j$ entonces $|f_j(x) - f(x)| \geq \varepsilon$ para infinitos j, lo cual sólo puede ocurrir en un conjunto de medida cero (dado que f_n converge a f en casi todo punto). Por tanto

$$\mu\left(\{x \in X : |f_n(x) - f(x)| \geq \varepsilon\}\right) = \mu(E_n) \leq \mu\left(\bigcup_{j=n}^{\infty} E_j\right) \to 0$$

cuando $n \to \infty$. □

Teorema 7.3 (Riesz)**.** *Si $f_n \xrightarrow{\mu} f$ entonces (f_n) tiene una subsucesión que converge a f en casi todo punto.*

Demostración. Para cada $j \in \mathbb{N}$ existe $n_j \in \mathbb{N}$ tal que

$$\mu\left(\{x \in X : |f_{n_j}(x) - f(x)| \geq \frac{1}{j}\}\right) := \mu(E_j) \leq \frac{1}{2^j}.$$

Podemos suponer que (n_j) es estrictamente creciente, y así (f_{n_j}) es una subsucesión de (f_n). Puesto que

$$\sum_{j=1}^{\infty} \mu(E_j) < \infty,$$

por el Lema de Borel-Cantelli (Problema 1.20) resulta que

$$\mu(E) := \mu\left(\{x \in X : x \in E_j \text{ para infinitos valores de } j\}\right) = 0,$$

y esto significa que para todo $x \in X \setminus E$ existe $j_x \in \mathbb{N}$ tal que, para todo $j \geq j_x$,

$$|f_{n_j}(x) - f(x)| < \frac{1}{j}.$$

Esto implica que $f_{n_j}(x)$ converge a $f(x)$ para casi todo $x \in X$. □

Corolario 7.4. *Si (f_n) converge a f en $L^1(\mu)$, entonces (f_n) tiene una subsucesión que converge a f en casi todo punto.*

Demostración. Se sigue de la Proposición 7.1 y del teorema anterior. □

Definición 7.1. Se dice que una familia de funciones $\mathcal{F} \subset L^1(\mu)$ es *uniformemente integrable* (o *equintegrable*) si para todo $\varepsilon > 0$ existe $\delta > 0$ tal que si $\mu(E) < \delta$ entonces $\int_E |f| d\mu < \varepsilon$.

Obsérvese que el Teorema 5.14 implica que toda familia finita de funciones integrables es equintegrable.[2] Más en general se tiene lo siguiente.

[2] Asimismo la unión de una cantidad finita de familias equintegrables es equintegrable

Proposición 7.5. *Si (f_n) converge a f en $L^1(\mu)$ entonces $\{f_n\}_{n\in\mathbb{N}} \cup \{f\}$ es uniformemente integrable.*

Demostración. Dado $\varepsilon > 0$, sea $n_0 \in \mathbb{N}$ suficientemente grande tal que

$$\int_X |f_n - f| d\mu < \varepsilon/2 \quad \text{para todo } n \geq n_0.$$

Por el Teorema 5.14 aplicado a $|f|+|f_1|+...+|f_{n_0}| \in L^1(\mu)$, existe $\delta > 0$ tal que si $\mu(E) < \delta$ entonces

$$\int_E (|f| + |f_1| + ... + |f_{n_0}|)\, d\mu < \varepsilon/2.$$

Luego, si $\mu(E) < \delta$ se tiene que, cuando $n \leq n_0$,

$$\max\left\{\int_E |f| d\mu, \int_E |f_n| d\mu\right\} \leq \int_E (|f| + |f_1| + ... + |f_{n_0}|)\, d\mu < \varepsilon;$$

y si $n > n_0$ es

$$\int_E |f_n| d\mu \leq \int_E |f_n - f| d\mu + \int_E |f| \leq \int_X |f_n - f| d\mu + \int_E |f| d\mu < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

En todo caso se tiene lo que queremos. □

Por tanto, si (f_n) converge en $L^1(\mu)$, también converge en medida y esta sucesión forma una familia uniformemente integrable. El siguiente resultado dice que el recíproco también es verdad si $\mu(X) < \infty$.

Teorema 7.6 (Vitali)**.** *Supongamos que $\mu(X) < \infty$, y sea $(f_n)_{n\in\mathbb{N}} \subset L^1(\mu)$, $f \in L^1(\mu)$. Las siguientes afirmaciones son equivalentes:*

1. *(f_n) converge a f en $L^1(\mu)$;*
2. *$f_n \xrightarrow{\mu} f$ y $\{f_n\}_{n\in\mathbb{N}}$ es uniformemente integrable.*

Demostración. Antes hemos probado (1) $\implies$ (2). Demostremos (2) $\implies$ (1): por reducción al absurdo, supongamos que existieran $\varepsilon > 0$ y una subsucesión (f_{n_k}) de (f_n) tales que

$$\int_X |f_{n_k} - f| d\mu \geq \varepsilon \quad \text{para todo } k \in \mathbb{N}. \tag{7.1}$$

Por el Teorema de Riesz, puesto que $f_n \xrightarrow{\mu} f$, podemos suponer (salvo tomar otra subsucesión que denotamos igual) que f_{n_k} converge a f en casi todo punto. Para dicho ε, por hipótesis existe $\delta > 0$ tal que

$$\mu(E) < \delta \implies \int_E (|f| + |f_n|) < \varepsilon/2 \quad \text{para todo } n \in \mathbb{N}.$$

Por otro lado, por el Teorema de Egoroff, existe $F \in \mathcal{A}$ tal que $\mu(X \setminus F) < \delta$ y f_{n_k} converge a f uniformemente en f en F. Como $\mu(F) < \infty$, por el Corolario 5.13 se tiene que

$$\lim_{k\to\infty} \int_F |f_{n_k} - f| = 0.$$

Entonces

$$\begin{aligned} \int_X |f_{n_k} - f| &= \int_F |f_{n_k} - f| + \int_{X\setminus F} |f_{n_k} - f| \\ &\leq \int_F |f_{n_k} - f| + \int_{X\setminus F} (|f_{n_k}| + |f|) \leq \int_F |f_{n_k} - f| + \varepsilon/2 \to \varepsilon/2 \end{aligned}$$

cuando $k \to \infty$, y esto contradice 7.1. □

7.3. Problemas

Problema 7.1. Dar ejemplos de sucesiones de funciones (f_n), (g_n), (h_n) tales que: (f_n) converge en casi todo punto pero no converge en media; (g_n) converge en media pero no converge en casi todo punto, y (h_n) converge en medida pero no converge en media.

Problema 7.2. Sea (A_n) una sucesión de conjuntos medibles de medida finita en un espacio de medida X, y supongamos que (χ_{A_n}) converge a f en $L^1(X)$. Demostrar que existe E medible tal que $f = \chi_E$ en casi todo punto.

Problema 7.3. Sea (A_n) una sucesión de conjuntos medibles en un espacio de medida X. Definamos $\limsup_n A_n = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m$, y $\liminf_n A_n = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m$. Observar que $\limsup_n \chi_{A_n} = \chi_{\limsup_n A_n}$, y $\liminf_n \chi_{A_n} = \chi_{\liminf_n A_n}$. Demostrar que:

1. (χ_{A_n}) converge a 0 en medida si y sólo si $\lim_{n\to\infty} \mu(A_n) = 0$.
2. (χ_{A_n}) converge a 0 en casi todo punto si y sólo si $\mu\left(\limsup_{n\to\infty} A_n\right) = 0$.
3. Si A es un conjunto medible, (χ_{A_n}) converge a χ_A en casi todo punto si y sólo si los conjuntos A, $\limsup_n A_n$ y $\liminf_n A_n$ únicamente difieren en conjuntos de medida nula.

Problema 7.4. Si (f_n) y (g_n) son sucesiones de funciones medibles que convergen en medida a f y g respectivamente, probar que $(f_n + g_n)$ converge en medida a $f + g$. Si además $\mu(X) < \infty$, probar también que $(f_n g_n)$ converge en medida a fg.

Problema 7.5. Si $(f_n) \subset L^1(X)$, $f \in L^1(X)$ y $\sum_{n=1}^{\infty} \int_X |f_n - f| < \infty$, demostrar que (f_n) converge en casi todo punto a f.

Problema 7.6. Probar que (f_n) converge en medida a f en el espacio $(\mathbb{N}, 2^{\mathbb{N}}, \sharp)$ si y sólo si (f_n) converge uniformemente a f.

Problema 7.7. Si $\mu(X) < \infty$, probar que una sucesión de funciones medibles (f_n) converge en medida a una función medible f si y sólo si cada subsucesión de (f_n) tiene una subsucesión que converge a f en casi todo punto.

Problema 7.8. Probar que el límite en medida es único en casi todo punto.

Problema 7.9. Estudiar los diferentes tipos de convergencia para las siguientes sucesiones de funciones: $f_n = \frac{1}{n^2}\chi_{[0,n]}$ en $\mathbb{R}$; $g_n(x) = \frac{1}{\sqrt{x}}\chi_{(0,1/n]}(x)$ en $(0, 1]$; $h_n(x) = \frac{1}{x}\chi_{(0,1/n]}(x)$ en $(0, 1]$; $\varphi_n(x) = ne^{-nx}$ en $[0, 1]$.

Problema 7.10. Si $f_n \geq 0$ converge a f en medida, probar que $\int_X f \leq \liminf_{n\to\infty} \int_X f_n$.

Capítulo 8

Medidas producto. Teorema de Fubini-Tonelli

En este capítulo definiremos el producto de dos medidas y demostraremos el teorema de Fubini que nos permite calcular la integral de una función de dos variables como una integral iterada.

8.1. Medidas producto

Definición 8.1. Sean $(X, \mathcal{S})$, $(Y, \mathcal{T})$ dos espacios medibles. Llamaremos *rectángulo medible* de $X \times Y$ a cualquier subconjunto de $X \times Y$ de la forma $S \times T$, donde $S \in \mathcal{S}$ y $T \in \mathcal{T}$.[1]

A las uniones finitas disjuntas de rectángulos medibles las llamaremos *conjuntos elementales*. Es decir, $A \subset X \times Y$ es un conjunto elemental si $A = \bigcup_{j=1}^{n} R_j$, donde cada $R_j = S_j \times T_j$ es un rectángulo medible, y se tiene $R_i \cap R_j = \emptyset$ si $i \neq j$. Denotemos

$$\mathcal{E} = \{A \subset X \times Y : A \text{ es un conjunto elemental}\}.$$

Definimos la σ-álgebra producto $\mathcal{S} \otimes \mathcal{T} := \sigma(\mathcal{E})$. Es decir, $\mathcal{S} \otimes \mathcal{T}$ es por definición la menor σ-álgebra en $X \times Y$ que contiene a todas las uniones finitas disjuntas de rectángulos medibles.

Definición 8.2. Una *clase monótona* es un conjunto $\mathcal{M}$ de conjuntos que tiene las siguientes propiedades:

1. Si $\{A_j\}_{j \in \mathbb{N}} \subset \mathcal{M}$ y $A_j \subset A_{j+1}$ para todo j, entonces $\bigcup_{j=1}^{\infty} A_j \in \mathcal{M}$;
2. Si $\{B_j\}_{j \in \mathbb{N}} \subset \mathcal{M}$ y $B_j \supset B_{j+1}$ para todo j, entonces $\bigcap_{j=1}^{\infty} B_j \in \mathcal{M}$;

es decir, si contiene a las uniones de sucesiones crecientes y a las intersecciones de sucesiones decrecientes de conjuntos de $\mathcal{M}$.

Es evidente que toda σ-álgebra es también una clase monótona. El recíproco no es cierto.

Definición 8.3. Si $E \subset X \times Y$, $x \in X$, definimos la sección de E por x como

$$E_x := \{y \in Y : (x, y) \in E\}.$$

Análogamente, si $y \in Y$, definimos la sección de E por x como

$$E^y := \{x \in X : (x, y) \in E\}.$$

Nótese que $E_x \subset Y$ y $E^y \subset X$.

[1]Esto no quiere decir que, si por ejemplo $X = Y = \mathbb{R}$, el conjunto $S \times T$ sea geométricamente un rectángulo en $\mathbb{R}^2$. Por ejemplo cuando $\mathcal{S} = \mathcal{T} = \mathcal{M}$, la σ-álgebra de los conjuntos medibles Lebesgue de $\mathbb{R}$, los conjuntos $S, T \in \mathcal{M}$ pueden ser extremadamente complicados, y $S \times T$ en general no va a parecerse a un rectángulo en el sentido usual (es decir, a un producto cartesiano de intervalos) de $\mathbb{R}^2$.

Obsérvese que, si $\{E_i\}_{i\in I}$ es una familia cualquiera de subconjuntos de $X \times Y$, se tiene

$$\left(\bigcup_{i\in I} E_i\right)_x = \bigcup_{i\in I} (E_i)_x, \quad \left(\bigcap_{i\in I} E_i\right)_x = \bigcap_{i\in I} (E_i)_x, \tag{8.1}$$

y análogamente

$$\left(\bigcup_{i\in I} E_i\right)^y = \bigcup_{i\in I} (E_i)^y, \quad \left(\bigcap_{i\in I} E_i\right)^y = \bigcap_{i\in I} (E_i)^y. \tag{8.2}$$

También es obvio que

$$(E_x)^c := Y \setminus E_x = ((X \times Y) \setminus E)_x := (E^c)_x,$$

y una igualdad similar es válida para la sección por y. Además, si $A \subset B \subset X \times Y$ entonces, obviamente $A_x \subset B_x$ y $A^y \subset B^y$ para todos $x \in X$, $y \in Y$.

Teorema 8.1. *Si $E \in \mathcal{S} \otimes \mathcal{T}$ entonces $E_x \in \mathcal{T}$ y $E^y \in \mathcal{S}$ para todo $x \in X$, $y \in Y$.*

Demostración. Sea

$$\Omega = \{E \in \mathcal{S} \otimes \mathcal{T} : E_x \in \mathcal{T} \text{ para todo } x \in X\}.$$

Es fácil ver que Ω es una σ-álgebra; en efecto:

1. $X \times Y \in \Omega$, ya que $(X \times Y)_x = Y \in \mathcal{T}$ para todo $x \in X$;
2. Si $E \in \Omega$ entonces $(E^c)_x = (E_x)^c \in \mathcal{T}$ para todo $x \in X$, luego $E^c \in \Omega$;
3. Si $\{E_j\}_{j\in\mathbb{N}} \subset \Omega$ y $E = \bigcup_{j=1}^\infty E_j$ entonces $E_x = \bigcup_{j=1}^\infty (E_j)_x \in \mathcal{T}$ para todo x (por definición de Ω y por ser $\mathcal{T}$ una σ-álgebra), luego $E \in \Omega$.

Además, si $E = A \times B$ es un rectángulo medible entonces

$$E_x = \begin{cases} B & \text{si } x \in A, \\ \emptyset & \text{si } x \notin A; \end{cases}$$

en cualquier caso es $E_x \in \mathcal{T}$, y por tanto $A \times B \in \Omega$. Como Ω es σ-álgebra, resulta que $\mathcal{E} \subset \Omega$. Luego $\sigma(\mathcal{E}) = \mathcal{S} \otimes \mathcal{T} \subset \Omega$. Esto quiere decir que $E_x \in \mathcal{T}$ para todo $E \in \mathcal{S} \otimes \mathcal{T}$.

Análogamente se prueba que $E^y \in \mathcal{S}$ para todo $y \in Y$, siempre que $E \in \mathcal{S} \otimes \mathcal{T}$. □

Teorema 8.2. *$\mathcal{S} \otimes \mathcal{T}$ es la menor clase monótona que contiene a $\mathcal{E}$.*

Demostración. Es fácil ver que la intersección $\mathcal{M}$ de todas las clases monótonas en $X \times Y$ que contienen a $\mathcal{E}$ es una clase monótona. Puesto que $\mathcal{S} \otimes \mathcal{T}$ es una clase monótona que contiene a $\mathcal{E}$, se tiene que $\mathcal{M} \subset \mathcal{S} \otimes \mathcal{T}$. Veamos que también es verdad el contenido opuesto.

Comencemos observando que la intersección de dos rectángulos medibles es un rectángulo medible; en efecto,

$$(A_1 \times B_1) \cap (A_2 \times B_2) = (A_1 \cap A_2) \times (B_1 \times B_2).$$

Por otra parte, la diferencia de dos rectángulos medibles es un conjunto elemental, ya que y

$$(A_1 \times B_1) \setminus (A_2 \times B_2) = [(A_1 \setminus A_2) \times B_1] \cup [(A_1 \cap A_2) \times (B_1 \setminus B_2)].$$

De esto se deduce fácilmente que la intersección de dos conjuntos elementales es un conjunto elemental, y su diferencia también lo es. Entonces, como $A \cup B = (A \setminus B) \cup B$, unión disjunta, se deduce que si $A, B \in \mathcal{E}$ entonces $A \cup B \in \mathcal{E}$. Por inducción, lo mismo ocurre para uniones finitas de conjuntos elementales.

Dado cualquier $A \subset X \times Y$, definamos $\Omega(A)$ como el conjunto de todos los $B \subset X \times Y$ tales que $A \setminus B \in \mathcal{M}$, $B \setminus A \in \mathcal{M}$, y $A \cup B \in \mathcal{M}$. Se tiene que:

(a) $B \in \Omega(A) \iff A \in \Omega(B)$;

(b) $\Omega(A)$ es una clase monótona.

La primera propiedad es obvia. Comprobemos la propiedad (b): si $\{A_n\}_{n\in\mathbb{N}}$, $\{B_n\}_{n\in\mathbb{N}} \subset \Omega(A)$ y para cada n es $A_n \subset A_{n+1}$ y $B_n \supset B_{n+1}$ entonces $A \setminus A_n$ y $B_n \setminus A$ forman sucesiones decrecientes de conjuntos, y $A_n \setminus A$, $A \setminus B_n$ son crecientes; todas estas sucesiones están en $\mathcal{M}$, y por ser esta una clase monótona se tiene que $\bigcap_n (A \setminus A_n) \in \mathcal{M}$, $(\bigcap_n B_n \setminus A) \in \mathcal{M}$, $\bigcup_n (A_n \setminus A) \in \mathcal{M}$, $\bigcup_n (A \setminus B_n) \in \mathcal{M}$, de lo que se deduce que $A \setminus \bigcup_n A_n \in \mathcal{M}$, $(\bigcap_n B_n) \setminus A \in \mathcal{M}$, $(\bigcup_n A_n) \setminus A \in \mathcal{M}$, $A \setminus \bigcap_n B_n \in \mathcal{M}$; además también se tiene $A \cup A_n \in \mathcal{M}$ y $A \cup B_n \in \mathcal{M}$ para todo n, y otra vez por ser $\mathcal{M}$ una clase monótona se obtiene $A \cup \bigcup_n A_n = \bigcup_n (A \cup A_n) \in \mathcal{M}$, y $A \cup \bigcap_n B_n = \bigcap_n (A \cup B_n) \in \mathcal{M}$. Por tanto $\bigcup_n A_n \in \Omega(A)$ y $\bigcap_n B_n \in \Omega(A)$.

Fijemos $A \in \mathcal{E}$. Como ya sabemos que la diferencia y la unión de conjuntos elementales es un conjunto elemental, es claro que $B \in \Omega(A)$ para todo $B \in \mathcal{E}$. Por tanto $\mathcal{E} \subset \Omega(A)$, y por la propiedad (b) y la definición de $\mathcal{M}$ se deduce que $\mathcal{M} \subset \Omega(A)$.

Si ahora fijamos $B \in \mathcal{M}$, como acabamos de ver, $B \in \Omega(A)$ si $A \in \mathcal{E}$. Por la propiedad (a), $A \in \Omega(B)$, luego $\mathcal{E} \subset \Omega(B)$, y usando (b) otra vez obtenemos que $\mathcal{M} \subset \Omega(B)$. Esto significa que si $A, B \in \mathcal{M}$ entonces $A \setminus B \in \mathcal{M}$ y $A \cup B \in \mathcal{M}$. Por inducción lo mismo ocurre para la unión de toda familia finita contenida en $\mathcal{M}$.

Ahora podemos ver que $\mathcal{M}$ es una σ-álgebra en $X \times Y$:

1. $X \times Y \in \mathcal{M}$, ya que $X \times Y \in \mathcal{E} \subset \mathcal{M}$.

2. Si $B \in \mathcal{M}$ entonces $B^c = (X \times Y) \setminus B \in \mathcal{M}$, puesto que ya hemos probado que la diferencia de dos conjuntos de $\mathcal{M}$ está en $\mathcal{M}$.

3. Si $\{A_j\}_{j\in\mathbb{N}} \subset \mathcal{M}$, para cada $n \in \mathbb{N}$ el conjunto $B_n = \bigcup_{j=1}^n A_j$ está en $\mathcal{M}$ porque ya hemos visto que $\mathcal{M}$ es estable por uniones finitas. Como la sucesión $\{B_n\}_{n\in\mathbb{N}}$ es creciente y $\mathcal{M}$ es una clase monótona, resulta que $\bigcup_{n=1}^\infty A_n = \bigcup_{n=1}^\infty B_n \in \mathcal{M}$.

Así pues, $\mathcal{M}$ es una σ-álgebra que contiene a $\mathcal{E}$, y como $\mathcal{S} \otimes \mathcal{T}$ es la menor σ-álgebra con esta propiedad, se concluye que $\mathcal{S} \otimes \mathcal{T} \subset \mathcal{M}$, como queríamos probar. □

Definición 8.4. Para cada función $f : X \times Y \to \mathbb{R}$ (o $\mathbb{C}$, o $[0, \infty]$, o un espacio métrico Z) y cada $x \in X$, definimos la sección de f por x como la función $f_x : Y \to \mathbb{R}$ definida por

$$f_x(y) := f(x, y).$$

Análogamente, para cada $y \in Y$ se define la sección de f por y como la función $f^y : X \to \mathbb{R}$ definida por

$$f^y(x) := f(x, y).$$

Teorema 8.3. *Si f es una función $\mathcal{S} \otimes \mathcal{T}$-medible en $X \times Y$, entonces, para cada $x \in X$ la sección f_x es una función $\mathcal{T}$-medible en Y, y para cada $y \in Y$ la sección f^y es una función $\mathcal{S}$-medible en X.*

Demostración. Sea V un abierto del espacio de llegada de la función f. El conjunto

$$A := f^{-1}(V) = \{(x, y) \in X \times Y : f(x, y) \in V\}$$

está en $\mathcal{S} \otimes \mathcal{T}$, por ser f medible. Entonces, por el Teorema 8.1 se tiene que

$$f_x^{-1}(V) = \{y \in Y : f(x, y) \in V\} = A_x \in \mathcal{T}.$$

La prueba de la medibilidad de la sección f^y es análoga. □

Teorema 8.4. *Sean $(X, \mathcal{S}, \mu)$, $(Y, \mathcal{T}, \lambda)$ dos espacios de medida σ-finitos, y sea $C \in \mathcal{S} \otimes \mathcal{T}$. Definamos*

$$\varphi(x) := \lambda(C_x), \ \ \psi(y) := \mu(C^y)$$

para cada $x \in X$, $y \in Y$. Entonces φ es $\mathcal{S}$-medible, ψ es $\mathcal{T}$-medible, y

$$\int_X \varphi d\mu = \int_Y \psi d\lambda.$$

Demostración. Sea Ω el conjunto de todos los $B \in \mathcal{S} \otimes \mathcal{T}$ para los cuales la conclusión del teorema es verdadera.

Lema 8.5. *Ω tiene las siguientes propiedades:*

(a) Todo rectángulo medible está en Ω.

(b) Si $\{C_j\}_{j\in\mathbb{N}} \subset \Omega$ es una sucesión creciente entonces $\bigcup_{j=1}^{\infty} C_j \in \Omega$.

(c) Si $\{C_j\}_{j\in\mathbb{N}} \subset \Omega$ son disjuntos dos a dos, entonces $\bigcup_{j=1}^{\infty} C_j \in \Omega$.

(d) Si $\mu(A) < \infty$, $\lambda(B) < \infty$, y $\{C_j\}_{j\in\mathbb{N}} \subset \Omega$ es una sucesión decreciente tal que $C_j \subset A \times B$ para todo j, entonces $\bigcap_{j=1}^{\infty} C_j \in \Omega$.

Supuesto probado el lema, en el caso en que X e Y son de medida finita, las propiedades (b) y (d) nos dicen que Ω es una clase monótona, y las propiedades (a) y (c) implican que Ω contiene a $\mathcal{E}$. Como además $\Omega \subset \mathcal{S} \otimes \mathcal{T}$, por el Teorema 8.2 se deduce que $\Omega = \mathcal{S} \otimes \mathcal{T}$, y esto equivale al enunciado que queremos probar. En el caso general en que μ y λ son σ-finitas, necesitamos usar un pequeño artificio adicional: por hipótesis podemos escribir X e Y como uniones numerables de conjuntos con medida finita, y podemos suponer que estas uniones son disjuntas; es decir, tenemos

$$X = \bigcup_{n=1}^{\infty} X_n, \ \ Y = \bigcup_{m=1}^{\infty} Y_m,$$

uniones disjuntas, con $\mu(X_n) < \infty$, $\lambda(Y_m) < \infty$ para cada $n, m \in \mathbb{N}$. Definamos $\mathcal{M}$ como el conjunto de todos los $C \in \mathcal{S} \otimes \mathcal{T}$ tales que

$$C_{n,m} := C \cap (X_n \times Y_m) \in \Omega$$

para todos $n, m \in \mathbb{N}$. Entonces (b) y (d) del lema implican que $\mathcal{M}$ es una clase monótona, y (a) y (c) que $\mathcal{E} \subset \mathcal{M}$. Por el Teorema 8.2 resulta que $\mathcal{M} = \mathcal{S} \otimes \mathcal{T}$. Así $C_{n,m} \in \Omega$ para todo $C \in \mathcal{S} \otimes \mathcal{T}$ y todos $n, m \in \mathbb{N}$. Puesto que C es unión disjunta de los $C_{n,m}$, la propiedad (c) del lema nos permite concluir que $C \in \Omega$. Es decir, $\mathcal{S} \otimes \mathcal{T} \subset \Omega$, lo que prueba el teorema.

Sólo nos queda probar el lema. Si $C = A \times B$, con $A \in \mathcal{S}$, $B \in \mathcal{T}$, entonces $C_x = B$ si $x \in A$, y es vacío en otro caso, mientras que $C^y = A$ si $y \in B$, y vacío en otro caso; por tanto

$$\lambda(C_x) = \lambda(B)\chi_A(x), \quad \text{y} \quad \mu(C^y) = \mu(A)\chi_B(y)$$

e integrando estas igualdades obtenemos

$$\int_X \varphi d\mu = \lambda(B)\mu(A) = \int_Y \psi d\lambda.$$

Esto prueba (a).

Para probar (b), sean $C = \bigcup_{j=1}^{\infty} C_j$, φ y ψ como en el enunciado del teorema (para este C), y

$$\varphi_j(x) = \lambda\left((C_j)_x\right), \ \ \psi_j(y) = \mu\left({C_j}^y\right);$$

entonces recordando (8.1) vemos que se cumplen

$$\varphi_j(x) \nearrow \lambda(C_x) = \varphi(x), \quad \psi_j(y) \nearrow \psi(y),$$

luego, como φ_j y ψ_j cumplen por hipótesis la conclusión del teorema, obtenemos (aplicando el Teorema de la Convergencia Monótona) que

$$\int_X \varphi d\mu = \lim_{j\to\infty} \int_X \varphi_j d\mu = \lim_{j\to\infty} \int_Y \psi_j d\lambda = \int_Y \psi d\lambda;$$

es decir, φ, ψ también cumplen la conclusión del teorema y por tanto $C \in \Omega$.

Para comprobar (c), definamos φ_j, ψ_j como en (b), pero teniendo en cuenta que ahora los C_j son disjuntos; entonces se tiene

$$\psi(y) = \mu\left((\bigcup_{j=1}^{\infty} C_j)^y \right) = \mu\left(\bigcup_{j=1}^{\infty} {C_j}^y \right) = \sum_{j=1}^{\infty} \mu({C_j}^y) = \sum_{j=1}^{\infty} \psi_j(y),$$

y análogamente

$$\varphi(x) = \sum_{j=1}^{\infty} \varphi_j(x);$$

entonces, usando que por hipótesis $\int_X \varphi_j d\mu = \int_Y \psi_j d\lambda$ para cada j, obtenemos

$$\int_X \varphi d\mu = \sum_{j=1}^{\infty} \int_X \varphi_j d\mu = \sum_{j=1}^{\infty} \int_Y \psi_j d\lambda = \int_Y \psi d\lambda,$$

y así $C \in \Omega$.

Finalmente demostremos (d). Por hipótesis $A \times B \supset C_1 \supset C_2 \supset C_3 ...$, y

$$\int_X \varphi_j d\mu = \int_Y \psi_j d\lambda \tag{8.3}$$

para todo $j \in \mathbb{N}$. Definamos $C = \bigcap_{j=1}^{\infty} C_j$. Obsérvese que ${C_j}^y \subset A$, luego $\mu({C_j}^y) \leq \mu(A) < \infty$ para todo j, y por tanto, recordando (8.2), tenemos

$$\psi_j(y) := \mu({C_j}^y) \searrow \mu(C^y) := \psi(y).$$

Análogamente, se ve que

$$\varphi_j(y) := \mu((C_j)_x) \searrow \mu(C_x) := \varphi(x).$$

Observemos también que ${C_j}^y = \emptyset$ si $y \notin B$, ya que $C_j \subset A \times B$; y en todo caso es ${C_j}^y \subset A$. Por tanto, para todo $j \in \mathbb{N}$,

$$\psi_j(y) = \mu({C_j}^y) \leq \mu(A)\chi_B(y) := g_1(y),$$

donde g_1 es integrable en Y ya que $\lambda(B) < \infty$. Análogamente,

$$\varphi_j(x) = \mu\left((C_j)_x\right) \leq \lambda(B)\chi_A(x) := g_2(x)$$

para todo j, con g_2 integrable en X. Entonces podemos aplicar el Teorema de la Convergencia Dominada para, tomando límites cuando $j \to \infty$ en (8.3), concluir que

$$\int_X \varphi d\mu = \int_Y \psi d\lambda,$$

es decir, $C \in \Omega$. Esto completa la prueba del lema y del teorema. $\square$

Definición 8.5. Con la notación e hipótesis del teorema anterior, este nos permite definir, para cada $C \in \mathcal{S} \otimes \mathcal{T}$,

$$(\mu \times \lambda)(C) = \int_X \lambda(C_x) d\mu(x) = \int_Y \mu(C^y) d\lambda(y).$$

Usando las propiedades de la integral, es inmediato comprobar que $\mu \times \lambda$ es una medida en la σ-álgebra producto $\mathcal{S} \otimes \mathcal{T}$. A esta medida se le llama *medida producto de μ y λ.*

Obsérvese que $\mu \times \lambda$ es σ-finita; en efecto, con la notación empleada en la demostración del teorema anterior, tenemos $X \times Y = \bigcup_{n,m} X_n \times Y_m$, con $(\mu \times \lambda)(X_n \times Y_m) = \mu(X_n)\lambda(Y_m) < \infty$.

8.2. El teorema de Fubini-Tonelli

Teorema 8.6 (Fubini-Tonelli). *Sean $(X, \mathcal{S}, \mu)$, $(Y, \mathcal{T}, \lambda)$ dos espacios de medida σ-finitos, y f una función $\mathcal{S} \otimes \mathcal{T}$-medible en $X \times Y$.*

(a) Si $f : X \times Y \to [0, \infty]$, entonces las funciones

$$x \mapsto \varphi(x) := \int_Y f_x d\lambda, \quad y \mapsto \psi(y) := \int_X f^y d\mu \tag{8.4}$$

son $\mathcal{S}$-medible y $\mathcal{T}$-medible, respectivamente, y se tiene

$$\int_X \varphi d\mu = \int_{X \times Y} f d(\mu \times \lambda) = \int_Y \psi d\lambda. \tag{8.5}$$

(b) Si f toma valores reales o complejos y si alguna de las integrales

$$\int_X \left(\int_Y |f|_x d\lambda \right) d\mu(x), \quad \int_Y \left(\int_X |f|_y d\mu \right) d\lambda(y)$$

es finita, entonces ambas son finitas e iguales a $\int_{X \times Y} |f| d(\mu \times \lambda)$, y por tanto $f \in L^1(\mu \times \lambda)$.

(c) Si $f \in L^1(\mu \times \lambda)$, entonces $f_x \in L^1(\lambda)$ para casi todo $x \in X$, $f^y \in L^1(\mu)$ para casi todo $y \in Y$, las funciones φ y ψ definidas por (8.4) *están en $L^1(\mu)$ y $L^1(\lambda)$ respectivamente, y se cumple* (8.5), *es decir,*

$$\int_X \left(\int_Y f(x, y) d\lambda(y) \right) d\mu(x) = \int_{X \times Y} f d(\mu \times \lambda) = \int_Y \left(\int_X f(x, y) d\mu(x) \right) d\lambda(y). \tag{8.6}$$

Demostración. Demostremos (a). Por el Teorema 8.3 las funciones φ y ψ están bien definidas. Dado $C \in \mathcal{S} \otimes \mathcal{T}$, consideremos el caso $f = \chi_C$. En vista de la Definición 8.5, la igualdad (8.5) del teorema es justo la conclusión del Teorema 8.4. Por tanto (a) se cumple para funciones características de conjuntos $\mathcal{S} \otimes \mathcal{T}$-medibles. Por la linealidad de la integral entonces es claro que se cumple también para toda función simple medible $s : X \times Y \to [0, \infty)$. En el caso general, dada $f : X \times Y \to [0, \infty]$ medible, existe una sucesión de funciones simples medibles $s_n : X \times Y \to [0, \infty)$ tales que $s_n(x, y) \nearrow f(x, y)$ para todo $(x, y) \in X \times Y$. Por el Teorema de la Convergencia Monótona

$$\varphi_n(x) := \int_Y (s_n)_x d\lambda \nearrow \int_Y f_x d\lambda := \varphi(x) \tag{8.7}$$

cuando $n \to \infty$ para todo $x \in X$. Y por lo ya sabido tenemos

$$\int_X \varphi_n d\mu = \int_{X \times Y} s_n d(\mu \times \lambda)$$

para todo n. Entonces podemos aplicar otra vez el Teorema de la Convergencia Monótona en la igualdad anterior, con las funciones de (8.7), para concluir que

$$\int_X \varphi d\mu = \lim_{n\to\infty} \int_X \varphi_n d\mu = \lim_{n\to\infty} \int_{X\times Y} s_n d(\mu\times\lambda) = \int_{X\times Y} f d(\mu\times\lambda);$$

es decir se cumple la primera igualdad de (a). La segunda igualdad de (a) se demuestra análogamente cambiando $x \in X$ por $y \in Y$, y μ por λ.

Para probar (b) basta aplicar (a) con $|f|$ en lugar de f.

Para demostrar (c) basta hacerlo en el caso de valores reales (el caso complejo se deduce inmediatamente aplicando el caso real a las partes real e imaginaria de f). Sea pues $f : X \times Y \to \mathbb{R}$, y escribamos $f = f^+ - f^-$, con $f^+ \geq 0$, $f^- \geq 0$. Sean

$$\varphi_1(x) = \int_Y (f^+)_x d\lambda, \quad \varphi_2(y) = \int_Y (f^-)_x d\lambda.$$

Como $f \in L^1(\mu\times\lambda)$ y $f^+ \leq |f|$ y (a) se cumple para f^+, se ve que $\varphi_1 \in L^1(\mu)$. Análogamente, $\varphi_2 \in L^1(\mu)$. Puesto que

$$f_x = (f^+)_x - (f^-)_x,$$

tenemos que $f_x \in L^1(\lambda)$ para todo x para el que se tenga $\varphi_1(x) < \infty$ y $\varphi_2(x) < \infty$, y esto ocurre para casi todo $x \in X$ ya que $\varphi_1, \varphi_2 \in L^1(\mu)$; además para estos x se tiene también $\varphi(x) = \varphi_1(x) - \varphi_2(x)$ por linealidad de la integral. Por tanto $\varphi = \varphi_1 - \varphi_2 \in L^1(\mu)$. Aplicando (8.5) con φ_1 y f^+, se tiene

$$\int_X \left(\int_Y f^+(x,y)d\lambda(y)\right) d\mu(x) = \int_{X\times Y} f^+ d(\mu\times\lambda),$$

y haciendo lo mismo con φ_2 y f^-,

$$\int_X \left(\int_Y f^-(x,y)d\lambda(y)\right) d\mu(x) = \int_{X\times Y} f^- d(\mu\times\lambda).$$

Restando las dos últimas igualdades y usando la linealidad de la integral obtenemos la primera igualdad de (8.6). La segunda igualdad de (8.6) se prueba de forma análoga, con f^y y ψ en lugar de f_x y φ. $\Box$

En la sección de problemas al final del capítulo se encontrarán ejemplos que muestran que las hipótesis del teorema de Fubini no pueden relajarse.

8.3. El teorema de Fubini para la compleción de la medida producto

La medida producto tiene un importante defecto: si $(X, \mathcal{S}, \mu)$, $(Y, \mathcal{T}, \lambda)$ son espacios de medida completos, no es cierto en general que el espacio producto $(X \times Y, \mathcal{S}\otimes\mathcal{T}, \mu\times\lambda)$ sea completo. De hecho este fenómeno ocurre siempre que existan $\emptyset \neq A \in \mathcal{S}$ con $\mu(A) = 0$ y $B \subset Y$ con $B \notin \mathcal{T}$: en esta situación se tiene $A \times B \subset A \times Y$, con $(\mu\times\lambda)(A\times Y) = 0$, pero $A \times B \notin \mathcal{S}\otimes\mathcal{T}$ (porque las secciones de un conjunto medible para la medida producto son siempre conjuntos medibles gracias al Teorema 8.1).

Claramente esto es lo que sucede cuando se consideran $X = Y = \mathbb{R}$, $\mu = \lambda = \mathcal{L}_1$: por tanto la medida producto $\mathcal{L}_1 \times \mathcal{L}_1$ en $\mathbb{R}^2$ no es completa, y en particular no coincide con la medida de Lebesgue $\mathcal{L}_2$ en $\mathbb{R}^2$. Sin embarso sí que es cierto que la compleción de la primera coincide con la segunda. Más en general, se tiene lo siguiente.

Teorema 8.7. *Si $n, m, k \in \mathbb{N}$ y $n = m+k$ entonces $\mathcal{L}_n$ es la compleción de la medida producto $\mathcal{L}_m \times \mathcal{L}_k$ en $\mathbb{R}^m \times \mathbb{R}^k \equiv \mathbb{R}^n$.*

Demostración. Denotemos $\mathcal{B}_k = \mathcal{B}(\mathbb{R}^k)$, sea $\mathcal{M}_k$ la σ-álgebra de los subconjuntos medibles Lebesgue de $\mathbb{R}^k$. Usaremos una notación similar con n o m en lugar de k en $\mathbb{R}^n$ o $\mathbb{R}^m$. Veamos primero que

$$\mathcal{B}_n \subset \mathcal{M}_m \otimes \mathcal{M}_k \subset \mathcal{M}_n. \tag{8.8}$$

En efecto, todo producto de rectángulos está en $\mathcal{M}_m \otimes \mathcal{M}_k$, y $\mathcal{B}_n$ es la σ-álgebra en $\mathbb{R}^n \equiv \mathbb{R}^m \times \mathbb{R}^k$ generada por los productos de rectángulos abiertos. Luego $\mathcal{B}_n \subset \mathcal{M}_m \otimes \mathcal{M}_k$. Por otra parte, si $E \in \mathcal{M}_m$, $F \in \mathcal{M}_k$, es fácil ver, usando el Teorema 3.17 que $E \times \mathbb{R}^k, \mathbb{R}^m \times F \in \mathcal{M}_n$; por tanto su intersección $E \times F$ está también en $\mathcal{M}_n$. Así pues $\mathcal{M}_n$ es una σ-álgebra que contiene a los productos cartesianos de conjuntos $\mathcal{L}_m$ medibles y $\mathcal{L}_k$ medibles, luego por definición de la σ-álgebra producto, se deduce que $\mathcal{M}_m \otimes \mathcal{M}_k \subset \mathcal{M}_n$.

Ahora, dado cualquier $C \in \mathcal{M}_m \otimes \mathcal{M}_k$, ya sabemos que $C \in \mathcal{M}_n$, luego por el Teorema 3.17 existen $C_1, C_2 \in \mathcal{B}_n$ tales que

$$C_1 \subset C \subset C_2 \quad \text{y} \quad \mathcal{L}_n(C_2 \setminus C_1) = 0.$$

Además tanto $\mathcal{L}_n$ como $\mathcal{L}_m \times \mathcal{L}_k$ (restringidas a $\mathcal{B}_n$) son medidas de Borel invariantes por traslaciones y que asignan el valor 1 al cubo unidad, luego por el Teorema 3.19 son iguales sobre los borelianos de $\mathbb{R}^n$. En particular,

$$(\mathcal{L}_m \times \mathcal{L}_k)(C \setminus C_1) \leq (\mathcal{L}_m \times \mathcal{L}_k)(C_2 \setminus C_1) = \mathcal{L}_n(C_2 \setminus C_1) = 0,$$

y por tanto

$$(\mathcal{L}_m \times \mathcal{L}_k)(C) = (\mathcal{L}_m \times \mathcal{L}_k)(C_1) = \mathcal{L}_n(C_1) = \mathcal{L}_n(C).$$

Esto prueba que $\mathcal{L}_m \times \mathcal{L}_k$ coincide con $\mathcal{L}_n$ en $\mathcal{M}_m \otimes \mathcal{M}_k$. Como consecuencia de esto y en vista de (8.8) es claro que la compleción de $\mathcal{M}_m \otimes \mathcal{M}_k$ está contenida en $\mathcal{M}_n$. Por otra parte, $\mathcal{M}_n$ también está contenida en la compleción de $\mathcal{M}_m \otimes \mathcal{M}_k$, por el Teorema 3.17 y el primer contenido de (8.8). Por tanto $(\mathbb{R}^n, \mathcal{M}_n, \mathcal{L}_n)$ es la compleción de $(\mathcal{R}^n, \mathcal{M}_m \otimes \mathcal{M}_k, \mathcal{L}_m \times \mathcal{L}_k)$. [2] □

Ahora podemos enunciar una versión del Teorema de Funibi-Tonelli para la compleción de la medida producto, y su corolario para la medida de Lebesgue en $\mathbb{R}^n$.

Teorema 8.8 (Fubini-Tonelli)**.** *Sean $(X, \mathcal{S}, \mu)$, $(Y, \mathcal{T}, \lambda)$ dos espacios de medida completos y σ-finitos, denotemos $(\mathcal{S} \otimes \mathcal{T})^*$ la compleción de $\mathcal{S} \otimes \mathcal{T}$ relativa a la medida $\mu \times \lambda$, y sea f una función $(\mathcal{S} \otimes \mathcal{T})^*$-medible en $X \times Y$.*

1. *La función f_x es $\mathcal{T}$-medible para casi todo $x \in X$, y la función f^y es $\mathcal{S}$-medible para casi todo $y \in Y$.*

2. *Si $f : X \times Y \to [0, \infty]$, entonces las funciones*

$$x \mapsto \varphi(x) := \int_Y f_x d\lambda, \quad y \mapsto \psi(y) := \int_X f^y d\mu \tag{8.9}$$

están definidas en casi todo punto, son $\mathcal{S}$-medible y $\mathcal{T}$-medible, respectivamente, y se tiene

$$\int_X \varphi d\mu = \int_{X \times Y} f d(\mu \times \lambda) = \int_Y \psi d\lambda. \tag{8.10}$$

(b) Si f toma valores reales o complejos y si alguna de las integrales

$$\int_X \left(\int_Y |f|_x d\lambda \right) d\mu(x), \quad \int_Y \left(\int_X |f|_y d\mu \right) d\lambda(y)$$

es finita, entonces ambas son finitas e iguales a $\int_{X \times Y} |f| d(\mu \times \lambda)$, y por tanto $f \in L^1(\mu \times \lambda)$.

[2] Un argumento similar prueba que $(\mathbb{R}^n, \mathcal{M}_n, \mathcal{L}_n)$ es la compleción de $(\mathbb{R}^n, \mathcal{B}_n, \mathcal{L}_n)$.

(c) *Si $f \in L^1(\mu \times \lambda)$, entonces $f_x \in L^1(\lambda)$ para casi todo $x \in X$, $f^y \in L^1(\mu)$ para casi todo $y \in Y$, las funciones φ y ψ definidas por* (8.9) *están en $L^1(\mu)$ y $L^1(\lambda)$ respectivamente, y se cumple* (8.10)*, es decir,*

$$\int_X \left(\int_Y f(x,y) d\lambda(y) \right) d\mu(x) = \int_{X \times Y} f d(\mu \times \lambda) = \int_Y \left(\int_X f(x,y) d\mu(x) \right) d\lambda(y). \quad (8.11)$$

Corolario 8.9. *Sean $n = m + k$ y f una función medible en $\mathbb{R}^m \times \mathbb{R}^k \equiv \mathbb{R}^n$. Entonces:*

1. *La función f_x es medible para casi todo $x \in \mathbb{R}^m$, y la función f^y es medible para casi todo $y \in \mathbb{R}^k$.*

2. *Si $f : \mathbb{R}^n \to [0, \infty]$, entonces las funciones*

$$x \mapsto \varphi(x) := \int_{\mathbb{R}^k} f_x dx, \quad y \mapsto \psi(y) := \int_{\mathbb{R}^m} f^y dy \quad (8.12)$$

están definidas en casi todo punto, son medibles, y se tiene

$$\int_{\mathbb{R}^m} \varphi(x) dx = \int_{\mathbb{R}^n} f = \int_{\mathbb{R}^k} \psi(y) dy. \quad (8.13)$$

(b) *Si f toma valores reales o complejos y si alguna de las integrales*

$$\int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^k} |f(x,y)| dy \right) dx, \quad \int_{\mathbb{R}^k} \left(\int_{\mathbb{R}^m} |f(x,y)| dx \right) dy$$

es finita, entonces ambas son finitas e iguales a $\int_{\mathbb{R}^n} |f|$, y por tanto $f \in L^1(\mathbb{R}^n)$.

(c) *Si $f \in L^1(\mathbb{R}^n)$, entonces $f_x \in L^1(\mathbb{R}^k)$ para casi todo $x \in \mathbb{R}^m$, $f^y \in L^1(\mathbb{R}^m)$ para casi todo $y \in \mathbb{R}^k$, las funciones φ y ψ definidas por* (8.12) *están en $L^1(\mathbb{R}^m)$ y $L^1(\mathbb{R}^k)$ respectivamente, y se cumple* (8.13)*, es decir,*

$$\int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^k} f(x,y) dy \right) dx = \int_{\mathbb{R}^n} f = \int_{\mathbb{R}^k} \left(\int_{\mathbb{R}^m} f(x,y) dx \right) dy. \quad (8.14)$$

Demostración del Teorema 8.8. La demostración depende de los dos lemas siguientes.

Lema 8.10. *Si ν es una medida en una σ-álgebra $\mathcal{A}$, $\mathcal{A}^*$ denota la compleción de $\mathcal{A}$ relativa a la medida ν, y f es una función $\mathcal{A}^*$-medible, entonces existe una función $\mathcal{A}$-medible g tal que $f = g$ en ν-casi todo punto.*

Lema 8.11. *Sean μ, λ medidas completas en $(X, \mathcal{S})$ e $(Y, \mathcal{T})$ respectivamente, y sea h una función $(\mathcal{S} \otimes \mathcal{T})^*$-medible en $X \times Y$. Supongamos que $h = 0$ en $(\mu \times \lambda)$-casi todo punto de $X \times Y$. Entonces, para casi todo $x \in X$ se cumple que $h(x,y) = 0$ para casi todo $y \in Y$; en particular, la sección h_x es $\mathcal{T}$-medible para casi todo $x \in X$. Y un enunciado análogo es verdad para la sección h^y.*

Supuestos estos dos lemas, la demostración del Teorema 8.8 es fácil: si f es como en el enunciado del teorema, el Lema 8.10 (con $\nu = \mu \times \lambda$) muestra que $f = g + h$, con $h = 0$ en casi todo punto de $X \times Y$, y g es $\mathcal{S} \otimes \mathcal{T}$-medible. El Teorema 8.6 se aplica a g, y el Lema 8.11 nos dice que, para casi todo $x \in X$, se tiene $f_x(y) = g_x(y)$ para casi todo $y \in Y$, y similarmente que $f^y = g^y$ en casi todo punto, para casi todo $y \in Y$ (en particular, como todas las secciones de g son medibles, casi todas las secciones de f serán también medibles). Por lo tanto las integrales iteradas de f, así como la integral total (doble), son las mismas que las de g, y puesto que las igualdades del teorema se cumplen para g, entonces también son verdad para f.

Demostración del Lema 8.10. Si $f : X \to [0, \infty]$ es $\mathcal{A}^*$-medible, existe una sucesión monótona creciente (s_n) de funciones simples $\mathcal{A}^*$-medibles positivas que converge a f puntualmente. Así $f = \sum_{n=1}^{\infty}(s_{n+1} - s_n)$, y como cada $s_{n+1} - s_n$ es combinación lineal finita de funciones características (necesariamente con coeficientes positivos, si escribimos la función simple $s_{n+1} - s_n$ en forma canónica), se sigue que

$$f(x) = \sum_{j=1}^{\infty} c_j \chi_{E_j}(x), \quad x \in X$$

para ciertas constantes $c_j > 0$ y conjuntos $E_j \in \mathcal{A}^*$ (no necesariamente disjuntos). Por definición de $\mathcal{A}^*$, para cada $j \in \mathbb{N}$ existen $A_j, B_j \in \mathcal{A}$ tales que $A_j \subset E_j \subset B_j$ y $\nu(B_j \setminus A_j) = 0$. Definamos

$$g(x) = \sum_{j=1}^{\infty} c_j \chi_{A_j}(x), \quad x \in X.$$

Entonces g es $\mathcal{A}$-medible, y $g(x) = f(x)$ para todo $x \in X \setminus \bigcup_{j=1}^{\infty}(E_j \setminus A_j)$. Puesto que $\bigcup_{j=1}^{\infty}(E_j \setminus A_j) \subset \bigcup_{j=1}^{\infty}(B_j \setminus A_j)$ y $\nu(B_j \setminus A_j) = 0$ para todo j, concluimos que $g = f$ en ν-casi todo punto.

En el caso en que f toma valores o complejos se deduce fácilmente del caso $f \geq 0$ por el procedimiento estándar (considerar $f = f^+ - f^-$, aplicar el resultado anterior a $f^{\pm}$, etc). □

Demostración del Lema 8.10. Sea $E = \{(x, y) \in X \times Y : h(x, y) \neq 0\}$. El conjunto E pertenece a $(\mathcal{S} \otimes \mathcal{T})^*$, y $(\mu \times \lambda)(E) = 0$, luego existe un conjunto $Z \in \mathcal{S} \otimes \mathcal{T}$ tal que $E \subset Z$ y $(\mu \times \lambda)(Z) = 0$. Por el Teorema 8.4,

$$\int_X \lambda(Z_x) d\mu(x) = 0,$$

luego, si $N := \{x \in X : \lambda(Z_x) > 0\}$, se tiene $\mu(N) = 0$. Como $E_x \subset Z_x$ y λ es una medida completa en $(Y, \mathcal{T})$, se sigue que $E_x \in \mathcal{T}$ para todo $x \in X \setminus N$. En consecuencia para casi todo $x \in X$ el conjunto $E_x = \{y \in Y : h(x, y) \neq 0\} = \{y \in Y : h_x(y) \neq 0\}$ tiene medida cero en $(Y, \mathcal{T}, \lambda)$, que es lo que queríamos demostrar. Además, puesto que λ es completa, que h_x sea igual a cero en casi todo punto implica que es medible (sin necesidad de redefinirla en un conjunto de medida cero). La prueba del enunciado para h^y es análoga. □

Esto completa la demostración del Teorema 8.8. □

8.4. Problemas

Problema 8.1. Dar un ejemplo de clase monótona que no sea σ-álgebra.

Problema 8.2. Sean μ, λ medidas σ-finitas en X e Y respectivamente, y E medible en $X \times Y$. Si para casi todo $x \in X$ es $\lambda(E_x) = 0$, probar que para casi todo $y \in Y$ es $\mu(E^y) = 0$.

Problema 8.3. Sean $(X, \mathcal{S})$, $(Y, \mathcal{T})$ espacios de medida, y $f : X \to \mathbb{R}$, $g : Y \to \mathbb{R}$ medibles. Probar que la función $h : X \times Y \to \mathbb{R}$ definida por $h(x, y) = f(x)g(y)$ es $\mathcal{S} \otimes \mathcal{T}$-medible.

Problema 8.4. Sean $(X, \mathcal{A})$ un espacio medible, y $f : X \to [0, \infty)$ una función. Demostrar que $\{(x, y) \in X \times \mathbb{R} : 0 < y < f(x)\}$ es $\mathcal{A} \otimes \mathcal{B}(\mathbb{R})$-medible si y sólo si f es $\mathcal{A}$-medible.

Problema 8.5. Sean $(X, \mathcal{A})$ un espacio medible, y $f : X \to \mathbb{R}$ una función. Demostrar que la gráfica de f es $\mathcal{A} \otimes \mathcal{B}(\mathbb{R})$-medible si f es $\mathcal{A}$-medible.

Problema 8.6. Probar que $\mathcal{B}(\mathbb{R}^n) \otimes \mathcal{B}(\mathbb{R}^m) = \mathcal{B}(\mathbb{R}^{n+m})$.

Problema 8.7. Sean $X = Y = [0,1]$, $\mu = \lambda = \mathcal{L}_1$ en $[0,1]$. Sea $\{\delta_n\}_{n\in\mathbb{N}} \subset [0,1]$ una sucesión tal que $\delta_1 = 0$, $\delta_n < \delta_{n+1}$ para todo n, y $\lim_{n\to\infty} \delta_n = 1$. Sea g_n una función continua con soporte en (δ_n, δ_{n+1}) tal que $\int_0^1 g_n(t)dt = 1$ para todo n, y definamos

$$f(x,y) = \sum_{n=1}^{\infty} (g_n(x) - g_{n+1}(x))\, g_n(y).$$

Comprobar que f es continua en $(X \times Y) \setminus \{(1,1)\}$, y que $\int_0^1 \left(\int_0^1 f(x,y)dy\right) dx = 1 \neq 0 = \int_0^1 \left(\int_0^1 f(x,y)dx\right) dy$.

Problema 8.8. Sea $f : [0,1]\times[0,1] \to \mathbb{R}$ definida por $f(x,y) = (x^2-y^2)/(x^2+y^2)^2$ si $(x,y) \neq (0,0)$, y $f(0,0) = 0$. Comprobar que las itegrales iteradas de f existen y son distintas.

Problema 8.9. Sea $f : \mathbb{R}^2 \to \mathbb{R}$ definida por $f(x,y) = 1$ si $x \geq 0$, $x \leq y < x+1$; por $f(x,y) = -1$ si $x \geq 0$, $x+1 \leq y < x+2$; y por $f(x,y) = 0$ en cualquier otro caso. Comprobar que las itegrales iteradas de f existen y son distintas.

Problema 8.10. Si $f : X \to \mathbb{R}$, $g : Y \to \mathbb{R}$ son integrables (donde X, Y son espacios de medida σ-finitos), demostrar que $h(x,y) = f(x)g(y)$ es integrable en $X \times Y$, y calcular su integral.

Problema 8.11. Sea $f(x,y) = xy/(x^2+y^2)^2$. Demostrar que las integrales iteradas de f existen y son iguales en $[-1,1] \times [-1,1]$, pero f no es integrable.

Problema 8.12. Integrando $e^{-sxy} \operatorname{sen} x$, probar que para $s > 0$ se tiene

$$\int_0^{\infty} e^{-sx} \frac{\operatorname{sen} x}{x} dx = \frac{1}{s} \arctan(1/s).$$

Problema 8.13. Usar el teorema de Fubini, y el hecho de que $1/x = \int_0^\infty e^{-xt}dt$ si $x > 0$, para demostrar que

$$\lim_{a\to\infty} \int_0^a \frac{\operatorname{sen} x}{x} dx = \frac{\pi}{2}.$$

Problema 8.14. Encontrar un subconjunto convexo $C \subset \mathbb{R}^2$ tal que las secciones C_x y C^y son borelianos de $\mathbb{R}$ para todos $x, y \in \mathbb{R}$, pero $C \notin \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$.

Problema 8.15. Si $f : \mathbb{R}^n \to \mathbb{R}$ es medible, probar que $\mathcal{L}^{n+1}\left(\{(x,y) \in \mathbb{R}^{n+1} : y = f(x)\}\right) = 0$.

Problema 8.16. Sean μ una medida σ-finita en $(X, \mathcal{A})$, y $f : X \to [0,\infty]$ $\mathcal{A}$-medible. Demostrar que $\int_X f d\mu = \int_0^\infty \mu\left(\{x : f(x) \geq t\}\right) dt$. *Indicación:* Considerar $E := \{(x,s) \in X \times [0,\infty) : f(x) \geq s\}$, y calcular $(\mu \times \mathcal{L}^1)(E)$.

Problema 8.17. Transcribir el enunciado del teorema de Fubini para $X = Y = \mathbb{N}$, $\mathcal{S} = \mathcal{T} = 2^{\mathbb{N}}$, $\mu = \lambda = \sharp$.

Problema 8.18. Sean $X = Y = [0,1]$, $\mu = \mathcal{L}^1$ en X, $\lambda = \sharp$ en Y, y definamos $f : X \times Y \to [0,\infty)$ por $f(x,y) = 1$ si $x = y$, y $f(x,y) = 0$ en otro caso. Comprobar que el teorema de Fubini falla para f. ¿Qué hipótesis de este teorema no se cumple en este caso?

Problema 8.19. Sean $(X, \mathcal{S}, \mu)$, $(Y, \mathcal{T}, \lambda)$ espacios de medida σ-finitos, y supongamos que ν es una medida en $\mathcal{S} \otimes \mathcal{T}$ tal que $\nu(A \times B) = \mu(A)\lambda(B)$ para todos $A \in \mathcal{S}$, $B \in \mathcal{T}$. Probar que $\nu = \mu \times \lambda$.

Problema 8.20. Sea $\mathbb{B}_n = \{x \in \mathbb{R}^n : |x| \leq 1\}$, es decir la bola unidad (euclídea) de $\mathbb{R}^n$. Demostrar que

$$\mathcal{L}_n(\mathbb{B}_n) = \begin{cases} \frac{\pi^{n/2}}{(n/2)!} & \text{si } n \text{ es par,} \\ \frac{2^{(n+1)/2}\pi^{(n-1)/2}}{1\cdot 3\cdot 5\cdot\dots\cdot n} & \text{si } n \text{ es impar.} \end{cases}$$

Indicación: Para $n = 1$ se tiene, obviamente, $\mathcal{L}_1(\mathbb{B}_1) = 2$. Para $n = 2$,

$$\mathcal{L}_2(\mathbb{B}_2) = 2\int_{-1}^{1} \sqrt{1-x^2}dx = \pi,$$

y por tanto la fórmula se cumple para $n = 1, 2$. Por inducción, suponer que la fórmula es verdad para todos los números naturales menores que n, donde $n \geq 3$, y demostrar que entonces es verdad para n usando el teorema de Fubini como sigue:

$$\begin{aligned}
&\mathcal{L}_n(\mathbb{B}_n) = \int_{\mathbb{R}^2} \left(\int_{\mathbb{R}^{n-2}} \chi_{\mathbb{B}_n}(x, y)dy \right) dx = \int_{\mathbb{R}^2} \mathcal{L}_{n-2} \left((1 - x_1^2 - x_2^2)^{1/2} \mathbb{B}_{n-2} \right) \chi_{\mathbb{B}_2}(x_1, x_2) dx_1 dx_2 \\
&= \int_{\mathbb{R}^2} (1 - x_1^2 - x_2^2)^{(n-2)/2} \mathcal{L}_{n-2}\left(\mathbb{B}_{n-2}\right) \chi_{\mathbb{B}_2}(x_1, x_2) dx_1 dx_2 \\
&= \mathcal{L}_{n-2}\left(\mathbb{B}_{n-2}\right) \int_{\mathbb{B}_2} (1 - x_1^2 - x_2^2)^{(n-2)/2} dx_1 dx_2 = \mathcal{L}_{n-2}\left(\mathbb{B}_{n-2}\right) \frac{2\pi}{n}.
\end{aligned}$$

Para evaluar la última integral en $\mathbb{B}_2$ puede emplearse también el teorema de Fubini y las técnicas de cálculo de primitivas en $\mathbb{R}$ (si no quiere usarse el teorema del cambio de variable con coordenadas polares, lo que resulta obviamente más fácil).

Problema 8.21. Usando la función Γ y el problema anterior, probar que

$$\mathcal{L}_n(\mathbb{B}_n) = \frac{\pi^{n/2}}{\Gamma\left(1 + \frac{n}{2}\right)}$$

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

Capítulo 9

Espacios L^p y convolución integral

9.1. Los espacios L^p

Sea $(X, \mathcal{A}, \mu)$ un espacio de medida. Para $1 \leq p < \infty$, sea

$$L^p(X) = \{f : X \to \mathbb{C} \mid f \text{ es medible y } \int_X |f|^p d\mu < \infty\},$$

y consideremos la relación de equivalencia $f \sim g \iff f = g$ en casi todo punto de X. Por abuso de notación, denotaremos también $L^p(X)$ como el conjunto cociente de $\mathcal{L}^p(X)$ por esta relación de equivalencia.[1]

Para cada $f : X \to \mathbb{C}$ medible ponemos

$$\|f\|_p = \left(\int_X |f|^p d\mu \right)^{1/p},$$

número bien definido en $[0, \infty]$, que será finito siempre y cuando $f \in L^p(X)$.

Si $p = \infty$, definimos $L^\infty(X)$ como el conjunto de funciones medibles *esencialmente acotadas*, es decir, las funciones medibles $f : X \to \mathbb{C}$ para las que existe $M > 0$ tal que

$$|f(x)| \leq M \text{ en ctp } x \in X. \tag{9.1}$$

Nótese si definimos $\|f\|_\infty$ como el ínfimo de tales $M > 0$, el número $\|f\|_\infty$ aún cumple (9.1): en efecto, por definición de ínfimo existe una sucesión $(M_j)_{j\in\mathbb{N}}$ tal que $M_j \searrow \|f\|_\infty$ y $|f(x)| \leq M_j$ para todo $x \in X \setminus A_j$, donde $\mu(A_j) = 0$, y entonces $|f(x)| \leq M_j$ para todo $x \in X \setminus \bigcup_{i=1}^\infty A_i$ y todo $j \in \mathbb{N}$, luego $|f(x)| \leq \lim_{j\to\infty} M_j = \|f\|_\infty$ para todo $x \in X \setminus \bigcup_{i=1}^\infty A_i$, donde $\mu\left(\bigcup_{i=1}^\infty A_i\right) = 0$. Obsérvese también que, por definición de ínfimo, para todo $\varepsilon > 0$ se tiene

$$\mu\left(\{x \in X : |f(x)| \geq \|f\|_\infty - \varepsilon\}\right) > 0.$$

En el caso $X = \mathbb{N}$ con la medida de contar, es tradicional denotar $\ell_p = L^p(\mathbb{N})$ para todo $p \in [1, \infty]$. Obsérvese que en este caso las funciones de $\mathbb{N}$ en $\mathbb{C}$ se identifican con sucesiones de $\mathbb{C}$, y se obtiene

$$\ell_p = \{z = (z_n)_{n\in\mathbb{N}} : \|z\|_p = \left(\sum_{n=1}^\infty |z_n|^p \right)^{1/p} < \infty\},$$

y

$$\ell_\infty = \{z = (z_n)_{n\in\mathbb{N}} : \|z\|_\infty = \sup_{n\in\mathbb{N}} |z_n| < \infty\},$$

A veces también usaremos $L^p(X)$ para denotar los mismos conjuntos de funciones pero con valores reales; por el contexto quedará claro si los valores son complejos o reales. Por ejemplo, el siguiente teorema se aplica sólo a funciones con valores reales.

[1] Si no decimos nada, por el contexto quedará claro si estamos o no identificando funciones que coinciden en casi todo punto (o dará lo mismo si lo hacemos o no porque las integrales de funciones que difieren en conjuntos de medida cero son iguales).

Teorema 9.1 (Desigualdad de Jensen). *Supongamos que $\mu(X) = 1$, φ es una función convexa definida en un intervalo abierto (posiblemente infinito) $I \subseteq \mathbb{R}$, y que $f \in L^1(X)$ cumple $f(x) \in I$ para casi todo $x \in X$. Entonces*

$$\varphi\left(\int_X f d\mu\right) \leq \int_X (\varphi \circ f) d\mu.$$

Demostración. Como φ es convexa, para cada $p \in I$ existe $r \in \mathbb{R}$ tal que

$$\varphi(q) \geq \varphi(p) + r(q - p) \text{ para todo } q \in \mathbb{R}.$$

Pongamos $q = f(x)$ y $p = \int_X f d\mu$; obsérvese que $p \in I$ ya que $f(x) \in I$ para casi todo $x \in X$ y $\mu(X) = 1$. Entonces de la desigualdad anterior obtenemos

$$\varphi(f(x)) \geq \varphi\left(\int_X f\right) + r\left(f(x) - \int_X f\right),$$

e integrando esta desigualdad concluimos que $\int_X (\varphi \circ f) d\mu \geq \varphi\left(\int_X f d\mu\right)$, ya que $\mu(X) = 1$. $\square$

Para cada $A \subset X$ medible con $0 < \mu(A) < \infty$, denotemos

$$⨍_A f = \frac{1}{\mu(A)} \int_A f.$$

Corolario 9.2. *Supongamos que $0 < \mu(A) < \infty$, φ es una función convexa definida en un intervalo abierto (posiblemente infinito) $I \subseteq \mathbb{R}$, y que $f \in L^1(A)$ cumple $f(x) \in I$ para casi todo $x \in A$. Entonces*

$$\varphi\left(⨍_A f\right) \leq ⨍_X \varphi \circ f.$$

Demostración. Basta sustituir la medida μ por $\nu = \frac{1}{\mu(A)}\mu$ y aplicar el teorema anterior con esta última medida, observando que $\int_A f d\nu = \frac{1}{\mu(A)} \int_A f d\mu$. $\square$

Lema 9.3 (Desigualdad de Young). *Si $p, q \in (1, \infty)$, $\frac{1}{p} + \frac{1}{q} = 1$ y $a, b \geq 0$ entonces*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Demostración. Podemos suponer $a, b > 0$ ya que si uno de los dos es 0 entonces la desigualdad es trivial. Puesto que la función exponencial es convexa en $\mathbb{R}$, se tiene

$$ab = e^{\log a + \log b} = e^{\frac{1}{p}\log a^p + \frac{1}{q}\log b^q} \leq \frac{1}{p} e^{\log a^p} + \frac{1}{q} e^{\log b^q} = \frac{a^p}{p} + \frac{b^q}{q}.$$

$\square$

Teorema 9.4 (Desigualdad de Hölder). *Si $p, q \in (1, \infty)$, $\frac{1}{p} + \frac{1}{q} = 1$, entonces para todas $f, g : X \to \mathbb{C}$ medibles se tiene*

$$\int_X |fg| d\mu \leq \|f\|_p \|g\|_q.$$

Demostración. Podemos suponer $\|f\|_p, \|g\|_g \in (0, \infty)$ ya que en los demás casos la desigualdad es trivial. Por el lema anterior tenemos

$$\frac{|f(x)g(x)|}{\|f\|_p \|g\|_q} \leq \frac{1}{p}\left(\frac{|f(x)|}{\|f\|_p}\right)^p + \frac{1}{q}\left(\frac{|g(x)|}{\|g\|_q}\right)^q,$$

e integrando esta desigualdad concluimos

$$\int_X |fg| d\mu \leq \|f\|_p \|g\|_q \left(\frac{1}{p\|f\|_p^p} \int_X |f|^p d\mu + \frac{1}{q\|g\|_q^q} \int_X |g|^q d\mu\right) = \|f\|_p \|g\|_q$$

ya que $1/p + 1/q = 1$. $\square$

Teorema 9.5 (Desigualdad de Minkowski). *Si $1 \leq p \leq \infty$ se tiene, para cualesquiera $f, g : X \to \mathbb{C}$ medibles,*

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p.$$

Demostración. Si $p = 1$ la desigualdad se deduce inmediatamente de la desigualdad triangular para el módulo y las propiedades elementales de la integral:

$$\int_X |f+g| \leq \int_X (|f|+|g|) = \int_X |f| + \int_X |g|.$$

Para $p = \infty$ se tiene también

$$|f(x)+g(x)| \leq |f(x)| + |g(x)| \leq \|f\|_\infty + \|g\|_\infty$$

para casi todo $x \in X$, de donde

$$\|f+g\|_\infty \leq \|f\|_\infty + \|g\|_\infty.$$

El caso $1 < p < \infty$ lo vamos a deducir de la desigualdad de Hölder:

$$\begin{aligned}
&\int_X |f+g|^p \leq \int_X |f||f+g|^{p-1} + \int_X |g||f+g|^{p-1} \\
&\leq \left(\int_X |f|^p\right)^{1/p} \left(\int_X |f+g|^{(p-1)q}\right)^{1/q} + \left(\int_X |g|^p\right)^{1/p} \left(\int_X |f+g|^{(p-1)q}\right)^{1/q} \\
&\leq (\|f\|_p + \|g\|_p) \left(\int_X |f+g|^p\right)^{1/q},
\end{aligned}$$

donde hemos usado que $(p-1)q = p$. Se deduce entonces que

$$\left(\int_X |f+g|^p\right)^{1-1/q} \leq \|f\|_p + \|g\|_p,$$

que es justo lo que queremos porque $1 - 1/q = 1/p$. □

Recordemos que un espacio vectorial normado $(E, \|\cdot\|)$ se dice que es de Banach cuando es completo, es decir, cuando toda sucesión de Cauchy en E es convergente.

Teorema 9.6. *Para cualquier $p \in [1, \infty]$, el espacio $L^p(X)$ es un espacio de Banach.*

Demostración. Para $p = \infty$ la demostración es mucho más fácil y se deja como ejercicio. Supongamos pues $p \in [1, \infty)$. Por la desigualdad de Minkowski, es obvio que $L^p(X)$ es un espacio vectorial normado. La propiedad $\|f\|_p = 0 \iff f = 0$ se tiene porque identificamos funciones que coinciden en casi todo punto. La parte no obvia del teorema es probar que $(L^p(X), \|\cdot\|_p)$ es completo.

Sea (f_n) una sucesión de Cauchy en $L^p(X)$. Tomemos una subsucesión (f_{n_j}) de (f_n) tal que

$$\|f_{n_{j+1}} - f_{n_j}\|_p < \frac{1}{2^j}$$

para todo $j \in \mathbb{N}$. Definamos

$$g_k = \sum_{j=1}^{k} |f_{n_{j+1}} - f_{n_j}|, \text{ y } g = \sum_{j=1}^{\infty} |f_{n_{j+1}} - f_{n_j}|.$$

Por la desigualdad de Minkowski se tiene que $\|g_k\|_p < 1$ para todo k, y aplicando el Lema de Fatou a la sucesión g_k^p, que converge puntualmente a g^p, se deduce

$$\int_X g^p \leq \liminf_{k\to\infty} \int_X g_k^p \leq 1,$$

luego también $\|g\|_p \leq 1$. En particular g es finita en casi todo punto, y esto implica que la serie de funciones

$$f_{n_1}(x) + \sum_{j=1}^{\infty} \big(f_{n_{j+1}}(x) - f_{n_j}(x)\big)$$

converge absolutamente para casi todo $x \in X$. Puesto que

$$f_{n_1}(x) + \sum_{j=1}^{k} \big(f_{n_{j+1}}(x) - f_{n_j}(x)\big) = f_{n_{k+1}}(x),$$

se deduce que $(f_{n_k}(x))$ converge para casi todo $x \in X$. Definamos

$$f(x) = \lim_{k\to\infty} f_{n_k}(x)$$

donde dicha subsucesión de funciones converge (que es en casi todo $x \in X$), y por ejemplo $f(x) = 0$ en el resto de puntos.

Ahora vamos a probar que $f \in L^p(X)$ y que (f_n) converge a f en $L^p(X)$. Por ser (f_n) de Cauchy, para cada $\varepsilon > 0$ existe $N \in \mathbb{N}$ tal que $\|f_n - f_m\|_p \leq \varepsilon$ si $n, m \geq N$. Dejando $m \geq N$ fijo por el momento y tomando $n = n_j$ y haciendo $j \to \infty$, el Lema de Fatou implica que

$$\int_X |f - f_m|^p \leq \liminf_{j\to\infty} \int_X |f_{n_j} - f_m|^p \leq \varepsilon^p;$$

en particular $f - f_m \in L^p(X)$, y como $f_m \in L^p(X)$ se obtiene $f \in L^p(X)$. Además la desigualdad anterior equivale a

$$\|f - f_m\|_p \leq \varepsilon$$

para todo $m \geq N$, y el argumento anterior muestra así también que (f_m) converge a f en $L^p(X)$. □

Corolario 9.7. *Si (f_n) converge a f en $L^p(X)$, entonces existe una subsucesión (f_{n_j}) que converge a f en casi todo punto $x \in X$.*

Demostración. Esto se ha visto en la demostración del teorema anterior. □

En lo que sigue, si X es un espacio métrico localmente compacto, denotaremos

$$C_c(X) = \{f : X \to \mathbb{C} : f \text{ es continua con soporte compacto}\},$$

donde el soporte de una función f se define como la adherencia del conjunto $\{x \in X : f(x) \neq 0\}$.

Teorema 9.8. *Sean X un espacio métrico localmente compacto, y μ una medida de Radon en X. Si $1 \leq p < \infty$ entonces el espacio $C_c(X)$ es denso en $L^p(X)$.*

Demostración. Definamos S como el conjunto de las funciones simples medibles $s : X \to \mathbb{C}$ tales que $\mu(\{x \in X : s(x) \neq 0\}) < \infty$.

Lema 9.9. *Para $1 \leq p < \infty$, S es denso en $L^p(X)$.*

Demostración. Es claro que $S \subset L^p(X)$. Si $f \in L^p$ y $f \geq 0$, sabemos que existe una sucesión de funciones simples medibles $0 \leq s_n \nearrow f$. Puesto que $0 \leq \int_X s_n^p \leq \int_X f^p < \infty$, se tiene $s_n \in L^p(X)$, y como s_n sólo toma una cantidad finita de valores, necesariamente los conjuntos donde estos valores son distintos de 0 tienen medida finita; por tanto, $s_n \in S$ para cada $n \in \mathbb{N}$. Además $0 \leq |f - s_n|^p \leq f^p$, y entonces por el Teorema de la Convergencia Dominada se deduce que $\int_X |f - s_n|^p \to 0$, es decir, (s_n) converge a f en $L^p(X)$. Esto prueba lo que queremos cuando $f \geq 0$.

En el caso general escribimos $f = (u^+ - u^-) + i(v^+ - v^-)$ y aplicamos el mismo argumento a u^+, u^-, v^+, v^- separadamente, obteniendo sucesiones de funciones $(\varphi_n), (\psi_n), (h_n), (k_n)$ en S que convergen, respectivamente, a u^+, u^-, v^+, v^-. Entonces la sucesión $(\varphi_n - \psi_n) + i(h_n - k_n)$ está en S y converge a f en L^p. □

Seguimos con la demostración del teorema. Usando el lema precedente, basta probar el resultado en el caso $f = \chi_E$, donde $E \subset X$ es medible y $\mu(E) < \infty$. En este caso, dado $\varepsilon > 0$, como μ es de Radon podemos encontrar un conjunto compacto $K \subset E$ tal que $\mu(E \setminus K) < (\varepsilon/2)^p$, y un conjunto abierto U tal que $K \subset U$, $\overline{U}$ es compacto, y $\mu(U \setminus K) < (\varepsilon/2)^p$. Por el Lema de Urysohn existe entonces $\varphi \in C_c(X)$ con soporte contenido en U, $0 \leq \varphi \leq 1$, y $\varphi(x) = 1$ para todo $x \in K$, y se concluye que

$$\begin{aligned}\|\chi_E - \varphi\|_p &= \left(\int_{X\setminus K} |\chi_E - \varphi|^p\right)^{1/p} \leq \left(\int_{X\setminus K} |\chi_E|^p\right)^{1/p} + \left(\int_{X\setminus K} |\varphi|^p\right)^{1/p} \\ &\leq \mu(E \setminus K)^{1/p} + \mu(U \setminus K)^{1/p} < \varepsilon.\end{aligned}$$

Esto completa la demostración del teorema. □

Corolario 9.10. *El espacio $C_c(\mathbb{R}^n)$ es denso en $L^p(\mathbb{R}^n)$.*

9.2. Convolución integral de funciones

La convolución de dos funciones $f, g : \mathbb{R}^n \to \mathbb{C}$ se define por

$$f * g(x) = \int_{\mathbb{R}^n} f(x-y)g(y)dy = \int_{\mathbb{R}^n} f(y)g(x-y)dy$$

cuando esta integral existe.

En esta sección, salvo mención expresa en sentido contrario, L^p denotará $L^p(\mathbb{R}^n)$, y C^0 el espacio de las funciones continuas sobre $\mathbb{R}^n$.

Al tratar con convoluciones la siguiente desigualdad resulta bastante útil.

Teorema 9.11 (Desigualdad de Minkowski para integrales)**.** *Sea $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{C}$ una función medible, y $1 \leq p < \infty$. Entonces:*

1. *Si $f \geq 0$, se tiene*

$$\left(\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^m} f(x,y)dy\right)^p dx\right)^{\frac{1}{p}} \leq \int_{\mathbb{R}^m}\left(\int_{\mathbb{R}^n} f(x,y)^p dx\right)^{\frac{1}{p}} dy.$$

2. *Si las secciones $y \mapsto f(x,y)$ son integrables para casi todo x, entonces*

$$\left(\int_{\mathbb{R}^n}\left|\int_{\mathbb{R}^m} f(x,y)dy\right|^p dx\right)^{\frac{1}{p}} \leq \int_{\mathbb{R}^m}\left(\int_{\mathbb{R}^n} |f(x,y)|^p dx\right)^{\frac{1}{p}} dy.$$

Demostración. Es claro que (2) es conseuencia de (1). Por otra parte, en el caso $p = 1$, (1) no es más que el teorema de Fubini para funciones positivas. Podemos suponer pues $f \geq 0$ y $1 < p < \infty$. Denotemos el exponente conjugado de p por q; es decir, $\frac{1}{p} + \frac{1}{q} = 1$ (y en particular $p - 1 = p/q$). Definamos la función $g(x) = \int_{\mathbb{R}^m} f(x,y)dy$. Tenemos entonces, aplicando sucesivamente el teorema de Fubini y la desigualdad de Hölder,

$$\begin{aligned}&\int_{\mathbb{R}^n} g(x)^p dx = \int_{\mathbb{R}^n} g(x)g(x)^{p-1}dx = \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^m} f(x,y)dy\right) g(x)^{p-1}dx = \\ &\int_{\mathbb{R}^m}\left(\int_{\mathbb{R}^n} f(x,y)g(x)^{p-1}dx\right) dy \leq \int_{\mathbb{R}^m}\left(\int_{\mathbb{R}^n} f(x,y)^p dx\right)^{\frac{1}{p}}\left(\int_{\mathbb{R}^n} g(x)^p dx\right)^{\frac{1}{q}} dy = \\ &\left(\int_{\mathbb{R}^n} g(x)^p dx\right)^{\frac{1}{q}}\int_{\mathbb{R}^m}\left(\int_{\mathbb{R}^n} f(x,y)^p dx\right)^{\frac{1}{p}} dy,\end{aligned}$$

de donde se obtiene la desigualdad deseada. □

Con la misma demostración puede obtenerse una versión del resultado anterior para un producto de espacios de medida σ-finitos.

Teorema 9.12 (Desigualdad de Minkowski para integrales)**.** *Sean $(X, \mathcal{S}, \mu)$ e $(Y, \mathcal{T}, \lambda)$ espacios de medida σ-finitos, y sea $f : X \times Y \to \mathbb{C}$ una función $\mathcal{S} \otimes \mathcal{T}$-medible (o bien, más en general, medible respecto de la compleción de la medida producto), y $1 \leq p < \infty$. Entonces:*

1. *Si $f \geq 0$, se tiene*
$$\left(\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^m} f(x,y)dy\right)^p dx\right)^{\frac{1}{p}} \leq \int_{\mathbb{R}^m}\left(\int_{\mathbb{R}^n} f(x,y)^p dx\right)^{\frac{1}{p}} dy.$$
2. *Si las secciones $y \mapsto f(x,y)$ son integrables para μ-casi todo x, entonces*
$$\left(\int_{\mathbb{R}^n}\left|\int_{\mathbb{R}^m} f(x,y)dy\right|^p dx\right)^{\frac{1}{p}} \leq \int_{\mathbb{R}^m}\left(\int_{\mathbb{R}^n} |f(x,y)|^p dx\right)^{\frac{1}{p}} dy.$$

El siguiente resultado nos será también muy útil.

Teorema 9.13. *Si $f \in L^1$ y $g \in L^p$, con $1 \leq p \leq \infty$ entonces $f * g$ está bien definida en casi todo punto, y $f * g \in L^p$. Además, $\|f * g\|_p \leq \|f\|_1 \|g\|_p$*

Demostración. Trataremos primero el caso $p = 1$. Consideremos la función $h(x,y) = f(x-y)g(y)$, que es medible.[2] Puesto que
$$\int(\int |h(x,y)|dx)dy = \int(\int |f(x-y)g(y)|dx)dy =$$
$$\int(|g(y)| \int |f(x-y)|dx)dy = \int |g(y)|(\int |f(z)|dz)dy =$$
$$\int |g(y)|\|f\|_1 dy = \|f\|_1 \|g\|_1 < \infty,$$
(donde en la tercera igualdad hemos usado la invariancia de la medida de Lebesgue por traslaciones), resulta de los teoremas de Tonelli y Fubini que $h \in L^1$ y las secciones $x \mapsto \int h(x,y)dx$, $y \mapsto \int h(x,y)dy$ están bien definidas en casi todo punto y son integrables. En particular $f * g(x) = \int h(x,y)dy$ está bien definida para casi todo x, y $f * g \in L^1$, con $\|f * g\|_1 = \int |\int h(x,y)dy|dx \leq \int\int |h(x,y)|dxdy = \|f\|_1 \|g\|_1$.

El caso $p = \infty$ es muy fácil y se deja como ejercicio para el lector. Finalmente, supongamos $1 < p < \infty$, y definamos q como el exponente conjugado de p (es decir, $1 = \frac{1}{p} + \frac{1}{q}$). Por lo anterior sabemos que, para casi todo x, la función $y \mapsto f(x-y)|g(y)|^p$ es integrable, o dicho de otro modo, la función $y \mapsto |f(x-y)|^{\frac{1}{p}} g(y)$ está en L^p. Como la función $y \mapsto |f(x-y)|^{\frac{1}{q}}$ está en L^q, resulta de la desigualdad de Hölder que $y \mapsto f(x-y)g(y)$ es integrable, para casi todo x. En particular $f * g(x)$ está bien definida para casi todo x. Ahora, aplicando la desigualdad de Minkowski para integrales se tiene que
$$\left(\int\left|\int f(y)g(x-y)dy\right|^p dx\right)^{\frac{1}{p}} \leq$$
$$\int\left(\int |f(y)g(x-y)|^p dx\right)^{\frac{1}{p}} dy = \int\left(|f(y)|^p \int |g(x-y)|^p dx\right)^{\frac{1}{p}} dy =$$
$$\int |f(y)|\left(\int |g(x-y)|^p dx\right)^{\frac{1}{p}} dy = \int |f(y)|\, \|g\|_p dy = \|f\|_1 \|g\|_p.$$
Esto demuestra que $f * g \in L^p$ y la relación $\|f * g\|_p \leq \|f\|_1 \|g\|_p$. □

[2]Puede escribirse $h(x,y) = F(x,y)G(x,y)$, donde $F = f \circ P_1 \circ T$, y $G = g \circ P_2$, con $P_1(x,y) = x$, $P_2(x,y) = y$, y $T(x,y) = (x-y, x+y)$, que es un isomorfismo lineal. Basta ver que F y G son medibles, y para ello puede suponerse que f y g toman valores reales. Para cada $t \in \mathbb{R}$ es $E := f^{-1}[-\infty, t)$ medible, $A := P_1^{-1}(E) = E \times \mathbb{R}^d$ medible (por ser producto de medibles), luego también $F^{-1}[-\infty, t) = T^{-1}(A)$ medible (porque T es una isometría lineal). Análogamente G es medible.

La siguiente proposición, cuya demostración se propone como ejercicio, resume las propiedades más básicas de la convolución de funciones.

Proposición 9.14. *Para todas $f, g, h \in L^1$ se tiene:*

1. $f * (g + h) = f * g + f * h$
2. $f * g = g * f$
3. $f * (g * h) = (f * g) * h.$

Una de las propiedades más importantes de la convolución de dos funciones integrables es que, en general, conserva las mejores propiedades de cada una de esas dos funciones: por ejemplo, si una de ellas es continua, o diferenciable, entonces la convolución también lo es. Lo mismo sucede con otras propiedades interesantes como la convexidad y la Lipschitzianidad.

Proposición 9.15. *Si $f \in L^1$ y g tiene derivadas continuas y acotadas hasta el orden k, entonces $f * g \in C^k$, y*

$$D^j(f * g)(x) = (f * D^j g)(x)$$

para todo x y para todo $j \in 0, 1, ..., k$.

Aquí, en el caso en que f y g están definidas en $\mathbb{R}^n$, y si $\{e_1, ..., e_n\}$ es la base canónica de $\mathbb{R}^n$, denotamos por $f * D^j g(x)$ la aplicación j-lineal cuyo valor en cada vector de la forma $(e_{i_1}, ..., e_{i_j})$ es $f * (\partial g/\partial x_{i_1}...\partial x_{i_j})(x)$. Dicho de otro modo y empleando la notación de multi-índices, se tiene que

$$\partial^\alpha(f * g)(x) = (f * \partial^\alpha g)(x)$$

para cada multi-índice α con $|\alpha| = j$.

Demostración: Para $k = 0$, fijando x, se tiene que, para toda sucesión (x_j) convergente a x, la sucesión de funciones $f(y)g(x_j - y)$ converge a $f(y)g(x - y)$ para todo y. Además esta sucesión de funciones está acotada por la función $y \mapsto \|g\|_\infty f(y)$, que es integrable. Por el teorema de la convergencia dominada se tiene entonces que $f * g(x_j) = \int f(y)g(x_j - y)dy$ converge a $f * g(x) = \int f(y)g(x - y)$.

Para $k = 1$, fijemos x y denotemos por y_i la enésima coordenada de la variable y. Dada una sucesión de números reales (t_j) que tienda a 0, consideremos la sucesión de funciones $h_j(y)$ definida por

$$f(y_1, ..., y_n)\frac{g(x_1 - y_1, ..., x_i + t_j - y_i, ..., x_n - y_n) - g(x_1 - y_1, ..., x_n - y_n)}{t_j}.$$

Aplicando el teorema del valor medio se ve que

$$|h_j(y)| \leq |f(y)| \, \|\frac{\partial g}{\partial x_i}\|_\infty \leq \|Dg\|_\infty |f(y)|,$$

y como la función $\|Dg\|_\infty f$ es integrable, se sigue del teorema de la convergencia dominada que

$$\lim_{j\to\infty} \frac{f * g(x_1..., x_i + t_j, ..., x_n) - f * g(x_1, ..., x_n)}{t_j} = \lim_{j\to\infty} \int h_j(y)dy$$
$$= \int \lim_{j\to\infty} h_j(y) = \int f(y)\frac{\partial g}{\partial x_i}(x - y)dy = f * \frac{\partial g}{\partial x_i}(x).$$

Esto prueba que

$$\frac{\partial(f * g)}{\partial x_i}(x) = f * \frac{\partial g}{\partial x_i}(x),$$

y como las funciones $\partial g/\partial x_i$ son continuas se sigue del caso $k = 0$ que también lo son las derivadas parciales $\partial f * g/\partial x_i$ para $i \in \{1, ..., n\}$. Por tanto $f * g \in C^1$ y $D(f * g)(x) = (f * Dg)(x)$. [3]

En el caso $k \geq 2$ la demostración se hace usando la misma técnica y un sencillo argumento de inducción. □

[3] Alternativamente, la prueba del caso $k = 1$ se obtiene aplicando un teorema de derivación bajo el signo integral; lo que acabamos de hacer esencialmente consiste en demostrar uno de estos teoremas.

9.3. Identidades aproximadas

Teniendo en cuenta las propiedades anteriores resulta que $(L^1, +, *)$ es un *álgebra de Banach*. Esta álgebra no posee elemento identidad para la operación $*$ (es decir, no existe ninguna función δ tal que $f * \delta = f$ para toda f; inténtese dar una demostración de este hecho). Sin embargo sí existen *identidades aproximadas*, que resultan muy útiles en diversas circunstancias.

Definición 9.1. Se dice que una sucesión de funciones $\{\delta_n\}$ de $L^1(\mathbb{R}^m)$ es una identidad aproximada (o aproximación de la identidad) en $L^1(\mathbb{R}^m)$ si

1. $\sup_{n\in\mathbb{N}} \|\delta_n\|_1 < \infty$
2. $\lim_{n\to\infty} \int_{\mathbb{R}^m} \delta_n = 1$
3. $\lim_{n\to\infty} \int_{\mathbb{R}^m\setminus B(0,r)} |\delta_n| = 0$ para todo $r > 0$,

y si además se tiene $\int_{\mathbb{R}^m} \delta_n = 1$ para todo $n \in \mathbb{N}$ entonces se dice que $\{\delta_n\}$ es un *núcleo de sumabilidad*.

También puede cambiarse en esta definición la familia numerable de funciones $\{\delta_n\}$ por una familia no numerable $\{\delta_t\}$ con $t \in (a, b) \subseteq \mathbb{R}$, y tomar límites cuando $t \to a^+$ (o cuando $t \to b^-$). Por ejemplo, en el caso $(a, b) = (0, \infty)$, $t \to 0^+$, una familia $\{\delta_t\}_{t>0}$ es un núcleo de sumabilidad en $L^1(\mathbb{R}^m)$ si

1. $\sup_{t>0} \|\delta_t\|_1 < \infty$
2. $\int_{\mathbb{R}^m} \delta_t = 1$ para todo $t > 0$, y
3. $\lim_{t\to 0^+} \int_{\mathbb{R}^m\setminus B(0,r)} |\delta_t| = 0$ para todo $r > 0$.

Es fácil comprobar que una familia $\{\delta_t\}_{t>0}$ es un núcleo de sumabilidad si y sólo si para cada sucesión (t_n) de números positivos que converja a 0 se tiene que la sucesión de funciones $\{\delta_{t_n}\}_{n\in\mathbb{N}}$ es un núcleo de sumabilidad.

Como ejemplo básico de núcleo de sumabilidad en $L^1(\mathbb{R}^m)$ tenemos el *núcleo del calor*, definido por

$$\delta_t(x) = \frac{1}{(4\pi t)^{n/2}} e^{-|x|^2/4t},$$

y llamado así porque $u(t, x) = f * \delta_t(x)$ es la solución de la ecuación del calor

$$\begin{cases} u_t - \Delta_x u = 0 & \text{en } (0, +\infty) \times \mathbb{R}^n, \\ u(0, x) = f(x) & \text{para } x \in \mathbb{R}^n \end{cases}$$

(ver la siguiente sección). En este ejemplo $t \in (0, \infty)$, y los límites se toman cuando $t \to 0^+$. Nótese que en el caso de *núcleos positivos* $\{\delta_t\}_{t>0}$ como es el del calor, la definición de núcleo de sumabilidad equivale a decir que:

1. $\int_{\mathbb{R}^m} \delta_t = 1$ para todo $t > 0$, y
2. $\lim_{t\to 0^+} \int_{\mathbb{R}^m\setminus B(0,r)} \delta_t = 0$ para todo $r > 0$.

Estas propiedades del núcleo del calor se comprobarán en la siguiente sección.

El término *identidad aproximada* tiene su justificación en el hecho de que la aplicación $f \mapsto \delta_n * f$ converge a la aplicación identidad cuando $n \to \infty$, como veremos a continuación, en diversos sentidos y contextos. Enunciaremos todos estos resultados sólo en el caso de identidades aproximadas sucesivas $\{\delta_n\}_{n\in\mathbb{N}}$, dejando a cuenta del lector el enunciado de los mismos en el caso de identidades aproximadas del tipo $\{\delta_t\}_{t\in(a,b)}$.

Proposición 9.16. *Si $\{\delta_n\}$ es una identidad aproximada en $L^1(\mathbb{R}^m)$ y f es uniformemente continua y acotada, entonces*

$$\lim_{n\to\infty} \|f * \delta_n - f\|_\infty = 0.$$

Si f es sólamente continua y acotada, se tiene aún que $f * \delta_n \to f$ uniformemente sobre cada compacto. Más aún, si $f \in C^k(\mathbb{R}^m)$ con $D^k f$ continua y acotada, entonces

$$\lim_{n\to\infty} D^k(f * \delta_n)(x) = D^k f(x)$$

uniformemente sobre cada compacto (y si $D^k f$ es uniformemente continua y acotada entonces la convergencia es uniforme sobre todo el espacio).
Finalmente, en el caso especial en que todas las funciones δ_n tengan soporte contenido en un mismo conjunto acotado, puede prescindirse de la suposición de que f y sus derivadas estén acotadas.

Demostración: Dado $\varepsilon > 0$, al ser f uniformemente continua existe $r > 0$ tal que $|f(x-y) - f(x)| \leq \varepsilon/3 \sup_n \|\delta_n\|_1$ si $|y| \leq r$. Por otro lado, como $\{\delta_n\}$ es una identidad aproximada podemos escribir $\int \delta_n = 1 + \nu_n$, con $\lim_{n\to\infty} \nu_n = 0$, y también se tiene que $\lim_{n\to\infty} \int_{|y|>r} \delta_n(y)dy = 0$. Por tanto existe $N \in \mathbb{N}$ tal que $|\nu_n| \leq \varepsilon/3(1 + \|f\|_\infty)$ y $\int_{|y|>r} |\delta_n(y)|dy \leq \varepsilon/6(1 + \|f\|_\infty)$ para todo $n \geq N$. Entonces, para todo $n \geq N$ y todo x se tiene

$$\begin{aligned}
&|f * \delta_n(x) - f(x)| = |\int (f(x-y) - f(x))\,\delta_n(y)dy + f(x)\nu_n| \leq \\
&\int |f(x-y) - f(x)|\,|\delta_n(y)|dy + |f(x)\nu_n| \leq \\
&\int_{|y|\leq r} |f(x-y) - f(x)|\,|\delta_n(y)|dy + \int_{|y|>r} |f(x-y) - f(x)|\,|\delta_n(y)|dy + \|f\|_\infty|\nu_n| \\
&\leq \varepsilon \sup_{n\in\mathbb{N}} \|\delta_n\|_1 + 2\|f\|_\infty \int_{|y|>r} |\delta_n(y)|dy + \|f\|_\infty|\nu_n| \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.
\end{aligned}$$

La demostración en el caso $f \in L^\infty \cap C^0$ es similar, usando el hecho de que las funciones continuas son uniformemente continuas en cada compacto.

Por otro lado, cuando todas las δ_n se anulan fuera de una bola $B(0,R)$, sustituyendo $\|f\|_\infty$ por $\sup_{y\in B(z,2R)} |f(y)|$ en las desigualdades anteriores, se ve que, aún cuando f no sea acotada, la continuidad uniforme de f en el compacto $B(z,2R)$ es suficiente para permitirnos deducir que $f * \delta_n$ tiende a f uniformemente en $B(z,R)$.

Por último, cambiando f por $D^k f$ y usando el hecho de que $D^k(f * \delta_n) = (D^k f) * \delta_n$ se obtiene el resultado en el caso en que $f \in C^k$. □

Lema 9.17. *Sea $1 \leq p < \infty$ y sea $f \in L^p$. Para cada x denotamos por f_x la función de L^p definida por*

$$y \longmapsto f(y - x).$$

La función $x \mapsto f_x \in L^p$ es uniformemente continua.

Demostración. Dado $\varepsilon > 0$, existe g continua y con soporte compacto que supondremos contenido en una bola $B(0,R)$, tal que $\|f - g\|_p < \frac{\varepsilon}{4}$. Como g es uniformemente continua, existe un $\delta \in (0,1)$ tal que si $|x - x'| < \delta$ entonces

$$\left|g(x) - g(x')\right| < \frac{\varepsilon}{2\mathcal{L}_n(B(0,R+1))^{\frac{1}{p}}}.$$

Así, si $|x - x'| < \delta$ se tiene

$$\begin{aligned}
\|f_x - f_{x'}\|_p \leq & \|f_x - g_x\|_p + \|g_x - g_{x'}\|_p + \|g_{x'} - f_{x'}\|_p = \\
=& 2\,\|f - g\|_p + \|g_x - g_{x'}\|_p \leq \\
\leq & \frac{\varepsilon}{2} + \left[\int |g(y-x) - g(y-x')|^p\, dy\right]^{1/p} = \frac{\varepsilon}{2} + \left[\int |g(z) - g(z + x' - x)|^p\, dz\right]^{1/p} \\
=& \frac{\varepsilon}{2} + \left[\int_{B(0,R+1)} |g(z) - g(z + x' - x)|^p\, dz\right]^{1/p} \\
\leq & \frac{\varepsilon}{2} + \left[\int_{B(0,R+1)} \left|\frac{\varepsilon}{2\mathcal{L}_n(B(0,R+1))^{\frac{1}{p}}}\right|^p dz\right]^{1/p} \\
=& \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

□

Proposición 9.18. *Si $\{\delta_n\}$ es una identidad aproximada en L^1 y $f \in L^p$, con $1 \leq p < \infty$, entonces $f * \delta_n \in L^p$, y*

$$\lim_{n\to\infty} \|f * \delta_n - f\|_p = 0.$$

Demostración. Sabemos por el Teorema 9.13 que $f * \delta_n \in L^p$, y de hecho $\|f * \delta_n\|_p \leq \|f\|_p \|\delta_n\|_1$. En particular, como $\{\delta_n\}_{n\in\mathbb{N}}$ es una identidad aproximada, si denotamos $M = \sup_n \|\delta_n\|_1$, se tiene $\|f * \delta_n\|_p \leq M\|f\|_p$ para todo $n \in \mathbb{N}$.

Para probar el resultado de convergencia, dado $\varepsilon > 0$, usando el lema anterior podemos encontrar $\delta > 0$ tal que

$$\|f_y - f\|_p \leq \varepsilon/3M \text{ si } |y| \leq \delta$$

y, para este δ, usando la tercera propiedad de la identidad aproximada, existe $n_0 \in \mathbb{N}$ tal que

$$\int_{|y|>\delta} |\delta_n(y)| dy \leq \frac{\varepsilon}{3(1 + 2\|f\|_p)}, \text{ y}$$

$$|\nu_n| \leq \frac{\varepsilon}{3\,(\|f\|_p + 1)}$$

para todo $n \geq n_0$, donde

$$\nu_n = \int \delta_n - 1$$

como en la demostración de la proposición anterior. Entonces, aplicando las desigualdades de Minkowski (tanto para la norma $\|\cdot\|_p$ como para integrales), y usando que $\|f_y\|_p = \|f\|_p$ para todo y, tenemos

$$\begin{aligned}
&\|f * \delta_n - f\|_p = \left(\left|\int (f(x-y) - f(x))\delta_n(y)dy + \nu_n f(x)\right|^p dx\right)^{\frac{1}{p}} \leq \\
&\left(\left|\int (f(x-y) - f(x))\delta_n(y)dy +\right|^p dx\right)^{\frac{1}{p}} + \nu_n\|f\|_p \leq \\
&\int \left(\int |(f(x-y) - f(x))\delta_n(y)|^p dx\right)^{\frac{1}{p}} dy + \nu_n\|f\|_p = \\
&\int |\delta_n(y)| \|f_y - f\|_p dy + \nu_n\|f\|_p \leq \\
&\int_{|y|\leq\delta} |\delta_n(y)| \|f_y - f\|_p dy + \int_{|y|>\delta} |\delta_n(y)| \|f_y - f\|_p dy + \nu_n\|f\|_p \leq \\
&\int_{|y|\leq\delta} |\delta_n(y)| \frac{\varepsilon}{3M} dy + 2\|f\|_p \int_{|y|>\delta} |\delta_n(y)| dy + \nu_n\|f\|_p \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon
\end{aligned}$$

siempre que $n \geq n_0$. □

9.4. Regularización de funciones

De las propiedades anteriores se sigue que la convolución de una función continua f con una identidad aproximada $\{\delta_n\}$ produce una sucesión $f_n = f * \delta_n$ de funciones que aproximan a f y poseen las mismas buenas propiedades de diferenciabilidad que tengan las δ_n. A continuación mostramos una elección típica de un núcleo de sumabilidad que permite obtener resultados muy interesantes sobre aproximación de funciones por funciones más regulares. Véase también en el capítulo siguiente la sección sobre el efecto regularizante de la ecuación del calor y la demostración del teorema de Weierstrass.

Definamos $\delta \in C^\infty(\mathbb{R}^n)$ por

$$\delta(x) = \begin{cases} C \exp\left(\frac{1}{|x|^2-1}\right) & \text{si } |x| < 1, \\ 0 & \text{si } |x| \geq 1, \end{cases}$$

donde C está elegida de modo que $\int_{\mathbb{R}^n} \delta = 1$. Definamos también, para cada $\varepsilon > 0$, la función

$$\delta_\varepsilon(x) = \frac{1}{\varepsilon^n}\delta\left(\frac{x}{\varepsilon}\right).$$

Es fácil ver que cada función δ_ε es de clase $C^\infty(\mathbb{R}^n)$, tiene soporte en $B(0,\varepsilon)$, e integral 1. Por tanto, es evidente que para cualquier sucesión decreciente de números positivos $\{\varepsilon_n\}$ que converja a 0 se tiene que $\{\delta_{\varepsilon_n}\}$ es un núcleo de sumabilidad.

Usando estas propiedades, es inmediato comprobar que para toda función continua $f : \mathbb{R}^n \to \mathbb{R}$, la familia de funciones $f_\varepsilon := f * \delta_\varepsilon$ es de clase C^∞ y converge uniformemente a f sobre cada subconjunto compacto de $\mathbb{R}^n$ cuando $\varepsilon \to 0$. Más en general, si f es de clase C^p (con $p \geq 1$), entonces las derivadas de f_ε convergen a las derivadas de f (hasta el orden p), uniformemente sobre acotados, cuando $\varepsilon \to 0$. Si además $D^k f$ es uniformemente continua en $\mathbb{R}^n$, se tiene que $D^k f_\varepsilon \to D^k f$ uniformemente en $\mathbb{R}^n$.

Por otro lado, este procedimiento de regularización conserva también otras buenas propiedades que pueda tener f. Por ejemplo, si f es convexa entonces f_ε también lo es. Igualmente, si f es M-Lipschitz, es fácil comprobar que f_ε también es M-Lipschitz. Ver los problemas 9.9 y 9.10.

9.5. Bonus: El efecto regularizante de la ecuación del calor y la demostración del teorema de Weierstrass.

En esta sección daremos una demostración del famoso teorema de Weierstrass sobre aproximación polinómica: dado cualquier compacto K de $\mathbb{R}^n$, los polinomios forman un conjunto denso de funciones en el espacio $(C(K), \|\cdot\|_\infty)$ de las funciones continuas en K. Dicha demostración es esencialmente la misma que dio Weierstrass, pero adaptada al lenguaje matemático moderno y usando ventajosamente los resultados sobre convolución de funciones e identidades aproximadas estudiados en el capítulo anterior.

Como motivación de esta técnica recordemos que para una función $f \in L^p(\mathbb{R}^n)$, $1 \leq p \leq \infty$, la única solución de la ecuación del calor

$$\begin{cases} u_t - \Delta_x u = 0 & \text{en } (0,+\infty) \times \mathbb{R}^n, \\ u(0,x) = f(x) & \text{para } x \in \mathbb{R}^n \end{cases}$$

viene dada por la fórmula

$$u(t,x) = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}} f(y)dy,$$

es decir

$$u(t,x) = f^t(x) := (f * \delta_t)(x),$$

donde

$$\delta_t(x) = \frac{1}{(4\pi t)^{n/2}} e^{-|x|^2/4t}.$$

Proposición 9.19. *La familia* $\{\delta_t\}_{t>0}$ *es un núcleo de sumabilidad en* $L^1(\mathbb{R}^n)$*, cuando* $t \to 0^+$.

Demostración. Como $\delta_t > 0$ para todo $t > 0$ sólo hay que comprobar que

1. $\int_{\mathbb{R}^n} \delta_t(x)dx = 1$ para todo $t > 0$, y que
2. $\lim_{t\to 0^+} \int_{\mathbb{R}^n\setminus B(0,r)} \delta_t = 0$ para todo $r > 0$.

La primera propiedad se verifica fácilmente haciendo el cambio de variable $y = x/\sqrt{4\pi t}$ y recordando que $\int_{\mathbb{R}^n} e^{-\pi|y|^2}dy = 1$. La segunda propiedad se puede comprobar, por ejemplo, haciendo el mismo cambio de variable:

$$\lim_{t\to 0^+} \int_{|x|>r} \delta_t(x)dx = \lim_{t\to 0^+} \int_{|y|>r/\sqrt{4\pi t}} e^{-\pi|y|^2}dy = 0,$$

y usando el teorema de la convergencia dominada para justificar la última igualdad. Los detalles de dichas comprobaciones se proponen como ejercicio. □

Teorema 9.20. *Para toda* $f \in L^\infty(\mathbb{R}^n)\cap C(\mathbb{R}^n)$*, la función* f^t *es real analítica para cualquier* $t > 0$*, su serie de Taylor en* 0 *tiene radio de convergencia infinito, y se tiene* $\lim_{t\to 0^+} f^t(x) = f(x)$ *uniformemente sobre acotados. Si además* f *es uniformemente continua entonces* $\lim_{t\to 0^+} f^t(x) \to f$ *uniformemente en* $x \in \mathbb{R}^n$.

Así, la ecuación del calor regulariza inmediatamente cualquier dato inicial continuo y acotado.
Demostración: Definamos, para $z \in \mathbb{C}^n$, una extensión compleja de la función f^t por

$$\widetilde{f}^t(z) = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} f(y)e^{-\frac{1}{4t}\sum_{j=1}^n (z_j^2 - 2y_jz_j + y_j^2)}dy,$$

donde $z \in \mathbb{C}^n$. Escribiendo $z_k = u_k + iv_k$ para $k = 1, ..., n$, es fácil comprobar que la integral es convergente y por tanto $\widetilde{f}^t(z)$ está bien definida para todo $z \in \mathbb{C}^n$. Vamos a probar que $\widetilde{f}^t$ es analítica en $\mathbb{C}^n$. En particular se deducirá que su restricción a $\mathbb{R}^n$, es decir f^t, es real analítica.

Tenemos

$$\widetilde{f}^t(z) = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} f(y)e^{-\frac{1}{4t}\sum_{j=1}^n (z_j^2 - 2y_jz_j + y_j^2)}dy =$$
$$\frac{e^{-\frac{1}{4t}\sum_{j=1}^n z_j^2}}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} f(y)e^{-\frac{1}{4t}\sum_{j=1}^n y_j^2} e^{\frac{1}{2t}\sum_{j=1}^n y_jz_j}dy,$$

La expresión que aparece fuera de la segunda integral es obviamente analitica respecto de $z \in \mathbb{C}^n$, así que basta probar que la función

$$G(z) := \int_{\mathbb{R}^n} f(y)e^{-\frac{1}{4t}\sum_{j=1}^n y_j^2} e^{\frac{1}{2t}\sum_{j=1}^n y_jz_j}dy$$

es analítica en $\mathbb{C}^n$. Tenemos

$$G(z) = \int_{\mathbb{R}^n} \sum_{k=0}^{\infty} g_k(y, z)dy,$$

donde

$$g_k(y, z) = f(y)e^{-\frac{1}{4t}\|y\|^2} \frac{\left(\frac{1}{2t}\sum_{j=1}^n y_jz_j\right)^k}{k!}.$$

Obsérvese que cada función $z \mapsto g_k(y, z)$ es un polinomio k-homogéneo, y por consiguiente también lo es $z \mapsto \int_{\mathbb{R}^n} g_k(y, z)dy$. Vamos a comprobar que podemos sacar la suma fuera de la integral, con lo que

$$G(z) = \sum_{k=0}^{\infty} \int_{\mathbb{R}^n} g_k(y, z)dy$$

es analítica, puesto que se expresa como suma de los polinomios k-homogéneos $\int_{\mathbb{R}^n} g_k(y,z)dy$.

Fijemos $R > 0$ y consideremos $z \in B_{\mathbb{C}^n}(0,R)$. Tenemos

$$\begin{aligned}\sum_{k=0}^{\infty} |g_k(y,z)|dy &\leq \sum_{k=0}^{\infty} \|f\|_\infty e^{-\frac{1}{4t}\|y\|^2} \frac{\left(\frac{1}{2t}\|y\|\,\|z\|\right)^k}{k!} \\ &= \|f\|_\infty e^{-\frac{1}{4t}\|y\|^2} \sum_{k=0}^{\infty} \frac{\left(\frac{1}{2t}\|y\|\,\|z\|\right)^k}{k!} \\ &= \|f\|_\infty e^{-\frac{1}{4t}(\|y\|^2 - 2\|y\|\,\|z\|)} \\ &\leq \|f\|_\infty e^{-\frac{1}{4t}(\|y\|^2 - 2R\|y\|)} := H_R(y)\end{aligned}$$

y puesto que la función $H_R(y)$ es integrable en $\mathbb{R}^n$ y no depende de $z \in B_{\mathbb{C}^n}(0,R)$, puede aplicarse el teorema de la convergencia dominada para concluir que

$$\sum_{k=0}^{\infty} \int_{\mathbb{R}^n} g_k(y,z)dy = \int_{\mathbb{R}^n} \sum_{k=0}^{\infty} g_k(y,z)dy = G(z)$$

uniformemente en $z \in B_{\mathbb{C}^n}(0,R)$.

Como $R > 0$ es arbitrario, este argumento también muestra que la serie $G(z) = \sum_{k=0}^{\infty} \int_{\mathbb{R}^n} g_k(y,z)dy$ tiene radio de convergencia infinito.

Finalmente, por ser $\{\delta_t\}_{t>0}$ un núcleo de sumabilidad, los resultados de este capítulo aseguran que si f es continua, entonces $f^t \to f$ uniformemente en acotados cuando $t \to 0^+$. Y si f es uniformemente continua entonces $f^t \to f$ uniformemente en $\mathbb{R}^n$. □

Corolario 9.21 (Weierstrass). *Si $K \subset \mathbb{R}^n$ es compacto y $f : K \to \mathbb{R}$ es continua, para todo $\varepsilon > 0$ existe un polinomio $P : \mathbb{R}^n \to \mathbb{R}$ tal que $|f(x) - P(x)| \leq \varepsilon$ para todo $x \in K$.*

Demostración: Como f es continua en el compacto K, tiene una extensión uniformemente continua y acotada definida en todo $\mathbb{R}^n$, que seguiremos denotando f por comodidad. Por el teorema anterior, dado $\varepsilon > 0$, existe $r > 0$ tal que si $t \in (0,r]$ se tiene

$$|f^t(x) - f(x)| \leq \varepsilon/2$$

para todo $x \in \mathbb{R}^n$, y f^t es real analítica con radio de convergencia infinito. Por tanto, fijado $R > 0$ tal que $K \subset B(0,R)$, existe $N_0 \in \mathbb{N}$ tal que si $N \geq N_0$ se tiene

$$|f^r(x) - \sum_{k=0}^{N} \frac{D^k f^r(0)(x)^k}{k!}| \leq \frac{\varepsilon}{2}$$

para todo $x \in B(0,R)$. Por tanto, si ponemos

$$P(x) = \sum_{k=0}^{N_0} \frac{D^k f^r(0)(x)^k}{k!},$$

se deduce que $P : \mathbb{R}^n \to \mathbb{R}$ es un polinomio tal que $|f(x) - P(x)| \leq \varepsilon$ para todo $x \in K$. □

9.6. Problemas

Problema 9.1. Construir una función $f \in L^1(\mathbb{R})$ tal que toda g medible que coincida en casi todo punto con f sea discontinua en todo punto. *Indicación:* definir $f(x) = \sum_{n=1}^{\infty} 2^{-n}\varphi(x - r_n)$, donde (r_n) es una enumeración de los racionales y $\varphi(x) = |x|^{-1/2}\chi_{(0,1)}(x)$. Probar que esta serie converge en casi todo punto, pero la función f no está acotada en ningún intervalo no vacío.

Problema 9.2. Transcribir las desigualdades de Hölder y de Minkowski en los casos $X = \{1, ..., n\}$ y $X = \mathbb{N}$, con μ la medida de contar.

Problema 9.3. Probar que $L^\infty(X)$ es un espacio de Banach.

Problema 9.4. Sea $1 \le p < \infty$. Si $f_n, f \in L^p(X)$ y $f_n \to f$ en ctp, probar que $\|f_n - f\|_p \to 0$ si y sólo si $\|f_n\|_p \to \|f\|_p$. *Indicación:* Considerar $\bigcup_{m=1}^{\infty} B_m$, donde $B_m := \{x : |f_n(x) - f(x)| \le |f(x)|\}$ e intentar ponerse en las condiciones del teorema de la convergencia dominada.

Problema 9.5. Comprobar que si (f_n) y (g_n) son sucesiones de funciones que convergen a f y g, respectivamente, en L^1, entonces $(f_n * g_n)$ converge a $f * g$ en L^1.

Problema 9.6. Demostrar que si $f \in L^p$ y $g \in L^q$ con $\frac{1}{p} + \frac{1}{q} = 1$ entonces $f * g(x)$ está bien definida para todo x, y $f * g$ es continua y acotada (de hecho $\|f * g\|_\infty \le \|f\|_p \|g\|_q$).

Problema 9.7. Si $f \in L^\infty$ es continua en un punto x_0 y $\{\delta_n\}_{n\in\mathbb{N}}$ es una aproximación de la identidad, demostrar que $\lim_{n\to\infty} f * \delta_n(x_0) = f(x_0)$.

Problema 9.8. Si $f \in L^1$ es continua en un punto x_0 y $\{\delta_n\}_{n\in\mathbb{N}}$ es una aproximación de la identidad que además satisface que, para cada $r > 0$ es $\lim_{n\to\infty} \delta_n(y) = 0$ uniformemente en $|y| \ge r$, demostrar que $\lim_{n\to\infty} f * \delta_n(x_0) = f(x_0)$.

Problema 9.9. Demostrar que si $f : \mathbb{R}^n \to \mathbb{R}$ es convexa entonces $f * g$ también lo es, siempre que esté bien definida en $\mathbb{R}^n$, para cualquier $g \ge 0$.

Problema 9.10. Demostrar que si $f : \mathbb{R}^n \to \mathbb{C}$ es Lipschitz entonces $f * g$ también lo es, siempre que esté bien definida en $\mathbb{R}^n$, para cualquier $g : \mathbb{R}^n \to \mathbb{C}$ integrable.

Problema 9.11. Sea Y un subconjunto no vacío de un espacio métrico X , y sea $f : Y \to \mathbb{R}$ un función L-Lipschitz. Definamos

$$\widetilde{f}(x) = \inf_{y\in Y}\{f(y) + Ld(x, y)\}$$

para cada $x \in X$. Demostrar que $\widetilde{f}$ es también L-Lipschitz y extiende a f, es decir $\widetilde{f}(y) = f(y)$ para todo $y \in Y$.

Problema 9.12. Sea X un espacio métrico, y $f : X \to \mathbb{R}$ uniformemente continua y acotada. Para cada $n \in \mathbb{N}$ definamos

$$f_n(x) = \inf_{y\in X}\{f(y) + nd(x, y)\}.$$

Demostrar que:

1. f es n-Lipschitz;
2. $f_n \nearrow f$ uniformemente en X cuando $n \to \infty$.

Indicación: si $nr \ge 2\|f\|_\infty$ entonces $\inf_{y\in X}\{f(y) + nd(x, y)\} = \inf_{y\in B(x,r)}\{f(y) + nd(x, y)\}$.

Problema 9.13. Sea K un subconjunto compacto de un espacio métrico X, y sea $f : K \to \mathbb{R}$ un función continua. Demostrar que para todo $\varepsilon > 0$ existe una función Lipschitz $F : X \to \mathbb{R}$ tal que $|F_{|_K} - f| \le \varepsilon$, y $\|F\|_\infty := \sup_{x\in X} |F(x)| \le \sup_{z\in K} |f(z)| := \|f\|_\infty$.

Problema 9.14. Demostrar que si K es un compacto de $\mathbb{R}^n$ y $f : K \to \mathbb{R}$ es continua entonces existe $F : \mathbb{R}^n \to \mathbb{R}$ uniformemente continua tal que $F = f$ en K.
Indicación: Usando el problema anterior, existe $F_1 : X \to \mathbb{R}$ tal que $|F_1 - f| \le 1/2$ en K y $\|F_1\|_\infty = \|f\|_\infty$. Denotando $g_1 = F_{1|_K}$ y aplicando el mismo argumento a $f - g_1$, existe $F_2 : X \to \mathbb{R}$ tal que $|F_2 - (f - g_1)| \le 1/2^2$ en K y $\|F_2\|_\infty = \|f - g_1\|_\infty$. Continuando el proceso de esta manera puede encontrarse una sucesión de funciones uniformemente continuas $F_n : X \to \mathbb{R}$, cuyas restricciones a K denotamos g_n, tales que $\|F_n\|_\infty = \|f - g_1 - g_2 - ... - g_{n-1}\|_\infty \le 1/2^{n-1}$ para todo n. Comprobar que $F = \sum_{n=1}^{\infty} F_n$ tiene las propiedades deseadas.

Problema 9.15 (Desigualdad de Hardy)**.** Si $1 < p < \infty$ y $f \in L^p((0,\infty))$, definamos

$$F(x) = \frac{1}{x}\int_0^x f(t)dt, \quad x > 0.$$

Probar que $\|F\|_p \leq \frac{p}{p-1}\|f\|_p$.
Indicación: usar un cambio de variable y la desigualdad de Minkowski para integrales en $(0,\infty) \times (0,1)$.

Capítulo 10

Medidas con signo. Teorema de Radon-Nikodym

10.1. Medidas con signo

Definición 10.1. Sea $\mathcal{A}$ una σ-álgebra en un conjunto X. Se dice que $\mu : \mathcal{A} \to [-\infty, \infty]$ es una medida con signo si:

1. $\mu(\emptyset) = 0$;
2. μ alcanza a lo sumo uno de los valores $\pm\infty$;
3. μ es numerablemente aditiva; es decir, si (A_j) es una familia de conjuntos $\mathcal{A}$-medibles disjuntos dos a dos, entonces $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$.

Ejercicio 10.1. Como ejemplos básicos de medidas con signo tenemos:

1. $\mu = \mu_1 - \mu_2$, donde μ_1 y μ_2 son medidas positivas y una de ellas es finita.
2. Si $(X, \mathcal{A}, \nu)$ es un espacio de medida y $f \in L^1(X)$ entonces $\mu(A) := \int_A f d\nu$ define una medida con signo (y finita).

Vamos a probar que cualquier medida con signo es diferencia de dos medidas positivas. A tal fin necesitaremos una definición y un lema previos.

Definición 10.2. Si μ es una medida con signo en un espacio medible $(X, \mathcal{A})$, un conjunto $E \in \mathcal{A}$ se dice que es *positivo* si $\mu(A) \geq 0$ para cada subconjunto $A \in \mathcal{A}$ tal que $A \subseteq E$. Análogamente pueden definirse conjuntos negativos.

Lema 10.1. *Si μ es una medida con signo y $0 < \mu(E) < \infty$ entonces existe $A \in \mathcal{A}$ positivo tal que $A \subset E$ y $\mu(A) > 0$.*

Demostración. Si E es positivo puede tomarse $A = E$. En caso contrario, existe $B \subset E$ tal que $\mu(B) < 0$. Sea n_1 el menor número natural para el que existe $B_1 \subset E$ con

$$\mu(B_1) \leq -\frac{1}{n_1}.$$

Si $A_1 := E \setminus B_1$ es positivo, podemos tomar $A = A_1$ (ya que $\mu(A_1) > 0$ porque en caso contrario $\mu(E) < 0$). Si A_1 no es positivo, definimos n_2 como el menor natural tal que existe $B_2 \subset E \setminus B_1$ con

$$\mu(B_2) \leq -\frac{1}{n_2}.$$

Si $A_2 := E \setminus (B_1 \cup B_2)$ es positivo, podemos tomar $A = A_2$ y acabar aquí la prueba (obsérvese que $\mu(A_2) > 0$ por el mismo motivo que antes). En caso contrario tomamos n_3 como el primer natural para el que existe $B_3 \subset E \setminus (B_1 \cup B_2)$ tal que

$$\mu(B_3) \leq -\frac{1}{n_3},$$

definimos $A_3 = E \setminus (B_1 \cup B_2 \cup B_3)$, y continuamos este proceso por inducción.

Si $A_m = E \setminus \bigcup_{j=1}^m B_j$ es positivo para algún $m \in \mathbb{N}$, entonces acabamos la prueba tomando $A = A_m$; nótese que $\mu(A_m) > 0$ (si no tendríamos $\mu(E) < 0$). En caso contrario, se obtiene una sucesión infinita de conjuntos B_j disjuntos dos a dos con las propiedades señaladas, y definimos

$$A = E \setminus \bigcup_{j=1}^{\infty} B_j.$$

Veamos que A es positivo y $\mu(A) > 0$. Tenemos que

$$0 < \mu(E) = \mu(A) + \sum_{j=1}^{\infty} \mu(B_j) \leq \mu(A) - \sum_{j=1}^{\infty} \frac{1}{n_j},$$

luego

$$0 < \sum_{j=1}^{\infty} \frac{1}{n_j} < \mu(A) < \infty$$

(la última desigualdad se tiene porque de lo contrario $\mu(E) = \mu(A) + \mu(E \setminus A) = \infty$). En particular $\lim_{j\to\infty} n_j = \infty$. Sólo queda por ver que A es positivo. Si $C \subset A$ es medible, entonces para todo $m \in \mathbb{N}$, $C \subset E \setminus \bigcup_{j=1}^m B_j$, y por definición de n_{m+1} obtenemos que

$$\mu(C) > -\frac{1}{n_{m+1} - 1} \to 0$$

cuando $m \to \infty$, de lo que se deduce que $\mu(C) \geq 0$. □

Ahora ya podemos probar el resultado de descomposición al que hemos aludido antes.

Teorema 10.2 (de descomposición de Hahn)**.** *Si μ es una medida con signo en un espacio medible $(X, \mathcal{A})$, entonces existen $X^+, X^- \in \mathcal{A}$ tales que:*

1. *$X = X^+ \cup X^-$;*
2. *$X^+ \cap X^- = \emptyset$;*
3. *μ es positiva en X^+ (es decir, $\mu(E) \geq 0$ para todo $E \in \mathcal{A}$ con $E \subset X^+$);*
4. *μ es negativa en X^- (es decir, $\mu(E) \leq 0$ para todo $E \in \mathcal{A}$ con $E \subset X^-$).*

Por tanto toda medida con signo es de la forma $\mu = \mu^+ - \mu^-$, donde $\mu^\pm$ son medidas positivas concentradas en dos conjuntos disjuntos cuya unión es X; a saber, $\mu^+(E) := \mu(E \cap X^+)$ y $\mu^-(E) := -\mu(E \cap X^-)$ para cada $E \in \mathcal{A}$. Puede probarse (y queda propuesto como ejercicio) que esta descomposición es única salvo conjuntos de medida cero: es decir, si $X = H^+ \cup H^-$ es otra descomposición como la del enunciado, entonces

$$\mu(X^+ \setminus H^+) = \mu(H^+ \setminus X^+) = \mu(X^- \setminus H^-) = \mu(H^- \setminus X^-) = 0.$$

Por ejemplo, si $\mu(E) := \int_E f d\nu$, donde $f \in L^1(X, \nu)$, entonces $\mu(E) = \int_E f^+ d\nu - \int_E f^- d\nu := \mu^+(E) - \mu^-(E)$, y puede tomarse $X^+ = f^{-1}([0, +\infty))$ (aunque también podría tomarse $X^+ = f^{-1}((0, +\infty)))$, y $X^- = X \setminus X^+$.

A la medida

$$|\mu| := \mu^+ + \mu^-$$

se le llama la *variación total* de μ, y tiene la propiedad de que

$$|\mu(A)| \leq |\mu|(A)$$

para todo $A \in \mathcal{A}$.

Demostración del Teorema de descomposición de Hahn. Supongamos que μ no alcanza el valor $+\infty$. Sea

$$M := \sup\{\mu(A) : A \in \mathcal{A},\ A \text{ es positivo}\},$$

y tomemos una sucesión creciente (A_j) de conjuntos positivos (A_j) tales que

$$M = \lim_{j\to\infty} \mu(A_j)$$

(esta sucesión se obtiene a partir de la definición de M y usando que la unión de dos conjuntos positivos es un conjunto positivo, para poder hacer la sucesión creciente). Definamos $A := \bigcup_{j=1}^{\infty} A_j$. Es fácil ver que A es positivo, y se cumple que $M = \mu(A) < \infty$; nótese que la restricción de μ a A define una medida positiva, y por tanto podemos usar que $\lim_{j\to\infty} \mu(A_j) = \mu(\bigcup_{j=1}^{\infty} A_j)$.

Veamos que $X \setminus A$ es negativo; así podemos tomar $X^+ = A$, $X^- = X \setminus A$ para acabar la demostración. Y en efecto, de lo contrario existiría $E \subset X \setminus A$ con $0 < \mu(E) < \infty$, y podríamos aplicar el lema anterior para obtener $C \subset E$ positivo con $\nu(C) > 0$; pero entonces $A \cup C$ es positivo y

$$\mu(A \cup C) = \mu(A) + \mu(C) > M,$$

contradiciendo la definición de M.

En el caso en que μ no alcanza el valor $-\infty$ la demostración es análoga (o puede aplicarse la demostración ya hecha a la medida con signo $-\mu$). □

10.2. El teorema de Radon-Nikodym

Definición 10.3. Sean μ, ν medida positivas en un espacio medible $(X, \mathcal{A})$. Decimos que ν es *absolutamente continua* con respecto a μ siempre y cuando

$$E \in \mathcal{A}, \mu(E) = 0 \implies \nu(E) = 0,$$

es decir, si los conjuntos de medida cero para μ también lo son para ν. En este caso escribimos

$$\nu \ll \mu.$$

Ejemplo 10.1. Si $f \geq 0$ es medible y

$$\nu(E) := \int_E f d\mu,$$

entonces $\nu \ll \mu$.

El teorema de Radon-Nikodym nos asegura que de hecho este es el único ejemplo que puede haber, siempre que μ y ν sean σ-finitas.

Teorema 10.3 (de Radon-Nikodym)**.** *Sean λ, μ medidas (positivas) σ-finitas en un espacio medible $(X, \mathcal{A})$, y supongamos que $\lambda \ll \mu$. Entonces existe $h : X \to [0, \infty)$ medible tal que*

$$\lambda(E) = \int_E h d\mu \tag{10.1}$$

para todo $E \in \mathcal{A}$.

Si se tiene $\lambda(X) < \infty$, entonces además $h \in L^1(X, \mu)$.

A la función h se le llama la derivada de Radon-Nikodym de λ, y suele denotarse

$$h = \frac{d\lambda}{d\mu}.$$

Es fácil ver que h es única como función de $L^1(X)$ (o en el caso $\lambda(X) = \infty$, salvo conjuntos de medida cero). Con esta notación, (10.1) se reescribe

$$\lambda(E) = \int_E \frac{d\lambda}{d\mu} d\mu \text{ para todo } E \in \mathcal{A}.$$

Demostración. Supongamos primero que $\mu(X) < \infty$, $\lambda(X) < \infty$. Definamos

$$\Phi = \{\varphi : X \to [0, \infty] : \varphi \text{ es medible y } \int_E \varphi d\mu \leq \lambda(E) \text{ para todo } E \in \mathcal{A}\}.$$

Se tiene $\Phi \neq \emptyset$ (porque $0 \in \Phi$). Si $\varphi_1, \varphi_2 \in \Phi$, entonces $\max\{\varphi_1, \varphi_2\} \in \Phi$; en efecto,

$$\int_E \max\{\varphi_1, \varphi_2\} d\mu = \int_{E\cap\{\varphi_1 \geq \varphi_2\}} \varphi_1 d\mu + \int_{E\cap\{\varphi_1 < \varphi_2\}} \varphi_2 d\mu$$
$$\leq \lambda(E \cap \{\varphi_1 \geq \varphi_2\}) + \lambda(E \cap \{\varphi_1 < \varphi_2\}) = \lambda(E).$$

Por tanto, por inducción, lo mismo sucede con el máximo de cualquier subconjunto finito de Φ. Definamos ahora

$$M = \sup_{\varphi\in\Phi} \int_X \varphi d\mu.$$

Obsérvese que $M \leq \lambda(X) < \infty$. Tomemos una sucesión $(\varphi_n) \subset \Phi$ tal que

$$\lim_{n\to\infty} \int_X \varphi_n d\mu = M,$$

y definamos $h_n = \max\{\varphi_1, ..., \varphi_n\} \in \Phi$. Obviamente (h_n) es creciente y por tanto converge a una función $h : X \to [0, \infty]$. Puesto que $\int_E h_n d\mu \leq \lambda(E)$ para todo $E \in \mathcal{A}$, aplicando el teorema de la convergencia monótona se deduce que

$$\int_E h d\mu \leq \lambda(E) \text{ para todo } E \in \mathcal{A},$$

y que

$$\int_X h d\mu = \lim_{n\to\infty} \int_X h_n d\mu = M,$$

ya que $\int_X \varphi_n \leq \int_X h_n \leq M$. En particular $h \in L^1(X, \mu)$.

Vamos a probar que

$$\int_E h d\mu = \lambda(E) \text{ para todo } E \in \mathcal{A}. \tag{10.2}$$

Es claro que

$$\eta(E) := \lambda(E) - \int_E h d\mu$$

define una medida positiva en $\mathcal{A}$, y nuestro objetivo es probar que $\eta = 0$. Y en efecto, de lo contrario existiría $A \in \mathcal{A}$ con $\eta(A) > 0$, lo que implica que $\lambda(A) > 0$, y por tanto también $\mu(A) > 0$ (ya que $\lambda \ll \mu$). Luego existe $\varepsilon > 0$ tal que $\eta(A) - \varepsilon\mu(A) > 0$. Definamos

$$\xi := \eta - \varepsilon\mu.$$

Entonces ξ es una medida con signo y $\xi(A) > 0$, luego por el Lema 10.1, salvo tomar un subconjunto de A, podemos suponer que A es positivo para ξ, es decir $\xi(A \cap E) \geq 0$ para todo $E \in \mathcal{A}$, y por tanto

$$0 \leq \xi(A \cap E) = \eta(A \cap E) - \varepsilon\mu(A \cap E) = \lambda(A \cap E) - \int_{A \cap E} h d\mu - \varepsilon\mu(A \cap E),$$

luego

$$\int_{A \cap E} h d\mu + \varepsilon\mu(A \cap E) \leq \lambda(A \cap E),$$

pero también se tiene

$$\int_{E \setminus A} h d\mu \leq \lambda(E \setminus A)$$

ya que $h \in \Phi$. Sumando las dos últimas desigualdades obtenemos

$$\int_E (h + \varepsilon\chi_A) d\mu = \int_E h d\mu + \varepsilon\mu(A \cap E) \leq \lambda(E),$$

lo que implica que $h + \varepsilon\chi_A \in \Phi$. Pero

$$\int_X (h + \varepsilon\chi_A) d\mu = M + \varepsilon\mu(A) > M,$$

lo que contradice la definición de M. Esto completa la prueba del teorema en el caso en que $\mu(X) < \infty$, $\lambda(X) < \infty$. Obsérvese que podemos suponer $h(x) < \infty$ para todo x, porque al ser h integrable es finita en casi todo punto, y por tanto si es necesario podemos redefinirla en el conjunto donde toma el valor ∞ (poniendo que valga 1 o cualquier otro valor finito en ese conjunto) sin alterar el resultado de que h es medible y $\lambda(E) = \int_E h d\mu$ para todo $E \in \mathcal{A}$.

En el caso general, puesto que μ y λ son σ-finitas, se tiene $\bigcup_{n \in \mathbb{N}} A_n = X = \bigcup_{n \in \mathbb{N}} B_n$, con $\mu(A_n) < \infty$, $\lambda(B_n) < \infty$ y $A_k \cap A_n = \emptyset = B_k \cap B_n$ si $n \neq k$, luego también $X = \bigcup_{n,m \in \mathbb{N}} C_{n,m}$, donde $C_{n,m} := A_n \cap B_m$ son disjuntos dos a dos y tienen medida finita tanto para μ como para λ. Aplicando el caso anterior a cada conjunto $C_{n,m}$ se obtiene $h_{n,m} \in L^1(C_{n,m}, \mu)$ tal que $\lambda(E \cap C_{n,m}) = \int_{E \cap C_{n,m}} h_{n,m} d\mu$ para todo $E \in \mathcal{A}$. Definiendo $h_{n,m} = 0$ en $X \setminus C_{n,m}$ y $h = \sum_{n,m \in \mathbb{N}} h_{n,m}$, es obvio que $h : X \to [0, \infty)$ es medible y se comprueba fácilmente que $\lambda(E) = \int_E h d\mu$ para todo $E \in \mathcal{A}$. $\square$

10.3. El teorema de descomposición de Lebesgue

Vamos ahora a estudiar un teorema que complementa el de Radon-Nikodym y que nos da una pauta para comparar medidas σ-finitas cualesquiera en un mismo espacio medible.

Definición 10.4. Sea λ una medida (positiva) en un espacio medible $(X, \mathcal{A})$. Se dice que λ está *concentrada* en A (o que *vive* en A) si

$$\lambda(E) = \lambda(E \cap A) \text{ para todo } E \in \mathcal{A};$$

equivalentemente, si $\lambda(X \setminus A) = 0$.

Definición 10.5. Sean λ_1, λ_2 medidas (positivas) en un espacio medible $(X, \mathcal{A})$. Si existen $A, B \in \mathcal{A}$ tales que $A \cap B = \emptyset$, λ_1 está concentrada en A, y λ_2 está concentrada en B, entonces diremos que las medidas λ_1, λ_2 son *mutuamente singulares*, y escribiremos

$$\lambda_1 \perp \lambda_2.$$

Lema 10.4. *Sean λ_1, λ_2 medidas en un espacio $(X, \mathcal{A})$. Si existe $E \in \mathcal{A}$ tal que $\lambda_1(E) = 0$ y λ_2 está concentrada en E, entonces $\lambda_1 \perp \lambda_2$.*

Demostración. Basta tomar $A = X \setminus E$ y $B = E$ y se verifica la definición. □

Ejemplos 10.6.

1. Para cualquier medida λ en $\mathcal{A}$ y cualquier subconjunto propio A de X, con $A \in \mathcal{A}$, si definimos $\lambda_1(E) = \lambda(E \cap A)$ y $\lambda_2(E) = \lambda(E \cap A^c)$ para cada $E \in \mathcal{A}$, se tiene $\lambda_1 \perp \lambda_2$.

2. $\delta_a \perp \mathcal{L}^n$ para todo $a \in \mathbb{R}^n$, $n \geq 1$.

3. Si M es una subvariedad diferenciable de dimensión $k < n$ de $\mathbb{R}^n$ y $\lambda(E) := \mathcal{H}^k(E \cap M)$ para cada E medible Lebesgue, se tiene que $\mathcal{L}^n \perp \lambda$.

4. Si C es el conjunto ternario de Cantor y $\lambda(E) := \mathcal{H}^{\log 2/\log 3}(E \cap C)$, entonces $\lambda \perp \mathcal{L}^1$.

Teorema 10.5 (Lebesgue)**.** *Si λ y μ son medidas σ-finitas (positivas) en un espacio medible $(X, \mathcal{A})$, entonces existen medidas únicas λ_a y λ_s tales que*

$$\lambda = \lambda_a + \lambda_s, \quad \lambda_a \ll \mu, \quad y \quad \lambda_s \perp \mu.$$

Demostración. Sea $\nu = \mu + \lambda$, entonces $\mu \ll \nu$ y $\lambda \ll \nu$. Por el Teorema 10.3, existen $f, g : X \to [0, \infty]$ medibles tales que

$$\mu(E) = \int_E f d\nu, \quad \lambda(E) = \int_E g d\nu \tag{10.3}$$

para todo $E \in \mathcal{A}$. Definamos

$$\lambda_s(E) = \lambda\left(E \cap f^{-1}(\{0\})\right), \quad \lambda_a(E) = \lambda\left(E \cap f^{-1}((0, \infty])\right).$$

Es obvio que $\lambda = \lambda_a + \lambda_s$. Además λ_s está concentrada en $f^{-1}(\{0\})$, y puesto que, de acuerdo con (10.3), es $\mu(f^{-1}(\{0\})) = 0$, se tiene que $\lambda_s \perp \mu$.

Veamos que $\lambda_a \ll \mu$. Supongamos $\mu(E) = 0$; entonces, otra vez por (10.3), $f(x) = 0$ en ν-ctp $x \in E$, es decir, $\nu(E \cap f^{-1}((0, \infty])) = 0$, y por tanto

$$\lambda_a(E) = \lambda(E \cap f^{-1}((0, \infty])) = \int_{E \cap f^{-1}((0,\infty])} g d\nu = 0.$$

La prueba de la unicidad de λ_a y λ_s se deja como ejercicio. □

10.4. Los duales de los espacios L^p, $1 \leq p < \infty$

En esta sección vamos a aplicar el teorema de Radon-Nikodym para caracterizar el conjunto de los funcionales lineales continuos de $L^p(X)$ en $\mathbb{R}$ (o en $\mathbb{C}$).

Comenzamos con un lema cuya sencilla demostración se deja a cargo del lector (y que en todo caso se verá en la asignatura de Análisis Funcional).

Lema 10.6. *Sea $L : E \to F$ una aplicación lineal entre espacios vectoriales normados (sobre $\mathbb{R}$ o sobre $\mathbb{C}$). Son condiciones equivalentes:*

1. *L es continua;*

2. *L es continua en 0;*

3. *existe $C > 0$ tal que*

$$\|L(x)\| \leq C\|x\| \quad \text{para todo } x \in E. \tag{10.4}$$

Cuando se cumple (3) se dice que L es un funcional o un operador acotado. Esto, en el contexto de la teoría de operadores lineales del Análisis Funcional no significa que su imagen sea un conjunto acotado, sino que la imagen de cualquier conjunto acotado es acotada. Si L es continua, se define su norma

$$\|L\| := \sup_{\|x\|\leq 1} \|L(x)\| = \text{ínf}\{C > 0 : \text{ se cumple (10.4)}\}$$

(la igualdad entre este supremo y este ínfimo también queda como ejercicio pára el lector)

Definición 10.7. Si E es un espacio de Banach (sobre $\mathbb{R}$ o sobre $\mathbb{C}$), se define su espacio dual como

$$E^* := \{L : E \to \mathbb{R} \text{ (ó } \mathbb{C}) : L \text{ es lineal y continua}\}.$$

En E^* se considera la norma

$$\|L\| := \sup_{|x\|\leq 1} |L(x)|,$$

con la que se puede probar que E^* es un espacio de Banach (esto también se verá en Análisis Funcional).

Por ejemplo, si $p, q \in (1, \infty)$, con $\frac{1}{p} + \frac{1}{q} = 1$ y $g \in L^q$ entonces $L_g : L^p(\mu) \to \mathbb{R}$ (o $\mathbb{C}$) definida por

$$L_g(f) = \int_X fgd\mu$$

es un funcional lineal acotado, ya que por la desigualdad de Hölder se tiene

$$\left|\int_X fgd\mu\right| \leq \|g\|_q \|f\|_p;$$

en particular $\|L_g\| \leq \|g\|_q$. Análogamente, si $g \in L^\infty(\mu)$ entonces

$$L_g(f) := \int_X fgd\mu$$

también define un funcional lineal y continuo, con $\|L_g\| \leq \|g\|_\infty$. El siguiente teorema nos dice que, bajo condiciones muy razonables sobre $(X, \mathcal{A}, \mu)$, estos son todos los funcionales lineales y continuos que hay de $L^p(X)$ en $\mathbb{R}$ (o $\mathbb{C}$), para $p \in [1, \infty)$.

Teorema 10.7. *Supongamos que $(X, \mathcal{A}, \mu)$ es un espacio de medida σ-finito, $p \in [1, \infty]$, y que $L : L^p(\mu) \to \mathbb{R}$ (o $\mathbb{C}$) es un funcional lineal acotado. Entonces existe una única función $g \in L^q(\mu)$, donde $\frac{1}{p} + \frac{1}{q} = 1$, tal que*

$$L(f) = \int_X fgd\mu$$

para toda $f \in L^p(\mu)$. Además, $\|L\| = \|g\|_q$.

Demostración. Comencemos por la unicidad: si existen $g_1, g_2 \in L^q(\mu)$ tales que $\int_X fg_1 d\mu = L(f) = \int_X fg_2 d\mu$ para toda $f \in L^p(X)$, entonces, poniendo $g = \chi_E$, donde $E \in \mathcal{A}$ es cualquier conjunto medible con $\mu(E) < \infty$, se deduce que

$$\int_E (g_1 - g_2) = 0$$

para todo tal conjunto E, y esto implica que $g_1 = g_2$ en casi todo punto.

Ahora probaremos la existencia bajo la hipótesis adicional de que $\mu(X) < \infty$. El caso de medidas σ-finitas se deduce fácilmente de este y queda como ejercicio. Definamos, para cada $E \in \mathcal{E}$,

$$\lambda(E) = L(\chi_E).$$

Como L es lineal y $\chi_{A\cup B} = \chi_A + \chi_B$ siempre que $A \cap B = \emptyset$, es obvio que se tiene

$$\lambda(A \cup B) = \lambda(A) + \lambda(B) \quad \text{para todos } A, B \in \mathcal{A},\ \ A \cap B = \emptyset,$$

y por inducción se ve que λ es finitamente aditiva. Veamos que de hecho λ es numerablemente aditiva: si $E = \bigcup_{i\in\mathbb{N}} E_i$ unión disjunta de conjuntos en $\mathcal{A}$, poniendo $A_k = E_1 \cup ... \cup E_k$, se tiene que

$$\|\chi_E - \chi_{A_k}\|_p = \mu(E \setminus A_k)^{1/p} \to 0 \quad \text{cuando } k \to \infty,$$

y como L es continuo se deduce que

$$\sum_{i=1}^{k} \lambda(E_i) = L(\chi_{A_k}) \to L(\chi_E) \quad \text{cuando } k \to \infty,$$

lo que significa que $\sum_{i=1}^{\infty} \lambda(E_i) = \lambda(E)$ y prueba la aditividad numerable de λ. Obviamente también se tiene $\lambda(\emptyset) = 0$. Nótese que λ no es necesariamente positiva, y por tanto λ es una medida con signo finita, en el caso real, o compleja, en el caso complejo. En el caso real, por el teorema de descomposición de Hahn, $\lambda = \lambda^+ - \lambda^-$, donde λ^+ y λ^- son medidas positivas finitas que están concentradas en conjuntos disjuntos; por ello $\mu(E) = 0$ implica que $\lambda(E) = 0$ y también $\lambda^+(E) = 0 = \lambda^-(E)$. Es decir, λ^+ y λ^- son absolutamente continuas respecto de μ. Luego por el teorema de Radon-Nikodym, existen $g^{\pm} \in L^1(\mu)$ tales que

$$\lambda^{\pm}(E) = \int_E g^{\pm} d\mu.$$

para todo $E \in \mathcal{A}$. Y por tanto, para todo $E \in \mathcal{A}$,

$$\lambda(E) = \int_E g d\mu,$$

donde $g = g^+ - g^- \in L^1(\mu)$. Esto significa que

$$L(\chi_E) = \int_X \chi_E g d\mu$$

para todo tal E. Por linealidad de la integral se sigue que para toda función simple medible s también se tiene

$$L(s) = \int_X s g d\mu.$$

Como toda función $f \in L^\infty(\mu)$ es límite uniforme de una sucesión de funciones simples medibles y $\mu(X) < \infty$, por el teorema de la convergencia dominada se sigue que

$$L(f) = \int_X f g d\mu \tag{10.5}$$

para toda $f \in L^\infty(\mu)$. En el caso complejo se llega a la misma conclusión aplicando el argumento precedente a las partes reales e imaginarias de la medida λ (que son medidas reales con signo, finitas).

Nos quedan por probar las siguientes tres propiedades:

(a) $g \in L^q(\mu)$ (sabemos que $g \in L^1(\mu)$);

(b) $L(f) = \int_X f g d\mu$ para toda $f \in L^p(\mu)$ (lo sabemos para $p = \infty$);

(c) $\|L\| = \|g\|_q$ (suponiendo (a) y (b), sabemos que $\|L\| \leq \|g\|_p$ por la desigualdad de Hölder).

Haremos primero las demostraciones de estas afirmaciones en el caso $1 < p < \infty$. Sean

$$\alpha(x) = \begin{cases} |g(x)|/g(x) \text{ si } g(x) \neq 0, \\ 0 \text{ si } g(x) = 0, \end{cases}$$

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

$$E_n = \{x : |g(x)| \leq n\}.$$

Definamos también

$$f = |g|^{q/p} \alpha \chi_{E_n},$$

que está en L^∞ (ya que $|f| \leq n^{q/p}$). Además

$$|f|^p = |g|^q \chi_{E_n} = fg.$$

Por tanto

$$\int_{E_n} |g|^q d\mu = \int_X fg d\mu = L(f) \leq \|L\| \|f\|_p = \|L\| \left(\int_{E_n} |g|^q d\mu \right)^{1/p},$$

lo que nos da

$$\left(\int_{E_n} |g|^q d\mu \right)^{1-1/p} \leq \|L\|.$$

haciendo $n \to \infty$ en esta última desigualdad obtenemos que

$$\|g\|_q \leq \|L\| < \infty,$$

lo que en particular prueba (a) y la mitad de (c). Por otro lado, dada $f \in L^p(\mu)$, como las funciones simples son densas en $L^p(\mu)$ y están en $L^\infty(\mu)$, existe una sucesión $(f_k) \subset L^\infty(\mu)$ que converge a f en $L^p(\mu)$, y por tanto

$$\begin{aligned} \left| L(f) - \int_X fg d\mu \right| &\leq |L(f) - L(f_k)| + \left| L(f_k) - \int_X fg d\mu \right| \\ &= |L(f) - L(f_k)| + \left| \int_X (f_k - f) g d\mu \right| \\ &\leq |L(f) - L(f_k)| + \|f_k - f\|_p \|g\|_q \to 0 \quad \text{cuando } k \to \infty, \end{aligned}$$

lo que prueba (b). Finalmente, de (a) y (b) y la desigualdad de Hölder se deduce inmediatamente la otra mitad de (c).

Consideremos ahora el caso $p = 1$, $q = \infty$. Veamos primero que $\|g\|_\infty \leq \|L\|$: si no fuera así, existirían $r > \|L\|$ y A con $\mu(A) > 0$ tal que $|g(x)| \geq r$ para todo $x \in A$. Luego, considerando la función

$$f = \frac{1}{\mu(A)} \frac{|g|}{g} \chi_A,$$

que está en $L^\infty(\mu) \cap L^1(\mu)$ y cumple $\|f\|_1 = 1$, tendríamos

$$\|L\| \geq L(f) = \int_X fg d\mu = \frac{1}{\mu(A)} \int_A |g| d\mu \geq r > \|L\|,$$

lo que es absurdo. En particular se tiene que $g \in L^\infty(\mu)$. Ya sabemos así (a) y la mitad de (c). Para probar (b), dada $f \in L^1(\mu)$, tomamos una sucesión de funciones simples medibles (s_n) que converjan a f en $L^1(\mu)$ y cumplan $|s_n| \leq |f|$, y como $g \in L^\infty$, luego $|fg| \in L^1(\mu)$, podemos aplicar la continuidad de L y el teorema de la convergencia dominada para obtener que

$$L(f) = \lim_{n \to \infty} L(s_n) = \lim_{n \to \infty} \int_X s_n g d\mu = \int_X fg d\mu.$$

Una vez sabidas (a), (b) y una mitad de (c), aplicamos la desigualdad de Hölder para obtener $\|L\| \leq \|g\|_\infty$, es decir, la otra mitad de (c). $\square$

10.5. Problemas

Problema 10.1. Probar, como se dice en la sección 1 de este capítulo, que la descomposición de Hahn de una medida con signo es única salvo conjuntos de medida cero

Problema 10.2. Probar que la derivada de Radon-Nikodym de λ dada por el Teorema 10.3 es única cuando se identifican funciones que son iguales en casi todo punto.

Problema 10.3. Probar que si μ es una medida, y ν es una medida con signo tal que $\nu \ll \mu$ y $\mu \perp \nu$ entonces $\nu = 0$.

Problema 10.4. Probar la parte de unicidad del Teorema 10.5. *Indicación:* comenzar con el caso de medidas finitas y usar el problema anterior.

Problema 10.5. Sea μ una medida σ-finita en $(X, \mathcal{A})$. Si ν es una medida con signo finita en $(X, \mathcal{A})$ y $|\nu| \ll \mu$, demostrar que existe $f \in L^1(X, \mu)$ tal que $\nu(A) = \int_A f$ para todo $A \in \mathcal{A}$.

Problema 10.6. Sean $X = \mathbb{R}$, $\nu = \mathcal{L}^1$, y μ la medida de contar restringida a los conjuntos medibles Lebesgue de $\mathbb{R}$. Probar que $\nu \ll \mu$ y que el teorema de Radon-Nikodym falla para estas medidas. ¿Cuál es la hipótesis de este teorema que no se cumple?

Problema 10.7. Sean μ, ν medidas σ-finitas en $(X, \mathcal{A})$ tales que $\nu \ll \mu$, y sea $g = d\nu/d\mu$. Demostrar que $f \in L^1(X, \nu)$ si y sólo si $fg \in L^1(X, \mu)$ y, en este caso, $\int_E f d\nu = \int_E fg d\mu$ para cada $E \in \mathcal{A}$.

Problema 10.8. Probar que si ν es una medida finita y μ es σ-finita, entonces $\nu \ll \mu$ si y sólo si para todo $\varepsilon > 0$ existe $\delta > 0$ tal que $\mu(A) < \delta \implies \nu(A) < \varepsilon$.

Problema 10.9. Consideremos la medida ν definida en $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ definida por $\mu(A) = \int_A |x| dx$. Demostrar que $\nu \ll \mathcal{L}^1$ y sin embargo para ningún $\varepsilon > 0$ existe $\delta > 0$ tal que $|A| < \delta \implies \nu(A) < \varepsilon$.

Problema 10.10. Sea (r_n) una enumeración de $\mathbb{Q}$. Para cada $n \in \mathbb{N}$, sea $f_n : \mathbb{R} \to [0, \infty)$ medible con $f = f\chi_{[r_n - 2^{-n}, r_n + 2^{-n}]}$ y $\int_{\mathbb{R}} f_n(x) dx = 1$. Demostrar que $\sum_{n \in \mathbb{N}} f_n(x) < \infty$ para casi todo $x \in \mathbb{R}$. Si μ es la medida definida en los borelianos de $\mathbb{R}$ por $\mu(A) = \int_A \sum_{n \in \mathbb{N}} f_n(x) dx$, probar que μ es σ-finita, $\mu \ll \mathcal{L}^1$, y que $\mu(U) = \infty$ para todo U abierto no vacío.

Problema 10.11. Sean $\nu_j \ll \mu_j$, $j = 1, 2$, medidas σ-finitas. Demostrar que $\nu_1 \times \nu_2 \ll \mu_1 \times \mu_2$, y

$$\frac{d(\nu_1 \times \nu_2)}{d(\mu_1 \times \mu_2)}(x, y) = \frac{d\nu_1}{d\mu_1}(x) \frac{d\nu_2}{d\mu_2}(y).$$

Problema 10.12. Sean μ, ν dos medidas σ-finitas en $(X, \mathcal{A})$. Demostrar que las condiciones siguientes son equivalentes:

1. $\nu \ll \mu$ y $\mu \ll \nu$;
2. μ y ν tienen los mismos conjuntos de medida cero;
3. existe $g : X \to (0, \infty)$ tal que $\nu(A) = \int_A g d\mu$ para todo $A \in \mathcal{A}$.

Problema 10.13. Demostrar que si μ es una medida σ-finita en $(X, \mathcal{A})$ entonces existe una medida finita ν en $(X, \mathcal{A})$ tal que $\nu \ll \mu$ y $\mu \ll \nu$.

Problema 10.14. Sabemos que existen homeomorfismos $\varphi : \mathbb{R} \to \mathbb{R}$ con la propiedad de que para algunos conjuntos $E \subset \mathbb{R}$ con $|E| = 0$ el conjunto $\varphi(E)$ puede tener medida positiva o incluso no ser medible Lebesgue. Sin embargo, demuéstrese lo siguiente: si $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ es un homeomorfismo entonces existe $Z \subset \mathbb{R}^n$ de medida cero tal que, para todo conjunto medible $E \subset \mathbb{R}^n \setminus Z$, si $|E| = 0$ entonces $|\varphi(E)| = 0$. *Indicación:* probar que $\mu(A) = |\varphi(A)|$ define una medida en los borelianos de $\mathbb{R}^n$ y usar el teorema de descomposición de Lebesgue.

Capítulo 11

Diferenciación

En cierta medida, la integración y la derivación son procesos inversos: el teorema fundamental del cálculo que se estudia en primero y nos permite calcular con precisión muchas integrales es la muestra más elemental de este principio. Sin embargo, dicho teorema (en su versión estudiada en primero) no caracteriza las funciones $F : [a, b] \to \mathbb{R}$ para las cuales existe $f : [a, b] \to \mathbb{R}$ tal que

$$F(x) = \int_a^x f \text{ para todo } x \in [a, b]; \tag{11.1}$$

simplemente nos dice que si f es integrable en $[a, b]$ y continua en $x \in [a, b]$ entonces definiendo F mediante esta fórmula se tiene que $F'(x) = f(x)$. Así pues en la demostración de este teorema se ve que

$$\frac{F(x+h) - F(x)}{h} = \frac{1}{h} \int_x^{x+h} f \to f(x) \text{ cuando } h \to 0 \tag{11.2}$$

si f es continua en x.

Entre los propósitos principales de este capítulo están: 1) caracterizar la clase de funciones F para las que se cumple (11.1) para alguna $f \in L^1[a, b]$; y 2) estudiar cuándo se cumple (11.2) y sus versiones análogas para funciones de varias variables, lo que tiene aplicaciones importantes en el Análisis Real.

Empezaremos investigando la segunda cuestión, y luego volveremos al caso de $\mathbb{R}$ para resolver la primera cuestión. El siguiente lema generaliza el clásico argumento de la demostración del teorema fundamental del cálculo al contexto de funciones de varias variables.

Lema 11.1. *Si $f \in L^1(\mathbb{R}^d)$ es continua en x, entonces*

$$\lim_{x \in B,\ |B| \to 0} \frac{1}{|B|} \int_B f(y)dy = f(x),$$

donde el límite se toma cuando el volumen de las bolas B que contienen a x tiende a 0.

Demostración. Como f es continua en x, dado $\varepsilon > 0$ existe $\delta > 0$ tal que si $|y - x| \leq \delta$ entonces $|f(y) - f(x)| \leq \varepsilon$. Entonces, si B es una bola de radio menor que $\delta/2$ que contiene a x, se tiene

$$\left| f(x) - \frac{1}{|B|} \int_B f(y)dy \right| = \left| \frac{1}{|B|} \int_B (f(x) - f(y))dy \right| \leq \frac{1}{|B|} \int_B |f(x) - f(y)|dy \leq \varepsilon,$$

ya que para todo $y \in B$ es $|x - y| \leq \delta$. $\square$

Sin embargo, la condición de continuidad es innecesariamente fuerte. Como veremos más tarde, el mismo resultado vale para casi todo $x \in \mathbb{R}^n$, pidiendo únicamente que $f \in L^1(\mathbb{R}^d)$. Este es el contenido del teorema de diferenciación de Lebesgue.

11.1. La función maximal de Hardy-Littlewood

Para averiguar bajo qué condiciones puede tenerse

$$\lim_{r\to 0} \frac{1}{|B(x,r)|}\int_{B(x,r)} f(y)dy = f(x),$$

resulta conveniente usar la función maximal de Hardy-Littlewood, definida, para cada $f \in L^1(\mathbb{R}^d)$, por

$$\mathcal{M}(f)(x) = \sup\left\{\frac{1}{|B|}\int_B |f(y)|dy : x \in B \text{ bola en } \mathbb{R}^d\right\}, \quad x \in \mathbb{R}^d. \tag{11.3}$$

Es fácil ver que el valor de $\mathcal{M}(f)(x)$ es el mismo independientemente de si en la definición se consideran bolas abiertas o cerradas.

A su vez, para establecer las propiedades básicas de la función maximal, necesitaremos un sencillo lema de recubrimiento por bolas.

Lema 11.2. *Si W es la unión de una colección finita de bolas $B(x_i, r_i)$ en $\mathbb{R}^d$, $1 \le i \le N$, entonces existe un subconjunto $S \subset \{1, ..., N\}$ tal que*

1. *las bolas $B(x_i, r_i)$, $i \in S$, son disjuntas;*
2. *$W \subset \bigcup_{i\in S} B(x_i, 3r_i)$;*
3. *$|W| \le 3^d \sum_{i\in S} |B(x_i, r_i)|$.*

En este lema las bolas pueden suponerse tanto cerradas como abiertas; la demostración es igual en ambos casos.

Demostración. Reordenamos, si es preciso, las bolas $B_i = B(x_i, r_i)$ de forma que $r_1 \ge r_2 \ge ... \ge r_N$. Definamos $i_1 = 1$. Eliminamos todas las bolas B_j que corten a B_{i_1}, y sea B_{i_2} la primera de las B_j que quedan, si es que queda alguna. A continuación eliminamos todas las B_j con $j > i_2$ que corten a B_{i_2}, y sea B_{i_3} la primera de las que quedan. Continuamos este proceso mientras sea posible. Después de una cantidad finita de etapas el proceso se concluye, y ponemos $S = \{i_1, i_2, ...\}$.

Es claro que se cumple (1). Cada bola B_j de las que se han eliminado en alguna etapa del proceso corta a $B(x_i, r_i)$ para algún $i \in S$, y por tanto está contenida en $B(x_i, 3r_i)$ (ya que, si $r' \le r$ y $B(x', r')$ corta a $B(x, r)$, entonces $B(x', r')) \subset B(x, 3r)$). Esto prueba (2), y (3) se sigue inmediatamente de (2), porque $|B(x, 3r)| = 3^d|B(x, r)|$. □

Teorema 11.3. *Si $f \in L^1(\mathbb{R}^d)$ entonces:*

1. *$\mathcal{M}(f) : \mathbb{R}^n \to [0, \infty]$ es medible;*
2. *$\mathcal{M}(f)(x) < \infty$ para casi todo $x \in \mathbb{R}^d$;*
3. *$\mathcal{M}(f)$ cumple la desigualdad*

$$\Big|\{x \in \mathbb{R}^d : \mathcal{M}(f)(x) > \alpha\}\Big| \le \frac{3^d}{\alpha}\|f\|_1 \quad \textit{para todo } \alpha > 0. \tag{11.4}$$

Demostración. Para cada $\alpha \in [0, \infty]$, el conjunto

$$E_\alpha := \{x \in \mathbb{R}^d : \mathcal{M}(f)(x) > \alpha\}$$

es abierto (y en particular medible): en efecto, si $x_0 \in E_\alpha$ entonces existe una bola abierta $B \ni x_0$ tal que

$$\frac{1}{|B|}\int_B |f(y)|dy > \alpha.$$

Entonces, si $0 < \delta < \operatorname{dist}(x_0, \partial B)$, se tiene que $B(x_0, \delta) \subset B$, y por tanto

$$\mathcal{M}(f)(x) \geq \frac{1}{|B|} \int_B |f(y)| dy > \alpha$$

para todo $x \in B(x_0, \delta)$, es decir, $B(x_0, \delta) \subset E_\alpha$. Esto prueba (1).

Veamos ahora (3). Como antes, si $x \in E_\alpha$ existe una bola B_x que contiene a x y tal que

$$\frac{1}{|B_x|} \int_{B_x} |f(y)| dy > \alpha.$$

Así, para cada bola B_x se tiene

$$|B_x| < \frac{1}{\alpha} \int_{B_x} |f(y)| dy. \tag{11.5}$$

Sea K un subconjunto compacto cualquiera de E_α. Puesto que $K \subset \bigcup_{x \in E_\alpha} B_x$ y K es compacto, existe un subrecubrimiento finito, digamos $K \subset \bigcup_{\ell=1}^N B_\ell$. Entonces, por el lema anterior existe una subcolección $\{B_{i_1}, ..., B_{i_k}\}$ de bolas disjuntas tales que

$$\left| \bigcup_{\ell=1}^N B_\ell \right| \leq 3^d \sum_{j=1}^k |B_{i_j}|. \tag{11.6}$$

Puesto que las bolas $B_{i_1}, ..., B_{i_k}$ son disjuntas y satisfacen (11.5) y (11.6), se obtiene que

$$|K| \leq \left| \bigcup_{\ell=1}^N B_\ell \right| \leq 3^d \sum_{j=1}^k |B_{i_j}| \leq \frac{3^d}{\alpha} \sum_{j=1}^k \int_{B_{i_j}} |f(y)| dy = \frac{3^d}{\alpha} \int_{\bigcup_{j=1}^k B_{i_j}} |f(y)| dy \leq \frac{3^d}{\alpha} \|f\|_1.$$

Ahora, tomando supremos sobre todos los compactos K contenidos en E_α, se obtiene que

$$|E_\alpha| = \sup\{|K| : K \subset E_\alpha,\ K \text{ compacto}\} \leq \frac{3^d}{\alpha} \|f\|_1,$$

lo cual termina la prueba de (3).

Finalmente, (2) se deduce de (3) haciendo $\alpha \to \infty$. $\square$

Observación 11.1. La misma demostración prueba que, si μ es una medida de Borel positiva y finita, y se define

$$\mathcal{M}(\mu)(x) := \sup_B \frac{1}{|B|} \mu(B), \tag{11.7}$$

donde el supremo se toma sobre todas las bolas que contienen a x, entonces $\mathcal{M}(\mu)$ es medible, finita en casi todo punto, y cumple la desigualdad

$$\left| \{x \in \mathbb{R}^d : \mathcal{M}(\mu)(x) > \alpha\} \right| \leq \frac{3^d}{\alpha} \mu(\mathbb{R}^d).$$

Basta sustituir, en esta demostración, $\int_A |f(y)| dy$ por $\mu(A)$, donde A es una unión finita de bolas (en particular medible Borel). Más tarde haremos uso de esta observación.

11.2. El teorema de diferenciación de Lebesgue

Ahora, usando la cota para el operador maximal hallada en el teorema anterior, junto con la densidad del espacio $C_c(\mathbb{R}^d)$ en $L^1(\mathbb{R}^d)$ y el Lema 11.1, podemos dar una demostración sencilla del teorema siguiente.

Teorema 11.4 (de diferenciación de Lebesgue)**.** *Si $f \in L^1(\mathbb{R}^d)$, entonces*

$$\lim_{x\in B,\ |B|\to 0} \frac{1}{|B|}\int_B f(y)dy = f(x) \ \ \textit{para casi todo } x \in \mathbb{R}^d. \tag{11.8}$$

Demostración. Definamos, para cada $\alpha > 0$, el conjunto

$$E_\alpha = \left\{ x \in \mathbb{R}^d : \limsup_{x\in B,\ |B|\to 0} \left| \frac{1}{|B|}\int_B f(y)dy - f(x) \right| > 2\alpha \right\}.$$

Bastará probar que $|E_\alpha| = 0$ para cada $\alpha > 0$, ya que entonces $E := \bigcup_{n=1}^\infty E_{1/n}$ tendrá medida cero, y (11.8) se cumplirá en $\mathbb{R}^d \setminus E$.

Fijemos $\alpha > 0$, $\varepsilon > 0$, y con ayuda del Teorema 9.8 encontremos $g \in C_c(\mathbb{R}^d)$ tal que $\|f - g\|_1 < \varepsilon$. Por el lema 11.1 sabemos que

$$\lim_{x\in B,\ |B|\to 0} \frac{1}{|B|}\int_B g(y)dy = g(x) \ \ \text{para todo } x \in \mathbb{R}^d. \tag{11.9}$$

Puesto que

$$\frac{1}{|B|}\int_B f(y)dy - f(x) = \frac{1}{|B|}\int_B (f(y) - g(y))dy + \frac{1}{|B|}\int_B g(y)dy - g(x) + g(x) - f(x),$$

se tiene, usando también (11.9), que

$$\limsup_{x\in B,\ |B|\to 0} \left| \frac{1}{|B|}\int_B f(y)dy - f(x) \right| \le \mathcal{M}(f-g)(x) + |g(x) - f(x)|.$$

Luego, si definimos

$$F_\alpha := \{x \in \mathbb{R}^d : \mathcal{M}(f-g)(x) > \alpha\}, \quad \text{y} \quad G_\alpha = \{x \in \mathbb{R}^d : |f(x) - g(x)| > \alpha\},$$

se tiene que $E_\alpha \subset F_\alpha \cup G_\alpha$ (porque si u_1, u_2 son números positivos, sólo puede ocurrir que $u_1 + u_2 > 2\alpha$ si se tiene $u_i > \alpha$ para al menos uno de los u_i). Además, por la desigualdad de Tchebychev,

$$|G_\alpha| \le \frac{1}{\alpha}\|f - g\|_1,$$

y por (3) del teorema anterior,

$$|F_\alpha| \le \frac{3^d}{\alpha}\|f - g\|_1.$$

Deducimos por tanto que

$$|E_\alpha| \le |F_\alpha| + |G_\alpha| \le \frac{1}{\alpha}\|f - g\|_1 + \frac{3^d}{\alpha}\|f - g\|_1 \le \frac{3^d}{\alpha}\varepsilon + \frac{1}{\alpha}\varepsilon,$$

y como ε es arbitrario se concluye que $|E_\alpha| = 0$. □

Corolario 11.5. *Si λ es una medida finita tal que $\lambda \ll \mathcal{L}_d$, entonces su derivada de Radon-Nikodym cumple*

$$\frac{d\lambda}{d\mathcal{L}_d}(x) = \lim_{r\to 0} \frac{1}{|B(x,r)|}\lambda\left(B(x,r)\right)$$

para casi todo $x \in \mathbb{R}^n$.

Demostración. Por el teorema de Radon-Nikodym existe $\frac{d\lambda}{d\mathcal{L}_d} \in L^1(\mathbb{R}^d)$, y, por el teorema de diferenciación de Lebesgue,

$$\frac{\lambda\left(B(x,r)\right)}{|B(x,r)|} = \frac{1}{|B(x,r)|}\int_{B(x,r)} \frac{d\lambda}{d\mathcal{L}_d} d\mathcal{L}_d \to \frac{d\lambda}{d\mathcal{L}_d}(x)$$

cuando $r \to 0$, para casi todo $x \in \mathbb{R}^d$. □

La demostración del siguiente corolario es fácil y se deja como ejercicio. Recuérdese que $L^1_{\text{loc}}(\mathbb{R}^d)$ es el espacio vectorial de las funciones localmente integrables, esto es, las funciones medibles f tales que para toda bola B $f\chi_B$ es integrable.

Corolario 11.6. *Si $f \in L^1_{loc}(\mathbb{R}^d)$, entonces*

$$\lim_{x\in B,\ |B|\to 0} \frac{1}{|B|}\int_B f(y)dy = f(x) \quad \textit{para casi todo } x \in \mathbb{R}^d. \tag{11.10}$$

11.3. Puntos de Lebesgue

Una aplicación muy interesante y útil del teorema de diferenciación de Lebesgue es cuando se considera $f = \chi_E$, siendo E un subconjunto medible cualquiera de $\mathbb{R}^d$. Nótese que $\chi_E \in L^1_{\text{loc}}(\mathbb{R}^d)$ por ser medible y acotada. Se obtiene entonces que

$$\lim_{x\in B,\ |B|\to 0} \frac{|B\cap E|}{|B|} = 1 \tag{11.11}$$

para casi todo $x \in E$. Geométricamente, esto significa que para cualquier $\alpha < 1$ tan cerca de 1 como se quiera, y para cualquier bola que contenga a x y tenga radio suficientemente pequeño, se tiene que $|B\cap E| \geq \alpha|B|$, y por tanto E cubre al menos una proporción α de la bola B. Cuando (11.11) se cumple para un x, se dice que x es un *punto de densidad de Lebesgue* de E. Tenemos así, considerando también puntos $x \in \mathbb{R}^d \setminus E$ y aplicando el corolario anterior con la misma función χ_E, lo siguiente

Corolario 11.7. *Sea $E \subset \mathbb{R}^d$ medible. Entonces:*

1. *Casi todo $x \in E$ es un punto de densidad de Lebesgue de E. Es decir,*

$$\lim_{x\in B,\ |B|\to 0} \frac{|B\cap E|}{|B|} = 1$$

para casi todo $x \in E$.

2. *Para casi todo $x \in \mathbb{R}^d \setminus E$ se tiene que*

$$\lim_{x\in B,\ |B|\to 0} \frac{|B\cap E|}{|B|} = 0.$$

También se habla de puntos de Lebesgue de funciones: si $f \in L^1_{\text{loc}}(\mathbb{R}^d)$, se dice que x *es un punto de Lebesgue de* f si

$$\lim_{x\in B,\ |B|\to 0} \frac{1}{|B|}\int_B |f(y) - f(x)|dy = 0,$$

lo que obviamente implica $\lim_{x\in B,\ |B|\to 0} \frac{1}{|B|}\int_B f(y)dy = f(x)$.

Corolario 11.8. *Sea $f \in L^1_{loc}(\mathbb{R}^d)$. Entonces casi todo $x \in \mathbb{R}^d$ es un punto de Lebesgue de f.*

Demostración. Para cada $c \in \mathbb{R}$, aplicando el Corolario 11.6 se tiene que

$$\lim_{x\in B,\ |B|\to 0} \frac{1}{|B|}\int_B |f(y) - c|dy = |f(x) - c| \text{ para todo } x \in \mathbb{R}^d \setminus E_c, \tag{11.12}$$

donde $|E_c| = 0$. Sea $E := \bigcup_{c\in\mathbb{Q}} E_c$. Como $\mathbb{Q}$ es numerable, se tiene que $|E| = 0$. Ahora, para todo $x \in \mathbb{R}^d \setminus E$ con $|f(x)| < \infty$ (o sea, para casi todo $x \in \mathbb{R}^d$), dado $\varepsilon > 0$ existe $c \in \mathbb{Q}$ tal que $|f(x) - c| < \varepsilon$, y como

$$\frac{1}{|B|}\int_B |f(y) - f(x)|dy \leq \frac{1}{|B|}\int_B |f(y) - c|dy + |f(x) - c|,$$

se deduce que

$$\limsup_{x\in B,\ |B|\to 0} \frac{1}{|B|}\int_B |f(y)-f(x)|dy \le 2\varepsilon,$$

y al ser ε arbitrario se concluye que

$$\lim_{x\in B,\ |B|\to 0} \frac{1}{|B|}\int_B |f(y)-f(x)|dy = 0.$$

□

Observación 11.2. Recuérdese que los elementos del espacio de Banach $L^1(\mathbb{R}^d)$ son en realidad clases de equivalencia de funciones (estando dos funciones en la misma clase de equivalencia si y sólo si coinciden en casi todo punto). Es interesante observar que el conjunto de puntos x para los que existe

$$\lim_{x\in B,\ |B|\to 0} \frac{1}{|B|}\int_B f(y)dy$$

es independiente del representante de la clase de equivalencia escogido, puesto que si dos funciones f y $\tilde{f}$ coinciden en casi todo punto entonces sus promedios integrales sobre cualquier bola son los mismos.

Sin embargo, el conjunto de Lebesgue de f sí que depende del representante escogido. Por ello, para evitar ambigüedades, cuando se habla de puntos de Lebesgue, suele escogerse como representante la función

$$x \mapsto \limsup_{r\to 0} \frac{1}{|B(x,r)|}\int_{B(x,r)} f(y)dy,$$

que está bien definida para todo $x\in\mathbb{R}^n$. A esta función, por abuso de notación, también la llamamos f.

Definición 11.3. Se dice que una familia $\mathcal{F}$ de subconjuntos medibles de $\mathbb{R}^n$ es *regular en un punto* $x\in\mathbb{R}^d$ si existe una constante $C>0$ tal que para cada $S\in\mathcal{F}$ existe una bola $B(x,r_S)$ tal que $S\subset B(x,r_S)$ y $|B(x,r_S)|\le C|S|$, y si para todo $\varepsilon>0$ existe $S\in\mathcal{F}$ tal que $|S|<\varepsilon$.

Obsérvese que no se pide que $x\in S$, ni siquiera que x esté en la adherencia de S; únicamente que cada S ocupe una porción significativa de la bola $B(x,r_S)$ (y que haya conjuntos S de medida arbitrariamente pequeña).

Teorema 11.9. *Si x es un punto de Lebesgue de $f\in L^1_{loc}(\mathbb{R}^d)$ y $\mathcal{F}$ es regular en x, entonces*

$$\lim_{S\in\mathcal{F},|S|\to 0} \frac{1}{|S|}\int_S f(y)dy = f(x).$$

En particular esto es verdad para casi todo $x\in\mathbb{R}^n$.

Demostración. Puesto que $|B(x,r_S)|\le C|S|$, se tiene que si $|S|\to 0$ entonces $r_S\to 0$. Entonces, por definición de r_S y usando el Teorema 11.8 obtenemos

$$\begin{aligned}\left|\frac{1}{|S|}\int_S f(y)dy - f(x)\right| &\le \frac{1}{|S|}\int_S |f(y)-f(x)|dy \le \frac{1}{|S|}\int_{B(x,r_S)} |f(y)-f(x)|dy\\ &\le \frac{|B(x,r_S)|}{|S|}\frac{1}{|B(x,r_S)|}\int_{B(x,r_S)} |f(y)-f(x)|dy \le C\frac{1}{|B(x,r_S)|}\int_{B(x,r_S)} |f(y)-f(x)|dy \to 0\end{aligned}$$

cuando $|S|\to 0$. □

11.4. Propiedades de las partes singulares de medidas de Borel en $\mathbb{R}^d$.

En esta sección λ denotará una medida finita de Borel en $\mathbb{R}^d$, y μ la restricción de la medida de Lebesgue a $\mathcal{B}(\mathbb{R}^d)$. Por el teorema de descomposición de Lebesgue, sabemos que existen medidas de Borel σ-finitas λ_a y λ_s tales que

$$\lambda = \lambda_a + \lambda_s, \;\; \lambda_a \ll \mu, \;\; \lambda_s \perp \mu. \tag{11.13}$$

A su vez, por el teorema de Radon-Nikodym, existe $f \in L^1(\mathbb{R}^d)$ tal que

$$\lambda_a(E) = \int_E f d\mu \quad \text{para todo } E \text{ boreliano.}$$

Aplicando el Teorema 11.9 a la parte absolutamente continua de λ obtenemos que, si $\mathcal{F}_x$ es una familia regular en x para cada x, entonces

$$\lim_{S\in\mathcal{F}_x,\, |S|\to 0} \frac{1}{\mu(S)}\lambda_a(S) = f(x) \text{ para } \mu\text{-casi todo } x \in \mathbb{R}^n. \tag{11.14}$$

El siguiente resultado nos dice que este tipo de límite, aplicado a la parte singular λ_s de la medida λ da cero en μ-casi todo punto.

Teorema 11.10. *Sean ν una medida de Borel finita (positiva) en $\mathbb{R}^d$, y μ es la restricción de la medida de Lebesgue a los borelianos de $\mathbb{R}^d$: Si $\nu \perp \mu$, entonces se tiene que*

$$\lim_{x\in B,\, |B|\to 0} \frac{1}{\mu(B)}\nu(B) = 0 \quad \text{para } \mu\text{-casi todo } x \in \mathbb{R}^d.$$

Similarmente, si para cada $x \in \mathbb{R}^d$ $\mathcal{F}_x$ es una familia de subconjuntos medibles de $\mathbb{R}^n$ que es regular en x, entonces para μ-casi todo $x \in \mathbb{R}^d$ es

$$\lim_{S\in\mathcal{F}_x,\, |S|\to 0} \frac{1}{\mu(S)}\nu(S) = 0.$$

Demostración. Sean $\alpha, \varepsilon > 0$. Puesto que $\nu \perp \mu$, ν está concentrada en un conjunto de medida de Lebesgue 0. Como ν es de Radon, existe un compacto K con $\nu(K) > \nu(\mathbb{R}^d) - \varepsilon$, y necesariamente $\mu(K) = 0$.

Definamos

$$\nu_1(E) = \nu(K \cap E) \quad \text{para todo boreliano } E,$$

y sea

$$\nu_2 = \nu - \nu_1,$$

que es obviamente una medida de Borel positiva y finita, con $\nu_2(\mathbb{R}^d) < \varepsilon$. Para todo $x \in \mathbb{R}^d \setminus K$ se tiene

$$\limsup_{x\in B,\, |B|\to 0} \frac{\nu(B)}{\mu(B)} = \limsup_{x\in B,\, |B|\to 0} \frac{\nu_2(B)}{\mu(B)} \le \mathcal{M}(\nu_2)(x).$$

Luego

$$\{x \in \mathbb{R}^d : \limsup_{x\in B,\, |B|\to 0} \frac{\nu(B)}{\mu(B)} > \alpha\} \subset K \cup \{x \in \mathbb{R}^d : \mathcal{M}(\nu_2)(x) > \alpha\},$$

y por la Observación 11.1 se verifica (teniendo en cuenta el contenido anterior y el hecho de que $\mu(K) = 0$) que

$$\mu\left(\left\{x \in \mathbb{R}^d : \limsup_{x\in B,\, |B|\to 0} \frac{\nu(B)}{\mu(B)} > \alpha\right\}\right) \le \frac{3^d}{\alpha}\nu_2(\mathbb{R}^d) < \frac{3^d}{\alpha}\varepsilon.$$

Como lo anterior es cierto para todos los $\alpha, \varepsilon > 0$, se concluye primero, tomando $\alpha = 1/j$ y haciendo $\varepsilon \to 0$, que

$$\mu\left(\left\{x \in \mathbb{R}^d : \limsup_{x\in B,\, |B|\to 0} \frac{\nu(B)}{\mu(B)} > \frac{1}{j}\right\}\right) = 0$$

para todo $j \in \mathbb{N}$, y entonces también

$$\mu\left(\left\{x \in \mathbb{R}^d : \limsup_{x\in B,\, |B|\to 0} \frac{\nu(B)}{\mu(B)} > 0\right\}\right) = \mu\left(\bigcup_{j=1}^{\infty}\left\{x \in \mathbb{R}^d : \limsup_{x\in B,\, |B|\to 0} \frac{\nu(B)}{\mu(B)} > \frac{1}{j}\right\}\right) = 0,$$

es decir,

$$\lim_{x\in B,\ |B|\to 0} \frac{1}{\mu(B)}\nu(B) = 0$$

para μ-casi todo $x \in \mathbb{R}^n$.

El enunciado para familias regulares en x es una consecuencia inmediata de lo que acabamos de demostrar, ya que

$$\frac{\nu(S)}{\mu(S)} \leq \frac{\nu(B(x, r_S))}{\mu(S)} \leq C\frac{\nu(B(x, r_S))}{\mu(B(x, r_S))}.$$

□

En contraste con el resultado anterior, puede probarse también que, si ν es una medida de Borel positiva en $\mathbb{R}^d$ y $\nu \perp \mu$ (donde μ es la medida de Lebesgue restringida a los borelianos de $\mathbb{R}^d$), entonces

$$\lim_{r\to 0} \frac{\nu\left(B(x,r)\right)}{\mu\left(B(x,r)\right)} = \infty \quad \text{para } \nu\text{-casi todo punto } x \in \mathbb{R}^d.$$

No usaremos este otro resultado. Su demostración puede consultarse en [5, Teorema 7.15].

Combinando (11.13) y (11.14) con el Teorema 11.10 (aplicado con $\nu = \lambda_s$) se obtiene inmediatamente lo siguiente.

Corolario 11.11. *Si λ es una medida de Borel positiva y finita en $\mathbb{R}^d$, y μ denota la restricción de la medida de Lebesgue a los borelianos de $\mathbb{R}^d$, entonces existe $f \in L^1(\mathbb{R}^d)$ tal que, para cualquier conjunto $\{\mathcal{F}_x\}_{x\in\mathbb{R}^n}$ de familias regulares se tiene que*

$$\lim_{S\in\mathcal{F}_x,\ |S|\to 0} \frac{1}{\mu(S)}\lambda(S) = f(x) \quad \textit{para } \mu\textit{-casi todo } x \in \mathbb{R}^n.$$

De hecho f es la derivada de Radon-Nikodym de la parte absolutamente continua de la descomposición de Lebesgue de la medida λ.

11.5. El teorema fundamental del cálculo

En esta sección vamos a caracterizar la clase de funciones F para las que vale el teorema fundamental del cálculo, es decir,

$$F(x) = F(a) + \int_a^x f(t)dt \tag{11.15}$$

para todo $x \in [a, b]$ y alguna f integrable tal que $F'(x) = f(x)$ en casi todo punto. La parte más fácil de este trabajo es una simple consecuencia del Teorema 11.4: puesto que en una dimensión las bolas son intervalos, se tiene lo siguiente.

Teorema 11.12. *Si $f \in L^1([a, b])$ y $F(x) := \int_a^x f(y)dy$, entonces F es derivable en casi todo punto $x \in [a, b]$, con $F'(x) = f(x)$.*

Demostración. Sea x un punto para el que vale el Teorema 11.4. Sea $(h_n)_{n\in\mathbb{N}}$ una sucesión de números estrictamente positivos que tienda a 0. Como la sucesión de las longitudes de los intervalos $[x, x+h_n]$ tiende a 0 y todos ellos contienen a x, el Teorema 11.4 nos dice que

$$\lim_{n\to\infty} \frac{F(x+h_n)-F(x)}{h_n} = \lim_{n\to\infty} \frac{\int_{[x,x+h_n]} f(y)dy}{h_n} = f(x).$$

Como (h_n) es arbitraria, esto prueba que existe la derivada lateral por la derecha de F en x y que coincide con $f(x)$. Análogamente, considerando los intervalos $[x-h_n, x]$, se ve que existe la derivada lateral por la izquierda de F en x y también coincide con $f(x)$. Por tanto existe $F'(x) = f(x)$. $\square$

Sabemos que cualquier función F de la forma

$$F(x) = c + \int_a^x f(y)dy, \quad x \in [a,b], \tag{11.16}$$

donde $f \in L^1([a,b])$, cumple que para todo $\varepsilon > 0$ existe $\delta > 0$ tal que si $\mu(E) < \delta$ entonces $\int_E |f| < \varepsilon$. Por tanto, si se toma E como la unión de una colección finita de intervalos abiertos disjuntos (a_i, b_i), $i = 1, ..., n$ cuyas longitudes sumen menos que δ, se tendrá que

$$\sum_{j=1}^n |F(b_j) - F(a_j)| = \sum_{j=1}^n |\int_{a_j}^{b_j} f(y)dy| \leq \sum_{j=1}^n \int_{a_j}^{b_j} |f(y)|dy = \int_E |f| < \varepsilon.$$

En lo que sigue vamos a probar que esta propiedad, además de necesaria para que se cumpla (11.16), también es suficiente. La clase de funciones con esta propiedad recibe un nombre especial.

Definición 11.4. Se dice que una función $F : [a,b] \to \mathbb{R}$ es *absolutamente continua* si para todo $\varepsilon > 0$ existe $\delta > 0$ tal que si (a_i, b_i), $i = 1, ..., n$, es una colección finita de subintervalos abiertos disjuntos dos a dos con $\sum_{j=1}^n (b_i - a_i) < \delta$ entonces

$$\sum_{j=1}^n |F(b_i) - F(a_i)| < \varepsilon.$$

Nótese que toda función F que verifique la definición anterior es uniformemente continua: tómese $n = 1$ y $(a_1, b_1) = (x, y) \cup (y, x)$ en dicha definición para constatarlo.

Para ver que toda función absolutamente continua F satisface (11.15), vamos a comenzar estudiando el caso en que F es además creciente.

Teorema 11.13. *Sea $F : [a,b] \to \mathbb{R}$ continua y creciente. Son equivalentes:*

(a) F es absolutamente continua;

(b) F aplica conjuntos de medida cero en conjuntos de medida cero;

(c) F es diferenciable en casi todo punto, con $F' \in L^1([a,b])$, y

$$F(x) = F(a) + \int_a^x F'(y)dy$$

para todo $x \in [a,b]$.

Demostración. $(a) \implies (b)$: Dado $\varepsilon > 0$, sea $\delta > 0$ como en la definición anterior. Dado E con $|E| = 0$, para probar que $|F(E)| = 0$, podemos suponer sin pérdida de generalidad que ni a ni b están en E. Existe V abierto con $|V| < \delta$ y $E \subset V \subset [a,b]$, y V es unión numerable de intervalos abiertos disjuntos (a_i, b_i), $i \in \mathbb{N}$. Entonces $\sum_{i=1}^{\infty}(b_i - a_i) < \delta$, luego también $\sum_{i=1}^{n}(b_i - a_i) < \delta$ para cualquier $n \in \mathbb{N}$; por tanto $\sum_{i=1}^{n} |F(b_i) - F(a_i)| < \varepsilon$ para cualquier $n \in \mathbb{N}$, y así, dado que F es creciente,

$$|F(E)| \leq |F(V)| = \left| \bigcup_{i=1}^{\infty} (F(a_i), F(b_i)) \right| = \sum_{i=1}^{\infty} |F(b_i) - F(a_i)| \leq \varepsilon.$$

Como ε es arbitrario, se deduce que $|F(E)| = 0$.

$(b) \implies (c)$: Definamos

$$g(x) = x + F(x), \quad x \in [a, b],$$

y observemos que g es estrictamente creciente, y en particular inyectiva; de hecho es un homeomorfismo sobre su imagen. Observemos que si la imagen por F de un intervalo de longitud η tiene longitud η', entonces la imagen por g del mismo intervalo tendrá longitud $\eta + \eta'$. De este hecho se puede deducir que g también verifica (b); esta tarea queda como ejercicio para el lector. Ahora, dado $E \subset [a, b]$ medible Lebesgue, podemos escribir $E = E_0 \cup E_1$, donde $|E_0| = 0$ y E_1 es un F_σ; por tanto E_1 es unión numerable de cerrados contenidos en $[a, b]$, y entonces $g(E_1)$ también lo es, por ser g homeomorfismo. Puesto que g satisface (b), se tiene $|g(E_0)| = 0$, y como $g(E) = g(E_1) \cup g(E_0)$, se deduce que $g(E)$ es medible Lebesgue. Por lo tanto, podemos definir

$$\mu(E) = |g(E)|$$

para cada $E \subset [a, b]$ medible Lebesgue. Dado que g es inyectiva, conjuntos disjuntos en $[a, b]$ tienen imágenes disjuntas, y la aditividad numerable de la medida de Lebesgue muestra entonces que μ es una medida (positiva y finita) en la σ-álgebra de los subconjuntos medibles Lebesgue de $[a, b]$. Esta medida μ es absolutamente continua respecto de $\mathcal{L}_1$, ya que g satisface (b). Luego, por el teorema de Radon-Nikodym, existe $h \in L^1([a, b])$ tal que

$$\mu(E) = \int_E h(y)dy$$

para cada $E \subset [a, b]$ medible Lebesgue. Tomando $E = [a, x]$, se tiene $g(E) = [g(a), g(x)]$ por ser g estrictamente creciente, y así

$$g(x) - g(a) = |g(E)| = \mu(E) = \int_E h(y)dy = \int_a^x h(y)dy.$$

Como $g(x) = x + F(x)$, se concluye que

$$F(x) - F(a) = \int_a^x (h(y) - 1)dy$$

para todo $x \in [a, b]$, y por el Teorema 11.12, $F'(x) = h(x) - 1$ en casi todo $x \in [a, b]$, luego se cumple (c).

$(c) \implies (a)$: esto se ha visto en la discusión que precede la Definición 11.4. □

Ahora veremos cómo cualquier función absolutamente continua es diferencia de dos funciones absolutamente continua y crecientes, a las que podremos aplicar el teorema anterior.

Teorema 11.14. *Si $F : [a, b] \to \mathbb{R}$ es absolutamente continua y definimos, para cada $x \in [a, b]$,*

$$V_F(x) = \sup \sum_{i=1}^{N} |F(t_i) - F(t_{i-1})|,$$

donde el supremo se toma sobre todos los $N \in \mathbb{N}$ y todas las elecciones de puntos $\{t_i\}$ tales que

$$a = t_0 < t_1 < ... < t_N = x,$$

entonces se tiene que las funciones V_F, $V_F + F$, $V_F - F$ son crecientes y absolutamente continuas en $[a, b]$.

A V_F se le llama *función variación total de F*. Si F es cualquier función en $[a, b]$ (absolutamente continua o no), y $V_F(b) < \infty$, se dice que F es de *variación acotada*, y a $V_F(b)$ se le llama la *variación total de F* en $[a, b]$. Es obvio que toda función creciente en un intervalo cerrado y acotado $[a, b]$ es de variación acotada, con variación total $V_F(b) = F(b) - F(a)$. También es prácticamente obvio que la suma (o la diferencia) de dos funciones de variación acotada es de variación acotada. Por consiguiente, toda función absolutamente continua es de variación acotada, pero hay funciones de variación acotada que no son absolutamente continuas.

Demostración. Sea una colección de puntos $\{t_i\}$ como la del enunciado, y supongamos que $a \le x < y \le b$, entonces

$$V_F(y) \ge |F(y) - F(x)| + \sum_{i=1}^{N} |F(t_i) - F(t_{i-1})|.$$

Por lo tanto, tomando supremos sobre tales $\{t_i\}$, se tiene que $V_F(y) \ge |F(y) - F(x)| + V_F(x)$. En particular

$$V_F(y) \ge F(y) - F(x) + V_F(x), \quad V_F(y) \ge F(x) - F(y) + V_F(x).$$

Esto prueba que V_F, $V_F + F$ y $V_F - F$ son crecientes. Puesto que las sumas y restas de funciones absolutamente continuas son (casi obviamente) absolutamente continuas, sólo hay que probar que V_F es absolutamente continua. Si $(\alpha, \beta) \subset [a, b]$ entonces es fácil ver que

$$V_F(\beta) - V_F(\alpha) = \sup \sum_{i=1}^{n} |F(t_i) - F(t_{i-1})|, \tag{11.17}$$

donde el supremo se toma sobre todos los $\{t_i\}$ tales que $\alpha = t_0 < t_1 < ... < t_n = \beta$. Obsérvese que $\sum_{i=1}^{n}(t_i - t_{i-1}) = \beta - \alpha$. Ahora, dado $\varepsilon > 0$ elijamos $\delta > 0$ satisfaciendo la Definición 11.4. Para cualquier colección finita de intervalos disjuntos $(\alpha_1, \beta_1), ..., (\alpha_N, \beta_N) \subset [a, b]$ con $\sum_{j=1}^{N}(\beta_j - \alpha_j) < \delta$, aplicando (11.17) a cada intervalo (α_j, β_j), tenemos que existen $\alpha_j = t_0^j < t_1^j < ... < t_{n_j}^j = \beta_j$ con

$$V_F(\beta_j) - V_F(\alpha_j) \le \frac{\varepsilon}{N} + \sum_{i=1}^{n_j} |F(t_i^j) - F(t_{i-1}^j)|,$$

luego, teniendo en cuenta la elección de δ y el hecho de que $\sum_j \sum_i (t_i^j - t_{i-1}^j) = \sum_j (\beta_j - \alpha_j) < \delta$, se deduce que

$$\sum_{j=1}^{N}(V_F(\beta_j) - V_F(\alpha_j)) \le \varepsilon + \sum_{j=1}^{N}\sum_{i=1}^{n_j} |F(t_i^j) - F(t_{i-1}^j)| \le 2\varepsilon.$$

Como V_F es creciente, esto prueba que V_F es absolutamente continua en $[a, b]$. □

Observación 11.5. Si F es de variación acotada, el mismo argumento de la primera parte de la demostración anterior prueba que V_F, $V_F + F$ y $V_F - F$ son funciones crecientes. Por tanto

$$F = \frac{1}{2}(V_F + F) - \frac{1}{2}(V_F - F)$$

es diferencia de dos funciones crecientes. Es decir, toda función de variación acotada (y en particular toda función absolutamente continua) es diferencia de funciones crecientes.

Ahora ya podemos demostrar que las funciones absolutamente continuas son precisamente las que cumplen el teorema fundamental del cálculo.

Teorema 11.15 (Lebesgue). *Sea $F : [a, b] \to \mathbb{R}$. Las siguientes afirmaciones son equivalentes:*

1. *F es absolutamente continua;*

2. *F es derivable en casi todo punto de $[a,b]$, $F' \in L^1([a,b])$, y*

$$F(x) = F(a) + \int_a^x F'(t)dt$$

para todo $x \in [a,b]$;

3. *Existe $g \in L^1([a,b])$ tal que*

$$F(x) = F(a) + \int_a^x g(t)dt.$$

Demostración. (1) $\implies$ (2): Por la observación anterior, $F = F_1 - F_2$, donde F_1 y F_2 son crecientes y absolutamente continuas. Aplicando el Teorema 11.13 a F_1 y F_2 obtenemos que F_1 y F_2 son diferenciables en casi todo punto, con $F_1', F_2' \in L^1([a,b])$, y

$$F_i(x) = F_i(a) + \int_a^x F_i'(y)dy,$$

$i = 1, 2$. Luego F también es diferenciable en casi todo punto, con $F' = F_1' - F_2' \in L^1([a,b])$, y

$$F(x) = F_1(x) - F_2(x) = F_1(a) + \int_a^x F_1'(t)dt - F_2(a) - \int_a^x F_2'(t)dt$$
$$= F(a) + \int_a^x \big(F_1'(t) - F_2'(t)\big)\, dt = F(a) + \int_a^x F'(t)dt.$$

(2) $\implies$ (3) es trivial, y (3) $\implies$ (1): esto está demostrado en la motivación de la Definición 11.4. □

11.6. Diferenciabilidad en casi todo punto de las funciones de variación acotada

Como aplicación de los teoremas sobre diferenciación de medidas de las secciones anteriores, vamos a dar una prueba rápida del hecho de que toda función creciente es diferenciable en casi todo punto.

Teorema 11.16 (Lebesgue). *Si $f : [a,b] \to \mathbb{R}$ es creciente, entonces f es diferenciable en casi todo punto.*

Obviamente el mismo resultado es verdad cambiando creciente por decreciente.

Demostración. Como f es diferenciable en casi todo punto si y sólo $x \mapsto x + f(x)$ es diferenciable en casi todo punto, podemos suponer que f es estrictamente creciente. Sabemos que el conjunto

$$D := \{x \in [a,b] : f \text{ es discontinua en } x\}$$

es a lo sumo numerable, con

$$f(x^-) := \lim_{y \to x^-} f(x) < \lim_{y \to x^+} f(x) := f(x^+) \quad \text{y} \quad f(x^-) \le f(x) \le f(x^+) \quad \text{para todo } x \in D.$$

Definamos

$$h(x) = \begin{cases} f(x) & \text{si } x \in [a,b] \setminus D; \\ f(x^-) & \text{si } x \in D; \end{cases}$$

es claro que $\{x : f(x) \ne h(x)\}$ es a lo sumo numerable y que h es continua por la izquierda. Además, si h es derivable en $x \in [a,b] \setminus D$ entonces f también lo es: en efecto, dado $\varepsilon > 0$, existe $\delta > 0$ tal que si $0 < |y - x| < \delta$ entonces

$$|h(y) - h(x) - h'(x)(y - x)| < \varepsilon |y - x|.$$

Entonces, si $x < y < x + \delta$ se tiene que existe $r_y > 0$ tal que $y + r_y < x + \delta$, $y + r_y \in [a, b] \setminus D$, $|h'(x)r_y| < \varepsilon|y - x|$ y $r_y < \varepsilon|x - y|$, y por tanto

$$\begin{aligned} &f(y) - f(x) - h'(x)(y - x) \leq f(y + r_y) - f(x) - h'(x)(y - x) \\ &= h(y + r_y) - h(x) - h'(x)(y + r_y - x) + h'(x)r_y < \varepsilon|y + r_y - x| + \varepsilon|y - x| < 3\varepsilon|y - x|, \end{aligned}$$

es decir $x < y < x + \delta \implies |f(y) - f(x) - h'(x)(y - x)| < 3\varepsilon|y - x|$. Análogamente se ve que si $x - \delta < y < x$ entonces

$$|f(y) - f(x) - h'(x)(y - x)| < 3\varepsilon|y - x|.$$

Esto prueba que f es derivable en x, con $f'(x) = h'(x)$ en todos los puntos x donde h es derivable.

Por tanto, en lo que sigue, cambiando si fuera preciso f por h, podemos suponer que f es continua por la izquierda. Veamos que $f(E)$ es boreliano para todo $E \subset [a, b]$ boreliano.[1] En efecto, sea $\mathcal{A} = \{E \subset [a, b] : f(E) \text{ es boreliano}\}$. Para cada $a \leq s < t \leq b$, por ser f estrictamente creciente, se tiene

$$f((s,t)) = (f(s), f(t)) \setminus \bigcup_{x \in D} \left(f(x^-), f(x^+)\right], \tag{11.18}$$

que es de Borel por ser un intervalo menos una unión numerable de intervalos. Luego $\mathcal{A}$ contiene a los intervalos abiertos de $\mathbb{R}$ intersecados con $[a, b]$, y en particular al propio intervalo $[a, b]$. Por otro lado, si $A \in \mathcal{A}$, entonces $f(A)$ es de Borel, luego, por ser f inyectiva, $f([a, b] \setminus A) = [f(a), f(b)] \setminus f(A)$ es también de Borel. Además, si $(A_i)_{i \in \mathbb{N}} \subset \mathcal{A}$, entonces $f(A_i)$ es de Borel para cada $i \in \mathbb{N}$, y por tanto $f\left(\bigcup_{i=1}^{\infty} A_i\right) = \bigcup_{i=1}^{\infty} f(A_i)$ es de Borel. Es decir, $\mathcal{A}$ es una σ-álgebra que contiene a los intervalos, y por tanto $\mathcal{B}([a, b]) \subset \mathcal{A}$; es decir, $f(E)$ es boreliano siempre que E lo sea.

Podemos entonces definir $\lambda : \mathcal{B}([a, b]) \to \mathbb{R}$ por

$$\lambda(E) = |f(E)| + \sum_{x \in D \cap E} (f(x^+) - f(x^-))$$

y, usando nuevamente que f es inyectiva en combinación con (11.18), es inmediato comprobar que λ es una medida positiva tal que

$$\lambda([a, x)) = f(x) - f(a) \text{ para todo } x \in [a, b], \tag{11.19}$$

y en particular $\lambda([a, b]) = f(b) - f(a)$, luego λ es finita. Podemos extender λ a $\mathbb{R}$ poniendo

$$\lambda(E) = \lambda(E \cap [a, b])$$

para cualquier $E \subset \mathbb{R}$ boreliano. Entonces, aplicando el Corolario 11.11, obtenemos que existe $g \in L^1(\mathbb{R})$ tal que

$$\lim_{I \in \mathcal{F}_x,\ |I| \to 0} \frac{1}{\mu(I)} \lambda(I) = g(x) \quad \text{para } \mu\text{-casi todo } x \in \mathbb{R},$$

donde μ denota la medida de Lebesgue restringida a los borelianos de $\mathbb{R}$ y $\mathcal{F}_x$ es la familia de los intervalos I (de cualquier tipo) cuyas adherencias contienen a x (es obvio que esta familia es regular en x). Usando (11.19) concluimos entonces que

$$\lim_{y \to x^+} \frac{f(y) - f(x)}{y - x} = \lim_{y \to x^+} \frac{\lambda([a, y)) - \lambda([a, x))}{\mu([x, y))} = \lim_{y \to x^+} \frac{\lambda([x, y))}{\mu([x, y))} = g(x)$$

para μ-casi todo $x \in [a, b]$. Análogamente, para μ-casi todo $x \in [a, b]$,

$$\lim_{y \to x^-} \frac{f(y) - f(x)}{y - x} = \lim_{y \to x^-} \frac{\lambda([a, x)) - \lambda([a, y))}{\mu([y, x))} = \lim_{y \to x^-} \frac{\lambda([y, x))}{\mu([y, x))} = g(x).$$

[1] Sin embargo, no es en general cierto que $f(E)$ sea medible Lebesgue si E lo es. Adviértase lo delicado de la situación: f podría ser un homeomorfismo que transforma el conjunto ternario de Cantor en un conjunto de Cantor gordo, y por tanto podría llevar un conjunto de medida cero (pero no boreliano) en un conjunto no medible Lebesgue.

Por tanto, existe
$$\lim_{y\to x} \frac{f(y)-f(x)}{y-x} = g(x)$$
para μ-casi todo $x \in [a,b]$. □

Corolario 11.17. *Si $F : [a,b] \to \mathbb{R}$ es de variación acotada entonces F es diferenciable en casi todo punto.*

Demostración. Como por la Observación 11.5 toda función de variación acotada es diferencia de dos crecientes, este resultado es consecuencia inmediata del teorema anterior. □

Una función f creciente no tiene por qué satisfacer $\int_a^b f'(t)dt = f(b) - f(a)$. Por ejemplo, si f es la función *escalera de Cantor*, entonces $f : [0,1] \to \mathbb{R}$ es creciente y continua, pero $f'(t) = 0$ para casi todo t, y por tanto $\int_0^1 f'(t)dt = 0 < 1 = f(1) - f(0)$. No obstante, la siguiente desigualdad es siempre cierta.

Teorema 11.18. *Si $f : [a,b] \to \mathbb{R}$ es creciente, entonces*
$$\int_a^b f'(t)dt \leq f(b) - f(a).$$

Demostración. Como en la demostración del Teorema 11.16,podemos redefinir f en sus puntos de discontinuidad para que sea continua por la izquierda; esto no afecta en nada a la conclusión del enunciado. También, considerando $g(x) = x + f(x)$, puesto que la identidad es absolutamente continua y diferenciable en todo punto, la desigualdad del enunciado se tendrá si y sólo si esta función g la verifica. Por tanto, podemos suponer sin pérdida de generalidad que f es estrictamente creciente, y en particular inyectiva. En esta situación y en la misma demostración del Teorema 11.16 se ha visto que existe λ, medida de Borel finita en $\mathbb{R}$, tal que
$$f(x) - f(a) = \lambda\left([a,x)\right)$$
para cualquier $x \in [a,b]$. Por el teorema de descomposición de Lebesgue,
$$\lambda = \lambda_a + \lambda_s,$$
donde λ_a, λ_s son medidas de Borel finitas mutuamente singulares, y $\lambda_a \ll \mu$, siendo μ la restricción de la medida de Lebesgue en $\mathbb{R}$ a los borelianos de $\mathbb{R}$. Por el Corolario 11.11 se tiene
$$f'(x) = \lim_{y\to x^+} \frac{\lambda\left([x,y)\right)}{\mu\left([x,y)\right)} = g(x)$$
para casi todo x, donde
$$g = \frac{d\lambda_a}{d\mu}.$$
Entonces
$$f(b) - f(a) = \lambda\left([a,b)\right) = \lambda_a\left([a,b)\right) + \lambda_s\left([a,b)\right) \geq \lambda_a\left([a,b)\right) = \int_a^b g(t)dt = \int_a^b f'(t)dt.$$
□

Terminamos este breve estudio de las funciones crecientes con un resultado que resulta muy útil cuando se consideran sumas infinitas de funciones de este tipo.

Teorema 11.19 (Fubini)**.** *Si $f_n : [a,b] \to \mathbb{R}$ es creciente para cada $n \in \mathbb{N}$ y $\sum_{n=1}^\infty f_n(x)$ converge puntualmente a una función $f(x)$, entonces*
$$f'(x) = \sum_{n=1}^\infty f_n'(x)$$
para casi todo $x \in [a,b]$.

Demostración. Escribamos

$$f(x) = \sum_{k=1}^{n} f_k(x) + R_n(x),$$

donde

$$R_n(x) = \sum_{k=n+1}^{\infty} f_k(x).$$

Como f_n, f y R_n son crecientes (por ser sumas de funciones crecientes), son diferenciables en casi todo punto, con derivadas mayores o iguales que 0, y se tiene

$$f'(x) = \sum_{k=1}^{n} f_k'(x) + R_n'(x) \geq \sum_{k=1}^{n} f_k'(x)$$

para casi todo x. Haciendo $n \to \infty$ en la desigualdad anterior obtenemos

$$f'(x) - \sum_{k=1}^{\infty} f_k'(x) \geq 0 \tag{11.20}$$

para casi todo x. Por otro lado, aplicando a R_n el teorema anterior, tenemos

$$0 \leq \int_a^b R_n'(x)dx \leq R_n(b) - R_n(a) = \sum_{k=n+1}^{\infty} (f_k(b) - f_k(a)) \to 0,$$

cuando $n \to \infty$, es decir

$$\lim_{n\to\infty} \int_a^b R_n'(x)dx = 0.$$

Luego, combinando este hecho con

$$\int_a^b f'(x)dx = \int_a^b \sum_{k=1}^{n} f_k'(x)dx + \int_a^b R_n'(x)dx$$
$$\leq \int_a^b \sum_{k=1}^{\infty} f_k'(x)dx + \int_a^b R_n'(x)dx,$$

deducimos que

$$\int_a^b f'(x)dx \leq \int_a^b \sum_{k=1}^{\infty} f_k'(x)dx,$$

es decir

$$\int_a^b \left(f'(x) - \sum_{k=1}^{\infty} f_k'(x) \right) dx \leq 0,$$

y en vista de (11.20) la última desigualdad sólo puede ser cierta si

$$f'(x) - \sum_{k=1}^{\infty} f_k'(x) = 0$$

para casi todo x, que es lo que queríamos probar. □

11.7. Diferenciación de funciones Lipschitz. El teorema de Rademacher

En esta sección probaremos que las funciones Lipschitz de $\mathbb{R}^n$ en $\mathbb{R}^m$ son diferenciables en casi todo punto de $\mathbb{R}^n$.

Recordemos que una función $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ se dice que es de Lipschitz si existe $C \geq 0$ tal que

$$|f(x) - f(y)| \leq C|x - y|$$

para todos $x, y \in A$. También se dice que f es C-Lipschitz. Al ínfimo de tales C se le llama la constante de Lipschitz, y se denota $\text{Lip}(f)$. Visto de otra forma,

$$\text{Lip}(f) = \sup_{x,y \in A,\ x \neq y} \frac{|f(x) - f(y)|}{|x - y|}.$$

Se dice que f es localmente lipschitziana si la anterior condición se cumple localmente, es decir, para todo $z \in A$ existen $\delta > 0$ y $C \geq 0$ tales que $|f(x) - f(y)| \leq |x - y|$ para todos $x, y \in B(z, \delta) \cap A$.

La fórmula de McShane nos permite suponer cuando resulte oportuno que el dominio de una función Lipschitz $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ es todo $\mathbb{R}^n$.

Teorema 11.20 (Fórmula de McShane)**.** *Sean X un espacio métrico, y $f : A \subset X \to \mathbb{R}$ una función C-Lipschitz: Entonces la función $F : X \to \mathbb{R}$ definida por*

$$F(x) = \inf_{a \in A} \{f(a) + Cd(x, a)\}, \ \ x \in X,$$

es C-Lipschitz y extiende a f, es decir, $F(a) = f(a)$ para todo $a \in A$.

La prueba de este resultado es fácil y se deja como ejercicio. Aplicando el resultado anterior a las funciones coordenadas de una función $f : A \subset X \to \mathbb{R}^m$ se obtiene la siguiente proposición, cuya demostración también se deja como ejercicio para el lector.

Proposición 11.21. *Si X es un espacio métrico y $f : A \subset X \to \mathbb{R}^m$ es C-Lipschitz, entonces existe una función $\sqrt{m}C$-Lipschitz $F : X \to \mathbb{R}^m$ que extiende a f.*

El teorema de Kirszbraun (que es más difícil de demostrar y no necesitaremos aquí) afirma que si A es un subconjunto cualquiera de un espacio de Hilbert X y $f : A \to Y$ es de Lipschitz, donde Y es otro espacio de Hilbert, entonces existe una función de Lipschitz $F : X \to Y$ que extiende a f y satisface $\text{Lip}(F) = \text{Lip}(f)$.

Teorema 11.22 (de Rademacher)**.** *Si $f : \mathbb{R}^n \to \mathbb{R}^m$ es localmente Lipschitz, entonces f es diferenciable para casi todo $x \in \mathbb{R}^n$.*

Demostración. Podemos suponer $m = 1$. Para cada $v \in \mathbb{S}^{n-1}$, consideremos la derivada direccional

$$D_v f(x) = \frac{d}{dt} f(x + tv)|_{t=0}$$

siempre que exista.

Lema 11.23. *Para cada $v \in \mathbb{S}^{n-1}$, la derivada $D_v f(x)$ existe para casi todo $x \in \mathbb{R}^n$.*

Demostración. Al ser la diferenciabilidad una propiedad local, podemos suponer que f es Lipschitz en $\mathbb{R}^n$. Como f es continua,

$$\overline{D}_v f(x) := \limsup_{t \to 0} \frac{f(x + tv) - f(x)}{t} \lim_{k \to \infty} \sup_{0 < |t| < \frac{1}{k},\ t \in \mathbb{Q}} \frac{f(x + tv) - f(x)}{t}$$

es medible Borel, y análogamente

$$\underline{D}_v f(x) := \liminf_{t \to 0} \frac{f(x + tv) - f(x)}{t}$$

es medible Borel. Por tanto

$$A_v := \{x \in \mathbb{R}^n : D_v f(x) \text{ no existe}\} = \{x \in \mathbb{R}^n : \underline{D}_v f(x) < \overline{D}_v f(x)\}$$

es medible Borel, y la función $\mathbb{R}^n \setminus A_v \ni x \mapsto D_v f(x)$ es medible Borel. La función $t \mapsto f(x+tv)$ es absolutamente continua (por ser Lipschitz), y por tanto es diferenciable en casi todo punto. Luego la intersección de A_v con cada línea paralela a v tiene medida de Lebesgue 1 dimensional igual a cero, y por el teorema de Fubini (usado en un sistema de coordenadas paralelo a v; recuérdese que la medida de Lebesgue es invariante por traslaciones y rotaciones, así que no hay problema en hacer esto) se concluye que $\mathcal{L}_n(A_v) = 0$. □

En particular, las derivadas parciales de f existen en casi todo punto de $\mathbb{R}^n$, y por tanto el vector gradiente $\nabla f(x) = \left(\frac{\partial f}{\partial x_1}(x), ..., \frac{\partial f}{\partial x_1}(x)\right)$ existe para casi todo $x \in \mathbb{R}^n$.

La demostración del siguiente lema se deja como ejercicio.

Lema 11.24. *Si $g \in L^1_{loc}(\mathbb{R}^n)$ y $\int_{\mathbb{R}^n} g(x)\varphi(x)dx = 0$ para toda $\varphi \in C^\infty_c(\mathbb{R}^n)$, entonces $g = 0$ en casi todo punto.*

Lema 11.25. *Para todo $v \in \mathcal{S}^{n-1}$ se tiene $D_v f(x) = \langle \nabla f(x), v\rangle$ para casi todo $x \in \mathbb{R}^n$.*

Demostración. Por la invariancia de la integral respecto de traslaciones, para todo $h \in \mathbb{R}$ tenemos

$$\int_{\mathbb{R}^n} f(x+hv)\varphi(x)dx = \int_{\mathbb{R}^n} f(x)\varphi(x-hv)dx,$$

luego también

$$\int_{\mathbb{R}^n} \frac{f(x+h_k v) - f(x)}{h_k}\varphi(x)dx = -\int_{\mathbb{R}^n} \frac{\varphi(x-h_k v) - \varphi(x)}{-h_k} f(x)dx$$

para toda sucesión $(h_k) \subset \mathbb{R} \setminus \{0\}$ con $h_k \to 0$. Aplicando el teorema de la convergencia dominada obtenemos entonces

$$\int_{\mathbb{R}^n} \varphi(x) D_v f(x)dx = -\int_{\mathbb{R}^n} f(x) D_v\varphi(x)dx.$$

En particular

$$\int_{\mathbb{R}^n} \varphi(x)\frac{\partial f}{\partial x_i}(x)dx = -\int_{\mathbb{R}^n} f(x)\frac{\partial \varphi}{\partial x_i}(x)dx$$

para todo $i = 1, ..., n$. Entonces

$$\begin{aligned}
&\int_{\mathbb{R}^n} \varphi(x) D_v f(x)dx = -\int_{\mathbb{R}^n} f(x) D_v\varphi(x)dx = -\int_{\mathbb{R}^n} f(x)\langle\nabla\varphi(x), v\rangle dx \\
&-\sum_{i=1}^n \int_{\mathbb{R}^n} f(x)\frac{\partial\varphi}{\partial x_i}(x)v_i dx = \sum_{i=1}^n \int_{\mathbb{R}^n} \varphi(x)\frac{\partial f}{\partial x_i}(x)v_i dx \\
&= \int_{\mathbb{R}^n} \varphi(x)\langle\nabla f(x), v\rangle dx
\end{aligned}$$

para toda $\varphi \in C^\infty_c(\mathbb{R}^n)$, y por el lema anterior se deduce que $D_v f(x) = \langle\nabla f(x), v\rangle$ para casi todo $x \in \mathbb{R}^n$. □

Ya estamos preparados para terminar la demostración del teorema. Sea $\{v_k\}_{k\in\mathbb{N}}$ un subconjunto denso numerable de $\mathcal{S}^{n-1}$, definamos

$$B_k := \{x \in \mathbb{R}^n : \nabla f(x),\ D_{v_k} f(x) \text{ existen y } D_{v_k} f(x) = \langle\nabla f(x), v_k\rangle\}$$

para cada $k \in \mathbb{N}$, y pongamos

$$B := \bigcap_{k\in\mathbb{N}} B_k.$$

Nótese que B es medible Borel. Como por el primer lema $|\mathbb{R}^n \setminus B_k|$ para todo $k \in \mathbb{N}$, es obvio que $|\mathbb{R}^n \setminus B| = 0$. Por tanto bastará probar que f es diferenciable en todo $x \in B$. Para ello, dados $x \in B$, $v \in \mathcal{S}^{n-1}$, $h > 0$, definimos

$$Q(x, v, h) = \frac{f(x + hv) - f(x)}{h} - \langle \nabla f(x), v \rangle,$$

y tenemos que demostrar que para todo $\varepsilon > 0$ existe $\delta > 0$ tal que

$$|Q(x, v, h)| < \varepsilon \text{ para todo } h \in (0, \delta) \text{ y todo } v \in \mathcal{S}^{n-1}$$

Denotemos $L = \mathrm{Lip}(f)$. Entonces se tiene $|\partial f/\partial x_i| \leq L$ para todo $x \in B$, y por tanto también $|\nabla f(x)| \leq \sqrt{n}L$ para todo $x \in B$. Por tanto para todos $x \in B$, $v, v' \in \mathcal{S}^{n-1}$, $h > 0$, se tiene

$$|Q(x, v, h) - Q(x, v', h)| \leq \left(1 + \sqrt{n}\right) L|v - v'|.$$

Fijado $\varepsilon > 0$, podemos encontrar un $N \in \mathbb{N}$ suficientemente grande tal que para todo $v \in \mathcal{S}^{n-1}$ existe $k = k_v \in \{1, 2, ..., N\}$ tal que

$$|v - v_k| < \frac{\varepsilon}{2\,(1 + \sqrt{n})\,L}.$$

Por otra parte, como $\lim_{h\to 0^+} Q(x, v_i, h) = 0$ para todo $x \in B$ y todo $i \in \mathbb{N}$, existe $\delta > 0$ tal que

$$|Q(x, v_i, h)| < \frac{\varepsilon}{2} \text{ para todo } h \in (0, \delta) \text{ y todo } i \in \{1, 2, ..., N\}.$$

Entonces para todos $h \in (0, \delta)$, $v \in \mathcal{S}^{n-1}$ se cumple

$$|Q(x, v, h)| \leq |Q(x, v_k, h)| + |Q(x, v_k, h) - Q(x, v, h)| < \frac{\varepsilon}{2} + \left(1 + \sqrt{n}\right) L|v_k - v| < \varepsilon,$$

lo que completa la demostración del teorema. □

Observación 11.6. Se ha probado que f es diferenciable en todo punto de B, que es un boreliano cuyo complementario tiene medida cero. Por otro lado, si f es diferenciable en x, entonces es claro, por definición de B, que $x \in B$. Por tanto, la demostración también nos dice que el conjunto de puntos de diferenciabilidad de una función Lipschitz f es un boreliano, y que la función derivada, definida en este conjunto por $x \mapsto Df(x)$ es medible Borel.

Por último presentamos una consecuencia del teorema de Rademacher que usaremos más adelante en la demostración de la fórmula de la coárea.

Teorema 11.26. *Sean $f, g : \mathbb{R}^n \to \mathbb{R}^m$ localmente Lipschitz, y definamos*

$$Z = \{x \in \mathbb{R}^n : f(x) = 0\},$$

y

$$Y = \{x \in \mathbb{R}^n : g(f(x)) = x\}.$$

Entonces:

(1) $Df(x) = 0$ para casi todo $x \in Z$;

(2) $Dg(f(x)) \circ Df(x) = I$ para casi todo $x \in Y$.

Demostración. Para probar (1) podemos suponer $m = 1$. Por el teorema de Rademacher f es diferenciable en casi todo punto, y casi todo punto es un punto de densidad de Z. Así, para casi todo $x \in Z$ se tiene que $Df(x)$ existe y

$$\lim_{r\to 0} \frac{|Z \cap B(x, r)|}{|B(x, r)|} = 1. \tag{11.21}$$

Supongamos que $a := Df(x) \neq 0$, y sea

$$S := \{v \in \mathcal{S}^{n-1} : \langle a, v\rangle \geq \frac{1}{2}|a|.\}$$

Como f es diferenciable en x y $f(x) = 0$, tomando $\varepsilon = |a|/4 >$, existe $\delta > 0$ tal que si $0 < t < \delta$ y $v \in \mathcal{S}^{n-1}$ entonces

$$|f(x+tv) - t\langle a, v\rangle| < \varepsilon = \frac{|a|}{4}t,$$

y se deduce que para todo $v \in S$ y todo $t \in (0, \delta)$ es

$$f(x+tv) \geq t\langle a, v\rangle - \frac{|a|}{4}t \geq \frac{1}{4}|a|t > 0;$$

luego $x + tv \notin Z$ para todo $v \in S$ y todo $t \in (0, \delta)$, lo cual contradice (11.21). Esto prueba (1).

Para demostrar (2), sea A el conjunto de puntos donde f es diferenciable, y B el conjunto de puntos donde g es diferenciable. Sean también $X := Y \cap A \cap f^{-1}(B)$. Se tiene que

$$Y \setminus X \subset (\mathbb{R}^n \setminus A) \cup g(\mathbb{R}^n \setminus B);$$

en efecto, si $x \in Y \setminus f^{-1}(B)$ entonces $f(x) \in \mathbb{R}^n \setminus B$, y por tanto $x = g(f(x)) \in g(\mathbb{R}^n \setminus B)$. Por tanto, como $|\mathbb{R}^n \setminus A| = 0$ y $|\mathbb{R}^n \setminus B| = 0$ (por el teorema de Rademacher) y g es Lipschitz, también es $|g(\mathbb{R}^n \setminus B)| = 0$, y se deduce que

$$|Y \setminus X| = 0.$$

Además, si $x \in X$ entonces $Dg(f(x))$ y $Df(x)$ existen, y por tanto, por la regla de la cadena,

$$D(g \circ f)(x) = Dg(f(x)) \circ Df(x)$$

también existe. Y puesto que $(g\circ f)(x) - x = 0$ en Y, por la propiedad (1) se concluye que $D(g\circ f)(x) = I$ para casi todo $x \in Y$; por tanto, como $|Y \setminus X| = 0$, también es $Dg(f(x)) \circ Df(x) = I$ para casi todo $x \in Y$. □

11.8. Problemas

Problema 11.1. Probar que el valor de $\mathcal{M}(f)(x)$ es el mismo independientemente de si en la definición (11.3) se consideran bolas abiertas o cerradas.

Problema 11.2. Demostrar que si $f, g \in L^1(\mathbb{R}^d)$ entonces $\mathcal{M}(f+g) \leq \mathcal{M}(f) + \mathcal{M}(g)$.

Problema 11.3. Usar el Teorema 11.9 para demostrar que si $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ y x es un punto de Lebesgue de f, entonces, para toda sucesión $(Q_i)_{i\in\mathbb{N}}$ de cubos tales que $x \in Q_i$ y $\text{diam}(Q_i) \to 0$ se tiene $\lim_{i\to\infty} \frac{1}{|Q_i|}\int_{Q_i} f(y)dy = f(x)$.

Problema 11.4. Demostrar que si $\varepsilon > 0$, no existen conjuntos $E \subset \mathbb{R}$ medibles tales que

$$\varepsilon < \frac{|E \cap I|}{|I|} < 1 - \varepsilon$$

para todo intervalo abierto I.

Problema 11.5. Si $\mu \ll \mathcal{L}_d$ y $\lim_{r\to 0^+} r^{-d}\mu(B(x,r)) = 0$ para todo $x \in \mathbb{R}^d$, probar que $\mu = 0$.

Problema 11.6. Sea $K \subset \mathbb{R}^d$ compacto, y definamos $K_r := \{x \in \mathbb{R}^d : d(x, K) = r\}$. Probar que $|K_r| = 0$ para cada $r > 0$. *Indicación:* Considerar los puntos de densidad de Lebesgue de K_r.

Problema 11.7. Demostrar que si f y $\mathcal{M}(f)$ están ambas en $L^1(\mathbb{R}^d)$, entonces $f = 0$ en casi todo punto. *Indicación:* A cualquier otra función f le corresponde una constante $c = c(f) > 0$ tal que $\mathcal{M}(f)(x) \geq c|x|^{-d}$ cuando $|x|$ es suficientemente grande.

Problema 11.8. Si $f(x) = x^{-1}(\log x)^{-2}$, para todo $x \in (0, 1/2)$ y $f(x) = 0$ en otro caso, demostrar que $f \in L^1(\mathbb{R})$ pero $\mathcal{M}(f)(x) \geq |2x \log(2x)|^{-1}$ para todo $x \in (0, 1/4)$. Deducir que $\mathcal{M}(f)$ puede no ser localmente integrable incluso si f es integrable. Sin embargo, ver el problema 11.10.

Problema 11.9. Usar el teorema de Fubini para demostrar que si μ es σ-finita y $p < \infty$ entonces $\int_x |f|^p d\mu = p \int_0^\infty t^{p-1} \mu(\{x : |f(x)| > t\}) dt$.

Problema 11.10. Si $p > 1$ y $f \in L^p(\mathbb{R}^d)$, probar que $\mathcal{M}(f) \in L^p(\mathbb{R}^n)$. *Indicación:* El caso $p = \infty$ es muy fácil. Para $1 < p < \infty$, para cada $t > 0$ considerar $f = f_1 + f_2$ con $f_1 = f\chi_{\{|f|>t/2\}}$ y $f_2 = f\chi_{\{|f|\leq t/2\}}$. Observar que $\mathcal{M}(f) \leq \mathcal{M}(f_1) + t/2$, luego $\{\mathcal{M}(f) > t\} \subset \{\mathcal{M}(f_1) > t/2\}$. Aplicar el Teorema 12.3(3) a f_1 y luego estimar $\int_{\mathbb{R}^d} \mathcal{M}(f)^p$ usando el problema 11.9.

Problema 11.11. Si $f, g : [a, b] \to \mathbb{R}$ son absolutamente continuas, demostrar que $f + g$ y fg también lo son. Deducir que el teorema de integración por partes vale para las funciones absolutamente continuas:

$$\int_a^b fg' = fg|_a^b - \int_a^b f'g.$$

Problema 11.12. Demostrar que si $f : [a, b] \to \mathbb{R}$ es absolutamente continua y $f'(x) = 0$ en casi todo punto entonces f es constante.

Problema 11.13. Probar que toda función de Lipschitz $f : [a, b] \to \mathbb{R}$ es absolutamente continua, con $f' \in L^\infty([a, b])$, y $\|f'\|_\infty = \mathrm{Lip}(f)$.

Problema 11.14. Si $E \subset [a, b]$, $|E| = 0$, construir una función absolutamente continua $f : [a, b] \to \mathbb{R}$ tal que $f'(x) = \infty$ para todo $x \in E$. *Indicación:* $E \subset \bigcap_{n=1}^\infty V_n$, donde los V_n son abiertos con $|V_n| < 2^{-n}$. Considerar $f(x) = \int_a^x \sum_{n\in\mathbb{N}} \chi_{V_n}(t) dt$.

Problema 11.15. Demostrar que la *escalera de Cantor* no es absolutamente continua, aunque sí de variación acotada en $[0, 1]$.

Problema 11.16. Si $\{K_\delta\}_{\delta>0}$ es un núcleo de sumabilidad tal que $\lim_{\delta\to0^+} \|K_\delta \chi_{\mathbb{R}^d\setminus B(0,r)}\|_\infty = 0$ para cada $r > 0$, y $f \in L^1(\mathbb{R}^d)$, demostrar que $\lim_{\delta\to0^+} f * K_\delta(x) = f(x)$ para cada punto de Lebesgue x de f (y en particular para casi todo x).

Problema 11.17. Sea E un subconjunto cerrado de $\mathbb{R}^d$. Demostrar que $\mathrm{dist}(x + y, E) = o(|y|)$ para casi todo $x \in E$. *Indicación:* suponer que x es un punto de densidad de E.

Problema 11.18. Construir una función creciente $f : \mathbb{R} \to \mathbb{R}$ cuyo conjunto de discontinuidades sea precisamente $\mathbb{Q}$.

Problema 11.19. Construir una función estrictamente creciente $f : [0, 1] \to \mathbb{R}$ tal que $f'(x) = 0$ en casi todo punto. *Indicación:* partiendo de la función escalera de Cantor $g : [0, 1] \to [0, 1]$, extendámosla a $\mathbb{R}$ poniendo $g = 1$ en $(-\infty, 0]$ y $g = 1$ en $[1, \infty)$. Si $[a_n, b_n]$ es la sucesión de los intervalos diádicos $[0, 1]$, $[0, 1/2]$, $[1/2, 1]$, $[0, 1/4]$, $[1/4, 1/2]$, ..., $[0, 1/8]$, ..., definir

$$f(x) = \sum_{n=1}^\infty \frac{1}{2^n} g\left(\frac{x - a_n}{b_b - a_n}\right)$$

para todo $x \in [0, 1]$.

Problema 11.20. Sea $f(x) = x^2 \operatorname{sen}(1/x^2)$ si $x \neq 0$, $f(0) = 0$. Demostrar que f es derivable en todo punto pero f' no es integrable en $[-1, 1]$. En particular f no es absolutamente continua.

Problema 11.21. Demostrar que si $F : [a, b] \to \mathbb{R}$ es absolutamente continua entonces lleva conjuntos de medida cero en conjuntos de medida cero.

Problema 11.22. Demostrar que si $f : [a, b] \to \mathbb{R}$ es absolutamente continua entonces lleva conjuntos medibles Lebesgue en conjuntos medibles Lebesgue.

Problema 11.23. Demostrar que si $f : [a, b] \to \mathbb{R}$ es absolutamente continua y $g : \mathbb{R} \to \mathbb{R}$ es Lipschitz, entonces $g \circ f : [a, b] \to \mathbb{R}$ es absolutamente continua.

Problema 11.24. Sea K el complementario en $[0, 1]$ de un conjunto de Cantor gordo. Probar que la función $f(x) = \int_0^x \chi_K(y)dy$ es estrictamente creciente, absolutamente continua, y cumple que $f'(x) = 0$ en un conjunto de medida positiva.

Problema 11.25. Demostrar el lema 11.24.

Problema 11.26. Demostrar el teorema de Stepanov: sean $\Omega \subset \mathbb{R}^n$ abierto y $f : \Omega \to \mathbb{R}$ medible. Entonces f es diferenciable en casi todo punto si y sólo si

$$\limsup_{y \to x} \frac{|f(y) - f(x)|}{|y - x|} < \infty$$

para casi todo $x \in \Omega$.
Indicación: encontrar un conjunto A grande en sentido de medida donde f sea Lipschitz, y extender $f|_A$ a $\mathbb{R}^n$ aplicando la fórmula de McShane; entonces aplicar el teorema de Rademacher a dicha extensión F_A, y demostrar que en todos los puntos de densidad de A donde F_A es diferenciable la función original f también lo es.

Problema 11.27. Sean U un abierto de $\mathbb{R}^n$, y $f : U \to \mathbb{R}$ una función convexa. Demostrar que f es localmente Lipschitz. Deducir que f es diferenciable en casi todo punto.[2]

Problema 11.28. Si $E \subset \mathbb{R}^n$ es cerrado y no vacío, definamos $U(E) = \{x \in \mathbb{R}^n : \exists! y \in E\, |x - y| = d(x, E)\}$, es decir el conjunto de puntos de $\mathbb{R}^n$ cuya distacia a E se alcanza en un único punto. Para cada $x \in U(E)$, denotemos $\pi(x)$ como el único punto de E donde se alcanza $d(x, E)$. A la aplicación $\pi : U(E) \to E$ se le llama la *proyección métrica* sobre E. Demostrar que $\mathbb{R}^n \setminus U(E)| = 0$, es decir, π está definida en casi todo punto.
Indicación: como la función $d(\cdot, E)$ es 1-Lipschitz, por el Teorema de Rademacher es diferenciable en casi todo punto. Probar que si $d(\cdot, E)$ es diferenciable en $x \in \mathbb{R}^n \setminus E$ y $p_x \in E$ es tal que $d(x, p_x) = d(x, E)$, entonces $p = d(x, E)\nabla d(\cdot, E)(x)$.

Problema 11.29. Sean $f : [0, 1] \to \mathbb{R}$ continua, y $D = \{x \in [0, 1] : \exists f'(x) \in \mathbb{R}\}$. Demostrar que si $|E| = 0$ entonces $|f(E \cap D)| = 0$. Luego, suponiendo que $f_{|_D}$ es además inyectiva, deducir que existe $g : [0, 1] \to [0, \infty)$ integrable tal que $|f(A \cap D)| = \int_A g(x)dx$ para todo $A \subset \mathbb{R}$ medible Lebesgue.

[2]Hay un resultado mucho más fuerte, debido a Alexandroff, que dice que f es dos veces diferenciable en casi todo punto (en el sentido de que tiene desarrollos de Taylor de orden 2 en casi todo punto).

Capítulo 12

Desigualdades de Brunn-Minkowski e isodiamétrica. $\mathcal{L}_n = \mathcal{H}^n$

En este capítulo veremos algunas desigualdades para la medida de Lebesgue que, aparte de ser importantes y útiles de por sí, nos permitirán demostrar que la medida de Hausdorff n-dimensional en $\mathbb{R}^n$ coincide con la de Lebesgue.

12.1. La desigualdad de Brunn-Minkowski

Ya sabemos que la medida de Lebesgue en $\mathbb{R}^n$ (y también la de Hausdorff) es invariante por por isometrías; en particular lo es por traslaciones. Ahora vamos a estudiar lo que ocurre con la medida de la suma de dos conjuntos medibles $A, B \subset \mathbb{R}^n$, definida por

$$A + B = \{x \in \mathbb{R}^d : x = a + b, a \in A, b \in B\}.$$

Lo primero que debe observarse es que $A + B$ no tiene por qué ser medible aunque A y B lo sean. Por ejemplo, en $\mathbb{R}^2$, si $A = \{0\} \times [0, 1]$ y $B = E \times \{0\}$, donde $E \subset [0, 1]$ no es medible en $\mathbb{R}$, entonces $A + B = \{(x, y) \in \mathbb{R}^2 : x \in E, y \in [0, 1]\} = E \times [0, 1]$ no es medible (porque, como veremos más adelante, si lo fuera, por el teorema de Fubini casi todas sus secciones horizontales serían medibles, y ello significaría que E sería medible). Sin embargo, este tipo de situaciones no se dan si A y B son cerrados, o si uno de ellos es abierto. Dicho esto, en las situaciones en que $A+B$ es medible, es natural preguntarse si hay alguna estimación general de la medida de Lebesgue de $A + B$ en terminos de las medidas de A y de B. Es imposible dar una cota superior de $|A + B|$ en esos términos, ya que puede suceder que $|A| = 0 = |B|$ y sin embargo $|A + B| > 0$ (por ejemplo, en $\mathbb{R}^2$, si $A = [0, 1] \times \{0\}$ y $B = \{0\} \times [0, 1]$ entonces $A \times B = [0, 1] \times [0, 1]$). Pero sí que cabe la posibilidad de tener una desigualdad del tipo

$$|A + B|^\alpha \geq c_\alpha \left(|A|^\alpha + |B|^\alpha\right),$$

donde $\alpha > 0$ y c_α es una constante independiente de A y B. Obviamente lo mejor que puede esperarse es $c_\alpha = 1$. Además, si A es convexo y tomamos $B = \lambda A$, entonces $A + B = (1 + \lambda)A$. Puesto que $|(1+\lambda)A| = (1+\lambda)^d|A$, el tipo de desigualdad que buscamos sólo puede darse si $(1+\lambda)^{n\alpha} \geq 1+\lambda^{n\alpha}$ para todo $\lambda > 0$. Además, sabemos que es

$$(a + b)^\gamma \geq a^\gamma + b^\gamma \tag{12.1}$$

para todo $\gamma \geq 1$ y $a, b \geq 0$, mientras que la desigualdad opuesta es verdad cuando $0 \leq \gamma \leq 1$. Por tanto lo que buscamos sólo puede cumplirse si $\alpha \geq 1/n$. Además, si se cumple para $\alpha = 1/n$ entonces como consecuencia de (12.1) también se cumple para $\alpha \geq 1/n$. Así que lo mejor que podemos esperar en este sentido es tener

$$|A + B|^{1/n} \geq |A|^{1/n} + |B|^{1/n}$$

para todos $A, B \subset \mathbb{R}^n$ medibles. Y en efecto, el siguiente teorema nos asegura que esto es verdad.

Teorema 12.1. *Sean $A, B \subset \mathbb{R}^n$ medibles Lebesgue, y supongamos que $A+B$ también es medible Lebesgue. Entonces se tiene*

$$|A+B|^{1/n} \geq |A|^{1/n} + |B|^{1/n}. \tag{12.2}$$

Demostración. Consideremos primero el caso en que A y B son rectángulos con lados paralelos a los ejes, de longitudes $a_1, ..., a_n$, en el caso de A, y $b_1, ..., b_n$, en el caso de B. Entonces $A+B$ es un rectángulo cuyos lados tienen longitudes $a_1+b_1, ..., a_n+b_n$, y (12.2) equivale a

$$\left(\prod_{j=1}^{n}(a_j+b_j)\right)^{1/n} \geq \left(\prod_{j=1}^{n} a_j\right)^{1/n} + \left(\prod_{j=1}^{n} b_j\right)^{1/n}. \tag{12.3}$$

Esta desigualdad es una consecuencia fácil de la desigualdad aritmética-geométrica

$$\frac{1}{n}\sum_{j=1}^{n} x_j \geq \left(\prod_{j=1}^{n} x_j\right)^{1/n}$$

para todos $x_1, ..., x_n \geq 0$: tomando primero $x_j = a_j/(a_j+b_j)$ y aplicando esta desigualdad, y luego sumando al resultado la conclusión de aplicarla con $x_j = b_j/(a_j+b_j)$ obtenemos

$$\left(\prod_{j=1}^{n}\frac{a_j}{a_j+b_j}\right)^{1/n} + \left(\prod_{j=1}^{n}\frac{b_j}{a_j+b_j}\right)^{1/n} \leq \frac{1}{n}\sum_{j=1}^{n}\frac{a_j}{a_j+b_j} + \frac{1}{n}\sum_{j=1}^{n}\frac{b_j}{a_j+b_j} = 1;$$

entonces multiplicando ambos miembros por $\left(\prod_{j=1}^{n}(a_j+b_j)\right)^{1/n}$ obtenemos (12.3).

Ahora consideremos el caso en que A y B son uniones finitas de rectángulos con interiores disjuntos. Probaremos (12.2) en este caso por inducción sobre el número total de rectángulos que componen A y B. Denotemos este número por N. Para $N=2$, el resultado es verdad porque es justo lo que acabamos de probar. Supongamos que la desigualdad es verdad para esta clase de conjuntos cuando $N \geq 2$, y veamos que entonces también lo es para $N+1$.

Es importante observar que la desigualdad que queremos demostrar no cambia si trasladamos A y B independientemente (en efecto, si cambiamos A por $h+A$ y B por $h'+B$ entonces lo que estamos haciendo es cambiar $A+B$ por $(h+h')+(A+B)$, y como la medida de Lebesgue es invariante por traslaciones, todos estos cambios dejan la desigualdad invariante). Como el número total de rectángulos de A y B es $N+1 \geq 3$, alguno de los conjuntos A o B tiene por lo menos dos rectángulos; podemos suponer sin pérdida de generalidad que A tiene al menos dos rectángulos. Elijamos pues R_1, R_2 rectángulos con interiores disjuntos en la colección de rectángulos que forman A, y observemos que pueden separarse por un hiperplano $x_j = a_j$ definido por alguna de las coordenadas. Después de aplicar una traslación por un vector adecuado h, esto significa que $R_1 \subset A_- := A \cap \{x_j \leq 0\}$, y $R_2 \subset A_+ := A \cap \{x_j \geq 0\}$. Observemos también que tanto A_- como A_+ tienen por lo menos un rectángulo menos de los que tiene A, y que $A = A_- \cup A_+$.

Sea

$$\lambda = |A_+|/|A|.$$

Como la función $t \mapsto |B \cap \{x_k \leq t\}$ es continua, toma todos los valores entre 0 y $|B|$, luego podemos trasladar B de manera que $B_+ := B \cap \{x_j \geq 0\}$ satisfaga

$$|B_+|/|B| = \lambda,$$

y por tanto también, definiendo $B_- = B \cap \{x_j \leq 0\}$, se tiene

$$\frac{|B_-|}{|B|} = (1-\lambda) = \frac{|A_-|}{|A|}.$$

Además
$$A+B\supset (A_+ + B_+)\cup(A_- + B_-),$$
y la unión de la derecha es esencialmente disjunta, puesto que $A_-+B_-\subset\{x_j\leq 0\}$ y $A_++B_+\subset\{x_j\geq 0\}$. Como el número toral de rectángulos en $A_+\cup B_+$ es a lo sumo N (porque al menos un rectángulo de A está en A^-), y lo mismo es cierto para $A_-\cup B_-$, podemos aplicar la hipótesis de inducción y concluir que

$$\begin{aligned}|A+B| &\geq |A_++B_+|+|A_-+B_-|\\ &\geq \left(|A_+|^{1/n}+|B_+|^{1/n}\right)^n+\left(|A_-|^{1/n}+|B_-|^{1/n}\right)^n\\ &=\lambda\left(|A|^{1/n}+|B|^{1/n}\right)^n+(1-\lambda)\left(|A|^{1/n}+|B|^{1/n}\right)^n\\ &=\left(|A|^{1/n}+|B|^{1/n}\right),\end{aligned}$$

lo que prueba (12.2) para uniones finitas de rectángulos con interiores disjuntos.

Supongamos ahora que A y B son conjuntos abiertos con medida finita; entonces para cada $\varepsilon>0$ existen $A_\varepsilon, B_\varepsilon$, uniones finitas de cubos con interiores disjuntos, tales que $A_\varepsilon\subset A$, $B_\varepsilon\subset B$, y $|A|\leq |A_\varepsilon|+\varepsilon$, $|B|\leq|B_\varepsilon|+\varepsilon$. Como $A_\varepsilon+B_\varepsilon\subset A+B$, aplicando lo demostrado en el caso anterior a A_ε y B_ε y haciendo después $\varepsilon\to 0$, obtenemos (12.2) para abiertos de medida finita.

El penúltimo caso es el de dos conjuntos compactos A, B. Obsérvese entonces que $A+B$ también es compacto, y definiendo $A^\varepsilon=\{x: d(x,A)<\varepsilon\}$ se tiene que A^ε es abierto, y $A^\varepsilon\searrow A$ cuando $\varepsilon\searrow 0$. Definiendo B^ε y $(A+B)^\varepsilon$ de forma análoga, se tiene que $A+B\subset A^\varepsilon+B^\varepsilon\subset(A+B)^{2\varepsilon}$. Entonces, haciendo $\varepsilon\to 0$ en (12.2) aplicada a $A^\varepsilon, B^\varepsilon$ se obtiene (12.2) para A, B compactos.

Por último, el caso general de dos conjuntos medibles A, B cualesquiera, se deduce del caso de dos compactos usando el hecho de que $|E|=\sup\{|K|: K\subset E,\ K \text{ compacto}\}$ para todo $E\subset\mathbb{R}^n$ medible. Los detalles de los dos últimos casos se dejan como ejercicios. □

12.2. La desigualdad isodiamétrica

Observemos que si B es una bola (euclídea) en $\mathbb{R}^n$ entonces
$$|B|=\frac{\omega_n}{2^n}\mathrm{diam}(E)^n.$$
La desigualdad isodiamétrica, que deduciremos del Teorema 12.1, nos dice que si E es cualquier compacto de $\mathbb{R}^n$ entonces su medida de Lebesgue es menor o igual que la de una bola con el mismo diámetro que E. Por tanto, las bolas son los conjuntos de $\mathbb{R}^n$ que, a igual diámetro, maximizan la medida.

Teorema 12.2. *Para todo compacto $E\subset\mathbb{R}^n$ se tiene que*
$$|E|\leq\frac{\omega_n}{2^n}\,\mathit{diam}(E)^n.$$

Demostración. Vamos a ver que existe un conjunto compacto F, simétrico con respecto a 0 (o sea, $x\in F\iff -x\in F$), tal que

(1) $|F|\geq|E|$, y

(2) $\mathrm{diam}(F)\leq\mathrm{diam}(E)$.

Una vez hecho esto, el resultado se deduce fácilmente: como F es simétrico, para todo $x\in F$ se tiene
$$|x|=\frac{1}{2}|x-(-x)|\leq\frac{1}{2}\mathrm{diam}(F),$$

y por tanto

$$F \subset \overline{B}\left(0, \frac{\operatorname{diam}(F)}{2}\right),$$

con lo que (1) nos da

$$|E| \leq |F| \leq \omega_n \left(\frac{\operatorname{diam}(F)}{2}\right)^n \leq \frac{\omega_n}{2^n} \operatorname{diam}(E)^n,$$

que es lo que queremos probar.

Definamos

$$F := \frac{1}{2}(E - E) := \{\frac{1}{2}(x' - x'') : x', x'' \in E\}.$$

Como E es compacto, es inmediato ver que F también es compacto. También es obvio por la definición que f es simétrico respecto del origen. Al ser F compacto, podemos encontrar $x, y \in F$ tales que $|x - y| = \operatorname{diam}(F)$. Por definición de F, existen $x', x'', y', y'' \in E$ tales que

$$x = \frac{1}{2}(x' - x''), \;\; y = \frac{1}{2}(y' - y''),$$

y por tanto

$$\begin{aligned}
&\operatorname{diam}(F) = |x - y| = \frac{1}{2}|(x' - x'') - (y' - y'')| \\
&\leq \frac{1}{2}\left(|x' - y'| + |y'' - x''|\right) \leq (\operatorname{diam}(E) + \operatorname{diam}(E)) \\
&= \operatorname{diam}(E),
\end{aligned}$$

lo que prueba (2). Para demostrar que F cumple (1) usamos la desigualdad de Brunn-Minkowski:

$$\begin{aligned}
&|F|^{1/n} = \left|\frac{1}{2}E + \frac{1}{2}(-E)\right|^{1/n} \geq \left|\frac{1}{2}E\right|^{1/n} + \left|\frac{1}{2}(-E)\right|^{1/n} \\
&= \frac{1}{2}|E|^{1/n} + \frac{1}{2}|-E|^{1/n} = |E|^{1/n},
\end{aligned}$$

y por tanto $|F| \geq |E|$. Esto completa la demostración del teorema. □

12.3. $\mathcal{L}_n^* = \mathcal{H}_\delta^n = \mathcal{H}^n$ en $\mathbb{R}^n$.

En esta sección probaremos que la medida exterior de Hausdorff de dimensión n coincide con la de Lebesgue en $\mathbb{R}^n$. Para ello, además de la desigualdad isodiamétrica, necesitaremos dos resultados auxiliares.

Lema 12.3. *Sea $U \subset \mathbb{R}^n$ un abierto. Entonces, para todo $\delta > 0$ existe una familia numerable de bolas cerradas disjuntas $B_i \subset U$ tales que diam$(B_i) < \delta$ para todo $i \in \mathbb{N}$, y*

$$|U \setminus \bigcup_{i=1}^{\infty} B_i| = 0.$$

Demostración. Supongamos primero que $|U| < \infty$. Sabemos que U es unión numerable de cubos con interiores disjuntos: $U = \bigcup_{j=1}^{\infty}$, y subdividiendo convenientemente cada uno de esos cubos podemos suponer sin pérdida de generalidad que $\operatorname{diam}(Q_j) < \delta$ para todo $j \in \mathbb{N}$. Existe $\beta = \beta(n) \in (0, 1)$ tal que para todo Q cubo de $\mathbb{R}^n$ existe B bola cerrada tal que $B \subset \operatorname{int}(Q)$ y $|B| \geq \beta|Q|$. Por tanto para todo $k \in \mathbb{N}$ existe una bola cerrada $B_k \subset \operatorname{int}(Q_k)$ tal que $|B_k| \geq \beta|Q_k|$. Puesto que

$$|U \setminus \bigcup_{j=1}^{h} B_j| = |U \setminus \bigcup_{j=1}^{h} Q_j| + \sum_{j=1}^{h} (|Q_j| - |B_j|) \leq |U \setminus \bigcup_{j=1}^{h} Q_j| + (1-\beta) \sum_{j=1}^{h} (|Q_j| - |B_j|) \to (1-\beta)|G|,$$

existe $h_0 \in \mathbb{N}$ suficientemente grande tal que

$$|U \setminus \bigcup_{j=1}^{h_0} B_j| \leq \left(1 - \frac{\beta}{2}\right) |U|.$$

Al conjunto $U_1 = U \setminus \bigcup_{j=1}^{h_0}$, que es abierto con medida finita, se le puede aplicar el mismo hecho que a G, obteniendo $h_1 > h_0$ tal que

$$|U_1 \setminus \bigcup_{j=h_0+1}^{h_1} B_j| \leq \left(1 - \frac{\beta}{2}\right) |U_1|.$$

y por tanto también

$$|U \setminus \bigcup_{j=1}^{h_1} B_j| \leq \left(1 - \frac{\beta}{2}\right)^2 |U|.$$

Reiterando este proceso, se obtiene una sucesión (B_k) de bolas cerradas disjuntas dos a dos tales que $\bigcup_{k=1}^\infty B_k \subset U$, $\mathrm{diam}(B_k) < \delta$ para todo k, y

$$\left| U \setminus \bigcup_{k=1}^{\infty} B_k \right| \leq \left| U \setminus \bigcup_{k=1}^{j} B_k \right| \leq \left(1 - \frac{\beta}{2}\right)^j |U|$$

para todo $j \in \mathbb{N}$, de donde

$$\left| U \setminus \bigcup_{k=1}^{\infty} B_k \right| = 0.$$

Esto prueba el resultado en el caso $|U| < \infty$.

Supongamos ahora que $|U| = \infty$. Nuevamente podemos escribir $U = \bigcup_{i=1}^\infty Q_i$, donde los Q_i son cubos con interiores disjuntos y diámetros menores que δ. Podemos aplicar el caso anterior al interior de cada Q_i, obteniendo una familia de bolas cerradas disjuntas $B_{ij} \subset Q_i$ tales que

$$\left| \mathrm{int}(Q_j) \setminus \bigcup_{j=1}^{\infty} B_{ij} \right| = 0.$$

Puesto que $\left| \bigcup_{j=1}^\infty \partial Q_j \right| = 0$, se obtiene inmediatamente que $\left| U \setminus \bigcup_{i,j \in \mathbb{N}} B_{ij} \right| = 0$. □

Lema 12.4. *Si $Z \subset \mathbb{R}^n$ cumple $\mathcal{L}_n(Z) = 0$, entonces también se tiene $\mathcal{H}^n_\delta = 0$ para todo $\delta \in (0, \infty]$.*

Demostración. Para cada $\varepsilon \in (0, (\delta/2)^n)$, existe una sucesión de bolas (B_k) tal que $Z \subset \bigcup_{k=1}^\infty B_k$, $B_k = B(x_k, r_k)$, $\sum_{k=1}^\infty r_k^n < \varepsilon$. Entonces

$$\frac{\omega_n}{2^n} \sum_{k=1}^{\infty} \mathrm{diam}(B_k)^n = \omega_n \sum_{k=1}^{\infty} r_k^n < \omega_n \varepsilon;$$

además, para cualquier k es

$$\frac{\omega_n}{2^n} \mathrm{diam}(B_k)^n < \omega_n \varepsilon < \omega_n (\delta/2)^n,$$

luego $\mathrm{diam}(B_k) < \delta$, y por definición de $\mathcal{H}^n_\delta$ es inmdeiato que $\mathcal{H}^n_\delta(Z) < \omega_n \varepsilon$, y al ser ε arbitrariamente pequeño, también $\mathcal{H}^n_\delta(Z) = 0$. □

Teorema 12.5. *Si $A \subset \mathbb{R}^n$ y $0 < \delta \leq \infty$, entonces*

$$\mathcal{L}^*_n(A) = \mathcal{H}^n_\delta(A) = \mathcal{H}^n(A).$$

Demostración. Basta probar que $\mathcal{L}_n^*(A) = \mathcal{H}_\delta^n(A)$, porque una vez hecho esto, haciendo $\delta \to 0$, se obtiene también $\mathcal{L}_n^*(A) = \mathcal{H}^n(A)$.

Veamos primero que $\mathcal{L}_n^*(A) \leq \mathcal{H}_\delta^n(A)$. Recordemos que

$$\mathcal{H}_\delta^n(A) = \text{ínf}\, \frac{\omega_n}{2^n} \sum_{i=1}^{\infty} \text{diam}(E_i)^n,$$

donde el ínfimo se toma sobre todos los recubrimientos $\bigcup_{i=1}^{\infty} E_i \supset A$ con $\text{diam}(E_i) < \delta$. Puesto que la adherencia de cualquier conjunto tiene el mismo diámetro que dicho conjunto, podemos suponer que los conjuntos E_i son cerrados y acotados, es decir compactos. Para tales recubrimientos, la desigualdad isodiamétrica nos da

$$\mathcal{L}_n^* \leq \sum_{i=1}^{\infty} |E_i| \leq \sum_{i=1}^{\infty} \frac{\omega_n}{2^n} \text{diam}(E_i)^n,$$

y al tomar el ínfimo sobre dichos recubrimientos se obtiene $\mathcal{L}_n^*(A) \leq \mathcal{H}_\delta^n(A)$.

Veamos ahora que $\mathcal{L}_n^*(A) \geq \mathcal{H}_\delta^n(A)$. Si $\mathcal{L}_n^*(A) = \infty$ no hay nada que probar, así que podemos suponer $\mathcal{L}_n^*(A) < \infty$. Dado $\varepsilon > 0$, existe un recubrimiento de A por una familia numerable de rectángulos $\{R_i\}$ tales que

$$\sum_{i=1}^{\infty} |R_i| \leq \mathcal{L}_n^*(A) + \varepsilon,$$

y para cada $i \in \mathbb{N}$, por el Lema 12.3 podemos hallar una familia $\{B_{ij}\}_{j\in\mathbb{N}}$ de bolas cerradas disjuntas contenidas en R_i y con diámetros menores que δ de manera que $Z_i := R_i \setminus \bigcup_{j=1}^{\infty} B_{ij}$ tiene medida de Lebesgue 0. Entonces $R_i = Z_i \cup \bigcup_{j=1}^{\infty} B_{ij}$, y

$$\mathcal{H}_\delta^n(R_i) \leq \mathcal{H}_\delta^n(Z_i) + \sum_{j=1}^{\infty} \mathcal{H}_\delta^n(B_{ij}) \leq \sum_{j=1}^{\infty} \frac{\omega_n}{2^n} \text{diam}(B_{ij})^n = \sum_{j=1}^{\infty} |B_{ij}| = |R_i|,$$

donde en la segunda desigualdad hemos usado el Lema 12.4. Por tanto

$$\mathcal{H}_\delta^n(A) \leq \sum_{i=1}^{\infty} \mathcal{H}_\delta^n(R_i) \leq \sum_{i=1}^{\infty} |R_i| \leq \mathcal{L}_n^*(A) + \varepsilon,$$

y como $\varepsilon > 0$ es arbitrario se concluye que $\mathcal{H}_\delta^n(A) \leq \mathcal{L}_n^*(A)$. □

12.4. Problemas

Problema 12.1. Si $A, B \subset \mathbb{R}^n$ y uno de ellos es abierto, o los dos son cerrados, probar que $A + B$ es medible.

Problema 12.2. Probar la desigualdad aritmética-geométrica:

$$\frac{1}{n} \sum_{j=1}^{n} x_j \geq \left(\prod_{j=1}^{n} x_j \right)^{1/n}$$

para todos $x_1, ..., x_n \geq 0$. *Indicación:* demostrarla primero en el caso en que n es potencia de dos. Luego demostrar que si la desigualdad es verdad para n entonces también lo es para $n-1$.

Problema 12.3. Escribir los detalles de los dos últimos casos de la demostración del Teorema 12.1.

Problema 12.4. Para cada compacto $A \subset \mathbb{R}^n$ se define el *contenido de Minkowski* de A por

$$\mu_+(A) = \liminf_{h \to 0} \frac{|A_h| - |A|}{h},$$

donde $A_h := \{x \in \mathbb{R}^n : d(x, A) \leq h\}$. Este número nos da una medida de ∂A.[1] Demostrar esta versión de la *desigualdad isoperimétrica*:

$$|A|^{\frac{n-1}{n}} \leq \frac{1}{n\omega_n} \mu_+(A).$$

Indicación: usar la desigualdad de Brunn-Minkowski con $A_h = A + B(0, h)$.

[1] Es cierto, aunque muy difícil de demostrar, que $\mu_+(A) \leq \mathcal{H}^{n-1}(A)$.

Capítulo 13

Las fórmulas del área y la coárea

En este capítulo estudiaremos las fórmulas del área y de la co-area debidas a H. Federer, que además de darnos herramientas de cálculo para las medidas de Hausdorff de conjuntos de nivel o de subvariedades lipschitzianas, son potentes generalizaciones de los teoremas del cambio de variable y de Fubini en $\mathbb{R}^n$.

13.1. Repaso de álgebra lineal. Jacobiano generalizado

Recordemos que una aplicación lineal $O : \mathbb{R}^n \to \mathbb{R}^m$ se dice que es *ortogonal* si

$$\langle Ox, Oy \rangle = \langle x, y \rangle$$

para todos $x, y \in \mathbb{R}^n$. Una aplicación lineal $S : \mathbb{R}^n \to \mathbb{R}^n$ es *simétrica* si

$$\langle x, Sy \rangle = \langle Sx, y \rangle$$

para todos $x, y \in \mathbb{R}^n$, lo que equivale a decir que tiene una matriz simétrica respecto de la base canónica de $\mathbb{R}^n$. Diremos que $D : \mathbb{R}^n \to \mathbb{R}^n$ es diagonal si su matriz respecto de la base canónica de $\mathbb{R}^n$ es diagonal. Para cualquier aplicación lineal $A : \mathbb{R}^n \to \mathbb{R}^m$ se define la adjunta de A como la aplicación lineal $A^* : \mathbb{R}^m \to \mathbb{R}^n$ definida por

$$\langle x, A^*y \rangle = \langle Ax, y \rangle$$

para todos $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$. Si usamos las bases canónicas nuevamente, esto es lo mismo que decir que la matriz de A^* es la que se obtiene de A al cambiar filas por columnas.

A continuación consignamos un par de resultados básicos de álgebra lineal que emplearemos para definir y utilizar el jacobiano generalizado.

Teorema 13.1. *En lo que sigue A, B, O, S, D denotan aplicaciones lineales. Se tiene:*

1. $A^{**} = A$.

2. $(AB)^* = B^*A^*$.

3. *Si $O : \mathbb{R}^n \to \mathbb{R}^n$ es ortogonal, $O^* = O^{-1}$.*

4. *Si $S : \mathbb{R}^n \to \mathbb{R}^n$ es simétrica, existen una aplicación ortogonal $O : \mathbb{R}^n \to \mathbb{R}^n$ y una aplicación diagonal $D : \mathbb{R}^n \to \mathbb{R}^n$ tales que*
$$S = ODO^{-1}.$$

5. *Si $O : \mathbb{R}^n \to \mathbb{R}^m$ es ortogonal entonces $n \leq m$, y*
$$O^*O = I \text{ en } \mathbb{R}^n, \text{ y } OO^* = I \text{ en } O(\mathbb{R}^n).$$
Además: O es una isometría sobre su imagen, $Q := O^*|_{O(\mathbb{R}^n)} : O(\mathbb{R}^n) \to \mathbb{R}^n$ también es una isometría, y $O^* = QP$, donde $P : \mathbb{R}^m \to O(\mathbb{R}^n)$ es la proyección ortogonal de $\mathbb{R}^m$ sobre su subespacio $O(\mathbb{R}^n)$.

Teorema 13.2 (de descomposición polar)**.** *Sea $L : \mathbb{R}^n \to \mathbb{R}^m$ una aplicación lineal. Se tiene:*

(a) Si $n \leq m$, existen una aplicación lineal simétrica $S : \mathbb{R}^n \to \mathbb{R}^n$ y una aplicación ortogonal $O : \mathbb{R}^n \to \mathbb{R}^m$ tales que
$$L = OS.$$

(b) Si $n \geq m$, existen $S : \mathbb{R}^m \to \mathbb{R}^m$ simétrica y $O : \mathbb{R}^m \to \mathbb{R}^n$ ortogonal tales que
$$L = SO^*.$$

Demostración. La prueba de (a) consiste en: 1) definir $C = L^*L : \mathbb{R}^n \to \mathbb{R}^n$, comprobar que es simétrica y definida no negativa, y diagonalizarla para encontrar una base ortonormal $\{x_k\}$ de $\mathbb{R}^n$ tal que $Cx_k = \mu_k x_k$, donde μ_k son los autovalores de C; 2) definir λ_k como la raíz cuadrada positiva de μ_k, y si $\lambda_k \neq 0$ poner $z_k = \frac{1}{\lambda_k} Lx_k$; se ve fácilmente que si $\lambda_k, \lambda_j \neq 0$ entonces $\langle z_k, z_j \rangle = \frac{\lambda_k}{\lambda_j}\delta_{kj}$, luego podemos completar el sistema ortonormal $\{z_k : \lambda_k \neq 0\}$ con vectores ortonormales hasta encontrar un sistema ortonormal total $\{z_k\}_{n=1}^n$ en $\mathbb{R}^n$; 3) definir $S : \mathbb{R}^n \to \mathbb{R}^n$ por $Sx_k = \lambda_k x_k$, y $O : \mathbb{R}^n \to \mathbb{R}^m$ por $Ox_k = z_k$. Los detalles quedan a cargo del lector, o pueden consultarse en cualquier texto estándar de álgebra lineal. La prueba de (b) se deduce aplicando (a) a $L^* : \mathbb{R}^m \to \mathbb{R}^n$. □

Este teorema nos permite hacer la siguiente definición.

Definición 13.1. Sea $L : R^n \to \mathbb{R}^m$ lineal.

1. Si $n \leq m$, escribimos $L = OS$ como en el teorema anterior y definimos el jacobiano de L como
$$[[L]] = |\text{det} S|.$$

2. Si $n \geq m$, escribimos $L = SO^*$ como en el teorema anterior y definimos el jacobiano de L como
$$[[L]] = |\text{det} S|.$$

La siguiente proposición nos da una forma práctica de calcular $[[L]]$ y además nos asegura que la definición de $[[L]]$ no depende de las elecciones de O y S en las anteriores descomposiciones.

Proposición 13.3. *Sea $L : \mathbb{R}^n \to \mathbb{R}^m$ lineal.*

1. Si $n \leq m$ entonces
$$[[L]]^2 = det(L^*L).$$

2. Si $n \geq m$ entonces
$$[[L]]^2 = det(LL^*).$$

Demostración. Probaremos (1) (la prueba de (2) es similar). Si $n \leq m$ tenemos
$$L = OS, \;\; L^* = SO^*,$$
luego
$$L^*L = SO^*OS = S^2$$
ya que O es ortogonal y por tanto $O^*O = I$. Así,
$$\det(L^*L) = (\det S)^2 = [[L]]^2.$$
□

Observación 13.2. Nótese que $[[L]] = [[L^*]]$.

Otro método útil para calcular $[[L]]$ nos lo proporciona la fórmula de Cauchy-Binet, para cuyo enunciado usaremos las siguientes definiciones y notaciones: si $n \leq m$ definimos

$$\Lambda(m,n) = \{\lambda : \{1,...,n\} \to \{1,...,m\} \mid \lambda \text{ es creciente.}\}$$

Para cada $\lambda \in \Lambda(m,n)$ definimos la proyección

$$P_\lambda(x_1,...,x_m) := (x_{\lambda(1)},...,x_{\lambda(n)}),$$

y también

$$S_\lambda := \text{span}\{e_{\lambda(1)},...,e_{\lambda(n)}\} \subset \mathbb{R}^m.$$

Identificando $\mathbb{R}^n$ con S_λ, P_λ es la proyección ortogonal de $\mathbb{R}^m$ en $S_\lambda \subset \mathbb{R}^m$.

Teorema 13.4 (Fórmula de Cauchy-Binet)**.** *Supongamos que $n \leq m$ y $L : \mathbb{R}^n \to \mathbb{R}^m$ es lineal. Entonces*

$$[[L]]^2 = \sum_{\lambda \in \Lambda(m,n)} (\mathit{det}(P_\lambda L))^2.$$

Es decir, para calcular $[[L]]^2$ basta sumar los cuadrados de los determinantes de cada submatriz $(n \times n)$ de la matriz $(m \times n)$ que representa a L respecto de las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$.

Demostración. Sea $A := L^*L : \mathbb{R}^n \to \mathbb{R}^n$. Identificando las aplicaciones lineales con sus matrices respecto de las bases canónicas, escribamos

$$L = (l_{ij}), \ \ A = (a_{ij}),$$

de forma que

$$a_{ij} = \sum_{k=1}^m l_{ki} l_{kj}; \ \ i,j = 1,...,n.$$

Entonces

$$[[L]]^2 = \det A = \sum_{\sigma \in \Sigma} \text{signo}(\sigma) \prod_{i=1}^n a_{i,\sigma(i)},$$

donde Σ denota el conjunto de permutaciones de $\{1,...,n\}$. Así,

$$[[L]]^2 = \sum_{\sigma \in \Sigma} \text{signo}(\sigma) \prod_{i=1}^n \sum_{k=1}^m l_{ki} l_{k\sigma(i)} = \sum_{\sigma \in \Sigma} \text{signo}(\sigma) \sum_{\phi \in \Phi} \prod_{i=1}^n l_{\phi(i)i} l_{\phi(i)\sigma(i)},$$

donde Φ denota el conjunto de aplicaciones inyectivas de $\{1,...,n\}$ en $\{1,...,m\}$. Para cada $\phi \in \Phi$, podemos esribir de forma única $\phi = \lambda \circ \theta$, donde $\theta \in \Sigma$ y $\lambda \in \Lambda(m,n)$. Por tanto

$$\begin{aligned}
[[L]]^2 &= \sum_{\sigma \in \Sigma} \text{signo}(\sigma) \sum_{\lambda \in \Lambda(m,n)} \sum_{\theta \in \Sigma} \prod_{i=1}^n l_{\lambda\circ\theta(i),i} l_{\lambda\circ\theta(i),\sigma(i)} \\
&= \sum_{\sigma \in \Sigma} \text{signo}(\sigma) \sum_{\lambda \in \Lambda(m,n)} \sum_{\theta \in \Sigma} \prod_{i=1}^n l_{\lambda(i),\theta^{-1}(i)} l_{\lambda(i),\sigma\circ\theta^{-1}(i)} \\
&= \sum_{\lambda \in \Lambda(m,n)} \sum_{\theta \in \Sigma} \sum_{\sigma \in \Sigma} \text{signo}(\sigma) \prod_{i=1}^n l_{\lambda(i),\theta(i)} l_{\lambda(i),\sigma\circ\theta(i)} \\
&= \sum_{\lambda \in \Lambda(m,n)} \sum_{\rho \in \Sigma} \sum_{\theta \in \Sigma} \text{signo}(\sigma)\text{signo}(\rho) \prod_{i=1}^n l_{\lambda(i),\theta(i)} l_{\lambda(i),\rho(i)} = \\
&= \sum_{\lambda \in \Lambda(m,n)} \left(\sum_{\theta \in \Sigma} \text{signo}(\theta) \prod_{i=1}^n l_{\lambda(i),\theta(i)} \right)^2 = \sum_{\lambda \in \Lambda(m,n)} (\det(P_\lambda L))^2,
\end{aligned}$$

donde en la cuarta igualdad pusimos $\rho = \sigma \circ \theta$. Esto concluye la prueba. $\square$

Ya estamos preparados para definir el jacobiano (generalizado) de una aplicación Lipschitz de $\mathbb{R}^n$ en $\mathbb{R}^m$.

Definición 13.3. Sea $f : \mathbb{R}^n \to \mathbb{R}^m$ de Lipschitz. Por el teorema de Rademacher sabemos que f es diferenciable en $\mathcal{L}_n$-casi todo punto. Para cada punto x donde $Df(x)$ existe, definimos

$$Jf(x) := [[Df(x)]],$$

al que llamaremos jacobiano de f.

13.2. La fórmula del área

Recordemos que el teorema del cambio de variables que se estudia en segundo dice que si $f : U \to f(U)$ es un difeomorfismo de clase C^1 entre dos abiertos de $\mathbb{R}^n$ y $g : f(U) \to [0, \infty]$ es medible entonces $g \circ f$ es medible y

$$\int_{f(U)} g(y)dy = \int_U g(f(x))Jf(x)dx. \tag{13.1}$$

Como consecuencia (que en realidad es una versión equivalente del enunciado, porque este se deduce rápidamente de ella, pasando de conjuntos a funciones simples, y de estas, gracias al teorema de la convergencia monótona, a funciones positivas), se tiene que, para todo $A \subset U$ medible, $f(A)$ es medible, y

$$|f(A)| = \int_A Jf(x)dx. \tag{13.2}$$

La versión más sencilla de la fórmula del área nos dice que si A es medible y $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ es de Lipschitz e inyectiva y $n \leq m$ entonces

$$\mathcal{H}^n(f(A)) = \int_A Jf(x)dx. \tag{13.3}$$

Es claro que esta fórmula generaliza en varios sentidos a (13.2) (incluso si $n = m$, lo único que pedimos es que f sea de Lipschitz e inyectiva, no hace falta que sea un difeomorfismo); además, cuando $n < m$, nos permite calcular la medida de Hausdorff n-dimensional de conjuntos parametrizados por funciones Lipschitz (por ejemplo, subvariedades diferenciables de dimensión n en $\mathbb{R}^m$. Como en el caso del teorema del cambio de variable, este tipo de fórmula nos permite obtener fácilmente otra más general para funciones medibles Borel $g : f(A) \to [0, \infty]$: se cumple que $g \circ f : A \to [0, \infty]$ es medible Borel y

$$\int_{f(A)} g(y)d\mathcal{H}^n(y) = \int_A g(f(x))Jf(x)dx. \tag{13.4}$$

La versión más general de la fórmula del área nos permite incluso prescindir de la hipótesis de que f sea inyectiva.

Teorema 13.5 (Fórmula del área)**.** *Si $n \leq m$, $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ es localmente Lipschitz (donde U es un abierto de $\mathbb{R}^n$), y $A \subset U$ es medible Lebesgue, entonces*

$$\int_{\mathbb{R}^m} \mathcal{H}^0(A \cap f^{-1}(\{y\})d\mathcal{H}^n(y) = \int_A Jf(x)dx.$$

En particular, si $\mathcal{L}_n(A) < \infty$ y f es de Lipschitz entonces $\int_A Jf(x)dx$ es finita y se ve que $f^{-1}(\{y\}) \cap A$ es finito para $\mathcal{H}^n$-casi todo $y \in \mathbb{R}^n$. Obsérvese también que si f es inyectiva entonces $\mathcal{H}^0(A \cap f^{-1}(\{y\}) = 1$ si $y \in f(A)$, y 0 en otro caso, con lo que se recupera (13.3). En el caso general, la fórmula nos dice que la medida $\mathcal{H}^n$ de la imagen de f, contando multiplicidades, se calcula integrando el jacobiano de f sobre A.

Como antes, pasando de funciones características, de estas a funciones simples, y usando el teorema de la convergencia monótonca para llegar al caso de funciones positivas medibles, se puede deducir lo siguiente.

Corolario 13.6 (Cambio de variables generalizado)**.** *Si* $n \leq m$ *y* $f : \mathbb{R}^n \to \mathbb{R}^m$ *es Lipschitz, para cualquier función* $\mathcal{L}_n$*-medible* $g : \mathbb{R}^n \to [0, \infty]$ *(o bien* $g \in L^1(\mathbb{R}^n)$*) se tiene*

$$\int_{\mathbb{R}^m} \left(\sum_{x \in f^{-1}(\{y\})} g(x) \right) d\mathcal{H}^n(y) = \int_{\mathbb{R}^n} g(x) Jf(x) dx.$$

Antes de demostrar la fórmula del área veamos algunas aplicaciones importantes: todos los resultados ya conocidos vistos en otras asignaturas para medidas de curvas, superficies y en general subvariedades diferenciables de clase C^1 se recuperan, y se extienden al caso de parametrizaciones localmente lipschitzianas.

13.2.1. Longitud de una curva lipschitziana

Sea $f : [a, b] \subset \mathbb{R} \to \mathbb{R}^m$ una función Lipschitz e inyectiva. Entonces se tiene que, para casi todo $t \in [a, b]$,

$$Jf^2 = [[Df]]^2 = \sum_{j=1}^{m} \left(\frac{d}{dt} f_j \right)^2,$$

luego $Jf(t) = |f'(t)|$ y por tanto, por la fórmula del área obtenemos

$$\mathcal{H}^1 \left(f([a, b]) \right) = \int_a^b |f'(t)| dt,$$

y ambos números son finitos. Análogamente, si $f : (a, b) \to \infty$ es localmente Lipschitz, es fácil deducir de lo anterior que

$$\mathcal{H}^1 \left(f((a, b)) \right) = \int_a^b |f'(t)| dt,$$

aunque en este caso ambos números pueden ser infinitos.

13.2.2. Medida de una gráfica de una función Lipschitz

Supongamos que $f : U \subset \mathbb{R}^n \to \mathbb{R}$ es localmente Lipschitz, y consideremos su gráfica $G_f := \{(x, y) \in \mathbb{R}^n \times \mathbb{R} : y = f(x), x \in U\}$. Vamos a calcular $\mathcal{H}^n(G_f)$. Definamos $g : \mathbb{R}^n \to \mathbb{R}^{n+1}$ por $g(x) = (x, f(x))$, y observemos que $G_f = g(U)$. Entonces, en $\mathcal{L}_n$-casi todo punto de U se tiene

$$Dg = \begin{pmatrix} 1 & ... & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \\ \frac{\partial f}{\partial x_1} & \dots & \frac{\partial f}{\partial x_n} \end{pmatrix},$$

matriz de tamaño $(n + 1) \times n$. Usando la fórmula de Cauchy-Binet, vemos que $(Jg)^2 = [[Dg]]^2$ es la suma de los cuadrados de los subdeterminantes de tamaño $n \times n$ de la anterior matriz, lo que da exactamente $1 + |\nabla f|^2$. Aplicando la fórmula del área obtenemos entonces

$$\mathcal{H}^n(G_f) = \mathcal{H}^n(g(U)) = \int_U \sqrt{1 + |\nabla f(x)|^2} dx.$$

13.2.3. Medida de una hipersuperficie parametrizada

Sean $n \geq 1$, $m = n + 1$. Supongamos que $f : U \subset \mathbb{R}^n \to \mathbb{R}^{n+1}$ es localmente Lipschitz e inyectiva. Entonces

$$Df = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_{n+1}}{\partial x_1} & \dots & \frac{\partial f_{n+1}}{\partial x_n} \end{pmatrix},$$

matriz de tamaño $(n+1) \times n$. Usando la fórmula de Cauchy-Binet, en $\mathcal{L}_n$-casi todo punto de U es

$$(Jf)^2 = \sum_{k=1}^{n+1} \left[\frac{\partial(f_1, ..., f_{k-1}, f_{k+1}, ..., f_{n+1})}{\partial(x_1, ..., x_n)} \right]^2,$$

y aplicando la fórmula del área concluimos que

$$\mathcal{H}^n(f(U)) = \int_U \sqrt{\sum_{k=1}^{n+1} \left[\frac{\partial(f_1, ..., f_{k-1}, f_{k+1}, ..., f_{n+1})}{\partial(x_1, ..., x_n)} \right]^2} dx.$$

13.2.4. Medida de una subvariedad diferenciable de $\mathbb{R}^m$

Sea (M, g) una variedad riemanniana de dimensión n. Por el teorema de Nash de incrustación isométrica (isometric embedding), podemos suponer que M es una subvariedad diferenciable de $\mathbb{R}^m$, para algún $m \geq n$. Para toda carta $f : U \subset \mathbb{R}^n \to M$ de M, y todo $A \subset f(U)$ boreliano en $\mathbb{R}^m$, $B := f^{-1}(A)$ es boreliano en $\mathbb{R}^n$, y se tiene

$$g_{ij} := \langle f_{x_i}, f_{x_j} \rangle.$$

Luego

$$(Df)^*(Df) = (g_{ij}),$$

y por tanto

$$Jf = |\text{det}(g_{ij})|^{1/2},$$

con lo cual

$$\mathcal{H}^n(A) = \int_B |\text{det}(g_{ij})|^{1/2}$$

Pero obsérvese que este cálculo sigue siendo válido si sólo se supone f localmente Lipschitz e inyectiva, con lo que se obtiene una generalización importante de lo ya conocido: pueden definirse subvariedades localmente Lipschitz de $\mathbb{R}^n$, y calcular su medida de Hausdorff $\mathcal{H}^n$ de acuerdo con la anterior fórmula.

13.3. Demostración de la fórmula del área

Dividiremos la prueba en varios lemas. Comenzamos estudiando el caso de una aplicación lineal. En toda esta sección n, m serán números naturales con $n \leq m$.

Lema 13.7. *Si $T : \mathbb{R}^n \to \mathbb{R}^n$ es un isomorfismo lineal y $A \subset \mathbb{R}^n$ es abierto entonces*

$$\mathcal{L}_n(T(A)) = |det(T)|\mathcal{L}_n(A).$$

Demostración. **Caso 1.** Supongamos primero que T es ortogonal; entonces T es una isometría, es decir, T y T^{-1} son 1-Lipschitz. Por tanto $\mathcal{H}^n(A) = \mathcal{H}^n(T^{-1}(T(A))) \leq \mathcal{H}^n(T(A)) \leq \mathcal{H}^n(A)$, y como $\mathcal{L}_n = \mathcal{H}^n$ y $|\text{det}(T)| = 1$ se deduce el resultado.

Caso 2. Supongamos que T es simétrica. Entonces T es diagonalizable, y podemos escribir $T = PDP^{-1}$, donde P tiene matriz formada por una colección de autovectores de T que forman una base ortonormal de $\mathbb{R}^n$; en particular P y P son ortogonales (teorema espectral real). Como A es abierto, podemos escribir también $A = \bigcup_{j \in \mathbb{N}} Q_j$, donde los $Q_j = [a_1^j, b_1^j] \times ... \times [a_n^j, b_n^j]$ son cubos diádicos y la unión es disjunta (los cubos tienen interiores disjuntos dos a dos). Si los autovalores de T, y por tanto la diagonal de D, son $\lambda_1, ..., \lambda_n$, entonces para cada $j \in \mathbb{N}$ se tiene que $D(Q_j) = [c_1^j, d_1^j] \times ... \times [c_n^j, d_n^j]$, donde $[c_i^j, d_i^j] = [\lambda_i a_i^j, \lambda_i b_i^j]$ si $\lambda_i > 0$, y $[c_i^j, d_i^j] = [\lambda_i b_i^j, \lambda_i a_i^j]$ si $\lambda_i < 0$; por tanto

$$\mathcal{L}_n\left(D(Q_j)\right) = \prod_{i=1}^{n} |\lambda_i|(b_i^j - a_i^j) = \left(\prod_{i=1}^{n} |\lambda_i|\right)\left(\prod_{i=1}^{n}(b_i^j - a_i^j)\right) = \text{det}(T)\mathcal{L}_n(Q_j).$$

Luego

$$\mathcal{L}_n(D(A)) = \sum_{j\in\mathbb{N}} \mathcal{L}_n(Q_j) = \sum_{j\in\mathbb{N}} \det(T)\mathcal{L}_n(Q_j) = \det(T)\mathcal{L}_n(A).$$

Entonces, como P y P^{-1} son ortogonales, usando el caso anterior obtenemos

$$\mathcal{L}_n(T(A)) = \mathcal{L}_n(PDP^{-1}(A)) = \mathcal{L}_n(DP^{-1}(A)) = \det(T)\mathcal{L}_n(P^{-1}(A)) = \det(T)\mathcal{L}_n(A).$$

Caso 3. En el caso general, por el teorema de descomposición polar, se tiene $T = OS$, donde O es ortogonal y S es simétrica. Combinando los dos casos anteriores obtenemos

$$\mathcal{L}_n(T(A)) = \mathcal{L}_n(O(S(A))) = \mathcal{L}_n(S(A)) = |\det(S)|\mathcal{L}_(A) = |\det(T)|\mathcal{L}_n(A),$$

lo que completa la demostración. □

Lema 13.8. *Si $L : \mathbb{R}^n \to \mathbb{R}^m$ es lineal entonces*

$$\mathcal{H}^n(L(A)) = [[L]]\mathcal{L}_n^*(A)$$

para todo $A \subset \mathbb{R}^n$.

Demostración. Supongamos primero que A es abierto. Escribamos $L = OS$ con $S : \mathbb{R}^n \to \mathbb{R}^n$ simétrica y $O : \mathbb{R}^n \to \mathbb{R}^m$ ortogonal, como en el teorema 13.2. Usando la notación del Teorema 13.1(5), tenemos $O^* = QP$, donde $P : \mathbb{R}^m \to O(\mathbb{R}^n)$ es la proyección ortogonal y Q es una isometría; por tanto O^* es 1-Lipschitz, como también lo es O (por ser isometría sobre su imagen). Entonces (usando también el lema anterior) se tiene

$$\mathcal{H}^n(L(A)) = \mathcal{H}^n(O(S(A))) \le \mathcal{H}^n(S(A)) = \mathcal{L}_n(S(A)) = |\det(S)|\mathcal{L}_n(A),$$

y por otro lado, puesto que $O^*O = I$ en $\mathbb{R}^n$, también

$$\mathcal{L}_n(S(A)) = \mathcal{H}^n(O^*OS(A)) \le \mathcal{H}^n(OS(A)) = \mathcal{H}^n(L(A)).$$

Combinando estas desigualdades se obtiene que

$$\mathcal{H}^n(L(A)) = |\det(S)|\mathcal{L}_n(A) = [[L]]\mathcal{L}_n(A)$$

para cualquier abierto $A \subset \mathbb{R}^n$. Como las medidas exteriores $\mathcal{L}_n^*$ y $\mathcal{H}^n$ son exteriormente regulares, se deduce fácilmente que el resultado es cierto para todo $A \subset \mathbb{R}^n$. □

El siguiente lema nos asegura que la función que aparece en el integrando de la izquierda de la fórmula del área es medible.

Lema 13.9. *Sean $n \le m$, $f : \mathbb{R}^n \to \mathbb{R}^m$ Lipschitz, $A \subset \mathbb{R}^n$ medible Lebesgue. Entonces:*

1. *$f(A)$ es $\mathcal{H}^n$-medible;*

2. *la aplicación $y \mapsto \mathcal{H}^0(A \cap f^{-1}(\{y\}))$ es $\mathcal{H}^n$-medible en $\mathbb{R}^m$, y*

3. $$\int_{\mathbb{R}^m} \mathcal{H}^0(A \cap f^{-1}(\{y\}))d\mathcal{H}^n(y) \le (Lip(f))^n\, \mathcal{L}_n(A).$$

(en este contexto, a la función $y \mapsto \mathcal{H}^0(A \cap f^{-1}(\{y\}))$ se le llama función de multiplicidad).

Demostración. Podemos suponer sin pérdida de generalidad que A es acotado. Para cada $i \in \mathbb{N}$ existe $K_i \subset A$ compacto tal que

$$\mathcal{L}_n(K_i) < \mathcal{L}_n(A) - \frac{1}{i},$$

luego $\mathcal{L}_n(A \setminus K_i) < 1/i$. Como f es continua, $f(K_i)$ es compacto, y en particular $\mathcal{H}^n$-medible. Por tanto

$$f(\bigcup_{i=1}^{\infty} K_i) = \bigcup_{i=1}^{\infty} f(K_i)$$

también es medible, y como

$$\mathcal{H}^n\left(f(A) \setminus f\left(\bigcup_{i=1}^{\infty} K_i\right)\right) \leq \mathcal{H}^n\left(f\left(A \setminus \bigcup_{i=1}^{\infty} K_i\right)\right) \leq (\mathrm{Lip}(f))^n \, \mathcal{L}_n\left(A \setminus \bigcup_{i=1}^{\infty} K_i\right) = 0,$$

se deduce que $f(A)$ es $\mathcal{H}^n$-medible. Esto prueba (1).

Para probar (2), definamos B_k como la familia de todos los cubos Q del tipo

$$Q = (a_1, b_1] \times \ldots \times (a_n, b_n],$$

con $a_i = c_i/2^k$, $b_i = (c_i + 1)/2^k$, $c_i \in \mathbb{Z}$, y observemos que $\mathbb{R}^n = \bigcup_{Q \in B_k} Q$, unión disjunta. Entonces la función

$$g_k := \sum_{Q \in B_k} \chi_{f(A \cap Q)}$$

es $\mathcal{H}^n$-medible por (1), y

$$g_k(y) = \sharp\{Q \in B_k : f^{-1}(\{y\}) \cap A \cap Q \neq \emptyset\}.$$

Por tanto,

$$\mathcal{H}^0(A \cap f^{-1}(\{y\})) = \lim_{k \to \infty} g_k(y)$$

es $\mathcal{H}^n$-medible, al ser límite de una sucesión de funciones $\mathcal{H}^n$-medibles. Esto demuestra (2). Además, la sucesión g_k es creciente, luego por el Teorema de la Convergencia Monótona tenemos

$$\int_{\mathbb{R}^m} \mathcal{H}^0(A \cap f^{-1}(\{y\})) d\mathcal{H}^n(y) = \lim_{k \to \infty} \int_{\mathbb{R}^m} g_k d\mathcal{H}^n = \lim_{k \to \infty} \sum_{Q \in B_k} \mathcal{H}^n(f(A \cap Q))$$
$$\leq \limsup_{k \to \infty} \sum_{Q \in B_k} (\mathrm{Lip}(f))^n \, \mathcal{L}_n(A \cap Q) = (\mathrm{Lip}(f))^n \, \mathcal{L}_n(A).$$

□

El siguiente lema es el meollo técnico de todo el asunto: nos permite descomponer el conjunto donde $Jf(x)$ es no nulo en una familia numerable de borelianos tal que la restricción de f a cada uno de ellos se comporta de manera aproximadamente lineal.

Lema 13.10. *Sean $n \leq m$, $F : \mathbb{R}^n \to \mathbb{R}^m$ Lipschitz, $t > 1$, y*

$$B := \{x \in \mathbb{R}^n : Df(x) \text{ existe}, \; Jf(x) > 0\}.$$

Entonces existe una familia $\{E_k\}_{k \in \mathbb{N}}$ de borelianos de $\mathbb{R}^n$ tales que:

1. $B = \bigcup_{k=1}^{\infty} E_k$;
2. $f|_{E_k}$ *es inyectiva para cada* $k \in \mathbb{N}$, *y*

3. *Para cada $k \in \mathbb{N}$ existe un isomorfismo lineal simétrico $T_k : \mathbb{R}^n \to \mathbb{R}^n$ tal que*

$$\textit{Lip}\left((f|_{E_k}) \circ T_k^{-1}\right) \leq t, \quad \textit{Lip}\left(T_k \circ (f|_{E_k})^{-1}\right) \leq t, \quad y$$

$$t^{-n}|\textit{det}(T_k)| \leq Jf|_{E_k} \leq t^n|\textit{det}(T_k)|.$$

Demostración. Elijamos $\varepsilon > 0$ tal que

$$t^{-1} + \varepsilon < 1 < t - \varepsilon.$$

Sean C un subconjunto denso numerable de B, y $\mathcal{S}$ un subconjunto denso numerable del conjunto de los isomorfismos lineales simétricos de $\mathbb{R}^n$. Para cada $c \in C$, $T \in \mathcal{S}$, $i \in \mathbb{N}$, definamos $E(c, T, i)$ como el conjunto de todos los puntos $b \in B \cap B(c, 1/c)$ que cumplen

$$\left(t^{-1} + \varepsilon\right)|Tv| \leq |Df(b)v| \leq (t - \varepsilon)|Tv| \tag{13.5}$$

para todo $v \in \mathbb{R}^n$, y que también cumplen

$$|f(a) - f(b) - Df(b)(a - b)| \leq \varepsilon|T(a - b)| \tag{13.6}$$

para todo $a \in B(b, 2/i)$. De estas dos condiciones se sigue la desigualdad

$$t^{-1}|T(a - b)| \leq |f(a) - f(b)| \leq t|T(a - b)| \tag{13.7}$$

para todos los $b \in E(c, T, i)$, $a \in B(b, 2/i)$.

Nótese que los conjuntos $E(c, T, i)$ son borelianos ya que Df es medible Borel; véase la Observación 11.6.

Como el conjunto $C \times \mathcal{S} \times \mathbb{N}$ es numerable, podemos escribir la familia $\{E(c, T, i)\}_{c \in C, T \in \mathcal{S}, i \in \mathbb{N}}$ en forma de una sucesión de conjuntos $\{E_k\}_{k \in \mathbb{N}}$: estos son los conjuntos del enunciado, y puesto que $E(c, T, i) \subset B(c, 1/i) \subset B(b, 2/i)$ y T es isomorfismo, la desigualdad (13.7) nos da directamente la propiedad (2) del enunciado, es decir que la restricción de f a cada conjunto $E(c, T, i)$ es inyectiva. Y no sólo eso, sino que también se obtiene

$$\text{Lip}\left(f|_{E(c,T,i)} \circ T^{-1}\right) \leq t, \quad \text{y} \quad \text{Lip}\left(T \circ (f|_{E(c,T,i)})^{-1}\right) \leq t,$$

lo que prueba la primera mitad de la condición (3) del enunciado.

Veamos ahora que se cumple la propiedad (1) del enunciado. Dado cualquier $b \in B$, escribamos $Df(b) = OS$, con O ortogonal y S simétrica, y elijamos $T \in \mathcal{S}$ tal que

$$\text{Lip}(T \circ S^{-1}) \leq (t^{-1} + \varepsilon)^{-1}, \quad \text{y} \quad \text{Lip}(S \circ T^{-1}) \leq t - \varepsilon.$$

Escojamos ahora $i \in \mathbb{N}$, $c \in C$ tales que $|b - c| < 1/i$ y

$$|f(a) - f(b) - Df(b)(a - b)| \leq \frac{\varepsilon}{\text{Lip}(T^{-1})}|a - b| \leq \varepsilon|T(a - b)|$$

para todo $a \in B(a, 2/i)$; esto último puede hacerse gracias a la diferenciabilidad de f en b. Entonces se cumple (13.6), y además

$$|Df(b)v| = |OS(v)| = |S(v)| = |ST^{-1}T(v)| \leq (t - \varepsilon)|Tv|,$$

y

$$(t^{-1} + \varepsilon)|Tv| = (t^{-1} + \varepsilon)|TS^{-1}Sv| \leq (t^{-1} + \varepsilon)(t^{-1} + \varepsilon)^{-1}|Sv| = |Df(b)v|,$$

luego también se cumple (13.5). Es decir, se tiene $b \in E(c, T, i)$. Esto prueba (1).

Finalmente, veamos que también se cumple la segunda mitad de la condición (3) del enunciado. Si $b \in E(c, T, i)$, escribiendo $Df(b) = L = OS$ como antes, se tiene

$$Jf(b) = [[L]] = |\text{det}(S)|.$$

Por (13.5) sabemos que

$$(t^{-1}+\varepsilon)|Tv| \leq |OSv| = |Sv| \leq (t-\varepsilon)|Tv|$$

para todo $v \in \mathbb{R}^n$, y por tanto

$$(t^{-1}+\varepsilon)|u| \leq |ST^{-1}u| \leq (t-\varepsilon)|u|$$

para todo $u \in \mathbb{R}^n$. Esto implica que

$$(ST^{-1})(B(0,1)) \subset B(0,t-\varepsilon),$$

y por consiguiente

$$|\text{det}(ST^{-1})|\alpha(n) \leq \mathcal{L}_n\left(B(0,t-\varepsilon)\right) = \alpha(n)(t-\varepsilon)^n,$$

de donde

$$Jf(b) = |\text{det}(S)| \leq (t-\varepsilon)^n|\text{det}(T)|.$$

La otra desigualdad se prueba análogamente. □

Ya estamos preparados técnicamente para terminar la demostración de la fórmula del área.

Demostración del teorema 13.5. Consideremos primero el caso en que $f:\mathbb{R}^n \to \mathbb{R}^m$ es Lipschitz. Gracias al teorema de Rademacher[1] y al Lema 13.9(3), salvo descontar un boreliano de medida 0, podemos suponer que f es diferenciable en todo punto de A. También podemos suponer sin pérdida de gererализ dad que $\mathcal{L}_n(A) < \infty$ (una vez probado para este tipo de conjuntos, se deduce sin dificultad para los de medida infinita usando el teorema de la convergencia monótona). Consideraremos dos casos.

Caso 1. Supongamos que $A \subset \{x : Jf(x) > 0\}$. Fijemos $t > 1$ y elijamos una familia $\{E_k\}_{k\in\mathbb{N}}$ de borelianos como en el Lema 13.10. Por el procedimiento habitual (es decir, definir $\widetilde{E}_1 = E_1$, y $\widetilde{E}_k = E_k \setminus \bigcup_{j=1}^{k-1} E_j$ para $k \geq 2$, y reemplazar una familia por otra), podemos suponer sin pérdida de generalidad que los E_k son disjuntos dos a dos. Definamos también B_k como en la prueba del lema 13.9, y pongamos

$$F_{i,j,k} := E_j \cap Q_j \cap A,$$

donde $\{Q_j\}_{j\in\mathbb{N}}$ es una enumeración de B_k. Para cada $k \in \mathbb{N}$ es $F_{i,j,k} \cap F_{i',j',k} = \emptyset$ si $(i,j) \neq (i',j')$, y se tiene que $A = \bigcup_{i,j\in\mathbb{N}} F_{i,j,k}$ para cada k. Definiendo

$$g_k := \sum_{i,j\in\mathbb{N}} \chi_{f(F_j^i)},$$

observamos que $g_k(y)$ es el número de conjuntos $F_{i,j,k}$ para los cuales $F_{i,j,k} \cap f^{-1}(\{y\}) \neq \emptyset$, y por tanto se tiene que

$$\lim_{k\to\infty} g_k(y) = \mathcal{H}^0(A \cap f^{-1}(\{y\})).$$

Aplicando el Teorema de la Convergencia Monótona a la sucesión de funciones g_k obtenemos que

$$\lim_{k\to\infty} \sum_{i,j\in\mathbb{N}} \mathcal{H}^n(f(F_{i,j,k})) = \int_{\mathbb{R}^m} \mathcal{H}^0(A \cap f^{-1}(\{y\}))d\mathcal{H}^n(y). \tag{13.8}$$

Ahora, por el Lema 13.10, tenemos que

$$\mathcal{H}^n\left(f(F_{i,j,k})\right) = \mathcal{H}^n\left(f|_{E_j} \circ T_j^{-1} \circ T_j(F_{i,j,k})\right) \leq t^n \mathcal{L}_n\left(T_j(F_{i,j,k})\right),$$

y

$$\mathcal{H}^n\left(T_j(F_{i,j,k})\right) = \mathcal{H}^n\left(T_j \circ (f|_{E_j})^{-1} \circ f(F_{i,j,k})\right) \leq t^n \mathcal{L}_n\left(f(F_{i,j,k})\right).$$

[1]Recordemos otra vez que en la demostración del teorema de Rademacher se ve que el conjunto de puntos donde f es diferenciable es de Borel

Luego, volviendo a usar el Lema 13.10 y también el Lema 13.8, obtenemos

$$
\begin{aligned}
&t^{-2n}\mathcal{H}^n\left(f(F_{i,j,k})\right) \le t^{-n}T_j(F_{i,j,k}) = t^{-n}|\det(T_j)|\mathcal{L}_n(F_{i,j,k})\\
&\le \int_{F_{i,j,k}} Jf(x)dx \le t^n|\det(T_j)|\mathcal{L}_n(F_{i,j,k}) = t^n\mathcal{L}_n\left(T_j(F_{i,j,k})\right) \le t^{2n}\mathcal{H}^n\left(f(F_{i,j,k})\right).
\end{aligned}
$$

Sumando en $i,j\in\mathbb{N}$, deducimos que

$$
t^{-2n}\sum_{i,j\in\mathbb{N}}\mathcal{H}^n\left(f(F_{i,j,k})\right) \le \int_A Jf(x)dx \le t^{2n}\sum_{i,j\in\mathbb{N}}\mathcal{H}^n\left(f(F_{i,j,k})\right),
$$

y haciendo ahora $k\to\infty$ y recordando (13.8), concluimos que

$$
t^{-2n}\int_{\mathbb{R}^m}\mathcal{H}^0(A\cap f^{-1}(\{y\})d\mathcal{H}^n(y) \le \int_A Jf(x)dx \le t^{2n}\int_{\mathbb{R}^m}\mathcal{H}^0(A\cap f^{-1}(\{y\})d\mathcal{H}^n(y).
$$

Finalmente, como $t>1$ es arbitrario, haciendo $t\to 1^+$ se obtiene

$$
\int_{\mathbb{R}^m}\mathcal{H}^0(A\cap f^{-1}(\{y\})d\mathcal{H}^n(y) = \int_A Jf(x)dx
$$

Caso 2. Supongamos que $A\subset\{x : Jf(x)=0\}$. Fijemos $\varepsilon\in(0,1]$, y escribamos

$$
f = P\circ g,
$$

donde $g:\mathbb{R}^n\to\mathbb{R}^m\times\mathbb{R}^n$ está definida por

$$
g(x) := (f(x),\varepsilon x),
$$

y $P:\mathbb{R}^m\times\mathbb{R}^n\to\mathbb{R}^m$ es la proyección

$$
P(x,y) := y.
$$

Veamos que existe una constante $C>0$ tal que

$$
0 < Jg(x) \le C\varepsilon
$$

para todo $x\in A$, donde C sólo depende de n, m, y $\text{Lip}(f)$. En efecto,

$$
g(x) = (f_1(x),...,f_m(x),\varepsilon x_1,...,\varepsilon x_n),
$$

luego

$$
Dg(x) = \begin{pmatrix} Df(x)\\ \varepsilon I\end{pmatrix},
$$

matriz de tamaño $(n+m)\times n$. Por la fórmula de Cauchy-Binet, $Jg(x)^2$ es la suma de los cuadrados de los subdeterminantes de tamaño $n\times n$ de $Dg(x)$, es obvio que $Jg(x)^2\ge\varepsilon^{2n}>0$. Por otra parte, como $|Df|\le\text{Lip}(f)<\infty$, también tenemos

$$
Jg(x)^2 = Jf(x)^2 + \left\{\begin{array}{c}\text{suma de cuadrados de } c(n,m) \text{ términos,}\\ \text{cada uno de los cuales tiene al menos un } \varepsilon\end{array}\right\} \le C\varepsilon^2,
$$

para todo $x\in A$, donde $c(n,m)\in\mathbb{N}$ sólo depende de n,m, y C sólo depende de n,m y $\text{Lip}(f)$.

Ahora, como P es 1-Lipschitz, tenemos, usando el Caso 1,

$$
\begin{aligned}
&\mathcal{H}^n(f(A)) = \mathcal{H}^n(P(g(A))) \le \mathcal{H}^n(g(A))\\
&\le \int_{\mathbb{R}^{n+m}}\mathcal{H}^0(A\cap g^{-1}(\{(y,z)\})d\mathcal{H}^n(y,z) = \int_A Jg(x)dx \le \varepsilon C\mathcal{L}_n(A).
\end{aligned}
$$

Entonces, haciendo $\varepsilon \to 0$ deducimos que

$$\mathcal{H}^n(f(A)) = 0,$$

y por tanto también

$$\int_{\mathbb{R}^m} \mathcal{H}^0(A \cap f^{-1}(\{y\}) d\mathcal{H}^n(y) = 0,$$

ya que la función $y \mapsto \mathcal{H}^0(A \cap f^{-1}(\{y\})$ es cero fuera de $f(A)$. Luego se tiene

$$\int_{\mathbb{R}^m} \mathcal{H}^0(A \cap f^{-1}(\{y\}) d\mathcal{H}^n(y) = 0 = \int_A Jf(x)dx$$

también en este caso.

Caso 3. En el caso general, podemos escribir $A = A_1 \cup A_2$, donde $A_1 = \{x : Jf(x) > 0\}$, $A_2 = \{x : Jf(x) = 0\}$, y aplicar el Caso 1 a A_1 y el caso 2 a A_2, concluyendo que

$$\begin{aligned}
&\int_{\mathbb{R}^m} \mathcal{H}^0(A \cap f^{-1}(\{y\}) d\mathcal{H}^n(y) \\
&= \int_{\mathbb{R}^m} \mathcal{H}^0(A_1 \cap f^{-1}(\{y\}) d\mathcal{H}^n(y) + \int_{\mathbb{R}^m} \mathcal{H}^0(A_2 \cap f^{-1}(\{y\}) d\mathcal{H}^n(y) \\
&= \int_{A_1} Jf(x)dx + \int_{A_2} Jf(x)dx = \int_A Jf(x)dx.
\end{aligned}$$

Esto concluye la prueba en el caso en que $f : \mathbb{R}^n \to \mathbb{R}^m$ es Lipschitz.

Caso 4. Como toda función Lipschitz $f : V \subset \mathbb{R}^n \to \mathbb{R}^m$ admite una extensión Lipschitz a todo $\mathbb{R}^n$, también se tiene el resultado para funciones Lipschitz definidas en subconjuntos medibles propios V de $\mathbb{R}^n$. Veamos cómo esto también implica el resultado para funciones $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ localmente Lipschitz, donde U es un abierto cualquiera.

Como f es localmente Lipschitz, para todo $x \in \mathbb{R}^n$, existe $B(x, r_x) \subset U$ tal que f es Lipschitz en esta bola; y poniendo $U = \bigcup_{x \in \mathbb{R}^n} B(x, r_x)$, como U es separable existe una sucesión $\{x_k\}_{k \geq 1} \subset U$ de manera que $U = \bigcup_{k \geq 1} B(x_k, r_{x_k})$. Más aún, poniendo $\tilde{B}_k = \bigcup_{j=1}^k B(x_j, r_{x_j})$ obtenemos una sucesión de abiertos $\left\{\tilde{B}_k\right\}_{k \geq 1}$ creciente a U y de manera que f es Lipschitz en cada $\tilde{B}_k$ (porque es una unión finita de abiertos en los que f es Lipschitz). Además, para cada k existe una extensión Lipschitz $\tilde{f}_k$ de $f|_{\tilde{B}_k}$ a todo $\mathbb{R}^n$. Para todo conjunto $\mathcal{L}^n$-medible A se tiene que $A \cap \tilde{B}_k \nearrow A$. Como $J(f) \geq 0$ y $J(f)\chi_{A \cap \tilde{B}_k}$ es una sucesión creciente a $J(f)\chi_A$ de funciones positivas, el Teorema de la Convergencia Monótona nos dice que

$$\int_A J(f)(x)\, dx = \lim_{k \to \infty} \int_{A \cap \tilde{B}_k} J(f)(x)\, dx = \lim_{k \to \infty} \int_{A \cap \tilde{B}_k} J(\tilde{f}_k)(x)\, dx.$$

Aplicando el Caso 3 a la función globalmente Lipschitz $\tilde{f}_k$ sobre el $\mathcal{L}^n$-medible $A \cap \tilde{B}_k$ obtenemos que

$$\begin{aligned}
&\lim_k \int_{A \cap \tilde{B}_k} J(\tilde{f}_k)(x)\, dx = \lim_k \int_{\mathbb{R}^m} \mathcal{H}^0\left(A \cap \tilde{B}_k \cap \tilde{f}_k^{\,-1}(y)\right) d\mathcal{H}^n(y) \\
&= \lim_k \int_{\mathbb{R}^m} \mathcal{H}^0\left(A \cap \tilde{B}_k \cap f^{-1}(y)\right) d\mathcal{H}^n(y) = \int_{\mathbb{R}^m} \mathcal{H}^0\left(A \cap f^{-1}(y)\right) d\mathcal{H}^n(y),
\end{aligned}$$

donde la penúltima igualdad se debe a que $f|_{\tilde{B}_k} = \tilde{f}_k|_{\tilde{B}_k}$, y la última al Teorema de la Convergencia Monótona aplicado con las funciones $\mathcal{H}^0\left(A \cap \tilde{B}_k \cap f^{-1}(y)\right) \nearrow \mathcal{H}^0\left(A \cap f^{-1}(y)\right)$. Juntando las dos últimas cadenas de igualdades tenemos lo que queremos:

$$\int_A Jf(x)dx = \int_{\mathbb{R}^m} \mathcal{H}^0\left(A \cap f^{-1}(y)\right) d\mathcal{H}^n(y).$$

Esto concluye la demostración de la fórmula del área. □

Demostración del Corolario 13.6. Haremos sólo el caso $g \geq 0$ (el caso $g \in L^1(\mathbb{R}^n)$ se deduce fácilmente de este poniendo $g = g^+ - g^-$). Sabemos que toda función medible positiva se puede escribir como combinación lineal infinita, con coeficientes positivos, de funciones características de conjuntos medibles (no necesariamente disjuntos), digamos

$$g = \sum_{i=1}^{\infty} c_i \chi_{A_i},$$

donde $c_i \in [0, \infty)$ y A_i es $\mathcal{L}_n$-medible para cada $i \in \mathbb{N}$. Entonces, usando el Teorema de la Convergencia Monótona y la fórmula del área se tiene

$$\begin{aligned}
&\int_{\mathbb{R}^n} g(x)Jf(x)dx = \sum_{i=1}^{\infty} c_i \int_{\mathbb{R}^n} \chi_{A_i}(x)Jf(x)dx = \sum_{i=1}^{\infty} c_i \int_{\mathbb{R}^n} \chi_{A_i}(x)Jf(x)dx \\
&= \sum_{i=1}^{\infty} c_i \int_{\mathbb{R}^m} \mathcal{H}^0(A_i \cap f^{-1}(\{y\}))d\mathcal{H}^n(y) = \int_{\mathbb{R}^m} \sum_{i=1}^{\infty} c_i \mathcal{H}^0(A_i \cap f^{-1}(\{y\}))d\mathcal{H}^n(y) \\
&= \int_{\mathbb{R}^m} \sum_{i=1}^{\infty} c_i \sum_{x \in f^{-1}(\{y\})} \chi_{A_i}(x)d\mathcal{H}^n(y) = \int_{\mathbb{R}^m} \sum_{x \in f^{-1}(\{y\})} \sum_{i=1}^{\infty} c_i \chi_{A_i}(x)d\mathcal{H}^n(y) \\
&= \int_{\mathbb{R}^m} \left(\sum_{x \in f^{-1}(\{y\})} g(x) \right) d\mathcal{H}^n(y).
\end{aligned}$$

□

13.4. La fórmula de la coárea

En esta sección comenzaremos a estudiar la fórmula de la coárea, un resultado que generaliza a la vez el teorema de Fubini y el del cambio de variables.

Empecemos con una versión algo simplificada de la fórmula.

Teorema 13.11 (Fórmula de la coárea, versión 1.0)**.** *Si $m \leq n$, $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ es localmente Lipschitz y $g : f(U) \to [0, \infty]$ es medible, entonces*

$$\int_{f(U)} g(y)\mathcal{H}^{n-m}(f^{-1}(\{y\}))d\mathcal{H}^m(y) = \int_U (g \circ f)Jf(x)dx.$$

Nótese que cuando $n = m$ y f es un difeomorfismo C^1 se obtiene como corolario el teorema del cambio de variables (sin embargo, incluso en este caso $n = m$ el teorema anterior sólo pide que f sea localmente Lipschitz, con lo que es más general).

Por otra parte, cuando $g = 1$, se obtiene:

Corolario 13.12. *Si $m \leq n$ y $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ es localmente Lipschitz entonces*

$$\int_U Jf(x)dx = \int_{f(U)} \mathcal{H}^{n-m}(f^{-1}(\{y\}))d\mathcal{H}^m(y) = \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(f^{-1}(\{y\}))dy,$$

entendiendo que $f^{-1}(\{y\}) = \emptyset$ si $y \notin f(U)$.

La versión más general de la fórmula del área es la siguiente.

Teorema 13.13 (Fórmula de la coárea, versión 2.0)**.** *Si $m \leq n$, $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ es localmente Lipschitz y $g : U \to [0, \infty]$ es medible (o bien $g \in L^1(U)$), entonces*

$$\int_{\mathbb{R}^m} \left(\int_{f^{-1}(\{y\})} g(x)d\mathcal{H}^{n-m}(x) \right) d\mathcal{H}^m(y) = \int_U g(x)Jf(x)dx.$$

La diferencia entre la segunda versión y la primera es que en esta g está definida en $f(U)$ y en la segunda está definida en U. La primera versión de la fórmula se sigue de la segunda cuando la aplicamos con $g \circ f$ en lugar de g.

Obsérvese que cuando $f : \mathbb{R}^n \to \mathbb{R}^m$ es una proyección lineal se tiene

$$(Df)(Df)^* = I,$$

con lo cual $Jf(x) \equiv 1$; además, denotando $\mathbb{R}^n = \mathbb{R}^{n-m} \times \mathbb{R}^m = \{(x,y) : x \in \mathbb{R}^{n-m}, y \in \mathbb{R}^m\}$, se tiene

$$f^{-1}(\{y\}) = \mathbb{R}^{n-m} \times \{y\}.$$

Entonces para cualquier función integrable $g : \mathbb{R}^n \to \mathbb{R}$ se tiene

$$\int_{f^{-1}(\{y\})} g \, d\mathcal{H}^{n-m} = \int_{\mathbb{R}^{n-m}} g(x,y) d\mathcal{H}^{n-m}(x),$$

y por tanto la segunda versión de la fórmula del área nos da

$$\begin{aligned}\int_{\mathbb{R}^n} g \, d\mathcal{H}^n &= \int_{\mathbb{R}^n} gJf d\mathcal{H}^n = \int_{\mathbb{R}^m} \left(\int_{f^{-1}(\{y\})} g d\mathcal{H}^{n-m} \right) d\mathcal{H}^m \\ &= \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^{n-m}} g(x,y) d\mathcal{H}^{n-m}(x) \right) d\mathcal{H}^m(y),\end{aligned}$$

lo que, puesto que $\mathcal{H}^d = \mathcal{L}_d$ en $\mathbb{R}^d$, es el enunciado del teorema de Fubini en $\mathbb{R}^n = \mathbb{R}^{n-m} \times \mathbb{R}^m$.

Es por esto que, en el caso en que $f : \mathbb{R}^n \to \mathbb{R}^m$ es Lipschitz y $n > m$, la fórmula de la coárea puede considerarse una potente generalización curvilínea del teorema de Fubini.[2]

13.4.1. Conjuntos de nivel

En el caso $m = 1$ se obtiene una fórmula con un bonito sentido geométrico:

Corolario 13.14 (Fórmula de la coárea para conjuntos de nivel)**.** *Si $f : U \subset \mathbb{R}^n \to \mathbb{R}$ es Lipschitz, entonces*

$$\int_U |\nabla f(x)| dx = \int_{-\infty}^{\infty} \mathcal{H}^{n-1}\left(f^{-1}(\{t\})\right) dt.$$

Demostración. En efecto, en este caso $(Df)(Df)^* = |\nabla f|^2$, luego $Jf = |\nabla f|$ y aplicando el Corolario 13.12 se obtiene la fórmula del enunciado. □

13.4.2. Coordenadas esféricas

Por otra parte, en el caso de la función $f : \mathbb{R}^n \to \mathbb{R}$ definida por $f(x) = |x|$ (la norma euclídea de x), denotando

$$S^{n-1}(0,t) = \{x \in \mathbb{R}^n : |x| = t\},$$

obtenemos lo siguiente.

Corolario 13.15 (Coordenadas esféricas)**.** *Para toda $g \in L^1(\mathbb{R}^n)$, se tiene*

$$\int_{\mathbb{R}^n} g(x) dx = \int_0^{\infty} \left(\int_{S^{n-1}(0,t)} g(v) d\mathcal{H}^{n-1}(v) \right) dt.$$

Demostración. En efecto, si $f(x) = |x|$, se tiene $f^{-1}(\{t\}) = S^{n-1}(0,t)$ para todo $t > 0$, y $f^{-1}(\{t\}) = \emptyset$ si $t < 0$; además $Jf(x) = |\nabla f(x)| = 1$ para todo $x \in \mathbb{R}^n \setminus \{0\}$. Luego aplicando el teorema (13.13) se obtiene la fórmula del enunciado. □

[2] Curvilínea en el sentido de que los conjuntos $f^{-1}(\{y\})$ ya no necesariamente son subespacios afines de dimensión $n-m$, sino (para casi todo valor de y) subconjuntos de $\mathbb{R}^n$ con dimensión de Hausdorff a lo sumo $n-m$.

13.5. Demostración de la fórmula de la coárea

En toda esta sección suponemos $n \geq m$. Usaremos tres lemas auxiliares. El primero de ellos demuestra el resultado en el caso más simple: cuando f es una aplicación lineal.

Lema 13.16. *Si $L : \mathbb{R}^n \to \mathbb{R}^m$ es lineal y $A \subset \mathbb{R}^n$ es $\mathcal{L}_n$-medible, entonces:*

1. *La función $\mathbb{R}^m \ni y \mapsto \mathcal{H}^{n-m}(A \cap L^{-1}(\{y\}))$ es $\mathcal{L}_m$-medible;*
2. $$\int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap L^{-1}(\{y\}))dy = [[L]]\mathcal{L}_n(A).$$

Demostración. **Caso 1.** Supongamos que $\dim L(\mathbb{R}^n) < m$. Entonces, para $\mathcal{L}_m$-casi todo $y \in \mathbb{R}^m$ se tiene $A \cap L^{-1}(\{y\}) = \emptyset$, y en particular $\mathcal{H}^{n-m}(A \cap L^{-1}(\{y\})) = 0$, con lo cual resulta que la función $y \mapsto \mathcal{H}^{n-m}(A \cap L^{-1}(\{y\}))$ es $\mathcal{L}_m$-medible. Además, si escribimos $L = SO^*$ como en el Teorema 13.2, tenemos $L(\mathbb{R}^n) = S(\mathbb{R}^m)$, luego $\dim S(\mathbb{R}^m) < m$ y por tanto $[[L]] = |\det S| = 0$, con lo que la fórmula del apartado (2) se cumple.

Caso 2. Supongamos ahora que $L = P : \mathbb{R}^n \to \mathbb{R}^m$ es la proyección ortogonal

$$P(x_1, ..., x_n) = (x_1, ..., x_m).$$

Entonces para cada $y \in \mathbb{R}^m$, $P^{-1}(\{y\})$ es un subespacio afín de dimensión $n - m$ de $\mathbb{R}^n$, una traslación de $P^{-1}(\{0\})$. Puesto que $\mathcal{H}^k = \mathcal{L}_k$ en $\mathbb{R}^k$ para todo k (y ambas medidas son invariantes por isometrías), por el Teorema de Fubini deducimos que para casi todo $y \in \mathbb{R}^m$ la sección $A \cap P^{-1}(\{y\}) = y + A_y$ es $\mathcal{L}_{n-m}$-medible, y la aplicación

$$y \mapsto \mathcal{L}_{n-m}(A_y) = \mathcal{H}^{n-m}(A_y) = \mathcal{H}^{n-m}(A \cap P^{-1}(\{y\}))$$

(definida en casi todo punto) es $\mathcal{L}_m$-medible, con

$$\mathcal{L}_n(A) = \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap P^{-1}(\{y\}))dy. \tag{13.9}$$

Caso 3. Supongamos que $L : \mathbb{R}^n \to \mathbb{R}^m$ tiene rango m, es decir, $\dim L(\mathbb{R}^n) = m$. Por el Teorema 13.2,

$$L = SO^*,$$

donde $S : \mathbb{R}^m \to \mathbb{R}^m$ es simétrica y $O : \mathbb{R}^m \to \mathbb{R}^n$ es ortogonal, y $[[L]] = |\det S|$. Sea $Q : \mathbb{R}^n \to \mathbb{R}^n$ una aplicación ortogonal tal que

$$Q^*(x_1, ..., x_m, 0, ..., 0) = O(x_1, ..., x_m)$$

para todo $x \in \mathbb{R}^n$, y observemos que, si $P : \mathbb{R}^m \to \mathbb{R}^n$ es la proyección ortogonal $P(x_1, ..., x_n) := (x_1, ..., x_m)$ entonces

$$P^*(x_1, ..., x_m) = (x_1, ..., x_m, 0, ..., 0) \in \mathbb{R}^n$$

para todo $x \in \mathbb{R}^m$. Luego $O = Q^*P^*$, y por tanto

$$O^* = PQ.$$

Ahora observemos que $L^{-1}(\{0\})$ es un subespacio de dimensión $n - m$ de $\mathbb{R}^n$, y $L^{-1}(\{y\})$ es una traslación paralela de $L^{-1}(\{0\})$ para cada $y \in \mathbb{R}^m$. Por tanto, por el teorema de Fubini[3] la aplicación $y \mapsto \mathcal{H}^{n-m}(A \cap L^{-1}(\{y\}))$ es medible, y se tiene

$$\mathcal{L}_n(A) = \mathcal{L}_n(Q(A)) = \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(Q(A) \cap P^{-1}(\{y\}))dy = \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap (Q^{-1}P^{-1}(\{y\}))dy,$$

[3]El teorema de Fubini aplicado en un sistema de coordenadas paralelo a $L^{-1}(\{0\})$; recuérdese que la medida de Lebesgue es invariante por isometrías.

donde en la primera y la tercera igualdad hemos usado que Q es una isometría, y en la segunda (13.9). Haciendo el cambio de variable $z = Sy$, $z \in \mathbb{R}^m$, y aplicando la fórmula del área deducimos que

$$\int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap Q^{-1}P^{-1}S^{-1}(\{z\}))dz = \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap Q^{-1}P^{-1}(\{y\}))|\mathrm{det}S|dy = |\mathrm{det}S|\mathcal{L}_n(A).$$

Como $L = SO^* = SPQ$, esto significa que

$$[[L]]\mathcal{L}_n(A) = \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap L^{-1}(\{z\}))dz.$$

□

Ahora veamos que la aplicación que interviene en la fómula de la coárea (además del jacobiano) es medible. Recordemos que ω_k denota $\mathcal{L}_k(B(0,1))$.

Lema 13.17. *Sea $A \subset \mathbb{R}^n$ medible Lebesgue, y $f : \mathbb{R}^n \to \mathbb{R}^m$ Lipschitz. Entonces:*

1. *$A \cap f^{-1}(\{y\})$ es $\mathcal{H}^{n-m}$-medible para $\mathcal{L}_m$-casi todo $y \in \mathbb{R}^m$;*
2. *la función $\mathbb{R}^m \ni y \mapsto \mathcal{H}^{n-m}(A \cap f^{-1}(\{y\}))$ es $\mathcal{L}_m$-medible, y*
3. $$\int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap f^{-1}(\{y\}))dy \leq \frac{\omega_{n-m}\,\omega_m}{\omega_n}\,(Lip(f))^m\,\mathcal{L}_n(A).$$

Demostración. **Caso 1.** Supongamos A compacto. Entonces (1) se cumple porque $A \cap f^{-1}(\{y\})$ es cerrado para todo $y \in \mathbb{R}^m$. Fijemos $t \geq 0$, y para cada $i \in \mathbb{N}$ sea U_j el conjunto de los puntos $y \in \mathbb{R}^m$ para los que existen cantidades finitas de conjuntos abiertos $S_1, ..., S_\ell$ tales que

$$A \cap f^{-1}(\{y\}) \subset \bigcup_{j=1}^{\ell} S_j;$$
$$\mathrm{diam}(S_j) \leq \frac{1}{i} \text{ para cada } j = 1, ..., \ell, \text{ y}$$
$$\sum_{j=1}^{\ell} \omega_{n-m}\left(\frac{\mathrm{diam}(S_j)}{2}\right)^{n-m} \leq t + \frac{1}{i}.$$

Veamos que dicho conjunto U_i es abierto para cada $i \in \mathbb{N}$. En efecto, si y cumple las propiedades anteriores, entonces $A \cap f^{-1}(\{z\}) \subset \bigcup_{j=1}^{\ell} S_j$ para todo z que esté suficientemente cerca de y, porque de lo contrario existirían una sucesión $z_k \to y$ y puntos $x_k \in A \cap f^{-1}(\{z_k\})$ tales que $(x_k)_{k\in\mathbb{N}} \subset \mathbb{R}^n \setminus \bigcup_{j=1}^{\ell} S_j$; al ser A compacto podemos suponer, salvo tomar una subsucesión, que $\lim_{k\to\infty} x_k = x_0$ para algún $x_0 \in A$, y como f es continua se tiene $z_k = f(x_k) \to f(x_0)$, luego, por la unicidad del límite $y = f(x_0)$; pero $x_0 \in \mathbb{R}^n \setminus \bigcup_{j=1}^{\ell} S_j$ por ser este conjunto cerrado, y esto contradice que $A \cap f^{-1}(\{y\}) \subset \bigcup_{j=1}^{\ell} S_j$.

Veamos ahora que el conjunto

$$\{y \in \mathbb{R}^m : \mathcal{H}^{n-m}(A \cap f^{-1}(\{y\})) \leq t\} = \bigcap_{i=1}^{\infty} U_i,$$

y por tanto la función $y \mapsto \mathcal{H}^{n-m}(A \cap f^{-1}(\{y\}))$ es medible Borel. Si $\mathcal{H}^{n-m}(A \cap f^{-1}(\{y\})) \leq t$ entonces, para cada $\delta > 0$ es $\mathcal{H}^{n-m}_\delta(A \cap f^{-1}(\{y\})) \leq t$. Dado $i \in \mathbb{N}$, elijamos $\delta \in (0, 1/i)$. Existen conjuntos S_j, que podemos suponer abiertos, tales que

$$A \cap f^{-1}(\{y\}) \subset \bigcup_{j=1}^{\infty} S_j;$$
$$\operatorname{diam}(S_j) \leq \delta < \frac{1}{i} \text{ para cada } j = 1, ..., \ell, \text{ y}$$
$$\sum_{j=1}^{\ell} \omega_{n-m} \left(\frac{\operatorname{diam}(S_j)}{2} \right)^{n-m} < t + \frac{1}{i}.$$

Como $A \cap f^{-1}(\{y\})$ es compacto, existe $\ell \in \mathbb{N}$ tal que $A \cap f^{-1}(\{y\}) \subset \bigcup_{j=1}^{\ell} S_j$, y por tanto $y \in U_i$. Así pues

$$\{y : \mathcal{H}^{n-m}(A \cap f^{-1}(\{y\})) \leq t\} \subset \bigcap_{i=}^{\infty} U_i.$$

Por otra parte, si $y \in \bigcap_{i=1}^{\infty} U_i$ entonces para todo i es

$$\mathcal{H}^{n-m}_{1/i}(A \cap f^{-1}(\{y\})) \leq t + \frac{1}{i},$$

de donde se deduce que $\mathcal{H}^{n-m}(A \cap f^{-1}(\{y\}) \leq t$. Por tanto

$$\{y : \mathcal{H}^{n-m}(A \cap f^{-1}(\{y\})) \leq t\} = \bigcap_{i=}^{\infty} U_i$$

es un conjunto G_δ, y en particular boreliano.

Caso 2. Supongamos que A es abierto; entonces $A \cap f^{-1}(\{y\})$ es también medible para cada $y \in \mathbb{R}^m$ (por ser intersección de un abierto con un cerrado, ambos conjuntos medibles), y podemos escribir

$$A = \bigcup_{j=1}^{\infty} K_j,$$

donde $K_i \subset K_{i+1}$ son compactos para todo $j \in \mathbb{N}$. Por tanto

$$y \mapsto \mathcal{H}^{n-m}(A \cap f^{-1}(\{y\})) = \lim_{j \to \infty} \mathcal{H}^{n-m}(K_j \cap f^{-1}(\{y\}))$$

es medible Borel por ser límite de funciones medibles Borel.

Caso 3. En todos los casos que ya sepamos que $y \mapsto \mathcal{H}^{n-m}(A \cap f^{-1}(\{y\}))$ es una función medible, veamos que se cumple la propiedad (3) del enunciado. Para cada $j \in \mathbb{N}$, existe una sucesión de bolas cerradas $\{B_i^j\}_{i \in \mathbb{N}}$ tales que

$$A \subset \bigcup_{i=1}^{\infty} B_i^j, \quad , \operatorname{diam}(B_i^j) \leq \frac{1}{j} \quad \text{y} \quad \sum_{i=1}^{\infty} \mathcal{L}_n(B_i^j) \leq \mathcal{L}_n(A) + \frac{1}{j}.$$

Definamos

$$g_i^j := \omega_{n-m} \left(\frac{\operatorname{diam}(B_i^j)}{2} \right)^{n-m} \chi_{f(B_i^j)},$$

que obviamente es una función medible para cada i, j, y observemos que

$$\mathcal{H}^{n-m}_{1/j}(A \cap f^{-1}(\{y\})) \leq \sum_{i=1}^{\infty} g_i^j(y)$$

para todo $y \in \mathbb{R}^m$. Entonces, usando el Lema de Fatou y la desigualdad isodiamétrica, tenemos

$$\begin{aligned}
&\int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap f^{-1}(\{y\}))dy = \int_{\mathbb{R}^m} \lim_{j\to\infty} \mathcal{H}^{n-m}_{1/j}(A \cap f^{-1}(\{y\}))dy \\
&\leq \int_{\mathbb{R}^m} \liminf_{j\to\infty} \sum_{i=1}^{\infty} g_i^j(y) \leq \liminf_{j\to\infty} \sum_{i=1}^{\infty} \int_{\mathbb{R}^m} g_i^j(y)dy \\
&= \liminf_{j\to\infty} \sum_{i=1}^{\infty} \omega_{n-m} \left(\frac{\text{diam}(B_i^j)}{2}\right)^{n-m} \mathcal{L}_m\left(f(B_i^j)\right) \\
&\leq \liminf_{j\to\infty} \sum_{i=1}^{\infty} \omega_{n-m} \left(\frac{\text{diam}(B_i^j)}{2}\right)^{n-m} \omega_m \left(\frac{\text{diam}(f(B_i^j)}{2}\right)^{m} \\
&\leq \frac{\omega_{n-m}\,\omega_m}{\omega_n} (\text{Lip}(f))^m \liminf_{j\to\infty} \sum_{i=}^{\infty} \mathcal{L}_n(B_i^j) \\
&\leq \frac{\omega_{n-m}\,\omega_m}{\omega_n} (\text{Lip}(f))^m \mathcal{L}_n(A).
\end{aligned}$$

Caso 4. Supongamos que $\mathcal{L}_n(A) < \infty$. Sean $\{V_j\}_{j\in\mathbb{N}}$ una sucesión decreciente de abiertos de medidas finitas, y $\{K_j\}_{j\in\mathbb{N}}$ una sucesión creciente de compactos tales que

$$K_j \subset A \subset V_j \quad \text{y} \quad \mathcal{L}_n(V_j \setminus K_j) < 1/j$$

para todo $j \in \mathbb{N}$. Entonces

$$\mathcal{H}^{n-m}(K_j \cap f^{-1}(\{y\})) \leq \mathcal{H}^{n-m}(A \cap f^{-1}(\{y\})) \leq \mathcal{H}^{n-m}(V_j \cap f^{-1}(\{y\})), \tag{13.10}$$

y además la sucesión $\mathcal{H}^{n-m}(V_j \cap f^{-1}(\{y\})) - \mathcal{H}^{n-m}(K_j \cap f^{-1}(\{y\}))$ es monótona decreciente y por tanto tiene límite, para cada $y \in \mathbb{R}^m$; límite que es mayor o igual que cero. Luego, usando el Lema de Fatou y la propiedad (3) del enunciado con $V_j \setminus K_j$ en lugar de A, tenemos

$$\begin{aligned}
&\int_{\mathbb{R}^m} \lim_{j\to\infty} \left(\mathcal{H}^{n-m}(V_j \cap f^{-1}(\{y\})) - \mathcal{H}^{n-m}(K_j \cap f^{-1}(\{y\}))\right) dy \\
&\leq \liminf_{j\to\infty} \int_{\mathbb{R}^m} \left(\mathcal{H}^{n-m}(V_j \cap f^{-1}(\{y\})) - \mathcal{H}^{n-m}(K_j \cap f^{-1}(\{y\}))\right) dy \\
&\leq \limsup_{j\to\infty} \int_{\mathbb{R}^m} \mathcal{H}^{n-m}((V_j \setminus K_j) \cap f^{-1}(\{y\}))dy \\
&\leq \limsup_{j\to\infty} \frac{\omega_{n-m}\,\omega_m}{\omega_n} (\text{Lip}(f))^m \mathcal{L}_n(V_j \setminus K_j) = 0.
\end{aligned}$$

Por tanto se tiene que

$$\lim_{j\to\infty} \left(\mathcal{H}^{n-m}(V_j \cap f^{-1}(\{y\})) - \mathcal{H}^{n-m}(K_j \cap f^{-1}(\{y\}))\right) = 0$$

para casi todo $y \in \mathbb{R}^m$; esto en particular implica que $A \cap f^{-1}(\{y\})$ es $\mathcal{H}^{n-m}$-medible para $\mathcal{L}_m$-casi todo y, y además, combinando con (13.10) y usando el Caso 2 obtenemos que

$$y \mapsto \mathcal{H}^{n-m}(A \cap f^{-1}(\{y\})) = \lim_{j\to\infty} \mathcal{H}^{n-m}(V_j \cap f^{-1}(\{y\}))$$

es $\mathcal{L}_m$-medible por ser límite en casi todo punto de una sucesión de funciones medibles. Por supuesto, una vez que sabemos que esta función es medible, también se cumple la desigualdad (3) del enunciado por lo visto en el Caso anterior.

Caso 5. Supongamos que $\mathcal{L}_n(A) = \infty$. Entonces

$$A \cap f^{-1}(\{y\}) = \bigcup_{k=1}^{\infty} A_k \cap f^{-1}(\{y\}),$$

donde $A_k = A \cap B(0,k)$, $A_k \subset A_{k+1}$ y $\mathcal{L}_n(A_k) < \infty$ para todo k. Por el Caso 4 todos los A_k son $\mathcal{H}^{n-m}$-medibles, luego también lo es su unión. Además, $\{y \mapsto \mathcal{H}^{n-m}(A_k \cap f^{-1}(\{y\}))\}_{k\in\mathbb{N}}$ es una sucesión de funciones $\mathcal{L}_m$-medibles que converge monótonamente a $y \mapsto \mathcal{H}^{n-m}(A \cap f^{-1}(\{y\}))$, luego esta función también es medible. Así pues se cumplen (1) y (2), y por tanto también (3). □

Observación 13.4. Cálculos análogos a los de la demostración de la propiedad (3) del lema anterior muestran que

$$\int_{\mathbb{R}^m} \mathcal{H}^k(A \cap f^{-1}(\{y\}))d\mathcal{H}^\ell(y) \leq \frac{\omega_k\,\omega_\ell}{\omega_{k+\ell}}\,(\mathrm{Lip}(f))^\ell\,\mathcal{H}^{k+\ell}(A)$$

para todo boreliano $A \subset \mathbb{R}^n$, o más en general para todo conjunto A tal que $y \mapsto \mathcal{H}^k(A \cap f^{-1}(\{y\}))$ sea $\mathcal{H}^\ell$-medible.

Lema 13.18. *Sean $t > 1$, $h : \mathbb{R}^n \to \mathbb{R}^n$ Lipschitz, y consideremos*

$$B := \{x \in \mathbb{R}^n : Dh(x) \text{ existe y } Jh(x) > 0\}.$$

Entonces existe una sucesión de conjuntos borelianos $\{D_k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ tales que:

1. *$\mathcal{L}_n\left(B \setminus \bigcup_{k=1}^\infty D_k\right) = 0$;*
2. *$h|_{D_k}$ es inyectiva para todo $k \in \mathbb{N}$, y*
3. *Para cada $k \in \mathbb{N}$ existe un automorfismo lineal simétrico $S_k : \mathbb{R}^n \to \mathbb{R}^n$ tal que*

$$\mathit{Lip}\left(S_k^{-1} \circ (h|_{D_k})\right) \leq t,\ \ \mathit{Lip}\left((h|_{D_k})^{-1} \circ S_k\right) \leq t,$$

$$t^{-n}|\mathit{det}(S_k)| \leq Jh|_{D_k} \leq t^n|\mathit{det}(S_k)|.$$

Demostración. Por el Lema 13.10 existe una sucesión $\{E_k\}_{k\in\mathbb{N}}$ de borelianos de $\mathbb{R}^n$ y una sucesión $\{T_k\}_{k\in\mathbb{N}}$ de automorfismos lineales simétricos de $\mathbb{R}^n$ tales que

(a) $B = \bigcup_{k=1}^\infty E_k$;

(b) $h|_{E_k}$ es inyectiva para todo k, y

(c) para todo k,

$$\mathrm{Lip}\left((h|_{E_k}) \circ T_k^{-1}\right) \leq t,\ \ \mathrm{Lip}\left(T_k \circ (h|_{E_k})^{-1}\right) \leq t,$$

$$t^{-n}|\det(T_k)| \leq Jh|_{E_k} \leq t^n|\det(T_k)|.$$

Se sigue de (c) que $(h|_{E_k})^{-1}$ es Lipschitz, y por tanto, usando la fórmula de McShane (o el teorema de Kirszbraun) existe $h_k : \mathbb{R}^n \to \mathbb{R}^n$ Lipschitz y tal que $h_k = (h|_{E_k})^{-1}$ en $h(E_k)$.

Veamos que $Jh_k(x) > 0$ en $\mathcal{L}_n$-casi todo punto $x \in h(E_k)$. En efecto, como $h_k \circ h(x) = x$ para todo $x \in E_k$, por el Teorema 11.26 se tiene

$$Dh_k(h(x)) \circ Dh(x) = 1 \quad \text{para } \mathcal{L}_n\text{-casi todo } x \in E_k,$$

y por tanto

$$Jh_k(h(x))\,Jh(x) = 1 \quad \text{para } \mathcal{L}_n\text{-casi todo } x \in E_k.$$

En vista de (c), esto implica que $Jh_k(h(x)) > 0$ para casi todo $x \in E_k$, y como h es Lipschitz (y por ello lleva conjuntos de medida cero en conjuntos de medida cero) también resulta que $Jh_k(y) > 0$ para casi todo $y \in h(E_k)$.

Entonces podemos aplicar el Lema 13.10 a la función h_k en E_k, para obtener una sucesión $\{F_j^k\}_{j\in\mathbb{N}}$ de borelianos de $\mathbb{R}^n$ y una sucesión $\{R_j^k\}_{j\in\mathbb{N}}$ de automorfismos simétricos tales que

(d) $\mathcal{L}_n\left(h(E_k)\setminus\bigcup_{j=1}^{\infty}F_j^k\right)=0$;

(e) $h_k|_{F_j^k}$ es inyectiva, y

(f) Para todos $k,j\in\mathbb{N}$,

$$\text{Lip}\left((h_k|_{F_j^k})\circ(R_j^k)^{-1}\right)\leq t,\;\; \text{Lip}\left((R_j^k)\circ(h_k|_{F_j^k})^{-1}\right)\leq t,$$

$$t^{-n}|\text{det}(R_j^k)|\leq Jh_k|_{F_j^k}\leq t^n|\text{det}(R_j^k)|.$$

Definamos

$$D_j^k:=E_k\cap h^{-1}(F_j^k),\;\; S_j^k:=(R_j^k)^{-1}$$

para cada $k,j\in\mathbb{N}$, y veamos que

$$\mathcal{L}_n\left(B\setminus\bigcup_{k,j=1}^{\infty}D_j^k\right)=0.$$

En efecto,

$$h_k\left(h(E_k)\setminus\bigcup_{j=1}^{\infty}F_j^k\right)=h^{-1}\left(h(E_k)\setminus\bigcup_{j=1}^{\infty}F_j^k\right)=E_k\setminus\bigcup_{j=1}^{\infty}D_j^k,$$

luego, usando la propiedad (d) y la lipschitzianidad de h_k, deducimos que

$$\mathcal{L}_n\left(E_k\setminus\bigcup_{j=1}^{\infty}D_j^k\right)=0$$

para todo k, lo que combinado con (a) nos da que $\mathcal{L}_n\left(B\setminus\bigcup_{k,j=1}^{\infty}D_j^k\right)=0$. Si renombramos la familia numerable conjuntos $\{D_j^k\}_{j,k\in\mathbb{N}}$ en forma de una sucesión $\{D_i\}_{i\in\mathbb{N}}$, esto prueba (1)

Es obvio que (b) implica que $h|_{D_j^k}$ es inyectiva para todo $j,k\in\mathbb{N}$, y por tanto se cumple también (2), salvo renombración de los conjuntos D_j^k.

Sólo falta probar (3); hagámoslo. Tenemos que

$$\text{Lip}\left((S_j^k)^{-1}\circ(h|_{D_j^k})\right)=\text{Lip}\left(R_j^k\circ(h|_{D_j^k})\right)\leq\text{Lip}\left(R_j^k\circ(h_k|_{F_j^k})^{-1}\right)\leq t$$

por la propiedad (f). Análogamente,

$$\text{Lip}\left((h|_{D_j^k})^{-1}\circ S_j^k\right)=\text{Lip}\left((h|_{D_j^k})^{-1}\circ(R_j^k)^{-1}\right)\leq\text{Lip}\left((h_k|_{F_j^k})\circ(R_j^k)^{-1}\right)\leq t,$$

y así tenemos las dos primeras desigualdades de (3). Finalmente, usando (f) y el hecho de que, como hemos observado antes,

$$Jh_k(h(x))Jh(x)=1\;\;\text{para casi todo } x\in D_j^k,$$

obtenemos que

$$t^{-n}|\text{det}(S_j^k)|=t^{-n}|\text{det}(R_j^k)|^{-1}\leq Jh|_{D_j^k}\leq t^n|\text{det}(R_j^k)|^{-1}=t^n|\text{det}(S_j^k)|,$$

y esto completa la demostración del lema. □

Ahora ya estamos preparados para probar la fórmula de la coárea. Empezaremos demostrando el caso relativamente simple, pero esencialmente equivalente, en el que g es la función característica de un conjunto A y $f:\mathbb{R}^n\to\mathbb{R}^m$ es (globalmente) Lipschitz, es decir, el siguiente resultado.

Teorema 13.19 (Fórmula de la coárea)**.** *Si $n \geq m$ y $f : \mathbb{R}^n \to \mathbb{R}^m$ es Lipschitz entonces, para todo $A \subset \mathbb{R}^n$ medible Lebesgue se tiene que*

$$\int_A Jf(x)dx = \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap f^{-1}(\{y\}))dy. \tag{13.11}$$

Demostración. Gracias al Teorema de Rademacher y al Lema 13.17, y salvo quitar un conjunto de medida cero (lo que, por la propiedad (3) de este lema, no va a afectar a la igualdad que queremos probar), podemos suponer que f es diferenciable en todos los puntos de A, y en particular que $Df(x)$ y $J(x)$ existen en cada punto de A. También podemos suponer que $\mathcal{L}_n(A) < \infty$, porque una vez probado el teorema para este tipo de conjuntos escribiendo A como unión de una sucesión creciente de conjuntos A_k de medida finita y usando el Teorema de la Convergencia Monótona, de la igualdad

$$\int_{A_k} Jf(x)dx = \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A_k \cap f^{-1}(\{y\}))dy$$

para todo k se obtiene en el límite la validez de (13.11) incluso si $\mathcal{L}_n(A) = \infty$.

Caso 1. Supongamos que $A \subset \{x : Jf(x) > 0\}$. Para cada $\lambda \in \Lambda(n, n-m)$, escribamos

$$f = q \circ h_\lambda,$$

donde $h_\lambda : \mathbb{R}^n \to \mathbb{R}^m \times \mathbb{R}^{n-m}$ y $q : \mathbb{R}^m \times \mathbb{R}^{n-m} \to \mathbb{R}^m$ están definidas por

$$h_\lambda(x) := (f(x), P_\lambda(x)),$$

$$q(y, z) := y,$$

y $P_\lambda : \mathbb{R}^n \to \mathbb{R}^{n-m}$ es la proyección ortogonal definida por $P_\lambda(x_1, ..., x_n) = (x_{\lambda(1)}, ..., x_{\lambda(n-m)})$. Definamos también

$$A_\lambda := \{x \in A : \det(Dh_\lambda) \neq 0\}.$$

Por la fórmula de Cauchy-Binet,

$$0 < |Jf|^2 = [[(Df)^*]]^2 = \left\{ \begin{array}{c} \text{suma de los cuadrados de los subdeterminantes} \\ m \times m \text{ de la matriz de } (Df)^* \end{array} \right\} = \sum_{\lambda \in \Lambda(n,n-m)} |Jh_\lambda|^2,$$

y por tanto

$$A = \bigcup_{\lambda \in \Lambda(n,n-m)} A_\lambda = \bigcup_{j=1}^{N} A_{\lambda_j},$$

donde $N = \binom{n}{m}$ y $\{\lambda_1, ..., \lambda_N\}$ es una ordenación de $\Lambda(n, n-m)$. Luego, escribiendo $B_1 = A_{\lambda_1}$ y $B_k = A_{\lambda_k} \setminus \bigcup_{j=1}^{k-1} A_{\lambda_j}$ si $k \geq 2$, se tiene

$$A = \bigcup_{j=1}^{N} B_j \quad \text{unión disjunta, con } B_j \subset A_{\lambda_j},$$

y por tanto basta probar el resultado para cada B_j. En consecuencia podemos suponer sin pérdida de generalidad que $A = A_\lambda$ para algún $\lambda \in \Lambda(n, n-m)$.

Fijemos $t > 1$ y apliquemos el Lema 13.5 a la función $h = h_\lambda$ para obtener una colección $\{D_k\}_{k \in \mathbb{N}}$ de borelianos disjuntos y una sucesión $\{S_k\}_{k \in \mathbb{N}}$ de automorfismos lineales simétricos que tienen las propiedades (1)-(3) de dicho lema. Definamos

$$G_k = A \cap D_k$$

para cada $k \in \mathbb{N}$.

Aserto 13.20. *Se tiene* $t^{-n}[[qS_k]] \leq Jf|_{G_k} \leq t^n[[qS_k]]$ *para todo* k.

Demostración. Como $f = q \circ h$, se tiene, en $\mathcal{L}_n$-casi todo punto

$$\begin{aligned} Df &= q \circ Dh = qS_kS_k^{-1} \circ Dh \\ &= qS_k \circ D(S_k^{-1} \circ h) = qS_k \circ C, \end{aligned}$$

donde $C := D(S_k^{-1} \circ h)$. Por el Lema 13.5,

$$t^{-1} \leq \mathrm{Lip}(S_k^{-1} \circ h) = \mathrm{Lip}(C) \leq t \quad \text{en } G_k. \tag{13.12}$$

Ahora escribamos

$$Df(x) = S(x) \circ O(x)^*, \quad q \circ S_k = TP^*,$$

donde $S(x), T : \mathbb{R}^m \to \mathbb{R}^m$ son simétricas y $O(x), P : \mathbb{R}^m \to \mathbb{R}^n$ son ortogonales para todo $x \in G_k$. Tenemos entonces

$$S \circ O^* = T \circ P^* \circ C, \tag{13.13}$$

y en consecuencia

$$S = T \circ P^* \circ C \circ O.$$

Como $G_k \subset A \subset \{x : Jf(x) > 0\} \cap \{x : Jh(x) \neq 0\}$, se tiene $\det(S) \neq 0$, y por tanto también $\det(T) \neq 0$. Luego, para todo $v \in \mathbb{R}^m$,

$$|T^{-1}S(v)| = |P^* \circ C \circ O(v)| \leq |C \circ O(v)| \leq t|O(v)| = t|v|,$$

donde en la segunda desigualdad hemos usado (13.12). Así pues

$$(T^{-1} \circ S)(B(0,1)) \subset B(0,t),$$

y en consecuencia

$$Jf = |\det(S)| \leq t^n |\det(T)| = t^n[[qS_k]].$$

Análogamente, para todo $v \in \mathbb{R}^n$, de (13.12) y (13.13) deducimos que

$$|S^{-1} \circ T(v)| = |O^* \circ C^{-1} \circ P(v)| \leq |C^{-1} \circ P(v)| \leq t|P(v)| = t|v|,$$

y por tanto

$$[[qS_k]] = |\det(T)| \leq t^n |\det(S)| = t^n Jf.$$

□

Ahora podemos estimar, usando las desigualdades anteriores y el Lema 13.16,

$$\begin{aligned} & t^{-3n+m} \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(G_k \cap f^{-1}(\{y\}))dy = t^{-3n+m} \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(h^{-1}(h(G_k) \cap q^{-1}(\{y\})))dy \\ & \leq t^{-2n} \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(S_k^{-1}(h(G_k) \cap q^{-1}(\{y\})))dy \\ & = t^{-2n} \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(S_k^{-1} \circ h(G_k) \cap (q \circ S_k)^{-1}(\{y\}))dy \\ & = t^{-2n}[[qS_k]]\mathcal{L}_n(S_k^{-1} \circ h(G_k)) \leq t^{-n}[[qS_k]]\mathcal{L}_n(G_k) \\ & \leq \int_{G_k} Jf(x)dx \leq t^n[[qS_k]]\mathcal{L}_n(G_k) \leq t^{2n}\mathcal{L}_n(S_k^{-1} \circ h(G_k)) \\ & = t^{2n} \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(S_k^{-1} \circ h(G_k) \cap (q \circ S_k)^{-1}(\{y\})))dy \\ & \leq t^{3n-m} \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(h^{-1}(h(G_k) \cap q^{-1}(\{y\})))dy = t^{3n-m} \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(G_k \cap f^{-1}(\{y\}))dy. \end{aligned}$$

Es decir,

$$t^{-3n+m}\int_{\mathbb{R}^m}\mathcal{H}^{n-m}(G_k\cap f^{-1}(\{y\}))dy \le \int_{G_k} Jf(x)dx \le t^{3n-m}\int_{\mathbb{R}^m}\mathcal{H}^{n-m}(G_k\cap f^{-1}(\{y\}))dy.$$

Puesto que $\mathcal{L}_n(A\setminus\bigcup_{k=1}^{\infty}G_k)=0$, podemos sumar en $k\in\mathbb{N}$ las desigualdades anteriores, usar la desigualdad (3) del Lema 13.17, y tomar límites cuando $t\to 1^+$ para concluir que

$$\int_{\mathbb{R}^m}\mathcal{H}^{n-m}(A\cap f^{-1}(\{y\}))dy = \int_A Jf(x)dx.$$

Caso 2. Supongamos ahora que $A\subset\{x : Jf(x)=0\}$. Fijemos $0<\varepsilon\le 1$, y definamos

$$g(x,y) := f(x)+\varepsilon y,\;\; p(x,y) := y \quad \text{para } x\in\mathbb{R}^n, y\in\mathbb{R}^m.$$

Entonces

$$Dg = (Df, \varepsilon I),$$

matriz de tamaño $m\times(n+m)$, y

$$\varepsilon^m \le Jg = [[Dg]] = [[(Dg)^*]] \le C\varepsilon,$$

donde $C>0$ es una constante que sólo depende de n, m y $\mathrm{Lip}(f)$.

Aserto 13.21. *Si $y, u\in\mathbb{R}^m$ y $B := A\times B(0,1)\subset\mathbb{R}^{n+m}$, entonces*

$$B\cap g^{-1}(\{y\})\cap p^{-1}(\{u\}) = \begin{cases}\emptyset & \text{si } u\notin B(0,1),\\ (A\cap f^{-1}(\{y-\varepsilon u\}))\times\{u\} & \text{si } u\in B(0,1).\end{cases}$$

Demostración. Se tiene $(x,z)\in B\cap g^{-1}(\{y\})\cap p^{-1}(\{u\})$ si y sólo si

$$x\in A,\; z\in B(0,1),\; f(x)+\varepsilon z = y,\; z=u$$

si y sólo si

$$x\in A,\; z=u\in B(0,1)\; f(x) = y-\varepsilon u$$

si y sólo si

$$u\in B(0,1),\; (x,z)\in(A\cap f^{-1}(\{y-\varepsilon u\}))\times\{u\}.$$

□

Usando el aserto, la Observación 13.4 (con p en lugar de f), y lo probado en el Caso 1 (con g en lugar de f), tenemos que

$$\int_{\mathbb{R}^m}\mathcal{H}^{n-m}(A\cap f^{-1}(\{y\}))dy = \int_{\mathbb{R}^m}\mathcal{H}^{n-m}(A\cap f^{-1}(\{y-\varepsilon u\})dy \quad \text{para todo } u\in\mathbb{R}^m$$
$$= \frac{1}{\omega_m}\int_{B(0,1)}\int_{\mathbb{R}^m}\mathcal{H}^{n-m}(A\cap f^{-1}(\{y-\varepsilon u\}))dydu$$
$$= \frac{1}{\omega_m}\int_{\mathbb{R}^m}\int_{\mathbb{R}^m}\mathcal{H}^{n-m}(B\cap g^{-1}(\{y\})\cap p^{-1}(\{u\}))dudy \le \frac{\omega_{n-m}}{\omega_n}\int_{\mathbb{R}^m}\mathcal{H}^n(B\cap g^{-1}(\{y\})dy$$
$$= \frac{\omega_{n-m}}{\omega_n}\int_B Jg(x,z)dxdz \le \frac{\omega_{n-m}\,\omega_m}{\omega_n}\sup_B Jg\mathcal{L}_n(A) \le C\varepsilon\mathcal{L}_n(A).$$

Haciendo $\varepsilon\to 0$ en la desigualdad anterior, deducimos que

$$\int_{\mathbb{R}^m}\mathcal{H}^{n-m}(A\cap f^{-1}(\{y\}))dy = 0 = \int_A Jf(x)dx,$$

lo que concluye la prueba en el Caso 2.

Caso 3. En el caso general tenemos $A=A_1\cup A_2$, donde $A_1 := \{x\in A : Jf(x)>0\}$ y $A_2 := \{x\in A : Jf(x)=0\}$ son disjuntos. Aplicando los Casos 1 y 2 a A_1 y A_2 respectivamente, y sumando las igualdades resultantes, se obtiene la fórmula (13.11). □

Demostremos ahora la siguiente versión más general de la fórmula de la coárea.

Teorema 13.22. *Si $m \leq n$, $f : \mathbb{R}^n \to \mathbb{R}^m$ es Lipschitz y $g : \mathbb{R}^n \to [0,\infty]$ es medible (o bien $g \in L^1(\mathbb{R}^n)$), entonces $g|_{f^{-1}(\{y\})}$ es $\mathcal{H}^{n-m}$-medible para $\mathcal{L}_n$-casi todo $y \in \mathbb{R}^m$, y*

$$\int_{\mathbb{R}^m} \left(\int_{f^{-1}(\{y\})} g(x) d\mathcal{H}^{n-m}(x) \right) dy = \int_U g(x) Jf(x) dx. \tag{13.14}$$

Demostración. Nótese que el teorema anterior (junto con el Lema 13.17 que nos asegura la medibilidad de los conjuntos y funciones involucrados en la fórmula de su enunciado) equivale a este cuando g es la función característica de un conjunto medible Lebesgue $A \subset \mathbb{R}^n$. En particular, en este caso la sección $g|_{f^{-1}(\{y\})}$ es $\mathcal{H}^{n-m}$-medible para $\mathcal{L}_n$-casi todo $y \in \mathbb{R}^m$.

Caso 1. Supongamos $g \geq 0$ medible. Entonces g es límite puntual de una sucesión de funciones simples medibles positivas $(s_j)_{j\in\mathbb{N}}$. Para cada s_j podemos usar la linealidad de la integral y el hecho de que el teorema es verdad para las funciones características de los conjuntos $\mathcal{L}_n$ medibles para concluir que

$$\int_{\mathbb{R}^m} \left(\int_{f^{-1}(\{y\})} s_j(x) d\mathcal{H}^{n-m}(x) \right) dy = \int_U s_j(x) Jf(x) dx.$$

Entonces, tomando límite cuando $j \to \infty$ y usando el Teorema de la Convergencia Monótona se deduce la fórmula (13.14) (la $\mathcal{H}^{n-m}$-medibilidad de la sección $g|_{f^{-1}(\{y\})}$ para $\mathcal{L}_n$-casi todo $y \in \mathbb{R}^m$ se deduce del hecho de que es límite puntual de una de sucesión formada por combinaciones lineales finitas de funciones con esta propiedad).

Caso 2. Si $f \in L^1(\mathbb{R}^n)$, puede escribirse $g = g^+ - g^-$, aplicar el Caso 1 a g^+ y g^- y sumar las igualdades resultantes para obtener (13.14) en este caso. □

Finalmente, la versión para f localmente Lipschitz de la fórmula de la coárea dada en el Teorema 13.13 se deduce del Teorema 13.22 mediante un argumento análogo al de la demostración de la fórmula del área para f localmente Lipschitz. Los detalles se dejan como ejercicio para el lector. □

13.6. Problemas

Problema 13.1. Para cada $n \in \mathbb{N}$ sea E_n el conjunto $\{k/2^n : k = 0, 1, ..., 2^n\} \cup (\mathbb{R} \setminus [0,1])$. Definamos $f_n(x) = d(x, E_n)$, y $f : [0,1] \to [0,1]$ por $f = \sum_{n=1}^{\infty} 2^{-n} f_n$. Demostrar que f es de Lipschitz, y usando la fórmula del área calcular $\mathcal{H}^1(G_f)$, donde G_f denota la gráfica de f en $\mathbb{R}^2$.

Problema 13.2. Si $f : [0,\infty) \to \mathbb{R}$ es Lipschitz, calcular

$$\mathcal{H}^2 \left(\{(x,y,z) \in \mathbb{R}^3 : z = f(x^2+y^2),\, x^2+y^2 \leq a^2\} \right).$$

Problema 13.3. Sea $B_n(0,r)$ la bola euclídea de centro 0 y radio r en $\mathbb{R}^n$. Sin calcularla explícitamente, demostrar que la función $[0,\infty) \ni r \mapsto \mathcal{H}^{n-1}(\partial B_n(0,r))$ es continua.

Problema 13.4. Sea $F(r) = |B(0,r)| = \omega_n r^n$, donde $B(0,r)$ es la bola euclídea de centro 0 y radio r en $\mathbb{R}^n$. Usando la fórmula de la coárea con la función $f(x) = |x|$ en el conjunto $f^{-1}((r, r+t))$, demostrar que $\mathcal{H}^{n-1}(\partial B(0,r)) = F'(r) = n\omega_n r^{n-1}$.

Problema 13.5. Dada cualquier función localmente Lipschitz $f : \mathbb{R}^n \to \mathbb{R}$ tal que $\lim_{|x|\to\infty} f(x) = \infty$, generalizar la estrategia del ejercicio anterior para hallar una fórmula para $\mathcal{H}^{n-1}(f^{-1}(\{r\}))$ en casi todo $r \in \mathbb{R}$.

Problema 13.6. Dar un ejemplo que pruebe que, incluso suponiendo $f \in C^\infty(\mathbb{R}^n, \mathbb{R}^m)$ la función $\mathbb{R}^m \ni y \mapsto \mathcal{H}^{n-m}(f^{-1}(\{y\}))$ no tiene por qué ser continua (aunque por supuesto es medible gracias al Lema 13.17).

Problema 13.7. Si $g \in L^1(\mathbb{R}^n)$, $f : \mathbb{R}^n \to \mathbb{R}$ es Lipschitz, y existe $a > 0$ tal que $|\nabla f(x)| \geq a$ para casi todo $x \in \mathbb{R}^n$, probar que

$$\int_{\{f>t\}} g(x)dx = \int_t^\infty \left(\int_{\{f=s\}} \frac{g(u)}{|\nabla f(u)|} d\mathcal{H}^{n-1}(u) \right) ds,$$

y deducir que

$$\frac{d}{dt}\left(\int_{\{f>t\}} g \right) = -\int_{\{f=t\}} \frac{g}{|\nabla f|} d\mathcal{H}^{n-1}$$

para casi todo $t \in \mathbb{R}$.

Indicación: usar la fórmula de la coárea (por ejemplo en su versión del Teorema 13.22).

Bibliografía

La originalidad que puede tener un manual de estas características es escasa. En este caso se reduce a variaciones en pruebas y enunciados, a la selección y el orden de exposición de ciertos contenidos, y algunos ejercicios. Las referencias que hemos manejado con mayor intensidad durante la preparación de este texto son las siguientes: [4] para los capítulos 1–7 y 10; [5] para los capítulos 8 y 11; [6] para los capítulos 11 y 12, y [2] para el capítulo 13. Muchos de los ejercicios se han tomado de [6, 5, 1]; agradezco también a Javier Gómez Gil y Mar Jiménez Sevilla los enuciados de otros tantos ejercicios.

[1] D.L. Cohn, *Measure Theory*, Birkhäuser, 1980.

[2] L.C. Evans y R.F. Gariepy, *Measure Theory and Fine Properties of Functions*, Studies in Advanced Mathematics. Taylor & Francis, 1991.

[3] J. Gómez Gil, *Teoría de la Medida*, Universidad Complutense de Madrid, 2022.

[4] P. Hajlasz, *Analysis I*, disponible en https://sites.pitt.edu/ hajlasz/Notatki/Analysis

[5] W. Rudin, *Real and complex analysis*. Third edition. McGraw-Hill Book Co., New York, 1987.

[6] E.M. Stein, R. Shakarchi, *Real Analysis*. Princeton Lectures in Analysis, 3. Princeton University Press, Princeton, NJ, 2005.