ISSN 1851-1317

Fascículo 4 | Cursos de grado

Norberto Fava

Felipe Zó

Medida e integral de Lebesgue

Departamento de Matemática

Facultad de Ciencias Exactas y Naturales

Universidad de Buenos Aires

2013

Cursos de grado

Fascículo 4

ISSN 1851-1317 (Versión Electrónica)
ISSN 1851-1295 (Versión Impresa)

Departamento de Matemática
Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires
Ciudad Universitaria – Pabellón I
(1428) Ciudad de Buenos Aires
Argentina.
http://www.dm.uba.ar
e-mail. secre@dm.uba.ar
tel/fax: (+54-11)-4576-3335

INDICE GENERAL

Capítulo III. **Medida de Lebesgue**

Capítulo IV. **Funciones medibles**

Capítulo V. **Integral de Lebesgue**

Capítulo VI. **Cambio de variables**

Capítulo VII. **Espacios de funciones clásicos**

Capítulo VIII. **Diferenciación de la integral**

Capítulo IX. **El teorema de Radon-Nikodym**

PRESENTACION

El concepto moderno de integral, construido por Lebesgue a principios de este siglo sobre la base de una profunda revisión de la teoría de la medida, ha permitido superar las dificultades que presentaba la noción clásica de integral debida a Cauchy y Riemann (la misma que se estudia en los primeros cursos de "Cálculo" o Análisis Elemental).

Las ventajas del método de Lebesgue son tan evidentes, y su uso se encuentra tan difundido, que uno se siente tentado de aceptar literalmente la opinión de J. Dieudonné:

"Es razonable creer que la integral de Riemann, si no fuera por el prestigioso nombre que la acompaña, habría desaparecido desde hace tiempo de la enseñanza universitaria o –con el debido respeto al genio de Riemann– habría quedado relegada a la categoría de un modesto ejercicio dentro de la teoría general de la medida y la integración".

El presente libro está destinado a quienes deseen iniciarse en la teoría de la medida y la integral de Lebesgue con vistas a sus aplicaciones: Series e Integrales de Fourier, Ecuaciones Diferenciales, Teoría de Probabilidades, etc.

Hemos eludido en lo posible todos los aspectos que no sean "instrumentales": el libro está escrito para que pueda ser leído con facilidad por estudiantes de Matemática, Física o Ingeniería que necesiten manejar el concepto de integral con la soltura y generalidad que requiere el Análisis moderno.

Después de haber enseñado el tema bajo formas diversas, nos ha parecido que la más conveniente para quienes deben estudiarlo por vez primera es la presentación intuitiva que sólo puede hacerse en los espacios euclidianos, postergando la inevitable generalización para cuando se hayan comprendido las ideas fundamentales y se advierta su necesidad. Tradición que reconoce como insignes antecedentes los cursos de Análisis Real que durante varios años tuvieron a su cargo Antoni Zygmund en la Universidad de Chicago y Mischa Cotlar en la de Buenos Aires.

En relación con dicha forma de presentación, confiamos en haber superado el engorro de algunas demostraciones que no satisfacían los criterios normales de rigor (teorema 3.1 y proposiciones 2.19 y 2.21). Dificultad que llevó a veces a preferir el planteo abstracto, aparentemente más "limpio".

La introducción de las normas y los espacios de Orlicz ayudan a comprender el papel de la convexidad en ciertos espacios de funciones y presentan el marco adecuado para la solución de algunos problemas del "Análisis Real", del que forma una parte importante el contenido de esta obra.

La relación entre la teoría de la integral y la teoría de la diferenciación ha sido tratada con especial atención: "El mérito de Lebesgue –ha escrito S. Saks– no consiste solamente en haber creado una nueva y más general noción de integral, ni siquiera en haber establecido su íntima conexión con la teoría de la medida. El valor de su trabajo consiste primordialmente en su teoría de la derivación, que es paralela a la de integración. Esto permitió a su descubrimiento hallar numerosas aplicaciones en las más diversas ramas del Análisis y, desde el punto de vista metodológico, hizo posible unificar dos concepciones fundamentales de la integral, a saber, la de integral definida y la de primitiva, que parecen quedar definitivamente separadas tan pronto como la integración se sale del dominio de las funciones continuas".

Los Autores

Buenos Aires, 1995

PLAN DE LA OBRA

Los dos primeros capítulos proveen los conceptos necesarios sobre la teoría de conjuntos y la topología de los espacios euclidianos. Las nociones de intervalo y conjunto elemental se desarrollan en el segundo capítulo –párrafos 6, 8 y 9–, únicas secciones de dicho capítulo cuya lectura no conviene omitir.

El tema principal del libro comienza en el tercer capítulo, a partir del cual se construyen los conceptos de la teoría de Lebesgue: medida exterior y conjuntos medibles, medida, funciones medibles, integral, teoremas de paso al límite, teoremas de Tonelli y Fubini y fórmula del cambio de variables (capítulos III al VI).

A partir del capítulo VII la lectura es más exigente, porque se estudian los aspectos más profundos y generales de la teoría: los espacios de funciones; la diferenciación de integrales y de medidas; las medidas en los espacios abstractos y el teorema de Radon-Nikodym.

Los conceptos de medida e integral en espacios abstractos son indispensables en capítulos importantes del Análisis y la Teoría de Probabilidades. No es posible, por lo tanto, instalarse definitivamente en los espacios euclidianos, aunque en ellos todavía sean posibles algunas generalizaciones útiles.

Los párrafos cuyo título se señala con una estrella pueden omitirse sin perjuicio en primera lectura.

AGRADECIMIENTOS

Los autores desean agradecer el estímulo que les ha brindado el Director del Instituto Argentino de Matemática, Dr. Carlos Segovia Fernández.

El texto y los dibujos han estado a cargo de la Sra. Blanca Alvarez, quien con su dedicación y eficacia hizo posible que esta obra llegara por fin a concluirse, por lo que los autores desean expresarle su agradecimiento.

CAPITULO I

INTRODUCCION

1. Números reales.

Supondremos al lector familiarizado con las propiedades algebraicas y de orden del cuerpo $I\!R$ formado por todos los números reales; en particular, con el hecho de que todo conjunto no vacío y acotado superiormente $A \subset I\!R$ posee una cota superior mínima llamada el **extremo superior** o bien el **supremo** de A y denotada comúnmente por el símbolo $\sup A$. Simétricamente, si el conjunto $A \subset I\!R$ es no vacío y acotado inferiormente, entonces entre sus cotas inferiores hay una que es máxima, llamada el **extremo inferior** o **ínfimo** de A y denotada por $\inf A$.

El número real s es el supremo de A si y sólo si 1°) cada $x \in A$ verifica $x \leq s$ y 2°) para cada número $\varepsilon > 0$ existe al menos un elemento $x \in A$ tal que $x > s - \varepsilon$. Y en forma simétrica se enuncian las propiedades que caracterizan al número $\inf A$ en el caso de que exista.

Dados dos números reales a y b tales que $a \leq b$ hay cuatro intervalos con extremo izquierdo a y extremo derecho b, a saber, $[a, b] = \{x \in I\!R : a \leq x \leq b\}$, $(a, b) = \{x \in I\!R : a < x < b\}$, $(a, b] = \{x \in I\!R : a < x \leq b\}$ y $[a, b) = \{x \in I\!R : a \leq x < b\}$. El primero se llama un **intervalo cerrado**; el segundo, un **intervalo abierto**, y los dos últimos son intervalos semicerrados (o semiabiertos).

Es conveniente incluir entre los intervalos al conjunto vacío, con lo que se logra eludir la consideración de casos excepcionales en la siguiente proposición que, aunque elementalísima, jugará más adelante un importante papel:

(1.1) **Proposición.** *La intersección de dos intervalos es un intervalo. La diferencia de dos intervalos es la unión de dos intervalos disjuntos.*

Consideremos, para fijar ideas, los intervalos $I = (a, b)$ y $J = (c, d]$. Luego, $I \cap J = \{x : a < x,\ x < b,\ c < x,\ x \leq d\}$. Ahora bien; la conjunción de las dos desigualdades del mismo sentido $x < b$ y $x \leq d$ es equivalente a una sola de ellas: a la primera si $b \leq d$ o la la segunda si $b > d$; y análogamente para las relaciones del mismo sentido $a < x$ y $c < x$, lo cual demuestra que $I \cap J$ es un intervalo, eventualmente vacío. En cuanto a la diferencia $I - J = \{x : x \in I,\ x \notin J\}$, sus elementos son los que satisfacen $a < x < b$ y alguna de las relaciones incompatibles $x \leq c$, $x > d$. Por consiguiente,

$$I - J = \{x : a < x < b,\ x \leq c\} \cup \{x : a < x < b,\ x > d\}$$

es la unión de dos intervalos disjuntos en virtud de una consideración análoga a la anterior.

La **medida** o **longitud** de cualquier intervalo con extremos a y b es el número $b - a$.

De fundamental importancia es el llamado **principio de encaje** de intervalos cerrados, enunciado en la siguiente proposición.

(1.2) **Proposición.** *Si* $I_n = [a_n, b_n]$, $n = 1, 2, 3, \ldots$, *es una sucesión de intervalos cerrados que verifican las inclusiones*

$$I_1 \supset I_2 \supset I_3 \supset \ldots \supset I_n \supset I_{n+1} \supset \ldots,$$

entonces existe un número real x *que pertenece a cada intervalo de la sucesión.*

En efecto, de la inclusión $I_{n+1} \subset I_n$ se deducen las relaciones $a_n \leq a_{n+1} \leq b_{n+1} \leq b_n$, lo cual muestra que los números a_n forman una sucesión creciente y los b_n una decreciente:

$$a_1 \leq a_2 \leq a_3 \leq \ldots, \qquad b_1 \geq b_2 \geq b_3 \geq \ldots.$$

Por consiguiente, si $m < n$ tendremos $a_m \leq a_n \leq b_n$ y también $a_n \leq b_n \leq b_m$; es decir, $a_n \leq b_m$ cualesquiera que sean los índices m y n. Luego,

$$\sup_n a_n = \sup\{a_1, a_2, a_3, \ldots\} \leq b_m \qquad (m = 1, 2, 3, \ldots),$$

de donde $\sup_n a_n \leq \inf_m b_m$ y el número $x = \sup_n a_n$ verifica $a_n \leq x \leq b_n$ $(n = 1, 2, 3, ...)$; es decir, $x \in I_n$ para cada n.

Q.E.D.

Otra importante consecuencia del hecho de que todo conjunto no vacío y acotado superiormente de números reales posea una cota superior mínima es que el conjunto $I\!N = \{1, 2, 3, ...\}$ no admite ninguna cota superior. La demostración es por el absurdo: si $I\!N$ admite un cota superior, entonces, como no es vacío, posee una cota superior mínima $s = \sup I\!N$ y como $s - 1 < s$, existe un elemento $n \in I\!N$ tal que $n > s - 1$; pero entonces el número natural $n + 1$ es mayor que s, lo cual es absurdo.

(1.3) **Corolario.** *Todo intervalo abierto no vacío contiene números racionales.*

Para demostrarlo, tomemos dos números reales a y b tales que $a < b$ y supongamos primero que a es positivo. En virtud de lo anterior, existe un número natural n tal que $n > (b - a)^{-1}$. Considerando el mínimo entero positivo m tal que $a < m/n$, tendremos

$$a < m/n = (m - 1)/n + 1/n < a + (b - a) = b$$

y el número racional $r = m/n$ verifica $a < r < b$.

Si a no es positivo, existe un número natural k tal que $a+k$ es positivo; y como $a + k < b + k$, la demostración se reduce al caso anterior.

Para indicar que el conjunto $A \subset I\!R$ no es acotado superiormente escribiremos $\sup A = +\infty$; la notación $\inf A = -\infty$ se usa para señalar que A no es acotado inferiormente.

Con la introducción de los elementos $+\infty$ y $-\infty$, que no son números, y la convención de que para cada número real x, $-\infty < x < +\infty$, podemos afirmar que todo conjunto no vacío $A \subset I\!R$ posee un supremo y un ínfimo. El conjunto $\overline{I\!R} = I\!R \cup \{-\infty, +\infty\}$ se llama la **recta real extendida**; y los números reales, por oposición a los nuevos elementos introducidos, se llaman los elementos **finitos** de $\overline{I\!R}$.

Las operaciones de suma y producto entre números reales se extienden a los elementos de $\overline{I\!R}$ con arreglo a las convenciones siguientes:

$$\begin{aligned}
(+\infty)+(+\infty) &= +\infty, \\
(-\infty)+(-\infty) &= -\infty, \\
(+\infty)+(-\infty) &= (-\infty)+(+\infty)=0, \\
(+\infty)\cdot(+\infty) &= (-\infty)\cdot(-\infty)=+\infty, \\
(+\infty)\cdot(-\infty) &= (-\infty)\cdot(+\infty)=-\infty;
\end{aligned}$$

si x es un número real,

$$x+(+\infty)=(+\infty)+x=+\infty, \qquad x+(-\infty)=(-\infty)+x=-\infty,$$

$$x\cdot(\pm\infty)=(\pm\infty)\cdot x=\begin{cases}\pm\infty & \text{si } x>0\\ 0 & \text{si } x=0\\ \mp\infty & \text{si } x<0.\end{cases}$$

Obsérvese que $0\cdot(\pm\infty)=(\pm\infty)\cdot 0=0$.

La utilidad de estas convenciones se irá percibiendo cada vez más claramente a medida que se desarrolle el tema principal de este libro.

El **límite inferior** ℓ y el **límite superior** L de una sucesión (a_k) de números reales se definen por medio de las fórmulas

$$\ell=\lim\inf{}_{k\to\infty} a_k=\sup_j\left(\inf_{k\ge j} a_k\right),$$

$$L=\lim\sup{}_{k\to\infty} b_k=\inf_j\left(\sup_{k\ge j} a_k\right).$$

De la definición se deducen fácilmente las dos propiedades que caracterizan al elemento $L\in\overline{I\!R}$; a saber,

1^a) para cada $\beta>L$, existe un entero positivo j tal que si $k\ge j$, entonces $a_k<\beta$;

2^a) para cada $\alpha<L$ y para cada entero positivo j, existe un entero $k\ge j$, tal que $\alpha<a_k$.

La primera afirma que si $\beta>L$, entonces $a_k<\beta$ para k suficientemente grande; la segunda que si $\alpha<L$, entonces $\alpha<a_k$ para infinitos índices k.

Ejercicio: Probar que si el elemento $\Lambda \in \overline{I\!R}$ posee las dos propieda- des anteriores, entonces $\Lambda = L$. Sugerencia: probar las relaciones $\Lambda \leq L$ y $L \leq \Lambda$.

En forma simétrica, las dos propiedades que caracterizan al elemento $\ell \in \overline{I\!R}$ son:

1^a) para cada $\alpha < \ell$ existe un entero positivo j tal que si $k \geq j$, entonces $\alpha < a_k$;

2^a) para cada $\beta > \ell$ y para cada entero positivo j, existe un entero $k \geq j$, tal que $a_k < \beta$.

Definiendo b_j y c_j por medio de las expresiones $b_j = \inf\limits_{k \geq j} a_k$, $c_j = \sup\limits_{k \geq j} a_k$, tendremos

$$b_j \leq c_j, \quad b_1 \leq b_2 \leq b_3 \leq ..., \quad c_1 \geq c_2 \geq c_3 \geq ...,$$

$$\ell = \sup_j \ b_j, \qquad L = \inf_j \ c_j.$$

Por consiguiente, $\ell \leq L$.

Por definición, la sucesión (a_k) **tiende** al **límite** $a \in \overline{I\!R}$ si se verifica $\ell = L = a$; y en tal caso escribimos $a = \lim\limits_{k \to \infty} \ a_k$. Si el límite a existe y es finito, decimos que la sucesión es **convergente** o que **converge** al valor a.

Toda sucesión monótona tiende a un límite en $\overline{I\!R}$, igual al supremo de todos sus valores si es creciente e igual al ínfimo de todos sus valores si es decreciente.

Si $a_n \geq 0, \ n = 1, 2, 3, ...,$ entonces la sucesión $s_n = a_1 + a_2 + \ldots + a_n$ es monótona creciente; de donde se sigue que el límite

$$\sum_{k=1}^{\infty} a_k = \lim_{n \to \infty} \ s_n = \lim_{n \to \infty} \ \sum_{k=1}^{n} \ a_k$$

tiene un sentido bien preciso en la recta real extendida. Por ejemplo, $1 + 1 + 1 + \ldots = +\infty$.

Consideremos ahora dos sucesiones monótonas con valores no negativos:

$$0 \leq a_1 \leq a_2 \leq a_3 \leq \ldots \quad \text{y} \quad 0 \leq b_1 \leq b_2 \leq b_3 \leq \ldots$$

Es fácil probar que cualesquiera que sean los límites de estas sucesiones (finitos o no) se cumplen las relaciones

$$\lim \ (a_n + b_n) = \lim \ a_n + \lim \ b_n, \qquad \lim \ (a_n b_n) = (\lim \ a_n)(\lim \ b_n).$$

Considerando el ejemplo $a_n = 0$ para todo n, $b_n = n$, se comprende la necesidad del convenio $0 \cdot (+\infty) = 0$ para que se mantenga la validez de la segunda.

2. Familias y sucesiones de conjuntos.

Frecuentemente será necesario considerar conjuntos cuyos elementos son conjuntos. Tales conjuntos se llaman **clases** o **colecciones** de conjuntos. Por ejemplo, si X es un conjunto, la clase $\mathcal{P}(X)$ formada por todos los subconjuntos de X se llama la clase de **partes** de X. En símbolos,

$$\mathcal{P}(X) = \{A : A \subset X\}.$$

Debe tenerse presente que X y $\emptyset$ son miembros de $\mathcal{P}(X)$.

Consideremos ahora un conjunto I y una función $F : I \to \mathcal{P}(X)$ que asigna a cada elemento $i \in I$ un conjunto $A_i = F(i)$ que es un subconjunto de X. En estas condiciones, la función F se llama una **familia de conjuntos** y se escribe $F = (A_i,\ i \in I)$ para destacar los conjuntos A_i que se llaman los **miembros** de la familia; los elementos de I se llaman los **índices** de la familia F.

La **unión** y la **intersección** de una familia $F = (A_i,\ i \in I)$ se definen, respectivamente, por medio de las fórmulas

$$\bigcup F = \bigcup_{i \in I} A_i = \{x : x \in A_i \text{ para algún } i \in I\},$$

$$\bigcap F = \bigcap_{i \in I} A_i = \{x : x \in A_i \text{ para todo } i \in I\}.$$

Una familia $(A_k,\ k = 1, 2, 3, ...)$ cuyos índices son los enteros positivos se llama una **sucesión** de conjuntos. En este caso, la unión y la intersección se representan por medio de los símbolos

$$\bigcup_{k=1}^{\infty} A_k, \qquad \bigcap_{k=1}^{\infty} A_k;$$

y en el caso de una **sucesión finita** $(A_k,\ 1 \leq k \leq n)$, escribiremos

$$\bigcup_{k=1}^{n} A_k, \qquad \bigcap_{k=1}^{n} A_k,$$

en lugar de $A_1 \cup A_2 \cup \ldots \cup A_n$ y $A_1 \cap A_2 \cap \ldots \cap A_n$, respectivamente.

Diremos que los conjuntos de la familia $F = (A_i,\ i \in I)$ son disjuntos o bien, con más propiedad, que F es una **familia disjunta** si para cada par de índices $i \neq j$ se verifica $A_i \cap A_j = \emptyset$.

Todos los conjuntos que se consideran en una teoría son subconjuntos de un determinado conjunto X llamado **espacio** con el fin de sugerir una interpretación geométrica. Los elementos de X se llaman entonces los **puntos** del espacio; y para cada $A \subset X$, el conjunto de todos los puntos del espacio que no pertenecen a A se llama el **complemento** de A y se denota por A^c o bien por CA.

Si A y B son conjuntos, definimos $A - B$ por $A - B = A \cap B^c$; $A \triangle B$ por $A \triangle B = (A - B) \cup (B - A)$. El conjunto $A \triangle B$ es la **diferencia simétrica** entre A y B.

Dejaremos a cargo del lector la demostración de las fórmulas

$$\left(\bigcup_{i \in I} A_i \right)^c = \bigcap_{i \in I} A_i^c, \quad \left(\bigcap_{i \in I} A_i \right)^c = \bigcup_{i \in I} A_i^c ,$$

que se conocen con el nombre de **leyes de complementación** o de De Morgan (para demostrar que dos conjuntos son idénticos hay que mostrar que cada elemento de uno de ellos pertenece al otro y viceversa).

Dada una sucesión de conjuntos $(E_k,\ k = 1, 2, 3, ...)$, los conjuntos

$$\text{lim inf } E_k = \bigcup_{j=1}^{\infty} \left(\bigcap_{k \geq j} E_k \right), \quad \text{lim sup } E_k = \bigcap_{j=1}^{\infty} \left(\bigcup_{k \geq j} E_k \right),$$

se llaman, respectivamente, el **límite inferior** y el **límite superior** de la sucesión (E_k).

Poniendo $A = \text{lim inf } E_k$, $B = \text{lim sup } E_k$, se tiene:

1°) $x \in A$ si y sólo si existe un entero positivo j tal que para todo $k \geq j$, $x \in E_k$.

2°) $x \in B$ si y sólo si para cada j existe un entero $k \geq j$ tal que $x \in E_k$.

Por consiguiente, $A \subset B$.

En el caso de se verifique $A = B$, se dice que la sucesión (E_k) **converge** al **límite** $L = A = B$.

Si la sucesión (E_k) verifica $E_1 \subset E_2 \subset E_3 \subset \ldots \subset E_k \subset E_{k+1} \subset \ldots$, decimos que es una **sucesión creciente**; y en este caso el límite de la sucesión existe y es igual a la unión de todos los conjuntos E_k.

Si la sucesión (E_k) verifica $E_1 \supset E_2 \supset E_3 \supset \ldots \supset E_k \supset E_{k+1} \supset \ldots$, decimos que es **decreciente**; en tal caso también existe el límite de la sucesión y es igual a la intersección de todos los conjuntos E_k.

Si $(A_i,\ i \in I)$ y $(B_j,\ j \in J)$ son dos familias de conjuntos, su cumple la **ley distributiva**

$$\left(\bigcup_{i\in I} A_i\right) \cap \left(\bigcup_{j\in J} B_j\right) = \bigcup_{i\in I}\bigcup_{j\in J} (A_i \cap B_j)\,.$$

Además, si ambas familias son disjuntas, entonces la familia $(A_i \cap B_j,\ (i,j) \in I \times J)$, cuya unión es precisamente el miembro derecho de la última igualdad, es también disjunta.

A modo de ejercicio útil sugerimos que el lector demuestre las siguientes fórmulas (las dos primeras generalizan las leyes de De Morgan):

$$B - \bigcap_{i\in I} A_i = \bigcup_{i\in I}(B - A_i)$$

$$B - \bigcup_{i\in I} A_i = \bigcap_{i\in I}(B - A_i)$$

$$\bigcup_{i\in I} A_i - B = \bigcup_{i\in I}(A_i - B).$$

Si $\mathcal{C}$ es una colección de conjuntos, la función $F : \mathcal{C} \to \mathcal{C}$ definida por $F(A) = A$ para cada $A \in \mathcal{C}$ nos permite pensar a $\mathcal{C}$ como una familia F en la que el conjunto de índices es el mismo conjunto $\mathcal{C}$. La unión y la intersección de esta familia se designan, en este caso, por medio de los símbolos

$$\bigcup \mathcal{C} = \bigcup\{A : A \in \mathcal{C}\} \quad \text{y} \quad \bigcap \mathcal{C} = \bigcap\{A : A \in \mathcal{C}\},$$

respectivamente. Por consiguiente, el punto x pertenece a la unión de $\mathcal{C}$ si y sólo si x pertenece al menos a un miembro de $\mathcal{C}$ y x pertenece a la intersección de $\mathcal{C}$ si y sólo si x pertenece a cada miembro de $\mathcal{C}$.

Obsevemos de paso que, por lo visto, la noción de familia es más general que la de colección o clase de conjuntos.

3. Imágenes directas e inversas.

Sea $f : X \to Y$ una función (o aplicación) de X en Y. Para cada $A \subset X$, el conjunto

$$f(A) = \{y \in Y : y = f(x) \text{ para algún } x \in A\}$$

es la **imagen** de A por f y para cada $B \subset Y$, el conjunto

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

se llama la **imagen inversa** de B por f.

Es fácil ver que $A \subset f^{-1}(f(A))$, $f(f^{-1}(B)) \subset B$ y $f^{-1}(Y - B) = Y - f^{-1}(B)$; además, la igualdad vale en la primera de estas relaciones si f es inyectiva y en la segunda si f es suryectiva. Si $(B_j,\ j \in J)$ es una familia de subconjuntos de Y, entonces

$$f^{-1}\left(\bigcup_{j\in J} B_j\right) = \bigcup_{j\in J} f^{-1}(B_j) \ \text{ y } \ f^{-1}\left(\bigcap_{j\in J} B_j\right) = \bigcap_{j\in J} f^{-1}(B_j).$$

Estas fórmulas muestran que la aplicación $f^{-1} : \mathcal{P}(Y) \to \mathcal{P}(X)$ tiene la importante propiedad de preservar todas las operaciones entre subconjuntos de Y.

4. Productos cartesianos.

El producto cartesiano de una familia finita $(A_1, A_2, ..., A_n)$ se define como el conjunto $A_1 \times A_2 \times ... \times A_n$ formado por todas las sucesiones finitas $(a_1, a_2, ..., a_n)$ tales que $a_i \in A_i$ para cada entero i que verifique $1 \leq i \leq n$. De acuerdo con esta definición, cada elemento del producto cartesiano no es más que una función f cuyo dominio es el conjunto $I_n = \{1, 2, ..., n\}$ con la propiedad de que $f(i) \in A_i$ para cada $i \in I_n$. La única particularidad, puramente notacional, consiste en la convención de escribir f en la forma $(f(1), f(2), ..., f(n))$. En particular, el producto cartesiano $A \times B$ está formado por todos los pares ordenados (x, y), tales que $x \in A$ e $y \in B$.

Para que el producto cartesiano resulte asociativo es necesario establecer una convención: si $a = (a_1, ..., a_n)$ es un elemento de $A_1 \times ... \times A_n$ y $b = (b_1, ..., b_m)$ un elemento de $B_1 \times ... \times B_m$, interpretaremos el par ordenado

(a, b) como la sucesión finita $(a_1, ..., a_n, b_1, ..., b_m)$. Con esta convención resulta $(A_1 \times ... \times A_n) \times (B_1 \times ... \times B_m) = A_1 \times ... \times A_n \times B_1 \times ... \times B_m$.

Por su sencillez, dejaremos a cargo del lector la demostración de las fórmulas

$$(A_1 \times ... \times A_n) \cap (B_1 \times ... \times B_n) = (A_1 \cap B_1) \times ... \times (A_n \cap B_n),$$

$$\left(\bigcup_{i=1}^{n} A_i\right) \times \left(\bigcup_{j=1}^{m} B_j\right) = \bigcup_{i=1}^{n} \bigcup_{j=1}^{m} (A_i \times B_j).$$

El miembro derecho de la segunda es la unión de la familia $F = (A_i \times B_j,\ (i, j) \in I_n \times I_m)$ y será útil observar que si los conjuntos A_i son disjuntos del mismo modo que los B_j, entonces F es una familia disjunta.

El siguiente lema tiene un valor auxiliar en la demostración de la proposición (2.21) del siguiente capítulo.

(1.4) **Lema.** *Si el producto cartesiano* $A \times B$ *es la unión de los productos no vacíos y disjuntos* $A_1 \times B_1$ *y* $A_2 \times B_2$, *entonces, o bien* $A = A_1 \cup A_2$ *con* $A_1 \cap A_2 = \emptyset$ *y* $B = B_1 = B_2$, *o bien* $B = B_1 \cup B_2$ *con* $B_1 \cap B_2 = \emptyset$ *y* $A = A_1 = A_2$.

Demostración: De la hipótesis se sigue fácilmente que ninguno de los conjuntos del enunciado es vacío y además, $A = A_1 \cup A_2$ y $B = B_1 \cup B_2$.

Puesto que

$$(A_1 \cap A_2) \times (B_1 \cap B_2) = (A_1 \times B_1) \cap (A_2 \times B_2) = \emptyset,$$

debe cumplirse el menos una de las relaciones

$$A_1 \cap A_2 = \emptyset, \qquad B_1 \cap B_2 = \emptyset.$$

Suponiendo que se verifica la primera, probaremos que $B = B_1$, para lo cual es suficiente probar que $B \subset B_1$. Si esto último no fuera cierto, entonces existiría un elemento $b \in B$, tal que $b \notin B_1$.

Eligiendo un elemento $a \in A_1$, tendríamos $a \notin A_2$, pues A_1 y A_2 son disjuntos. Ahora bien; $(a, b) \in A \times B$ y sin embargo, $(a, b) \notin A_1 \times B_1$ y $(a, b) \notin A_2 \times B_2$, lo cual contradice la hipótesis. Luego, $B = B_1$ y análogamente se demuestra que $B = B_2$.

Q.E.D.

La notación A^n se usa para designar el producto cartesiano $A \times A \times ... \times A$ (n factores); es decir, el conjunto de todas las sucesiones finitas $(a_1, a_2, ..., a_n)$, tales que $a_i \in A$ $(i = 1, 2, ..., n)$.

5. Conjuntos numerables.

Diremos que el conjunto X es **equivalente** al conjunto Y y escribiremos $X \sim Y$ si existe una aplicación biyectiva $f : X \to Y$. En tal caso también se dice que X e Y tienen **igual potencia** o bien que son coordinables.

Como la aplicación idéntica de X en sí mismo es biyectiva, se tiene siempre $X \sim X$; puesto que la inversa de una aplicación biyectiva es biyectiva, si $X \sim Y$ entonces $Y \sim X$; finalmente, puesto que la composición de dos aplicaciones biyectivas es biyectiva, si $X \sim Y$ e $Y \sim Z$, entonces $X \sim Z$.

Demotaremos por $I\!N$ al conjunto de los enteros positivos y llamaremos una **sección inicial** a cualquier conjunto I_n de la forma $I_n = \{1, 2, ..., n\} = \{k \in I\!N : k \leq n\}$, donde n representa un entero positivo.

Por definición, un conjunto es **finito** si es vacío o si es equivalente a alguna sección inicial y es **infinito** en caso contrario.

De acuerdo con la definición precedente, las secciones iniciales son los conjuntos finitos **típicos**.

Por inducción sobre n se demuestran fácilmente las siguientes afirmaciones:

(a) Cualquier subconjunto de I_n es finito.

(b) No existe ninguna aplicación inyectiva de I_{n+1} en I_n.

De la primera se deduce que todo subconjunto de un conjunto finito es finito; de la segunda, que si $m < n$, no existe ninguna aplicación inyectiva de I_n en I_m y por consiguiente, que todo conjunto finito no vacío es equivalente a una única sección inicial.

Otra consecuencia de (b) es que el conjunto $I\!N$ es infinito, pues si existiera una aplicación biyectiva $f : I\!N \to I_n$, la restricción de f a I_{n+1} nos daría una aplicación inyectiva de este último conjunto en I_n, en contradicción con (b).

Confiando en su evidencia, aceptaremos sin demostración la siguiente proposición:

(c) Todo conjunto infinito contiene un conjunto equivalente a $I\!N$. En otras palabras, si X es un conjunto infinito, existe una aplicación inyectiva $f : I\!N \to X$.

Un conjunto A se llama **numerable** si es finito o bien si $A \sim I\!N$. De acuerdo con esta definición, el conjunto A es numerable si y sólo si es posible representar la totalidad de sus elementos en la forma de una sucesión $a_1, a_2, a_3, ...$ finita o infinita.

(1.5) **Proposición.** *Cualquier subconjunto de un conjunto numerable es numerable.*

Sea B un subconjunto del conjunto numerable A. Si A es finito, entonces B es numerable en virtud de (a). Supongamos que $A = \{a_1, a_2, a_3, ...\}$ es equivalente a $I\!N$ y que B es un subconjunto infinito de A. Sea n_1 el mínimo de los enteros positivos n tales que $a_n \in B$; y suponiendo definido n_k , llamemos n_{k+1} al mínimo entero $n > n_k$, tal que $a_n \in B$.

Con este procedimiento inductivo queda definida una sucesión creciente de enteros positivos (n_k) , tal que $a_{n_k} \in B$, $k = 1, 2, 3, ...$. Consideremos ahora el conjunto numerable

$$C = \{a_{n_1}, a_{n_2}, a_{n_3}, ...\}.$$

Sabemos que $C \subset B$ y la proposición quedará demostrada si probamos que $B = C$.

Si $x \in B$, entonces $x = a_n$ para algún $n \geq n_1$. Consideremos el mínimo entero k tal que $n \leq n_k$. Si fuera $n < n_k$ sería $k > 1$ y $n_{k-1} < n < n_k$, en contradicción con lo supuesto. Luego $n = n_k$ y $x \in C$. Hemos probado que $B \subset C$ y junto con ello la proposición.

(1.6) **Proposición.** *Si A es numerable y $f : A \to B$ es suryectiva, entonces B es numerable.*

Sin restricción de generalidad, podemos suponer que $A = I\!N$, pues si A es numerable, existe una aplicación suryectiva de $I\!N$ sobre A cuya composición con f nos da una aplicación suryectiva de $I\!N$ sobre B . Suponiendo pues que $A = I\!N$, para cada $b \in B$ sea $g(b)$ el mínimo entero $n \in I\!N$ tal que $f(n) = b$; de modo que $f(g(b)) = b$ para cada $b \in B$. De aquí se deduce que g es una aplicación inyectiva de B en $I\!N$ y por consiguiente,

$B \sim g(B)$; pero $g(B)$ es numerable por ser un subconjunto de $\mathbb{N}$. Luego, B es numerable como queríamos demostrar.

(1.7) **Proposición.** $\mathbb{N} \times \mathbb{N} \sim \mathbb{N}$.

Teniendo en cuenta que todo número natural admite una única descomposición como producto de una potencia de dos por un número impar, se sigue que la función $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, definida por

$$f(i,j) = 2^{i-1}(2j-1)$$

es biyectiva.

Q.E.D.

Si A y B son numerables, entonces existe una aplicación suryectiva de $\mathbb{N} \times \mathbb{N}$ sobre $A \times B$. Luego, $A \times B$ es numerable en virtud de (1.6) y (1.7).

Por inducción sobre n resulta que si $(A_1, A_2, ..., A_n)$ es una sucesión finita de conjuntos numerables, entonces $A_1 \times A_2 \times ... \times A_n$ es numerable.

Ejemplos.

1) El conjunto $\mathbb{Z}$ formado por los números enteros es numerable, pues la función $f : \mathbb{Z} \to \mathbb{N}$ definida por $f(a) = 2a$ si $a > 0$ y $f(a) = 2|a| + 1$ si $a \leq 0$, es biyectiva.

2) El conjunto $\mathbb{Q}$ formado por los números racionales es numerable, pues la función $f : \mathbb{Z} \times \mathbb{N} \to \mathbb{Q}$ definida por $f(a,n) = a/n$ es suryectiva.

3) El conjunto $\mathbb{Q}^n$ formado por todas las n-uplas $(r_1, r_2, ..., r_n)$ de números racionales es numerable.

4) El conjunto $\mathbb{Q}[X]$ formado por todos los polinomios con coeficientes racionales es numerable. En efecto, si a cada polinomio $A = a_0 + a_1X + ... + a_nX^n$ le asignamos la sucesión $f(A) = (a_0, a_1, ..., a_n) \in \mathbb{Q}^{n+1}$, tendremos una aplicación inyectiva f de $\mathbb{Q}[X]$ en el conjunto $\bigcup\limits_{n=1}^{\infty} \mathbb{Q}^n$ que es numerable por ser la unión de una sucesión de conjuntos numerables.

(1.8) **Proposición.** *La unión de cualquier sucesión de conjuntos numerables es numerable.*

Sea $(A_n,\ n = 1, 2, 3, ...)$ una sucesión de conjuntos numerables. Para cada n, existe una función suryectiva $f_n : I\!N \to A_n$.

Si A es la unión de todos los conjuntos A_n, la fórmula

$$g(n,k) = f_n(k) \qquad\qquad (n \in I\!N\,,\ k \in I\!N\,),$$

define una aplicación suryectiva $g : I\!N \times I\!N \to A$, lo cual prueba que A es numerable.

6. Potencia del continuo.

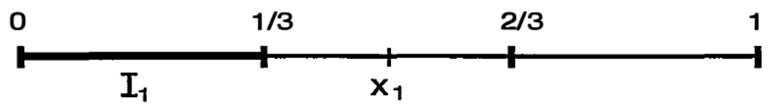
La teoría de los conjuntos infinitos no tendría mucho interés si todos los conjuntos infinitos fueran equivalentes. El descrubrimiento de que no es así representa uno de los momentos cruciales en la historia de las Matemáticas como cuando los griegos descubrieron que la diagonal de un cuadrado es inconmensurable con el lado del mismo.

(1.9) **Proposición.** *El intervalo unitario* $U = [0, 1]$ *es un conjunto no numerable.*

Suponiendo por el absurdo que sea posible representar los elementos de U en la forma de una sucesión

$$U = \{x_1, x_2, x_3, ...\},$$

dividamos el intervalo U en tres intervalos cerrados de igual longitud: $[0, 1/3], [1/3, 2/3], [2/3, 1]$. Entonces es claro que uno de estos intervalos ce-



rrados, al cual llamaremos I_1, tiene la propiedad de que $x_1 \notin I_1$. Subdividiendo I_1 en forma análoga, podemos elegir un intervalo cerrado $I_2 \subset I_1$, tal que $x_2 \notin I_2$.

Si continuamos indefinidamente con este proceso obtendremos una sucesión decreciente de intervalos cerrados $U \supset I_1 \supset I_2 \supset ... \supset I_k \supset ...,$ con la propiedad de que para cada entero positivo k, $x_k \notin I_k$. Pero entonces, ningún punto de U pertenece a cada intervalo de la sucesión, en contradicción con el principio de "encaje" de intervalos cerrados. Obviamente, la contradicción provino de suponer que U es numerable.

Q.E.D.

Diremos que un conjunto tiene la **potencia del continuo** si es equivalente al intervalo unitario U. Veamos algunos ejemplos.

1) El intervalo abierto $V = (0,1)$ tiene la potencia del continuo, pues si ponemos $A = \mathbb{Q} \cap U$, $B = \mathbb{Q} \cap V$, tendremos:

$$U = (U - A) \cup A, \qquad V = (U - A) \cup B.$$

Puesto que A y B son subconjuntos infinitos de $\mathbb{Q}$, se tiene $A \sim B$ en virtud de (1.5). Luego, existe una aplicación biyectiva $f : A \to B$.

Si ahora definimos $g : U \to V$ por medio de la expresión

$$g(x) = \begin{cases} x & \text{si } x \in U - A \\ f(x) & \text{si } x \in A, \end{cases}$$

es fácil ver que g es biyectiva, de donde $U \sim V$. Análogamente se demuestra que los intervalos $(0,1]$ y $[0,1)$ tienen la potencia del continuo.

2) Si $a < b$, el intervalo $[a,b]$ tiene la potencia del continuo, pues la función lineal

$$f(x) = a + (b-a)x \qquad (0 \leq x \leq 1)$$

es una aplicación biyectiva de U sobre $[a,b]$. Por consiguiente, cualquier intervalo que no sea vacío ni se reduzca a un solo punto tiene la potencia del continuo.

3) El conjunto $\mathbb{R}$ formado por los números reales tiene la potencia del continuo, pues la función estrictamente creciente $f : \mathbb{R} \to (-1,1)$, definida por $f(x) = x/(1+|x|)$ es biyectiva.

Para estudiar el siguiente ejemplo es necesario recordar que si b es un entero mayor que 1, cada número $x \in U$ admite un **desarrollo "b-ario"**

$$x = \sum_{k=1}^{\infty} \frac{u_k}{b^k} \ ,$$

donde cada u_k es un **dígito b-ario**; es decir, igual a uno de los números $0, 1, ..., b-1$. Además, si x admite dos desarrollos b-arios distintos $\sum u_k/b^k$ y $\sum v_k/b^k$, entonces para k suficientemente grande se cumple $u_k = 0$ y $v_k = b-1$ o viceversa: $u_k = b-1$ y $v_k = 0$. Como consecuencia, ningún número admite más que, a lo sumo, dos desarrollos b-arios distintos y si x admite dos desarrollos, entonces x es un número de la forma n/b^m. Ahora sí veamos el ejemplo.

4) El conjunto T formado por todas las sucesiones $u = (u_1, u_2, u_3, ...)$ de dígitos binarios (cada u_k igual a cero o uno) tiene la potencia del continuo.

Para demostrarlo, comencemos observando que hay exactamente 2^n elementos de T, tales que $u_k = 0$ para todo $k > n$. Luego, el conjunto T_0 formado por todos los elementos $u \in T$, tales que $u_k = 0$ para k suficientemente grande es numerable por ser la unión de una sucesión de conjuntos finitos. Análogamente, el conjunto T_1 formado por los $u \in T$, tales que $u_k = 1$ para k suficientemente grande es numerable.

Poniendo $T_2 = T - (T_0 \cup T_1)$, sea A el conjunto de todos los números de la forma $n/2^m$ dentro del intervalo $U = [0, 1]$; es decir, aquellos que admiten dos desarrollos binarios distintos. Puesto que A es numerable, tenemos

$$A \sim T_0 \cup T_1.$$

Por otro lado, cada $x \in U - A$ posee un único desarrollo binario

$$x = \sum_{k=1}^{\infty} \frac{u_k}{2^k},$$

donde $u = (u_1, u_2, u_3, ...) \in T_2$. Como la función que asigna al elemento x la sucesión u dada por su único desarrollo binario es biyectiva, resulta

$$U - A \sim T_2.$$

Finalmente, puesto que $U = A \cup (U - A)$ y $T = (T_0 \cup T_1) \cup T_2$, concluimos que $U \sim T$ como habíamos anunciado.

5) El conjunto $S = I\!R - \mathbb{Q}$ formado por todos los números irracionales tiene la potencia del continuo. En efecto, puesto que $I\!R = \mathbb{Q} \cup S$, vemos que S es no numerable (si S fuera numerable, resultaría $I\!R$ numerable).

Por ser infinito, S contiene un conjunto A equivalente a $I\!N$ y de las igualdades obvias $S = (S-A) \cup A$, $I\!R = S \cup \mathbb{Q} = (S-A) \cup (A \cup \mathbb{Q})$, teniendo en cuenta que $A \sim A \cup \mathbb{Q}$, resulta $S \sim I\!R$.

Con el mismo método se prueba que si X es infinito y B numerable, entonces $X \cup B \sim X$ (hágalo el lector como ejercicio).

6) Un número (real o complejo) se llama **algebraico** si es raíz de algún polinomio con coeficientes enteros. Por ejemplo, el número racional a/b es algebraico por ser raíz del polinomio $a - bx$; $\sqrt{2}$ es algebraico por ser raíz de $x^2 - 2$. Los números que no son algebraicos se llaman **trascendentes**.

El problema de exhibir un número trascendente no es trivial. El primero en lograrlo fue Liouville, quien probó que el número

$$0,1100010... = 10^{-1!} + 10^{-2!} + ... + 10^{-n!} + ...$$

es trascendente.

La demostración de que π es trascendente se debe a Lindemann (1882); y el teorema de Lindemann generalizado afirma que si α es un número algebraico no nulo, entonces e^α es trascendente. Pero la demostración de estos resultados que pueden estudiarse en el libro de Niven [10], dista de ser fácil. En cambio, la solución de Cantor al problema de la existencia de números trascendentes es sencilla: el conjunto A de los números algebraicos es numerable, pues es numerable el conjunto de los polinomios con coeficientes enteros y cada uno de éstos tiene un número finito de raíces. Luego $\mathbb{R} - A$ (el conjunto de los números trascendentes) tiene la potencia del continuo; en particular, no es vacío.

Aun siendo infinito, el conjunto de los números algebraicos está esparcido sobre la recta "como las estrellas en el firmamento" cuya densa obscuridad corresponde, en esta metáfora de E.T. Bell, a los números trascendentes.

6) El conjunto $\mathbb{R}^n$ formado por todas las n-uplas de números reales tiene la potencia del continuo.

Puesto que $\mathbb{R} \sim T$ (ejemplo 4), es suficiente probar que $T \sim T^n$. Con este propósito definimos una aplicación $f : T \to T^n$ de acuerdo con el esquema siguiente:

$$(u_1, u_2, u_3, ...) \xrightarrow{f} \begin{cases} (u_1, u_{n+1}, u_{2n+1}, ...) \\ (u_2, u_{n+2}, u_{2n+2}, ...) \\ ... \quad ... \quad ... \\ (u_n, u_{2n}, u_{3n}, ...). \end{cases}$$

Observando que para cada $u \in T$ las n sucesiones que componen $f(u)$

permiten reconstruir el elemento u, concluimos que f es biyectiva; y por consiguiente, $T \sim T^n$.

7) El conjunto $\mathbb{R}^\infty$ formado por todas las sucesiones $(x_1, x_2, x_3, ...)$ de números reales tiene la potencia del continuo.

Por lo mismo que antes, si designamos por T^∞ el conjunto de todas las sucesiones de elementos de T, es suficiente que probemos que $T \sim T^\infty$. Con este fin definimos una aplicación g de T en T^∞ haciendo corresponder a la sucesión $u = (u_1, u_2, u_3, ...) \in T$ la sucesión de elementos de T dados por

$$\begin{array}{l} (u_1, u_3, u_5, ...), \\ (u_2, u_6, u_{10}, ...), \\ (u_4, u_{12}, u_{20}, ...), \\ \ \ ... \quad ... \quad ... \ , \end{array}$$

donde en la i-ésima fila figuran aquellos elementos de u con subíndices iguales a $2^{i-1}(2k-1), k = 1, 2, 3, ...$; en otras palabras, $g(u) = ((u_1, u_3, u_5, ...), (u_2, u_6, u_{10}, ...), ...)$.

Puesto que a partir de la sucesión $g(u)$ es posible reconstruir u, se deduce que g es biyectiva; de donde $T \sim T^\infty$.

Q.E.D.

7. Cardinales transfinitos.

En la teoría de conjuntos, a cada conjunto X se le asigna un objeto $\text{Card}(X)$, llamado el **cardinal** o **potencia** de X, con la propiedad de que $\text{Card}(X) = \text{Card}(Y)$ si y sólo si $X \sim Y$. La definición se da de tal modo que resulta $\text{Card}(\emptyset) = 0$, $\text{Card}(I_n) = n$; de manera que el cardinal de un conjunto representa una generalización del concepto de **número de elementos** de un conjunto finito.

El cardinal de un conjunto infinito se llama un **cardinal transfinito** y se demuestra que hay infinitos cardinales transfinitos distintos. Además, entre los números cardinales se establece en forma natural una relación de orden y se desarrollan ciertas nociones aritméticas; a saber, suma, producto y potenciación de cardinales.

El lector interesado en estudiar estos temas puede consultar la obra clásica de Sierpinski [14]. El apéndice a la obra de Kelley [7] desarrolla la teoría de los conjuntos desde un punto de vista axiomático, sin exigir del lector más que conocimientos rudimentarios de lógica y cierta familiaridad con el estudio de una teoría en un marco formal y abstracto. Superada esta barrera inicial, la exposición de Kelley es de notable elegancia; pero conviene advertir que el estudio de los aspectos formales conviene postergarlo hasta haber adquirido cierta madurez matemática, como la que resulta de haber estudiado el tema de este libro o bien un curso de Topología.

8. Partición de dominios; el teorema de Schröeder-Bernstein.

El teorema de Schröeder-Bernstein establece que si el conjunto X es equivalente a una parte de Y e Y es equivalente a una parte de X, entonces X es equivalente a Y. Para demostrarlo utilizaremos el siguiente teorema, debido al matemático Stefan Banach, que se conoce con el nombre de "teorema de partición de dominios".

(1.10) **Teorema.** *Si f es una aplicación de X en Y, y g una aplicación de Y en X, entonces existe un subconjunto A de X y un subconjunto B de Y, tales que*

$$f(A) = B, \qquad g(Y - B) = X - A.$$

Es decir, f aplica A sobre B y g aplica el complemento de B relativo a Y sobre el complemento de A relativo a X.

Demostración: Consideremos la aplicación $\Phi : \mathcal{P}(X) \to \mathcal{P}(X)$, definida para cada subconjunto A de X por medio de la fórmula

$$\Phi(A) = X - g(Y - f(A)), \qquad A \in \mathcal{P}(X).$$

Se verifica fácilmente que Φ es "creciente"; es decir, que si A_1 y A_2 son subconjuntos de X, la relación $A_1 \subset A_2$ implica $\Phi(A_1) \subset \Phi(A_2)$. Denotemos por Γ la colección formada por todos los subconjuntos G de X, tales que $\Phi(G) \subset G$.

Puesto que $\Phi(X) \subset X$, se sigue que $X \in \Gamma$, de modo que Γ no es vacía.

Llamemos A a la intersección de la colección Γ; es decir, definamos el conjunto A poniendo $A = \bigcap \Gamma$.

Si $G \in \Gamma$, entonces $A \subset G$ y por consiguiente, $\Phi(A) \subset \Phi(G) \subset G$, lo cual muestra que $\Phi(A)$ está contenido en cada miembro de Γ; de donde se infiere que $\Phi(A) \subset A$.

Aplicando Φ a cada miembro de la última relación, obtenemos

$$\Phi(\Phi(A)) \subset \Phi(A),$$

la cual significa que $\Phi(A)$ es un miembro de Γ y por consiguiente, $A \subset \Phi(A)$. Esta inclusión, juntamente con la inclusión opuesta que ya habíamos obtenido, nos da la igualdad $\Phi(A) = A$; es decir,

$$X - g(Y - f(A)) = A.$$

Poniendo $B = f(A)$, tendremos $X - g(Y - B) = A$. Esta relación muestra que $g(Y - B)$ es el complemento de A relativo a X, que es precisamente lo que queríamos probar.

Q.E.D.

Afirmar que X es equivalente a una parte de Y equivale a decir que existe una aplicación inyectiva $f : X \to Y$. Por este motivo es conveniente dar al teorema de Schröeder-Bernstein el siguiente enunciado:

(1.11) **Teorema.** *Si f es una aplicación inyectiva de X en Y, y g es una aplicación inyectiva de Y en X, entonces existe una aplicación biyectiva $h : X \to Y$.*

Demostración: Consideremos los conjuntos $A \subset X$ y $B \subset Y$ cuya existencia garantiza el teorema anterior.

Puesto que g es inyectiva, para cada elemento x del conjunto $X - A = g(Y - B)$, existe un único elemento y de $Y - B$, tal que $x = g(y)$. En forma natural escribiremos $y = g^{-1}(x)$.

Definiendo la aplicación $h : X \to Y$ por medio de la fórmula:

$$h(x) = \begin{cases} f(x) & \text{si } x \in A, \\ g^{-1}(x) & \text{si } x \in X - A, \end{cases}$$

se comprueba muy fácilmente que h es biyectiva.

Q.E.D.

Del teorema demostrado extraemos un corolario que resultará útil en el capítulo siguiente: si un subconjunto E de $I\!R^n$ contiene un conjunto con la potencia del continuo, entonces E tiene la potencia del continuo.

En efecto, puesto que $\mathbb{R}^n \sim \mathbb{R}$, se sigue que E es equivalente a una parte de $\mathbb{R}$. Por otro lado, la hipótesis nos dice que $\mathbb{R}$ es equivalente a una parte de E; luego, $E \sim \mathbb{R}$.

EJERCICIOS

1. Probar las fórmulas

$$\begin{aligned} A - B &= A - A \cap B \\ A \cap (B - C) &= A \cap B - A \cap C \\ (A - B) - C &= A - (B \cup C) \\ A \triangle (B \triangle C) &= (A \triangle B) \triangle C \\ A \cap (B \triangle C) &= (A \cap B) \triangle (A \cap C) \\ A - B^c &= B - A^c \\ \bigcap_{i \in I} A_i - B &= \bigcap_{i \in I} (A_i - B) \end{aligned}$$

2. ¿Cómo debe se el conjunto X para que se cumpla $A \triangle X = A$?

3. Sea f una aplicación de X en Y.
(a) Probar que si A_1 y A_2 son subconjuntos de X, entonces

$$f(A_1 \cup A_2) = f(A_1) \cup f(A_2) \qquad \text{y}$$

$$f(A_1 \cap A_2) \subset f(A_1) \cap f(A_2).$$

Generalizar estas relaciones y dar un ejemplo donde la inclusión de la segunda sea propia.
(b) Probar que f es inyectiva si y sólo si $f^{-1}(f(A)) = A$ para cada $A \subset X$.
(c) Probar que f es suryectiva si y sólo si $f(f^{-1}(B)) = B$ para cada $B \subset Y$.

4. Probar que si $f : X \to Y$, $g : Y \to Z$ y $h = g \circ f$, entonces para cada $C \subset Z$, $h^{-1}(C) = f^{-1}\left(g^{-1}(C)\right)$.

5. Mostrar que la sucesión de conjuntos

$$E_k = \begin{cases} (0, 1 - 1/k) & \text{si } k \text{ es par} \\ [1/k, 1) & \text{si } k \text{ es impar} \end{cases}$$

no es monótona, pero tiende a un límite.

6. Probar que si (a_k) y (b_k) son sucesiones de números reales y $c \in I\!R$, entonces
a) $\limsup (a_k + b_k) \leq \limsup a_k + \limsup b_k$
b) $\liminf (a_k + b_k) \geq \liminf a_k + \liminf b_k$
c) $\limsup (c + a_k) = c + \limsup a_k$
d) $\liminf (c - a_k) = c - \limsup a_k$
e) si $c_k \to c$ $(k \to \infty)$, entonces $\limsup (c_k + a_k) = c + \limsup a_k$
f) si $c \geq 0$, $\limsup (c a_k) = c \limsup a_k$.

7. Si $(x_\alpha,\ \alpha \in A)$ es una familia arbitraria (no necesariamente numerable) de valores no negativos de la recta extendida $\overline{I\!R}$, entonces definimos la **suma desordenada** de dicha familia por medio de la fórmula

$$\sum_{\alpha \in A} x_\alpha = \sup_F \sum_{\alpha \in F} x_\alpha,$$

donde el supremo se toma sobre todos los subconjuntos finitos F de A. Probar las siguientes afirmaciones:
a) Si $A = \{\alpha_1, \alpha_2, ...\}$ es numerable, entonces

$$\sum_{\alpha \in A} x_\alpha = x_{\alpha_1} + x_{\alpha_2} + ... = \sum_{k=1}^{\infty} x_{\alpha_k}.$$

b) Si $\sum_{\alpha \in A} x_\alpha < \infty$, entonces el conjunto de todos los α, tales que $x_\alpha > 0$ es numerable.
c) Si $(A_i,\ i \in I)$ es una **partición** de A, es decir, una familia disjunta cuya unión es A, entonces

$$\sum_{\alpha \in A} x_\alpha = \sum_{i \in I} \sum_{\alpha \in A_i} x_\alpha$$

(fórmula de la asociatividad).

8. Usar la fórmula de asociatividad del ejercicio precedente para probar que si (a_{nk}) es una sucesión doble de números no negativos entonces, denotando por $I\!N$ el conjunto de los números naturales,

$$\sum_{(n,k)\in I\!N \times I\!N} a_{nk} = \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} a_{nk} = \sum_{k=1}^{\infty}\sum_{n=1}^{\infty} a_{nk}$$

(una serie doble con términos no negativos puede sumarse en cualquier orden).

9. Probar que los siguientes conjuntos son numerables:
a) el conjunto de todos los intervalos con extremos racionales
b) cualquier colección de intervalos abiertos disjuntos
c) el conjunto de todas las sucesiones finitas de elementos de un conjunto numerable
d) el conjunto $\mathcal{F}(A)$ formado por todos los subconjuntos finitos de un conjunto numerable A
e) cualquier colección de lemniscatas disjuntas del plano (se llama **lemniscata** a cualquier curva continua cerrada que se autointersecta en uno de sus puntos, en forma de ocho)
f) el conjunto de puntos de discontinuidad de una función monótona.

10. Probar que los siguientes conjuntos tienen la potencia del continuo:
a) el conjunto de todas las sucesiones estrictamente crecientes $n_1 < n_2 < n_3 < ...$ de enteros positivos
b) el conjunto de todas ls sucesiones de enteros positivos
c) la colección de todos los intervalos
d) el conjunto C formado por todas las funciones continuas sobre un intervalo. Sugerencia: denotando por $r_1, r_2, ...$ los puntos racionales del intervalo, asignemos a cada función continua f una sucesión de números reales, de acuerdo con el esquema

$$f \mapsto (f(r_1), f(r_2), ...).$$

Dicha asignación define una aplicación inyectiva de C en el conjunto $I\!R^{\infty}$ formado por las sucesiones de números reales (equivalente a $I\!R$). Puesto que por otro lado las funciones constantes forman un subconjunto de C con la potencia del continuo, la afirmación resulta del teorema de Schröeder-Bernstein.

Nota: Los ejercicios siguientes requieren el uso del **axioma de elección**, el cual establece que para cualquier conjunto X, si denotamos por $\mathcal{P}_0(X)$ la colección formada por los subconjuntos no vacíos de X, entonces existe una aplicación $e : \mathcal{P}_0(X) \to X$, tal que $e(A) \in A$ para cada A en $\mathcal{P}_0(X)$.
Obsérvese que la función e "elige" un punto dentro de cada subconjunto no vacío de X.

11. Probar que si existe una aplicación suryectiva f de X sobre Y, entonces $Card\, Y \leq Card\, X$. Sugerencia: usar el axioma de elección para definir una aplicación $g : Y \to X$, tal que $f \circ g = id_Y =$ aplicación idéntica de Y en sí mismo.

12. Si X es un conjunto infinito y $I\!N$ el conjunto de los números naturales, entonces existe una aplicación inyectiva $f : I\!N \to X$. Sugerencia: definir f por medio de las fórmulas recursivas

$$f(1) = e(X)$$
$$f(n+1) = e(X - \{f(1), ..., f(n)\}).$$

13. Probar que si $A \cup B$ tiene la potencia del continuo, entonces al menos uno de los dos conjuntos tiene la potencia del continuo. Pista: partiendo de una aplicación biyectiva $f : A \cup B \to I\!R \times I\!R$, razonar por el absurdo suponiendo que ni A ni B tienen la potencia del continuo. Entonces las imágenes $\pi_1(f(A))$ y $\pi_2(f(B))$, donde π_1 y π_2 son las proyecciones de $I\!R \times I\!R$ sobre $I\!R$, son subconjuntos propios de $I\!R$ en virtud del ejercicio 11.

CAPITULO II

ESPACIOS EUCLIDIANOS

1. Espacio $\mathbb{R}^n$.

Como es usual en el estudio del álgebra lineal y de las funciones de varias variables, interpretaremos cada n-upla de números reales

$$x = (x_1, \ldots, x_n)$$

como un punto o bien como un vector de un espacio n-dimensional al que llamaremos el **espacio euclidiano** $\mathbb{R}^n$, que no es otra cosa que el producto cartesiano $\mathbb{R} \times \ldots \times \mathbb{R}$ (n factores) provisto de una estructura de espacio vectorial con un producto interno, la cual extiende en forma natural ciertas nociones geométricas familiares del plano y del espacio tridimensional.

La suma de los vectores $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$ y el producto del vector x por el número real λ se definen, respectivamente, por medio de las fórmulas

$$x + y = (x_1 + y_1, \ldots, x_n + y_n), \qquad \lambda x = (\lambda x_1, \ldots, \lambda x_n).$$

El número real

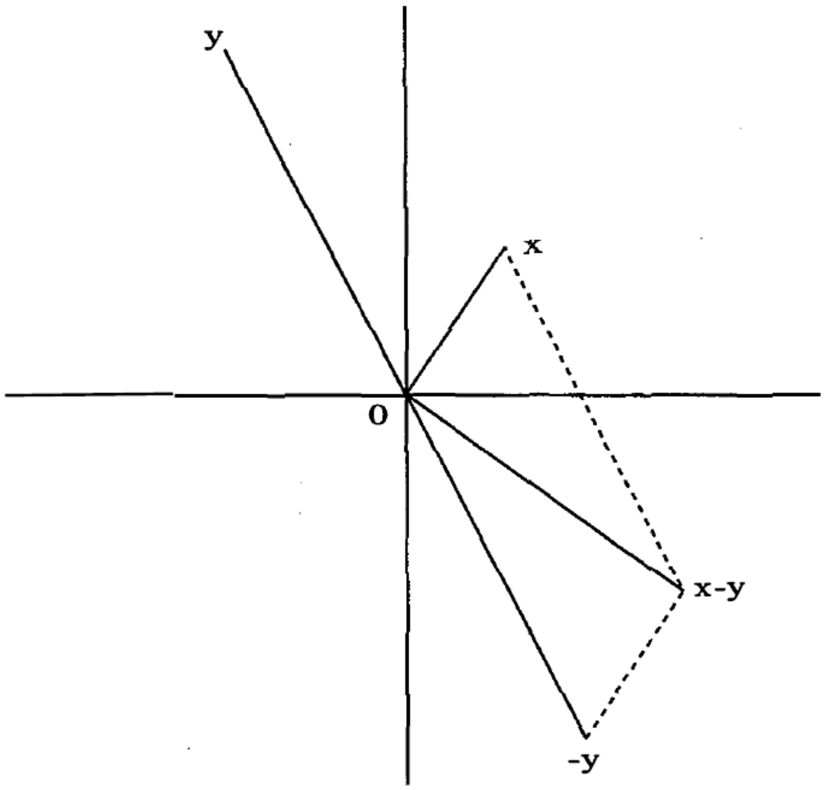
$$x \cdot y = x_1 y_1 + \ldots + x_n y_n$$

se llama el **producto escalar** o **producto interno** de los vectores x e y. La n-upla $0 = (0, \ldots, 0)$ se llama el **vector nulo** u **origen** y el vector $-x = (-x_1, \ldots, -x_n)$ es el vector **opuesto** o **simétrico** de x. Por definición, $x - y = x + (-y)$.

Llamaremos **módulo** o **norma** del vector x al número no negativo

$$|x| = \sqrt{x \cdot x} = \sqrt{x_1^2 + \ldots + x_n^2}.$$

En muchas cuestiones la dimensión n desempeña un papel secundario y entonces es conveniente imaginar los elementos de $\mathbb{R}^n$ en un diagrama como el de la figura, que tiene el mérito de adaptarse al plano del dibujo.



Decimos que dos vectores x e y son **ortogonales** cuando su producto escalar es igual a cero. Los vectores de módulo igual a uno se llaman **vectores unitarios**.

Por el **segmento** que une a los puntos x e y entendemos el conjunto de todos los puntos z de la forma $z = (1-\lambda)x + \lambda y \quad (0 \leq \lambda \leq 1)$.

El teorema fundamental sobre productos escalares es la llamada **desigualdad de Cauchy-Schwarz**:

(2.1) **Teorema.**

$$|x \cdot y| \leq |x|\,|y|$$

Demostración. Notemos que para cualquier número real λ se cumple

$$0 \le (\lambda x - y) \cdot (\lambda x - y) = \lambda^2 (x \cdot x) - 2\lambda (x \cdot y) + (y \cdot y).$$

Poniendo $A = x \cdot x$, $B = x \cdot y$, $C = y \cdot y$, podemos escribir estas relaciones en la forma

$$0 \le |\lambda x - y|^2 = A\lambda^2 - 2B\lambda + C.$$

Si $A = 0$, entonces $x = 0$ y la desigualdad es trivialmente verdadera en este caso. Suponiendo $A > 0$, podemos completar cuadrados en el miembro derecho, obteniendo

$$(1) \qquad 0 \le |\lambda x - y|^2 = A(\lambda - B/A)^2 + (AC - B^2)/A.$$

Si en esta relación elegimos $\lambda = B/A$, resulta $0 \le AC - B^2$; o sea, $|B| \le \sqrt{A}\sqrt{B}$, que es precisamente la afirmación del teorema.

La desigualdad es estricta, a menos que los vectores x e y sean linealmente dependientes. En efecto, si $AC - B^2 = 0$ y $x \ne 0$, entonces para $\lambda = B/A$, se deduce de (1) que $\lambda x - y = 0$.

(2.2) **Teorema** *(desigualdad de Minkowski).*

$$|x + y| \le |x| + |y|.$$

Demostración. $|x + y|^2 = (x + y) \cdot (x + y) = |x|^2 + 2x \cdot y + |y|^2 \le |x|^2 + 2|x \cdot y| + |y|^2 \le |x|^2 + 2|x|\,|y| + |y|^2 = (|x| + |y|)^2$.

Q.E.D.

Puesto que $|\lambda x|^2 = (\lambda x) \cdot (\lambda x) = \lambda^2 (x \cdot x) = |\lambda|^2 |x|^2 = (|\lambda|\,|x|)^2$, se deduce

$$(2.3) \qquad |\lambda x| = |\lambda|\,x|.$$

Las relaciones (2.2) y (2.3), juntamente con el hecho de que $|x| = 0$ si y sólo si $x = 0$, constituyen las propiedades fundamentales de la función numérica $|x|$ cuyo dominio es $I\!R^n$.

Si observamos que para cualquier par de vectores x e y se tiene $|x| = |(x - y) + y| \le |x - y| + |y|$, concluimos que $|x| - |y| \le |x - y|$. Análogamente, $|y| - |x| \le |y - x| = |x - y|$. Por consiguiente,

$$(2.4) \qquad \big|\, |x| - |y| \,\big| \leq |x - y|.$$

Ahora estamos en condiciones de introducir la noción fundamental de este capítulo, a saber, la **distancia** del punto x al punto y se define como el número no negativo

$$|x - y| = \sqrt{(x_1 - y_1)^2 + \ldots + (x_n - y_n)^2}.$$

La relación $|x-y| \leq |x-z| + |z-y|$, que resulta de aplicar la desigualdad de Minkowski a la suma $x - y = (x - z) + (z - y)$, se llama **desigualdad triangular**, pues expresa el hecho geométrico de que en un triángulo de vértices x, y, z, la longitud de cada lado es menor que la suma de las longitudes de los otros dos.

La distancia es una función simétrica del par (x, y); es decir, $|x - y| = |y - x|$. Además, $|x - y| = 0$ si y sólo si $x = y$.

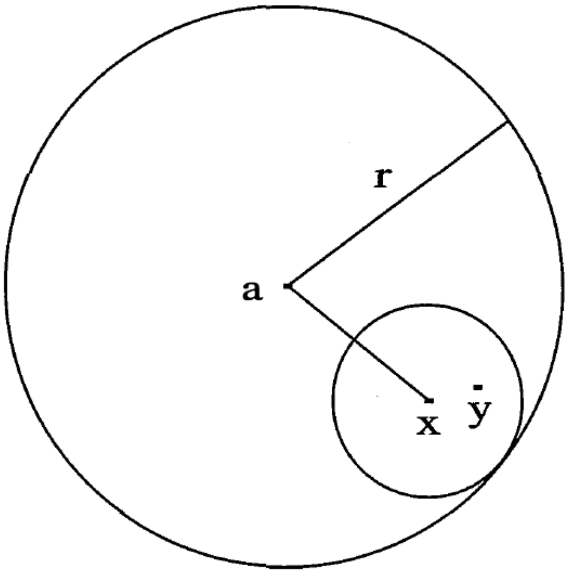
2. Conceptos topológicos.

Se dice que la sucesión $x_k = (x_{k1}, x_{k2}, \ldots, x_{kn})$, $k = 1, 2, 3, \ldots$, de puntos de $I\!R^n$ tiende al punto a del mismo espacio si $|x_k - a| \to 0$ cuando $k \to \infty$. En tal caso se dice que a es el **límite** de la sucesión y se escribe $a = \lim x_k$ o bien $x_k \to a$.

La **bola abierta** $B(a, r)$ con centro en el punto a y radio $r > 0$ se define como el conjunto de todos los puntos x tales que $|x - a| < r$.

Un conjunto $G \subset I\!R^n$ se llama **abierto** si para cada $x \in G$, existe $r > 0$, tal que $B(x, r) \subset G$.

Cada bola abierta $B(a, r)$ es un conjunto abierto. En efecto, si $x \in B(a, r)$, se tiene $\rho = |x - a| < r$ y podemos probar que $B(x, r - \rho) \subset B(a, r)$, pues si $|y - x| < r - \rho$, entonces $|y - a| \leq |y - x| + |x - a| < (r - \rho) + \rho = r$ (ver figura).



El espacio $I\!R^n$ y el conjunto vacío $\emptyset$ son abiertos. En el caso del conjunto vacío, la afirmación se basa en el hecho de que cualquier proposición de la forma "si $x \in \emptyset$, entonces ..." es verdadera por ser falso el antecedente. El lector enemistado con este tipo de razonamiento puede aceptar que el vacío es abierto por convención.

Un conjunto $F \subset I\!R^n$ se llama **cerrado** si su complemento $CF = I\!R^n - F$ es abierto.

(2.5) **Proposición.**

(a) La unión de cualquier familia de conjuntos abiertos es un conjunto abierto;

(b) La intersección de una familia finita de conjuntos abiertos es un conjunto abierto;

(c) La intersección de cualquier familia de conjuntos cerrados es un conjunto cerrado;

(d) La unión de una familia finita de conjuntos cerrados es un conjunto cerrado.

Demostración. Llamando U a la unión de la familia $(G_i, i \in I)$, cuyos miembros son abiertos, para cada $x \in U$, existe $i \in I$, tal que $x \in G_i$; y como G_i es abierto, existe $r > 0$, tal que $B(x,r) \subset G_i \subset U$, lo que prueba que U es abierto.

Sea V la intersección de la familia finita $(G_1, G_2, \ldots, G_m)$ cuyos miembros son abiertos.

Si $x \in V$, entonces $x \in G_i \quad (i = 1, 2, \ldots, m)$ y como cada G_i es abierto, existe $r_i > 0$, tal que $B(x, r_i) \subset G_i$. Llamando r al mínimo de los números r_i, tendremos $B(x,r) \subset B(x, r_i) \subset G_i$ $(1 \leq i \leq m)$, de donde $B(x,r) \subset V$, lo que prueba que V es abierto.

Habiendo probado (a) y (b), notemos que (c) y (d) se obtienen de las dos anteriores por una aplicación de las leyes de complementación.

Cualquier conjunto abierto V, tal que $x \in V$, se llama un **entorno** del punto x. En particular, cada bola $B(x,r)$, $r > 0$, es un entorno de x.

Intuitivamente, dar un entorno de x equivale a fijar un grado de proximidad a dicho punto.

Diremos que x es un **punto interior** del conjunto A si existe un entorno V de x, tal que $V \subset A$. El conjunto de los puntos interiores de A se llama el **interior** de A y se donota por A°.

De acuerdo con la definición, A° es la unión de todos los conjuntos abiertos V tales que $V \subset A$. Luego, A° es un conjunto abierto incluido en A. Más aún: si V es abierto y $V \subset A$, entonces $V \subset A^\circ$. Es decir, A° es el más grande de los conjuntos abiertos incluidos en A.

El punto x se llama un **punto de adherencia** (o de **clausura**) de A si cada entorno V de x contiene al menos un punto de A; es decir, si para cada entorno V de x, $A \cap V \neq \emptyset$. El conjunto de los puntos de adherencia de A se llama la **adherencia** (o **clausura**) de A y se denota por $\bar{A}$.

Intuitivamente, $x \in \bar{A}$ si y sólo si hay puntos de A tan próximos a x como se desee. Notemos que $A \subset \bar{A}$.

Si $x \in \bar{A}$, entonces eligiendo un punto $a_k \in A \cap B(x, 1/k)$, $k = 1, 2, 3, \ldots$, tendremos una sucesión (a_k) de puntos de A, tal que $a_k \to x$; de modo que $x \in \bar{A}$ si y sólo si existe una sucesión de puntos de A que converge al punto x.

La afirmación $x \notin \bar{A}$ es equivalente a afirmar que existe un entorno V de x, tal que $A \cap V = \emptyset$, o sea, $V \subset CA$, lo cual, a su vez, equivale a la afirmación $x \in (CA)^\circ$. Hemos probado la fórmula

$$C\bar{A} = (CA)^\circ \tag{2.6}$$

es decir, el complemento de la clausura es el interior del complemento.

Puesto que el interior de cualquier conjunto es un subconjunto abierto del mismo, de (2.6) se desprende que $\bar{A}$ es un conjunto cerrado que contiene a A. Por otra parte, si F es cerrado y $F \supset A$, entonces $CF \subset CA$ y por consiguiente, $CF \subset (CA)^\circ = C\bar{A}$, pues CF es abierto. Tomando nuevamente complementos, resulta $F \supset \bar{A}$.

Hemos probado que $\bar{A}$ es el más pequeño conjunto cerrado que contiene a A. En consecuencia, A es cerrado si y sólo si $\bar{A} = A$, para lo cual es suficiente que se verifique $\bar{A} \subset A$, pues la inclusión opuesta se verifica para cualquier conjunto. Resumiendo: un conjunto A es cerrado si y sólo si A contiene a todos sus puntos de adherencia.

Ejemplos.

1) El conjunto $\mathbb{Z}$ formado por los números enteros es cerrado en $\mathbb{R}^1$, pues su complemento es la unión de los intervalos abiertos $(a, a+1)$, $a \in \mathbb{Z}$.
2) Cualquier intervalo cerrado $[a, b]$ es un subconjunto cerrado de $\mathbb{R}^1$, del mismo modo que cualquier semirrecta $[a, \infty) = \{x \in \mathbb{R} : x \geq a\}$ o $(-\infty, b] = \{x \in \mathbb{R} : x \leq b\}$.
3) Cualquier "bola cerrada" $K(a, r) = \{x \in \mathbb{R}^n : |x - a| \leq r\}$ es un subconjunto cerrado de $\mathbb{R}^n$. En efecto, si $x \notin K(a, r)$, entonces $|x - a| = \rho > r$ y se verifica fácilmente que $B(x, \rho - r)$ está contenida en el complemento de $K(a, r)$, el cual, por consiguiente, es abierto.
4) El conjunto $A = \{1, 1/2, 1/3, 1/4, \ldots\}$ no es cerrado, pues $0 \in \bar{A}$, pero $0 \notin A$.
5) Cualquier conjunto finito $F \subset \mathbb{R}^n$ es cerrado, pues si $x \in CF$, llamando r al mínimo de los números $|x - y|$, $y \in F$, es claro que $B(x, r) \subset CF$.
6) $\mathbb{Q}$ no es cerrado en $\mathbb{R}^1$, pues la adherencia de $\mathbb{Q}$ es $\mathbb{R}^1$, ya que cualquier intervalo abierto $(a - r, a + r)$, $r > 0$, contiene puntos racionales.
7) Si $A \subset \mathbb{R}^1$ es no vacío y acotado superiormente, entonces $sup\, A \in \bar{A}$.

En el siguiente teorema la palabra "intervalo" se usa por única vez para designar un conjunto de la forma $\{x \in \mathbb{R} : a < x < b\}$, donde a y b son elementos de la recta extendida; por ejemplo, se aceptan como intervalos las semirrectas $(-\infty, b)$, (a, ∞) y aun el mismo conjunto $\mathbb{R} = (-\infty, \infty)$.

(2.7) **Teorema.** *Todo conjunto abierto de* $\mathbb{R}^1$ *es la unión de una sucesión de intervalos abiertos disjuntos.*

DEMOSTRACIÓN. Sea G un subconjunto abierto de $\mathbb{R}^1$. Un intervalo $I = (a, b)$ se llama un **intervalo componente** de G si $I \subset G$, $a \notin G$, $b \notin G$. Comenzaremos probando que cada punto de G pertenece a un intervalo componente. En efecto, si $x \in G$, pongamos $a = \sup\{y : y \leq x, y \in CG\} = \sup[(-\infty, x] \cap CG]$, con la convención de que $a = -\infty$ si el conjunto es vacío. Análogamente, definamos $b = \inf\{y : y \geq x, y \in CG\}$, con la convención de que $b = +\infty$ si el último conjunto es vacío.

Por ser x un punto de G (abierto), se sigue que $a < x < b$. Puesto que los conjuntos $(-\infty, x] \cap CG$ y $[x, \infty) \cap CG$ son cerrados, vemos que $a \notin G$ y $b \notin G$. Finalmente, es claro que $(a, b) \subset G$. Hemos probado que el intervalo $I = (a, b)$ es un intervalo componente de G, tal que $x \in I$.

Luego, G es la unión de la colección $\mathcal{C}$ formada por los intervalos componentes de G y sólo nos queda probar que $\mathcal{C}$ es disjunta y numerable. En efecto, si dos intervalos componentes (a, b) y (c, d) tuvieran un punto común, entonces necesariamente, $a = c$ y $b = d$, lo cual muestra que $\mathcal{C}$ es una colección disjunta.

Si ahora definimos una función $f : \mathcal{C} \to \mathbb{Q}$ eligiendo para cada $I \in \mathcal{C}$ un número racional $f(I) \in I$, es claro que f resulta inyectiva, en vista de que $\mathcal{C}$ es disjunta, de donde se sigue que $\mathcal{C}$ es equivalente a un subconjunto de $\mathbb{Q}$ y por lo tanto, $\mathcal{C}$ es numerable. Luego, $\mathcal{C} = \{(a_k, b_k); k = 1, 2, 3, \ldots\}$ y por todo lo dicho más arriba,

$$G = \bigcup_{k=1}^{\infty} (a_k, b_k),$$

con $(a_k, b_k) \cap (a_j, b_j) = \emptyset$ si $k \neq j$.

3. Funciones.

Sean A un subconjunto de $\mathbb{R}^n$ y x_0 un punto de A. Una función o aplicación $f : A \to \mathbb{R}^m$ se llama **continua** en el punto x_0 si para cada número $\varepsilon > 0$, existe un número $\delta > 0$, tal que las relaciones $x \in A$, $|x - x_0| < \delta$ implican $|f(x) - f(x_0)| < \varepsilon$. Es decir, si a cada ε positivo corresponde un δ también positivo, tal que $f\,(A \cap B(x_0, \delta)) \subset B\,(f(x_0), \varepsilon)$.

De la definición precedente resulta fácilmente que f es continua en x_0 si y sólo si para cada entorno U de $f(x_0)$ en el espacio $\mathbb{R}^m$, existe un entorno V de x_0 en $\mathbb{R}^n$, tal que $f(V \cap A) \subset U$.

Si f es continua en cada punto de A, decimos que f es continua sobre A, o bien que f es **continua**.

(2.8) *Para una función $f : \mathbb{R}^n \to \mathbb{R}^m$ son equivalentes las afirmaciones:*

(a) f es continua;

(b) para cada conjunto abierto U de $\mathbb{R}^m$, la imagen inversa $f^{-1}(U)$ es un subconjunto abierto de $\mathbb{R}^n$;

(c) para cada conjunto cerrado A de $\mathbb{R}^m$, la imagen inversa $f^{-1}(A)$ es un subconjunto cerrado de $\mathbb{R}^n$.

La equivalencia entre las dos primeras se demuestra fácilmente, teniendo en cuenta que la relación $f(V) \subset U$ equivale a $V \subset f^{-1}(U)$. La equivalencia entre las dos últimas se obtiene inmediatamente tomando complementos.

La función $f : A \to \mathbb{R}^m$ se llama **uniformemente continua** sobre A si para $\varepsilon > 0$, existe $\delta > 0$, tal que las relaciones $x \in A$, $y \in A$, $|x - y| < \delta$, implican $|f(x) - f(y)| < \varepsilon$.

Más adelante resultará conveniente considerar funciones f con valores en la recta extendida $\overline{\mathbb{R}}$, de modo que (además de los números reales) $-\infty$ y $+\infty$ son valores posibles para f. En tal caso diremos que f es **finita** sobre E si $|f(x)| < \infty$ para cada $x \in E$. Una función se llama **finita** cuando es finita en cualquier punto.

Tratándose de una función $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ con valores en la recta extendida, para cada punto de x_0 de $\mathbb{R}^n$ y cada $\delta > 0$, definimos:

$$M'_\delta(x_0) = \sup\{f(x) : 0 < |x - x_0| < \delta\},$$

$$m'_\delta(x_0) = \inf\{f(x) : 0 < |x - x_0| < \delta\}.$$

La "prima" se usa para destacar la exclusión del punto x_0 (comparar con el ejercicio 10).

El **límite superior** L y el **límite inferior** l de la función f en el punto x_0 se definen ahora por medio de las fórmulas:

$$L = \limsup_{x \to x_0} f(x) = \inf_{\delta > 0} M'_\delta(x_0),$$

$$l = \liminf_{x \to x_0} f(x) = \sup_{\delta > 0} m'_\delta(x_0).$$

El número L se caracteriza por la siguiente propiedad: si $s < L < t$, entonces 1°) existe $\delta > 0$, tal que $0 < |x - x_0| < \delta$ implica $f(x) < t$ y 2°) para cada $\delta > 0$, existe al menos un punto x, tal que $0 < |x - x_0| < \delta$ y además, $f(x) > s$.

Simétricamente, el número l se caracteriza por la siguiente propiedad: si $s < l < t$, entonces 1°) existe $\delta > 0$, tal que $0 < |x - x_0| < \delta$ implica $f(x) > s$ y 2°) para cada $\delta > 0$, existe al menos un punto x, tal que $0 < |x - x_0| < \delta$ y además, $f(x) < t$.

La demostración de estas propiedades es un ejercicio muy instructivo sobre supremos e ínfimos que dejaremos a cargo del lector.

De las propiedades enunciadas se deduce que $l \le L$. Cuando estos dos valores coinciden, es decir, cuando se verifica $l = L$, decimos que f tiene **límite** en el punto x_0 o bien que $f(x)$ tiende al valor $l = L$ cuando x tiende a x_0 y escribimos $\lim_{x \to x_0} f(x) = L$.

Afirmar que f es continua en el punto x_0 equivale a decir que $f(x_0)$ es finito y además, $\lim_{x \to x_0} f(x) = f(x_0)$.

Diremos que la función $f : I\!R^n \to \overline{I\!R}$ es **semicontinua superiormente** (abreviado, s.s.) en el punto x_0, si se verifica

$$\limsup_{x \to x_0} f(x) \le f(x_0).$$

Si f es s.s. en cada punto de $I\!R^n$, entonces decimos que f es s.s.

Por ejemplo, la función de Dirichlet $f : I\!R \to I\!R$, definida por $f(x) = 1$ si x es racional y $f(x) = 0$ si x es irracional, es s.s. en cada punto racional, pero no lo es en puntos irracionales; la función $g : I\!R \to I\!R$ definida por $g(x) = 1$ si $x \ge 0$ y $g(x) = 0$ si $x < 0$ es s.s.

Afirmar que f es s.s. en el punto x_0, equivale a decir que para cada $t > f(x_0)$, existe un entorno V de x_0, tal que la relación $x \in V$ implica $f(x) < t$.

Simétricamente, diremos que f es **semicontinua inferiormente** (abreviado s.i.) en el punto x_0, si se verifica

$$\lim_{x \to x_0} \inf f(x) \geq f(x_0).$$

Si f es s.i. en cada punto de $\mathbb{R}^n$, entonces decimos que f es s.i.

Si $f(x_0)$ es finito, entonces f es continua en x_0 si y sólo si f es s.s. y s.i. en dicho punto. Por consiguiente, una función finita f es continua si y sólo si f es s.s. y s.i.

(2.9) *Para una función* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *son equivalentes las afirmaciones:*

(a) f *es semicontinua superiormente;*

(b) para cada $t \in \overline{\mathbb{R}}$, $\{x \in \mathbb{R}^n : f(x) < t\}$ *es abierto;*

(c) para cada $t \in \overline{\mathbb{R}}$, $\{x \in \mathbb{R}^n : f(x) \geq t\}$ *es cerrado.*

Dejaremos a cargo del lector la fácil demostración y el enunciado de la proposición análoga para la semicontinuidad en el otro sentido.

En el caso de una función $f : \mathbb{R} \to \mathbb{R}$, decimos que f tiene **límite por la derecha** igual a t en el punto x_0, si para cada $\varepsilon > 0$, existe un número $\delta > 0$, tal que la relación $x_0 < x < x_0 + \delta$ implica $|f(x) - t| < \varepsilon$. En tal caso, el número t se denota por $f(x_0+)$ o bien por $\lim_{x \to x_0+} f(x)$.

Análogamente, el número s se llama el **límite por la izquierda** de f en el punto x_0, si para cada $\varepsilon > 0$, existe un número $\delta > 0$, tal que la relación $x_0 - \delta < x < x_0$ implica $|f(x) - s| < \varepsilon$. Y en este caso, el número s se denota por $f(x_0-)$ o bien por $\lim_{x \to x_0-} f(x)$.

Una función f con valores reales, definida sobre un intervalo de la recta se llama **monótoma creciente** en dicho intervalo, si la relación $x_1 < x_2$ implica $f(x_1) \leq f(x_2)$. Simétricamente se define el concepto de función monótona decreciente.

Para una función monótona (creciente o decreciente), los límites laterales $f(x+)$ y $f(x-)$ existen en cada punto interior de su dominio. Por ejemplo, si f es creciente, se verifica fácilmente que

$$f(x-) = \sup_{s<x} f(s) \quad \text{y} \quad f(x+) = \inf_{t>x} f(t);$$

además, $f(x-) \leq f(x) \leq f(x+)$; de modo que f es continua en x si y sólo si estos tres valores coinciden.

Suponiendo que x e y son dos puntos de discontinuidad de la función creciente f, tales que $x < y$, tendremos $f(x-) < f(x+) \leq f(y-) < f(y+)$, lo cual implica que los intervalos abiertos no vacíos

$$(f(x-), f(x+)) \quad \text{y} \quad (f(y-), f(y+))$$

no tienen ningún punto común. Por consiguiente, llamando $D(f)$ al conjunto formado por todos los puntos donde f es discontinua, la colección formada por todos los intervalos abiertos no vacíos

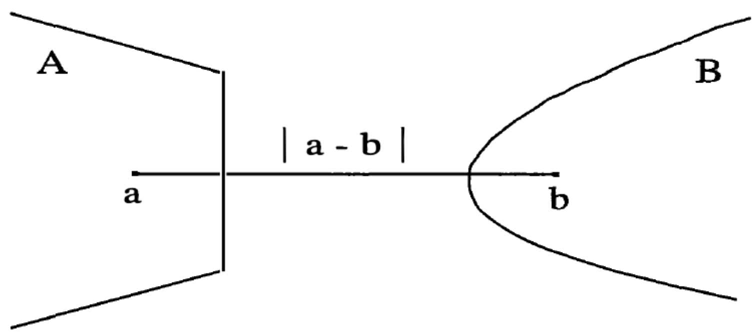
$$(f(x-), f(x+)) \qquad (x \in D(f)),$$

es disjunta; pero sabemos que una colección de intervalos con estas propiedades es forzosamente numerable. Luego, $D(f)$ es numerable.

Hemos demostrado que el conjunto formado por todos los puntos de discontinuidad de una función monótona es numerable.

4. Distancia y diámetro; conjuntos acotados.

La **distancia** entre dos conjuntos no vacíos A y B se define como el ínfimo del conjunto de números $|a - b|$, tales que $a \in A$ y $b \in B$.



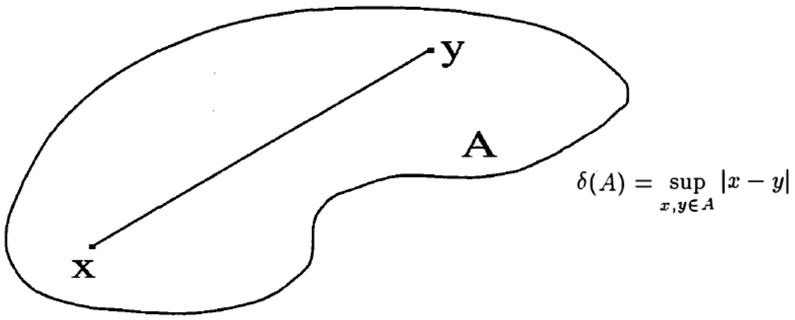
Para denotarla, usaremos el símbolo $d(A, B)$. En particular, la distancia del punto x al conjunto no vacío A es, por definición, el número $d(x, A) = d(\{x\}, A) = \inf\{|x-a| : a \in A\}$. Notemos que $x \in \bar{A}$ si y sólo si $d(x, A) = 0$.

Es útil conocer la relación

$$|d(x,A) - d(y,A)| \leq |x-y|, \tag{2.8}$$

la cual implica que $d(x,A)$ es una función uniformemente continua de x. Para demostrarla, observemos que para cualquier $a \in A$, $|x-a| \leq |x-y| + |y-a|$; luego, $d(x,A) \leq |x-y| + |y-a|$, de donde, tomando el ínfimo del miembro derecho para todos los $a \in A$, resulta $d(x,A) \leq |x-y| + d(y,A)$; o sea, $d(x,A) - d(y,A) \leq |x-y|$. Si en esta relación permutamos x con y, obtenemos $d(y,A) - d(x,A) \leq |x-y|$, la cual, juntamente con la anterior, demuestra (2.8).

El **diámetro** de un conjunto no vacío A se define como el supremo del conjunto de todos los números de la forma $|x-y|$, donde x e y son elementos de A. El símbolo $\delta(A)$ será usado para designar el diámetro del conjunto A.



Por ejemplo, el diámetro de la bola $B(a,r)$ es $2r$. En efecto, si x e y son puntos de dicha bola, entonces $|x-y| \leq |x-a| + |a-y| < 2r$; de modo que $\delta\,(B(a,r)) \leq 2r$. Por otro lado, si consideramos el vector $u = (1, 0, \ldots, 0)$, es claro que $|u| = 1$ y para cualquier entero positivo k, los vectores $a \pm k^{-1}(k-1)ru$ son puntos de $B(a,r)$ cuya diferencia tiene módulo igual a $2(k-1)r/k$. Luego, $\delta\,(B(a,r)) \geq \sup_k 2(k-1)r/k = 2r$.

Si A y B son no vacíos, se tiene

$$\delta(A \cup B) \leq \delta(A) + \delta(B) + d(A,B). \tag{2.9}$$

En efecto, si x e y son puntos de $A \cup B$, $a \in A$, $b \in B$, suponiendo primero que $x \in A$ e $y \in B$, tendremos $|x-y| \leq |x-a| + |a-b| + |b-y| \leq$

$\delta(A) + \delta(B) + |a - b|$ y esta desigualdad se mantiene válida aunque x e y pertenezcan a un mismo conjunto (A o B). Luego, $\delta(A \cup B) \leq \delta(A) + \delta(B) + |a - b|$ para cualquier par de puntos $a \in A$, $b \in B$, de donde resulta (2.9).

El conjunto no vacío A se llama **acotado** si $\delta(A) < \infty$, lo que equivale a afirmar que A está contenido en alguna bola $B(a, r)$.

La relación (2.9) muestra que la unión de dos conjuntos acotados es acotada y por inducción se demuestra que cualquier unión finita de conjuntos acotados es un conjunto acotado. En particular, cualquier conjunto finito es acotado.

5. Conjuntos convexos.

Un conjunto A se llama **convexo** si juntamente con cada par de puntos $x \in A$, $y \in A$, el conjunto A contiene a todo el segmento que une x con y.

Cada bola $B(a, r)$ es un conjunto convexo, pues si $z = (1 - \lambda)x + \lambda y$, donde $x, y \in B(a, r)$, $0 \leq \lambda \leq 1$, entonces $|z - a| = |(1 - \lambda)(x - a) + \lambda(y - a)| \leq |(1 - \lambda)(x - a)| + |\lambda(y - a)| = (1 - \lambda)|x - a| + \lambda|y - a| < (1 - \lambda)r + \lambda r = r$.

Cualquier intersección de conjuntos convexos es un conjunto convexo (probarlo). La intersección de todos los conjuntos convexos que contienen al conjunto A se llama la **cápsula convexa** de A y se denota por $\hat{A}$. De modo que $\hat{A}$ es un conjunto convexo que contiene a A y si C es convexo y $A \subset C$, entonces $\hat{A} \subset C$.

Por ejemplo, si $F = \{a_0, a_1, \ldots, a_n\}$ es un conjunto formado por $n + 1$ puntos de $I\!R^n$, tales que los vectores $a_k - a_0 \quad (k = 1, 2, \ldots, n)$ son linealmente independientes, la cápsula convexa de F se llama el **simple** con vértices $a_k \quad (k = 0, 1, \ldots, n)$. Este conjunto consta de todos los puntos x de la forma $x = t_0 a_0 + t_1 a_1 + \ldots + t_n a_n$, donde cada t_k es no negativo y $\sum t_k = 1$ (ejercicio 5).

El siguiente ejemplo lo consignamos aquí por parecernos a la vez sencillo e interesante.

(2.10) *El diámetro de cada conjunto es igual al diámetro de su cápsula convexa.*

Puesto que $A \subset \hat{A}$, $\delta(A) \leq \delta(\hat{A})$. Por otro lado, si $r > \delta(A)$, entonces para cualquier par x, y de elementos de A, $|y-x| < r$ de donde $A \subset B(x,r)$ y como la bola es convexa, tenemos

$$\hat{A} \subset B(x,r) \qquad (x \in A).$$

De aquí se sigue que para cualquier $u \in \hat{A}$ y cualquier $x \in A$, se tiene $|u-x| < r$, de donde $A \subset B(u,r)$ y por consiguiente,

$$\hat{A} \subset B(u,r) \qquad (u \in \hat{A}).$$

Luego $|v-u| < r$ para cualquier par de puntos $v \in \hat{A}$, $u \in \hat{A}$, lo cual prueba que $\delta(\hat{A}) \leq r$ para cada $r > \delta(A)$; es decir, $\delta(\hat{A}) \leq \delta(A)$.

Más detalles sobre cápsulas y conjuntos convexos hallará el lector en los ejercicios al final del capítulo.

6. Intervalos.

Llamaremos **intervalo** a cada conjunto $I \subset \mathbb{R}^n$ que pueda expresarse como el producto cartesiano de n intervalos lineales. Por consiguiente, I es un intervalo de $\mathbb{R}^n$ si y sólo si existen intervalos de la recta $I_1, \dots, I_n$, tales que $I = I_1 \times \dots \times I_n$.

Un intervalo consta de todos los puntos $x = (x_1, \dots, x_n)$ que satisfacen en cada coordenada una desigualdad de alguno de los siguientes tipos:

(1) $a_i \leq x_i \leq b_i$

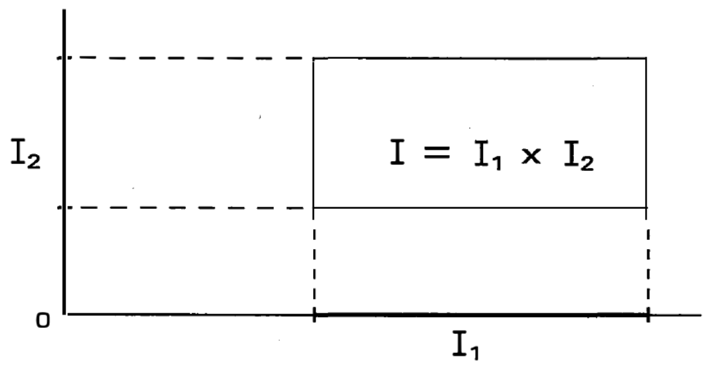
(2) $a_i < x_i \leq b_i$

(3) $a_i \leq x_i < b_i$

(4) $a_i < x_i < b_i$

Si todas las desigualdades son de la forma (1), el intervalo se llama **cerrado**; si todas ellas son de la forma (4), el intervalo se llama **abierto**.

En el caso de la recta $(n = 1)$, la representación geométrica de un intervalo cerrado es un segmento; en el caso del plano $(n = 2)$, un rectángulo de lados paralelos a los ejes, como se muestra en la figura; y en el caso del espacio tridimensional $(n = 3)$, un paralelepípedo de aristas paralelas a los ejes.



Cada uno de los intervalos lineales $I_1, \dots, I_n$ se llama un **lado** del intervalo $I = I_1 \times \dots \times I_n$. Obviamente I es vacío si y sólo si lo es alguno de sus lados; por consiguiente, el conjunto vacío es un intervalo de $I\!R^n$.

Un intervalo cuyos lados tienen todos la misma longitud se llama un **cubo**.

Dados un punto $a = (a_1, \dots, a_n)$ y un número $\varepsilon > 0$, el conjunto

$$Q(a, \varepsilon) = \{x \in I\!R^n : |x_i - a_i| < \varepsilon \ \ (i = 1, 2, \dots, n)\}$$

representa un cubo abierto cuyos lados son los intervalos lineales $(a_i - \varepsilon, a_i + \varepsilon)$, cada uno con longitud igual a 2ε. El punto a se llama el **centro** del cubo $Q(a, \varepsilon)$.

(2.11) *Toda bola abierta* $B(a,r)$ *contiene un cubo* $Q(a,\varepsilon)$ *con el mismo centro y viceversa.*

En efecto, tratemos de encontrar un número $\varepsilon > 0$ de modo tal que se cumpla $Q(a,\varepsilon) \subset B(a,r)$. Si $x \in Q(a,\varepsilon)$, tendremos

$$|x-a|^2 = (x_1-a_1)^2 + \ldots + (x_n-a_n)^2 < n\varepsilon^2.$$

Poniendo $n\varepsilon^2 = r^2$, resulta $\varepsilon = r/\sqrt{n}$. Luego, $Q(a, r/\sqrt{n}) \subset B(a,r)$.

Puesto que para cada $i = 1,2,\ldots,n$ se cumple $|x_i - a_i| \leq |x-a|$, se sigue que $B(a,\varepsilon) \subset Q(a,\varepsilon)$.

(2.12) **Corolario.** *Todo intervalo abierto es un conjunto abierto. Todo intervalo cerrado es un conjunto cerrado.*

Sea $z = (z_1,\ldots,z_n)$ un punto del intervalo abierto $J = (a_1,b_1) \times \ldots \times (a_n,b_n)$. Luego, $a_i < z_i < b_i \quad (i = 1,2,\ldots,n)$ y podemos elegir un número $\varepsilon > 0$, tal que

$$a_i < z_i - \varepsilon < z_i + \varepsilon < b_i \qquad (i = 1,2,\ldots,n),$$

de donde se sigue que $Q(z,\varepsilon) \subset J$ y por consiguiente, $B(z,\varepsilon) \subset J$, lo cual prueba que J es un conjunto abierto.

Para probar que el intervalo cerrado $I = [a_1,b_1] \times \ldots \times [a_n,b_n]$ es un conjunto cerrado, lo más sencillo es mostrar que su complemento CI es un conjunto abierto, tarea que dejaremos como ejercicio a cargo del lector.

El principio de encaje de intervalos cerrados es válido para intervalos de $I\!R^n$: dada una sucesión decreciente de intervalos cerrados no vacíos

$$I_1 \supset I_2 \supset I_3 \supset \ldots \tag{1}$$

hay un punto $x \in I\!R^n$ que pertenece a cada intervalo de la sucesión. En efecto, si $I_k = I_k^1 \times \ldots \times I_k^n$, donde cada I_k^j es un intervalo lineal cerrado, en virtud de la hipótesis, para cada $j = 1,2,\ldots,n$, tendremos $I_1^j \supset I_2^j \supset I_3^j \supset \ldots$ y en virtud del principio de encaje para intervalos de la recta, existe un número real x_j que pertenece a cada intervalo de esta última sucesión. Luego, el punto $x = (x_1,\ldots,x_n)$ pertenece a cada intervalo de la sucesión (1).

Definición. *Un intervalo cuyo interior es vacío se llamará* **degenerado.**

Por ejemplo, en el plano $\mathbb{R}^2$, cada intervalo degenerado es vacío, o bien un segmento paralelo a uno de los ejes, o bien un conjunto unitario.

7. Cubrimientos abiertos; conjuntos compactos.

Una colección de conjuntos Γ se llama un **cubrimiento** del conjunto A si cada punto de A pertenece al menos a un miembro de Γ; es decir, si A está contenido en la unión de Γ. Un cubrimiento Γ se llama **abierto** si cada miembro de Γ es un conjunto abierto.

Diremos que el conjunto $K \subset \mathbb{R}^n$ es un conjunto **compacto** si para cada cubrimiento abierto Γ de K, existe una colección finita $\Gamma_0 \subset \Gamma$, tal que Γ_0 es un cubrimiento de K. En otras palabras: K es compacto si y sólo si cada cubrimiento abierto de K contiene un cubrimiento finito de dicho conjunto.

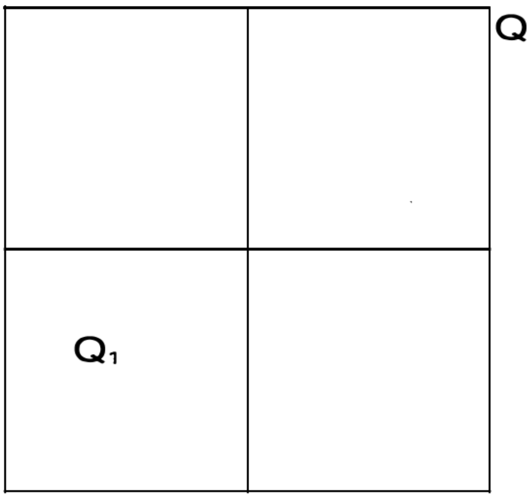
(2.13) *Todo cubo cerrado es un conjunto compacto.*

Sea Q un cubo cerrado en el espacio $\mathbb{R}^n$ y sea Γ un cubrimiento abierto de Q. Debemos probar que existe una colección finita $\Gamma_0 \subset \Gamma$ que es también un cubrimiento de Q.

Supongamos lo contrario; es decir, supongamos que ninguna colección finita de miembros de Γ sea suficiente para cubrir Q.

Dividiendo cada lado de Q en dos intervalos cerrados de igual longitud (con el punto medio de cada segmento como único punto común entre ambos subintervalos) podemos expresar Q como la unión de 2^n cubos cerrados con diámetro igual a la mitad del diámetro de Q.

Si cada uno de estos cubos más pequeños pudiera cubrirse con una colección finita de miembros de Γ, entonces todo Q podría cubrirse con una colección finita $\Gamma_0 \subset \Gamma$, contrariamente a lo que hemos supuesto. Luego, al menos uno de estos cubos más pequeños, al cual llamaremos Q_1, tiene la propiedad de que ninguna colección finita de miembros de Γ es un cubrimiento de Q_1.



Subdividiendo Q_1 en la misma forma que el cubo anterior, encontraremos un cubo cerrado $Q_2 \subset Q_1$ con diámetro $\delta(Q_2) = (1/2)\delta(Q_1) = (1/4)\delta(Q)$ y con la propiedad de que ninguna colección finita de miembros de Γ es un cubrimiento de Q_2.

Subdividiendo Q_2 en la misma forma y prosiguiendo con el mismo proceso de selección, tendremos una sucesión decreciente de cubos cerrados:

$$Q \supset Q_1 \supset Q_2 \supset Q_3 \supset \dots \tag{1}$$

ninguno de los cuales se puede cubrir con una colección finita de miembros de Γ; además, $\delta(Q_k) \to 0$ cuando $k \to 0$.

En virtud del principio de encaje de intervalos cerrados, existe un punto x que pertenece a cada cubo de la sucesión (1); en particular $x \in Q$. Ahora bien; por ser Γ un cubrimiento abierto de Q, existe un conjunto abierto $G \in \Gamma$, tal que $x \in G$ y por ser G un conjunto abierto, existe un número $r > 0$, tal que $B(x, r) \subset G$.

Eligiendo k suficientemente grande, tendremos $\delta(Q_k) < r$ y como $x \in Q_k$, se verifica $Q_k \subset B(x, r)$, de donde $Q_k \subset G$: el cubo Q_k cubierto por un solo miembro de Γ, en contradicción con la forma en que elegimos los cubos de la sucesión (1).

La contradicción provino de suponer que ninguna colección finita $\Gamma_0 \subset \Gamma$ es un cubrimiento de Q. Luego, alguna colección finita $\{G_1, \ldots, G_k\} \subset \Gamma$ verifica $Q \subset G_1 \cup \ldots \cup G_k$, que es lo que queríamos demostrar.

(2.14) *Cualquier conjunto cerrado y acotado es un conjunto compacto.*

Consideremos un conjunto A cerrado y acotado y sea Γ un cubrimiento abierto de A. Por ser A acotado, existe un cubo cerrado Q, tal que $A \subset Q$ y la colección $\Gamma' = \Gamma \cup \{CA\}$ es un cubrimiento abierto de Q (el complemento de A es abierto por ser A un conjunto cerrado).

Siendo Q compacto, existe una colección finita $\{G_1, \ldots G_k\} \subset \Gamma$, tal que

$$Q \subset G_1 \cup \ldots \cup G_k \cup CA,$$

de donde se sigue que $A \subset G_1 \cup \ldots \cup G_k$.

(2.15) *Un conjunto no puede ser compacto a menos que sea cerrado y acotado.*

Si A no es cerrado, existe un punto b, tal que $b \in \bar{A}$, $b \notin A$.

Puesto que $|x - b|$ (la distancia del punto x al punto b) es una función continua de x, los conjuntos

$$U_k = \left\{ x : |x - b| > \frac{1}{k} \right\} \qquad (k = 1, 2, 3, \ldots)$$

son abiertos, forman una sucesión creciente: $U_1 \subset U_2 \subset U_3 \subset \ldots$, y cubren A. Sin embargo, ninguno de ellos cubre totalmente al conjunto A, en vista de que $b \in \bar{A}$. Luego, el cubrimiento abierto (U_k) no contiene ningún cubrimiento finito de A.

Si A no es acotado, los conjuntos $B(0, k) = \{x : |x| < k\}$ (bola con centro en el origen y radio k) forman un cubrimiento abierto de todo el espacio, en particular de A, que no contiene ningún cubrimiento finito de A.

(2.16) *Si A es un conjunto compacto del espacio $I\!R^n$ y $f : A \to I\!R^m$ una función continua, entonces $f(A)$ es compacto.*

En efecto si Γ es un cubrimiento abierto de $f(A)$, entonces para cada $x \in A$ existe un conjunto $U \in \Gamma$, tal que $f(x) \in U$, y como f es continua, existe un entorno V del punto x, tal que $f(V \cap A) \subset U$. Luego, la colección Γ' formada por todos los conjuntos abiertos V de $I\!R^n$, tales que $f(V \cap A)$ está contenido en algún miembro de Γ es un cubrimiento abierto de A. Puesto que A es compacto, existe una colección finita $\{V_1, \ldots, V_r\} \subset \Gamma'$ que es un cubrimiento de A, de donde se sigue que

$$A = (V_1 \cap A) \cup (V_2 \cap A) \cup \ldots \cup (V_r \cap A).$$

Por otra parte, para cada i, $1 \leq i \leq r$, existe un conjunto $U_i \in \Gamma$, tal que $f(V_i \cap A) \subset U_i$. Por consiguiente (ejercicio 3, cap. I),

$$f(A) = \bigcup_{i=1}^{r} f(V_i \cap A) \subset \bigcup_{i=1}^{r} U_i.$$

Hemos demostrado que la colección finita $\{U_1, \ldots, U_r\} \subset \Gamma$ es un cubrimiento de $f(A)$, lo cual prueba que este conjunto es compacto.

Corolario. *Toda función continua con valores reales cuyo dominio es un conjunto compacto es acotada y alcanza en sendos puntos del dominio un valor máximo y uno mínimo.*

En efecto, si A es un subconjunto compacto de $I\!R^n$ y $f : A \to I\!R$ es continua, el conjunto $B = f(A)$ es un subconjunto compacto de $I\!R$; por consiguiente, acotado y cerrado. Luego, $\sup B \in B$ e $\inf B \in B$.

(2.17) *Toda función continua cuyo dominio es un subconjunto compacto de $I\!R^n$ es uniformemente continua en dicho conjunto.*

Sea A un subconjunto compacto de $I\!R^n$ y $f : A \to I\!R^m$ continua. Debemos probar que para cada $\varepsilon > 0$ existe un número $\delta > 0$, tal que las relaciones $x \in A$, $y \in A$, $|x - y| < \delta$, implican $|f(x) - f(y)| < \varepsilon$.

Puesto que f es continua, para cada punto x de A existe un número $r(x) > 0$, tal que las relaciones $y \in A$, $|y - x| < r(x)$, implican $|f(y) - f(x)| < \varepsilon$.

Puesto que la colección formada por todas las bolas $B(x, r(x)/2)$, $x \in A$, es un cubrimiento abierto de A, existe una sucesión finita $x_1, x_2, \ldots, x_N$

de puntos de A, tal que

$$A \subset \bigcup_{k=1}^{N} B(x_k, r(x_k)/2).$$

Llamando δ al mínimo de los números $r(x_k)/2 \quad (k = 1, 2, \ldots, N)$, supongamos que x e y son puntos de A que verifican $|x - y| < \delta$.

Puesto que $y \in A$, existe un índice k, $1 \leq k \leq N$, tal que

$$|y - x_k| < r(x_k)/2,$$

de donde $|x - x_k| \leq |x - y| + |y - x_k| < r(x_k)/2 + r(x_k)/2 = r(x_k)$. Luego

$$|f(x) - f(x_k)| < \varepsilon, \qquad |f(y) - f(x_k)| < \varepsilon$$

y por consiguiente, $|f(x) - f(y)| < 2\varepsilon$.

8. Conjuntos elementales.

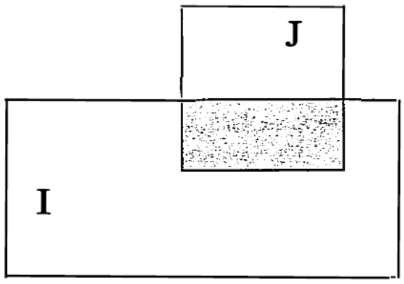
La teoría de la medida en espacios euclidianos, que desarrollaremos en el capítulo siguiente, se basa en las propiedades de los intervalos y de los conjuntos elementales que vamos a definir en esta sección. Por este motivo nos proponemos hacer un estudio detallado de dichas propiedades.

La intersección de dos intervalos de $\mathbb{R}^n$ es un intervalo de dicho espacio. En efecto, si $I = I_1 \times \ldots \times I_n$ y $J = J_1 \times \ldots \times J_n$ son dos intervalos de $\mathbb{R}^n$, tendremos

$$I \cap J = (I_1 \cap J_1) \times \ldots \times (I_n \cap J_n).$$

Ahora bien; en virtud de la proposición (1.1), cada una de las intersecciones $I_k \cap J_k \quad (k = 1, 2, \ldots, n)$ es un intervalo de la recta, lo cual prueba que $I \cap J$ es un intervalo de $\mathbb{R}^n$.

En la figura que sigue hemos representado la intersección de dos intervalos I y J del plano $\mathbb{R}^2$.



Si I y J son intervalos de $I\!R^n$, la diferencia $I - J$ no es en general un intervalo, como puede apreciarse enseguida a simple vista. Sin embargo, siempre es posible expresar $I - J$ como una unión finita de intervalos disjuntos.

Es un hecho curioso que la demostración rigurosa de esta propiedad intuitivamente clara no sea tan inmediata como podría esperarse.

Dada la importancia que este hecho tiene en el capítulo que sigue, vamos a dar una demostración correcta; pero antes debemos introducir la siguiente definición.

Definición. *Un conjunto $A \subset I\!R^n$ se llama un* **conjunto elemental** *si existen intervalos disjuntos $I_1, \ldots, I_N$ del espacio $I\!R^n$, tales que*

$$A = \bigcup_{k=1}^{N} I_k.$$

De la definición se sigue que la unión de dos conjuntos elementales disjuntos es un conjunto elemental. Además, si

$$A = \bigcup_{k=1}^{N} I_k \quad \text{y} \quad B = \bigcup_{i=1}^{M} J_i$$

son conjuntos elementales, la fórmula

$$A \cap B = \bigcup_{k=1}^{N} \bigcup_{i=1}^{M} (I_k \cap J_i)$$

nos muestra que la intersección de dos conjuntos elementales es siempre otro conjunto elemental. Más generalmente, cualquier intersección finita de conjuntos elementales es un conjunto elemental.

(2.18) *Si A es un conjunto elemental de $\mathbb{R}^n$ y B un conjunto elemental de $\mathbb{R}^m$, entonces $A \times B$ es un conjunto elemental de $\mathbb{R}^{n+m}$.*

En efecto, sean

$$A = \bigcup_{k=1}^{N} I_k \quad \text{y} \quad B = \bigcup_{i=1}^{M} J_i,$$

donde los I_k son intervalos disjuntos de $\mathbb{R}^n$, en tanto que los J_i son intervalos disjuntos de $\mathbb{R}^m$. Por un lado tenemos

$$A \times B = \bigcup_{k=1}^{N} \bigcup_{i=1}^{M} (I_k \times J_i).$$

Por otro, los productos $I_k \times J_i$ son intervalos de $\mathbb{R}^{n+m}$ que forman una familia disjunta, pues si $k \neq k'$ o bien $i \neq i'$, se verifica

$$(I_k \times J_i) \cap (I_{k'} \times J_{i'}) = (I_k \cap I_{k'}) \times (J_i \cap J_{i'}) = \emptyset.$$

Ahora estamos en condiciones de probar la siguiente proposición.

(2.19) *Si I y J son interevalos de $\mathbb{R}^n$, la diferencia $I - J$ es un conjunto elemental.*

Demostración. La desmostración es por inducción sobre la dimensión n. El caso $n = 1$ está contenido en la proposición (1.1).

Si I y J son intervalos de $\mathbb{R}^n$, podemos representarlos en la forma

$$I = I_1 \times I' \quad \text{y} \quad J = J_1 \times J',$$

donde I_1 y J_1 son intervalos lineales, en tanto que I' y J' son intervalos de $\mathbb{R}^{n-1}$. Escribiendo cada punto x de $\mathbb{R}^n$ en la forma $x = (x_1, x')$ con $x_1 \in \mathbb{R}$ y $x' \in \mathbb{R}^{n-1}$, es fácil deducir la fórmula

$$I - J = [(I_1 - J_1) \times I'] \cup [(I_1 \cap J_1) \times (I' - J')] .$$

Ahora bien; si (2.19) es verdadera en el espacio $I\!R^{n-1}$, la diferencia $I' - J'$ es un conjunto elemental de este espacio y en virtud de (2.18), la última fórmula nos muestra que $I - J$ es la unión de dos conjuntos elementales disjuntos, de donde resulta que $I - J$ es un conjunto elemental.

(2.20) *Si A y B son conjuntos elementales del espacio $I\!R^n$, la diferencia $A - B$ es también un conjunto elemental.*

Demostración. Supongamos que A es la unión de los intervalos disjuntos $I_1, \ldots, I_N$, en tanto que B es la unión de los intervalos disjuntos $J_1, \ldots, J_M$. Luego, $A - B$ es la unión de los conjuntos disjuntos $I_k - B \quad (k = 1, 2, \ldots, N)$ y en virtud de lo dicho anteriormente será suficiente con mostrar que cada uno de estos conjuntos es un conjunto elemental. En efecto, en virtud de las leyes de complementación,

$$I_k - B = I_k - \bigcup_{i=1}^{M} J_i = \bigcap_{i=1}^{M} (I_k - J_i).$$

Luego, $A - B$ es un conjunto elemental.

Corolario. *Si A y B son conjuntos elementales de $I\!R^n$, entonces la unión $A \cup B$ es también un conjunto elemental.*

En efecto, $A \cup B = A \cup (B - A)$ es la unión de dos conjuntos elementales disjuntos. Los conjuntos elementales más sencillos son los intervalos. Cada conjunto elemental es acotado, de donde se sigue que el complemento de un conjunto elemental no puede ser nunca otro conjunto elemental.

Consideremos ahora un intervalo I de $I\!R^n$ que representaremos en la forma $I = I_1 \times I'$ donde como antes, I_1 es un intervalo lineal e I' un intervalo de $I\!R^{n-1}$. Si J_1 y J_2 son dos intervalos lineales disjuntos cuya unión es I_1, tendremos

$$I = (J_1 \cup J_2) \times I' = (J_1 \times I') \cup (J_2 \times I') = H \cup L,$$

donde $H = J_1 \times I'$ y $L = J_2 \times I'$ son dos intervalos disjuntos del espacio $I\!R^n$. La siguiente proposición afirma que esta es, esencialmente, la única forma de partir un intervalo de $I\!R^n$ en dos intervalos disjuntos.

(2.21) *Cualquier partición de un intervalo I en dos intervalos disjuntos H y L, resulta de dividir un lado de I en dos intervalos lineales disjuntos.*

Demostración. La demostración se realiza nuevamente por inducción. Cuando la dimensión n es igual a uno, la verdad de (2.21) no requiere demostración.

Supongamos que el intervalo I del espacio $I\!R^n$ es la unión de los intervalos no vacíos y disjuntos H y L. Escribiendo al modo de antes

$$I = I_1 \times I', \quad H = H_1 \times H', \quad L = L_1 \times L',$$

el lema (1.4) nos dice que debe verificarse una de las siguientes alternativas:

1a) $I_1 = H_1 \cup L_1$ con $H_1 \cap L_1 = \emptyset$ y $H' = L' = I'$,

2a) $I' = H' \cup L'$ con $H' \cap L' = \emptyset$ y $H_1 = L_1 = I_1$.

En caso de verificarse la primera, el lado dividido resulta ser I_1. Supongamos que se verifica la segunda. Si (2.21) es verdadera en el espacio $I\!R^{n-1}$, puesto que los lados de $I' = I_2 \times \ldots \times I_n$ son lados de I, la proposición resulta verdadera también en el espacio $I\!R^n$, lo cual completa la inducción.

9. Hiperplanos y semiespacios.

El conjunto de todos los puntos $x = (x_1, \ldots, x_n)$ del espacio $I\!R^n$ que satisfacen una ecuación de la forma

$$f(x) = \sum_{k=1}^{n} a_k x_k = c,$$

donde c y los a_k son números reales fijos y al menos uno de los a_k es distinto de cero, se llama un **hiperplano**. Cada uno de los subconjuntos de $I\!R^n$ determinados por las relaciones

$$f(x) \leq c, \quad f(x) > c, \quad f(x) < c, \quad f(x) \geq c,$$

se llama un **semiespacio** correspondiente a dicho hiperplano.

Los semiespacios $f(x) \leq c$ y $f(x) > c$ se llaman **complementarios** por ser cada uno de ellos el complemento del otro; análogamente para los semiespacios $f(x) < c$ y $f(x) \geq c$.

La ecuación $x_k = c$ representa un hiperplano ortogonal al eje x_k; más precisamente, ortogonal al vector unitario $e_k = (0, \ldots, 0, 1, 0, \ldots, 0)$, con el uno en la k-ésima coordenada.

(2.22) *La intersección de un intervalo I de $\mathbb{R}^n$ con uno cualquiera de los semiespacios $x_k \leq c, x_k > c, x_k < c, x_k \geq c$ es un intervalo.*

Consideremos, para fijar ideas, la intersección de $I = I_1 \times I_2 \times \ldots \times I_n$ con el semiespacio $S = \{x : x_1 \leq c\}$. Claramente,

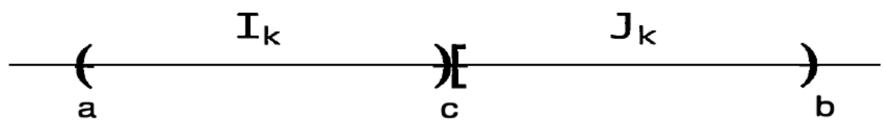
$$I \cap S = [(-\infty, c] \cap I_1] \times I_2 \times \ldots \times I_n,$$

lo cual demuestra nuestra afirmación, pues la intersección de un intervalo lineal con una semirrecta es un intervalo.

(2.23) *Si I y J son dos intervalos disjuntos de $\mathbb{R}^n$, entonces existe un hiperplano H de ecuación $x_k = c$, tal que I está contenido en uno de los semiespacios correspondientes a H y J en el semiespacio complementario.*

Demostración. Sean $I = I_1 \times \ldots \times I_n$ y $J = J_1 \times \ldots \times J_n$. Puesto que $I \cap J = \emptyset$, existe un entero k, $1 \leq k \leq n$, tal que $I_k \cap J_k = \emptyset$.

En el caso más desfavorable, los intervalos I_k y J_k pueden tener un extremo común, como se muestra en la figura. Pero en tal caso, I está contenido en el semiespacio $x_k < c$ y J en el semiespacio $x_k \geq c$, que es precisamente el semiespacio complementario del anterior.



Análogamente se resuelven los otros casos que podrían presentarse y que no vale la pena enumerar.

En general, decimos que un hiperplano H **separa** a los conjuntos A y B si A está contenido en uno de los semiespacios correspondientes a H, en tanto que B está contenido en el semiespacio complementario. La proposición (2.23) se expresa diciendo que si dos intervalos de $\mathbb{R}^n$ son disjuntos, entonces existe un hiperplano H de ecuación $x_k = c$ que los separa.

10. Puntos de acumulación; conjuntos perfectos.

Todos los conjuntos de esta sección son subconjuntos del espacio euclidiano $\mathbb{R}^n$.

Un punto x se llama un **punto de acumulación** del conjunto A si cada entorno V de x contiene al menos un punto de A distinto de x.

Si x es un punto de acumulación de A, entonces cada bola $B(x,r)$ contiene infinitos puntos de A. En efecto, si $A \cap B(x,r)$ fuera finito, denotando por $a_1, a_2, \ldots, a_N$ todos los puntos de A dentro de la bola $B(x,r)$ que son distintos de x, y llamando δ al mínimo de los números $|a_k - x|$ $(k = 1, 2, \ldots, N)$, es claro que dentro de la bola $B(x,\delta)$ no hay ningún punto de A, con la posible excepción del mismo punto x en el caso de que éste sea un elemento de A. Pero esto contradice la afirmación de que x es un punto de acumulación de A. Luego, x contiene infinitos puntos de A.

Un punto de A que no sea un punto de acumulación de A se llama un **punto aislado** de dicho conjunto.

(2.24) *Todo conjunto infinito y acotado posee al menos un punto de acumulación.*

En efecto, si A es infinito y acotado, entonces existe un cubo cerrado Q, tal que $A \subset Q$. Si ningún punto de Q es un punto de acumulación de A, entonces para cada $x \in Q$, existe un entorno V del punto x, tal que $A \cap V$ es finito; y como Q es compacto, bastará una colección finita $\{V_1, V_2, \ldots, V_N\}$ de dichos entornos para cubrir Q; pero entonces, la relación

$$A = A \cap \left(\bigcup_{k=1}^{N} V_k \right) = \bigcup_{k=1}^{N} (A \cap V_k)$$

implica que por ser una unión finita de conjuntos finitos, el conjunto A es finito, en contra de lo supuesto. La contradicción provino de suponer que ningún punto de Q es un punto de acumulación de A.

Corolario. *Toda sucesión acotada contiene una subsucesión convergente.*

Si (a_k) es una sucesión acotada de puntos de $I\!R^n$, consideremos el conjunto

$$A = \{a_1, a_2, a_3, \ldots\}$$

formado por todos los valores de la sucesión (este conjunto no debe confundirse nunca con la sucesión misma). Si A es finito, hay un valor de la sucesión que se repite infinitamente:

$$x = a_{k_1} = a_{k_2} = a_{k_3} = \ldots \qquad (k_1 < k_2 < k_3 < \ldots);$$

y entonces, la subsucesión constante (a_{k_j}) converge al valor x.

Si A es infinito, como es acotado por hipótesis, tiene un punto de acumulación. Supongamos que x es un punto de acumulación de A. Luego, para cada entorno V de x, existen infinitos índices k, tales que $a_k \in V$.

Comencemos por seleccionar un índice k_1, tal que $|a_{k_1} - x| < 1$ y a continuación, un índice $k_2 > k_1$, tal que

$$|a_{k_2} - x| < 1/2;$$

luego un índice $k_3 > k_2$, tal que

$$|a_{k_3} - x| < 1/3;$$

y así siguiendo, es claro que obtendremos una subsucesión (a_{k_j}) que converge al punto x cuando $j \to \infty$.

El conjunto de todos los puntos de acumulación del conjunto A se llama el **conjunto derivado** de A y se denota por A'.

Vamos a probar la fórmula

$$\bar{A} = A \cup A' \tag{2.25}$$

Supongamos que $x \in \bar{A}$. Si $x \notin A$, entonces cada entorno V de x contiene un punto de A que no puede ser el mismo x, ya que éste no es un elemento de A; luego, $x \in A'$.

Hemos demostrado que el miembro izquierdo de (2.25) está contenido en el miembro derecho; y como la inclusión opuesta es obvia, queda demostrada la igualdad.

De la fórmula demostrada se deduce que A es cerrado si y sólo si $A' \subset A$.

Si cada punto de A es un punto de acumulación de A, es decir, si $A \subset A'$, decimos que el conjunto A es **denso en sí mismo**. Un conjunto cerrado denso en sí mismo se llama un conjunto **perfecto**. Por consiguiente, el conjunto A es perfecto si y sólo si $A = A'$.

Es claro que un conjunto finito no puede tener ningún punto de acumulación; en particular, $\emptyset' = \emptyset$. Luego, el conjunto vacío es un conjunto perfecto.

Cualquier intervalo cerrado que no se reduzca a un único punto es otro ejemplo de conjunto perfecto.

Al final de la próxima sección daremos la demostración de la siguiente proposición:

(2.26) *Todo conjunto perfecto no vacío tiene la potencia del continuo.*

11. Conjunto ternario de Cantor.

En esta sección construiremos un conjunto perfecto en la recta real $\mathbb{R}^1$, al cual le asignamos mucha importancia por tratarse de un ejemplo que sirve para dar respuesta a muchas preguntas interesantes.

Dividamos el intervalo cerrado $[0,1]$ en tres intervalos de igual longitud y substraigamos los puntos del intervalo abierto $(1/3, 2/3)$ que representa el tercio central. Nos queda el conjunto cerrado

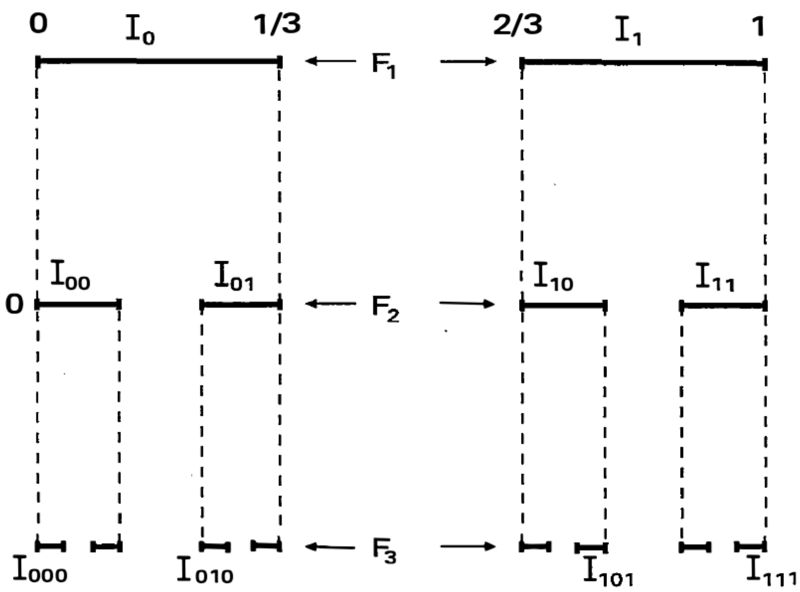
$$F_1 = [0, 1/3] \cup [2/3, 1]$$

que es la unión de dos intervalos cerrados de longitud $1/3$, a los que llamaremos, respectivamente, I_0 e I_1.

Subdividiendo a cada uno de estos intervalos en la misma forma y substrayendo de cada uno de ellos un intervalo abierto que representa su tercio central, nos queda el conjunto cerrado

$$F_2 = [0, 1/9] \cup [2/9, 1/3] \cup [2/3, 7/9] \cup [8/9, 1]$$

que es la unión de cuatro intervalos cerrados de longitud $1/9$, a los que llamaremos, respectivamente, I_{00}, I_{01}, I_{10}, I_{11}.



Continuando indefinidamente con este procedimiento, obtendremos para cada n un conjunto cerrado F_n que es la unión de 2^n intervalos cerrados disjuntos de longitud $1/3^n$:

$$(1) \qquad I_{u_1 u_2 \ldots u_n} \qquad \text{(cada } u_i \text{ igual a cero o uno).}$$

Por conveniencia, los subíndices se colocan de tal manera que resulte $I_{u_1 \ldots u_n u_{n+1}} \subset I_{u_1 \ldots u_n}$. Por ejemplo, I_{010} e I_{011} son subintervalos de I_{01}.

El conjunto ternario de Cantor se define por la fórmula

$$P = \bigcap_{n=1}^{\infty} F_n,$$

la cual muestra claramente que P es cerrado.

Para demostrar que P es perfecto, comencemos observando que los extremos de cada intervalo de la familia (1) son elementos de P.

Si x es un elemento de P, para cada entorno $(x-\varepsilon, x+\varepsilon)$ del punto x, podemos elegir un número n, tal que $1/3^n < \varepsilon$; y como $x \in F_n$, existe un intervalo I de la familia (1), tal que

$$x \in I \subset (x-\varepsilon, x+\varepsilon).$$

Puesto que ambos extremos de I pertenecen a P, el entorno considerado contiene al menos un elemento de P distinto de x. Luego, x es un punto de acumulación de P. Hemos demostrado que $P \subset P'$, lo cual prueba que P es perfecto.

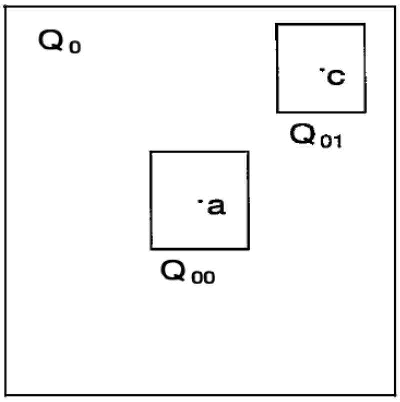
Vamos a probar que P tiene la potencia del continuo: con ese fin, para cada sucesión de dígitos binarios $u = (u_1, u_2, u_3, \ldots)$, sea $x = f(u)$ el único punto que pertenece a cada intervalo de la sucesión decreciente

$$I_{u_1} \supset I_{u_1 u_2} \supset I_{u_1 u_2 u_3} \supset \ldots$$

El punto x, cuya existencia está asegurada por el principio de encaje de intervalos cerrados del mismo Cantor, es un elemento de P, pues para cada n, $x \in I_{u_1 u_2 \ldots u_n} \subset F_n$. Además es bien claro que la función f establece una correspondecia biunívoca entre los elementos del conjunto T formado por todas las sucesiones de dígitos binarios y los elementos de P, de donde resulta que P tiene la potencia del continuo.

La demostración de (2.26) se inspira en lo que acabamos de ver y por esa razón la hemos postergado hasta este momento: si P es un conjunto perfecto no vacío del espacio $\mathbb{R}^n$, eligiendo dos puntos distintos a y b del conjunto P, podemos construir dos cubos cerrados disjuntos Q_0 y Q_1 con centros en dichos puntos y diámetro menor que uno.

Puesto que $a \in P'$, en el interior del cubo Q_0 existe un punto c del conjunto P, tal que $c \neq a$. Por consiguiente, dentro del cubo Q_0 podemos construir (entiéndase elegir) dos cubos cerrados disjuntos Q_{00} y Q_{01} con centros en los puntos a y c y diámetro menor que $1/2$. Análogamente, dentro del cubo Q_1 existen dos cubos cerrados disjuntos Q_{10} y Q_{11} con centros en sendos puntos de P y diámetro menor que $1/2$.



Continuando inductivamente con este proceso, habremos asignado a cada k-upla de dígitos binarios un cubo cerrado $Q_{u_1 u_2 \ldots u_k}$ cuyo centro es un elemento de P y cuyo diámetro es menor que $1/k$; además, los cubos correspondientes a dos k-uplas distintas son disjuntos, y los índices se eligen de tal manera que $Q_{u_1 u_2 \ldots u_k u_{k+1}} \subset Q_{u_1 u_2 \ldots u_k}$.

Ahora, para cada sucesión de dígitos binarios $u = (u_1, u_2, u_3, \ldots)$, llamemos $f(u)$ al único punto de $\mathbb{R}^n$ que pertenece a cada cubo de la sucesión Q_{u_1}, $Q_{u_1 u_2}$, $Q_{u_1 u_2 u_3}$, $\ldots$. Es claro que $f(u)$ es un punto de adherencia del conjunto cerrado P, de donde, $f(u) \in P$. Puesto que la aplicación $f : T \to P$ es inyectiva, se sigue que P tiene la potencia del continuo, como queríamos demostrar.

Utilizando la misma idea de la demostración que acabamos de ver, es inmediato probar que si x es un elemento del conjunto perfecto P, entonces, para cada entorno V de x, el conjunto $P \cap V$ tiene la potencia del continuo.

Volviendo al conjunto de Cantor P, tanto sus elementos como los intervalos de la familia (1) pueden describirse en términos puramente aritméticos, pues es fácil probar por inducción sobre n que el punto x pertenece al intervalo $I_{u_1 \ldots u_n}$ si y sólo si

$$0 \leq x - \sum_{k=1}^{n} (2u_k)/3^k \leq 1/3^n.$$

Por consiguiente, $x \in P$ si y sólo si x admite un desarrollo ternario

$$x = \sum_{k=1}^{\infty} a_k/3^k = 0, a_1 a_2 a_3 \ldots \qquad \text{(base tres)}$$

tal que para cada k, $a_k = 0$ ó $a_k = 2$, es decir, un desarrollo ternario que no use el dígito uno (un desarrollo ternario de este tipo particular es siempre único). Así , por ejemplo, los puntos $1/3$ y $1/4$ que admiten respectivamente los desarrollos

$$0,0222\ldots \ \text{ y } \ 0,020202\ldots \qquad \text{(base tres)},$$

son elementos del conjunto de Cantor.

12. Puntos de condensación.

En esta sección vamos a denotar por Γ la colección numerable formada por todas las bolas $B(a, \rho)$ del espacio $I\!R^n$, tales que $a = (a_1, \ldots, a_n)$ es un punto con coordenadas racionales y ρ un número racional positivo. Para lo que sigue, será esencial que probemos la siguiente proposición.

(2.27) *Para cada bola $B(x, r)$, existe una bola $B(a, \rho) \in \Gamma$, tal que $x \in B(a, \rho) \subset B(x, r)$.*

La idea de la demostración consiste en recordar que cualquier intervalo abierto no vacío de la recta contiene números racionales.

Para cada índice $k = 1, 2, \ldots, n$, comencemos por elegir un número racional a_k, tal que $|a_k - x_k| < r/2\sqrt{n}$. Entonces, el punto de coordenadas racionales $a = (a_1, \ldots, a_n)$ verifica

$$|a - x| = \sqrt{\sum_{k=1}^{n} (a_k - x_k)^2} \ < \ r/2.$$

Si ahora elegimos un número racional ρ, tal que $|a - x| < \rho < r/2$, es claro que $x \in B(a, \rho)$ y vamos a probar que además, $B(a, \rho) \subset B(x, r)$. En efecto, si $y \in B(a, \rho)$, entonces

$$|y - x| \le |y - a| + |a - x| < 2\rho < r,$$

lo que significa que y es un elemento de $B(x,r)$.

Diremos que un punto x es un **punto de condensación** del conjunto E, si para cada entorno V de x, la intersección $E \cap V$ es un conjunto infinito no numerable. Por ejemplo, cada punto de un conjunto perfecto no vacío es un punto de condensación del mismo conjunto.

Denotaremos por E^s el conjunto formado por todos los puntos de condensación de E.

(2.28) *Si $E \cap E^s = \emptyset$, es decir, si ningún punto de E es un punto de condensación del mismo conjunto, entonces E es numerable.*

Por hipótesis, para cada punto x de E, existe una bola $B(x,r)$ tal que $E \cap B(x,r)$ es numerable, y en virtud de (2.27), existe una bola $B \in \Gamma$, tal que $x \in B \subset B(x,r)$, de modo que $E \cap B$ es numerable. Hemos demostrado que la colección $\Gamma(E)$ formada por todas las bolas B de la colección Γ, tales que $E \cap B$ es numerable, es un cubrimiento del conjunto E. Puesto que $\Gamma(E)$ es numerable, podemos escribir $\Gamma(E) = \{B_1, B_2, B_3, \ldots\}$. Ahora bien; la relación

$$E = E \cap \left(\bigcup_k B_k \right) = \bigcup_k (E \cap B_k)$$

nos muestra que el conjunto E es la unión de una familia numerable de conjuntos numerables, de donde se sigue que E es numerable.

Corolario. *Para cualquier conjunto E, el conjunto $A = E - E \cap E^s$ es numerable.*

En efecto, si A no fuera numerable, existiría al menos un punto x perteneciente a la intersección $A \cap A^s$. El punto x, por ser de condensación de A, sería también de condensación de E y por esta razón, no podría ser un elemento de A, lo cual es absurdo.

(2.29) *Para cualquier conjunto E, el conjunto E^s es perfecto.*

Probaremos solamente la inclusión $E^s \subset (E^s)'$, ya que la inclusión opuesta es muy fácil. Si $x \in E^s$, entonces para cualquier entorno V de x, el conjunto $H = E \cap (V - \{x\})$ es no numerable y por consiguiente, existe un punto $y \in H \cap H^s$. Este punto y verifica las relaciones $y \in V$, $y \in E^s$,

$y \neq x$. Hemos demostrado que cada entorno de x contiene un punto de E^s distinto de x, es decir, $x \in (E^s)'$.

Llegamos por fin a la proposición más importante de esta sección, que se conoce con el nombre de "descomposición de Cantor-Bendixon":

(2.30) *Para cualquier conjunto cerrado F, existen un conjunto perfecto P y un conjunto numerable A, tales que*

$$F = P \cup A, \qquad P \cap A = \emptyset.$$

En otras palabras, todo conjunto cerrado es la unión disjunta de un conjunto perfecto y un conjunto numerable. Para demostrarlo, no hay más que escribir $P = F^s$ y $A = F - P$, teniendo en cuenta que $P \subset F$ por ser F un conjunto cerrado.

Corolario. *Todo conjunto cerrado no numerable tiene la potencia del continuo.*

13. Espacios métricos.

Las nociones fundamentales de límite y continuidad, así como las de interior, adherencia, conjunto abierto y conjunto cerrado, entorno, etc., pueden desarrollarse en forma abstracta en los llamados **espacios métricos**. Este proceso de abstracción permite distinguir con claridad los principios en que reposa el Análisis y ampliar notablemente el campo de aplicación de los mismos.

Supongamos que a cada par de elementos x e y de un cierto conjunto E se le ha asignado un número real no negativo $d(x, y)$ que llamaremos la "distancia entre x e y", de modo tal que para cualquier terna x, y, z de elementos de E se verifiquen las siguientes propiedades:

(M1) $\quad d(x, y) = d(y, x)$,

(M2) $\quad d(x, y) \leq d(x, z) + d(z, y)$,

(M3) $\quad d(x, y) = 0$ si y sólo si $x = y$.

En tal caso, la función $d : E \times E \to \mathbb{R}$ se llama una **métrica** o bien una **distancia** sobre el conjunto E y el par (E, d) se llama un **espacio**

métrico. Los elementos de un espacio métrico E se llaman los **puntos** de dicho espacio.

La propiedad M2 se conoce bajo el nombre de "desigualdad triangular", por expresar en forma abstracta el hecho geométrico de que cada lado de un triángulo es menor o igual que la suma de los otros lados.

Ejemplos.

Entre los numerosos ejemplos existentes elegiremos unos pocos que permitan intuir el significado y el alcance de la noción introducida.

1) El conjunto $\mathbb{R}^n$ con la distancia $d(x,y) = |x-y|$ es un espacio métrico al cual nos referimos como el **espacio euclidiano** $\mathbb{R}^n$.
2) El mismo conjunto $\mathbb{R}^n$ con la distancia $d(x,y) = \max\limits_{1\leq i\leq n} |x_i - y_i|$, donde x_i e y_i representan las i-ésimas coordenadas de los puntos x e y, es un espacio métrico distinto del anterior.
3) El conjunto $C[a,b]$ formado por todas las funciones continuas $f(t)$ sobre el intervalo cerrado $a \leq t \leq b$, con la distancia
$$d(f,g) = \max_{a\leq t\leq b} |f(t) - g(t)|$$
es un espacio métrico muy útil para el estudio de ciertas propiedades relacionadas con la convergencia uniforme de funciones continuas.
4) El mismo conjunto $C[a,b]$ del ejemplo anterior con la distancia d' definida por la fórmula
$$d'(f,g) = \int_a^b |f(t) - g(t)| dt$$
es un espacio métrico distinto del que hemos considerado en aquel ejemplo.
5) Sea M el conjunto formado por todas las sucesiones acotadas de números reales y si $x = (x_k)$ e $y = (y_k)$ son dos elementos de M, definamos la distancia entre x e y por medio de la fórmula $d(x,y) = \sup\limits_k |x_k - y_k|$.
6) Siendo E un conjunto cualquiera, definamos la distancia entre elementos de E poniendo $d(x,y) = 1$ si $x \neq y$, $d(x,y) = 0$ si $x = y$.

En todos los ejemplos propuestos el lector puede probar sin dificultad que la función $d(x,y)$ que se define satisface los tres axiomas de una distancia; es decir, las propiedades M1, M2, M3.

Diremos que una sucesión (x_k) de puntos de E **converge** a un punto x de dicho espacio si la distancia $d(x_k, x)$ tiende a cero cuando k tiende a infinito. En tal caso también se dice que x es el **límite** de la sucesión y se escribe $x = \lim x_k$ o bien $x_k \to x$.

Una misma sucesión (x_k) no puede converger a dos límites distintos x e y, pues si $x_k \to x$ y $x_k \to y$, la desigualdad triangular nos da $d(x, y) \leq d(x, x_k) + d(x_k, y)$, y haciendo que k tienda a infinito, resulta $d(x, y) = 0$, de donde $x = y$.

La **bola abierta** $B(x, r)$ con centro en un punto x de E y radio $r > 0$ se define como el conjunto formado por todos los puntos y de E que verifican la relación $d(y, x) < r$. Las nociones de conjunto abierto, conjunto cerrado, entorno de un punto, interior y adherencia de un conjunto, se desarrollan en el espacio métrico E del mismo modo que lo hicimos en el espacio euclidiano $\mathbb{R}^n$.

Si (E, d) y (E', d') son dos espacios métricos, una función f de E en E' se llama **continua** en un punto x_0 de E si para cada $\varepsilon > 0$ existe un $\delta > 0$, tal que la relación $d(x, x_0) < \delta$ implica $d'(f(x), f(x_0)) < \varepsilon$.

El espacio métrico E se llama **separable** si existe un conjunto numerable $A \subset E$ cuya adherencia es E. La teoría desarrollada en la sección 12 de este capítulo hasta la proposición (2.30) inclusive (descomposición de Cantor-Bendixon) es válida no sólo en $\mathbb{R}^n$ sino más generalmente, en cualquier espacio métrico separable.

Una sucesión (x_k) de puntos de un espacio métrico E se llama una **sucesión fundamental** o **de Cauchy** si $d(x_k, x_j)$ tiende a cero cuando k y j tienden a infinito. Toda sucesión convergente es una sucesión fundamental.

El espacio métrico E se llama un espacio **completo** si toda sucesión fundamental es convergente.

Las nociones de separabilidad y completitud antes definidas juegan un papel muy destacado en la teoría de los espacios métricos. El lector interesado en estudiar la teoría de los espacios métricos puede consultar la obra de J. Dieudonné, Foundations of Modern Analysis, Academic Press, New York (1960), o bien la obra de I. Kaplansky, Set theory and metric spaces, Allyn an Bacon, Boston (1972), más especialmente dedicada al desarrollo de dicho tema.

* 14. Espacios normados, normas de Orlicz sobre $\mathbb{R}^n$.

Sea X un espacio vectorial sobre el cuerpo de los números reales o bien sobre el cuerpo de los números complejos. Diremos que X es un **espacio normado** si a cada vector x de dicho espacio se le ha asignado un número no negativo $\|x\|$, llamado la **norma** de x, de modo tal que se satisfagan las propiedades siguientes:

(N1) $\|x+y\| \leq \|x\| + \|y\|,$

(N2) $\|\lambda x\| = |\lambda|\, \|x\|,$

(N3) $x = 0$ si y sólo si $\|x\| = 0,$

para cualquier par de vectores x, y del espacio X y cualquier elemento λ del cuerpo de escalares ($\mathbb{R}$ o $\mathbb{C}$).

Una función $\|\cdot\| : X \to \mathbb{R}$ que verifique dichas propiedades se llama una **norma** sobre el espacio X.

El ejemplo más conocido es el de la **norma euclididana** o "norma dos", definida para cada vector $x = (x_1, x_2, \ldots, x_n)$ del espacio $\mathbb{R}^n$ por medio de la fórmula

$$\|x\|_2 = |x| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}.$$

A veces es necesario considerar distintas normas definidas sobre un mismo espacio vectorial X. Por ejemplo, las fórmulas

$$\|x\|_1 = |x_1| + |x_2| + \ldots + |x_n|,$$

$$\|x\|_\infty = \max(|x_1|, |x_2|, \ldots |x_n|),$$

definen otras dos normas distintas sobre el mismo espacio $\mathbb{R}^n$ a las cuales llamaremos la "norma uno" y la "norma infinito", respectivamente. El lector puede verificar fácilmente que estas funciones verifican las propiedades N1, N2 y N3 del comienzo y son, por lo tanto, normas sobre el espacio $\mathbb{R}^n$.

Notemos también que la norma euclidiana o norma dos no es otra cosa que el módulo o longitud del vector x.

Si $\|\cdot\|$ es una norma sobre el espacio X, la función

$$d(x,y) = \|x - y\|$$

es una distancia entre elementos de X, como puede demostrarse muy fácilmente, de modo que el par (X, d) es un espacio métrico. Por consiguiente, todo espacio normado es también un espacio métrico. En particular, podemos decir que una sucesión de puntos (x_k) del espacio X **converge** al punto x de mismo espacio, con respecto a la norma $\|\cdot\|$, si la sucesión $\|x - x_k\|$ tiende a cero cuando k tiende a infinito.

Se comprende entonces que en un espacio normado también se definen las nociones de conjunto abierto y conjunto cerrado, como en cualquier espacio métrico.

La bola con centro x y radio r con respecto a la norma $\|\cdot\|$, es decir, el conjunto

$$\{y : y \in X,\ \|y - x\| < r\}$$

se denota a veces en la forma $B_{\|\ \|}(x, r)$, especialmente cuando se han definido varias normas sobre el mismo espacio y es necesario señalar a cuál de ellas se refiere el concepto.

Es ilustrativo representar gráficamente las bolas con centro en el origen $0 = (0,0)$ y radio $r = 1$ del espacio $I\!R^2$ con respecto a las normas

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2}, \qquad \|x\|_1 = |x_1| + |x_2|, \qquad \|x\|_\infty = \max(|x_1|, |x_2|).$$

A continuación veremos que cualquier norma sobre el espacio euclidiano $I\!R^n$ genera los mismos conjuntos abiertos que hemos definido en la sección 2, lo cual se expresa diciendo que todas las normas sobre $I\!R^n$ son equivalentes. Más precisamente, demostraremos que si $\|\cdot\|$ es una norma sobre $I\!R^n$, entonces existen constantes positivas k_1 y k_2, tales que para cualquier $x \in I\!R^n$, se verifica

$$k_1|x| \leq \|x\| \leq k_2|x|, \tag{2.31}$$

donde las constantes k_1 y k_2 dependen de la norma $\|\cdot\|$ dada.

Veamos primero la segunda desigualdad en (2.31). Pongamos e_i para el vector en $I\!R^n$ cuya i-ésima componente es uno y las restantes cero. Un repetido uso de las propiedades (N1) y (N2) nos da

$$\|x\| \leq \sum_{i=1}^{n} |x_i|\ \|e_i\|.$$

Si k_2 denota la constante $\left(\|e_1\|^2 + \ldots + \|e_n\|^2\right)^{1/2}$, usando la desigualdad de Cauchy-Schwarz, obtenemos

$$\|x\| \leq k_2|x|. \tag{2.32}$$

Dado que (N3) implica $|\,\|x\| - \|y\|\,| \leq \|x-y\|$ tenemos, por (2.32),

$$| \, \|x\| - \|y\| \, | \leq k_2 |x - y|. \tag{2.33}$$

Observe que (2.33) nos dice que la función $x \to \|x\|$ es uniformemente continua sobre $I\!R^n$. Así, por el teorema de Weierstrass, el ínfimo k_1 de $\|x\|$ para $x \in S = \{x \in I\!R^n : |x| = 1\}$ es alcanzado sobre ese conjunto, pero como $\|\cdot\|$ es estrictamente positiva sobre S, tenemos que $k_1 > 0$. En particular si $x \neq 0$ el vector $x/|x| \in S$ y por consiguiente

$$\left\| \frac{x}{|x|} \right\| \geq k_1$$

o bien

$$k_1 |x| \leq \|x\|. \tag{2.34}$$

Ahora (2.32) y (2.34) dan (2.31).

Las desigualdades (2.31) nos dicen que la noción de convergencia de una sucesión (x_k) en $I\!R^n$ hacia un vector x es independiente de la norma usada para definir la convergencia. Además si $B_{|,|}(x,r)$ denota la bola centrada en x y radio r definida con la distancia $|x-y|$, las desigualdades (2.31) nos aseguran

$$\begin{cases} B_{\|\,\|}(x,r) \subseteq B_{|\,|}(x, r/k_1) \\ B_{|\,|}(x,r) \subseteq B_{\|\,\|}(x, k_2 r) \end{cases} \tag{2.35}$$

Así hemos demostrado que la métrica generada por una norma da origen exactamente a los mismos conjuntos abiertos que la métrica euclideana $d(x,y) = |x-y|$.

El concepto de norma se puede definir sobre un espacio vectorial E cualquiera cuyo campo de escalares son los números reales o los complejos. Por lo tanto diremos que $\|\cdot\|$ es una **norma sobre** E si se cumplen las condiciones (N1) a (N3), donde naturalmente ahora x, y son vectores en E y λ un número real o complejo según E sea un espacio vectorial real o complejo. En esta situación decimos que E es un **espacio normado**. En los ejemplos 3) y 4) de la sección anterior se han definido dos distancias para el espacio vectorial real $C[a,b]$ y claramente aquellas distancias provienen de normas.

Dejamos como ejercicio ver que dichas normas no son equivalentes. Esto nos previene de que en espacios normados de dimensión infinita la convergencia depende fuertemente de la norma usada para definirla. Importantes ejemplos de espacios normados de dimensión infinita serán analizados en detalle en el Capítulo VIII.

Por el momento insistiremos sobre normas definidas en $\mathbb{R}^n$. Dada una norma $\|\cdot\|$ pongamos

$$C = \{x \in \mathbb{R}^n : \|x\| \leq 1\}.$$

Señalamos las siguientes propiedades del conjunto C, todas fácilmente verificables:

(1) El conjunto C es convexo y $0 \in C$.

(2) El conjunto C **absorbe** cualquier punto de $\mathbb{R}^n$, i.e. dado $x \in \mathbb{R}^n$ existe $\varepsilon > 0$ tal que $\varepsilon x \in C$.

(3) El conjunto C es **simétrico** respecto al origen, i.e. $x \in C$ si y sólo si $-x \in C$.

(4) No existe ninguna recta contenida en C.

Hacemos notar que las propiedades (1) a (4) son puramente algebraicas.

Sea $C \subseteq \mathbb{R}^n$ un conjunto no vacío que cumpla (1) a (4) y definamos para $x \in \mathbb{R}^n$

$$p_C(x) = p(x) = \inf \left\{ \lambda > 0 : \frac{x}{\lambda} \in C \right\}.$$

La función $p(x)$ es conocida con el nombre de **funcional de Minkowski** del convexo C. De la propiedad (2) se sigue que $p(x)$ es una función con valores reales no negativos.

(2.36) **Teorema.** *Dado un conjunto $C \subseteq \mathbb{R}^n$ con las propiedades (1) a (4), su funcional de Minkowski $p_C(x)$ es una norma sobre $\mathbb{R}^n$.*

DEMOSTRACIÓN. Si $x = 0$ como $0 \in C$ entonces $p(0) = 0$. Sea $p(x) = 0$; luego existe una sucesión $\lambda_k \searrow 0$ tal que $x/\lambda_k \in C$ y si x fuera no nulo el conjunto C contendría la recta $\{tx : t \in \mathbb{R}\}$ lo que contradice (4). Así $x = 0$ si y sólo si $p(x) = 0$.

Veamos que se cumple la desigualdad triangular $p(x+y) \leq p(x) + p(y)$. Sean (λ_k) y (μ_k) sucesiones tales que $\lambda_k \searrow p(x)$ y $\mu_k \searrow p(y)$ con $x/\lambda_k \in C$

e $y/\mu_k \in C$. Entonces por ser C convexo tenemos

$$\frac{x+y}{\lambda_k+\mu_k} = \frac{\lambda_k}{\lambda_k+\mu_k}\,\frac{x}{\lambda_k} + \frac{\mu_k}{\lambda_k+\mu_k}\,\frac{y}{\mu_k} \in C.$$

Así $p(x+y) \leq \lambda_k + \mu_k$ y haciendo tender $k \to \infty$ obtenemos la desigualdad.

Demostraremos ahora que $p(cx) = |c|p(x)$ cuando $c \neq 0$ (el caso $c = 0$ es trivial). Sea $\lambda_k \searrow p(x)$ con $x/\lambda_k \in C$. Por (3)

$$\frac{c}{|c|}\,\frac{x}{\lambda_k} = \frac{cx}{|c|\lambda_k} \in C,$$

de donde $p(cx) \leq |c|\lambda_k$, y haciendo $k \to \infty$, $p(cx) \leq |c|p(x)$.

Ahora bien, por la desigualdad que acabamos de probar, también tendremos

$$p(x) = p\left(\frac{1}{c}\ cx\right) \leq \frac{1}{|c|}p(cx);$$

es decir $p(cx) \geq |c|p(x)$.

Dejamos como ejercicio demostrar las siguientes afirmaciones:

(2.37) $$\{x \in \mathbb{R}^n : p_C(x) < 1\} \subset C \subset \{x \in \mathbb{R}^n : p_C(x) \leq 1\}.$$

(2.38) *Si C es además un conjunto cerrado en $\mathbb{R}^n$*

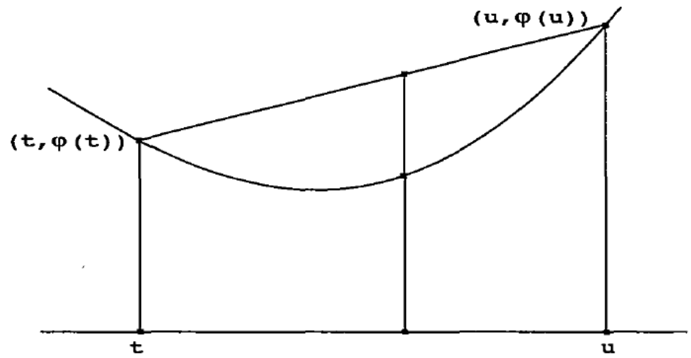
$$C = \{x \in \mathbb{R}^n : p_C(x) \leq 1\}.$$

Gracias al teorema (2.36) se nos amplían considerablemente los ejemplos de normas en $\mathbb{R}^n$, pues basta dar un convexo C con las propiedades adecuadas. A continuación consideraremos algunos ejemplos que más adelante tendrán sus análogos en espacios de dimensión infinita.

Una función numérica $\varphi : I \to \mathbb{R}$ definida sobre un intervalo I de la recta se llama **convexa** si para cualquier par de puntos t, u de dicho intervalo y cualquier número λ que verifique $0 \leq \lambda \leq 1$, se cumple

$$\varphi(\lambda t + (1-\lambda)u) \leq \lambda\varphi(t) + (1-\lambda)\varphi(u).$$

Gráficamente esto significa que el arco de la gráfica de φ comprendido entre los puntos $(t, \varphi(t))$ y $(u, \varphi(u))$ está por debajo de la cuerda determinada por dichos puntos (ver figura)



En lo que resta de esta sección, los símbolos φ, ψ con o sin subíndices denotarán funciones convexas no negativas en el intervalo $[0,1]$, iguales a cero en el punto $t = 0$ e iguales a uno en el punto $t = 1$. En particular, de la relación $t = (1-t) \cdot 0 + t \cdot 1$ resulta $0 \leq \varphi(t) \leq t$, lo cual muestra que φ es continua en el origen.

Más adelante se demostrará que la convexidad de φ implica necesariamente su continuidad en $(0,1)$, pero este hecho no será utilizado por el momento. Sin embargo, φ puede no ser continua en el extremo derecho del intervalo, como lo muestra el caso de la función convexa φ_∞ que definimos así:

$$\varphi_\infty(t) = \begin{cases} 0 & \text{si } 0 \leq t < 1 \\ 1 & \text{si } t = 1. \end{cases}$$

Además, para cada $p \geq 1$ definimos la función convexa φ_p por la fórmula

$$\varphi_p(t) = t^p \qquad\qquad (0 \leq t \leq 1).$$

Entonces es claro que en cualquier punto t se cumple

$$\varphi_\infty(t) = \lim_{p \to \infty} \varphi_p(t).$$

Sea Q el cubo cerrado de $I\!R^n$ con centro en el origen y radio 1, es decir, el conjunto de los puntos x que verifican $\|x\|_\infty \leq 1$. Es inmediato verificar que para cada φ el conjunto

$$C_\varphi = \{x \in Q : \sum_{i=1}^{n} \varphi(|x_i|) \leq 1\}$$

cumple las propiedades (1) a (4). La norma asociada a este convexo que será denotada por $\|\cdot\|_\varphi$ se llama la **norma de Orlicz** correspondiente a φ.

Dejamos como ejercicio la demostración de las desigualdades siguientes:

$$\|x\|_{\varphi_p} = \|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}, \tag{2.39}$$

$$\|x\|_{\varphi_\infty} = \|x\|_\infty$$

Notemos que para $1 \leq p < \infty$ se cumple $C_{\varphi_p} = \{x \in I\!\!R^n : \|x\|_p \leq 1\}$. Sin embargo, esta relación no se mantiene válida para $p = \infty$; por ejemplo, para $n = 2$, C_{φ_∞} es el cuadrado Q menos sus cuatro vértices.

Para finalizar ilustraremos las desigualdades (2.31) con las normas $\|\cdot\|_\varphi$. Notemos que si $\varphi \leq \psi$, es decir, si $\varphi(t) \leq \psi(t)$ para cada t, entonces $C_\psi \subset C_\varphi$, y en consecuencia $\|x\|_\varphi \leq \|x\|_\psi$. En particular, puesto que para cualquier función convexa φ se cumple $\varphi_\infty \leq \varphi \leq \varphi_1$ podemos enunciar la siguiente proposición.

(2.41) *Si* $\varphi \leq \psi$ *entonces para cada* $x \in I\!\!R^n$ *tenemos* $\|x\|_\infty \leq \|x\|_\varphi \leq \|x\|_\psi \leq \|x\|_1$.

Un caso particular de (2.41) merece un enunciado especial:

(2.42) *Si* $1 \leq p < q \leq \infty$ *entonces para cada* x *de* $I\!\!R^n$ *tenemos* $\|x\|_\infty \leq \|x\|_q \leq \|x\|_p \leq \|x\|_1$.

Las desigualdades en (2.42) así como en (2.41) son óptimas, lo que significa que para algunos x se alcanza la igualdad. Por ejemplo $\|(1, 0, \ldots, 0)\|_p = 1$ para cualquier p. Es fácil obtener desigualdades en sentido inverso al de las (2.42). Por ejemplo $\|x\|_1 \leq n\|x\|_\infty$ y esta desigualdad es también óptima, como se ve eligiendo $x = (1, 1, \ldots, 1)$. Así (2.42) junto con la última observación implica que si $q < p$ entonces $\|x\|_p \leq n\|x\|_q$. Por consiguiente, para este tipo de normas hemos encontrado explícitamente las constantes k_1 y k_2 de (2.31) a saber:

(2.43) *Si* $p < q$ *,se tiene*

$$\frac{1}{n}\|x\|_p \leq \|x\|_q \leq \|x\|_p.$$

La constante n de la última proposición no es óptima; más adelante veremos que como consecuencia de teoremas más generales, se cumple $\|x\|_p \leq n^{\frac{1}{p}-\frac{1}{q}}\|x\|_q$ y que aquí sí, a veces, se alcanza la igualdad.

EJERCICIOS

1. Probar que en la desigualdad de Minkowski (teorema 2.2) la igualdad se verifica si y sólo si uno de los dos vectores es igual al producto del otro por un número no negativo.
2. Probar que $|x-y| = |x-z|+|z-y|$ si y sólo si z pertenece al segmento que une x con y.
3. Probar que la bola cerrada $K(a,r) = \{x : |x-a| \leq r\}$ es la adherencia de la bola abierta $B(a,r)$.
4. Probar que si A es un subconjunto convexo de $I\!R^n$, entonces cada vector x de la forma

   $$x = t_1a_1 + t_2a_2 + \ldots + t_ka_k,$$

   donde k es cualquier entero positivo, cada a_j es un punto de A y los t_j son números no negativos que verifican $t_1 + t_2 + \ldots + t_k = 1$, es un punto de A. Cada vector x de dicha forma se llama una **combinación convexa** de puntos de A. Sugerencia: inducción sobre k.
5. Probar que si A es un subconjunto de $I\!R^n$, la cápsula convexa de A está formada por todas las combinaciones convexas de puntos de A.
6. Siendo $A \subset I\!R^n$, probar que $f : A \to I\!R^m$ es continua si y sólo si para cada conjunto abierto $U \subset I\!R^m$, existe un conjunto abierto G de $I\!R^n$, tal que $f^{-1}(U) = G \cap A$. Sugerencia: llamar G a la unión de todos los abiertos V de $I\!R^n$, tales que $f(V \cap A) \subset U$.
7. Probar que $f : I\!R^n \to R$ es s.s. si y sólo si $-f$ es s.i.

8. Probar las relaciones

$$\limsup_{x\to x_0}\{f(x)+g(x)\} \le \limsup_{x\to x_0} f(x) + \limsup_{x\to x_0} g(x),$$

$$\liminf_{x\to x_0}\{f(x)+g(x)\} \ge \liminf_{x\to x_0} f(x) + \liminf_{x\to x_0} g(x),$$

9. Si f y g son ambas s.s., también lo son $f+g$ y λf, siempre que λ sea un número no negativo.

10. Dada una función $f : I\!R^n \to I\!R$, para cada x de $I\!R^n$ definimos

$$M_\delta(x) = \sup\{f(y) : |y-x| < \delta\},$$

$$m_\delta(x) = \inf\{f(y) : |y-x| < \delta\},$$

$$M(x) = \inf_{\delta>0} M_\delta(x), \qquad m(x) = \sup_{\delta>0} m_\delta(x).$$

Probar que $M(x)$ es s.s., en tanto que $m(x)$ es s.i.; que en cada punto x se verifica $m(x) \le f(x) \le M(x)$ y finalmente, que f es continua en x si y sólo si estos tres valores coinciden.

11. Probar que para cualquier función $f : I\!R^n \to R$, el conjunto $D(f)$ formado por todos los puntos x donde f es discontinua (conjunto de discontinuidad de f) es una unión numerable de conjuntos cerrados. Sugerencia: observar que

$$D(f) = \bigcup_{k=1}^{\infty}\{x : M(x) - m(x) \ge 1/k\}.$$

12. Sean $A \subset I\!R^n$, $B \subset I\!R^m$ y consideremos dos aplicaciones

$$f : A \to B \quad \text{y} \quad g : B \to I\!R^k.$$

Probar que si f es continua en el punto x_0 de A y g es continua en el punto $f(x_0)$, entonces la función compuesta $h = g \circ f$ es continua en x_0. Concluir que la composición de dos funciones continuas es siempre una función continua.

13. El diámetro de un cubo $Q(a, \rho/2)$ cuyos lados tienen longitud ρ, es igual a $\rho\sqrt{n}$ (proporcional a ρ).

14. Probar que cada intervalo de $\mathbb{R}^n$ es un conjunto convexo.

15. El diámetro de cualquier conjunto es igual al diámetro de su adherencia.

16. Probar que $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ es semicontinua superiormente en el punto x_0 si y sólo si para cada sucesión (a_k), tal que $a_k \to x_0$ cuando $k \to \infty$, se verifica $\limsup f(a_k) \leq f(x_0)$.

17. Probar que si $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ es s.s., entonces f alcanza un valor máximo sobre cada conjunto compacto $A \subset \mathbb{R}^n$. Sugerencia: si M es el supremo de todos los valores de f sobre A, considerar una sucesión (a_k) de elementos de A, tal que $f(a_k)$ tiende a M cuando k tiende a infinito.

18. Para $f \in C[0,1]$ definimos

$$\|f\|_\infty = \max_{0 \leq t \leq 1} |f(t)|, \qquad \|f\|_1 = \int_0^1 |f(t)| dt.$$

    Las expresiones anteriores nos dan dos normas en $C[0,1]$ que no son equivalentes.

19. Una norma es **estrictamente convexa** si cada vez que $x \neq y$ y $1 = \|x\| = \|y\|$ se verifica $\|x+y\| < 2$. Las normas $\| \|_1$ y $\| \|_\infty$ no son estrictamente convexas sobre $\mathbb{R}^n$ pero sí lo es $\|x\|_2 = |x|$.

20. Demuestre las inclusiones en (2.37) y la igualdad en (2.38).

21. Demuestre (2.39) y (2.40). Dibuje en $\mathbb{R}^2$ el conjunto $B_p = \{x : \|x\|_p \leq 1\}$ para $p = 1, \frac{3}{2}, 2, 3,$ y ∞.

22. Una norma $\|\cdot\|$ sobre $\mathbb{R}^n$ es **monótona** si x, $y \in \mathbb{R}^n$ con $|x_i| \leq |y_i|$, $i = 1, \ldots, n$, entonces $\|x\| \leq \|y\|$. Las normas de Orlicz $\|\cdot\|_\varphi$ son monótonas. Dé ejemplos de normas que no son monótonas.

23. El Teorema (2.36) vale, con la misma demostración, si $\mathbb{R}^n$ es reemplazado por un espacio vectorial E sobre el campo de los reales. En las propiedades (1) a (4) cada vez que aparece $\mathbb{R}^n$ debe ser reemplazado por E.

24. Sean E un espacio vectorial real con una norma $\|\cdot\|$ y $e_1, \ldots, e_n$, n vectores linealmente independientes en E y sea F el subespacio generado por $e_1, e_2, \ldots, e_n$. Entonces F es un conjunto cerrado en E. (Sugerencia: en $\mathbb{R}^n$ podemos definir una norma $\| \|_*$ por medio de

$\|x\|_* = \|(x_1, \ldots, x_n)\|_* = \|(x_1 e_1 + \ldots + x_n e_n)\|$. La función $Lx = x_1 e_1 + \ldots + x_n e_n$ es continua y tiene inversa continua, use (2.31)).

CAPITULO III

MEDIDA DE LEBESGUE

1. Introducción.

La definición de integral que se estudia en los primeros cursos de Análisis Matemático, debida a los matemáticos Cauchy y Riemann, corresponde al procedimiento usado en la Geometría elemental para definir el área de una figura plana: el área de un rectángulo se define como el producto de las longitudes de sus lados, y el de una **figura elemental**, es decir, una unión finita de rectángulos disjuntos, como la suma de las áreas de los rectángulos que la componen.

En el caso de una figura plana acotada F de forma arbitraria, definimos el **área interior** de F como el supremo de las áreas de todas las figuras elementales contenidas en F y el **área exterior** de F como el ínfimo de las áreas de todas las figuras elementales que contienen a F. Si los dos números así obtenidos coinciden, decimos que F es medible y que tiene un área igual al valor común de dichos números.

La teoría clásica de la medida se desarrolla sobre la base de esta definción. Denotando por $m(F)$ la medida o área de F, el teorema fundamental de dicha teoría afirma que si $F_1, F_2, \dots, F_k$ son figuras medibles disjuntas, entonces $F_1 \cup F_2 \cup \dots \cup F_k$ es medible y además,

$$m(F_1 \cup F_2 \cup \dots \cup F_k) = m(F_1) + m(F_2) + \dots + m(F_k).$$

Sin embargo, este teorema deja de ser cierto al considerar uniones numerables y pueden darse ejemplos muy sencillos de conjuntos que carecen de área, es decir, que no son medibles en el sentido clásico. Por ejemplo, el conjunto formado por todos los puntos (x, y) del cuadrado unitario $0 \leq x \leq 1$, $0 \leq y \leq 1$ cuyas coordenadas son ambas racionales, tiene área exterior igual a uno y área interior igual a cero.

A principios del presente siglo, los matemáticos franceses E. Borel y H. Lebesgue consiguieron superar las limitaciones de la vieja teoría de la medida reemplazando los conjuntos elementales usados para definirla por conjuntos σ-elementales, que son los que pueden representarse como una unión numerable de rectángulos disjuntos.

Esta fructífera idea, en unión con la teoría de los conjuntos que había comenzado a desarrollarse a fines del siglo anterior, conduce naturalmente a una nueva y más general noción de integral: la integral de Lebesgue, cuyas propiedades demostraron adaptarse a las necesidades del Análisis Moderno, particularmente bien en los procesos de paso al límite dentro de la integral, que sólo pueden justificarse bajo hipótesis demasiado restrictivas cuando se trabaja con la definición clásica de Cauchy y de Riemann.

El presente capítulo se dedica a exponer desde sus fundamentos la teoría de la medida de Lebesgue.

2. Medida de intervalos.

Recordemos que dados dos números reales a y b, tales que $a \leq b$, la **medida** o **longitud** de cualquier intervalo lineal J con extremo izquierdo a y con extremo derecho b se define como el número no negativo

$$m(J) = b - a.$$

De modo que los cuatro intervalos (a, b), $[a, b]$, $(a, b]$ y $[a, b)$ tienen la misma longitud $b - a$; en particular, $m(\emptyset) = 0$.

Es inmediato verificar que si J_1 y J_2 son dos intervalos lineales disjuntos cuya unión es el intervalo J, entonces $m(J) = m(J_1) + m(J_2)$.

El siguiente paso consiste en extender la noción de medida a cualquier intervalo I del espacio $\mathbb{R}^n$.

Si $J_1, J_2, \ldots, J_n$ son intervalos lineales tales que $I = J_1 \times J_2 \times \ldots \times J_n$, la **medida** o **volumen** de I se define por medio de la fórmula

$$m(I) = m(J_1) \cdot m(J_2) \cdot \ldots \cdot m(J_n).$$

Es decir, la medida de un intervalo de $I\!R^n$ es el producto de las longitudes de sus lados.

3.1 **Teorema.** *Si el intervalo* I *es la unión de los intervalos disjuntos* $I_1, I_2, \ldots, I_N$ *, entonces* $m(I) = m(I_1) + m(I_2) + \ldots + m(I_N)$ *.*

La demostración es muy fácil si $N = 2$, pues la única manera de descomponer el intervalo $I = J_1 \times J_2 \times \ldots \times J_n$ como unión de dos intervalos disjuntos I_1 e I_2 consiste en dividir uno de sus lados J_k en dos intervalos disjuntos J_k' y J_k''. Si por simplicidad en la notación suponemos $k = 1$, entonces

$$I_1 = J_1' \times J_2 \times \ldots \times J_n \quad \text{y} \quad I_2 = J_1'' \times J_2 \times \ldots \times J_n,$$

de donde

$$\begin{aligned} m(I) &= m(J_1)\, m(J_2)\, \ldots\, m(J_n) \\ &= [\, m(J_1') + m(J_1'')\,] \cdot\ m(J_2)\, \ldots\, m(J_n) \\ &= \ m(I_1) + m(I_2). \end{aligned}$$

La demostración en general se realiza por inducción sobre N, del modo siguiente: puesto que I_1 e I_2 son disjuntos, existe un hiperplano de ecuación $x_k = c$ que deja al intervalo I_1 completamente contenido en un semiespacio S_1 de los que él determina, y al intervalo I_2 en el semiespacio complementario $S_2 = I\!R^n - S_1$.

Si para cada intervalo J del espacio $I\!R^n$, ponemos

$$J' = J \cap S_1 \quad \text{y} \quad J'' = J \cap S_2,$$

tendremos $m(J) = m(J') + m(J'')$; además,

$$I_1' = I_1, \quad I_1'' = \emptyset, \quad I_2' = \emptyset, \quad I_2'' = I_2,$$

de donde, recordando que $I = I_1 \cup I_2 \cup \ldots \cup I_N$, resulta

$$I' = I_1 \cup I_3' \cup \ldots \cup I_N', \quad I'' = I_2 \cup I_3'' \cup \ldots \cup I_N''.$$

Suponiendo que (3.1) ha sido demostrado para descomposiciones de un intervalo cualquiera en $N - 1$ intervalos disjuntos, tendremos

$$\begin{aligned} m(I) &= m(I') + m(I'') \\ &= m(I_1) + m(I_3') + \ldots + m(I_N') + m(I_2) + m(I_3'') + \ldots + m(I_N'') \\ &= m(I_1) + m(I_2) + \ldots + m(I_N). \end{aligned}$$

con lo cual queda demostrada la propiedad aditiva de la medida de intervalos.

3. Medida de conjuntos elementales.

Si el conjunto elemental A del espacio $\mathbb{R}^n$ es la unión de los intervalos disjuntos $I_1, I_2, \ldots, I_N$ de dicho espacio, llamaremos **medida** de A al número

$$m(A) = m(I_1) + m(I_2) + \ldots + m(I_N).$$

La definición es correcta, pues si $J_1, J_2, \ldots, J_M$ es otra sucesión finita de intervalos disjuntos cuya unión es A, entonces cada intervalo I_k de la descomposición anterior es la unión de los intervalos disjuntos $I_k \cap J_i$ $(i = 1, 2, \ldots, M)$ y análogamente, cada intervalo J_i es la unión de los intervalos disjuntos $I_k \cap J_i$ $(k = 1, 2, \ldots, N)$, y en virtud de (3.1),

$$\begin{aligned} \sum_{k=1}^{N} m(I_k) &= \sum_{k=1}^{N} \sum_{i=1}^{M} m(I_k \cap J_i) = \sum_{i=1}^{M} \sum_{k=1}^{N} m(I_k \cap J_i) \\ &= \sum_{i=1}^{M} m(J_i). \end{aligned}$$

Es decir, el número $m(A)$ no depende de la representación particular de A usada para calcularlo.

Como consecuencia inmediata de la definición, tenemos:

(3.2) *Si el conjunto elemental A es la unión de los conjuntos elementales disjuntos $A_1, A_2, \ldots, A_N$, entonces*

$$m(A) = m(A_1) + m(A_2) + \ldots + m(A_N).$$

La siguiente propiedad es una consecuencia también inmediata de la que acabamos de enunciar.

(3.3) *Si el conjunto elemental A está contenido en el conjunto elemental B, entonces $m(B - A) = m(B) - m(A)$; en particular, $m(A) \leq m(B)$.*

En efecto, puesto que $B = A \cup (B - A)$, en virtud de (3.2) tendremos

$$m(B) = m(A) + m(B - A),$$

lo cual demuestra la proposición.

(3.4) *Si* A *y* B *son conjuntos elementales, entonces*

$$m(A \cup B) + m(A \cap B) = m(A) + m(B);$$

en particular, $m(A \cup B) \leq m(A) + m(B)$.

La demostración se obtiene sumando miembro a miembro las igualdades

$$m(A) = m(A - B) + m(A \cap B) \quad \text{y} \quad m(B) = m(B - A) + m(A \cap B)$$

y teniendo en cuenta que $m(A - B) + m(B - A) + m(A \cap B) = m(A \cup B)$.

Por inducción sobre N obtenemos el siguiente corolario.

(3.5) *Si* $A_1, A_2, \ldots, A_N$ *son conjuntos elementales cualesquiera, entonces*

$$m\left(\bigcup_{k=1}^{N} A_k\right) \leq \sum_{k=1}^{N} m(A_k).$$

La propiedad (3.5) se llama **subaditividad** de la medida.

(3.6) *Si* I *es un intervalo, entonces para cada* $\varepsilon > 0$ *, existen un intervalo cerrado* $H \subset I$ *y un intervalo abierto* $J \supset I$ *, tales que*

$$m(H) > m(I) - \varepsilon, \qquad m(J) < m(I) + \varepsilon.$$

Comencemos observando que si I es un intervalo lineal $(n = 1)$ con extremos a y b , y δ un número positivo, el intervalo cerrado $H = [a+\delta, b-\delta]$ y el intervalo abierto $J = (a-\delta, b+\delta)$ verifican $H \subset I \subset J$ y además, cuando δ tiende a cero, los números $m(H)$ y $m(J)$ tienden a $m(I)$. Utilizando la continuidad del producto, se puede hacer una consideración análoga para cualquier intervalo I del espacio $\mathbb{R}^n$.

(3.7) *Si* A *es un conjunto elemental, entonces, para cada* $\varepsilon > 0$ *, existen un conjunto elemental cerrado* $C \subset A$ *y un conjunto elemental abierto* $B \supset A$ *, tales que*

$$m(C) > m(A) - \varepsilon, \qquad m(B) < m(A) + \varepsilon.$$

Para demostrarlo, supongamos que A es la unión de los intervalos disjuntos $I_1, I_2, \ldots I_N$. Para cada intervalo I_k existen un intervalo cerrado $H_k \subset I_k$ y un intervalo abierto $J_k \supset I_k$, tales que $m(H_k) > m(I_k) - \varepsilon/N$ y $m(J_k) < m(I_k) + \varepsilon/N$. Poniendo

$$C = \bigcup_{k=1}^{N} H_k \quad \text{y} \quad B = \bigcup_{k=1}^{N} J_k,$$

tendremos $C \subset A \subset B$. Además,

$$m(C) = \sum_{k=1}^{N} m(H_k) > \sum_{k=1}^{N} \{m(I_k) - \varepsilon/N\} = m(A) - \varepsilon,$$

$$m(B) \leq \sum_{k=1}^{N} m(J_k) < \sum_{k=1}^{N} \{m(I_k) + \varepsilon/N\} = m(A) + \varepsilon,$$

Puesto que C es cerrado y B abierto, estos conjuntos satisfacen todas las afirmaciones del enunciado.

Llegamos ahora al primer resultado que no pertenece a la teoría clásica de la medida.

(3.8) **Teorema.** *Si un conjunto elemental A está contenido en la unión de una sucesión de conjuntos elementales A_k $(k = 1, 2, 3 \ldots)$, entonces*

$$m(A) \leq \sum_{k=1}^{\infty} m(A_k).$$

Si la serie de la derecha es divergente, su suma es $+\infty$ y la desigualdad no requiere demostración. Supongamos pues, que dicha serie es convergente.

Dado $\varepsilon > 0$, existe un conjunto elemental cerrado $C \subset A$, tal que $m(C) > m(A) - \varepsilon$. Por otra parte, para cada índice k, existe un conjunto elemental abierto $B_k \supset A_k$, tal que $m(B_k) < m(A_k) + \varepsilon/2^k$. Además, las inclusiones

$$C \subset A \subset \bigcup_{k=1}^{\infty} A_k \subset \bigcup_{k=1}^{\infty} B_k$$

muestran que los conjuntos abiertos B_k forman un cubrimiento del conjunto cerrado y acotado C. Puesto que C es compacto, existe un entero positivo

s, tal que $C \subset \bigcup_{k=1}^{s} B_k$ y por consiguiente,

$$m(A) - \varepsilon < m(C) \leq m\left(\bigcup_{k=1}^{s} B_k\right) \leq \sum_{k=1}^{s} m(B_k) < \sum_{k=1}^{s} \{m(A_k) + \varepsilon/2^k\} \leq$$
$$\leq \sum_{k=1}^{\infty} \{m(A_k) + \varepsilon/2^k\} = \sum_{k=1}^{\infty} m(A_k) + \varepsilon.$$

En vista de que el número ε ha sido elegido en forma arbitraria, la desigualdad del enunciado se obtiene haciendo que ε tienda a cero.

De la propiedad demostrada se obtiene el siguiente corolario.

(3.9) *Si el conjunto elemental* A *es la unión de una sucesión de conjuntos elementales* A_k $(k = 1, 2, 3, \ldots)$, *disjuntos dos a dos, entonces*

$$m(A) = \sum_{k=1}^{\infty} m(A_k).$$

La hipótesis implica que $A_k \cap A_j = \emptyset$ si $k \neq j$. Puesto que A está contenido en la unión de los conjuntos A_k, en virtud de (3.8),

$$m(A) \leq \sum_{k=1}^{\infty} m(A_k). \tag{1}$$

Por otra parte, para cada entero positivo s, $\bigcup_{k=1}^{s} A_k \subset A$ y por consiguiente,

$$\sum_{k=1}^{s} m(A_k) = m\left(\bigcup_{k=1}^{s} A_k\right) \leq m(A).$$

Haciendo $s \to \infty$ en la última relación, obtenemos

$$\sum_{k=1}^{\infty} m(A_k) \leq m(A) \tag{2}$$

y la igualdad del enunciado resulta de comparar las relaciones (1) y (2).

La propiedad (3.9) se conoce con el nombre de σ**-aditividad** de la medida. Aclaremos que en esta teoría, la letra griega sigma se usa para referirse a uniones o descomposiciones numerables.

4. Conjuntos $\boldsymbol{\sigma}$-elementales.

Diremos que un conjunto U es σ**-elemental** si existe una sucesión de conjuntos elementales A_k $(k = 1, 2, 3, \ldots)$, disjuntos dos a dos, tal que

$$U = \bigcup_{k=1}^{\infty} A_k.$$

La **medida** del conjunto σ-elemental U se define por la fórmula

$$m(U) = \sum_{k=1}^{\infty} m(A_k) = \lim_{N\to\infty} \sum_{k=1}^{N} m(A_k),$$

de modo que $m(U)$ es un elemento no negativo de la recta extendida (posiblemente igual a $+\infty$).

La definición es correcta, pues si B_j $(j = 1, 2, 3, \ldots)$ es otra sucesión de conjuntos elementales disjuntos cuya unión es el mismo conjunto U, entonces, cada A_k es la unión de los conjuntos elementales disjuntos $A_k \cap B_j$ $(j = 1, 2, 3, \ldots)$ y análogamente, cada B_j es la unión de los conjuntos elementales disjuntos $A_k \cap B_j$ $(k = 1, 2, 3, \ldots)$; y en virtud de (3.9),

$$\sum_{k=1}^{\infty} m(A_k) = \sum_{k=1}^{\infty}\sum_{j=1}^{\infty} m(A_k \cap B_j) = \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} m(A_k \cap B_j) = \sum_{j=1}^{\infty} m(B_j)$$

(3.10) *La unión de cualquier sucesión de conjuntos elementales es un conjunto σ-elemental.*

Debemos probar que si A_k $(k = 1, 2, 3, \ldots)$ es una sucesión de conjuntos elementales, entonces, el conjunto $U = \bigcup_{k=1}^{\infty} A_k$ es σ-elemental, lo que significa exhibir una sucesión de conjuntos elementales disjuntos cuya unión sea U. Con este fin, consideremos la sucesión de conjuntos elementales B_k definidos por las fórmulas

$$\begin{aligned}
B_1 &= A_1, \\
B_2 &= A_2 - A_1, \\
B_3 &= A_3 - (A_1 \cup A_2), \\
B_4 &= A_4 - (A_1 \cup A_2 \cup A_3),
\end{aligned}$$

y en general, para cada entero $k > 1$,

$$B_k = A_k - (A_1 \cup A_2 \cup \ldots \cup A_{k-1}).$$

Los conjuntos B_k son disjuntos y para cada índice k, $B_k \subset A_k$, de modo que llamando V a la unión de todos los B_k, es claro que el conjunto σ-elemental V está contenido en U. Por otra parte, si x es un elemento de U, considerando el mínimo índice k, tal que $x \in A_k$, resultará $x \in B_k$. Hemos probado que $U \subset V$, lo cual junto con lo anterior nos da $U = V$. Luego, U es σ-elemental.

La proposición que acabamos de probar puede expresarse diciendo que cualquier unión numerable de conjuntos elementales es un conjunto σ-elemental.

(3.11) *La unión de cualquier sucesión de conjuntos σ-elementales es un conjunto σ-elemental. La intersección de dos conjuntos σ-elementales es un conjunto σ-elemental.*

En efecto, si $U_k = \bigcup_{j=1}^{\infty} A_{kj}$ $(k = 1, 2, 3, \ldots)$, las fórmulas

$$\bigcup_{k=1}^{\infty} U_k = \bigcup_{k=1}^{\infty} \bigcup_{j=1}^{\infty} A_{kj}, \qquad U_1 \cap U_2 = \bigcup_{k=1}^{\infty} \bigcup_{j=1}^{\infty} (A_{1k} \cap A_{2j}),$$

en vista de (3.10), prueban las afirmaciones de (3.11).

(3.12) *Todo conjunto abierto es un conjunto σ-elemental.*

Demostración. Si $I = J_1 \times \ldots J_n$ es un intervalo de $\mathbb{R}^n$, un punto $v = (v_1, \ldots v_n)$ se llama **vértice** de I, si para cada índice i, v_i es un extremo del intervalo lineal J_i.

Es claro que la colección K formada por todos los cubos de $\mathbb{R}^n$ cuyos vértices son puntos con coordenadas racionales es una colección numerable.

Comenzaremos probando que para cada cubo

$$Q(x, \delta) = (x_1 - \delta, x_1 + \delta) \times \ldots \times (x_n - \delta, x_n + \delta)$$

existe un cubo $Q' \in K$, tal que $x \in Q' \subset Q(x, \delta)$. En efecto, sea r un número racional que verifica $\delta/2 < r < \delta$, y para cada índice i $(1 \le i \le n)$

elijamos un número racional a_i, tal que $x_i - \delta/2 < a_i < x_i$. En estas condiciones, el intervalo con extremos racionales $J_i = (a_i, a_i + r)$ verifica $x_i \in J_i \subset (x_i - \delta, x_i + \delta)$, de modo que el cubo $Q' = J_1 \times \ldots \times J_n$ es un miembro de K que satisface $x \in Q' \subset Q(x, \delta)$.

Si G es un conjunto abierto del espacio $\mathbb{R}^n$, llamemos $K(G)$ a la colección numerable formada por todos los miembros de K que están contenidos en G, y sea U la unión de todos los miembros de la colección $K(G)$, de modo que U es σ-elemental y además, $U \subset G$.

Por otra parte, si $x \in G$, entonces existe un número $\delta > 0$, tal que $Q(x, \delta) \subset G$ (recuérdese que toda bola contiene un cubo) y en virtud de lo que demostramos anteriormente, existe un cubo $Q' \in K$, tal que $x \in Q' \subset Q(x, \delta)$. Por estar contenido en G, Q' es un miembro de $K(G)$, de donde se sigue que el punto x pertenece a U. Luego $G \subset U$, lo cual, junto con la inclusión opuesta que ya teníamos, nos da $G = U$ y por lo tanto, G es un conjunto σ-elemental.

(3.13) *Si el conjunto σ-elemental U está contenido en la unión de una sucesión de conjuntos σ-elementales U_k $(k = 1, 2, 3, \ldots)$, entonces*

$$m(U) \leq \sum_{k=1}^{\infty} m(U_k).$$

En efecto, supongamos que

$$U = \bigcup_{k=1}^{\infty} A_i \quad \text{y} \quad U_k = \bigcup_{j=1}^{\infty} B_{kj},$$

donde los conjuntos elementales A_i son disjuntos y para cada índice k, los conjuntos elementales B_{kj} $(j = 1, 2, 3, \ldots)$ son también disjuntos. En vista de la hipótesis, para cada entero positivo s, tenemos

$$\bigcup_{i=1}^{s} A_i \subset U \subset \bigcup_{k=1}^{\infty} U_k = \bigcup_{k=1}^{\infty} \bigcup_{j=1}^{\infty} B_{kj}$$

y en virtud de (3.8),

$$\sum_{i=1}^{s} m(A_i) = m\left(\bigcup_{i=1}^{s} A_i\right) \leq \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} m(B_{kj}) = \sum_{k=1}^{\infty} m(U_k).$$

Por consiguiente,

$$m(U) = \sum_{i=1}^{\infty} m(A_i) = \lim_{s\to\infty} \sum_{i=1}^{s} m(A_i) \leq \sum_{k=1}^{\infty} m(U_k).$$

Para todo lo que sigue, conviene que enunciemos en general la siguiente definición.

Definición. *Diremos que una función* $\phi : \mathcal{C} \to \overline{I\!R}$ *definida sobre una clase de conjuntos* $\mathcal{C}$ *es* σ**-aditiva***, si para cada sucesión disjunta* (E_k) *de miembros de* $\mathcal{C}$ *cuya unión* E *pertenezca a* $\mathcal{C}$, *se verifica*

$$\phi(E) = \phi\left(\bigcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} \phi(E_k).$$

Por ejemplo, la proposición (3.9) establece que la medida m es σ-aditiva sobre la clase de los conjuntos elementales.

Observaciones complementarias y ejemplos.

(1) La medida de un conjunto elemental es siempre finita; en cambio hemos observado que la medida de un conjunto σ-elemental no siempre lo es.

(2) Puesto que cada conjunto elemental es una unión finita de intervalos disjuntos, se sigue que el conjunto U es σ-elemental si y sólo si existe una sucesión (I_k) de intervalos disjuntos, cuya unión es U.

(3) Todo conjunto numerable es un conjunto σ-elemental con medida igual a cero, pues cada conjunto unitario $\{x\}$ es un intervalo degenerado, cuyos lados son intervalos lineales reducidos a un punto único.

(4) Puesto que el conjunto vacío es un intervalo, todo conjunto elemental es también un conjunto σ-elemental.

(5) El conjunto de Cantor P no es σ-elemental: si existiera una sucesión de intervalos I_k cuya unión es P, entonces, puesto que P tiene interior vacío, cada I_k debería ser un conjunto unitario y P resultaría numerable, lo que nos es cierto. Sin embargo, por la misma construcción del conjunto de Cantor, sabemos que existe una sucesión de conjuntos

elementales F_k cuya intersección es P. Por consiguiente, la intersección de una sucesión de conjuntos σ-elementales puede no ser un conjunto σ-elemental.

(6) Llamando U a la unión de todos los intervalos que substraemos de $[0,1]$ para construir el conjunto de Cantor P, se tiene que $P = [0,1] - U$. Luego, la diferencia entre dos conjuntos σ-elementales puede no ser otro conjunto σ-elemental.

(7) Si U y V son conjuntos σ-elementales que verifican $U \subset V$, entonces $m(U) \leq m(V)$. Esto se deduce de (3.13) escribiendo $V = V \cup \emptyset \cup \emptyset \cup \ldots$.

5. Medida exterior de Lebesgue.

Para cada subconjunto E del espacio $I\!R^n$, definimos la **medida exterior** de E por medio de la fórmula

$$m_e(E) = \inf\{m(U) : U \supset E\},$$

donde el ínfimo se toma sobre todos los conjuntos σ-elementales U que contienen al conjunto E.

La definición es correcta, pues al menos el espacio entero, que es un conjunto σ-elemental, contiene al conjunto E.

De la definición se deduce que para cada número $\varepsilon > 0$, existe un conjunto σ-elemental U que contiene a E, tal que $m(U) \leq m_e(E) + \varepsilon$.

(3.14) *La medida exterior goza de las siguientes propiedades:*

(1) $0 \leq m_e(E) \leq +\infty, \qquad m_e(\emptyset) = 0$ *;*

(2) la relación $E_1 \subset E_2$ *implica* $m_e(E_1) \leq m_e(E_2)$ *;*

(3) $m_e\left(\bigcup_{k=1}^{\infty} E_k\right) \leq \sum_{k=1}^{\infty} m_e(E_k)$ *;*

(4) si V *es un conjunto* σ*-elemental, entonces* $m_e(V) = m(V)$ *;*

(5) si E_1 *y* E_2 *son dos conjuntos con medida exterior finita, denotando por* Δ *la diferencia simétrica, se verifica*

$$|m_e(E_1) - m_e(E_2)| \leq m_e(E_1 \Delta E_2).$$

Las dos primeras son una consecuencia inmediata de la definición. En cuanto a la tercera propiedad, que se conoce con el nombre de σ**-subaditividad**, sólo hace falta probarla en el caso de que la serie de la derecha sea convergente, pues de otro modo es trivialmente verdadera. Suponiendo esto y dado un número positivo ε, para cada índice k existe un conjunto σ-elemental U_k, tal que

$$E_k \subset U_k \quad \text{y} \quad m(U_k) < m_e(E_k) + \varepsilon/2^k.$$

Poniendo

$$E = \bigcup_{k=1}^{\infty} E_k \quad \text{y} \quad U = \bigcup_{k=1}^{\infty} U_k,$$

tendremos $E \subset U$ y por consiguiente,

$$m_e(E) \leq m(U) \leq \sum_{k=1}^{\infty} m(U_k) < \sum_{k=1}^{\infty} \{m_e(E_k) + \varepsilon/2^k\} = \sum_{k=1}^{\infty} m_e(E_k) + \varepsilon.$$

Puesto que ε fue elegido en forma arbitraria, la propiedad (3) resulta haciendo que ε tienda a cero.

Si V es σ-elemental, para cualquier conjunto σ-elemental U que contenga a V, se verifica que $m(V) \leq m(U)$. Luego, $m(V) \leq m_e(V)$. Por otra parte, uno de los conjuntos σ-elementales que contienen a V es el mismo V, de donde $m_e(V) \leq m(V)$ y por lo tanto, $m_e(V) = m(V)$.

La propiedad (5) es una consecuencia de las tres primeras. En efecto de la inclusión $E_1 \subset E_2 \cup (E_1 \Delta E_2)$, deducimos que $m_e(E_1) \leq m_e(E_2) + m_e(E_1 \Delta E_2)$ y análogamente $m_e(E_2) \leq m_e(E_1) + m_e(E_1 \Delta E_2)$, lo cual demuestra (5).

Corolario. *Si (E_k) es una sucesión de conjuntos con medida exterior finita, tal que $m_e(E_k \Delta E)$ tiende a cero cuando k tiende a infinito, entonces la medida exterior de E es finita, y además,*

$$m_e(E) = \lim_{k \to \infty} m_e(E_k).$$

En virtud de (5), sólo hay que probar que la medida exterior de E es finita; pero esto sigue inmediatamente de aplicar la hipótesis en la relación $m_e(E) \leq m_e(E_k) + m_e(E_k \Delta E)$.

La medida exterior está bien definida para cualquier subconjunto E del espacio $I\!R^n$, pero no es σ-aditiva. En realidad, ni siquiera es aditiva, pues según probaremos más adelante, existen dos conjuntos disjuntos E_1 y E_2, tales que $m_e(E_1 \cup E_2) \neq m_e(E_1) + m_e(E_2)$.

Para recuperar la propiedad de σ-aditividad nos veremos obligados a restringir la medida exterior m_e a una clase especial de conjuntos que estudiaremos en la próxima sección.

6. Conjuntos medibles

Diremos que un subconjunto E del espacio $I\!R^n$ es **medible**, si para cada número $\varepsilon > 0$, existe un conjunto σ-elemental U, tal que

$$E \subset U \quad \text{y} \quad m_e(U - E) < \varepsilon.$$

Si E es medible, la medida exterior de E se llama simplemente la **medida** de E y se donota por cualquiera de los símbolos

$$m(E), \quad mE \quad \text{o bien} \quad |E|.$$

De modo que para un conjunto medible cualquiera, "medida" es sinónimo de "medida exterior".

La notación de las barras verticales para indicar la medida de un conjunto es muy cómoda y se usa con frecuencia.

Veamos algunos corolarios inmediatos de la definición:

(1) Todo conjunto σ-elemental es medible, pues si U es σ-elemental y ε un número positivo, entonces $U \subset U$ y $m_e(U - U) = m_e(\emptyset) = 0 < \varepsilon$.

(2) Todo conjunto de medida exterior nula es medible, pues si $m_e(E) = 0$, entonces para cada $\varepsilon > 0$, existe un conjunto σ-elemental U que contiene a E y verifica $m(U) < \varepsilon$. Luego, $m_e(U - E) \leq m(U) < \varepsilon$, lo cual muestra que E es medible.

(3.15) *La unión de cualquier sucesión de conjuntos medibles es un conjunto medible. La intersección de dos conjuntos medibles es medible.*

Sea (E_k) una sucesión de conjuntos medibles y llamemos E a la unión de todos los E_k. Si ε es un número positivo, para cada índice k existe un conjunto σ-elemental U_k, tal que

$$E_k \subset U_k, \qquad m_e(U_k - E_k) < \varepsilon/2^k.$$

Llamando U a la unión de todos los U_k, tendremos

$$E \subset U, \qquad U - E \subset \bigcup_{k=1}^{\infty} (U_k - E_k);$$

por consiguiente,

$$m_e(U - E) \leq \sum_{k=1}^{\infty} m_e(U_k - E_k) < \sum_{k=1}^{\infty} \varepsilon/2^k = \varepsilon,$$

lo cual demuestra que E es medible.

Si E_1 y E_2 son dos conjuntos medibles y ε un número positivo, existen dos conjuntos σ-elementales U_1 y U_2, tales que

$$E_1 \subset U_1, \qquad E_2 \subset U_2,$$

$$m_e(U_1 - E_1) < \varepsilon, \qquad m_e(U_2 - E_2) < \varepsilon.$$

Poniendo $E = E_1 \cap E_2$ y $U = U_1 \cap U_2$, tendremos

$$E \subset U, \qquad U - E \subset (U_1 - E_1) \cup (U_2 - E_2),$$

de donde

$$m_e(U - E) \leq m_e(U_1 - E_1) + m_e(U_2 - E_2) < 2\varepsilon,$$

lo cual, en vista de que ε se eligió arbitrariamente, demuestra que la intersección E es medible.

Definición. *Diremos que el conjunto E es* **finitamente medible** *si es medible y su medida $m(E)$ es finita.*

El siguiente teorema, que caracteriza a los conjuntos finitamente medibles, constituye uno de los resultados fundamentales de la teoría que estamos desarrollando.

(3.16) **Teorema.** *Las siguientes afirmaciones son equivalentes:*

(1) El conjunto E es finitamente medible;

(2) Para cada número $\varepsilon > 0$, existe un conjunto elemental A, tal que $m_e(A \Delta E) < \varepsilon$.

DEMOSTRACIÓN. Con el fin de probar que (1) implica (2), supongamos que E es medible y $m(E) < \infty$. Entonces, dado $\varepsilon > 0$, existe un conjunto σ-elemental U, tal que

$$E \subset U, \qquad m_e(U - E) < \varepsilon;$$

además, de la igualdad $U = E \cup (U - E)$ deducimos que la medida de U es finita, pues

$$m(U) = m_e(U) \leq m(E) + m_e(U - E) < m(E) + \varepsilon < \infty.$$

Teniendo en cuenta que U es σ-elemental, existe una sucesión (I_k) de intervalos disjuntos, tal que

$$U = \bigcup_{k=1}^{\infty} I_k.$$

Puesto que

$$\sum_{k=1}^{\infty} m(I_k) = m(U) < \infty,$$

existe un entero positivo N, que verifica

$$\sum_{k=N+1}^{\infty} m(I_k) < \varepsilon.$$

Ahora, si definimos los conjuntos

$$A = \bigcup_{k=1}^{N} I_k, \qquad V = \bigcup_{k=N+1}^{\infty} I_k,$$

es claro que A es elemental, V σ-elemental, y además, $U = A \cup V$, de donde

$$A - E \subset U - E \quad \text{y} \quad E - A \subset U - A = V.$$

Luego, $m_e(A - E) \leq m_e(U - E) < \varepsilon$ y también

$$m_e(E - A) \leq m(V) = \sum_{k=N+1}^{\infty} m(I_k) < \varepsilon.$$

Finalmente, tomando medidas exteriores en la relación $A \Delta E = (A - E) \cup (E - A)$, obtenemos

$$m_e(A \Delta E) \leq m_e(A - E) + m_e(E - A) < 2\varepsilon,$$

lo cual, en vista de que ε puede elegirse arbitrariamente, demuestra que (1) implica (2).

Supongamos ahora que la afirmación (2) es verdadera. Entonces, dado $\varepsilon > 0$, existe un conjunto elemental A, tal que $m_e(A\Delta E) < \varepsilon$, y en virtud de la definición de medida exterior, existe un conjunto σ-elemental V, que verifica las relaciones

$$A\Delta E \subset V, \qquad m(V) < \varepsilon.$$

El conjunto σ-elemental $U = A \cup V$ contiene al conjunto E, pues $E \subset A \cup (E - A) \subset A \cup V = U$, y su medida $m(U)$ es finita, de donde, en primer lugar, $m_e(E) < \infty$. Además,

$$U - E \subset (U - A) \cup (A - E) \subset V \cup V = V,$$

de donde, $m_e(U - E) \leq m(V) < \varepsilon$. Luego, E es medible y su medida $m(E)$ es finita, con lo cual el teorema queda completamente demostrado.

Q.E.D.

Consideremos una sucesión (ε_k) de números positivos, tal que $\varepsilon_k \to 0$ cuando k tiende a infinito. Si E es finitamente medible, entonces, en virtud del último teorema, para cada índice k existe un conjunto elemental A_k, tal que $m_e(A_k\Delta E) < \varepsilon_k$. Esto muestra que la afirmación (2) de dicho teorema es equivalente al siguiente enunciado:

($2'$) *Existe una sucesión* (A_k) *de conjuntos elementales, tal que* $m_e(A_k\Delta E)$ *tiende a cero cuando* k *tiende a infinito.*

A su vez, este enunciado adopta una forma intuitivamente más clara si introducimos la siguiente definición:

Definición. *Diremos que una sucesión* (E_k) *de conjuntos con medida exterior finita* **converge** *al conjunto* E, *y escribiremos* $E_k \to E$, *si* $m_e(E_k\Delta E) \to 0$ *cuando* $k \to \infty$.

Con la ayuda de esta definición, la afirmación ($2'$) se puede expresar del modo siguiente:

($2''$) *Existe una sucesión de conjuntos elementales* (A_k), *tal que* $A_k \to E$;

y el teorema (3.16) adopta la siguiente forma que consideramos útil enunciar explícitamente:

(3.17) *El conjunto* E *es finitamente medible si y sólo si existe una sucesión de conjuntos elementales* (A_k), *tal que* $A_k \to E$.

La utilidad de este teorema reside en el hecho de que las operaciones de unión, intersección y diferencia son "continuas" con respecto a la noción de convergencia que hemos introducido. Más precisamente:

(3.18) **Teorema.** *Si* $E_k \to E$ *y* $F_k \to F$, *entonces*

$$E_k - F_k \to E - F, \quad E_k \cup F_k \to E \cup F \quad \text{y} \quad E_k \cap F_k \to E \cap F.$$

Demostración. La demostración se basa en las siguientes inclusiones:

(i) $(E_k - F_k)\Delta(E - F) \subset (E_k \Delta E) \cup (F_k \Delta F)$,

(ii) $(E_k \cup F_k)\Delta(E \cup F) \subset (E_k \Delta E) \cup (F_k \Delta F)$,

(iii) $(E_k \cap F_k)\Delta(E \cap F) \subset (E_k \Delta E) \cup (F_k \Delta F)$,

cuya fácil verificación dejamos a cargo del lector.

Para completar la demostración, no hay más que tomar medidas exteriores en cada miembro de estas relaciones, teniendo en cuenta las propiedades de la medida exterior, establecidas en la proposición (3.14).

(3.19) **Teorema.** *Si* E *y* F *son conjuntos medibles, entonces* $E - F$ *es medible.*

Demostración. Supongamos primero que E y F son finitamente medibles. Entonces existen dos sucesiones de conjuntos elementales (A_k) y (B_k), tales que $A_k \to E$ y $B_k \to F$; y en virtud de (3.18), la sucesión de conjuntos elementales $(A_k - B_k)$ verifica $A_k - B_k \to E - F$. Luego, $E - F$ es finitamente medible y en particular, es medible.

Sean ahora E y F dos conjuntos medibles cualesquiera, y denotemos por $Q_k = Q(0,k)$ el cubo con centro en el origen y lados de longitud igual a $2k$ $(k = 1, 2, 3, \ldots)$. Puesto que el espacio $I\!R^n$ es la unión de todos los cubos Q_k, tendremos:

$$E - F = \bigcup_{k=1}^{\infty} [(E - F) \cap Q_k] = \bigcup_{k=1}^{\infty} [(E \cap Q_k) - (F \cap Q_k)].$$

Ahora bien; el cubo Q_k es finitamente medible y por consiguiente, $E \cap Q_k$ y $F \cap Q_k$ son también finitamente medibles, en virtud de (3.15). Luego, para cada índice k, la diferencia $(E \cap Q_k) - (F \cap Q_k)$ es medible, de donde resulta que $E - F$ es medible por ser la unión de una sucesión de conjuntos medibles, nuevamente en virtud de (3.15).

(3.20) **Corolario.** *El complemento de cada conjunto medible es medible.*

En efecto, si E es medible, entonces $CE = \mathbb{R}^n - E$ es la diferencia de dos conjuntos medibles.

(3.21) **Teorema.** *La intersección de cualquier sucesión de conjuntos medibles es un conjunto medible.*

La demostración resulta inmediatamente de las propiedades establecidas y de la fórmula de complementación

$$\bigcap_{k=1}^{\infty} E_k = C\left(\bigcup_{k=1}^{\infty} CE_k\right).$$

Resumiendo, hemos demostrado que la unión y la intersección de cualquier sucesión de conjuntos medibles son conjuntos medibles, y que el complemento de cada conjunto medible es medible.

Nuestro próximo paso será probar que la medida es σ-aditiva sobre la clase de los conjuntos medibles; en otras palabras, que si (E_k) es una sucesión de conjuntos medibles disjuntos, entonces

$$m(E_1 \cup E_2 \cup \ldots) = m(E_1) + m(E_2) + \ldots$$

Comenzamos probando que si E y F son dos conjuntos medibles cualesquiera, entonces se verifica

$$m(E \cup F) + m(E \cap F) = m(E) + m(F). \tag{3.22}$$

Si alguno de los conjuntos tiene medida infinita, entonces ambos miembros son iguales a $+\infty$ y la relación es trivialmente cierta. Supongamos, pues, que E y F son finitamente medibles. Entonces, existen dos sucesiones de conjuntos elementales (A_k) y (B_k), tales que $A_k \to E$ y $B_k \to F$, de donde se sigue que $A_k \cup B_k \to E \cup F$ y $A_k \cap B_k \to E \cap F$. Ahora bien; en virtud de (3.4),

$$m(A_k \cup B_k) + m(A_k \cap B_k) = m(A_k) + m(B_k).$$

Si en esta relación tomamos el límite de cada miembro cuando k tiende a infinito, obtenemos la igualdad (3.22).

Q.E.D.

En particular, si E y F son dos conjuntos medibles disjuntos, entonces $m(E \cap F) = m(\emptyset) = 0$, y la fórmula (3.22) nos da la igualdad

$$m(E \cup F) = m(E) + m(F) \qquad (E \cap F = \emptyset).$$

Por inducción sobre N, obtenemos fácilmente el siguiente cololario.

(3.23) **Corolario.** *Si $E_1, E_2, \ldots, E_N$ son conjuntos medibles disjuntos, entonces*

$$m(E_1 \cup E_2 \cup \ldots \cup E_N) = m(E_1) + m(E_2) + \ldots + m(E_N).$$

Llegamos ahora a uno de los resultados más importantes de la teoría de Lebesgue:

(3.24) **Teorema.** *Si el conjunto E es la unión de una sucesión (E_k) formada por conjuntos medibles disjuntos, entonces*

$$m(E) = \sum_{k=1}^{\infty} m(E_k).$$

Demostración. Puesto que para cada conjunto medible E, $m(E) = m_e(E)$, en el caso que estamos considerando, tendremos

$$m(E) = m_e\left(\bigcup_{k=1}^{\infty} E_k\right) \leq \sum_{k=1}^{\infty} m_e(E_k) = \sum_{k=1}^{\infty} m(E_k).$$

Por otra parte, teniendo en cuenta que los conjuntos medibles E_k son disjuntos, para cada entero positivo N,

$$\sum_{k=1}^{N} m(E_k) = m\left(\bigcup_{k=1}^{N} E_k\right) \leq m(E),$$

de modo que haciendo que N tienda a infinito, resulta la relación

$$\sum_{k=1}^{\infty} m(E_k) \leq m(E),$$

y el teorema queda demostrado.

Observando la última demostración, notemos de paso que para cualquier sucesión de conjuntos medibles (E_k) – disjuntos o no –, se verifica

$$m\left(\bigcup_{k=1}^{\infty} E_k\right) \leq \sum_{k=1}^{\infty} m(E_k). \tag{3.25}$$

Esta propiedad se expresa diciendo que la medida es σ-subaditiva.

(3.26) *Si E es medible, F es finitamente medible y $F \subset E$, entonces $m(E-F) = m(E) - m(F)$.*

La demostración resulta de aplicar el corolario (3.23) en la relación $E = F \cup (E - F)$.

7. Sucesiones monótonas de conjuntos medibles.

Veremos en esta sección que, como consecuencia de la σ-aditividad, la medida de Lebesgue tiene ciertas propiedades de continuidad con respecto a las sucesiones monótonas de conjuntos medibles.

(3.27) *Para cualquier sucesión creciente de conjuntos medibles:*

$$E_1 \subset E_2 \subset \ldots \subset E_k \subset E_{k+1} \subset \ldots,$$

se cumple

$$m\left(\bigcup_{k=1}^{\infty} E_k\right) = \lim_{k\to\infty} m(E_k).$$

Para demostrarlo, notemos que el conjunto E igual a la unión de todos los E_k se puede representar como la unión de una sucesión de conjuntos medibles disjuntos de la siguiente manera:

$$E = E_1 \cup (E_2 - E_1) \cup (E_3 - E_2) \cup \ldots \cup (E_k - E_{k-1}) \cup \ldots,$$

y análogamente

$$E_k = E_1 \cup (E_2 - E_1) \cup (E_3 - E_2) \cup \ldots \cup (E_k - E_{k-1}).$$

Luego

$$\begin{aligned} m(E) &= m(E_1) + m(E_2 - E_1) + \ldots + m(E_k - E_{k-1}) + \ldots \\ &= \lim_{k\to\infty} \{m(E_1) + m(E_2 - E_1) + \ldots + m(E_k - E_{k-1})\} \\ &= \lim_{k\to\infty} m(E_k). \end{aligned}$$

(3.28) *Si* $E_1 \supset E_2 \supset E_3 \supset \ldots$ *es una sucesión decreciente de conjuntos medibles y si para algún* k *,* $m(E_k) < \infty$ *, entonces*

$$m\left(\bigcap_{k=1}^{\infty} E_k\right) = \lim_{k\to\infty} m(E_k).$$

Comencemos observando que el límite del miembro derecho existe y es finito, en razón de que la sucesión $m(E_k)$ es monótona decreciente.

Sin restricción de generalidad, supongamos que $m(E_1) < \infty$. Los conjuntos $E_1 - E_k$ ($k = 1, 2, 3, \ldots$) forman una sucesión creciente, a la cual es posible aplicar (3.27), de la manera siguiente:

$$\begin{aligned} m(E_1) - m\left(\bigcap_{k=1}^{\infty} E_k\right) &= m\left(E_1 - \bigcap_{k=1}^{\infty} E_k\right) = m\left(\bigcup_{k=1}^{\infty}(E_1 - E_k)\right) \\ &= \lim_{k\to\infty} m(E_1 - E_k) = m(E_1) - \lim_{k\to\infty} m(E_k), \end{aligned}$$

y la igualdad del enunciado se obtiene cancelando el término $m(E_1)$ y cambiando los signos en la nueva igualdad.

8. Conjuntos de medida nula.

Hemos señalado que si Z es un subconjunto de $\mathbb{R}^n$ que verifica $m_e(Z) = 0$, entonces Z es medible y podemos escribir $m(Z) = 0$. Tales conjuntos se llaman **conjuntos de medida nula**.

Si (Z_k) es una sucesión de conjuntos de medida nula, llamando Z a la unión de todos los Z_k, tendremos

$$m(Z) \leq \sum_{k=1}^{\infty} m(Z_k) = 0.$$

Luego, cualquier unión numerable de conjuntos de medida nula, es también un conjunto de medida nula.

Todo conjunto numerable es un conjunto de medida nula; pero existen conjuntos de medida nula no numerables: esto es fácil de ver si la dimensión n es mayor que uno, pues cualquier segmento paralelo a uno de los ejes tiene medida nula.

En el caso de la recta $(n = 1)$, el conjunto de Cantor P tiene medida nula, pues para cada entero positivo k, P está contenido en un conjunto elemental F_k que es la unión de 2^k intervalos cerrados disjuntos de longitud $1/3^k$, de donde se sigue que

$$m(P) \leq m(F_k) = (2/3)^k,$$

de modo que haciendo que k tienda a infinito, resulta $m(P) = 0$.

9. Estructura de los conjuntos medibles.

En la teoría de la medida la letra griega delta se usa para simbolizar intersecciones numerables, así como la letra sigma se usa para referirse a uniones numerables, en la forma que vamos a ejemplificar.

Diremos que un conjunto H es **de clase** G_δ si existe una sucesión de conjuntos abiertos $G_1, G_2, G_3, \ldots,$ tal que

$$H = \bigcap_{k=1}^{\infty} G_k.$$

Un conjunto E se llama **de clase** F_σ si existe una sucesión de conjuntos cerrados (F_k) tal que

$$E = \bigcup_{k=1}^{\infty} F_k.$$

De las fórmulas de complementación resulta que el complemento de cada conjunto de clase G_δ es un conjunto de clase F_σ y viceversa.

Cada conjunto abierto es medible por ser un conjunto σ-elemental. Luego, cada conjunto cerrado es medible por ser el complemento de un conjunto medible. Además, de las definiciones anteriores se sigue que todos los conjuntos de clase G_δ y todos los conjuntos de clase F_σ son conjuntos medibles.

El siguiente teorema y su corolario ilustran sobre la utilidad de las clases de conjuntos que acabamos de definir.

(3.29) **Teorema.** *Las siguientes afirmaciones son equivalentes:*

(a) el conjunto E es medible;

(b) para cada $\varepsilon > 0$, existe un conjunto abierto G, tal que

$$E \subset G \quad y \quad m_e(G - E) < \varepsilon;$$

(c) existen un conjunto H de clase G_δ y un conjunto Z de medida nula, tales que $E = H - Z$.

Demostración. Supongamos que E es medible y sea ε un número positivo. Entonces existe un conjunto σ-elemental U, tal que

$$E \subset U \quad \text{y} \quad m_e(U - E) < \varepsilon.$$

Por ser U un conjunto σ-elemental, existe una sucesión de intervalos disjuntos (I_k), tal que

$$U = \bigcup_{k=1}^{\infty} I_k,$$

y en virtud de (3.6), para cada índice k, existe un intervalo abierto J_k, tal que

$$I_k \subset J_k, \qquad m(J_k) < m(I_k) + \varepsilon/2^k.$$

El conjunto abierto

$$G = \bigcup_{k=1}^{\infty} J_k$$

verifica las relaciones

$$G \supset U \supset E, \qquad G - U \subset \bigcup_{k=1}^{\infty} (J_k - I_k),$$

de donde

$$\begin{aligned} m(G-U) &\le \sum_{k=1}^{\infty} m(J_k - I_k) = \sum_{k=1}^{\infty} \{m(J_k) - m(I_k)\} \\ &< \sum_{k=1}^{\infty} \varepsilon/2^k = \varepsilon. \end{aligned}$$

Puesto que $G - E = (G - U) \cup (U - E)$, por la aditividad de la medida, tendremos

$$\begin{aligned} m_e(G-E) = m(G-E) &= m(G-U) + m(U-E) \\ &< \varepsilon + \varepsilon = 2\varepsilon. \end{aligned}$$

Queda demostrado que (a) implica (b).

Supongamos ahora que la afirmación (b) es verdadera. Entonces, para cada número natural k, existe un conjunto abierto G_k, tal que $E \subset G_k$, $m_e(G_k - E) < 1/k$.

El conjunto de clase G_δ

$$H = \bigcap_{k=1}^{\infty} G_k$$

contiene a E, pues cada G_k contiene a E. Además, para cada k, $H - E \subset G_k - E$. Luego,

$$m_e(H-E) \le m_e(G_k - E) < 1/k,$$

y haciendo que k tienda a infinito, resulta $m_e(H - E) = 0$.

Si ahora definimos $Z = H - E$, es claro que $E = H - Z$, donde H es de clase G_δ y $m(Z) = 0$. Hemos probado que (b) implica (c).

Finalmente, si la afirmación (c) es verdadera, entonces E es medible por ser la diferencia entre dos conjuntos medibles, lo cual demuestra que (c) implica (a) y el teorema queda demostrado.

(3.30) **Corolario.** *Todo conjunto medible es la unión de un conjunto de clase F_σ con un conjunto de medida nula.*

En efecto, si E es medible, también lo es su complemento. Luego, existen un conjunto H de clase G_δ y un conjunto Z de medida nula, tales que $CE = H - Z = H \cap CZ$, de modo que $E = CH \cup Z$; pero CH es un conjunto de clase F_σ.

Veamos otra consecuencia útil de (3.29).

(3.31) *Si E es medible, entonces, para cada $\varepsilon > 0$, existen un conjunto cerrado F y un conjunto abierto G, tales que*

$$F \subset E \subset G \qquad m(G - F) < \varepsilon.$$

Para demostrar esta proposición, comencemos por elegir un conjunto abierto G, tal que $G \supset E$, $m(G - E) < \varepsilon/2$. Además, puesto que el complemento de E es también medible, existe un conjunto abierto G', tal que $G' \supset CE$, $m(G' - CE) < \varepsilon/2$.

Llamando F al complemento de G', tendremos $F = CG' \subset E$ y también $E - F = G' - CE$, de donde $m(E - F) = m(G' - CE) < \varepsilon/2$. Finalmente, de la igualdad $G - F = (G - E) \cup (E - F)$, deducimos

$$m(G - F) = m(G - E) + m(E - F) < \varepsilon,$$

lo cual demuestra la proposición.

Q.E.D.

Si E es medible, existen un conjunto H de clase G_δ y un conjunto Z de medida nula, tales que $E = H - Z$. Ahora bien; en virtud de (3.26), $m(E) = m(H) - m(Z) = m(H)$. Hemos demostrado la siguiente afirmación:

(3.32) *Todo conjunto medible está contenido en un conjunto de clase G_δ con igual medida; en particular, todo conjunto de medida nula está contenido en un conjunto de clase G_δ de medida nula.*

Otra consecuencia sencilla de (3.29) es que la medida exterior puede definirse por medio de los conjuntos abiertos. Más explícitamente, vamos a probar que para cada subconjunto E de $\mathbb{R}^n$, se verifica

(3.33) $$m_e(E) = \inf\{m(G) : G \supset E\},$$

donde el ínfimo se toma sobre todos los conjuntos abiertos G que contienen a E.

En efecto, para cada número $\varepsilon > 0$, existe un conjunto σ-elemental U, tal que

$$E \subset U, \qquad m(U) \leq m_e(E) + \varepsilon.$$

Además, puesto que U es medible, existe un conjunto abierto G, tal que

$$U \subset G, \qquad m(G - U) < \varepsilon.$$

De la relación $G = U \cup (G - U)$, deducimos

$$m(G) = m(U) + m(G - U) \leq m_e(E) + 2\varepsilon$$

y por consiguiente, $\inf\{m(G) : G \supset E\} \leq m_e(E)$. Como la desigualdad opuesta se verifica aún más fácilmente, pues $E \subset G$ implica $m_e(E) \leq m(G)$, la fórmula (3.33) queda demostrada.

10. Conjuntos borelianos.

Una clase de conjuntos Σ se llama una σ**-álgebra**, si verifica las siguientes condiciones

(1) $\emptyset \in \Sigma$;

(2) La unión de cualquier sucesión (E_k) de miembros de Σ es un miembro de Σ;

(3) El complemento de cada miembro de Σ es un miembro Σ.

Si Σ es una σ-álgebra, entonces

i) $\mathbb{R}^n = C\emptyset$ es un miembro de Σ;

ii) La intersección de cualquier sucesión (E_k) de miembros de Σ es un miembro de Σ.

La segunda afirmación se deduce de la fórmula de complementación

$$\bigcap_{k=1}^{\infty} E_k = C\left(\bigcup_{k=1}^{\infty} CE_k\right).$$

y de las propiedades (2) y (3).

Ejemplos de σ-álgebras.

1) La clase $\mathcal{M}$ formada por todos los conjuntos medibles .
2) La clase $\mathcal{T} = \{\emptyset, I\!R^n\}$ cuyos únicos miembros son el vacío y el espacio.
3) La clase $\mathcal{P}(I\!R^n)$ formada por todos los subconjuntos del espacio $I\!R^n$.
4) La clase formada por todos los conjuntos E, tales que E es numerable, o bien el complemento de E es numerable.
5) La clase formada por todos los conjuntos E, tales que $m(E) = 0$ o bien $m(CE) = 0$.

(3.33) *Si $\Gamma = (\Sigma)$ es una colección de σ-álgebras, entonces la intersección*

$$\Sigma_0 = \bigcap \Gamma = \bigcap \{\Sigma : \Sigma \in \Gamma\}$$

es una σ-álgebra.

Es decir, la intersección de cualquier colección de σ-álgebras es una σ-álgebra. La demostración es un sencillo ejercicio si se tiene presente que por definición de intersección, $E \in \Sigma_0$ si y sólo si E pertenece a cada miembro de Γ y que además, cada miembro de Γ es una σ-álgebra.

Si $\mathcal{C}$ es una clase de subconjuntos de $I\!R^n$, denotemos por $\sigma(\mathcal{C})$ a la intersección de todas las σ-álgebras Σ, tales que $\mathcal{C} \subset \Sigma$ (por lo menos hay una, ya que $\mathcal{C} \subset \mathcal{P}(I\!R^n)$).

En virtud de (3.33) la clase $\sigma(\mathcal{C})$ es una σ-álgebra que posee las siguientes propiedades:

1a) $\mathcal{C} \subset \sigma(\mathcal{C})$;

2a) Si Σ es una σ-álgebra que verifica $\mathcal{C} \subset \Sigma$, entonces $\sigma(\mathcal{C}) \subset \Sigma$.

Es decir, $\sigma(\mathcal{C})$ es la mínima σ-álgebra que contiene a $\mathcal{C}$.

Definición 1. *La clase $\sigma(\mathcal{C})$ se llama la σ-álgebra* **generada** *por la colección $\mathcal{C}$.*

Es inmediato que si $\mathcal{C}_1 \subset \mathcal{C}_2$, entonces $\sigma(\mathcal{C}_1) \subset \sigma(\mathcal{C}_2)$.

En lo que sigue, denotaremos por $\mathcal{I}$, $\mathcal{E}$, $\mathcal{U}$ y $\mathcal{G}$, respectivamente, a las clases formadas por todos los intervalos de $I\!R^n$, los conjuntos elementales, los conjuntos σ-elementales y los conjuntos abiertos. Claramente

$$\mathcal{I} \subset \mathcal{E} \subset \mathcal{U} \quad \text{y} \quad \mathcal{G} \subset \mathcal{U} \qquad \text{(proposición (3.12))}.$$

Definición 2. *La σ-álgebra $\mathcal{B} = \sigma(\mathcal{I})$, generada por la clase de los intervalos, se llama la σ-álgebra de Borel. Los miembros de $\mathcal{B}$ se llaman* **conjuntos borelianos**.

Por ser la unión de una sucesión de intervalos, cada conjunto σ-elemental es un conjunto boreliano; en particular, todo conjunto abierto es boreliano. Luego, todo conjunto cerrado es un conjunto boreliano.

Los conjuntos de clase G_δ y los de clase F_σ son también conjuntos borelianos, del mismo modo que los conjuntos de clase $G_{\delta\sigma}$ (uniones numerables de conjuntos de clase G_δ) y los conjuntos de clase $F_{\sigma\delta}$ (intersecciones numerables de conjuntos de clase F_σ).

Otras clases de conjuntos borelianos más generales son los conjuntos de clase $G_{\delta\sigma\delta}$, $F_{\sigma\delta\sigma}$, $G_{\delta\sigma\delta\sigma}$, $F_{\sigma\delta\sigma\delta}$, etc.

La σ-álgebra de Borel $\mathcal{B}$ se puede construir por este proceso de etapas sucesivas; pero para ello se requiere el conocimiento de los números ordinales y la inducción transfinita, conocimiento que no presuponemos en el lector. Nos contentamos con señalar que la inducción finita no es suficiente para construir la σ-álgebra de Borel. En cambio, un hecho importante que está inmediatamente a nuestro alcance el el siguiente:

(3.34) *Todo conjunto boreliano es medible.*

En efecto, puesto que $\mathcal{I} \subset \mathcal{M}$ y la segunda clase es una σ-álgebra, se deduce que $\mathcal{B} = \sigma(\mathcal{I}) \subset \mathcal{M}$.

Q.E.D.

Aparte de los intervalos, hay otras clases de conjuntos que generan la σ-álgebra de Borel $\mathcal{B}$. Para comprenderlo, comenzaremos probando la siguiente proposición

(3.35) *Todo intervalo es un conjunto de clase G_δ.*

Esto es inmediato en la recta. Por ejemplo, $(a, b] = \bigcap_{k=1}^{\infty}(a, b + 1/k)$ es la intersección de una sucesión de intervalos abiertos, y análogamente se procede para los otros tipos de intervalos.

Si $I = I_1 \times \ldots \times I_n$ es un intervalo de $I\!R^n$, consideremos para cada índice $i = 1, 2, \ldots, n$, una sucesión de intervalos abiertos $J_i^{(k)}$ $(k = 1, 2, 3, \ldots)$

cuya intersección sea el intervalo I_i. Llamando $J^{(k)}$ al producto cartesiano $J_1^{(k)} \times \ldots \times J_n^{(k)}$, es claro que I es la intersección de los intervalos abiertos $J^{(k)}$.

Q.E.D.

En virtud de (3.35), $\mathcal{I} \subset \sigma(\mathcal{G})$ y por consiguiente, $\mathcal{B} = \sigma(\mathcal{I}) \subset \sigma(\mathcal{G})$. Por otra parte, habíamos visto que $\mathcal{G} \subset B$, de donde $\sigma(\mathcal{G}) \subset \mathcal{B}$; es decir, $\mathcal{B} = \sigma(\mathcal{G})$. En otras palabras: la σ-álgebra de Borel es idéntica a la σ-álgebra generada por la clase de los conjuntos abiertos.

La clase $\mathcal{F}$ formada por todos los conjuntos cerrados, así como las clases $\mathcal{E}$ y $\mathcal{U}$ son otras clases que generan la σ-álgebra de Borel. La fácil demostración de estos hechos queda a cargo del lector.

Cuando exista alguna posibilidad de confusión, usaremos el símbolo $\mathcal{B}_n$ para designar la σ-álgebra de Borel del espacio euclidiano n-dimensional $\mathbb{R}^n$. En particular, $\mathcal{B}_1$ es la σ-álgebra de Borel de la recta, que jugará un papel muy destacado en el capítulo siguiente.

11. Invariancia bajo translaciones

Dado un vector $z = (z_1, \ldots, z_n)$ del espacio $\mathbb{R}^n$, la aplicación T_z de $\mathbb{R}^n$ en sí mismo, definida por

$$T_z(x) = x + z \qquad (x \in \mathbb{R}^n)$$

se llama la **translación** según el vector z.

La aplicación T_z, así como su inversa $T_z^{-1} = T_{-z}$ son ambas continuas; de modo que T_z aplica cada conjunto abierto en otro conjunto abierto.

Para cada conjunto $E \subset \mathbb{R}^n$, escribiremos $E + z$ en lugar de $T_z(E)$.

Si I es un intervalo definido por las desigualdades $a_i < x_i \leq b_i$ ($i = 1, \ldots, n$), entonces $I + z$ es el intervalo J definido por las desigualdades $a_i + z_i < x_i \leq b_i + z_i$ ($i = 1, \ldots, n$). Por consiguiente, T_z aplica cada intervalo en otro intervalo de la misma medida.

Si G es un conjunto abierto, entonces existe una sucesión de intervalos disjuntos (I_k) cuya unión es G, y el conjunto abierto $G + z$ es la unión de los intervalos $I_k + z$, también disjuntos, pues la translación es una aplicación biyectiva. Luego,

$$m(G+z)=\sum_{k=1}^{\infty} m(I_k+z)=\sum_{k=1}^{\infty} m(I_k)=m(G).$$

Vamos a probar que para cualquier conjunto E, se verifica

$$m_e(E+z)=m_e(E) \tag{3.36}$$

En efecto, si G es un conjunto abierto que contiene a E, entonces $G+z$ contiene a $E+z$, y por consiguiente,

$$m_e(E+z)\leq m(G+z)=m(G),$$

y en virtud de (3.33), $m_e(E+z)\leq m_e(E)$. Luego,

$$m_e(E)=m_e((E+z)+(-z))\leq m_e(E+z),$$

de donde resulta la igualdad (3.36).

Q.E.D.

Si E es medible, entonces, para cada $\varepsilon>0$, existe un conjunto abierto G, tal que $E\subset G$ y $m_e(G-E)<\varepsilon$. Ahora bien; de estas relaciones se deduce que $E+z\subset G+z$, y además,

$$m_e((G+z)-(E+z))=m_e((G-E)+z)=m_e(G-E)<\varepsilon.$$

Luego, $E+z$ es medible y podemos resumir todas las consideraciones precedentes en el siguiente teorema.

(3.37) **Teorema.** *Si E es medible, entonces el conjunto transladado $E+z$ es también medible y además, $m(E+z)=m(E)$.*

El teorema (3.37) expresa una importante propiedad de la medida de Lebesgue, que se conoce con el nombre **invariancia bajo translaciones**.

En general, vamos a decir que una función de conjunto $\phi(E)$ definida para cada conjunto E de cierta clase, con valores en la recta extendida, es **invariante bajo translaciones**, si cualquier transladado $E+z$ de un conjunto de la clase dada es un conjunto de la misma clase y además, $\phi(E+z)=\phi(E)$. Por ejemplo, la medida exterior m_e es invariante bajo translaciones.

12. Conjuntos no medibles; conjunto de Vitali.

En esta sección demostraremos la existencia de conjuntos no medibles en el sentido de Lebesgue y juntamente con ello la imposibilidad de extender la medida de Lebesgue a todos los subconjuntos de $\mathbb{R}^n$, preservando las propiedades fundamentales de σ-aditividad e invariancia bajo translaciones. Al mismo tiempo veremos que la medida exterior m_e no es aditiva; es decir, que existen dos conjuntos disjuntos S y T, tales que $m_e(S \cup T) < m_e(S) + m_e(T)$.

Por razones de simplicidad nos limitaremos a trabajar en el espacio $\mathbb{R}^1$; pero el lector comprenderá sin dificultad que en lo principal los enunciados se mantienen válidos cualquiera que sea la dimensión.

Los razonamientos de esta sección estarán basados en un axioma de la teoría de los conjuntos que se conoce con el nombre de **axioma de elección** o de Zermelo:

> **Axioma de Zermelo.** *Para cualquier colección de conjuntos no vacíos y disjuntos, existe un conjunto que contiene exactamente un elemento de cada conjunto de la colección dada.*

El siguiente teorema se debe al matemático Giuseppe Vitali.

(3.39) **Teorema.** *Existe un conjunto $V \subset \mathbb{R}$ que no es medible en el sentido de Lebesgue.*

Demostración. Convengamos en decir que dos elementos cualesquiera x e y del intervalo $E = (0, 1)$ son **equivalentes** si la diferencia $x - y$ es un número racional.

El lector verificará fácilmente que la relación así definida es, efectivamente, una relación de equivalencia (reflexiva, simétrica y transitiva) entre elementos de E.

Las clases de equivalencia C_x ($x \in E$) forman una partición de E, y en virtud del axioma de Zermelo, exite un conjunto V que contiene exactamente un elemento de cada clase de equivalencia. Nos proponemos demostrar que el conjunto $V \subset E$, comúnmente llamado **conjunto de Vitali**, no es medible en el sentido de Lebesgue.

Si x es un elemento de E, entonces existe un único elemento v del conjunto V, tal que $x - v = r$ es un número racional del intervalo $(-1, 1)$.

Denotando por A el conjunto de los números racionales de dicho intervalo, cada elemento x de E admite una única representación de la forma $x = v+r$, donde v es un elemento de V y r un elemento de A, de donde se siguen las inclusiones

$$E \subset \bigcup_{r \in A} (V+r) \subset (-1,2). \tag{1}$$

Los conjuntos $V + r$ $(r \in A)$ son disjuntos, pues si $V + r$ tuviera un punto común z con el conjunto $V + s$, entonces existirían dos elementos v y w del conjunto V, tales que $z = v+r = w+s$, de donde se deduce que v es equivalente a w, y por consiguiente, $v = w$ y $r = s$, pues dos elementos de V no pueden ser equivalentes, a menos que sean idénticos.

Suponiendo que V sea medible, pongamos $a = m(V)$. Entonces tendremos $m(V+r) = m(V) = a$, en virtud de (3.37), y como A es numerable, tomando medidas en las inclusiones (1), obtenemos

$$1 \leq a + a + a + \ldots \leq 3.$$

Estas relaciones muestran que no puede ser $a = 0$, ni tampoco $a > 0$. Llegamos así a una contradicción que provino de suponer que V es medible.

Q.E.D.

En la construcción anterior se puede substituir el intervalo $(0,1)$ por cualquier conjunto de medida positiva $E \subset (0,1)$ o más generalmente, por cualquier conjunto acotado de medida positiva. El resultado sería que todo conjunto acotado de medida positiva contiene un conjunto no medible; y aun la condición de acotación puede suprimirse, como se verá en uno de los ejercicios al final del capítulo.

El siguiente corolario es una consecuencia inmediata de lo anterior.

(3.40) **Corolario.** *No existe ninguna función de conjunto $m'(S)$ definida para cada conjunto $S \subset \mathbb{R}$, tal que m' sea σ-aditiva, invariante bajo translaciones, y verifique $m'(J) = m(J)$ para cualquier intervalo J.*

En efecto, si tal función existiera, volviendo a las inclusiones (1) y poniendo $a = m'(V) = m'(V+r)$, obtendríamos nuevamente la misma contradicción.

El corolario que acabamos de demostrar afirma la imposibilidad de extender la medida de Lebesgue a todos lo subconjuntos de $I\!R$, preservando las propiedades de σ-aditividad e invariancia bajo translaciones.

(3.41) *Existen dos conjuntos disjuntos S y T, tales que*

$$m_e(S \cup T) \neq m_e(S) + m_e(T).$$

En efecto, si tales conjuntos no existieran, la medida exterior m_e sería aditiva y como es σ-subaditiva, resultaría ser σ-aditiva en virtud de una demostración completamente análoga a la del teorema (3.24). Puesto que además m_e es invariante bajo translaciones y para cada conjunto medible E se cumple $m_e(E) = m(E)$, estaríamos en contradicción con el corolario (3.40).

13. Medidas de Lebesgue-Stieljes.

Imitando con un poco de cuidado el procedimiento que hemos usado para definir la medida de Lebesgue, es posible introducir un concepto de medida más general que tiene interés por sus importantes aplicaciones, particularmente en la teoría de las probabilidades. El cuidado se refiere a que mientras la medida de Lebesgue de un conjunto unitario es igual a cero, no ocurrirá así, necesariamente, con la noción que vamos a introducir.

Con esta nueva noción de medida se gana en generalidad y se conserva la propiedad de σ-aditividad, pero se abandona la propiedad de invariancia bajo translaciones.

En esta sección, la palabra **intervalo** designa exclusivamente un conjunto de la forma

$$I = (a_1, b_1] \times \ldots \times (a_n, b_n] \qquad (a_i \leq b_i);$$

es decir, un producto cartesiano de n intervalos lineales que en el caso de no ser vacíos incluyen el extremo derecho, pero no el extremo izquierdo. En particular, un intervalo de la recta será, exclusivamente, un intervalo de la forma $(a, b]$.

La intersección de dos intervalos es un intervalo. Además, por inducción sobre la dimensión n se demuestra que la diferencia entre dos intervalos es un **conjunto elemental**, es decir, una unión finita de intervalos disjuntos. De estas propiedades se deduce igual que antes que si A y B son conjuntos elementales, entonces $A \cap B$, $A - B$ y $A \cup B$ son también conjuntos elementales.

Supongamos ahora que a cada intervalo I se le ha asignado un número real no negativo $\mu(I)$, al cual llamaremos la **medida** de I, de modo tal que se cumplen las siguientes propiedades:

1a) (**aditividad.**) Si el intervalo I es la unión de los intervalos disjuntos $I_1, I_2, \ldots, I_N$, entonces

$$\mu(I) = \mu(I_1) + \mu(I_2) + \ldots + \mu(I_N);$$

2a) (**regularidad.**) Si I es un intervalo, entonces, para cada $\varepsilon > 0$, existen dos intervalos H y J, tales que $\overline{H} \subset I \subset J^\circ$ y además,

$$\mu(H) > \mu(I) - \varepsilon \quad \text{y} \quad \mu(J) < \mu(I) + \varepsilon.$$

Adviértase que la adherencia de un intervalo, o bien su interior, no son generalmente intervalos en el sentido que usamos en esta sección. En cambio, sí lo es el conjunto vacío; además, de la aditividad y de la descomposición $\emptyset = \emptyset \cup \emptyset$, se deduce que $\mu(\emptyset) = 0$.

Una función de conjunto μ definida sobre la clase de los intervalos y que posea las propiedades de aditividad y regularidad, se llama una **medida de Lebesgue-Stieljes**.

Ejemplo (medidas de Lebesgue-Stieljes en $\mathbb{R}^1$).

Consideremos una función monótona creciente y continua por la derecha $F : \mathbb{R} \to \mathbb{R}$, de modo que en cada punto $x \in \mathbb{R}$, tenemos $F(x+) = F(x)$. Una función con estas propiedades se llama una **función de distribución**.

Para cada intervalo $I = (a, b]$, definamos

$$\mu(I) = \mu_F(I) = F(b) - F(a).$$

Dejaremos a cargo del lector, la fácil verificación de que la función μ es aditiva.

Por otra parte, puesto que F es continua por la derecha, dado $\varepsilon > 0$, existen dos números a' y b', tales que

$$a < a' < b, \qquad b' > b,$$

$$F(a') < F(a) + \varepsilon \quad \text{y} \quad F(b') < F(b) + \varepsilon,$$

de donde se deduce que si ponemos $H = (a', b]$ y $J = (a, b']$, entonces tendremos $\overline{H} \subset I \subset J^\circ$ y además,

$$\mu(H) = F(b) - F(a') > F(b) - F(a) - \varepsilon = \mu(I) - \varepsilon,$$

$$\mu(J) = F(b') - F(a) < F(b) - F(a) - \varepsilon = \mu(I) + \varepsilon,$$

lo cual muestra que μ es regular y por consiguiente, μ es una medida de Lebesgue-Stieljes sobre los intervalos de la recta.

Si en el ejemplo precedente se elige la función de distribución $F(x) = x$, entonces se obtiene $\mu(I) = F(b) - F(a) = b - a$, que es la medida de Lebesgue del intervalo I.

Supongamos ahora que μ es una medida de Lebesgue-Stieljes sobre los intervalos de $\mathbb{R}^n$. Si el conjunto elemental A es la unión de los intervalos disjuntos $I_1, I_2, \ldots, I_N$, pongamos

$$\mu(A) = \mu(I_1) + \mu(I_2) + \ldots + \mu(I_N).$$

De la definición resulta que si los conjuntos elementales $A_1, A_2, \ldots, A_N$ son disjuntos, entonces

$$\mu(A_1 \cup A_2 \cup \ldots \cup A_N) = \mu(A_1) + \mu(A_2) + \ldots + \mu(A_N),$$

de donde se sigue que si A y B son dos conjuntos elementales cualesquiera, entonces

$$\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B);$$

en particular, $\mu(A \cup B) \leq \mu(A) + \mu(B)$. Por inducción sobre N, se demuestra que si $A_1, A_2, \ldots, A_N$ es una sucesión finita de conjuntos elementales, entonces

$$\mu(A_1 \cup A_2 \cup \ldots \cup A_N) \leq \mu(A_1) + \mu(A_2) + \ldots + \mu(A_N).$$

Si el conjunto elemental A está contenido en el conjunto elemental B, entonces $\mu(A) \leq \mu(B)$, pues $\mu(B) = \mu(A) + \mu(B - A)$.

La regularidad es esencial para probar el siguiente resultado, análogo al teorema (3.8): Si el intervalo I está contenido en la unión de una sucesión de intervalos (I_k), entonces $\mu(I) \leq \sum_{k=1}^{\infty} \mu(I_k)$. En efecto, dado $\varepsilon > 0$, existe un intervalo H, cuya adherencia está contenida en I, tal que $\mu(H) > \mu(I) - \varepsilon$. Además, para cada entero positivo k, existe un intervalo J_k, cuyo interior contiene a I_k, tal que $\mu(J_k) < \mu(I_k) + \varepsilon/2^k$. Puesto que la adherencia de H es un conjunto compacto contenido en la unión de los conjuntos abiertos J_k° $(k = 1, 2, 3, \ldots)$, se deduce que existe un entero positivo N, tal que $\overline{H} \subset J_1^\circ \cup J_2^\circ \cup \ldots \cup J_N^\circ$, de donde $H \subset J_1 \cup J_2 \cup \ldots \cup J_N$, y el resto de la demostración es completamente análoga a la del teorema (3.8).

De lo demostrado se deduce como en (3.9) que si el intervalo I es la unión de los intervalos disjuntos I_k ($k = 1, 2, 3, \ldots$), entonces $\mu(I) = \sum_{k=1}^{\infty} \mu(I_k)$.

Diremos que un conjunto U es σ**-elemental**, si existe una sucesión de intervalos disjuntos (I_k) cuya unión es U; y en tal caso, definimos la **medida** de U por medio de la fórmula

$$\mu(U) = \sum_{k=1}^{\infty} \mu(I_k).$$

Para cada subconjunto E del espacio $\mathbb{R}^n$, definimos la **medida exterior** de E como el número de la recta extendida

$$\mu_e(E) = \inf\{\mu(U) : U \supset E\},$$

y diremos que el conjunto E es μ**-medible**, si para cada $\varepsilon > 0$, existe un conjunto σ-elemental U que contiene a E, tal que $\mu_e(U - E) < \varepsilon$. En el caso de que E sea medible, escribimos $\mu(E) = \mu_e(E)$ y llamamos a este número la **medida** de E.

La clase $\mathcal{M}_\mu$ formada por todos los conjuntos μ-medibles es una σ-álgebra que contiene a la clase de los conjuntos σ-elementales y entre estos, a todos los conjuntos abiertos, de donde se sigue que $\mathcal{M}_\mu$ contiene a la σ-álgebra de Borel de $\mathbb{R}^n$. Además, μ es no negativa y σ-aditiva sobre la clase de los conjuntos μ-medibles; luego, también lo es sobre la σ-álgebra de Borel.

Los detalles que hemos omitido siguen exactamente los mismos pasos que dimos para construir la medida de Lebesgue y por esta razón los dejaremos como un excelente ejercicio de repaso a cargo del lector.

Resumiendo todo lo dicho, podemos enunciar el siguiente teorema:

(3.42) **Teorema.** *Toda medida de Lebesgue-Stieljes μ se extiende a una función de conjunto no negativa y σ-aditiva sobre una σ-álgebra $\mathcal{M}_\mu$ que contiene a la σ-álgebra de Borel de $\mathbb{R}^n$.*

En particular, toda medida de Lebesgue-Stieljes se extiende a la σ-álgebra de Borel, preservando la propiedad de ser no negativa y σ-aditiva.

También se prueba que esta extensión es única: si μ' es otra función de conjunto no negativa y σ-aditiva sobre la σ-álgebra de Borel, tal que $\mu'(I) = \mu(I)$ para cada intervalo I, entonces $\mu'(E) = \mu(E)$ para cada conjunto boreliano E.

EJERCICIOS

1. Mostrar que cualquier conjunto con medida exterior positiva contiene un conjunto acotado con medida exterior positiva.

2. Probar que si $m_e(E) > 0$ y $0 < \alpha < 1$, entonces existe un intervalo I, tal que $m_e(E \cap I) > \alpha m(I)$. Sugerencia: suponiendo primero que $0 < m_e(E) < \infty$, considerar un conjunto σ-elemental $U = \bigcup_{k=1}^{\infty} I_k$, tal que $U \supset E$ y $m(U) < \alpha^{-1} m_e(E)$; entonces al menos uno de los intervalos I_k debe satisfacer la desigualdad del enunciado.

3. Probar que si E es un subconjunto medible de la recta que verifica $m(E) > 0$, entonces el conjunto $D(E)$ formado por todas las diferencias entre elementos de E incluye un entorno del origen. Sugerencia: considere un intervalo $I = (a, b)$, tal que $m(E \cap I) > (3/4)m(I)$ y sea $\delta = (1/4)m(I)$. Si $|x| < \delta$, el conjunto $F = E \cap I$ y su transladado $F + x$ tienen al menos un punto en común, en vista de que ambos están incluidos en $(a - \delta, b + \delta)$.

4. El disco cerrado $x^2 + y^2 \leq 1$ no es un conjunto σ-elemental en el plano $\mathbb{R}^2$.

5. La intersección de una sucesión de conjuntos σ-elementales y la diferencia entre dos conjuntos σ-elementales pueden no ser σ- elementales. Sugerencia: considere el conjunto ternario de Cantor.

6. (Condición de Caratheodory). Decimos que un conjunto $E \subset I\!R^n$ **satisface la condición de Caratheodory** si para cualquier conjunto $S \subset I\!R^n$,
$$m_e(S \cap E) + m_e(S - E) = m_e(S).$$
(a) Probar que todo conjunto medible E satisface la condición de Caratheodory. Sugerencia: es suficiente probar que para cualquier conjunto S se cumple $m_e(S \cap E) + m_e(S - E) \leq m_e(S)$, en vista de que la desigualdad opuesta se cumple en cualquier circunstancia, por la subaditividad de la medida exterior.
(b) Probar que si E_1 y E_2 satisfacen la condición de Caratheodory, entonces la intersección de ambos conjuntos también la satisface. Sugerencia: el conjunto $S - E_1 \cap E_2$ es la unión de los conjuntos disjuntos $S \cap E_1 - E_2$, $(S - E_1) \cap E_2$ y $(S - E_1) - E_2$.
(c) Probar que si E satisface la condición de Caratheodory, entonces E es medible. Sugerencia: en virtud de (a) y (b) se puede suponer que E es acotado.
La moraleja del problema es que **los conjuntos medibles son exactamente los que satisfacen la condición de Caratheodory**.

7. Para cualquier conjunto $E \subset I\!R^n$, existe un conjunto H de clase G_δ, tal que $E \subset H$ y $m_e(E) = m(H)$.

8. Si el conjunto E es la unión de una sucesión creciente de conjuntos $E_1 \subset E_2 \subset E_3 \subset \ldots$, entonces
$$m_e(E) = \lim_{k\to\infty} m_e(E_k).$$
Sugerencia: para cada k elíjase un conjunto H_k de clase G_δ, tal que $E_k \subset H_k$, $m(H_k) = m_e(E_k)$; los conjuntos de clase G_δ
$$D_k = H_k \cap H_{k+1} \cap H_{k+2} \cap ... \qquad (k = 1, 2, 3, ...)$$
forman una sucesión creciente que verifica $E_k \subset D_k$, $m(D_k) = m_e(E_k)$ y el conjunto E está incluido en la unión de los D_k, de donde se sigue que $m_e(E) \leq \lim_{k\to\infty} m_e(E_k)$. La desigualdad opuesta es inmediata.

9. Probar que para cualquier conjunto $E \subset \mathbb{R}^n$ existe un conjunto H de clase G_δ, tal que $E \subset H$ y para cualquier conjunto medible M se cumple

$$m_e(E \cap M) = m(H \cap M)$$

Sugerencia: suponiendo primero que $m_e(E) < \infty$, considérese un conjunto H de clase G_δ como en le problema 7. El conjunto M satisface la condición de Caratheodory con respecto a E.

10. Probar que si E_1 y E_2 son conjuntos medibles disjuntos, entonces para cualquier conjunto $S \subset \mathbb{R}^n$ se cumple

$$m_e(S \cap E_1) + m_e(S \cap E_2) = m_e(S \cap (E_1 \cup E_2)).$$

Generalizar a cualquier sucesión (E_k) de conjuntos medibles disjuntos.

11. Probar la equivalencia de las afirmaciones siguientes:
(a) el conjunto E es medible;
(b) para cada $\varepsilon > 0$ existe un conjunto cerrado F tal que $F \subset E$, $m_e(E - F) < \varepsilon$;
(c) $E = H \cup Z$, donde H es de clase F_σ y Z de medida nula.

12. Mostrar que existe un conjunto H incluido en el intervalo unitario [0,1], de clase F_σ, de medida uno, formado exclusivamente por puntos irracionales. Mostrar que H es unión numerable de conjuntos cerrados con interior vacío.

13. Sea (E_k) una sucesión de conjuntos medibles. Probar que
(a) $m(\liminf E_k) \le \liminf m(E_k)$;
(b) si para algún j, $m(\bigcup_{k \ge j} E_k) < \infty$, entonces $\limsup m(E_k) \le m(\limsup E_k)$;
(c) si la sucesión (E_k) tiende a un límite y todos los E_k son subconjuntos de un conjunto fijo A de medida finita, entonces $m(\lim E_k)$.
Exhibir una sucesión de conjuntos E_k en el intervalo unitario $[0,1]$, tal que $m(\liminf E_k) < \liminf m(E_k) < \limsup m(E_k) < m(\limsup E_k)$. Todas las desigualdades estrictas.

14. Probar que para cualquier conjunto medible E vale la fórmula

$$m(E) = \sup\{m(K) : \ K \subset E\},$$

donde el supremo se toma sobre todos los conjuntos compactos K incluidos en E.

15. ¿Cómo son los subconjuntos medibles del conjunto de Vitali ?

16. Mostrar que existen 2^c conjuntos medibles, donde c denota la potencia del continuo. Sugerencia: considere todos los subconjuntos de un conjunto de medida nula con la potencia del continuo. Muestre que la clase de los conjuntos no medibles también tiene cardinal 2^c.

17. Exhibir una sucesión decreciente de conjuntos acotados $E_1 \supset E_2 \supset E_3 \supset \ldots$, cuya intersección sea vacía y que sin embargo verifique

$$\lim_{k\to\infty} m_e(E_k) > 0.$$

Sugerencia: denotando por V el conjunto de Vitali en el intervalo $(0,1)$, considere la sucesión $E_k = \bigcup_{j\geq k} V_j$, donde $V_j = V + 1/j$.

18. Probar que si A y B son subconjuntos borelianos de la recta, entonces $A \times B$ es un conjunto boreliano en el plano $\mathbb{R}^2$.

19. Probar que cualquier conjunto con medida positiva tiene la potencia del continuo.

Nota: Para indicar que los conjuntos E_k $(k = 1, 2, 3, ...)$ forman una sucesión creciente cuya unión es E (límite de la sucesión), se usa la notación $E_k \nearrow E$. Con esta notación, el problema 8 del presente capítulo se enuncia brevemente así :

$$\text{si } E_k \nearrow E, \text{ entonces } m_e(E_k) \nearrow m_e(E).$$

Análogamente, la notación $E_k \searrow E$ se usa para indicar que (E_k) es una sucesión decreciente cuya intersección es $E = \lim E_k$.

CAPITULO IV

FUNCIONES MEDIBLES

1. El concepto de función medible.

En este capítulo vamos a considerar funciones f definidas sobre el espacio $I\!R^n$, con valores en la recta extendida $\overline{I\!R}$, de modo que $-\infty$ y $+\infty$ son posibles valores de f.

Para cada número real a, indicaremos por $\{f > a\}$ (léase: "el conjunto donde f es mayor que a") al conjunto formado por todos los puntos del espacio donde el valor de f es mayor que a. Es decir,

$$\{f > a\} = \{x \in I\!R^n : f(x) > a\} = f^{-1}((a, \infty]).$$

Análogamente se definen los conjuntos

$$\{f \geq a\}, \quad \{f < a\}, \quad \{f \leq a\}, \tag{1}$$

que corresponden, respectivamente, a las desigualdades $f(x) \geq a$, $f(x) < a$ y $f(x) \leq a$. El símbolo $\{f = a\}$ indica el conjunto formado por todos los puntos donde f toma el valor a.

Diremos que f es una función **medible** si para cada número real a, el conjunto donde f es mayor que a es un subconjunto medible del espacio $I\!R^n$; es decir, si para cada número real a, se verifica

$$\{f > a\} \in \mathcal{M}.$$

Si f es medible, entonces los tres conjuntos (1) son medibles cualquiera que sea el número a, en virtud de las fórmulas

$$\{f \geq a\} = \bigcap_{k=1}^{\infty} \{f > a - 1/k\},$$

$$\{f < a\} = I\!\!R^n - \{f \geq a\}, \qquad \{f \leq a\} = I\!\!R^n - \{f > a\}.$$

Para que f sea medible es suficiente que cualquiera de los conjuntos (1) sea medible para todo número a, en virtud de una consideración análoga que dejaremos a cargo del lector.

Si f es medible y $J = (a, b]$ un intervalo de la recta, entonces $f^{-1}(J) = \{f > a\} \cap \{f \leq b\}$ es medible por ser la intersección de dos conjuntos medibles. En general, si f es medible y J un intervalo cualquiera de $I\!\!R$, entonces $f^{-1}(J)$ es un subconjunto medible de $I\!\!R^n$.

Puesto que $I\!\!R$ es la unión de una sucesión de intervalos (J_k), se sigue que si f es medible, entonces también lo es $f^{-1}(I\!\!R)$, por ser la unión de la sucesión de conjuntos medibles $f^{-1}(J_k)$. El siguiente teorema amplía estas consideraciones:

(4.1) **Teorema.** *Si f es medible, entonces para cada conjunto boreliano $H \subset I\!\!R$, el conjunto $f^{-1}(H)$ es medible.*

Para demostrarlo, observemos que la clase de conjuntos

$$\mathcal{S} = \{S : S \subset I\!\!R, f^{-1}(S) \text{ es medible }\}$$

es una σ-álgebra de subconjuntos de $I\!\!R$, en virtud de las fórmulas

$$f^{-1}\left(\bigcup_{k=1}^{\infty} S_k\right) = \bigcup_{k=1}^{\infty} f^{-1}(S_k), \qquad f^{-1}(I\!\!R - S) = f^{-1}(I\!\!R) - f^{-1}(S).$$

Puesto que según vimos anteriormente, $\mathcal{S}$ contiene a la clase $\mathcal{I}_1$ formada por todos los intervalos de la recta, se sigue que la σ-álgebra de Borel $\mathcal{B}_1 = \sigma(\mathcal{I}_1)$ está contenida en $\mathcal{S}$.

Q.E.D.

Los conjuntos de la forma $H \cup A$, donde H es un subconjunto boreliano de $I\!\!R$ y A un subconjunto del conjunto de dos elementos $\{-\infty, +\infty\} = \overline{I\!\!R} - I\!\!R$, se llaman **conjuntos borelianos de la recta extendida**.

(4.2) *La función* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *es medible si y sólo si para cada conjunto boreliano* M *de la recta extendida, el conjunto* $f^{-1}(M)$ *es medible.*

En efecto, si f es medible, los conjuntos

$$\{f = +\infty\} = \bigcap_{k=1}^{\infty} \{f > k\}, \qquad \{f = -\infty\} = \bigcap_{k=1}^{\infty} \{f < -k\}$$

son ambos medibles y por consiguiente, si $M = H \cup A$ es un conjunto boreliano de la recta extendida, el conjunto $f^{-1}(M) = f^{-1}(H) \cup f^{-1}(A)$ es medible.

Recíprocamente, si para todo conjunto boreliano M de la recta extendida, $f^{-1}(M)$ es un conjunto medible, tomando $M = (a, -\infty]$ se deduce que $\{f > a\}$ es medible para cualquier número real a, lo cual prueba que f es medible.

La condición expresada en la última proposición puede tomarse como definición del concepto de función medible. Si se procede en esa forma, el requerimiento de que $\{f > a\}$ sea medible para cada número real a se convierte en un criterio de medibilidad.

Más generalmente, si Σ es una σ-álgebra de subconjuntos de $\mathbb{R}^n$, diremos que la función $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ es **medible con respecto a** Σ, si para cada número real a, se verifica $\{f > a\} \in \Sigma$.

Imitando paso a paso las etapas anteriores, el lector no tendrá dificultad en probar por sí mismo la siguiente proposición:

(4.3) *La función* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *es medible con respecto a* Σ *si y sólo si para cada conjunto boreliano* M *de la recta extendida,* $f^{-1}(M)$ *es un miembro de* Σ.

Las funciones medibles con respecto a la σ-álgebra $\mathcal{M}$ formada por los conjuntos medibles del espacio $\mathbb{R}^n$, son las que llamamos **funciones medibles**, a secas.

Las funciones medibles con respecto a la σ-álgebra de Borel $\mathcal{B} = \mathcal{B}_n$ se llaman **funciones medibles Borel**, o bien **funciones borelianas**.

Puesto que $\mathcal{B} \subset \mathcal{M}$, concluimos que toda función boreliana es medible.

Si f es semicontinua inferiormente, $\{f > a\}$ es un conjunto abierto y por consiguiente boreliano. Luego, toda función semicontinua inferiormente es una función boreliana. Análogamente se prueba que toda función semicontinua superiormente es boreliana. En particular, toda función continua es una función boreliana.

En lo que resta de esta sección nos proponemos analizar la siguiente cuestión: supongamos que $f : I\!R^n \to \overline{I\!R}$ es medible y que $g : \overline{I\!R} \to \overline{I\!R}$ es una función de la recta extendida en sí misma. ¿Bajo qué hipótesis podremos afirmar que la función compuesta $h = g \circ f$ es medible ?.

Para responder a esta pregunta comenzaremos introduciendo la siguiente definición: diremos que una función g de $\overline{I\!R}$ en sí mismo es una **función boreliana** si para cada conjunto boreliano M de la recta extendida, la imagen inversa $g^{-1}(M)$ es un conjunto de la misma clase.

Para que una función g de $\overline{I\!R}$ en sí mismo sea una función boreliana, es suficiente que la restricción $g_\circ$ de g a $I\!R$, definida por $g_\circ(x) = g(x)$ para cada número real x, sea una función medible con respecto a la σ-álgebra de Borel de la recta. Esto sigue inmediatamente de la fórmula

$$g^{-1}(M) = g_\circ^{-1}(M) \cup [g^{-1}(M) - I\!R].$$

Como veremos en los ejemplos, esta observación provee un criterio de fácil aplicación.

(4.4) *Si $f : I\!R^n \to \overline{I\!R}$ es medible con respecto a Σ y $g : \overline{I\!R} \to \overline{I\!R}$ es una función boreliana, entonces la función compuesta $h = g \circ f$ es medible con respecto a Σ.*

La demostración resulta inmediatamente de la fórmula $h^{-1}(M) = f^{-1}(g^{-1}(M))$ y de la proposición (4.3).

Ejemplos.

Si f es medible con respecto a Σ, también lo son las funciones $|f|$, f^2 (cuadrado de f), e^f, $|f|^p$ (p un número real positivo), $\log|f|$, etc.

Con respecto al último ejemplo, aclaremos que por $\log|t|$ entendemos la función g de $\overline{I\!R}$ en sí mismo, definida de la siguiente manera:

$$g(t) = \begin{cases} \log|t| & \text{si } t \in I\!R \text{ y } t \neq 0, \\ -\infty & \text{si } t = 0, \\ +\infty & \text{si } t = +\infty \text{ o } t = -\infty. \end{cases}$$

Esta función es boreliana, pues su restricción a $I\!R$ es semicontinua en ambos sentidos, como se ve enseguida analizando lo que ocurre en el origen.

Observaciones complementarias.

1) Si f es una función constante, para cada subconjunto M de $\overline{I\!R}$ la imagen inversa $f^{-1}(M)$ es o bien vacía, o bien igual a $I\!R^n$. Por consiguiente, toda función constante es medible con respecto a cualquier σ-álgebra Σ, pues $I\!R^n$ y $\emptyset$ son miembros de Σ.
2) Si f es medible con respecto a Σ y c es un número real, el conjunto $\{f = c\} = f^{-1}(\{c\})$ es un miembro de Σ, pues el conjunto unitario $\{c\}$ es un conjunto cerrado y por consiguiente, boreliano.
3) Si f y g son medibles con respecto a Σ, entonces el conjunto $\{f < g\}$ formado por todos los puntos x tales que $f(x) < g(x)$ es un miembro de Σ, pues si Q es el conjunto de los números racionales, podemos probar muy fácilmente que

$$\{f < g\} = \bigcup_{r \in Q} (\{f < r\} \cap \{g > r\}).$$

En efecto, si $f(x) < g(x)$, entonces existe un número racional r, tal que $f(x) < r$ y $g(x) > r$.

2. Operaciones algebraicas.

En esta sección demostraremos que al realizar operaciones algebraicas (suma, producto y cociente) entre funciones medibles con respecto a una σ-álgebra Σ, resultan siempre funciones medibles con respecto a la misma σ-álgebra. Sin embargo, las demostraciones se harán suponiendo que las funciones consideradas son finitas, y se completarán al final de la siguiente sección (recordemos que una función se llama finita cuando todos sus valores son elementos de $I\!R$).

(4.5) *Si f y g son funciones medibles con respecto a Σ y c es una constante real, entonces $f+g$, cf y fg son medibles con respecto a Σ.*

Demostración. Denotando por Q el conjunto de los números racionales, se tiene

$$\{f+g>a\} = \bigcup_{r\in Q} (\{f>r\}\cap\{g>a-r\}).$$

En efecto, si en un punto x se verifica $f(x)+g(x)>a$, entonces $f(x)>a-g(x)$ y existe un número racional r, tal que $f(x)>r>a-g(x)$; es decir, $f(x)>r$ y $g(x)>a-r$. Recíprocamente, si para algún r, $f(x)>r$ y $g(x)>a-r$, entonces $f(x)+g(x)>a$. Puesto que f y g son medibles con respecto a Σ, para cualquier número racional r, $\{f>r\}$ y $\{g>a-r\}$ son miembros de Σ, de donde se infiere que también lo es $\{f+g>a\}$, pues Q es numerable.

Si $c>0$, $\{cf>a\}=\{f>ac^{-1}\}$; si $c<0$, $\{cf>a\}=\{f<ac^{-1}\}$ y finalmente, $cf=0$ si $c=0$. Luego, cf es medible con respecto a Σ.

Finalmente, la fórmula

$$fg = \frac{1}{4}\{(f+g)^2-(f-g)^2\},$$

en vista de que el cuadrado de cada función medible con respecto a Σ es medible con respecto a la misma σ-álgebra, prueba que fg es medible con respecto a Σ.

Para estudiar el cociente entre funciones medibles, denotemos por $1/f$ la función que toma el valor $1/f(x)$ si $f(x)\neq 0$ y el valor cero si $f(x)=0$. Además, si para cada número real a, ponemos

$$E_a = (\{f>0\}\cap\{af<1\})\cup(\{f<0\}\cap\{af>1\}),$$

se verifica fácilmente que

$$\{1/f>a\} = \begin{cases} E_a & \text{si } a\geq 0,\\ E_a\cup\{f=0\} & \text{si } a<0,\end{cases}$$

lo cual muestra que si f es medible con respecto a Σ, entonces también lo es $1/f$.

(4.6) *Si f y g son medibles con respecto a Σ, entonces también lo es el cociente $f/g = f\times(1/g)$.*

3. Sucesiones de funciones medibles.

En esta sección veremos que la clase formada por todas las funciones medibles con respecto a una σ-álgebra Σ es cerrada bajo límites puntuales. El paso previo es la siguiente proposición:

(4.7) *Si* (f_k) *es una sucesión de funciones medibles con respecto a* Σ, *entonces las funciones*

$$g(x) = \inf_k f_k(x), \qquad h(x) = \sup_k f_k(x)$$

son también medibles con respecto a Σ.

La demostración se deduce de las fórmulas

$$\{h > a\} = \bigcup_{k=1}^{\infty} \{f_k > a\}, \qquad \{g < a\} = \bigcup_{k=1}^{\infty} \{f_k < a\},$$

y del hecho de que para cada índice k, $\{f_k > a\}$ y $\{f_k < a\}$ son miembros de Σ.

(4.8) *Si* (f_k) *es una sucesión de funciones medibles con respecto a* Σ, *entonces las funciones*

$$g(x) = \liminf_{k\to\infty} f_k(x) \quad y \quad h(x) = \limsup_{k\to\infty} f_k(x)$$

son también medibles con respecto a Σ.

La demostración resulta de aplicar (4.7) en las relaciones

$$g(x) = \sup_j \inf_{k\geq j} f_k(x), \qquad h(x) = \inf_j \sup_{k\geq j} f_k(x).$$

En particular, si la sucesión (f_k) converge puntualmente a una función f, es decir, si en cada punto x de $I\!R^n$ se verifica

$$f(x) = \lim_{k\to\infty} f_k(x),$$

entonces f es medible con respecto a Σ.

Las funciones g y h de la proposición (4.8) suelen designarse en la forma abreviada

$$g = \liminf_{k\to\infty} f_k, \qquad h = \limsup_{k\to\infty} f_k.$$

Análogamente, si (f_k) converge puntualmente al límite f, escribimos $f = \lim_{k\to\infty} f_k$.

Completemos ahora la demostración de (4.5) que se había realizado bajo la hipótesis adicional de que f y g sólo tomaban valores finitos. Para librarnos de esa restricción, consideremos la sucesión de funciones (φ_k) de $\overline{I\!R}$ en sí mismo, definidas por

$$\varphi_k(t) = \begin{cases} t & \text{si } |t| \le k \\ k & \text{si } t > k \\ -k & \text{si } t < -k. \end{cases}$$

Cada una de ellas es una función boreliana, pues la restricción de φ_k a $I\!R$ es continua. Además, para cada $t \in \overline{I\!R}$, $\varphi_k \to t$ cuando $k \to \infty$. Por consiguiente, las funciones

$$f_k = \varphi \circ f \quad \text{ y } \quad g_k = \varphi_k \circ g$$

son medibles con respecto a Σ, finitas, y convergen puntualmente a f y a g, respectivamente, cuando k tiende a infinito. Luego las funciones

$$f + g = \lim_{k\to\infty}(f_k + g_k) \quad \text{ y } \quad fg = \lim_{k\to\infty} f_k g_k$$

son medibles en virtud de (4.8).

4. Funciones simples.

La **función característica** χ_E de un subconjunto E de $I\!R^n$ se define por medio de la fórmula

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

Es fácil demostrar que χ_E es medible con respecto a Σ si y sólo si $E \in \Sigma$.

Una función medible y finita $\varphi : I\!R^n \to I\!R$ se llama una **función simple** si el conjunto de todos sus valores es finito; es decir, si φ es medible y la imagen $\varphi(I\!R^n)$ es un subconjunto finito de $I\!R$.

De la definición se sigue que si φ y ψ son funciones simples y c un número real, entonces $\varphi + \psi$, $c\varphi$ y $\varphi\psi$ son también funciones simples.

La función característica de un conjunto medible es el ejemplo más sencillo de lo que entendemos por una función simple.

Si $\{\alpha_1, \ldots, \alpha_N\}$ es el conjunto formado por los distintos valores que toma la función simple φ, los conjuntos

$$E_i = \{\varphi = \alpha_i\} = \varphi^{-1}(\{\alpha_i\}) \qquad (1 \le i \le N)$$

son medibles y forman una partición de $I\!R^n$. Además, se compueba fácilmente que

$$\varphi = \sum_{i=1}^{N} \alpha_i \chi_{E_i}.$$

Luego, toda función simple es una combinación lineal de funciones características de conjuntos medibles y recíprocamente, toda función de esta forma es una función simple.

Por ejemplo, la función característica de los números racionales (función de Dirichlet) es una función simple sobre $I\!R$.

La notación $f \le g$ se usará para indicar que en cada punto x de $I\!R^n$ se verifica $f(x) \le g(x)$. En particular, las funciones f que verifican $f \ge 0$ se llaman **no negativas**.

Las funciones simples desempeñan un papel muy importante en la teoría de la integración que desarrollaremos en el próximo capítulo, en virtud del siguiente teorema.

(4.9) **Teorema.** *Si $f : I\!R^n \to \overline{I\!R}$ es una función medible no negativa, entonces existe una sucesión (φ_k) de funciones simples no negativas, tal que*

$$\varphi_1 \le \varphi_2 \le \varphi_3 \le \ldots \qquad f(x) = \lim_{k\to\infty} \varphi_k(x)$$

en cada punto x de $I\!R^n$.

En otras palabras, toda función medible no negativa es el límite puntual de una sucesión creciente de funciones simples no negativas.

DEMOSTRACIÓN. Para cada entero positivo k, dividamos el intervalo $[0,k)$ en $k2^k$ intervalos diádicos disjuntos:

$$\left[\frac{i-1}{2^k}, \frac{i}{2^k}\right) \qquad (i = 1, 2, \dots, k2^k)$$

y definamos la función g_k de $\overline{I\!R}$ en sí mismo por medio de la fórmula

$$g_k(t) = \begin{cases} (i-1)/2^k & \text{si } 0 \le t < k, \quad (i-1)/2^k \le t < i/2^k \\ k & \text{si } t \ge k \\ 0 & \text{si } t < 0. \end{cases}$$

Las funciones g_k son borelianas, no negativas, y verifican

$$0 \le g_1 \le g_2 \le g_3 \le \dots, \qquad \lim_{k\to\infty} g_k(t) = t \quad (0 \le t \le +\infty).$$

La relación $g_k(t) \le g_{k+1}(t)$ es obvia si t es negativo; y si $t \ge 0$ se realiza distinguiendo tres casos:

1°) $0 \le t < k$. Entonces $g_k(t) = (i-1)/2^k$, donde

$$(i-1)/2^k \le t < i/2^k \quad \text{o sea,} \quad (2i-2)/2^{k+1} \le t < 2i/2^{k+1},$$

y para determinar el valor $g_{k+1}(t)$ hay que considerar dos casos:

$$t < (2i-1)/2^{k+1}, \qquad (2i-1)/2^{k+1} \le t.$$

En el primero, $g_{k+1}(t) = (2i-2)/2^{k+1} = (i-1)/2^k = g_k(t)$; en el segundo $g_{k+1}(t) = (2i-1)/2^{k+1} > (2i-2)/2^{k+1} = g_k(t)$.

2°) $k \le t < k+1$. En este caso, $g_k(t) = k$ y $g_{k+1}(t) = (j-1)/2^{k+1}$, donde

$$(j-1)/2^{k+1} \le t < j/2^{k+1};$$

pero entonces, $j > t2^{k+1}$, de donde $j - 1 \ge k2^{k+1}$ y por consiguiente, $g_{k+1}(t) \ge k = g_k(t)$.

3°) $t \ge k+1$. Entonces, $g_{k+1}(t) = k+1 > k = g_k(t)$.

Por último, si $0 \le t < +\infty$, entonces $0 \le t < k$ para todo k suficientemente grande y por lo tanto, $g_k(t) = (i-1)/2^k$, donde $(i-1)/2^k \le t < i/2^k$, de donde $0 \le t - g_k(t) < 1/2^k$ para todo k suficientemente grande. Puesto que $g_k(+\infty) = k$ para todo k, se sigue que $g_k(+\infty) \to +\infty$ cuando k

tiende a infinito y nuestras afirmaciones acerca de la sucesión (g_k) quedan completamente demostradas.

Teniendo en cuenta que g_k toma exactamente $k2^k+1$ valores distintos, resulta que si $f : I\!R^n \to \overline{I\!R}$ es medible y no negativa, las funciones

$$\varphi_k = g_k \circ f \qquad\qquad (k = 1, 2, 3, \ldots)$$

son simples y verifican las afirmaciones del teorema.

Q.E.D.

Observaciones.

1) Si f es medible con respecto a una σ-álgebra $\Sigma \subset \mathcal{M}$, las funciones φ_k del teorema percedente son también medibles con respecto a Σ. En particular, si f es medible con respecto a la σ-álgebra de Borel, también lo son las funciones de la sucesión (φ_k).
2) Multiplicando cada función φ_k por la función característica de la bola $B(0,k)$ resulta una nueva sucesión de funciones simples (ψ_k) que verifica todas las afirmaciones del teorema (4.9) y además, cada ψ_k es nula fuera del conjunto acotado $B(0,k)$.
3) Revisando la última parte de la demostración precedente, se ve que la sucesión $g_k(t)$ converge a la función t uniformemente sobre cada intervalo finito $0 \le t \le M$, de donde se sigue que si la función no negativa f es acotada, entonces φ_k converge a f uniformemente.

5. Partes positiva y negativa.

Si f es una función medible con respecto a Σ, también lo son las funciones no negativas

$$f^+ = \sup(0, f) \quad \text{y} \quad f^- = \sup(0, -f)$$

llamadas, respectivamente, la **parte positiva** y la **parte negativa** de f. La primera toma en cada punto x el máximo entre los dos valores cero y $f(x)$; la segunda, el máximo entre los valores cero y $-f(x)$.

El hecho de que f^+ y f^- sean medibles con respecto a Σ puede verse como un caso particular de (4.7).

Analizando por separado cada una de las dos alternativas $f(x) \geq 0$, $f(x) < 0$, se verifica fácilmente que en cada punto x, $f(x) = f^+(x) - f^-(x)$ y $|f(x)| = f^+(x) + f^-(x)$; es decir,

$$f = f^+ - f^-, \qquad |f| = f^+ + f^-.$$

Entre todas las descomposiciones de f como diferencia de dos funciones no negativas, la descomposición que hemos analizado se distingue por la siguiente propiedad de minimalidad:

(4.10) *Si* f_1 *y* f_2 *son dos funciones no negativas, tales que* $f = f_1 - f_2$, *entonces* $f^+ \leq f_1$ *y* $f^- \leq f_2$.

En efecto, $f \leq f_1$ y $f \geq -f_2$, de donde $f^+ = \sup(0, f) \leq \sup(0, f_1) = f_1$ y además $f^- = \sup(0, -f) \leq \sup(0, f_2) = f_2$.

Si f es medible, aplicando el teorema (4.9) a las funciones f^+ y f^-, concluimos que existen dos sucesiones crecientes (φ_k) y (ψ_k) de funciones simples no negativas que convergen puntualmente a f^+ y a f^-, respectivamente. Luego, la sucesión de funciones simples $(\varphi_k - \psi_k)$ converge puntualmente a f y además, $|\varphi_k - \psi_k| \leq \varphi_k + \psi_k \leq f^+ + f^- = |f|$.

6. Propiedades verdaderas en casi todo punto.

La σ-álgebra $\mathcal{M}$ formada por todos los conjuntos medibles del espacio $\mathbb{R}^n$ posee la siguiente propiedad, cuyas consecuencias nos proponemos analizar en esta sección: si $m(E) = 0$, entonces todo subconjunto de E es un miembro de $\mathcal{M}$. Recordemos que para una función f, “medible” significa medible con respecto a $\mathcal{M}$.

Si todos los puntos x del espacio $\mathbb{R}^n$ con la posible excepción de un conjunto de medida nula poseen una cierta propiedad $P(x)$, entonces diremos que P es verdadera **en casi todo punto**, o bien que $P(x)$ es verdadera para casi todo x.

Por ejemplo, si f y g son dos funciones definidas sobre el espacio $\mathbb{R}^n$ y si el conjunto de todos los puntos x tales que $f(x) \neq g(x)$ tiene medida nula, entonces decimos que $f = g$ en casi todo punto, o bien que $f(x) = g(x)$ para casi todo x.

(4.11) *Si la función h es igual a cero en casi todo punto, entonces h es medible.*

Por hipótesis, el conjunto $Z = \{x : h(x) \neq 0\}$ es de medida nula. Si $a \geq 0$, el conjunto $\{h > a\}$ es de medida nula por estar contenido en Z; si $a < 0$, entonces $\{h > a\}$ es el complemento del conjunto $\{h \leq a\}$, el cual es medible por estar contenido en Z.

Q.E.D.

(4.12) **Corolario.** *Si f es una función medible y $g = f$ es casi todo punto, entonces g es medible.*

Para demostrarlo, pongamos $Z = \{x : g(x) \neq f(x)\}$, de modo que por la hipótesis, Z es un conjunto de medida nula. La identidad

$$g = f\chi_{CZ} + g\chi_Z,$$

en la que el segundo término del miembro derecho es medible por ser igual a cero en casi todo punto, muestra que g es medible por ser la suma de dos funciones medibles.

De acuerdo con la definición dada al comienzo, diremos que una sucesión de funciones (f_k) **converge en casi todo punto** a la función f, si la relación

$$f(x) = \lim_{k\to\infty} f_k(x)$$

se verifica en todos los puntos x del espacio $\mathbb{R}^n$, con la posible excepción de un conjunto de medida nula.

(4.13) *Si la sucesión de funciones medibles (f_k) converge en casi todo punto a la función f, entonces f es medible.*

En efecto, la función $g(x) = \liminf_{k\to\infty} f_k(x)$ es medible, en virtud de (4.9). Además $f = g$ en casi todo punto.

En la teoría de Lebesgue es conveniente identificar dos funciones f y g cuyos valores coincidan en casi todo punto. Para comprender esta identificación, notemos que la relación "$f = g$ en casi todo punto" es una relación de equivalencia (reflexiva, simétrica y transitiva). Además si f_1 es equivalente a f_2 y g_1 es equivalente a g_2, entonces $f_1 + g_1$ es equivalente a $f_2 + g_2$ y $f_1 g_1$ es equivalente a $f_2 g_2$.

Si $f = g$ en casi todo punto, decimos que f es **esencialmente igual** a g.

7. Convergencia en medida.

Denotaremos por E un subconjunto medible del espacio $I\!R^n$. Si f es una función definida sobre este espacio y δ un número positivo, escribiremos $E(|f| \geq \delta)$ para denotar el conjunto de todos los puntos x de E, tales que $|f(x)| \geq \delta$.

Más generalmente, si para cada punto x del espacio $I\!R^n$, tenemos una proposición, enunciado o afirmación $P(x)$ que puede ser tildada de verdadera o falsa, escribiremos $E(P)$ para denotar el conjunto $E \cap \{x : P(x)\}$.

Vamos a decir que la sucesión de funciones medibles (f_k) **converge en medida** a la función medible f sobre el conjunto E, si para cada número positivo δ, la medida del conjunto $E(|f_k - f| \geq \delta)$ tiende a cero cuando k tiende a infinito.

Para afirmar que (f_k) converge en medida a f sobre E, escribiremos $f_k \overset{\mathrm{m}}{\to} f$.

Una misma sucesión (f_k) no puede converger en medida a dos funciones f y g esencialmente diferentes dentro de E; pues si $f_k \overset{\mathrm{m}}{\to} f$ y $f_k \overset{\mathrm{m}}{\to} g$, tomando medidas en la inclusión

$$E(|f - g| \geq \delta) \subset E(|f_k - f| \leq \delta/2) \cup E(|f_k - g| \geq \delta/2)$$

y haciendo $k \to \infty$, resulta $mE(|f - g| \geq \delta) = 0$ para cada $\delta > 0$. Puesto que

$$E(f \neq g) = \bigcup_{k=1}^{\infty} E(|f - g| \geq 1/k),$$

se sigue que $mE(f \neq g) = 0$; es decir, $f = g$ en casi todo punto de E.

(4.14) *Si la medida de E es finita y si la sucesión de funciones medibles (f_k) converge a la función f en casi todo punto de E, entonces $f_k \overset{\mathrm{m}}{\to} f$.*

Si Z es el conjunto de los puntos de E donde f_k no converge a f, entonces por hipótesis, $mZ = 0$. Luego, dado $\delta > 0$, los conjuntos

$$B_j = \bigcup_{k=j}^{\infty} E(|f_k - f| \geq \delta) \qquad (j = 1, 2, 3, \ldots)$$

forman una sucesión decreciente de conjuntos de medida finita cuya intersección es de medida nula por estar contenida en Z. Por lo tanto, la medida de B_j tiende a cero cuando j tiende a infinito, y como para todo $k \geq j$, $E(|f_k - f| \geq \delta) \subset B_j$, se sigue que

$$\lim_{k\to\infty} mE(|f_k - f| \geq \delta) = 0,$$

lo cual significa que $f_k \overset{\mathrm{m}}{\to} f$.

Q.E.D.

Por otro lado, es relativamente fácil dar ejemplo de una sucesión de funciones que converge en medida sobre un intervalo finito de la recta pero que, sin embargo, no converge ni siquiera en un solo punto de dicho intervalo. A tal efecto, dividamos el intervalo $[0, 1]$ primero en dos, luego en cuatro, después en ocho, y así siguiendo, intervalos de igual longitud y consideremos la sucesión de **intervalos diádicos**:

$$[0, 1/2),\ [1/2, 1],\ [0, 1/4),\ [1/4, 1/2),\ [1/2, 3/4),$$
$$[3/4, 1],\ [0, 1/8),\ [1/8, 1/4),\ \ldots$$

Denotando por f_k la función característica del k-ésimo intervalo de esta sucesión, una sencilla inspección a los gráficos de estas funciones muestra que $f_k \overset{\mathrm{m}}{\to} 0$ sobre $[0, 1]$. Sin embargo $f_k(x)$ no converge en ningún punto x de dicho intervalo.

Por consiguiente, si $m(E) < \infty$, la noción de convergencia en medida sobre E es más amplia (menos restrictiva) que la noción de convergencia en casi todo punto de E.

Diremos que una sucesión de funciones medibles (f_k) es **fundamental en medida** sobre el conjunto E, si para cada $\delta > 0$, la medida del conjunto

$$E(|f_k - f_j| \geq \delta)$$

tiende a cero cuando k y j tienden a infinito; es decir, si para cada $\delta > 0$ y cada $\varepsilon > 0$, existe un número natural N, tal que

$$mE(|f_k - f_j| \geq \delta) < \varepsilon,$$

a condición de elegir ambos números k y j no menores que N.

(4.15) **Teorema.** *Si la sucesión de funciones* (f_k) *es fundamental en medida sobre* E*, entonces existe una subsucesión* (f_{k_i}) *de la sucesión dada que converge a una función finita* f *en casi todo punto de* E*. Además,* $f_k \overset{\mathrm{m}}{\to} f$ *sobre* E*.*

DEMOSTRACIÓN. Por hipótesis, para cada entero positivo i, existe un entero positivo k_i, tal que

$$mE(|f_k - f_j| \geq 1/2^i) < 1/2^i,$$

a condición de elegir ambos índices k y j no menores que k_i. Sin restricción de generalidad, podemos suponer que $k_1 < k_2 < k_3 < \ldots$. Por consiguiente, poniendo

$$E_i = E(|f_{k_{i+1}} - f_{k_i}| \geq 1/2^i),$$

tendremos $mE_i < 1/2^i$.

El conjunto $Z \subset E$, definido por

$$Z = \bigcap_{j=1}^{\infty} \bigcup_{i \geq j} E_i = \limsup E_i \qquad \text{(Cap. I, sec. 2)}$$

es un conjunto de medida nula, pues para cada j, $Z \subset \bigcup_{i \geq j} E_i$, de donde

$$mZ \leq \sum_{i=j}^{\infty} mE_i < \sum_{i=j}^{\infty} \frac{1}{2^i} = \frac{1}{2^{j-1}}.$$

Si $x \in E - Z$, entonces existe un j, tal que para todo $i \geq j$, $x \notin E_i$; es decir,

$$|f_{k_{i+1}}(x) - f_{k_i}(x)| < 1/2^i \qquad (i \geq j).$$

Luego, la serie funcional

(1) $$f_{k_1}(x) + \{f_{k_2}(x) - f_{k_1}(x)\} + \{f_{k_3}(x) - f_{k_2}(x)\} + \ldots$$

es absolutamente convergente (y por consiguiente convergente) fuera de Z.

Llamando $f(x)$ a la suma de la serie (1) si $x \notin Z$ y poniendo $f(x) = 0$ en puntos de Z, es claro que f es una función finita. Además, puesto que la suma de una serie es el límite de sus sumas parciales, para cada $x \in E - Z$, tenemos

$$f(x) = \lim_{i \to \infty} f_{k_i}(x),$$

lo cual pueba que la subsucesión (f_{k_i}) converge a f en casi todo punto de E cuando i tiende a infinito.

Probemos ahora que $f_{k_i} \overset{\mathrm{m}}{\to} f$ cuando $i \to \infty$. En efecto, dado $\delta > 0$, comencemos por tomar j suficientemente grande como para que se cumpla $1/2^{j-1} < \delta$. Puesto que fuera de Z,

$$f = f_{k_j} + (f_{k_{j+1}} - f_{k_j}) + (f_{k_{j+2}} - f_{k_{j+1}}) + \dots,$$

deducimos que

$$E(|f_{k_j} - f| \geq \delta) \subset Z \cup \left(\bigcup_{i \geq j} E_i \right),$$

de donde

$$mE(|f_{k_j} - f| \geq \delta) \leq \sum_{i \geq j} mE_i < 1/2^{j-1} \tag{2}$$

para todo j suficientemente grande, lo cual muestra que $f_{k_i} \overset{\mathrm{m}}{\to} f$.

Finalmente, probemos que $f_k \overset{\mathrm{m}}{\to} f$ cuando $k \to \infty$. De la inclusión

$$E(|f_k - f| \geq \delta) \subset E(|f_k - f_{k_j}| \geq \delta/2) \cup E(|f_{k_j} - f| \geq \delta/2),$$

deducimos

$$mE(|f_k - f| \geq \delta) \leq mE(|f_k - f_{k_j}| \geq \delta/2) + mE(|f_{k_j} - f| \geq \delta/2).$$

Ahora bien; puesto que (f_k) es fundamental en medida, dado $\varepsilon > 0$, existe un número natural N, tal que

$$mE(|f_k - f_{k_j}| \geq \delta/2) < \varepsilon \qquad (k \geq N, j \geq N).$$

Por otra parte, en virtud de (2) existe un número natural M, tal que si $j \geq M$, entonces

$$mE(|f_{k_j} - f| \geq \delta/2) < \varepsilon.$$

Por consiguiente, eligiendo un j mayor que los números N y M, tendremos

$$mE(|f_k - f| \geq \delta) < 2\varepsilon,$$

a condición de elegir $k \geq N$. Hemos probado que $f_k \overset{\mathrm{m}}{\to} f$.

Q.E.D.

8. Función singular de Cantor

El conjunto de Cantor P (Cap. II, sec. 11) se define por medio de la fórmula

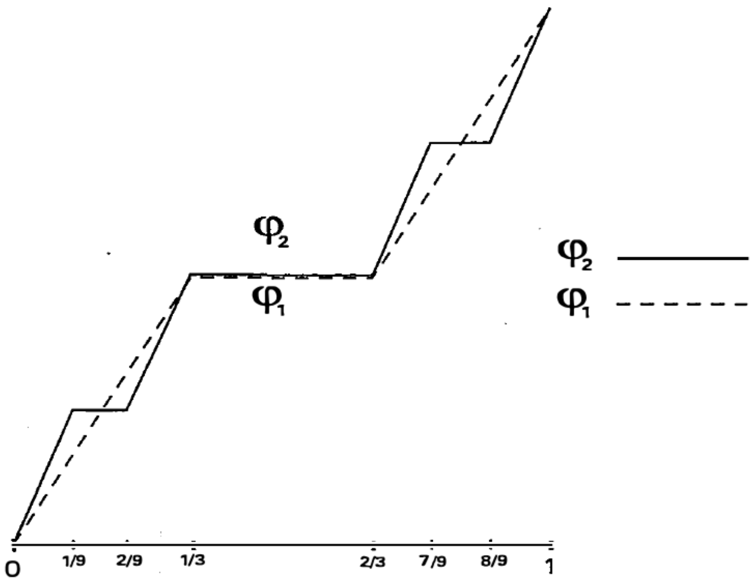
$$P = \bigcap_{n=1}^{\infty} F_n,$$

donde F_n es la unión de 2^n intervalos cerrados disjuntos. El complemento de F_n relativo al intervalo $[0,1]$ es la unión de $2^n - 1$ intervalos abiertos, que numerados de izquierda a derecha podemos expresar en la forma $J_{n,i}$ $(i = 1, 2, ..., 2^n - 1)$. Entre los intervalos de dos etapas sucesivas existe la relación $J_{n,i} = J_{n+1,2i}$.

Así , por ejemplo,

$$J_{1,1} = (1/3, 2/3), \qquad J_{2,1} = (1/9, 2/9),$$
$$J_{2,2} = (1/3, 2/3), \qquad J_{2,3} = (7/9, 8/9).$$

Para cada n, definamos ahora una función continua $\varphi_n(x)$ en el intervalo $0 \leq x \leq 1$, poniendo $\varphi_n(0) = 0$, $\varphi_n(1) = 1$, $\varphi_n(x) = i/2^n$ si $x \in J_{n,i}$, completando la gráfica de φ_n mediante segmentos rectilíneos de que modo que resulte una gráfica poligonal continua. En el diagrama siguiente se muestran las gráficas de φ_1 y φ_2.



Nótese que φ_{n+1} coincide con φ_n en cada unos de los $J_{n,i}$ y además $|\varphi_{n+1} - \varphi_n| < 1/2^{n+1}$ en cualquier punto del intervalo. Luego, la serie de funciones continuas

$$\varphi_1 + (\varphi_2 - \varphi_1) + (\varphi_3 - \varphi_2) + ...,$$

cuyas sumas parciales son las funciones φ_n, converge uniformemente a una función continua φ que se conoce con el nombre de **función singular de Cantor**. Es claro que φ es monótona creciente y su restricción a cualquiera de los intervalos $J_{n,i}$ es constante.

La unión de dichos intervalos es un conjunto abierto G que representa el complemento de P relativo a $[0,1]$. En los ejercicios al final del capítulo se usa la función de Cantor para exhibir algunos ejemplos interesantes. A veces es conveniente extender la función φ a toda la recta, poniendo $\varphi(x) = 0$ si $x < 0$, $\varphi(x) = 1$ si $x > 1$.

EJERCICIOS

1. Probar que si $f(x)$ es una función medible y h un vector de $I\!R^n$, la "función transladada" $f(x+h)$ es también medible.
2. Demuestre que una función f continua en casi todo punto es medible.
3. Sea (f_k) una sucesión de funciones medibles que converge a una función finita f en casi todo punto de un conjunto E de medida finita. Demuestre que para cada $\delta > 0$ existe un conjunto $E_\delta \subset E$ de medida menor que δ, tal que f_k converge a f uniformemente sobre $E - E_\delta$ (teorema de Egorov). Pista: considérese para cada i la sucesión decreciente de conjuntos
$$E^i_m = \bigcup_{k \geq m} E(|f_k - f| \geq 1/i) \qquad (m = 1, 2, 3, \ldots)$$
y elíjase un índice m_i tal que la medida de $E^i_{m_i}$ sea menor que $\delta/2^i$. Llámese E_δ a la unión de estos conjuntos.
4. Exhibir una función no medible f, tal que $|f|$ es medible.
5. Probar que toda función medible f coincide en casi todo punto con una función boreliana. Sugerencia: suponer primero que f es la función característica de un conjunto medible; luego que f es una función simple; suponiendo $f \geq 0$ aplicar el teorema (4.9) y usar el límite superior.
6. Sea E un subconjunto medible de $I\!R^n$ y sea f una función medible no negativa sobre $I\!R^n$. Probar que el conjunto
$$O_E(f) = \{(x,t)\} : \ x \in E, \ t \in I\!R, \ 0 < t < f(x)\}$$
es medible en $I\!R^{n+1}$. Sugerencia: suponer primero que f es simple; usar luego el teorema (4.9).
Mostrar que si (f_k) es una sucesión creciente de funciones medibles no negativas que tiende puntualmente a f, entonces los conjuntos $O_E(f_k)$ forman una sucesión creciente cuya unión es $O_E(f)$.
Nota: algunos autores definen la **integral** de f sobre E como la medida $(n+1)$-dimensional de $O_E(f)$.
7. Siendo E un subconjunto de medida finita de $I\!R^n$, para cada función medible f y cada número real t, pongamos
$$\begin{aligned} \lambda_f(t) &= m(\{x : \ x \in E, \ f(x) > t\}) \\ &= m\ E(f > t). \end{aligned}$$

Probar que (i) λ_f es decreciente y continua por la derecha; (ii) si $f_k \xrightarrow{\text{m}} f$, entonces $\lambda_{f_k}(t) \longrightarrow \lambda_f(t)$ en cada punto t donde λ_f sea continua. Sugerencia: de la inclusión

$$E(f_k > t) \subset E(f > t - \varepsilon) \cup E(|f_k - f| > \varepsilon)$$

se deduce $\limsup_{k\to\infty} \lambda_{f_k}(t) \leq \lambda_f(t-)$ y análogamente se prueba que $\lambda_f(t) \leq \liminf_{k\to\infty} \lambda_{f_k}(t)$.

8. **(Ejemplo de conjunto medible no boreliano).** Considérese la función ψ definida por la fórmula $\psi(x) = x + \varphi(x)$, donde φ es la función singular de Cantor definida en la última sección.
Mostrar que:
   (i) ψ es estrictamente creciente y continua; su inversa ψ^{-1} tiene las mismas propiedades.
   (ii) para cada conjunto boreliano $B \subset R$, $\psi(B)$ y $\psi^{-1}(B)$ son borelianos.
   (iii) $\psi([0,1]) = [0,2]$
   (iv) poniendo $G = [0,1] - P$ (complemento del conjunto ternario relativo a $[0,1]$) se tiene $m(\psi(G)) = m(G) = 1$.
   (v) $m(\psi(P)) = 1$.
   (vi) existe un conjunto no medible $V \subset \psi(P)$.
   (vii) $\psi^{-1}(V)$ es medible Lebesgue pero no es medible Borel.

9. Dar ejemplo de una función medible f y una función boreliana g (ambas de $I\!R$ en sí mismo), tales que $f \circ g$ no sea medible.

10. Dar ejemplo de una función medible $f : I\!R \to I\!R$ y un conjunto medible Lebesgue $E \subset I\!R$, tales que $f^{-1}(E)$ no sea medible.

11. Si $f_k \xrightarrow{\text{m}} f$ y $g_k \xrightarrow{\text{m}} g$, entonces $f_k + g_k \xrightarrow{\text{m}} f + g$; si además todas las funciones son finitas y $m(E) < \infty$, entonces $f_k g_k \xrightarrow{\text{m}} fg$ dentro del conjunto E.

12. Probar que si $f_k \xrightarrow{\text{m}} f$, entonces existe una subsucesión (f_{k_i}) que converge a f en casi todo punto.

13. Probar que una función medible $f : I\!R \to I\!R$ que satisfaga la ecuación funcional $f(x+y) = f(x) + f(y)$ es de la forma $f(x) = cx$ (c una constante). Sugerencia: Probar que f es continua en el origen.

CAPITULO V

INTEGRAL DE LEBESGUE

1. Integral de funciones no negativas.

En este capítulo las palabras "conjunto" y "función" se usan como sinónimos de "conjunto medible" y "función medible", respectivamente.

Si E es un subconjunto del espacio euclidiano $\mathbb{R}^n$ y $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ una función no negativa sobre E, para cada descomposición del conjunto E como unión finita de conjuntos disjuntos $E_1, E_2, \dots, E_N$, calculemos la suma

$$\sum_{i=1}^{N} v_i m(E_i), \tag{1}$$

donde v_i es el ínfimo de los valores que toma la función f sobre el conjunto E_i.

El supremo de tales sumas se llama la **integral** de f sobre E y se denota por cualquiera de los símbolos

$$\int_E f, \qquad \int_E f\,dx, \qquad \int_E f(x)\,dx$$

o bien por

$$\int \dots \int_E f(x_1, \dots, x_n)\,dx_1 \dots dx_n.$$

De manera que

$$\int_E f = \int_E f(x)\,dx = \sup \sum_{i=1}^{N} v_i m(E_i),$$

donde el supremo se toma sobre todas las posibles descomposiciones de E como unión finita de conjuntos (medibles) disjuntos y $v_i = \inf f(E_i)$.

Es conveniente aclarar que si en algún término de la suma (1) uno de los factores fuera cero y el otro $+\infty$, debe usarse la convención $0 \cdot (+\infty) = (+\infty) \cdot 0 = 0$ que hemos introducido en la sección 1 del primer capítulo.

De la definición se deduce que si la función f toma el valor constante $c \geq 0$ sobre el conjunto E, entonces

$$\int_E f = c\, m(E),$$

pues en tal caso, para cualquier descomposición de E, cada uno de los números v_i es igual a c y por consiguiente, cada una de las sumas (1) toma el valor $cm(E_1) + cm(E_2) + \ldots + cm(E_N) = cm(E)$.

(5.1) *Si A y B son dos conjuntos disjuntos cuya unión es E, entonces para cualquier función f no negativa sobre E,*

$$\int_E f = \int_A f + \int_B f.$$

DEMOSTRACIÓN. Para cada descomposición de E como unión finita de conjuntos disjuntos $E_1, \ldots, E_N$, pongamos $A_i = E_i \cap A$, $B_i = E_i \cap B$ $(i = 1, 2, \ldots, N)$. Si $v_i = \inf f(E_i)$, tendremos

$$\begin{aligned}\sum_{i=1}^{N} v_i m(E_i) = \sum_{i=1}^{N} v_i [m(A_i) + m(B_i)] &= \sum_{i=1}^{N} v_i m(A_i) + \sum_{i=1}^{N} v_i m(B_i) \\ &\leq \int_A f + \int_B f,\end{aligned}$$

pues los conjuntos A_i forman una descomposición del conjunto A y $v_i \leq \inf f(A_i)$, y análogamente para el conjunto B. Luego,

$$\int_E f \leq \int_A f + \int_B f. \tag{2}$$

Supongamos ahora que los conjuntos A_1, A_2, $\ldots$, A_N forman una descomposición de A, mientras que los conjuntos $B_1, B_2, \ldots, B_M$ forman una descomposición de B. Si $v_i = \inf f(A_i)$ y $w_j = \inf f(B_j)$, entonces

$$\sum_{i=1}^{N} v_i m(A_i) + \sum_{j=1}^{M} w_j m(B_j) \leq \int_E f,$$

pues los conjuntos $A_1, \ldots, A_N$, $B_1, \ldots, B_M$ forman una descomposición de E.

Tomando el supremo en el miembro izquierdo y teniendo en cuenta que para cualquier par U y V de subconjuntos de R, $\sup(U+V) = \sup U + \sup V$, obtenemos la desigualdad

$$\int_A f + \int_B f \leq \int_E f,$$

que juntamente con (2) nos da la igualdad del enunciado.

Q.E.D.

Es interesante destacar que la demostración precedente no usa el hecho de que f sea medible.

Otra consecuencia inmediata de la definición de integral es que si $0 \leq f \leq g$ en puntos de E, entonces $\int_E f \leq \int_E g$, pues para cada descomposición de E en conjuntos disjuntos E_i se cumple $\inf f(E_i) \leq \inf g(E_i)$.

Nótese también que si f es no negativa sobre E y A es un subconjunto de E, entonces $\int_A f \leq \int_E f$, pues en virtud de (5.1),

$$\int_E f = \int_A f + \int_{E-A} f.$$

Si $m(E) = 0$, cualquier función no negativa tiene integral nula sobre el conjunto E, pues en tal caso, todas las sumas (1) son nulas.

De la última observación se deduce que si las funciones no negativas f y g coinciden en casi todo punto de E, entonces $\int_E f = \int_E g$, pues llamando A al conjunto formado por todos los puntos de E donde el valor de f no coincide con el valor de g, tendremos $m(A) = 0$ y por consiguiente, $\int_E f = \int_{E-A} f = \int_{E-A} g = \int_E g$.

Por último señalemos que para cualquier función no negativa f y cualquier conjunto E se cumple

$$\int_E f = \int_{\mathbb{R}^n} f \chi_E, \tag{3}$$

como se ve enseguida aplicando (5.1) a la función $f\chi_E$, con la descomposición $\mathbb{R}^n = E \cup CE$.

Cuando una integral se extiende a todo el espacio (es decir, cuando $E = \mathbb{R}^n$) convendremos en omitir el conjunto que acompaña al símbolo de la integral. Así, por ejemplo, la igualdad (3) se escribe

$$\int_E f = \int f \chi_E.$$

En el caso de una función $f : I\!R \to \overline{I\!R}$, la integral de f sobre el intervalo (a, b) se escribe

$$\int_a^b f(x)\,dx \quad \text{o bien} \quad \int_a^b f.$$

2. Integral de funciones simples.

Veamos cómo se calcula la integral de una función simple no negativa sobre el conjunto E.

(5.2) *Sean $B_1, \ldots, B_N$ conjuntos disjuntos cuya unión es E, y sean $\beta_1, \ldots, \beta_N$ números no negativos. Si la función f toma el valor constante β_i sobre el conjunto B_i, entonces*

$$\int_E f = \beta_1 m(B_1) + \ldots + \beta_N m(B_N).$$

En efecto, en virtud de (5.1),

$$\int_E f = \int_{B_1} f + \ldots + \int_{B_N} f.$$

Además, puesto que f toma el valor constante β_i sobre el conjunto B_i, la integral de f sobre este conjunto es $\beta_i m(B_i)$, lo cual demuestra la proposición.

(5.3) *Si φ y ψ son dos funciones simples no negativas sobre E y c es un número no negativo, entonces*

$$\int_E (\varphi + \psi) = \int_E \varphi + \int_E \psi \quad \text{y} \quad \int_E c\varphi = c \int_E \varphi.$$

Sean $\alpha_1, \ldots, \alpha_M$ los distintos valores que toma la función φ sobre el conjunto E y sean $\beta_1, \ldots, \beta_N$ los distintos valores que toma ψ sobre el mismo conjunto E. Los conjuntos $A_i = \{x \in E : \varphi(x) = \alpha_i\}$ forman una descomposición de E, al igual que los conjuntos $B_j = \{x \in E : \psi(x) = \beta_j\}$. Luego, los conjuntos $A_i \cap B_j$ ($1 \le i \le M$, $1 \le j \le N$) forman una nueva descomposición de E.

Puesto que $\varphi + \psi = \alpha_i + \beta_j$ sobre $A_i \cap B_j$, en virtud de (5.2),

$$\begin{aligned}\int_E (\varphi + \psi) &= \sum_{i=1}^{M} \sum_{j=1}^{N} (\alpha_i + \beta_j) m(A_i \cap B_j) \\ &= \sum_{i=1}^{M} \alpha_i \sum_{j=1}^{N} m(A_i \cap B_j) + \sum_{j=1}^{N} \beta_j \sum_{i=1}^{M} m(A_i \cap B_j) \\ &= \sum_{i=1}^{M} \alpha_i m(A_i) + \sum_{j=1}^{N} \beta_j m(B_j) \\ &= \int_E \varphi + \int_E \psi.\end{aligned}$$

Puesto que $c\varphi$ toma el valor constante $c\alpha_i$ sobre el conjunto A_i,

$$\int_E c\varphi = \sum_{i=1}^{M} (c\alpha_i) m(A_i) = c \int_E \varphi.$$

(5.4) *Si f es una función no negativa, entonces*

$$\int_E f = \sup_{0 \le \varphi \le f} \int_E \varphi,$$

donde el supremo se toma sobre todas las funciones simples φ, tales que $0 \le \varphi \le f$.

Para demostrarlo, consideremos una descomposición de E en conjuntos disjuntos $E_1, \ldots, E_N$ y sea $v_i = \inf f(E_i)$. Si todos los v_i son finitos, entonces la suma

$$\sum_{i=1}^{N} v_i m(E_i)$$

es la integral de la función simple

$$\varphi = \sum_{i=1}^{N} v_i \chi_{E_i},$$

la cual verifica $0 \leq \varphi \leq f$.

En general, definamos para cada $i = 1, 2, \ldots, N$ la sucesión de números

$$v_{ik} = \min(k, v_i) \qquad (k = 1, 2, \ldots),$$

y pongamos

$$\varphi_k = \sum_{i=1}^{N} v_{ik} \chi_{E_i}.$$

Es claro que $0 \leq \varphi_k \leq f$, $0 \leq v_{i1} \leq v_{i2} \leq v_{i3} \leq \ldots$ y además, $\lim_{k\to\infty} v_{ik} = v_i$, de donde

$$\begin{aligned} \sum_{i=1}^{N} v_i m(E_i) &= \lim_{k\to\infty} \sum_{i=1}^{N} v_{ik} m(E_i) \\ &= \lim_{k\to\infty} \int_E \varphi_k = \sup_k \int_E \varphi_k \\ &\leq \sup_{0\leq\varphi\leq f} \int_E \varphi. \end{aligned}$$

Luego, $\displaystyle\int_E f \leq \sup_{0\leq\varphi\leq f} \int_E \varphi$ y como la desigualdad opuesta es inmediata, la proposición queda demostrada.

(5.5) *Sea $f_1 \leq f_2 \leq f_3 \leq \ldots$ una sucesión creciente de funciones no negativas. Si φ es una función simple no negativa que verifica*

$$\varphi \leq \lim_{k\to\infty} f_k,$$

entonces para cualquier conjunto E,

$$\int_E \varphi \leq \lim_{k\to\infty} \int_E f_k.$$

DEMOSTRACIÓN. Consideremos ordenados en forma creciente $0 \leq \alpha_1 < \alpha_2 < \ldots < \alpha_M$ los distintos valores que toma φ sobre el conjunto E y sea $A_i = \{x \in E : \varphi(x) = \alpha_i\}$.

Sin restringir generalidad podemos suponer que $\alpha_1 > 0$, pues si la proposición estuviera probada con esta restricción y fuera $\alpha_1 = 0$, entonces la restricción se cumple sobre $E - A_1$ y por consiguiente,

$$\int_E \varphi = \int_{E-A_1} \varphi \leq \lim_{k\to\infty} \int_{E-A_1} f_k \leq \lim_{k\to\infty} \int_E f_k.$$

Suponiendo pues que $\alpha_1 > 0$, sea ε un número que verifica $0 < \varepsilon < \alpha_1$, de manera que $\varphi - \varepsilon$ es una función simple que toma el valor positivo $\alpha_i - \varepsilon$ sobre el conjunto A_i. Poniendo

$$E_k = \{x \in E : f_k(x) > \varphi(x) - \varepsilon\},$$

en virtud de la hipótesis, tendremos:

$$E_1 \subset E_2 \subset E_3 \subset \ldots, \qquad E = \bigcup_{k=1}^{\infty} E_k.$$

Luego, $m(E_k) \to m(E)$ cuando $k \to \infty$, en virtud de (3.27).

Para la demostración distinguiremos dos casos:

1°) $m(E) < \infty$. En este caso $m(E - E_k) \to 0$ cuando $k \to \infty$ y además,

$$\int_E f_k \geq \int_{E_k} f_k \geq \int_{E_k} (\varphi - \varepsilon) = \int_E (\varphi - \varepsilon) - \int_{E - E_k} (\varphi - \varepsilon)$$

$$= \int_E \varphi - \varepsilon m(E) - \int_{E - E_k} (\varphi - \varepsilon) \geq \int_E \varphi - \varepsilon m(E) - \alpha_M m(E - E_k).$$

Haciendo $k \to \infty$, obtenemos

$$\lim_{k \to \infty} \int_E f_k \geq \int_E \varphi - \varepsilon m(E)$$

y la desigualdad del enunciado resulta haciendo que ε tienda a cero.

2°) $m(E) = \infty$. En este caso, $m(E_k) \to \infty$ y además

$$\int_E f_k \geq \int_{E_k} f_k \geq \int_{E_k} (\varphi - \varepsilon) \geq (\alpha_1 - \varepsilon) m(E_k).$$

Luego, $\lim_{k \to \infty} \int_E f_k = \infty$ y la desigualdad del enunciado se verifica trivialmente.

Q.E.D.

3. Paso al límite bajo el signo integral.

El siguiente teorema debe considerarse el resultado fundamental de la teoría de la integral.

(5.6) **(teorema de Beppo Levi).** *Sea* $f_1 \leq f_2 \leq f_3 \leq \dots$ *una sucesión creciente de funciones no negativas. Si* f *es el límite puntual de la sucesión* (f_k)*, entonces*

$$\int_E f = \lim_{k\to\infty} \int_E f_k.$$

Demostración. Ante todo, observemos que la existencia del límite puntual

$$f(x) = \lim_{k\to\infty} f_k(x)$$

está garantizada por ser creciente la sucesión de funciones (f_k).

Para cualquier función simple φ que verifique $0 \leq \varphi \leq f$, tendremos

$$\varphi \leq \lim_{k\to\infty} f_k;$$

y en virtud de (5.5),

$$\int_E \varphi \leq \lim_{k\to\infty} \int_E f_k.$$

Luego,

$$\int_E f = \sup_{0\leq\varphi\leq f} \int_E \varphi \leq \lim_{k\to\infty} \int_E f_k.$$

Por otro lado, puesto que $0 \leq f_k \leq f$, deducimos que

$$\lim_{k\to\infty} \int_E f_k \leq \int_E f,$$

lo cual, junto con la relación anterior nos da la igualdad del teorema.

Q.E.D.

Veamos los más importantes corolarios del teorema de Beppo-Levi:

(5.7) *Si* f *y* g *son funciones no negativas y* c *un número real no negativo, entonces*

$$\int_E (f+g) = \int_E f + \int_E g, \qquad \int_E cf = c\int_E f.$$

Demostración. En virtud de (4.9), existen dos sucesiones crecientes (φ_k) y (ψ_k) de funciones simples no negativas que convergen puntualmente a f

y a g, respectivamente. Luego, la sucesión de funciones simples $(\varphi_k + \psi_k)$ converge puntualmente y en forma creciente hacia $f+g$; y por el teorema (5.6) y la proposición (5.3),

$$\int_E (f+g) = \lim_{k\to\infty} \int_E (\varphi_k + \psi_k) = \lim_{k\to\infty} \left\{ \int_E \varphi_k + \int_E \psi_k \right\} = \int_E f + \int_E g.$$

Análogamente

$$\int_E cf = \lim_{k\to\infty} \int_E c\varphi_k = \lim_{k\to\infty} c \int_E \varphi_k = c \int_E f.$$

(5.8) *Si* (f_k) *es una sucesión de funciones no negativas, se tiene*

$$\int_E \left(\sum_{k=1}^{\infty} f_k \right) = \sum_{k=1}^{\infty} \int_E f_k.$$

Demostración. Poniendo

$$s_N = \sum_{k=1}^{N} f_k, \qquad s = \sum_{k=1}^{\infty} f_k = \lim_{N\to\infty} s_N,$$

es claro que $0 \le s_1 \le s_2 \le s_3 \le \dots$ y en virtud de (5.6) y (5.7),

$$\int_E s = \lim_{N\to\infty} \int_E s_N = \lim_{N\to\infty} \sum_{k=1}^{N} \int_E f_k = \sum_{k=1}^{\infty} \int_E f_k.$$

(5.9) **(Lema de Fatou).** *Si* (f_k) *es una sucesión de funciones no negativas,*

$$\int_E (\liminf_{k\to\infty} f_k) \le \liminf_{k\to\infty} \int_E f_k.$$

Demostración. Poniendo

$$g(x) = \liminf_{k\to\infty} f_k(x), \qquad g_k(x) = \inf_{j\ge k} f_j(x),$$

es claro que $0 \le g_1 \le g_2 \le g_3 \le \dots$ y además,

$$g(x) = \sup_k g_k(x) = \lim_{k\to\infty} g_k(x).$$

Puesto que $g_k \leq f_k$, tendremos

$$\begin{aligned}\int_E g &= \lim_{k\to\infty}\int_E g_k = \liminf_{k\to\infty}\int_E g_k\\ &\leq \liminf_{k\to\infty}\int_E f_k,\end{aligned}$$

que es precisamente lo que había que demostrar.

En particular, si la sucesión de funciones no negativas (f_k) converge puntualmente al límite f dentro del conjunto E, entonces $\displaystyle\int_E f \leq \liminf_{k\to\infty}\int_E f_k$.

4. Integral de funciones con valores de distinto signo.

En la sección 5 del capítulo anterior hemos visto que toda función f se puede expresar como diferencia de dos funciones no negativas en la forma $f = f^+ - f^-$. Diremos que f es **integrable** sobre E si las integrales

$$\int_E f^+ \qquad \text{e} \qquad \int_E f^-$$

son ambas finitas y en tal caso definimos la **integral** de f sobre E como el número real

$$\int_E f = \int_E f^+ - \int_E f^-.$$

Igual que antes, cuando $E = I\!R^n$ se conviene en omitir el conjunto que acompaña al símbolo integral; de modo que

$$\int f = \int_{I\!R^n} f$$

siempre que f sea integrable sobre $I\!R^n$.

En el caso de una función no negativa se tiene $f^+ = f$, $f^- = 0$ y el hecho de que f sea integrable sobre E equivale a decir que la integral de f sobre E es finita.

(5.10) *La función f es integrable sobre E si y sólo si $|f|$ es integrable sobre E; y en tal caso,*

$$\left|\int_E f\right| \leq \int_E |f|.$$

Demostración. Puesto que $|f| = f^+ + f^-$, el teorema (5.7) nos da la relación

$$\int_E |f| = \int_E f^+ + \int_E f^-.$$

Luego, $|f|$ es integrable sobre E si f es integrable sobre E. Recíprocamente, si $|f|$ es integrable sobre E, puesto que $f^+ \leq |f|$ y $f^- \leq |f|$, concluimos que f es integrable sobre E y además,

$$\left|\int_E f\right| \leq \int_E f^+ + \int_E f^- = \int_E |f|.$$

(5.11) *Si f_1 y f_2 son dos funciones no negativas e integrables sobre E, tales que $f = f_1 - f_2$, entonces f es integrable sobre E y además,*

$$\int_E f = \int_E f_1 - \int_E f_2.$$

Demostración. En virtud de (4.10),

$$f^+ \leq f_1 \quad \text{y} \quad f^- \leq f_2,$$

de donde se sigue que f es integrable sobre E. Puesto que $f = f^+ - f^- = f_1 - f_2$, tenemos $f^+ + f_2 = f^- + f_1$, de donde

$$\int_E f^+ + \int_E f_2 = \int_E f^- + \int_E f_1$$

y la proposición se obtiene por un simple pasaje de términos (todos finitos).

(5.12) *Si f y g son integrables sobre E y c es un número real, entonces $f+g$ y cf son integrables sobre E y además,*

$$\int_E (f+g) = \int_E f + \int_E g, \qquad \int_E cf = c\int_E f,$$

Demostración. Como $f+g = (f^+ + g^+) - (f^- + g^-)$ es la diferencia de dos funciones no negativas integrables sobre E, en virtud de (5.11) concluimos que $f+g$ es integrable sobre E y además,

$$\begin{aligned}\int_E (f+g) &= \int_E (f^+ + g^+) - \int_E (f^- + g^-)\\ &= \int_E f^+ + \int_E g^+ - \int_E f^- - \int_E g^-\\ &= \int_E f + \int_E g.\end{aligned}$$

Si $c \geq 0$, $cf = cf^+ - cf^-$ es la diferencia entre dos funciones no negativas e integrables sobre E y además,

$$\int_E cf = \int_E cf^+ - \int_E cf^- = c\int_E f^+ - c\int_E f^- = c\int_E f.$$

Si $c < 0$, $cf = (-c)f^- - (-c)f^+$ y aplicando nuevamente (5.11) se deduce que cf es integrable sobre E y además,

$$\int_E cf = \int_E (-c)f^- - \int_E (-c)f^+ = (-c)\int_E f^- - (-c)\int_E f^+ = c\int_E f.$$

Llamando $L(E)$ al conjunto de todas las funciones integrables sobre E, la proposición (5.12) afirma que $L(E)$ es un espacio vectorial sobre $\mathbb{R}$ y que la aplicación de $L(E)$ en $\mathbb{R}$ definida por

$$f \mapsto \int_E f$$

es una aplicación lineal.

5. Convergencia mayorada.

El hecho de que una sucesión de funciones (f_k) integrables sobre E converja a una función límite f en cada punto de E no basta para asegurar la validez de la relación

$$\int_E f = \lim_{k\to\infty} \int_E f_k, \tag{1}$$

ni aun suponiendo que f sea también integrable sobre E. Considérese, por ejemplo, la sucesión de funciones $f_k = k\chi_{(0,1/k)}$ sobre el intervalo lineal $(0,1)$.

Es claro que f_k converge puntualmente a cero y sin embargo,

$$\int_0^1 f_k(x)\,dx = 1 \qquad (k = 1, 2, 3, \ldots).$$

¿Bajo qué condiciones es válida la igualdad (1) ?. Un paso decisivo hacia la respuesta es el siguiente teorema:

(5.13) **(teorema de Fatou-Lebesgue).** *Sea* Φ *una función integrable sobre* E. *Si una sucesión de funciones* (f_k) *verifica*

$$|f_k| \leq \Phi \qquad (k = 1, 2, 3, \ldots)$$

en puntos de E, *entonces las funciones*

$$g = \liminf_{k\to\infty} f_k \quad y \quad h = \limsup_{k\to\infty} f_k$$

son integrables sobre E *y además,*

$$\int_E g \leq \liminf_{k\to\infty} \int_E f_k \leq \limsup_{k\to\infty} \int_E f_k \leq \int_E h.$$

Demostración. Sin pérdida de generalidad podemos suponer que todas las funciones del enunciado son nulas fuera de E (para lograrlo bastaría multiplicar a cada una de ellas por la función característica de E, lo cual no altera la validez de las hipótesis).

Puesto que $-\Phi \leq f_k \leq \Phi$, tendremos $-\Phi \leq g \leq h \leq \Phi$ y en consecuencia

$$|g| \leq \Phi \qquad \text{y} \qquad |h| \leq \Phi,$$

lo cual muestra que g y h son integrables sobre E. Además, las funciones $f_k + \Phi$ y $\Phi - f_k$ son no negativas y verifican

$$\liminf_{k\to\infty}(f_k + \Phi) = \liminf_{k\to\infty} f_k + \Phi = g + \Phi,$$

$$\liminf_{k\to\infty}(\Phi - f_k) = \Phi - \limsup_{k\to\infty} f_k = \Phi - h.$$

En virtud de la linealidad de la integral (proposición 5.12) y del lema (5.9),

$$\begin{aligned}\int_E g + \int_E \Phi = \int_E (g + \Phi) &\leq \liminf_{k\to\infty} \int_E (f_k + \Phi) \\ &= \liminf_{k\to\infty} \int_E f_k + \int_E \Phi,\end{aligned}$$

de donde por cancelación,

$$\int_E g \leq \liminf_{k\to\infty} \int_E f_k.$$

Análogamente,

$$\begin{aligned}\int_E \Phi - \int_E h = \int_E (\Phi - h) &\leq \liminf_{k\to\infty} \int_E (\Phi - f_k) \\ &= \int_E \Phi - \limsup_{k\to\infty} \int_E f_k\end{aligned}$$

y por consiguiente,

$$\limsup_{k\to\infty} \int_E f_k \leq \int_E h,$$

lo cual completa la demostración del teorema.

La función Φ del teorema precedente se llama una **mayorante integrable** de la sucesión (f_k), de donde deriva el nombre del siguiente corolario, que se conoce como **teorema de la convergencia mayorada**.

(5.14) **Corolario.** *Sea (f_k) una sucesión de funciones que converge a una función f en cada punto de E. Si existe una función Φ integrable sobre E, tal que*

$$|f_k| \leq \Phi \qquad (k = 1, 2, 3, \ldots)$$

en puntos de E, entonces f es integrable sobre E y además,

$$\int_E f = \lim_{k\to\infty} \int_E f_k.$$

Demostración. En efecto, con las mismas notaciones del teorema anterior se tiene en este caso $g = h = f$ y las relaciones de aquel teorema nos dan la conclusión del enunciado. Q.E.D.

6. La integral y los conjuntos de medida nula.

Comencemos recordando que si f y g son funciones no negativas que coinciden en casi todo punto de E, entonces

$$\int_E f = \int_E g.$$

En otras palabras, la integral no distingue entre dos funciones que coinciden en casi todo punto o lo que es lo mismo, **los conjuntos de medida nula**

equivalen al conjunto vacío desde el punto de vista de la integral. En particular, sobre un conjunto de medida nula cualquier función resulta integrable y su integral sobre dicho conjunto vale cero, pues si $m(E) = 0$, entonces cualquier función f verifica $f(x) = 0$ en casi todo punto de E, y la integral de la función nula es cero.

Para lo que sigue recordemos que el símbolo $E(g > \lambda)$ representa el conjunto de todos los puntos x de E, tales que $g(x) > \lambda$. En esta sección usaremos la **desigualdad de Chebyshev**, la cual establece que para cualquier número positivo λ, cualquier función f y cualquier conjunto E, se cumple

$$mE(|f| > \lambda) \leq \frac{1}{\lambda} \int_E |f|. \tag{5.15}$$

Es decir, **la medida del conjunto de los puntos de E donde se verifica $|f| > \lambda$ es menor o igual que el recíproco de λ por la integral de $|f|$ sobre el conjunto E.**

Su demostración es muy fácil:

$$\int_E |f| \geq \int_{E(|f|>\lambda)} |f| \geq \lambda \cdot mE(|f| > \lambda).$$

(5.16) *Si la función no negativa f tiene integral nula sobre el conjunto E, entonces f es nula en casi todo punto de E.*

Demostración. Llamando Z al conjunto de puntos de E donde f no es nula, es claro que Z es la unión de los conjuntos

$$Z_k = E(f > 1/k) \qquad (k = 1, 2, 3, \ldots).$$

Por la desigualdad de Chebyshev,

$$mZ_k \leq k \int_E f = 0.$$

Luego, $mZ = 0$, lo cual prueba que f es nula en casi todo punto de E.

(5.17) *Si f es integrable sobre E, entonces f es finita en casi todo punto de E.*

DEMOSTRACIÓN. Poniendo $Z = E(|f| = +\infty)$, $Z_k = E(|f| > k)$ ($k = 1, 2, 3, \dots$), es claro que para todo k, $Z \subset Z_k$ y en virtud de (5.15),

$$mZ \leq mZ_k \leq \frac{1}{k}\int_E |f|.$$

Haciendo $k \to \infty$, resulta $mZ = 0$, que es precisamente lo que queríamos demostrar.

7. Integral de funciones con valores complejos.

Si $f : \mathbb{R}^n \to \mathbb{C}$ es una función con valores complejos definida sobre $\mathbb{R}^n$, pondremos

$$f_1(x) = \text{Re } f(x), \qquad f_2(x) = \text{Im } f(x),$$

de modo que f_1 y f_2 (parte real y parte imaginaria de f) son funciones con valores reales y además, $f = f_1 + if_2$.

Diremos que f es **medible** si lo son f_1 y f_2. Todas las funciones que se consideren en lo sucesivo son funciones medibles.

Se dirá que f es **integrable** sobre E si lo son f_1 y f_2, y en tal caso definimos la **integral** de f sobre E por medio de la fórmula

$$\int_E f = \int_E f_1 + i\int_E f_2.$$

Dejaremos a cargo del lector la demostración de la siguiente proposición (linealidad de la integral):

(5.18) *Si f y g son integrables sobre E y c es un número complejo, entonces $f + g$ y cf son integrables sobre E y además,*

$$\int_E (f+g) = \int_E f + \int_E g, \qquad \int_E cf = c\int_E f.$$

La siguiente propiedad se conoce con el nombre de **integrabilidad absoluta**.

(5.19) *La función f con valores complejos es integrable sobre E si y sólo si lo es $|f|$; y en tal caso,*

$$\left|\int_E f\right| \leq \int_E |f|.$$

DEMOSTRACIÓN. Ya hemos visto que el enunciado (5.19) es válido para funciones con valores reales (proposición (5.10)).

En primer lugar observemos que la función $|f| = (f_1^2+f_2^2)^{1/2}$ es medible.

Puesto que para $k = 1, 2$ se tiene $|f_k| \leq |f| \leq |f_1| + |f_2|$, se sigue que f es integrable sobre E si y sólo si lo es $|f|$.

Suponiendo que f sea integrable sobre E, pongamos

$$\int_E f = re^{i\theta} \qquad \text{(forma polar de un número complejo).}$$

La función $g = e^{-i\theta} f$ cuyas partes real e imaginaria llamaremos g_1 y g_2, verifica

$$\int_E g = \int_E g_1 + i\int_E g_2 = e^{-i\theta}\int_E f = r \geq 0.$$

Luego, la integral de g_2 sobre E es nula y en virtud de (5.10),

$$\left|\int_E f\right| = r = \int_E g_1 \leq \int_E |g_1| \leq \int_E |g| = \int_E |f|.$$

El corolario (5.14) vale con idéntico enunciado para funciones con valores complejos, pues si

$$f(x) = \lim_{k\to\infty} f_k(x)$$

en cada punto de E y Φ es una función integrable sobre E, tal que

$$|f_k| \leq \Phi \qquad (k = 1, 2, 3, \ldots),$$

entonces $|f| = \lim_{k\to\infty} |f_k| \leq \Phi$, de donde se sigue que f es integrable sobre E. Además,

$$\left|\int_E f_k - \int_E f\right| = \left|\int_E (f_k - f)\right| \leq \int_E |f_k - f|.$$

Ahora bien; $f_k - f$ tiende a cero en cada punto de E y por otro lado, $|f_k - f| \leq |f_k| + |f| \leq 2\Phi$ en cada punto de E. Puesto que 2Φ es integrable sobre E, se sigue que

$$\int_E |f_k - f| \to 0 \quad \text{cuando} \quad k \to 0$$

y nuestra afirmación queda demostrada.

De los teoremas de paso al límite bajo el signo integral: (5.6), (5.8), (5.9), (5.13) y (5.14), el último es el único que tiene sentido y se mantiene válido para funciones con valores complejos. Agreguemos que todos estos teoremas se mantienen válidos si sus respectivas hipótesis se verifican en casi todo punto de E.

8. Invariancia bajo translaciones.

El teorema (3.37) trae como consecuencia la siguiente propiedad de la integral:

(5.20) *Si f es una función no negativa, entonces para cualquier vector h de $\mathbb{R}^n$ se verifica*

(a) $\displaystyle\int f(x+h)\,dx = \int f(x)\,dx$, *donde las integrales se extienden a todo el espacio $\mathbb{R}^n$. Además, para cada conjunto E, se tiene*

(b) $\displaystyle\int_E f(x+h)\,dx = \int_{E+h} f(x)\,dx.$

Demostración. Comencemos demostrando (a). Si $f = \chi_E$ es la función característica de un conjunto E, entonces $f(x+h) = \chi_E(x+h) = \chi_{E-h}(x)$ y en virtud de (3.37),

$$\int f(x+h)\,dx = m(E-h) = m(E) = \int f(x)\,dx.$$

Luego, (a) es verdadera si f es la función característica de un conjunto.

Si f es una función simple no negativa, entonces

$$f(x) = \sum_{i=1}^{N} \alpha_i f_i(x), \qquad \text{donde } \alpha_i \geq 0, f_i = \chi_{E_i}.$$

Luego, por la linealidad de la integral,

$$\int f(x+h)\,dx = \sum_{i=1}^{N} \alpha_i \int f_i(x)(x+h)\,dx = \sum_{i=1}^{N} \alpha_i \int f_i\,dx = \int f(x)\,dx$$

lo cual prueba que (a) es correcta si f es una función simple no negativa.

Finalmente, si f es una función no negativa, en virtud de (4.9) existe una sucesión creciente de funciones simples no negativas (f_k) que converge puntualmente a f, de modo que $f(x+h) = \lim_{k\to\infty} f_k(x+h)$ en cada punto x, y en virtud del teorema de Beppo Levi (5.6) y el caso antes considerado,

$$\int f(x+h)\,dx = \lim_{k\to\infty} \int f_k(x+h)\,dx = \lim_{k\to\infty} \int f_k(x)\,dx = \int f(x)\,dx$$

y la fórmula (a) queda así demostrada.

La fórmula (b) es una consecuencia casi inmediata de (a). En efecto,

$$\begin{aligned}\int_E f(x+h)\,dx &= \int \chi_E(x) f(x+h)\,dx = \int \chi_{E+h}(x+h) f(x+h)\,dx \\ &= \int \chi_{E+h}(x) f(x)\,dx = \int_{E+h} f(x)\,dx.\end{aligned}$$

El método usado en la demostración de (a) es típico en la teoría de la integral y exhibe una vez más la importancia del teorema de Beppo Levi y del teorema (4.9).

La fórmula (b) es correcta para cualquier función f que sea integrable sobre $E+h$, pues si f es una función con valores reales podemos aplicar dicha fórmula a las funciones no negativas f^+ y f^-, obteniendo

$$\int_E f^+(x+h)\,dx = \int_{E+h} f^+(x)\,dx, \qquad \int_E f^-(x+h)\,dx = \int_{E+h} f^-(x)\,dx.$$

Puesto que $f(x+h) = f^+(x+h) - f^-(x+h)$, no hay más que restar miembro a miembro las igualdades anteriores para obtener (b) en esta nueva situación.

Si f es una función con valores complejos, (b) sigue siendo válida bajo la condición de que f sea integrable sobre $E+h$.

Los últimos argumentos, que reducen la demostración de una fórmula general al caso en que f es no negativa, son también típicos en la teoría de la integral y es conveniente para lo que sigue que el lector los asimile con atención.

9. La integral como función de conjunto.

Si f es una función integrable sobre $I\!R^n$, la función de conjunto

$$\phi(E) = \int_E f \qquad\qquad (E \in \mathcal{M})$$

se llama la **integral indefinida** de f.

Su propiedad más importante es la llamada σ**-aditividad**:

(5.21) *Si el conjunto E es la unión de los conjuntos disjuntos E_k ($k = 1, 2, 3, \dots$), entonces* $\phi(E) = \sum_{k=1}^{\infty} \phi(E_k)$.

Para demostrarla es suficiente considerar el caso en que f es no negativa, aplicando (5.8) en la igualdad $f\chi_E = \sum_{k=1}^{\infty} f\chi_{E_k}$.

La siguiente propiedad de la integral indefinida se llama **continuidad absoluta**.

(5.22) *Si f es integrable sobre $I\!R^n$, entonces para cada $\varepsilon > 0$ existe un número $\delta > 0$, tal que la relación $m(E) < \delta$ implica $|\phi(E)| < \varepsilon$.*

Demostración. Por las mismas razones que antes podemos suponer que f es no negativa.

Poniendo $f_k = \min(k, f)$, es claro que la sucesión (f_k) es creciente y tiende puntualmente a f; y en virtud del teorema de Beppo Levi, $\int f = \lim_{k\to\infty} \int f_k$. Luego, dado $\varepsilon > 0$, existe un índice k, tal que

$$\int (f - f_k) < \varepsilon.$$

Puesto que $f_k \le k$, si $m(E) < \varepsilon$, tendremos

$$\phi(E) = \int_E f = \int_E (f - f_k) + \int_E f_k < \varepsilon + k \cdot m(E)$$

y por consiguiente, $\phi(E) < 2\varepsilon$ si $m(E) < \varepsilon/k$.

Q.E.D.

Resumiendo: si ϕ es la integral indefinida de una función integrable, entonces ϕ es σ-aditiva y absolutamente continua. La recíproca de esta afirmación, llamada el teorema de Radón-Nikodym, es una de las propiedades más profundas de la teoría de la integral, pero de esto nos ocuparemos en otro capítulo.

Digamos finalmente que en este libro, la palabra "integral" se usa siempre para referirse a la **integral en el sentido de Lebesgue**, definida en la primera sección del presente capítulo.

10. Comparación con la integral de Riemann.

Sea f una función acotada definida sobre el intervalo $[a, b]$ de la recta. Para cada partición de dicho intervalo en subintervalos por medio de los puntos $x_0, x_1, \ldots, x_n$, tales que $a = x_0 < x_1 < x_2 < \ldots < x_n = b$, formemos las **sumas de Riemann**:

$$(1) \qquad s = \sum_{i=1}^{n} m_i(x_i - x_{i-1}), \qquad S = \sum_{i=1}^{n} M_i(x_i - x_{i-1}),$$

donde m_i y M_i representan respectivamente el ínfimo y el supremo de los valores de f sobre el intervalo $J_i = [x_{i-1}, x_i]$.

La función f es **integrable Riemann** sobre $[a, b]$ si para cada $\varepsilon > 0$, existe una partición del intervalo, tal que $S - s < \varepsilon$; y en tal caso, el límite común de las sumas (1) cuando $\max(x_i - x_{i-1}) \to 0$ se llama la **integral de Riemann** de f sobre $[a, b]$ y se denotará en esta sección por el símbolo

$$(\mathrm{R}) \int_a^b f(x)\, dx,$$

para distinguirla de la integral en el sentido de Lebesgue de f sobre $[a, b]$, que será consistentemente representada por

$$\int_a^b f(x)\, dx.$$

(5.23) *Si la función acotada f es integrable Riemann sobre $[a, b]$, entonces f es integrable en el sentido de Lebesgue y además,*

$$\int_a^b f(x)\, dx = \ (R) \int_a^b f(x)\, dx.$$

DEMOSTRACIÓN. Para demostrarlo, comencemos observando que las sumas (1) son las integrales en el sentido de Lebesgue de las **"funciones escalonadas"**

$$\varphi(x) = \sum_{i=1}^{n} m_i \chi_{J_i}(x), \qquad \psi(x) = \sum_{i=1}^{n} M_i \chi_{J_i}(x),$$

las cuales verifican las relaciones $\varphi \le f \le \psi$ con excepción de un conjunto finito. Por consiguiente, si f es integrable Riemann sobre $[a,b]$, para cada entero positivo k, existen funciones escalonadas φ_k y ψ_k, tales que $\varphi_k \le f \le \psi_k$ y además

$$\int_a^b (\psi_k - \varphi_k) < 1/k.$$

Las funciones borelianas

$$g(x) = \sup_k \varphi_k(x), \qquad h(x) = \inf_k \psi_k(x)$$

verifican las relaciones $g \le f \le h$ con la posible excepción de un conjunto numerable y además, para cada k, $0 \le h - g \le \psi_k - \varphi_k$, de donde

$$\int_a^b (h-g) \le \int_a^b (\psi_k - \varphi_k) < 1/k.$$

Luego, la función no negativa $h - g$ tiene integral nula sobre $[a,b]$ y por lo tanto, $g = f = h$ en casi todo punto, de donde se sigue que f es medible.

Puesto que para cada índice k,

$$\int_a^b \varphi_k \le \int_a^b f \le \int_a^b \psi_k$$

y en virtud de la definición de integral de Riemann,

$$\int_a^b \varphi_k \le \text{ (R)} \int_a^b f \le \int_a^b \psi_k,$$

se sigue que

$$\left| \int_a^b f - \text{ (R)} \int_a^b f \right| \le \int_a^b \psi_k - \int_a^b \varphi_k < 1/k,$$

y la igualdad del enunciado se obtiene haciendo que k tienda a infinito.

11. Integración parcial; el teorema de Fubini.

En los cursos de Análisis el lector habrá aprendido a reducir el cálculo de una integral múltiple al de las integrales iteradas, por medio de la fórmula

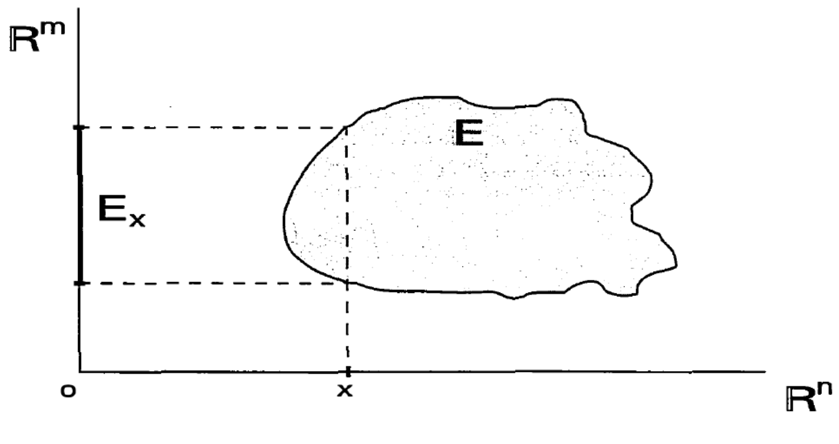
$$\int_c^d \int_a^b f(x,y)\,dx\,dy = \int_a^b dx \int_c^d f(x,y)\,dy.$$

En esta sección estudiaremos la validez de esta fórmula para integrales de Lebesgue, pero para comenzar es necesario dar algunas definiciones.

Cada punto u del espacio euclidiano $I\!R^{n+m}$ será escrito en la forma $u=(x,y)$, donde x es un punto de $I\!R^n$ e y un punto de $I\!R^m$. Si E es un subconjunto del espacio $I\!R^{n+m}$ y x un punto de $I\!R^n$, el conjunto

$$E_x = \{y \in I\!R^m : (x,y) \in E\}$$

se llama la **sección** de E en x. Nótese que $E_x \subset I\!R^m$.



Muy fácilmente se demuestran las fórmulas

$$\left(\bigcup_{k=1}^{\infty} E_k\right)_x = \bigcup_{k=1}^{\infty} (E_k)_x, \qquad \left(\bigcap_{k=1}^{\infty} E_k\right)_x = \bigcap_{k=1}^{\infty} (E_k)_x$$

para cualquier sucesión de conjuntos E_k contenidos en $I\!R^{n+m}$. Además, si $E_1 \subset E_2$, entonces $(E_1)_x \subset (E_2)_x$ y si $E_1 \cap E_2 = \emptyset$, entonces $(E_1)_x \cap (E_2)_x = \emptyset$.

Finalmente, la fórmula

$$(E_1 - E_2)_x = (E_1)_x - (E_2)_x$$

termina de probar que la operación de tomar la sección en un punto x preserva todas las operaciones entre conjuntos, así como las relaciones de inclusión y la disyunción.

Por simplicidad en la notación usaremos el mismo símbolo m para denotar la medida de Lebesgue en cualquier espacio euclidiano; la dimesión correspondiente surgirá claramente del contexto. Téngase presente que el símbolo $\int f$ representa la integral de f sobre todo el espacio del que se trate.

Nuestro primer enunciado corresponde de manera evidente al **principio de Cavalieri** en la geometría del espacio: el volumen de un cuerpo está determinado por las áreas de todas sus secciones planas paralelas a un plano fijo.

(5.23) *Si E es un subconjunto medible del espacio $I\!R^{n+m}$, entonces*

(a) la sección E_x es medible para casi todo x;

(b) la medida de E_x es una función medible de x;

(c) $m(E) = \int m(E_x)\, dx$.

DEMOSTRACIÓN. La demostración se realiza en varias etapas comenzando por los conjuntos de estructura más simple hasta llegar a los conjuntos medibles más generales:

1°) Si E es un intervalo de $I\!R^{n+m}$, entonces existen un intervalo I de $I\!R^n$ y un intervalo J de $I\!R^m$, tales que $E = I \times J$. La sección E_x es igual a J si $x \in I$ y es vacía si $x \notin I$. Por consiguiente, E_x es medible para cada x y además, $m(E_x) = m(J)\chi_I(x)$ es obviamente una función medible de x. Puesto que

$$\int m(E_x)\, dx = m(J)m(I) = m(E),$$

las afirmaciones del enunciado se verifican si E es un intervalo.

2°) Si E es un conjunto abierto, entonces existe una sucesión de intervalos disjuntos (I_k) cuya unión es E y por consiguiente,

$$E_x = \bigcup_{k=1}^{\infty} (I_k)_x$$

es medible para cada x, en virtud del caso anterior. Además,

$$m(E_x) = \sum_{k=1}^{\infty} m((I_k)_x)$$

es una función medible de x por ser la suma de una serie de funciones medibles no negativas, y en virtud de (5.8),

$$\int m(E_x)\, dx = \sum_{k=1}^{\infty} \int m((I_k)_x)\, dx = \sum_{k=1}^{\infty} m(I_k) = m(E)$$

y las afirmaciones del enunciado se verifican si E es abierto.

3°) Supongamos que E es un conjunto acotado de clase G_δ. Entonces existen una bola abierta B de $\mathbb{R}^{n+m}$ y una sucesión de conjuntos abiertos (G_k) de dicho espacio, tales que

$$E \subset B \quad \text{y} \quad E = \bigcap_{k=1}^{\infty} G_k.$$

Puesto que E es la intersección de los conjuntos abiertos $G'_k = B \cap G_1 \cap \ldots \cap G_k$, podemos suponer que se verifica

$$B \supset G_1 \supset G_2 \supset G_3 \supset \ldots .$$

El conjunto $E_x = \bigcap_{k=1}^{\infty} (G_k)_x$ es medible para cada x en virtud del caso anterior y como además, $B_x \supset (G_1)_x \supset (G_2)_x \supset \ldots$, teniendo en cuenta que B_x es acotado, resulta

$$m(E_x) = \lim_{k\to\infty} m((G_k)_x).$$

Luego $m(E_x)$ es una función medible de x por ser el límite puntual de una sucesión de funciones medibles. Por otra parte, la función $m(B_x)$ es integrable, pues por ser B un conjunto abierto, $\int m(B_x)\,dx = m(B) < \infty$. Como además, $m((G_k)_x) \leq m(B_x)$, el teorema de la convergencia mayorada (5.14) nos da

$$\int m(E_x)\,dx = \lim_{k\to\infty} \int m((G_k)_x)\,dx = \lim_{k\to\infty} m(G_k) = m(E)$$

y las afirmaciones del enunciado se verifican si E es un conjunto acotado de clase G_δ.

4°) Si E es un conjunto de clase G_δ, llamando B_k a la bola abierta de $\mathbb{R}^{n+m}$ con centro en el origen y radio k, los conjuntos $E_k = E \cap B_k$ $(k = 1, 2, 3, \ldots)$ forman una sucesión creciente de conjuntos acotados de clase G_δ cuya unión es E. Luego la sección

$$E_x = \bigcup_{k=1}^{\infty} (E_k)_x$$

es medible para cada x en virtud del caso anterior. Además,

$$m(E_x) = \lim_{k\to\infty} m((E_k)_x)$$

es una función medible de x por ser el límite puntual de una sucesión creciente de funciones medibles. Finalmente, el teorema de Beppo-Levi (5.6) nos da la relación

$$\int m(E_x)\,dx = \lim_{k\to\infty} \int m((E_k)_x)\,dx = \lim_{k\to\infty} m(E_k) = m(E)$$

y las afirmaciones (a), (b) y (c) se verifican para cualquier conjunto de clase G_δ.

5°) Si E es un conjunto de medida nula, entonces existe un conjunto H de clase G_δ, tal que $E \subset H$, $m(H) = 0$. Puesto que en virtud de la etapa anterior,

$$\int m(H_x)\,dx = m(H) = 0,$$

de (5.16) se sigue que $m(H_x) = 0$ para casi todo x, y como $E_x \subset H_x$ resulta que E_x es un conjunto de medida nula para casi todo x, luego, medible para casi todo x, como afirma (a).

La función $m(E_x)$ es medible por ser nula en casi todo punto. Además,

$$\int m(E_x)\,dx = 0 = m(E)$$

y las afirmaciones del enunciado se verifican si E es un conjunto de medida nula.

Llegamos por fin al caso más general:

6°) Si E es un conjunto medible, entonces existen un conjunto H de clase G_δ y un conjunto Z de medida nula, tales que $E = H - Z$. Luego, $E_x = H_x - Z_x$ es medible para casi todo x en virtud de lo anterior. Puesto que $m(E_x) = m(H_x)$ para casi todo x, la función $m(E_x)$ es medible en virtud de (4.12) (en los puntos donde E_x no es medible ponemos $m(E_x) = 0$). Finalmente,

$$\int m(E_x)\,dx = \int m(H_x)\,dx = m(H) = m(E)$$

y la proposición (5.23) queda completamente demostrada.

Ejemplos.

(1) Recordemos que un hiperplano H del espacio $I\!R^n$ está formado por todos los puntos x cuyas coordenadas satisfacen una ecuación de la forma $a_1x_1 + a_2x_2 + \ldots + a_nx_n = c$, donde los a_k no son todos nulos. Vamos a probar que $m(H) = 0$.

La demostración es por inducción sobre n: suponiendo que entre los coeficientes a_k $(2 \le k \le n)$ haya al menos uno distinto de cero, si x_1 es un número real, la sección H_{x_1} consta de todas las $(n-1)$-uplas $(x_2, \ldots, x_n)$ tales que $a_2x_2 + \ldots + a_nx_n = c - a_1x_1$; es decir, H_{x_1} es un hiperplano en $I\!R^{n-1}$ y si suponemos que nuestra afirmación es verdadera en este espacio, tendremos $m(H_{x_1}) = 0$ y por consiguiente,

$$m(H) = \int m(H_{x_1})\,dx_1 = 0.$$

Puesto que cada hiperplano de $I\!R^1$ es un conjunto unitario, nuestra afirmación queda demostrada.

(2) Dado un número $a \ge 0$, el **simple** de altura a es el conjunto S formado por todos los puntos de $I\!R^n$ cuyas coordenadas x_i son no negativas y verifican la relación

$$x_1 + x_2 + \ldots + x_n \le a.$$

Por inducción sobre n probaremos que la medida de S es $a^n/n!$. En primer lugar, esto es cierto si $n = 1$. Ahora bien; si $n > 1$, la sección S_{x_1} es vacía si x_1 está fuera del intervalo $0 \leq x_1 \leq a$; y si x_1 pertenece a dicho intervalo, entonces S_{x_1} está formado por todas las $(n-1)$-uplas $(x_2, \ldots, x_n)$ de coordenadas no negativas, tales que $x_2+\ldots+x_n \leq a-x_1$ (el simple de altura $a - x_1$ en $I\!R^{n-1}$).

Si nuestra afirmación es correcta en el espacio $I\!R^{n-1}$, entonces tendremos

$$m(S) = \int_0^a m(S_{x_1})\,dx_1 = \int_0^a \frac{(a-x_1)^{n-1}}{(n-1)!}\,dx_1 = a^n/n!$$

como queríamos demostrar.

El siguiente teorema, debido a los matemáticos G. Fubini y L. Tonelli, comprende a la proposición (5.23) como un caso particular. Recordemos que $u = (x, y)$ denota un punto del espacio $I\!R^{n+m}$.

(5.24) **(teorema de Fubini-Tonelli).** *Si $f(u) = f(x, y)$ es una función medible no negativa sobre $I\!R^{n+m}$, entonces*

(1) $f(x, y)$ es una función medible de y para casi todo x ;

(2) la función $g(x) = \int f(x, y)\,dy$ es medible sobre $I\!R^n$;

(3) $\int g(x)\,dx = \int dx \int f(x, y)\,dy = \int f(u)\,du$.

Antes de dar la demostración, digamos que la última integral se escribe a veces en la forma

$$\iint f(x, y)\,dx\,dy.$$

Demostración. La demostración se hace en tres etapas:

1°) Si $f = \chi_E$ es la función característica de un conjunto medible E del espacio $I\!R^{n+m}$, la igualdad

$$\chi_E(x, y) = \chi_{E_x}(y)$$

permite reducir las tres afirmaciones del enunciado a las afirmaciones (a), (b) y (c) de (5.23). Por consiguiente, (1), (2) y (3) se verifican si f es la función característica de un conjunto.

2°) Si f es una función simple no negativa, entonces

$$f(x,y) = \sum_{k=1}^{N} \alpha_k f_k(x,y),$$

donde $\alpha_k \geq 0$ y $f_k = \chi_{E_k}$ es la función característica de un conjunto medible.

En virtud de la etapa anterior, $f_k(x,y)$ es una función medible de y siempre que x esté fuera de un cierto conjunto Z_k de medida nula. Llamando Z a la unión de dichos conjuntos, es claro que $m(Z) = 0$ y además, $f(x,y)$ es una función medible de y si x no pertenece a Z, lo cual muestra que $f(x,y)$ es una función medible de y para casi todo x.

Para cada x fuera de Z, se cumple

$$g(x) = \int f(x,y)\,dy = \sum_{k=1}^{N} \alpha_k \int f_k(x,y)\,dy,$$

lo cual muestra que g es una función medible, sin importar cómo se la defina en puntos de Z. Por último, la linealidad de la integral, junto con la etapa anterior, nos dan

$$\begin{aligned}\int g(x)\,dx = \sum_{k=1}^{N} \alpha_k \int dx \int f_k(x,y)\,dy \\ = \sum_{k=1}^{N} \alpha_k \int f_k(u)\,du = \int f(u)\,du,\end{aligned}$$

y las afirmaciones del teorema se verifican si f es una función simple.

3°) Si f es una función (medible) no negativa, entonces existe una sucesión creciente de funciones simples no negativas $f_1 \leq f_2 \leq f_3 \leq \dots$, definidas sobre $\mathbb{R}^{n+m}$, tal que en cada punto $u = (x,y)$ de dicho espacio se cumple

$$f(x,y) = \lim_{k\to\infty} f_k(x,y). \tag{4}$$

Puesto que f_k es simple, en virtud de la etapa precedente, $f_k(x,y)$ es una función medible de y, siempre que x esté fuera de un cierto conjunto Z_k de medida nula. LLamando Z a la unión de dichos conjuntos, es claro que $m(Z) = 0$. Además, la relación (4) muestra que si x no pertenece a Z, entonces $f(x,y)$ es una función medible de y. Por otro lado, el teorema de Beppo Levi (5.6) nos da

$$g(x) = \int f(x,y)\,dy = \lim_{k\to\infty} \int f_k(x,y)\,dy \qquad (x \notin Z),$$

lo cual muestra que g es medible, independientemente de cómo se la defina en puntos de Z. Finalmente, una nueva aplicación del teorema de Beppo Levi nos permite afirmar que

$$\int g(x)\,dx = \lim_{k\to\infty} \int dx \int f_k(x,y)\,dy = \lim_{k\to\infty} \int f_k(u)\,du = \int f(u)\,du$$

(nótese que la integral de g sobre $I\!\!R^n$ coincide con la integral sobre el complemento de Z, por ser nula la medida de Z). El teorema queda así completamente demostrado.

El siguiente corolario del teorema anterior se conoce bajo el nombre de teorema de Fubini.

(5.25) **(teorema de Fubini).** *Si* $f(u) = f(x,y)$ *es integrable sobre* $I\!\!R^{n+m}$, *entonces*

(1) para casi todo x, $f(x,y)$ *es una función integrable de* y;

(2) la función $g(x) = \int f(x,y)\,dy$ *es integrable;*

(3) $\int g(x)\,dx = \int dx \int f(x,y)\,dy = \int f(u)\,du$.

Demostración. Si bien el teorema de Fubini es aplicable a funciones con valores complejos, es suficiente considerar el caso en que f toma valores reales.

Considerando las funciones no negativas f^+ y f^- definidas en la sección 5 del capítulo 4, en virtud del teorema anterior, tenemos

$$\int dx \int f^+(x,y)\,dy = \int f^+(u)\,du < \infty,$$

$$\int dx \int f^-(x,y)\, dy = \int f^-(u)\, du < \infty,$$

pues f es integrable.

Las funciones no negativas

$$g_1(x) = \int f^+(x,y)\, dy, \qquad g_2(x) = \int f^-(x,y)\, dy$$

por ser integrables son ambas finitas en casi todo punto, lo cual prueba que $f(x,y)$ es para casi todo x una función integrable de y. La relación

$$g(x) = \int f(x,y)\, dy = g_1(x) - g_2(x)$$

muestra que g es integrable sobre $\mathbb{R}^n$ por ser la diferencia entre dos funciones integrables. Finalmente,

$$\begin{aligned}\int g(x)\, dx = \int g_1(x)\, dx - \int g_2(x)\, dx \\ = \int f^+(u)\, du - \int f^-(u)\, du = \int f(u)\, du\end{aligned}$$

y el teorema queda demostrado.

Observaciones.

La única hipótesis del teorema de Fubini es la integrabilidad de f. Para que esta condición se cumpla, es suficiente que alguna de las integrales

$$\int dx \int |f(x,y)|\, dy, \qquad \int dy \int |f(x,y)|\, dx$$

sea finita, pues por el teorema de Fubini-Tonelli, cada una de ellas es igual a

$$\int |f(u)|\, du.$$

Como caso particular, vamos a probar que si $f(x,y)$ es una función integrable sobre el conjunto E del espacio euclidiano $\mathbb{R}^{n+m}$, la integral de f sobre E se calcula por medio de la fórmula

$$\iint_E f(x,y)\, dx\, dy = \int dx \int_{E_x} f(x,y)\, dy. \tag{5.26}$$

En efecto, $f\chi_E$ es integrable sobre todo el espacio $I\!R^{n+m}$ y por consiguiente,

$$\begin{aligned}\int\int_E f(x,y)\,dx\,dy &= \int\int \chi_E(x,y)f(x,y)\,dx\,dy\\ &= \int dx\int \chi_{E_x}(y)f(x,y)\,dy\\ &= \int dx\int_{E_x} f(x,y)\,dy.\end{aligned}$$

12. La Convolución.

El teorema (5.25) es uno de los más útiles de la teoría de la integral y será usado con frecuencia en los capítulos siguientes. En este párrafo lo usaremos para probar la existencia de la convolución de dos funciones integrables.

Dadas dos funciones f y g medibles sobre $I\!R^n$, definimos la **convolución** de ambas por medio de la fórmula

$$(f*g)(x) = \int f(x-y)g(y)dy$$

en cada punto x de $I\!R^n$ donde la integral exista. En otras palabras, la convolución $f*g$ es una función cuyo valor en cada punto x donde está definida está dado por la fórmula anterior.

De antemano, salvo que se den ciertas condiciones, no hay razones para creer que la función $f(x-y)g(y)$ sea integrable con respecto a y sobre todo el espacio, ni siquiera en algún punto x. Por eso nos proponemos demostrar que si f y g son integrables, la convolución $f*g$ existe en casi todo punto y es integrable sobre $I\!R^n$.

La demostración es muy fácil si aceptamos que la función

$$(x,y)\mapsto f(x-y)g(y) \tag{1}$$

es medible sobre $I\!R^{2n} = I\!R^n \times I\!R^n$.

En tal caso tendremos

$$\begin{aligned}\int_{I\!R^{2n}} |f(x-y)g(y)|\;dxdy &= \int dy\int |f(x-y)g(y)|dx\\ &= \int dy|g(y)|\int |f(x-y)|dx\\ &= \left(\int |f(x)|dx\right)\left(\int |g(x)|dx\right),\end{aligned}$$

lo cual muestra que la función (1) es integrable sobre $\mathbb{R}^{2n}$. Entonces, en virtud del teorema de Fubini, la integral

$$h(x) = \int f(x-y)g(y)dy$$

existe para casi todo punto x y es integrable. Además,

$$\begin{aligned}\int |h(x)|dx &= \int \left| \int f(x-y)g(y)dy \right| dx \\ &\leq \int dx \int |f(x-y)g(y)|dy \\ &= \left(\int |f(x)|dx \right) \left(\int |g(x)|dx \right).\end{aligned}$$

Falta probar que la función (1) es medible sobre $\mathbb{R}^{2n}$. Con este fin definamos $\varphi : \mathbb{R}^{2n} \to \mathbb{R}^n$ por medio de la fórmula $\varphi(x,y) = x - y$, de manera que $f(x-y) = f \circ \varphi(x,y)$.

Si M es un conjunto boreliano de la recta extendida, tendremos

$$(f \circ \varphi)^{-1}(M) = \varphi^{-1}(E),$$

donde $E = f^{-1}(M)$ es un subconjunto medible de $\mathbb{R}^n$, por ser f una función medible. Necesitamos el siguiente lema:

(5.27) *Para cada conjunto medible $E \subset \mathbb{R}^n$, el conjunto $\varphi^{-1}(E)$ es medible en $\mathbb{R}^{2n}$.*

La verdad de la afirmación es inmediata si E es abierto (resp. de clase G_δ), pues en tal caso $\varphi^{-1}(E)$ es abierto (resp. de clase G_δ).

Si E es de medida nula, existe un conjunto H de clase G_δ en $\mathbb{R}^n$, de medida nula, tal que $E \subset H$. Puesto que $\varphi^{-1}(E) \subset \varphi^{-1}(H)$, bastará probar que el segundo conjunto es de medida nula para que $\varphi^{-1}(E)$ resulte medible por ser de medida nula. En efecto, para cada $y \in \mathbb{R}^n$, la sección

$$\begin{aligned}\varphi^{-1}(H)^y &= \{x : (x,y) \in \varphi^{-1}(H)\} \\ &= H + y\end{aligned}$$

es de medida nula, de donde se sigue que $\varphi^{-1}(H)$ es de medida nula, en virtud de (5.23).

En el caso de un conjunto medible cualquiera usamos la representación $E = H - Z$, donde H es de clase G_δ y Z de medida nula, de donde resulta que $\varphi^{-1}(E) = \varphi^{-1}(H) - \varphi^{-1}(Z)$ es medible en virtud de lo anterior.

Q.E.D.

Luego, $f(x-y)$ como función de las variables x e y es medible sobre $\mathbb{R}^{2n}$, como asimismo lo es la función

$$(x, y) \mapsto g(y),$$

en virtud del problema 19 del presente capítulo. De todo lo dicho resulta que la función (1) es medible. Resumiendo, podemos enunciar el siguiente teorema.

(5.28) **Teorema.** *Sean f y g dos funciones integrables sobre $\mathbb{R}^n$, entonces la convolución $f * g$ existe en casi todo punto de $\mathbb{R}^n$, es integrable y verifica la relación*

$$\int |(f * g)(x)| dx \leq \left(\int |f(x)| dx \right) \left(\int |g(x)| dx \right).$$

La convolución de funciones es una operación conmutativa sobre la clase de las funciones integrables, es decir si f y g son dos funciones integrables, las funciones $f * g$ y $g * f$ son dos funciones integrables e iguales en casi todo punto. También dejaremos al lector la demostración de las siguientes propiedades: $f * (g * h) = (f * g) * h$, $f * (g + h) = f * g + f * h$, $\alpha(f * g) = (\alpha f) * g = f * (\alpha g)$.

Los resultados de este párrafo serán usados con frecuencia en el capítulo 8.

EJERCICIOS

1. Mostrar que la función $x^{p-1} e^{-x}$ es integrable sobre $(0, \infty)$ si y sólo si $p > 0$.

2. La función $\operatorname{sen} x / x$ no es integrable sobre $(0, \infty)$, aunque existe el límite

$$\lim_{R \to \infty} \int_0^R \frac{\operatorname{sen} x}{x} \, dx.$$

3. Usar el teorema de la convergencia mayorada (5.14) para probar la fórmula
$$\lim_{n\to\infty} \int_0^1 \frac{n^{3/2}x}{1+n^2x^2}\, dx = 0.$$

4. Probar que si $p > 0$, se cumple
$$\lim_{n\to\infty} \int_0^n x^{p-1}\left(1-\frac{x}{n}\right)^n dx = \int_0^\infty x^{p-1}e^{-x}\, dx.$$

5. Usando integración término a término, probar las fórmulas
a) $\displaystyle\int_0^\infty \frac{x\, dx}{e^x - 1} = \sum_{n=1}^\infty \frac{1}{n^2}$
b) $\displaystyle\int_0^1 \frac{x^p}{1-x}\log\frac{1}{x}\, dx = \sum_{n=1}^\infty \frac{1}{(p+n)^2}$,
siempre que $p > -1$. Aquí $\log$ denota el logaritmo neperiano.

6. Considere la sucesión de funciones $f_n(x) = n\chi_n(x)$ en el intervalo $0 \le x \le 1$, donde χ_n es la función característica del intervalo $(0, 1/n)$. ¿Es posible que exista una función $g(x)$ integrable en dicho intervalo, tal que $f_n(x) \le g(x)$ para cualquier n y cualquier x?

7. Sea f una función medible no negativa sobre $\mathbb{R}^n$ y sea (E_k) una sucesión creciente de conjuntos cuya unión es E.
(a) Probar que $\displaystyle\int_E f = \lim_{k\to\infty}\int_{E_k} f$;
(b) extender a cualquier función f integrable sobre E.

8. En el corolario (5.14) la hipótesis $f_k \to f$ en cada punto de E puede reemplazarse por $f_k \xrightarrow{m} f$.

9. Sean f y f_k $(k = 1, 2, ...)$ funciones medibles no negativas e integrables sobre E. Si $f_k \to f$ en cada punto de E y $\displaystyle\int_E f_k \to \int_E f$, entonces para cualquier conjunto medible $A \subset E$, $\displaystyle\int_A f_k \to \int_A f$.

10. Sea D un disco cerrado del plano complejo y sea f una función con valores complejos, integrable sobre $\mathbb{R}^n$, tal que
$$\frac{1}{m(E)}\int_E f(x)dx \in D$$
para cualquier conjunto E de medida positiva y finita. Probar que $f(x) \in D$ para casi todo x de $\mathbb{R}^n$.

11. Si $\varphi(x)f(x)$ es integrable sobre E para cualquier función f integrable sobre E, entonces existe una constante finita C, tal que $|\varphi(x)| \leq C$ en casi todo punto x de E.

12. Sea f una función medible no negativa sobre $I\!R^1$, tal que

$$\int_a^b f(x)dx > 0$$

siempre que sea $a < b$. ¿Puede concluirse que $f(x) > 0$ en casi todo punto x?

13. Probar que el teorema de la convergencia mayorada (5.14) se extiende a una familia de funciones medibles $f_t(x)$, $a < t < b$, dependiente de un parámetro real t, de la manera siguiente: supongamos que $\tau \in (a, b)$ y que en cada punto de E existe el límite

$$f(x) = \lim_{t \to \tau} f_t(x).$$

Supongamos que existe una función $\Phi(x)$ integrable sobre E, tal que

$$|f_t(x)| \leq \Phi(x) \qquad (x \in E,\ a < t < b).$$

Entonces f es integrable sobre E y además,

$$\int_E f(x)dx = \lim_{t \to \tau} \int_E f_t(x)dx.$$

14. (derivación de una integral paramétrica). Supongamos que la integral

$$\varphi(t) = \int_E f(t, x)dx \qquad (a < t < b)$$

existe para cada $t \in (a, b)$; que $f(t, x)$ es derivable con respecto a t y existe una función $g(x)$ integrable sobre E, tal que

$$\left|\frac{\partial f(t, x)}{\partial t}\right| \leq g(x) \qquad (x \in E,\ a < t < b).$$

Probar que φ es derivable y además,

$$\varphi'(t) = \int_E \frac{\partial f(t, x)}{\partial t}\, dx \qquad (a < t < b).$$

Sugerencia: escribir el cociente $\{\varphi(t+h)-\varphi(t)\}/h$ y usar el teorema del valor medio del cálculo diferencial y el ejercicio precedente.

15. (Transformada de Fourier) Si f es integrable sobre $I\!R^1$, la función

$$g(t)=\int_{-\infty}^{\infty} e^{itx} f(x)dx$$

es acotada y uniformemente continua. Si $x^k f(x)$ es integrable, entonces g es de clase C^k y además,

$$g^{(k)}(t)=\int_{-\infty}^{\infty} e^{itx}(ix)^k f(x)dx.$$

16. (Función gamma). Demostrar que la función

$$\Gamma(p)=\int_0^{\infty} x^{p-1}e^{-x}dx \qquad (p>0)$$

es infinitamente derivable en la semirecta $(0,\infty)$ y además,

$$\Gamma^{(k)}(p)=\int_0^{\infty} x^{p-1}(\log x)^k e^{-x}\ dx.$$

17. (Condición necesaria y suficiente de integrabilidad Riemann). Sea f una función acotada sobre el intervalo $[a,b]$. Con las notaciones de la sección 10, para cada partición π de dicho intervalo

$$\pi: \quad a=x_0<x_1<x_2<...<x_n=b,$$

sean $\varphi=\sum m_i \chi_{J_i}$ y $\psi=\sum M_i \chi_{J_i}$ las funciones escalonadas correspondientes a π. El número $N(\pi)$ igual al máximo de los números x_i-x_{i-1} $(i=1,2,...,n)$ se llama **norma** de la partición π. Probar las afirmaciones siguientes:

(a) si (π_k) es una sucesión de particiones que verifica $N(\pi_k)\to 0$ cuando $k\to\infty$, entonces las correspondientes funciones escalonadas (φ_k) y (ψ_k) verifican

$$\lim_{k\to\infty}\varphi_k(x)=m(x), \qquad \lim_{k\to\infty}\psi_k(x)=M(x),$$

donde $m(x)$ y $M(x)$ son las **envolventes semicontinuas** de f, definidas en el ejercicio 10 del segundo capítulo, siempre que x no sea un punto

de división de alguna de las π_k.
(b) si $s(\pi)$ y $S(\pi)$ son las sumas de Riemann correspondientes a π, se cumplen las relaciones

$$\lim_{N(\pi)\to 0} s(\pi) = \int_a^b m(x)dx, \qquad \lim_{N(\pi)\to 0} S(\pi) = \int_a^b M(x)dx.$$

Sugerencia: aplicar la parte (a) y (5.14).
(c) f es integrable Riemann sobre $[a,b]$ si y sólo si

$$\int_a^b \{M(x) - m(x)\}dx = 0,$$

lo cual equivale a afirmar que f es continua en casi todo punto del intervalo.

18. Sean f y g no negativas e integrables sobre un conjunto E. Si para cada $y \geq 0$ ponemos

$$E_y = E(g > y), \qquad \varphi(y) = \int_{E_y} f(x)dx,$$

entonces

$$\int_E f(x)g(x)dx = \int_0^\infty \varphi(y)dy.$$

19. En este problema E y F son, respectivamente, un subconjunto de $I\!R^n$ y un subconjunto de $I\!R^m$. Probar las siguientes afirmaciones
(a) $|E \times F|_e \leq |E|_e \ |F|_e$
(b) Si E es medible, también lo es $E \times I\!R^m$
(c) Se E y F son medibles, también lo es $E \times F$.

20. Mostrar que la función $f(x,y) = xy/(x^2+y^2)^2$ verifica

$$\int_{-1}^1 dx \int_{-1}^1 f(x,y)dy = \int_{-1}^1 dy \int_{-1}^1 f(x,y)dx,$$

a pesar de que f no es integrable sobre el cuadrado $Q = [-1,1]\times[-1,1]$.

21. Mostrar que si $f(x,y) = (x^2-y^2)/(x^2+y^2)^2$, entonces

$$\int_0^1 dx \int_0^1 f(x,y)dy \neq \int_0^1 dy \int_0^1 f(x,y)dx.$$

22. Una función no negativa f, definida sobre $I\!R^1$, se llama una **densidad** (de probabilidad) si su integral sobre toda la recta es igual a uno. Probar que
(a) La convolución de dos densidades es otra densidad.
(b) Para cada $p > 0$, la función
$$f_p(x) = \begin{cases} \dfrac{1}{\Gamma(p)} x^{p-1} e^{-x} & \text{si } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$$
es una densidad.
(c) Si $p > 0$ y $q > 0$, entonces $f_p * f_q = f_{p+q}$ y además,
$$\int_0^1 (1-t)^{p-1} t^{q-1} dt = \frac{\Gamma(p)\,\Gamma(q)}{\Gamma(p+q)}.$$
La integral biparamétrica del primer miembro se denota por $B(p,q)$ y se llama **función beta**.

23. Si $\alpha > 0$ y $f(x)$ es integrable en el intervalo $[0,b]$, donde $0 < b < \infty$, entonces la integral
$$h(x) = \int_0^x (x-t)^{\alpha-1} f(t) dt$$
existe para casi todo x del intervalo $[0,b]$ y es integrable sobre el mismo. Sugerencia: suponer primero que f es no negativa.

24. Si $\alpha > 0$ y $f(x)$ es integrable sobre $[0,b]$, donde $0 < b < \infty$, la **integral fraccionaria** de orden α de f se define por medio de la fórmula
$$I_\alpha f(x) = \frac{1}{\Gamma(\alpha)} \int_0^x (x-t)^{\alpha-1} f(t) dt$$
que tiene sentido para casi todo x de $[0,b]$. Probar que si $\alpha > 0$ y $\beta > 0$, entonces $I_\beta I_\alpha f(x) = I_{\alpha+\beta} f(x)$ en casi todo punto x del intervalo $[0,b]$. Nótese que $I_1 f(x) = \int_0^x f(t) dt$.

25. Probar que no existe una función integrable u tal que $u * f = f$ para toda función integrable f. Sugerencia: considere funciones del tipo $f = \chi_{(-\delta,\delta)}$ con $\delta > 0$ pequeño.

26. Sea $f = \chi_{(0,1)}$ y definamos inductivamente $f_1 = f$, $f_m = f_{m-1} * f$. Grafique las funciones f_2, f_3 y f_4. Para $\varepsilon > 0$, pongamos $f_\varepsilon(x) = \frac{1}{\varepsilon} \chi_{(0,\varepsilon)}(x)$ y definamos $f_{1,\varepsilon} = f_\varepsilon$, $f_{m,\varepsilon} = f_{m-1,\varepsilon} * f_\varepsilon$. Demuestre que $f_{m,\varepsilon}(x) = \frac{1}{\varepsilon} f_m\left(\frac{x}{\varepsilon}\right)$ y grafique las funciones $f_{1,\varepsilon}$, $f_{2,\varepsilon}$, $f_{3,\varepsilon}$ y $f_{4,\varepsilon}$ para valores pequeños de ε.

CAPITULO VI

CAMBIO DE VARIABLES

1. Imagen de un conjunto medible por una transformación lineal.

En esta sección estudiaremos la forma en que actúa una tranformación lineal T del espacio $\mathbb{R}^n$ sobre un conjunto medible E.

Siendo T una aplicación lineal de $\mathbb{R}^n$ en sí mismo, escribiremos Tx en lugar de $T(x)$ y TE en lugar de $T(E)$. Al hablar de la matriz de T nos referiremos exclusivamente a la matriz de T en la base canónica de $\mathbb{R}^n$. El símbolo $\det T$ indica el determinante de la transformación T.

Si $a = (a_{ij})$ es la matriz de la transformación lineal T e $y = Tx$, la relación entre las coordenadas de los puntos x e y se puede escribir en la forma de un sistema de ecuaciones

$$(1) \qquad y_i = \sum_{j=1}^{n} a_{ij} x_j \qquad (i = 1, 2, \ldots, n)$$

o bien en la forma de un producto matricial

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Toda transformación lineal de $\mathbb{R}^n$ es una aplicación continua. Si T es invertible, entonces T transforma cada conjunto abierto G en otro conjunto abierto, pues $TG = (T^{-1})^{-1}(G)$ y T^{-1} es continua. Luego, T aplica cada conjunto de clase G_δ en otro conjunto de la misma clase; en particular, si I es un intervalo, entonces TI es un conjunto de clase G_δ.

Necesitaremos considerar tres tipos especiales de transformaciones lineales, a las cuales llamaremos **aplicaciones elementales**, a saber: permutar dos coordenadas, multiplicar una coordenada por un número real $\lambda \neq 0$, sumar a una coordenada el producto de otra coordenada por un factor fijo λ. Más explícitamente, consideremos las aplicaciones definidas por medio de las fórmulas

$$\begin{aligned}
T_\alpha(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) &= (x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n),\\
T_\beta(x_1, \ldots, x_i, \ldots, x_n) &= (x_1, \ldots, \lambda x_i, \ldots, x_n), \qquad (\lambda \neq 0),\\
T_\gamma(x_1, \ldots, x_i, \ldots, x_n) &= (x_1, \ldots, x_i + \lambda x_j, \ldots, x_n), \qquad (i \neq j).
\end{aligned}$$

Las matrices que corresponden a estas aplicaciones se llaman **matrices elementales**. Escribiendo estas matrices, se comprueba fácilmente que

$$\det T_\alpha = -1, \quad \det T_\beta = \lambda, \quad \det T_\gamma = 1.$$

De las fórmulas anteriores también se ve que la inversa de cada aplicación elemental es también elemental. En efecto, es evidente que $T_\alpha^{-1} = T_\alpha$ y además,

$$\begin{aligned}
T_\beta^{-1}(x_1, \ldots, x_i, \ldots, x_n) &= (x_1, \ldots, \lambda^{-1} x_i, \ldots, x_n),\\
T_\gamma^{-1}(x_1, \ldots, x_i, \ldots, x_n) &= (x_1, \ldots, x_i - \lambda x_j, \ldots, x_n).
\end{aligned}$$

(6.1) **Teorema.** *Toda aplicación lineal invertible T de $\mathbb{R}^n$ en sí mismo es un producto de aplicaciones elementales $T_1, \ldots, T_k$.*

DEMOSTRACIÓN. Será suficiente esbozar la demostración, que suele estudiarse en los cursos de Algebra Lineal. El teorema equivale a probar que toda matriz no singular $a = (a_{ij})$ es un producto de matrices elementales.

Recordemos que la matriz a se puede transformar en la matriz unitaria

$$1_n = \begin{pmatrix} 1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

al cabo de un número finito de operaciones de fila, a saber:

1°) permutar dos filas,

2°) multiplicar una fila por un coeficiente no nulo,

3°) sumar a una fila un múltiplo de otra fila.

Por otra parte, cualquier operación de fila sobre una matriz a se puede realizar multiplicando a izquierda la matriz a por la correspondiente matriz elemental. Luego, si a es no singular, existen matrices elementales $e_1, \ldots, e_k,$, tales que

$$e_k \ldots e_1 a = 1_n,$$

es decir, $a = (e_k \ldots e_1)^{-1} = e_1^{-1} \ldots e_k^{-1}$ y como la inversa de una matriz elemental es otra matriz elemental, el teorema queda demostrado.

(6.2) **Teorema.** *Si T es una transformación lineal del espacio $I\!R^n$, entonces para cada subconjunto medible E de dicho espacio, la imagen TE es medible y además,* $m(TE) = |\det T|\, mE$.

Demostración. Pongamos $\delta = |\det T|$. Si T es singular (no invertible), entonces $\delta = 0$ y $T\,I\!R^n$ es un subespacio vectorial propio de $I\!R^n$, lo cual implica que el conjunto TE tiene medida nula y el teorema es trivialmente cierto en este caso.

Para estudiar el caso en que T sea no singular, comenzaremos verificando las afirmaciones del teorema en el caso especial en que T es una aplicación elemental y $E = I$ un intervalo de $I\!R^n$. Para fijar ideas supondremos que

$$I = \{x \in I\!R^n : a_i < x_i \leq b_i \ (i = 1, \ldots, n)\}$$

La aplicación T_α transforma I en otro intervalo de la misma medida, mientras que T_β transforma I en otro intervalo cuyos lados coinciden con los de I, con la excepción de uno solo de ellos, que se transforma según una homotecia de razón λ (distinguir los casos $\lambda > 0$ y $\lambda < 0$), lo que hace que la verificación sea muy fácil en el caso de estas aplicaciones elementales.

En cuanto a una aplicación elemental del tipo T_γ consideremos, a modo de ejemplo, la aplicación definida por

$$T(x_1, \ldots, x_n) = (x_1, \ldots, x_{n-1}, x_n + \lambda x_1).$$

La aplicación T y su inversa T^{-1} están dadas por los sistemas de ecuaciones

$$\begin{cases} y_1 & = & x_1 \\ & \dots & \\ y_{n-1} & = & x_{n-1} \\ y_n & = & x_n + \lambda x_1 \end{cases} \qquad \begin{cases} x_1 & = & y_1 \\ & \dots & \\ x_{n-1} & = & y_{n-1} \\ x_n & = & y_n - \lambda y_1, \end{cases}$$

donde $y = Tx$, $x = T^{-1}y$. De dichas ecuaciones y de la definición de I se sigue que TI está formado por todos los puntos $y = (y_1, \dots, y_n)$ que verifican las relaciones

$$a_1 < y_1 \le b_1, \dots, \quad a_{n-1} < y_{n-1} \le b_{n-1}, \quad a_n + \lambda y_1 < y_n \le b_n + \lambda y_1.$$

Por consiguiente, la sección

$$(TI)_{y_1} = \{(y_2, \dots, y_n) : (y_1, y_2, \dots, y_n) \in TI\}$$

es el intervalo de $I\!R^{n-1}$ formado por todos los puntos $(y_2, \dots, y_n)$ que satisfacen

$$a_2 < y_2 \le b_2, \dots, \quad a_{n-1} < y_{n-1} \le b_{n-1}, \quad a_n + \lambda y_1 < y_n \le b_n + \lambda y_1,$$

siempre que $a_1 < y_1 \le b_1$, y es vacía en caso contrario. Ahora bien; en virtud de (5.23),

$$m(TI) = \int_{a_1}^{b_1} m(TI)_{y_1} \, dy_1 = (b_2 - a_2) \dots (b_n - a_n)(b_1 - a_1) = mI.$$

Puesto que en este caso, $\delta = |\det T| = 1$, queda demostrado que si T es una aplicación elemental e I un intervalo de $I\!R^n$, entonces $m(TI) = \delta m I$.

Manteniendo la hipótesis de que T es una aplicación elemental, sea G un conjunto abierto de $I\!R^n$. Entonces existe una sucesión de intervalos disjuntos (I_k) cuya unión es G y por consiguiente,

$$m(TG) = \sum_{k=1}^{\infty} m(TI_k) = \sum_{k=1}^{\infty} \delta m I_k = \delta m G.$$

Por la propiedad multiplicativa del determinante: $\det(T_1T_2) = \det T_1 \cdot \det T_2$, de (6.1) se sigue que si T es invertible y G un conjunto abierto, entonces

$$m(TG) = \delta mG \qquad (\delta \neq 0).$$

Probaremos ahora que para cualquier conjunto E, se cumple

$$m_e(TE) = \delta m_e(E). \tag{2}$$

En efecto, si G es un conjunto abierto que contiene a E, entonces TG es abierto y $TG \supset TE$, de donde $m_e(TE) \leq m(TG) = \delta mG$. Como esto vale para cada conjunto abierto G que contenga a E, en virtud de (3.33) concluimos que

$$m_e(TE) \leq \delta m_e(E)$$

y de esta misma desigualdad, obtenemos

$$m_e(E) = m_e(T^{-1}(TE)) \leq \delta^{-1} m_e(TE),$$

es decir, $\delta m_e(E) \leq m_e(TE)$ que junto con la desigualdad anterior demuestra (2).

Finalmente, si E es medible, entonces para cada $\varepsilon > 0$, existe un conjunto abierto G, tal que $G \supset E$, $m_e(G - E) < \varepsilon$. Luego, $TG \supset TE$ y además,

$$m_e(TG - TE) = m_e(T(G - E)) = \delta m_e(G - E) < \delta\varepsilon,$$

lo cual prueba que TE es medible y la demostración de (6.2) está completa.

Recordemos que una transformación lineal T se llama **ortogonal** si preserva la longitud de los vectores; y que para una tal transformación se cumple $|\det T| = 1$. Luego, una transformación ortogonal transforma cada conjunto medible en otro conjunto de igual medida. En particular, toda rotación tiene esta propiedad.

Si T es una transformación lineal invertible del espacio $\mathbb{R}^n$ y f una función medible sobre dicho espacio, entonces $f \circ T$ es medible, pues para cada conjunto boreliano M de la recta extendida, $(f \circ T)^{-1}(M) = T^{-1} \circ f^{-1}(M)$. Vamos a probar que si f es no negativa, entonces se cumple

$$\int f(x)\,dx = |\det T| \int f(Tx)\,dx. \tag{6.3}$$

La verificación es inmediata si f es la función característica de un conjunto medible E, en virtud de la fórmula

$$\chi_E(Tx) = \chi_{T^{-1}E}(x)$$

y el teorema (6.2); y por la linealidad de la integral se deduce que (6.3) es válida para cualquier función simple no negativa. Finalmente, si f es una función medible no negativa, entonces existe una sucesión creciente (f_k) de funciones simples no negativas que convergen puntualmente a f y el resto de la demostración sigue fácilmente por el teorema de Beppo Levi.

En el resto del presente capítulo se requieren ciertos conocimientos muy elementales sobre el cálculo en varias variables: el concepto de aplicaciones diferenciables, la regla de la cadena y el teorema de la aplicación inversa.

2. Aplicaciones diferenciables.

Recordemos que una aplicación φ de $\mathbb{R}^n$ en sí mismo se llama **diferenciable** en el punto x, si existe una transformación lineal A del espacio $\mathbb{R}^n$, tal que

$$\lim_{h\to 0} \frac{|\varphi(x+h) - \varphi(x) - Ah|}{|h|} = 0. \tag{1}$$

La única aplicación lineal A que verifica (1) se llama **transformación jacobiana** o **derivada** de φ en x, y se denota por $D\varphi(x)$. Si $y = \varphi(x)$, podemos escribir φ en la forma

$$y_i = \varphi_i(x_1, \dots, x_n) \qquad (i = 1, \dots, n).$$

La matriz de la transformación $D\varphi(x)$ es la llamada **matriz jacobiana**:

$$\varphi'(x) = \left(\frac{\partial \varphi_i}{\partial x_k}\right) = \left(\frac{\partial y_i}{\partial x_k}\right);$$

es decir, aquella que en el lugar (i,k) tiene inscripto el valor de la derivada $D_k\varphi_i(x)$, donde D_k denota la derivada parcial con respecto a la k-ésima variable.

El determinante

$$\frac{\partial \varphi}{\partial x} = \frac{\partial(\varphi_1, \dots, \varphi_n)}{\partial(x_1, \dots, x_n)} = \det \varphi'(x)$$

se llama el **determinante jacobiano** de φ en el punto x.

Siendo $U \subset I\!R^n$ un conjunto abierto, diremos que φ es **diferenciable** en U si φ es diferenciable en cada punto de U y decimos que φ es de clase C^1 en U si las derivadas parciales $D_k\varphi_i$ existen y son continuas en U.

Toda función de clase C^1 en el conjunto abierto U es diferenciable en dicho conjunto.

Sean U, V y W conjuntos abiertos de $I\!R^n$ y sea x un punto de U. Si $\varphi : U \to V$ es diferenciable en el punto x, y $\psi : V \to W$ es diferenciable en el punto $\varphi(x)$, entonces la **regla de la cadena** establece que la función compuesta $\psi \circ \varphi$ es diferenciable en x, y además

$$D(\psi \circ \varphi)(x) = D\psi(\varphi(x)) \circ D\varphi(x).$$

Mencionemos por último que si φ es lineal, entonces $D\varphi(x) = \varphi$ en cada punto x, como se comprueba directamente a partir de la definición de derivada.

Llamaremos **norma** del vector $x = (x_1, \ldots, x_n)$ al número

$$\|x\| = \max_{1 \le i \le n} |x_i|, \tag{6.4}$$

y si A es una transformación lineal con matriz (a_{ik}), llamamos **norma** de A al número

$$\|A\| = \max_{1 \le i \le n} \sum_{k=1}^{n} |a_{ik}|. \tag{6.5}$$

Para cualquier par de tranformaciones lineales A y B, se verifican las relaciones

$$\|A + B\| \le \|A\| + \|B\|, \qquad \|AB\| \le \|A\| \cdot \|B\|.$$

De la primera de ellas se deduce fácilmente que $\big|\, \|A\| - \|B\| \,\big| \le \|A - B\|$. Además, llamando e_n a la aplicación idéntica de $I\!R^n$ en sí mismo, se comprueba que $\|e_n\| = 1$.

3. Fórmula del cambio de variables.

En lo que sigue supondremos permanentemente que G y H son dos subconjuntos abiertos de $I\!R^n$ y que $\varphi : G \to H$ es una aplicación biyectiva de clase C^1 con determinante jacobiano distinto de cero en cada punto de G, de modo que la aplicación inversa $\varphi^{-1} : H \to G$ es también de clase C^1, en virtud del teorema de la función inversa (en particular, ambas aplicaciones φ y φ^{-1} son diferenciables en sus respectivos dominios).

Suponiendo que $y = \varphi(x)$, pondremos

$$j(x) = D\varphi(x), \qquad j^{-1}(y) = D\varphi^{-1}(y),$$

y también

$$J(x) = |\det j(x)|, \qquad J^{-1}(y) = |\det j^{-1}(y)|.$$

Aplicando la regla de la cadena en la relación $\varphi^{-1} \circ \varphi = \mathrm{id}_G$, donde id_G es la aplicación idéntica del conjunto G, obtenemos

$$j^{-1}(y) j(x) = e_n = \mathrm{id}_{I\!R^n} \qquad (y = \varphi(x)).$$

y tomando determinantes en esta relación, resulta $J^{-1}(y)J(x) = 1$, siempre que $y = \varphi(x)$.

El resultado central de este capítulo estará basado en el siguiente lema:

(6.6) **Lema.** *Si Q es un cubo de $I\!R^n$ cuya adherencia está contenida en G, entonces*

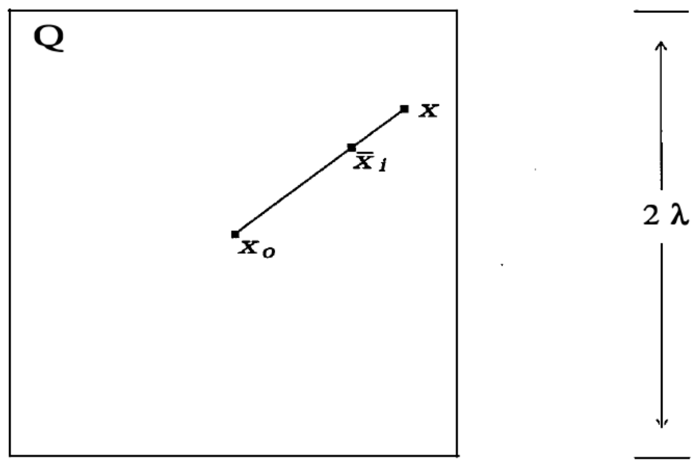
$$m(\varphi(Q)) \leq \int_Q J(x)\, dx.$$

Demostración. Puesto que φ y φ^{-1} son ambas continuas, se deduce que φ aplica cada subconjunto abierto de G en otro conjunto abierto y cada conjunto de clase G_δ contenido en G en otro conjunto de la misma clase. Luego, $\varphi(Q)$ es un conjunto de clase G_δ. Además, puesto que $\varphi(Q) \subset \varphi(\overline{Q})$, podemos suponer que Q es compacto.

Llamando $x_0 = (x_{01}, x_{02}, \ldots, x_{0n})$ al centro de Q, con la ayuda de la norma (6.4) podemos escribir Q en la forma

$$Q = \{x : \|x - x_0\| \leq \lambda\},$$

donde λ representa la mitad de la longitud de cada lado de Q, cuya medida es entonces $mQ = (2\lambda)^n$.



En virtud del teorema del valor medio, para cada x perteneciente a Q tenemos

$$\varphi_i(x) - \varphi_i(x_0) = \sum_{k=1}^{n} D_k\varphi_i(\overline{x}_i)(x_k - x_{0k}),$$

donde $\bar{x}_i = x_0 + \theta_i(x - x_0)$ con $0 < \theta_i < 1$, de modo que $\bar{x}_i \in Q$ por ser Q un conjunto convexo. Luego,

$$\begin{aligned}|\varphi_i(x) - \varphi_i(x_0)| &\leq \sum_{k=1}^{n} |D_k\varphi_i(\bar{x}_i)| \cdot |x_k - x_{0k}| \leq \lambda \sum_{k=1}^{n} |D_k\varphi_i(\bar{x}_i)| \leq \lambda \|j(\bar{x}_i)\| \\ &\leq \lambda \max_{x \in Q} \|j(x)\|,\end{aligned}$$

es decir,

$$\|\varphi(x) - \varphi(x_0)\| \leq \lambda \max_{x \in Q} \|j(x)\|.$$

Esto muestra que $\varphi(Q)$ está contenido en el cubo

$$Q^* = \{y \in I\!R^n : \|y - \varphi(x_0)\| \leq \lambda \max_{x \in Q} \|j(x)\|\},$$

de donde resulta

$$m(\varphi(Q)) \leq mQ^* \leq (2\lambda \max_{x\in Q} \|j(x)\|)^n;$$

es decir,

$$m(\varphi(Q)) \leq (\max_{x\in Q} \|j(x)\|)^n \cdot mQ. \tag{1}$$

Por otro lado, para cualquier transformación lineal invertible A del espacio $\mathbb{R}^n$, la aplicación $\psi = A^{-1} \circ \varphi$ tiene derivada dada por

$$j_\psi(x) = D\psi(x) = A^{-1} \circ j(x) \tag{2}$$

y aplicando la desigualdad (1) con la función ψ en lugar de φ y j_ψ en lugar de j, obtenemos

$$m(A^{-1}\varphi(Q)) \leq (\max_{x\in Q} \|A^{-1} \circ j(x)\|)^n \cdot mQ,$$

y en virtud del teorema (6.2), resulta

$$m(\varphi(Q)) \leq |\det A|(\max_{x\in Q} \|A^{-1}j(x)\|)^n \cdot mQ \tag{3}$$

para cualquier cubo cerrado Q contenido en G y cualquier transformación lineal invertible A del espacio $\mathbb{R}^n$. Pongamos ahora

$$M = \max_{y\in\varphi(Q)} \|j^{-1}(y)\|.$$

Dado $\varepsilon > 0$, existe $\delta > 0$, tal que las relaciones $u \in Q$, $v \in Q$, $\|u - v\| < \delta$ implican $\|j(u) - j(v)\| < \varepsilon$.

Dividamos Q en cubos cerrados no rampantes $Q_1, Q_2, \ldots, Q_N$ (es decir, tales que no exista ningún punto interior a dos de estos cubos) y supongamos que el diámetro de cada uno de estos cubos es menor que un número positivo $\eta \leq \delta$.

Llamando x_k al centro del cubo Q_k, pongamos $y_k = \varphi(x_k)$ $(k = 1, 2, \ldots, N)$ y apliquemos la desigualdad (3) al cubo Q_k con $A = j(x_k)$, de manera que $A^{-1} = j^{-1}(y_k)$ y por lo tanto,

$$m(\varphi(Q_k)) \leq J(x_k)\{\max_{x\in Q_k} \|j^{-1}(y_k)j(x)\|\}^n \cdot mQ_k.$$

Por otra parte, denotando por e_n la aplicación idéntica de $I\!R^n$ en sí mismo, para cada $x \in Q_k$, tenemos

$$\begin{aligned}\|j^{-1}(y_k)j(x)\| - 1 &= \|j^{-1}(y_k)j(x)\| - \|e_n\| \\ &\leq \|j^{-1}(y_k)j(x) - e_n\| = \|j^{-1}(y_k)\{j(x) - j(x_k)\}\| \\ &\leq \|j^{-1}(y_k)\| \cdot \|j(x) - j(x_k)\| < M\varepsilon.\end{aligned}$$

Por consiguiente,

$$\|j^{-1}(y_k)j(x)\| < 1 + M\varepsilon \qquad (x \in Q_k),$$

de donde

$$m(\varphi(Q_k)) \leq (1 + M\varepsilon)^n J(x_k) \cdot mQ_k.$$

Luego,

$$m(\varphi(Q)) \leq \sum_{k=1}^{N} m(\varphi(Q_k)) \leq (1 + M\varepsilon)^n \sum_{k=1}^{N} J(x_k) mQ_k$$

y haciendo que η tienda a cero, obtenemos

$$m(\varphi(Q)) \leq (1 + M\varepsilon)^n \int_Q J(x)\, dx$$

y como ε se eligió arbitrariamente, haciendo $\varepsilon \to 0$ resulta la desigualdad del lema.

(6.7) **Corolario.** *Si E es un subconjunto medible de G, entonces $\varphi(E)$ es medible y además,*

$$m(\varphi(E)) \leq \int_E J(x)\, dx.$$

Demostración. La demostración se realiza en varias etapas:

1°) Si E es abierto, entonces $\varphi(E)$ es abierto. Además, existe una sucesión de cubos disjuntos (Q_k) cuya unión es E, tal que para cada k, $\overline{Q}_k \subset G$; y en virtud de (6.6),

$$m(\varphi(E)) = \sum_{k=1}^{\infty} m(\varphi(Q_k)) \leq \sum_{k=1}^{\infty} \int_{Q_k} J(x)\, dx = \int_E J(x)\, dx.$$

2 °) Si E es un conjunto de clase G_δ acotado y situado a distancia positiva del complemento de G, pongamos $\rho = d(E, CG)$ y sean U y K, respectivamente, los conjuntos definidos por las relaciones

$$d(x,E) < \rho/2, \qquad d(x,E) \leq \rho/2.$$

Es claro que U es abierto, K es compacto y además, $E \subset U \subset K \subset G$. Por otro lado, puesto que E es de clase G_δ, existe una sucesión decreciente de conjuntos abiertos (G_i), tales que

$$U \supset G_1 \supset G_2 \supset \ldots, \qquad E = \bigcap_{i=1}^{\infty} G_i.$$

Por consiguiente,

$$\varphi(K) \supset \varphi(G_1) \supset \varphi(G_2) \supset \ldots, \qquad \varphi(E) = \bigcap_{i=1}^{\infty} \varphi(G_i),$$

y en virtud de la etapa anterior y el teorema de la convergencia mayorada,

$$m(\varphi(E)) = \lim_{i\to\infty} m(\varphi(G_i)) \leq \lim_{i\to\infty} \int_{G_i} J(x)\,dx = \int_E J(x)\,dx.$$

3 °) Si E es un conjunto de clase G_δ, consideremos la sucesión de conjuntos

$$E_k = \{x \in E : |x| < k, \quad d(x, CG) > 1/k\}.$$

Cada E_k es de clase G_δ, acotado, y situado a distancia positiva del complemento de G. Además

$$E_1 \subset E_2 \subset E_3 \subset \ldots, \qquad E = \bigcup_{k=1}^{\infty} E_k,$$

y en virtud de la etapa anterior,

$$m(\varphi(E)) = \lim_{k\to\infty} m(\varphi(E_k)) \leq \lim_{k\to\infty} \int_{E_k} J(x)\,dx = \int_E J(x)\,dx.$$

4 °) Si $mE = 0$, entonces existe un conjunto D de clase G_δ, tal que $E \subset D$ y $mD = 0$, de donde

$$m(\varphi(D)) \leq \int_D J(x)\,dx = 0.$$

Puesto que $\varphi(E) \subset \varphi(D)$, se sigue que $\varphi(E)$ tiene medida nula. Luego, φ transforma cada conjunto de medida nula dentro de G en otro conjunto de medida nula.

5°) Si E es un subconjunto medible de G, entonces existen un conjunto D de clase G_δ y un conjunto Z de medida nula (ambos contenidos en G), tales que $E = D - Z$. El conjunto $\varphi(E) = \varphi(D) - \varphi(Z)$ es medible en virtud de todo lo anterior; además,

$$m(\varphi(E)) = m(\varphi(D)) \leq \int_D J(x)\,dx = \int_E J(x)\,dx,$$

y el corolario queda completamente demostrado.

La primera consecuencia de (6.7) es que si $f : H \to \overline{I\!R}$ es una función medible sobre H, entonces la función compuesta $f \circ \varphi$ es medible sobre G. En efecto, para cada conjunto boreliano M de la recta extendida, $(f \circ \varphi)^{-1}(M) = \varphi^{-1}(f^{-1}(M)) = \varphi^{-1}(F)$, donde $F = f^{-1}(M)$ es un subconjunto medible de H. Puesto que φ^{-1} tiene las mismas propiedades que φ, la primera afirmación de (6.7) implica que $\varphi^{-1}(F)$ es medible, lo cual demuestra que $f \circ \varphi$ es una función medible.

Estamos ahora en condiciones de enunciar y probar el resultado principal de este capítulo.

(6.8) **Teorema.** *Si $f : H \to \overline{I\!R}$ es una función medible no negativa, entonces $f \circ \varphi$ es medible sobre G y además,*

$$\int_H f(y)\,dy = \int_G f(\varphi(x))J(x)\,dx.$$

Demostración. Comenzaremos probando la desigualdad

$$\int_H f(y)\,dy \leq \int_G f(\varphi(x))J(x)\,dx. \tag{4}$$

Si $f = \chi_F$ es la función característica de un conjunto medible $F \subset H$, poniendo $E = \varphi^{-1}(F)$, es claro que $F = \varphi(E)$ y $f(\varphi(x)) = \chi_E(x)$. Por consiguiente, la desigualdad (4) se reduce a la desigualdad del corolario (6.7).

De la linealidad de la integral se deduce que (4) se mantiene válida si f es una función simple no negativa.

Finalmente, si f es una función medible no negativa, entonces existe una sucesión creciente de funciones simples no negativas (f_k) que converge puntualmente a f y la desigualdad (4) resulta ser cierta en general, en virtud del teorema de Beppo Levi.

Si ahora aplicamos la desigualdad (4) a la función $g(x) = f(\varphi(x))J(x)$ permutando H con G y poniendo φ^{-1} en lugar de φ, obtenemos

$$\begin{aligned}\int_G g(x)\,dx &\leq \int_H g(\varphi^{-1}(y))J^{-1}(y)\,dy = \int_H f(y)J(\varphi^{-1}(y))J^{-1}(y)\,dy \\ &= \int_H f(y)\,dy\end{aligned}$$

que es precisamente la desigualdad opuesta a (4), y el teorema queda así demostrado.

Recurriendo a la descomposición $f = f^+ - f^-$, obtenemos inmediatamente el siguiente corolario.

(6.9) **Corolario.** *La función medible $f(y)$ es integrable sobre H si y sólo si $f(\varphi(x))J(x)$ es integrable sobre G, y en tal caso,*

$$\int_H f(y)\,dy = \int_G f(\varphi(x))J(x)\,dx.$$

Recordando que $J(x) = |\partial\varphi/\partial x|$, la última fórmula se puede escribir en la forma más sugestiva

$$\int_H f(y)\,dy = \int_G f(\varphi(x))\left|\frac{\partial\varphi}{\partial x}\right|dx$$

en notoria analogía con la correspondiente fórmula para intervalos de la recta en el caso unidimensional ($n = 1$), de la cual la fórmula que acabamos de demostrar representa una muy amplia generalización. En particular, obsérvese que los conjuntos G y H pueden no ser acotados.

EJERCICIOS

1. Sean $v_1, v_2, ..., v_n$ vectores linealmente independientes de $\mathbb{R}^n$. Mostrar que el paralelepípedo P formado por los puntos de la forma

$$x = t_1v_1 + t_2v_2 + ... + t_nv_n \qquad (0 \leq t_k \leq 1)$$

tiene medida $m(P) = |\det a|$, donde a es la matriz cuyas filas son los vectores dados.

2. La aplicación φ definida por las ecuaciones

$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases} \qquad (r > 0,\ 0 < \theta < 2\pi)$$

transforma biyectivamente el rectángulo infinito abierto $G = \{(r,\theta) \in I\!R^2 : r > 0, 0 < \theta < 2\pi\}$ en el conjunto abierto $H = I\!R^2 - \{(x,y) : x \geq 0, y = 0\}$. Notese que H es todo $I\!R^2$ con excepción de una semirrecta cerrada (un conjunto de medida nula).
Probar que para cualquier función medible no negativa $f(x,y)$, definida sobre $I\!R^2$, se cumple

$$\iint_{I\!R^2} f(x,y)dxdy = \iint_G f(r\cos\theta,\ r\mathrm{sen}\theta)r\ drd\theta.$$

Aplicar esta fórmula y el teorema de Fubini-Tonelli (5.24) a la función $f(x,y) = e^{-(x^2+y^2)}$ para obtener la fórmula $\int_{-\infty}^{\infty} e^{-x^2}dx = \sqrt{\pi}$.

3. Si $f(x)$ es no negativa (o bien integrable) sobre $I\!R^n$ y λ un número real distinto de cero, entonces

$$\int f(\lambda x)dx = |\lambda|^{-n}\int f(x)dx.$$

4. Demuestre que si $a = (a_{ij})$ es una matriz simétrica y $Q(x) = xax^\tau$, donde x^τ es el vector transpuesto de x, la correspondiente forma cuadrática, entonces la función $f(x) = e^{-Q(x)}$ es integrable sobre $I\!R^n$ si y sólo si todos los autovalores de a son positivos, y en tal caso,

$$\int f(x)dx = \frac{\pi^{n/2}}{(\det a)^{1/2}}.$$

5. (coordenadas polares en $I\!R^n$). Consideremos la aplición S de $I\!R^n$ en sí mismo, dada por las ecuaciones

$$\begin{array}{lcl} x_1 & = & r\cos\theta_1 \\ x_2 & = & r\ \mathrm{sen}\theta_1\cos\theta_2 \\ x_3 & = & r\ \mathrm{sen}\theta_1\ \mathrm{sen}\theta_2\cos\theta_3 \\ \ldots & & \ldots\quad\ldots\quad\ldots \\ x_{n-1} & = & r\ \mathrm{sen}\theta_1\ldots\ \mathrm{sen}\theta_{n-2}\cos\theta_{n-1} \\ x_n & = & r\ \mathrm{sen}\theta_1\ldots\ \mathrm{sen}\theta_{n-2}\ \mathrm{sen}\theta_{n-1}, \end{array}$$

de manera tal que $x = (x_1, ..., x_{n-1}, x_n) = S(r, \theta_1, ..., \theta_{n-1})$.
Probar las siguientes afirmaciones
(a) la aplicación S transforma biyectivamente el conjunto abierto G, definido por las relaciones

$$r > 0,\ 0 < \theta_1 < \pi, ..., 0 < \theta_{n-2} < \pi,\ 0 < \theta_{n-1} < 2\pi$$

en el conjunto H formado por los puntos x que satisfacen alguna de las relaciones $x_n \neq 0$ o $x_{n-1} < 0$.
Nótese que el complemento de H es un "semihiperplano" y por consiguiente un conjunto de medida nula en $I\!R^n$. Obsérvese también que el punto $\theta = (\theta_1, ..., \theta_{n-1})$ varía en un intervalo L del espacio $I\!R^{n-1}$, caracterizado por las relaciones

$$0 < \theta_1 < \pi, ..., 0 < \theta_{n-2} < \pi,\ 0 < \theta_{n-1} < 2\pi.$$

(b) si $x = S(r, \theta)$, entonces

$$r^2 = x_1^2 + x_2^2 + ... + x_n^2$$

(c) el determinante jacobiano de la transformación S está dado por la fórmula $J = r^{n-1} g(\theta)$ donde $g(\theta) = \mathrm{sen}^{n-2}\theta_1\ \mathrm{sen}^{n-3}\theta_2 ...\ \mathrm{sen}\theta_{n-2}$.
Sugerencia: expresar S como el producto (composición) $S_2 \circ S_1$, donde S_1 está dada por las ecuaciones

$$x_1 = r\cos\theta_1,\ r_1 = r\ \mathrm{sen}\theta_1,\ \theta_2 = \theta_2, ..., \theta_{n-1} = \theta_{n-1},$$

en tanto que S_2 está dada por

$$x_1 = x_1,\ x_2 = r_1\cos\theta_2,\ x_3 = r_1\ \mathrm{sen}\theta_2\cos\theta_3,$$
$$...,\ x_n = r_1\ \mathrm{sen}\theta_2\ ...\ \mathrm{sen}\theta_{n-1}.$$

Usar inducción con respecto a n.

6. (continuación) Poniendo $x' = S(1, \theta)$, la transformación S puede escribirse en la forma $x = rx'$, donde $r = |x| = (x_1^2 + ... + x_n^2)^{1/2}$ y x' es un punto de la esfera unitaria $\sum = \{x : |x| = 1\}$. Mostrar que para cualquier función medible no negativa se cumple

$$\int f(x)dx = \int_0^\infty dr\ r^{n-1} \int_L f(rx')g(\theta)d\theta.$$

Aplicar esta fórmula a la función $f(x) = e^{-r^2}$ para evaluar la constante

$$C_n = \int_L g(\theta) d\theta,$$

expresándola en términos de la función gamma: $\Gamma(p) = \int_0^\infty t^{p-1}\, e^{-t} dt.$

7. Calcular la medida (o volumen) de la bola unitaria $B = \{x : |x| \leq 1\}$.
8. Probar que $\Gamma(1/2) = \sqrt{\pi}$
9. La función $f(x)$ se llama una **función radial** si existe una función $f_0(t)$ definida sobre la semirecta $t \geq 0$, tal que $f(x) = f_0(|x|)$. Mostrar que si f es una función radial, entonces

$$\int f(x) dx = C_n \int_0^\infty r^{n-1} f_0(r) dr.$$

10. ¿Para qué valores de p es $|x|^p$ integrable sobre la bola unitaria $|x| \leq 1$?
11. Demostrar que la integral biparamétrica

$$\int_0^1 x^{p-1} |\log x|^{q-1} dx$$

es finita si $p > 0$ y $q > 0$; expresar su valor en términos de la función gamma. Sugerencia: considere el cambio de varible $x = e^{-t}$ $(0 \leq t < \infty)$.

12. Calcular la integral de la función $(1+|x|^2)^{-(n+1)/2}$ sobre el espacio $\mathbb{R}^n$.
13. Sea A un subconjunto boreliano de $\mathbb{R}^n$ con la siguiente propiedad: para cada $v \in \mathbb{R}^n$ que verifique $|v| = 1$, el conjunto $A_v = \{t \in \mathbb{R} : tv \in A\}$ tiene medida nula. Probar que A tiene medida nula.
14. Si M es un conjunto convexo en $\mathbb{R}^n$, probar que la frontera de M tiene medida nula y que M es medible.

CAPITULO VII

ESPACIOS DE FUNCIONES CLASICOS

1. El espacio de las funciones integrables.

En este capítulo consideraremos funciones medibles f con valores reales o complejos, definidas sobre un subconjunto medible E del espacio euclidiano $\mathbb{R}^n$.

El conjunto de las funciones integrables sobre E será denotado por $L^1(E)$ o simplemente por L^1, si no hay necesidad de hacer referencia específica a E o bien si este conjunto queda sobreentendido por el contexto.

El espacio L^1 es un espacio vectorial de funciones; y dado que la integral no distingue entre dos funciones que sean iguales en casi todo punto, éstas serán identificadas en una misma clase . En otras palabras, aceptaremos que $f = g$ si $f(x) = g(x)$ en casi todo punto x de E. Así, por ejemplo, la igualdad $f = 0$ significa que $f(x) = 0$ en casi todo punto x de E.

El lector no tendrá dificultad en probar que la relación

"*$f = g$ en casi todo punto de E*"

es una relación de equivalencia entre funciones medibles que respeta las operaciones algebraicas habituales de suma, producto y multiplicación por un número, lo cual permite definir unas operaciones algebraicas homólogas entre las clases de equivalencia, tal como se hace con las estructuras cocientes en el álgebra.

Así, por ejemplo, si denotáramos por (f) la clase de equivalencia de f, podríamos definir las operaciones vectoriales por medio de las fórmulas

$$(f) + (g) = (f + g), \quad \lambda(f) = (\lambda f).$$

En adelante cada elemento de L^1 será una clase de equivalencia de las definidas por aquella relación. Además, el símbolo f se usará indistintamente para representar a la función f o a su clase de equivalencia (pronto veremos que este abuso de notación y de lenguaje resulta muy saludable).

Para cada f en $L^1(E)$ pongamos

$$\|f\|_1 = \int_E |f(x)|\ dx.$$

Nótese que todas las funciones que se encuentran en una misma clase dan el mismo valor para la integral.

De esta manera hemos definido una función no negativa sobre L^1 que cumple

$$\begin{array}{lrcll} \text{N1} & \|f+g\|_1 & \leq & \|f\|_1 + \|g\|_1 & \\ \text{N2} & \|\lambda f\|_1 & = & |\lambda|\ \|f\|_1 & \\ \text{N3} & \|f\|_1 & = & 0 & \text{si y sólo si } f=0, \end{array}$$

para cualquier par de vectores f y g en L^1 y cualquier escalar λ. Luego, $L^1(E)$ es un espacio normado con la norma $\|\cdot\|_1$, en el sentido que hemos definido en la sección 14 del capítulo II.

La distancia entre los vectores f y g se define entonces por medio de la fórmula

$$d(f,g) = \|f-g\|_1 = \int_E |f(x)-g(x)|\ dx$$

y la noción de convergencia con respecto a esta métrica se llama **convergencia en norma**, más específicamente convergencia en norma $\|\ \|_1$ o también conocida como convergencia en L^1.

Seguidamente veremos que como espacio métrico $L^1(E)$ es un espacio completo, o sea que toda sucesión de Cauchy en L^1 converge en norma hacia un vector f del mismo espacio.

Un espacio normado, completo con respecto a la distancia inducida por la norma, se llama un **espacio de Banach**.

(7.1) **Teorema.** $L^1(E)$ *es un espacio de Banach.*

DEMOSTRACIÓN. Resta ver que este espacio es completo, para lo cual demostraremos lo siguiente:

Sea (f_i) una sucesión en L^1 tal que

$$\sum_{i\geq 1} \|f_i\|_1 < \infty ,$$

entonces existe una función S en L^1 tal que

$$\|S_i - S\|_1 \to 0 \quad \text{para } n \to \infty ,$$

donde $S_i = \sum_{1\leq j\leq i} f_j$.

Dejamos como ejercicio demostrar que la afirmación de arriba implica la completitud del espacio L^1 .

Sea ahora $\Phi(x)$ la suma de la serie

$$\sum_{1\leq i} |f_i(x)| .$$

Dado que la serie numérica $\sum \|f_i\|_1$ es convergente, el teorema de Beppo-Levi nos asegura que Φ es integrable, luego finita en casi todo punto. Así está bien definida la función

$$S(x) = \sum_{1\leq i} f_i(x) ,$$

excepto sobre un conjunto de medida nula. Como $|S| \leq \Phi$, la función S está en L^1 . Por otro lado la función $|S_i - S|$ tiende a cero en casi todo punto y está dominada por 2Φ , luego el teorema de la convergencia dominada nos permite afirmar que $\|S - S_i\|_1$ tiende a cero cuando i tiende a infinito.

Q.E.D.

Creemos que la demostración dada es bastante directa y en ella se usan principalmente dos hechos: la completitud del campo de escalares y el teorema de la convergencia dominada. Nos parece instructivo esquematizar una segunda demostración.

Sea (f_i) una sucesión de Cauchy en L^1 . Luego por la desigualdad de Chebyshev es de Cauchy en medida y entonces existe una función medible f tal que f_i converge a f en medida (véase Teorema 4.15). Además existe una subsucesión (f_{i_j}) de (f_i) que converge a f en casi todo punto. Usando el teorema de Fatou tenemos

$$\|f - f_i\|_1 \leq \liminf_{j\to\infty} \|f_{i_j} - f_i\|_1 .$$

Esta última desigualdad nos asegura que f está en L^1 y que f_i converge hacia f en norma $\| \|_1$.

Es útil tener presente que los argumentos utilizados arriba dan también una demostración del siguiente hecho.

Sea (f_i) una sucesión de funciones integrables que converge en L^1 hacia una función f. Entonces existe una subsucesión (f_{i_j}) de (f_i) que converge a f en casi todo punto.

Para lo que sigue conviene recordar que se llama **soporte** de una función f a la adherencia del conjunto formado por los puntos x tales que $f(x) \neq 0$.

Un principio general, muchas veces útil, es reemplazar una función "buena" g de tal manera que el "resto" $\|f - g\|_1$ sea chico. Veremos a continuación algunas de las clases de funciones buenas que se usan con frecuencia.

Sea $S = S(E)$ el conjunto de las funciones simples e integrables sobre E. Este conjunto S es **denso** en L^1, lo que significa que para cada f en L^1 y cada $\varepsilon > 0$ existe g en S tal que $\|f - g\|_1 < \varepsilon$.

Las funciones escalonadas, combinaciones lineales finitas de características de intervalos acotados, son densas en $L^1(\mathbb{R}^n)$. Esta afirmación es consecuencia del hecho de que S es denso en L^1 y de la definición de medida de Lebesgue.

Por otro lado no es difícil convencerse de que la función característica de un intervalo acotado de $\mathbb{R}^n$ es aproximable en norma L^1 por funciones continuas con soporte compacto; luego esta última clase de funciones es también densa en L^1. Nótese que hemos prácticamente demostrado al pasar que L^1 es **separable**, i.e. tiene un subconjunto numerable denso.

El mismo tipo de argumento que hemos dado nos permite afirmar que $C_0^m(\mathbb{R}^n)$ es denso en L^1. Esta última clase consta de las funciones con soporte compacto y con derivadas parciales continuas hasta el orden m. En particular, $C_0(\mathbb{R}^n)$ denotará la clase de las funciones continuas con soporte compacto.

Para referencia enunciamos el siguiente teorema:

(7.2) **Teorema.** *La clase formada por las funciones continuas con soporte compacto y la clase de las funciones escalonadas son densas en el espacio $L^1(\mathbb{R}^n)$.*

2. Las funciones esencialmente acotadas.

Una función medible f es esencialmente acotada sobre E si existe una constante finita M tal que

$$|\{x \in E : |f(x)| > M\}| = 0 \ ,$$

en esta situación M recibe el nombre de **cota esencial**. En otras palabras M es cota esencial para f sobre E sii $|f(x)| \leq M$ para casi todo $x \in E$. Pondremos $\|f\|_\infty$ para el ínfimo de las cotas esenciales. Se demuestra fácilmente que $\|f\|_\infty$ es una cota esencial para f.

Nosotros hemos definido $\|f\|_\infty$ cuando f es esencialmente acotada; si este no es el caso pondremos $\|f\|_\infty = \infty$, así f es esencialmente acotada sii $\|f\|_\infty < \infty$. El conjunto de las funciones esencialmente acotadas sobre E es denotado por L^∞. En L^∞ identificamos las funciones iguales en casi todo punto; con esta convención $\|\cdot\|_\infty$ es una norma sobre este espacio vectorial de funciones. Más todavía, L^∞ es un espacio de Banach. Dejamos como ejercicio demostrar esta afirmación.

Nuevamente los resultados del Capítulo IV nos aseguran que las funciones simples son densas en L^∞. Por otro lado ni las funciones escalonadas ni las funciones continuas acotadas son densas en $L^\infty(\mathbb{R}^n)$.

Nótese que el Teorema (7.2) nos asegura que no existe ningún espacio vectorial X estrictamente comprendido entre $C_0(\mathbb{R}^n)$ y $L^1(\mathbb{R}^n)$ tal que X con $\|\ \|_1$ sea un espacio de Banach, dicho de otra manera uno podría definir $L^1(\mathbb{R}^n)$ como la completación de $C_0(\mathbb{R}^n)$ con la norma $\|\cdot\|_1$. No tenemos la misma situación con la norma $\|\ \|_\infty$, pues $C_0(\mathbb{R}^n)$ con esta norma no es un espacio de Banach y el mínimo espacio normado completo que lo contiene no es L^∞. Por otro lado las funciones continuas acotadas sobre $\mathbb{R}^n$ forman un espacio de Banach con la norma $\|\ \|_\infty$.

Veremos a continuación de qué manera están relacionados los espacios L^1 y L^∞.

Notemos que si $f \in L^1$ y $g \in L^\infty$ el producto fg es integrable. Más todavía, tenemos que

$$\int_E |fg| \ dx \leq \|f\|_1 \ \|g\|_\infty \ .$$

Luego para cada función $g \in L^\infty$ está bien definida la siguiente función sobre L^1

$$\ell_g(f) = \int_E f \ \overline{g} \ dx \ ,$$

donde $\overline{g}(x) = \overline{g(x)}$ denota el complejo conjugado de $g(x)$.

Esta función ℓ_g es lineal y verifica

$$|\ell_g(f)| \leq \|g\|_\infty \ \|f\|_1 \ ,$$

para cada $f \in L^1$. Este es un ejemplo de ciertas funciones generales que describiremos brevemente.

Diremos que una transformación lineal T de un espacio normado X en un espacio normado Y es **acotada** si existe un número no negativo $M < \infty$ tal que

$$\|Tx\| \leq M\|x\| \ ,$$

para cada $x \in X$.

Obsérvese que en la desigualdad anterior hemos usado el mismo símbolo $\| \ \|$ para denotar tanto la norma en X como en Y, pero está claro que éstas pueden ser de naturaleza completamente distinta.

Al conjunto de transformaciones lineales (u operadores lineales) acotados de X en Y lo designaremos con $L(X,Y)$. Este es un espacio vectorial de funciones en el que podemos definir la siguiente norma:

Si $T \in L(X,Y)$ entonces $\|T\| = \sup\{\|Tx\| \ : \ \|x\| \leq 1\}$.

En los ejercicios se verá que si Y es un espacio completo también lo es $L(X,Y)$; en particular $L(X,\mathbb{C})$ o $L(X,\mathbb{R})$ son espacios de Banach. Para cualquiera de estos dos usaremos la notación X^* y diremos que es el **espacio dual** de X. Para nosotros X^* siempre denotará el dual topológico de X cuyos elementos son las **funcionales lineales acotadas**.

(7.3) **Teorema.** *Si $g \in L^\infty$ entonces $\ell_g \in (L^1)^*$ y $\|\ell_g\| = \|g\|_\infty$. Más aún, cada elemento de $(L^1)^*$ es de esta forma.*

Usualmente se enuncia este teorema diciendo que el dual de L^1 es L^∞ y se escribe $(L^1)^* = L^\infty$.

Demostración. Veamos la primera parte del teorema para $g \in L^\infty$ no esencialmente nula. Por un lado es claro que $\|\ell_g\| \leq \|g\|_\infty$. Solamente es de interés el caso $m(E) > 0$. Luego, dado $\varepsilon > 0$ existe $A \subseteq E$, $0 < m(A) < \infty$, tal que $|g(x)| \geq \|g\|_\infty - \varepsilon$ si $x \in A$.

Pongamos ahora $f(x) = sgn(g(x)) \ \frac{1}{m(A)} \ \chi_A(x)$, donde $sgnx = x/|x|$ si $x \neq 0$ y $sgn0 = 0$. Claramente tenemos

$$\|\ell_g\| \geq \ell_g(f) \geq \|g\|_\infty - \varepsilon \ ,$$

con lo que $\|\ell_g\| = \|g\|_\infty$. La demostración de la segunda parte del teorema es más complicada y necesitamos postergarla hasta el final del siguiente parágrafo.

Sabemos que si $g \in L^\infty$, entonces $gf \in L^1$ para cada $f \in L^1$. Veamos que la afirmación recíproca también es cierta.

(7.4) **Teorema.** *Sea g una función medible tal que $gf \in L^1$ para cada $f \in L^1$. Entonces $g \in L^\infty$.*

DEMOSTRACIÓN. Las hipótesis del teorema nos garantizan que g es una función finita en casi todo punto. Nosotros demostraremos que si g es una función finita en casi todo punto tal que $g \notin L^\infty$ entonces se puede construir una función $f \in L^1$ tal que $gf \notin L^1$. En efecto, sea (a_i) una sucesión numérica tal que

$$0 < a_1 < a_2 < \ldots \text{ y } \sum_{1 \le i} a_i^{-1} < \infty \ .$$

Sea ahora A_i, $0 < m(A_i) < \infty$ contenido en el conjunto

$$\{x \in E : a_i < |g(x)|\} \ .$$

La siguiente función

$$f = \sum_{i=1}^{\infty} (a_i \ m(A_i))^{-1} \chi_{A_i}$$

es la que mencionamos en un principio.

Q.E.D.

3. Funciones de cuadrado integrable.

Denotamos con $L^2(E)$ o L^2 al espacio formado por todas las funciones medibles f con valores en los complejos tales que $|f|^2 \in L^1(E)$. Consideraciones similares a las que haremos serán válidas si las funciones toman valores en la recta real extendida.

El espacio L^2 es un espacio vectorial y si $f, g \in L^2$ entonces $fg \in L^1$ (recuerde que el producto de dos números positivos es menor o igual que el promedio de sus cuadrados). Luego está bien definida la expresión

$$(f,g) = \int_E f(x)\ \overline{g(x)}\ dx\ .$$

Vemos que $(\cdot\,,\cdot)$ es un **producto escalar**, i.e. una función con valores en los complejos con las siguientes propiedades: es lineal en la primera variable, $(f,g) = \overline{(g,f)}$, $(f,f) \geq 0$ y $(f,f) = 0$ sii $f = 0$.

Si ponemos $\|f\|_2 = (f,f)^{1/2}$ y seguimos los pasos realizados para el caso de la norma euclídea en $\mathbb{R}^n$, ver Capítulo II, parágrafo 1, tenemos que L^2 es un espacio normado con la norma $\|\cdot\|_2$, la cual proviene de un producto escalar. Además se cumple la desigualdad de Cauchy-Schwarz

$$\int_E |f(x)g(x)|\ dx \leq \|f\|_2\ \|g\|_2\ ,$$

para $f, g \in L^2$.

En forma similar a lo realizado en L^1 se demuestra que L^2 es completo; o sea que estamos en presencia de un espacio normado completo cuya norma proviene de un producto escalar. Esto se resume diciendo que L^2 es un **espacio de Hilbert**.

Haremos a continuación una pequeña incursión por la teoría abstracta de los espacios de Hilbert.

Sea H un espacio vectorial sobre los números complejos provisto de un producto escalar $(\ ,\ )$ y una norma $\|x\| = (x,x)^{1/2}$. Supondremos además que es completo como espacio normado, es decir H es un espacio de Hilbert.

(7.5) **Teorema.** *Sea H un espacio de Hilbert y C un subconjunto no vacío de H cerrado y convexo. Entonces C posee un único elemento de norma mínima, i.e. existe un único $c_0 \in C$ tal que*

$$\|c_0\| = \inf_{c \in C}\ \|c\|\ .$$

DEMOSTRACIÓN. Sea (c_i) una **sucesión minimizante** en C, i.e. cada c_i está en C y $\|c_i\| \to d = \inf\limits_{c \in C} \|c\|$ cuando $i \to \infty$.

De la siguiente igualdad **(ley del paralelogramo)**

$$\left\| \frac{c_i + c_j}{2} \right\|^2 + \left\| \frac{c_i - c_j}{2} \right\|^2 = \frac{1}{2} \left(\|c_i\|^2 + \|c_j\|^2 \right) ,$$

teniendo en cuenta la convexidad de C se obtiene la desiguadad

$$4d^2 + \|c_i - c_j\|^2 \leq 2 \left(\|c_i\|^2 + \|c_j\|^2 \right) .$$

Esta nos asegura que (c_i) es de Cauchy en H; luego existe c_0 tal que $\|c_i - c_0\| \to 0$ para $i \to \infty$, y como C es cerrado se tiene que $c_0 \in C$. Teniendo en cuenta que (c_i) es minimizante llegamos a $\|c_0\| = d$. La unicidad de este c_0 es consecuencia directa de la ley del paralelogramo y de la convexidad del conjunto C.

Q.E.D.

El teorema demostrado nos permite asegurar la existencia y unicidad de la **proyección** de un vector x sobre un conjunto convexo cerrado C. Dicha proyección, que será denotada por $p(x) = p(x/C)$, se define como el único elemento en C que verifica

$$\|x - p(x)\| \leq \|x - c\| , \qquad c \in C .$$

Observe que el conjunto $x - C = \{x - c : c \in C\}$ es convexo y cerrado.

(7.6) **Teorema.** *Sean H un espacio de Hilbert, C un subconjunto de H no vacío cerrado y convexo, y sean x, p dos vectores en H. Entonces las siguientes afirmaciones son equivalentes:*
(1) El vector p es la proyección de x sobre C.
(2) $p \in C$ y $Re(x - p, p - c) \geq 0$, para cada $c \in C$.
(3) $p \in C$ y $\|x - c\|^2 \geq \|x - p\|^2 + \|p - c\|^2$, para cada $c \in C$.

DEMOSTRACIÓN. Para ver que (1) implica (2) se considera la función

$$f(\varepsilon) = \|x - (1 - \varepsilon)p - \varepsilon c\|^2 , \qquad 1 \geq \varepsilon \geq 0 ,$$

luego se observa que $f'(0) \geq 0$ por ser p una proyección. Calculando $f'(0)$ se obtiene la parte real de $2(x - p, p - c)$.

La parte (3) es consecuencia de (2) y de desarrollar la expresión

$$\|x - c\|^2 = \|(x - p) + (p - c)\|^2 .$$

Claramente (3) implica (1).

Q.E.D.

La proyección $p(x/C)$ tiene particular interés cuando C es un subespacio cerrado de H, en cuyo caso la condición (2) de (7.6) toma la forma:

$$p \in C \text{ y } (x-p, c) = 0 \text{ para cada } c \in C\ .$$

Esto último se suele expresar diciendo que un vector p en C es la proyección de x sii $x-p$ **es perpendicular** a C.

Observemos que la afirmación (3) de (7.6) implica la unicidad de la proyección. En efecto si p y q son dos vectores en C que satisfacen (3), tendremos

$$\|x-q\|^2 \geq \|x-p\|^2 + \|p-q\|^2$$
$$\|x-p\|^2 \geq \|x-q\|^2 + \|p-q\|^2\ ,$$

pero estas dos inecuaciones aseguran que $\|p-q\|^2 = 0$, o sea $p = q$.

Veamos a continuación cómo son la funcionales lineales de un espacio de Hilbert H.

Si $v \in H$ y ponemos $\ell_v(u) = (u,v)$ entonces $\ell_v \in H^*$. Además, como fácilmente se ve, $\|\ell_v\| = \|v\|$. El teorema de representación de Riesz, que probaremos a continuación, nos asegura que aparte de éstos no hay otros ejemplos de funcionales lineales continuas sobre H.

(7.7) **Teorema.** *Sea H un espacio de Hilbert y ℓ una funcional lineal continua en H. Entonces existe un único vector $v \in H$ tal que $\ell(u) = (u,v)$ para cada $u \in H$. Además $\|\ell\| = \|v\|$.*

Demostración. supondremos de entrada que existe $u \in H$ tal que $\ell(u) \neq 0$; si no, la afirmación es obvia. Definamos

$$N = \{x \in H : \ell(x) = 0\}\ .$$

Por ser ℓ lineal y continua el conjunto N es un subespacio cerrado y además no es todo H. Sea ahora $y_0 \in H$, $y_0 \notin N$, luego por (7.6) existe un único $x \in N$ tal que el vector $y_0 - x$ es perpendicular a N. En otras palabras, siempre podemos encontrar un vector $y \notin N$ perpendicular a N tal que $\|y\| = 1$. Cada vector $x \in H$ se puede escribir así :

$$x = \left(x - \frac{\ell(x)}{\ell(y)}\, y\right) + \frac{\ell(x)}{\ell(y)}\, y\ ,$$

o sea que todo vector del espacio admite una representación de la forma $x = u + \alpha y$, $u \in N$, α complejo.

Pongamos ahora $v = \overline{\ell(y)}\ y$. Claramente $\ell(\alpha y) = (\alpha y, v)$. Además si $x \in N$, entonces $(v, x) = 0$. Por lo tanto tenemos que para todo $x \in H$ se verifica que

$$\ell(x) = \ell(\alpha y + u) = \ell(\alpha y) = (\alpha y, v) = (x, v) \ .$$

Es obvio que el vector v es único.

Q.E.D.

Usaremos el teorema de representación de Riesz para completar la demostración del Teorema (7.3).

Sea ℓ una funcional lineal continua definida sobre $L^1(E)$. Nos interesa ver que existe $g \in L^\infty(E)$ tal que

$$\ell(f) = \int_E f\ \overline{g}\ dx = \ell_g(f)\ , \qquad f \in L^1(E)\ .$$

Supondremos primero que $m(E) < \infty$. Recordemos que ℓ verifica

$$|\ell(f)| \leq \|\ell\|\ \|f\|_1\ , \qquad f \in L^1(E)\ .$$

Usando Cauchy-Schwarz tenemos que

$$|\ell(f)| \leq \|\ell\|[m(E)]^{1/2}\ \|f\|_2\ , \qquad f \in L^2(E)\ .$$

O sea que podemos pensar a ℓ como una funcional lineal sobre el espacio de Hilbert $L^2(E)$, así usando el Teorema (7.7) vemos que existe $g \in L^2(E)$ tal que

$$\ell(f) = \int_E f\overline{g}\ dx = \ell_g(f)\ , \qquad f \in L^2(E)\ .$$

De la anterior igualdad y teniendo en cuenta que ℓ is una funcional lineal continua tenemos que

$$\int_E |fg|\ dx \leq \|f\|_1\ \|\ell\|\ , \qquad f \in L^2(E)\ .$$

Si f is una función integrable, la última desigualdad vale para $f_n = \min\ (|f|, n)$ y haciendo tender n a infinito Beppo-Levi nos garantiza que ella es cierta para f integrable. Usando (7.4) se obtiene que $g \in L^\infty(E)$.

Ahora bien; las funcionales lineales continuas ℓ y ℓ_g coinciden en L^2 (un subconjunto denso de L^1). Luego son idénticas.

Si el conjunto E tiene medida infinita se lo puede reducir al caso ya demostrado poniendo a este conjunto como una unión numerable disjunta de conjuntos F_i de medida finita. El resto de la demostración se deja a cargo del lector. Q.E.D.

4. Funciones convexas.

En el Capítulo II, Sección 14 hemos visto la utilidad de las funciones convexas como generadoras de normas en $\mathbb{R}^n$. Es nuestra intención estudiar con algún detalle propiedades de estas funciones.

Denotamos por I un intervalo de la recta con extremos a y b, el cual puede contener alguno de sus extremos y ser no acotado.

Una función φ de I en $\mathbb{R}$ se llama **convexa** sobre I si las relaciones $s, t \in I$ y $0 \leq \lambda \leq 1$ implican

$$\varphi(\lambda s + (1-\lambda)t) \leq \lambda\varphi(s) + (1-\lambda)\varphi(t) \ . \tag{1}$$

Es siempre conveniente tener la imagen geométrica de la desigualdad anterior. Si P_1 y P_2 son dos puntos sobre la curva $y = \varphi(x)$, los puntos del arco de curva $P_1\ P_2$ deben estar por debajo o sobre la cuerda determinada por los puntos $P_1\ P_2$.

Una manera alternativa de definir función convexa es suponer que φ es continua sobre I y que verifica

$$\varphi\left(\frac{s+t}{2}\right) \leq \frac{1}{2}\,\varphi(s) + \frac{1}{2}\,\varphi(t) \ , \qquad (s,t \in I) \ . \tag{2}$$

En este parágrafo veremos que (1) implica la continuidad de la función φ en el interior del intervalo I. El hecho de que la continuidad φ conjuntamente con (2) implican la convexidad de φ se deja como ejercicio. La caracterización con (2) es más fácil de verificar.

A continuación damos un ejemplo de función convexa. Sea $p \in L^1(I)$ una función creciente. Luego para cada $t \in I$ ponemos

$$\varphi(t) = \int_a^t p(s)\ ds \ . \tag{3}$$

Por consiguiente φ es continua sobre I. El hecho de que la función definida por (3) cumple (2) es un ejercicio que el lector debería realizar.

Más adelante veremos que la fórmula (3) nos da esencialmente todas las funciones convexas.

Usando (3) obtenemos rápidamente la convexidad de las siguientes funciones x^p $p \geq 1$, e^x, $xlnx$ sobre $x \geq 0$.

La desigualdad (1) equivale a afirmar que para cualquier terna de puntos $x < y < z$ en el intervalo I se verifica alguna de las relaciones siguientes:

$$\frac{\varphi(y)-\varphi(x)}{y-x} \leq \frac{\varphi(z)-\varphi(x)}{z-x} . \tag{4}$$

$$\frac{\varphi(z)-\varphi(x)}{z-x} \leq \frac{\varphi(z)-\varphi(y)}{z-y} . \tag{5}$$

$$\frac{\varphi(y)-\varphi(x)}{y-x} \leq \frac{\varphi(z)-\varphi(y)}{z-y} . \tag{6}$$

Todas las desigualdades tienen un claro significado geométrico que en la mayoría de los casos es suficiente para convencernos de su validez. A modo de ejemplo damos un razonamiento analítico para ver que (1) es equivalente a (4).

A partir de (1) obtenemos

$$\varphi(y) \leq \frac{z-y}{z-x}\,\varphi(x) + \frac{y-x}{z-x}\,\varphi(z) .$$

Si a esta última desigualdad le sumamos miembro a miembro $-\varphi(x)$ obtenemos (4). El proceso claramente se puede revertir, de modo que (4) es equivalente a (1). Las dos equivalencias restantes no ofrecen dificultad adicional.

Podemos usar (6) para demostrar que si $\varphi \in C^1(a,b)$ y φ' es creciente entonces φ es convexa sobre (a,b), para lo cual basta usar el teorema del valor medio.

Pongamos ahora

$$D^+\varphi(x) = \lim_{h\to 0+} \frac{\varphi(x+h)-\varphi(x)}{h}$$

$$D^- \varphi(x) = \lim_{h \to 0+} \frac{\varphi(x) - \varphi(x-h)}{h} \;,$$

cuando estos límites existan.

Si φ es una función convexa sobre I valen las afirmaciones (a)-(c) siguientes:

(a) $D^+\varphi(x)$ existe y es menor que infinito para $a \leq x < b$. [use (4)].

(b) $D^-\varphi(x)$ existe y es mayor que menos infinito para $a < x \leq b$. [use (5)].

(c) $D^-\varphi(x) \leq D^+\varphi(x)$ para $a < x < b$. [use (6)].

En particular hemos demostrado que si φ es convexa sobre I, entonces es continua en el interior de I. Nos interesa ahora analizar la monotonía de las derivadas.

Para una función convexa, como consecuencia de (6), tenemos

$$\text{(7)} \qquad \frac{\varphi(x) - \varphi(x_1)}{x - x_1} \leq \frac{\varphi(x_2) - \varphi(y)}{x_2 - y} \;, \qquad x_1 < x < y < x_2 \;.$$

Si en esta desigualdad hacemos $x \to x_1$, $y \to x_2$, obtenemos $D^+\varphi(x_1) \leq D^-\varphi(x_2)$ y usando (6) resulta

$$\text{(8)} \qquad D^+\varphi(x_1) \leq D^+\varphi(x_2) \qquad \text{para } x_1 < x_2 \;.$$

Así hemos demostrado que $D^+\varphi$ es monótona creciente. Sabemos que una función monótona tiene a lo más una cantidad numerable de discontinuidades. Sea ahora x un punto de continuidad de $D^+\varphi$; como $D^+\varphi(y) \leq D^-\varphi(x) \leq D^+\varphi(x)$, si $y < x$, haciendo $y \to x$ tenemos que $D^-\varphi(x) = D^+\varphi(x)$.

Todo esto nos dice que si x es un punto de continuidad para $D^+\varphi$ entonces $\varphi'(x)$ existe. En forma análoga uno puede demostrar que $D^-\varphi$ es monótona y que si x es un punto de continuidad para $D^-\varphi$ entonces $\varphi'(x)$ existe. El enunciado del siguiente teorema resume nuestros razonamientos anteriores.

(7.8) **Teorema.** *Sea φ convexa sobre I. Entonces φ es continua en el interior del I y para cada punto interior existen las derivadas laterales a derecha y a izquierda de φ, las cuales son funciones crecientes. Si x es un punto de continuidad para la derivada lateral derecha o izquierda,* *entonces $\varphi'(x)$ existe. La derivada $\varphi'(x)$ existe salvo en un conjunto numerable de puntos.*

Diremos que una recta $y = \ell(x)$ **soporta** a $y = \varphi(x)$ en el punto x_0 si $\ell(x_0) = \varphi(x_0)$ y además $\varphi(x) \geq \ell(x)$ para todo x. Para una función convexa φ las rectas soportes existen en cada punto interior de I y están dadas por

$$\ell(x) = \varphi(x_0) + m(x - x_0) \ ,$$

donde $D^-\varphi(x_0) \leq m \leq D^+\varphi(x_0)$. Para ver esto use la definición de derivadas laterales.

El lector no tendrá dificultad en demostrar que si φ es una función convexa sobre I, entonces existen dos sucesiones de números reales (a_i) y (b_i) tal que

$$\varphi(x) = \sup(a_i x + b_i) \ , \qquad a < x < b \ .$$

Por otro lado si (φ_i) es una sucesión de funciones convexas sobre I y definimos

$$\varphi(x) = \sup_i \varphi_i(x) \ ,$$

entonces la función φ es convexa, suponiéndola finita. Por lo tanto también podemos pensar a una función convexa como el supremo de una sucesión de funciones lineales.

5. Los espacios L^p.

Dado un conjunto medible $E \subset \mathbb{R}^n$, para cualquier función medible f y cualquier $p > 0$, llamaremos **norma *p*** de f sobre E al número

$$\|f\|_p = \|f\|_{p,E} = \left(\int_E |f(x)|^p \ dx \right)^{1/p} \ ,$$

que puede ser igual a $+\infty$.

Las funciones f que verifican $\|f\|_p < \infty$ forman una clase muy especial que se denota por $L^p(E)$ o simplemente L^p, si no hay posibilidad de confusión. Luego, $f \in L^p(E)$ si y sólo si $|f|^p \in L^1(E)$. Como antes, convendremos en identificar dos funciones cualesquiera que coincidan en casi todo punto de E; y por razones que enseguida se comprenderán nos interesaremos principalmente en el caso $p \geq 1$.

El espacio L^p, $1 \leq p < \infty$ es un espacio normado con la norma $\|\cdot\|_p$. La única propiedad a demostrar que no es obvia es la desigualdad triangular o también llamada desigualdad de Minkowski:

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p \ , \qquad\qquad f,g \in L^p \ .$$

Para demostrar la desigualdad de Minkowski supondremos, sin pérdida de generalidad, que $\|f\|_p$ y $\|g\|_p$ son positivas. Dado que t^p es una función convexa para $p \geq 1$, para cada $x \in E$ se cumple

$$\left(\frac{|f(x)|+|g(x)|}{\|f\|_p+\|g\|_p}\right)^p \leq \frac{\|f\|_p}{\|f\|_p+\|g\|_p}\,\frac{|f(x)|^p}{\|f\|_p^p} + \frac{\|g\|_p}{\|f\|_p+\|g\|_p}\,\frac{|g(x)|^p}{\|g\|_p^p} \ .$$

Integrando sobre E obtenemos

$$\int_E \left(\frac{|f(x)|+|g(x)|}{\|f\|_p+\|g\|_p}\right)^p dx \leq 1 \ .$$

Por lo tanto

$$\int_E |f+g|^p dx \leq \int_E (|f|+|g|)^p \, dx \leq (\|f\|_p + \|g\|_p)^p \ ,$$

y esta última desigualdad implica la desigualdad de Minkowski.

En forma similar a lo hecho en L^1 se demuestra el teorema siguiente.

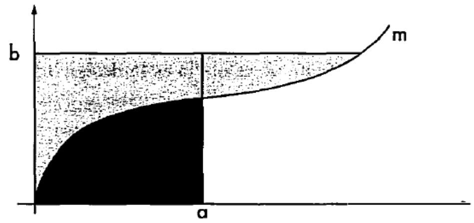
(7.9) **Teorema.** *Para $1 \leq p \leq \infty$, $L^p(E)$ es un espacio de Banach.*

Observemos además, como consecuencia de la desigualdad de Chebyshev, que convergencia en norma L^p implica convergencia en medida. Así, si una sucesión (f_i) converge hacia una función f en norma L^p, entonces existe una subsucesión (f_{i_j}) de (f_i) que converge a f en casi todo punto.

En L^p existe una desigualdad análoga a la de Cauchy-Schwarz. Para obtener este tipo de desigualdad pongamos $m(t) = t^{p-1}$, $t \geq 0$ y $1 < p < \infty$ fijo, y sea $n(t)$ la función inversa de $m(t)$, o sea $n(t) = t^{1/(p-1)}$. Tenemos la siguiente desigualdad, válida para $a, b \geq 0$

$$ab \leq \int_0^a m(t)\ dt + \int_0^b n(t)\ dt \ ,$$

para cuya demostración es suficiente analizar la figura.



Más específicamente tenemos

$$ab \leq \frac{1}{p}\, a^p + \frac{1}{\frac{1}{p-1}+1}\, b^{\frac{1}{p-1}+1}, \qquad a,b \geq 0\ .$$

O bien, poniendo $p' = p/(p-1)$, obtenemos la desigualdad

$$ab \leq \frac{1}{p}\, a^p \ + \ \frac{1}{p'}\, b^{p'}\ ,$$

donde a y b son no negativos y $1 < p < \infty$. Notemos que $1/p \ + \ 1/p' \ = 1$. El número p' se llama el **exponente conjugado** de p. Si $p = 1$ su conjugado será $p' = \infty$; si $p = \infty$ definimos $p' = 1$.

Supongamos ahora que tenemos $f \in L^p$ y $g \in L^{p'}$ dos funciones esencialmente no nulas, i.e. $\|f\|_p\|g\|_{p'} > 0$. Si en la última desigualdad ponemos

$$a = |f(x)|/\|f\|_p \quad \text{y} \quad b = |g(x)|/\|g\|_{p'}$$

e integramos sobre el conjunto E, obtenemos

$$\int_E \frac{|f(x)|\ |g(x)|}{\|f\|_p\|g\|_{p'}}\, dx \leq \frac{1}{p} \ + \ \frac{1}{p'} = 1\ .$$

Así hemos demostrado la llamada desigualdad de Hölder.

Desigualdad de Hölder: Sean $f \in L^p(E)$, $g \in L^{p'}(E)$, $1 \leq p \leq \infty$, $(1/p + 1/p') = 1$. Entonces

$$\int_E |f(x)g(x)|dx \leq \|f\|_p\|g\|_{p'}\ .$$

Los detalles faltantes en la demostración anterior no son difíciles de completar.

Una vez conocida la desigualdad de Hölder, la desigualdad de Minkowski se obtiene fácilmente a partir de ella (véanse los ejercicios).

Otra desigualdad importante es la siguiente.

Desigualdad de Jensen: Sea $f \in L^1(E)$ una función con valores reales, $0 < m(E) < \infty$. Si φ es una función convexa definida sobre $I\!R$, entonces

$$\varphi\left(\frac{1}{m(E)} \int_E f(x)\, dx\right) \leq \frac{1}{m(E)} \int_E \varphi(f(x))\, dx\ .$$

La integral de la derecha está bien definida pudiendo valer más infinito.

En efecto, por ser φ una función convexa existen dos sucesiones (a_i) y (b_i) tales que $\varphi(x) = \sup_i (a_i x + b_i)$.

Luego

$$\begin{aligned}
\varphi\left(\frac{1}{m(E)} \int_E f(x)\, dx\right) &= \sup_i \left(\frac{a_i}{m(E)} \int_E f(x)\, dx + b_i\right) \\
&= \sup_i \frac{1}{m(E)} \int_E (a_i\, f(x) + b_i) dx \\
&\leq \frac{1}{m(E)} \int_E \sup_i (a_i f(x) + b_i) dx \\
&= \frac{1}{m(E)} \int_E \varphi(f)\, dx
\end{aligned} .$$

El siguiente teorema nos da una manera muy útil de expresar la norma L^p de una función.

(7.10) **Teorema.** *Sea f una función medible con valores reales definida sobre E, y $1 \leq p \leq \infty$. Entonces*
(a) Si $f \in L^p$, su norma está dada por

$$\|f\|_p = \sup_g \int_E fg\, dx\ ,$$

donde el supremo se toma sobre toda las funciones g que verifican $\|g\|_{p'} \leq 1$.
(b) Si el supremo anterior es finito la función f pertenece a L^p y dicho supremo es $\|f\|_p$.

DEMOSTRACIÓN: Analizaremos el caso $1 \le p < \infty$; para $p = \infty$ véase (7.3) y (7.4).

Para demostrar (a) pongamos

$$N_p(f) = \sup_{\|g\|_{p'} \le 1} \int_E fg \, dx .$$

Por la desigualdad de Hölder tenemos que

$$N_p(f) \le \|f\|_p .$$

Puesto que para $f = 0$ tenemos la igualdad (a), supondremos $\|f\|_p > 0$. Poniendo $g = \frac{|f|^{p-1}}{a} \mathrm{sgn} f$, con $a = \|f\|_p^{p-1}$, se tiene $\|g\|_{p'} = 1$ y $N_p(f) \ge \int_E fg \, dx = \|f\|_p$. De esta manera hemos demostrado la parte (a).

Para demostrar (b) veremos que si $\|f\|_p = \infty$ para una función no negativa f, entonces $N_p(f) = \infty$. Sea f_i definida en cada punto de E como sigue:

$$f_i(x) = \begin{cases} 0 & \text{si } |x| > i \\ \min(f(x), i) & \text{si } |x| \le i \, . \end{cases}$$

Claramente $f_i \in L^p$ y el teorema de Beppo-Levi nos asegura que $\|f_i\|_p \to \infty$ cuando $i \to \infty$. Sea $g_i \in L^{p'}$, $\|g_i\|_{p'} = 1$ tal que

$$\int_E f_i g_i = \|f_i\|_p .$$

Luego, para cualquier n se tiene

$$N_p(f) \ge \int_E f g_i \ge \int_E f_i g_i = \|f_i\|_p ,$$

o sea $N_p(f) = \infty$.

Q.E.D.

En el teorema anterior no hace falta toda la bola de $L^{p'}$ para definir la norma L^p. Sea $D \subseteq L^{p'}$ una clase de funciones **densa** en $L^{p'}$, i.e. para cada $g \in L^{p'}$ existe una sucesión $(g_i) \subseteq D$ tal que $\|g - g_i\|_{p'} \to 0$ cuando $i \to \infty$.

(7.10)' **Teorema.** *(a) Si* $f \in L^p$, $1 \le p \le \infty$, *entonces*

$$\|f\|_p = \sup \left\{ \int_E fg \, dx : \|g\|_{p'} \le 1 , \qquad g \in D \right\}$$

(b) Si f es una función medible y c una constante finita tal que

$$\int_E |f|\ |g|\ dx \leq c\|g\|_{p'}\ , \qquad g \in D\ .$$

Entonces $f \in L^p$.

DEMOSTRACIÓN: Ver ejercicios

Nuevamente podemos preguntarnos cuáles son lass funcionales lineales continuas definidas sobre L^p .

Sea $g \in L^p$ y pongamos como antes

$$\ell_g(f) = \int_E f\ \overline{g}\ dx \qquad (f \in L^p)\ .$$

Entonces tenemos $\ell_g \in (L^p)^*$ y $\|\ell_g\| = \|g\|_{p'}$. Aquí $1 \leq p \leq \infty$.

Enunciamos ahora el siguiente teorema:

(7.11) **Teorema de representación de Riesz.** *Sea $1 \leq p < \infty$ y ℓ una funcional lineal continua sobre L^p . Entonces existe una única $g \in L^{p'}$ tal que $\ell = \ell_g$.*

Con las herramientas que poseemos en este momento es muy fácil demostrar (7.11) para $1 \leq p \leq 2$ (ver ejercicios). Más adelante daremos una demostración para los restantes p ; pero para ello deberemos usar una de los teoremas más importantes de la teoría de la medida, a saber, el teorema de Radon - Nikodym.

6. La función de distribución.

Para cada función f medible no negativa y cada $t \geq 0$, definamos

$$\lambda(t) = \lambda_{f,E}(t) = m\left(\{x \in E : f(x) > t\}\right)\ .$$

Llamaremos a $\lambda(t)$ **la función de distribución** de f sobre E .

El lector debería verificar que $\lambda(t)$ es una función monótona decreciente y continua por la derecha. Además si suponemos que f es finita en casi todo punto, entonces $\lambda(t) \to 0$ cuando $t \to \infty$, a menos que la función $\lambda(t)$ sea idénticamente infinita.

Si $f \in L^p(E)$, $p < \infty$ y $\lambda(t) = \lambda_{|f|,E}(t)$ entonces tenemos

(1) $$t^p\ \lambda(t) \le \int_{\{|f|>t\}\cap E} |f|^p\ dx\ .$$

La desigualdad de arriba brinda información sobre el comportamiento de $\lambda(t)$ para $t \to +\infty$.

(7.12) **Teorema.** *Sea* $p < \infty$, $f \in L^p(E)$ *y* $\lambda(t) = \lambda_{|f|,E}(t)$. *Entonces*

(a) $t^p\ \lambda(t) \to 0$ *cuando* $t \to \infty$.

(b) $t^p\ \lambda(t) \to 0$ *cuando* $t \to 0$.

DEMOSTRACIÓN: La desigualdad (1) y el teorema de la convergencia dominada implican (a). Si la medida de E es finita la parte (b) es obvia. Para el caso general pongamos $E = E_1 \cup E_2$, donde

$$m(E_1) < \infty \qquad \text{y} \qquad \int_{E_2} |f(t)|^p\ dt < \varepsilon\ ,$$

para un $\varepsilon > 0$ dado.

Luego,

$$t^p\ \lambda(t)\ \le\ t^p m(E_1)\ + \int_{E_2} |f(t)|^p\ dt\ .$$

Q.E.D.

Ahora estamos en condiciones de expresar la integral de $|f|^p$ usando la función de distribución $\lambda(t) = \lambda_{|f|,E}(t)$. En efecto para $1 \le p < \infty$ y $x \in E$ se tiene

$$|f(x)|^p = p \int_0^{|f(x)|} t^{p-1}\ dt\ .$$

Integrando la igualdad anterior sobre E y usando el teorema de Tonelli llegamos a

(2) $$\int_E |f(x)|^p\ dx = p \int_0^\infty t^{p-1}\ \lambda(t)\ dt\ ,$$

donde $\lambda(t) = m(\{x \in E : |f(x)| > t\})$.

La integral de Lebesgue de la fórmula (2) también se puede expresar como una integral de Riemann - Stieltjes. Para ello basta integrar por partes en la integral

$$p \int_\varepsilon^B t^{p-1}\ \lambda(t)\ dt\ ,$$

haciendo luego $B \to \infty$, $\varepsilon \to 0$ y usando el Teorema (7.12), se obtiene la fórmula

$$\int_E |f(x)|^p \, dx = -\int_0^\infty t^p \, d\lambda(t) \; . \tag{3}$$

En el futuro utilizaremos asiduamente las igualdades (2) y (3).

* 7. Espacios de Orlicz.

Sea φ una función monótona creciente y continua de $[0,\infty)$ en sí mismo tal que $\varphi(0)=0$ y $\varphi(t)$ tiende a infinito si $t \to \infty$. Si E es un subconjunto medible de $I\!R^n$, consideremos el siguiente conjunto de funciones medibles:

$$L^\varphi(E) = L^\varphi = \left\{ f : \text{ existe } \lambda > 0 \text{ tal que } \int_E \varphi(\lambda|f(x)|) \, dx < \infty \right\} \; .$$

Como en el caso de L^p los elementos de L^φ son clases de funciones medibles con valores complejos o reales extendidos. En cualquier caso es fácil ver que L^φ es un espacio vectorial cerrado para las operaciones de máximo y mínimo entre dos funciones. Abreviamos esto diciendo que estamos en presencia de un **espacio vectorial reticulado**.

En la literatura también se usa el siguiente espacio

$$L^\varphi_\infty(E) = L^\varphi_\infty = \left\{ f : \text{ para todo } \lambda > 0, \; \int_E \varphi(\lambda|f(x)|) \, dx < \infty \right\} \; .$$

Fácilmente se demuestra que L^φ_∞ es un espacio vectorial reticulado incluido en L^φ.

El conjunto de funciones

$$D^\varphi = \left\{ f : \int_E \varphi(|f(x)|) \, dx < \infty \right\} \; ,$$

está comprendido entre los dos anteriores definidos, pero dado que en general no es un espacio vectorial se usa con menos frecuencia que aquellos.

Diremos que la función φ cumple una condición Δ_2 si existe una constante no negativa c tal que $\varphi(2x) \le c\varphi(x)$ para todo $x \ge 0$. Si esta desigualdad se cumple para todo x mayor que un cierto x_0, decimos que φ cumple una condición Δ_2 para valores grandes de x.

(7.13) **Teorema.** *Si φ cumple una condición Δ_2, entonces tenemos $L^\varphi_\infty = L^\varphi$. Si $m(E) < \infty$ obtenemos la igualdad entre ambos espacios con sólo pedir que se cumpla una condición Δ_2 para valores grandes de x.*

DEMOSTRACIÓN: ejercicio.

(7.14) **Teorema.** *Si φ no cumple una condición Δ_2 para valores grandes de x, existe una función f en D^φ tal que $\beta f \notin D^\varphi$ para todo $\beta > 1$.*

DEMOSTRACIÓN: Sin pérdida de generalidad supondremos $m(E) < \infty$. Dado que φ no cumple Δ_2 en el infinito, existe u_1 tal que $\varphi(u_1) > 1$ y $\varphi((1+\frac{1}{1})u_1) > 2\varphi(u_1)$. También existe $u_2 > u_1$ tal que $\varphi((1+\frac{1}{2})u_2) > 2^2\varphi(u_2)$. En general podemos elegir una sucesión $u_i \nearrow \infty$ tal que para todo i se verifica

$$\varphi\left(\left(1+\frac{1}{i}\right)u_i\right) > 2^i\varphi(u_i)\ .$$

Elijamos ahora una sucesión disjunta de conjuntos (E_i) tal que para cada i,

$$E_i \subset E, \qquad m(E_i) = m(E)/2^i\varphi(u_i)\ .$$

Pongamos ahora

$$f(x) = \begin{cases} u_i & \text{si } x \in E_i \\ 0 & \text{si } x \notin \bigcup\limits_i E_i\ . \end{cases}$$

Claramente

$$\int_E \varphi(f(x))\,dx = \sum \frac{m(E)}{2^i\varphi(u_i)}\,\varphi(u_i) = m(E) < \infty\ .$$

Si $\beta > 1$, tendremos

$$\int_E \varphi(\beta f(x))\,dx \geq \sum_{(\beta-1)i\geq 1} \frac{m(E)}{2^i\varphi(u_i)}\,\varphi\left((1+\frac{1}{i})u_i\right) = \infty\ .$$

Q.E.D.

Seguidamente introducimos una métrica en L^φ_∞ que es la análoga de $d_p(f,g) = \int_E |f-g|^p\,dx$ cuando $0 < p \leq 1$, para lo cual supondremos que φ es una función cóncava (i.e. $-\varphi$ convexa) diferenciable, además de las condiciones pedidas al principio de esta sección. En esta situación la función

φ resulta ser subaditiva sobre $[0,\infty)$ o sea $\varphi(s+t) \leq \varphi(s)+\varphi(t)$ para $s,t \geq 0$ (ver ejercicios).

Para $f,g \in L^{\varphi}_{\infty}$ ponemos

$$d_{\varphi}(f,g) = \int_E \varphi(|f-g|)\ dx\ .$$

Es fácil ver que

d_1) $d_{\varphi}(f,g) = 0$ sii $f = g$.

d_2) $d_{\varphi}(f,g) \leq d_{\varphi}(f,h) + d_{\varphi}(h,g)$.

d_3) $d_{\varphi}(f,g) = d_{\varphi}(g,f)$.

Hemos demostrado que cuando φ es cóncava, L^{φ}_{∞} es un espacio métrico con la distancia d_{φ}.

Notemos que d_{φ} no es, en general, una métrica sobre L^{φ}. Más aún, $d_{\varphi}(f,0)$ puede valer infinito si $f \notin L^{\varphi}_{\infty}$.

Para lo que resta de esta sección supondremos que $\varphi : [0,\infty) \to [0,\infty)$ es una función convexa no idénticamente nula con $\varphi(0) = 0$. Nótese que las condiciones anteriores implican que $\varphi(x)$ tiende a infinito si x tiende a infinito y además φ es estrictamente creciente desde cierto punto en adelante.

Definimos ahora

$$C_{\varphi} = \left\{ f \in L^{\varphi} : \int_E \varphi(|f|)\ dx \leq 1 \right\}\ .$$

(7.14) **Teorema.** *Se verifican las siguientes propiedades:*

(1) C_{φ} *es convexo y* $0 \in C_{\varphi}$ *.*

(2) C_{φ} *es simétrico respecto al origen :* $f \in C_{\varphi}$ *implica* $-f \in C_{\varphi}$ *.*

(3) Si $f \in L^{\varphi}$ *existe* $\varepsilon > 0$ *tal que* $\varepsilon f \in C_{\varphi}$*, i.e.* C_{φ} ***absorbe*** *cualquier elemento de* L^{φ} *.*

(4) C_{φ} *no contiene ninguna recta que pase por el origen.*

Demostración: Las dos primeras propiedades son triviales. La propiedad (3) es consecuencia de la continuidad de φ, del teorema de convergencia dominada de Lebesgue y del hecho que $\varphi(0) = 0$.

Ahora, si f es una función no nula sabemos que para cierto $a > 0$ y un conjunto A medible de medida positiva, $|f| \geq a\ \chi_A$. Luego,

$$\int_E \varphi(\lambda |f|)\ dx \geq \varphi(\lambda a)\ m(A)$$

y recordando que $\varphi(t)$ tiende a infinito cuando $t \to \infty$, obtenemos la demostración de (4).

Q.E.D.

En forma análoga a lo realizado en el Cap.II, Sección 14 podemos ver que la siguiente funcional de Minkowski

$$\|f\|_\varphi = \inf\left\{\lambda > 0 : \frac{f}{\lambda} \in C_\varphi\right\} ,$$

define una norma sobre L^φ. Esta norma se conoce con el nombre de **norma de Luxemburg**.

Note (use el teorema de Fatou) que si f es no nula, tenemos

$$\int_E \varphi\left(\frac{|f|}{\|f\|_\varphi}\right)\, dx \leq 1 .$$

Así el ínfimo que define $\|f\|_\varphi$ es en realidad un mínimo. Esto nos permite dar una rápida demostración de la desigualdad triangular. En efecto, si $\|f\|_\varphi \cdot \|g\|_\varphi > 0$ tendremos

$$\frac{f+g}{\|f\|_\varphi + \|g\|_\varphi} = \frac{\|f\|_\varphi}{\|f\|_\varphi + \|g\|_\varphi}\,\frac{f}{\|f\|_\varphi} + \frac{\|g\|_\varphi}{\|f\|_\varphi + \|g\|_\varphi}\,\frac{g}{\|g\|_\varphi} \in C_\varphi ,$$

o sea $\|f+g\|_\varphi \leq \|f\|_\varphi + \|g\|_\varphi$.

(7.15) *Si φ cumple una condición Δ_2, para cada $f \neq 0$ tenemos la igualdad*

$$\int_E \varphi\left(\frac{|f|}{\|f\|_\varphi}\right)\, dx = 1 .$$

Si φ no cumple la propiedad Δ_2 no podemos asegurar la igualdad en (7.15). No obstante tenemos

(7.16) $$\lambda \geq \|f\|_\varphi \text{ si y sólo si } \int_E \varphi\left(\frac{|f|}{\lambda}\right)\, dx \leq 1 .$$

Para estudiar la convergencia en norma es útil tener en cuenta el siguiente resultado.

(7.17) **Teorema.** *Para una sucesión* (f_i) *en* L^φ *son equivalentes las afirmaciones*

(1) $\|f_i\|_\varphi \to 0$ *cuando* $i \to \infty$ *;*

(2) *para cada* $\lambda > 0$ $\int_E \varphi(\lambda |f_i|)\, dx \to 0$ *cuando* $i \to \infty$ *.*

DEMOSTRACIÓN: Veamos que (1) implica (2). Dado que $\varphi(tx) \le t\ \varphi(x)$, $0 \le t \le 1$, $x \ge 0$, para $\lambda \|f_i\|_\varphi \le 1$ y f_i no nula tenemos las estimaciones

$$\begin{aligned}\int_E \varphi(\lambda |f_i(x)|)\, dx &= \int_E \varphi\left(\lambda \|f_i\|_\varphi \frac{|f_i(x)|}{\|f_i\|_\varphi}\right) dx \\ &\le \lambda \|f_i\|_\varphi \int_E \varphi\left(\frac{|f_i(x)|}{\|f_i\|_\varphi}\right) dx \\ &\le \lambda\ \|f_i\|_\varphi\ .\end{aligned}$$

Supongamos ahora que (2) es cierta. Si tuviéramos para una subsucesión $\big(f_{i_j}\big)$ y un número $a > 0$ las desigualdades $\|f_{i_j}\|_\varphi > a$, tendríamos por (7.16)

$$\int_E \varphi\left(\frac{|f_{i_j}|}{a}\right) dx > 1 \text{ para todo } j\ .$$

Q.E.D.

(7.18) **Teorema.** *Consideremos una sucesión* (f_i) *en* L^φ *que converge puntualmente a una función* f *. Entonces* $\|f\|_\varphi \le \liminf_{i\to\infty} \|f_i\|_\varphi$ *.*

DEMOSTRACIÓN: Sea $\alpha = \liminf_{i\to\infty} \|f_i\|_\varphi$ y supongamos $\alpha > 0$, si $\alpha = 0$ se puede demostrar que $f = 0$. Luego existe una subsucesión (f_{i_j}) tal que $0 < \|f_{i_j}\|_\varphi \to \alpha$ si $j \to \infty$. Usando el teorema de Fatou, tenemos

$$\begin{aligned}\int \varphi\left(\frac{|f|}{\alpha}\right) dx &= \int \lim_{j\to\infty} \varphi\left(\frac{|f_{i_j}|}{\|f_{i_j}\|_\varphi}\right) dx \\ &\le \liminf_{j\to\infty} \int \varphi\left(\frac{|f_{i_j}|}{\|f_{i_j}\|_\varphi}\right) dx\ \le 1\ .\end{aligned}$$

O sea hemos demostrado $\|f\|_\varphi \le \alpha$.

Q.E.D.

(7.19) **Teorema.** *El espacio* L^φ *con la norma de Luxemburg* $\|\cdot\|_\varphi$ *es un espacio de Banach.*

DEMOSTRACIÓN: Hemos visto que $\|\cdot\|_\varphi$ es una norma y sólo resta demostrar que el espacio es completo. Sea (f_k) una sucesión en L^φ tal que $\sum_k \|f_k\|_\varphi < \infty$ y pongamos $f^* = \sum_k |f_k|$. Por (7.18) tenemos que $f^* \in L^\varphi$ y por lo tanto la función

$$f = \sum f_k \ ,$$

está bien definida en casi todo punto y pertenece a L^φ, puesto que $|f| \le f^*$. Veamos ahora que $h_i = \sum_{k=1}^i f_k$ converge a f en norma $\|\cdot\|_\varphi$. En efecto,

$$\|f - h_i\|_\varphi = \left\| \sum_{i+1}^\infty f_k \right\|_\varphi \le \left\| \sum_{n+1}^\infty |f_k| \ \right\|_\varphi \le \sum_{i+1}^\infty \|f_k\|_\varphi \ ,$$

donde la última desigualdad es nuevamente consecuencia de (7.18). Como $\sum_{i+1} \|f_k\|_\varphi \to 0$ cuando $i \to \infty$, se sigue que el espacio L^φ es completo.

Q.E.D.

Una consecuencia directa del teorema de convergencia dominada de Lebesgue es que si $f \in L^p$, $p < \infty$, entonces $\|f \ \chi_A\|_p$ tiende a cero cuando $m(A)$ tiende a cero.

Diremos que una función $f \in L^\varphi$ tiene **norma $\|\cdot\|_\varphi$ absolutamente continua** si $\|f \ \chi_A\|_\varphi \to 0$ cuando $m(A) \to 0$ (para abreviar escribiremos que f tiene n.a.c.).

A continuación exploramos brevemente el concepto de n.a.c. y veremos el papel que desempeña L^φ_∞ en relación con este concepto.

(7.20) **Teorema.** *Cada $f \in L^\varphi_\infty$ tiene norma absolutamente continua. Si $f \in L^\varphi$ tiene n.a.c., entonces para cada conjunto medible A de medida finita, $f \ \chi_A \in L^\varphi_\infty$.*

DEMOSTRACIÓN: Observemos primero que si $f \in L^\varphi_\infty$ esta función puede ser aproximada en norma $\|\cdot\|_\varphi$ por funciones acotadas con soporte acotado. En efecto, sea (A_i) una sucesión creciente de conjuntos acotados tal que $E = \bigcup A_i$, donde E es el conjunto medible sobre el cual está definido el espacio de Orlicz L^φ. Para cada $f \in L^\varphi_\infty$ pongamos

$$f_i = \chi_{A_i} \ f \ \chi_{\{|f| \le i\}} \ .$$

La sucesión (f_i) converge puntualmente hacia f y por otro lado $f_i \in L^\infty$ y es nula fuera de A_i. Por ser $f \in L^\varphi_\infty$ tenemos que para cada $\lambda > 0$,

$$\int_E \varphi(\lambda |f_i - f|) \ dx \to 0 \ ,$$

para $i \to \infty$. Asi que por (7.17) vale $\|f - f_i\|_\varphi \to 0$ si $i \to \infty$.

Sea ahora $f \in L^\varphi_\infty$ y $\varepsilon > 0$ y tomemos $g \in L^\infty$ tal que $\|f - g\|_\varphi < \varepsilon$. Luego

$$\|f \ \chi_A\|_\varphi \leq \varepsilon + \|g \ \chi_A\|_\varphi \leq \varepsilon + \|g\|_\infty \ \|\chi_A\|_\varphi \ .$$

Pero (ver ejercicios)

$$\|\chi_A\|_\varphi = 1/\varphi^{-1}(1/m(A)) \ .$$

Con lo que se ha demostrado la primera parte del teorema.

Recíprocamente, sea $f \in L^\varphi$ con n.a.c. y A un conjunto medible de medida finita. Sean $B_i = \{x \in A : |f(x)| \leq i\}$, $A_i = A - B_i$ y $f_i = f \ \chi_{B_i}$. Como $m(A_i) \to 0$ cuando $i \to \infty$, tenemos

$$\|f \ \chi_A - f_i\|_\varphi = \|f \ \chi_{A_i}\|_\varphi \to 0 \qquad \text{si } i \to \infty \ .$$

Puesto que f_i está en L^φ_∞ y este último espacio es cerrado en L^φ (ejercicio 35), hemos demostrado que $f \ \chi_A \in L^\varphi_\infty$.

Q.E.D.

Inspirados por la demostración del teorema precedente definiremos la clase B^φ como la adherencia en L^φ del conjunto

$$\{f \in L^\infty : m(\text{ soporte } f) < \infty\} \ .$$

Con esta notación tenemos

$$L^\varphi_\infty = B^\varphi \tag{7.21}$$

La demostración de la afirmación (7.21) está incluida en la demostración del teorema (7.20).

No es difícil ver que si $\varphi : [0, \infty) \to [0, \infty)$ es una función convexa no nula tal que $\varphi(0) = 0$, existen funciones no acotadas en el espacio de Orlicz L^φ. Así , si queremos obtener L^∞ como en un caso particular de espacio de Orlicz debemos modificar nuestro concepto de función convexa permitiéndole tomar valores infinitos. En consecuencia definimos

$$\varphi_\infty(t) = \begin{cases} 0 & \text{si } 0 \leq t \leq 1 \\ \infty & \text{si } t > 1 \ . \end{cases}$$

Claramente tenemos $\int_E \varphi_\infty(|f|)\, dx < \infty$ si y sólo si $|f(x)| \leq 1$ en casi todo punto $x \in E$. Por lo tanto $L^\varphi_\infty = L^\infty$ y $\|f\|_{\varphi_\infty} = \|f\|_\infty$. Nótese que para verificar estas últimas afirmaciones, es necesario usar el convenio $0 \cdot \infty = 0$.

EJERCICIOS

1. Sea $(X, \|\cdot\|)$ un espacio normado y (x_n) una sucesión en X, diremos que la serie formal $\sum_{1\leq n} x_n$ es **absolutamente convergente** si la serie numérica $\sum_{1\leq n} \|x_n\|$ es convergente.
   Demuestre que un espacio normado es completo si y sólo si toda serie absolutamente convergente es convergente.

2. Dé un ejemplo de una función f en $L^1([0,1])$ tal que para todo $p > 1$ $|f|^p \notin L^1([0,1])$.

3. El espacio $L^1(\mathbb{R}^n)$ es separable, o sea tiene un conjunto denso numerable.

4. Probar que existe una sucesión monótona creciente de funciones no negativas $\varphi_n \in C_0^\infty(\mathbb{R})$ que converge puntualmente a la función característica del intervalo abierto $(0,1)$.
   Sugerencia: mostrar que la función
   $$\psi(t) = \begin{cases} e^{-1/t} & \text{si } t > 0 \\ 0 & \text{si } t \leq 0 \end{cases}$$
   es infinitamente diferenciable y tiende a uno cuando $t \to \infty$. Considerar la sucesión $\varphi_n(t) = \psi(nt)\psi(n(1-t))$. Demuestre que $C_0^\infty(\mathbb{R})$ es denso en $L^1(\mathbb{R})$.

5. Demuestre el siguiente resultado, conocido como el **teorema de Riemann - Lebesgue**: si $f \in L^1(\mathbb{R})$, entonces
   $$\lim_{n\to\infty} \int_{-\infty}^{\infty} f(x) \cos nx\, dx = \lim_{n\to\infty} \int_{-\infty}^{\infty} f(x) \operatorname{sen} nx\, dx = 0.$$
   Sugerencia: demostrar el resultado para una función escalera y luego usar el teorema (7.2).

6. La función $\operatorname{sen} x/x$ no es integrable Lebesgue en $(0,\infty)$. Observe que el $\lim_{n\to\infty} \int_0^n (\operatorname{sen} x)/x \, dx$ existe.

7. El espacio $L^\infty(0,1)$ con la norma $\|\cdot\|_\infty$ es un espacio de Banach no separable.

8. Para $0 < p \leq 1$ y $f, g \in L^p(E)$ ponemos $d_p(f,g) = \int_E |f-g|^p \, dx$. Entonces $(L^p(E), d_p)$ es un espacio métrico completo. Note que L^p es un espacio vectorial.

9. Sea $\mathcal{M}(E)$ el conjunto de las funciones medibles definidas sobre E, suponemos $0 < m(E) < \infty$. Para $f, g \in \mathcal{M}(E)$ sea $d(f,g) = \int_E |f-g|/(1+|f-g|) \, dx$. Entonces $(\mathcal{M}(E), d)$ es un espacio métrico completo.
   Sugerencia: demuestre primero que $d(f_n, f) \to 0$ si y sólo si $f_n \to f$ en medida. También se puede demostrar este ejercicio sin recurrir al concepto de convergencia en medida, recuerde que hemos demostrado la completitud de L^1 de dos formas diferentes.

10. Sean $1 \leq p < \infty$ y $f \in L^p(\mathbb{R}^n)$. Para cada $t \geq 0$ definimos la siguiente función

$$\omega_p(f,t) = \sup_{|x|\leq t} \left(\int_{\mathbb{R}^n} |f(x+y) - f(y)|^p \, dy \right)^{1/p} .$$

   Demuestre que $\omega_p(f,t) \to 0$ cuando $t \to 0$. Es cierto el resultado para $p = \infty$?
   Sugerencia: La afirmación anterior es fácil de demostrar cuando f es una función continua con soporte compacto. Recuerde que la clase de estas funciones es densa en L^p y use además las desigualdades siguientes:

$$\begin{aligned} &\omega_p(f,t) \leq 2\|f\|_p , \\ &\omega_p(f,t) \leq \omega_p(f-g,t) + \omega_p(g,t) . \end{aligned}$$

11. Demuestre la desigualdad de Minkowski usando la desigualdad de Hölder.

12. Si $m(E) < \infty$ y $1 \leq p \leq q \leq \infty$, entonces $L^q(E) \subset L^p(E)$.
    Sugerencia: Sea $q < \infty$, luego use Jensen para verificar que

$$\|f/|E|^{1/p}\|_p \leq \|f/|E|^{1/q}\|_q .$$

    También se puede demostrar este ejercicio usando la desigualdad de Hölder.

13. Para $r \leq p \leq s$ se tiene que $L^p(E) \subset L^r(E) + L^s(E)$, donde el conjunto de la derecha es $\{f + g : f \in L^r(E) \text{ y } g \in L^s(E)\}$.

14. Si $m(E) < \infty$ y $f \in L^\infty(E)$, entonces $\|f\|_p \to \|f\|_\infty$ cuando $p \to \infty$.
Sugerencia: La versión discreta de la afirmación anterior es muy fácil de demostrar. En efecto sea

$$f_p(x) = f_p(x_1, \ldots, x_n) = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p} , \qquad \text{y}$$

$$|x_{i_0}| = \max_{1 \leq i \leq n} |x_i| \ .$$

Luego $f_p(x) = |x_{i_0}| \ f_p\left(\frac{1}{|x_{i_0}|} x\right)$ y tenemos $f_p\left(\frac{1}{|x_{i_0}|} x\right) \to 1$ para $p \to \infty$.
Así la afirmación es cierta para funciones simples. Para el caso general use el hecho de que las funciones simples son densas en $L^\infty(E)$.¿Podría dar una demostración alternativa sin pasar por el caso discreto?

15. Sean $f \in L^p(\mathbb{R}^n)$ y $g \in L^{p'}(\mathbb{R}^n)$, $1/p + 1/p' = 1$, entonces la convolución $f * g$ es una función acotada uniformemente continua.
Sugerencia: Use el ejercicio 10.

16. Demuestre la siguiente desigualdad conocida como la desigualdad de Minkowski para integrales.

$$\left(\int_F \left(\int_E f(x,y)\ dx \right)^p dy \right)^{1/p} \leq \int_E \left(\int_F f^p(x,y)\ dy \right)^{1/p} dx \ ,$$

donde f es una función medible no negativa, $E \subset \mathbb{R}^n$, $F \subset \mathbb{R}^m$ y $1 \leq p < \infty$.
Sugerencia: Use el teorema (7,10) y el teorema de Tonelli.

17. Si $f \in L^1(\mathbb{R}^n)$ y $g \in L^p(\mathbb{R}^n)$, $1 \leq p \leq \infty$. Entonces $f * g \in L^p(\mathbb{R}^n)$ y se verifica $\|f * g\|_p \leq \|f\|_1 \|g\|_p$.
Sugerencia: Use la desigualdad de Minkowski para integrales.

18. Sean X e Y espacios normados y T una transformación lineal de X en Y. Las siguientes afirmaciones son equivalentes
   (i) Existe un número real no negativo M tal que

$$\|Tx\| \leq M\|x\| \ , \qquad x \in X \ .$$

(ii) T es continua en $x = 0$.

iii) T es continua sobre X.

19. Sea T una transformación lineal continua de un espacio normado X en un espacio normado Y. Pongamos

$$\|T\| = \sup_{\|x\| \leq 1} \|Tx\|$$

(a) La expresión anterior define una norma sobre el espacio vectorial $L(X, Y)$ de todas las transformaciones lineales continuas de X en Y.
(b) Si Y es un espacio de Banach también lo es $L(X, Y)$.
(c) Las siguientes son expresiones alternativas para $\|T\|$, es decir iguales a ésta.
(c_1) $\sup\{\|Tx\| : \|x\| = 1\}$
(c_2) $\sup\{\|Tx\|/\|x\| : x \neq 0\}$
(c_3) $\inf\{M : \|Tx\| \leq M\|x\| \ , \quad x \in X\}$.

20. Sean $K \in L^2(I\!\!R^n \times I\!\!R^n)$, $f \in L^2(I\!\!R^n)$ y definimos

$$T\ f(x) = \int_{I\!\!R^n} K(x, y) f(y)\ dy\ .$$

La función T es una transformación lineal continua de L^2 y se verifica $\|T\| \leq \|K\|_2$.

21. Sea $C_0(I\!\!R^n)$ el espacio de las funciones continuas con soporte compacto definidas sobre $I\!\!R^n$. El espacio normado $(C_0(I\!\!R^n), \|\ \|_\infty)$ no es completo. El espacio anterior es denso en el espacio de las funciones continuas que se anulan en el infinito (funciones $f(x)$ que tienden a cero cuando $|x| \to \infty$).

22. Si $f \in L^1(I\!\!R^n)$ definimos la **transformada de Fourier** de f como

$$\mathcal{F}(f)(y) = (2\pi)^{-n/2} \int_{I\!\!R^n} e^{ix \cdot y}\ f(x)\ dx$$

$\mathcal{F}$ es una transformación lineal continua de $L^1(I\!\!R^n)$ en el espacio de las funciones continuas que se anulan en el infinito.

23. Sea φ una función continua de $I\!\!R$ en $I\!\!R$ que cumple $2\varphi(s + t) \leq \varphi(2s) + \varphi(2t)$. Entonces φ es una función convexa.

24. Sea $\varphi : [0, \infty) \to I\!R$ tal que $\varphi(0) = 0$. Si φ es convexa y $0 \leq \lambda \leq 1$ tenemos que $\varphi(\lambda t) \leq \lambda \varphi(t)$, $t \in I\!R$. Si φ es cóncava, i.e. $-\varphi$ convexa, y $\lambda \geq 1$ se verifica $\varphi(\lambda t) \leq \lambda \varphi(t)$, $t \in I\!R$.

25. Sea $\varphi : [0, \infty) \to I\!R$, $\varphi(0) = 0$ una función cóncava y derivable. Entonces $\varphi(s + t) \leq \varphi(s) + \varphi(t)$ para $s, t \geq 0$.
Sugerencia: Use el ejercicio anterior para verificar que la desigualdad se cumple para $s = t$, recuerde luego que φ' es una función monótona decreciente. Muestre que la hipótesis de derivabilidad puede omitirse.

26. Sea $\varphi : [0, \infty) \to [0, \infty)$, $\varphi(0) = 0$, una función cóncava derivable. Supongamos además $\varphi(t) > 0$ si $t > 0$. Definimos $L^{\varphi}(E)$ como el conjunto de las funciones medibles f tal que $\int_E \varphi(|f|)\, dx$ es finita.
Para $f, g \in L^{\varphi}(E)$ ponemos
$$d_{\varphi}(f, g) = \int_E \varphi(|f - g|)\, dx\ .$$
Pruebe que $(L^{\varphi}(E), d_{\varphi})$ es un espacio métrico completo. Nótese que $L^{\varphi}(E)$ es también un espacio vectorial.

27. Sea H un espacio de Hilbert, C un conjunto convexo cerrado en H y $p(x) = p(x/C)$ la proyección de un vector x sobre C. La función $p : H \to C$ es lineal sii C es un subespacio.

28. Demuestre el Teorema (7.11) para $1 \leq p \leq 2$.
Sugerencia: Suponga que el espacio E tiene medida finita y con la desigualdad de Hölder reduzca el problema al caso $p = 2$. Véase la demostración del Teorema (7.3) dada el final del parágrafo 3.

29. Si $\|f - f_n\|_p \to 0$, entonces la sucesión f_n converge en medida hacia f.

30. Teorema de convergencia de Vitali. Sea (f_n) una sucesión en $L^p(E)$, $0 < p < \infty$, que converge en casi todo punto hacia una función f la cual es finita en casi todo punto. Entonces $f \in L^p(E)$ y $\|f - f_n\|_p \to 0$ cuando $n \to \infty$ si y sólo si se cumple:
(i) Para cada $\varepsilon > 0$ existe un conjunto A de medida finita tal que para todo n vale
$$\int_{E-A} |f_n(x)|^p\, dx \leq \varepsilon\ .$$
(ii) Dado $\varepsilon > 0$ existe $\delta > 0$ tal que si $|A| < \delta$ entonces
$$\int_A |f_n(x)|^p\, dx < \varepsilon\ ,$$

para todo n.
Sugerencia: Es más fácil ver que las anteriores son condiciones necesarias. Para demostrar que son suficientes si $1 \leq p < \infty$, primero use Fatou y Minkowski para reducir al caso $m(E) < \infty$, y lluego use Egorov.

31. Sea $\varphi : [0, \infty) \to [0, \infty)$ una función continua creciente $\varphi(0) = 0$ tal que para una constante finita c se cumple $\varphi(s + t) \leq c(\varphi(s) + \varphi(t))$ para todo par de números no negativos s, t. Sea (f_n) una sucesión de funciones medibles que converge en casi todo punto hacia f, y supongamos que $\int_E \varphi(|f|)\ dx$ es finita. Si $\int_E \varphi(|f_n|)\ dx \to \int_E \varphi(|f|)\ dx$ cuando $n \to \infty$, entonces $\int_E \varphi(|f - f_n|)\ dx \to 0$ si $n \to \infty$.
Sugerencia: Use Fatou para la sucesión $c(\varphi(|f|) + \varphi(|f_n|)) - \varphi(|f_n - f|)$.

32. Demuestre que un espacio de Banach real es un espacio de Hilbert si y sólo si para cualquier par de vectores x, y se cumple la "ley del paralelogramo" : $\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$. Extienda el resultado para espacios de Banach complejos.

33. Sea $1 \leq p < \infty$ y supongamos que para todo par de funciones $f, g \in L^p$ se cumple la ley del paralelogramo: $\|f+g\|_p^2 + \|f-g\|_p^2 = 2(\|f\|_p^2 + \|g\|_p^2)$. Demuestre que $p = 2$.

34. Sea $\varphi(t)$ una función creciente continua sobre la semirrecta $0 \leq t < \infty$, con derivada continua para $t > 0$ y tal que $\varphi(0) = 0$. Probar que para cualquier función medible no negativa $f : I\!R^n \to I\!R$ y cualquier conjunto medible $E \subset I\!R^n$,

$$\int_E \varphi(f(x))dx = \int_0^\infty m(E \cap \{f > t\})\varphi'(t)dt.$$

35. L^φ_∞ es un subespacio cerrado de L^φ.

36. Si E tiene medida finita verifique $L^\infty(E) \subset L^\varphi(E) \subseteq L^1(E)$. Observe que la primera inclusión es propia.

37. Sea A un conjunto de medida finita tal que $\varphi(m(A)) > 0$. Demuestre que

$$\|\chi_A\|_\varphi = 1/\varphi^{-1}(1/m(A))$$

CAPITULO VIII

DIFERENCIACION DE LA INTEGRAL

1. La función maximal de Hardy-Littlewood.

En este capítulo consideraremos funciones con valores en la recta real extendida. Para una función **localmente integrable** f (integrable sobre cualquier cubo $Q \subset I\!\!R^n$), definimos el **promedio** de f sobre Q por medio de la fórmula

$$f_Q = \frac{1}{m(Q)} \int_Q f(x)\ dx\ .$$

Cuando f es continua en un punto x_0, tenemos que

$$\lim_{m(Q)\to 0} f_Q = f(x_0)\ ,$$

donde se han considerado aquellos cubos que contienen al punto x_0.

El primer objetivo de este capítulo será analizar la validez del límite anterior cuando la función f es integrable. A lo sumo podemos esperar un resultado válido en casi todo punto, ya que para f y g iguales en casi todo punto tenemos $f_Q = g_Q$ para cada cubo Q.

A continuación definimos una función que mayora puntualmente a todos los promedios. Para f localmente integrable ponemos

$$Mf(x) = \sup_Q\ |f|_Q = \sup_Q\ \frac{1}{m(Q)} \int_Q |f(y)|\ dy\ ,$$

donde el supremo se toma sobre todos los cubos que contienen al punto x.

La función Mf recibe el nombre de **función maximal de Hardy-Littlewood**. Veamos que Mf es semicontinua inferiormente y en consecuencia medible.

En efecto, para cada número real positivo t definimos:

$$E_t = \{x \in \mathbb{R}^n : Mf(x) > t\} \ .$$

El conjunto E_t es abierto puesto que la integral de Lebesgue es absolutamente continua.

Hemos definido un operador M sobre las funciones localmente integrables con valores en las funciones medibles. Este operador es **sublineal**, lo que significa que:

$$M(af + bg)(x) \leq |a|Mf(x) + |b|Mg(x) \ ,$$

para cada par f, g de funciones, números a, b cualesquiera y cada $x \in \mathbb{R}^n$.

El operador M transforma el espacio L^∞ en sí mismo, más precisamente tenemos $\|Mf\|_\infty \leq \|f\|_\infty$. Por otro lado M no preserva la integrabilidad de funciones. En efecto, $Mf \notin L^1(\mathbb{R}^n)$ si $f \neq 0$ (ver ejercicios).

La integrabilidad de f no implica que Mf sea localmente integrable, como surgirá de la siguiente consideración.

Sea f una función no negativa definida sobre la recta real. Entonces

$$Mf(x) \geq \frac{1}{x} \int_0^x f(y)\ dy \qquad\qquad (x > 0) \ .$$

Integrando la expresión anterior sobre el intervalo $[0, 1]$ y usando el teorema de Tonelli, obtenemos:

$$\int_0^1 \log(1/x)\ f(x)\ dx \leq \int_0^1 Mf(x)\ dx \ .$$

En consecuencia Mf no será integrable sobre $[0, 1]$ si $\log(1/x)\ f(x)$ no lo es en dicho intervalo. Por otra parte, existen funciones f integrables en $[0, 1]$, tales que $\log(1/x)\ f \notin L^1[0, 1]$.

Observemos que si a una función integrable f le imponemos que $\log(1/x)f \in L^1[0, \frac{1}{2})$, solamente estamos pidiendo una condición extra de integrabilidad en la cercanía del punto $x = 0$. Condiciones semejantes para f, pero globales, están dadas por expresiones del tipo

$$\int_0^1 |f(x)| \, \log^+ |f(x)| \, dx < \infty ,$$

donde $\log^+ t$ es cero si $0 \leq t \leq 1$ y vale $\log t$ para $t \geq 1$.

Si para una función f la integral anterior es finita, diremos que f pertenece a la clase $L \log^+ L[0,1]$.

El lector puede verificar que si $f \in L \log^+ L[0,1]$, entonces $f(x) \, \log(1/x)$ es integrable a $[0,1]$. La recíproca de esta afirmación es cierta si f es no negativa y decreciente.

Más adelante veremos que si una función f está en la clase $L \log^+ L$, entonces su función maximal es localmente integrable.

El siguiente teorema es de suma importancia en la teoría de diferenciación de la integral.

(8.1) **Teorema de Hardy-Littlewood.** *Para cada* $f \in L^1(\mathbb{R}^n)$ *y* $t > 0$ *se tiene*

$$m(\{x : Mf(x) > t\}) \leq \frac{c}{t} \int_{\mathbb{R}^n} |f(x)| \, dx ,$$

donde c *es una constante finita que no depende ni de* f *ni de* t.

El teorema anterior se obtendrá fácilmente usando el siguiente lema de cubrimiento.

(8.2) **Lema de cubrimiento.** *Sea* K *un subconjunto compacto de* $\mathbb{R}^n$ *y* $\mathcal{C}$ *un cubrimiento de* K *por cubos abiertos. Entonces se puede seleccionar de* $\mathcal{C}$ *una familia finita disjunta de cubos* $Q_1, Q_2, \ldots, Q_m$ *tal que*

$$K \subset \bigcup_{i=1}^{m} 3Q_i ,$$

donde $3Q_i$ *denota el cubo concéntrico con* Q_i *cuyo lado tiene tres veces la longitud del lado de* Q_i.

DEMOSTRACIÓN DEL LEMA: Por el teorema de Heine-Borel podemos seleccionar cubos $I_1, I_2, \ldots, I_\ell$ de la clase $\mathcal{C}$ cuya unión cubre a K. Denotemos por Q_1 un cubo I_j que tenga longitud de lado máxima entre los ℓ cubos anteriores.

Observemos que si un cubo I_j tiene intersección no vacía con el cubo Q_1, entonces $I_j \subset 3Q_1$. Así nos quedamos con los cubos I_j que no intersecan a Q_1 y dentro de ellos elegimos como Q_2 a un cubo de longitud de lado máxima.

Repitiendo este proceso obtenemos una familia $Q_1, \ldots, Q_m$, $m \leq \ell$ de cubos disjuntos con la propiedad requerida por el lema.

Q.E.D.

DEMOSTRACIÓN DEL TEOREMA DE HARDY-LITTLEWOOD: Sea K un subconjunto compacto del conjunto $\{x \in \mathbb{R}^n : Mf(x) > t\}$. Para cada $x \in K$ elegimos un cubo abierto Q que contenga a x, tal que el promedio $|f|_Q$ sea mayor que t.

Teniendo en cuenta la compacidad de K y el lema de cubrimiento (8.2), existen cubos disjuntos $Q_1, \ldots, Q_m$ tal que:

$$K \subset \bigcup_{i=1}^{m} 3Q_i \ , \qquad |f|_{Q_i} > t \qquad (i = 1, \ldots, m).$$

Ahora obtenemos fácilmente una estimación para $m(K)$. En efecto,

$$m(K) \leq \sum_{i=1}^{m} m(3Q_i) = 3^n \sum_{i=1}^{m} m(Q_i) \leq$$
$$\frac{3^n}{t} \sum_{i=1}^{m} \int_{Q_i} |f(x)| \, dx = \frac{3^n}{t} \int_{\bigcup_1^m Q_1} |f(x)| \, dx \leq \frac{3^n}{t} \|f\|_1 \ .$$

Usando la regularidad de la medida de Lebesgue (ver 3.31), obtenemos la afirmación del teorema con $c = 3^n$.

Q.E.D.

Diremos que una función medible f pertenece a $\boldsymbol{L^1(\mathbb{R}^n)}$ **débil** si existe una constante $c(f)$ tal que $t\lambda_{|f|}(t) \leq c(f)$ para cada $t > 0$, donde $\lambda_{|f|}(t) = m(\{x : |f(x)| > t\})$.

Por la desigualdad de Chebyshev se tiene que si f es integrable, pertenece a $L^1(\mathbb{R}^n)$ débil. La afirmación recíproca no es cierta, por ejemplo tenemos que sobre la recta vale $t\lambda_{\frac{1}{|x|}}(t) = 2$.

El teorema (8.1) nos dice que la función maximal Mf pertenece a $L^1(\mathbb{R}^n)$ débil cuando f es una función integrable.

2. El teorema de diferenciación de Lebesgue.

Nos interesa ver bajo qué condiciones se cumple la igualdad

$$\lim_{Q\to x} f_Q = f(x) \ , \tag{1}$$

donde el límite se toma sobre la familia de los cubos Q que contienen al punto x cuando $m(Q)\to 0$.

Si f es la función característica de un intervalo I, fácilmente se ve que la igualdad (1) se cumple para $x\in I^\circ$ ó $x\notin \overline{I}$. Luego dicha relación se cumple en casi todo punto si f is una función escalonada. Por otro lado, la relación (1) se cumple si x es un punto de continuidad para la función f. Por lo tanto tenemos dos clases densas en L^1 (las funciones escalonadas y las funciones continuas con soporte compacto) para las cuales la igualdad (1) se cumple en todo punto o bien al menos en casi todo punto.

(8.3) **Teorema de diferenciación de Lebesgue.** *Sea f una función localmente integrable en $\mathbb{R}^n$. Entonces los promedios*

$$\frac{1}{m(Q)}\int_Q |f(y)-f(x)|\ dy$$

tienden a cero cuando $Q\to x$, para casi todo $x\in\mathbb{R}^n$.

DEMOSTRACIÓN: Podemos suponer, sin pérdida de generalidad, que $f\in L^1(\mathbb{R}^n)$. Definamos ahora el siguiente operador sobre $L^1(\mathbb{R}^n)$:

$$\Gamma f(x) = \limsup_{Q\to x} \frac{1}{m(Q)}\int_Q |f(y)-f(x)|\ dy \ ,$$

o más precisamente

$$\Gamma f(x) = \inf_{\delta>0}\ \sup\left\{\frac{1}{m(Q)}\int_Q |f(y)-f(x)|\ dy\ :\ m(Q)\le\delta,\ Q\ni x\right\}\ .$$

Debemos demostrar que Γf es cero en casi todo punto, o lo que es equivalente veremos que para todo $t>0$ el siguiente conjunto tiene medida cero

$$E_t(f) = \{x:\ \Gamma f(x) > t\}\ .$$

Dado que Γ es un operador sublineal y que $\Gamma g(x)$ es cero en casi todo punto siempre que g sea una función escalonada, tenemos que $m_e(E_t(f)) \leq m_e(E_t(f-g))$.

Puesto que $\Gamma f \leq Mf + |f|$, obtenemos

$$\begin{aligned} m_e(E_t(f)) &\leq m(\{M(f-g) + |f-g| > t\}) \\ &\leq m(\{M(f-g) > t/2\}) + m(\{|f-g| > t/2\}) \ . \end{aligned}$$

Usando el teorema de Hardy-Littlewood y la desigualdad de Chebyshev, obtenemos

$$m_e(E_t(f)) \leq \frac{2}{t}(c+1) \ \|f-g\|_1 \ ,$$

para cada g en la clase de las funciones escalonadas. Puesto que dicha clase es densa en $L^1(\mathbb{R}^n)$, tenemos que $m(E_t(f)) = 0$.

Q.E.D.

Los promedios f_Q en el teorema de diferenciación de Lebesgue se consideraron sobre cubos Q de aristas paralelas a los ejes coordenados. A continuación extenderemos el teorema para promedios sobre conjuntos más generales.

Supongamos que para cada $x \in \mathbb{R}^n$ se ha definido una clase $\mathcal{C}_x$ de conjuntos medibles, de modo que se cumplen las condiciones siguientes:

(a) Para cada $E \in \mathcal{C}_x$ existe un cubo Q tal que $x \in Q$, $Q \supseteq E$ y $m(Q) \leq c\, m(E)$, siendo c una constante independiente de x y E.

(b) Para cada $\varepsilon > 0$ y $x \in \mathbb{R}^n$ existe $E \in \mathcal{C}_x$ tal que $m(E) < \varepsilon$.

Un ejemplo de clase $\mathcal{C}_x$ con las condiciones (a) y (b) es el de una familia de bolas $B_r(x) = \{y \in \mathbb{R}^n : |x-y| < r\}$ con radios r tendiendo a cero.

Si A es un conjunto medible acotado de medida positiva ponemos $\mathcal{C}_x = \{x + rA : r > 0\}$, con $rA = \{ra : a \in A\}$. Es fácil ver que esta clase $\mathcal{C}_x$ cumple (a) y (b).

Dada una clase $\mathcal{C}_x$ nos interesa nuevamente averiguar si se verifica la igualdad

$$\lim_{E \to x} f_E = f(x) \ ,$$

donde

$$f_E = \frac{1}{m(E)} \int_E f(x) \ dx$$

y al escribir $E \to x$ debe entenderse que nos aproximamos al punto x con conjuntos E de la clase $\mathcal{C}_x$.

(8.4) **Teorema.** *Sea* $\{\mathcal{C}_x\}_{x\in I\!R^n}$ *una familia que cumple las condiciones (a) y (b) y* f *una función localmente integrable. Entonces para casi todo* $x \in I\!R^n$ *, tenemos*

$$\lim_{E\to x} \frac{1}{m(E)} \int_E |f(y) - f(x)| \, dy = 0 \ .$$

DEMOSTRACIÓN: Definimos

$$\Gamma^* f(x) = \limsup_{E\to x} \int_E |f(y) - f(x)| \, dy$$

Luego $\Gamma^* f(x) \leq c\Gamma f(x)$, siendo c la constante que aparece en la condición (a) de la definición de la familia $\mathcal{C}_x$ y Γf la función que se introdujo en la demostración del teorema (8.3).

Dado que Γf es cero en casi todo punto, también lo es $\Gamma^* f$. Más aún hemos demostrado que si para un punto x vale la afirmación del teorema (8.3) para ese mismo punto es cero el límite en (8.4).

Q.E.D

Decimos que x es un **punto de Lebesgue** de la función f si

$$\lim_{r\to 0} r^{-n} \int_{|y|\leq r} |f(x+y) - f(x)| \, dy = 0 \ .$$

Una breve meditación muestra que x es un punto de Lebesgue de f si y sólo si el límite en el teorema (8.3) es cero. Por lo tanto si f es localmente integrable, entonces casi todo punto x es un punto de Lebesgue de f .

Si E es un conjunto medible, para casi todo $x \in I\!R^n$ vale

$$\lim_{r\to 0} \frac{m(E \cap B_r(x))}{m(B_r)} = \chi_E(x) \ .$$

Un punto x para el cual el límite anterior vale 1 se llama **punto de densidad** del conjunto E . Por lo tanto casi todo punto de E es punto de densidad del mismo conjunto. El resto de esta sección puede omitirse en una primera lectura.

A continuación damos una aplicación del concepto de punto de densidad.

Sea $d : I\!R^n \to I\!R$ una función continua y denotemos por E el conjunto de sus ceros. Dado que E es cerrado este conjunto es medible.

Supongamos ahora que la función d satisface la siguiente condición: para cada $x \in E$,

$$d(x+y) = O(|y|) \qquad \text{cuando} \qquad |y| \to 0 \ .$$

Esto es, para cada $x \in E$ existe una constante M (que puede depender de x) y un número $\delta > 0$ tal que $|d(x+y)| \leq M|y|$ si $|y| \leq \delta$.

Un ejemplo de función d con estas propiedades se obtiene tomando la distancia de un punto x a un conjunto cerrado E. En este caso la constante M vale 1 y δ es cualquier número positivo.

Demostraremos que para casi todo punto $x \in E$ se tiene $d(x+y) = o(|y|)$ cuando $|y| \to 0$. O sea, dado $\varepsilon > 0$ existe $\delta > 0$ tal que $|d(x+y)| \leq \varepsilon |y|$ si $|y| \leq \delta$.

En otras palabras, bajo las suposiciones hechas para la función d, veremos que en casi todo punto de E la función d es diferenciable y su diferencial vale cero.

Necesitamos demostrar primero la proposición siguiente:

(c) Sea x un punto de densidad para un conjunto medible F, es decir,

$$\lim_{r \to 0} \frac{m(F \cap B_r(x))}{m(B_r)} = 1 \ .$$

Entonces para cada $\varepsilon > 0$, existe $\delta > 0$ tal que $F \cap B_{\varepsilon|y|}(x+y) \neq \emptyset$ si $|y| \leq \delta$.

En efecto, siempre tenemos la siguiente inclusión $B_{\varepsilon|y|}(x+y) \subset B_{|y|+\varepsilon|y|}(x)$. Si para algún valor de y ocurre que $B_{\varepsilon|y|}(x+y) \cap F = \emptyset$, entonces tendremos

$$\begin{aligned}\frac{m\left(F \cap B_{|y|+\varepsilon|y|}(x)\right)}{m\left(B_{|y|+\varepsilon|y|}\right)} &\leq \frac{m\left(B_{|y|+\varepsilon|y|}(x)\right) - m\left(B_{\varepsilon|y|}(x+y)\right)}{m\left(B_{|y|+\varepsilon|y|}(x)\right)} \\ &= 1 - \left(\frac{\varepsilon}{1+\varepsilon}\right)^n \ .\end{aligned}$$

Para $\varepsilon > 0$ fijo, haciendo $|y| \to 0$, la desigualdad anterior contradice la existencia del límite en el enunciado de (c).

Ahora estamos en condiciones de demostrar que $d(x+y) = o(|y|)$, para casi todo $x \in E$. Con este fin definimos

$$E_k = \{x \in E : |d(x+y)| \leq k|y| \text{ si } |y| \leq 1/k\} \ .$$

Los conjuntos E_k son cerrados por ser d una función continua; además tenemos $E = \bigcup_{k\geq 1} E_k$.

Para k fijo, sea $x \in E_k$ un punto de densidad para este conjunto. Usando (c) tenemos que para $\varepsilon > 0$ existe $\delta > 0$ y $z \in E_k \cap B_{\varepsilon|y|}(x+y)$, si $|y| \leq \delta$. Supongamos además que $\varepsilon\delta \leq 1/k$.

Luego tenemos las desigualdades

$$|d(z+u)| \leq k|u| \qquad \text{si } |u| \leq 1/k \ ,$$
$$|z-(x+y)| \leq \varepsilon|y| \ .$$

Por lo tanto $|d(x+y)| = |d(z+x+y-z)| \leq k|x+y-z| \leq k\varepsilon|y|$.

Q.E.D.

3. Medidas abstractas.

En el parágrafo siguiente usaremos en $I\!R^n$ medidas diferentes a la de Lebesgue y aprovechamos esta oportunidad para introducir el concepto de medida en forma general.

El concepto de σ-álgebra en $I\!R^n$ se introdujo en Capítulo III, parágrafo 10. Ahora reemplazaremos $I\!R^n$ por un conjunto no vacío X arbitrario.

Decimos que una clase de conjuntos $\Sigma \subset \mathcal{P}(X)$ es una **σ- álgebra** de subconjuntos de X si se verifican las siguientes condiciones:

(1) $X \in \Sigma$.

(2) Toda unión numerable de miembros de Σ es un miembro de Σ.

(3) El complemento de cada miembro de Σ es un miembro de Σ.

Lamaremos **conjuntos medibles** a los miembros de Σ.

Ejemplos triviales de σ-álgebras para un conjunto X son $\Sigma = \{\emptyset, X\}$ y $\Sigma = \mathcal{P}(X)$. Para $X = I\!R^n$ hemos considerado las σ- álgebras de Lebesgue y de Borel.

Si X es un espacio métrico, la mínima σ-álgebra que contiene a todos los conjuntos abiertos de X recibe también el nombre de **σ-álgebra de Borel**. Esta σ-álgebra se obtiene como la intersección de las σ-álgebras que incluyen a los conjuntos abiertos.

En σ-álgebras concretas hemos analizado en detalle la medida de Lebesgue y también hemos introducido la medida de Lebesgue-Stieltjes. Damos ahora el concepto axiomático de medida.

Sea X un conjunto y Σ una σ-álgebra de subconjuntos de X. Decimos que μ es una **medida sobre Σ** si es una función con valores en $[0,\infty]$ que cumple:

$$\mu(\bigcup_j E_j) = \sum_j \mu(E_j) \ ,$$

para cada sucesión (E_j) en Σ de conjuntos disjuntos dos a dos. Además suponemos que $\mu(\emptyset) = 0$.

Insistimos que en la suma anterior, y también en la unión, solamente permitimos, a lo más, una cantidad infinita numerable de conjuntos medibles. De la definición de medida obtenemos: $A, B \in \Sigma$ y $A \subset B \implies \mu(A) \leq \mu(B)$.

Cuando μ es una medida sobre una σ-álgebra Σ la terna (X, Σ, μ) recibe el nombre de **espacio de medida**.

Veamos otro ejemplo de medida sobre $I\!R^n$ equipado con la σ-álgebra de Lebesgue. Sea f una función medible no negativa tal que sobre un conjunto de medida positiva es no nula y finita. Para un conjunto medible E ponemos

$$\mu_f(E) = \int_E f(x) \ dx \ .$$

La función μ_f es una medida sobre la σ-álgebra de Lebesgue de $I\!R^n$. Además μ_f toma valores distintos de 0 e ∞. Notemos al pasar que si $m(A) = 0$ entonces $\mu_f(A) = 0$. Esta propiedad se enuncia diciendo que la medida μ_f es **absolutamente continua con respecto a la medida** m.

En general, dadas dos medidas μ, ν sobre una σ-álgebra Σ, decimos que ν es **absolutamente continua con respecto a** μ si $\nu(A) = 0$ cada vez que $\mu(A) = 0$. En tal caso escribimos $\nu \ll \mu$. Medite el lector cómo deber ser una función no-negativa f para que valga simultáneamente $\mu_f \ll m$ y $m \ll \mu_f$.

Siendo Σ una σ-álgebra en X y x_0 un punto de X, para cada $A \in \Sigma$ definamos

$$\delta_{x_0}(A) = \begin{cases} 1 & \text{si } x_0 \in A \\ 0 & \text{si } x_0 \notin A. \end{cases}$$

Esta función δ_{x_0} define una medida sobre Σ que se llama la **delta de Dirac** concentrada en x_0.

Una generalización de la delta de Dirac es la función

$$\mu(A) = \sum_{j=1}^{\infty} a_j \; \delta_{x_j}(A) \ ,$$

donde (x_j) es una sucesión de puntos de X y (a_j) una sucesión de números reales positivos. Esta función μ es una medida sobre Σ (ver ejercicios).

Damos un último ejemplo de medida μ sobre una σ-álgebra abstracta Σ. Para cada $A \in \Sigma$ ponemos

$$\mu(A) = \begin{cases} \text{número de elementos de } A & \text{si } A \text{ es finito} \\ \infty & \text{si } A \text{ es infinito} \end{cases}$$

La función anterior es una medida ("counting measure") que se usa con cierta frecuencia.

Sea μ una medida sobre una σ-álgebra Σ en el espacio X. Si A y B son conjuntos medibles, tales que $A \subset B$ y $\mu(A) < \infty$, entonces es fácil ver que $\mu(B - A) = \mu(B) - \mu(A)$.

(8.5) **Teorema.** *Sea* (X, Σ, μ) *un espacio de medida y* (E_j) *una sucesión en* Σ. *Entonces valen las afirmaciones siguientes:*

(a) Si (E_j) *es monótona creciente, entonces*

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

(b) Si (E_j) *es monótona decreciente y* $\mu(E_1) < \infty$, *entonces*

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

(c) Si $\mu(X) < \infty$, *se verifican las desigualdades* $\mu(\liminf E_j) \leq \liminf \mu(E_j) \leq \limsup \mu(E_j) \leq \mu(\limsup E_j)$.

La demostración de estas propiedades es análoga a la ya realizada para la medida de Lebesgue y se deja como ejercicio. Observe que en *(c)* la finitud de $\mu(X)$ se necesita sólo para demostrar la última desigualdad.

Recordemos que si E es un conjunto medible Lebesgue en $I\!R^n$, podemos aproximar $m(E)$ con las medidas de los conjuntos abiertos que lo contienen y con las medidas de los conjuntos compactos contenidos en él. A continuación extenderemos esta propiedad a medidas más generales.

Sea X un espacio métrico y μ una **medida de Borel** sobre X, o sea una medida definida sobre la σ-álgebra de Borel de X. Decimos que μ es una **medida regular** si se cumplen las tres condiciones siguientes:

(1) $\mu(E) = \inf\{\mu(G) : G \supset E, G$ abierto $\}$, para cada conjunto medible E.

(2) $\mu(E) = \sup\{\mu(F) : F \subset E, F$ compacto$\}$, para cada abierto E o cada conjunto medible E de medida finita.

(3) $\mu(E) < \infty$ si E es compacto.

Observe que la siguiente condición es claramente equivalente a (1) si la medida del espacio total es finita.

(2') $\mu(E) = \sup\{\mu(F) : F \subset E, F$ cerrado $\}$, para cada conjunto medible E.

En los casos más interesantess (2) puede ser reemplazada por (2'). Si X es un espacio **$\boldsymbol{\sigma}$-compacto** (o sea que se puede expresar como una unión numerable de conjuntos compactos) y además vale (3), entonces (2) es equivalente a (2').

(8.6) **Teorema.** *Sea X un espacio métrico σ-compacto y μ una medida de Borel finita sobre cada conjunto compacto. Entonces μ es una medida regular.*

Demostración: Suponemos primero que X es un conjunto compacto y definimos

$$\mathcal{C} = \left\{ E \in \Sigma : \mu(E) = \inf_{G \supset E} \mu(G) \text{ y } \mu(E) = \sup_{F \subset E} \mu(F) \right\} ,$$

donde Σ es la σ-álgebra de Borel, G abierto y F cerrado.

La clase $\mathcal{C}$ contiene a los conjuntos abiertos de X. En efecto, para cada $E \subset X$ abierto pongamos

$$E_j = \{x \in X : d(x, CE) \geq 1/j\} ,$$

con $d(x, CE) = \inf\{d(x,y) : y \notin E\}$ y $d(\,\cdot\,,\,\cdot\,)$ la métrica sobre X.

Los conjuntos E_j son compactos, no vacíos desde cierto j en adelante, y crecen monotonamente hacia el conjunto E. Luego cada conjunto abierto está en $\mathcal{C}$.

Para ver que $\mathcal{C}$ es una σ-álgebra es conveniente hacerlo en el siguiente orden: $\mathcal{C}$ es cerrada bajo complementos, $\mathcal{C}$ es cerrada bajo uniones e intersecciones finitas [es consecuencia de la fórmula $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$], $\mathcal{C}$ es cerrada bajo unión numerable disjunta.

Luego $\mathcal{C} = \Sigma$ si X es un conjunto compacto y el teorema queda demostrado en este caso. El lector debería aportar los detalles faltantes en la demostración de este teorema

Q.E.D.

Como consecuencia del teorema (8.6) tenemos que toda medida de Borel sobre $I\!R^n$, finita sobre cada cubo, es necesariamente una medida regular.

4. Diferenciación de una medida de Borel con respecto a la medida de Lebesgue.

En este parágrafo μ denotará una medida de Borel sobre $I\!R^n$, tal que $\mu(Q) < \infty$ para cada cubo Q. Con este tipo de medida estudiaremos el límite siguiente:

$$\lim_{Q \to x} \frac{\mu(Q)}{m(Q)} .$$

En la sección 2 hemos considerado promedios sobre cubos, bolas euclídeas y otros conjuntos más generales (ver teorema (8.4)); lo mismo puede hacerse con la medida μ.

Para cubrir una amplia clase de ejemplos, consideraremos sobre $I\!R^n$ una métrica d que genere los mismos conjuntos abiertos que la distancia euclídea. Como es usual, definimos con esta distancia d las bolas

$$B = B_r(x) = \{y \in I\!R^n : d(x,y) < r\} .$$

Supondremos de ahora en adelante que existe una constante finita k tal que

$$(a) \qquad m(B_{2r}(x)) \leq k\ m(B_r(x)) ,$$

para todo $x \in I\!R^n$ y $r > 0$. Ver ejercicios 6 a 8.

Nos interesará estudiar el comportamiento de los promedios $\mu(B)/m(B)$ para $B \to x$; y esto nos induce a definir el siguiente operador maximal

$$M\mu(x) = \sup_{B \ni x} \frac{\mu(B)}{m(B)} \ ,$$

donde el supremo se toma sobre todas las bolas que contienen a x. Para tener el límite mencionado al comienzo de esta capítulo hay que tomar las bolas definidas por la métrica $d(x,y) = \max_{1 \le i \le n} |x_i - y_i|$, donde x_i, y_i son las i-ésimas coordenadas de x e y. Por otro lado, podemos estudiar promedios no comprendidos en la clase descripta por el teorema (8.4). Ver Ejercicio 8.

Insistimos en que en el resto de esta sección, cuando hablemos de bolas será en el sentido general señalado anteriormente. Sea $\mathcal{C}$ una colección de bolas en $I\!R^n$ entonces existe una subfamilia maximal $\mathcal{D}$ de bolas disjuntas, esto es $\mathcal{D}$ es una subfamilia disjunta de bolas y no existe ninguna bola en $\mathcal{C}$ disjunta con todas las bolas de $\mathcal{D}$. La afirmación anterior es una consecuencia inmediata del lema de Zorn, véase por ejemplo la obra de Kelley [7].

(8.7) **Lema de cubrimiento.** *Sea $\mathcal{C}$ una colección de bolas generadas por una métrica d en $I\!R^n$ tal que*

$$\sup \{ \ \delta(B) \ : \ B \in \mathcal{C} \ \} < \infty \ .$$

Entonces existe una subfamilia numerable $\mathcal{D} \subseteq \mathcal{C}$ de bolas disjuntas que verifica:

$$\bigcup_{B \in \mathcal{C}} \subseteq \bigcup_{B \in \mathcal{D}} 5B$$

Demostración: Sean $d = \sup\{\delta(B) : B \in \mathcal{C}\} < \infty$, y $\mathcal{C}_j = \{B \in \mathcal{C} : d2^{-j} < \delta(B) \le d2^{-j+1}\}, (j = 1, 2, \ldots)$. Definamos $\mathcal{D}_j$ inductivamente como sigue:

(a) $\mathcal{D}_1$ es una colección maximal de bolas disjuntas en $\mathcal{C}_1$.

(b) Supongamos que las familias de bolas $\mathcal{D}_j$, $(j = 1, \ldots, k-1)$ han sido seleccionadas. Elegimos como $\mathcal{D}_k$ cualquier familia maximal de bolas disjuntas contenida en

$$\{B \in \mathcal{C}_k : \ B \cap B' = \emptyset \text{ para toda } B' \in \cup_{j=1}^{k-1} \mathcal{D}_j\} \ .$$

Pongamos finalmente $\mathcal{D} = \cup_{j=1}^{\infty} \mathcal{D}_j$. Claramente $\mathcal{D}$ es una colección numerable de bolas disjuntas contenidas en $\mathcal{C}$. Nos resta demostrar que para cada $B \in \mathcal{C}$ existe una bola $B' \in \mathcal{D}$ tal que $B \cap B' \neq \emptyset$ y que $B \subset 5B'$.

En efecto, para $B \in \mathcal{C}$ existe j tal que $B \in \mathcal{D}_j$, y teniendo en cuenta la maximalidad de $\mathcal{D}_j$ existe una bola $B' \subset \cup_{k=1}^{j} \mathcal{D}_j$ tal que $B \cap B' \neq \emptyset$. Como $\delta(B') \geq d2^{-j}$ y $\delta(B) \leq d2^{1-j}$, resulta que $\delta(B) \leq 2\delta(B')$ y por lo tanto $B \subset 5B'$.

Q.E.D.

El siguiente teorema es una generalización en dos aspectos del teorema (8.1). Por un lado nos da una acotación para la función maximal de una medida en lugar de una función; y por otro lado los promedios se toman sobre bolas en $I\!R^n$ generadas por una métrica d (que cumplan la condición (a) señalada anteriormente).

(8.8) **Teorema.** *Sea μ una medida de Borel sobre $I\!R^n$, finita sobre cubos, y sea*

$$M\mu(x) = \sup_{B \ni x} \frac{\mu(B)}{m(B)} \ .$$

Entonces existe una constante c independiente de la medida μ, tal que

$$t \ m_e(\{x \in I\!R^n : M\mu(x) > t\}) \leq c \ \mu(I\!R^n) \ ,$$

para cada $t > 0$.

Demostración: La demostración de este teorema es similar a la del Teorema (8.1) pero debemos usar el lema (8.7). Los detalles se dejan al lector interesado.

Se dice que una medida μ está **concentrada** sobre un conjunto medible E si $\mu(A) = \mu(A \cap E)$ para cada conjunto medible A. Dos medidas ν, μ definidas sobre una misma σ-álgebra se llaman **mutuamente singulares** (escribimos $\nu \perp \mu$) si están concentradas en conjuntos medibles disjuntos.

La delta del Dirac es mutuamente singular con respecto a la medida de Lebesgue. Dadas dos funciones no negativas f, g localmente integrables en $I\!R^n$, definimos las siguientes medidas de Borel sobre $I\!R^n$:

$$\mu_f(A) = \int_A f(x) \ dx, \qquad \mu_g(A) = \int_A g(x) \ dx \ .$$

Claramente μ_f, μ_g son medidas de Borel regulares y $\mu_f \perp \mu_g$ si las funciones f y g tienen soportes disjuntos.

(8.9) **Teorema.** *Sea μ una medida de Borel sobre $I\!R^n$, finita sobre cubos, y mutuamente singular con respecto de la medida de Lebesgue. Entonces*

$$\lim_{B\to x} \frac{\mu(B)}{m(B)} = 0 \ ,$$

para casi todo $x \in I\!R^n$.

Nota: En el límite de arriba se toman todas las bolas (formadas con una métrica dada d) que contienen al punto x.

DEMOSTRACIÓN: Para $t > 0$ definamos

$$E(t) = \left\{ x \in I\!R^n : \limsup_{B\to x} \frac{\mu(B)}{m(B)} > t \right\} \ .$$

Sabemos que existe un conjunto medible A tal que $m(A) = \mu(I\!R^n - A) = 0$. Por ser la medida μ regular, dado $\varepsilon > 0$ existe un conjunto abierto G tal que

$$G \supset I\!R^n - A \qquad \text{y} \qquad \mu(G) < \varepsilon \ .$$

Pongamos ahora $E_1 = E(t) \cap (I\!R^n - A)$ y para cada $x \in E_1$ elegimos una bola B_x tal que $x \in B_x \subset G$ y $\mu(B_x) > tm(B_x)$.

Por el lema (8.7) existe una sucesión (B_{x_i}) de bolas disjuntas contenidas en G tal que $E_1 \subset \bigcup\limits_{i=1}^{\infty} 5B_{x_i}$ y $\mu(B_{x_i}) > tm(B_{x_i})$.

Por lo tanto, con k la constante que aparece en (a),

$$\begin{aligned} m(E_1) &\le \sum_{i=1}^{\infty} m(5B_{x_i}) \le k^3 \sum_{i=1}^{\infty} m(B_{x_i}) \le \frac{k^3}{t} \sum_{i=1}^{\infty} \mu(B_{x_i}) \\ &= \frac{k^3}{t} \mu\left(\bigcup_{i=1}^{\infty} B_{x_i} \right) \le \frac{k^3}{t}\mu(G) < \frac{\varepsilon k^3}{t} \ . \end{aligned}$$

Luego $m_e(E_1) = 0$, lo que implica $m_e(E(t)) = 0$.

Q.E.D.

El teorema (8.9) nos permite demostrar (como se hizo en el teorema (8.3)) que si μ es una medida de Borel de la forma

$$\mu(A) = \mu_f(A) = \int_A f(x)\, dx \ ,$$

donde f es no negativa localmente integrable. Entonces se cumple

$$\lim_{B \to x} \frac{\mu(B)}{m(B)} = f(x) \ ,$$

para casi todo $x \in I\!R^n$.

Si λ y μ son dos medidas sobre una σ-álgebra, definimos la suma $\lambda + \mu$ por medio de la fórmula $(\lambda + \mu)(E) = \lambda(E) + \mu(E)$.

Teniendo en cuenta el teorema (8.9) hemos demostrado que si μ es una medida de Borel regular de la forma $\mu = \mu_f + \nu$ con f localmente en $L^1(I\!R^n)$ y $\nu \perp m$, entonces el límite

$$\lim_{B \to x} \frac{\mu(B)}{m(B)} \ ,$$

existe y es igual a $f(x)$ para casi todo $x \in I\!R^n$.

Más adelante veremos que toda medida de Borel regular se puede expresar de esta forma, en virtud del teorema de Radon-Nikodym.

Es interesante señalar que se puede demostrar la existencia en casi todo punto del límite anterior sin recurrir al teorema de Radon-Nikodym. Esto no es difícil, pero no se hará aquí . Consultar "Differentiation of Integrals in $I\!R^n$", por Miguel de Guzmán, Springer Verlag, (1975).

5. Continuidad en L^p del operador maximal de Hardy-Littlewood.

Sea T un operador definido en $L^p(I\!R^n)$, $1 \le p \le \infty$, con valores en $\mathcal{M}(I\!R^n)$ (el espacio de las funciones medibles Lebesgue). Decimos que T es **subaditivo** si

$$|T(f+g)(x)| \le |T\ f(x)| + |T\ g(x)| \ ,$$

para cada f, g en L^p.

Cuando hablemos de operadores definidos en L^p siempre supondremos que son al menos subaditivos. Por ejemplo el operador maximal de Hardy-Littlewood es subaditivo.

Recordemos que, para un operador lineal $T : L^p \to L^q$ la continuidad de T es equivalente a afirmar que existe una constante c, tal que

$$(a) \qquad \|Tf\|_q \le c\|f\|_p \qquad (f \in L^p) \; .$$

Usando la desigualdad de Chebyshev en (a) obtenemos que

$$t^q m(\{|Tf| > t\}) \le (c\|f\|_p)^q \; ,$$

para toda $f \in L^p$ y cada $t > 0$. En la desigualdad anterior hemos supuesto $1 \le q < \infty$.

Se dice que un operador $T : L^p(\mathbb{R}^n) \to \mathcal{M}(\mathbb{R}^n)$ **es de tipo débil** $\boldsymbol{(p,q)}$ si cumple la última desigualdad con una constante c independiente de f y de t. Por ejemplo, el operador maximal de Hardy-Littlewood es de tipo débil $(1,1)$.

Si un operador T cumple la desigualdad (a) decimos que **es de tipo fuerte** $\boldsymbol{(p,q)}$. Aquí $q = \infty$ está permitido. Hemos demostrado que si un operador es de tipo fuerte (p,q), entonces es de tipo débil (p,q).

Decimos que una función medible f pertenece a $\boldsymbol{L^p(\mathbb{R}^n)}$ **débil** si existe c tal que

$$t^p m(\{|f| > t\}) \le c^p \; ,$$

para todo $t > 0$. Hemos supuesto $p < \infty$.

Es un ejercicio fácil ver que $L^p(\mathbb{R}^n)$ débil contiene estrictamente a $L^p(\mathbb{R}^n)$.

Sabiendo que un operador T es simultáneamente de tipo débil (p_1, p_1) y (p_2, p_2), se puede inferir que T es de tipo fuerte para "valores intermedios" de p. Esta es una situación típica dentro de una vasta teoría conocida con el nombre de **interpolación de operadores**.

Pongamos $L^1(\mathbb{R}^n) + L^r(\mathbb{R}^n) = \{f + g : f \in L^1, g \in L^r\}$. Claramente $L^1(\mathbb{R}^n) + L^r(\mathbb{R}^n) \supset L^p(\mathbb{R}^n)$ si $1 < p < r$.

El siguiente teorema nos bastará para nuestros propósitos.

(8.10) **Teorema de Interpolación de Marcinkiewicz.** *Sean* $1 < r \le \infty$ *y* T *un operador subaditivo definido en* $L^1(\mathbb{R}^n) + L^r(\mathbb{R}^n)$ *con valores en el espacio de las funciones medibles, el cual es simultáneamente de*

tipo débil $(1,1)$ *y* (r,r). *Entonces* T *es de tipo fuerte* (p,p) *para* $1 < p < r$.
Más explicitamente, supongamos que para f, g *se verifican las desigualdades:*
(1) $|T(f+g)(x)| \leq |Tf(x)| + |Tg(x)|$.
(2) $tm(\{|Tf| > t\}) \leq c_1 \; \|f\|_1$.
(3) $t^r m(\{|Tf| > t\}) \leq c_r^r \; \|f\|_r^r$.
(para el caso $r = \infty$ *suponemos* $\|Tf\|_\infty \leq c_\infty \|f\|_\infty$ *). Entonces*

$$\|Tf\|_p \leq c_p \|f\|_p \qquad\qquad (f \in L^p(I\!R^n)) ,$$

para todo p *en el intervalo* $(1,r)$. *Siendo* c_p *una constante que depende de* c_1, c_r, p *y* r, *pero no de la función* f.

Demostración: Supondremos $r < \infty$. El caso $r = \infty$ es más fácil y lo dejamos como ejercicio.

Para $f \in L^p$, $1 < p < r$, deseamos estimar $\|Tf\|_p$ para lo cual usaremos la fórmula

$$\|Tf\|_p^p = p \int_0^\infty t^{p-1} \; \lambda(t) \; dt , \tag{4}$$

donde $\lambda(t) = m(\{|Tf| > t\})$. Ver capítulo VII, sección 6.

Para cada $t > 0$ definimos

$$f_1(x) = \begin{cases} f(x) & \text{si } |f(x)| > t \\ 0 & \text{si } |f(x)| \leq t \end{cases}$$

y $f_2 = f - f_1$. Claramente $f_1 \in L^1$ y $f_2 \in L^r$ y tenemos que

$$\begin{aligned} \lambda(t) &\leq m(\{|Tf_1| + |Tf_2| > t\}) \\ &\leq m(\{|Tf_1| > t/2\}) + m(\{|Tf_2| > t/2\}) . \end{aligned}$$

En la primera desigualdad anterior hemos usado (1). Ahora usando (2) y (3) llegamos a la desigualdad siguiente

$$\lambda(t) \leq \frac{2c_1}{t} \int_{|f(x)|>t} |f(x)| dx + \left(\frac{2c_r}{t}\right)^r \int_{|f(x)|\leq t} |f(x)|^r dx .$$

Poniendo en (4) la estimación obtenida arriba para $\lambda(t)$ y usando el teorema de Tonelli obtenemos el teorema con una constante c_p dada por

$$c_p^p = 2c_1 \frac{p}{p-1} + (2c_r)^r \frac{p}{r-p} .$$

Para el caso $r = \infty$ es conveniente poner

$$f_2(x) = \begin{cases} f(x) & \text{si } |f(x)| \leq t/2c \\ 0 & \text{si } |f(x)| > t/2c \ , \end{cases}$$

$f_1 = f - f_2$, siendo $\|Tf\|_\infty \leq c\|f\|_\infty$.

En este caso la constante c_p estará dada por

$$c_p = 2c_1^{1/p} \, c^{1/p'} \left(\frac{p}{p-1}\right)^{1/p} \ , \qquad 1/p + 1/p' = 1 \ .$$

Q.E.D.

(8.11) **Teorema.** *Si M es el operador maximal de Hardy-Littlewood, entonces valen las siguientes desigualdades*

(i) $tm(\{Mf > t\}) \leq c_1\|f\|_1 \qquad f \in L^1$.

(ii) $\|Mf\|_p \leq c_p\|f\|_p \qquad f \in L^p \ , \qquad 1 < p \leq \infty$,

(iii) $\|\chi_E Mf\|_p \leq a_p(m(E))^{\frac{1-p}{p}}\|f\|_1 \qquad f \in L^1 \ ,$

donde $0 < p < 1$ *y* $E \subset I\!R^n$ *es de medida finita.*

(iv) $\|\chi_E Mf\|_1 \leq 2m(E) + 2c_1 \int |f(x)| \log^+ |f(x)| dx \ ,$

con $m(E) < \infty$.

Nota: Las constantes que aparecen en el teorema tienen el siguiente comportamiento: $c_p = O\left(\frac{1}{p-1}\right)$ cuando $p \to 1$,

$c_p = O(1)$ si $p \to \infty$.

$a_p^p = 2 + 2c_1\frac{p}{1-p}$ para $0 < p < 1$.

Demostración: La parte *(i)* es consecuencia del teorema (8.1) o del teorema (8.8).

Dado que el operador maximal Mf es de tipo débil $(1,1)$ y fuerte (∞,∞), el teorema de interpolación (8.10) nos asegura la validez de *(ii)*

Para demostrar *(iii)* ponemos

$$\lambda(t) = m(E \cap \{Mf > t\}) \ .$$

Luego, en forma análoga a lo hecho para $p \geq 1$, tenemos

$$\begin{aligned} \int_E |Mf(x)|^p \, dx &= p\int_0^\infty t^{p-1}\lambda(t) \, dt \\ &= p\int_0^s t^{p-1}\lambda(t) \, dt + p\int_s^\infty t^{p-1}\lambda(t) \, dt \\ &= I_1 + I_2 \ , \end{aligned}$$

siendo $s > 0$ un valor a elegir dentro de un momento.

Por un lado $I_1 \leq s^p m(E)$. Por otro lado, usando la parte *(i)* del teorema, se tiene

$$I_2 \leq p \int_s^\infty t^{p-1} c_1 \; t^{-1} \|f\|_1 dt = c_1 \frac{p}{1-p} s^{p-1} \|f\|_1 \; .$$

La demostración de *(iii)* finaliza eligiendo $s = \|f\|_1 / m(E)$.

Para demostrar *(iv)* utilizamos la siguiente descomposición para f (ya usada en la demostración del teorema de interpolación),

$$f_2(x) \begin{cases} f(x) & \text{si } |f(x)| \leq t \\ 0 & \text{si } |f(x)| > t, \end{cases}$$

$f = f_1 + f_2$.
Luego

$$m(\{Mf > 2t\}) \leq m(\{Mf_1 > t\}) \leq \frac{c_1}{t} \int_{|f(x)|>t} |f(x)| \; dx \; .$$

Usando la función $\lambda(t)$ introducida en la demostración de *(iii)* y la desigualdad anterior, tenemos

$$\begin{aligned} \int_E Mf(x) \; dx &= \int_0^\infty \lambda(t) \; dt \\ &\leq 2m(E) + \int_2^\infty \lambda(t) \; dt \\ &\leq 2m(E) + \int_2^\infty \frac{2c_1}{t} \left(\int_{2|f(x)|>t} |f(x)| dx \right) dt \; . \end{aligned}$$

Usando el teorema de Tonelli en la integral iterada anterior se concluye la demostración del teorema.

Llegado a este punto es muy probable que el lector haya observado que la demostración de *(ii)* a *(iv)* del teorema (8.11) sólo depende de las siguientes propiedades del operador maximal M : es de tipo débil (1,1), $\|Mf\|_\infty \leq \|f\|_\infty$, es sublineal y **homogéneo** $(M(af) = |a|Mf)$. La demostración no depende de ninguna propiedad especial de la función maximal de Hardy-Littlewood fuera de las ya mencionadas.

6. Aproximaciones de la identidad.

Consideremos los siguientes promedios para una función f localmente integrable:

$$f_\varepsilon(x) = \frac{1}{m(B_\varepsilon)} \int_{|x-y|\leq\varepsilon} f(y)\ dy \qquad (\varepsilon > 0)\ ,$$

donde $m(B_\varepsilon)$ representa la medida de Lebesgue de una bola de radio ε.

Poniendo $B = \{y : |y| \leq 1\}$, $K(x) = [m(B)]^{-1}\chi_B(x)$, $K_\varepsilon(x) = \varepsilon^{-n}K(x/\varepsilon)$, es conveniente expresar f_ε como la siguiente convolución

$$f_\varepsilon(x) = K_\varepsilon * f(x) = \int K_\varepsilon(x-y)\ f(y)\ dy\ ,$$

Obsevemos que para cualquier $\varepsilon > 0$, $\int K_\varepsilon(x)\ dx = \int K(x)\ dx = 1$.

Hemos demostrado en secciones anteriores que $f_\varepsilon(x)$ converge puntualmente a $f(x)$ cuando $\varepsilon \to 0$. O sea, usando la expresión de arriba, que convolucionando la función f con ciertos “núcleos” K_ε obtenemos una aproximación de esta función.

Nuestro siguiente objetivo será obtener una aproximación para f usando núcleos más generales.

Decimos que una familia de funciones $\{K_\varepsilon\}_{\varepsilon>0}$ es una **aproximación de la identidad** sobre $\mathbb{R}^n$ si se cumplen las siguientes condiciones:

I_1. $\int |K_\varepsilon(x)|dx \leq c < \infty$ $\qquad (\varepsilon > 0)$.

I_2. $\int K_\varepsilon(x)dx = 1$ $\qquad (\varepsilon > 0)$.

I_3. para cada $\delta > 0$ tenemos

$$\lim_{\varepsilon\to 0} \int_{|x|\geq\delta} |K_\varepsilon(x)|dx = 0\ .$$

El ejemplo típico de aproximaciones de la identidad se obtiene tomando $K \in L^1(\mathbb{R}^n)$ con $\int K(x)dx = 1$ y poniendo $K_\varepsilon(x) = \varepsilon^{-n}K(x/\varepsilon)$. Para otros ejemplos ver los ejercicios.

(8.12) **Teorema.** *Sea $\{K_\varepsilon\}_{\varepsilon>0}$ una aproximación de la identidad y $1 \leq p < \infty$. Entonces, para cada $f \in L^p(\mathbb{R}^n)$ tenemos que $\|K_\varepsilon * f - f\|_p \to 0$ cuando $\varepsilon \to 0$. Más todavía, vale la siguiente desigualdad*

$$\|f - K_\varepsilon * f\|_p \leq c\omega(\delta) + 2\|f\|_p \int_{|x|\geq\delta} |K_\varepsilon(x)|dx\ ,$$

donde c *es la constante que aparece en* I_1 , $\delta > 0$ *es arbitrario y*

$$\omega(\delta) = \sup_{|x|\leq\delta} \left\{ \int |f(y-x) - f(y)|^p dy \right\}^{1/p}$$

Demostración: Será suficiente demostrar la desigualdad anterior, ver el ejercicio 10 del capítulo VII.

Desde que vale I_2 , tenemos las siguientes igualdades

$$\begin{aligned} K_\varepsilon * f(x) - f(x) &= \int K_\varepsilon(x-y)(f(x-y) - f(x))dy \\ &= \int_{|y|\leq\delta} K_\varepsilon(y)(f(x-y) - f(x))dy + \\ &\quad + \int_{|y|>\delta} K_\varepsilon(y)(f(x-y) - f(x))dy \ . \end{aligned}$$

Usando la desigualdad de Minkowski para integrales (ejercicio 16 del capítulo VII) obtenemos la desigualdad del teorema.

Q.E.D.

Desde ahora en adelante nos restringiremos al caso de aproximaciones de la identidad de la forma $K_\varepsilon(x) = \varepsilon^{-n}K(x/\varepsilon)$ con K integrable y de integral igual a uno. Nos interesará buscar condiciones suficientes sobre K que aseguren la existencia en casi todo punto del límite

$$\lim_{\varepsilon\to 0} K_\varepsilon * f(x) \ , \qquad 1 \leq p \leq \infty \ , \ f \in L^p(\mathbb{R}^n) \ .$$

Se podría demostrar, aunque no se hará aquí , que la integrabilidad de la función K no es condición suficiente para la convergencia puntual de $K_\varepsilon * f$ cuando $f \in L^p$ y $1 \leq p < \infty$. El caso $f \in L^\infty$ es excepcional, como lo demuestra el siguiente teorema.

(8.13) **Teorema.** *Sea* K *una función integrable cuya integral vale uno y* f *una función en* $L^\infty(\mathbb{R}^n)$ *. Entonces, para casi todo* x *tenemos que* $K_\varepsilon * f(x) \to f(x)$ *cuando* $\varepsilon \to 0$ *. Más todavía, dicha relación límite se cumple en cada punto de Lebesgue de* f *.*

Demostración: Descompongamos adecuadamente al núcleo $K(x)$. Para un número positivo N definimos

$$\begin{aligned} K_N^1(x) &= K(x)\chi_{\{|x|\leq N\}}(x), \\ K_N^2(x) &= K(x) - K_N^1(x). \end{aligned}$$

Usando los núcleos K_N^1, K_N^2 escribimos $K_\varepsilon * f(x)$ de la forma

$$K_\varepsilon * f(x) = J_1 + J_2 + J_3,$$

donde

$J_1 = \int_{|K_N^1(y)|\geq T} K_N^1(y)(f(x-\varepsilon y) - f(x))\, dy$,
$J_2 = \int_{|K_N^1(y)|<T} K_N^1(y)(f(x-\varepsilon y) - f(x))\, dy$,
$J_3 = \int K_N^2(y)(f(x-\varepsilon y) - f(x))\, dy$,

y T es un número que luego será elegido.

Sea ahora x un punto de Lebesgue de la función f y $\delta > 0$ arbitrario. Elijamos el número N tal que

$$|J_3| \leq (\|f\|_\infty + |f(x)|) \int_{|y|>N} |K(y)| dy < \delta \ .$$

Fijado N elegimos T de modo que cumpla

$$|J_1| \leq (\|f\|_\infty + |f(x)|) \int_{|K_N^1(y)|\geq T} |K_N^1(y)| dy < \delta \ .$$

Ahora, puesto que x es un punto de Lebesgue de f, tenemos

$$\limsup_{\varepsilon\to 0} |J_2| \leq T \limsup_{\varepsilon\to 0} \int_{|y|\leq N} |f(x-\varepsilon y) - f(x)| dy = 0 \ .$$

Resumiendo, acabamos de demostrar, que para todo $\delta > 0$ vale

$$\limsup_{\varepsilon\to 0} |K_\varepsilon * f(x) - f(x)| \leq 2\delta \ ,$$

cuando x es un punto de Lebesgue de f.

Q.E.D.

A cada función K definida sobre $I\!R^n$ con valores en la recta real extendida le asociamos la llamada **mínima mayorante radial no creciente** y la designaremos con ℓ. Más precisamente, para $x \in I\!R^n$ definimos

$$\ell(x) = \sup_{|x|\leq|y|} |K(y)| \ .$$

La función ℓ es **radial** ($\ell(x) = \ell(y)$ si $|x| = |y|$). Usaremos ℓ_0 para la función definida sobre la recta real positiva por $\ell_0(r) = \ell(x)$ con $|x| = r$. La función ℓ_0 es decreciente.

(8.14) **Teorema.** *Sea* K *una función medible definida sobre* $\mathbb{R}^n$ *. Suponemos que la mínima mayorante radial no creciente* ℓ *asociada a* K *es una función integrable y la integral de la función* K *vale uno. Entonces valen las afirmaciones siguientes:*
(a) $\sup_{\varepsilon>0} |f * K_\varepsilon(x)| \le c_n \left(\int \ell(y)\ dy\right) Mf(x)$ *, donde* c_n *es una constante que depende sólo de la dimensión y* Mf *es la función maximal de Hardy-Littlewood.*
(b) Para $f \in L^p(\mathbb{R}^n)$ *,* $1 \le p \le \infty$ *y para casi todo* $x \in \mathbb{R}^n$ *se cumple que* $f * K_\varepsilon(x) \to f(x)$ *cuando* $\varepsilon \to 0$ *.*
(c) Si $f \in L^p$ *,* $1 \le p < \infty$ *, la convergencia de* $f * K_\varepsilon$ *hacia* f *cuando* $\varepsilon \to 0$ *, es en la norma* L^p *.*

Demostración: La parte *(c)* fue demostrada en (8.12) bajo condiciones más débiles. No obstante, observe que *(c)* es una consecuencia directa de *(a)*, *(b)*, el teorema (8.11)*(ii)* y el teorema de la convergencia dominada de Lebesgue cuando $1 < p < \infty$.

Pasamos a demostrar *(a)* y por simplicidad en la notación suponemos $f \ge 0$. Luego tenemos

$$\begin{aligned}
|K_\varepsilon * f(x)| &\le \varepsilon^{-n} \int \ell_0(|y|/\varepsilon) f(x-y)\ dy \\
&= \sum_{-\infty}^{+\infty} \varepsilon^{-n} \int_{\varepsilon 2^j \le |y| \le \varepsilon 2^{j+1}} \ell_0(|y|/\varepsilon) f(x-y) dy \\
&\le \sum_{-\infty}^{\infty} \varepsilon^{-n} \ell_0(2^j) \int_{|y| \le \varepsilon 2^{j+1}} f(x-y) dy \\
&\le c \sum_{-\infty}^{\infty} 2^{jn} \ell_0(2^j) Mf(x) \\
&\le c\ Mf(x) \int \ell(x) dx\ ,
\end{aligned}$$

donde hemos utilizado la misma letra c para dos constantes distintas que dependen sólo de la dimensión.

La demostración de *(b)* es muy similar a la del teorema de diferenciación de Lebesgue. Pese a ello la realizaremos en detalle para el caso $p < \infty$.

Si f es una función continua con soporte compacto, *(b)* se cumple en todo punto. Esto es consecuencia del teorema (8.13) o de un cálculo muy sencillo.

Para $f \in L^p(I\!R^n)$ ponemos

$$\Gamma f(x) = \limsup_{\varepsilon \to 0} |K_\varepsilon * f(x) - f(x)| \ .$$

Demostraremos que para todo $t > 0$ se cumple $m_e(\{\Gamma f > t\}) = 0$, y con esto finalizaremos la demostración del teorema.

Por la parte *(a)* resulta que

$$\Gamma f(x) \leq |f(x)| + c\ Mf(x) \ .$$

Usando el teorema (8.11) y la desigualdad de Chebyshev obtenemos la siguiente desigualdad para una constante c independiente de f y de t:

$$m_e(\{\Gamma f > t\}) \leq c\ t^{-p} \|f\|_p^p \ .$$

Por otro lado si g es continua con soporte compacto $\Gamma f \leq \Gamma(f - g)$.

Combinando estas dos últimas desigualdades llegamos a

$$m_e(\{\Gamma f > t\}) \leq c\ t^{-p} \|f - g\|_p^p \ ,$$

para cada g continua con soporte compacto y usando la densidad de estas funciones en L^p, obtenemos que $m_e(\{\Gamma f > t\}) = 0$.

Q.E.D.

El teorema (8.14) da condiciones suficientes para la convergencia en casi todo punto de una aproximación de la identidad, obtenida a partir de un núcleo K que puede no ser acotado en el origen. Más todavía, estas condiciones no permiten singularidades para K fuera del origen. En conexión con esto último compare el lector con el teorema (8.13) y el ejercicio 20.

Quizá no escapó a la observación del lector que el teorema (8.14), si bien nos asegura la convergencia en casi todo punto para una función f, no nos dice cuáles son los puntos donde hay convergencia. El próximo teorema subsanará esta situación.

(8.15) **Teorema.** *Sea K una función medible sobre $I\!R^n$, tal que su mínima mayorante radial no creciente es integrable. Sean $1 \leq p \leq \infty$, $f \in L^p(I\!R^n)$ y x un punto de Lebesgue de la función f. Entonces se tiene*

$$\lim_{\varepsilon \to 0} \int_{I\!R^n} |K_\varepsilon(y)|\ |f(x - y) - f(y)| dy = 0 \ ,$$

donde $K_\varepsilon(x) = \varepsilon^{-n} K(x/\varepsilon)$.

DEMOSTRACIÓN: Nos restringiremos a $1 \le p < \infty$; el caso $p = \infty$ se analizó en el teorema (8.13).

Si x es un punto de Lebesgue de f, dado $\rho > 0$ existe $\delta > 0$ tal que si $0 < r \le \delta$ entonces

$$r^{-n} \int_{|y| \le r} |f(x-y) - f(x)| dy \le \rho \ .$$

Definamos

$$g(y) = \begin{cases} |f(x+y) - f(x)| & \text{si } |y| \le \delta \\ 0 & \text{si } |y| > \delta \end{cases}$$

Es fácil ver que $Mg(0) \le c\rho$, con una constante c que depende sólo de la dimensión.

Ahora acotamos convenientemente la función $A_\varepsilon f$ definida por

$$A_\varepsilon f(x) = \int |K_\varepsilon(y)| \, |f(x-y) - f(x)| dy \ .$$

Es claro que

$$\begin{aligned} A_\varepsilon f(x) \le & \int_{|y| \le \delta} |K_\varepsilon(y)| g(-y) dy + \int_{|y| > \delta} |K_\varepsilon(y)| \, |f(x-y)| dy + \\ & + |f(x)| \int_{|y| > \delta} |K_\varepsilon(y)| dy \\ = & J_1 + J_2 + J_3 \ . \end{aligned}$$

Claramente $J_3 \to 0$ cuando $\varepsilon \to 0$.

Como $J_1 = |K|_\varepsilon * g(0)$ resulta, usando la parte *(a)* del teorema (8.14), que

$$J_1 \le c_n \int \ell(x) dx \ Mg(0) \le c \ c_n \rho \ .$$

Para estimar la integral J_2 observamos las siguientes desigualdades

$$\begin{aligned} J_2 & \le \left(\int_{|y| \ge \delta} |K_\varepsilon(y)| dy \right)^{1/p'} \left(\int_{|y| \ge \delta} |K_\varepsilon(y)| \, |f(x-y)|^p dy \right)^{1/p} \\ & \le \|K\|_1^{1/p'} \left(\varepsilon^{-n} \ell_0(\delta/\varepsilon) \right)^{1/p} \|f\|_p \ . \end{aligned}$$

Puesto que $t^n \ell_0(t) \to 0$ cuando $t \to \infty$ (ver ejercicio 18), tenemos que $J_2 \to 0$ cuando $\varepsilon \to 0$.

Resumiendo, hemos demostrado que para cada $\rho > 0$,

$$\limsup_{\varepsilon \to 0} A_\varepsilon f(x) \leq c\, c_n \rho \ .$$

Q.E.D.

7. Ejemplos y aplicaciones de aproximaciones de la identidad.

El núcleo de Poisson para el semi espacio superior.

Llamamos **núcleo de Poisson** a la siguiente función

$$P(x,t) = c_n \ \frac{t}{(|x|^2 + t^2)^{(n+1)/2}} \qquad\qquad (x \in I\!R^n,\ t > 0) \ .$$

Observemos que $P(x,t)$ es de la forma $t^{-n}K(x/t)$ con $K(x) = c_n(|x|^2+1)^{-(n+1)/2}$. El valor de la constante c_n se calcula a partir de la igualdad

$$\int K(x)dx = 1 \ .$$

Para $f \in L^p(I\!R^n)$, $1 \leq p \leq \infty$ la siguiente función se llama **integral de Poisson** de f:

$$u(x,t) = \int P(x-y,t)f(y)dy \ .$$

La función $u(x,t)$ es armónica en $I\!R^{n+1}_+ = \{(x,t) : x \in I\!R^n,\ t > 0\}$, esto es $\sum_{i=1}^n \frac{\partial^2 u}{\partial x_i^2} + \frac{\partial^2 u}{\partial t^2} = 0$, en dicho conjunto (ver ejercicio 21).

Si la función f es continua con soporte compacto, su integral de Poisson $u(x,t)$ es armónica en $I\!R^{n+1}_+$, continua en la adherencia de este conjunto y $u(x,0) = f(x)$ para cada $x \in I\!R^n$. Ver ejercicio 22.

En otras palabras, dada una función continua f, con soporte compacto, su integral de Poisson $u(x,t)$ es una solución al **problema de Dirichlet** en el semiespacio superior $I\!R^{n+1}_+$ con dato f sobre $I\!R^n$.

Si sólo suponemos que $f \in L^p$, la integral de Poisson $u(x,t)$ tenderá al dato $f(x)$ cuando $t \to 0$, siempre que x sea un punto de Lebesgue de la función f. En efecto, el teorema (8.15) se puede usar para este caso particular de aproximación de la identidad.

El núcleo del calor.

Este núcleo está dado por la función

$$\Gamma(x,t) = c_n \ t^{-n/2}\exp(-|x|^2/4t)$$

para $x \in I\!R^n$ y $t > 0$. La constante c_n se determina por la igualdad

$$c_n \int \exp(-1|x|^2/4) \ dx = 1 \ .$$

Poniendo $\Gamma(x) = c_n\exp(-|x|^2/4)$ resulta $\Gamma(x,t) = \Gamma_{\frac{1}{\sqrt{t}}}(x) = \frac{1}{t^{n/2}}\Gamma(x/\sqrt{t})$.

Para $f \in L^p(I\!R^n)$, $1 \le p \le \infty$, ponemos

$$H(x,t) = \int \Gamma(x-y,t)f(y) \ dy \ .$$

No es difícil verificar que la función $H(x,t)$ es solución de la **ecuación del calor** $\frac{\partial u}{\partial t} = \sum_{j=1}^n \frac{\partial^2 u}{\partial x_j^2}$ en la región $I\!R^{n+1}_+$. Además, si x es un punto de Lebesgue de f, entonces $H(x,t) \to f(x)$ cuando $t \to 0$.

A continuación daremos algunas aplicaciones de las aproximaciones de la identidad. Cuando se quiere verificar cierta propiedad para una función f, a veces es más conveniente demostrar dicha propiedad para **las funciones regularizadas** o **las funciones modificadas**, $f_\varepsilon = K_\varepsilon * f$, donde K es un cierto núcleo suave. Posteriormente uno espera obtener la propiedad para f haciendo $\varepsilon \to 0$.

Para $f \in L^p(I\!R)$ hemos definido su transformada de Fourier $\widehat{f}$ por medio de la fórmula

$$\widehat{f}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ixy} f(y) \ dy \ .$$

Sabemos que $\widehat{f}$ es una función continua que se anula en el infinito, ver ejercicios del capítulo VII.

La **antitransformada de Fourier** de una función integrable f la denotamos por $\check{f}$, y viene dada por la siguiente expresión

$$\check{f}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ixy} f(y) \ dy \ .$$

(8.16) **Teorema.** (Inversión de la transformada de Fourier.)
Sea $f \in L^1(\mathbb{R})$ *tal que* $\widehat{f} \in L^1(\mathbb{R})$. *Entonces la antitransformada de Fourier de* $\widehat{f}$ *coincide en casi todo punto con* f.

Demostración: Usaremos la siguiente expresión para el núcleo de Poisson

$$P_\varepsilon(x) = \frac{1}{\pi}\,\frac{\varepsilon}{\varepsilon^2 + x^2} = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\varepsilon|y|} e^{ixy}\, dy \ .$$

La igualdad anterior se obtiene integrando por partes la siguiente integral

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\varepsilon|y|} e^{ixy}\, dy = \frac{1}{\pi}\int_{0}^{\infty} e^{-\varepsilon y} \cos yx\, dy \ .$$

Usando el teorema de Fubini y la expresión de arriba para P_ε, la función

$$h_\varepsilon(u) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \widehat{f}(x) e^{-\varepsilon|x|} e^{iux}\, dx \ ,$$

se puede expresar como

$$\begin{aligned}
&\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{ixy} f(y)\, dy\right) e^{-\varepsilon|x|} e^{-ixu} dx \\
&= \int_{-\infty}^{\infty} f(y)\left(\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\varepsilon|x|} e^{ix(y-u)} dx\right) dy \\
&= P_\varepsilon * f(u) \ .
\end{aligned}$$

Como $\widehat{f}$ es integrable, $h_\varepsilon(u)$ tiende uniformemente en u a la integral

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \widehat{f}(x)\, e^{-ixu}\, dx \ ,$$

cuando $\varepsilon \to 0$.

Por el teorema (8.14) $P_\varepsilon * f(u) \to f(u)$ cuando $\varepsilon \to 0$ para casi todo u. Consecuentemente la igualdad siguiente se verifica en casi todo punto.

$$f(y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-ixy} \widehat{f}(x)\, dx \ .$$

Q.E.D.

(8.17) **Teorema.** (Titchmarsh). *Sea f una función integrable tal que*

$$\int_{-\infty}^{\infty} |f(x+t) - f(x)| \; dx = o(|t|) \; , \quad t \to 0 \; .$$

Entonces f es constante en casi todo punto.

DEMOSTRACIÓN: Suponemos primero que f tiene segunda derivada continua en $I\!R$.

Integrando sobre un conjunto E la siguiente desigualdad

$$|f'(x)t| \leq \frac{t^2}{2} \sup_{x \in E} |f''(x)| + |f(x+t) - f(x)| \; ,$$

tenemos que

$$\int_E |f'(x)| \; dx \leq m(E) \frac{t}{2} \sup_{x \in E} |f''(x)| + \frac{1}{t} \int_{-\infty}^{\infty} |f(x+t) - f(x)| \; dx \; .$$

Haciendo $t \to 0$ obtenemos $\int_E |f'(x)| dx = 0$. Por lo tanto f es una constante en este caso.

Si $f \in L^1(I\!R)$ consideramos una aproximación de la identidad $f_\varepsilon(x) = K_\varepsilon * f(x)$ con un núcleo $K \in C_0^2(I\!R)$. La función $f_\varepsilon \in L^1(I\!R)$, tiene segunda derivada continua y además cumple

$$\int_{-\infty}^{\infty} |f_\varepsilon(x-t) - f_\varepsilon(x)| dx \leq \|K\|_1 \int_{-\infty}^{\infty} |f(x+t) - f(x)| dx = o(|t|) \; .$$

Por lo tanto, para cada $\varepsilon > 0$ existe una constante c_ε tal que $f_\varepsilon = c_\varepsilon$.

Por un lado $f_\varepsilon(x)$ converge a una constante c en todo punto x . Por otro lado $f_\varepsilon(x) \to f(x)$ en casi todo x cuando $\varepsilon \to 0$. Luego, $f = c$.

Q.E.D.

Recordemos que $f : I\!R \to I\!R$ es una **función de Lipschitz** con una constante de Lipschitz c si se cumple

$$|f(x) - f(y)| \leq c|x - y| \; , \qquad x, y \in I\!R \; .$$

(8.18) **Teorema.** *Sea f una función continua sobre $I\!R$. Entonces f es una función de Lipschitz con una constante c si y sólo si se cumple*

$$\left| \int_{-\infty}^{\infty} f(x) \varphi'(x) dx \right| \leq c \int_{-\infty}^{\infty} |\varphi(x)| dx \; ,$$

para toda $\varphi \in C_0^1(\mathbb{R})$.

DEMOSTRACIÓN: Veamos primero que la condición anterior es suficiente. Sea $\varphi \in C_0^1(\mathbb{R})$ no negativa con integral igual a uno y pongamos $f_\varepsilon = \varphi_\varepsilon * f$. Para cada $t > 0$, tenemos que

$$\begin{aligned} f_\varepsilon(y+t) - f_\varepsilon(y) &= \int_{-\infty}^{\infty} f(x) \left(\varphi_\varepsilon(y+t-x) - \varphi_\varepsilon(y-x)\right) dx \\ &= \int_{-\infty}^{\infty} f(x) \int_0^t \varphi_\varepsilon'(s+y-x) ds dx \ . \end{aligned}$$

Luego

$$|f_\varepsilon(y+t) - f_\varepsilon(y)| \le \int_0^t \ c \ \int_{-\infty}^{\infty} \varphi_\varepsilon(x) dx dt = ct \ .$$

Teniendo en cuenta la continuidad de f, tenemos que $|f(y+t) - f(y)| \le ct$.

Para demostrar que la condición es necesaria supongamos, sin pérdida de generalidad, que la función f es integrable.

Observemos que si f es derivable, trivialmente se verifica

$$\left| \int_{-\infty}^{\infty} \varphi'(x) f(x) dx \right| \le \sup_x \ |f'(x)| \int_{-\infty}^{\infty} |\varphi(x)| dx \ ,$$

para toda $\varphi \in C_0^1(\mathbb{R})$.

Sea ahora $f_\varepsilon(x) = K_\varepsilon * f(x)$, donde $\{K_\varepsilon\}$ es una aproximación de la identidad con $K \in C_0^1(\mathbb{R})$, $K > 0$.

Si f tiene una constante de Lipschitz c, entonces para cada $\varepsilon > 0$ tenemos $|f_\varepsilon(x+t) - f_\varepsilon(x)| \le ct$.

Por lo tanto $|f_\varepsilon'(x)| \le c$ para cada $\varepsilon > 0$ y $x \in \mathbb{R}$. Así , usando la desigualdad de arriba, vemos que

$$\left| \int_{-\infty}^{\infty} \varphi'(x) f_\varepsilon(x) dx \right| \le c \int_{-\infty}^{\infty} |\varphi(x)| dx \ ,$$

para cada $\varphi \in C_0^1(\mathbb{R})$.

Como $f_\varepsilon \to f$ en L^1 cuando $\varepsilon \to 0$, se tiene que

$$\int_{-\infty}^{\infty} \varphi'(x) f_\varepsilon(x) dx \to \int_{-\infty}^{\infty} f(x) \varphi'(x) dx \ , \qquad \varepsilon \to 0 \ .$$

Q.E.D.

*8. Límite no tangencial.

Nos interesa puntualizar la siguiente situación geométrica para una aproximación de la identidad $\{K_\varepsilon\}$ sobre $\mathbb{R}^n$.

Sea f una función continua con soporte compacto y definamos

$$u(x,\varepsilon)=\begin{cases} K_\varepsilon * f(x) & \text{si } \varepsilon>0 \\ f(x) & \text{si } \varepsilon=0. \end{cases}$$

Esta función u es continua en el semiespacio superior

$$\overline{\mathbb{R}^{n+1}_+}=\{(x,t): x\in \mathbb{R}^n,\ t\geq 0\}.$$

Si $f\in L^p(\mathbb{R}^n)$ y $\{K_\varepsilon\}$ es de la forma $\varepsilon^{-n}K(x/\varepsilon)$ hemos demostrado, con condiciones bastantes generales para K, que $u(x,\varepsilon)$ se aproxima a $u(x,0)$ si x es un punto de Lebesgue de f, y la aproximación se realiza a lo largo de la semirrecta perpendicular a $\mathbb{R}^n$ que pasa por el punto $(x,0)$.

Será nuestra tarea en lo que sigue obtener teoremas sobre aproximaciones de la identidad, cuando el acercamiento se hace dentro de regiones más substanciosas que la recta vertical antes mencionada. A tal efecto introducimos algunas notaciones.

Sea ℓ una función con valores no negativos definida sobre $[0,\infty)$, decreciente y acotada. Para $\delta>0$ y $x\in\mathbb{R}^n$ ponemos

$$\widetilde{\ell}_\delta(x)=\sup_{|y|\leq\delta}\ell(|x-y|)\ .$$

La afirmación siguiente queda como ejercicio.

(8.19) Sean $\ell,\widetilde{\ell}_\delta$ las funciones descriptas arriba. Entonces tenemos:

(i) $\widetilde{\ell}_\delta$ es una función radial y verifica

$$\widetilde{\ell}_\delta(x)\leq \ell(0)\,\chi_{\{|x|\leq\delta\}}(0)+\ell(|x|-\delta)\,\chi_{\{|x|\geq\delta\}}(x).$$

(ii) Si llamamos $\ell_\delta(x)$ al miembro derecho de la última desigualdad, entonces ℓ_δ es radial y decreciente pensada sobre $[0,\infty)$. Además la integral

$$\int_0^\infty r^{n-1}\ell_\delta(r)dr$$

es finita cuando lo es la integral

$$\int_0^\infty r^{n-1}\ell(r)dr\ .$$

En $I\!R^{n+1}_+$ definimos **cono con vértice** $(x_0, 0)$ y **abertura** $\delta > 0$ al conjunto

$$\Gamma_\delta(x_0) = \{(x,t) \in I\!R^{n+1}_+ : |x - x_0| < t\delta\} \ .$$

Sea u una función definida sobre $I\!R^{n+1}_+$ y $\Gamma_\delta(x_0)$ un cono. Definimos **la función maximal no-tangencial** de u mediante la fórmula

$$M_\delta(u)(x_0) = \sup\{|u(x,t)| : (x,t) \in \Gamma_\delta(x_0)\} \ .$$

Si $f \in L^p$ y $\{K_\varepsilon\}$ es una aproximación de la identidad denotamos por $M_\delta(f)(x_0)$ a la función $M_\delta(u)(x_0)$, donde $u(x,\varepsilon) = K_\varepsilon * f(x)$.

(8.20) **Teorema.** *Sea K una función medible tal que su mínima mayorante radial no creciente es integrable y acotada. Entonces*

$$M_\delta(f)(x) \leq A \ Mf(x) \ ,$$

donde la constante A es independiente de la función f, pero depende de n, K y δ. Una expresión para esta constante se obtendrá durante la demostración del teorema.

Demostración: Sea ℓ la mínima mayorante radial no creciente de K y ℓ_δ la función definida en (8.19) usando la función $\widetilde{\ell_\delta}$ definida en esta sección.

Fácilmente se pueden verificar las siguientes desigualdades en la región $|y| \leq \varepsilon\delta$:

$$\begin{aligned}
|K_\varepsilon * f(x_0 + u)| &\leq \int_{I\!R^n} |K_\varepsilon(x_0 + y - x)| \ |f(x)| \ dx \\
&\leq \varepsilon^{-n} \int_{I\!R^n} \widetilde{\ell_\delta}\left(\frac{x_0 - x}{\varepsilon}\right) \ |f(x)| \ dx \\
&\leq \varepsilon^{-n} \int_{I\!R^n} \ell_\delta\left(\frac{x_0 - x}{\varepsilon}\right) \ |f(x)| \ dx \\
&= (\ell_\delta)_\varepsilon * |f|(x_0) \\
&\leq c_n \int_0^\infty r^{n-1}\ell_\delta(r) \ dr Mf(x_0) \ .
\end{aligned}$$

La siguiente es una expresión para la constante A

$$A = c_n\left(\frac{\delta^n}{n}\ell(0) + \int_0^\infty (r+\delta)^{n-1}\ell(r) \ dr\right) \ ,$$

donde c_n depende sólo de la dimensión.

De ahora en adelante supondremos que $M_\delta(f)$ es una función medible. El lector no tendrá dificultad en demostrar la medibilidad de $M_\delta(f)$ para los ejemplos dados en la sección anterior.

La siguiente afirmación es consecuencia inmediata del teorema (8.20) y las propiedades del operador maximal de Hardy-Littlewood Mf.

(8.21) **Teorema.**
(i) $\|M_\delta f\|_p \leq c_p \|f\|_p \qquad 1 < p \leq \infty$,
donde c_p *depende de* δ *pero no de* f.
(ii) $t\, m(\{M_\delta(f) > t\}) \leq c_1 \|f\|_1$.

(8.22) **Teorema.** *Sea* K *una función medible, tal que su mínima mayorante radial no creciente es integrable y acotada. Si* $1 \leq p < \infty$, $f \in L^p(I\!R^n)$ *y* x_0 *es un punto de Lebesgue de* f, *entonces para cada* $\delta > 0$ *tenemos*

$$\int |K_\varepsilon(x-y)|\; |f(y) - f(x_0)|dy \to 0 \ ,$$

cuando $(x,\varepsilon) \to (x_0, 0)$, $(x,\varepsilon) \in \Gamma_\delta(x_0)$.

Demostración: Si $u(x,\varepsilon) = K_\varepsilon * f(x)$ y $|x| \leq \varepsilon\delta$, se verifica

$$\begin{aligned}
|u(x_0 + x, \varepsilon) - u(x_0, 0)| &\leq \int |K_\varepsilon(x_0 + x - y)|\; |f(y) - f(x_0)|dy \\
&\leq \varepsilon^{-n} \int \widetilde{\ell}_\delta \left(\frac{x_0 - y}{\varepsilon} \right) |f(y) - f(x_0)|dy \\
&\leq \int (\ell_\delta)_\varepsilon(y) |f(x_0 - y) - f(x_0)|dy
\end{aligned}$$

Por (8.19) podemos usar el teorema (8.15) para $p < \infty$ y obtenemos el teorema.

Q.E.D.

Veamos que la afirmación del teorema (8.21) también es cierta para $p = \infty$. Primero observamos:

(8.23) Si $1 \leq p \leq \infty$, $f \in L^p(I\!R^n)$ y x_0 es un punto de Lebesgue de f, entonces

$$\int_{|y| \leq N} |f(x + ty) - f(x_0)|dy \to 0$$

cuando $(x,t) \to 0$, $(x,t) \in \Gamma_\delta(x_0)$. Donde $\delta > 0$ y N es un número fijo.

El límite en (8.23) se obtiene usando la definición de punto de Lebesgue y la de cono $\Gamma_\delta(x_0)$ y significa que cada punto de Lebesgue de f es también un "punto de Lebesgue no tangencial" de f.

Con la observación (8.23) en mente, es fácil repetir los argumentos del teorema (8.13), y así obtener el siguiente teorema.

(8.24) **Teorema.** *Sean $K \in L^1(\mathbb{R}^n)$, $f \in L^\infty(\mathbb{R}^n)$ y x_0 un punto de Lebesgue de la función f. Entonces tenemos*

$$\int |K_\varepsilon(x-y)| \; |f(y) - f(x_0)| dy \to 0 \;,$$

cuando $(x,\varepsilon) \to (x_0, 0)$, $(x,\varepsilon) \in \Gamma_\delta(x_0)$.

EJERCICIOS

1. La función maximal de Hardy-Littlewood Mf no es integrable sobre $\mathbb{R}^n$, si f no es cero en casi todo punto.

2. Sea (x_j) una sucesión de puntos de un conjunto X y a_j una sucesión de números reales positivos. La siguiente función define una medida sobre $\mathcal{P}(X)$
$$\mu(A) = \sum_{j=1}^{\infty} a_j \; \delta_{x_j}(A) \;,$$
donde δ_{x_j} es la delta de Dirac concentrada en x_j.

3. Demuestre el teorema (8.5). ¿Qué pasa con las desigualdades de la parte *(c)* del teorema para el caso $\mu(X) = \infty$?

4. Para cada función f localmente integrable sobre $\mathbb{R}^n$ ponemos
$$MF(f)(x) = \sup_{I \ni x} \frac{1}{m(I)} \int_I |f(y)| dy \;,$$
donde I es un intervalo acotado en $\mathbb{R}^n$ que contiene al punto x. Demuestre que si $1 < p \leq \infty$ tenemos
$$\|MF(f)\|_p \leq c_p \; \|f\|_p \qquad\qquad (f \in L^p(\mathbb{R}^n)) \;,$$

donde c_p es una constante independiente de f.
La función $MF(f)$ definida arriba se conoce como **función maximal fuerte**.
Sugerencia: Considere, por ejemplo, $n = 2$. El promedio de una función sobre un rectángulo en el plano está acotado por la iteración de dos funciones maximales unidimensionales. Use que la función maximal de Hardy- Littlewood es de tipo fuerte (p, p), $1 < p \leq \infty$.

5. El siguiente ejemplo demuestra que el operador maximal fuerte MF no es de tipo débil $(1, 1)$. Para $1 < k$ y $1 \leq j \leq k$ ponemos $I_{k,j} = [0, j] \times [0, k/j]$, $I_k = \bigcup_{j=1}^{k} I_{k,j}$, $I = \bigcap_{j=1}^{k} I_{k,j}$.
Si $f = \chi_I$, entonces claramente se tiene
$$I_k \subseteq \{MF(f) > 1/2k\} .$$
Por lo tanto
$$k^{-1}\ m(\{MF(f) > 1/k\}) \to \infty, \quad \text{cuando } k \to \infty .$$

6. Sea $\|\cdot\|$ una norma sobre $\mathbb{R}^n$ y $B_r(x) = \{y \in \mathbb{R}^n : \|x - y\| < r\}$. Entonces tenemos $m(B_r(x)) = r^n\ m(B_1(0))$. Luego se cumple $m(B_{2r}(x)) = 2^n\ m(B_r(x))$.

7. Sea $0 < \alpha \leq 1$ y $B_r(x)$ la bola
$$\{(y_1, \ldots, y_n) : \sum_{i=1}^{n} |x_i - y_i|^\alpha < r\} .$$
Entonces, existe una constante $c = c(n, \alpha)$ tal que $m(B_{2r}(x) \leq c\ m(B_r(x))$, para todo $x \in \mathbb{R}^n$ y $r > 0$. ¿Puede generalizar este resultado para otras funciones cóncavas $\varphi(t)$ distintas de t^α?

8. Sea φ una función convexa sobre $[0, \infty)$ tal que $\varphi(0) = 0$ y $\varphi(x) > 0$ para $x > 0$. Para $\lambda > 0$ definimos la siguiente elipse
$$C_\lambda = \{(x, y) \in \mathbb{R}^2 : \frac{x^2}{\lambda^2} + \frac{y^2}{\varphi^2(\lambda)} = 1\} .$$
Si $P \in \mathbb{R}^2$, $P \neq 0$, entonces denotamos por $\|P\|$ el único número $\lambda > 0$ tal que $P \in C_\lambda$. Cuando $P = 0$, ponemos $\|P\| = 0$.

(a) Demuestre que $d(P,Q) = \|P - Q\|$ es una métrica sobre $I\!R^2$ que genera los mismos abiertos que la métrica euclídea.
(b) Sea $B_r(P)$ la bola centrada en P de radio r generada por la métrica anterior. Supongamos que existe una constante c tal que $\varphi(2x) \le c\,\varphi(x)$ para cada $\alpha \in I\!R^2$. Entonces $m(B_{2r}(P)) \le 2c\; m(B_r(P))$.
(c) Encuentre funciones convexas φ tales que $m(B_{2r}(0))/m(B_r(0)) \to \infty$ cuando $r \to \infty$ $(r \to 0)$.
Sugerencias:
(a) Para demostrar la desigualdad triangular pongamos: $\varepsilon_1 = \|P\|$, $\varepsilon_2 = \|Q\|$ y $(s,t) = P + Q$. Entonces

$$\left(\frac{s}{\varepsilon_1 + \varepsilon_2}\right)^2 + \left(\frac{t}{\varphi(\varepsilon_1 + \varepsilon_2)}\right)^2 \le 1\ .$$

En la demostración de la desigualdad anterior use la convexidad de la función t^2 y además que $(\varphi(\varepsilon_1)/\varphi(\varepsilon_1 + \varepsilon_2)) \le \varepsilon_1/(\varepsilon_1 + \varepsilon_2)$.
(b) Observe que $B_r(0)$ es precisamente el interior de la elipse C_r.

9. Demuestre el teorema (8.8).

10. Demuestre el teorema (8.10) para el caso $r = \infty$.

11. La siguiente desigualdad, *(iv)* en teorema (8.11), $\|Mf\|_{L^1(E)} \le 2m(E) + 2c_1 \int_{I\!R^n} |f(x)| \log^+ |f(x)| dx$, puede ser reformulada usando la norma de Luxemburg $\|\cdot\|_\varphi$ con $\varphi(t) = t\ \log^+ t$.
Demuestre que para todo $f \in L^\varphi$ se tiene

$$\|Mf\|_{L^1(E)} \le 2(m(E) + c_1)\|f\|_\varphi\ .$$

La desigualdad anterior dice que el operador maximal de Hardy-Littlewood es un operador continuo de $L^\varphi(I\!R^n)$ en $L^1(E)$, si $m(E) < \infty$.
Sugerencia: Observe que si $\int \varphi(|f|)dx = 1$, entonces $Mf \le 2(m(E) + c_1)$.

12. La condición (I_2) en la definición de una aproximación de la identidad $\{K_\varepsilon\}_{\varepsilon>0}$ puede ser reemplazada por $(I_2')\int K_\varepsilon(x)dx \to 1$ cuando $\varepsilon \to 0$. Demuestre que el teorema (8.12) es válido con esta nueva condición.

13. Sea (E_ε) una familia de conjuntos en $I\!R^n$ medibles y acotados. Supongamos que para cada $\delta > 0$ existe ε_0 tal que $0 < \varepsilon \le \varepsilon_0$ implica $E_\varepsilon \subset B_\delta(0)$.

Demuestre que $K_\varepsilon = [m(E_\varepsilon)]^{-1}\chi_{E_\varepsilon}$ es una aproximación de la identidad.

14. Sea $\{K_\varepsilon\}_{\varepsilon>0}$ una aproximación de la identidad y f una función medible acotada. Para $F \subseteq I\!R^n$ y $\delta > 0$ ponemos

$$\Omega_F(f,\delta) = \sup_{x\in F, |y|\leq\delta} |f(x-y) - f(x)| \ .$$

Verifique que

$$\sup_{x\in F} |f(x) - K_\varepsilon * f(x)| \leq c\ \Omega_F(f,\delta) + 2\|f\|_\infty \int_{|x|>\delta} |K_\varepsilon(x)|dx \ ,$$

donde c es la constante que aparece en la definición de una aproximación de la identidad.
Convengamos, para este ejercicio, en decir que f es uniformemente continua sobre una región F si $\Omega_F(f,\delta)\to 0$ cuando $\delta \to 0$. Por lo tanto, la última desigualdad nos asegura que f_ε tiende a f uniformemente en la región donde f es uniformemente continua.

15. Existe una función $K \in C_0^\infty(I\!R^n)$ no negativa y de integral igual a uno.

Sugerencia: considere la función $\exp(1/(|x|^2-1))$ sobre $|x|\leq 1$ e igual a cero fuera de la bola unidad.

16. El conjunto $C_0^\infty(I\!R^n)$ es denso en $L^p(I\!R^n)$, $1\leq p<\infty$.

17. Sea $F \subseteq I\!R^n$ un conjunto compacto y G un abierto que lo contiene. Entonces existe una función $\varphi \in C_0^\infty(I\!R^n)$ indénticamente igual a 1 sobre F, con soporte contenido en G y $0 \leq \varphi \leq 1$.
Sugerencia: Sea G_1 un conjunto abierto acotado tal que $K \subseteq G_1 \subseteq \overline{G}_1 \subseteq G$. Para $K \in C_0^\infty(B_1(0))$, con integral igual a uno ponemos $f_\varepsilon = \chi_{\overline{G}_1} * K_\varepsilon$. Es fácil ver que existe ε_0 para el cual f_{ε_0} cumple las condiciones requeridas.

18. Sea ℓ una función nonegativa, decreciente en $(0,\infty)$, que verifica

$$\int_0^\infty r^{n-1}\ell(r)dr < \infty \ .$$

Entonces $r^n\ell(r) \to 0$ cuando $r\to 0$ o $r\to\infty$.

19. Sea ℓ una función como en el ejercicio 18. Entonces para cada $j \geq 1$ existe una función $\ell_j \in C^j(0,\infty)$ tal que

$$\ell(r) \leq \ell_j(r) \leq \ell(r/2^j) \ , \qquad\qquad r > 0 \ .$$

Más aún, la función ℓ_j puede ser elegida decreciente en $(0,\infty)$.
Sugerencia: Para $j = 1$ considere la función definida por

$$\ell_1(x) = \frac{n2^n}{2^n - 1} \, \frac{1}{x^n} \int_{\frac{x}{2}}^{x} t^{n-1} \, \ell(t) \, dt \ .$$

Para analizar la monotonía de $\ell_1(x)$ es conveniente hacerlo suponiendo primero que $\ell(x) = a\,\chi_{(0,b)}(x)$, y luego aproximar la función por sumas de funciones de este último tipo.

20. Sea $K \in L^p(I\!R^n)$, $1 < p < \infty$, una función son soporte compacto. Pongamos

$$Tf(x) = \sup_{\varepsilon > 0} |K_\varepsilon * f(x)| \ ,$$

con $K_\varepsilon(x) = \varepsilon^{-n} K(x/\varepsilon)$. Demostrar:
(a) Si $f \in L^q(I\!R^n)$, $p/(p-1) \leq q$, entonces la función Tf es medible. En efecto, para cada $\varepsilon > 0$, $K_\varepsilon * f$ es una función continua.
(b) Para toda $f \in L^q(I\!R^n)$, $p/(p-1) \leq q \leq \infty$ se tiene que

$$\lim_{\varepsilon \to 0} K_\varepsilon * f(x) = f(x) \int K(x) \, dx \ ,$$

para casi todo $x \in I\!R^n$.
Sugerencia: Dado que el soporte de K es acotado, por la desigualdad de Hölder se tiene

$$|K_\varepsilon * f(x)| \leq c \|K\|_p (M|f|^{p'}(x))^{1/p'} \ ,$$

donde $\frac{1}{p} + \frac{1}{p'} = 1$, donde la constante c es independiente de f y M es el operador maximal de Hardy-Littlewood.

21. El núcleo de Poisson $P(x,t) = c_n \ t \ (|x|^2 + t^2)^{-(n+1)/2}$ es una función armónica en la región $I\!R^{n+1}_+$. Las derivadas del núcleo de Poisson son funciones integrables sobre $I\!R^n$, para cada $t > 0$ fijo.

22. Sea f una función continua con soporte compacto y $u(x,t)$ su integral de Poisson. Demuestre que la función $u(x,t)$ es armónica en $I\!R^{n+1}_+$, continua en la adherencia de este conjunto y $u(x,0) = f(x)$ para cada $x \in I\!R^n$.

CAPITULO IX

EL TEOREMA DE RADON-NIKODYM

1. La integral en espacios abstractos.

En la sección 3 del capítulo VIII hemos definido el concepto de espacio de medida, formado por una terna (X, Σ, μ) en la que X es un conjunto, Σ una σ-álgebra de subconjuntos de X y μ una medida sobre Σ.

Diremos que una función $f : X \to \overline{I\!R}$ **es medible** (con respecto a Σ) si para cada número real a, el conjunto

$$\{f > a\} = \{x \in X : f(x) > a\}$$

pertenece a la σ-álgebra Σ. En tal caso también se dice que f is Σ-medible.

Las funciones con valores en $I\!R$ medibles con respecto a Σ forman un **álgebra** de funciones (clase cerrada bajo suma, producto y multiplicación por constantes) que es cerrada bajo supremo e ínfimo de sucesiones y por consiguiente cerrada para el límite puntual de sucesiones.

Si consideramos la clase de funciones medibles con valores en $\overline{I\!R}$, valen consideraciones similares a las enunciadas para funciones medibles con valores en $I\!R$ excepto en lo referente a la suma de funciones medibles. En efecto la suma $f + g$ de dos funciones medibles con valores en $\overline{I\!R}$ no está definida en el conjunto medible

$$D = \{f = +\infty \text{ y } g = -\infty\} \cup \{f = -\infty \text{ y } g = +\infty\} .$$

Es fácil demostrar que si f y g son dos funciones medibles con valores en $\overline{I\!R}$ y $f+g$ se define arbitrariamente como una función medible sobre D, entonces $f+g$ es medible (por ejemplo si $f+g$ es una función constante sobre D).

Las demostraciones son completamente análogas a las del capítulo IV y por esta razón se dejan a cargo del lector. Todos los conceptos introducidos en aquel capítulo, por ejemplo, función característica, función simple, convergencia en casi todo punto, convergencia en medida, pueden desarrollarse en la misma forma en esta situación más general, y además se mantienen válidos entre otros los teoremas (4.9) y (4.15), con idénticas demostraciones. En particular, diremos que una propiedad $P(x)$ se verifica para casi todo x (con respecto a μ) si el conjunto E de los puntos donde $P(x)$ es falsa cumple $\mu(E)=0$.

El concepto más importante es el de **integral** de una función medible f sobre un conjunto $E \in \Sigma$. Si f es no negativa, la **integral** de f sobre el conjunto E con respecto a la medida μ se define, análogamente, por medio de la fórmula

$$\int_E f\ d\mu = \sup \sum_{i=1}^{N} v_i\ \mu(E_i)\ ,$$

donde el supremo se toma sobre todas las particiones finitas del conjunto E en conjuntos medibles disjuntos $E_1, E_2, \ldots, E_N$ y v_i denota el ínfimo de f sobre E_i.

La integral de una función medible f con valores de distinto signo, se define como en el capítulo V, por medio de la fórmula

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

siempre que las dos integrales del miembro derecho sean finitas, en cuyo caso decimos que f es **integrable** sobre E.

A veces la integral se denota por el símbolo

$$\int_E f(x)d\mu(x) = \int_E f d\mu.$$

Cuando $E=X$ suele omitirse el símbolo E en el extremo inferior del símbolo integral. Así ,

$$\int f(x)d\mu(x) = \int f d\mu$$

representa la integral de f sobre X.

Los teoremas de linealidad y convergencia desarrollados en las secciones 1 a 7 del capítulo V mantienen su validez.

Con algún detalle analizaremos la validez de los teoremas de Fubini y Tonelli para la integral en espacios abstractos.

Sean (X, Σ) e (Y, Γ) un par de espacios medibles, i.e. conjuntos abstractos con sus correspondientes σ-álgebras de conjuntos. Sobre el producto cartesiano $X \times Y$ definimos la **$\boldsymbol{\sigma}$-álgebra producto de $\boldsymbol{\Sigma}$ por $\boldsymbol{\Gamma}$** como la mínima σ-álgebra que contiene a todos los conjuntos $A \times B$ con A en Σ y B en Γ; para esta σ-álgebra usamos la notación $\Sigma \otimes \Gamma$. Los conjuntos R de la forma $R = A \times B$ con $A \in \Sigma$ y $B \in \Gamma$ se llaman rectángulos medibles.

Seguidamente definiremos una **medida producto** $\mu \otimes \nu$ sobre $\Sigma \otimes \Gamma$, siendo μ y ν medidas sobre Σ y Γ respectivamente, de modo tal que se cumpla

$$\mu \otimes \nu(R) = \mu(A)\ \nu(B)$$

para cada rectángulo medible $R = A \times B$.

Para evitar complicaciones innecesarias, cada vez que consideremos un espacio de medida abstracto (X, Σ, μ), supondremos que éste es **$\boldsymbol{\sigma}$-finito**, es decir, que existe una sucesión creciente (X_n) de conjuntos en Σ tal que $X = \bigcup_{n=1}^{\infty} X_n$ y $\mu(X_n)$ es finito para cada n.

La existencia de la medida producto se obtendrá como una consecuencia del siguiente resultado. El lector debería meditar sobre las diferencias entre los resultados (5.23) y (9.1).

(9.1) **Teorema.** *Sea E un conjunto medible en la σ-álgebra producto $\Sigma \otimes \Gamma$. Entonces*

(i) $E_x = \{y : (x, y) \in E\} \in \Gamma$ para todo $x \in X$.

(ii) $E^y = \{x : (x, y) \in E\} \in \Sigma$ para todo $y \in Y$.

(iii) $\nu(E_x)$ y $\mu(E^y)$ son funciones medibles sobre X e Y, respectivamente.

(iv) $\int_X \nu(E_x)\ d\mu(x) = \int_Y \mu(E^y)\ d\nu(y)$.

Demostración: Ver el ejercicio 6.

Veamos que el teorema (9.1) nos permite construir una medida producto $\mu \otimes \nu$ sobre $\Sigma \otimes \Gamma$. En efecto, es fácilmente verificable que si $E = A \times B$ es un rectángulo medible, ambos miembros de (iv) nos dan el mismo resultado

$\mu(A)\ \nu(B)$. Además la función de conjunto

$$\gamma(E) = \int_X \nu(E_x)\ d\mu(x) = \int_Y \mu(E^y)\ d\nu(y)\ ,$$

es una medida σ-finita sobre $\Sigma \otimes \Gamma$.

El siguiente teorema nos asegura la unicidad de la medida producto.

(9.2) **Teorema.** *La medida producto es única. Más precisamente, dados dos espacios de medida σ-finitos (X, Σ, μ) e (Y, Γ, ν) existe a lo sumo una medida σ-finita γ sobre la σ-álgebra producto $\Sigma \otimes \Gamma$, tal que para cada rectángulo medible $R = A \times B$ verifica $\gamma(R) = \mu(A)\ \nu(B)$.*

DEMOSTRACIÓN: Para entender la demostración sugerimos primero hacer los ejercicios 2 a 5 del presente capítulo.

Sean γ_1 y γ_2 dos medidas σ-finitas sobre $\Sigma \otimes \Gamma$ que tienen el mismo valor sobre cada rectángulo medible. Tomemos dos sucesiones (X_n), (Y_n) de conjuntos medibles tales que $X = \bigcup_n X_n$, $Y = \bigcup_n Y_n$ y para cada n ambas medidas $\mu(X_n)$ y $\nu(Y_n)$ son finitas.

Para n fijo definimos

$$\mathcal{F}_n = \{E \in \Sigma \otimes \Gamma : \gamma_1(E \cap X_n \times Y_n) = \gamma_2(E \cap X_n \times Y_n)\}\ .$$

Se ve fácilmente que $\mathcal{F}_n$ es una clase monótona que contiene a la semiálgebra de los rectángulos medibles y también el álgebra generada por ellos, por lo tanto $\mathcal{F}_n$ es la σ-álgebra producto $\Sigma \otimes \Gamma$. Luego para cada n y $E \in \Sigma \otimes \Gamma$ tenemos

$$\gamma_1(E \cap X_n \times Y_n) = \gamma_2(E \cap X_n \times Y_n)\ .$$

Haciendo n tender a ∞ obtenemos $\gamma_1(E) = \gamma_2(E)$.

Q.E.D.

En definitiva los teoremas (9.1) y (9.2) aseguran que la medida producto $\mu \otimes \nu$ existe, es única y está dada por la expresión

$$\mu \otimes \nu(E) = \int_X \nu(E_x)\ d\mu(x) = \int_Y \mu(E^y)\ d\nu(y)\ .$$

Si $f(x, y)$ es una función definida sobre $X \times Y$ con valores en $\overline{R}$, usaremos la notación $f(x, \cdot)$ para indicar la función $y \mapsto f(x, y)$ de Y

en $\overline{R}$, llamada la **sección de $\boldsymbol{f}$ en $\boldsymbol{x}$**. Análogamente se define la **sección de $\boldsymbol{f}$ en $\boldsymbol{y}$**, $f(\,\cdot\,, y)$, que es una función de X en $\overline{R}$.

El siguiente teorema es una inmediata extensión del teorema (9.1).

(9.3) **Teorema de Fubini-Tonelli.** *Sean (X, Σ, μ) e (Y, Γ, ν) dos espacios de medida σ-finitos. Sea $f(x, y)$ una función no negativa medible con respecto a la σ-álgebra producto $\Sigma \otimes \Gamma$. Entonces*
(i) $f(x, \cdot\,)$ es Γ-medible para cada $x \in X$.
(ii) $f(\,\cdot\,, y)$ es Σ-medible para cada $y \in Y$.
(iii) las funciones

$$\int_Y f(\,\cdot\,, y) d\nu(y), \int_X f(x, \cdot\,) d\mu(x)$$

son Σ y Γ-medibles, respectivamente.
(iv)

$$\begin{aligned}\int_X \left(\int_Y f(x, y)\ d\nu(y) \right) d\mu(x) &= \int_Y \left(\int_X f(x, y)\ d\mu(x) \right) d\nu(y) \\ &= \int_{X \times Y} f(x, y)\ d(\mu \otimes \nu)(x, y)\ .\end{aligned}$$

Demostración: Poniendo

$$E = \{(x, y) : f(x, y) > a\}\ ,$$

las afirmaciones *(i)* y *(ii)* son consecuencia directa de las respectivas afirmaciones del teorema (9.1).

Sea $\mathcal{F}$ la clase de funciones no negativas y $\Sigma \otimes \Gamma$-medibles que cumplen las afirmaciones *(iii)* y *(iv)* del teorema. Claramente $\mathcal{F}$ contiene a las funciones simples no negativas y esta clase es cerrada para el límite puntual creciente. Por lo tanto $\mathcal{F}$ es precisamente la clase de las funciones $\Sigma \otimes \Gamma$-medibles y no negativas.

Q.E.D.

Conocido el teorema de Fubini-Tonelli o también llamado teorema de Tonelli, el siguiente teorema se obtiene fácilmente repitiendo lo realizado en (5.25).

(9.4) **Teorema de Fubini.** *Sean* (X, Σ, μ) *e* (Y, Γ, ν) *dos espacios de medida* σ*-finitos y* f *una función integrable en el espacio de medida* $(X \times Y, \Sigma \otimes \Gamma, \mu \otimes \nu)$*. Entonces*
(i) para casi todo x*, la función* $f(x, \cdot)$ *es una función integrable en el espacio de medida* (Y, Γ, ν)*.*
(ii) para casi todo y*, la función* $f(\cdot, y)$ *es una función integrable en el espacio de medida* (X, Σ, μ)*.*
(iii) las funciones

$$\int_Y f(\cdot, y)\ d\nu(y), \qquad \int_X f(x, \cdot)\ d\mu(x)$$

son funciones integrables en X *e* Y *respectivamente.*
(iv)

$$\begin{aligned} \int_{Y \times X} f(x,y)\ d(\mu \times \nu)(x,y) &= \int_Y \left(\int_X f(x,y)\ d\mu(x) \right) d\nu(y) \\ &= \int_X \left(\int_Y f(x,y)\ d\nu(y) \right) d\mu(x)\ . \end{aligned}$$

Todos los conceptos y propiedades del capítulo VII permanecen válidos enunciados en un espacio de medida abstracto (X, Σ, μ). Por ejemplo, para $1 \leq p < \infty$, decimos que una función f, definida sobre X y medible con respecto a Σ, pertenece al espacio $L^p(\mu) = L^p(X, \Sigma, \mu)$ si $|f|^p$ es integrable sobre todo el espacio con respecto a la medida μ.

El espacio $L^p(\mu)$ es un espacio de Banach con la norma

$$\|f\|_p = \left(\int |f|^p\ d\mu \right)^{1/p} .$$

El espacio L^∞ se define de la misma manera que se hizo en $I\!R^n$. En espacios de medida abstractos valen las correspondientes versiones de las desigualdades de Minkowski, Hölder y Jensen y también es cierto el importante teorema (7.10). Para la demostración de los resultados citados no se necesitan nuevas ideas.

Más adelante completaremos en este capítulo la demostración del teorema (7.11).

2. El teorema de Radon-Nikodym.

Recordemos dos conceptos ya definidos en el capítulo VIII en una situación particular. Sea Σ una σ-álgebra de subconjuntos de X y supongamos definidas dos medidas μ y ν sobre Σ. Decimos que ν **es absolutamente continua** con respecto a μ, y escribimos $\nu \ll \mu$ si y sólo si $\nu(E) = 0$ cada vez que $\mu(E) = 0$. Las medidas μ y ν son **mutuamente singulares**, y denotamos este hecho por $\nu \perp \mu$, si y sólo si existen conjuntos medibles disjuntos A y B tales que $\nu(E) = \nu(E \cap A)$ y $\mu(E) = \mu(E \cap B)$ para cada $E \in \Sigma$, o de otra manera si y sólo si μ y ν están concentradas en conjuntos disjuntos. La suma de dos medidas λ y μ sobre Σ se define naturalmente por la fórmula

$$(\lambda + \mu)(E) = \lambda(E) + \mu(E) \qquad\qquad (E \in \Sigma).$$

Por lo tanto para cualquier función medible no negativa f, se cumple

$$\int f \; d(\lambda + \mu) = \int f \; d\lambda + \int f \; d\mu,$$

donde las integrales se extienden a todo espacio X. La fórmula mantiene su validez si $f \in L^1(\lambda + \mu)$.

Como hemos visto, el ejemplo típico de dos medidas absolutamente continuas es el siguiente: si (X, Σ, μ) es un espacio de medida y f una función medible no negativa sobre X, poniendo

$$\mu_f(E) = \int_E f \; d\mu \qquad\qquad (E \in \Sigma) \;,$$

se tiene $\mu_f \ll \mu$.

(9.5) **Teorema de Radon-Nikodym.** *Sean μ y ν dos medidas finitas sobre el espacio medible (X, Σ). Entonces existe un único par de medidas ν_a, ν_s sobre Σ tal que:*
(1) $\nu = \nu_a + \nu_s$, $\quad \nu_a \ll \mu$, $\quad \nu_s \perp \mu$.
(2) existe una única función $h \in L^1(\mu)$ que verifica

$$\nu_a(E) = \int_E h(x) \; d\mu(x) \qquad\qquad (E \in \Sigma) \;.$$

El par (ν_a, ν_s) se llama la ***descomposición de Lebesgue*** *de ν con respecto a μ. Además ν_a y ν_s son mutuamente singulares.*

DEMOSTRACIÓN: La demostración de la unicidad de ν_a, ν_s y h en *(1)* y *(2)* es un sencillo ejercicio que dejamos a cargo del lector. Nosotros nos ocuparemos de demostrar la existencia de la descomposición. La funcional lineal que a cada función f le hace corresponder $\int f\ d\nu$ es continua sobre $L^2(\mu+\nu)$, por lo tanto existe g en este espacio tal que

$$\int f\ d\nu = \int f\ g\ d(\mu+\nu) \qquad\qquad (f \in L^2(\nu+\mu))\ .$$

Dado que para cada $E \in \Sigma$ se verifica la desigualdad

$$0 \le \int_E g\ d(\mu+\nu) \le \mu(E) + \nu(E)\ ,$$

obtenemos inmediatamente que $0 \le g \le 1$ en casi todo punto con respecto a la medida $\mu+\nu$, por lo tanto supondremos que siempre g toma valores en $[0,1]$.

La igualdad anterior la escribimos de la forma

$$\int f(1-g)\ d\nu = \int f\ g\ d\mu \qquad\qquad (f \in L^2(\nu+\mu))\ .$$

Poniendo $f = (1+g+\ldots+g^n)\chi_E$, tenemos

$$(a) \qquad \int_E (1-g^{n+1}) d\nu = \int_E g(1+g+\ldots+g^n) d\mu \qquad (E \in \Sigma)\ .$$

La función h del teorema se obtiene poniendo $h(x) = \sum_{n=1}^{\infty} g^n(x)$. Esta última serie es convergente salvo sobre el conjunto $B = \{g = 1\}$, pero $\mu(B) = 0$ por *(a)*, luego $h \in L^1(\mu)$ y verifica

$$(b) \qquad \nu(E \cap A) = \int_{E\cap A} h(x)\ d\mu(x) \qquad (E \in \Sigma)\ ,$$

donde hemos puesto $A = \{g < 1\}$.

Ahora si ponemos $\nu_a(E) = \nu(A \cap E)$ y $\nu_s(E) = \nu(B \cap E)$, claramente se tiene $\nu = \nu_a + \nu_s$ y $\nu_a \perp \nu_s$.

La parte *(2)* del teorema es consecuencia de (b) y como $\mu(B) = 0$ se tiene que $\nu_s \perp \mu$.

Q.E.D.

La demostración anterior se debe esencialmente a J. von Neumann, 1903-1957.

El teorema (9.5) admite generalizaciones que son casi inmediatas. Así el teorema es válido cuando μ es una medida σ-finita y ν una medida finita. En efecto, basta considerar X como una unión disjunta de conjuntos X con $\mu(X_n)$ finita y usar el teorema en cada X_n.

La mayor parte del teorema se puede extender al caso de que ambas medidas μ y ν sean σ-finitas. La única parte del teorema que no vale en su totalidad es la afirmación *(2)*, donde ahora la función h no es necesariamente integrable sobre todo X, pero sí lo es sobre cada X_n, donde $\bigcup X_n = X$ y $\mu(X_n) < \infty$.

3. Medidas signadas.

Siendo f una función con valores reales, integrable sobre un espacio de medida (X, Σ, μ), pongamos

$$\mu_f(E) = \int_E f \; d\mu \qquad\qquad (E \in \Sigma) \ .$$

La función μ_f está definida sobre Σ con valores en los reales y dado que es una función σ-aditiva decimos que μ_f es una **medida real** o **medida signada**.

Por cierto, si hubiéramos partido de una función f con valores en los números complejos, μ_f sería una función σ-aditiva con valores complejos y por lo tanto un ejemplo de **medida compleja**. En cualquiera de los dos casos, la función μ_f cumple

$$|\mu_f|(E) \le \int_E |f| \; d\mu = \mu_{|f|}(E) \ ,$$

donde hemos usado $|\mu_f|(E)$ para denotar la expresión

$$\sup \left\{ \sum_{i=1}^{\infty} |\mu_f(E_i)| : \bigcup_{i=1}^{\infty} E_i = E, \quad E_i \text{ medibles y disjuntos} \right\} \ .$$

Más aún, si f es una función real es fácil ver que $|\mu_f|(E) = \int_E |f| \, d\mu$, por lo tanto $|\mu_f|$ es una medida finita sobre Σ.

Motivados por el ejemplo de arriba daremos la siguiente definición. Una función $\mu : \Sigma \to I\!R$ se llama una **medida real** o **medida signada** si es σ-aditiva, es decir,

$$\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu(E_i) \ ,$$

para cada sucesión disjunta (E_i) incluida en Σ.

Es importante notar que la serie $\sum_{i=1}^{\infty} \mu(E_i)$ es absolutamente convergente. En efecto, el valor de esta serie no se altera por un reordenamiento arbitrario de sus términos.

En forma análoga se define el concepto de medida compleja.

Dada una medida signada μ sobre Σ, para cada $E \in \Sigma$ ponemos

$$|\mu|(E) = \sup\left\{ \sum_{i=1}^{\infty} |\mu(E_i)| : \bigcup_{i=1}^{\infty} E_i = E, \quad (E_i) \subset \Sigma, \text{ disjunta} \right\} .$$

La función $|\mu|$ así definida se llama la **variación total** de μ.

Se puede demostrar con relativa facilidad (ver ejercicios) que la variación total del μ es una medida sobre Σ. Más aún, la variación total de μ es una medida finita. Este hecho es menos evidente y lo demostraremos a conti nuación.

(9.6) **Teorema.** *La variación total de una medida signada es una medida finita.*

DEMOSTRACIÓN: Suponemos ya demostrado que la variación total $|\mu|$ es una medida; luego resta ver que $|\mu|(X) < \infty$.

Demostraremos primero que si para algún conjunto medible E tenemos $|\mu|(E) = \infty$, entonces existen conjuntos medibles disjuntos A y B tal que $A \cup B = E$, $|\mu(A)| > 1$ y $|\mu|(B) = \infty$. Bastará encontrar conjuntos disjuntos A y B con $A \cup B = E$, tales que $|\mu(A)| > 1$ y $|\mu(B)| > 1$, pues la variación total es una medida y la variación total de μ sobre alguno de los dos conjuntos es infinita.

Para probar la existencia de A y B consideremos una partición (E_i) de E tal que

$$\sum_{i=1}^{\infty} |\mu(E_i)| > 2(1 + |\mu(E)|) \ .$$

Luego para algún n se verifica

$$\sum_{i=1}^{n} |\mu(E_i)| > 2(1 + |\mu(E)|) \ .$$

Es fácil ver que existe un conjunto de índices $S \subset \{1, 2, 3, \ldots, n\}$ tal que

$$2\left|\sum_{i \in S} \mu(E_i)\right| \geq \sum_{i=1}^{n} |\mu(E_i)| \ .$$

Pongamos $A = \bigcup_{i \in S} E_i$. Claramente $|\mu(A)| > 1$ y si $B = E - A$ se verifica $|\mu(B)| = |\mu(E) - \mu(A)| \geq |\mu(A)| - |\mu(E)| > (1 + |\mu(E)|) - |\mu(E)| = 1$.

Ahora es fácil concluir la demostración del teorema: supongamos que $|\mu|(X) = \infty$. Luego existen conjuntos E_1 y F_1 que parten a X con $|\mu|(E_1) = \infty$ y $|\mu|(F_1) > 1$. Pero como $|\mu|(E_1) = \infty$, existe una partición de E_1, digamos $E_1 = E_2 \cup F_2$, tal que $|\mu|(E_2) = \infty$ y $|\mu(F_2)| > 1$. Así siguiendo, construimos una sucesión (F_i) de conjuntos disjuntos tal que $|\mu(F_i)| > 1$ para todo i. Esto último no es posible pues la serie $\sum_{i=1}^{\infty} \mu(F_i)$ es convergente.

Q.E.D.

Para una medida signada μ definimos, con ayuda de la variación total $|\mu|$, las siguientes medidas:

$$\mu^+ = (|\mu| + \mu)/2 \qquad\qquad \mu^- = (|\mu| - \mu)/2 \ .$$

Observe que $\mu = \mu^+ - \mu^-$ y $|\mu| = \mu^+ + \mu^-$.

Las medidas μ^+ y μ^- son no negativas y finitas y la fórmula $\mu = \mu^+ - \mu^-$ se conoce como la **descomposición de Jordan** de la medida μ.

Definimos la integral f con respecto a una medida signada μ mediante la expresión:

$$\int f \ d\mu = \int f \ d\mu^+ - \int f \ d\mu^- \ ,$$

si f es integrable simultáneamente con respecto a μ^+ y μ^-. Nótese que con la definición anterior tenemos que si

$$\mu_f(E) = \int_E f \ d\mu \ ,$$

entonces se verifica
$$|\mu_f(E)| \leq \int_E |f| \; d|\mu| \; .$$
Por lo tanto si μ es una medida signada y $f \in L^1(|\mu|)$, entonces la función μ_f definida arriba es nuevamente una medida real y verifica
$$|\mu_f|(E) \leq \int_E |f| \; d|\mu| \qquad\qquad (E \in \Sigma) \; .$$
Más aún, en la desigualdad anterior vale la igualdad para todo $E \in \Sigma$ (ver ejercicio 12).

El teorema (9.5) se puede extender fácilmente al caso en que la medida μ es finita y la medida ν es signada. La noción de mutua singularidad para medidas signadas tiene el mismo sentido que para medidas, y el concepto de ν absolutamente continua respecto de μ se transcribe textualmente para el caso de que μ sea una medida y ν una medida signada.

Veamos una primera aplicación del teorema de Radon-Nikodym.

Sea (X, Σ) un espacio medible y μ una medida signada sobre Σ. De acuerdo con el ejercicio 11 existe una función h tal que
$$(c) \qquad\qquad \mu(E) = \int_E h \; d|\mu| \qquad\qquad (E \in \Sigma) \; .$$
y además $|h| = 1$ salvo en un conjunto de medida $|\mu|$ nula. Luego supondremos $|h| = 1$ sobre todo el espacio X. Poniendo
$$P = \{h = 1\} \qquad\qquad \text{y} \qquad\qquad N = \{h = -1\} \; ,$$
tenemos que
$$\begin{array}{lll} \mu(A) \geq 0 & \qquad \text{si} \qquad & A \subseteq P \quad \text{y} \\ \mu(A) \leq 0 & \qquad \text{si} \qquad & B \subseteq N \; . \end{array}$$
Por razones obvias, el conjunto P es llamado **un conjunto positivo** para μ y N **un conjunto negativo** para μ. Las medidas definidas por:
$$\begin{aligned} \mu_1(E) &= \mu(E \cap P) \\ \mu_2(E) &= -\mu(E \cap N) \; , \end{aligned}$$
son mutuamente singulares y además se verifica $\mu_1 - \mu_2 = \mu$, $\mu_1 + \mu_2 = |\mu|$. El lector no tendrá dificultad en probar que $\mu^+ = \mu_1$ y $\mu^- = \mu_2$.

Una descomposición del conjunto X en dos conjuntos medibles disjuntos P y N, tales que μ es positiva sobre cualquier subconjunto medible de P y negativa sobre cualquier subconjunto medible de N se llama una **descomposición de Hahn** para (X, Σ, μ).

Si μ es una medida signada, la expresión (c) anterior nos sugiere la siguiente definición de la integral de una función f con respecto a la medida μ

$$\int_X f \; d\mu = \int_X f \; h \; d|\mu| \; ,$$

para cada $f \in L^1(|\mu|)$.

Dejamos al lector verificar que esta nueva definición de $\int f \; d\mu$, coincide con la dada anteriormente en términos de las medidas μ^+ y μ^-.

4. Aplicaciones del teorema de Radon-Nikodym.

A. Esperanza Condicional. Sea (X, Σ, μ), $\mu \geq 0$, un espacio de medida σ-finito y Σ_0 una σ-álgebra contenida en Σ, tal que μ es σ-finita cuando la restringimos a Σ_0. Veremos que para cada $f \in L^1(X, \Sigma, \mu)$ existe una única función $E\,[f/\Sigma_0] = E[f]$ en $L^1(X, \Sigma_0, \mu)$ tal que

$$\int_A E[f] \; d\mu = \int_A f \; d\mu \qquad\qquad (A \in \Sigma_0) \; .$$

La función $E\,[f/\Sigma_0]$ se conoce con el nombre de **esperanza condicional** de f dada la σ-álgebra Σ_0.

La existencia y unicidad de $E\,[f/\Sigma_0]$ es consecuencia inmediata del teorema de Radon-Nikodym. En efecto, bastaría usar el teorema (9.5) con

$$\nu(A) = \int_A f \; d\mu \qquad\qquad (A \in \Sigma_0) \; ,$$

y el hecho de que el teorema es válido para una medida σ-finita.

Supongamos $\mu(X)$ finito y f en $L^2(X, \Sigma, \mu)$. Usando la definición de esperanza condicional y la desigualdad de Cauchy-Schwarz obtenemos

$$\left| \int E\,[f/\Sigma_0]\, g \; d\mu \right| \leq \|f\|_2 \; \|g\|_2 \; ,$$

para cada función simple g medible con respecto a Σ_0.

Luego, por (7.10)', la esperanza condicional $E[f/\Sigma_0]$ está en $L^2(X,\Sigma_0,\mu)$ si $f \in L^2(X,\Sigma,\mu)$. Además tenemos que

$$\int E[f/\Sigma_0]\, g\, d\mu = \int f\ g\ d\mu \qquad \big(g \in L^2(X,\Sigma_0,\mu)\big)\ .$$

Esta última igualdad se puede reformular diciendo que si f pertenece a $L^2(X,\Sigma,\mu)$, la función $E[f/\Sigma_0]$ es la proyección de f sobre el subespacio cerrado $L^2(X,\Sigma_0,\mu)$ del espacio de Hilbert $L^2(X,\Sigma,\mu)$.

Pensar la esperanza condicional como una proyección en un espacio de Hilbert, nos da un camino para definirla sin usar el teorema de Radon-Nikodym (ver ejercicio 13).

Para evitar algunas complicaciones supondremos $\mu(X) < \infty$ en el siguiente teorema, lo cual es habitual cuando se trabaja con la esperanza condicional.

(9.7) **Teorema.** *Sea* (X,Σ,μ) *un espacio de medida finito y* Σ_0 *una sub* σ*-álgebra de* Σ*. Entonces*

(1) $0 \le f$ *implica* $0 \le E[f/\Sigma_0]$.

(2) $|E[f/\Sigma_0]| \le E[|f|/\Sigma_0]$.

(3) Sea $f \in L^p(X,\Sigma,\mu)$, $1 \le p \le \infty$ *y* $\frac{1}{p}+\frac{1}{p'}=1$. *Entonces*

$$\int f\ g\ d\mu = \int E[f/\Sigma_0]\, g\, d\mu \qquad \Big(g \in L^{p'}(X,\Sigma_0,\mu)\Big)\ .$$

(4) $\|E[f/\Sigma_0]\|_p \le \|f\|_p\,, \qquad 1 \le p \le \infty\,.$

(5) Si $\mu(X)=1$ *y* φ *es una función convexa cuyo dominio contiene al rango de* f*, tal que* f *y* $\varphi(f)$ *son integrables. Entonces*

$$\varphi\left(E[f/\Sigma_0]\right) \le E[\varphi(f)/\Sigma_0]\ .$$

DEMOSTRACIÓN: Ver ejercicio 14.

B. Dualidad de los espacios L^p. Consideraremos espacios L^p sobre un espacio de medida σ-finito (X,Σ,μ). Para $1 \le p < \infty$ y $g \in L^{p'}$ con $1/p + 1/p' = 1$, ponemos

$$\ell_g(f) = \int f\ g\ d\mu\ .$$

Claramente ℓ_g es una funcional lineal continua definida sobre L^p, más todavía $\|\ell_g\| = \|g\|_{p'}$.

Para verificar la afirmación anterior basta repetir los argumentos rea lizados en la sección 5 del capítulo VII.

Recordemos que por el dual de L^p, denotado por $(L^p)^*$, entendemos el dual topológico, o sea el espacio de las funcionales lineales continuas definidas sobre L^p.

Nos proponemos ahora completar las demostración del teorema (7.11), cuya versión para espacios de medida abstracta es la siguiente.

(9.8) **Teorema.** *Sea $1 \leq p < \infty$ y ℓ en el dual de L^p. Entonces existe una única $g \in L^{p'}$ tal que $\ell = \ell_g$.*

DEMOSTRACIÓN: Supondremos $\mu(X) < \infty$, dejando como ejercicio el caso en que X es σ-finito.

Dada $\ell \in (L^p)^*$, definimos la siguiente función sobre Σ

$$\nu(E) = \ell(\chi_E) \ .$$

Suponiendo que ℓ toma valores reales, ν es una medida con signo sobre Σ y además es absolutamente continua con respecto a la medida μ. Por el teorema de Radon-Nikodym, existe $g \in L^1(X, \Sigma, \mu)$, tal que

$$\ell(\chi_E) = \int_E g \, d\mu \ , \qquad\qquad (E \in \Sigma) \ .$$

Usando la continuidad de ℓ y (7.10)' obtenemos que $g \in (L^p)^*$ y de aquí fácilmente resulta

$$\ell(f) = \int f \ g \ d\mu \ ,$$

para cada $f \in L^p$, y esto concluye la demostración del teorema.

Q.E.D.

C. Diferenciación de una medida con respecto a la medida de Lebesgue. Sea μ una medida sobre los conjuntos medibles Lebesgue de $\mathbb{R}^n$ y supongamos que es finita sobre cada conjunto acotado. Luego, por el teorema (8.6), μ es una medida regular y por el teorema de Radon-Nikodym podemos escribir $\mu = \mu_s + \mu_a$ donde μ_s es singular con respecto a la medida de Lebesgue y μ_a absolutamente continua con respecto a dicha medida. Además existe una función h localmente integrable sobre $I\!R^n$, tal que

$$\mu_a(E) = \int_E h \ dx \ .$$

Usando los teoremas (8.3) y (8.9) obtenemos que

$$\lim_{Q \to x} \frac{\mu(Q)}{m(Q)} = h(x) \ ,$$

para casi todo x.

Los cocientes $\mu(Q)/m(Q)$, pueden ser reemplazados por $\mu(B)/m(B)$, donde B son bolas determinadas por una adecuada métrica en $I\!R^n$, (véase el parágrafo 4 del capítulo 8).

5. Convergencia débil en L^p.

Hemos visto las nociones de convergencia en norma, puntual, en medida, etc. A ellas añadimos el siguiente importante concepto de convergencia.

Dada (f_n) una sucesión de funciones en $L^p(E)$, $1 \leq p < \infty$, decimos que f_n **converge débilmente** hacia una función $f \in L^p(E)$ si para cada $g \in L^{p'}(E)$ tenemos que

$$\int_E f_n \ g \to \int_E f \ g \ , \qquad\qquad n \to \infty \ .$$

Observe que la definición anterior tiene sentido para $p = \infty$, siendo ahora $L^{p'}(E) = L^1(E)$. En este caso la convergencia débil es llamada por algunos autores convergencia débil estrella.

Nosotros usaremos la definición de convergencia débil anterior aun para el caso $p = \infty$. El lector interesado en comprender la diferencia existente entre convergencia débil y convergencia débil estrella en el espacio L^∞, debería consultar algún texto de análisis funcional.

En lo que resta de esta sección nos restringiremos a $L^p = L^p(E)$ con $E \subseteq I\!R^n$ un conjunto medible Lebesgue. La medida subyacente es la medida de Lebesgue.

(9.9) **Teorema.** *Sea* $1 < p \leq \infty$ *y* (f_n) *una sucesión de funciones acotadas en* L^p *, es decir* $\sup_n \|f_n\|_p$ *es finito. Entonces existe una subsucesión* (f_{n_k}) *de* (f_n) *que converge débilmente hacia una función* $f \in L^p$ *.*

DEMOSTRACIÓN: Sea F un conjunto numerable denso en $L^{p'}$. Por el proceso de diagonalización de Cantor existe una subsucesión (f_{n_k}) de (f_n) tal que para cada $g \in F$ existe el siguiente límite

$$\ell(g) = \lim_{k\to\infty} \int_E f_{n_k}\, g \ .$$

Como la sucesión (f_n) es acotada en L^p existe una constante finita c tal que:

$$|\ell(g)| \leq c\|g\|_{p'} \qquad\qquad (g \in F) \ .$$

Observemos que siempre podemos elegir el subconjunto denso F de modo tal que sea un espacio vectorial sobre el cuerpo de los números racionales.

Es fácil ver que podemos extender la función ℓ a todo $L^{p'}$ en forma lineal y continua. En efecto sea $g \in L^{p'}$ y (g_n) una sucesión en F tal que $\|g_n - g\|_{p'} \to 0$ cuando $n \to \infty$. Luego

$$|\ell(g_n) - \ell(g_m)| = |\ell(g_n - g_m)| \leq c\|g_n - g_m\|_{p'} \ .$$

Por lo tanto podemos definir $\ell(g)$ como el límite de la sucesión $(\ell(g_n))$. Es fácil ver que $\ell(g)$ no depende de la sucesión (g_n) elegida.

Por la definición de ℓ obtenemos que

$$|\ell(g)| \leq c\|g\|_{p'} \qquad\qquad (g \in L^{p'}) \ .$$

Para ver que ℓ es una función lineal sobre $L^{p'}$ recordemos que ℓ es aditiva sobre F y homogénea con respecto a escalares racionales. Resumiendo, $\ell \in (L^{p'})^*$ y por el teorema (9.8) existe $f \in L^p$ tal que

$$\ell(g) = \int_E f\, g \qquad\qquad (g \in L^{p'}) \ .$$

Nos resta demostrar que

$$\ell(g) = \lim_{k\to\infty} \int_E f_{n_k}\, g \qquad\qquad (g \in L^{p'}) \ .$$

Sea $g \in L^{p'}$ y $\varepsilon > 0$ elegimos $g_{n_0} \in F$ tal que $|\ell(g) - \ell(g_{n_0})| < \varepsilon$ y además $\|g - g_{n_0}\|_{p'} < \varepsilon$. Sea k_0 tal que si $k \geq k_0$ se tiene

$$\left| \ell(g_{n_0}) - \int_E g_{n_0} \, f_{n_k} \right| < \varepsilon \; .$$

Por lo tanto

$$\begin{aligned} \left| \ell(g) - \int_E g \, f_{n_k} \right| &\leq |\ell(g) - \ell(g_{n_0})| \; + \\ &\quad + \left| \ell(g_{n_0}) - \int_E g_{n_0} \, f_{n_k} \right| \; + \; \left| \int_E (g_{n_0} - g) f_{n_k} \right| \\ &\leq 2\varepsilon \; + \; \varepsilon c \; , \end{aligned}$$

cuando $k \geq k_0$.

Q.E.D.

La siguiente afirmación es útil.

(9.10) *Si* f_n *converge débilmente a* f *en* L^p, $1 \leq p \leq \infty$. *Entonces*

$$\|f\|_p \leq \mathop{\mathit{lim\,inf}}_{n\to\infty} \|f_n\|_p \; .$$

Demostración: En efecto para cada $g \in L^{p'}$, $\|g\|_{p'} \leq 1$ tenemos

$$\begin{aligned} \int_E f \; g &= \lim_{n\to\infty} \int_E f_n \; g \\ &\leq \liminf_{n\to\infty} \|f_n\|_p \; \|g\|_{p'} \\ &\leq \liminf_{n\to\infty} \|f_n\|_p \; . \end{aligned}$$

Ahora, la afirmación (9.10) resulta de la igualdad

$$\|f\|_p = \sup_{\|g\|_{p'} \leq 1} \int_E f \; g \; .$$

Q.E.D.

En la afirmación (9.10) la convergencia débil puede ser reemplazada por convergencia en medida, así tenemos:

(9.11) *Sean* (f_n), f *medibles finitas en casi todo punto de* E, *tal que* f_n *converge a* f *en medida cuando* n *tiende a infinito. Entonces para cualquier* $p \geq 1$

$$\|f\|_p \leq \mathop{lim\,inf}_{n\to\infty} \|f_n\|_p \ .$$

DEMOSTRACIÓN: Sin pérdida de generalidad supondremos f_n, f no negativas. Sea

$$\alpha = \liminf_{n\to\infty} \|f_n\|_p \ .$$

Entonces existe una sucesión $n_1 < n_2 < n_3 < \dots$ tal que $\|f_{n_k}\|_p \to \alpha$. Consideremos ahora una nueva sucesión $(f_{n_{k_i}})$ que converge a f en casi todo punto cuando $i \to \infty$.

Para $g \in L^{p'}$, $g \geq 0$ y $\|g\|_{p'} \leq 1$ tenemos

$$\int_E |f| \ g \leq \liminf_{i\to\infty} \int_E |f_{n_{k_i}}| g \leq \liminf_{i\to\infty} \|f_{n_{k_i}}\|_p = \alpha.$$

Tomando supremo sobre las anteriores g se obtiene (9.11).

Q.E.D.

(9.12) **Teorema.** *Sean* (f_n) *y* f *funciones en* $L^p(E)$, $1 \leq p < \infty$, *tal que* $\|f_n\|_p \to \|f\|_p$ *cuando* $n \to \infty$ *y* f_n *converge a* f *en medida. Entonces*

(a) Para cada $A \subseteq E$ *medible*

$$\| \chi_A \ f_n\|_p \to \| \chi_A \ f\|_p \ .$$

(b) $\|f - f_n\|_p \to 0$ *cuando* $n \to \infty$.

DEMOSTRACIÓN: Veamos primero *(a)*. Teniendo en cuenta (9.11) se verifica

(1) $$\| \chi_A \ f\|_p^p \leq \liminf_{n\to\infty} \| \chi_A \ f_n\|_p^p \ .$$

Poniendo $B = E - A$, por (9.11) tenemos:

$$\begin{aligned} \|\chi_B \ f\|_p^p &\leq \liminf_{n\to\infty} \|\chi_B \ f_n\|_p^p \\ &= \liminf_{n\to\infty} \left(\|f_n\|_p^p - \|\chi_A \ f_n\|_p^p\right) \\ &= \|f\|_p^p \ - \ \limsup_{n\to\infty} \|\chi_A \ f_n\|_p^p \ . \end{aligned}$$

Luego

$$\limsup_{n\to\infty} \|\chi_A f_n\|_p^p \le \|\chi_A f\|_p^p \, . \tag{2}$$

Las desigualdades (1) y (2) implican *(a)*.

Para $\varepsilon > 0$ elegimos $A \subseteq E$ con $m(A) < \infty$, tal que

$$\|\chi_{E-A} f\|_p < \varepsilon \, .$$

En virtud de *(a)* existe n_0, tal que para todo $n \ge n_0$

$$\|\chi_{E-A} f_n\|_p < 2\varepsilon \, .$$

Por lo tanto para demostrar *(b)* podemos suponer sin pérdida de generalidad que $m(E) < \infty$. Además bastará demostrar que existe una subsucesión (f_{n_k}) de (f_n) tal que $\|f_{n_k} - f\|_p \to 0$ cuando $k \to \infty$.

Existe (f_{n_k}) tal que $f_{n_k} \to f$ en casi todo punto de E, y por el teorema de Egorov existe $B \subseteq E$ tal que f_{n_k} converge uniformemente a f sobre B y $m(E - B) < \delta$, donde $\delta > 0$ es elegido de forma tal que se cumpla:

$$\text{si } m(A) < \delta \text{ entonces } \|\chi_A f\|_p < \varepsilon \, .$$

Teniendo en cuenta *(a)* existe n_0 tal que si $n \ge n_0$ se verifica

$$\|\chi_A f_n\|_p < 2\varepsilon \, .$$

Ahora es fácil estimar la norma de $f - f_{n_k}$. En efecto, tenemos

$$\begin{aligned}\|f - f_{n_k}\|_p &\le \|\chi_B(f - f_{n_k})\|_p + \|\chi_{E-B} f\|_p + \|\chi_{E-B} f_{n_k}\|_p \\ &< \|\chi_B(f - f_{n_k})\|_p + 3\varepsilon \, .\end{aligned}$$

Teniendo en cuenta que $f_{n_k} \to f$ uniformemente sobre B tenemos *(b)*.

Q.E.D.

A continuación extenderemos el teorema (7.5) a los espacios L^p con $1 < p < \infty$, para lo cual usaremos los resultados (9.9) y (9.10).

Suponemos para el siguiente teorema que C es un subconjunto convexo en $L^p(E)$ que cumple:*Si* (f_n) *es una sucesión de elementos en* C *que* *converge débilmente hacia una función* $f \in L^p(E)$, *entonces* $f \in C$. Un conjunto C con esta propiedad se llama **débilmente cerrado**.

Se puede demostrar que para conjuntos convexos la propiedad de ser débilmente cerrado es equivalente a ser cerrado con respecto a la norma en L^p. Por ejemplo véase el libro Nociones de Espacios Normados por Mischa Cotlar y Roberto Cignoli.

En algunas situaciones se demuestra que un conjunto convexo es débilmente cerrado sin usar el resultado anteriormente mencionado (véanse los ejercicios).

(9.13) **Teorema.** *Sea* $C \subseteq L^p(E)$ *un conjunto convexo débilmente cerrado,* $1 < p < \infty$. *Entonces existe un único elemento* $f_0 \in C$ *de norma mínima, o sea*

$$\|f_0\|_p = \inf_{f \in C} \|f\|_p \ .$$

DEMOSTRACIÓN: Sean $a = \inf_{f \in C} \|f\|_p$ y (f_n) una sucesión en C tal que $\|f_n\|_p \to a$ cuando $n \to \infty$.

Por (9.9) existe una subsucesión (f_{n_k}) de (f_n) tal que f_{n_k} converge débilmente a una función $f_0 \in L^p$. Luego por (9.10) $\|f_0\|_p \leq a$ y como C es débilmente cerrado $f_0 \in C$ y por lo tanto $\|f_0\|_p = a$.

Para demostrar la unicidad del elemento de norma mínima referimos al lector a los ejercicios.

Q.E.D.

6. Diferenciación de funciones.

En los parágrafos 2 y 4 del capítulo VIII hemos visto resultados sobre diferenciación de funciones de conjunto, donde se consideró el límite de cocientes de dos medidas. En esta sección haremos una breve incursión dentro del tema de diferenciación de funciones, dando condiciones suficientes para que una función tenga derivada en casi todo punto.

Sea O un conjunto abierto en $\mathbb{R}^n$ y f una función real definida sobre O. Recordemos que f es **diferenciable** en $x_0 \in O$ (o que tiene derivada en ese punto) si existe un vector $f'(x_0) \in \mathbb{R}^n$ tal que

$$f(x_0 + h) - f(x_0) = (f'(x_0), h) + o(h) \quad \text{para} \quad h \to 0 \ ,$$

donde $(\,\cdot\,,\,\cdot\,)$ denota el producto escalar en $I\!R^n$ y $o(h)$ es una "o chica de h" (o sea $o(h)/h \to 0$ cuando $h \to 0$).

Se demostró en el capítulo anterior que si $d(x)$ es la distancia del punto x a un conjunto cerrado $E \subseteq I\!R^n$, entonces $d(x+h) = o(h)$ cuando $h \to 0$, para casi todo $x \in E$. En otras palabras $d(x)$ tiene derivada cero en casi todo punto de E.

El teorema siguiente es una extensión del resultado anteriormente mencionado.

(9.14) **Teorema.** *Sea E un subconjunto de $I\!R^n$ no necesariamente medible y f una función medible definida en un entorno del conjunto E. Suponemos que para cada $x \in E$ existen números reales M y $\delta > 0$ tales que $|f(x+h)| \le M|h|$ si $|h| \le \delta$. Entonces $f(x+h) = o(h)$ cuando $h \to 0$ para casi todo punto $x \in E$.*

DEMOSTRACIÓN: Uniformamos las acotaciones definiendo los conjuntos siguientes:

$$F_k = \{x \in E : |f(x+h)| \le k|h| \text{ para } |h| \le 1/k\}\ .$$

Sea $\widetilde{F}_k$ un conjunto medible tal que $\widetilde{F}_k \supset F_k$ y para cada conjunto medible A se verifica $m_e(F_k \cap A) = m(\widetilde{F}_k \cap A)$, ver ejercicio 9 del capítulo III. Es suficiente probar la afirmación del teorema para cada $x_0 \in F_k$ tal que

$$\frac{m(\widetilde{F}_k \cap B_r(x_0))}{m(B_r(x_0))} \to 1 \quad \text{cuando} \quad r \to 0\ .$$

A continuación repetiremos argumentos ya usados en el parágrafo 2 del capítulo VIII. No obstante creemos conveniente realizar la demostración en detalle.

Sean $\varepsilon > 0$ e $y \in I\!R^n$ y supongamos por un momento que $m(\widetilde{F}_k \cap B_{\varepsilon|y|}(x_0+y)) = 0$. Teniendo en cuenta que $B_{\varepsilon|y|}(x_0+y) \subseteq B_{|y|(1+\varepsilon)}(x_0)$, obtenemos

$$\begin{aligned}\frac{m\left(\widetilde{F}_k \cap B_{|y|+\varepsilon|y|}(x_0)\right)}{m(B_{|y|+\varepsilon|y|})} &\le \frac{m(B_{|y|+\varepsilon|y|}(x_0)) - m(B_{\varepsilon|y|}(x_0))}{m(B_{|y|+\varepsilon|y|})} \\ &= 1 - \left(\frac{\varepsilon}{1+\varepsilon}\right)^n\ .\end{aligned}$$

Puesto que x_0 es punto de densidad 1 para $\widetilde{F}_k$, se obtiene una contradicción haciendo $y \to 0$ en la desigualdad anterior. Luego para cada $\varepsilon > 0$ existe $\delta > 0$ tal que si $|y| \leq \delta$, entonces $m(\widetilde{F}_k \cap B_{\varepsilon|y|}(x_0 + y)) > 0$.

Dado que

$$m(\widetilde{F}_k \cap B_{\varepsilon|y|}(x_0 + y)) = m(F_k \cap B_{\varepsilon|y|}(x_0 + y)) \ ,$$

si $|y| \leq \delta$ existe $z \in F_k$ tal que $|x_0 + y - z| \leq \varepsilon|y|$.

Por otra parte como $|h| \leq 1/k$ implica $|f(z + h)| \leq k|h|$, tenemos que

$$|f(x_0 + t)| = |f(z + x_0 - z + t)| \leq k|x_0 - z + t| \leq k \ \varepsilon|t| \ ,$$

siempre que $|t| \leq \delta$ y $\delta\varepsilon \leq 1/k$.

Q.E.D.

Una función f de $\mathbb{R}$ en $\mathbb{R}$ es de **variación acotada sobre $\mathbb{R}$** si existe una constante finita c_f tal que

$$\sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| \leq c_f \ ,$$

para cualquier elección finita de puntos $x_1 < x_2 < \ldots < x_n$ en $\mathbb{R}$. Llamamos **variación total** de f al supremo de las sumas anteriores y éste es denotado por $V(f, \mathbb{R})$ o simplemente por $V(f)$.

En forma análoga se define la variación total de f sobre un intervalo I (acotado o no) la cual será denotada por $V(f, I)$.

Se puede ver, sin mucha dificultad, que si una función es de variación acotada el conjunto de sus puntos de discontinuidad es a lo sumo numerable.

Para verificar la afirmación anterior observe que si ponemos

$$\omega(x) = \limsup_{\delta \to 0} |f(x + \delta) - f(x)| \ ,$$

el conjunto $\{x : \omega(x) \geq 1/k\}$ tiene a lo sumo $k \ V(f)$ elementos.

Las funciones monótonas acotadas y las funciones derivables con derivada acotada son de variación acotada sobre un intervalo acotado I.

Sea μ una medida real definida sobre los conjuntos medibles Lebesgue de la recta y pongamos

$$f_\mu(x) = \mu((-\infty, x)) \ .$$

La función f_μ es de variación acotada sobre $I\!R$, pues si $x_0 < x_1 < \ldots < x_n$ entonces tenemos

$$\begin{aligned}\sum_{i=1}^{n} |f_\mu(x_i) - f_\mu(x_{i-1})| &\leq \sum_{i=1}^{n} |\mu([x_{i-1}, x_i))| \\ &\leq |\mu|\ (I\!R)\ ,\end{aligned}$$

luego $V(f_\mu) \leq |\mu|(I\!R)$, y por el teorema (9.6) tenemos $|\mu|(I\!R) < \infty$. Más todavía, se puede ver que $V(f_\mu) = |\mu|(I\!R)$ (ver ejercicios).

Notemos que la función f_μ definida arriba es continua por la izquierda en todo punto y que $f_\mu(x) \to 0$ cuando $x \to -\infty$. Decimos que una **función de variación acotada f está normalizada** o abreviadamente que **f es VAN**, si la función es continua por la izquierda y tiende a cero cuando su argumento tiende a menos infinito.

Hemos demostrado que una medida real μ genera una función f_μ que es VAN. Veremos a continuación que si f es una función de variación acotada, una traslación de ésta la transforma "practicamente" en VAN. Necesitamos primero demostrar la siguiente propiedad:

(9.15) *Para toda función de variación acotada f existen sus límites laterales en cualquier punto, así como los números reales:*

$$c_f = \lim_{x\to-\infty} f(x)\ , \qquad \mathcal{C}_f = \lim_{x\to+\infty} f(x)\ .$$

DEMOSTRACIÓN: a modo de ejemplo veamos que si $-\infty < x_0 \leq +\infty$ entonces existe el límite lateral izquierdo $f(x_0-)$. El caso restante se demuestra análogamente.

Sea (x_n) una sucesión que converge a x_0 en forma creciente. Como f es de variación acotada tenemos que

$$\sum_{n=1}^{\infty} |f(x_{n+1}) - f(x_n)| < \infty\ .$$

Pero si la serie anterior es finita, entonces $(f(x_n))$ es una sucesión de Cauchy. También se ve que le límite es independiente de la sucesión (x_n).

Q.E.D.

Sea f una función de variación acotada y definamos $g(x) = f(x-) - c_f$. Dejamos al lector demostrar que g es VAN. Notemos que salvo un conjunto numerable de puntos $f(x) - g(x) = c_f$. Veremos como una aplicación del teorema (9.14) que la función $f(x) - g(x)$ tiene derivada cero en casi todo punto.

(9.16) **Teorema.** *Sea f una función de variación acotada que se anula en casi todo punto de la recta. Entonces para casi todo x la derivada $f'(x)$ existe y vale cero.*

Demostración: Sobre el conjunto $E = \{f = 0\}$ definimos la siguiente función:

$$\Delta(x) = \limsup_{h \to 0} \frac{|f(x+h)|}{|h|} \ .$$

Si demostramos que Δ es finita en casi todo punto el teorema será consecuencia de (9.14).

Definamos $E_\infty = \{\Delta = \infty\}$ y sea $\ell > 0$. Para cada $x \in E_\infty$ existe $\varepsilon > 0$ tal que

$$\ell\, \varepsilon \le |f(x-\varepsilon)| \ \text{ ó } \ \ell\, \varepsilon \le |f(x+\varepsilon)| \ .$$

Usando el lema de cubrimiento (8.7), existe una sucesión disjunta de intervalos $((x_i - \varepsilon_i, x_i + \varepsilon_i))$ tal que $E_\infty \subset \cup_i (x_i - 5\varepsilon_i, x_i + 5\varepsilon_i)$.

Por consiguiente,

$$m_e(E_\infty) \le \frac{10}{\ell} \sum_i |f(x_i \pm \varepsilon_i)| \le \frac{10\ V(f)}{\ell} \ .$$

Por lo tanto $m_e(E_\infty) \le 10\ V(f)/\ell$ para cada ℓ, esto es $m(E_\infty) = 0$.

Q.E.D.

(9.17) **Teorema.** *Si f es una función de variación acotada existe una única función g VAN y una constante c, tal que $f(x) = g(x) + c$ en cada punto de continuidad de f. Además en casi todo punto la función $f - g$ tiene derivada igual a cero.*

Demostración: Ya hemos construido la función g y la segunda parte de (9.17) es consecuencia de (9.16).

Para verificar la unicidad basta recordar que g por ser VAN es continua por la izquierda y coincide con $f(x) - c$ en cada punto de continuidad de f.

Q.E.D.

Sabemos que si μ es una medida de Borel signada sobre $I\!R$, entonces la función $f_\mu(x) = \mu((-\infty, x))$ es VAN. Recíprocamente tenemos al siguiente teorema.

(9.18) **Teorema.** *Si f es una función VAN existe una única medida de Borel signada μ sobre $I\!R$ tal que $f_\mu = f$.*

DEMOSTRACIÓN: Primero el lector debería hacer los ejercicios 24, 25 y 26 al final de este capítulo.

Si la función f del teorema es además no-decreciente la construcción de la medida μ está dada en el ejercicio 25. La igualdad $f_\mu = f$ es consecuencia inmediata de la definición y de que estas funciones son continuas por la izquierda.

Para obtener la unicidad de la medida μ use que $f_\mu = f$ y el ejercicio 27 de este capítulo.

Dado que podemos diferenciar una medida de Borel regular con respecto a la medida de Lebesgue (ver capítulo VIII, sección 4) tenemos como consecuencia de (9.18) que si una función f es VAN entonces tiene derivada en casi todo punto. Más todavía hemos demostrado el siguiente teorema.

(9.19) **Teorema.** *Una función de variación acotada es derivable en casi todo punto.*

Volvamos al caso en que f es una función VAN. Recordemos la igualdad $\mu(-\infty, x) = f(x)$, donde μ es la medida de Borel signada asociada a f .

Usando el teorema de Radon-Nikodym podemos escribir

$$\mu(E) = \mu_s(E) + \int_E h(t)\ dt\ ,$$

donde μ_s es una medida singular con respecto a la medida de Lebesgue y h una función integrable Lebesgue.

Poniendo $E = (-\infty, x)$ y $f_s(x) = \mu_s(-\infty, x)$ en la igualdad anterior, resulta:

$$f(x) = f_s(x) + \int_{-\infty}^{x} h(t)\ dt\ . \tag{9.20}$$

En (9.20) la función f_s tiene derivada cero en casi todo punto por (9.16). Además $f'(x) = h(x)$ en casi todo x, como consecuencia del teorema de diferenciación de la integral.

Una función $f : I\!R \to I\!R$ es **absolutamente continua** si para cada $\varepsilon > 0$ existe $\delta > 0$ tal que para cualquier sucesión $x_1 < y_1 < x_2 < y_2 < \ldots < x_n < y_n$, la relación $\sum_{i=1}^n (y_i - x_i) < \delta$ implica $\sum_{i=1}^n |f(y_i) - f(x_i)| < \varepsilon$.

Toda función absolutamente continua es uniformemente continua y además es de variación acotada sobre todo intervalo acotado (probarlo). Siendo h integrable sobre $I\!R$ con respecto a la medida de Lebesgue, definamos

$$f(x) = \int_{-\infty}^{x} h(t)\ dt \ .$$

Por las conocidas propiedades de la integral, f es absolutamente continua.

Sea ahora f una función VAN absolutamente continua y consideremos la medida μ asociada a f por medio de $f(x) = \mu(-\infty, x)$. Se puede demostrar que en esta situación (ver ejercicios) la medida μ es absolutamente continua con respecto a la medida de Lebesgue. Usando (9.20) obtenemos

(9.21) $$f(x) = \int_{-\infty}^{x} h(t)\ dt \qquad (x \in I\!R) \ .$$

Resumimos lo anterior en el siguiente teorema.

(9.21) **Teorema.** *Sea f VAN. Entonces f es una función absolutamente continua si y sólo si existe h integrable Lebesgue sobre $I\!R$ que verifica (9.21).*

Se comprende sin dificultad que el concepto de función absolutamente continua puede definirse en la misma forma para una función f definida sobre un intervalo acotado I. En tal caso podemos suponer que I es cerrado, pues toda función absolutamente continua en el interior de un intervalo I es uniformemente continua y puede extenderse por continuidad hasta los extremos de I.

El siguiente teorema se obtiene fácilmente de (9.22).

(9.23) **Teorema.** *Sea f definida sobre un intervalo acotado $[a,b]$. Entonces f es absolutamente continua si y sólo si existe una función h integrable Lebesgue en $[a,b]$ tal que*

$$f(x) = f(a) + \int_{a}^{x} h(t)\ dt \qquad (a \le x \le b) \ .$$

EJERCICIOS

1. Sea (X, Σ) un espacio medible y f una función sobre X con valores en $\overline{I\!R}$ tal que $\{f > a\} \in \Sigma$ para todo $a \in D$, siendo D un conjunto denso en $I\!R$. Demuestre que f es medible.

   Para lo que sigue es necesario establecer dos definiciones:

(a) Una clase $\mathcal{S}$ de subconjuntos de X se llama una **semiálgebra** en el espacio X si tiene las siguientes propiedades:
   (1) $\emptyset \in \mathcal{S}$, $\quad X \in \mathcal{S}$
   (2) $S_1, S_2 \in \mathcal{S} \Rightarrow S_1 \cap S_2 \in \mathcal{S}$
   (3) $S \in \mathcal{S} \Rightarrow$ existen $S_1, ..., S_n \in \mathcal{S}$ disjuntos dos a dos, tales que $S^c = X - S = S_1 \cup ... \cup S_n$ (el complemento de cada miembro de $\mathcal{S}$ es una unión finita disjunta de miembros de $\mathcal{S}$)

(b) Una colección $\mathcal{A}$ de subconjuntos de X se llama un **álgebra** si
   (1) $\emptyset \in \mathcal{A}$
   (2) $A \in \mathcal{A} \Rightarrow A^c = X - A \in \mathcal{A}$
   (3) $A_1, A_2 \in \mathcal{A} \Rightarrow A_1 \cap A_2 \in \mathcal{A}$.

2. Probar que si $\mathcal{S}$ es una semiálgebra de subconjuntos de X, entonces la clase $\mathcal{A}$ formada por las uniones finitas disjuntas de miembros de $\mathcal{S}$ es un álgebra de subconjuntos de X; más aún, $\mathcal{A}$ es la mínima álgebra que incluye a $\mathcal{S}$ o como se dice, el **álgebra generada** por $\mathcal{S}$.

3. Dados dos espacios medibles (X, Σ) e (Y, Γ), sea $\mathcal{R}$ la clase de todos los conjuntos R de la forma $R = A \times B$ con $A \in \Sigma$ y $B \in \Gamma$, llamados **rectángulos medibles**. Probar que $\mathcal{R}$ es una semiálgebra en el producto $X \times Y$.

4. Para cualquier colección $\mathcal{C} \subset \mathcal{P}(X)$ existe una mínima σ-álgebra de subconjuntos de X que incluye a $\mathcal{C}$. Dicha σ-algebra, que denotaremos por $\boldsymbol{\sigma}(\mathcal{C})$, es la **$\boldsymbol{\sigma}$-álgebra generada** por $\mathcal{C}$ y se obtiene formando la intersección de todas las σ-álgebras Σ, tales que $\mathcal{C} \subset \Sigma$.
   Probar que si $\mathcal{A}$ es el álgebra generada por una semiálgebra $\mathcal{S}$ de subconjuntos de X, entonces $\boldsymbol{\sigma}(\mathcal{S}) = \boldsymbol{\sigma}(\mathcal{A})$.

5. Una clase de conjuntos $\mathcal{M} \subset P(X)$ se llama **monótona** si la unión de cualquier sucesión creciente de miembros de $\mathcal{M}$ y la intersección de

cualquier sucesión decreciente de miembros de $\mathcal{M}$ son miembros de $\mathcal{M}$ (clase cerrada bajo límites monótonos).

Probar que para cualquier colección $\mathcal{C} \subset \mathcal{P}(X)$ existe una mínima clase monótona $\boldsymbol{m}(\mathcal{C})$ que incluye a $\mathcal{C}$, a la que llamaremos la **clase monótona generada** por $\mathcal{C}$.

Probar las afirmaciones siguientes:

(i) toda σ-álgebra es una clase monótona.

(ii) toda álgebra monótona es una σ-álgebra.

(iii) si $\mathcal{A}$ es un álgebra, entonces $\boldsymbol{m}(\mathcal{A}) = \boldsymbol{\sigma}(\mathcal{A})$.
Sugerencia para (iii): en virtud de (i), $\boldsymbol{m}(\mathcal{A}) \subset \boldsymbol{\sigma}(\mathcal{A})$. En virtud de (ii), para probar la inclusión opuesta basta demostrar que $\boldsymbol{m}(\mathcal{A})$ es un álgebra y para ello conviene seguir los siguientes pasos:

(a) $\boldsymbol{m}(\mathcal{A})$ es cerrada bajo complementos, como resulta de comprobar que la colección

$$\mathcal{M}^* = \{E : E^c \in \boldsymbol{m}(\mathcal{A})\}$$

es una clase monótona que incluye al álgebra $\mathcal{A}$.

(b) $\boldsymbol{m}(\mathcal{A})$ es cerrada bajo intersecciones finitas. En efecto, para cada $E \subset X$,

$$\mathcal{M}_E = \{F : E \cap F \in \boldsymbol{m}(\mathcal{A})\}$$

es una clase monótona.
Claramente, $F \in \mathcal{M}_E$ si y sólo si $E \in \mathcal{M}_F$ y si $A \in \mathcal{A}$, entonces $\mathcal{A} \subset \mathcal{M}_A$, de donde $\boldsymbol{m}(\mathcal{A}) \subset \mathcal{M}_A$. Luego, si $E \in \boldsymbol{m}(\mathcal{A})$ y $A \in \mathcal{A}$, tendremos $E \in \mathcal{M}_A$; o sea, $A \in \mathcal{M}_E$, de donde $\mathcal{A} \subset \mathcal{M}_E$ y por lo tanto, $\boldsymbol{m}(\mathcal{A}) \subset \mathcal{M}_E$.

6. Complete los detalles de la demostración del teorema (9.2).

7. Sea $f \in L^p(X, \Sigma, \mu)$, $1 \leq p < \infty$, donde el espacio de medida no es necesariamente σ-finito. Probar que $E = \{f \neq 0\}$ es unión numerable de conjuntos de medida finita.

8. Demuestre las afirmaciones sobre unicidad en (9.5).

9. Sea μ una medida real sobre (X, Σ) y $|\mu|$ la variación total de μ. Entonces $|\mu|$ es una medida sobre Σ.

10. Sea Σ una σ-álgebra de conjuntos sobre X y $M(X)$ el conjunto de medidas reales definidas sobre Σ. Demuestre que $M(X)$ es un espacio normado completo con $\|\mu\| = |\mu|(X)$.

11. Sea μ una medida real sobre (X, Σ). Entonces existe una función medible f tal que $|f| = 1$ en cada punto y para cada $E \in \Sigma$ verifica

$$\mu(E) = \int_E f \; d|\mu| \; .$$

Sugerencia: de la desigualdad $|\mu|(E) \leq \int_E |f| \; d|\mu|$ se obtiene $|f| \geq 1$ en casi todo punto con respecto a $|\mu|$; y de $\left|\int_E f \; d|\mu|\right| \leq |\mu|(E)$ obtenemos que $|f| \leq 1$ en casi todo punto.

12. Sea μ una medida real y $f \in L^1(|\mu|)$. Para cada $E \in \Sigma$ pongamos

$$\mu_f(E) = \int_E f \; d\mu = \int_E f \; d\mu^+ - \int_E f \; d\mu^- \; ,$$

donde $\mu^+ - \mu^- = \mu$ es la descomposición de Jordan de la medida μ. Demuestre que la variación total de la medida real μ_f está dada por

$$|\mu_f|(E) = \int_E |f| \; d|\mu| \qquad\qquad (E \in \Sigma) \; .$$

13. Dado un espacio de medida σ-finito (X, Σ, μ) y una sub σ-álgebra Σ_0, demuestre sin usar el teorema de Radon-Nikodym que para cada $f \in L^1(X, \Sigma, \mu)$ existe una única función $g \in L^1(X, \Sigma_0, \mu)$ tal que

$$\int_E f \; d\mu = \int_E g \; d\mu \qquad\qquad (E \in \Sigma_0) \; .$$

Sugerencia: Suponga primero que $\mu(X) < \infty$, y use el teorema de representación de Riesz para la funcional sobre $L^2(\Sigma_0)$ definida por

$$\ell(g) = \int f \; g \; d\mu \; .$$

14. Demuestre el teorema (9.7).

15. Complete la demostración del teorema (9.8) para una medida σ-finita.

16. El teorema (9.9) no vale para $p = 1$.
Sugerencia: considere $f_n = n \; \chi_{[0, \frac{1}{n}]}$.

17. Sea M_p el conjunto de las funciones de $[0, 1]$ en $I\!R$ crecientes contenidas en $L^p[0, 1]$. Demuestre:

(a) M_p es un subconjunto convexo y cerrado en L^p, $1 \leq p < \infty$.
(b) M_p es débilmente cerrado en L^p, $1 \leq p < \infty$.
Sugerencia: para la demostración de (b) puede usarse el teorema (8.14) parte *(b)*.

18. Sea (f_n) una sucesión de funciones crecientes de $(0,1)$ en $I\!R$ tal que para cada $0 < x < 1$ $\sup_n |f_n(x)| < \infty$. Entonces existe unas subsucesión (f_{n_k}) y una función creciente f tal que para cada $x \in (0,1)$ la sucesión $(f_{n_k}(x))$ tiende a $f(x)$ cuando $k \to \infty$.
Sugerencia:
(a) Usando el principio de diagonalización de Cantor se puede extraer una subsucesión (f_{n_k}) de (f_n) tal que $f_{n_k}(x) \to f(x)$ cuando $k \to \infty$ para todo racional en $(0,1)$.
(b) La función f obtenida en (a) se puede extender a una función creciente sobre todo $(0,1)$. Esta función f es continua salvo en un conjunto $F \subset (0,1)$ a lo más numerable. Demuestre que si $x \notin F$ la sucesión $f_{n_k}(x) \to f(x)$ cuando $k \to \infty$.
(c) Usando nuevamente el proceso de diagonalización de Cantor existe una subsucesión $\left(f_{n_{k_i}}\right)$ de (f_{n_k}) tal que $\left(f_{n_{k_i}}(x)\right)$ converge para cada $x \in F$.

19. Sea (f_n) una sucesión de funciones crecientes tal que
$$\int_0^1 |f_n(x)|\, dx \leq 1 \;,$$
para todo n. Entonces existe una subsucesión (f_{n_k}) de (f_n) tal que para cada $x \in (0,1)$ la sucesión $(f_{n_k}(x))$ converge hacia una función creciente $f(x)$.

20. Una norma $\|\cdot\|$ se llama **estrictamente convexa** si $\|x\| = \|y\| = 1$, $x \neq y$ implica $\|1/2(x+y)\| < 1$. Demuestre que $\|\cdot\|_p$, $1 < p < \infty$ son normas estrictamente convexas.
Sugerencia: revise la demostración de la desigualdad de Minkowski dada por nosotros. Compare con el ejercicio 19 del Capítulo II.

21. Sea $E \subset I\!R$ un conjunto cerrado y $d(x) = \inf\{|x-y| : \; y \in E\}$, entonces d es una función diferenciable en casi todo punto de la recta.

22. Sea μ una medida real sobre $I\!R$ y $f_\mu(x) = \mu(-\infty, x)$. Entonces
$$V(f_\mu) = |\mu|(I\!R) \;.$$

23. Sea $f : I\!R \to I\!R$ una función de variación acotada. Pongamos (ver (9.15)) $f(x-) = \lim\limits_{y\to x, y<x} f(x)$, $c_f = \lim\limits_{x\to -\infty} f(x)$ $g(x) = f(x-) - c_f$. La función g es una función de variación acotada sobre $I\!R$. Poniendo $V(x)$ para la variación total de f sobre el conjunto $(-\infty, x]$, demuestre que
(a) $x < y \Longrightarrow |f(x) - f(y)| \le V(y) - V(x)$.
(b) $f = f_1 - f_2$ donde f_1 y f_2 son funciones crecientes. Por ejemplo $f_1 = (V + f)/2$, $f_2 = (V - f)/2$.
(c) Si f es VAN entonces V es VAN.

25. Sea f una función creciente en $I\!R$, acotada y continua por izquierda. Para cada conjunto $E \subseteq I\!R$ pongamos

$$\Phi(E) = \bigcup_{x\in E} [f(x), f(x+)] ,$$

$\Sigma = \{E \subset R : E$ y $\Phi(E)$ son borelianos $\}$. Demuestre que Σ es la σ-álgebra de Borel y que $\mu(E) = m(\Phi(E))$ es una medida sobre Σ.

26. Sea f VAN y μ la medida asociada a f por medio de $f(x) = \mu(-\infty, x)$. Entonces μ es absolutamente continua con respecto a la medida de Lebesgue si sólo si f es absolutamente continua.

27. Demuestre que si dos medidas borelianas reales μ y ν verifican la relación $\mu((-\infty, x)) = \nu((-\infty, x))$ para cada $x \in I\!R$, entonces $\mu = \nu$.

    Sugerencia: Considere la clase de los conjuntos borelianos E tales que $\mu(E) = \nu(E)$ y use el ejercicio 5.

28. Sea φ de $[0, \infty)$ en sí mismo convexa, con $\varphi(0) = 0$. Entonces existe una función $f \ge 0$ creciente, tal que

$$\varphi(x) = \int_0^x f(t)\; dt .$$

Sugerencia:
$\varphi(y) - \varphi(x) \le (y - x) \cdot D^- \varphi(n)$ si $0 < x < y \le n$, y use (9.23).

BIBLIOGRAFIA

[1] COTLAR, M. y CIGNOLI, R., *Nociones de Espacios Normados*, EUDEBA, Buenos Aires, 1967.

[2] DE GUZMÁN, M. y RUBIO, B., *Integración: teoría y técnicas*, Alhambra, Madrid, 1979.

[3] DIEUDONNÉ, J., *Foundations of Modern Analysis*, Academic Press, New York, 1960.

[4] FRIEDMAN, A., *Foundations of Modern Analysis*, Holt, Rinehart and Winston, Inc., 1970.

[5] HALMOS, P.R., *Measure Theory*, Van Nostrand, Princeton, 1950.

[6] KAPLANSKY, I., *Set Theory and Metric Spaces*, Allyn and Bacon, Boston, 1972.

[7] KELLEY, J., *Topología General*, EUDEBA, Buenos Aires, 1962.

[8] KOLMOGOROV, A.N. y FOMIN, S.V., *Elementos de la teoría de funciones y del análisis funcional*, Ed. MIR, Moscú, 1972.

[9] NATANSON, I.P., *Theory of functions of a real variable, Vols. I y II*, Ungar Publ. Co., New York, 1955.

[10] NIVEN, I., *Irrational numbers*, Math. Assoc. of America, 1967.

[11] ROYDEN, H.L., *Real Analysis*, The MacMillan Co., New York, 1963.

[12] RUDIN, W., *Real and Complex Analysis*, McGraw-Hill, New York, 1966.

[13] SAKS, S., *Theory of the Integral*, Monografie Matematyczne, Varsovia, 1937.

[14] SIERPINSKY, W., *Lecons sur les nombres transfinis*, Gauthier-Villars, Paris, 1950.

[15] SPIVAK, M., *Calculus on Manifolds*, Benjamin, New York, 1965.

[16] TITCHMARSH, E.C., *The Theory of Functions*, Oxford University Press, Londres, 1939.

[17] WHEEDEN, R. y ZYGMUND, A., *Measure and Integral*, Marcel Dekker, Inc., New York, 1977.

[18] ZAANEN, A.C., *An introduction to the theory of integration*, North-Holland, Amsterdam, 1961.

GLOSARIO

F

H

I

Medida e Integral de Lebesgue

Norberto Fava

Nació en Buenos Aires en 1936, y cursó la Licenciatura en Ciencias Matemáticas en la Universidad de Buenos Aires y estudios de doctorado en la Universidad de Minnesota (E.U.A.), donde se graduó en 1969. Ha sido profesor de Matemática en las Universidades Nacionales de San Luis y Río Cuarto y actualmente se desempeña como profesor titular en el Departamento de Matemática de la Facultad de Ciencias Exactas de la U.B.A.

Ha publicado varios trabajos científicos relacionados con la Teoría de Funciones de Variable Real (teoría de la diferenciación), en Análisis Funcional (teoría ergódica) y la Teoría de Probabilidades (procesos estocásticos). Es miembro titular de la Academia Nacional de Ciencias Físicas, Exactas y Naturales.

Felipe Zó

Nació en Mendoza el 29 de noviembre de 1945, y cursó en esa provincia sus estudios primarios y secundarios. Posteriormente se trasladó a San Luis donde obtuvo el título de Licenciado en Matemáticas de la Universidad de Cuyo en 1968. Realizó sus estudios de postgrado en la Universidad de Minnesota donde se doctoró bajo la dirección de Néstor Riviere en 1975. Es actualmente profesor titular de la Universidad Nacional de San Luis e investigador del Conicet y ha sido profesor visitante en varias universidades nacionales y extranjeras. Su interés en investigación está centrado en el análisis real y la teoría de aproximación de funciones, temas donde principalmente ha realizado sus publicaciones científicas.