A. N. KOLMOGOROV
S. V. FOMIN

ELEMENTOS DE LA TEORIA DE FUNCIONES Y DEL ANALISIS FUNCIONAL

TRADUCIDO DEL RUSO POR
CARLOS VEGA,
Candidato a Doctor en ciencias físico-matemáticas

Segunda edición

EDITORIAL MIR MOSCU

Impreso en la URSS, 1975

Traducción al español. Editorial Mir, 1975

на испанском языке

INDICE

PREFACIO

La primera edición rusa de *Elementos de la teoría de funciones y del análisis funcional* apareció en dos fascículos en 1954 y 1960. La publicación de estos dos fascículos se debió a que, a finales de la década del 40, fue incluido en el programa de la Facultad de Mecánica y Matemáticas de la Universidad de Moscú el curso de *Análisis III* que comprendía elementos de la teoría de medida y de la teoría de funciones, ecuaciones integrales, nociones del análisis funcional y, más tarde, cálculo de variaciones. Este curso, dictado en la Universidad de Moscú primero por Andréi Kolmogórov y luego por otros profesores, entre ellos Serguéi Fomín, integró posteriormente también los programas de otras Universidades. Debido a su reducida tirada, la edición de nuestro libro se agotó rápidamente y hace tiempo que surgió la necesidad de reeditarlo.

La sustitución de los cursos de la teoría de funciones de variable real, de ecuaciones integrales y de cálculo de variaciones por el curso unificado de *Análisis III* en la Universidad de Moscú, dio lugar a grandes discusiones en su tiempo. El curso tenía por objeto habituar a los estudiantes a una visión doble: por una parte, seguir la lógica interna del desarrollo de la teoría de conjuntos, de la teoría general de aplicaciones continuas en espacios métricos y topológicos, de la teoría general de espacios lineales y de funcionales y operadores en ellos y de la teoría pura de medida e integración en «espacios generales provistos de medida», y por otra parte, no perder de vista los problemas del análisis clásico y del aplicado, a los que prestan servicio estas ramas más abstractas de las Matemáticas.

Para resolver esta tarea, en la planificación del libro damos preferencia a la línea abstracta de estructuración del curso. De la teoría general de conjuntos (capítulo I) se puede pasar o bien a los espacios métricos y topológicos y sus aplicaciones continuas (capítulo II), o bien, directamente, a los espacios provistos de medida (sin topología) y a la integración en ellos (capítulo VI). En los capítulos III y IV se estudian los espacios lineales y los funcionales y operadores lineales en ellos. De estos capítulos se puede pasar directamente al capítulo V (operadores y funcionales diferenciables no lineales). En el capítulo VIII se estudian los espacios lineales de funciones sumables. Solamente en los capítulos VII y IX se concentra, de hecho, la atención en las funciones de variable real. La exposición de la teoría de ecuaciones integrales en el capítulo X está formalmente vinculada con el segmento $[a, b]$; pero, se le puede dar, sin modificaciones esenciales, una forma más general.

Aunque en nuestro libro se exponen, en primer lugar, los conceptos generales de la teoría de funciones y del análisis funcional, el lector podrá advertir, en casi todos los capítulos, la atención que se presta a los problemas clásicos contiguos. El haber incluido en nuestro libro los capítulos VII (teoría de diferenciación), IX (series trigonométricas e integral de Fourier) y X (ecuaciones integrales lineales) hace que abarque ahora todo el programa del curso de *Análisis III* adoptado en la Universidad de Moscú, menos el cálculo de variaciones. No hemos incluido este último en nuestro libro, limitándonos a exponer en el capítulo V los rudimentos del análisis funcional no lineal.

En la nueva edición, lo mismo que en la primera, ocupa un lugar considerable la teoría general de medida. En los últimos tiempos han aparecido varias exposiciones de la teoría de integración a base del esquema de Daniell, que no utiliza el aparato de la teoría de medida. Consideramos, sin embargo, que la teoría de medida tiene por sí sola suficiente importancia, independientemente de si se introduce o no el concepto de integral, y merece ser incluida en el curso universitario.

Al revisar el libro e incluir en él nuevas secciones hemos procurado, sin embargo, conservar el estilo relativamente elemental de exposición que, según nos parece, tenía la primera edición. Esperamos que éste hallará su lugar natural en la enseñanza universitaria a la par de otros textos.

A. Kolmogórov
S. Fomín

CAPITULO
I

ELEMENTOS DE LA TEORIA DE CONJUNTOS

§ 1. CONCEPTO DE CONJUNTO. OPERACIONES SOBRE CONJUNTOS

1°. Definiciones principales. En las Matemáticas tropezamos constantemente con distintos *conjuntos*. Podemos hablar del conjunto de facetas de un poliedro, de puntos de una recta, del conjunto de números naturales, etc. El concepto de conjunto es tan amplio que resulta difícil darle una definición que no se reduzca a sustituir simplemente la palabra «conjunto» por expresiones sinónimas: cúmulo, colección de elementos, etc.

El concepto de conjunto desempeña en las Matemáticas modernas un papel de extraordinaria importancia no sólo porque la propia teoría de conjuntos ha pasado a ser en la actualidad una disciplina sumamente vasta y enjundiosa, sino, principalmente, en virtud de la influencia que la teoría de conjuntos, nacida a fines del siglo pasado, ha ejercido y ejerce sobre todas las Matemáticas. Vamos a enunciar aquí las notaciones fundamentales y a exponer brevemente los conceptos primarios de la teoría de conjuntos que serán utilizados en los capítulos sucesivos.

Designaremos los conjuntos con letras mayúsculas A, B, ... y sus elementos con minúsculas a, b, ... La afirmación de que «el elemento a pertenece al conjunto A» se denota simbólicamente así: $a \in A$ o bien $A \ni a$; $a \overline{\in} A$ (o bien $A \overline{\ni} a$) significa que el elemento a no pertenece a A. Si todos los elementos que componen el conjunto A pertenecen también al conjunto B (con la particularidad de que el caso $A = B$ no está excluido), decimos que A es *subconjunto* del conjunto B y escribimos $A \subset B$. Por ejemplo, los números enteros forman un subconjunto del conjunto de todos los números reales.

A veces no sabemos de antemano si un conjunto (por ejemplo, el conjunto de las raíces de una ecuación) contiene o no por lo menos un elemento. De ahí, la conveniencia de introducir el llamado conjunto *vacío*, es decir, el conjunto que no contiene ni un elemento. Lo designaremos con el símbolo $\varnothing$. Cualquier conjunto contiene $\varnothing$ como subconjunto.

2°. Operaciones sobre conjuntos. Sean A y B conjuntos arbitrarios; se llama *suma* o *unión* $A \cup B$ de estos conjuntos al conjunto compuesto de todos los elementos pertenecientes por lo menos a uno de los conjuntos A ó B (fig. 1).

De manera análoga se define la suma de cualquier número (finito o infinito) de conjuntos: si A_α son conjuntos arbitrarios, su suma $\bigcup\limits_\alpha A_\alpha$ es la colección de elementos, cada uno de los cuales pertenece por lo menos a uno de los conjuntos A_α.

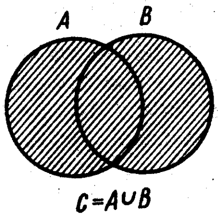


FIG. 1

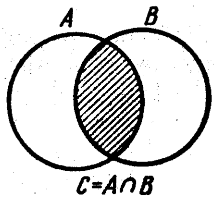


FIG. 2

Llamaremos *intersección* $A \cap B$ de los conjuntos A y B al conjunto compuesto por todos los elementos pertenecientes tanto al conjunto A como al conjunto B (fig. 2). Por ejemplo, la intersección del conjunto de todos los números pares y del conjunto de todos los números divisibles por tres está compuesta por todos los números enteros divisibles por seis. La intersección de un número cualquiera (finito o infinito) de conjuntos A_α es la colección $\bigcap\limits_\alpha A_\alpha$ de elementos pertenecientes a cada uno de los conjuntos A_α.

Por su propia definición las operaciones de suma e intersección son conmutativas y asociativas, es decir, $A \cup B = B \cup A$, $(A \cup B) \cup C = A \cup (B \cup C)$, $A \cap B = B \cap A$, $(A \cap B) \cap C = A \cap (B \cap C)$. Además, verifican las siguientes relaciones distributivas:

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C), \tag{1}$$

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C). \tag{2}$$

En efecto, comprobemos, por ejemplo, la primera de estas igualdades[1)]. Supongamos que el elemento x pertenece al conjunto que figura en la parte izquierda de la igualdad (1): $x \in (A \cup B) \cap C$. Esto significa que x pertenece a C y, además, por lo menos a uno de los conjuntos A ó B. Pero entonces x pertenece siquiera a uno de los conjuntos $A \cap C$ o $B \cap C$, es decir, figura en la parte derecha de la igualdad considerada. Viceversa, supongamos que $x \in (A \cap C) \cup (B \cap C)$. Entonces, $x \in A \cap C$ o bien $x \in B \cap C$. Por consiguiente, $x \in C$ y, además, x figura en A o en B, es decir, $x \in A \cup B$. De manera que $x \in (A \cup B) \cap C$, y la igualdad (1) queda demostrada. Análogamente se verifica la igualdad (2).

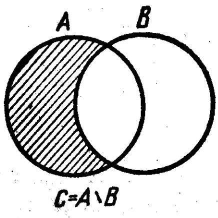


FIG. 3

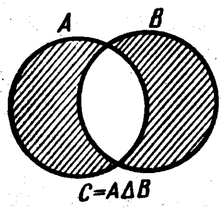


FIG. 4

Definamos la operación de resta de conjuntos. Llamaremos *diferencia* $A \setminus B$ de los conjuntos A y B a la colección de aquellos elementos de A que no pertenecen a B (fig. 3). Señalemos que aquí no se supone que $A \supset B$. A veces en lugar de $A \setminus B$ se escribe $A - B$. En algunos casos (por ejemplo, en la teoría de la medida) conviene introducir la llamada *diferencia simétrica* de dos conjuntos A y B que se define como la suma de las diferencias $A \setminus B$ y $B \setminus A$ (fig. 4). Denotaremos la diferencia simétrica de los conjuntos A y B con el símbolo $A \triangle B$. De manera que según la definición,

$$A \triangle B = (A \setminus B) \cup (B \setminus A).$$

EJERCICIO. Demostrar que

$$A \triangle B = (A \cup B) \setminus (A \cap B).$$

En lo sucesivo deberemos considerar con frecuencia distintos conjuntos, que todos son subconjuntos de un conjunto principal S, por ejemplo, diferentes conjuntos de puntos sobre la recta nu-

---

[1] La igualdad de dos conjuntos $A = B$ se entiende como una igualdad idéntica, es decir, significa que cada elemento de A pertenece a B y viceversa. En otras palabras, la igualdad $A = B$ equivale a que se verifican ambas inclusiones: $A \subset B$ y $B \supset A$.

mérica. En este caso, para todo conjunto A la diferencia $S \setminus A$ se llama *complemento* del conjunto A y se denota frecuentemente mediante CA o bien A'.

En la teoría de los conjuntos y sus aplicaciones desempeña un papel muy importante el llamado principio de dualidad, que se basa en las dos siguientes relaciones:

1. *El complemento de la suma es igual a la intersección de los complementos*

$$S \setminus \bigcup_{\alpha} A_{\alpha} = \bigcap_{\alpha} (S \setminus A_{\alpha}). \tag{3}$$

2. *El complemento de la intersección es igual a la suma de los complementos*

$$S \setminus \bigcap_{\alpha} A_{\alpha} = \bigcup_{\alpha} (S \setminus A_{\alpha}). \tag{4}$$

El principio de dualidad consiste en que de cualquier teorema referente a un sistema de subconjuntos de un conjunto fijo S se puede deducir de manera automática otro teorema, el teorema dual, sustituyendo los conjuntos considerados por sus complementos, la suma de conjuntos, por su intersección y la intersección, por la suma. Un ejemplo de la aplicación de este principio nos lo da el teorema 3' del § 2 del capítulo II.

Demostremos la relación (3).

Supongamos que $x \in S \setminus \bigcup_{\alpha} A_{\alpha}$. Esto significa que x no pertenece a la unión $\bigcup_{\alpha} A_{\alpha}$, es decir, no figura en ninguno de los conjuntos A_{α}. Por consiguiente, x aparece en cada uno de los complementos $S \setminus A_{\alpha}$ y por eso $x \in \bigcap_{\alpha} (S \setminus A_{\alpha})$. Viceversa, supongamos que $x \in \bigcap_{\alpha} (S \setminus A_{\alpha})$, es decir, que x pertenece a cada $S \setminus A_{\alpha}$; entonces, x no figura en ninguno de los conjuntos A_{α}, es decir, no pertenece a la suma $\bigcup_{\alpha} A_{\alpha}$, y por eso $x \in S \setminus \bigcup_{\alpha} A_{\alpha}$. La igualdad (3) queda demostrada. De manera análoga se demuestra la relación (4). (Realícese la demostración).

La expresión «diferencia simétrica» que se emplea para la operación $A \triangle B$ no es del todo acertada; esta operación es, en muchos aspectos, análoga a la suma de conjuntos $A \cup B$. En efecto, $A \cup B$ significa que unimos dos afirmaciones con el «o» *alternativo*: «el elemento pertenece al conjunto A» o «el elemento pertenece al conjunto B», mientras que $A \triangle B$ significa que unimos las mismas afirmaciones con el «o» *no alternativo*: el elemento x

pertenece a $A \triangle B$ si, y sólo si, figura o bien *solamente* en el conjunto A o bien *solamente* en el conjunto B. El conjunto $A \triangle B$ podría llamarse «suma módulo dos» de los conjuntos A y B (se toma la unión de estos dos conjuntos pero los elementos que figuran en ambos se excluyen).

§ 2. EQUIVALENCIA DE CONJUNTOS. CONCEPTO DE POTENCIA DE UN CONJUNTO

1°. Conjuntos finitos e infinitos. Al considerar diferentes conjuntos observamos que para algunos de ellos es posible señalar—aunque sea de una manera general y no de hecho—la cantidad de elementos que los componen. De este tipo es, por ejemplo, el conjunto de todos los vértices de un poliedro, el conjunto de todos los números primos inferiores a un número dado, el conjunto de todas las moléculas de agua en la Tierra, etc. Cada uno de estos conjuntos contiene un número finito, que posiblemente desconocemos, de elementos. Por otra parte, existen conjuntos compuestos por un número infinito de elementos. De este tipo son, por ejemplo, el conjunto de todos los números naturales, el conjunto de todos los puntos de una recta, el conjunto de todos los círculos del plano, el conjunto de todos los polinomios de coeficientes racionales, etc. Vale subrayar que al decir que uno u otro conjunto es infinito entendemos que se puede escoger de él un elemento, dos elementos, etc., y después de cada una de estas operaciones en el conjunto quedarán aún otros elementos.

Dos conjuntos finitos los podemos comparar por el número de elementos que los componen y decidir si este número es el mismo o si uno de los conjuntos posee más elementos que otro. ¿Es posible comparar de manera análoga los conjuntos infinitos? En otras palabras, ¿tiene sentido preguntar qué hay más: círculos sobre el plano o puntos racionales sobre la recta, funciones definidas sobre el segmento [0, 1] o rectas en el espacio, etc.?

Veamos cómo comparamos entre sí dos conjuntos finitos. Podemos proceder de dos maneras: en primer lugar, podemos contar el número de elementos de cada uno de estos conjuntos y comparar así ambos conjuntos. Pero podemos actuar de modo distinto, tratando de establecer una correspondencia biunívoca entre los elementos de estos conjuntos, es decir, una correspondencia que asigne a cada elemento de un conjunto un elemento, y sólo uno, del otro y viceversa. Está claro que una correspondencia biunívoca entre dos conjuntos finitos se puede establecer si, y sólo si, el número de elementos en ambos conjuntos es el mismo. Por ejemplo, para ver si coinciden el número de alumnos en el grupo y la cantidad de sillas en el aula, podemos, sin contar el número de alumnos y de sillas, sentar a cada alumno en una silla determinada. Si hay lugar para todos y no queda ningún asiento sobrante, es decir, si se establece una correspondencia biunívoca entre estos dos conjuntos, ello significará que tienen el mismo número de elementos.

Señalemes ahora que el primer camino (calculando el número de elementos) es válido sólo si se comparan conjuntos finitos, mientras que el segundo (estableciendo una correspondencia biunívoca) se puede aplicar también a conjuntos infinitos.

2°. **Conjuntos numerables.** El conjunto infinito más elemental es el conjunto de los números naturales. Llamaremos *conjunto numerable* a todo conjunto cuyos elementos se puedan poner en correspondencia biunívoca con todos los números naturales. En otras palabras, un conjunto numerable es un conjunto cuyos elementos se pueden colocar en una sucesión infinita: $a_1, a_2, \ldots$ $\ldots, a_n, \ldots$ Veamos algunos ejemplos de conjuntos numerables.

1. El conjunto de todos los números enteros. Establezcamos la correspondencia entre todos números enteros y todos los números naturales según el esquema siguiente:

$$\begin{array}{cccccc} 0 & -1 & 1 & -2 & 2 & \ldots \\ 1 & 2 & 3 & 4 & 5 & \ldots \end{array}$$

En general, pongamos en correspondencia a cada número no negativo $n \geqslant 0$ el número impar $2n+1$ y a cada número negativo $n < 0$ el número par $2|n|$:

$$n \leftrightarrow 2n+1, \text{ cuando } n \geqslant 0,$$
$$n \leftrightarrow 2|n|, \text{ cuando } n < 0.$$

2. El conjunto de todos los números pares positivos. La correspondencia es evidente: $n \leftrightarrow 2n$.

3. El conjunto 2, 4, 8, $\ldots$, 2^n, $\ldots$ de potencias de dos. Aquí la correspondencia es también evidente. A cada número 2^n se pone en correspondencia el número n.

4. Consideremos un ejemplo más complejo demostrando que el conjunto de todos los números racionales es numerable. Cada número racional se puede escribir, de manera única, en forma de una fracción irreducible $\alpha = \frac{p}{q}$, $q > 0$. Llamemos *altura* del número racional α a la suma $|p| + q$. Está claro que el número de fracciones de altura dada n es finito. Por ejemplo, la altura 1 la tiene sólo el número $\frac{0}{1} = 0$, la altura 2 la tienen sólo los números $\frac{1}{1}$ y $\frac{-1}{1}$, la altura 3, la tienen sólo los números $\frac{2}{1}$, $\frac{1}{2}$, $\frac{-2}{1}$ y $\frac{-1}{2}$, etc. Coloquemos ahora todos los números raciona-

les según su altura, es decir, primero los números de altura 1, después los de altura 2, etc. Cada número racional tendrá entonces su número, es decir, quedará establecida una correspondencia biunívoca entre todos los números naturales y todos los números racionales.

Un conjunto infinito que no sea numerable se llama conjunto *no numerable*. Demostremos algunas propiedades generales de conjuntos numerables.

1. *Todo subconjunto de un conjunto numerable es finito o numerable.*

DEMOSTRACION. Sea A un conjunto numerable y B un subconjunto suyo. Numeremos todos los elementos del conjunto A: a_1, $a_2, \dots, a_n, \dots$ Sean $a_{n_1}, a_{n_2}, \dots$ aquellos elementos que figuran en B. Si entre los números $n_1, n_2, \dots$ existe el máximo, B es finito; en el caso contrario B es numerable.

2. *La suma de cualquier conjunto finito o numerable de conjuntos numerables es también un conjunto numerable.*

DEMOSTRACION. Sean $A_1, A_2, \dots$ conjuntos numerables. Podemos suponer que son disjuntos (sin elementos comunes) dos a dos, ya que, de lo contrario, consideraríamos en su lugar los conjuntos A_1, $A_2 \setminus A_1$, $A_3 \setminus (A_1 \cup A_2)$, ... que son a lo sumo numerables y que tienen la misma suma que los conjuntos $A_1, A_2, \dots$ Todos los elementos de los conjuntos $A_1, A_2, \dots$ pueden escribirse en forma de la siguiente tabla infinita:

$$\begin{array}{ccccc}
a_{11} & a_{12} & a_{13} & a_{14} & \dots \\
a_{21} & a_{22} & a_{23} & a_{24} & \dots \\
a_{31} & a_{32} & a_{33} & a_{34} & \dots \\
a_{41} & a_{42} & a_{43} & a_{44} & \dots \\
\cdot & \cdot & \cdot & \cdot & \cdot
\end{array}$$

en cuya primera fila aparecen los elementos del conjunto A_1, en la segunda fila, los elementos del conjunto A_2, etc. Numeremos ahora todos estos elementos desplazándonos «en diagonales», es decir, tomando por primero el elemento a_{11}; por segundo, el elemento a_{12}; por tercero, el elemento a_{21}, etc., siguiendo el sentido que indican las flechas del siguiente cuadro:

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

Está claro que cada elemento de cada conjunto recibirá entonces un número determinado, es decir, quedará establecida una correspondencia biunívoca entre todos los elementos de todos los conjuntos A_1, A_2, ... y todos los números naturales. Nuestra afirmación resulta demostrada.

EJERCICIOS. 1. Demostrar que el conjunto de todos los polinomios con coeficientes racionales es numerable.

2. El número ξ se denomina *algebraico* si es raíz de un polinomio con coeficientes racionales. Demostrar que el conjunto de todos los números algebraicos es numerable.

3. Demostrar que el conjunto de todos los intervalos racionales (es decir, intervalos con extremos racionales) sobre la recta es numerable

4. Demostrar que el conjunto de todos los puntos del plano que tienen coordenadas reales es numerable.

Sugerencia. Empléese la propiedad 2.

3. *Todo conjunto infinito contiene un subconjunto numerable.*

DEMOSTRACION. Sea M un conjunto infinitio. Tomemos en él un elemento cualquiera a_1. Por ser M un conjunto infinito encontraremos en él un elemento a_2 distinto de a_1, después el elemento a_3 distinto de a_1 y a_2, etc. Continuando este proceso (que no podrá interrumpirse por «falta» de elementos ya que M es infinito) obtendremos un subconjunto numerable

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

del conjunto M. Hemos demostrado la afirmación.

Este resultado señala que los conjuntos numerables son los «más pequeños» de los conjuntos infinitos. El problema sobre la existencia de conjuntos infinitos no numerables será considerado más adelante.

3°. Equivalencia de conjuntos. Buscando una correspondencia biunívoca entre unos u otros conjuntos infinitos y los números naturales, hemos llegado al concepto de conjunto numerable.

Está claro que de manera análoga pueden compararse los conjuntos no sólo con el conjunto de números naturales: este procedimiento permite comparar entre sí dos conjuntos cualesquiera. Introduzcamos la siguiente definición.

DEFINICION. Dos conjuntos M y N se llaman *equivalentes* (lo que se denota mediante $M \sim N$) si entre sus elementos se puede establecer una correspondencia biunívoca.

El concepto de equivalencia puede aplicarse a cualesquiera conjuntos tanto finitos como infinitos. Dos conjuntos finitos son equivalentes entre sí si (y sólo si) tienen el mismo número de elementos. El conjunto numerable se puede ahora definir del siguiente modo:

un conjunto se llama numerable si es equivalente al conjunto de los números naturales.

Está claro que dos conjuntos, equivalentes cada uno a un tercer conjunto, son equivalentes entre sí; en particular, todos los conjuntos numerables son equivalentes entre sí.

Ejemplos. 1. Los conjuntos de puntos de dos cualesquiera segmentos $[a, b]$ y $[c, d]$ son equivalentes entre sí. En la fig. 5 se señala cómo se puede establecer una correspondencia biunívoca entre ellos: los puntos p y q corresponden uno al otro si se hallan sobre un mismo radio que parte del punto O donde se cruzan las rectas ac y bd.

2. El conjunto de todos los puntos del plano complejo es equivalente al conjunto de todos los puntos sobre una esfera.

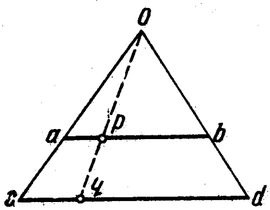


FIG. 5

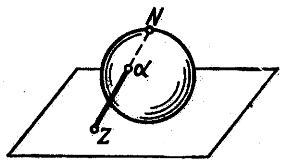


FIG. 6

La correspondencia biunívoca $\alpha \leftrightarrow z$ puede establecerse, por ejemplo, mediante la proyección estereográfica (fig. 6).

3. El conjunto de todos los números del intervalo $(0, 1)$ es equivalente al conjunto de todos los puntos de la recta. La correspondencia se puede establecer, por ejemplo, mediante la función

$$y = \frac{1}{\pi} \operatorname{arctg} x + \frac{1}{2}.$$

Al considerar los ejemplos de este punto y del punto 2° podemos observar que, a veces, un conjunto infinito puede resultar equivalente a una parte propia. Por ejemplo, resulta haber «tantos» números naturales como números enteros o incluso racionales; el intervalo $(0, 1)$ tiene «tantos» puntos como los tiene la recta, etc. Esta situación es característica para todos los conjuntos infinitos. En efecto, en el punto 2° (propiedad 3) hemos demostrado que de todo conjunto infinito M se puede elegir un subconjunto numerable; supongamos que éste sea el conjunto

$$A = \{a_1, a_2, \ldots, a_n, \ldots\}.$$

Dividamos A en dos subconjuntos numerables

$$A_1 = \{a_1, a_3, a_5 \ldots\} \text{ y } A_2 = \{a_2, a_4, a_6, \ldots\}.$$

Entre los conjuntos numerables A y A_1 se puede establecer una correspondencia biunívoca que se puede extender luego hasta una correspondencia biunívoca entre los conjuntos $A \cup (M \setminus A) = M$ y $A_1 \cup (M \setminus A) = M \setminus A_2$ refiriendo a cada elemento de $M \setminus A$ ese mismo elemento. El conjunto $M \setminus A_2$ es un subconjunto propio del conjunto M. Llegamos de esta forma a la siguiente proposición:

Todo conjunto infinito es equivalente a un subconjunto propio.

Esta propiedad se puede tomar como la definición de un conjunto infinito.

EJERCICIO. Demostrar que siendo M un conjunto infinito arbitrario y A numerable, resulta $M \sim A \cup A$.

4°. Innumerabilidad del conjunto de los números reales. En el punto 2° hemos visto varios ejemplos de conjuntos numerables. La cantidad de estos ejemplos se puede ampliar considerablemente. Además, hemos demostrado que tomando la suma finita o numerable de conjuntos numerables obtendremos de nuevo conjuntos numerables. Surge naturalmente la pregunta ¿existen conjuntos infinitos no numerables? La respuesta afirmativa la da el siguiente teorema.

TEOREMA 1. *El conjunto de números reales comprendidos entre el cero y la unidad es innumerable.*

DEMOSTRACION. Supongamas que existe una lista (de todos o una parte) de los números reales α, pertenecientes al segmento $[0, 1]$:

$$\left.\begin{array}{l} \alpha_1 = 0, a_{11}a_{12}a_{13} \ldots a_{1n} \ldots, \\ \alpha_2 = 0, a_{21}a_{22}a_{23} \ldots a_{2n} \ldots, \\ \alpha_3 = 0, a_{31}a_{32}a_{33} \ldots a_{3n} \ldots, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \alpha_n = 0, a_{n1}a_{n2}a_{n3} \ldots a_{nn} \ldots, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \end{array}\right\} \quad (1)$$

Aquí a_{ik} es la k-ésima cifra decimal del número α_i. Consideremos la fracción

$$\beta = 0, b_1 b_2 \ldots b_n \ldots$$

construida del siguiente modo: b_1 es una cifra arbitraria distinta de a_{11}; b_2 es una cifra arbitraria distinta de a_{22}, etc., en general, b_n es una cifra arbitraria distinta de a_{nn}. Esta fracción decimal no puede coincidir con ninguna de las que figuran en la lista (1). En efecto, la fracción β se distingue de la fracción α_1 al menos por su primera cifra; de la segunda fracción, por su segunda cifra, etc.; en general, puesto que $b_n \neq a_{nn}$ para todo n, la fracción β se distingue de cualquiera de las fracciones α_i que figuran en la lista (1). Luego, ninguna lista de números reales pertenecientes al segmento [0, 1] puede consumirlo.

La demostración expuesta necesita una pequeña precisión puesto que algunos números $\left(\text{concretamente los números de la forma } \frac{p}{10^q}\right)$ pueden tener dos representaciones decimales: bien con un número infinito de ceros o bien con un número infinito de nueves; por ejemplo:

$$\frac{1}{2} = \frac{5}{10} = 0{,}5000\ldots = 0{,}4999\ldots$$

Por consiguiente, el hecho de que dos fracciones decimales no coincidan no significa todavía que representan números distintos.

Sin embargo, si la fracción β se construye de manera que no contenga ni ceros ni nueves, (tomando, por ejemplo, $b_n = 2$, si $a_{nn} = 1$, y $b_n = 1$, si $a_{nn} \neq 1$) esta objeción quedará superada.

EJERCICIO. Demostrar que los números que tienen dos distintas representaciones decimales forman un conjunto numerable.

De manera que el segmento [0, 1] ofrece un ejemplo de un conjunto infinito no numerable. Veamos algunos ejemplos de conjuntos equivalentes al conjunto formado por los puntos del segmento [0, 1].

1. El conjunto de todos los puntos pertenecientes a un segmento $[a, b]$ o intervalo (a, b) cualquiera.
2. El conjunto de todos los puntos de la recta.
3. Los conjuntos de todos los puntos del plano, del espacio, de la superficie de una esfera, de los puntos que se encuentran dentro de una esfera, etc.
4. El conjunto de todas las rectas del plano.
5. El conjunto de todas las funciones continuas de una o varias variables.

En los casos 1 y 2 la demostración no ofrece dificultades (véanse los ejemplos 1 y 3 del punto 3°). En los demás casos la demostración directa no es tan sencilla.

EJERCICIO. Demostrar, aprovechando los resultados de este punto y del ejercicio 2 del punto 2°, la existencia de números trascendentes, es decir, de números no algebraicos.

5°. Concepto de potencia de un conjunto. Si dos conjuntos finitos son equivalentes, tienen el mismo número de elementos. Cuando dos conjuntos equivalentes entre sí M y N son arbitrarios se dice que M y N tienen la misma *potencia*. Así pues, la potencia es aquello común que tienen todos los conjuntos equivalentes entre sí. En el caso de conjuntos finitos el concepto de potencia coincide con el concepto habitual del número de elementos del conjunto. La potencia del conjunto de los números naturales (es decir, de cualquier conjunto numerable) se denota mediante el símbolo $\aleph_0$ (se lee "alef cero"). Los conjuntos equivalentes al conjunto de todos los números naturales comprendidos entre 0 y 1 se dice que tienen potencia de *continuo*. Esta potencia se denota mediante el símbolo c (o el símbolo $\aleph$).

En las observaciones que concluyen este capítulo tocamos el problema, muy profundo, de la existencia de potencias intermedias entre $\aleph_0$ y c. Como regla general, los conjuntos infinitos que se emplean en el análisis son numerables o tienen potencia de continuo. En el caso de las potencias de conjuntos finitos, es decir, en el caso de los números naturales, tenemos, además del concepto de igualdad, los conceptos de «más» y «menos». Veamos cómo pueden extenderse estos últimos al caso de potencias infinitas.

Sean A y B dos conjuntos arbitrarios y $m(A)$ y $m(B)$ sus potencias. Si A es equivalente a B, $m(A)=m(B)$ por definición. Si A es equivalente a una parte del conjunto B y A no contiene ninguna parte equivalente a B, se acepta, naturalmente, que $m(A)$ es menor que $m(B)$, es decir, $m(B)$ es mayor que $m(A)$. Sin embargo, existen, lógicamente, otras dos posibilidades, además de las mencionadas:

a) B contiene una parte equivalente a A y A contiene una parte equivalente a B.

b) A y B no son equivalentes y ninguno posee una parte equivalente al otro.

Del teorema de Cantor — Bernstein, que se expone en el siguiente punto, se desprende que en el caso a) los conjuntos A y B son equivalentes, es decir tienen potencias iguales. En cuanto al caso b), que equivaldría a la existencia de potencias incomparables, resulta que no puede realizarse. Esto se deduce del teorema de Zermelo que enunciamos en el § 4.

De lo anterior resulta (si aceptamos sin demostración los teoremas de Cantor — Bernstein y de Zermelo) que dos conjuntos arbitrarios A y B o bien tienen la misma potencia o bien verifican una de las relaciones

$$m(A) < m(B) \quad \text{o} \quad m(A) > m(B).$$

Hemos señalado más arriba que los conjuntos numerables son los «más pequeños» entre los conjuntos infinitos y hemos demostrado luego que existen conjuntos infinitos de potencia superior: los conjuntos de potencia de continuo. ¿Existen potencias infinitas superiores a la potencia de continuo? En general, ¿existe o no una potencia «máxima»? Resulta que tiene lugar el siguiente teorema:

TEOREMA 2. *Sea M un conjunto cualquiera y $\mathfrak{M}$ el conjunto formado por todos los subconjuntos del conjunto M. Entonces, la potencia de $\mathfrak{M}$ es superior a la potencia del conjunto inicial M.*

DEMOSTRACION. Es fácil ver que la potencia $\mathfrak{m}$ del conjunto $\mathfrak{M}$ no puede ser inferior a la potencia m del conjunto inicial M; en efecto, los subconjuntos de M formados por un solo elemento representan en $\mathfrak{M}$ un subconjunto equivalente al conjunto M. Basta demostrar que las potencias $\mathfrak{m}$ y m no coinciden. Supongamos que entre los elementos $a, b, \ldots$ del conjunto M y algunos elementos $A, B, \ldots$ del conjunto $\mathfrak{M}$ (es decir, algunos subconjuntos de M) se ha logrado establecer una correspondencia biunívoca:

$$a \leftrightarrow A, \quad b \leftrightarrow B, \quad \ldots$$

Demostremos que esta correspondencia no puede consumir todos los subconjuntos del conjunto M, es decir, todos los elementos del conjunto $\mathfrak{M}$. Sea X la colección de elementos de M que no pertenecen a los subconjuntos que les corresponden. Más detalladamente: si $a \leftrightarrow A$ y $a \in A$, el elemento a no se incluye en X, pero si $a \leftrightarrow A$ y $a \overline{\in} A$, el elemento a se incluye en X. Es evidente que X representa un subconjunto de M, es decir, un elemento de $\mathfrak{M}$. Demostremos que al subconjunto X no le puede corresponder ningún elemento de M. Supongamos que tal elemento $x \leftrightarrow X$ existe y veamos si pertenece o no al subconjunto X. Aceptemos que $x \overline{\in} X$; por definición, en X figura t o d o elemento que no pertenece al subconjunto que le corresponde y, por lo tanto, x debe ser incluido en X. Viceversa, si se acepta que el elemento x pertenece a X, deduciremos que x no puede figurar en X, ya que éste contiene sólo los elementos que no aparecen en los subconjuntos que les corresponden. De manera que el elemento x correspondiente al subconjunto X debe simultáneamente pertenecer y no pertenecer a X. De aquí se desprende que tal elemento no existe, es decir, que no se puede establecer una correspondencia biunívoca entre los elementos del conjunto M y todos sus subconjuntos. El teorema queda demostrado.

De este modo, para cualquier potencia podemos construir efectivamente un conjunto de potencia superior, después otro de potencia aún mayor, etc., obteniendo así una escala de potencias no acotada superiormente.

EJERCICIO. Demostrar que la totalidad de las funciones numéricas (o, en general de funciones que toman valores en un conjunto compuesto por lo menos de dos elementos) definidas sobre un conjunto M tiene una potencia superior a la de M.

Sugerencia. Aprovéchese que el conjunto de todas las funciones características sobre M (es decir, de funciones que sólo toman valores 0 y 1) es equivalente al conjunto formado por todos los subconjuntos de M.

6°. Teorema de Cantor-Bernstein. Demostraremos ahora el siguiente teorema importante al que nos hemos referido ya en el punto anterior.

TEOREMA (CANTOR — BERNSTEIN). *Sean A y B dos conjuntos arbitrarios. Si en A existe un subconjunto A_1 equivalente a B y en B un subconjunto B_1 equivalente a A, los conjuntos A y B son equivalentes.*

DEMOSTRACION. Sea f la aplicación biyectiva[1] de A en B_1 y sea g la aplicación biyectiva de B en A_1, es decir,

$$f(A)=B_1\subset B, \quad g(B)=A_1\subset A.$$

Entonces, $gf(A)=g(f(A))=g(B_1)$ es un conjunto, que denotaremos con A_2, contenido en A_1 y equivalente al conjunto A. De manera análoga $fg(B)=f(g(B))=f(A_1)=B_2$ es un conjunto contenido en B_1 y equivalente a B. Sea ahora A_3 aquel subconjunto de A en el que se transforma el conjunto A_1 mediante la aplicación gf y sea A_4 aquel conjunto en el que se transforma A_2 mediante la misma aplicación gf. En general, sea A_{k+2} aquel conjunto en el que se transforma el conjunto A_k mediante la aplicación gf $(k=1, 2, \ldots)$. Está claro que

$$A\supset A_1\supset A_2\supset \ldots \supset A_k\supset A_{k+1}\supset \ldots$$

Si ponemos

$$D=\bigcap_{k=1}^{\infty} A_k,$$

podremos representar A mediante la siguiente suma de conjuntos disjuntos dos a dos:

$$A=(A\setminus A_1)\cup(A_1\setminus A_2)\cup(A_2\setminus A_3)\cup \ldots \cup(A_k\setminus A_{k+1})\cup \ldots \cup D.$$

---

[1] A veces se emplea el término «aplicación biunívoca». (*N. del T.*)

De un modo análogo el conjunto A_1 se puede representar en la forma:

$$A_1 = D \cup (A_1 \setminus A_2) \cup (A_2 \setminus A_3) \cup \ldots \cup (A_k \setminus A_{k+1}) \cup \ldots$$

Evidentemente, estas dos fórmulas pueden escribirse así:

$$A = D \cup [(A_1 \setminus A_2) \cup (A_3 \setminus A_4) \cup \ldots] \cup \\ \cup [(A \setminus A_1) \cup (A_2 \setminus A_3) \cup \ldots], \quad (2)$$

$$A_1 = D \cup [(A_1 \setminus A_2) \cup (A_3 \setminus A_4) \cup \ldots] \cup \\ \cup [(A_2 \setminus A_3) \cup (A_4 \setminus A_5) \cup \ldots]. \quad (3)$$

Observemos ahora que el conjunto $A \setminus A_1$ es equivalente al conjunto $A_2 \setminus A_3$ (ya que el primero se transforma en el segundo

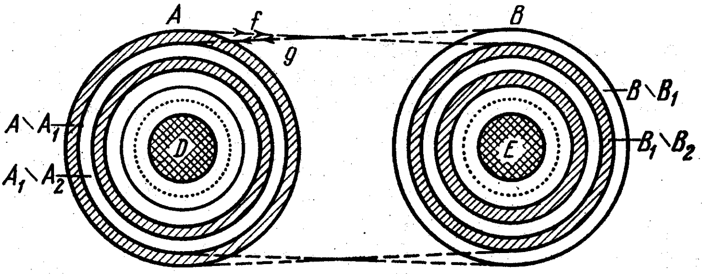


FIG. 7

mediante la aplicación gf); del mismo modo $A_2 \setminus A_3$ es equivalente a $A_4 \setminus A_5$, etc. Por eso, los conjuntos que figuran en las segundas líneas de las fórmulas (2) y (3) son equivalentes. En cuanto a las primeras líneas de estas fórmulas, son sencillamente idénticas. De aquí se desprende que entre los elementos de los conjuntos A y A_1 se puede establecer una correspondencia biunívoca. Pero A_1 es equivalente a B por hipótesis. De manera que A es equivalente a B. El teorema queda demostrado.

Podemos incluso «permitirnos el lujo» (aunque no hay necesidad de ello) de escribir explícitamente la correspondencia biunívoca φ que transforma A en B. Es la siguiente:

$$\varphi(a) = \begin{cases} g^{-1}(a), & \text{si } a \in D \cup (A_1 \setminus A_2) \cup (A_3 \setminus A_4) \cup \ldots \\ f(a), & \text{si } a \in (A \setminus A_1) \cup (A_2 \setminus A_3) \cup \ldots \end{cases}$$

(fig. 7)

§ 3. APLICACIONES. PARTICIÓN EN CLASES

1°. Aplicaciones de conjuntos. Concepto general de función. En el Análisis el concepto de función se introduce del siguiente modo. Sea X un conjunto sobre la recta numérica. Se dice que sobre este conjunto está definida una función f si a cada elemento $x \in X$ se le ha puesto en correspondencia un número determinado $y = f(x)$. El conjunto X se denomina en este caso *campo de deficinión* de esta función y el conjunto Y formado por todos los valores que toma esta función, *campo de valores* de la misma.

Si en vez de conjuntos numéricos consideramos conjuntos de naturaleza arbitraria, llegaremos al concepto más general de función: sean M y N dos conjuntos arbitrarios. Se dice que está definida sobre M una función f con valores en N si a cada elemento $x \in M$ se le pone en correspondencia un elemento, y sólo uno, y de N. En el caso de conjuntos de naturaleza arbitraria (y a veces también en el caso de conjuntos numéricos) en lugar del término "función" se usa «aplicación» y se habla de la aplicación de un conjunto en otro [1].

Si a es un elemento de M, el elemento $b = f(a)$ de N que le corresponde se denomina *imagen* del elemento a (para la aplicación f). La colección de todos los elementos de M que tienen por imagen el elemento $b \in N$ se denomina *imagen recíproca* [2] (én términos más precisos, *imagen recíproca completa*) del elemento b y se denota mediante $f^{-1}(b)$.

Sea A un conjunto de M; la totalidad $\{f(a): a \in A\}$ de todos los elementos del tipo $f(a)$, donde $a \in A$, se denomina *imagen de* A y se designa $f(A)$. Para todo conjunto B de N se puede, a su vez, definir la imagen recíproca $f^{-1}(B)$; a saber: $f^{-1}(B)$ es la totalidad de todos aquellos elementos de M cuyas imágenes pertenecen a B. Puede ocurrir que ningún elemento b de B tiene imagen recíproca y en este caso la imagen recíproca completa $f^{-1}(B)$ será el conjunto vacío.

Aquí nos limitaremos a exponer las propiedades más generales de las aplicaciones.

Convendremos en emplear la siguiente terminología. Diremos que f es una aplicación del conjunto M «sobre» el conjunto N, si $f(M) = N$; en el caso general, es decir, cuando $f(M) \subset N$, se dice que f es una aplicación de M «en» N.

Señalemos las propiedades principales de las aplicaciones.

---

[1] De hecho hemos tropezado ya anteriormente con el concepto de aplicaciones de conjuntos (por ejemplo, al introducir el concepto de equivalencia de conjuntos, al demostrar el teorema de Cantor—Bernstein, etc.).

[2] En algunos libros de texto se emplea el término «preimagen». (*Nota del T.*)

TEOREMA 1. *La imagen recíproca de la unión de dos conjuntos es igual a la unión de sus imágenes recíprocas:*

$$f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B).$$

DEMOSTRACION. Supongamos que el elemento x pertenece al conjunto $f^{-1}(A \cup B)$. Ello significa que $f(x) \in A \cup B$, o sea, $f(x) \in A$ o bien $f(x) \in B$. En este caso x debe pertenecer al menos a uno de los dos conjuntos $f^{-1}(A)$ y $f^{-1}(B)$, es decir, $x \in f^{-1}(A) \cup f^{-1}(B)$. Viceversa, si $x \in f^{-1}(A) \cup f^{-1}(B)$, entonces x pertenece al menos a uno de los conjuntos $f^{-1}(A)$ y $f^{-1}(B)$, es decir, $f(x)$ figura al menos en uno de los conjuntos A y B; en otras palabras, $f(x) \in A \cup B$ y, por lo tanto, $x \in f^{-1}(A \cup B)$.

TEOREMA 2. *La imagen recíproca de la intersección de dos conjuntos es igual a la intersección de sus imágenes recíprocas:*

$$f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B).$$

DEMOSTRACION. Si $x \in f^{-1}(A \cap B)$, debe ser $f(x) \in A \cap B$, o sea, $f(x) \in A$ y $f(x) \in B$, y, por lo tanto, $x \in f^{-1}(A)$ y $x \in f^{-1}(B)$, es decir, $x \in f^{-1}(A) \cap f^{-1}(B)$.

Viceversa, si $x \in f^{-1}(A) \cap f^{-1}(B)$, es decir, $x \in f^{-1}(A)$ y $x \in f^{-1}(B)$, entonces, $f(x) \in A$ y $f(x) \in B$, en otras palabras, $f(x) \in A \cap B)$. Por lo tanto, $x \in f^{-1}(A \cap B)$.

Los teoremas 1 y 2 se verifican también en el caso de la unión o intersección de un número arbitrario (finito o infinito) de conjuntos.

TEOREMA 3. *La imagen de la unión de dos conjuntos es igual a la unión de sus imágenes:*

$$f(A \cup B) = f(A) \cup f(B).$$

DEMOSTRACION. Si $y \in f(A \cup B)$, esto significa que $y = f(x)$, donde x pertenece al menos a uno de los conjuntos A y B. Por consiguiente, $y = f(x) \in f(A) \cup f(B)$. Viceversa, si $y \in f(A) \cup f(B)$, entonces, $y = f(x)$, donde x pertenece al menos a uno de los conjuntos A y B, es decir, $x \in A \cup B$ y, por lo tanto, $y = f(x) \in f(A \cup B)$.

Vale subrayar que *la imagen de la intersección de dos conjuntos no coincide, en general, con la intersección de sus imágenes.* Supongamos, por ejemplo, que la aplicación considerada representa la proyección del plano sobre el eje x. En este caso los segmentos

$$0 \leqslant x \leqslant 1, \quad y = 0,$$
$$0 \leqslant x \leqslant 1, \quad y = 1$$

no se intersecan y, sin embargo, sus imágenes coinciden.

2°. Partición en clases. Relación de equivalencia. En diferentes cuestiones tropezamos con la partición de unos u otros conjuntos en subconjuntos disjuntos dos a dos. Por ejemplo, se puede partir el plano (considerado como un conjunto de puntos) en rectas paralelas al eje x; podemos imaginarnos el espacio de tres dimensiones como un conjunto de esferas concéntricas de distintos radios; se puede dividir a los habitantes de una ciudad determinada en grupos según el año de nacimiento, etc.

Cada vez que un conjunto M se representa, de uno u otro modo, como la unión de subconjuntos disjuntos dos a dos, hablamos de la *partición del conjunto M en clases.*

Comúnmente nos encontramos con particiones realizadas de acuerdo con uno u otro criterio, según el cual se unen en clases los elementos del conjunto M. Por ejemplo, la totalidad de los triángulos del plano se puede partir en clases de triángulos iguales o en clases de triángulos de la misma área; todas las funciones de x pueden partirse en clases agrupando en una misma clase todas las funciones que tienen el mismo valor en el punto dado x, etc.

Los criterios, según los cuales se realiza la partición en clases de los elementos de uno u otro conjunto, pueden ser muy diversos. Sin embargo, no son del todo arbitrarios. Supongamos, por ejemplo, que queremos partir en clases todos los números reales, incluyendo el número b en la misma clase que el número a si, y sólo si, $b > a$. Está claro que de esta forma no podremos obtener ninguna partición de los números reales en clases, ya que si $b > a$, es decir, si el elemento b debe ser incluido en la misma clase que a, entonces, $a < b$, es decir, el número a *no se puede incluir* en la misma clase que el número b. Además, como a no es mayor de sí mismo, ¡*a no debe figurar* en la clase con sí mismo! Otro ejemplo. Veamos si se puede partir en clases los puntos del plano incluyendo dos puntos en una misma clase si, y sólo si, la distancia entre ellos es menor de 1. Está claro que es imposible realizar esta partición, puesto que si la distancia entre a y b y entre b y c es inferior a 1, ello no significa, de ningún modo, que la distancia de a a c es menor de 1. Por eso, al incluir a en la misma clase con b y b en la misma clase con c, resultará que en una misma clase pueden aparecer dos puntos, la distancia entre los cuales es mayor a 1.

Los ejemplos dados sugieren las condiciones que aseguran que uno u otro criterio permite verdaderamente realizar la partición de los elementos de un conjunto en clases.

Sea M un conjunto y sean «marcados» algunos de los pares (a, b) de elementos de este conjunto[1]. Si (a, b) es un par

---

[1] Con la particularidad de que los elementos a y b se toman en un orden determinado, es decir, (a, b) y (b, a) son, en general, pares distintos

«marcado», diremos que el elemento a está ligado al elemento b por la relación φ, lo que denotaremos con el símbolo $a \underset{\varphi}{\sim} b$. Por ejemplo, si se trata de la partición de los triángulos en clases de triángulos de la misma área, $a \underset{\varphi}{\sim} b$ significa que «el triángulo a tiene la misma área que el triángulo b». Una relación φ se llama *relación de equivalencia* si, por sus propiedades, es:

1) *reflexiva*: $a \underset{\varphi}{\sim} a$ para todo elemento $a \in M$;
2) *simétrica*: si $a \underset{\varphi}{\sim} b$, entonces $b \underset{\varphi}{\sim} a$;
3) *transitiva*: si $a \underset{\varphi}{\sim} b$ y $b \underset{\varphi}{\sim} c$, entonces $a \underset{\varphi}{\sim} c$.

Estas condiciones son necesarias y suficientes para que la relación φ (¡el criterio!) permita partir en clases el conjunto M. En efecto, *toda* partición de un conjunto en clases determina una relación de equivalencia entre sus elementos: si $a \underset{\varphi}{\sim} b$ significa que «a se encuentra en la misma clase que b», es fácil probar que la relación φ será reflexiva, simétrica y transitiva. Viceversa, si φ es una relación de equivalencia entre los elementos del conjunto M, siempre obtendremos una partición de este conjunto en clases incluyendo en una misma clase aquellos elementos de M, y sólo aquellos, que son equivalentes.

Efectivamente, sea K_a la clase de elementos de M equivalentes a un elemento fijado a. De la propiedad reflexiva se desprende que el propio elemento a pertenece a la clase K_a. Probemos ahora que dos clases K_a y K_b o bien coinciden o bien no tienen elementos comunes. Supongamos que existe un elemento c que pertenece simultáneamente a K_a y K_b, es decir, $c \underset{\varphi}{\sim} a$ y $c \underset{\varphi}{\sim} b$.

Entonces $a \underset{\varphi}{\sim} c$, debido a la propiedad simétrica, y

$$a \underset{\varphi}{\sim} b \tag{1}$$

debido a la propiedad transitiva.

Si x es ahora un elemento arbitrario de K_a, es decir, $x \underset{\varphi}{\sim} a$, de (1) y de la propiedad transitiva se sigue que $x \underset{\varphi}{\sim} b$, es decir, $x \in K_b$.

De la misma forma se demuestra que todo elemento $y \in K_b$ figura en K_a. Por consiguiente, dos clases K_a y K_b que tienen al menos un elemento común coinciden. De manera que a partir de la relación de equivalencia dada hemos obtenido efectivamente una partición del conjunto M en clases.

El concepto de partición de un conjunto en clases está estrechamente vinculado al concepto de aplicación considerado en el punto anterior.

Sea f una aplicación del conjunto A en el conjunto B. Si unimos en una misma clase todos aquellos elementos de A cuyas imágenes en B coinciden, obtendremos evidentemente una partición del conjunto A. Viceversa, consideremos un conjunto arbitrario A y alguna partición de este conjunto en clases. Sea B

la totalidad de clases en las que se ha partido el conjunto A. Si a cada elemento $a \in A$ ponemos en correspondencia aquella clase (es decir, aquel elemento de B) a la que a pertenece, obtendremos una aplicación del conjunto A sobre el conjunto B.

Ejemplos. 1. Consideremos la proyección del plano xy sobre el eje x. Las imágenes recíprocas de los puntos del eje x son las rectas verticales. Por consiguiente, a esta aplicacion le corresponde la partición del plano en rectas paralelas.

2. Dividamos todos los puntos del espacio de tres dimensiones en clases incluyendo en una misma clase los puntos equidistantes del origen de coordenadas. De este modo, cada clase representa una esfera de un radio determinado. La totalidad de estas clases puede identificarse con el conjunto de todos los puntos pertenecientes al rayo $[0, \infty)$. Por lo tanto, a la partición del espacio tridimensional en esferas concéntricas le corresponde la aplicación de este espacio sobre una semirrecta.

3. Agrupemos en una misma clase todos los números reales que tienen la misma parte fraccionaria. La aplicación que corresponde a esta partición representa la aplicación de la recta sobre la circunferencia de longitud unidad.

El concepto de equivalencia es un caso particular del concepto más general de *relación binaria* que se define del siguiente modo. Sea M un conjunto arbitrario. Designemos mediante M^2 el conjunto de todos los pares ordenados (a, b), donde $a, b \in M$. Se dice que en M se tiene una relación binaria φ, si en M^2 se escoge un subconjunto R_φ arbitrario. Más exactamente, diremos que el elemento a está en relación φ con el elemento b, y escribiremos $a\varphi b$, si, y sólo si, (a, b) pertenece a R_φ. Un ejemplo de relación binaria es la relación de identidad E consistente en que aEb si, y sólo si, $a=b$; en otras palabras, ésta es la relación determinada por el subconjunto de pares de tipo (a, a). Está claro que toda relación de equivalencia φ en un conjunto M es una relación binaria; pero no es arbitraria sino que verifica las siguientes condiciones:

1) Todo par de tipo (a, a), donde $a \in M$, pertenece a R_φ (condición reflexiva).

2) Si $(a, b) \in R_\varphi$ y $(b, c) \in R_\varphi$, entonces también $(a, c) \in R_\varphi$ (condición transitiva).

3) Si $(a, b) \in R_\varphi$, entonces también $(b, a) \in R_\varphi$ (condición simétrica).

De manera que la relación de equivalencia es una relación binaria que cumple las condiciones reflexiva, transitiva y simétrica. En el párrafo siguiente estudiaremos otro caso particular importante de relación binaria, la ordenación parcial.

§ 4. CONJUNTOS ORDENADOS. NÚMEROS TRANSFINITOS

En este párrafo exponemos varios conceptos relacionados con la idea de ordenación de los elementos de los conjuntos, limitándonos a dar las nociones elementales; una exposición más detallada se puede encontrar en la bibliografía que señalamos al final del libro.

1°. Conjuntos parcialmente ordenados. Sea M un conjunto arbitrario y φ una relación binaria en él. Se dice que ella es una relación de orden parcial si verifica la propiedad reflexiva $[(a, a) \in R_\varphi]$, la propiedad transitiva [si $(a, b \in R_\varphi$ y $(b, c) \in R$, entonces $(a, c) \in R_\varphi]$ y la siguiente condición conocida como propiedad *antisimétrica:* si $a\varphi b$ y $b\varphi a$, entonces $a = b$. El orden parcial se denota generalmente por medio del símbolo $\leqslant$. De manera que $a \leqslant b$ significa que el par (a, b) pertenece al conjunto R_φ correspondiente. Entonces se dice que el elemento a no supera a b o bien que está contenido en b. Un conjunto en el que está dado un orden parcial se dice *parcialmente ordenado*.

Veamos ejemplo de conjuntos parcialmente ordenados.

1. Todo conjunto puede considerarse, de un modo trivial, como un conjunto parcialmente ordenado aceptando que $a \leqslant b$ si, y sólo si, $a = b$. En otras palabras, el orden parcial puede determinarse en cualquier conjunto mediante la relación binaria de identidad E. Este ejemplo no es, por supuesto, de gran interés.

2. Sea M el conjunto de todas las funciones continuas sobre el segmento $[\alpha, \beta]$. Obtendremos evidentemente un orden parcial aceptando que $f \leqslant g$ si, y sólo si, $f(t) \leqslant g(t)$ para todo t, $\alpha \leqslant t \leqslant \beta$.

3. El conjunto de todos los subconjuntos de un conjunto dado resulta parcialmente ordenado si $M_1 \leqslant M_2$ significa que $M_1 \subset M_2$.

4. El conjunto de todos los números naturales resulta parcialmente ordenado si $a \leqslant b$ significa «b es divisible por a».

Sea M un conjunto arbitrario parcialmente ordenado. En el caso de que $a \leqslant b$ y $a \neq b$ emplearemos el símbolo $<$, es decir, escribiremos $a < b$, y diremos que *a es menor que b* o bien *a está contenido estrictamente en b*. A veces en lugar de $a \leqslant b$ emplearemos la denotación equivalente $b \geqslant a$ y diremos en este caso que b es no menor que a (o mayor que a, si $b \neq a$) o bien que b *sigue* a a. El elemento a de un conjunto parcialmente ordenado se denomina *maximal* si de $a \leqslant b$ se deduce que $b = a$. El elemento a se llama *minimal* si de $c \leqslant a$ se sigue que $c = a$.

Un conjunto parcialmente ordenado se llama *dirigido* si para cualesquiera dos de sus puntos a y b existe un tercer punto c que los sigue ($a \leqslant c$, $b \leqslant c$).

2°. Aplicaciones que conservan el orden. Sean M y M' dos conjuntos parcialmente ordenados y f una aplicación biunívoca de M sobre M'. Diremos que esta aplicación *conserva el orden*, si de $a \leqslant b$, donde $a, b \in M$, se deduce que $f(a) \leqslant f(b)$ (en M'). La aplicación f se llama *isomorfismo* de los conjuntos parcialmente ordenados M y M' si $f(a) \leqslant f(b)$ se cumple *cuando, y sólo cuando*, $a \leqslant b$. Los propios conjuntos M y M' se denominan en este caso isomorfos.

Sea, por ejemplo, M el conjunto de los números naturales con el orden parcial señalado en el ejemplo 4 del punto 1 y sea M' el mismo conjunto, pero ordenado del modo natural, es decir, de manera que $b > a$ si $b - a$ es un número positivo. Entonces, la aplicación de M sobre M', que a cada número n le pone en correspondencia ese mismo número, conserva el orden (pero no representa un isomorfismo).

La propia relación de isomorfismo entre conjuntos parcialmente ordenados es, evidentemente, una relación de equivalencia (ya que es simétrica, transitiva y reflexiva). Por lo tanto, si tenemos una colección de conjuntos parcialmente ordenados, todos los conjuntos de esta colección[1] pueden dividirse en clases de conjuntos isomorfos. Cuando lo que nos interesa no es la naturaleza de los elementos de los conjuntos, sino sólo el orden parcial que en ellos existe, se puede, claro está, considerar como idénticos dos conjuntos parcialmente ordenados isomorfos.

3°. Conjuntos ordenados. Tipos ordinales. Puede suceder que siendo a y b elementos de un conjunto parcialmente ordenado no se cumpla ninguna de las relaciones $a \leqslant b$ y $b \leqslant a$. En este caso se dice que los elementos a y b son *incomparables*. De manera que la relación de orden resulta definida sólo para algunos pares de elementos y por eso hablamos del orden *parcial*. En cambio, si el conjunto parcialmente ordenado M no posee elementos incomparables, decimos que M es un *conjunto ordenado* (*linealmente ordenado, totalmente ordenado*). Es decir, M es un conjunto ordenado, si está parcialmente ordenado y si para cualesquiera dos elementos distintos $a, b \in M$ se tiene obligatoriamente que o bien $a < b$ o bien $b < a$.

Los conjuntos considerados en los ejemplos 1, 2, 3 y 4 del

---

[1] Nos abstenemos de emplear conceptos como «todos los conjuntos parcialmente ordenados» porque son, al igual que el concepto del «conjunto de todos los conjuntos», internamente contradictorios por su esencia y no pueden incluirse en concepciones matemáticas rigurosas.

primer punto son conjuntos parcialmente ordenados pero no ordenados. Ejemplos elementales de conjuntos ordenados son los números naturales, el conjunto de todos los números racionales, el conjunto de todos los números reales del segmento [0, 1], etc. (con las relaciones naturales de «mayor» y «menor» que existen en estos conjuntos).

Está claro que cualquier subconjunto de un conjunto ordenado también está ordenado.

El orden es un caso particular del orden parcial y por eso a los conjuntos ordenados se puede aplicar el concepto de aplicación que conserva el orden y, en particular, el concepto de isomorfismo.

Se dice que unos conjuntos tienen el mismo *tipo ordinal* si son ordenados e isomorfos. De manera que el tipo ordinal es lo común que tienen todos los conjuntos ordenados isomorfos, de la misma forma que la potencia es lo común que tienen todos los conjuntos equivalentes (considerados independientemente de cualquier relación de orden que pueda existir en ellos).

La serie de números naturales 1, 2, 3, ... con la relación natural de orden entre sus elementos es el ejemplo más elemental de conjunto ordenado. Su tipo ordinal se acostumbra a denotar con el símbolo ω.

Dos conjuntos ordenados isomorfos tienen, por supuesto, la misma potencia (ya que el isomorfismo es una correspondencia *biunívoca*) y por eso se puede hablar de la potencia que responde al tipo ordinal dado (por ejemplo, al tipo ω le corresponde la potencia $\aleph_0$). La afirmación recíproca, sin embargo, no es cierta: un conjunto de una potencia dada puede ser ordenado, en general, de diferentes modos. Sólo en el caso de conjuntos finitos el tipo ordinal queda determinado unívocamente por el número n de sus elementos (y se denota también mediante n). Pero ya en el conjunto numerable de los números naturales es posible, además de su tipo «natural» ω, considerar, por ejemplo, el siguiente tipo:

$$1, 3, 5, \ldots, 2, 4, 6, \ldots,$$

cuando cualquier número par sigue a cualquier impar y los números pares e impares se ordenan entre sí por su magnitud. Se puede probar que a la potencia $\aleph_0$ le corresponde una cantidad infinita, e incluso innumerable, de diferentes tipos ordinales.

4°. Suma ordenada de conjuntos ordenados. Sean M_1 y M_2 dos conjuntos ordenados y θ_1 y θ_2 sus tipos ordinales, respectivamente. En la unión $M_1 \cup M_2$ de los conjuntos M_1 y M_2 se puede introducir un orden aceptando que cada par de elementos de M_1 está ordenado igual que en M_1, que cada par de elementos de M_2 tiene el mismo orden que en M_2 y que *cualquier* elemento de M_1 *precede* a cualquier elemento de M_2. (¡Compruébese que así queda efectivamente establecido un orden!) El conjunto ordenado obtenido de esta forma lo llamaremos *suma ordenada* de los sonjuntos M_1 y M_2 y lo designaremos M_1+M_2. Subrayemos que es importante el orden de los sumandos: la suma M_2+M_1 no es isomorfa, en general, a la suma M_1+M_2.

Se dice que el tipo ordinal de la suma M_1+M_2 es la *suma ordenada* de los tipos ordinales θ_1 y θ_2 y se denota mediante $\theta_1+\theta_2$.

Esta definición se puede fácilmente extender al caso de un número finito arbitrario de sumandos $\theta_1, \theta_2, \ldots, \theta_m$.

Ejemplo. Consideremos los tipos ordinales ω y n. Es fácil ver que $n+\omega=\omega$; en efecto, si a la serie de los números naturales $1, 2, 3, \ldots, k, \ldots$ agregamos a su izquierda un número finito de elementos, obtendremos el mismo tipo ordinal ω. Sin embargo, el tipo ordinal $\omega+n$, es decir, el tipo ordinal del conjunto

$$1, 2, 3, \ldots, k, \ldots, a_1, a_2, \ldots a_n,$$

no es, evidentemente, igual a ω.

5°. Conjuntos bien ordenados. Números transfinitos. Hemos introducido anteriormente los conceptos de orden parcial y de orden. Ahora introduciremos un concepto más restringido, pero muy importante, el concepto de buen orden.

DIFINICION. Se dice que un conjunto ordenado *está bien ordenado* si cualquiera de sus subconjuntos no vacíos tiene un elemento mínimo (es decir, un elemento que precede a todos los del subconjunto).

Si un conjunto ordenado es finito, obviamente, está bien ordenado. Un ejemplo de un conjunto ordenado, pero no bien ordenado, nos ofrece la totalidad de los números racionales del segmento $[0, 1]$. Este conjunto, en sí, tiene un elemento mínimo, el número 0, pero el subconjunto suyo formado por los números racionales *positivos* no tiene elemento mínimo.

Está claro que todo subconjunto (no vacío) de un conjunto bien ordenado está bien ordenado.

El tipo ordinal de un conjunto bien ordenado se llama *número ordinal (número ordinal transfinito* o más brevemente *transfinito ordinal*, particularmente si se quiere subrayar que se trata de un conjunto infinito).

La serie de los números naturales (con la relación de orden corriente) representa no sólo un conjunto ordenado sino bien ordenado. De manera que su tipo ordinal ω es número ordinal

(¡transfinito!). También lo es $\omega+k$, es decir, el tipo del conjunto

$$1,\ 2,\ \ldots,\ n,\ \ldots,\ a_1,\ a_2,\ \ldots,\ a_k.$$

Al contrario, el conjunto

$$\ldots,\ -n,\ \ldots,\ -3,\ -2,\ -1 \qquad (1)$$

está ordenado, pero no está bien ordenado. Aquí todo subconjunto no vacío tiene un elemento máximo (es decir, un elemento que sigue a todos) pero no tiene, en general, un elemento mínimo (por ejemplo, el conjunto (1), en sí, no lo tiene). El tipo ordinal (¡que no es un número ordinal!) del conjunto (1) suele denotarse mediante el símbolo ω^*.

Demostremos la siguiente proposiicón, simple pero importante.

LEMA 1. *La suma ordenada de un número finito de conjuntos bien ordenados es un conjunto bien ordenado.*

En efecto, sea M un subconjunto arbitrario de la suma ordenada $M_1+M_2+\ldots+M_n$ de conjuntos bien ordenados; consideremos el primero de los conjuntos M_k que contiente elementos de M. La parte del conjunto M que figura en M_k representa un subconjunto del conjunto bien ordenado M_k y, por lo tanto, tiene el primer elemento. Este será también el primer elemento de M.

COROLARIO. *La suma ordenada de números ordinales es un número ordinal.*

Podemos, entonces, a partir de ciertos números ordinales construir números ordinales nuevos. Por ejemplo, partiendo de los números naturales (es decir, de los números ordinales finitos) y del número ordinal ω, se pueden obtener los números ordinales

$$\omega+n,\ \omega+\omega,\ \omega+\omega+n,\ \omega+\omega+\omega, \text{ etc.}$$

El lector podrá construir sin dificultad los conjuntos bien ordenados correspondientes a estos transfinitos ordinales.

Además de la suma ordenada de tipos ordinales se puede introducir el *producto ordenado*. Sean M_1 y M_2 dos conjuntos ordenados según los tipos θ_1 y θ_2. Tomemos varios ejemplares del conjunto M_1, uno por cada elemento de M_2, y sustituyamos los elementos de M_2 por estos ejemplares de M_1. El conjunto obtenido de este modo se llama producto ordenado de M_1 y M_2 y se denota mediante $M_1 \cdot M_2$. Desde el punto de vista formal el conjunto $M_1 \cdot M_2$ se compone de los pares (a, b), donde $a \in M_1$ y $b \in M_2$, aceptándose que $(a_1, b_1) < (a_2, b_2)$ siempre que $b_1 < b_2$ (cualesquiera que sean a_1, a_2) y $(a_1, b) < (a_2, b)$ si $a_1 < a_2$.

De un modo análogo se define el producto ordenado de un número finito arbitrario de factores $M_1 \cdot M_2 \cdot \ldots \cdot M_p$. El tipo ordinal θ del producto $M_1 \cdot M_2$ de conjuntos ordenados se llama *producto de los tipos ordinales* θ_1 y θ_2:

$$\theta=\theta_1 \cdot \theta_2.$$

El producto ordenado, al igual que la suma ordenada, no es conmutativo.

LEMA 2. *El producto de conjuntos bien ordenados es un conjunto bien ordenado.*

Sea M un subconjunto cualquiera del producto $M_1 \cdot M_2$; el conjunto M está compuesto por pares (a, b). Consideremos los segundos elementos b de todos los pares que figuran en M. Ellos forman un subconjunto de M_2. Puesto que M_2 está bien ordenado, este subconjunto tiene el primer elemento. Denotémoslo mediante b_1 y consideremos todos los pares de M del tipo (a, b_1). Los primeros elementos a de estos pares forman un subconjunto en M_1. Puesto que M_1 está bien ordenado, entre ellos existe el primer elemento. Sea éste a_1. Es fácil ver que el par (a_1, b_1) será entonces el primer elemento de M.

COROLARIO. *El producto ordenado de números ordinales es un número ordinal.*

Ejemplos. Se ve fácilmente que $\omega+\omega=\omega\cdot 2$, $\omega+\omega+\omega=\omega\cdot 3$. Es fácil también construir los conjuntos ordenados según los tipos $\omega\cdot n$, ω^2, $\omega^2\cdot n$, ω^3, ..., ω^p, ... Todos estos conjuntos tienen potencia numerable.

Se puede introducir, así mismo, otras operaciones con los tipos ordinales como, por ejemplo, la potencia y considerar entonces números ordinales como, por ejemplo, ω^ω, ω^{ω^ω}, etc.

6°. Comparación de números ordinales. Si n_1 y n_2 son dos números ordinales finitos, o bien coinciden o bien uno es mayor que otro. Veamos cómo se puede extender esta relación de orden a los números ordinales transfinitos.

Con este fin introduciremos los siguientes conceptos. Todo elemento a de un conjunto bien ordenado M determina el *segmento inicial* $P=$(la totalidad de elementos $<a$) y el *resto* $Q=$ $=$(la totalidad de elementos $\geqslant a$).

Sean α y β dos números ordinales y M y N dos conjuntos de tipo α y β, respectivamente. Diremos que $\alpha=\beta$ si los conjuntos M y N son isomorfos, que $\alpha<\beta$ si M es isomorfo a algún segmento inicial de N y que $\alpha>\beta$ si, al contrario, N es isomorfo a un segmento inicial de M.

TEOREMA. *Dos números ordinales α y β arbitrarios verifican una, y sólo una, de las relaciones:*

$$\alpha=\beta,\ \alpha<\beta \text{ o bien } \alpha>\beta.$$

Para demostras esta proposición estableceremos, ante todo, el siguiente lema.

LEMA. *Si f es una aplicación isomorfa de un conjunto bien ordenado A sobre algún subconjunto suyo B, entonces $f(a)\geqslant a$ para todo $a\in A$.*

En efecto, si existen elementos $a\in A$ tales que $f(a)<a$, entre ellos podremos encontrar el primero (¡tenemos buen orden!). Sea éste a_0 y sea $b_0=f(a_0)$. Entonces, $b_0<a_0$ y, por ser f un isomorfismo, $f(b_0)<f(a_0)=b$, es decir, a_0 no sería el primero de los elementos con esta propiedad.

De este lema se deduce inmediatamente que un conjunto bien ordenado no puede ser isomorfo a un segmento suyo, ya que si A fuese isomorfo al segmento determinado por el elemento a, tendríamos que $f(a) < a$. En otras palabras, las relaciones

$$\alpha = \beta \text{ y } \alpha < \beta$$

no pueden tener lugar simultáneamente. Del mismo modo no puede ser a la vez $\alpha = \beta$ y $\alpha > \beta$. Por otro lado, tampoco pueden verificarse simultáneamente las relaciones

$$\alpha < \beta \text{ y } \alpha > \beta,$$

ya que si esto fuese así, tendríamos (¡por la propiedad transitiva!) que $\alpha < \alpha$ lo que, como hemos visto, es imposible. Hemos demostrado de esta forma que si se verifica una de las relaciones $\alpha \gtreqless \beta$ las otras dos no se cumplen. Probemos ahora que una de estas relaciones siempre tiene lugar, es decir, que cualesquiera dos números ordinales son comparables.

Consideremos para cada número ordinal α el conjunto $W(\alpha)$ de los números ordinales $< \alpha$. Los números que figuran en $W(\alpha)$ son comparables y el propio conjunto $W(\alpha)$ (ordenado según la magnitud de los números ordinales) es de tipo α. En efecto, si el conjunto

$$A = \{\ldots, a, \ldots, b, \ldots\}$$

es de tipo α, los números ordinales inferiores a α corresponden biunívocamente, por definición, a los segmentos iniciales del conjunto A y, por consiguiente, a los elementos de este conjunto. En otras palabras, los elementos de un conjunto de tipo α pueden ser numerados mediante los números ordinales inferiores a α:

$$A = (a_0, a_1, \ldots, a_p, \ldots).$$

Sean ahora α y β dos números ordinales; entonces, $A = W(\alpha)$ y $B = W(\beta)$ son conjuntos de tipo α y β, respectivamente. Sea, además, $C = A \cap B$ la intersección de los conjuntos A y B, es decir, la totalidad de números ordinales inferiores simultáneamente a α y β. El conjunto C está bien ordenado; sea γ su tipo. Demostremos que $\gamma \leqslant \alpha$. En efecto, si $C = A$, entonces $\gamma = \alpha$; si $C \neq A$ tendremos que C es un segmento del conjunto A y por eso

$$\gamma < \alpha.$$

Efectivamente, para todos los $\xi \in C$, $\eta \in A \diagdown C$ los números ξ y η son comparables, es decir, $\xi \gtrless \eta$. Pero es imposible que sea $\eta < \xi < \alpha$, ya que en este caso tendríamos $\eta \in C$. Por lo tanto,

$\xi < \eta$, lo cual significa que C es un segmento del conjunto A y que $\gamma < \alpha$. Además, γ es el primer elemento del conjunto $A \setminus C$. De manera que

$$\gamma \leqslant \alpha \text{ y, análogamente, } \gamma \leqslant \beta.$$

Ahora bien, no puede ocurrir que $\gamma < \alpha$ y $\gamma < \beta$, ya que en este caso tendríamos que

$$\gamma \in A \setminus C, \quad \gamma \in B \setminus C,$$

es decir, por una parte, $\gamma \overline{\in} C$ y, por otra, $\gamma \in A \cap B = C$. Por consiguiente, sólo pueden darse los siguientes casos

$$\gamma = \alpha, \quad \gamma = \beta, \quad \alpha = \beta,$$
$$\gamma = \alpha, \quad \gamma < \beta, \quad \alpha < \beta,$$
$$\gamma < \alpha, \quad \gamma = \beta, \quad \alpha > \beta,$$

es decir, α y β son comparables. El teorema queda demostrado.

Puesto que a cada número ordinal le corresponde una potencia determinada y puesto que la comparabilidad de los números ordinales implica, obviamente, la comparabilidad de las potencias correspondientes, llegamos al siguiente resultado:

Si A y B son dos conjuntos bien ordenados, entonces, o bien son equivalentes (tienen la misma potencia) o bien la potencia de uno es mayor que la potencia del otro (en otras palabras, los conjuntos bien ordenados no pueden tener potencias incomparables).

Consideremos la totalidad de todos los números ordinales correspondientes a la potencia finita o numerable. Ellos forman un conjunto bien ordenado. Es fácil ver que este conjunto es, en sí, no numerable. En efecto denotemos, según se acostumbra comúnmente, mediante ω_1 el tipo ordinal del conjunto de todos los transfinitos numerables. Si la potencia correspondiente a este tipo fuese numerable, también sería numerable el conjunto de tipo ordinal $\omega_1 + 1$. Pero es obvio que el número ω_1 sigue a *todos* los transfinitos correspondientes a la potencia finita o numerable.

Designemos mediante $\aleph_1$ la potencia correspondiente al transfinito ordinal ω_1. Es fácil ver que no puede existir ni una potencia m que verifique la desigualdad

$$\aleph_0 < m < \aleph_1.$$

Verdaderamente, si hubiese tal m, en el conjunto $W(\omega_1)$ de todos los transfinitos ordinales precedentes a ω_1 tendría que existir un subconjunto de potencia m. Este subconjunto estaría bien ordenado y sería no numerable. Pero entonces su tipo ordinal α tendría que preceder a ω_1 y, al mismo tiempo, seguir a todos los transfinitos numerables. Esto contradecería a la definición de ω_1.

7°. Axioma de elección, teorema de Zermelo y otras proposiciones equivalentes a ellos. El hecho de que las potencias de dos cualesquiera conjuntos bien ordenados son obligatoriamente comparables sugiere el siguiente planteamiento: ¿es posible establecer de alguna manera un buen orden en un conjunto cualquiera? Esto permitiría afirmar que no existen potencias incomparables. Una respuesta positiva a este problema fue dada por Zermelo quien demostró que *todo conjunto puede ser bien ordenado*. Con este teorema está vinculada una serie de problemas profundos y de principio. Esto se debe a que su demostración (que no la daremos aquí) se basa, de un modo esencial, en el llamado *axioma de elección* que consiste en lo siguiente:

Si tenemos un conjunto cualquiera M, existe una función φ que pone en correspondencia a todo subconjunto no vacío $A \subset M$ un elemento determinado $\varphi(A)$ de este subconjunto.

Uno de los problemas más complejos y discutidos que surgen al fundamentar la teoría de los conjuntos es la legalidad de aplicar este axioma a conjuntos arbitrarios en los razonamientos que se hacen. No tenemos posibilidad de entrar aquí en su discusión. Observaremos, sin embargo, que la renuncia al axioma de elección restringe, de manera muy esencial, la posibilidad de diferentes construcciones en la teoría de los conjuntos.

Enunciemos algunas proposiciones, cada una de ellas equivalente al axioma de elección (es decir, cualquiera de ellas puede ser demostrada si se acepta el axioma de elección y, viceversa, el axioma de elección puede ser demostrado si se toma por verdadera alguna de estas proposiciones). Es obvio, ante todo, que una de estas proposiciones es el propio teorema de Zermelo. En efecto, si suponemos que el conjunto M está bien ordenado, es suficiente para construir la función $\varphi(A)$, cuya existencia se afirma en el axioma de elección, tomar el primer elemento en cada subconjunto $A \subset M$.

Antes de enunciar las demás proposiciones equivalentes al axioma de elección introduciremos los siguientes conceptos. Sea M un conjunto parcialmente ordenado. Todo subconjunto suyo A cuyos dos elementos cualesquiera son comparables (en el sentido del orden parcial existente en M) se llama *cadena*. Una cadena se llama *maximal* si no forma parte propia de ninguna otra cadena perteneciente a M. Finalmente, diremos que el elemento a del conjunto parcialmente ordenado M es una cota superior del subconjunto $M' \subset M$ si cualquier elemento $a' \in M'$ es menor que a.

TEOREMA DE HAUSDORFF. *Toda cadena de un conjunto parcialmente ordenado está contenida en alguna de las cadenas maximales de este conjunto.*

El siguiente teorema es, posiblemente, la forma más conveniente de las proposiciones equivalentes al axioma de elección.

TEOREMA DE ZORN. *Si toda cadena de un conjunto parcialmente ordenado M admite una cota superior, todo elemento de M precede a un elemento maximal.*

No reproduciremos aquí la demostración de la equivalencia de todas estas proposiciones (el axioma de elección, el teorema de Zermelo, el teorema de Hausdorff y el teorema de Zorn).

Si el conjunto de las cotas superiores del subconjunto A tiene elemento mínimo a, se dice que a es la *cota superior mínima* del subconjunto A; de un modo análogo se define la *cota inferior máxima*. Un conjunto parcialmente ordenado se llama *retículo* o *estructura* si todo subconjunto *finito* no vacío de él admite cota superior mínima y cota inferior máxima.

8°. Inducción transfinita. Un método ampliamente extendido de demostración de unas u otras proposiciones es el método de inducción matemática. Como se sabe, consiste en lo siguiente. Supongamos que se tiene una proposición $P(n)$ que se enuncia para cada número natural n y supongamos que:

1) La proposición $P(1)$ es cierta.

2) Bajo la hipótesis de la validez de $P(k)$ para todo $k \leqslant n$, se deduce que $P(n+1)$ es cierta.

Entonces, la proposición $P(n)$ es válida para todos los $n = 1, 2, \ldots, n, \ldots$. En efecto, en el caso contrario podríamos encontrar entre aquellos n para los cuales $P(n)$ no es cierta el número mínimo, digamos, n_1. Es obvio que $n_1 > 1$, es decir, $n_1 - 1$ es también un número natural, y esto contradice a la condición 2).

Se puede emplear un método análogo sustituyendo la serie de los números naturales por un conjunto bien ordenado cualquiera. En este caso de habla de la *inducción transfinita.* De manera que el método de inducción transfinita consiste en lo siguiente. Sea A un conjunto bien ordenado (se puede considerar, si se quiere, que es el conjunto de todos los transfinitos ordinales inferiores a uno dado) y sea $P(a)$ una proposición que se enuncia para todo $a \in A$ y tal que $P(a)$ es cierta para el primer elemento de A y es cierta para a si es válida para todos los elementos precedentes al elemento a. Entonces, $P(a)$ es cierta para todo $a \in A$. Efectivamente, si en A existiesen elementos para los cuales $P(a)$ no se cumple, podríamos encontrar en el conjunto formado por ellos el primer elemento, digamos a^*, y obtendríamos una contradicción ya que para todos los elementos $a < a^*$ la proposición $P(a)$ sería cierta.

La inducción transfinita puede, en principio, aplicarse a un conjunto cualquiera, ya que éste puede ser bien ordenado de acuerdo con el teorema de Zermelo. Sin embargo, en la práctica conviene más, casi siempre, sustituirlo por el teorema de Zorn, donde sólo se necesita la existencia de un orden parcial en el conjunto considerado.

§ 5. SISTEMAS DE CONJUNTOS [1]

1°. Anillo de conjuntos. Se llama *sistema de conjuntos* a todo conjunto cuyos elementos son, en sí, ciertos conjuntos. Si no se especifica lo contrario, estudiaremos sistemas cuyos elementos son cada uno un subconjunto de cierto conjunto fijado X. Denotaremos los sistemas de conjuntos con letras góticas mayúsculas. Serán de interés principal para nosotros aquellos sistemas de conjuntos que resultan cerrados respecto a las operaciones introducidas en el § 1.

DEFINICION 1. Un sistema no vacío de conjuntos $\mathfrak{R}$ se llama *anillo* si de $A\in\mathfrak{R}$ y $B\in\mathfrak{R}$ se deduce $A\triangle B\in R$ y $A\cap B\in\mathfrak{R}$.

Puesto que para cualesquiera dos conjuntos A y B

$$A\cup B=(A\triangle B)\triangle(A\cap B)$$

y

$$A\setminus B=A\triangle(A\cap B),$$

resulta que si $A\in\mathfrak{R}$ y $B\in\mathfrak{R}$, también pertenecen a $\mathfrak{R}$ los conjuntos $A\cup B$ y $A\setminus B$. Consecuentemente, un anillo de conjuntos es un sistema de conjuntos invariante respecto a las operaciones de unión e intersección y de resta y diferencia simétrica. Es obvio que todo anillo es, además, invariante respecto a la operación de toda unión o intersección finitas, es decir, respecto a las operaciones del tipo

$$C=\bigcup_{k=1}^{n}A_k,\quad D=\bigcap_{k=1}^{n}A_k.$$

Todo anillo contiene el conjunto vacío $\varnothing$ ya que siempre $A\setminus A=\varnothing$. El sistema que consta sólo del conjunto vacío representa el menor anillo de conjuntos posible.

Un conjunto E se llama *unidad* del sistema de conjuntos $\mathfrak{S}$ si pertenece a $\mathfrak{S}$ y si, además, para todo $A\in\mathfrak{S}$ se verifica la igualdad

---

[1] Los conceptos que se introducen en este párrafo serán empleados en el capítulo VI al exponer la teoría general de la medida. Por eso este párrafo puede ser estudiado más tarde. Aquellos lectores que piensan, al estudiar la teoría de la medida, limitarse sólo a la medida sobre el plano (§ 1 del capítulo VI) pueden omitirlo completamente.

$$A \cap E = A.$$

De manera que la unidad de un sistema de conjuntos $\mathfrak{S}$ no es otra cosa que el conjunto maximal de este sistema que contiene todos los demás conjuntos que figuran en $\mathfrak{S}$.

Un anillo de conjuntos provisto de la unidad se denomina *álgebra* de conjuntos.

Ejemplos. 1. Cualquiera que sea el conjunto A, el sistema $\mathfrak{M}(A)$ de todos subconjuntos suyos representa un álgebra de conjuntos siendo la unidad el conjunto $E = A$.

2. Cualquiera que sea el conjunto no vacío A, el sistema $\{\varnothing, A\}$ formado por el conjunto A y el conjunto vacío $\varnothing$ representa un álgebra de conjuntos siendo la unidad el conjunto $E = A$.

3. El sistema de todos los subconjuntos finitos de un conjunto arbitrario A es un anillo de conjuntos. Este anillo representará un álgebra si, y sólo si, el propio conjunto A es finito.

4. El sistema de todos los subconjuntos acotados de la recta númerica es un anillo de conjuntos sin unidad.

Directamente de la definición de un anillo de conjuntos se desprende el teorema siguiente:

TEOREMA 1. *La intersección* $\mathfrak{R} = \bigcup\limits_{\alpha} \mathfrak{R}_\alpha$ *de un conjunto cualquiera de anillos es también un anillo.*

Demostremos un resultado simple pero de importancia para lo sucesivo.

TEOREMA 2. *Cualquiera que sea el sistema no vacío de conjuntos* $\mathfrak{S}$ *existe un anillo y sólo uno,* $\mathfrak{R}(\mathfrak{S})$ *que contiene* $\mathfrak{S}$ *y está contenido en cualquier anillo* $\mathfrak{R}$ *que contiene* $\mathfrak{S}$.

DEMOSTRACION. Es fácil ver que el anillo $\mathfrak{R}(\mathfrak{S})$ se determina unívocamente por el sistema $\mathfrak{S}$. Para demostrar la existencia de este anillo, consideremos la unión $X = \bigcup\limits_{A \in \mathfrak{S}} A$ de todos los conjuntos A que figuran en $\mathfrak{S}$ y el anillo $\mathfrak{M}(X)$ de todos los subconjuntos del conjunto X. Sea $\sum$ la totalidad de los anillos de conjuntos que están contenidos en $\mathfrak{M}(X)$ y contienen $\mathfrak{S}$. La intersección

$$\mathfrak{P} = \bigcap_{\mathfrak{R} \in \Sigma} \mathfrak{R}$$

de todos estos anillos será precisamente el anillo $\mathfrak{R}(\mathfrak{S})$.

En efecto, cualquiera que sea el anillo $\mathfrak{R}^*$ conteniendo $\mathfrak{S}$, la intersección $\mathfrak{R} = \mathfrak{R}^* \cap \mathfrak{M}(X)$ será un anillo de $\sum$ y, por consiguiente,

$$\mathfrak{S} \subset \mathfrak{R} \subset \mathfrak{R}^*,$$

es decir, $\mathfrak{P}$ verifica realmente la condición de ser el anillo minimal. Este anillo se llama anillo minimal sobre el sistema $\mathfrak{S}$ y se denota $\mathfrak{R}(\mathfrak{S})$.

2°. **Semianillo de conjuntos.** En varios problemas, por ejemplo, en la teoría de medida, desempeña un papel importante, además del concepto de anillo, el concepto más general de semianillo de conjuntos.

DEFINICION 2. Un sistema de conjuntos $\mathfrak{S}$ se llama *semianillo* si contiene el conjunto vacío $\varnothing$, está cerrado respecto a la operacción de intersección y cumple la siguiente propiedad: si A y $A_1 \subset A$ pertenecen a $\mathfrak{S}$, se puede representar el conjunto A en la forma $A = \bigcup\limits_{k=1}^{n} A_k$, donde A_k son conjuntos de $\mathfrak{S}$ disjuntos dos a dos y el primero de ellos es el conjunto dado A_1.

Todo sistema de conjuntos disjuntos dos a dos $A_1, A_2, \ldots, A_n$ cuya unión es el conjunto dado A se llamará, en lo sucesivo, *descomposición finita* del conjunto A.

Todo anillo de conjuntos $\mathfrak{R}$ es un semianillo ya que si A y $A_1 \subset A$ figuran en $\mathfrak{R}$, tiene lugar la descomposición

$$A = A_1 \cup A_2, \text{ donde } A_2 = A \setminus A_1 \in \mathfrak{R}.$$

Un ejemplo de un semianillo que no es anillo de conjuntos obtendremos al tomar la totalidad de los intervalos (a, b), los segmentos $[a, b]$ y los semisegmentos $[a, b)$ y $(a, b]$ de la recta numérica[1]).

Veamos algunas propiedades de los semianillos de conjuntos.

LEMA 1. *Supongamos que los conjuntos $A_1, A_2, \ldots, A_n, A$ pertenecen al semianillo $\mathfrak{S}$ y que, además, los conjuntos A_i son subconjuntos de A disjuntos dos a dos. Entonces, los conjuntos A_i $(i = 1, 2, \ldots, n)$ se pueden tomar como los n primeros miembros en la descomposición finita*

$$A = \bigcup_{k=1}^{s} A_k, \quad s \geqslant n,$$

del conjunto A siendo todos los $A_k \in \mathfrak{S}$.

DEMOSTRACION. Emplearemos el método de inducción matemática. Para $n = 1$ la afirmación del lema es válida de acuerdo con la definición de semianillo. Supongamos que esta proposición es

---

[1]) En los intervalos se incluye, por supuesto, el intervalo « vacío » (a, a) y en los segmentos, el segmento compuesto por un solo punto $[a, a]$.

cierta para $n=m$ y consideremos $m+1$ conjuntos $A_1, \ldots, A_m, A_{m+1}$ que verifican las condiciones del lema. Por hipótesis:

$$A = A_1 \cup A_2 \cup \ldots \cup A_m \cup B_1 \cup B_2 \cup \ldots \cup B_p,$$

donde todos los conjuntos $B_q (q=1, 2, \ldots, p)$ pertenecen a $\mathfrak{S}$. Tomemos

$$B_{q1} = A_{m+1} \cap B_q.$$

De acuerdo con la definición de semianillo tiene lugar la descomposición

$$B_q = B_{q1} \cup B_{q2} \cup \ldots \cup B_{qr_q},$$

donde todos los B_{qj} pertenecen a $\mathfrak{S}$. Es fácil ver que

$$A = A_1 \cup \ldots \cup A_m \cup A_{m+1} \cup \bigcup_{q=1}^{p} \bigcup_{j=2}^{r_q} B_{qj}.$$

De modo que hemos demostrado la afirmación del lema para $n=m+1$ y, por consiguiente, para todo n.

LEMA 2. *Cualquiera que sea el sistema finito de conjuntos $A_1, \ldots, A_n$ pertenecientes al semianillo $\mathfrak{S}$, existe en $\mathfrak{S}$ un sistema finito de conjuntos $B_1, \ldots, B_t$, disjuntos dos a dos, tal que todo A_k se puede representar mediante la unión*

$$A_k = \bigcup_{s \in M_k} B_s$$

de algunos de los conjuntos B_s.

DEMOSTRACION. Para $n=1$ el lema resulta trivial, ya que es suficiente tomar, en este caso, $t=1$ y $B_1 = A_1$. Supongamos que el lema es cierto para $n=m$ y consideremos en $\mathfrak{S}$ un sistema de conjuntos $A_1, \ldots, A_m, A_{m+1}$ cualquiera. Sean $B_1, B_2, \ldots, B_t$ los conjuntos de $\mathfrak{S}$ que verifican las condiciones del lema respecto a $A_1, A_2, \ldots, A_m$. Tomemos

$$B_{s1} = A_{m+1} \cap B_s.$$

En virtud del lema 1 tiene lugar la descomposición

$$A_{m+1} = \bigcup_{s=1}^{t} B_{s1} \cup \bigcup_{p=1}^{q} B'_p, \quad B'_p \in \mathfrak{S}, \tag{1}$$

y de acuerdo con la definición propia de semianillo, la descomposición

$$B_s = B_{s1} \cup B_{s2} \cup \ldots \cup B_{sf_s}, \quad B_{sj} \in \mathfrak{S}.$$

fácil ver que

$$A_k = \bigcup_{s \in M_k} \bigcup_{j=1}^{f_s} B_{sj}, \quad k = 1, 2, \ldots, m$$

y que los conjuntos

$$B_{sj}, \; B'_p$$

son disjuntos dos a dos. Es decir, los conjuntos B_{sj}, B'_p verifican las condiciones del lema respecto a $A_1, \ldots, A_m, A_{m+1}$. El lema queda demostrado.

3°. Anillo engendrado por un semianillo. Hemos visto ya en el primer punto que para todo sistema de conjuntos $\mathfrak{S}$ existe un anillo minimal, y sólo uno, que contiene a $\mathfrak{S}$. Sin embargo, no es fácil construir de hecho a partir de $\mathfrak{S}$ el anillo $\mathfrak{R}(\mathfrak{S})$ en el caso de un sistema $\mathfrak{S}$ arbitrario. Esto resulta posible en el importante caso cuando $\mathfrak{S}$ representa un semianillo, como se ve del siguiente teorema.

TEOREMA 3. *Si $\mathfrak{S}$ es un semianillo, $\mathfrak{R}(\mathfrak{S})$ coincide con el sistema $\mathfrak{Z}$ de conjuntos A que admiten descomposiciones finitas*

$$A = \bigcup_{k=1}^{n} A_k$$

en conjuntos $A_k \in \mathfrak{S}$.

DEMOSTRACION. Comprobemos que el sistema $\mathfrak{Z}$ forma un anillo. Si A y B son conjuntos cualesquiera de $\mathfrak{Z}$, tienen las descomposiciones

$$A = \bigcup_{k=1}^{n} A_k, \quad B = \bigcup_{k=1}^{m} B_k, \quad A_k \in \mathfrak{S}, \; B_k \in \mathfrak{S}.$$

Puesto que $\mathfrak{S}$ es un semianillo, los conjuntos

$$C_{ij} = A_i \cup B_j$$

también pertenecen a $\mathfrak{S}$. En virtud del lema 1 existen las descomposiciones

$$A_i = \bigcup_{j} C_{ij} \cup \bigcup_{k=1}^{r_i} D_{ik}; \quad B_j = \bigcup_{i} C_{ij} \cup \bigcup_{l=1}^{s_j} E_{jl}, \tag{2}$$

donde $D_{ik}, E_{jl} \in \mathfrak{S}$. De (2) se desprende que los conjuntos $A \cap B$ y $A \triangle B$ admiten las descomposiciones siguientes:

$$A \cap B = \bigcup_{i,\, j} C_{ij}, \quad A \triangle B = \bigcup_{i,\, k} D_{ik} \cup \bigcup_{j,\, l} E_{jl},$$

es decir, pertenecen a $\mathfrak{Z}$. Por lo tanto, $\mathfrak{Z}$ es realmente un anilo; es obvio que entre los anillos que contienen $\mathfrak{S}$, éste es el anillo minimal.

4°. Algebras de Borel. En varios problemas, en la teoría de la medida, en particular, es preciso considerar la unión e intersección de una cantidad numerable, y no sólo finita, de conjuntos. Por eso conviene introducir los siguientes conceptos, además del de anillo de conjuntos.

DEFINICION 3. Un anillo de conjuntos se llama *σ-anillo* si junto con toda sucesión de conjuntos $A_1, A_2, \ldots, A_n, \ldots$ contiene la unión

$$S = \bigcup_n A_n.$$

DEFINICION 4. Un anillo de conjuntos se llama *δ-anillo* si junto con toda sucesión de conjuntos $A_1, A_2, \ldots, A_n, \ldots$ contiene la intersección

$$D = \bigcap_n A_n.$$

Es natural llamar *σ-álgebra* a todo σ-anillo con unidad y *δ-álgebra* a todo δ-anillo con unidad. Es fácil ver, sin embargo, que estos dos conceptos coinciden: toda σ-álgebra es al mismo tiempo una δ-álgebra y toda δ-álgebra, una σ-álgebra. Esto se deduce de las relaciones de dualidad:

$$\bigcup_n A_n = E \setminus \bigcup_n (E \setminus A_n),$$

$$\bigcap_n A_n = E \setminus \bigcup_n (E \setminus A_n)$$

(véase el § 1). Las δ-álgebras o, que es lo mismo, las σ-álgebras suelen llamarse *álgebras de Borel* o, simplemente, *B-álgebras*.

El ejemplo más sencillo de una B-álgebra es la totalidad de los subconjuntos de un conjunto A.

Si se tiene un sistema de conjuntos $\mathfrak{S}$, siempre existe al menos una B-álgebra que contiene este sistema. En efecto, pongamos

$$X = \bigcup_{A \in \mathfrak{S}} A$$

y consideremos el sistema $\mathfrak{B}$ de todos los subconjuntos del conjunto X. Está claro que $\mathfrak{B}$ es una B-álgebra que contiene $\mathfrak{S}$. Si $\tilde{\mathfrak{B}}$ es una B-álgebra arbitraria que contiene $\mathfrak{S}$

y $\tilde{X}$ es su unidad, todo $A \in \mathfrak{S}$ está contenido en $\tilde{X}$ y, por consiguiente, $X = \bigcup_{A \in \mathfrak{S}} A \subset \tilde{X}$. Una B-álgebra $\mathfrak{B}$ se llama *irreducible* (respecto al sistema $\mathfrak{S}$) cuando $\tilde{X} = \bigcup_{A \in \mathfrak{S}} A$. En otras palabras, una B-álgebra irreducible es una B-álgebra que no contiene puntos no pertenecientes a ninguno de los conjuntos $A \in \mathfrak{S}$. Es natural limitarse siempre a considerar sólo estas B-álgebras.

Para las B-álgebras irreducibles tiene lugar un teorema análogo al teorema 2 demostrado anteriormente para los anillos.

TEOREMA 4. *Cualquiera que sea el sistema de conjuntos no vacío $\mathfrak{S}$ existe una B-álgebra $\mathfrak{B}(\mathfrak{S})$ irreducible (respecto a este sistema) que contiene $\mathfrak{S}$ y está contenida en cualquier B-álgebra que contiene $\mathfrak{S}$.*

LA DEMOSTRACION se realiza exactamente del mismo modo que la demostración del teorema 2. La B-álgebra $\mathfrak{B}(\mathfrak{S})$ se llama *B-álgebra minimal sobre el sistema $\mathfrak{S}$* o *clausura boreliana del sistema $\mathfrak{S}$.*

En el Análisis desempeñan un papel importante los llamados *conjuntos de Borel* o *B-conjuntos*, es decir, los conjuntos de la recta numérica pertenecientes al B-álgebra minimal sobre la totalidad de los segmentos $[a, b]$.

5°. Sistemas de conjuntos y aplicaciones. Señalemos algunos resultados que nos harán falta al estudiar las funciones medibles.

Sea $y = f(x)$ una función definida sobre el conjunto M y que toma valores en el conjunto N y sea $\mathfrak{M}$ algún sistema de subconjuntos del conjunto M. Designemos mediante $f(\mathfrak{M})$ el sistema de todas las imágenes $f(A)$ de los conjuntos pertenecientes a $\mathfrak{M}$. Sea, además, $\mathfrak{N}$ algún sistema de conjuntos contenidos en N y $f^{-1}(\mathfrak{N})$ el sistema de todas las imágenes recíprocas $f^{-1}(A)$ de los conjuntos que figuran en $\mathfrak{N}$. Tienen lugar las siguientes proposiciones (dejamos a cargo del lector las demostraciones):

1) Si $\mathfrak{N}$ es un anillo, también lo es $f^{-1}(\mathfrak{N})$.
2) Si $\mathfrak{N}$ es un álgebra, también lo es $f^{-1}(\mathfrak{N})$.
3) Si $\mathfrak{N}$ es una B-álgebra, también lo es $f^{-1}(\mathfrak{N})$.
4) $\mathfrak{R}(f^{-1}, \mathfrak{N}) = f^{-1}(\mathfrak{R}(\mathfrak{N}))$.
5) $\mathfrak{B}(f^{-1}\mathfrak{N}) = f^{-1}(\mathfrak{B}(\mathfrak{N}))$.

¿Continuarán siendo ciertas estas proposiciones si sustituimos f^{-1} por f y $\mathfrak{N}$ por $\mathfrak{M}$?

Observaciones finales. La teoría de conjuntos surge como una rama de las Matemáticas en los trabajos del matemático alemán Georg Cantor. Las ideas de Cantor, recibidas primero con desconfianza, obtienen más tarde una amplia divulgación y en el siglo veinte el punto de vista de la teoría de conjuntos se convierte en base de las más diferentes ramas de las Matemáticas: conceptos tan fundamentales como grupo, anillo, cuerpo, espacio lineal, etc., se definen generalmente como conjuntos compuestos por elementos de una naturaleza arbitraria que verifican unos u otros axiomas complementarios. En el desarrollo sucesivo de la teoría de conjuntos surgieron varias dificultades lógicas y esto condujo, naturalmente, a tratar de sustituir la teoría de conjuntos «ingenua» por construcciones axiomáticas rigurosas. Aquí resultó que algunos problemas de la teoría de conjuntos que, al parecer, admiten una solución determinada de tipo «sí» o «no», tienen, en realidad, otra naturaleza. Uno de ellos es el famoso problema de continuo: ¿existen potencias no numerables inferiores al continuo? Basándose en cierta axiomática (en la discusión de la cual no podemos aquí entrar), Gödel demostró que la solución negativa de este problema no contradice a la teoría axiomática mencionada y, más tarde, Cohen demostró que su solución positiva no es contradictoria en el mismo sentido.

La teoría de conjuntos se expone, con mayor o menor detalle y generalidad, en varios libros de texto.

CAPITULO II

ESPACIOS METRICOS Y TOPOLOGICOS

§ 1. CONCEPTO DE ESPACIO MÉTRICO

1°. Definición y ejemplos principales. Una de las operaciones principales del Análisis es el paso al límite. Esta operación descansa sobre el hecho de que en la recta númerica está definida la distancia entre dos puntos. Muchos resultados principales del Análisis no tienen nada que ver con la naturaleza algebraica del conjunto de los números reales (es decir con el hecho de que forman un cuerpo) y sólo se apoyan en aquellas propiedades de los números reales que están relacionadas con el concepto de distancia. Generalizando la interpretación de los números reales como un conjunto en el que se ha definido la distancia entre sus elementos, llegamos al concepto de espacio métrico, uno de los conceptos más importantes de la matemática moderna. A continuación exponemos los resultados fundamentales de la teoría de los espacios métricos y de sus generalizaciones, los espacios topológicos. Los resultados de este capítulo son esenciales para toda la exposición ulterior.

DEFINICION. Un *espacio métrico* es un par (X, ρ), compuesto de un conjunto (espacio) X de elementos (puntos) y de una distancia, es decir, una función unívoca real y no negativa $\rho(x, y)$ definida para dos cualesquiera elementos x e y de X y que verifica las tres condiciones siguientes:

1) $\rho(x, y)=0$ si, y sólo si, $x=y$,
2) (axioma de simetría): $\rho(x, y)=\rho(y, x)$,
3) (axioma triangular): $\rho(x, z)\leqslant\rho(x, y)+\rho(y, z)$.

El propio espacio métrico, es decir, el par (x, ρ), lo denotaremos, como regla general, mediante una letra:

$$R=(X, \rho).$$

En los casos en que la equivocación esté excluida denotaremos frecuentemente el espacio métrico con el mismo símbolo X que describe la «reserva de puntos».

Señalemos ejemplos de espacios métricos. Algunos de estos espacios desempeñan un papel muy importante en el Análisis.

1. Tomando para los elementos de un conjunto arbitrario

$$\rho(x, y) = \begin{cases} 0, & \text{si } x = y, \\ 1, & \text{si } x \neq y, \end{cases}$$

obtendremos, evidentemente, un espacio métrico que puede ser denominado espacio de puntos aislados.

2. El conjunto de los números reales con la distancia

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

forma el espacio métrico R^1.

3. El conjunto de grupos ordenados de n números reales

$$x = (x_1, x_2, \ldots, x_n)$$

con la distancia

$$\rho(x, y) = \sqrt{\sum_{k=1}^{n} (y_k - x_k)^2} \tag{1}$$

se denomina *espacio aritmético euclídeo de n dimensiones* R^n. Es evidente que en R^n se verifican los axiomas 1 y 2. Demostremos que en R^n también se cumple el axioma triangular.

Sean $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ y $z = (z_1, \ldots, z_n)$; entonces el axioma triangular se puede escribir en la forma:

$$\sqrt{\sum_{k=1}^{n} (z_k - x_k)^2} \leqslant \sqrt{\sum_{k=1}^{n} (y_k - x_k)^2} + \sqrt{\sum_{k=1}^{n} (z_k - y_k)^2}. \tag{2}$$

Si tomamos $y_k - x_k = a_k$ y $z_k - y_k = b_k$, tendremos $z_k - x_k = a_k + b_k$ y la desigualdad (2) obtiene la forma:

$$\sqrt{\sum_{k=1}^{n} (a_k + b_k)^2} \leqslant \sqrt{\sum_{k=1}^{n} a_k^2} + \sqrt{\sum_{k=1}^{n} b_k^2}. \tag{3}$$

Pero esta desigualdad se deduce inmediatamente de la conocida desigualdad de Cauchy—Buniakovski[1]:

$$\left(\sum_{k=1}^{n} a_k b_k\right)^2 \leqslant \sum_{k=1}^{n} a_k^2 \cdot \sum_{k=1}^{n} b_k^2. \tag{4}$$

En efecto, de acuerdo con esta desigualdad tenemos

$$\sum_{k=1}^{n} (a_k+b_k)^2 = \sum_{k=1}^{n} a_k^2 + 2\sum_{k=1}^{n} a_k b_k + \sum_{k=1}^{n} b_k^2 \leqslant \sum_{k=1}^{n} a_k^2 + \\ + 2\sqrt{\sum_{k=1}^{n} a_k^2 \cdot \sum_{k=1}^{n} b_k^2} + \sum_{k=1}^{n} b_k^2 = \left(\sqrt{\sum_{k=1}^{n} a_k^2} + \sqrt{\sum_{k=1}^{n} b_k^2}\right)^2;$$

con esto queda demostrada la desigualdad (3) y, por consiguiente, la desigualdad (2).

4. Consideremos el mismo conjunto de grupos ordenados de n números reales $x=(x_1, \ldots, x_n)$, pero definiendo la distancia en él mediante la fórmula

$$\rho_1(x, y) = \sum_{k=1}^{n} |x_k - y_k|. \tag{5}$$

Se ve inmediatamente que se cumplen los axiomas 1, 2 y 3. Denotaremos este espacio métrico con el símbolo R_1^n.

5. Tomemos de nuevo el conjunto de los ejemplos 3 y 4 definiendo la distancia entre sus elementos mediante la fórmula

$$\rho_0(x, y) = \max_{1 \leqslant k \leqslant n} |y_k - z_k|. \tag{6}$$

Es evidente que se cumplen los axiomas 1, 2 y 3. Para muchas cuestiones del Análisis este espacio, que denotaremos R_0^n, es no menos cómodo que el espacio euclídeo R^n.

Los tres últimos ejemplos muestran que a veces es importante tener, en efecto, diferentes denotaciones para el espacio métrico y para el conjunto de sus puntos, ya que una misma reserva de puntos puede ser metrizada de diferentes maneras.

6. El conjunto $C_{[a, b]}$ de todas las funciones reales continuas definidas en el segmento $[a, b]$ con la distancia

$$\rho(f, g) = \max_{a \leqslant t \leqslant b} |g(t) - f(t)| \tag{7}$$

---

[1] La desigualdad de Cauchy—Buniakovski se sigue de la identidad

$$\left(\sum_{k=1}^{n} a_k b_k\right)^2 = \sum_{k=1}^{n} a_k^2 \sum_{k=1}^{n} b_k^2 - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} (a_i b_j - b_i a_j)^2,$$

que se puede comprobar directamente.

también forma un espacio métrico. Los axiomas 1, 2 y 3 se comprueban directamente. Este espacio desempeña un papel muy importante en el Análisis. Lo denotaremos con el mismo símbolo $C_{[a, b]}$ que empleamos para el conjunto de los puntos de este espacio. En lugar de $C_{[0, 1]}$ escribiremos simplemente C.

7. Designemos mediante l_2 el espacio métrico cuyos puntos son todas las sucesiones

$$x = (x_1, x_2, \ldots, x_n, \ldots)$$

de números reales que verifican la condición

$$\sum_{k=1}^{n} x_k^2 < \infty$$

y en el que la distancia viene dada por

$$\rho(x, y) = \sqrt{\sum_{k=1}^{\infty} (y_k - x_k)^2}. \tag{8}$$

La función $\rho(x, y)$ así definida tiene sentido para todos $x, y \in l_2$, ya que la serie $\sum\limits_{k=1}^{\infty} (y_k - x_k)^2$ converge siempre que $\sum\limits_{k=1}^{\infty} x_k^2 < \infty$ y $\sum\limits_{k=1}^{\infty} y_k^2 < \infty$; esto se desprende de la siguiente desigualdad elemental

$$(x_k \pm y_k)^2 \leqslant 2(x_k^2 + y_k^2).$$

Al mismo tiempo vemos que si $(x_1, x_2, \ldots, x_n, \ldots)$ e $(y_1, y_2, \ldots, y_n, \ldots)$ pertenecen a l_2, también $(x_1 + y_1, \ldots, x_n + y_n, \ldots) \in l_2$. Comprobemos ahora que la función (8) verifica los axiomas de espacio métrico. Los axiomas 1 y 2 son evidentes y el axioma 3 adquiere en este caso la forma

$$\sqrt{\sum_{k=1}^{\infty} (z_k - x_k)^2} \leqslant \sqrt{\sum_{k=1}^{\infty} (z_k - y_k)^2} + \sqrt{\sum_{k=1}^{\infty} (y_k - x_k)^2} \tag{9}$$

De acuerdo con lo dicho anteriormente converge cada una de las tres series que aquí figuran. Por otro lado para todo n se verifica la desigualdad

$$\sqrt{\sum_{k=1}^{n} (z_k - x_k)^2} \leqslant \sqrt{\sum_{k=1}^{n} (z_k - y_k)^2} + \sqrt{\sum_{k=1}^{n} (y_k - x_k)^2}$$

(véase el ejemplo 4). Pasando al límite para $n \to \infty$ obtenemos (9), es decir, la desigualdad triangular en l_2.

8. Consideremos, al igual que en el ejemplo 6, el conjunto de todas las funciones continuas en el segmento $[a, b]$, pero definiendo la distancia de otro modo, a saber

$$\rho(x, y)=\left(\int_a^b (x(t)-y(t))^2\,dt\right)^{1/2}. \tag{10}$$

Este espacio métrico lo denotaremos $C^2_{[a,\,b]}$ y lo llamaremos *espacio de funciones continuas con métrica cuadrática*. Los axiomas 1 y 2 del espacio métrico son otra vez evidentes, mientras que el axioma triangular se deduce directamente de la desigualdad de Cauchy—Buniakovski en su forma integral[1]

$$\left(\int_a^b x(t)\,y(t)\,dt\right)^2 \leqslant \int_a^b x^2(t)\,dt\cdot\int_a^b y^2(t)\,dt.$$

9. Consideremos el conjunto de todas las sucesiones acotadas $x=(x_1, x_2, \ldots, x_n, \ldots)$ de números reales. Admitiendo

$$\rho(x, y)=\sup_k |y_k-x_k|, \tag{11}$$

obtenemos un espacio métrico que denotaremos m. Los axiomas 1, 2 y 3 se verifican evidentemente.

10. El conjunto de todos los grupos ordenados de n números reales con la distancia

$$\rho_p(x, y)=\left(\sum_{k=1}^{n} |y_k-x_k|^p\right)^{1/p}, \tag{12}$$

donde p es un número fijo arbitrario $\geqslant 1$, representa un espacio métrico que denotaremos $R_p^{(n)}$. También en este caso es obvio que se verifican los axiomas 1 y 2. Comprobemos el axioma 3. Sean $x=(x_1, x_2, \ldots, x_n)$, $y=(y_1, y_2, \ldots, y_n)$ y $z=(z_1, z_2, \ldots, z_n)$ tres puntos de $R_p^{(n)}$. Pongamos

$$y_k-x_k=a_k, \quad z_k-y_k=b_k.$$

La desigualdad

$$\rho_p(x, z)\leqslant \rho_p(x, y)+\rho_p(y, z)$$

---

[1] Esta desigualdad puede obtenerse, por ejemplo, de la siguiente identidad que se comprueba fácilmente

$$\left(\int_a^b x(t)\,y(t)\,dt\right)^2=\int_a^b x^2(t)\,dt\int_a^b y^2(t)\,dt-\frac{1}{2}\int_a^b\int_a^b [x(s)\,y(t)-y(s)\,x(t)]^2\,ds\,dt.$$

que debemos demostrar puede representarse entonces en la forma

$$\left(\sum_{k=1}^{n} |a_k+b_k|^p\right)^{1/p} \leqslant \left(\sum_{k=1}^{n} |a_k|^p\right)^{1/p} + \left(\sum_{k=1}^{n} |b_k|^p\right)^{1/p} \tag{13}$$

Esta es la *desigualdad de Minkowski.* Para $p=1$ la desigualdad de Minkowski es obvia (el valor absoluto de la suma no sobrepasa la suma de los valores absolutos) y, por consiguiente, podemos limitarnos a considerar el caso $p>1$ [1].

La demostración de la desigualdad (13) para $p>1$ se basa en la *desigualdad de Hölder:*

$$\sum_{k=1}^{n} |a_k b_k| \leqslant \left(\sum_{k=1}^{n} |a_k|^p\right)^{1/p} \left(\sum_{k=1}^{n} |b_k|^q\right)^{1/q}, \tag{14}$$

donde los números $p>1$ y $q>1$ cumplen la condición

$$\frac{1}{p}+\frac{1}{q}=1. \tag{15}$$

Observemos que la desigualdad (14) es homogénea. Ello significa que si se cumple para dos vectores cualesquiera

$$a=(a_1, a_2, \ldots, a_n) \text{ y } b=(b_1, b_2, \ldots, b_n)$$

también se verifica para los vectores λa y μb, donde λ y μ son números arbitrarios. Por eso es suficiente demostrar la desigualdad (14) para el caso de que

$$\sum_{k=1}^{n} |a_k|^p = \sum_{k=1}^{n} |b_k|^q = 1. \tag{16}$$

Supongamos, pues, que se cumple la condición (16); demostremos que

$$\sum_{k=1}^{n} |a_k b_k| \leqslant 1. \tag{17}$$

Consideremos en el plano (ξ, η) la curva dada por la ecuación $\eta=\xi^{p-1}$ $(\xi>0)$, o, que es lo mismo, por la ecuación $\xi=\eta^{q-1}$ (fig. 8). Del dibujo se ve claramente que para cualesquiera valores positivos a y b será $S_1+S_2 \geqslant ab$. Calculemos las áreas S_1 y S_2:

$$S_1=\int_0^a \xi^{p-1}\,d\xi=\frac{a^p}{p}, \quad S_2=\int_0^b \eta^{q-1}\,d\eta=\frac{b^q}{q}.$$

---

[1] Para $p<1$ la desigualdad de Minkowski no tiene lugar. En otras palabras, si pretendiéramos considerar el espacio $R_p^{(n)}$ para $p<1$, en este espacio no se cumpliría el axioma triangular.

Por lo tanto, se verifica la siguiente desigualdad numérica

$$ab \leqslant \frac{a^p}{p} + \frac{b^q}{q}.$$

Poniendo $a = |a_k|$, $b = |b_k|$ y sumando respecto a k desde 1 hasta n, obtendremos tomando en cuenta (15) y (16)

$$\sum_{k=1}^{n} |a_k b_k| \leqslant 1.$$

Hemos demostrado la desigualdad (17) y, por consiguiente, la desigualdad general (14). Para $p = 2$ la desigualdad de Hölder (14) se convierte en la desigualdad de Cauchy—Buniakovski (4).

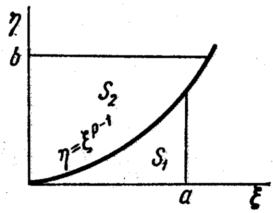


FIG. 8

Pasemos a demostrar ahora la desigualdad de Minkowski. Para ello consideremos la identidad

$$(|a| + |b|)^p = (|a| + |b|)^{p-1} |a| + (|a| + |b|)^{p-1} |b|.$$

Tomando en la identidad escrita $a = a_k$, $b = b_k$ y sumando respecto a k desde 1 hasta n, obtenemos

$$\sum_{k=1}^{n} (|a_k| + |b_k|)^p =$$
$$= \sum_{k=1}^{n} (|a_k| + |b_k|)^{p-1} |a_k| + \sum_{k=1}^{n} (|a_k| + |b_k|)^{p-1} |b_k|.$$

Aplicando ahora a cada una de las sumas que figuran a la derecha la desigualdad de Hölder y tomando en consideración que $(p-1)q = p$, encontramos

$$\sum_{k=1}^{n} (|a_k| + |b_k|)^p \leqslant$$
$$\leqslant \left(\sum_{k=1}^{n} (|a_k| + |b_k|)^p \right)^{1/q} \left(\left[\sum_{k=1}^{n} |a_k|^p \right]^{1/p} + \left[\sum_{k=1}^{n} |b_k|^p \right]^{1/p} \right).$$

Dividiendo ambas partes de esta desigualdad por

$$\left(\sum_{k=1}^{n}(|a_k|+|b_k|)^p\right)^{1/q},$$

obtenemos

$$\left(\sum_{k=1}^{n}(|a_k|+|b_k|)^p\right)^{1/p}\leqslant\left(\sum_{k=1}^{n}|a_k|^p\right)^{1/p}+\left(\sum_{k=1}^{n}|b_k|^p\right)^{1/p}$$

y de aquí se deduce inmediatamente la desigualdad (13). Con esto queda comprobado el axioma triangular para el espasio $R_p^{(n)}$.

La métrica ρ_p considerada en este ejemplo coincide para $p=2$ con la métrica euclídea (ejemplo 3) y para $p=1$ con la métrica del ejemplo 4. Se puede demostrar que la métrica

$$\rho_0(x, y)=\max_{1\leqslant k\leqslant n}|y_k-x_k|,$$

introducida en el ejemplo 5, es el caso límite de la métrica $\rho_p(x, y)$, es decir,

$$\rho_0(x, y)=\lim_{p\to\infty}\left(\sum_{k=1}^{n}|y_k-x_k|^p\right)^{1/p}.$$

De la desigualdad

$$ab\leqslant\frac{a^p}{p}+\frac{b^q}{q}\left(\frac{1}{p}+\frac{1}{q}=1\right),$$

establecida anteriormente, es fácil deducir la *desigualdad integral de Hölder*

$$\int_a^b x(t)\,y(t)\,dt\leqslant\left(\int_a^b|x(t)|^p dt\right)^{1/p}\left(\int_a^b|y(t)|^q dt\right)^{1/q}$$

que se verifica para cualesquiera funciones $x(t)$ e $y(t)$ siempre que las integrales que figuran a la derecha tengan sentido. A su vez, de aquí se obtiene la *desigualdad integral de Minkowski*

$$\left(\int_a^b|x(t)+y(t)|^p dt\right)^{1/p}\leqslant\left(\int_a^b|x(t)|^p dt\right)^{1/p}+\left(\int_a^b|y(t)|^p dt\right)^{1/p}.$$

11. Veamos otro ejemplo interesante de espacio métrico. Sus elementos son todas las sucesiones de números reales

$$x=(x_1,\ x_2,\ \ldots,\ x_n,\ \ldots)$$

tales que

$$\sum_{k=1}^{\infty}|x_k|^p<\infty,$$

donde $p \geqslant 1$ es un número fijo, y la distancia viene dada por la fórmula

$$\rho(x, y) = \left(\sum_{k=1}^{\infty} |y_k - x_k|^p \right)^{1/p}. \tag{18}$$

Denotaremos este espacio métrico mediante l_p.

De acuerdo con la desigualdad de Minkowski (13) tenemos para n cualquiera

$$\left(\sum_{k=1}^{n} |y_k - x_k|^p \right)^{1/p} \leqslant \left(\sum_{k=1}^{n} |x_k|^p \right)^{1/p} + \left(\sum_{k=1}^{n} |y_k|^p \right)^{1/p}.$$

Por hipótesis, las series

$$\sum_{k=1}^{\infty} |x_k|^p \quad \text{y} \quad \sum_{k=1}^{n} |y_k|^p$$

convergen; por eso, pasando al límite para $n \to \infty$, obtenemos

$$\left(\sum_{k=1}^{\infty} |y_k - x_k|^p \right)^{1/p} \leqslant \left(\sum_{k=1}^{\infty} |x_k|^p \right)^{1/p} + \left(\sum_{k=1}^{\infty} |y_k|^p \right)^{1/p} < \infty. \tag{19}$$

Esto demuestra que la fórmula (18) mediante la cual se define la distancia en l_p, tiene sentido para cualesquiera $x, y \in l_p$. Al mismo tiempo, la desigualdad (19) muestra que en l_p se verifica el axioma triangular. Los demás axiomas son evidentes.

Un número ilimitado de nuevos ejemplos proporciona la siguiente idea. Sea $R = (X, \rho)$ un espacio métrico y M cualquier subconjunto de X. En este caso el conjunto M con la misma función $\rho(x, y)$ pero definida ya para x e y de M también representa un espacio métrico, que se denomina *subespacio* del espacio R.

2°. Aplicaciones continuas de espacios métricos. Isometría. Sean X e Y dos espacios métricos y f una aplicación del espacio X en Y. Por lo tanto, a cada elemento $x \in X$ se pone en correspondencia un elemento $y = f(x)$ de Y. Esta aplicación se dice *continua* en el punto $x_0 \in X$, si para cada $\varepsilon > 0$ existe un $\delta > 0$ tal, que para todos $x \in X$ que cumplen la desigualdad

$$\rho(x, x_0) < \delta,$$

se verifica la desigualdad

$$\rho_1(f(x), f(x_0)) < \varepsilon$$

(aquí ρ es la distancia en X y ρ_1 la distancia en Y). Si la aplicación f es continua en todos los puntos del espacio X, se dice que f es continua sobre X. Si X e Y son dos conjuntos numéricos, es decir, f es una función numérica definida sobre un subconjunto X de la recta numérica, esta definición de la continuidad de una aplicación coincide con definición, conocida del Análisis elemental, de la continuidad de una función.

Siendo la aplicación f del espacio X sobre el espacio Y biunívoca, existe la aplicación inversa $x=f^{-1}(y)$ del espacio Y sobre el espacio X. Si la aplicación f es biunívoca y bicontinua (es decir, tanto f como f^{-1} son aplicaciones continuas), se la denomina aplicación homeomorfa o *homeomorfismo* y los espacios X e Y, entre los cuales se puede establecer una aplicación homeomorfa, se denominan espacios homeomorfos. Como ejemplo de espacios homeomorfos pueden servir toda la recta numérica $(-\infty, \infty)$ y un intervalo, por ejemplo, el intervalo $(-1, 1)$. En este caso el homeomorfismo se establece mediante la fórmula

$$y=\frac{2}{\pi}\operatorname{arctg} x.$$

Un caso importante particular de homeomorfismo en la así llamada aplicación isométrica de espacios métricos.

Una aplicación biunívoca f del espacio métrico $R=(X, \rho)$ sobre el espacio métrico $R'=(Y, \rho')$ se denomina aplicación *isométrica*, si

$$\rho(x_1, x_2)=\rho'(f(x_1), f(x_2))$$

para cualesquiera $x_1, x_2\in R$. Dos espacios R y R', entre los cuales se puede establecer una correspondencia de isometría, se denominan *isométricos*.

La isometría de dos espacios R y R' significa que las relaciones métricas entre sus elementos son las mismas; puede ser distinta sólo la naturaleza de los propios elementos, lo que, desde el punto de vista de la teoría de espacios métricos, no tiene importancia. En lo sucesivo los espacios isométricos serán considerados como idénticos.

Al final del § 5 de este capítulo volveremos a tratar, desde un punto de vista más general, los conceptos aquí introducidos (continuidad, homeomorfismo).

§ 2. CONVERGENCIA. CONJUNTOS ABIERTOS Y CERRADOS

1°. Puntos de acumulación. Adherencia. En este parágrafo expondremos algunos conceptos principales de la teoría de espacios métricos que emplearemos frecuentemente en lo sucesivo.

Una *bola abierta* $B(x_0, r)$ en el espacio métrico R es el conjunto de los puntos $x\in R$ que verifican la condición

$$\rho(x, x_0)<r.$$

El punto fijo x_0 se llama *centro* y el número r, *radio* de esta bola.

Una *bola cerrada* $B[x_0, r]$ es el conjunto de los puntos $x \in R$ que cumplen la condición

$$\rho(x, x_0) \leqslant r.$$

Una bola abierta de radio ε y centro en x_0 se denominará también *ε-vecindad*[1] del punto x_0 y se denotará $O_\varepsilon(x_0)$.

EJERCICIO. Dése un ejemplo de espacio métrico y dos bolas $B(x, \rho_1)$ y $B(y, \rho_2)$ en él, tales que $\rho_1 > \rho_2$, $B(x, \rho_1) \subset B(y, \rho_2)$.

Un punto $x \in R$ se denomina *punto de adherencia* del conjunto $M \subset R$, si cualquier vecindad suya contiene al menos un punto de M. La totalidad de los puntos de adherencia del conjunto M se denota mediante $[M]$ y se llama la *adherencia* (también clausura) de este conjunto. De esta manera hemos definido para los conjuntos de un espacio métrico la *operación de adherencia* que consiste en el paso del conjunto M a su adherencia $[M]$.

TEOREMA 1. *La operación de adherencia tiene las siguientes propiedades:*

1) $M \subset [M]$,
2) $[[M]] = [M]$,
3) *si* $M_1 \subset M_2$, *entonces* $[M_1] \subset [M_2]$,
4) $[M_1 \cup M_2] = [M_1] \cup [M_2]$.

DEMOSTRACION. La primera afirmación es evidente, ya que todo punto perteneciente a M es un punto de adherencia del conjunto M. Demostremos la segunda propiedad. Sea $x \in [[M]]$. En este caso, en cualquier vecindad $O_\varepsilon(x)$ de este punto existirá un punto $x_1 \in [M]$. Pongamos $\varepsilon - \rho(x, x_1) = \varepsilon_1$ y consideremos la bola $O_{\varepsilon_1}(x_1)$. Esta bola se encuentra íntegramente dentro de la bola $O_\varepsilon(x)$. En efecto, si $z \in O_{\varepsilon_1}(x_1)$, tendremos $\rho(z, x_1) < \varepsilon$ y puesto que $\rho(x, x_1) = \varepsilon - \varepsilon_1$, encontraremos, de acuerdo con el axioma triangular, que

$$\rho(z, x) < \varepsilon_1 + (\varepsilon - \varepsilon_1) = \varepsilon,$$

es decir, $z \in O_\varepsilon(x)$. Como $x_1 \in [M]$, existirá en $O_{\varepsilon_1}(x)$ un punto $x_2 \in M$. Pero en este caso $x_2 \in O_\varepsilon(x)$ y, puesto que $O_\varepsilon(x)$ es una vecindad arbitraria de x, tendremos $x \in [M]$. Hemos demostrado la segunda afirmación.

La tercera propiedad es evidente. Demostremos, finalmente, la cuarta.

---

[1] Algunos autores prefieren emplear en este caso el término «ε-entorno». (*N. del T.*)

Si $x \in [M_1 \cup M_2]$, x pertenece por lo menos a uno de los conjuntos $[M_1]$ o $[M_2]$, es decir

$$[M_1 \cup M_2] \subset [M_1] \cup [M_2].$$

Puesto que $M_1 \subset M_1 \cup M_2$ y $M_2 \subset M_1 \cup M_2$, la inclusión inversa se desprende de la propiedad 3).

Hemos demostrado completamente el teorema.

Un punto $x \in R$ se llama *punto de acumulación* del conjunto $M \subset R$, cuando en toda vecindad suya existe un número infinito de puntos de M.

El punto de acumulación puede pertenecer y puede no pertenecer a M. Por ejemplo, si M es el conjunto de los números racionales del segmento $[0, 1]$, todo punto de este conjunto es un punto de acumulación de M.

Un punto x, *perteneciente a* M, se llama *punto aislado* de este conjunto, cuando una vecindad suya $O_\varepsilon(x)$ suficientemente pequeña no contiene otros puntos de M distintos de x. Proponemos al lector demostrar, a título de ejercicio, la siguiente afirmación:

Todo punto de adherencia del conjunto M o bien es un punto de acumulación o bien un punto aislado de este conjunto.

De aquí se puede deducir que la adherencia $[M]$ se compone, en general, de tres tipos de puntos:

1) los puntos aislados del conjunto M;

2) los puntos de acumulación del conjunto M, pertenecientes a M;

3) los puntos de acumulación del conjunto M que no pertenecen a M; de esta forma la adherencia $[M]$ se obtiene agregando a M todos sus puntos de acumulación.

2°. Convergencia. Sea $x_1, x_2, \ldots$ una sucesión de puntos en el espacio métrico R. Se dice que esta sucesión *converge al punto* x, cuando toda vecindad $O_\varepsilon(x)$ del punto x contiene todos los puntos x_n, empezando desde alguno, es decir, cuando a cada número $\varepsilon > 0$ le corresponde un número N_ε tal que $O_\varepsilon(x)$ contiene todos los puntos x_n para $n > N_\varepsilon$. El punto x se llama *límite* de la sucesión $\{x_n\}$.

Es obvio que se puede enunciar esta definición también de la siguiente manera: la sucesión $\{x_n\}$ converge a x, cuando

$$\lim_{n \to \infty} \rho(x, x_n) = 0.$$

De esta definición se desprende directamente que 1) ninguna sucesión puede tener dos límites distintos y que 2) si la sucesión $\{x_n\}$ converge al punto x, toda sucesión parcial contenida en ella converge al mismo punto.

El siguiente teorema explica la relación estrecha existente entre los conceptos de punto de adherencia y de límite.

TEOREMA 2. *Para que el punto x sea un punto de adherencia del conjunto M, es necesario y suficiente que exista una sucesión $\{x_n\}$ de puntos de M que converja a x.*

DEMOSTRACION. La condición es necesaria, puesto que siendo x un punto de adherencia del conjunto M, en cada vecindad $O_{1/n}(x)$ existe por lo menos un punto $x_n \in M$. Estos puntos forman una sucesión que converge a x. La suficiencia es evidente.

Siendo x un punto de acumulación del conjunto M, los puntos $x_n \in O_{1/n}(x) \cap M$, correspondientes a diferentes n, pueden escogerse de manera que no coincidan entre sí. Es decir, *para que el punto x sea un punto de acumulación de M, es necesario y suficiente que en M exista una sucesión de puntos distintos dos a dos que converga a x.*

El concepto de continuidad de una aplicación del espacio métrico X en el espacio métrico Y, que hemos introducido en el § 1, puede enunciarse ahora en términos de convergencia de sucesiones: la aplicación $y=f(x)$ es continua en el punto x_0, cuando cualquiera que sea la sucesión $\{x_n\}$ convergente a x_0, la sucesión $\{y_n=f(x_n)\}$ converge a $y_0=f(x_0)$. La demostración de la equivalencia de esta definición a la dada en el § 1 no difiere en nada de la demostración de la equivalencia de las dos definiciones de la continuidad («en el lenguaje de ε, δ» y «en el lenguage de sucesiones») de las funciones de argumento numérico y el lector podrá realizarla.

3°. Subconjuntos densos. Sean A y B dos conjuntos de un espacio métrico R. El conjunto A se dice *denso* en B, cuando $[A]\supset B$. En particular, el conjunto A es *siempre denso* (en el espacio R), cuando su adherencia $[A]$ coincide con todo el espacio R. Por ejemplo, el conjunto de los números racionales es siempre denso en la recta numérica. Se dice que un conjunto A es *nunca denso*, cuando no es denso en ninguna bola.

Ejemplos de espacios provistos de un conjunto numerable siempre denso. Un espacio, que tiene un conjunto numerable siempre denso, se llama *separable*. Consideremos desde este punto de vista los ejemplos expuestos en el § 1.

1. El espacio «discreto» del ejemplo 1 del § 1 contiene un conjunto numerable siempre denso, si, y sólo si, está compuesto por un número numerable de puntos. Ello se debe a que la adherencia $[M]$ de cualquier conjunto M de este espacio coincide con M.

Todos los espacios dados en los ejemplos 2—8 del § 1 tienen conjuntos numerables siempre densos. Señalemos en cada uno de ellos un conjunto de este tipo, recomendando con insistencia al lector realizar la demostración detallada.

2. Sobre el eje real R^1, los puntos racionales.

3—5. En el espacio euclídeo de n dimensiones R^n y en los espacios R_1^n y R_0^n, el conjunto de vectores con coordenadas racionales.

6. En el espacio $C_{[a, b]}$, el conjunto de todos los polinomios de coeficientes racionales.

7. En el espacio l_2, el conjunto de sucesiones en cada una de las cuales todos los miembros son racionales y solamente un número finito (para cada sucesión el suyo) de estos miembros es distinto de cero.

8. En el espacio $C^2_{[a, b]}$, el conjunto de todos los polinomios de coeficientes racionales.

El espacio m de sucesiones acotadas (ejemplo 9 del § 1) no tiene ningún conjunto numerable siempre denso. En efecto, consideremos todas las sucesiones posibles compuestas de ceros y unidades. Ellas forman un conjunto de potencia de continuo (ya que se puede establecer una correspondencia biunívoca entre estas sucesiones y subconjuntos de la serie natural). La distancia entre dos de estos puntos, definida mediante la fórmula (11) del § 1, es igual a 1. Envolvamos cada uno de estos puntos en una bola abierta de radio $^1/_2$. Estas bolas no se intersecan. Si tenemos un conjunto siempre denso en este espacio, cada una de las bolas construidas deberá contener por lo menos un punto de este conjunto y, por consiguiente, él no puede ser numerable.

4°. Conjuntos abiertos y cerrados. Consideremos los tipos más importantes de conjuntos en espacios métricos, es decir los conjuntos abiertos y cerrados. El conjunto M, perteneciente a un espacio métrico R, se llama *cerrado*, cuando coincide con su adherencia: $[M] = M$. En otras palabras, el conjunto se llama cerrado, si contiene todos sus puntos de acumulación.

En virtud del teorema 1, la adherencia de cualquier conjunto M es un conjunto cerrado. Del mismo teorema se desprende que $[M]$ es el menor conjunto cerrado que contiene a M.

Ejemplos. 1. Todo segmento $[a, b]$ de la recta numérica es un conjunto cerrado.

2. Una bola cerrada representa un conjunto cerrado. En particular, el conjunto de funciones f del espacio $C_{[a, b]}$, que verifican la condición $|f(t)| \leqslant K$, es cerrado.

3. El conjunto de funciones de $C_{[a, b]}$, que verifican la condición $|f(t)| < K$ (bola abierta), no es cerrado; su adherencia es

el conjunto de funciones que cumplen la condición $|f(t)| \leqslant K$.

4. Cualquiera que sea el espacio métrico R, el conjunto vacío $\varnothing$ y todo el espacio R son cerrados.

5. Todo conjunto, compuesto por un número finito de puntos, es cerrado.

Las propiedades principales de los conjuntos cerrados pueden enunciarse por medio del siguiente teorema.

TEOREMA 3. *La intersección de cualquier número y la unión de un número finito de conjuntos cerrados son conjuntos cerrados.*

DEMOSTRACION. Sea $F = \bigcap F_\alpha$ la intersección de los conjuntos cerrados F_α y sea x un punto de acumulación del conjunto F. Ello significa que cualquier vecindad $O_\varepsilon(x)$ contiene un número infinito de puntos de F. Pero entonces $O_\varepsilon(x)$ contiene así mismo un número infinito de puntos de cada conjunto F_α y, puesto que todos los F_α son cerrados, el punto x pertenece a cada F_α; por consiguiente, $x \in F = \bigcap F_\alpha$, es decir, F es cerrado.

Sea ahora F la unión de un número finito de conjuntos cerrados: $F \bigcup\limits_{i=1}^{n} F_i$; supongamos que el punto x no pertenece a F. Demostremos que x no puede ser un punto de acumulación de F. En efecto, x no pertenece a ninguno de los conjuntos cerrados F_i y, por consiguiente, no es punto de acumulación de ninguno de ellos. Esto quiere decir que para todo i se puede encontrar una vecindad $O_{\varepsilon_i}(x)$ del punto x que contiene, a lo sumo, un número finito de puntos de F_i. Tomando la menor de las vecindades $O_{\varepsilon_1}(x), \ldots, O_{\varepsilon_n}(x)$, obtendremos una vecindad $O_\varepsilon(x)$ del punto x que contiene a lo sumo un número finito de puntos de F.

Resumiendo, si el punto x no pertenece a F, no puede ser punto de acumulación de F, es decir, F es cerrado.

El teorema queda demostrado.

Un punto x se llama *punto interior* del conjunto M, cuando hay una vecindad $O_\varepsilon(x)$ de este punto que pertenece íntegramente a M.

Un conjunto, todos los puntos del cual son interiores, se llama *conjunto abierto*.

Ejemplos. 6. Un intervalo (a, b) de la recta numérica R^1 es un conjunto abierto; en efecto, si $a < \alpha < b$, la vecindad $O_\varepsilon(\alpha)$, donde $\varepsilon = \min(\alpha - a, b - \alpha)$, está contenida íntegramente en el intervalo (a, b).

7. Una bola abierta $B(a, r)$ de cualquier espacio métrico R es un conjunto abierto. En efecto, si $x \in B(a, r)$, tenemos

$\rho(a, x) < r$. Tomemos $\varepsilon = r - \rho(a, x)$. Entonces $B(x, \varepsilon) \subset B(a, r)$.

8. El conjunto de funciones continuas sobre $[a, b]$ que verifican la condición $f(t) < g(t)$, donde $g(t)$ es una función continua determinada, representa un subconjunto abierto del espacio $C_{[a, b]}$.

TEOREMA 4. *Para que el conjunto M sea abierto es necesario y suficiente que su complemento $R \setminus M$ al espacio R sea cerrado.*

DEMOSTRACION. Si M es abierto, cada punto x de M posee una vecindad que pertenece íntegramente a M, es decir, que no tiene ningún punto común con $R \setminus M$. Por consiguiente, ninguno de los puntos que no pertenecen a $R \setminus M$ puede ser un punto de adherencia de $R \setminus M$, es decir, $R \setminus M$ es cerrado. Viceversa, si $R \setminus M$ es cerrado, cualquier punto de M posee una vecindad que pertenece íntegramente a M, es decir, M es abierto.

Puesto que el conjunto vacío y todo el espacio R son cerrados y al mismo tiempo son complementos uno del otro, del teorema demostrado se desprende el siguiente corolario.

COROLARIO. *El conjunto vacío y todo el espacio R son abiertos.*

Del teorema 3 y del principio de dualidad (la intersección de complementos es igual al complemento de la unión y la unión de complementos es igual al complemento de las intersecciones; véase pág. 16) se sigue el siguiente teorema importante, dual al teorema 3.

TEOREMA 3'. *La unión de un número cualquiera (finito o infinito) y la intersección de un número finito de conjuntos abiertos son conjuntos abiertos.*

5°. Conjuntos abiertos y cerrados sobre la recta. La estructura de los conjuntos abiertos y cerrados en uno u otro espacio métrico puede ser muy compleja. Esto se refiere incluso a los conjuntos abiertos y cerrados de un espacio euclídeo de dos o más dimensiones. Sin embargo, en el caso unidimensional, es decir, en el caso de la recta, no es difícil dar una descripción de todos los conjuntos abiertos (y, por consiguiente, de todos los cerrados). Ella viene dada por el siguiente teorema.

TEOREMA 5. *Todo conjunto abierto de una recta numérica representa la suma de un número finito o numerable de intervalos disjuntos dos a dos*[1].

---

[1] Los conjuntos de tipo $(-\infty, \infty)$, (α, ∞) y $(-\infty, \beta)$ también los onsideramos como intervalos.

DEMOSTRACION. Sea G un conjunto abierto sobre la recta y x un punto suyo cualquiera. De acuerdo con la definición de conjunto abierto, existe por lo menos un intervalo que contiene a x y está contenido en G. Denotemos mediante I_x la suma de todos estos intervalos y demostremos que I_x es también un intervalo. Tomemos para ello

$$a = \inf I_x, \quad b = \sup I_x$$

(puede ocurrir que $a = -\infty$ y $b = +\infty$) y comprobemos que $I_x = (a, b)$. Está claro, ante todo, que $I_x \subset (a, b)$. Viceversa, sea y un punto arbitrario de (a, b), distinto de x; probemos que $y \in I_x$. Aceptemos, para concretar, que $a < y < x$. En este caso, en I_x existirá un punto y' tal que $a < y' < y$. Esto significa que en G hay un intervalo que contiene los puntos y' y x. Pero entonces también contiene y, es decir, $y \in I_x$. Del mismo modo se puede considerar el caso $y > x$. El punto x pertenece a I_x por hipótesis. Por consiguiente, $I_x = (a. b)$. El intervalo (a, b) ha sido definido de manera que él pertenece a G y no está contenido en ningún intervalo mayor que pertenezca a G. Es evidente que los intervalos I_{x_1} e I_{x_2}, correspondientes a dos distintos puntos, o bien coinciden o bien no tienen puntos comunes (de lo contrario, I_{x_1} e I_{x_2} estarían contenidos en el intervalo $I_{x_1} \cup I_{x_2}$ mayor que cada uno de ellos y perteneciente a G). Un sistema tal de intervalos disjuntos es a lo sumo numerable: en efecto, escogiendo en cada uno de estos intervalos y de una manera arbitraria un punto racional, estableceremos una correspondencia biunívoca entre estos intervalos y ciertos subconjuntos del conjunto de números racionales. Está claro, finalmente, que la suma de estos intervalos coincide con G. El teorema queda demostrado.

Puesto que los conjuntos cerrados son complementos de los abiertos, de aquí se sigue que todo conjunto cerrado sobre la recta se obtiene omitiendo de la recta un número finito o numerable de intervalos.

Los ejemplos más simples de conjuntos cerrados son los segmentos, los puntos aislados y la suma de un número finito de conjuntos de estos tipos. Consideremos un ejemplo más complejo de conjunto cerrado sobre la recta, el así llamado *conjunto de Cantor*.

Sea F_0 el segmento $[0, 1]$. De él excluyamos el intervalo $\left(\frac{1}{3}, \frac{2}{3}\right)$ y designemos mediante F_1 el conjunto cerrado que queda. Después, omitamos de F_1 los intervalos $\left(\frac{1}{9}, \frac{2}{9}\right)$ y $\left(\frac{7}{9}, \frac{8}{9}\right)$ y designemos mediante F_2 el conjunto cerrado que queda (compuesto por cuatro segmentos). En cada uno de estos cuatro segmentos excluyamos el intervalo abierto central de longitud $\frac{1}{3^3}$, etc. (fig. 9). Continuando este proceso, obtendremos una sucesión decreciente de conjuntos cerrados F_n. Tomemos

$$F=\bigcap_{n=0}^{\infty} F_n.$$

F es un conjunto cerrado (como intersección de cerrados). Se obtiene omitiendo del segmento $[0, 1]$ un número numerable de

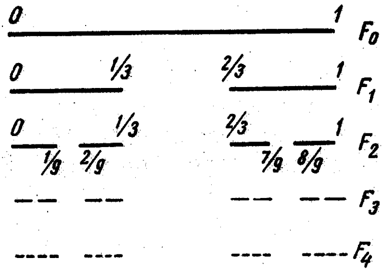


FIG. 9

intervalos. Examinemos la estructura del conjunto F. Pertenecen a él, evidentemente, los puntos

$$0,\ 1,\ \frac{1}{3},\ \frac{2}{3},\ \frac{1}{9},\ \frac{2}{9},\ \frac{7}{9},\ \frac{8}{9},\ \ldots \tag{1}$$

que representan los extremos de los intervalos omitidos. El conjunto F no se compone, sin embargo, solamente de estos puntos. En efecto, se puede describir los puntos del segmento $[0, 1]$, que pertenecen al conjunto F, de la siguiente forma. Representemos cada uno de los números x, $0 \leqslant x \leqslant 1$, en el sistema ternario:

$$x=\frac{a_1}{3}+\frac{a_2}{3^2}+\ldots+\frac{a_n}{3^n}+\ldots,$$

donde los números a_n pueden tomar los valores 0, 1 y 2. Al igual que en el caso de fracciones decimales, algunos números pueden tener dos representaciones. Por ejemplo,

$$\frac{1}{3}=\frac{1}{3}+\frac{0}{3^2}+\ldots+\frac{0}{3^n}+\ldots=\frac{0}{3}+\frac{2}{3^2}+\frac{2}{3^3}+\ldots+\frac{2}{3^n}+\ldots$$

Es fácil comprobar que al conjunto F pertenecen aquellos números, y solamente aquellos, x, $0 \leqslant x \leqslant 1$, que pueden representarse al menos de una forma mediante una fracción ternaria tal que en la sucesión $a_1, a_2, \ldots, a_n, \ldots$ no figura la unidad. Por consiguiente, a cada punto $x \in F$ se le puede poner en correspondencia la sucesión

$$a_1, a_2, \ldots, a_n, \ldots, \tag{2}$$

donde a_n es igual a 0 ó a 2. La totalidad de estas sucesiones forma un conjunto de potencia de continuo. Es fácil comprobar esto, si se pone en correspondencia a cada sucesión (2) la sucesión

$$b_1, b_2, \ldots, b_n, \ldots, \tag{2'}$$

donde $b_n = 0$, cuando $a_n = 0$, y $b_n = 1$, cuando $a_n = 2$. La sucesión (2') puede ser considerada como la representación de un número real y, $0 \leqslant y \leqslant 1$ mediante una fracción binaria. De este modo obtenemos una aplicación del conjunto F sobre todo el segmento $[0, 1]$. De aquí se desprende que F tiene potencia de continuo[1]. Puesto que el conjunto de los puntos (1) es numerable, el conjunto F no puede limitarse a ellos.

EJERCICIOS. 1. Demuéstrese directamente que el punto $\frac{1}{4}$ pertenece al conjunto F sin ser extremo de ninguno de los intervalos omitidos.

Sugerencia. El punto $\frac{1}{4}$ divide el segmento $[0, 1]$ en dos partes según la razón $\frac{1}{3}$. También divide en dos partes según la razón $\frac{1}{3}$ el segmento $\left[0, \frac{1}{3}\right]$ que queda después de la primera exclusión, etc.

Los puntos (1) se llaman *puntos de primer género* del conjunto F y los demás puntos suyos se denominan *puntos de segundo género*.

2. Demuéstrese que los puntos de primer género forman un conjunto siempre denso en F.

3. Demuéstrese que los números $t_1 + t_2$, donde $t_1, t_2 \in F$ llenan todo el segmento $[0, 2]$.

Hemos demostrado que el conjunto F es de potencia de continuo, es decir, tiene tantos puntos como todo el segmento $[0, 1]$. Es interesante comparar esto con el siguiente resultado: la suma de longitudes de todos los intervalos omitidos es igual a $\frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \ldots$, es decir, ¡exactamente a la unidad!

---

[1] La correspondencia establecida entre F y el segmento $[0, 1]$ es unívoca, pero no es biunívoca (porque un mismo número puede ser representado, a veces, por diferentes fracciones). De aquí se deduce que la potencia de F es no menos que la de continuo. Pero F es parte del segmento $[0, 1]$ y, por consiguiente, su potencia no puede ser mayor que la potencia de continuo.

Observaciones complementarias.

(1) Sea M un conjunto de un espacio métrico R y sea x un punto de este espacio. *La distancia del punto x al conjunto M* se define por el número

$$\rho(M, x) = \inf_{a \in M} \rho(a, x).$$

Si $x \in M$, tenemos $\rho(M, x) = 0$; en cambio, $\rho(M, x) = 0$ no implica que $x \in M$. De acuerdo con la definición de punto de adherencia, obtenemos inmediatamente que $\rho(M, x) = 0$ si, y sólo si, x es un punto de adherencia del conjunto M.

Por eso se puede decir que la operación de adherencia consiste en que al conjunto se añaden todos aquellos puntos cuyas distancias al conjunto sean igual a cero.

(2) De una manera análoga se define la distancia entre dos conjuntos. Si A y B son dos conjuntos del espacio métrico R, tenemos

$$\rho(A, B) = \inf_{\substack{a \in A \\ b \in B}} \rho(a, b).$$

Si $A \cap B \neq \varnothing$, entonces $\rho(A, B) = 0$; la afirmación contraria no tiene, en general, lugar.

(3) Sea M_K el conjunto de todas las funciones f de $C_{[a, b]}$ que verifican la condición de Lipschitz: para todos los $t_1, t_2 \in [a, b]$

$$|f(t_1) - f(t_2)| \leqslant K |t_2 - t_1|,$$

donde K es un número fijo. El conjunto M_K es cerrado. Coincide con la adherencia del conjunto de funciones diferenciables en $[a, b]$ y tales que $|f'(t)| \leqslant K$.

(4) El conjunto $M = \bigcup_K M_K$ de todas las funciones que verifican la condición de Lipschitz, cada una con su número K, no es cerrado. Su adherencia coincide con todo el espacio $C_{[a, b]}$.

(5) Un conjunto abierto G de un espacio euclídeo de n dimensiones se llama *conexo*, cuando dos cualesquiera puntos $x, y \in G$ pueden unirse mediante una quebrada que pertenece íntegramente a G. Por ejemplo, el conjunto formado por los puntos interiores del círculo $x^2 + y^2 < 1$ es conexo. Al contrario, la suma de dos círculos

$$x^2 + y^3 < 1 \text{ y } (x-2)^2 + y^2 < 1$$

es un conjunto desconexo (¡aun cuando estos círculos tienen un punto comun de adherencia!). Un subconjunto abierto H de un conjunto abierto G se llama *componente* del conjunto G, cuando es conexo sin ser parte de ningún otro subconjunto abierto conexo mayor del conjunto G. Valiéndose del teorema de Zorn, es fácil ver que tiene lugar la siguiente afirmación: todo conjunto abierto G de un espacio euclídeo de n dimensiones es la suma de un número a lo sumo numerable de componentes disjuntos dos a dos.

En el caso de $n = 1$, es decir, en la recta, todo conjunto abierto conexo es un intervalo (incluyéndose en los intervalos los intervalos infinitos $(-\infty, a)$, (b, ∞) y $(-\infty, \infty)$. Por consiguiente, el teorema 5 sobre la estructura de los conjuntos abiertos de la recta contiene dos afirmaciones: a) todo conjunto abierto de la recta es la suma de un número finito o numerable de componentes y b) un conjunto abierto conexo de la recta es un intervalo. La primera de estas afirmaciones se cumple también para los conjuntos de los espacios euclídeos de n dimensiones (y admite otras generalizaciones), mientras que la segunda afirmación se refiere solamente a la recta.

§ 3. ESPACIOS MÉTRICOS COMPLETOS

1°. Definición y ejemplos de espacios métricos completos. Desde los pasos iniciales en el estudio del Análisis Matemático, podemos persuadirnos del papel tan importante que desempeña en el Análisis la propiedad de complitud de la recta numérica, es decir, el hecho de que toda sucesión fundamental de números reales converge a un límite determinado. La recta numérica representa el ejemplo más sencillo de los así llamados espacios métricos *completos*, cuyas propiedades principales consideraremos en este punto.

Una sucesión $\{x_n\}$ de puntos de un espacio métrico R se llamará *fundamental*, cuando verifique el criterio de Cauchy, es decir, cuando para cualquier $\varepsilon>0$ exista un número N_ε tal, que $\rho(x_{n'}, x_{n''})<\varepsilon$ para cualesquiera $n'>N_\varepsilon$, $n''>N_\varepsilon$.

Del axioma triangular se sigue inmediatamente que toda sucesión convergente es fundamental. En efecto, si $\{x_n\}$ converge a x, cualquiera que sea $\varepsilon>0$ se puede encontrar un número N_ε tal que $\rho(x_n, x)<\frac{\varepsilon}{2}$ para todo $n>N_\varepsilon$. Por eso $\rho(x_{n'}, x_{n''})\leqslant \leqslant\rho(x_{n'}, x)+\rho(x_{n''}, x)<\varepsilon$ para cualesquiera $n'>N_\varepsilon$ y $n''>N_\varepsilon$.

DEFINICION 1. Un espacio R se llama *completo*, cuando toda sucesión fundamental de él converge.

Ejemplos. Todos los espacios considerados en el § 1, a excepción del señalado en el ejemplo 8, son completos. En efecto:

1. En el espacio de puntos aislados (ejemplo 1 del § 1) son fundamentales sólo las sucesiones estacionarias, es decir, en las que se repite constantemente, comenzando desde cierto número, un mismo punto. Toda sucesión de este tipo converge evidentemente y, por consiguiente, este espacio es completo.

2. La complitud del espacio R^1, compuesto por los números reales, es conocida del Análisis.

3. La complitud del espacio euclídeo R^n se sigue inmediatamente de la complitud de R^1. En efecto, sea $\{x^{(p)}\}$ una sucesión fundamental de puntos de R^n; esto significa que para todo $\varepsilon>0$ existe un $N=N_\varepsilon$ tal que

$$\sum_{k=1}^{n}(x_k^{(p)}-x_k^{(q)})^2<\varepsilon^2$$

para todos p, q mayores que N. Aquí $x^{(p)}=\{x_1^{(p)}, \ldots, x_n^{(p)}\}$. En este caso para todo $k=1, 2, \ldots, n$ obtendremos la desigualdad correspondiente a la coordenada $x_k^{(p)}$:

$$|x_k^{(p)}-x_k^{(q)}|<\varepsilon$$

siempre que $p, q > N$; por consiguiente, $\{x_k^{(p)}\}$ es una sucesión numérica fundamental. Tomemos

$$x_k = \lim_{p \to \infty} x_k^{(p)} \text{ y } x = (x_1, x_2, \ldots, x_n).$$

Es obvio entonces que

$$\lim_{p \to \infty} x^{(p)} = x.$$

4—5. La complitud de los espacios R_0^n y R_1^n se demuestra de una manera totalmente análoga.

6. Demostremos la complitud del espacio $C_{[a, b]}$. Sea $\{x_n(t)\}$ una sucesión fundamental de $C_{[a, b]}$. Ello significa que para todo $\varepsilon > 0$ existe un N tal que

$$|x_n(t) - x_m(t)| < \varepsilon$$

siempre que $n, m > N$ y para todo t, $a \leqslant t \leqslant b$. De aquí se deduce que la sucesión $\{x_n(t)\}$ converge uniformemente. En este caso, como se sabe, su límite $x(t)$ será una función continua. Haciendo tender m al infinito en la desigualdad anterior, obtendremos

$$|x_n(t) - x(t)| \leqslant \varepsilon$$

para todo t y para todo $n > N$ y esto significa precisamente que $\{x_n(t)\}$ converge a $x(t)$ en el sentido de la métrica del espacio $C_{[a, b]}$.

7. El espacio l_2. Sea $\{x^{(n)}\}$ una sucesión fundamental en l_2. Esto significa que para todo $\varepsilon > 0$ existe un N tal que

$$\rho^2(x^{(n)}, x^{(m)}) = \sum (x_k^{(n)} - x_k^{(m)})^2 < \varepsilon, \text{ para } n, m > N. \qquad (1)$$

Aquí

$$x^{(n)} = (x_1^{(n)}, x_2^{(n)}, \ldots, x_k^{(n)}, \ldots).$$

De (1) se desprende que cualquiera que sea k,

$$(x_k^{(n)} - x_k^{(m)})^2 < \varepsilon,$$

es decir, la sucesión de números reales $\{x_k^{(n)}\}$ es fundamental cualquiera que sea k y, por consiguiente, converge. Tomemos $x_k = \lim_{n \to \infty} x_k^{(n)}$. Sea x la sucesión $(x_1, x_2, \ldots, x_k, \ldots)$. Debemos demostrar que

a) $\sum_{k=1}^{\infty} x_k^2 < \infty$, es decir, $x \in l_2$,

b) $\lim_{n \to \infty} \rho(x^{(n)}, x) = 0$.

Demostrémoslo. De la desigualdad (1) se deduce que para todo número fijo M

$$\sum_{k=1}^{M} (x_k^{(n)} - x_k^{(m)})^2 < \varepsilon.$$

Fijando n, pasemos al límite para $m \to \infty$. Tendremos

$$\sum_{k=1}^{M} (x_k^{(n)} - x_k)^2 \leqslant \varepsilon.$$

Esta relación se verifica para cualquier M. Pasando en ella al límite para $M \to \infty$, obtendremos

$$\sum_{k=1}^{\infty} (x_k^{(n)} - x_k)^2 \leqslant \varepsilon. \tag{2}$$

De la convergencia de las series $\sum_{k=1}^{\infty} (x_k^{(n)})^2$ y $\sum_{k=1}^{\infty} (x_k^{(n)} - x_k)^2$ se deduce la convergencia de la serie $\sum_{k=1}^{\infty} x_k^2$ (véase el ejemplo 7 del § 1) y con esto queda demostrada la afirmación a). Ahora bien, puesto que ε es arbitrariamente pequeño, la desigualdad (2) significa que

$$\lim_{n\to\infty} \rho(x^{(n)}, x) = \lim_{n\to\infty} \sqrt{\sum_{k=1}^{\infty} (x_k^n - x_k)^2} = 0,$$

es decir, que $x^{(n)} \to x$ en la métrica de l_2. Queda también demostrada la afirmación b).

8. Es fácil ver que el espacio $C^2_{[a,\,b]}$ no es completo. Consideremos, por ejemplo, la siguiente sucesión de funciones continuas

$$\varphi_n(t) = \begin{cases} -1 & \text{para } -1 \leqslant t \leqslant -\frac{1}{n}, \\ nt & \text{para } -\frac{1}{n} \leqslant t \leqslant \frac{1}{n}, \\ 1 & \text{para } \frac{1}{n} \leqslant t \leqslant 1. \end{cases}$$

Es una sucesión fundamental de $C^2_{[-1,\,1]}$, puesto que

$$\int_{-1}^{1} (\varphi_n(t) - \varphi_m(t))^2\, dt \leqslant \frac{2}{\min(n, m)}.$$

Sin embargo, no converge a ninguna función de $C^2_{[-1,\,1]}$. En efecto, sea f una función de $C^2_{[-1,\,1]}$ y sea ψ la función discontinua, idéntica a -1 para $t < 0$ y $+1$ para $t \geqslant 0$.

De la desigualdad de Cauchy—Buniakovski (que, evidentemente, se verifica también, para las funciones continuas a trozos) se deduce que

$$\left(\int_{-1}^{1}(f(t)-\psi(t))^2\,dt\right)^{\frac{1}{2}}\leqslant\left(\int_{-1}^{1}(f(t)-\varphi_n(t))^2\,dt\right)^{\frac{1}{2}}+$$
$$+\left(\int_{-1}^{1}(\varphi_n(t)-\psi(t))^2\,dt\right)^{\frac{1}{2}}$$

Puesto que f es una función continua, la integral del miembro de la izquierda es distinta de cero. Además, es obvio que

$$\lim_{n\to\infty}\int_{-1}^{1}(\varphi_n(t)-\psi(t))^2\,dt=0.$$

Por eso, $\int_{-1}^{1}(f(t)-\varphi_n(t))^2\,dt$ no puede tender a cero para $n\to\infty$.

EJERCICIO. Demuéstrese que el espacio de todas las sucesiones acotadas (ejemplo 9 del § 1) es completo.

2°. Principio de bolas encajadas. En el Análisis se emplea con frecuencia el así llamado lema de intervalos encajados. En la teoría de espacios métricos desempeña un papel semejante el siguiente teorema, llamado principio de bolas encajadas.

TEOREMA 1. *Para que un espacio métrico R sea completo es necesario y suficiente que cualquier sucesión de bolas cerradas de este espacio, encajadas unas en otras y cuyos radios tienden a cero, tenga una intersección no vacía.*

DEMOSTRACION. NECESIDAD. Supongamos que el espacio R es completo y sea B_1, B_2, B_3, ... una sucesión de bolas cerradas, encajadas unas en otras. Sea r_n el radio y x_n, el centro de la bola B_n. La sucesión de centros $\{x_n\}$ es fundamental, ya que $\rho(x_n, x_m)<r_n$ para $m>n$ y $r_n\to 0$ para $n\to\infty$. Puesto que R es completo, existe $\lim\limits_{n\to\infty}x_n$. Tomemos

$$x=\lim_{n\to\infty}x_n;$$

entonces, $x\in\bigcap\limits_{n}B_n$. En efecto, la bola B_n contiene todos los puntos de la sucesión $\{x_k\}$, excepto, posiblemente, los puntos

$x_1, x_2, \ldots, x_{n-1}$. Por consiguiente, x es un punto de acumulación de toda bola B_n. Pero B_n es un conjunto cerrado y por eso $x \in B_n$ para todo n.

SUFICIENCIA. Sea $\{x_n\}$ una sucesión fundamental. Demostremos que tiene un límite. Puesto que es fundamental, podemos encontrar un punto x_{n_1} de ella tal que $\rho(x_n, x_{n_1}) < \frac{1}{2}$ para todo $n \geqslant n_1$. Consideremos la bola cerrada de radio 1 con centro en x_{n_1} y denotémosla mediante B_1. Escojamos luego un punto x_{n_2} de $\{x_n\}$ tal que sea $n_2 > n_1$ y $\rho(x_n, x_{n_2}) < \frac{1}{2^2}$ para todo $n \geqslant n_2$. Consideremos la bola B_2 de radio $\frac{1}{2}$ y centro en x_{n_2}. En general, si los puntos $x_{n_1}, x_{n_2}, \ldots, x_{n_k}$ ya se han escogido $(n_1 < n_2 < \ldots < n_k)$, escogeremos el punto $x_{n_{k+1}}$ de manera que sea $n_{k+1} > n_k$ y $\rho(x_n, x_{n_{k+1}}) < \frac{1}{2^{k+1}}$ para todo $n \geqslant n_{k+1}$ y lo envolveremos en una bola cerrada B_{k+1} de radio $\frac{1}{2^k}$. Continuando este proceso, obtendremos una sucesión de bolas cerradas B_k, encajadas unas en otras; la bola B_k es de radio $\frac{1}{2^{k-1}}$. Esta sucesión de bolas tiene, por hipótesis, un punto común; denotémoslo x. Es obvio que este punto x es el límite de la sucesión $\{x_{n_k}\}$. Pero una sucesión fundamental, que contiene una sucesión parcial convergente a x, converge al mismo límite. Por consiguiente, $x = \lim_{n \to \infty} x_n$. El teorema queda demostrado.

EJERCICIOS. 1. Dado un conjunto M de un espacio métrico el número

$$d(M) = \sup_{x,\, y \in M} \rho(x, y)$$

se llama diámetro del conjunto M. Demuéstrese que en un espacio métrico completo toda sucesión de conjuntos cerrados no vacíos, encajados unos en otros y cuyos diámetros tienden a cero, tiene una intersección no vacía.

2. Dese un ejemplo de un espacio métrico completo y de una sucesión de bolas cerradas de este espacio, encajadas unas en otras, que tiene una intersección *vacía*.

3. Demuéstrese que un subespacio de un espacio métrico completo R es completo si, y sólo si, es cerrado en R.

3°. Teorema de Baire. El siguiente teorema desempeña un papel fundamental en la teoría de los espacios métricos completos.

TEOREMA DE BAIRE. *Un espacio métrico completo R no puede representarse como la unión de un número numerable de conjuntos nunca densos*[1].

[1] Se dice que el conjunto M es nunca denso en R, cuando cada bola $B \subset R$ contiene otra bola B', que no tiene con M ningún punto común.

DEMOSTRACION. Supongamos lo contrario. Sea $R=\bigcup_{n=1}^{\infty} M_n$, donde cada uno de los conjuntos M_n es nunca denso. Sea S_0 una bola cerrada de radio 1. Puesto que M_1 no es denso en S_0, ya que es nunca denso, existe una bola cerrada S_1 de radio menor que $\frac{1}{2}$ y tal que $S_1 \subset S_0$ y $S_1 \cap M_1 = \varnothing$. El conjunto M_2 no es denso en S_1 y por eso la bola S_1 contiene una bola cerrada S_2 de radio menor que $\frac{1}{3}$ tal que $S_2 \cap M_2 = \varnothing$, etc. De esta forma obtenemos una sucesión de bolas cerradas $\{S_n\}$, encajadas unas en otras, cuyos radios tienden a cero, siendo, además, $S_n \cap M_n = \varnothing$. En virtud del teorema 1 del punto anterior, la intersección $\bigcap_{n=1}^{\infty} S_n$ contiene un punto x. De acuerdo con el procedimiento seguido, este punto no puede pertenecer a ninguno de los conjuntos M_n y, por consiguiente, $x \overline{\in} \bigcup_n M_n$, es decir, $R \neq \bigcup_n M_n$, lo que está en contradicción con la suposición hecha.

En particular, *todo espacio métrico completo sin puntos aislados es innumerable*. En efecto, en este espacio todo punto es nunca denso.

4°. Completación de un espacio. Si el espacio R no es completo, siempre puede ser incluido de cierta manera (y de hecho de una manera única) en un espacio completo.

DEFINICION 2. Sea R un espacio métrico. Un espacio métrico completo R^* se llama *completación* del espacio R, si:

1) R es un subespacio del espacio R^*;

2) R es siempre denso en R^*, es decir, $[R]=R^*$. (Aquí $[R]$ significa, claro está, la adherencia del espacio R en R^*).

Por ejemplo, el espacio de todos los números reales es completación del espacio de los números racionales.

TEOREMA 2. *Todo espacio métrico R posee una completación y esta completación es única, a menos de una aplicación isométrica que transforma los puntos de R en sí mismos.*

DEMOSTRACION. Comencemos por la *unicidad*. Debemos comprobar que si R^* y R^{**} son dos completaciones del espacio R, existe una aplicación biunívoca φ del espacio R^* sobre R^{**} tal que:

1) $\varphi(x)=x$ para todo $x \in R$;

2) si $x^* \leftrightarrow x^{**}$ e $y^* \leftrightarrow y^{**}$, entonces $\rho_1(x^*, y^*) = \rho_2(x^{**}, y^{**})$, donde ρ_1 es la distancia en R^* y ρ_2, la distancia en R^{**}.

La aplicación φ se construye del siguiente modo. Sea x^* un punto arbitrario de R^*. En este caso, de acuerdo con la definición de completación, existe una sucesión $\{x_n\}$ de puntos de R que converge a x^*. Los puntos $\{x_n\}$ pertenecen también a R^{**}. Puesto que R^{**} es completo, $\{x_n\}$ converge en R^{**} a un punto x^{**}. Está claro que x^{**} no depende de cómo se escoge la sucesión $\{x_n\}$, convergente al punto x^*. Tomemos $\varphi(x^*) = x^{**}$. La aplicación φ es la aplicación isométrica que necesitamos.

En efecto, está claro que $\varphi(x) = x$ para todo $x \in R$. Además, supongamos que

$$\{x_n\} \longrightarrow x^* \text{ en } R^* \text{ y } \{x_n\} \longrightarrow x^{**} \text{ en } R^{**},$$
$$\{y_n\} \longrightarrow y^* \text{ en } R^* \text{ y } \{y_n\} \longrightarrow y^{**} \text{ en } R^{**};$$

entonces,

$$\rho_1(x^*, y^*) = \lim_{n \to \infty} \rho_1(x_n, y_n) = \lim_{n \to \infty} \rho(x_n, y_n)$$

y al mismo tiempo

$$\rho_2(x^{**}, y^{**}) = \lim_{n \to \infty} \rho_2(x_n, y_n) = \lim_{n \to \infty} \rho(x_n, y_n).$$

Por consiguiente,

$$\rho_1(x^*, y^*) = \rho_2(x^{**}, y^{**}).$$

Demostremos ahora la *existencia* de la completación. La idea de esta demostración es la misma que en la teoría de Cantor de los números reales. La situación ahora es incluso más simple que en la teoría de los números reales, ya que allí era necesario, además, definir todas las operaciones aritméticas para los antes nuevos introducidos, es decir, para los números irracionales.

Sea R un espacio métrico arbitrario. Dos sucesiones fundamentales $\{x_n\}$ y $\{x'_n\}$ de R se llamarán equivalentes (denotación $\{x_n\} \sim \{x'_n\}$), cuando $\lim\limits_{n \to \infty} \rho(x_n, x'_n) = 0$. Esta relación de equivalencia es reflexiva, simétrica y transitiva. De aquí se desprende que todas las sucesiones fundamentales, que pueden formarse por medio de los puntos del espacio R, se dividen en clases de sucesiones equivalentes entre sí. Definamos ahora el espacio R^* del siguiente modo. Tomemos como puntos de este espacio todas las clases de sucesiones equivalentes entre sí y determinemos la distancia entre ellos como sigue. Sean x^* e y^* dos de estas clases. Escojamos en cada una de estas clases un representante, es decir, una sucesión fundamental $\{x_n\}$ e $\{y_n\}$. Pongamos

$$\rho(x^*, y^*) = \lim_{n \to \infty} (\rho(x_n, y_n). \tag{3}$$

Demostremos que es correcto definir así la distancia, es decir, que el límite (3) existe y no depende de cómo se escogen los representantes $\{x_n\} \in x^*$ e $\{y_n\} \in y^*$.

Puesto que las sucesiones $\{x_n\}$ e $\{y_n\}$ son fundamentales, obtenemos, por medio del axioma triangular, que para todos n y m suficientemente grandes

$$\begin{aligned}|\rho(x_n, y_n)-\rho(x_m, y_m)| &= \\ &= |\rho(x_n, y_n)-\rho(x_n y_m)+\rho(x_n, y_m)-\rho(x_m, y_m)| \leqslant \\ \leqslant |\rho(x_n, y_n)-\rho(x_n, y_m)|&+|\rho(x_n, y_m)-\rho(x_m, y_m)| \leqslant \\ &\leqslant \rho(y_n, y_m)+\rho(x_n, x_m) < \frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon. \quad (4)\end{aligned}$$

Por consiguiente, la sucesión de números reales $s_n=\rho(x_n, y_n)$ verifica el criterio de Cauchy y, por lo tanto, tiene un límite. Este límite no depende de la selección de $\{x_n\} \in x^*$ e $\{y_n\} \in y^*$. En efecto, sea

$$\{x_n\}, \ \{x'_n\} \in x^* \ \text{e} \ \{y_n\}, \ \{y'_n\} \in y^*.$$

Obtenemos por un razonamiento, análogo a (4),

$$|\rho(x_n, y_n)-\rho(x'_n, y'_n)| \leqslant \rho(x_n, x'_n)+\rho(y_n, y'_n).$$

Puesto que $\{x_n\} \sim \{x'_n\}$ e $\{y_n\} \sim \{y'_n\}$, de aquí se sigue que

$$\lim_{n\to\infty} \rho(x_n, y_n) = \lim_{n\to\infty} \rho(x'_n, y'_n).$$

Demostremos ahora que R^* se cumplen los axiomas de espacio métrico.

El axioma 1 se desprende inmediatamente de la definición de equivalencia de sucesiones fundamentales.

El axioma 2 es obvio.

Comprobemos ahora el axioma triangular. En el espacio inicial R este axioma se cumple y por eso

$$\rho(x_n, z_n) \leqslant \rho(x_n, y_n)+\rho(y_n, z_n).$$

Pasando al límite para $n \to \infty$, obtenemos

$$\lim_{n\to\infty} \rho(x_n, z_n) \leqslant \lim_{n\to\infty} \rho(x_n, y_n) + \lim_{n\to\infty} \rho(y_n, z_n$$

es decir,

$$\rho(x, z) \leqslant \rho(x, y)+\rho(y, z).$$

Demostremos ahora que R^* es una completación del espacio R.

A cada punto $x \in R$ le corresponde una clase de sucesiones fundamentales equivalentes entre sí, a saber, la totalidad de las sucesiones convergentes al punto x. Además, si

$$x = \lim_{n\to\infty} x_n \text{ e } y = \lim_{n\to\infty} y_n, \text{ tenemos } \rho(x, y) = \lim_{n\to\infty} \rho(x_n, y_n).$$

Por consiguiente, obtendremos una aplicación isométrica de R en el espacio R^*, si a todo punto $x \in R$ le ponemos en correspondencias la clase de sucesiones fundamentales convergentes a x.

En adelante podemos identificar el espacio R con su imagen en R^* y considerar R como un subconjunto de R^*.

Demostremos ahora que R es siempre denso en R^*. En efecto, sean x^* un punto de R^* y $\varepsilon > 0$ un número arbitrario. Tomemos en x^* un representante, esto es, una sucesión fundamental $\{x_n\}$. Sea N tal que $\rho(x_n, x_m) < \varepsilon$ para todo $n, m > N$. En este caso

$$\rho(x_n, x^*) = \lim_{m\to\infty} \rho(x_n, x_m) \leqslant \varepsilon$$

para $n \geqslant N$, es decir, una vecindad arbitraria del punto x^* contiene un punto de R. Por consiguiente, la adherencia de R en R^* es todo el espacio R^*.

Resta demostrar que el espacio R^* es completo. Observemos, ante todo, que R^* ha sido construido de manera que toda sucesión fundamental

$$x_1, x_2, \ldots, x_n, \ldots \tag{5}$$

compuesta de puntos pertenecientes a R, converge en R^* a un punto determinado, a saber, al punto $x^* \in R^*$ determinado por la sucesión (5). Además, puesto que R es denso en R^*, para toda sucesión fundamental $x_1^*, x_2^*, \ldots, x_n^*, \ldots$ de puntos de R^* se puede construir una sucesión equivalente $x_1, x_2, \ldots, x_n, \ldots$, compuesta de puntos, pertenecientes a R. Para ello es suficiente escoger a título de x_n cualquier punto de R tal que $\rho(x_n, x_n^*) < \frac{1}{n}$. La sucesión $\{x_n\}$ así obtenida es fundamental y, de acuerdo con lo demostrado, converge a un punto $x^* \in R^*$. Pero esto significa que la sucesión $\{x_n^*\}$ también converge a x^*. El teorema queda demostrado completamente.

§. 4 PRINCIPIO DE APLIGACIONES CONTRAÍDAS Y SUS APLIGACIONES

1°. Principio de aplicaciones contraídas. A título de aplicación del concepto de complitud consideremos el así llamado *principio de aplicaciones contraídas*. Representa un instrumento útil para la demostración de diferentes teoremas de existencia y unicidad (por ejemplo, en la teoría de ecuaciones diferenciales).

Sea R un espacio métrico. La aplicación A del espacio R en sí mismo se llama *contraída*, cuando existe un número $\alpha < 1$ tal que para cualesquiera dos puntos x, $y \in R$ se verifica la desigualdad

$$\rho(Ax, Ay) \leqslant \alpha\rho(x, y). \tag{1}$$

Toda aplicación contraída es continua. En efecto, si $x_n \longrightarrow x$, tenemos, de acuerdo con (1), $Ax_n \longrightarrow Ax$.

TEOREMA 1 (principio de aplicaciones contraídas). *Toda aplicación contraída, definida en un espacio métrico completo R, tiene un punto fijo, y sólo uno, (es decir, la ecuación $Ax = x$ tiene una solución, y solamente una).*

DEMOSTRACION. Sea x_0 un punto arbitrario de R. Pongamos $x_1 = Ax_0$, $x_2 = Ax_1 = A^2x_0$, etc.; en general, $x_n = Ax_{n-1} = A^n x_0$.

Demostremos que la sucesión $\{x_n\}$ es fundamental. En efecto, suponiendo, para concretar, que $m \geqslant n$, tenemos

$$\rho(x_n, x_m) = \rho(A^n x_0, A^m x_0) \leqslant \alpha^n \rho(x_0, x_{m-n}) \leqslant$$
$$\leqslant \alpha^n \{\rho(x_0, x_1) + \rho(x_1, x_2) + \ldots + \rho(x_{m-n-1}, x_{m-n})\}$$
$$\leqslant \alpha^n \rho(x_0, x_1)\{1 + \alpha + \alpha^2 + \ldots + \alpha^{m-n-1}\} \leqslant \alpha^n \rho(x_0, x_1) \frac{1}{1-\alpha}.$$

Puesto que $\alpha < 1$, esta magnitud resulta tan pequeña como se quiera, siempre que n sea suficientemente grande.

Debido a la complitud de R, la sucesión $\{x_n\}$, que es fundamental, tiene límite. Pongamos

$$x = \lim_{n \to \infty} x_n.$$

Pero la aplicación A es continua; por eso

$$Ax = A \lim_{n \to \infty} x_n = \lim_{n \to \infty} Ax_n = \lim_{n \to \infty} x_{n+1} = x.$$

Por consiguiente, queda demostrada la *existencia* del punto fijo. Probemos que es *único*. Si

$$Ax = x, \quad Ay = y,$$

la desigualdad (1) nos da

$$\rho(x, y) < \alpha\rho(x, y);$$

puesto que $\alpha < 1$, de aquí se deduce que

$$\rho(x, y) = 0, \text{ es decir, que } x = y.$$

EJERCICIO. Demuéstrese que para la existencia de un punto fijo no es suficiente que se cumpla la condición $\rho(Ax, Ay) < \rho(x, y)$ para todo $x \neq y$.

2°. Aplicaciones elementales del principio de aplicaciones contraídas. El principio de aplicaciones contraídas puede emplearse para demostrar teoremas de existencia y unicidad de soluciones de diferentes ecuaciones. Junto con la demostración de la existencia y de la unicidad de la solución de la ecuación $Ax=x$, el principio de aplicaciones contraídas ofrece un método aproximado para buscar esta solución (el método de aproximaciones sucesivas). Veamos los siguientes ejemplos elementales.

1. Sea f una función, definida en el segmento $[a, b]$, que verifica la condición de Lipschitz

$$|f(x_2)-f(x_1)| \leqslant K|x_2-x_1|$$

con una constante $K<1$ y que transforma el segmento $[a, b]$ en sí mismo. En este caso, f es una aplicación contraída y, de acuerdo con el teorema demostrado, la sucesión x_0, $x_1=f(x_0)$, $x_2=f(x_1)$, ... converge a la única raíz de la ecuación $x=f(x)$.

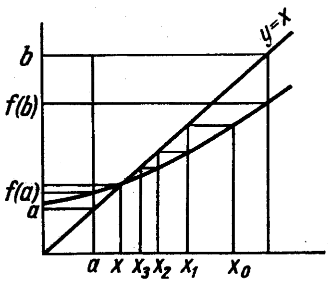


FIG. 10

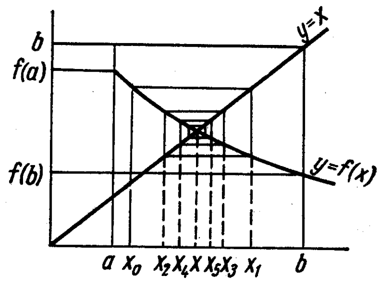


FIG. 11

En particular, la aplicación será contraída, si la función tiene en el segmento $[a, b]$ una derivada $f'(x)$ tal que

$$|f'(x)| \leqslant K \leqslant 1.$$

En las figuras 10 y 11 están representadas las aproximaciones sucesivas para el caso $0<f'(x)<1$ y para el caso $-1<f'(x)<0$, respectivamente.

Supongamos ahora que tenemos una ecuación de tipo $F(x)=0$, que $F(a)<0$, $F(b)>0$ y $0<K_1 \leqslant F'(x) \leqslant K_2$ sobre $[a, b]$. Para encontrar la raíz, introduzcamos la función $f(x)=x-\lambda F(x)$ y busquemos la solución de la ecuación $x=f(x)$, equivalente a la ecuación $F(x)=0$. Puesto que $f'(x)=1-\lambda F'(x)$, tenemos

$1-\lambda K_2 \leqslant f'(x) \leqslant 1-\lambda K_1$ y no costará ningún esfuerzo escoger el número λ de manera que se pueda emplear el método de aproximaciones sucesivas. La idea aquí expuesta es un método extendido que se emplea para buscar la raíz.

2. Consideremos la aplicación A del espacio de n dimensiones en sí mismo definida por el sistema de ecuaciones lineales

$$y_i=\sum_{j=1}^{n} a_{ij}x_j+b_i \quad (i=1,\ 2,\ \ldots,\ n).$$

Si A es una aplicación contraída, podemos aplicar el método de aproximaciones sucesivas para resolver la ecuación $x=Ax$.

¿Bajo qué condiciones la aplicación A será contraída? La respuesta depende de cómo se escoja la métrica en el espacio. Consideremos tres variantes.

a) El espacio R_0^n, es decir, $\rho(x,\ y)=\max\limits_{1\leqslant i\leqslant n}|x_i-y_i|$;

$$\rho(y',\ y'')=\max_i|y_i'-y_i''|=\max_i\left|\sum_j a_{ij}\ (x_j'-x_j'')\right|\leqslant$$
$$\leqslant\max_i\sum_j|a_{ij}|\,|x_j'-x_j''|\leqslant\max_i\sum_j|a_{ij}|\max|x_j'-x_j''|=$$
$$=\left(\max_i\sum_j|a_{ij}|\right)\rho(x',\ x'').$$

De aquí la condición de contracción

$$\sum_{j=1}^{n}|a_{ij}|\leqslant\alpha<1,\quad j=1,\ \ldots,\ n. \tag{2}$$

b) El espacio R_1^n, es decir, $\rho(x,\ y)=\sum\limits_{i=1}^{n}|x_i-y_i|$;

$$\rho(y',\ y'')=\sum_i|y_i'-y_i''|=\sum_i\left|\sum_j a_{ij}(x_j'-x_j'')\right|\leqslant$$
$$\leqslant\sum_i\sum_j|a_{ij}|\,|x_j'-x_j''|\leqslant\left(\max_j\sum_i|a_{ij}|\right)\rho(x',\ x'').$$

De aquí la condición de contracción

$$\sum_i|a_{ij}|\leqslant\alpha<1,\ j=1,\ \ldots,\ n. \tag{3}$$

c) El espacio R^n, es decir $(x,\ y)=\sqrt{\sum\limits_{i=1}^{n}(x_1-y_i)^2}$. En virtud de la desigualdad de Cauchy—Buniakovski, tenemos

$$\rho^2(y',\ y'')=\sum_i\left|\left(\sum_j a_{ij}(x_j'-x_j'')\right)^2\leqslant\left(\sum_i\sum_j a_{ij}^2\right)\rho^2(x',\ x'').$$

De aquí la condición de contracción

$$\sum_i \sum_j a_{ij}^2 \leqslant \alpha < 1. \tag{4}$$

Por consiguiente, si se verifica al menos una de las condiciones (2), (3) ó (4), existe un punto y sólo uno, $x=(x_1, x_2, \ldots, x_n)$ tal que $x_i = \sum_{j=1}^{n} a_{ij}x_j + b_i$ [1]; además las aproximaciones sucesivas de esta solución tienen la forma

$$x^{(0)} = (x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)},$$
$$x^{(1)} = (x_1^{(1)}, x_2^{(1)}, \ldots, x_n^{(1)},$$
$$\ldots\ldots\ldots\ldots$$
$$x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)},$$

donde

$$x_i^{(k)} = \sum_{j=1}^{n} a_{ij}x_j^{(k-1)} + b_i$$

y $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$ puede ser un punto cualquiera de R^n.

Cada una de las condiciones (2), (3) y (4) *es suficiente* para que la aplicación $y = Ax$ sea contraída. Respecto a la condición (2) se puede demostrar que es también necesaria para que la aplicación $y = Ax$ sea contraída (en el sentido de la métrica *a*)).

Ninguna de las condiciones (2), (3) y (4) es necesaria para que se pueda aplicar el método de aproximaciones sucesivas. Se pueden dar ejemplos, cuando una de estas condiciones se verifica, pero las otras dos no se cumplen.

Si $|a_{ij}| < \frac{1}{n}$ (en este caso se cumplen las tres condiciones), es aplicable el método de aproximaciones sucesivas.

Si $|a_{ij}| = \frac{1}{n}$ (en este caso las tres sumas son iguales a 1), es fácil ver que el método de aproximaciones sucesivas no puede aplicarse.

3°. Teoremas de existencia y unicidad de ecuaciones diferenciales. En el punto anterior hemos señalado algunos ejemplos

---

[1] En particular, de cualquiera de las condiciones (2), (3) ó (4) se deduce que

$$\begin{vmatrix} a_{11}-1 & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22}-1 & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{22} & \ldots & a_{nn}-1 \end{vmatrix} \neq 0.$$

elementales de aplicación del principio de aplicaciones contraídas en los espacios de una y n dimensiones. Sin embargo, las aplicaciones más importantes para el Análisis del principio de aplicaciones contraídas se refieren a los espacios funcionales de infinitas dimensiones. Ahora veremos cómo mediante este principio se puede obtener teoremas de existencia y unicidad de la solución de ciertos tipos de ecuaciones diferenciales e integrales.

1. Supongamos que se tiene una ecuación diferencial

$$\frac{dy}{dx}=f(x,\ y) \tag{5}$$

con la condición inicial

$$y(x_0)=y_0 \tag{6}$$

y que la función f, definida y continua en un recinto plano G que contiene el punto $(x_0,\ y_0)$ verifica la condición de Lipschitz respecto a y:

$$|f(x,\ y_1)-f(x,\ y_2)|\leqslant M|y_1-y_2|.$$

Demostremos que entonces existe en un segmento $|x-x_0|\leqslant d$ una solución, y sólo una, $y=\varphi(x)$ de la ecuación (1), que verifica la condición inicial (6) (teorema de Picard).

La ecuación (5) junto con la condición inicial (6) es equivalente a la ecuación integral

$$\varphi(x)=y_0+\int_{x_0}^{x}f(t,\ \varphi(t))\,dt. \tag{7}$$

Debido a la continuidad de la función f, tenemos $|f(x,\ y)|\leqslant K$ para un recinto $G'\subset G$ que contiene el punto $(x_0,\ y_0)$. Escogemos ahora el número $d>0$ de manera que se cumplan las condiciones:

1) $(x,\ y)\in G'$, siempre que $|x-x_0|\leqslant d$, $|y-y_0|\leqslant Kd$;

2) $Md<1$.

Designemos mediante C^* el espacio de funciones continuas φ, definidas sobre el segmento $|x-x_0|\leqslant d$ y tales que $|\varphi(x)-$ $-y_0|\leqslant Kd$, con la métrica $\rho(\varphi_1,\ \varphi_2)=\max\limits_{x}|\varphi(x_1)-\varphi_2(x)|$.

El espacio C^* es completo, ya que representa un subespacio cerrado del espacio completo de todas las funciones continuas sobre $[x_0-d,\ x_0+d]$. Consideremos la aplicación $\psi=A\varphi$ definida mediante la fórmula

$$\psi(x)=y_0+\int_{x_0}^{x}f(t,\ \varphi(t))\,dt,$$

donde $|x-x_0|\leqslant d$. Esta aplicación transforma el espacio completo C^* en sí mismo y es contraída en él. En efecto, sea $\varphi \in C^*$ y $|x-x_0| \leqslant d$. En este caso,

$$|\psi(x)-y_0| = \left| \int_{x_0}^{x} f(t, \varphi(t))\,dt \right| \leqslant Kd$$

y, por consiguiente, $A(C^*) \subset C^*$. Además,

$$|\psi_1(x)-\psi_2(x)| \leqslant \int_{x_0}^{x} [f(t, \varphi_1(t))-f(t, \varphi_2(t))|\,dt \leqslant$$
$$\leqslant Md \max_x |\varphi_1(x)-\varphi_2(x)|.$$

Puesto que $Md < 1$, la aplicación A es contraída.

De aquí se deduce que la ecuación $\varphi = A\varphi$ (es decir, la ecuación (7)) tiene una solución, y sólo una, en el espacio C^*.

2. Supongamos que se tiene un sistema de ecuaciones diferenciales

$$\varphi_i'(x) = f_i(x, \varphi_1(x), \ldots, \varphi_n(x)), \quad i = 1, 2, \ldots, n, \tag{8}$$

con las condiciones iniciales

$$\varphi_i(x_0) = y_{0i}, \quad i = 1, 2, \ldots, n \tag{9}$$

y que las funciones f_i, definidas y continuas en un recinto G del espacio R^{n+1} que contiene el punto $(x_0, y_0, \ldots, y_{0n})$, verifican la condición de Lipschitz

$$|f_i(x, y_1^{(1)}, \ldots, y_n^{(1)})-f(x, y_1^{(2)}, \ldots, y_n^{(2)})| \leqslant$$
$$\leqslant M \max_{1 \leqslant i \leqslant n} |y_i^{(1)}-y_i^{(2)}|.$$

Demostremos que entonces existe en un segmento $|x-x_0| \leqslant d$ una solución, y sólo una, del problema inicial (8) y (9), es decir, un sistema, y sólo uno, de funciones φ_i que verifican las ecuaciones (8) y las condiciones iniciales (9).

El sistema (8) junto con las condiciones iniciales (9) es equivalente al sistema de ecuaciones integrales

$$\varphi_i(x) = y_{0i} + \int_{x_0}^{x} f_i(t, \varphi_1(t), \ldots, \varphi_n(t))\,dt, \quad i = 1, \ldots, n. \tag{10}$$

Debido a la continuidad, las funciones f_i son acotadas en un recinto $G' \subset G$ que contiene el punto $(x_0, y_0, \ldots, y_{0n})$, es decir, existe una constante K tal que $|f_i(x, y_1, \ldots, y_n)| \leqslant K$.

Escojamos ahora el número $d>0$ de manera que se cumplan las condiciones:

1) $(x, y_1, \ldots, y_n) \in G'$, siempre que $|x-x_0| \leqslant d$, $|y_i - y_{0i}| \leqslant Kd$;

2) $Md<1$.

Consideremos ahora el espacio C_n^*, cuyos elementos son los sistemas ordenados $\overline{\varphi}=(\varphi_1, \ldots, \varphi_n)$ de n funciones, definidas y continuas para todo x, siempre que $|x-x_0| \leqslant d$, y tales que $|\varphi_i(x)-y_{0i}| \leqslant Kd$, y cuya métrica está definida por

$$\rho(\overline{\varphi}, \overline{\psi}) = \max_{x,\,i} |\varphi_i(x)-\psi_i(x)|.$$

Este espacio es completo. La aplicación $\overline{\psi}=A\overline{\varphi}$, dada mediante el sistema de ecuaciones integrales

$$\psi_i(x)=y_{0i}+\int_{x_0}^{x} f_i(t, \varphi_1(t), \ldots, \varphi_n(t)\, dt,$$

es una aplicación contraída del espacio completo C_n^* en sí mismo. En efecto,

$$\psi_i^{(1)}(x)-\psi_i^{(2)}(x) =$$
$$= \int_{x_0}^{x} [f_i(t, \varphi_1^{(1)}(t), \ldots, \varphi_n^{(1)}(t))-f_i(t, \varphi_1^{(2)}(t), \ldots, \varphi_n^{(2)}(t))]\, dt$$

y, por consiguiente,

$$\max_{x,\,i} |\psi_i^{(1)}(x)-\psi_i^{(2)}(x)| \leqslant Md \max |\varphi_i^{(1)}(x)-\varphi_i^{(2)}(x)|.$$

La aplicación A es contraída, ya que $Md<1$.

De aquí se deduce que la ecuación $\overline{\varphi}=A\overline{\varphi}$ tiene una solución, y sólo una, en el espacio C_n^*.

4°. Aplicación del principio de aplicaciónes contraídas a ecuaciones integrales. Empleamos ahora el método de aplicaciones contraídas para demostrar la existencia y unicidad de la solucíón de la ecuación integral lineal no homogénea de Fredholm de segunda especie, es decir de la ecuación

$$f(x)=\lambda \int_a^b K(x, y)\, f(y)\, dy+\varphi(x), \tag{11}$$

donde K (llamada *núcleo*) y φ son funciones dadas; f, la función incógnita y λ, un parámetro arbitrario.

Como veremos, nuestro método es aplicable solamente para valores suficientemente pequeños del parámetro λ.

Supongamos que $K(x, y)$ y $\varphi(x)$ son continuas cuando $a \leqslant x \leqslant b$, $a \leqslant y \leqslant b$, y, por consiguiente, $|K(x, y)| \leqslant M$. Consideremos la aplicación $g = Af$ del espacio completo $C_{[a, b]}$ en sí mismo, definida por la fórmula

$$g(x) = \lambda \int_a^b K(x, y) f(y) dy + \varphi(x).$$

Tenemos

$$\rho(g_1, g_2) = \max |g_1(x) - g_2(x)| \leqslant$$
$$\leqslant |\lambda| M(b-a) \max |f_1(x) - f_2(x)|.$$

Por consiguiente, la aplicación A es contraída, si $|\lambda| < \frac{1}{M(b-a)}$.

Del principio de aplicaciones contraídas deducimos que para todo λ, tal que $|\lambda| < \frac{1}{M(b-a)}$, la ecuación de Fredholm tiene una solución continua única. Las aproximaciones sucesivas $f_0, f_1, \ldots, f_n, \ldots$ de esta solución tienen la forma

$$f_n(x) = \lambda \int_a^b K(x, y) f_{n-1}(y) dy + \varphi(x),$$

donde $f_0(x)$ es una función continua cualquiera.

El principio de aplicaciones contraídas puede ser también empleado en el caso de una ecuación integral no lineal de tipo

$$f(x) = \lambda \int_a^b K(x, y; f(y)) dy + \varphi(x), \tag{12}$$

donde K y φ son continuas y, además, K verifica la condición de Lipschitz respecto a su argumento «funcional»:

$$|K(x, y; z_1) - K(x, y; z_2)| \leqslant M |z_1 - z_2|.$$

En este caso, para la aplicación $g = Af$ del espacio completo $C_{[a,b]}$ en sí mismo, definida por

$$g(x) = \lambda \int_a^b K(x, y; f(y)) dy + \varphi(x), \tag{13}$$

obtenemos la desigualdad

$$\max |g_1(x) - g_2(x)| \leqslant |\lambda| M(b-a) \max |f_1(x) - f_2(x)|,$$

donde $g_1 = Af_1$, $g_2 = Af_2$. Por consiguiente, la aplicación A será contraída, siempre que $|\lambda| < \frac{1}{M(b-a)}$.

Consideremos finalmente la ecuación integral del tipo Volterra

$$f(x)=\lambda\int_a^x K(x, y)\, f(y)\,dy+\varphi(x). \qquad (14)$$

En diferencia de las ecuaciones de Fredholm, el extremo superior de la integral es aquí la variable x. Formalmente esta ecuación puede considerarse como caso particular de la ecuación de Fredholm, completando la definición de la función K mediante la igualdad

$$K(x, y)=0 \text{ para } y>x.$$

Sin embargo, en el caso de la ecuación integral de Fredholm hemos tenido que limitarnos a pequeños valores del parámetro λ, mientras que en el caso de la ecuación de Volterra el principio de aplicaciones contraídas (y el método de aproximaciones sucesivas) puede emplearse para todos los valores de λ. Hablando con más precisión, se trata de la siguiente generalización del principio de aplicaciones contraídas:

Sea A una aplicación continua del espacio métrico completo R en sí mismo tal que la aplicación A^n es contraída para algún n; en este caso, la ecuación

$$Ax=x$$

tiene una solución, y sólo una.

En efecto, tomemos un punto arbitrario $x_0 \in R$ y consideremos la sucesión $A^{kn}x_0$ $(k=0, 1, 2, \ldots)$. Repitiendo la parte correspondientemente de la demostración del principio de aplicaciones contraídas, podemos probar que esta sucesión converge. Pongamos

$$x=\lim_{k\to\infty} A^{kn}x_0.$$

Afirmamos que

$$Ax=x.$$

Efectivamente, debido a la continuidad de A, tenemos

$$Ax=\lim_{k\to\infty} A^{kn}Ax_0.$$

Puesto que la aplicación A es contraída,

$$\rho(A^{kn}Ax_0,\ A^{kn}x_0)\leqslant\alpha\rho(A^{(k-1)n}Ax_0 \leqslant \ldots\ \alpha^k\rho(Ax_0,\ x_0).$$

Por consiguiente,

$$\lim_{k\to\infty}\rho(A^{kn}Ax_0,\ A^{kn}x_0)=0,$$

es decir, $Ax=x$.

Demostremos la unicidad del punto fijo. Como todo punto fijo respecto a A también será fijo respecto a la aplicación A^n y ésta es, a su vez, contraída, este punto fijo puede ser solamente único.

Probemos ahora que cierta potencia de la aplicación

$$Af(x) = \lambda \int_a^x K(x, y) f(y)\, dy + \varphi(x)$$

es una aplicación contraída. Sean f_1 y f_2 dos funciones continuas sobre el segmento $[a, b]$. Entonces

$$|Af_1(x) - Af_2(x)| = \left|\lambda \int_a^x K(x, y)(f_1(y) - f_2(y))\, dy\right| \leqslant$$
$$\leqslant |\lambda| M(x-a) \max |f_1(x) - f_2(x)|.$$

Aquí

$$M = \max |K(x, y)|.$$

De aquí

$$|A^2f_1(x) - A^2f_2(x)| \leqslant |\lambda|^2 M^2 \frac{(x-a)^2}{2} \max |f_1(x) - f_2(x)|$$

y, en general,

$$|A^nf_1(x) - A^nf_2(x)| \leqslant |\lambda|^n M^n \frac{(x-a)^n}{n!} m \leqslant |\lambda|^n M^n m \frac{(b-a)^n}{n!},$$

donde $m = \max |f_1(x) - f_2(x)|$.

Cualquiera que sea el valor λ, podemos escoger n tan grande que

$$\frac{|\lambda|^n M^n (b-a)^n}{n!} < 1,$$

es decir, la aplicación A^n será contraída para n suficientemente grandes. Por consiguiente, la ecuación de Volterra (14) tiene solución, y además única, cualquiera que sea λ.

§ 5. ESPACIOS TOPOLÓGICOS

1°. Definición y ejemplos de espacios topológicos. Hemos introducido los conceptos principales de la teoría de espacios métricos (punto de acumulación, punto de adherencia, adherencia de un conjunto, etc.) basándonos en el concepto de vecindad o, que de hecho es lo mismo, en el concepto de conjunto abierto. Estos últimos conceptos (vecindad y conjunto abierto) se definían, a su vez, mediante la métrica existente en el espacio considerado.

Podemos, sin embargo, escoger otro camino y, sin introducir métrica ninguna en el conjunto dado R, definir directamente en R, mediante axiomas, el sistema de conjuntos abiertos. Este camino conduce a los *espacios topológicos*; respecto a ellos los espacios métricos representan un caso, aunque muy importante, pero especial.

DEFINICION. Sea X un conjunto cualquiera. Se llama *topología* en X a todo sistema τ de subconjuntos G de X que verifica dos condiciones:

1°. El propio conjunto X y el conjunto vacío $\varnothing$ pertenecen a τ.

2°. La unión $\bigcup\limits_{\alpha} G_\alpha$ de un número cualquiera (finito o infinito) y la intersección $\bigcap\limits_{k=1}^{n} G_k$ de un número finito de conjuntos de τ pertenecen a τ.

El conjunto X junto con la topología τ, definida en él, es decir, el par (X, τ) se llama *espacio topológico*.

Los conjuntos, pertenecientes al sistema τ, se llaman *abiertos*.

Un espacio métrico está constituido por un conjunto de puntos y una métrica introducida en este conjunto; de la misma forma, un espacio topológico está constituido por un conjunto de puntos y una topología introducida en él. Por consiguiente, definir un espacio topológico significa definir un conjunto X y una topología τ en él, es decir, indicar aquellos subconjuntos que se consideran abiertos en X.

Está claro, que en un mismo conjunto X se puede introducir diferentes topologías, convirtiéndolo de esta forma en diferentes espacios topológicos. Sin embargo, denotaremos el espacio topológico, es decir, el par (X, τ), mediante *una* letra, digamos T. Llamaremos puntos a los elementos del espacio topológico.

Los conjuntos $T \setminus G$, complementarios a los abiertos, se llaman conjuntos *cerrados* del espacio topológico T. En virtud de las relaciones de dualidad (§ 1, capítulo I), de los axiomas 1° y 2° se deduce que:

1'. El conjunto vacío $\varnothing$ y todo el espacio T son cerrados.

2'. La intersección de un número cualquiera (finito o infinito) y la unión de un número finito de conjuntos cerrados son cerrados.

En todo espacio topológico se introducen, a partir de estas definiciones y de un modo natural, los conceptos de vecindad, puntos de adherencia, adherencia de conjuntos, etc.:

Se llama *vecindad* del punto $x \in T$ a todo conjunto abierto $G \subset T$ que contiene el punto x; un punto $x \in T$ se llama *punto de adherencia* del conjunto $M \subset T$, cuando toda vecindad del punto x contiene al menos un punto de M; un punto x se llama *punto de acumulación* del conjunto M, cuando toda vecindad del punto x contiene un número infinito de puntos de M. La totalidad de los puntos de adherencia del conjunto M se llama *adherencia* del conjunto M y se denota mediante el símbolo $[M]$. Es fácil ver (realícese la demostración) que los conjuntos cerrados, definidos más arriba como complementos de abiertos, *y solamente ellos, verifican la condición* $[M] = M$. *Al igual que en el caso de espacios métricos,* $[M]$ *es el menor conjunto cerrado que contiene a* M.

Ejemplos. 1. En virtud del teorema 3 del § 2, los conjuntos abiertos de cualquier espacio métrico verifican los axiomas 1° y 2° de la definición de un espacio topológico. Por consiguiente, todo espacio métrico es un espacio topológico.

2. Sea T un conjunto arbitrario. Consideremos como abiertos todos sus subconjuntos. Es obvio, entonces, que se cumplen los axiomas 1° y 2°, es decir, obtenemos efectivamente un espacio topológico. En él todos los conjuntos son a la vez abiertos y cerrados y, por eso, cada uno de ellos coincide con su adherencia. Esta topología trivial se observa, por ejemplo, en el espacio métrico, señalado en el ejemplo 1 del § 1.

3. Otro caso extremo se obtiene al considerar en un conjunto arbitrario X la topología compuesta sólo de dos conjuntos: el conjunto X y el conjunto vacío $\varnothing$. Aquí la adherencia de todo conjunto no vacío coincide con todo X. Este espacio topológico (que no representa, claro está, gran interés) puede ser llamado «espacio de puntos *pegados*».

4. Supongamos que T consta solamente de dos puntos a y b y que los conjuntos abiertos son todo el T, el conjunto vacío y el conjunto compuesto solamente del punto b. Los axiomas 1° y 2° se cumplen. En este espacio (que frecuentemente se llama *espacio de dos puntos conexos*) los conjuntos cerrados son: todo el T, el conjunto vacío y el punto a. La adherencia del conjunto puntual $\{b\}$ coincide con todo el T.

2°. Comparación de topologías. Supongamos que en un mismo conjunto X se tienen dos topologías τ_1 y τ_2 (con ello quedan definidos *dos* espacios topológicos: $T_1 = (X_1, \tau_1)$ y $T_2 = (X_1, \tau_2)$. Se dice que la topología τ_1 es *más fuerte* que la topología τ_2, cuando el sistema de conjuntos τ_2 está contenido en τ_1. De la topología τ_2 se dice en este caso que es más débil que τ_1.

En el conjunto de todas las topologías posibles del conjunto X se introduce, de manera natural, el orden parcial (la topología τ_2 precede a la τ_1, cuando es más débil que τ_1). Esta totalidad de topologías tiene el elemento maximal—la topología en la que son abiertos todos los subconjuntos (ejemplo 2)—y el elemento minimal—la topología en la que son abiertos solamente todo el X y $\varnothing$ (ejemplo 3).

TEOREMA 1. *La intersección* $\tau = \bigcap_{\alpha} \tau_\alpha$ *de un conjunto cualquiera de topologías de* X *es una topología de* X. *Esta topología* τ *es más débil que cualquiera de las topologías* τ_α.

DEMOSTRACION. Está claro que $\bigcap \tau_\alpha$ contiene a X y $\varnothing$. Además, cada sistema τ_α es cerrado respecto a cualesquiera sumas e intersecciones finitas; de aquí se deduce que $\tau = \bigcap_{\alpha} \tau_\alpha$ también posee esta propiedad.

COROLARIO. *Sea* B *un sistema arbitrario de subconjuntos del conjunto* X; *entonces existe una topología minimal de* X *que contiene a* B.

En efecto, existen topologías que contienen a B (por ejemplo, aquella en la que todo $A \subset X$ es abierto). La intersección de todas las topologías que contienen a B es la deseada. Esta topología minimal se llama topología generada por el sistema B y se denota con $\tau(B)$.

Sea X un conjunto arbitrario y A un subconjunto suyo. Llamaremos *traza* del sistema de conjuntos B sobre el subconjunto A al sistema B_A, compuesto de subconjuntos de tipo $A \cap B$, $B \in B$. Es fácil ver que la traza (sobre A) de la *topología* τ (definida en X) representa una topoloqia τ_A de A. Por consiguiente, todo subconjunto A de cualquier espacio topológico resulta ser un espacio topológico. El espacio topológico (A, τ_A) se llama *subespacio* del espacio topológico inicial (X, τ). Está claro que dos distintas topologías τ_1 y τ_2 de X pueden producir una misma topología en $A \subset X$.

3°. Sistemas determinantes de vecindades. Base. Axiomas de numerabilidad. Como hemos visto, para definir una topología en un espacio T hay que señalar en él el sistema de conjuntos abiertos. Sin embargo, en muchos casos concretos es más cómodo señalar no la totalidad de subconjuntos abiertos del espacio dado, sino un sistema determinante de subconjuntos que permite definir unívocamente la totalidad de los subconjuntos abiertos. Por ejemplo, en el espacio métrico hemos introducido primero el concepto de bola abierta ($=\varepsilon$-vecindad) y después hemos definido los conjuntos abiertos como aquellos que junto con cada punto contienen una vecindad suya en forma de bola. En otras palabras, en el espacio métrico son abiertos aquellos conjuntos y solamente aquellos que se pueden representar como la suma de bolas abiertas (en número finito o infinito). En particular son abiertos en la recta los conjuntos que se puede representar como la suma de un número de intervalos y solamente estos conjuntos. Estas consideraciones nos conducen al importante concepto de *base* de un espacio topológico.

DEFINICION. Una colección $\mathcal{G}$ de subconjuntos abiertos se llama *base* del espacio topológico T, cuando todo conjunto abierto de T se puede representar como suma de cierto número de conjuntos de $\mathcal{G}$.

Por ejemplo, la colección de todas las bolas abiertas (de todos los radios y centros posibles) constituye una base en un espacio métrico. En particular, el sistema de todos los intervalos es una base en la recta. Si se toman solamente los intervalos con extremos racionales, también constituyen una base en la recta, ya que mediante la suma de estos intervalos se puede representar cualquier intervalo y, por consiguiente, cualquier conjunto abierto sobre la recta.

De lo expuesto se desprende que la topología τ del espacio T queda definida, si se indica en este espacio una base $\mathcal{G}$. Esta topología τ coincide con la colección de conjuntos que pueden representarse como suma de conjuntos de $\mathcal{G}$. Para que esta forma de introducir la topología tenga un valor práctico es necesario señalar aquellas condiciones que debe cumplir un sistema $\mathcal{G}$ de subconjuntos del conjunto dado T para que la colección de todas las sumas posibles de conjuntos de $\mathcal{G}$ pueda ser considerada como la colección de conjuntos abiertos en T (es decir, para que estas sumas verifiquen los axiomas 1° y 2° de espacio topológico).

Estas condiciones vienen dadas por el siguiente teorema.

TEOREMA 2. *Supongamos que en un conjunto T se ha escogido un sistema $\mathcal{G}$ de subconjuntos G_α que verifica las siguientes condiciones:*

a) *Todo punto $x \in T$ está contenido al menos en un subconjunto $G_\alpha \in \mathcal{G}$.*

b) *Si $x \in G_\alpha$ y $x \in G_\beta$, existe un $G_\gamma \in \mathcal{G}$ tal que*

$$x \in G_\gamma \subset G_\alpha \cap G_\beta.$$

Si declaramos abiertos en T al conjunto vacío y a todos los conjuntos que se pueden representarse como la suma de determinados $G_\alpha \in \mathcal{G}$, el conjunto T resultará un espacio topológico (es decir, estas sumas verificarán los axiomas 1° y 2°) y el sistema $\mathcal{G}$ será una base de él.

DEMOSTRACION. De las condiciones del teorema se deduce inmediatamente que todo el T y el conjunto vacío son conjuntos abiertos y que la unión de cualquier número de conjuntos abiertos será abierta. Demostremos que la intersección de cualquier número finito de conjuntos abiertos será abierta. Es suficiente realizar la demostración para el caso de dos conjuntos. Sea $A=\bigcup\limits_{\alpha} G_\alpha$ y $B=\bigcup\limits_{\beta} G_\beta$; entonces, $A\cap B=\bigcup\limits_{\alpha,\beta}(G_\alpha\cap G_\beta)$. Por hipótesis, para todo punto $x\in G_\alpha\cap G_\beta$ existe un $G_\gamma\in\mathcal{G}$ tal que $x\in G_\gamma\subset G_\alpha\cap G_\beta$. Por consiguiente, el conjunto $G_\alpha\cap G_\beta$ es abierto, ya que puede ser representado como la suma de todos los G_γ contenidos en él. Pero en este caso es también abierto el conjunto $A\cap B=\bigcup(G_\alpha\cap G_\beta)$. El hecho de que $\mathcal{G}$ constituye una base del espacio topológico construido se desprende de la forma misma en que se han definido los conjuntos abiertos en T.

Para probar si una colección dada de conjuntos abiertos es base o no suele ser útil el siguiente criterio.

TEOREMA 3. *Para que un sistema $\{G_\alpha\}$ de conjuntos abiertos sea una base del espacio topológico T es necesario y suficiente que para todo conjunto abierto G y todo punto $x\in G$ exista un conjunto G_α de este sistema tal que $x\in G_\alpha\subset G$.*

DEMOSTRACION. *Si $\{G_\alpha\}$ es base, todo abierto $G\subset T$ es suma de determinados G_α:*

$$G=\bigcup_i G_{\alpha_i}$$

y, por consiguiente, todo punto $x\in G$ pertenece a algún G_α contenido en G. Viceversa, si se cumple la condición del teorema, $\{G_\alpha\}$ es base. En efecto, sea G un conjunto abierto arbitrario. Para todo punto $x\in G$ podemos encontrar un $G_\alpha(x)$ tal que $x\in G_\alpha(x)\subset G$. La suma de estos $G_\alpha(x)$, construidos para cada $x\in G$, coincide con G.

Es fácil ver, mediante este criterio, que la colección de todas las bolas abiertas de un espacio métrico constituye una base. La colección de todas las bolas de radio racional también constituye una base. En la recta es una base, por ejemplo, la colección de todos los intervalos racionales (es decir, de todos los intervalos de extremos racionales).

Una clase importante de espacios topológicos la constituyen los *espacios de base numerable*, es decir, los espacios en los que existe por lo menos una base, compuesta por un número, a lo sumo numerable, de conjuntos. Los espacios de base numerable suelen también llamarse *espacios con el segundo axioma de numerabilidad.*

Si un espacio topológico tiene una base numerable, en él existe obligatoriamente un conjunto numerable siempre denso, es decir, un conjunto numerable, cuya adherencia es todo el T. En efecto, sea $\{G_n\}$ una tal base. Tomemos en cada elemento de esta base un punto arbitrario x_n. El conjunto numerable $X=\{x_n\}$ es siempre denso en T, ya que, de lo contrario, el conjunto abierto no vacío $G=T\setminus[X]$ no contendría ningún punto de X y esto no puede ocurrir puesto que G es la suma de determinados conjuntos del sistema $\{G_n\}$ y $x\in G_n$.

Para los espacios métricos se cumple también la afirmación recíproca:

Si en el espacio métrico R existe un conjunto $\{x_n\}$ numerable siempre denso, también existe en R una base numerable.

En efecto, una tal base es, por ejemplo, la constituida por las bolas abiertas $B\left(x_n, \frac{1}{m}\right)$, donde n y m recorren todos los valores naturales. Por lo tanto, tiene lugar el siguiente teorema:

TEOREMA 4. *Un espacio métrico R tiene una base numerable si, y sólo si, existe en él un conjunto siempre denso.*

En virtud de este teorema, todos los ejemplos de espacios métricos, provistos de subconjuntos numerables siempre densos, ofrecen, al mismo tiempo, ejemplos de espacios métricos con el segundo axioma de numerabilidad. El espacio de las sucesiones acotadas (véase el ejemplo 9 del § 1) que no tiene ningún subconjunto numerable siempre denso, tampoco tiene base numerable.

Observación. El teorema 4 no se cumple, en general, para los espacios topológicos arbitrarios (no métricos): se puede dar ejemplos de espacios, provistos de un conjunto numerable siempre denso, que no tienen base numerable. Expliquemos la razón de este fenómeno. Para todo punto x de un espacio métrico R existe un sistema *numerable* $\mathcal{U}$ de vecindades $\left(\text{por ejemplo, el sistema de bolas abiertas } B\left(x, \frac{1}{n}\right)\right)$, que cumple la siguiente condición: cualquiera que sea el conjunto abierto G que contiene al punto x, existe una vecindad der sistema $\mathcal{U}$ que pertenece íntegramente a G. Un sistema tal de vecindades se llama *sistema determinante de vecindades del punto x.* Si para un punto x de un espacio topológico T existe un sistema determinante de vecindades, se dice que en este punto se cumple el *primer axioma de numerabilidad.* Si esto tiene lugar para cada punto del espacio T, el espacio T se llama espacio con el primer axioma de numerabilidad.

Sin embargo, en un espacio topológico arbitrario (aun cuando esté compuesto por un número numerable de puntos) puede no tener lugar el primer axioma de numerabilidad. Por eso no se puede extender al caso de un epsacio topológico arbitrario aquellos razonamientos que en el caso de espacio métrico nos permitieron deducir de la existencia de un conjunto numerable siempre denso la existencia en este espacio de una base numerable.

Un sistema $\{M_\alpha\}$ de conjuntos se llama *cubrimiento* del espacio topológico T, cuando $\bigcup\limits_{\alpha} M_\alpha = T$. Un cubrimiento, compuesto por conjuntos abiertos o cerrados, se llama cubrimiento abierto o cerrado, respectivamente. Si una parte $\{M_{\alpha_i}\}$ del cubrimiento M_α también constituye un cubrimiento del espacio T, se dice que $\{M_{\alpha_i}\}$ es un subcubrimiento del cubrimiento $\{M_\alpha\}$.

TEOREMA 5. *Si T es un espacio topológico de base numerable, de todo cubrimiento suyo abierto se puede extraer un subcubrimiento finito o numerable.*

DEMOSTRACION. Sea $\{O_\alpha\}$ un cubrimiento abierto del espacio T. Entonces, todo punto $x \in T$ está contenido en un O_α. Sea $\{G_n\}$ la base numerable de T. Para todo $x \in T$ existe un elemento $G_n(x)$ de esta base tal que $x \in G_n(x) \subset O_\alpha$. La totalidad de conjuntos $G_n(x)$, obtenidos de esta forma, es finita o numerable y cubre todo T. Escogiendo para cada $G_n(x)$ uno de los conjuntos O_α, que lo contienen, obtendremos un subcubrimiento finito o numerable del cubrimiento $\{O_\alpha\}$. El teorema queda demostrado.

De acuerdo con la definición de espario topológico el conjunto vacío y todo el espacio T son a la vez abiertos y cerrados. Todo espacio, que no tiene ningún otro conjunto a la vez abierto y cerrado, se llama *conexo*. La línea recta R^1 representa uno de los ejemplos más simples de espacios conexos. Pero si extraemos de R^1 aunque sea un sólo punto, el espacio que queda ya no será conexo.

4°. Sucesiones convergentes en T. El concepto de sucesión convergente, conocido de los espacios métricos, se extiende fácilmente al caso de los espacios topológicos. A saber, la sucesión $x_1, x_2, \ldots, x_n, \ldots$ de puntos de T se llama *convergente al punto* x, cuando toda vecindad del punto x contiene todos los puntos de esta sucesión, comenzando de alguno. Sin embargo, en los espacios topológicos este concepto de convergencia no desempeña ese papel fundamental que le corresponde en los espacios métricos. Esto se debe a que en un espacio métrico R un punto x es un punto de adherencia del conjunto $M \subset R$ si, y sólo si, en M existe una sucesión convergente a x, mientras que en un espacio topológico esto, como regla general, no se cumple. Si x es un punto de adherencia para M (es decir, pertenece a $[M]$) en un espacio topológico T, ello no implica la existencia en M de una sucesión convergente a x. Consideremos, a título de ejemplo, el segmento $[0, 1]$, llamando abiertos aquellos subconjuntos suyos (además del conjunto vacío) que se obtienen omitiendo de él un número finito o numerable de puntos. Es fácil ver que los axiomas 1° y 2° (pág. 90) se cumplen, es decir, que tenemos un espacio topológico. En este espacio serán convergentes sólo las sucesiones estacionarias, es decir, tales que sus elementos, empezando de cierto número, coinciden: $x_n = x_{n+1} = \ldots$ (¡demuéstrese esto!). Por otro lado, si cogemos, por ejemplo, a título de M el semisegmento $(0, 1]$, el punto 0 será para él un punto de adherencia (¡compruébese!), pero ninguna sucesión de puntos de M converge a 0 en nuestro espacio.

Las sucesiones convergentes «recobran su importancia», si en vez de considerar espacios topológicos arbitrarios nos limitamos a aquellos espacios, en los que se verifica el primer axioma de numerabilidad, es decir, cuando para todo punto x del espacio T existe un sistema numerable determinante de vecindades. En este caso, todo punto de adherencia x de un conjunto arbitrario $M \subset T$ puede ser representado como límite de cierta sucesión de puntos de M. En efecto, sea $\{O_n\}$ un sistema numerable determinante de vecindades del punto x. Podemos admitir que $O_{n+1} \subset O_n$ $\left(\text{de lo contrario, sustituiríamos } O_n \text{ por } \bigcap_{k=1}^{n} O_k\right)$. Sea x_k un punto arbitrario de M perteneciente a O_k $(k = 1, 2, \ldots)$. Está claro que un punto x_k así debe existir, ya que de lo contrario x no sería punto de adherencia para M. Es obvio que la sucesión $\{x_k\}$ converge a x.

Como hemos señalado, todos los espacios métricos verifican el primer axioma de numerabilidad. Es por eso que en el caso de espacios métricos hemos podido enunciar, en términos de convergencia de sucesiones, conceptos como adherencia, punto de adherencia, etc.

5°. Axiomas de separabilidad. Aunque muchos conceptos principales de la teoría de espacios métricos se extienden fácilmente a cualesquiera espacios topológicos, sin embargo, un espacio topológico arbitrario representa un ente demasiado general desde el punto de vista de los problemas del Análisis. En estos espacios se producen, a veces, situaciones que difieren de modo sustancial de lo que puede ocurrir en los espacios métricos. Así, hemos visto (ejemplo 4, pág. 91) que en un espacio topológico un conjunto finito de puntos puede no ser cerrado, etc.

Entre los espacios topológicos se pueden destacar espacios que por sus propiedades se aproximan a los espacios métricos. Para ello hay que agregar a los axiomas 1° y 2° de espacio topológico unas u otras condiciones adicionales. Condiciones de este tipo son, por ejemplo, los axiomas de numerabilidad que permiten estudiar la topología del espacio a partir del concepto de convergencia. Otro tipo importante de condiciones adicionales y de naturaleza distinta son los así llamados *axiomas de separabilidad*. Enunciaremos esta serie de axiomas en orden de generalización.

T_1. *Primer axioma de separabilidad*: para dos cualesquiera diferentes puntos x e y del espacio T existe una vecindad O_x del punto x, que no contiene al punto y, y una vecindad O_y del punto y, que no contiene al punto x.

Los espacios que verifican este axioma se llaman T_1-espacios. Un ejemplo de un espacio topológico, que no es T_1-espacio, lo ofrece el espacio de dos puntos conexos.

En un T_1-espacio todo punto es un conjunto cerrado. En efecto, si $x \neq y$, existe una vecindad O_y del punto y que no contiene a x, es decir, $y \overline{\in} [x]$. De manera que $[x] = x$. Por consiguiente, en un T_1-espacio resulta también cerrado todo conjunto compuesto de un número finito de puntos. Es más, se puede demostrar fácilmente que el axioma T_1 equivale a exigir que todos estos conjuntos sean cerrados.

El axioma T_2 es una acentuación del primer axioma de separabilidad.

T_2. *Segundo axioma de separabilidad o axioma de Hausdorff:* para dos cualesquiera puntos x e y del espacio topológico T existen vecindades O_x y O_y de intersección vacía.

Los espacios que verifican este axioma se llaman T_2-*espacios* o *espacios de Hausdorff*. Todo espacio de Hausdorff es un T_1-espacio, pero no viceversa. A título de ejemplo de un T_1-espacio que no es espacio de Hausdorff podemos señalar el segmento [0, 1], en el que se consideran abiertos el conjunto vacío y todos los conjuntos que se obtienen omitiendo del segmento a lo sumo un número numerable de puntos.

Generalmente, en el Análisis no se emplean espacios más generales que los de Hausdorff. Es más, como regla general, los espacios, que representan interés para el Análisis, verifican además la siguiente condición, más fuerte aún, llamada condición de normalidad del espacio.

Se llama *espacio normal* a un T_1-espacio en el que cualesquiera dos conjuntos cerrados tienen vecindades[1] de intersección vacía.

Todo espacio normal es de Hausdorff. Un ejemplo de un espacio de Hausdorff que no es normal lo ofrece el segmento $[0, 1]$, en el que las vecindades de todos los puntos, excepto el punto 0, se definen de la manera corriente, mientras que para las vecindades del cero se toman todos los semisegmentos $[0, \alpha)$ de los que se han excluido los puntos de tipo $\frac{1}{n}$ $(n=1, 2, \ldots)$. Esto es un espacio de Hausdorff: el punto 0 y la sucesión $\left\{\frac{1}{n}\right\}$ representan dos conjuntos cerrados de intersección vacía de este espacio que no pueden ser separados mediante dos vecindades de intersección vacía.

Son espacios normales, por ejemplo, todos los espacios métricos. En efecto, sean X e Y dos conjuntos cerrados de intersección vacía de un espacio métrico R. Todo punto $x \in X$ tiene una vecindad O_x que no se interseca con Y y, por consiguiente, está a una distancia positiva ρ_x de Y. De la misma forma todo punto $y \in Y$ está a una distancia positiva ρ_y de X. Consideremos los conjuntos abiertos[2]

$$U = \bigcup_{x \in X} B\left(x, \frac{\rho_x}{2}\right) \quad \text{y} \quad V = \bigcup_{y \in Y} B\left(y, \frac{\rho_y}{2}\right),$$

que contienen a X e Y respectivamente, y demostremos que la intersección de estos conjuntos es vacía. Supongamos que $z \in U \cap V$. En este caso, existe en X un punto x_0, tal que $\rho(x_0, z) < \frac{\rho_{x_0}}{2}$, y en Y un punto y_0 tal que $\rho(z, y_0) < \frac{\rho_{y_0}}{2}$. Supongamos, para concretar, que $\rho_{x_0} \leqslant \rho_{y_0}$. Entonces

$$\rho(x_0, y_0) \leqslant \rho(x_0, z) + \rho(z, y_0) < \frac{\rho_{x_0}}{2} + \frac{\rho_{y_0}}{2} \leqslant \rho_{y_0},$$

es decir, $x_0 \in B(y_0, \rho_{y_0})$; pero esto contradice a la definición de ρ_{y_0}. Hemos demostrado nuestra afirmación.

Todo subespacio de un espacio métrico es por sí mismo un espacio métrico y por eso también posee la propiedad de normalidad. Esto, como regla general, no tiene lugar para los espa-

---

[1] Se llama vecindad de un conjunto M de un espacio topológico T a todo conjunto abierto U que contiene a M.

[2] Aquí $B(x, r)$ representa, como siempre, una bola abierta de radio r y centro en x.

cios normales arbitrarios: un subespacio de un espacio normal no es necesariamente normal. De manera que la normalidad de un espacio no es una propiedad heredera[1].

Una propiedad heredera es la así llamada *regularidad total* de los espacios topológicos. Un T_1-espacio topológico se llama *totalmente regular*, cuando para todo conjunto cerrado $F \subset T$ y todo punto $x_0 \in T \setminus F$ existe una función continua real f, definida sobre T, que es igual a cero en x_0, a la unidad sobre F y que verifica la condición $0 \leqslant f(x) \leqslant 1$. Todo espacio normal es totalmente regular[2], pero no viceversa. Todo subespacio de un espacio totalmente regular (de un espacio normal, en particular) es totalmente regular. A. N. Tíjonov, a quien se debe el propio concepto de espacio totalmente regular, ha demostrado que la clase de espacios totalmente regulares coincide con la clase de todos los subespacios normales. Desde el punto de vista del Análisis, la importancia de los espacios totalmente regulares radica en que sobre cualquier espacio de esta índole se puede definir un número «suficientemente grande» de funciones continuas, ya que para cualesquiera dos puntos distintos x e y de un espacio totalmente regular T existe una función real continua, definida sobre T, que toma en estos puntos diferentes valores.

6°. Aplicaciones continuas. Homeomorfismo. El concepto de aplicación continua, que hemos introducido en el § 1 para los espacios métricos, es extensible, naturalmente, a espacios topológicos arbitrarios.

DEFINICION. Sean X e Y dos espacios topológicos. Una aplicación f del espacio X en el espacio Y se llama *continua en el punto* x_0, cuando para toda vecindad U_{y_0} del punto $y_0 = f(x_0)$ existe una vecindad V_{x_0} del punto x_0 tal que $f(V_{x_0}) \subset U_{y_0}$. Una *aplicación* $f\colon X \to Y$ se llama *continua*, cuando es continua para todo punto $x \in X$. En particular, una aplicación continua del espacio topológico X en la recta numérica se llama *función continua* sobre este espacio topológico. Es fácil ver, que en el caso de espacios métricos esta definición coincide, en efecto, con la definición de una aplicación continua de un espacio métrico en otro, que ha sido dada en el § 1 de este capítulo.

---

[1] Una propiedad P se llama *heredera*, si siendo justa para todo el espacio topológico T también se verifica para cualquiera de sus subespacios.

[2] Este resultado (lejos de ser evidente) se desprende del siguiente teorema de P. S. Urisón: si T es un espacio normal y F_1 y F_2 dos conjuntos suyos cerrados y de intersección vacía, existe una función f, $0 \leqslant f(x) \leqslant 1$, continua, definida sobre T, que es igual a cero sobre F_1 e igual a la unidad sobre F_2.

Enunciemos ahora el concepto de continuidad de una aplicación de un espacio topológico en otro en términos de conjuntos abiertos, es decir, en términos de las *topologías* de los dos espacios considerados. Notemos que la aplicación $f: X \longrightarrow Y$ puede ser considerada aplicación *sobre*, ya que siempre podemos considerar el subespacio $f(X) \subset Y$ en vez de Y.

TEOREMA 6. *Para que una aplicación f de un espacio topológico X en un espacio topológico Y sea continua, es necesario y suficiente que la imagen recíproca $\Gamma = f^{-1}(G)$ de todo conjunto abierto $G \subset Y$ sea abierta* (en X).

DEMOSTRACION. NECESIDAD. Sea f una aplicación continua y G un conjunto abierto en Y. Demostrèmos que $\Gamma = f^{-1}(G)$ es abierto. Sea x un punto cualquiera de Γ e $y = f(x)$. Entonces, G representa una vecindad del punto y. De acuerdo con la definición de continuidad, existe una vecindad V_x del punto x, tal que $f(V_x) \subset G$, es decir, $V_x \subset \Gamma$. En otras palabras, si $x \in \Gamma$, existe una vecindad V_x de este punto contenida en Γ. Pero esto significa precisamente que Γ es abierto.

SUFICIENCIA. Sea $\Gamma = f^{-1}(G)$ abierto, cuando $G \subset Y$ es abierto. Consideremos un punto arbitrario $x \in X$ y una vecindad arbitraria U_y del punto $y = f(x)$. Puesto que $y \in U_y$, el punto x pertenece al conjunto $f^{-1}(U_y)$. Este es un conjunto abierto y puede ser considerado como aquélla vecindad del punto x, cuya imagen está contenida en U_y. Hemos demostrado el teorema.

Observacion. Sean X e Y conjuntos arbitrarios y f una aplicación de X en Y. Supongamos que Y está provisto de una topología τ (es decir, de un sistema de conjuntos que contiene a Y y a $\varnothing$ y que resulta cerrado respecto a las operaciones de sumas cualesquiera e intersecciones finitas). Debido a que la imagen recíproca de la suma (intersección) de conjuntos es igual a la suma (intersección) de las imágenes recíprocas (teoremas 1 y 2 del § 3, cap. I), obtenemos que la imagen recíproca de la topología τ (es decir, la colección de todos los conjuntos $f^{-1}(G)$, donde $G \in \tau$) es una topología en X, que denotaremos $f^{-1}(\tau)$. Si ahora X e Y son espacios topológicos, con topologías τ_x y τ_y respectivamente, el teorema 6, que da la condición necesaria y suficiente de la continuidad de la aplicación $f: X \longrightarrow Y$, puede ser enunciado así: la aplicación $f: X \longrightarrow Y$ es continua si, y sólo si, la topología τ_x es *más fuerte* que la topología $f^{-1}(\tau_y)$.

Teniendo en cuenta que la imagen recíproca del complemento es igual al complemento de la imagen recíproca, obtenemos el teorema dual al teorema 6.

TEOREMA 6'. *Para que la aplicación f de un espacio topológico X en un espacio topológico Y sea continua, es necesario y suficiente que la imagen recíproca de todo conjunto cerrado de Y sea cerrada.*

Es fácil ver que la *imagen* de un conjunto abierto (cerrado) ofrecida por una aplicación continua no es necesariamente abierta (cerrada). Consideremos, por ejemplo, la aplicación del semisegmento $X=[0, 1)$ en una circunferencia de la misma longitud. El conjunto $\left[\frac{1}{2}, 1\right)$, cerrado en $[0, 1)$, se transforma en este caso en un conjunto no cerrado de la circunferencia (fig. 12).

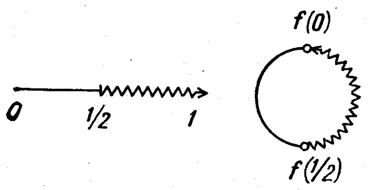


FIG. 12

Para las aplicaciones continuas se cumple el siguiente teorema, análogo al teorema, bien conocido del Análisis, sobre la continuidad de la función compuesta.

TEOREMA 7. *Sean X, Y, y Z espacios topológicos y sean f y φ aplicaciones continuas de X en Y y de Y en Z, respectivamente. Entonces la aplicación $x \dashrightarrow \varphi(f(x))$ del espacio X en Z es continua.*

La demostración de este teorema se obtiene del teorema 6.

Una aplicación f del espacio topológico X sobre el espacio topológico Y biunívoca y bicontinua a la vez se llama *homeomorfismo* y los espacios X e Y se llaman *homeomorfos*. Los espacios homeomorfos entre sí tienen las mismas propiedades topológicas y desde el punto de vista topológico pueden considerarse como dos ejemplares de un mismo espacio. Las topologías de dos espacios homeomorfos son imágenes e imágenes recíprocas una de la otra. La relación homeomorfa es reflexiva, simétrica y transitiva; por consiguiente, la totalidad de los espacios topológicos se divide en clases disjuntas de espacios homeomorfos entre sí.

Observación. El concepto de homeomorfismo ya fue introducido en el § 1 para los espacios métricos. Debe tenerse en cuenta, sin embargo, que las propiedades métricas de dos espacios métricos homeomorfos pueden ser distintas[1]. Así, uno de ellos puede ser completo y el otro no serlo. Por ejemplo, el intervalo $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ es homeomorfo a la recta numérica (el homeomorfismo correspondiente viene dado por la función $x \longrightarrow \operatorname{tg} x$) y, sin embargo, la recta es un espacio completo y el intervalo no lo es.

7°. Distintos métodos de definición de topologías en un espacio. Metrizabilidad. La forma más directa y, por su esencia, más elemental de introducir una topología en un espacio consiste en señalar directamente aquellos conjuntos que se consideran abiertos. La totalidad de estos conjuntos debe verificar las condiciones 1° y 2° (véase la pág. 90). Equivalente a ella es la forma dual que consiste en indicar la colección de conjuntos cerrados. Es obvio que esta colección debe verificar las condiciones 1′ y 2′ (pág. 90). Sin embargo, este método se puede aplicar, de hecho, en reducidos casos. Por ejemplo, ya en el caso del plano no se puede dar, por lo visto, una descripción directa de los subconjuntos abiertos (de la forma como se logra hacer esto para la recta (teorema 5 del § 2)).

Un método que se emplea frecuentemente para introducir una topología en un espacio consiste en indicar en él una base; de hecho, precisamente de este modo, se introduce la topología en los espacios métricos, donde, a partir de la métrica, se define la base, es decir, la colección de bolas abiertas.

Otra forma posible de definir una topología en un espacio consiste en introducir en él el concepto de convergencia. Sin embargo, hemos señalado ya en el punto 3 que este procedimiento no es universal: por medio de él sólo se pueden introducir topologías en espacios, en los que se cumple el primer axioma de numerabilidad. No obstante, desde el punto de vista del Análisis este método resulta frecuentemente útil[2].

En un espacio se puede introducir la topología definiendo en él axiomáticamente la operación de adherencia. Se dice que en el conjunto X se ha introducio la operación de adherencia, cuando a todo $A \subset X$ se le ha puesto en correspondencia un conjunto $[A] \subset X$, llamado adherencia de A, de manera que la operación

[1] La métrica del espacio R determina unívocamente su topología, pero no viceversa: una misma topología en $R = (X, \rho)$ se puede obtener introduciendo en X diferentes métricas.

[2] Más aun si tenemos en cuenta que generalizando el concepto de convergencia (convergencia respecto a los así llamados filtros) este método resulta aplicable también en el caso general.

consistente en pasar del conjunto A al conjunto $[A]$ verifica las propiedades 1), 2), 3) y 4) indicadas en el teorema 1 del § 2.

Uno de los métodos más importantes, aunque lejos de ser universal, de introducir una topología consiste en definir en el espacio una métrica. Como hemos visto, todo espacio métrico es normal y verifica el primer axioma de numerabilidad. Por eso si en algún espacio no se cumple una de estas dos condiciones, no se puede introducir la topología en él mediante ninguna métrica.

DEFINICION. Un espacio topológico se llama *metrizable*, cuando su topología puede ser introducida mediante alguna métrica.

De acuerdo con lo que acabamos de señalar, la normalidad del espacio y el primer axioma de numerabilidad son condiciones necesarias para que el espacio sea metrizable. Sin embargo, ninguna de estas dos condiciones por separado ni, incluso, ambas juntas son suficientes para que el espacio sea metrizable. No obstante, tiene lugar el siguiente teorema, que pertenece a P. S. Urisón:

para que en espacio topológico provisto de base numerable sea metrizable, es necesario y suficiente que sea normal.

La necesidad de esta condición es evidente.

§ 6. COMPACIDAD

1°. Concepto de compacidad. En el Análisis desempeña un papel fundamental el siguiente hecho, conocido como lema de Heine—Borel:

de cualquier cubrimiento del segmento $[a, b]$ de la recta numérica por medio de intervalos se puede extraer un subcubrimiento finito.

Esta afirmación continúa siendo válida, cuando en vez de intervalos se consideran cualesquiera conjuntos abiertos: de todo cubrimiento abierto del segmento $[a, b]$ se puede extraer un subcubrimiento finito.

Partiendo de esta propiedad del segmento de la recta numérica, introducimos el siguiente concepto importante.

DEFINICION. Un espacio topológico T se llama *compacto*, cuando cualquier cubrimiento abierto suyo contiene un subcubrimiento finito. Un espacio topológico compacto que verifica el axioma de separabilidad de Hausdorff se llama un *compacto*.

Como veremos más adelante, la propiedad de compacidad la tienen, además de los segmentos, todos los subconjuntos cerrados acotados de un espacio euclídeo de cualquier dimensión finita. Al contrario, la recta, el plano y el espacio de tres dimensiones son ejemplos elementales de espacios no compactos.

Un sistema de subconjuntos $\{A\}$ del conjunto T se llama sistema *centrado*, cuando cualquier intersección finita $\bigcap_{i=1}^{n} A_i$ de elementos de este sistema es no vacía. De la definición dada de compacidad y de las relaciones de dualidad se deduce el siguiente teorema.

TEOREMA 1. *Para que un espacio topológico T sea compacto, es necesario y suficiente que verifique la condición (R):*

todo sistema centrado de subconjuntos cerrados de este espacio tiene intersección no vacía.

En efecto, sea $\{F_\alpha\}$ un sistema centrado de subconjuntos cerrados de T y sea T un espacio compacto. Los conjuntos $G_\alpha = T \setminus F_\alpha$ son abiertos; además, como cualquier intersección finita $\bigcap_{i=1}^{n} F_i$ es no vacía, ningún sistema finito de conjuntos $G_i = T \setminus F_i$ puede cubrir todo T. Pero en este caso todos los G_α tampoco forman cubrimiento (compacidad); esto significa que $\bigcap F_\alpha \neq \varnothing$. Por consiguiente, si T es un espacio compacto, en él se verifica la condición (R). Viceversa, supongamos que T verifica la condición (R) y que $\{G_\alpha\}$ es un cubrimiento abierto del espacio T. Tomando $F_\alpha = T \setminus G_\alpha$, obtenemos que $\bigcap F_\alpha = \varnothing$ y de aquí se desprende (condición (R)) que el sistema $\{F_\alpha\}$ no puede ser centrado, es decir, existen $F_1, \ldots, F_n$ tales que $\bigcap_{i=1}^{n} F_i = \varnothing$. Pero en este caso los conjuntos correspondientes $G_i = T \setminus F_i$ constituyen un subcubrimiento finito del cubrimiento $\{G_\alpha\}$. Por consiguiente, de la condición (R) se deduce la compacidad.

Veamos algunas propiedades principales de los espacios compactos.

TEOREMA 2. *Todo subconjunto cerrado de un espacio compacto es compacto.*

DEMOSTRACION. Sea F un subconjunto cerrado de un espacio compacto T y sea $\{F_\alpha\}$ un sistema centrado cualquiera de subconjuntos cerrados del subespacio $F \subset T$. En tal caso, todo F_α es cerrado también en T, es decir, $\{F_\alpha\}$ es un sistema centrado de conjuntos cerrados de T. Por consiguiente, $\bigcap F_\alpha \neq \varnothing$. De aquí, de acuerdo con el teorema 1, se desprende la compacidad de F.

Puesto que todo subespacio de un espacio de Hausdorff es también de Hausdorff, obtenemos de aquí el siguiente corolario.

COROLARIO. *Todo subconjunto cerrado de un compacto es un compacto.*

TEOREMA 3. *Un compacto resulta cerrado en cualquier espacio de Hausdorff que lo contiene.*

DEMOSTRACION. Sea K un conjunto compacto en un espacio de Hausdorff T y sea $y \in K$. Entonces, para todo punto $x \in K$ existe una vecindad U_x del punto x y una vecindad V_x del punto y tales que

$$U_x \cap V_x = \varnothing .$$

Las vecindades U_x forman un cubrimiento abierto del conjunto K. Debido a la compacidad de K, se puede extraer de él un subcubrimiento finito $U_{x_1}, U_{x_2}, \ldots, U_{x_n}$. Pongamos

$$V = V_{x_1} \cap V_{x_2} \cap \ldots \cap V_{x_n}.$$

Entonces, V es una vecindad del punto y que tiene intersección vacía con $U_{x_1} \cup \ldots \cup U_{x_n} \supset K$. Por consiguiente, $y \overline{\in} [K]$, es decir, K es cerrado. El teorema queda demostrado.

Los teoremas 2 y 3 indican que en los espacios de Hausdorff la compacidad es una propiedad *interna* del espacio, es decir, que todo compacto continúa siendo un compacto, aunque sea sumergido en espacios de Hausdorff cada vez más amplios.

TEOREMA 4. *Todo compacto representa un espacio normal.*

DEMOSTRACION. Sean X e Y dos subconjuntos cerrados disjuntos de un compacto K. Es fácil ver, repitiendo los razonamientos realizados en la demostración del teorema anterior, que para todo punto $y \in Y$ existe una vecindad suya U_y y un conjunto abierto $O_y \supset X$ tales que $U_y \cap O_y = \varnothing$. Extraigamos del cubrimiento $\{U_y\}$ del conjunto Y un subcubrimiento finito $U_{y_1}, \ldots, U_{y_n}$. Entonces, los conjuntos abiertos

$$O^{(1)} = O_{y_1} \cap \ldots \cap O_{y_n}$$

y

$$O^{(2)} = U_{y_1} \cup U_{y_2} \cup \ldots \cup U_{y_n}$$

verificarán las condiciones necesarias:

$$O^{(1)} \supset X, \quad O^{(2)} \supset Y$$

y

$$O^{(1)} \cap O^{(2)} = \varnothing .$$

2°. Aplicaciones continuas de espacios compactos. Las aplicaciones continuas de espacios compactos, en particular, de compactos, tienen varias propiedades interesantes e importantes.

TEOREMA 5. *La imagen continua de un espacio compacto es un espacio compacto.*

DEMOSTRACION. Sea X un espacio compacto y f una aplicación continua de él sobre el espacio topológico Y. Consideremos algún cubrimiento $\{V_\alpha\}$ del espacio Y mediante conjuntos abiertos y tomemos $U_\alpha = f^{-1}(V_\alpha)$. Los conjuntos U_α son abiertos (como imágenes recíprocas de conjuntos abiertos en caso de una aplicación continua) y forman un cubrimiento del espacio X. De este cubrimiento se puede extraer, debido a la compacidad de X, un subcubrimiento finito $U_1, U_2, \ldots, U_n$. En este caso, los conjuntos $V_1, V_2, \ldots, V_n$, donde $V_i = f(U_i)$, cubrirán a todo el espacio Y.

TEOREMA 6. *Toda aplicación biunívoca y continua de un compacto X sobre otro compacto Y es un homeomorfismo.*

DEMOSTRACION. Debemos probar que de las condiciones del teorema se deduce la continuidad de la aplicación inversa φ^{-1}. Sea F un conjunto cerrado de X y sea $P = \varphi(F)$ su imagen en Y. De acuerdo con el teorema anterior, P es un compacto y, por consiguiente, P es cerrado en Y. De manera que la imagen recíproca por la aplicación φ^{-1} de todo conjunto cerrado $F \subset X$ es cerrada. Esto significa precisamente que la aplicación φ^{-1} es continua.

3°. Compacidad numerable.

TEOREMA 7. *Si T es un espacio compacto, todo subconjunto suyo infinito tiene al menos un punto de acumulación.*

DEMOSTRACION. Si T contiene un conjunto infinito X sin puntos de acumulación, se puede escoger en él un conjunto numerable

$$X_1 = (x_1, x_2, \ldots)$$

que tampoco tiene puntos de acumulación. Pero los conjuntos

$$X_n = (x_n, x_{n+1} \ldots)$$

forman entonces un sistema centrado de conjuntos cerrados de T, que tiene intersección vacía, es decir, T no es un espacio compacto. Introduzcamos la siguiente definición.

DEFINICION. Un espacio T se llama espacio *compacto numerable*, cuando todo subconjunto infinito suyo tiene al menos un punto de acumulación.

El teorema 7 significa, por lo tanto, que todo espacio compacto es compacto numerable. La recíproca, en general, no se cumple. He aquí un ejemplo «tradicional» de un espacio compacto numerable, pero no compacto. Consideremos el conjunto X de todos los números ordinales α menores que el primer número ordinal innumerable Ω. Llamemos intervalo (α, β) de X a la totalidad de números ordinales γ, que verifican las desigualdades $\alpha < \gamma < \beta$. Llamemos conjunto abierto en X a toda unión de un número arbitrario de intervalos. Es fácil comprobar que el espacio construido es compacto numerable, pero no compacto.

El siguiente teorema deja clara la relación existente entre los conceptos de compacidad y compacidad numerable.

TEOREMA 8. *Para que un espacio topológico sea compacto numerable, es necesaria y suficiente cada una de estas dos condiciones:*

1) Todo cubrimiento abierto numerable del espasio T contiene un subcubrimiento finito.

2) Todo sistema centrado numerable de conjuntos cerrados de T tiene una intersección no vacía.

DEMOSTRACION. La equivalencia de las condiciones 1) y 2) se desprende inmediatamente de las relaciones de dualidad. Ahora bien, si T no es compacto numerable, podemos demostrar, repitiendo los razonamientos empleados al demostrar el teorema 7, que en T existe un sistema centrado numerable de conjuntos cerrados, cuya intersección es vacía. Con esto queda establecido que la condición 2) (y, por consiguiente, la condición 1)) es suficiente. Demostremos la necesidad de la condición 2). Sea T un espacio compacto numerable y sea $\{F_n\}$ un sistema numerable centrado de conjuntos cerrados de T. Probemos que $\bigcap F_n \neq \varnothing$. Sea

$$\Phi_n = \bigcap_{k=1}^{n} F_k.$$

Está claro que todos los Φ_n son no vacíos (debido a que $\{F_n\}$ es un sistema centrado), que forman un sistema no creciente

$$\Phi_1 \supset \Phi_2 \supset \ldots$$

y que

$$\bigcap \Phi_n = \bigcap F_n.$$

Pueden darse dos casos:

1) Comenzando de un número n_0,

$$\Phi_{n_0} = \Phi_{n_0+1} = \ldots$$

Es evidente, entonces, que $\bigcap \Phi_n = \Phi_{n_0} \neq \varnothing$.

2) Entre los Φ_n hay un número infinito de conjuntos distintos dos a dos. Basta entonces, evidentemente, considerar el caso en que todos los Φ_n son distintos entre sí. Sea

$$x_n \in \Phi_n \setminus \Phi_{n+1}.$$

La sucesión $\{x_n\}$ representa un conjunto infinito de puntos diferentes de T; debido a la compacidad numerable de T, esta sucesión debe tener al menos un punto de acumulación, digamos, x_0. Puesto que Φ_n contiene todos los puntos $x_n, x_{n+1}, \ldots$, el punto x_0 es un punto de acumulación para Φ_n y, como Φ_n es cerrado, $x_0 \in \Phi_n$. Por consiguiente, $\bigcap\limits_n \Phi_n \ni x_0$, es decir, $\bigcap\limits_n \Phi_n \neq \varnothing$.

De manera que los espacios compactos numerables son aquellos espacios topológicos, en los que de cada cubrimiento abierto *numerable* se puede extraer un subcubrimiento finito, mientras que en un espacio compacto *todo* cubrimiento abierto contiene un subcubrimiento finito.

Aunque en el caso general la compacidad numerable no implica la compacidad, tiene lugar el siguiente resultado importante.

TEOREMA 9. *Para los espacios de base numerable los conceptos de compacidad y compacidad numerable coinciden.*

En efecto, del teorema 6 del § 5 se deduce que de cualquier cubrimiento abierto del espacio T, provisto de una base numerable, se puede extraer un subcubrimiento numerable. Si T es, además, compacto numerable, de este último se puede extraer, de acuerdo con el teorema anterior, un subcubrimiento finito. Con esto queda establecido que T es un espacio compacto.

Observación. De hecho, el concepto de compacidad numerable de un espacio topológico resulta (a diferencia del concepto de compacidad) poco acertado y poco natural. Surgió en las Matemáticas debido a una especie de «inercia». Como quedará demostrado en el punto siguiente, en el caso de espacios métricos estos dos conceptos coinciden (al igual que en el caso de espacios de base numerable). El concepto de compacidad en los espacios métricos significaba inicialmente la existencia de un punto de acumulación en todo subconjunto infinito, o sea, coincidía con la definición de compacidad numerable. La extensión «automática» de esta definición de los espacios métricos a los topológicos condujo precisamente al concepto de espacio topológico compacto numerable. En la literatura, especialmente anticuada, el término «compacidad» se entiende a veces como «compacidad numerable», mientras que un espacio topológico compacto en el sentido de la definición que hemos introducido (es decir, espacio en que todo cubrimiento abierto suyo contiene un subcubrimiento finito) se llama *bicompacto*. Además, un espacio de Hausdorff compacto se llama un *bicompacto*, reservándose el término de «un compacto» para los espacios métricos compactos. Nos atendremos a la terminología (compacidad, compacidad numerable) que hemos introducido más arriba; además, los espacios métricos compactos los llamaremos simplemente compactos y en los casos, cuando resulte deseable subrayar la presencia de la métrica, diremos «compactos métricos».

4°. Conjuntos relativamente compactos. Un conjunto M, perteneciente a un espacio de Hausdorff T, que no sea cerrado en T, no puede ser compacto. Por ejemplo, ningún subconjunto no cerrado de la recta numérica es un compacto. Puede, sin embargo, ocurrir que la adherencia $[M]$ de un tal conjunto M de T tenga ya la propiedad de compacidad. Por ejemplo, esto sucede para *todo subconjunto acotado* de la recta numérica o de un espacio de n dimensiones. Introduzcamos la siguiente definición.

DEFINICIÓN. Un conjunto M, perteneciente a un espacio topológico T, se llama *relativamente compacto* (en T), cuando su adherencia en T es compacta. De la misma forma se dice que M es *relativamente compacto numerable en* T cuando todo subconjunto infinito $A\subset M$ tiene al menos un punto de acumulación (que puede pertenecer, pero puede y no pertenecer a M).

El concepto de compacidad relativa (a diferencia del concepto de compacidad) está relacionado, evidentemente, con aquel espacio T, en el que se considera el conjunto dado. Por ejemplo, el conjunto de los puntos racionales del intervalo (0, 1) es relativamente compacto, si se considera como un subconjunto de la recta numérica, pero no será relativamente compacto si se considera como un subconjunto del espacio de todos los números racionales.

El concepto de compacidad relativa adquiere su mayor importancia en el caso de los espacios métricos que trataremos en el parágrafo siguiente.

§ 7. COMPACIDAD EN ESPASIONS MÉTRICOS

1°. Acotación total. Puesto que los espacios métricos representan un caso particular de los espacios topológicos, a ellos se extienden los resultados y definiciones, expuestos en el parágrafo anterior. En el caso de los espacios métricos la compacidad está estrechamente ligada al concepto de *acotación total* que ahora introduciremos.

Sea M un conjunto de un espacio métrico R y ε un número positivo. Se dice que el conjunto A de R es una ε-red de M, cuando para todo punto $x \in M$ existe al menos un punto $a \in A$ tal que

$$\rho(x, a) \leqslant \varepsilon.$$

Por ejemplo, los puntos de coordenadas enteras forman una $\frac{1}{\sqrt{2}}$-red del plano. Un conjunto M se llama *totalmente acotado*, cuando cualquiera que sea $\varepsilon > 0$ existe una ε-red *finita* suya. Está claro, que un conjunto totalmente acotado es necesariamente acotado, como la suma de un número finito de conjuntos acotados; la afirmación recíproca no es, en general, justa, como lo demuestra el ejemplo 2 que citamos más abajo.

Frecuentemente resulta útil la siguiente observación obvia: si el conjunto M es totalmente acotado, su adherencia $[M]$ es también totalmente acotada.

De la definición de acotación total se desprende inmediatamente que todo espacio métrico R totalmente acotado es *separable*. En efecto, construyamos para todo n una $\frac{1}{n}$-red finita de R. La suma, respesto a n, de todas estas redes representa un conjunto numerable siempre denso en R.

Ejemplos. 1. En el espacio euclídeo de n dimensiones la acotación total coincide con la acotación corriente, es decir, cor la posibilidad de sumergir el conjunto dado dentro de un cubo suficientemente grande. En efecto, si dividimos este cubo en cubos de dimensión ε, los vértices de estos últimos formarán una $\frac{\sqrt{n}}{2}$ ε-red finita en el cubo inicial y, por consiguiente, en cualquier conjunto, contenido en este cubo.

2. La esfera unitaria S del espacio l_2 ofrece un ejemplo de un conjunto acotado, pero no totalmente. En efecto, consideremos los siguientes puntos de S

$$\begin{gathered} e_1 = (1, 0, 0, \ldots), \\ e_2 = (0, 1, 0, \ldots), \\ \cdots\cdots\cdots \\ e_n = (0, 0, 0, \ldots, 1, 0, \ldots) \\ \cdots\cdots\cdots\cdots\cdots \end{gathered}$$

La distancia entre dos cualesquiera de estos puntos e_n y e_m $(n \neq m)$ es igual a $\sqrt{2}$. De aquí se desprende que en S no puede existir una ε-red finita para ningún $\varepsilon < \frac{\sqrt{2}}{2}$.

3. Consideremos en l_2 el conjunto Π de puntos

$$x=(x_1, x_2, \ldots, x_n, \ldots)$$

que verifican las siguientes condiciones

$$|x_1| \leqslant 1, \ |x_2| \leqslant \frac{1}{2}, \ \ldots, \ |x_n| \leqslant \frac{1}{2^{n-1}}, \ \ldots$$

Este conjunto se llama *paralelepípedo fundamental* (o «ladrillo de Hilbert») del espacio l_2. Representa un ejemplo de un conjunto totalmente acotado de infinitas dimensiones. Para demostrar que es totalmente acotado podemos proceder del siguiente modo.

Sea dado $\varepsilon > 0$. Escogemos n de manera que $\frac{1}{2^{n-1}} < \frac{\varepsilon}{2}$. A todo punto

$$x=(x_1, x_2, \ldots, x_n, \ldots) \tag{1}$$

de Π pongamos en correspondencia el punto

$$x^*=(x_1, x_2, \ldots, x_n, 0, 0, \ldots) \tag{2}$$

de este mismo conjunto. En este caso

$$\rho(x, x^*)=\sqrt{\sum_{k=n+1}^{\infty} x_k^2} \leqslant \sqrt{\sum_{k=n}^{\infty} \frac{1}{4^k}} < \frac{1}{2^{n-1}} < \frac{\varepsilon}{2}.$$

El conjunto Π^* de puntos de Π de tipo (2) es totalmente acotado (por ser un conjunto acotado de un espacio de n dimensiones). Tomemos en Π^* una $\frac{\varepsilon}{2}$-red finita. Está claro que será al mismo tiempo una ε-red de Π.

2°. Compacidad y acotación total.

TEOREMA 1. *Todo espacio métrico R compacto numerable es totalmente acotado.*

DEMOSTRACION. Supongamos que R no es totalmente acotado. Esto significa que para algún $\varepsilon_0 > 0$ no existe en R una ε_0-red finita. Tomemos en R un punto arbitrario a_1. Existe en R al menos un punto, digamos a_2, tal que $\rho(a_1, a_2) > \varepsilon_0$ (de lo contrario, el punto a_1 resultaría ser una ε_0-red de R). De la misma forma, en R existe un punto a_3 tal que

$$\rho(a_1, a_3) > \varepsilon_0 \quad \text{y} \quad \rho(a_2, a_3) > \varepsilon_0,$$

ya que, de lo contrario, el par de puntos a_1, a_2 representaría una ε_0-red. Determinados ya los puntos $a_1, \ldots, a_k$, escojamos el punto $a_{k+1} \in R$ de manera que

$$\rho(a_i, a_{k+1}) > \varepsilon_0, \quad i=1, 2, \ldots, k.$$

Mediante este proceso obtenemos una sucesión infinita $a_1, a_2, \ldots$ que no tiene ningún punto de acumulación, ya que $\rho(a_i, a_j) > \varepsilon_0$ para $i \neq j$. Pero en tal caso R no es un espacio compacto numerable, que es lo que queríamos demostrar.

COROLARIO 1. *Un espacio métrico R compacto numerable tiene un conjunto numerable siempre denso y una base numerable.*

En efecto, construyamos en R una $\frac{1}{n}$-red finita para todo $n=1, 2, \ldots$ y tomemos la unión de estas redes. Esta será precisamente el conjunto numerable siempre denso en R. La existencia de una base numerable en un espacio métrico, provisto de un conjunto numerable siempre denso, fue demostrada ya anteriormente (teorema 5 del § 5).

Recordando el teorema 9 del § 6, obtenemos el siguiente corolario importante.

COROLARIO 2. *Todo espacio métrico compacto numerable es compacto.*

Hemos demostrado que la acotación total es una condición necesaria de la compacidad de un espacio métrico. Esta condición no es suficiente; por ejemplo, la totalidad de los puntos racionales del segmento $[0, 1]$ con la definición corriente de la distancia entre ellos es un espacio totalmente acotado, pero no compacto: la sucesión de puntos de este espacio

$$0;\ 0{,}4;\ 0{,}41;\ 0{,}414;\ 0{,}4142;\ \ldots,$$

es decir, la sucesión de aproximaciones decimales del número $\sqrt{2}-1$ no tiene en este espacio punto de acumulación. Sin embargo, tiene lugar el siguiente teorema.

TEOREMA 2. *Para que un espacio métrico R sea un compacto, es necesario y suficiente que sea*

1) *totalmente acotado,*
2) *completo.*

DEMOSTRACION. La necesidad de la acotación total ya la hemos demostrado. La necesidad de la complitud es evidente: en efecto, si $\{x_n\}$ es una sucesión fundamental sin límite de R, esta sucesión no tiene en R ningún punto de acumulación. Probemos ahora que siendo R totalmente acotado y completo, es compacto. Para ello es suficiente demostrar que toda sucesión $\{x_n\}$ de puntos de R tiene al menos un punto de acumulación.

Construyamos una bola cerrada de radio 1 en torno a cada uno de los puntos que forman una 1-red en R. Puesto que estas bolas cubren todo el R y el número de ellas es finito, al menos una de estas bolas, llamémosla B_1, contiene una subsucesion (sucesión parcial) infinita $x_1^{(1)}, \ldots, x_n^{(1)}, \ldots$ de la sucesión $\{x_n\}$. Escojamos ahora en B_1 una 1/2-red y construyamos alrededor de todo punto de esta red una bola cerrada de radio 1/2. Al menos una de estas bolas, llamémosla B_2, contiene una subsucesión infinita $x_1^{(2)}, \ldots, x_n^{(2)}, \ldots$ de la sucesión $\{x_n^{(1)}\}$. Busquemos luego una bola cerrada B_3 de radio 1/4 y centro en B_2 que contiene una subsucesión infinita $x_1^{(3)}, \ldots. x_n^{(3)}, \ldots$ de la sucesión $\{x_n^{(2)}\}$, etc. Consideremos con toda bola B_n una bola cerrada A_n con centro en el mismo punto, pero de un radio dos veces mayor. Es fácil ver que las bolas A_n están encajadas unas en otras. Debido a la complitud del espacio R, la intersección $\bigcap_{n=1}^{\infty} A_n$ es no vacía y consta de un sólo punto x_0. Este punto es un punto de acumulación de la sucesión inicial $\{x_n\}$, ya que toda vecindad suya contiene una bola B_k y, por consiguiente, una subsucesión infinita $\{x_n^{(k)}\}$ de la sucesión $\{x_n\}$.

3°. Compacidad relativa de subconjuntos en un espacio métrico. El concepto de compacidad relativa, introducido en el parágrafo anterior para subconjuntos de un espacio topológico arbitrario, es aplicable, en particular, a los subconjuntos de un espacio métrico. Es evidente, además, que el concepto de compacidad relativa coincide en este caso con el concepto de compacidad numerable relativa. Destaquemos el siguiente resultado simple, pero importante.

TEOREMA 3. *Para que un conjunto M de un espacio métrico completo R sea relativamente compacto, es necesario y suficiente que sea totalmente acotado.*

La demostración se obtiene inmediatamente del teorema 2 y del hecho evidente de que todo subconjunto cerrado de un espacio métrico completo es también completo.

La importancia de este teorema estriba en que, como regla general, resulta más fácil establecer la acotación total de uno u otro conjunto, que demostrar directamente su compacidad relativa. Al mismo tiempo, para las aplicaciones del Análisis tiene importancia precisamente la compacidad.

Observación. Al demostrar la acotación total de uno u otro conjunto M (es decir, al construir en él una ε-red finita para todo $\varepsilon > 0$) de un espacio métrico R, no es necesario que esta ε-red pertenezca a M. Es suficiente que esta ε-red finita pueda ser construida mediante puntos, pertenecientes a R. En efecto, si los puntos

$$a_1, a_2, \ldots, a_n \in R$$

son tales que para todo punto x

$$\rho(x, a_k) \leqslant \varepsilon$$

para cierto k, obtendremos una 2ε-red en el conjunto M, compuesta por puntos de este conjunto, al sustituir todo punto a_k mediante un punto b_k que verifique las condiciones

$$\rho(a_k, b_k) \leqslant \varepsilon, \quad b_k \in M.$$

4°. Teorema de Arzelá. La demostración de la compacidad de un conjunto de un espacio métrico, es un problema que encontramos con bastante frecuencia en el Análisis. Al mismo tiempo, la aplicación directa del teorema 2 del punto 2 no siempre resulta simple. Para los conjuntos, situados en un espacio concreto, se puede dar criterios especiales de compacidad, que resultan más cómodos para la aplicación práctica.

En un espacio euclídeo de n dimensiones la compacidad de un conjunto es equivalente, como hemos visto, a su acotación. Sin embargo, esto ya no es cierto para espacios métricos más generales.

En el Análisis, uno de los espacios métricos más importantes es el espacio $C_{[a, b]}$. Para los subconjuntos de este espacio, un criterio importante y frecuentemente empleado de compacidad relativa lo ofrece el así llamado teorema de Arzelá.

Para poder enunciarlo, necesitamos los siguientes conceptos.

Una familia Φ de funciones φ, definidas sobre un segmento, se llama *equiacotada*, cuando existe un número K tal que

$$|\varphi(x)| < K$$

para todo $x \in [a, b]$ y toda $\varphi \in \Phi$.

Una familia $\Phi = \{\varphi\}$ se llama *equicontinua*, cuando para cada $\varepsilon > 0$ hay un $\delta > 0$ tal que

$$|\varphi(x_1) - \varphi(x_2)| < \varepsilon$$

para todas las funciones $\varphi \in \Phi$ y para todo par x_1, x_2 de $[a, b]$ tal que $\rho(x_1, x_2) < \delta$.

TEOREMA 4 (ARZELÁ). *Para que una familia Φ de funciones continuas, definidas sobre el segmento $[a, b]$, sea relativamente compacta en $C_{[a, b]}$, es necesario y suficiente que esta familia sea equiacotada y equicontinua.*

DEMOSTRACION. NECESIDAD. Sea la familia Φ relativamente compacta en $C_{[a, b]}$. Entonces, de acuerdo con el teorema 3 del punto anterior, para cada ε positivo existe en Φ una $\frac{\varepsilon}{3}$-red finita $\varphi_1, \varphi_2, \ldots, \varphi_k$. Cada una de las funciones φ_i, siendo continua sobre un segmento, es acotada:

$$|\varphi_i(x)| \leqslant K_i.$$

Pongamos $K = \max K_i + \frac{\varepsilon}{3}$. Por definición de $\frac{\varepsilon}{3}$-red, para todo $\varphi \in \Phi$ tenemos, al menos para un φ_i,

$$\rho(\varphi_1, \varphi_i) = \max_x |\varphi(x) - \varphi_i(x)| \leqslant \frac{\varepsilon}{3}.$$

Por consiguiente,

$$|\varphi(x)| \leqslant |\varphi_i(x)| + \frac{\varepsilon}{3} \leqslant K_i + \frac{\varepsilon}{3} \leqslant K.$$

Es decir, Φ es una familia equiacotada.

Luego, puesto que cada una de las funciones φ_i que forman la $\frac{\varepsilon}{3}$-red es continua y, por consiguiente, uniformemente continua sobre $[a, b]$, para un $\frac{\varepsilon}{3}$ dado existe un δ_i tal que

$$|\varphi_i(x_1) - \varphi_i(x_2)| < \frac{\varepsilon}{3},$$

siempre que $|x_1 - x_2| < \delta_i$.

Sea $\delta = \min \delta_i$. Entonces, para $|x_1 - x_2| < \delta$ y para cualquier función $\varphi \in \Phi$, tomando φ_i de manera que $\rho(\varphi, \varphi_i) < \frac{\varepsilon}{3}$, obtenemos

$$\begin{aligned} \varphi(x_1) - \varphi(x_2)| &\leqslant \\ &\leqslant |\varphi(x_1) - \varphi_i(x_1)| + |\varphi_i(x_1) - \varphi_i(x_2)| + |\varphi_i(x_2) - \varphi(x_2)| < \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Hemos demostrado con ello la equicontinuidad de la familia Φ.

SUFICIENCIA. Sea Φ una familia equiacotada y equicontinua de funciones. De acuerdo con el teorema 3, para demostrar su compacidad relativa en $C_{[a, b]}$, es suficiente probar que existe en $C_{[a, b]}$ una ε-red finita de ella cualquiera que sea $\varepsilon > 0$. Sea

$$|\varphi(x)| \leqslant K \text{ para todos } \varphi \in \Phi$$

y sea $\delta > 0$ escogido de manera que

$$|\varphi(x_1) - \varphi(x_2)| < \frac{\varepsilon}{5} \text{ para } |x_1 - x_2| < \delta$$

y para todos $\varphi \in \Phi$. Dividamos el segmento $[a, b]$ del eje x mediante los puntos $x_0 = a_1 < x_1 < x_2 < \ldots < x_n = b$ en intervalos de longitud menor que δ y construyamos rectas verticales a través de estos puntos. Dividamos el segmento $[-K, K]$ del eje y mediante los puntos $y_0 = -K < y_1 < y_2 < \ldots < y_m = K$ en intervalos de longitud $< \frac{\varepsilon}{5}$ y construyamos rectas horizontales a través de estos puntos. De esta forma el rectángulo $a \leqslant x \leqslant b$, $-K \leqslant y \leqslant K$ resultará dividido en células con lado horizontal de longitud $< \delta$ y con lado vertical de longitud $< \frac{\varepsilon}{5}$. Asignemos ahora a cada función $\varphi \in \Phi$ la quebrada $\psi(x)$ con vértices en los puntos (x_k, y_l), es decir, en los nodos de la red construida, y que diverge de la función $\varphi(x)$ en los puntos x_k en menos que $\frac{\varepsilon}{5}$ (la existencia de una quebrada de este tipo es evidente).

Puesto que, por la construcción,

$$|\varphi(x_k) - \psi(x_k)| < \frac{\varepsilon}{5},$$

$$|\varphi(x_{k+1}) - \psi(x_{k+1})| < \frac{\varepsilon}{5}$$

y

$$|\varphi(x_k) - \varphi(x_{k+1})| < \frac{\varepsilon}{5},$$

tenemos

$$|\psi(x_k) - \psi(x_{k+1})| < \frac{3\varepsilon}{5}.$$

Puesto que la función $\psi(x)$ es lineal entre los puntos x_k y x_{k+1}, tenemos

$$|\psi(x_k) - \psi(x)| < \frac{3\varepsilon}{5} \text{ para todos } x \in [x_k, x_{k+1}].$$

Sea ahora x un punto arbitrario del segmento $[a, b]$ y sea x_k el punto de la división escogida más próximo a x por la izquierda. Entonces

$$|\varphi(x) - \psi(x)| \leqslant |\varphi(x) - \varphi(x_k)| + |\varphi(x_k) - \psi(x_k)| + \\ + |\psi(x_k) - \psi(x)| \leqslant \varepsilon.$$

Por consiguiente, las quebradas $\psi(x)$ representan una ε-red respecto a Φ. El número de ellas es finito; luego, Φ es totalmente acotada. Hemos demostrado completamente el teorema.

5°. Teorema de Peano. El teorema de Arzelá tiene múltiples aplicaciones. A título de aplicación suyo veamos el siguiente teorema de existencia para ecuaciones diferenciales ordinarias con el miembro derecho continuo.

TEOREMA 5 (Peano). *Sea*

$$\frac{dy}{dx}=f(x, y) \tag{3}$$

una ecuación diferencial dada. Si la función f es continua en un recinto cerrado G, al menos una curva integral de la ecuación dada pasa por cada punto interior (x_0, y_0) de este recinto.

DEMOSTRACION. Puesto que la función f es continua en un recinto cerrado, es acotada:

$$|f(x, y)| < M = \text{const.}$$

Tracemos por el punto (x_0, y_0) las rectas con pendiente M y $-M$. Tracemos, además, las rectas verticales $x=a$ y $x=b$ de manera que los dos triángulos con vértice común en (x_0, y_0) que ellas producen pertenezcan íntegramente al interior de la región G.

Construyamos ahora para la ecuación dada las así llamadas quebradas de Euler del siguiente modo: tracemos por el punto (x_0, y_0) la recta de pendiente $f(x_0, y_0)$. Tomemos en esta recta un punto (x_1, y_1) y tracemos, a través de él, una recta de pendiente $f(x_1, y_1)$. En esta recta tomemos un punto (x_2, y_2) y tracemos, a través de él, una recta de pendiente $f(x_2, y_2)$, etc. Consideremos ahora la sucesión de quebradas de Euler $L_1, L_2, \ldots, L_n, \ldots$, que pasan por el punto (x_0, y_0) y tales que la longitud del mayor de los eslabones de la línea L_k tiende a cero para $k \to \infty$. Sea φ_k la función, cuya gráfica es la línea L_k. Las funciones $\varphi_1, \varphi_2, \ldots, \varphi_n, \ldots$ poseen las siguientes propiedades:

1) están definidas en un mismo segmento $[a, b]$,
2) son equiacotadas,
3) son equicontinuas.

En virtud del teorema de Arzelá, se puede extraer de la sucesión $\{\varphi_k\}$ una subsucesión uniformemente convergente. Sea $\varphi^{(1)}, \varphi^{(2)}, \ldots, \varphi^{(k)}, \ldots$ esta sucesión.

Pongamos $\varphi(x)=\lim \varphi^{(k)}(x)$ para $k \to \infty$. Está claro que $\varphi(x_0)=y_0$. Resta probar que φ verifica sobre el segmento $[a, b]$ la ecuación diferencial dada. Para ello es necesario demostrar que cualquiera que sea $\varepsilon > 0$

$$\left|\frac{\varphi(x'')-\varphi(x')}{x''-x'}-f(x', \varphi(x'))\right| < \varepsilon,$$

siempre que la magnitud $|x''-x'|$ sea suficientemente pequeña. A su vez, para demostrar esto es preciso establecer para k suficientemente grande

$$\left|\frac{\varphi^{(k)}(x'')-\varphi^{(k)}(x')}{x''-x'}-f(x', \varphi^{(k)}(x'))\right| < \varepsilon,$$

siempre que la diferencia $|x''-x'|$ sea suficientemente pequeña.

Puesto que f es continua en la región G, para cualquier $\varepsilon > 0$ se puede encontrar un $\eta > 0$ tal que

$$f(x', y')-\varepsilon < f(x, y) < f(x', y')+\varepsilon$$
$$(y'=\varphi(x')),$$

siempre que

$$|x-x'| < 2\eta \quad \text{e} \quad |y-y'| < 4M\eta.$$

El conjunto de puntos $(x, y) \in G$, que verifican estas dos desigualdades, constituye un rectángulo Q. Sea ahora K tan grande que para todo $k > K$

$$|\varphi(x)-\varphi^{(k)}(x)| < 4\eta$$

y todos los eslabones de la quebrada L_k tienen longitud menor que η. Entonces, si $|x-x'|<2\eta$, todas las quebradas de Euler $\varphi^{(k)}$, correspondientes a $k>K$, se encuentran íntegramente en el interior de Q.

Además, sean (a_0, b_0), (a_1, b_1), ..., (a_{n+1}, b_{n+1}) los vértices de la quebrada $\varphi^{(k)}$, donde

$$a_0 \leqslant x' < a_1 < a_2 < \ldots < a_n < x'' \leqslant a_{n+1}$$

(suponemos, para concretar, que $x''>x'$; el caso $x''<x'$ se considera análogamente). Entonces,

$$\begin{gathered}\varphi^{(k)}(a_1)-\varphi^{(k)}(x')=f(a_0, b_0)(a_1-x'),\\ \varphi^{(k)}(a_{i+1})-\varphi^{(k)}(a_i)=f(a_i, b_i)(a_{i+1}-a_i); \quad i=1, 2, \ldots, n-1,\\ \varphi^{(k)}(x'')-\varphi^{(k)}(a_n)=f(a_n, b_n)(x''-a_n).\end{gathered}$$

De aquí, para $|x''-x'|<\eta$, obtenemos

$$\begin{gathered}[f(x', y')-\varepsilon](a_1-x')<\varphi^{(k)}(a_1)-\varphi^{(k)}(x')<[f(x', y')+\varepsilon](a_1-x'),\\ [f(x', y')-\varepsilon](a_{i+1}-a_i)<\varphi^{(k)}(a_{i+1})-\varphi^{(k)}(a_i)<\\ <[f(x', y')+\varepsilon](a_{i+1}-a_i); \quad i=1, \ldots, n-1,\\ [f(x', y')-\varepsilon](x''-a_n)<\varphi^{(k)}(x'')-\varphi^{(k)}(a_n)<[f(x', y')+\varepsilon](x''-a_n).\end{gathered}$$

Sumando estas desigualdades, encontramos

$$[f(x', y')-\varepsilon](x''-x')<\varphi^{(k)}(x'')-\varphi^{(k)}(x')<[f(x', y')+\varepsilon](x''-x'),$$

que es lo que queríamos demostrar.

Diferentes subsucesiones de la sucesión de quebradas de Euler pueden converger a diferentes soluciones de la ecuación (3). Por eso, la solución φ que hemos obtenido no es, en general, la única solución de la ecuación $y'=f(x, y)$ que pasa por el punto (x_0, y_0).

6. Teorema generalizado de Arzelá. Sean X e Y dos compactos métricos y sea C_{XY} el conjunto de todas las aplicaciones continuas f del compacto X en Y. Definamos la distancia en C_{XY} mediante la fórmula

$$\rho(f, g)=\sup_{x\in X}\rho(f(x), g(x)).$$

Es fácil comprobar que C_{XY} se convierte de esta forma en un espacio métrico.

TEOREMA 6 (teorema generalizado de Arzelá). *Para que un conjunto $D\subset C_{XY}$ sea relativamente compacto, es necesario y suficiente que las funciones f que integran D sean equicontinuas, es decir, que para cualquier $\varepsilon>0$ exista un $\delta>0$ tal que de*

$$\rho(x', x'')<\delta \tag{4}$$

se deduzca

$$\rho(f(x'), f(x''))<\varepsilon \tag{5}$$

cualesquiera que sean f de D y x' y x'' de X.

DEMOSTRACION. La necesidad se demuestra igual que en el teorema 4.

Demostremos la suficiencia. Para ello sumerjamos C_{XY} en el espacio M_{XY} de todas las aplicaciones del compacto X en el compacto Y con la misma métrica

$$\rho(f, g) = \sup_{x \in X} \rho(f(x), g(x))$$

que ha sido introducida en C_{XY}, y demostremos la compacidad relativa del conjunto D en M_{XY}. Puesto que C_{XY} es cerrado en M_{XY}[1], la compacidad relativa del conjunto D en M_{XY} implica su compacidad relativa en C_{XY}.

Tomemos arbitrariamente $\varepsilon > 0$ y escojamos δ de manera que de (4) se siga (5) para todo f de D y todo x', x'' de X. Es fácil ver que X se puede representar como la unión de conjuntos disjuntos E_i tales que de $x', x'' \in E_i$ se deduce que $\rho(x', x'') < \delta$. En efecto, para ello es suficiente escoger los puntos $x_1, x_2, \ldots, x_n$ de manera que formen una $\frac{\delta}{2}$-red en X y tomar, por ejemplo,

$$E_i = S(x_i, \delta) - \bigcup_{j<i} S(x_j, \delta).$$

Consideremos ahora en el compacto Y una ε-red finita $y_1, y_2, \ldots, y_m$; denotemos mediante L la colección de funciones $g(x)$ que toman los valores y_j sobre los conjuntos E_i. El número de estas funciones es, evidentemente, finito. Probemos que forman una 2ε-red respecto a D en M_{XY}. En efecto, sea $f \in D$. Para todo punto x_i de $x_1, \ldots, x_n$ existe un punto y_j de $y_1, \ldots, y_m$ tal que

$$\rho(f(x_i), y_j) < \varepsilon.$$

Sea la función $g \in L$ escogida de manera que $g(x_i) = y_i$. Entonces,

$$\rho(f(x), g(x)) \leqslant \rho(f(x), f(x_i)) + \rho(f(x_i), g(x_i)) + \\ + \rho(g(x_i), g(x)) < 2\varepsilon,$$

si i se ha escogido de manera que $x \in E_i$.

De aquí se desprende que $\rho(f, g) < 2\varepsilon$ y, por consiguiente, la compacidad de D en M_{XY} y, por lo tanto, en C_{XY} queda demostrada.

---

[1] Esto se debe a que el límite de una sucesión uniformemente convergente de aplicaciones continuas es también una aplicación continua. La proposición enunciada representa una generalización directa del teorema conocido del Análisis y se demuestra igual que este teorema.

§ 8. Funciones reales sobre espacios métricos y topológicos

1°. Funciones y funcionales continuas y uniformemente continuas. Una función real sobre un espacio topológico (en particular, métrico) T es una aplicación del espacio T en el espacio R^1 (la recta numérica). Así, por ejemplo, una función real sobre el espacio R^n de n dimensiones es la función ordinaria de n variables.

Cuando el propio espacio T se compone de funciones, las funciones definidas sobre él se llaman *funcionales*. Veamos algunos ejemplos de funcionales de funciones x, definidas sobre el segmento $[0, 1]$:

$$F_1(x)=\sup x(t);$$
$$F_2(x)=\inf x(t);$$
$$F_3(x)=x(t_0),\ \text{donde}\ t_0\in[0, 1];$$
$$F_4(x)=\varphi[x(t_0),\ x(t_1),\ \dots,\ x(t_n)],\ \text{donde}\ t_i\in[0, 1]$$

y la función $\varphi(s_0, s_1, \dots, s_n)$ está definida para todo s_i real;

$$F_5(x)=\int_0^1\varphi[t,\ x(t)]\,dt,\ \text{donde}\ \varphi(t, s)\ \text{está definida y}$$

es continua para todo $0\leqslant t\leqslant 1$ y todo s real;

$$F_6(x)=x'(t_0),\ \text{donde}\ t_0\in[0, 1];$$
$$F_7(x)=\int_0^1\sqrt{1+x'^2(t)}\,dt;$$
$$F_8(x)=\int_0^1|x'(t)|\,dt.$$

Las funcionales F_1, F_2, F_3, F_4 y F_5 están definidas sobre el espacio C de todas las funciones continuas sobre el segmento $[0, 1]$; F_6 está definida sólo para las funciones diferenciables en el punto t_0; F_7 para aquellas funciones para las cuales la expresión $\sqrt{1+x'^2(t)}$ es integrable y F_8 para funciones para las cuales $|x'(t)|$ es integrable.

La funcional $F_1(x)$ es continua sobre C, ya que

$$\rho(x, y)=\sup|x-y|\ \text{y}\ |\sup x-\sup y|\leqslant\sup|x-y|.$$

Las funcionales F_2, F_3 y F_5 son también continuas sobre C; la funcional F_4 es continua sobre C, cuando la función φ que lo determina es continua respecto a todos sus argumentos. La funcional F_6 es discontinua en todo punto de C donde está definida. En efecto, sea $x(t)$ una función tal que $x'(t_0)=1$ y $|x(t)|<\varepsilon$ y sea $y=x_0+x$. Entonces, $y'(t_0)=x_0'(t_0)+1$, mientras que

$\rho(y, x_0) < \varepsilon$. Esta funcional F_6 resultará continua, si se considera sobre el espacio $C^{(1)}$, compuesto de funciones que tienen derivada continua y provisto de la métrica

$$\rho(x, y) = \sup_t [\,|x(t) - y(t)| + |x'(t) - y'(t)|\,].$$

La funcional F_7 es también discontinua sobre el espacio C. En efecto, sea $x_0(t) \equiv 0$ y $x_n(t) = \frac{1}{n}$ sen $2\pi nt$. Entonces, $\rho(x_n, x_0) = \frac{1}{n} \longrightarrow 0$; sin embargo, $F_7(x_n) > 4$ para todo n, mientras que $F_7(x_0) = 1$. Por consiguiente, $F_7(x_n)$ no tiende a $F_7(x_0)$, cuando $x_n \longrightarrow x_0$. El mismo ejemplo sirve para demostrar que la funcional F_8 es también discontinua sobre el espacio C. Ambas funcionales F_7 y F_8 son continuas en el espacio $C^{(1)}$.

Para las funciones, definidas sobre un espacio métrico, subsiste el concepto habitual de continuidad uniforme: la función $f(x)$ es uniformemente continua sobre un espacio métrico R, cuando para todo $\varepsilon > 0$ existe un $\delta > 0$ tal que

$$|f(x_1) - f(x_2)| < \varepsilon, \text{ siempre que } \rho(x_1, x_2) < \delta.$$

Por ejemplo, la funcional F_1 es uniformemente continua sobre el espacio C (¡compruébese esto!).

Para las funciones reales sobre compactos métricos tiene lugar el siguiente teorema, que generaliza el teorema bien conocido del curso elemental de Análisis acerca de las funciones continuas sobre un segmento.

TEOREMA 1. *Una función real continua sobre un compacto métrico es uniformemente continua.*

DEMOSTRACION. Supongamos que $f(x)$ es continua, pero no uniformemente continua, sobre un compacto métrico K. Entonces, para cierto ε positivo y cualquier n natural existirán x_n y x'_n de K tales que

$$\rho(x_n, x'_n) < \frac{1}{n}, \text{ mientras que } |f(x_n) - f(x'_n)| \geqslant \varepsilon.$$

De la sucesión $\{x_n\}$ se puede extraer, debido a la compacidad de K, una subsucesión $\{x_{n_k}\}$ convergente a un punto $x \in K$. En este caso $\{x'_{n_k}\}$ también converge a x; pero para cada k debe cumplirse al menos una de las desigualdades

$$|f(x) - f(x'_{n_k})| \geqslant \frac{\varepsilon}{2}, \quad |f(x) - f(x_{n_k})| \geqslant \frac{\varepsilon}{2}$$

y esto contradice a la continuidad de la función f en el punto x.

2°. Funciones continuas y semicontinuas sobre espacios compactos. Más arriba hemos visto que el teorema sobre la continuidad uniforme de una función continua sobre un segmento subsiste para las funciones definidas sobre cualesquiera compactos métricos. En cuanto a las demás propiedades de funciones continuas sobre un segmento, conocidas del Análisis, algunas de ellas se extienden, como veremos ahora, a espacios compactos cualesquiera (no necesariamente métricos).

TEOREMA 2. *Sea T un espacio compacto y f una función continua sobre él. Entonces, f es acotada sobre T y alcanza sobre T sus extrèmos superior e inferior.*

DEMOSTRACION. Una función continua es una aplicación continua de T en la recta numérica R^1. La imagen de T en R^1 es, de acuerdo con el teorema general 3 del § 6, compacta. Pero un subconjunto compacto de la recta numérica es cerrado y acotado y, por consiguiente, no sólo tiene extremos superior e inferior finitos, sino que contiene incluso estos extremos. El teorema queda demostrado.

EJERCICIO. Sea K un compacto métrico y A una aplicación de K en si mismo tal que $\rho(Ax, Ay) < \rho(x, y)$ para $x \neq y$. Demuéstrese que la aplicación A tiene en K un punto fijo único.

Las afirmaciones del último teorema admiten una generalización al caso de funciones de una clase más amplia, a saber, al caso de las así llamadas funciones semicontinuas.

Una función $f(x)$ se llama *semicontinua inferiormente (superiormente)* en el punto x_0, cuando para todo $\varepsilon > 0$ existe una vecindad del punto x_0 tal que $f(x) > f(x_0) - \varepsilon$ ($f(x) < f(x_0) + \varepsilon$).

Por ejemplo, la función «parte entera de x», $f(x) = E(x)$ es semicontinua superiormente. Si aumentamos (disminuimos) el valor $f(x_0)$ de una función continua en un punto x_0, obtendremos una función semicontinua superiormente (inferiormente). Si $f(x)$ es semicontinua superiormente, la función $-f(x)$ es semicontinua inferiormente. Estas dos observaciones permiten obtener inmediatamente un gran número de ejemplos de funciones semicontinuas.

Para el estudio de las propiedades de semicontinuidad de funciones reales conviene permitirles que tomen valores infinitos. Si $f(x_0) = -\infty$, consideraremos que la función f es semicontinua inferiormente en x_0; en cambio, si para todo $h > 0$ existe una vecindad del punto x_0 tal que $f(x) < -h$, admitiremos que la función f es semicontinua también superiormente en el punto x_0.

Si $f(x_0) = +\infty$, consideraremos que la función f es semicontinua superiormente en x_0; en cambio, si para todo $h > 0$ existe una vecindad del punto x_0 tal que $f(x) > h$, admitiremos que la función f es semicontinua también inferiormente en el punto x_0.

Sea $f(x)$ una función real sobre un espacio métrico R. Se llama *límite superior* $\overline{f}(x_0)$ de la función $f(x)$ en el punto x_0 a la magnitud (finita o infinita) $\lim_{\varepsilon\to 0}\left[\sup_{x\in S(x_0,\varepsilon)} f(x)\right]$. Se llama *límite inferior* $\underline{f}(x_0)$ a la magnitud (finita o infinita) $\lim_{\varepsilon\to 0}\left[\inf_{x\in S(x_0,\varepsilon)} f(x)\right]$. La diferencia $\omega f(x_0)=\overline{f}(x_0)-\underline{f}(x_0)$ (si es que tiene sentido, es decir, si al menos uno de los números $\overline{f}(x_0)$, $\underline{f}(x_0)$ es finito) se llama *oscilación* de la función $f(x)$ en el punto x_0. Es fácil ver que para la continuidad de $f(x)$ en el punto x_0 es necesario y suficiente que $\omega f(x_0)=0$, es decir, que $-\infty<\underline{f}(x_0)=\overline{f}(x_0)<+\infty$.

Para cualquier función $f(x)$, la función $\overline{f}(x)$ es semicontinua superiormente y la función $\underline{f}(x)$, semicontinua inferiormente. Esto se obtiene fácilmente de la definición de los límites superior e inferior.

Veamos un ejemplo importante de una funcional semicontinua.

Definamos la longitud de la curva $y=f(x)$ $(a\leqslant x\leqslant b)$ mediante la funcional

$$L_a^b(f)=\sup\sum_{i=1}^{n}\sqrt{(x_i-x_{i-1})^2+(f(x_i)-f(x_{i-1}))^2},$$

donde el extremo superior (que puede ser igual a $+\infty$) se toma respecto a todas las divisiones posibles del segmento $[a, b]$. Esta funcional está definida sobre todo el espacio M de las funciones reales acotadas. Para las funciones continuas coincide con el valor del límite

$$\lim_{\max|x_i-x_{i-1}|\to 0}\sum_{i=1}^{n}\sqrt{(x_i-x_{i-1})^2+(f(x_i)-f(x_{i-1}))^2}.$$

Finalmente, en caso de funciones con derivada continua puede ser representada en la forma

$$\int_a^b\sqrt{1+f'^2(x)}\,dx.$$

La funcional $L_a^b(f)$ es semicontinua inferiormente en M, como se deduce fácilmente de su definición.

El teorema, establecido anteriormente, se generaliza al caso de funciones semicontinuas.

TEOREMA 2a. *Una función finita semicontinua inferiormente (superiormente) sobre un espacio compacto T está acotada inferiormente (superiormente).*

En efecto, supongamos que $\inf f(x) = -\infty$. Entonces, existe una sucesión $\{x_n\}$ tal que $f(x_n) < -n$. Puesto que el espacio T es compacto, el subconjunto infinito $\{x_n\}$ suyo tiene al menos un punto de acumulación x_0. Por hipótesis la función f es finita y semicontinua inferiormente; por eso, existirá una vecindad U del punto x_0 tal que $f(x) > f(x_0) - 1$ para $x \in U$. Pero en este caso la vecindad U puede contener solamente un número finito de puntos del conjunto $\{x_n\}$ y esto contradice a que x_0 es un punto de acumulación de este conjunto.

De manera análoga se demuestra el teorema en el caso de una función semicontinua superiormente.

TEOREMA 2b. *Una función finita semicontinua inferiormente (superiormente) sobre un espacio compacto T alcanza su extremo inferior (superior).*

Supongamos que la función $f(x)$ es semicontinua inferiormente. Entonces, de acuerdo con el teorema 2a, tiene un extremo inferior finito y, además, existe una sucesión $\{x_n\}$ tal que $f(x_n) \leqslant \leqslant \inf f(x) + \frac{1}{n}$.

Debido a la compacidad de T, el conjunto $\{x_n\}$ tiene un punto de acumulación x_0. Si fuese $f(x_0) > \inf f$, existirían, en virtud de la semicontinuidad inferior de la función f, una vecindad U del punto x_0 y un $\delta > 0$ tales que $f(x) > \inf f + \delta$ para $x \in U$. Pero en este caso la vecindad U no podría contener ningún subconjunto infinito del conjunto $\{x_n\}$. Por consiguiente, $f(x_0) = \inf f$, que es lo que se quería demostrar.

§ 9. CURVAS CONTINUAS EN ESPACIOS MÉTRICOS[1]

Sea dada una aplicación continua

$$P = f(t)$$

del segmento

$$a \leqslant t \leqslant b$$

en un espacio métrico R. Cuando t «recorre» el segmento desde a hasta b, el punto correspondiente P «recorre» una «curva continua» en el espacido R. Nos proponemos dar unas definiciones rigurosas, relacionadas con la idea tosca que acabamos de exponer. Consideraremos que el orden en el que se recorren los puntos de la curva es una propiedad esencial de la propia curva. Un mismo conjunto, indicado en la fig. 13, recorrido en las direcciones señaladas en las figs. 14 y 15, será considerado como diferentes curvas. A título de un ejemplo más consideremos la función real, definida sobre el segmento [0, 1], que viene representada en la fig. 16. Representa una «curva», situada

[1]) Este parágrafo no está relacionado con la exposición sucesiva. El lector puede, si lo desea, omitirlo.

en el segmento [0, 1] del eje y y distinta de este segmento recorrido una vez desde el punto 0 hasta el punto 1, ya que el segmento $[A, B]$ se pasa tres veces (dos hacia arriba y una hacia abajo).

Sin embargo, si los puntos del espacio se recorren en el mismo orden, consideraremos que la selección del «parámetro» t no es esencial. Por ejemplo, las funciones representadas en las figs. 16 y 17, determinan una misma «curva», situada sobre el eje y, aun cuando los valores del parámetro t, correspondientes a algún punto de la curva, resulten distintos en los casos de la fig. 16 y la fig. 17. Por ejemplo, en el caso de la fig. 16, al punto A le corresponden sobre el eje t dos puntos aislados, mientras que en el caso de la fig. 17, corresponden sobre el eje t un punto aislado y un segmento, situado a su derecha (cuando t recorre este segmento, el punto de la curva se mantiene fijo) [1]).

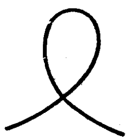


FIG. 13

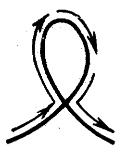


FIG. 14

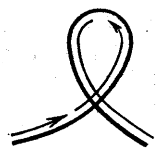


FIG. 15

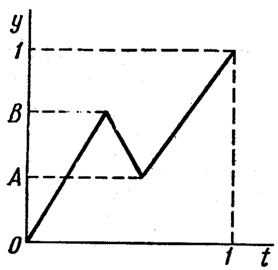


FIG. 16

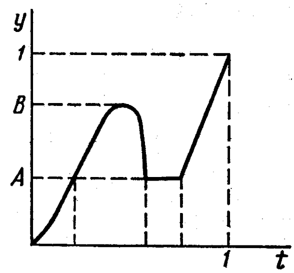


FIG. 17

Pasemos a las definiciones formales. Dos funciones continuas

$$P=f'(t') \quad \text{y} \quad P=f''(t'')$$

definidas, respectivamente, sobre los segmentos

$$a' \leqslant t' \leqslant b' \quad \text{y} \quad a'' \leqslant t'' \leqslant b''$$

y con valores en un espacio métrico R, se llaman *equivalentes*, cuando existen dos funciones continuas no decrecientes

$$t'=\varphi'(t) \quad \text{y} \quad t''=\varphi''(t),$$

definidas sobre un segmento

$$a \leqslant t \leqslant b,$$

[1]) En vista del estudio ulterior de la compacidad de sistemas de curvas conviene consentir estos segmentos de constancia del punto $P=f(t)$.

que poseen las propiedades

$$\varphi'(a)=a', \quad \varphi'(b)=b',$$
$$\varphi''(a)=a'', \quad \varphi''(b)=b'',$$
$$f'[\varphi'(t)]=f''[\varphi''(t)]$$

para todo $t \in [a, b]$.

Es fácil ver que la relación de equivalencia así introducida es reflexiva (f es equivalente a f), simétrica (si f' es equivalente a f'', también f'' es equivalente a f') y transitiva (de la equivalencia de f' y f'' y de la equivalencia de f'' y f''' se deduce la equivalencia de f' y f'''). Por eso, todas las funciones continuas del tipo considerado se dividen en clases de funciones equivalentes entre sí. Cada una de estas clases define una *curva continua* en el espacio R.

Es fácil ver que para toda función $P=f'(t')$, definida sobre un segmento $[a', b']$, existe una función equivalente a ella, definida sobre el segmento $[a'', b'']=[0, 1]$. En efecto, es suficiente tomar [1]

$$t'=\varphi'(t)=(b'-a')\,t+a', \quad t''=\varphi''(t)=t.$$

Por consiguiente, podemos suponer que toda curva viene dada en forma paramétrica mediante una función definida sobre el segmento $[0, 1]$.

Por eso resulta oportuno introducir el espacio $C(I, R)$ de aplicaciones continuas f del segmento $I=[0, 1]$ en el espacio R con la métrica

$$\rho(f, g)=\sup_t \rho(f(t), g(t)).$$

Admitiremos que la sucesión de curvas $L_1, L_2, \ldots, L_n, \ldots$ converge a la curva L, cuando las curvas L_n pueden ser representadas paramétricamente en la forma

$$P=f_n(t), \quad 0 \leqslant t \leqslant 1$$

y la curva L, en la forma

$$P=f(t), \quad 0 \leqslant t \leqslant 1$$

de manera que $\rho(f, f_n) \longrightarrow 0$ para $n \longrightarrow \infty$.

Aplicando el teorema generalizado de Arzelá (teorema 6 del § 7), es fácil demostrar el siguiente teorema.

TEOREMA 1. *Si la sucesión de curvas $L_1, L_2, \ldots, L_n, \ldots$, situadas en un compacto K, se puede representar paramétricamente mediante funciones equicontinuas sobre el segmento $[0, 1]$, se puede extraer de ella una subsucesión convergente.*

Definamos ahora la longitud de una curva, que en forma paramétrica viene dada por la función

$$P=f(t), \quad a \leqslant t \leqslant b,$$

como el extremo superior de las sumas de tipo

$$\sum_{i=1}^{n} \rho(f(t_{i-1}), f(t_i)),$$

[1]) Admitimos siempre que $a < b$. Sin embargo, no excluimos las «curvas» que constan de un solo punto y que se obtienen cuando la función $f(t)$ es constante sobre $[a, b]$. Este acuerdo también resulta oportuno para lo sucesivo.

donde los puntos t_i están sujetos solamente a las condiciones

$$a=t_0 \leqslant t_1 \leqslant \ldots \leqslant t_i \leqslant \ldots \leqslant t_n=b.$$

Es fácil ver que la longitud de una curva no depende de su representación paramétrica. Si nos limitamos a las representaciones paramétricas por medio de funciones, definidas sobre el segmento [0, 1], es fácil demostrar que la longitud de una curva es una funcional semicontinua inferiormente de f (en el espacio $C(I, R)$). En el lenguaje geométrico este resultado puede ser enunciado en forma del siguiente teorema sobre semicontinuidad.

TEOREMA 2. *Si la sucesión de curvas L_n converge a la curva L, la longitud de la curva L no es mayor que el límite inferior de las longitudes de las curvas L_n.*

Consideremos ahora especialmente las *curvas de longitud finita*. Supongamos que la curva está definida por la función paramétrica

$$P=f(t), \quad a \leqslant t \leqslant b.$$

La función f, considerada solamente en el segmento $[a, T]$, donde $a \leqslant T \leqslant b$, define el «segmento inicial» de la curva desde el punto

$$P_a=f(a)$$

hasta el punto

$$P_T=f(T).$$

Sea

$$s=\varphi(t)$$

su longitud. Es fácil probar que

$$P=g(s)=f[\varphi^{-1}(s)]$$

es una nueva representación paramétrica de la misma curva. Aquí s recorre el segmento

$$0 \leqslant s \leqslant S,$$

donde S es la longitud de toda la curva considerada. Esta representación verifica la exigencia

$$\rho(g(s_1), g(s_2)) \leqslant |s_2-s_1|$$

(la longitud del arco es no menos que la longitud de la cuerda).

Pasando al segmento [0, 1], obtenemos la representación paramétrica

$$P=F(\tau)=g(s), \tau=\frac{s}{S},$$

que verifica la condición de Lipschitz

$$\rho(F(\tau_1), F(\tau_2)) \leqslant S|\tau_1-\tau_2|.$$

Vemos, por consiguiente, que para todas las curvas de longitud

$$S \leqslant M,$$

donde M es una constante, es posible una representación paramétrica mediante funciones equicontinuas, definidas sobre ei segmento [0, 1]. Por lo tanto, a ellas es aplicable el teorema 1.

Mostremos el alcance de los resultados generales obtenidos aplicándolos a la demostración de la siguiente proposición importante.

TEOREMA 3. *Si dos puntos A y B de un compacto K pueden unirse por medio de una curva continua de longitud finita, entre estas curvas existe la de longitud mínima.*

En efecto, sea Y el extremo inferior de las longitudes de las curvas, que unen los puntos A y B del compacto K. Supongamos que las longitudes de las curvas $L_1, L_2, \ldots, L_n, \ldots$, que unen los puntos A y B, tienden a Y. De acuerdo con el teorema 1, de la sucesión L_n se puede extraer una subsucesión convergente. De acuerdo con el teorema 2, la curva límite de esta subsucesión no puede tener longitud mayor que Y.

Observemos, que incluso en el caso, en que K es una superficie cerrada suave (diferenciable suficiente número de veces) del espacio euclídeo de tres dimensiones, este teorema no se desprende directamente de los resultados que se establecen en el curso de Geometría Diferencial, donde se considera, generalmente, sólo el caso de puntos A y B suficientemente próximos.

Todo lo expuesto adquiriría mayor claridad, si hubiesemos provisto de una estructura de espacio métrico el conjunto de todas las curvas del espacio métrico dado R. Esto se puede hacer definiendo la distancia entre las curvas L_1 y L_2 mediante la fórmula

$$\rho(L_1, L_2) = \inf \rho(f_1, f_2),$$

donde el extremo inferior se toma respecto a todos los pares de representaciones paramétricas de la curva L_1 por medio de la función

$$P = f_1(t), \qquad 0 \leqslant t \leqslant 1,$$

y de la curva L_2 por medio de la función

$$P = f_2(t), \qquad 0 \leqslant t \leqslant 1.$$

La demostración de que ésta distancia verifica los axiomas corrientes es sencilla, a excepción de un momento: ofrece ciertas dificultades demostrar que de

$$\rho(L_1, L_2) = 0$$

se deduce la identidad de las curvas L_1 y L_2. Este resultado es consecuencia directa del hecho de que el extremo inferior en la fórmula, mediante la cual hemos definido la distancia $\rho(L_1, L_2)$ se alcanza, si se escogen adecuadamente las representaciones paramétricas f_1 y f_2. Pero la demostración de esta última proposición tampoco es sencilla.

CAPITULO III

ESPACIOS LINEALES NORMADOS Y TOPOLOGICOS

§ 1. ESPACIOS LINEALES

El concepto de espacio lineal es uno de los más importantes en las Matemáticas. Desempeñará un papel primordial no sólo en este capítulo, sino también en toda la exposición sucesiva.

1°. Definición y ejemplos de espacios lineales.

DEFINICION 1. Un conjunto no vacío L de elementos x, y, z, ... se llama *espacio lineal*, o *vectorial*, cuando satisface las siguientes condiciones:

I. Para cualesquiera dos elementos $x, y \in L$ está definido unívocamente un tercer elemento $z \in L$, llamado suma de ellos y denotado $x+y$, tal que

1) $x+y=y+x$ (conmutatividad),
2) $x+(y+z)=(x+y)+z$ (asociatividad),
3) en L existe un elemento 0 tal que $x+0=x$ para todo $x \in L$ (existencia del cero),
4) para todo $x \in L$ existe un elemento $-x$ tal que $x+(-x)=0$ (existencia del elemento opuesto).

II. Para cualquier número α y cualquier elemento $x \in L$ está definido el elemento $\alpha x \in L$ (producto del elemento x por el número α) de manera que

1) $\alpha(\beta x)=(\alpha\beta)x$,
2) $1 \cdot x=x$.

III. Las operaciones de adición y multiplicación están relacionadas entre sí mediante las leyes distributivas:

1) $(\alpha+\beta)x=\alpha x+\beta x$,
2) $\alpha(x+y)=\alpha x+\alpha y$.

En dependencia del conjunto de los números que se admite (todos los complejos o solamente los reales), se distinguen los espacios lineales complejos y reales[1]. A menos que no se diga lo contrario, nuestros razonamientos serán válidos tanto para los espacios complejos, como para los reales.

Observemos que todo espacio lineal complejo puede ser considerado como real, si nos limitamos a la multiplicación de los vectores por números reales.

Veamos algunos ejemplos de espacios lineales, dejando a cargo del lector la comprobación, en cada uno de ellos, de los axiomas enunciados anteriormente.

1. La recta numérica, es decir, el conjunto de los números reales con las operaciones habituales de adición y multiplicación, representa un espacio lineal.

2. El espacio vectorial de n dimensiones, es decir, el conjunto de todos los sistemas posibles de n números (reales o complejos) $x=(x_1, x_2, \ldots, x_n)$, en el que la adición y la multiplicación se definen mediante las fórmulas

$$(x_1, x_2, \ldots, x_n)+(y_1, y_2, \ldots, y_n)=(x_1+y_1, x_2+y_2, \ldots, x_n+y_n),$$
$$\alpha(x_1, x_2, \ldots, x_n)=(\alpha x_1, \alpha x_2, \ldots, \alpha x_n),$$

es también un espacio lineal. Se denomina espacio aritmético de n dimensiones[2] y se denota mediante R^n en el caso real y C^n en el caso complejo.

3. Las funciones continuas (reales o complejas) sobre un segmento $[a, b]$ con las operaciones habituales de adición de funciones y multiplicación de funciones por números constituyen el espacio lineal $C_{[a, b]}$, uno de los más importantes para el Análisis.

4. El espacio l_2, cuyos elementos son las sucesiones de números (reales o complejos)

$$x=(x_1, x_2, \ldots, x_n, \ldots)$$

que verifican la condición

$$\sum_{n=1}^{\infty} |x_n|^2 < \infty, \tag{1}$$

con las operaciones

$$(x_1, x_2, \ldots, x_n, \ldots)+(y_1, y_2, \ldots, y_n, \ldots)=$$
$$=(x_1+y_1, x_2+y_2, \ldots, x_n+y_n, \ldots),$$
$$\alpha(x_1, x_2, \ldots, x_n, \ldots)=(\alpha x_1, \alpha x_2, \ldots, \alpha x_n, \ldots),$$

es un espacio lineal. El hecho de que la suma de dos sucesiones, que satisfacen la condición (1), también verifica esta condi-

[1] Podrían considerarse también espacios lineales sobre un cuerpo cualquiera.

[2] Este término se explicará más en adelante.

ción, se desprende de la desigualdad elemental

$$(a_1+a_2)^2 \leqslant 2a_1^2+2a_2^2.$$

5. Las sucesiones convergentes $x=(x_1, x_2, \ldots)$, con la adición y multiplicación por números realizadas respecto a las coordenadas, forman un espacio lineal. Denotémoslo c.

6. Las sucesiones convergentes a 0, con las mismas operaciones de adición y multiplicación, forman también un espacio lineal. Denotémoslo c_0.

7. El conjunto m de todas las sucesiones numéricas acotadas, con las operaciones de adición y multiplicación por números definidas igual que en los ejemplos 4, 5 y 6, también representa un espacio lineal.

8. Finalmente, el conjunto R^∞ de todas las sucesiones numéricas, con las mismas operaciones de adición y multiplicación por números que en los ejemplos 4, 5, 6 y 7, es también un espacio lineal.

Puesto que las propiedades de un espacio lineal son las propiedades de adición y multiplicación por números de sus elementos, resulta natural introducir la siguiente definición.

DEFINICION 2. Dos espacios lineales L y L^* se llaman *isomorfos*, cuando se puede establecer entre sus elementos una correspondencia biunívoca compatible con las operaciones en L y L^*. Esto significa que de

$$x \leftrightarrow x^*$$

e

$$y \leftrightarrow y^*$$

$(x,\ y \in L,\ x^*,\ y^* \in L^*)$, se sigue

$$x+y \leftrightarrow x^*+y^*$$

y

$$\alpha x \leftrightarrow \alpha x^*$$

(α es un número arbitrario).

Conviene a veces considerar los espacios isomorfos como diferentes realizaciones de un mismo espacio. A título de ejemplo de espacios lineales isomorfos pueden servir el espacio aritmético de n dimensiones (real o complejo) y el espacio de todos los polinomios de potencia $\leqslant n-1$ (con coeficientes reales o complejos, respectivamente) (¡demuéstrese el isomorfismo de estos espacios!)

2°. Dependencia lineal. Los elementos $x, y, \ldots, w$ de un espacio lineal L se llaman *linealmente dependientes*, cuando existen unos números $\alpha, \beta, \ldots, \lambda$, no todos iguales a 0, tales que

$$\alpha x+\beta y+\ldots+\lambda w=0. \tag{2}$$

En el caso contrario, estos elementos se llaman linealmente independientes. En otras palabras, los elementos $x, y, \ldots, w$ son linealmente independientes, cuando de la igualdad (2) se sigue que

$$\alpha=\beta=\ldots=\lambda=0.$$

Un sistema infinito de elementos $x, y, \ldots$ del espacio L se llama linealmente independiente, cuando todo subsistema finito suyo es linealmente independiente.

Si en un espacio L se pueden encontrar n elementos linealmente independientes y cualesquiera $n+1$ elementos de este espacio son linealmente dependientes, se dice que el espacio L tiene *dimensión* n. En cambio, si en L se puede indicar un sistema, compuesto por un número finito cualquiera de elementos linealmente independientes, se dice que el espacio L *es de dimensión infinita*. Se llama *base* de un espacio L de n dimensiones a todo sistema de n elementos linealmente independientes. Es fácil comprobar que los espacios R^n en el caso real y C^n en el caso complejo son de dimensión n, justificando de esta forma su denominación.

En el curso del Algebra Lineal se consideran espacios lineales de dimensión finita. Nosotros, al contrario, nos dedicaremos, como regla general, a los espacios de dimensión infinita que, desde el punto de vista del Análisis, representan el mayor interés. Proponemos al lector comprobar que cada uno de los espacios, señalados en los ejemplos 3, 4, 5, 6, 7 y 8, es de dimensión infinita.

3°. Subespacios. Un subconjunto no vacío L' de un espacio lineal L se llama *subespacio*, cuando representa un espacio lineal respecto de las operaciones de adición y multiplicación por números definidas en L.

En otras palabras, $L' \subset L$ es un subespacio, cuando de $x \in L'$, $y \in L'$ se deduce que $\alpha x+\beta y \in L'$ cualesquiera que sean α y β.

En todo espacio lineal L existe el subespacio formado solamente del elemento cero, el subespacio nulo. Por otro lado, todo el L puede ser considerado como un subespacio suyo. Un subespacio, diferente de L, que contiene al menos un elemento no nulo, se llama *propio*.

Veamos ejemplos de subespacios propios.

1. Sea L un espacio lineal y x un elemento suyo no nulo. El conjunto de elementos $\{\lambda x\}$, donde λ toma todos los valores

numéricos (reales o complejos), forma, evidentemente, un subespacio unidimensional. Este subespacio es propio, si la dimensión de L es mayor que 1.

2. Consideremos el espacio de funciones continuas $C_{[a, b]}$ (ejemplo 3) y en él el conjunto de todos los polinomios $P_{[a, b]}$. Está claro que los polinomios forman en $C_{[a, b]}$ un subespacio (de dimensión infinita al igual que todo el $C_{[a, b]}$). Al mismo tiempo, el propio espacio $C_{[a, b]}$ puede ser considerado como un subespacio de un espacio más amplio de todas las funciones, tanto continuas, como discontinuas, sobre $[a, b]$.

3. Consideremos, finalmente, los espacios l_2, c_0, c, m y R^{∞} (ejemplos 4, 5, 6, 7 y 8 del punto 1). Cada uno de ellos es un subespacio propio del siguiente.

Sea $\{x_\alpha\}$ un conjunto no vacío cualquiera de elementos de un espacio lineal L. Entonces, existe en L un subespacio mínimo (posiblemente coincidente con L) que contiene $\{x_\alpha\}$. En efecto, existe en L al menos un subespacio que contiene $\{x_\alpha\}$: es todo el L. Además, está claro que *la intersección de cualquier conjunto $\{L_\gamma\}$ de subespacios es de nuevo un subespacio.* Efectivamente, si $L^* = \bigcap_\gamma L_\gamma$ y x, $y \in L^*$, también $\alpha x + \beta y \in L^*$ para todos los α y β. Tomemos ahora todos los subespacios, que contienen el sistema de vectores $\{x_\alpha\}$, y consideremos su intersección. Esta será precisamente el menor subespacio, que contiene el sistema dado de vectores $\{x_\alpha\}$. Este subespacio minimal se llamará *subespacio generado por el conjunto* $\{x_\alpha\}$ o *cápsula lineal* del conjunto $\{x_\alpha\}$. Denotaremos este subespacio mediante $L(\{x_\alpha\})$.

EJERCICIO. Un sistema linealmente independiente $\{x_\alpha\}$ de elementos de un espacio lineal L se llama *base de Hamel*, cuando su cápsula lineal coincide con L. Demuéstrense las siguientes proposiciones:

1) en todo espacio lineal existe una base de Hamel;

2) si $\{x_\alpha\}$ es una base de Hamel en L, todo vector $x \in L$ se representa de manera única mediante una combinación lineal finita de algunos vectores del sistema $\{x_\alpha\}$;

3) dos bases cualesquiera de Hamel de un espacio lineal L tienen la misma potencia; la potencia de una base de Hamel de un espacio lineal L suele llamarse *dimensión algebraica* de este espacio;

4) dos espacios lineales son isomorfos si, y sólo si, tienen la misma dimensión algebraica.

4°. Espacios cocientes. Sea L un espacio lineal y L' algún subespacio suyo. Diremos que dos elementos x e y de L pertenecen a una misma *clase de equivalencia* (según L'), cuando la diferencia de ellos $x - y$ pertenece a L'. El conjunto de todas estas clases se llamará *espacio cociente* de L según L' y se denotará con L/L'. En todo espacio cociente se introducen, de una manera natural, las operaciones de adición y multiplicación por números. A saber, sean ξ y η dos clases, representando elementos de L/L'. Tomemos en cada una de estas clases un elemento, digamos, x e y respectivamente, y llamemos suma de las clases ξ y η a aquella clase ζ, que contiene el elemento $x+y$, y producto de la clase ξ por el número α aquella clase que contiene el elemento αx. Es fácil probar que el resultado no cambiará, si los representantes x e y se sustituyen por cualesquiera otros representantes x' e y' de las mismas clases ξ y η. De esta forma quedan, efectivamente, definidas las operaciones lineales para los elementos del espacio cociente L/L'. Una comprobación directa demuestra que estas operaciones verifican todas las condiciones, contenidas en la definición de un espacio lineal. En otras palabras, *todo espacio cociente L/L'* (con las operaciones de adición y multiplicación por números que acabamos de definir en él) *representa un espacio lineal.*

Si L es un espacio de n dimensiones y el subespacio suyo L' tie' e dimensión k, el espacio cociente es de dimensión $n-k$ (¡demuéstrese esto!).

Sea L un espacio lineal arbitrario y L' algún subespacio suyo. La dimensión del espacio cociente L/L' se llama *codimensión* del subespacio L' del espacio L.

Si el subespacio $L' \subset L$ tiene codimensión finita n, se pueden escoger en L los elementos $x_1, x_2, \ldots, x_n$ de manera que todo elemento $x \in L$ quedará representado (unívocamente) en la forma

$$x = \alpha_1 x_1 + \ldots + \alpha_n x_n + y,$$

donde $\alpha_1, \ldots, \alpha_n$ son números e $y \in L'$. En efecto, sea n la dimensión del espacio cociente L/L'. Tomemos en este espacio cociente una base

$$\xi_1, \xi_2, \ldots, \xi_n$$

y escojamos en cada clase ξ_k arbitrariamente un elemento que designaremos con x_k. Sea ahora x un elemento cualquiera de L y ξ aquella clase en L/L' que contiene a x. Entonces

$$\xi = \alpha_1 \xi_1 + \ldots + \alpha_n \xi_n.$$

Esto significa, por definición, que todo elemento de ξ, en particular, el elemento x, difiere sólo en un elemento de L' de la combinación lineal construida con elementos, tomados por uno en cada clase $\xi_1, \ldots, \xi_n$, es decir,

$$x = \alpha_1 x_1 + \ldots + \alpha_n x_n + y.$$

Dejamos al lector la demostración de que esta representación es única.

5°. **Funcionales lineales.** Una función numérica f, definida sobre un espacio lineal L, se llamará *funcional*[1]. Una funcional f se llama *aditiva*, cuando

$$f(x+y)=f(x)+f(y) \text{ para todos los } x,\ y \in L;$$

se llama *homogénea*, cuando

$$f(\alpha x)=\alpha f(x) \quad (\alpha \text{ es un número}).$$

Una funcional f, definida en un espacio lineal *complejo*, se llama *conjugada homogénea*, cuando $f(\alpha x)=\overline{\alpha} f(x)$, donde $\overline{\alpha}$ es el número complejo conjugado de α.

Una funcional aditiva homogénea se llama *funcional lineal*. Una funcional aditiva conjugada homogénea se llama *conjugada lineal* (o *antilineal*).

Señalemos ejemplos de funcionales lineales.

1. Sea R^n el espacio aritmético de n dimensiones, compuesto por los elementos $x=(x_1, \ldots, x_n)$ y sea $a=(a_1, \ldots, a_n)$ un elemento determinado de R^n. Entonces,

$$f(x)=\sum_{i=1}^{n} x_i a_i$$

es una funcional lineal en R^n. La expresión

$$y(x)=\sum_{i=1}^{n} a_i \overline{x_i}$$

representa una funcional conjugada lineal en C^n.

2. Las integrales

$$I[x]=\int_a^b x(t)\,dt \quad \text{e} \quad \overline{I}[x]=\int_a^b \overline{x(t)}\,dt$$

representan, respectivamente, funcionales lineal y conjugada lineal en el espacio $C_{[a,\,b]}$.

3. Consideremos un ejemplo más general. Sea y_0 una función determinada continua sobre $[a, b]$. Tomemos para toda función $x \in C_{[a,\,b]}$

$$F(x)=\int_a^b x(t)\,y_0(t)\,dt.$$

[1] Aquí la palabra «funcional» se entiende en un sentido algo distinto que en el § 8 del capítulo II, donde hemos llamado funcional a una función numérica definida sobre un espacio métrico, cuyos elementos son funciones. En adelante, tendremos que tratar con espacios lineales, que al mismo tiempo son métricos y cuyos elementos son funciones. Por eso, no resultará esencial cierta discordancia en los términos de este capítulo y el anterior.

La linealidad de esta funcional se deduce de las propiedades principales de la operación de integración. La funcional

$$\overline{F}(x)=\int_a^b \overline{x(t)}\, y_0(t)\,dt$$

será conjugada lineal.

4. Consideremos en este mismo espacio $C_{[a,b]}$ una funcional lineal de otro tipo, tomando

$$\delta_{t_0}(x)=x(t_0),$$

es decir, haciendo igual el valor de la funcional δ_{t_0} para la función x al valor de esta función en un punto fijo t_0.

Frecuentemente, resulta necesario considerar esta funcional, por ejemplo, en la Mecánica Cuántica, donde suele escribirse en la forma

$$\delta_{t_0}(x)=\int_a^b x(t)\,\delta(t-t_0)\,dt,$$

entendiéndose por δ la «función» que es igual a cero en todo punto, excepto el punto $t=0$, y cuya integral es igual a la unidad (δ-función de Dirac). Como veremos en el capítulo siguiente, la δ-función se puede representar como límite, en cierto sentido, de una sucesión de funciones «auténticas» φ_n cada una de las cuales se anula fuera de una ε_n-vecindad ($\varepsilon_n \to 0$ para $n\to\infty$) del punto $t=0$ y verifica la condición $\int_a^b \varphi_n(t)\,dt=1$.

5. Veamos un ejemplo de una funcional lineal en el espacio l_2. Sea k un número entero positivo determinado. Para todo

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

de l_2 tomemos

$$f_k(x)=x_k.$$

Esta funcional es evidentemente lineal. Funcionales de este tipo pueden considerarse también en otros espacios de sucesiones, por ejemplo, en c_0, c, m y R^∞ (ejemplos 5, 6, 7 y 8 del primer punto).

6°. Interpretación geométrica de una funcional lineal. Sea f una funcional lineal, distinta del cero idéntico, en un espacio lineal L. El conjunto L_f de elementos x de L que satisfacen la condición

$$f(x)=0$$

representa un subespacio del espacio L, que se llama *subespacio* *de ceros* de la funcional f. En efecto, si x, $y \in L_f$ tenemos

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y) = 0.$$

El subespacio L_f tiene codimensión 1. En efecto, tomemos un elemento x_0 que no pertenece a L_f, es decir, un elemento tal que $f(x_0) \neq 0$. Tal elemento existe, ya que $f(x) \not\equiv 0$. Podemos admitir, sin perder generalidad, que $f(x_0) = 1$, ya que en el caso contrario podríamos dividir x_0 por $f(x_0)$ $\left(f\left(\frac{x_0}{f(x_0)}\right) = 1\right)$. Para un elemento x cualquiera tenemos $x = f(x) \cdot x_0 + y$, de manera que $f(y) = f(x - f(x)x_0) = 0$, es decir, $y \in L_f$.

Siendo x_0 un elemento fijo, el elemento x se representa de una manera única en la forma

$$x = \alpha x_0 + y, \text{ donde } y \in L_f.$$

En efecto, sea

$$x = \alpha x_0 + y, \quad y \in L_f,$$
$$x = \alpha' x_0 + y', \quad y' \in L_f.$$

Entonces,

$$(\alpha - \alpha')x_0 = y' - y.$$

Si $\alpha = \alpha'$, será evidentemente $y' = y$. En cambio, si $\alpha \neq \alpha'$, tendremos $x_0 = \frac{y' - y}{\alpha - \alpha'} \in L_f$ y esto contradice a la selección de x_0.

De aquí se deduce que dos elementos x_1 y x_2 pertenecen a una misma clase de equivalencia según el subespacio L_f si, y sólo si, $f(x_1) = f(x_2)$. En efecto, de

$$x_1 = f(x_1) \cdot x_0 + y_1,$$
$$x_2 = f(x_2) \cdot x_0 + y_2$$

se sigue que

$$x_1 - x_2 = (f(x_1) - f(x_2)) \cdot x_0 + (y_1 - y_2).$$

De aquí se ve que $x_1 - x_2 \in L_f$ si, y sólo si, el coeficiente de x_0, es decir, $f(x_1) - f(x_2)$, es igual a 0.

Toda clase ξ según el subespacio L_f se determina por cualquiera de sus representantes. A título de este representante podemos tomar el elemento de tipo $\alpha \cdot x_0$. De aquí se ve que el subespacio L/L_f es efectivamente unidimensional y que L_f tiene codimensión 1.

El subespacio L_f determina, salvo un factor constante, la funcional lineal que se anula en él.

En efecto, sean f y g dos funcionales que tienen el mismo espacio de ceros $L_f = L_g$. Tomemos, a partir de f, un elemento

x_0 de manera que $f(x_0)=1$. Afirmamos que $g(x_0)\neq 0$. En efecto,

$$x=f(x)x_0+y,\quad y\in L_f=L_g,$$

y

$$g(x)=f(x)g(x_0)+g(y)=f(x)g(x_0).$$

Si fuese el valor $g(x_0)$ igual a 0, la funcional g resultaría igual idénticamente a cero. De la igualdad $g(x)=g(x_0)f(x)$ se deduce precisamente que las funciones g y f son proporcionales.

Sea L' un subespacio de codimensión 1 en el espacio lineal L; entonces, toda clase de equivalencia del espacio L según el subespacio L' se llama *hiperplano* paralelo al subespacio L' (en particular, el propio subespacio L' es un hiperplano que contiene el 0, es decir, que «pasa por el origen de coordenadas»). En otras palabras, el hiperplano M' paralelo al subespacio L' es el conjunto que se obtiene a partir de L' mediante una traslación paralela a un vector $x_0\in L$:

$$M'=L'+x_0=\{y:y=x+x_0,\ x\in L'\}.$$

Está claro que, si $x_0\in L'$, tenemos $M'=L'$; en cambio, si $x_0\in L'$, tendremos $M'\neq L'$. Si f es una funcional lineal no trivial sobre el espacio L, el conjunto $M_f=\{x:f(x)=1\}$ es un hiperplano paralelo al subespacio L_f de ceros de la funcional f (en efecto, fijando un elemento x_0, para el cual $f(x_0)=1$, podemos representar todo vector $x\in M_f$ en la forma $x=x_0+y$, donde $y\in L_f$). Por otro lado, si M' es un hiperplano paralelo al subespacio L' (de codimensión 1) que no pasa por el origen de coordenadas, existe una funcional lineal f única tal que $M'=\{x:f(x)=1\}$. En efecto, sea $M'=L'+x_0$, $x_0\in L$; en este caso, todo elemento $x\in L$ se puede representar de manera única en la forma $x=\alpha x_0+y$, donde $y\in L'$. Tomando $f(x)=\alpha$, obtenemos la funcional lineal deseada; la unicidad se deduce de que, siendo $g(x)\equiv 1$ para $x\in M'$, tenemos $g(y)\equiv 0$ para $y\in L'$, de manera que

$$g(\alpha x_0+y)=\alpha=f(\alpha x_0+y).$$

Por consiguiente, hemos establecido una correspondencia biunívoca entre todas las funciones lineales no triviales, definidas sobre L, y todos los hiperplanos de L que no pasan por el origen de coordenadas.

EJERCICIO. Sean $f, f_1, \ldots, f_n$ funcionales lineales sobre un espacio lineal L tales que de $f_1(x)=\ldots=f_n(x)=0$ se deduce que $f(x)=0$. Entonces, existen unas constantes $a_1, \ldots, a_n$ tales que $f(x)=\sum_{k=1}^{n} a_k f_k(x)$ para todo $x\in L$.

§ 2. CONJUNTOS CONVEXOS Y FUNCIONALES CONVEXAS. TEOREMA DE HAHN — BANACH

1°. Conjuntos convexos y cuerpos convexos. Varios capítulos importantes de la teoría de espacios lineales tienen como base el concepto de *convexidad* que, apoyándose en ideas geométricas evidentes, admite, al mismo tiempo, un enunciado puramente analítico.

Sea L un espacio lineal real y x, y dos puntos suyos. Se llama *segmento (cerrado)* en L, que une los puntos x e y, al conjunto de todos los elementos de tipo

$$\alpha x+\beta y, \text{ donde } \alpha,\ \beta\geqslant 0,\ \alpha+\beta=1.$$

El segmento sin los puntos extremos x e y se llama *segmento abierto*.

Un conjunto $M\subset L$ se llama *convexo*, cuando junto con dos cualesquiera puntos suyos x e y contiene también al segmento que los une.

Llamaremos *núcleo* de un conjunto arbitrario $E\subset L$ al conjunto de puntos suyos x tales que para todo $y\in L$ existe un número $\varepsilon=\varepsilon(y)>0$ tal que $x+ty\in L$ para $|t|<\varepsilon$.

Un conjunto convexo, cuyo núcleo no es vacío se llama *cuerpo convexo*.

Ejemplos. 1. En el espacio euclídeo de tres dimensiones, el cubo, la bola, el tetraedro y el semiespacio representan cuerpos convexos. Un segmento, un plano o un triángulo del mismo espacio son conjuntos convexos, pero no cuerpos convexos.

2. Consideremos en el espacio de funciones continuas sobre el segmento $[a, b]$ el conjunto de funciones que verifican la condición

$$|f(t)|\leqslant 1.$$

Este conjunto es convexo; en efecto, si

$$|f(t)|\leqslant 1 \text{ y } |g(t)|\leqslant 1,$$

entonces, para $\alpha+\beta=1$, $\alpha,\ \beta\geqslant 0$

$$|\alpha f(t)+\beta g(t)|\leqslant\alpha+\beta=1.$$

3. La bola unitaria de l_2, es decir, el conjunto de puntos $x=(x_1, \ldots, x_n, \ldots)$ tales que $\sum x_n^2\leqslant 1$, es un cuerpo convexo. Su núcleo se compone de los elementos x que verifican la condición $\sum x_n^2<1$.

4. El paralelepípedo fundamental Π en l_2 es un conjun convexo, pero no un cuerpo convexo. En efecto, sea $x\in\Pi$; esto significa que $|x_n|\leqslant\frac{1}{2^{n-1}}$ para todo $n=1, 2, \ldots$ Tomemos

$y_0 = \left(1, \frac{1}{2}, \ldots, \frac{1}{n}, \ldots\right)$. Sea $x + ty_0 \in \Pi$, es decir, $\left|x_n + \frac{t}{n}\right| \leqslant$ $\leqslant \frac{1}{2^{n-1}}$; entonces, $\left|\frac{t}{n}\right| \leqslant \left|x_n + \frac{t}{n}\right| + |x_n| \leqslant \frac{1}{2^{n-1}} + \frac{1}{2^{n-1}} = \frac{1}{2^{n-2}}$, de donde se sigue que $t = 0$ y, por consiguiente, el núcleo del conjunto Π es vacío.

EJERCICIO. Sea Φ el conjunto de puntos $x = (x_1, \ldots, x_n, \ldots)$ de l_2 que verifican la condición $\sum n^2 x_n^2 \leqslant 1$. [Demuéstrese que Φ es un conjunto convexo, pero no un cuerpo convexo.

Si M es un conjunto convexo, su núcleo $I(M)$ es también convexo. En efecto, sean x, $y \in I(M)$ y $z = \alpha x + \beta y$, α, $\beta \geqslant 0$, $\alpha + \beta = 1$. Entonces, para un elemento dado $a \in L$ existen $\varepsilon_1 > 0$ y $\varepsilon_2 > 0$ tales que, siendo $|t_1| < \varepsilon_1$, $|t_2| < \varepsilon_2$, los puntos $x + t_1 a$ e $y + t_2 a$ pertenecen al conjunto M; por consiguiente, a él pertenece también el punto $\alpha(x + ta) + \beta(y + ta) = z + ta$, si $|t| \leqslant$ $\leqslant \varepsilon = \min(\varepsilon_1, \varepsilon_2)$, es decir, $z \in I(M)$.

Demostremos la siguiente propiedad sencilla, pero importante, de los conjuntos convexos.

TEOREMA 1. *La intersección de cualquier número de conjuntos convexos es un conjunto convexo.*

DEMOSTRACION. Sea $M = \bigcap_\alpha M_\alpha$, donde todos los M_α son conjuntos convexos. Sean, además, x e y dos puntos arbitrarios de M. En este caso, el segmento que une los puntos x e y pertenece a cada M_α y, por consiguiente, a M. Por lo tanto, M es efectivamente convexo. Observemos que la intersección de cuerpos convexos (que, de acuerdo con lo establecido, será un conjunto convexo) no es necesariamente un cuerpo convexo (dese un ejemplo).

Para todo conjunto A de un espacio lineal L existe el menor conjunto convexo que contiene A: éste será la intersección de todos los conjuntos convexos que contienen A (existe al menos un conjunto convexo que contiene A, éste es todo L). Este conjunto convexo minimal que contiene A se llama *cápsula convexa* del conjunto A.

Veamos un ejemplo importante de cápsula convexa. Sean $x_1, x_2, \ldots, x_{n+1}$ puntos de un espacio lineal. Diremos que estos puntos están en posición general, cuando no pertenecen a ningún subespacio de $(n-1)$ dimensiones. La cápsula convexa de los puntos $x_1, x_2, \ldots, x_{n+1}$ que se encuentran en posición general se denomina *símplice* de n dimensiones y los propios puntos x_1, $x_2, \ldots, x_{n+1}$ se llaman vértices de este símplice. Un símplice de dimensión cero es un punto. Un símplice unidimensional es un segmento; un bidimensional, un triángulo y un tridimensional, un tetraedro.

Si los puntos $x_1, x_2, \ldots, x_{n+1}$ se encuentran en posición general, cualesquiera $k+1$ de ellos ($k<n$) se encuentran también en posición general y, por consiguiente, generan un símplice k-dimensional que se denomina *faceta* k-dimensional del símplice n-dimensional dado. Por ejemplo, el tetraedro con vértices e_1, e_2, e_3 y e_4 tiene cuatro facetas bidimensionales, determinadas por las ternas de vértices (e_2, e_3, e_4), (e_1, e_3, e_4), (e_1, e_2, e_4) y (e_1, e_2, e_3) respectivamente, seis facetas unidimensionales y cuatro de dimensión cero.

TEOREMA 2. *Un símplice con vértices $x_1, x_2, \ldots x_{n+1}$ es el conjunto de todos los puntos que pueden representarse en la forma*

$$x=\sum_{k=1}^{n+1}\alpha_k x_k,\quad \alpha_k \geqslant 0,\quad \sum_{k=1}^{n+1}\alpha_k=1.$$

DEMOSTRACION. Es fácil probar que la totalidad de puntos de tipo (1) representa un conjunto convexo que contiene los puntos $x_1, x_2, \ldots, x_{n+1}$. Por otro lado, todo conjunto convexo que contenga los puntos $x_1, x_2, \ldots, x_{n+1}$ debe contener también los puntos de tipo (1); consecuentemente, estos puntos forman el menor conjunto convexo que contiene los puntos $x_1, x_2, \ldots, x_{n+1}$.

2°. Funcionales convexas. Al concepto de conjunto convexo está ligado estrechamente el concepto de funcional convexa.

DEFINICION. Una funcional no negativa p, definida sobre un espacio lineal real L, se llama *convexa*, si

1) $p(x+y) \leqslant p(x)+p(y)$ para todos los $x, y \in L$;

2) $p(\alpha x)=\alpha p(x)$ para todos los $\alpha \geqslant 0$.

No admitimos que el valor $p(x)$ es finito para todo $x \in L$, es decir, se admite el caso en que $p(x)=+\infty$ para algunos $x \in L$.

Señalemos ejemplos de funcionales convexas.

1. La longitud de un vector en el espacio euclídeo R^n de n dimensiones. La primera condición significa en este caso que la longitud de la suma de dos vectores no sobrepasa la suma de sus longitudes (desigualdad triangular), mientras que la segunda se deduce directamente de la definición de la longitud de un vector en R^n.

2. Sea M el espacio de funciones x acotadas sobre un conjunto S y sea s_0 un punto fijo de S. Entonces,

$$p_{s_0}(x)=|x(s_0)|$$

es una funcional convexa.

3. Sea m el espacio de sucesiones numéricas acotadas $x=(x_1, x_2, \ldots, x_n, \ldots)$. La funcional

$$p(x)=\sup_n |x_n|$$

es convexa.

3°. Funcional de Minkowski. Consideremos la relación existente entre las funcionales convexas y los conjuntos convexos.

TEOREMA 3. *Si p es una funcional convexa sobre un espacio lineal L y k un número positivo, el conjunto*

$$E=\{x:\ p(x)\leqslant k\}$$

es convexo. Si la funcional p es finita, el conjunto E representa un cuerpo convexo, cuyo núcleo es el conjunto

$$\{x:\ p(x)<k\}$$

(de manera que de antemano contiene el punto 0).

DEMOSTRACION. Si $x,\ y\in E$ y $\alpha+\beta=1$, $\alpha,\ \beta\geqslant 0$, tenemos

$$p(\alpha x+\beta y)\leqslant \alpha p(x)+\beta p(y)\leqslant k,$$

es decir, E es convexo. Supongamos ahora que la funcional p es finita, $p(x)<k$, que $t>0$ e $y\in L$; entonces,

$$p(x\pm ty)\leqslant p(x)+tp(\pm y).$$

Si $p(-y)=p(y)=0$, tenemos $x\pm ty\in E$ para todo t; en cambio, si al menos uno de los números $p(y)$, $p(-y)$ es distinto de cero, tendremos $x\pm ty\in E$ cuando

$$t<\frac{k-p(x)}{\max(p(y),\ p(-y))}.$$

Tomemos para k un valor determinado, digamos, $k=1$. En este caso, toda funcional finita convexa p determina unívocamente en L un cuerpo convexo E tal que $0\in I(E)$. Viceversa, sea E un cuerpo convexo, cuyo núcleo contiene el punto 0. Entonces,

$$p_E(x)=\inf\left\{r:\frac{x}{r}\in E,\ r>0\right\} \tag{2}$$

es una funcional convexa finita. Se llama *funcional de Minkowski* del cuerpo convexo E.

Probemos la convexidad de la funcional de Minkowski (2). Para todo $x\in L$, el elemento $\frac{x}{r}$ pertenece a E, cuando r es suficientemente grande; por eso, la magnitud $p_E(x)$, definida por la igualdad (2), es no negativa y finita. Si $t>0$ e $y=tx$, tenemos

$$p_E(y)=\inf\left\{r>0:\ \frac{y}{r}\in E\right\}=\inf\left\{r>0:\ \frac{tx}{r}\in E\right\}=$$
$$=\inf\left\{tr'>0:\ \frac{x}{r'}\in E\right\}=t\inf\left\{r'>0:\ \frac{x}{r'}\in E\right\}=$$
$$=tp_E(x). \tag{3}$$

Sean ahora $x_1,\ x_2\in L$ y sea $\varepsilon>0$ arbitrario. Escojamos los números r_i $(i=1,\ 2)$ de manera que $p_E(x_i)<r_i<p_E(x_i)+\varepsilon$; entonces, $\frac{x_i}{r_i}\in E$. Pongamos $r=r_1+r_2$; entonces, $\frac{x_1+x_2}{r}=\frac{r_1x_1}{rr_1}+$ $+\frac{r_2x_2}{rr_2}$ pertenece al segmento con extremos $\frac{x_1}{r_1}$ y $\frac{x_2}{r_2}$. Como E es convexo, este segmento y, por consiguiente, el punto $\frac{x_1+x_2}{r}$ pertenecen a E y por eso

$$p_E(x_1+x_2)\leqslant r=r_1+r_2<p_E(x_1)+p_E(x_2)+2\varepsilon.$$

Puesto que ε es aquí arbitrario, tenemos

$$p_E(x_1+x_2)\leqslant p_E(x_1)+p_E(x_2). \tag{4}$$

Las relaciones (3) y (4) significan precisamente que la funcional $p_E(x)$ es convexa.

4°. Teorema de Hahn—Banach. Sea L un espacio lineal real y sea L_0 un subespacio suyo. Supongamos, además, que sobre el subespacio L_0 se ha definido una funcional lineal f_0. Una funcional lineal f, definida sobre todo el espacio L, se llama *prolongación* de la funcional f_0, cuando

$$f(x)=f_0(x) \text{ para todo } x\in L_0.$$

El problema sobre la extensión de una funcional lineal, dada inicialmente sobre un subespacio, a un espacio mayor surge con frecuencia en el Análisis. El papel principal en estas cuestiones lo desempeña el siguiente teorema.

TEOREMA 4 (HAHN—BANACH). *Sea p una funcional convexa finita, definida sobre un espacio lineal real L, y sea L_0 un subespacio lineal de L. Si f_0 es una funcional lineal sobre L_0, que verifica sobre L_0 la condición*

$$f_0(x)\leqslant p(x), \tag{5}$$

la funcional f_0 puede ser prolongada a una funcional f sobre L, que verifica en todo L la condición (5).

DEMOSTRACION. Probemos que, siendo $L_0\neq L$, la funcional f_0 se puede prolongar de L_0 a un subespacio mayor L', conservando la condición (5). En efecto, sea z un elemento arbitrario de L

que no pertenece a L_0 y sea L' el subespacio generado por L_0 y el elemento z. Todo elemento de L' tiene la forma

$$tz+x, \text{ donde } x\in L_0.$$

Si f' es la prolongación deseada de la funcional f_0 sobre L', tenemos

$$f'(tz+x)=tf'(z)+f_0(x),$$

o, tomando $f'(z)=c$,

$$f'(tz+x)=tc+f_0(x).$$

Escojamos ahora c de manera que en todo L' se cumpla la condición de subordinación (5), es decir, que para todo $x\in L_0$ y cualesquiera t reales se verifique la desigualdad

$$f_0(x)+tc\leqslant p(x+tz).$$

Para $t>0$ esta condición es equivalente a

$$f_0\left(\frac{x}{t}\right)+c\leqslant p\left(\frac{x}{t}+z\right), \text{ o } c\leqslant p\left(\frac{x}{t}+z\right)-f_0\left(\frac{x}{t}\right), \tag{6}$$

y para $t<0$ es equivalente a

$$f_0\left(\frac{x}{t}\right)+c\geqslant -p\left(-\frac{x}{t}-z\right) \text{ o } c\geqslant -p\left(-\frac{x}{t}-z\right)-f_0\left(\frac{x}{t}\right). \tag{7}$$

Demostremos que siempre existe un c que cumple las condiciones (6) y (7). Sean y' e y'' elementos arbitrarios de L_0. Entonces,

$$-f_0(y'')+p(y''+z)\geqslant -f_0(y')-p(-y'-z). \tag{8}$$

En efecto, este resultado se obtiene de la desigualdad

$$f_0(y'')-f_0(y')\leqslant p(y''-y')=p((y''+z)-(y'+z))\leqslant$$
$$\leqslant p(y''+z)+p(-y'-z).$$

Tomemos

$$c''=\inf_{y''}(-f_0(y'')+p(y''+z)),\quad c'=\sup_{y'}(-f_0(y')-p(-y'-z)).$$

Debido a que y' e y'' son arbitrarios, de (8) se deduce que $c''\geqslant c'$. Escogiendo c de manera que

$$c''\geqslant c\geqslant c',$$

veremos que la funcional f', definida sobre L' por

$$f'(tz+x)=tc+f_0(x),$$

verifica la condición de subordinación (5). Por consiguiente, hemos demostrado que, si la funcional f_0 está definida sobre un subespacio $L_0\subset L$ y verifica en L_0 la condición (5), se puede extender f_0, conservando esta condición, a un subespacio mayor L'. En el caso en que se pueda escoger en L un sistema numerable de elementos $x_1, x_2, \ldots, x_n, \ldots$ que genera todo L, la funcional sobre L se construye por inducción, considerando la cadena creciente de subespacios

$$L^{(1)}=\{L_0, x_1\}, \quad L^{(2)}=\{L^{(1)}, x_2\}, \ldots$$

(aquí $\{L^{(k)}, x_{k+1}\}$ representa el subespacio lineal minimal de L que contiene $L^{(k)}$ y x_{k+1}). Entonces, todo elemento $x \in L$ entrará en algún $L^{(n)}$ y, por consiguiente, la funcional resultará prolongada a todo L.

En el caso general (es decir, cuando no existe un conjunto numerable generador de L), la demostración concluye aplicándose el lema de Zorn. El conjunto $\mathcal{F}$ de todas las prolongaciones posibles de la funcional f_0, que verifican la condición de subordinación (5), es un conjunto parcialmente ordenado y todo subconjunto suyo $\mathcal{F}_0$ linealmente ordenado tiene extremo superior; este extremo superior es la funcional, definida sobre la unión de los campos de definición de las funcionales $f' \in \mathcal{F}_0$ y coincidente con cada una de estas f' en su campo de definición. De acuerdo con el lema de Zorn, existe en $\mathcal{F}$ un elemento maximal f. Este elemento maximal f representa precisamente la funcional deseada. En efecto, es una prolongación de la funcional inicial f_0, verifica la condición (5) en su campo de definición y está definida sobre todo el espacio L, ya que de lo contrario sería posible prolongarla, empleando el método descrito anteriormente, del subespacio propio, en que esté definida, a un subespacio mayor, y, por consiguiente, f no sería un elemento maximal. El teorema queda demostrado.

Señalemos también la variante compleja del teorema de Hahn—Banach.

Una funcional no negativa p, definida sobre un espacio lineal *complejo* L, se llama *convexa*, cuando para todo $x, y \in L$ y cualesquiera números complejos λ

$$p(x+y) \leqslant p(x)+p(y),$$
$$p(\lambda x)=|\lambda| p(x).$$

TEOREMA 4a. *Sea p una funcional convexa finita sobre un espacio lineal complejo L y sea f_0 una funcional lineal, definida sobre un subespacio lineal $L_0 \subset L$, donde verifica la condición*

$$|f_0(x)| \leqslant p(x), \quad x \in L_0.$$

Entonces, existe una funcional lineal f, definida en todo L, que verifica las condiciones

$$|f(x)| \leqslant p(x), \quad x \in L,$$
$$f(x)=f_0(x), \quad x \in L_0.$$

DEMOSTRACION. Denotemos mediante L_R y L_{0R} los espacios L y L_0, considerados como espacios lineales reales. Está claro que p es una funcional convexa finita sobre L_R y que $f_{0R}(x) = \mathrm{Re}\, f_0(x)$ es una funcional lineal real sobre L_{0R} que verifica la condición

$$|f_{0R}(x)| \leqslant p(x)$$

y, es más, la condición

$$f_{0R}(x) \leqslant p(x).$$

En virtud del teorema 4, existe una funcional lineal real f_R, definida en todo L_R, que satisface las condiciones

$$f_R(x) \leqslant p(x), \quad x \in L_R \ (=L),$$
$$f_R(x) = f_{0R}(x), \quad x \in L_{0R} \ (=L_0).$$

Es evidente que $-f_R(x) = f_R(-x) \leqslant p(-x) = p(x)$ y, por lo tanto,

$$|f_R(x)| \leqslant p(x), \ x \in L_R \ (=L). \tag{9}$$

Definamos en L la funcional f, tomando

$$f(x) = f_R(x) - if_R(ix)$$

(aquí nos valemos de que L es un espacio lineal complejo, de manera que en él está definida la multiplicación por números complejos). Una comprobación directa muestra que f es una funcional lineal compleja sobre L y que, además,

$$f(x) = f_0(x) \quad \text{para } x \in L_0,$$
$$\mathrm{Re}\, f(x) = f_R(x) \quad \text{para } x \in L.$$

Resta demostrar que $|f(x)| \leqslant p(x)$ para todo $x \in L$. Supongamos lo contrario; entonces, para algún $x_0 \in L$ tendremos $|f(x_0)| > p(x_0)$. Representemos el número complejo $f(x_0)$ en la forma $f(x_0) = \rho e^{i\varphi}$, donde $\rho > 0$, y pongamos $y_0 = e^{-i\varphi}x_0$. Entonces, $f_R(y_0) = \mathrm{Re}\, f(y_0) = \mathrm{Re}\,[e^{-i\varphi}(x_0)] = \rho > p(x_0) = p(y_0)$ y esto contradice a la condición (9). El teorema queda demostrado.

EJERCICIO. Demuéstrese que en el teorema de Hahn—Banach se puede omitir la condición de que la funcional p sea finita.

5°. Separabilidad de conjuntos convexos en espacios lineales. Sea L un espacio real y M y N dos subconjuntos en él. Se dice que la funcional lineal f definida sobre L *separa* estos conjuntos, cuando existe un número C tal que

$$f(x) \geqslant C \text{ para } x(\in) M \text{ y } f(x) \leqslant C \text{ para } x \in N.$$

Las dos siguientes proposiciones se desprenden directamente de la definición dada.

1) Una funcional lineal f separa los conjuntos M y N si, y sólo si, separa los conjuntos $M-N$ y $\{0\}$ (es decir, el conjunto, compuesto por todos los elementos de tipo $x-y$, donde $x \in M$ e $y \in N$, y el punto 0).

2) Una funcional lineal f separa los conjuntos M y N si, y sólo si, separa los conjuntos $M-x$ y $N-x$ para todo $x \in L$.

Del teorema de Hahn—Banach se obtiene sin dificultad el siguiente teorema sobre la separabilidad de conjuntos convexos en un espacio lineal, que encuentra múltiples aplicaciones.

TEOREMA 5. *Sean M y N dos conjuntos convexos disjuntos en un espacio lineal real L, con la particularidad que al menos uno de ellos, digamos M, tiene un núcleo no vacío (esto es, representa un cuerpo convexo). Entonces, existe sobre L una funcional f lineal no nula que separa M y N.*

DEMOSTRACION. Sin perder generalidad, podemos admitir que el punto 0 pertenece al núcleo del conjunto M. (De lo contrario, consideraríamos los conjuntos $M-x_0$ y $N-x_0$, donde x_0 es algún punto del núcleo de M). Sea y_0 un punto del conjunto N; entonces, el punto $-y_0$ pertenece al núcleo del conjunto $M-N$, mientras que el punto 0 pertenece al núcleo del conjunto $K=M-N+y_0$. Como los conjuntos M y N no se intersecan, tenemos $0 \overline{\in} M-N$ e $y_0 \overline{\in} K$. Sea p la funcional de Minkowski del conjunto K. Entonces, $p(y_0) \geqslant 1$ (puesto que $y_0 \overline{\in} K$). Consideremos la funcional lineal

$$f_0(\alpha y_0) = \alpha p(y_0).$$

Está definida sobre un subespacio unidimensional, compuesto por elementos de tipo αy_0, y cumple la condición

$$f_0(\alpha y_0) \leqslant p(\alpha y_0),$$

ya que $p(\alpha y_0) = \alpha p(y_0)$ para $\alpha \geqslant 0$ y $f_0(\alpha y_0) = \alpha f_0(y_0) < 0 < p(\alpha y_0)$ para $\alpha < 0$. De acuerdo con el teorema de Hahn—Banach, la funcional f_0 puede extenderse hasta una funcional lineal f, definida en todo L, que verifica en L la condición $f(y) \leqslant p(y)$. De aquí se deduce que $f(y) \leqslant 1$ para $y \in K$ y que, al mismo tiempo, $f(y_0) \geqslant 1$. Es decir, f separa los conjuntos K e $\{y_0\}$ y, por consiguiente, f separa $M-N$ y $\{0\}$; pero, entonces, f separa los conjuntos M y N. El teorema queda demostrado.

§ 3. ESPACIOS NORMADOS

En el capítulo II hemos considerado los espacios topológicos, en particular, métricos, es decir, conjuntos en los que se ha introducido, de una u otra manera, el concepto de proximidad de elementos, mientras que en los parágrafos anteriores de este capítulo hemos tratado espacios lineales. Hasta ahora hemos considerado estos dos entes, los espacios topológicos y los espacios lineales, independientemente uno del otro. No obstante, en el Análisis tropezamos, casi siempre, con espacios provistos tanto de una topología como de operaciones de adición de elementos y multiplicación de éstos por números, es decir, tropezamos con los así llamados espacios topológicos lineales. Entre los espacios topológicos lineales constituyen una clase importante los espacios normados. La teoría de estos espacios fue desarrollada en los trabajos de S. Banach y de otros autores.

1°. Definición y ejemplos de espacios normados.

DEFINICION 1. Sea L un espacio lineal. Una funcional convexa finita p, definida sobre L, se llama *norma*, cuando verifica las siguientes condiciones adicionales (además de la de convexidad):

1) $p(x)=0$ sólo si $x=0$,
2) $p(\alpha x)=|\alpha|\,p(x)$ para todo α.

Por consiguiente, podemos decir, recordando la definición de convexidad, que se llama *norma* en L a una funcional finita que cumple las tres condiciones siguientes:

1) $p(x)\geqslant 0$, con la particularidad de que $p(x)=0$ sólo si $x=0$,
2) $p(x+y)\leqslant p(x)+p(y)$, $x, y\in L$,
3) $p(\alpha x)=|\alpha|\,p(x)$ cualquiera que sea el número α.

DEFINICION 2. Un espacio lineal L en el que se ha introducido una norma se llama *espacio normado*. La norma del elemento $x\in L$ se denotará mediante el símbolo $\|x\|$.

Todo espacio normado se convierte en un espacio métrico, si para dos cualesquiera elementos $x, y\in L$ se toma

$$\rho(x, y)=\|x-y\|.$$

La validez de los axiomas de espacio métrico se desprende directamente de las propiedades 1), 2) y 3) de la norma. En los espacios normados subsisten, por consiguiente, todos los conceptos y resultados expuestos en el capítulo II para los espacios métricos.

Un espacio normado *completo* se llama *espacio de Banach* o, brevemente, *B-espacio*.

Ejemplos de espacios normados. Muchos de los espacios, considerados en el capítulo II como ejemplos de espacios métricos (y en el § 1 de este capítulo, como ejemplos de espacios lineales), pueden proveerse de hecho de una estructura natural de espacio normado.

1. La recta numérica R^1 se convierte en un espacio normado, si se toma $\|x\| = |x|$ para todo número $x \in R^1$.

2. Si en el espacio real R^n de n dimensiones con elementos

$$x = (x_1, x_2, \ldots, x_n)$$

tomamos

$$\|x\| = \sqrt{\sum_{k=1}^{n} x_k^2}, \tag{1}$$

se comprobarán todos los axiomas de la norma. La fórmula

$$\rho(x, y) = \|x - y\| = \sum_{k=1}^{n} (x_k - y_k)^2$$

determina en R^n la misma métrica que hemos considerado ya en este espacio.

En este mismo espacio lineal se puede introducir la norma

$$\|x\|_1 = \sum_{k=1}^{n} |x_k| \tag{2}$$

y la norma

$$\|x\|_0 = \max_{1 \leqslant k \leqslant n} |x_k|. \tag{3}$$

Estas normas determinan en R^n las métricas que hemos considerado ya en los ejemplos 3 y 4 del § 1 del capítulo II. No ofrece dificultad comprobar que en cada uno de estos casos se cumplen efectivamente los axiomas de la norma.

En el espacio complejo C^n de n dimensiones se puede introducir la norma

$$\|x\| \sqrt{\sum_{k=1}^{n} |x_k|^2}$$

o cualquiera de las normas (2) ó (3).

3. Definamos la norma en el espacio $C_{[a, b]}$ de funciones continuas sobre el segmento $[a, b]$ mediante la fórmula

$$\|f\| = \max_{a \leqslant t \leqslant b} |f(t)|. \tag{4}$$

La distancia, correspondiente a esta norma, fue considerada ya en el ejemplo 6 del § 1 del capítulo II.

4. Sea m el espacio de sucesiones numéricas acotadas

$$x=(x_1, x_2, \ldots, x_n \ldots).$$

Pongamos

$$\|x\|=\sup_n |x_n|. \tag{5}$$

Es obvio que las condiciones 1), 2), y 3) de la definición de la norma se cumplen. La métrica que induce en m esta norma coincide con aquella que hemos considerado anteriormente (cap. II, § 1, ejemplo 9).

2°. **Subespacios de un espacio normado.** Hemos definido un subespacio de un espacio lineal L (desprovisto de topología cualquiera) como un conjunto no vacío L_0 tal que, si x, $y \in L_0$, se tiene $\alpha x+\beta y \in L_0$. En un espacio normado, son de interés principal los subespacios lineales *cerrados*, es decir, aquellos subespacios que contienen todos sus puntos de acumulación. En un espacio normado de dimensión finita todo subespacio es automáticamente cerrado (¡demuéstrese esto!). En el caso de un espacio de dimensión infinita esto no es así. Por ejemplo, en el espacio $C_{[a, b]}$ de las funciones continuas con la norma (4), los polinomios forman un subespacio, pero no cerrado[1].

Otro ejemplo: en el espacio m de las sucesiones acotadas, las sucesiones, que contienen solamente un número finito de elementos diferentes de cero, constituyen un subespacio. Sin embargo, no es cerrado: su adherencia contiene, por ejemplo, la sucesión $\left(1, \frac{1}{2}, \ldots, \frac{1}{n}, \ldots\right)$.

Puesto que en lo sucesivo consideraremos, como regla general, solamente subespacios cerrados, resulta lógico modificar algo la terminología establecida en el § 1. En este orden, entenderemos por subespacio de un espacio normado un subespacio *cerrado*; en particular, el subespacio generado por un sistema dado de elementos $\{x_\alpha\}$ será el menor subespacio *cerrado* que contiene $\{x_\alpha\}$. Llamaremos este subespacio *adherencia lineal* del sistema $\{x_\alpha\}$. El conjunto (no cerrado) de elementos que contiene junto con x e y cualquier combinación lineal $\alpha x+\beta y$ de ellos, se llamará *variedad lineal*.

Un sistema arbitrario de elementos, pertenecientes a un espacio normado E, se llamará *completo*, cuando el subespacio (¡cerrado!) generado por él es todo el E. Por ejemplo, en virtud del teorema de Weierstrass, el conjunto de todas las funciones 1, t,

---

[1] De acuerdo con el teorema de Weierstrass, según el cual *toda* función continua sobre un segmento es límite de una sucesión uniformemente convergente de polinomios, la adherencia del subespacio de polinomios en $C_{[a, b]}$ es todo $C_{[a, b]}$.

$t^2, \ldots, t^n, \ldots$ es completo en el espacio de funciones continuas $C_{[a,\,b]}$.

EJERCICIOS. 1. Sea R un espacio de Banach y sea $S_1 \supset S_2 \supset \ldots S_n \supset \ldots$ una sucesión de bolas cerradas encajadas en él. Demuéstrese que tiene una intersección no vacía (aquí no se supone que los radios de estas bolas tienden a 0; compárese con el ejercicio de la pág. 75). Dése un ejemplo de una sucesión de conjuntos encajados no vacíos, acotados, cerrados y convexos de un B-espacio, con intersección vacía.

2. Sea R un B-espacio de dimensión infinita; entonces, su dimensión algebraica (véase el ejercicio de la pág. 134) es innumerable.

3. Sea R un espacio de Banach y M un subespacio suyo cerrado. Consideremos el espacio cociente $P = R/M$ y definimos en él la norma, tomando para toda clase de equivalencia ξ

$$\|\xi\| = \inf_{x \in \xi} \|x\|.$$

Demuéstrese que la funcional así definida representa efectivamente una norma en P y que, además, el espacio P con esta norma es un espacio de Banach.

4. Sea R un espacio lineal normado; demuéstrese la validez de las siguientes proposiciones:

1) todo subespacio lineal de dimensión finita de R es cerrado;

2) si M es cerrado y N es un subespacio de R de dimensión finita, la suma

$$M + N = \{x : x = y + z,\ y \in M,\ z \in N\}$$

es un subespacio lineal cerrado; dése un ejemplo de dos subespacios lineales cerrados del espacio l_2, cuya suma no es cerrada;

3) sea Q un conjunto abierto convexo de R y sea $x_0 \overline{\in} Q$; entonces, existe un hiperplano cerrado que pasa por el punto x_0 y no se intersecta con Q.

5. Dos normas $\|\cdot\|_1$, $\|\cdot\|_2$ de un espacio lineal R se llaman *equivalentes*, cuando existen unas constantes $a, b > 0$ tales que $a\|x\|_1 \leqslant \|x\|_2 \leqslant b\|x\|_1$ para todos $x \in R$. Demuéstrese que, siendo el espacio R de dimensión finita, cualesquiera dos normas en él son equivalentes.

§ 4. ESPACIOS EUCLIDEOS

1°. Definición de espacios euclídeos. Un método bien conocido de introducir una norma en un espacio lineal es el de definir en éste el producto escalar. Recordemos que se llama *producto escalar* en un espacio lineal real R a una función real (x, y), definida para cada par de elementos $x, y \in R$, que verifica las siguientes condiciones:

1) $(x, y) = (y, x)$,
2) $(x, y_1 + y_2) = (x, y_1) + (x, y_2)$,
3) $(\lambda x, y) = \lambda (x, y)$,
4) $(x, x) \geqslant 0$ y $(x, x) = 0$ sólo si $x = 0$.

Un espacio lineal con un producto escalar definido en él se llama *espacio euclídeo*. En un espacio euclídeo R se introduce la norma mediante la fórmula

$$\|x\|=\sqrt{(x, x)}. \tag{1}$$

De las propiedades 1), 2), 3) y 4) del producto escalar se deduce que se cumplen todos los axiomas de la norma.

En efecto, el cumplimiento de los axiomas 1) y 3) de la norma (punto 1 del § 3) es evidente y la validez del axioma 2) (desigualdad triangular) se deduce de la *desigualdad de Cauchy—Buniakovski*

$$|(x, y)| \leqslant \|x\| \cdot \|y\|$$

que demostraremos ahora.

Consideremos el siguiente trinomio de segundo grado respecto a la variable real λ, no negativo para cualesquiera valores de λ:

$$\varphi(\lambda)=(\lambda x+y, \lambda x+y)=\lambda^2(x, x)+2\lambda(x, y)+(y, y)=$$
$$=\|x\|^2\lambda^2+2(x, y)\lambda+\|y\|^2 \geqslant 0.$$

La desigualdad de Cauchy—Buniakovski afirma simplemente que el discriminante de este trinomio de segundo grado es no positivo.

Observemos que en un espacio euclídeo todas las operaciones (adición, multiplicación por número, producto escalar) son continuas, es decir, si $x_n \to x$, $y_n \to y$ (en el sentido de convergencia respecto a la norma) y $\lambda_n \to \lambda$ (como una sucesión numérica), entonces,

$$x_n+y_n \to x+y,$$
$$\lambda_n x_n \to \lambda x,$$
$$(x_n, y_n) \to (x, y).$$

La demostración de este resultado se basa en la desigualdad de Cauchy—Buniakovski y queda, a título de ejercicio, a cargo del lector.

La existencia del producto escalar en R permite definir en este espacio no sólo la norma de un vector (es decir, su longitud), sino también el ángulo entre vectores; a saber, el ángulo φ entre dos vectores x e y se define mediante la fórmula

$$\cos\varphi=\frac{(x, y)}{\|x\|\cdot\|y\|}. \tag{2}$$

De la desigualdad de Cauchy—Buniakovski

$$|(x, y)| \leqslant \|x\|\cdot\|y\|$$

se deduce que el valor absoluto de la expresión que figura en el miembro derecho de (2) no sobrepasa 1, es decir, determina, efectivamente, un ángulo φ, $0 \leqslant \varphi \leqslant \pi$, cualesquiera que sean x e y.

Si $(x, y)=0$, tenemos de (2) $\varphi=\frac{\pi}{2}$; en este caso los vectores x e y se llaman *ortogonales*.

Un sistema $\{x_\alpha\}$ de vectores de R diferentes de cero se llama *ortogonal*, cuando

$$(x_\alpha, x_\beta)=0 \text{ para } \alpha \neq \beta.$$

Si los vectores x_α forman un sistema ortogonal, son linealmente independientes. En efecto, sea

$$a_1 x_{\alpha 1}+a_2 x_{\alpha_2}+\ldots+a_n x_{\alpha_n}=0;$$

si $\{x_\alpha\}$ es un sistema ortogonal,

$$(x_{\alpha_i}, a_1 x_{\alpha_1}+\ldots+a_n x_{\alpha_n})=a_i(x_{\alpha_i}, x_{\alpha_i})=0,$$

es decir, puesto que $(x_{\alpha_i}, x_{\alpha_i}) \neq 0$, tenemos $a_i=0$ para todos los $i=1, 2, \ldots, n$.

Un sistema ortogonal $\{x_\alpha\}$ *completo* (es decir, tal que el menor subespacio cerrado que lo contiene es todo el R) se llama *base ortogonal*. Si, además, la norma de cada elemento es igual a 1, el sistema $\{x_\alpha\}$ se llama *base ortonormal*. En general, un sistema $\{x_\alpha\}$ (completo o no) tal que

$$(x_\alpha, x_\beta)=\begin{cases} 0 \text{ para } \alpha \neq \beta, \\ 1 \text{ para } \alpha=\beta, \end{cases}$$

se llama sistema ortonormal. Está claro que siendo $\{x_\alpha\}$ un sistema ortogonal, el sistema $\left\{\frac{x_\alpha}{\|x_\alpha\|}\right\}$ resulta ortonormal.

2°. **Ejemplos.** Consideremos algunos ejemplos de espacios euclídeos y bases ortogonales en ellos.

1. El espacio de coordenadas R^n de n dimensiones, cuyos elementos son los sistemas de números reales

$$x=(x_1, x_2, \ldots, x_n)$$

con las operaciones habituales de adición y multiplicación en él y con el producto escalar

$$(x, y)=\sum_{i=1}^{n} x_i y_i \tag{3}$$

representa un ejemplo bien conocido de espacio euclídeo. Una base ortonormal en él (una del número infinito de bases ortonormales posibles que tiene) viene dada por los vectores

$$\begin{aligned} e_1 &= (1,\ 0,\ 0,\ \ldots,\ 0), \\ e_2 &= (0,\ 1,\ 0,\ \ldots,\ 0), \\ &\cdots\cdots\cdots\cdots \\ e_n &= (0,\ 0,\ 0,\ \ldots,\ 1). \end{aligned}$$

2. El espacio l_2 con los elementos

$$x = (x_1,\ x_2,\ \ldots,\ x_n,\ \ldots),\ \text{donde}\ \sum_{i=1}^{\infty} x_i^2 < \infty,$$

y con el producto escalar

$$(x,\ y) = \sum_{i=1}^{\infty} x_i y_i \tag{4}$$

es un espacio euclídeo. En efecto, la convergencia de la serie que figura en el miembro derecho de (4) fue demostrada ya en el cap. II, § 1. Las propiedades 1), 2), 3) y 4) del producto escalar se comprueban directamente. La base ortonormal más sencilla de l_2 es la formada por los vectores

$$\left.\begin{aligned} e_1 &= (1,\ 0,\ 0,\ \ldots), \\ e_2 &= (0,\ 1,\ 0,\ \ldots), \\ e_3 &= (0,\ 0,\ 1,\ \ldots), \\ &\cdots\cdots\cdots \end{aligned}\right\} \tag{5}$$

Son obvias la ortogonalidad y normalidad de este sistema; al mismo tiempo, el sistema (5) es completo: sea $x = (x_1,\ x_2,\ \ldots,\ x_n,\ \ldots)$ un vector cualquiera de l_2 y sea $x^{(n)} = (x_1,\ x_2,\ \ldots,\ x_n,\ 0,\ 0,\ \ldots)$. Entonces, $x^{(n)}$ es una combinación lineal de los vectores $e_1,\ \ldots,\ e_n$ y $\|x^{(n)} - x\| \to 0$ para $n \to \infty$.

3. El espacio $C_{[a,\ b]}^{(2)}$, compuesto por funciones continuas reales sobre $[a,\ b]$, con el producto escalar

$$(f,\ g) = \int_a^b f(t)\, g(t)\, dt \tag{6}$$

representa también un espacio euclídeo. Entre diferentes bases ortogonales que se pueden señalar en él, es de importancia principal el sistema trigonométrico, compuesto por las funciones

$$\frac{1}{2},\ \cos n\frac{2\pi t}{b-a},\ \operatorname{sen} n\frac{2\pi t}{b-a} \quad (n = 1,\ 2,\ \ldots). \tag{7}$$

La ortogonalidad de este sistema se comprueba directamente.

Si consideramos las funciones continuas sobre un segmento de longitud 2π, digamos sobre $[-\pi, \pi]$, el sistema trigonométrico correspondiente será $1/2$, $\cos nt$, $\operatorname{sen} nt$ $(n=1, 2, \ldots)$.

El sistema (7) es completo. En efecto, de acuerdo con el teorema de Weierstrass, toda función φ continua sobre el segmento $[a, b]$, que toma valores iguales en los puntos a y b, puede ser representada como límite de una sucesión uniformemente convergente de polinomios trigonométricos, es decir, de combinaciones lineales de elementos del sistema (7). Con más razón esta sucesión convergerá a φ según la norma del espacio $C^{(2)}_{[a, b]}$. Si f es una función arbitraria de $C^{(2)}_{[a, b]}$, se puede representarla como límite (según la norma del espacio $C^{(2)}_{[a, b]}$)

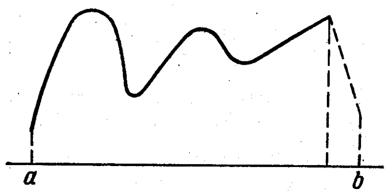


FIG. 18

de una sucesión de funciones φ_n cada una de las cuales coincide con f en el segmento $\left[a, b-\frac{1}{n}\right]$, es lineal en el segmento $\left[b-\frac{1}{n}, b\right]$ y toma en b el mismo valor que tiene en el punto a (fig. 18). Por consiguiente, todo elemento de $C^{(2)}_{[a, b]}$ se puede aproximar tanto como se quiera (en la métrica de este espacio) mediante combinaciones lineales de elementos del sistema (7) y esto demuestra precisamente su complitud.

3°. **Existencia de bases ortogonales, ortogonalización.** En la parte que queda de este parágrafo, nos limitaremos a espacios euclídeos separables (esto es, a los provistos de un conjunto numerable siempre denso). Cada uno de los espacios, señalados en el punto anterior, es separable (¡demuéstrese esto!). Un ejemplo de un espacio euclídeo, en el que no existe un conjunto numerable siempre denso, se puede construir de la siguiente manera. Consideremos sobre el segmento $[0, 1]$ todas las funciones posibles x, para cada una de las cuales el conjunto de puntos $t_1, t_2, \ldots,$ donde ella es diferente de cero, es a lo sumo numerable y la suma $\sum x^2(t)$, tomada respecto a estos puntos, es finita. Definamos en este espacio las operaciones de adición y multiplicación por números como las habituales adición y multiplicación de funciones y el producto escalar de x por y, de la siguiente manera:

$$(x, y) = \sum x(t_i)\, y(t_i),$$

donde la suma se toma según el conjunto de puntos t_i tales que $x(t_i)\, y(t_i) \neq 0$. Proponemos al lector demostrar que en este espacio no existe ningún subconjunto numerable siempre denso.

Sea, pues, R un espacio euclídeo separable. Demostremos que *todo sistema ortogonal de un tal espacio es a lo sumo numerable.*

En efecto, sin perder generalidad podemos admitir que el sistema $\{\varphi_\alpha\}$ no sólo es ortogonal, sino también normal $\left(\text{de lo contrario, podríamos sustituirlo por el sistema } \left\{\frac{\varphi_\alpha}{\|\varphi_\alpha\|}\right\}\right)$. En este caso,

$$\|\varphi_\alpha - \varphi_\beta\| = \sqrt{2} \quad \text{para } \alpha \neq \beta.$$

Consideremos el conjunto de bolas $B\left(\varphi_\alpha, \frac{1}{2}\right)$. Estas bolas no se intersecan. Si el conjunto numerable $\{\psi_n\}$ es siempre denso en R, en cada una de estas bolas hay al menos un elemento de $\{\psi_n\}$. Entonces, el número de estas bolas (y, por consiguiente, de los elementos φ_α también) es a lo sumo numerable.

Hemos señalado una base ortogonal en cada uno de los ejemplos expuestos de espacios euclídeos. Demostremos ahora el siguiente teorema general, análogo al teorema de existencia de una base ortogonal en el espacio euclídeo de n dimensiones.

TEOREMA 1 (sobre la ortogonalización). *Sea*

$$f_1, f_2, \ldots, f_n, \ldots \tag{8}$$

un sistema linealmente independiente de elementos de un espacio euclídeo R. Existe en R un sistema de elementos

$$\varphi_1, \varphi_2, \ldots, \varphi_n, \ldots \tag{9}$$

que satisface las siguientes condiciones:

1) *el sistema (9) es ortogonal y normal;*
2) *todo elemento φ_n es una combinación lineal de los elementos $f_1, f_2, \ldots, f_n$:*

$$\varphi_n = a_{n1} f_1 + \ldots + a_{nn} f_n,$$

con $a_{nn} \neq 0$;
3) *todo elemento f_n se representa en la forma $f_n = b_{n1}\varphi_1 + \ldots + b_{nn}\varphi_n$, donde $b_{nn} \neq 0$.*

Todo elemento del sistema (9) queda determinado por las condiciones 1), 2) y 3) *unívocamente, salvo un factor* ± 1.

DEMOSTRACION. Representemos el elemento φ_1 en la forma

$$\varphi_1 = a_{11} f_1;$$

entonces, a_{11} se determina por la condición

$$(\varphi_1, \varphi_1) = a_{11}^2 (f_1, f_1) = 1,$$

de donde

$$a_{11} = \frac{1}{b_{11}} = \frac{\pm 1}{\sqrt{(f_1, f_1)}}.$$

Está claro que φ_1 se determina unívocamente (salvo el signo). Supongamos que han sido construidos ya los elementos φ_k $(k < n)$ que verifican las condiciones 1), 2) y 3). En este caso, se puede representar f_n en la forma

$$f_n = b_{n1}\varphi_1 + \ldots + b_{n,\, n-1}\varphi_{n-1} + h_n,$$

donde

$$(h_n, \varphi_k) = 0 \text{ para } k < n.$$

En efecto, los coeficientes correspondientes b_{nk} y, por consiguiente, el elemento h_n quedan determinados unívocamente por las condiciones

$$(h_n, \varphi_k) = (f_n - b_{n1}\varphi_1 - \ldots - b_{n,\, n-1}\varphi_{n-1}, \varphi_k) =$$
$$= (f_n, \varphi_k) - b_{nk}(\varphi_k, \varphi_k) = 0.$$

Es evidente que $(h_n, h_n) > 0$ (la suposición $(h_n, h_n) = 0$ estaría en contradicción con la independencia lineal del sistema (8)). Pongamos

$$\varphi_n = \frac{h_n}{\sqrt{(h_n, h_n)}}.$$

De esta construcción inductiva se desprende que h_n y, por consiguiente, también φ_n se expresan mediante $f_1, \ldots, f_n$, es decir,

$$\varphi_n = a_{n1} f_1 + \ldots + a_{nn} f_n, \text{ donde } a_{nn} = \frac{1}{\sqrt{(h_n, h_n)}} \neq 0. \text{ Además}$$

$$(\varphi_n, \varphi_n) = 1, \quad (\varphi_n, \varphi_k) = 0 \quad (k < n).$$

y

$$f_n = b_{n1}\varphi_1 + \ldots + b_{nn}\varphi_n \quad (b_{nn} = \sqrt{(h_n, h_n)} \neq 0),$$

es decir, φ_n verifica las condiciones del teorema.

El paso del sistema (8) al sistema (9) que satisface las condiciones 1), 2) y 3) se llama *proceso de ortogonalización*.

Está claro, que los subespacios generados por los sistemas (8) y (9) coinciden. De suerte que estos sistemas son completos o no lo son simultáneamente.

COROLARIO. *En todo espacio euclídeo separable R existe una base ortonormal.*

En efecto, sea $\psi_1, \psi_2, \ldots, \psi_n, \ldots$ un conjunto numerable siempre denso en R. Escojamos de él un sistema completo de elementos linealmente independientes $\{f_n\}$. Para ello bastará con omitir de la sucesión $\{\psi_n\}$ todos aquellos elementos ψ_k que pueden representarse como una combinación lineal de ψ_i con $i < k$. Aplicando el proceso de ortogonalización al sistema completo de elementos linealmente independientes obtenido de esta forma, encontraremos la base ortonormal.

EJERCICIOS. 1. Dése un ejemplo de un espacio euclídeo (no separable) en el que no existe ninguna base ortogonal. Demuéstrese que en un espacio euclídeo completo (no necesariamente separable) existe una base ortonormal.

2. Demuéstrese que en un espacio euclídeo completo (no necesariamente separable) toda sucesión de conjuntos encajados no vacíos, convexos, cerrados y acotados, tiene una intersección no vacía (compárese con los ejercicios de las págs. 75 y 152).

4°. Desigualdad de Bessel. Sistemas ortogonales cerrados. Introduciendo en el espacio euclídeo R^n de n dimensiones una base ortonormal $e_1, e_2, \ldots, e_n$, todo vector $x \in R^n$ puede representarse en la forma

$$x = \sum_{k=1}^{n} c_k e_k, \tag{10}$$

donde

$$c_k = (x, e_k). \tag{11}$$

Veamos de qué forma puede generalizarse el desarrollo (10) al caso de un espacio euclídeo de dimensión infinita. Sea

$$\varphi_1, \varphi_2, \ldots, \varphi_n, \ldots \tag{12}$$

un sistema ortonormal en un espacio euclídeo R y sea f un elemento arbitrario de R. Pongamos en correspondencia a todo elemento $f \in R$ la sucesión de números

$$c_k = (f, \varphi_k), \quad k = 1, 2, \ldots, \tag{13}$$

que llamaremos *coordenadas* o *coeficientes de Fourier* del elemento f según el sistema $\{\varphi_k\}$, y la serie (por ahora formal)

$$\sum_k c_k \varphi_k, \tag{14}$$

que llamaremos *serie de Fourier* del elemento f según el sistema ortogonal $\{\varphi_n\}$.

Surgen lógicamente las siguientes preguntas: ¿converge la serie (14), es decir, tiende a algún límite (en el sentido de la métrica del espacio R) la sucesión de sus sumas parciales? y si converge ¿coincide su suma con el elemento inicial f?

Para responder a estas preguntas, consideremos primero el siguiente problema: para un n dado hay que escoger los coeficientes α_k $(k=1, 2, \ldots, n)$ de manera que la distancia entre f y la suma

$$S_n=\sum_{k=1}^{n}\alpha_k\varphi_k \tag{15}$$

sea minimal. Calculemos esta distancia. Como el sistema (12) es ortonormal, tenemos

$$\|f-S_n\|^2=\left(f-\sum_1^n\alpha_k\varphi_k,\ f-\sum_1^n\alpha_k\varphi_k\right)=$$
$$=(f, f)-2\left(f, \sum\alpha_k\varphi_k\right)+\left(\sum_{k=1}^{n}\alpha_k\varphi_k,\ \sum_{j=1}^{n}\alpha_j\varphi_j\right)=$$
$$=\|f\|^2-2\sum\alpha_kc_k+\sum\alpha_k^2=\|f\|^2-\sum_{k=1}^{n}c_k^2+\sum_{k=1}^{n}(\alpha_k-c_k)^2.$$

Está claro que esta expresión alcanza su mínimo cuando el último sumando es igual a 0, es decir, para

$$\alpha_k=c_k,\quad k=1, 2, \ldots, n. \tag{16}$$

En este caso,

$$\|f-S_n\|^2=\|f\|^2-\sum_{k=1}^{n}c_k^2. \tag{17}$$

Hemos demostrado que para un n prescrito entre todas las sumas de tipo (15) la de menor desviación de f es la suma parcial de la serie de Fourier del elemento f. Geométricamente este resultado se puede interpretar del siguiente modo. El elemento

$$f-\sum_{k=1}^{n}\alpha_k\varphi_k$$

es ortogonal a todas las combinaciones lineales de tipo

$$\sum_{k=1}^{n}\beta_k\varphi_k,$$

es decir, es ortogonal al subespacio generado por los elementos $\varphi_1, \varphi_2, \ldots, \varphi_n$ cuando, y sólo cuando, se cumple la condición (16) (¡compruébese esto!). Por consiguiente, el resultado obtenido representa una generalización del conocido teorema de la Geometría Elemental: la longitud de la perpendicular bajada de un punto dado a una recta o a un plano es menor que la longitud de cualquier oblicua trazada por el mismo punto.

Puesto que siempre $\|f-S_n\|^2 \geqslant 0$, de la igualdad (17) fluye que

$$\sum_{k=1}^{n} c_k^2 \leqslant \|f\|^2.$$

Aquí n es arbitrario y el miembro derecho no depende de n; por consiguiente, la serie $\sum_{k=1}^{\infty} c_k^2$ converge y

$$\sum_{k=1}^{\infty} c_k^2 \leqslant \|f\|^2. \tag{18}$$

Esta desigualdad se llama *desigualdad de Bessel*. Geométricamente significa lo siguiente; la suma de los cuadrados de las proyecciones de un vector f sobre direcciones mutuamente ortogonales no sobrepasa el cuadrado de la longitud del propio vector f.

Introduzcamos el siguiente concepto importante.

DEFINICION. Un sistema ortonormal (12) se llama *cerrado*, cuando para cualquier $f \in R$ se cumple la igualdad

$$\sum_{k=1}^{\infty} c_k^2 = \|f\|^2, \tag{19}$$

llamada *igualdad de Parseval*.

De la identidad (17) se deduce que el sistema (12) es cerrado si, y sólo si, para todo $f \in R$ las sumas parciales de la serie de Fourier $\sum c_n \varphi_n$ convergen a f.

El concepto de un sistema ortonormal cerrado está íntimamente ligado al concepto de un sistema completo introducido anteriormente.

TEOREMA 2. *En un espacio euclídeo separable R todo sistema ortonormal completo es cerrado y viceversa.*

DEMOSTRACION. Sea el sistema $\{\varphi_n\}$ cerrado; entonces, cualquiera que sea el elemento $f \in R$, la sucesión de las sumas parciales de su serie de Fourier converge a f. Esto significa que las combinaciones lineales de los elementos del sistema $\{\varphi_n\}$ son siempre densas en R, es decir, que el sistema $\{\varphi_n\}$ es completo. Viceversa, supongamos que el sistema $\{\varphi_n\}$ es completo, esto es, todo elemento $f \in R$ se puede aproximar con precisión arbitraria mediante combinaciones lineales

$$\sum_{k=1}^{n} \alpha_k \varphi_k$$

de elementos del sistema $\{\varphi_n\}$; la suma parcial

$$\sum_{k=1}^{n} c_k\varphi_k$$

de la serie de Fourier de f da una aproximación no menos exacta. Por consiguiente, la serie

$$\sum_{k=1}^{\infty} c_k\varphi_k$$

converge a f y tiene lugar la igualdad de Parseval.

En el punto anterior hemos demostrado la existencia de sistemas ortonormales completos en un espacio euclídeo separable. Como para los sistemas ortonormales los conceptos de sistema cerrado y completo coinciden, la existencia en R de sistemas ortonormales cerrados no necesita una nueva demostración y los ejemplos de sistemas ortonormales completos, señalados en el punto anterior, son al mismo tiempo, ejemplos de sistemas cerrados.

En la exposición anterior los sistemas ortogonales se suponían normales. Se puede enunciar los conceptos de coeficientes de Fourier, de serie de Fourier, etc., para cualesquiera sistemas ortogonales. Sea $\{\varphi_n\}$ un sistema ortogonal arbitrario. A partir de él podemos construir un sistema normal, compuesto por los elementos $\psi_n=\frac{\varphi_n}{\|\varphi_n\|}$. Para todo $f\in R$ tenemos

$$c_n=(f,\ \psi_n)=\frac{1}{\|\varphi_n\|}(f,\ \varphi_n)\ \text{ y }\ \sum c_n\psi_n=\sum\frac{c_n}{\|\varphi_n\|}\varphi_n=\sum a_n\varphi_n,$$

donde

$$a_n=\frac{c_n}{\|\varphi_n\|}=\frac{(f,\ \varphi_n)}{\|\varphi_n\|^2}. \tag{20}$$

Los coeficientes a_n, definidos mediante la fórmula (20), se llaman *coeficientes de Fourier* del elemento f según el sistema ortogonal (no normal) $\{\varphi_n\}$. Tomando en la desigualdad (18) en lugar de c_n sus expresiones $c_n=a_n\|\varphi_n\|$, deducidas de (20), obtenemos

$$\sum_n \|\varphi_n\|^2 a_n^2\leqslant\|f\|^2 \tag{21}$$

que representa la desigualdad de Bessel para un sistema ortogonal arbitrario.

5°. Espacios euclídeos completos. Teorema de Riesz—Fisher. Comenzando desde el punto 3, hemos considerado espacios euclideos separables; desde este momento vamos a suponer, además, que los espacios considerados son completos.

Sea, pues, R un espacio euclídeo separable completo y sea $\{\varphi_n\}$ un sistema ortonormal en él (no necesariamente completo). De la desigualdad de Bessel se deduce que para que los números $c_1, c_2, \ldots, c_n, \ldots$ representen los coeficientes de Fourier de un elemento $x \in R$ es necesario que la serie

$$\sum_n c_n^2$$

converja. Resulta que en un espacio completo esta condición no sólo es necesaria, sino también suficiente. Tiene lugar el siguiente teorema.

TEOREMA 3 (Riesz—Fisher). *Sea $\{\varphi_n\}$ un sistema ortonormal arbitrario en un espacio euclídeo completo R y sean los números*

$$c_1, c_2, \ldots, c_n, \ldots .$$

tales que la serie

$$\sum_{k=1}^{\infty} c_k^2 \tag{22}$$

converge. Entonces, existe un elemento $f \in R$ tal que

$$c_k = (f, \varphi_k)$$

y

$$\sum_{k=1}^{n} c_k^2 = (f, f) = \|f\|^2.$$

DEMOSTRACION. Pongamos

$$f_n = \sum_{k=1}^{n} c_k \varphi_k.$$

Entonces,

$$\|f_{n+p} - f_n\|^2 = \|c_{n+1}\varphi_{n+1} + \ldots + c_{n+p}\varphi_{n+p}\|^2 = \sum_{k=n+1}^{n+p} c_k^2;$$

como la serie (22) converge, de aquí se deduce, en virtud de la complitud de R, la convergencia de la sucesión $\{f_n\}$ a un elemento $f \in R$. Además,

$$(f, \varphi_i) = (f_n, \varphi_i) + (f - f_n, \varphi_i), \tag{23}$$

donde el primer sumando del miembro derecho es igual a c_i para $n \geqslant i$, mientras que el segundo sumando tiende a cero para $n \to \infty$, ya que

$$|(f - f_n, \varphi_i)| \leqslant \|f - f_n\| \cdot \|\varphi_i\|.$$

El miembro izquierdo de la igualdad (23) no depende de n; por eso, pasando al límite para $n \to \infty$, obtenemos

$$(f, \varphi_i) = c_i.$$

De acuerdo con la definición de f,

$$\|f - f_n\| \to 0 \text{ para } n \to \infty$$

y por eso

$$\sum_{k=1}^{\infty} c_k^2 = (f, f).$$

En efecto,

$$\left(f - \sum_{k=1}^{n} c_k \varphi_k, \ f - \sum_{k=1}^{n} c_k \varphi_k\right) = (f, f) - \sum_{k=1}^{n} c_k^2 \to 0$$

para $n \to \infty$. El teorema queda demostrado.

Demostremos, para concluir, el siguiente teorema útil.

TEOREMA 4. *Para que un sistema ortonormal $\{\varphi_n\}$ de un espacio euclídeo separable completo sea completo es necesario y suficiente que en R no exista ningún elemento diferente de cero que sea ortogonal a todos los elementos del sistema $\{\varphi_n\}$.*

DEMOSTRACION. Sea $\{\varphi_n\}$ un sistema completo y por consiguiente, cerrado. Si f es ortogonal a todos los elementos del sistema $\{\varphi_n\}$, todos sus coeficientes de Fourier se anulan. Entonces, de la igualdad de Parseval obtenemos

$$(f, f) = \sum_{k=1}^{\infty} c_k^2 = 0,$$

es decir, $f = 0$.

Viceversa, supongamos que el sistema $\{\varphi_n\}$ no es completo, esto es, existe en R un elemento $g \neq 0$ tal que

$$(g, g) > \sum_{k=1}^{n} c_k^2 \ (\text{donde } c_k = (g, \varphi_k)).$$

Entonces, de acuerdo con el teorema de Riesz—Fisher, existe un elemento $f \in R$ tal que

$$(f, \varphi_k) = c_k, \ (f, f) = \sum_{k=1}^{\infty} c_k^2.$$

El elemento $f - g$ es ortogonal a todos los φ_i. De la desigualdad

$$(f, f) = \sum_{k=1}^{\infty} c_k^2 < (g, g)$$

se desprende que $f - g \neq 0$. El teorema queda demostrado.

EJERCICIOS. 1. Sea H un espacio euclídeo completo (no necesariamente separable); entonces existe en él un sistema ortonormal completo $\{\varphi_\alpha\}$ (véase el ejercicio de la pág. 159). Demuéstrese que para todo vector $f \in H$ tienen lugar los desarrollos

$$f=\sum_{\alpha}(f,\ \varphi_\alpha)\varphi_\alpha,\quad \|f\|^2=\sum_{\alpha}|(f,\ \varphi_\alpha)|^2,$$

donde las sumas que figuran a la derecha tienen a lo sumo un numero numerable de sumandos diferentes de 0.

2. Un sistema $\{\varphi_\alpha\}$ de vectores de un espacio euclídeo R se llama *total*, cuando en R no existen elementos diferentes de 0 ortogonales a todos los $\{\varphi_\alpha\}$. El teorema 4 significa que en un espacio euclídeo completo la totalidad es equivalente a la complitud. Demuéstrese que en los espacios no completos pueden existir sistemas totales, pero no completos.

6°. Espacio de Hilbert. Teorema sobre el isomorfismo. Continuemos la consideración de espacios euclídeos completos. Nos interesarán, al igual que antes, los espacios de dimensión infinita y no de dimensión finita que se describen completamente en los cursos del Algebra lineal. Al igual que antes, admitiremos la existencia de un conjunto numerable siempre denso en los espacios considerados. Introduzcamos la siguiente definición.

DEFINICION. Un espacio euclídeo separable completo de dimensión infinita se llama *espacio de Hilbert*[1]. Es decir, un espacio de Hilbert es un conjunto H de elementos f, g, ... de naturaleza arbitraria que verifica las siguientes condiciones:

I. H es un espacio euclídeo (es decir, un espacio lineal con un producto escalar definido en él).

II. El espacio H es completo en el sentido de la métrica $\rho(f,\ g)=\|f-g\|$.

III. El espacio H es de dimensión infinita, esto es, cualquiera que sea n se puede encontrar en él n elementos linealmente independientes.

IV. H es separable, esto es, existe en él un conjunto numerable siempre denso.

Como ejemplo de un espacio de Hilbert podemos indicar el espacio real l_2.

Recordemos que dos espacios euclídeos R y R^* se llaman isomorfos, cuando se puede establecer entre sus elementos una correspondencia biunívoca de manera que si

$$x \leftrightarrow x^*,\qquad y \leftrightarrow y^*,$$

$(x,\ y \in R;\ x^*,\ y^* \in R^*)$, se tiene

[1] Por el apellido del famoso matemático alemán David Hilbert (1862—1943) que introdujo este concepto.

$$x+y \leftrightarrow x^* + y^*,$$
$$\alpha x \leftrightarrow \alpha x^*$$

y

$$(x, y) = (x^*, y^*).$$

En otras palabras, un isomorfismo de espacios euclídeos es una correspondencia biunívoca que conserva tanto las operaciones lineales, definidas en estos espacios, como el producto escalar.

Como se sabe, dos espacios euclídeos arbitrarios de n dimensiones son isomorfos y, por consiguiente, cualquier espacio de este tipo es isomorfo al espacio de coordenadas R^n (ejemplo 1). Los espacios euclídeos de infinita dimensión no son necesariamente isomorfos entre sí. Por ejemplo, los espacios l_2 y $C^2_{[a, b]}$ no son isomorfos. Esto se ve, por ejemplo, de que el primero de ellos es completo y el segundo no lo es.

Sin embargo, tiene lugar el siguiente teorema.

TEOREMA 5. *Dos espacios de Hilbert cualesquiera son isomorfos.*

DEMOSTRACION. Probemos que todo espacio de Hilbert H es isomorfo a l_2. Con esto quedará demostrada la afirmación del teorema. Escojamos en H un sistema ortonormal completo arbitrario $\{\varphi_n\}$ y pongamos en correspondencia a todo elemento $f \in H$ el conjunto $c_1, c_2, \ldots, c_n, \ldots$ de sus coeficientes de Fourier según este sistema. Puesto que $\sum c_k^2 < \infty$, la sucesión $(c_1, c_2, \ldots, \ldots, c_n, \ldots)$ es un elemento de l_2. Viceversa, en virtud del teorema de Riesz—Fisher, a todo elemento $(c_1, c_2, \ldots, c_n, \ldots)$ de l_2 le corresponde un elemento $f \in H$ para el cual los números $c_1, c_2, \ldots, c_n \ldots$ son sus coeficientes de Fourier. La correspondencia establecida entre los elementos de H y l_2 es biunívoca. Además, si

$$f^{(1)} \leftrightarrow (c_1^{(1)}, c_2^{(1)}, \ldots, c_n^{(1)}, \ldots)$$

y

$$f^2 \leftrightarrow (c_1^{(2)}, c_2^{(2)}, \ldots, c_n^{(2)}, \ldots),$$

tenemos

$$f^{(1)} + f^{(2)} \leftrightarrow (c_1^{(1)} + c_1^{(2)}, c_2^{(1)} + c_2^{(2)}, \ldots, c_n^{(1)} + c_n^{(2)}, \ldots)$$

y

$$kf^{(1)} \leftrightarrow (kc_1^{(1)}, kc_2^{(1)}, \ldots, kc_n^{(1)}, \ldots),$$

es decir, la suma se transforma en suma y el producto por un número, en el producto del elemento correspondiente por el mismo número. Finalmente, de la igualdad de Parseval se sigue que

$$(f^{(1)},\ f^{(2)})=\sum_{n=1}^{\infty} c_n^{(1)} c_n^{(2)}. \tag{24}$$

En efecto, de

$$(f^{(1)},\ f^{(1)})=\sum (c_n^{(1)})^2,\ (f^{(2)},\ f^{(2)})=\sum (c_n^{(2)})^2$$

y

$$(f^{(1)}+f^{(2)},\ f^{(1)}+f^{(2)})=(f^{(1)},\ f^{(1)})+2(f^{(1)},\ f^{(2)})+(f^{(2)},\ f^{(2)})=$$
$$=\sum (c_n^{(1)}+c_n^{(2)})^2=\sum (c_n^{(1)})^2+2\sum c_n^{(1)} c_n^{(2)}+\sum (c_n^{(2)})^2$$

se deduce (24). De esta forma, la correspondencia que hemos establecido entre los elementos de los espacios H y l_2 es efectivamente un isomorfismo; el teorema queda demostrado.

El teorema demostrado significa que, salvo un isomorfismo, existe sólo un espacio de Hilbert (es decir, que el sistema de axiomas I, II, III y IV es completo) y que el espacio l_2 puede considerarse como su «realización en coordenadas», de la misma forma que el espacio de coordenadas de n dimensiones con el producto escalar $\sum_{i=1}^{n} x_i y_i$ representa la realización en coordenadas del espacio euclídeo de n dimensiones definido axiomáticamente.

Otra realización del espacio de Hilbert la podemos obtener tomando el espacio funcional $C^2_{[a,\,b]}$ y considerando su completación. En efecto, es fácil ver que la completación R^* de todo espacio euclídeo R (en el sentido en el que hemos definido la completación de un espacio métrico en el § 3 del capítulo II) se convierte en un espacio euclídeo lineal, si las operaciones lineales y el producto escalar se definen en él prolongándolas por continuidad del espacio $R\subset R^*$, es decir, tomando

$$x+y=\lim_{n\to\infty}(x_n+y_n),\qquad \alpha x=\lim_{n\to\infty}\alpha x_n$$

y

$$(x,\ y)=\lim_{n\to\infty}(x_n,\ y_n)$$

donde $x_n\longrightarrow x$ e $y_n\longrightarrow y$, $x_n,\ y_n\in R$. (Es fácil ver que estos límites existen y no dependen de cómo se escojan las sucesiones $\{x_n\}$ e $\{y_n\}$). Entonces, la completación del espacio $C^2_{[a,\,b]}$ será un espacio euclídeo completo y, evidentemente, separable y de dimensión infinita, es decir, será un espacio de Hilbert. En el capítulo VII volveremos a tratar este tema y demostraremos que los elementos que se deben agregar a $C^2_{[a,\,b]}$ para obtener un espacio completo, también se pueden representar como funciones, pero ya discontinuas (más precisamente, como funciones de cuadrado integrable en el sentido de Lebesgue).

7°. Subespacios, complementos ortogonales, suma directa. De acuerdo con las definiciones generales del § 3, llamaremos *variedad lineal* en un espacio de Hilbert H a todo conjunto L de elementos de H tal que, si $f, g \in H$, también $\alpha f + \beta g \in L$ cualesquiera que sean los números α y β. Una variedad lineal cerrada se llamará *subespacio*.

Veamos algunos ejemplos de subespacios del espacio de Hilbert.

1. Sea h un elemento arbitrario de H. El conjunto de todos los elementos $f \in H$ ortogonales a h constituye un subespacio de H.

2. Supongamos que H está realizado mediante l_2, es decir, que sus elementos son las sucesiones $(x_1, x_2, \ldots, x_n, \ldots)$ tales que $\sum x_k^2 < \infty$. Los elementos que verifican la condición $x_1 = x_2$ forman un subespacio.

3. Supongamos de nuevo que H está realizado mediante l_2. Los elementos $x = (x_1, x_2, \ldots x_n, \ldots)$ para los cuales $x_n = 0$ cuando $n = 2, 4, 6, \ldots$ (mientras que x_n son arbitrarios para $n = 1, 3, 5, \ldots$) forman un subespacio.

Recomendamos al lector comprobar que los conjuntos de vectores de los ejemplos 1, 2 y 3 son efectivamente subespacios.

Todo subespacio de un espacio de Hilbert o bien es un espacio euclídeo de dimensión finita o bien representa un espacio de Hilbert. En efecto, la validez de los axiomas I, II y III para cualquiera de estos subespacios es evidente, y la validez del axioma IV se deduce del siguiente lema.

LEMA. *De la existencia de un conjunto numerable siempre denso en un espacio métrico R se deduce la existencia de un conjunto numerable siempre denso en cualquier subconjunto suyo R'.*

DEMOSTRACION. Sea

$$\xi_1, \xi_2, \ldots, \xi_n, \ldots$$

un conjunto numerable siempre denso en R y sea

$$a_n = \inf_{\eta \in R'} \rho(\xi_n, \eta).$$

Para cualesquiera n y m naturales existe un punto $\eta_{n,m} \in R'$ tal que

$$\rho(\xi_n, \eta_{n,m}) < a_n + \frac{1}{m}.$$

Sea $\varepsilon > 0$ y $\frac{1}{m} < \frac{\varepsilon}{3}$; para todo $\eta \in R'$ existe un n tal que

$$\rho(\xi_n, \eta) < \frac{\varepsilon}{3}$$

y, por consiguiente,

$$\rho(\xi_n, \eta_{n,m}) < a_n + \frac{1}{m} < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \frac{2\varepsilon}{3},$$

pero, entonces, $\rho(\eta, \eta_{n,m}) < \varepsilon$, es decir, el conjunto $\{\eta_{n,m}\}$ (n, $m = 1, 2 \ldots$), a lo sumo numerable, es siempre denso en R'.

Los subespacios del espacio de Hilbert poseen propiedades específicas (que no tienen lugar para los subespacios de un espacio normado arbitrario). Estas propiedades están relacionadas con la existencia en el espacio de Hilbert del producto escalar y del concepto de ortogonalidad, correspondiente a éste.

Aplicando el proceso de ortogonalización a una sucesión numerable siempre densa de elementos de un subespacio cualquiera del espacio de Hilbert, obtenemos el siguiente teorema.

TEOREMA 6. *En todo subespacio M del espacio H existe un sistema ortonormal $\{\varphi_n\}$ tal que su adherencia lineal coincide con M.*

Sea M un subespacio del espacio de Hilbert H. Denotemos mediante

$$M' = H \ominus M$$

el conjunto de elementos $g \in H$ ortogonales a todos los elementos $f \in M$ y demostremos que M' es también un subespacio del espacio H. La linealidad de M' es obvia, ya que de $(g_1, f) = (g_2, f) = 0$ fluye $(\alpha_1 g_1 + \alpha_2 g_2, f) = 0$. Para demostrar que es cerrado, admitamos que los elementos g_n pertenecen a M' y convergen a g. Entonces, para todo $f \in M$ tenemos

$$(g, f) = \lim_{n \to \infty} (g_n, f) = 0$$

y por eso g también pertenece a M'.

El subespacio M' se llama *complemento ortogonal* del subespacio M.

Del teorema 6 se deduce fácilmente que:

TEOREMA 7. *Si M es un subespacio lineal cerrado del espacio H, todo elemento $f \in H$ se puede representar de manera única en la forma $f = h + h'$, donde $h \in M$ y $h' \in M'$.*

DEMOSTRACION. Demostremos primero que existe esta descomposición. Para ello escojamos en M un sistema ortonormal completo $\{\varphi_n\}$ y pongamos

$$h = \sum_{h=1}^{\infty} c_n \varphi_n, \quad c_n = (f, \varphi_n).$$

Puesto que (debido a la desigualdad de Bessel) la serie $\sum_{n=1}^{\infty} c_n^2$

converge, el elemento h existe y pertenece a M. Tomemos

$$h' = f - h.$$

Es evidente que para todo n

$$(h', \varphi_n) = 0$$

y, como cualquier elemento de M se puede representar en la forma

$$\zeta = \sum a_n \varphi_n,$$

tenemos para $\zeta \in M$

$$(h', \zeta) = \sum_{n=1}^{\infty} a_n (h', \varphi_n) = 0,$$

es decir, $h' \in M'$.

Supongamos ahora que además de la descomposición obtenida $f = h + h'$ existe otra descomposición

$$f = h_1 + h_1', \quad h_1 \in M, \quad h_1' \in M'.$$

Entonces, tenemos para cualquier n

$$(h_1, \varphi_n) = (f, \varphi_n) = c_n$$

y de aquí se deduce que

$$h_1 = h, \quad h_1' = h'.$$

Del teorema 7 fluye el siguiente corolario.

COROLARIO 1. *El complemento ortogonal del complemento ortogonal de un subespacio lineal cerrado M coincide con M.*

De esta forma resulta posible hablar de subespacios recíprocamente complementarios del espacio H. Si M y M' son dos subespacios recíprocamente complementarios y si $\{\varphi_n\}$ y $\{\varphi_n'\}$ son dos sistemas ortogonales completos (de M y de M', respectivamente), la unión de estos sistemas $\{\varphi_n\}$ y $\{\varphi_n'\}$ ofrece un sistema ortogonal completo de todo el espacio H. Por eso, tiene lugar el siguiente corolario.

COROLARIO 2. *Todo sistema ortonormal puede ser extendido a un sistema completo de H.*

Siendo el sistema $\{\varphi_n\}$ finito, el número de elementos que lo componen coincide con la dimensión del subespacio M generado por $\{\varphi_n\}$ y con la codimensión del subespacio M'. Obtenemos, de esta forma, el siguiente corolario.

COROLARIO 3. *El complemento ortogonal a un subespacio de dimensión finita n tiene codimensión n y viceversa.*

Si todo vector $f \in H$ es representado en la forma $f=h+h'$, $h \in M$, $h' \in M'$ (donde M' es el complemento ortogonal de M), se dice que H es la suma directa de los espacios recíprocamente ortogonales M y M' y se escribe

$$H=M \oplus M'.$$

Está claro que el concepto de suma directa se puede generalizar inmediatamente a un número finito cualquiera o, incluso, a un número numerable de subespacios; a saber, se dice que H es la suma directa de sus subespacios $M_1, M_2, \ldots, M_n, \ldots$

$$H=M_1 \oplus M_2 \oplus \ldots \oplus M_n \oplus \ldots,$$

cuando

1) los subespacios M_i son ortogonales dos a dos, es decir, todo vector de M_i es ortogonal a cualquier vector de M_k para $i \neq k$;

2) todo elemento $f \in H$ se puede representar en la forma

$$f=h_1+h_2+\ldots+h_n+\ldots, \quad h_n \in M_n,$$

con la particularidad de que, si el número de subespacios M_n es infinito, la serie $\sum \|h_n\|^2$ converge. Es fácil probar, que si existe esta representación del elemento f, ella es única y que

$$\|f\|^2=\sum \|h_n\|^2.$$

Junto a la suma directa de subespacios, se puede considerar la suma directa de un número finito o numerable de espacios arbitrarios de Hilbert. Si H_1 y H_2 son dos espacios de Hilbert, la suma directa de ellos se define del siguiente modo: los elementos de H son todos los pares (h_1, h_2), donde $h_1 \in H_1$, $h_2 \in H_2$, y el producto escalar de dos de estos pares es igual a

$$((h_1, h_2), (h_1', h_2'))=(h_1, h_1')+(h_2, h_2').$$

El espacio H contiene, evidentemente, los subespacios recíprocamente ortogonales, compuestos por pares de tipo $(h_1, 0)$ y $(0, h_2)$ respectivamente; el primero de ellos puede ser identificado de un modo natural con el espacio H_1, y el segundo, con el espacio H_2.

Análogamente se define la suma de un número finito cualquiera de espacios. La suma $H=\sum \oplus H_n$ de un número numerable de espacios $H_1, H_2, \ldots, H_n, \ldots$ se define así: los elementos del espacio H son todas las sucesiones de tipo

$$h=(h_1, h_2, \ldots, h_n, \ldots), \quad (h_n \in H_n),$$

tales que

$$\sum \|h_n\|^2<\infty.$$

El producto escalar (h, g) de los elementos h y g de H es igual a

$$\sum (h_n, g_n).$$

8°. **Propiedad característica de los espacios euclídeos.** Analicemos el siguiente problema. Sea R un espacio normado. ¿Cuáles son las condiciones adicionales que debe verificar la norma, definida en R, para que el espacio R sea euclídeo, esto es, para que la norma en él se determine por un producto escalar? En otras palabras, ¿qué es lo que caracteriza los espacios euclídeos dentro de la clase de espacios normados? Esta característica viene dada por el siguiente teorema.

TEOREMA 8. *Para que un espacio normado R sea euclídeo es necesario y suficiente que para cualesquiera dos elementos f y g de él se cumpla la igualdad*

$$\|f+g\|^2+\|f-g\|^2=2(\|f\|^2+\|g\|^2). \tag{25}$$

Puesto que $f+g$ y $f-g$ son las diagonales del paralelogramo construido sobre los lados f y g, la igualdad (25) expresa la conocida propiedad de paralelogramo en un espacio euclídeo: la suma de los cuadrados de las diagonales de un paralelogramo es igual a la suma de los cuadrados de sus lados. Por consiguiente, la necesidad de esta condición es obvia. Demostremos que es suficiente. Tomemos

$$(f, g)=\frac{1}{4}(\|f+g\|^2-\|f-g\|^2) \tag{26}$$

y probemos que, si se cumple la igualdad (25), la función (26) satisface todos los axiomas del producto escalar. Puesto que para $f=g$ tenemos

$$(f, f)=\frac{1}{4}(\|2f\|^2-\|f-f\|^2)=\|f\|^2, \tag{27}$$

ésta será precisamente aquel producto escalar que induce en el espacio R la norma definida en él.

Ante todo, se ve inmediatamente de (26) que

$$(f, g)=(g, f),$$

es decir, se cumple la condición 1) de la definición del producto escalar. Además, se cumple también, en virtud de (27), la condición 4) de esta definición. Para demostrar la condición 2) consideremos la siguiente función de tres vectores

$$\Phi(f, g, h)=4[(f+g, h)-(f, h)-(g, h)].$$

es decir,

$$\Phi(f,\ g,\ h)=\|f+g+h\|^2-\|f+g-h\|^2-\|f+h\|^2+ \\ +\|f-h\|^2-\|g+h\|^2+\|g-h\|^2, \qquad (28)$$

y demostremos que es igual idénticamente a cero. De acuerdo con (25), tenemos

$$\|f+g\pm h\|^2=2\|f\pm h\|^2+2\|g\|^2-\|f\pm h-g\|^2.$$

Sustituyendo las correspondientes expresiones en $\Phi(f,\ g,\ h)$, encontramos

$$\Phi(f,\ g,\ h)=-\|f+h-g\|^2+\|f-h-g\|^2+ \\ +\|f+h\|^2-\|f-h\|^2-\|g+h\|^2+\|g-h\|^2. \qquad (29)$$

Tomando la suma media de (28) y (29), tenemos

$$\Phi(f,\ g,\ h)=\frac{1}{2}(\|g+h+f\|^2+\|g+h-f\|^2)- \\ -\frac{1}{2}(\|g-h+f\|^2+\|g-h-f\|^2)-\|g+h\|^2+\|g-h\|^2.$$

El primer paréntesis, en virtud de (25), es igual a

$$\|g+h\|^2+\|f\|^2$$

y el segundo, a

$$-\|g+h\|^2-\|f\|^2.$$

Es decir,

$$\Phi(f,\ g,\ h)\equiv 0.$$

Consideremos ahora, para cualesquiera f y g fijos, la función

$$\varphi(c)=(cf,\ g)-c(f,\ g).$$

De (26) se deduce inmediatamente que

$$\varphi(0)=\frac{1}{4}(\|g\|^2-\|g\|^2)=0;$$

además, $\varphi(-1)=0$, ya que $(-f,\ g)=-(f,\ g)$. Por eso para cualquier n entero

$$(nf,\ g)=(\operatorname{sign} n\,(f+\ldots+f),\ g)=\operatorname{sign} n\,[(f,\ g)+\ldots \\ \ldots+(f,\ g)]=|n|\operatorname{sign} n\,(f,\ g)=n(f,\ g),$$

es decir, $\varphi(n)=0$. Para p y q enteros y $q\neq 0$,

$$\left(\frac{p}{q}f,\ g\right)=p\left(\frac{1}{q}f,\ g\right)=\frac{p}{q}q\left(\frac{1}{q}f,\ g\right)=\frac{p}{q}(f,\ g),$$

es decir, $\varphi(c)=0$ para todo c racional; como la función φ es continua,

$$\varphi(c)\equiv 0.$$

Hemos demostrado con esto que la función (f, g) posee todas las propiedades del producto escalar.

Ejemplos. 1. Consideremos el espacio n-dimensional R_p^n en el que la norma se define por la fórmula

$$\|x\|_p = \left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}}.$$

Para $p \geqslant 1$ se cumplen todos los axiomas de la norma; sin embargo, R_p^n será un espacio euclídeo sólo cuando $p=2$. En efecto, consideremos en R_p^n dos vectores

$$f = (1, 1, 0, 0, \ldots, 0),$$
$$g = (1, -1, 0, 0, \ldots, 0);$$

tenemos

$$f+g = (2, 0, 0, \ldots, 0),$$
$$f-g = (0, 2, 0, \ldots, 0),$$

de donde

$$\|f\|_p = \|g\|_p = 2^{\frac{1}{p}}, \quad \|f+g\|_p = \|f-g\|_p = 2,$$

es decir, la identidad del paralelogramo (25) no se cumple para $p \neq 2$.

2. Consideremos el espacio de funciones continuas sobre el segmento $\left[0, \frac{\pi}{2}\right]$. Pongamos

$$f(t) = \cos t, \quad g(t) = \operatorname{sen} t.$$

Tenemos

$$\|f\| = 1, \quad \|g\| = 1$$

y

$$\|f+g\| = \max_{0 \leqslant t \leqslant \frac{\pi}{2}} |\cos t + \operatorname{sen} t| = \sqrt{2};$$

$$\|f-g\| = \max_{0 \leqslant t \leqslant \frac{\pi}{2}} |\cos t - \operatorname{sen} t| = 1.$$

De aquí se ve que

$$\|f+g\|^2 + \|f-g\|^2 \neq 2(\|f\|^2 + \|g\|^2).$$

Por consiguiente, la norma del espacio $C_{\left[0, \frac{\pi}{2}\right]}$ no puede ser definida mediante ningún producto escalar. Es fácil ver que el espacio $C_{[a, b]}$ de funciones continuas sobre cualquier segmento $[a, b]$ tampoco es un espacio euclídeo.

9°. **Espacios euclídeos complejos.** Junto al espacio real se puede introducir también el espacio euclídeo *complejo* (esto es, el espacio lineal complejo con producto escalar). Pero en el caso complejo resulta preciso modificar los axiomas mediante los cuales se define el producto escalar en el caso real, ya que los axiomas 1), 2), 3) y 4), enunciados al principio de este parágrafo, no pueden cumplirse simultáneamente en un espacio complejo. En efecto, de 1) y 3) se deduce

$$(\lambda x, \lambda x) = \lambda^2 (x, x),$$

de donde obtenemos para $\lambda = i$

$$(ix, ix) = -(x, x),$$

es decir, los cuadrados escalares de los vectores x e ix no pueden ser positivos al mismo tiempo. En otras palabras, los axiomas 1) y 3) son incompatibles con el axioma 4). Por eso, definiremos el producto escalar en un espacio complejo como una función (x, y) numérica (de valores complejos) de dos vectores que verifica las siguientes condiciones:

1) $(x, y) = \overline{(y, x)}$,
2) $(\lambda x, y) = \lambda (x, y)$,
3) $(x_1 + x_2, y) = (x_1, y) + (x_2, y)$,

4) $(x, x) \geqslant 0$ y, además, $(x, x) > 0$ cuando $x \neq 0$.

(Por consiguiente, modificamos el primer axioma conservando los tres restantes). De 1) y 2) se deduce que $(x, \lambda y) = \bar{\lambda} (x, y)$. En efecto,

$$(x, \lambda y) = \overline{(\lambda y, x)} = \lambda \overline{(y, x)} = \bar{\lambda} (x, y).$$

Un ejemplo bien conocido de un espacio euclídeo complejo de n dimensiones es el espacio lineal C^n (§ 1, ejemplo 2), en el que el producto escalar de los elementos

$$x = (x_1, x_2, \ldots, x_n) \text{ e } y = (y_1, y_2, \ldots, y_n)$$

se define por

$$(x, y) = \sum_{k=1}^{n} x_k \overline{y_k}.$$

Como se sabe, todo espacio euclídeo complejo de dimensión n es isomorfo a este espacio.

Como ejemplos de espacios euclídeos complejos de dimensión infinita pueden servir:

1) el espacio complejo l_2, cuyos elementos son sucesiones de números complejos

$$x=(x_1,\ x_2,\ \ldots,\ x_n,\ \ldots),$$

que verifican la condición

$$\sum_{n=1}^{\infty}|x_n|^2<\infty,$$

y en el que el producto escalar se define mediante la fórmula

$$(x,\ y)=\sum_{n=1}^{\infty}x_n\overline{y_n};$$

2) el espacio $C^2_{[a,\ b]}$ de funciones de valores complejos continuas sobre el segmento $[a,\ b]$ con el producto escalar

$$(f,\ g)=\int_a^b f(t)\overline{g(t)}\,dt.$$

En un espacio euclídeo complejo la longitud (norma) de un vector se define, al igual que en el caso real, mediante la fórmula

$$\|x\|=\sqrt{(x,\ x)}.$$

El concepto de ángulo entre vectores en el caso complejo no se introduce, ya que la magnitud $\frac{(x,\ y)}{\|x\|\cdot\|y\|}$ es, en general, compleja y puede no ser el coseno de ningún ángulo real; no obstante, se conserva el concepto de ortogonalidad: los elementos x e y se llaman recíprocamente ortogonales cuando $(x,\ y)=0$.

Si $\{\varphi_n\}$ es un sistema ortogonal de un espacio euclídeo complejo R y f es un elemento arbitrario de R, los números

$$a_n=\frac{1}{|\varphi_n|^2}(f,\ \varphi_n)$$

se llaman, al igual que en el caso real, coeficientes de Fourier y la serie

$$\sum_n a_n\varphi_n,$$

serie de Fourier del elemento f según el sistema ortogonal $[\varphi_n]$. Tiene lugar la desigualdad de Bessel

$$\sum_n\|\varphi_n\|^2|a_n|^2\leqslant(f,\ f).$$

En particular, si el sistema $[\varphi_n]$ es ortonormal, los coeficientes de Fourier según este sistema se definen mediante las fórmulas

$$c_n = (f, \varphi_n),$$

y la desigualdad de Bessel toma la forma

$$\sum_n |c_n|^2 \leqslant (f, f).$$

Un espacio euclídeo complejo separable y completo de dimensión infinita se llama espacio complejo de Hilbert. En el caso complejo subsiste el teorema sobre el isomorfismo de los espacios de Hilbert:

TEOREMA. 9. *Todos los espacios complejos de Hilbert son isomorfos entre sí.*

La realización más sencilla de un espacio complejo de Hilbert es el espacio complejo l_2. En el capítulo VII veremos otra realización, de carácter funcional, del espacio complejo de Hilbert.

Proponemos al lector comprobar que todos los teoremas demostrados anteriormente para los espacios reales euclídeos y de Hilbert, son válidos (con modificaciones insignificantes, debidas a la complejidad del producto escalar) también para los espacios complejos euclídeos y de Hilbert.

§ 5. ESPACIOS TOPOLOGICOS LINEALES

1°. Definición y ejemplos. La definición de la norma es sólo una de las formas posibles de introducir una topología en un espacio lineal. El desarrollo de ramas del Análisis Funcional como la teoría de funciones generalizadas (trataremos de ellas en el capítulo siguiente) ha demostrado que en muchos casos conviene considerar espacios lineales con una topología definida no mediante la norma, sino de alguna otra manera.

DEFINICION. Un conjunto E se llama *espacio topológico lineal*, cuando

I. E representa un espacio lineal (con la multiplicación de elementos por números reales o complejos);

II. E es un espacio topológico;

III. las operaciones de adición y multiplicación por números en E son continuas respecto a la topología existente en E. Más detalladamente la última condición significa lo siguiente:

1) si $z_0 = x_0 + y_0$, para toda vecindad U del punto z_0 se pueden indicar vecindades V y W de los puntos x_0 e y_0 tales que $x + y \in U$ para $x \in V$, $y \in W$;

2) si $\alpha_0 x_0 = y_0$, para toda vecindad U del punto y_0 existe una vecindad V del punto x_0 y un número $\varepsilon > 0$ tales que $\alpha x \in U$ para $[\alpha_0 - \alpha] < \varepsilon$ y $x \in V$.

De la relación, existente en un espacio topológico lineal entre las operaciones algebraicas y la topología, se deduce que la topología en este espacio queda totalmente determinada al dar el *sistema de vecindades del cero*. En efecto, sea x un punto de un espacio topológico E y sea U una vecindad del cero en E. Entonces, $U + x$, esto es, la «traslación» de esta vecindad paralelamente a x, es una vecindad del punto x; es evidente, que toda vecindad de cualguier punto $x \in E$ puede obtenerse de esta forma.

De la continuidad, en un espacio topológico lineal E, de las operaciones de adición y multiplicación por números, resultan inmediatamente las siguientes afirmaciones.

1) Si U y V son conjuntos abiertos de E, el conjunto $U + V$ (es decir, la totalidad de elementos de tipo $x + y$, $x \in U$, $y \in V$) es abierto.

2) Si U es abierto, el conjunto λU (es decir, la totalidad de elementos de tipo λx, $x \in U$) es también abierto para todo $\lambda \neq 0$.

3) Si $F \subset E$ es cerrado, el conjunto λF es también cerrado para todo λ.

Ejemplos. 1. Ante todo, son espacios topológicos lineales todos los espacios normados. En efecto, de las propiedades de la norma se deduce inmediatamente que las operaciones de adición de vectores y multiplicación de éstos por números son, en un espacio normado, continuas respecto a aquella topología que induce en él la norma.

2. En el espacio R^∞ de todas las sucesiones numéricas $x = (x_1, x_2, \ldots, x_n, \ldots)$ definamos el sistema de vecindades del cero como sigue: cada una de estas vecindades $U(k_1, \ldots, k_r; \varepsilon)$ se define por los números enteros $k_1, \ldots, k_r$ y el número $\varepsilon > 0$ y consta de todos los $x \in R^\infty$ que verifican las condiciones:

$$|x_{k_i}| < \varepsilon, \quad i = 1, 2, \ldots, r.$$

(El espacio R^∞ puede ser tanto real como complejo).

3. Sea $K_{[a, b]}$ el espacio de funciones indefinidamente diferenciables[1]) sobre el segmento $[a, b]$. Definamos la topología en $K_{[a, b]}$ por medio del siguiente sistema de vecindades del cero. Cada una de estas vecindades $U_{m, \varepsilon}$ se determina por el número m y por el número $\varepsilon > 0$ y consta de todas las funciones φ que verifican las desigualdades

$$|\varphi^{(k)}(x)| < \varepsilon, \quad k = 0, 1, 2, \ldots, m,$$

donde $\varphi^{(k)}$ es la derivada de orden k de la función φ.

---

[1]) Es decir, de funciones que tienen derivadas de todos los órdenes.

El hecho de que en un espacio lineal topológico la topología esté relacionada con las operaciones lineales definidas en él, impone limitaciones bastante rígidas a su topología. Resulta que *todo espacio topológico lineal verifica el así llamado axioma de separabilidad* T_3, esto es, que cualquier vecindad U de un punto arbitrario x contiene una vecindad menor W del mismo punto que junto con su adherencia pertenece a U.

Para demostrar esta proposición basta considerar las vecindades del cero. Sea U una vecindad cualquiera del cero. Debido a la continuidad en E de la operación de sustracción, existirá una vecindad W del cero tal que $W-W\subset U$. Probemos que la adherencia de la vecindad W está contenida en U. Sea $y\in[W]$. En este caso toda vecindad del punto y, en particular $y+W$, contiene un punto de W. Por consiguiente, existe un punto $z\in W$ tal que $y+z\in W$, es decir, $y\in W-W\subset U$, que es lo que se afirmaba.

Un espacio topológico lineal se llama *separable*, cuando verifica el axioma de separabilidad T_1, esto es, cuando todo subconjunto suyo compuesto de un punto resulta cerrado; es evidente que un espacio es separable cuando, y sólo cuando, la intersección de todas las vecindades del cero no contiene elementos diferentes de cero. Los espacios topológicos que verifican los axiomas de separabilidad T_1 y T_3 suelen llamarse *regulares*; de lo demostrado en el párrafo anterior se deduce que un *espacio topológico lineal separable es regular*.

En los espacios normados desempeña un papel importante el concepto de conjunto acotado. Aunque este concepto se introduce allí mediante la norma, puede ser, claro está, enunciado también para los espacios topológicos lineales arbitrarios.

Un conjunto M, situado en un espacio topológico lineal E, se llama *acotado*, cuando para toda vecindad U del cero existe un $n>0$ tal que $nU\supset M$.

Está claro que en el caso de los espacios normados este concepto de acotación coincide con la acotación según la norma (es decir, con la posibilidad de situar el conjunto dado dentro de una bola $\|x\|\leqslant R$). Un espacio E se llama *localmente acotado*, cuando existe en él al menos un conjunto acotado abierto no vacío. Todo espacio normado es localmente acotado. Como ejemplo de un espacio que no es localmente acotado, puede servir el espacio R^∞, señalado en el ejemplo 2 (¡demuéstrese esto!).

EJERCICIO. Sea E un espacio topológico lineal; demuéstrese la validez de las afirmaciones siguientes:

(a) un conjunto $M\subset E$ es acotado si, y sólo si, cualesquiera que sean la sucesión $\{x_n\}\subset M$ y la sucesión $\{\varepsilon_n\}$ de números positivos, convergente a 0, la sucesión $\{\varepsilon_n x_n\}$ también converge a cero;

(b) si $\{x_n\}_{n=1}^{\infty} \subset E$ y $x_n \longrightarrow x$, el conjunto $\{x_n\}$ es acotado;

(c) si E es localmente acotado, en él se cumple el primer axioma de numerabilidad.

2°. Convexidad local. Los espacios topológicos lineales arbitrarios pueden tener propiedades muy diferentes de las propiedades habituales de los espacios euclídeos o normados. Una clase importante de espacios, más generales que los normados, pero en los que subsisten muchas propiedades de los últimos, la forman los así llamados espacios *localmente convexos.*

DEFINICION. Un espacio topológico lineal E se llama *localmente convexo*, cuando todo conjunto abierto no vacío de él contiene un subconjunto abierto convexo no vacío.

Observemos que, si el espacio E es localmente convexo, para todo punto $x \in E$ y toda vecindad suya U existirá una vecindad convexa del cero V tal que $x \in V \subset U$. En efecto, basta comprobar la validez de esta afirmación para el punto $x = 0$. Sea U cualquier vecindad del cero. Existe una vecindad del cero V tal que $V - V \subset U$. Como E es localmente convexo, existe un conjunto abierto convexo no vacío $V' \subset V$; sea $y \in V'$; entonces, $V' - y$ es una vecindad convexa del cero contenida en U.

Todo espacio normado es localmente convexo. En efecto, todo conjunto abierto no vacío de él contiene una bola abierta. De esta forma, todo espacio normado es localmente acotado y localmente convexo. Se puede demostrar que la clase formada por espacios con estas dos propiedades se reduce, de hecho, a los espacios normados. Con más precisión: un espacio lineal topológico E se llamará *normable*, cuando la topología existente en E puede definirse mediante una norma; tiene lugar el siguiente teorema: *todo espacio topológico lineal separable, localmente convexo y localmente acotado, es normable.*

EJERCICIOS. 1. Demuéstrese que un conjunto abierto U de un espacio topológico lineal es convexo si, y sólo si, $U + U = 2U$.

2. Sea E un espacio lineal; un conjunto $U \subset E$ se llama *simétrico*, cuando $x \in U$ implica $-x \in U$. Sea $\mathfrak{B}$ la familia de todos los subconjuntos simétricos convexos del espacio E coincidentes con su núcleo (véase el § 2). Demuéstrese la validez de las afirmaciones siguientes.

(a) La familia $\mathfrak{B}$ es un sistema determinante de vecindades del cero respecto a una topología separable y localmente convexa del espacio E (esta topología se llama *convexa nuclear*).

(b) La topología convexa nuclear es la más fuerte de las topologías localmente convexas compatibles con las operaciones lineales en E.

(c) Toda funcional lineal sobre E es continua respecto a una topología convexa nuclear.

3°. Espacios normados numerables. Una clase de espacios topológicos lineales muy importante para el Análisis resulta ser la clase de los así llamados *espacios normados numerables*. Para poder dar la definición correspondiente necesitamos de un concepto auxiliar.

Supongamos que en un espacio lineal E se han introducido dos normas $\|\cdot\|_1$ y $\|\cdot\|_2$. Se llaman *compatibles*, cuando toda sucesión $\{x_n\}$ de E, fundamental respecto a cada una de estas normas y convergente a un límite $x \in E$ respecto a una de ellas, converge al mismo límite x respecto a la segunda norma.

Se dice que la norma $\|\cdot\|_1$ *es no más débil* que $\|\cdot\|_2$, cuando existe una constante $c > 0$ tal que $\|x\|_1 \geqslant c\|x\|_2$ para todo $x \in E$.

Si la primera norma es no más débil que la segunda y ésta, no más débil que la primera, estas dos normas se llaman *equivalentes*. Dos normas se llaman comparables, cuando una de ellas es no más débil que la otra.

DEFINICION. Un espacio *normado numerable* es un espacio lineal E provisto de un sistema numerable de normas $\|\cdot\|_n$ compatibles entre sí. Todo espacio normado numerable se convierte en un espacio topológico lineal, si se toma por sistema determinante de vecindades del cero la totalidad de conjuntos $U_{r,\varepsilon}$, cada uno de los cuales está definido por un número r y un número positivo ε y consta de todos los elementos $x \in E$ que verifican las condiciones

$$\|x\|_1 < \varepsilon, \ \ldots, \ \|x\|_r < \varepsilon.$$

Proponemos al lector comprobar que este sistema de vecindades del cero induce, efectivamente, en E una topología respecto a la cual resultan continuas las operaciones de adición de elementos y de multiplicación de éstos por números.

Observemos que todo espacio normado numerable verifica el primer axioma de numerabilidad, ya que el sistema de vecindades del cero $U_{r,\varepsilon}$ puede ser sustituido (sin que varíe la topología) por un subsistema numerable, en el que ε toma solamente los valores $1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots$. Por eso la topología de este espacio se puede describir en términos de convergencia de sucesiones. Es más, la topología de un espacio normado numerable se puede definir por medio de una métrica, por ejemplo, por ésta:

$$\rho(x, y) = \sum_{n=1}^{\infty} 2^{-n} \frac{\|x-y\|_n}{1+\|x-y\|_n}, \quad x, y \in E.$$

Proponemos al lector comprobar que la función $\rho(x, y)$ verifica todos los axiomas de distancia y es invariante respecto a las traslaciones (esto es, $\rho(x+z, y+z)=\rho(x, y)$; $x, y, z \in E$) y que la topología inducida por ella coincide con la inicial. De esta forma podemos hablar de la complitud de un espacio normado numerable, entendiéndose por ello la complitud respecto a la métrica introducida más arriba. Observemos, además, que la sucesión $[x_k]$ resulta fundamental respecto a la métrica ρ si, y sólo si, es fundamental respecto a cada una de las normas $\|\cdot\|_n$, y que converge (en la métrica ρ) al elemento $x \in E$ si, y sólo si, converge a x respecto a cada una de las normas $\|\cdot\|_n$. En otras palabras, la complitud de un espacio normado numerable significa que toda sucesión de él, fundamental respecto a cada una de las normas $\|\cdot\|_n$, converge.

Ejemplos. 1. Un ejemplo importante de un espacio normado numerable es el considerado anteriormente espacio $K_{[a, b]}$ de funciones indefinidamente diferenciables sobre un segmento, si admitimos que la norma $\|\cdot\|_m$ de este espacio se define mediante la fórmula

$$\|f\|_m = \sup_{\substack{a \leqslant t \leqslant b \\ 0 \leqslant k \leqslant m}} |f^{(k)}(t)|.$$

Es evidente que todas estas normas son compatibles entre sí y que definen en $K_{[a, b]}$ la misma topología que hemos descrito anteriormente.

2. Sea S el espacio de todas las funciones indefinidamente diferenciables sobre la recta, que tienden, junto con todas sus derivadas, en el infinito a cero más rápido que cualquier potencia de $\frac{1}{|t|}$ (esto es, que verifican la condición $t^k f^{(q)}(t) \to 0$ para $|t| \to \infty$ cualesquiera que sean los números fijos k y q). Definamos en este espacio un sistema numerable de normas tomando

$$\|f\|_m = \sup_{k, q \leqslant m} |t^k f^{(q)}(t)|, \quad m = 0, 1, 2, \ldots$$

Es fácil comprobar que estas normas son compatibles entre sí. Por consiguiente, S es un espacio normado numerable.

3. Un caso particular importante de los espacios normados numerables son los así llamados espacios numerables de Hilbert. Sea H un espacio lineal en el que se ha introducido un sistema numerable de productos escalares $(\varphi, \psi)_n$ y supongamos que las normas $\|\varphi\|_n = \sqrt{(\varphi, \varphi)_n}$, correspondientes a estos productos escalares, son compatibles entre sí. Si un espacio de esta índole es completo, se llama espacio *numerable de Hilbert*.

4. Como ejemplo concreto de un espacio numerable de Hilbert, puede servir el siguiente espacio. Sea Φ el conjunto de

sucesiones numéricas $[x_n]$ tales que para todo entero $k \geqslant 0$ la serie

$$\sum_{n=1}^{\infty} n^k x_n^2$$

converge. Introduzcamos en este espacio un sistema numerable de normas, tomando

$$\| x \|_k = \sqrt{\sum_{n=1}^{\infty} n^k x_n^2}.$$

Es fácil ver que estas normas son compatibles entre sí y que Φ es completo en el sentido señalado anteriormente. Está claro que cada una de las normas $|\cdot|_k$ puede ser definida mediante el producto escalar

$$(x, y)_k = \sum_{n=1}^{\infty} n^k x_n y_n,$$

es decir, que Φ es un espacio numerable de Hilbert. Se llama *espacio de sucesiones rápidamente decrecientes.*

Si E es un espacio normado numerable, se puede aceptar que las normas $\|\cdot\|_k$, definidas en él, verifican la condición

$$\| x \|_k \leqslant \| x \|_l \text{ para } k < l, \tag{1}$$

ya que, de lo contrario, podríamos sustituir las normas $\| x \|_k$ por las normas

$$\| x \|_k' = \sup (\| x \|_1, \| x \|_2, \ldots, \| x \|_k),$$

que definen en E la misma topología que define el sistema inicial de normas. Completando el espacio E según cada una de estas normas $\|\cdot\|_k$, obtendremos un sistema de espacios normados completos E_k. Además, de la relación (1) y de la compatibilidad de las normas, se deduce que tienen lugar las inclusiones naturales

$$E_k \supset E_l \text{ para } k < l.$$

De esta forma, a todo espacio normado numerable E se puede poner en correspondencia una cadena decreciente de espacios normados completos

$$E_1 \supset E_2 \supset \ldots \supset E_k \supset \ldots; \quad \bigcap_{k=1}^{\infty} E_k \supset E.$$

Se puede demostrar que el espacio E es completo cuando, y sólo cuando, $E = \bigcap_{k=1}^{\infty} E_k$ (¡demuéstrese esto!). Por ejemplo cuando el espacio $k_{[a, b]}$ de funciones indefinidamente diferenciables sobre el segmento $[a, b]$ es la intersección de los espacios normados completos $D^n (n=0, 1, 2, \ldots)$, donde el espacio D^n está compuesto por funciones, que tienen derivadas continuas de orden n inclusive, y la norma en él se define mediante la fórmula

$$\|f\|_n = \sup_{\substack{a \leq t \leq b \\ 0 \leq k \leq n}} |f^{(k)}(t)|.$$

En la década de los años 30, cuando fue construida, principalmente en los trabajos de Banach, la teoría de los espacios lineales normados, existía la opinión de que esta clase de espacios es lo suficientemente amplia para abastecer todas las necesidades concretas del Análisis. Más tarde se vio, sin embargo, que esto no es así. Resultó que en una serie de cuestiones desempeñan un papel importante espacios como el espacio de funciones indefinidamente diferenciables, el espacio R^∞ de todas las sucesiones numéricas, así como otros espacios, donde la topología natural no se puede definir mediante ninguna norma De esta forma, los espacios lineales topológicos, pero no normados, no son necesariamente «exotiquez» o «patalogía». Al contrario, algunos de estos espacios representan una generalización del espacio euclídeo de dimensión finita no menos natural e importante que, digamos, el espacio de Hilbert.

CAPITULO IV

FUNCIONALES LINEALES Y OPERADORES LINEALES

§ 1. FUNCIONALES LINEALES CONTINUAS

1°. Funcionales lineales continuas sobre espacios topológicos lineales. En el § 1 del cap. III hemos considerado ya funcionales lineales definidas sobre un espacio lineal. En el caso de funcionales definidas sobre un espacio topológico lineal, representan interés principal las funcionales lineales continuas; como de costumbre, una funcional f, definida sobre un espacio E, se llama *continua*, cuando para todo $\varepsilon > 0$ y todo $x_0 \in E$ existe una vecindad U del elemento x_0 tal que

$$|f(x) - f(x_0)| < \varepsilon \text{ para } x \in U. \tag{1}$$

Si E es un espacio topológico lineal de dimensión finita, toda funcional lineal sobre E es automáticamente continua. En el caso general, la linealidad de una funcional no implica su continuidad.

La siguiente proposición, aun siendo casi evidente, resulta esencial para lo sucesivo.

Si una funcional lineal f es continua en algún punto $x \in E$, es continua en todo E.

En efecto, sea y un punto arbitrario de E y sea $\varepsilon > 0$. Escojamos la vecindad U del punto x de manera que se cumpla la condición (1). Entonces el conjunto

$$V = U + (y - x)$$

será la vecindad deseada del punto y, ya que para $z \in V$, tenemos $z + x - y \in U$ y, por consiguiente,

$$|f(z) - f(y)| = |f(z - y + x) - f(x)| < \varepsilon.$$

De esta forma es suficiente comprobar la continuidad de una funcional solamente en un punto, por ejemplo, en el punto 0. Si E es un espacio con el primer axioma de numerabilidad, la continuidad de una funcional lineal sobre E se puede enunciar en términos de sucesiones, es decir, de la siguiente forma: una funcional f se llama continua en el punto $x \in E$, cuando de $x_n \longrightarrow x$ se deduce que $f(x_n) \longrightarrow f(x)$. Dejamos a cargo del lector la demostración de la equivalencia (en caso de que se verifique el primer axioma de numerabilidad) de esta definición con la dada anteriormente.

TEOREMA 1. *Para que una funcional lineal f sea continua sobre E, es necesario y suficiente que exista en E una vecindad del cero donde la funcional f sea acotada.*

DEMOSTRACION. Si la funcional f es continua en el punto 0, para todo $\varepsilon > 0$ existe una vecindad del cero donde

$$|f(x)| < \varepsilon.$$

Viceversa, sea U una vecindad del cero tal que

$$|f(x)| \leqslant C \text{ para } x \in U$$

y sea $\varepsilon > 0$. Entonces, $\frac{\varepsilon}{C} U$ es aquella vecindad del cero, donde $|f(x)| < \varepsilon$. Con esto queda demostrada la continuidad de f en el punto 0 y, por consiguiente, en todo el espacio.

EJERCICIO. Sea E un espacio topológico lineal; demuéstrese la validez de las siguientes proposiciones.

(a) Una funcional lineal f sobre E es continua cuando, y sólo cuando, existen un conjunto abierto $U \subset E$ y un número t tales que $t \overline{\in} f(U)$ (aquí $f(U)$ es el conjunto de los valores que toma f sobre U).

b) Una funcional lineal f sobre E es continua cuando, y sólo cuando, su subespacio nulo $\{x: f(x) = 0\}$ es cerrado en E.

c) Si toda funcional lineal sobre E es continua, la topología de E coincide con la topología convexa nuclear (véase el ejercicio en la pág. 180).

d) Si E es de dimensión infinita y normable, existe en él una funcional lineal discontinua (empléese la existencia en E de una base de Hamel; véase el ejercicio en la pág. 134).

(e) Supongamos que en E existe un sistema determinante de vecindades del cero y que la potencia de este sistema no sobrepasa la dimensión algebraica del espacio E (esto es, la potencia de una base de Hamel en E; véase el ejercicio en la pág. 134). Entonces, existe sobre E una funcional lineal no continua.

2°. Relación entre la continuidad de una funcional lineal y su acotación sobre conjuntos acotados. Recordemos que un conjunto M, situado en un espacio topológico lineal, ha sido llamado acotado, cuando para cualquier vecindad U del cero existe un número n

tal que $M \subset nU$. Demostremos el siguiente teorema que establece la relación existente entre la continuidad de una funcional lineal y su comportamiento sobre conjuntos acotados.

TEOREMA 2. *Para que una funcional lineal f sea continua sobre E, es necesario y, en caso de que E verifique el primer axioma de numerabilidad, también suficiente que sea acotada sobre todo conjunto acotado.*

DEMOSTRACION. NECESIDAD. Una funcional continua es acotada sobre cierta vecindad U del cero, de manera que

$$|f(x)| \leqslant C \text{ para } x \in U.$$

Si M es un conjunto acotado, tenemos que $M \subset nU$ para un n determinado y, por consiguiente,

$$|f(x)| \leqslant Cn \text{ sobre } M.$$

SUFICIENCIA. Sea $U_1 \supset U_2 \supset \ldots \supset U_n \supset \ldots$ un sistema numerable determinante de vecindades del cero en E. Si la funcional f no es continua, no está acotada en cada una de estas vecindades. Por eso, en cada una de las vecindades U_k se puede indicar un punto x_k en el cual $|f(x)| > k$. La sucesión $\{x_k\}$ tiende a cero y representa, por consiguiente, un conjunto acotado; la funcional f no está acotada sobre este conjunto. Es decir, si la funcional f no es continua en un espacio E que verifica el primer axioma de numerabilidad, existe en E un conjunto acotado sobre el cual la funcional no está acotada. El teorema queda demostrado.

Una funcional lineal acotada sobre todo conjunto acotado se llamará funcional lineal *acotada*. La acotación de una funcional lineal no implica, en general, su continuidad.

3°. **Funcionales lineales continuas sobre espacios normados.** Veamos más detalladamente el caso particular importante cuando E es un espacio normado. Como todo espacio normado verifica el primer axioma de numerabilidad, la continuidad de una funcional lineal en él equivale a su acotación. Pero en un espacio normado un conjunto es acotado si, y sólo si, está contenido en una bola con centro en el origen de coordenadas. Por eso la acotación de una funcional en un espacio normado significa que la funcional está acotada en cada bola. Debido a la linealidad, esta última condición equivale a que la funcional está acotada sobre la bola unitaria

$$\|x\| \leqslant 1$$

del espacio E.

Sea f una funcional lineal acotada (=continua) en un espacio normado E. El número

$$\|f\| = \sup_{\|x\| \leqslant 1} |f(x)|,$$

esto es, la cota superior mínima de los valores que toma $|f(x)|$ sobre la bola unitaria del espacio E, se llamará *norma de la funcional lineal* f. Señalemos las siguientes propiedades evidentes de $\|f\|$:

1) $$\|f\| = \sup_{x \neq 0} \frac{|f(x)|}{\|x\|}, \tag{2}$$

esto sigue directamente de que para todo $x \neq 0$

$$\frac{|f(x)|}{\|x\|} = \left| f\left(\frac{x}{\|x\|}\right) \right|.$$

2) Para cualquier $x \in E$

$$|f(x)| \leqslant \|f\| \cdot \|x\|. \tag{3}$$

En efecto, si $x \neq 0$, el elemento $\frac{x}{\|x\|}$ pertenece a la bola unitaria, es decir, por definición de la norma de una funcional

$$\left| f\left(\frac{x}{\|x\|}\right) \right| = \frac{|f(x)|}{\|x\|} \leqslant \|f\|,$$

de donde se deduce (3). En cambio, si $x = 0$, en los miembros de derecha y de izquierda de (3) figura 0.

En lo sucesivo sólo consideraremos funcionales lineales continuas; omitiremos, para abreviar, la palabra «continua».

Veamos algunos ejemplos de funcionales lineales en espacios normados.

1. Sea R^n el espacio euclídeo de n dimensiones y sea a un vector fijo de él. Entonces, el producto escalar

$$f(x) = (x, a),$$

donde x recorre todo R^n, representa, evidentemente, una funcional lineal sobre R^n. Debido a la desigualdad de Cauchy—Buniakovski, tenemos

$$|f(x)| = |(x, a)| \leqslant \|x\| \cdot \|a\|; \tag{4}$$

por consiguiente, esta funcional es acotada y, por lo tanto, continua sobre R^n. De la desigualdad (4) tenemos que

$$\frac{|f(x)|}{\|x\|} \leqslant \|a\|.$$

Puesto que el miembro derecho de esta desigualdad no depende de x, tendremos

$$\sup \frac{|f(x)|}{\|x\|} \leqslant \|a\|,$$

es decir, $\|f\| \leqslant \|a\|$. Pero tomando $x=a$, obtenemos

$$|f(a)| = (a, a) = \|a\|^2, \text{ es decir, } \frac{|f(a)|}{\|a\|} = \|a\|.$$

Por eso

$$\|f\| = \|a\|.$$

2. La integral

$$I(x) = \int_a^b x(t)\,dt,$$

donde $x(t)$ es una función continua sobre $[a, b]$, representa una funcional lineal en el espacio $C_{[a,b]}$. Esta funcional es acotada y su norma es igual a $b-a$. En efecto,

$$|I(x)| = \left| \int_a^b x(t)\,dt \right| \leqslant \max |x(t)|(b-a) = \|x\|(b-a)$$

y para $x \equiv \text{const}$ se alcanza la igualdad.

3. Consideremos un ejemplo más general. Sea $y_0(t)$ una función fija continua sobre $[a, b]$. Pongamos para toda función $x(t) \in C_{[a,b]}$

$$F(x) = \int_a^b x(t)\,y_0(t)\,dt. \tag{5}$$

Esta funcional es lineal. Además, es acotada, ya que

$$|F(x)| = \left| \int_a^b x(t)\,y_0(t)\,dt \right| \leqslant \|x\| \int_a^b |y_0(t)|\,dt. \tag{6}$$

De manera que la funcional (5) es lineal y acotada; luego, es continua. De (6) se deduce que

$$\|F\| \leqslant \int_a^b |y_0(t)|\,dt.$$

(Demuéstrese que de hecho tiene lugar la igualdad exacta).

4. Consideremos en el espacio $C_{[a,b]}$ la funcional lineal

$$\delta_{t_0}(x) = x(t_0)$$

ya mencionada en el punto 5 del § 1 del cap. III. Su valor en la función $x(t)$ se define como el valor de $x(t)$ en el punto dado t_0. Está claro que

$$|x(t_0)| \leqslant \|x\|$$

y que para $x \equiv \text{const}$ tiene lugar la igualdad. De aquí se deduce inmediatamente que la norma de la funcional δ_{t_0} es igual a 1.

5. En todo espacio euclídeo X, al igual que en R^n, se puede definir una funcional lineal, escogiendo un elemento fijo $a \in X$ y tomando para cualquier $x \in X$

$$F(x) = (x, a).$$

Al igual que en el caso de R^n, es fácil comprobar que

$$\|F\| = \|a\|.$$

Al concepto de la norma de una funcional lineal en un espacio normado se puede dar la siguiente interpretación geométrica clara. Hemos visto anteriormente (cap. III, § 1) que a toda funcional lineal se puede poner en correspondencia un hiperplano L dado por la ecuación

$$f(x) = 1.$$

Busquemos la distancia d de este hiperplano al punto 0. Por definición $d = \inf\limits_{f(x)=1} \|x\|$. Puesto que siempre

$$|f(x)| \leqslant \|f\| \cdot \|x\|,$$

de $f(x) = 1$ se deduce que $\|x\| \geqslant \frac{1}{\|f\|}$ para todo $x \in L$, es decir, que $d \geqslant \frac{1}{\|f\|}$. Por otro lado, de acuerdo con la definición de la norma de f, para cualquier $\varepsilon > 0$ existe un elemento x_ε que verifica la condición $f(x_\varepsilon) = 1$ y tal que

$$1 > (\|f\| - \varepsilon)\|x_\varepsilon\|;$$

por eso

$$d = \inf_{f(x)=1} \|x\| < \frac{1}{\|f\| - \varepsilon}.$$

Como $\varepsilon > 0$ es arbitrario, obtenemos

$$d = \frac{1}{\|f\|},$$

es decir, la norma de la funcional lineal $f(x)$ es el valor inverso de la distancia desde el hiperplano $f(x) = 1$ hasta el punto 0.

4°. Teorema de Hahn — Banach en un espacio normado. En el § 2 del cap. III hemos demostrado el teorema de Hahn — Banach, según el cual toda funcional lineal $f(x_0)$, que está definida sobre un subespacio L de un espacio lineal E y que verifica la condición

$$|f_0(x)| \leqslant p(x), \tag{7}$$

donde p es una funcional convexa fija sobre E, se puede prolongar a todo el E conservando la condición. En el caso de funcionales lineales acotadas en espacios normados, este teorema se puede enunciar de la siguiente manera:

Sean E un espacio normado real; L, un subespacio suyo y f_0, una funcional lineal acotada sobre L. Esta funcional lineal se puede prolongar a una funcional lineal f, definida sobre todo el espacio E, sin aumentar la norma, esto es, de manera que

$$\| f_0 \|_{\text{sobre } L} = \| f \|_{\text{sobre } E}.$$

En efecto, sea

$$\| f_0 \|_{\text{sobre } L} = k.$$

Está claro que $k\|x\|$ es una funcional convexa. Tomándola igual a p y aplicando el teorema de Hahn—Banach, demostrado en el § 2 del cap. III, obtendremos el resultado necesario.

En la forma que acabamos de dar, el teorema de Hahn—Banach admite la siguiente interpretación geométrica. La ecuación

$$f_0(x) = 1 \tag{8}$$

define en el subespacio L un hiperplano que se encuentra a la distancia $\frac{1}{\| f_0 \|}$ del cero. La posibilidad de prolongar la funcional f_0, sin incrementar su norma, hasta una funcional definida en todo E, significa que este hiperplano puede completarse hasta un hiperplano en todo E y de manera que la distancia hasta el cero de este hiperplano mayor sea la misma que la del hiperplano (8).

Aplicando la variante compleja del teorema de Hahn—Banach (teorema 4a del § 2 del cap. III), es fácil demostrar la validez de la siguiente proposición:

Sean E un espacio normado complejo y f_0 una funcional lineal acotada, definida sobre un subespacio $L \subset E$. Entonces existe una funcional lineal acotada f, definida en todo E, que verifica las condiciones:

$$f(x) = f_0(x), \quad x \in L,$$
$$\| f \|_{\text{sobre } E} = \| f_0 \|_{\text{sobre } L}.$$

5°. Funcionales lineales en espacios normados numerables. Sea E un espacio normado numerable con las normas $\|\cdot\|_k$

($k=1, 2, \ldots$); sin perder generalidad, se puede considerar que para todo $x \in E$

$$\|x\|_1 \leqslant \|x\|_2 \leqslant \ldots \leqslant \|x\|_n \leqslant \ldots \tag{9}$$

Sea f una funcional lineal continua sobre E; entonces, existe en E una vecindad U del cero en la cual la funcional f está acotada. De acuerdo con las desigualdades (9) y la definición de la topología en un espacio normado numerable, existen un número natural k y un $\varepsilon > 0$ tales que la bola $S_{k,\varepsilon} = \{x : \|x\|_k < \varepsilon\}$ está contenida en la vecindad U; entonces, la funcional f está acotada sobre esta bola y, por consiguiente, es acotada y continua respecto a la norma $\|\cdot\|_k$, esto es, existe un $C > 0$ tal que

$$f|(x)| \leqslant C \|x\|_k, \quad x \in E.$$

Por otro lado, es evidente que una funcional lineal acotada respecto a una de las normas $\|\cdot\|_n$ es continua sobre E. Por consiguiente, si E_n^* es el conjunto de todas las funcionales lineales sobre E, continuas respecto a la norma $\|\cdot\|_n$, y si E^* es el conjunto de todas las funcionales lineales continuas sobre E, tenemos

$$E^* = \bigcup_{n=1}^{\infty} E_n^*. \tag{10}$$

Además, de la condición (9) se deduce que

$$E_1^* \subset E_2^* \subset \ldots \subset E_n^* \subset \ldots$$

Siendo f una funcional lineal continua sobre E, esto es, $f \in E^*$, su *orden* se define como el menor de los números n tales que $f \in E_n^*$; en virtud de la igualdad (10), toda funcional lineal continua sobre E tiene un orden finito.

6°. Existencia de un número suficiente de funcionales lineales continuas. Si el espacio topológico lineal E no se sujeta a limitaciones adicionales, puede ocurrir que no exista sobre él ninguna funcional lineal continua diferente del cero idéntico. No obstante, puede señalarse una clase amplia de espacios para los cuales existe un número suficientemente grande de funcionales lineales continuas. Aceptemos la siguiente terminología. Diremos que sobre el espacio E existe un *número suficientemente grande* de funcionales lineales continuas, cuando para cualesquiera $x_1 \neq x_2$ de E existe una funcional lineal continua f, definida sobre E, tal que $f(x_1) \neq f(x_2)$. Esta condición equivale, evidentemente, a que para todo $x_0 \neq 0$ exista una funcional f tal que $f(x_0) \neq 0$.

Para cada espacio normado E existe un número suficientemente grande de funcionales lineales continuas.

En efecto, sea x_0 un elemento no nulo de E. Definamos para los elementos de tipo λx_0 la funcional f_0, tomando

$f_0(\lambda x_0)=\lambda$, y prolonguemos después esta funcional (valiéndonos, del teorema de Hahn—Banach) hasta una funcional continua, definida sobre todo el espacio E. Obtendremos una funcional f tal que $f(x_0)=1\neq 0$.

Si E es un espacio normado numerable con las normas $\|\cdot\|_n (n=1, 2, \ldots)$ y si $x_0 \in E$ y $x_0 \neq 0$, podemos construir, repitiendo el razonamiento anterior, una funcional lineal f, definida sobre E y continua respecto a la norma $\|\cdot\|_1$, tal que $f(x_0)\neq 0$; puesto que esta funcional será, evidentemente, continua sobre E, *para todo espacio normado numerable existe también un número suficientemente grande de funcionales lineales continuas.*

Finalmente, si E es un espacio real localmente convexo y separable, para cualquier elemento no nulo $x_0 \in E$ existe una vecindad U del cero convexa y simétrica ($U=-U$) tal que $x_0 \overline{\in} U$; sea p_U la funcional de Minkowski para la vecindad U, esto es,

$$p_U(x)=\inf\left\{t:\ t>0,\ \frac{x}{t}\in U\right\}.$$

De lo demostrado en el § 2 del cap. III se deduce que p_U es una funcional convexa sobre E finita y simétrica (es decir, tal que $p_U(-x)=p_U(x)$) y que, además,

$$p_U(x)<1, \quad x\in U,$$
$$p_U(x_0)\geqslant 1.$$

Consideremos en E el subespacio lineal unidimensional $L_0 = \{\lambda x_0\}$ y definamos en él una funcional lineal f_0, tomando $f_0(\lambda x_0)=\lambda$. Está claro que $|f_0(x)|\leqslant p_U(x)$ para $x\in L_0$ y que $f_0(x_0)=1$. De acuerdo con el teorema de Hahn — Banach, existe una prolongación f de la funcional f_0, definida en todo el E, que verifica la condición $|f(x)|\leqslant p_U(x)$ para todos los $x\in E$. En particular, para $x\in U$, tenemos $|f(x)|\leqslant p_U(x)<1$, de manera que la funcional f es acotada sobre U y, por consiguiente, continua sobre E.

Como $f(x_0)=1$, hemos demostrado que *para todo espacio real localmente convexo y separable existe un número suficientemente grande de funcionales lineales continuas* (una proposición análoga es válida también para un espacio complejo). En realidad, este hecho es el que determina en lo fundamental la importancia para el Análisis de la clase de espacios localmente convexos.

§ 2. ESPACIO DUAL

1°. Definición de espacio dual. Para funcionales lineales se puede definir las operaciones de adición y multiplicación por números. Sean f_1 y f_2 dos funcionales lineales definidas sobre un espacio topológico lineal E. La suma de ellas f_1+f_2 es la funcional lineal

$$f(x)=f_1(x)+f_2(x), \quad x\in E.$$

El producto αf_1 de la funcional lineal f_1 por el número α es la funcional

$$f(x)=\alpha f_1(x), \quad x\in E.$$

Está claro que la suma f_1+f_2 y el producto αf_1 representan funcionales lineales. Además, de la continuidad de las funcionales f_1 y f_2 se deduce que f_1+f_2 y αf_1 son también funcionales continuas sobre E.

Es fácil ver que las operaciones de adición y multiplicación de funcionales por números así definidas satisfacen todos los axiomas de un espacio lineal. En otras palabras, el conjunto de todas las funcionales lineales continuas, definidas sobre un espacio topológico lineal E, forma un espacio lineal. Se llama espacio *dual* a E y se denota mediante E^*.

Además de las operaciones lineales, en el espacio dual E^* se pueden introducir diferentes topologías. Consideremos primero el caso más sencillo cuando el espacio inicial E es normado.

2°. **Espacio dual a un espacio normado.** Para las funcionales lineales continuas, definidas sobre un espacio normado, hemos introducido el concepto de norma tomando

$$\|f\|=\sup_{x\neq 0}\frac{|f(x)|}{\|x\|}.$$

Esta norma verifica todas las condiciones contenidas en la definición de un espacio normado. En efecto,

1) $\|f\|>0$ para cualquier funcional lineal no nula f;

2) $\|\alpha f\|=|\alpha|\cdot\|f\|$;

3) $$\|f_1+f_2\|=\sup_{x\neq 0}\frac{|f_1(x)+f_2(x)|}{\|x\|}\leqslant\sup_{x\neq 0}\frac{|f_1(x)|}{\|x\|}+\sup_{x\neq 0}\frac{|f_2(x)|}{\|x\|}=\|f_1\|+\|f_2\|.$$

De esta forma, el espacio E^*, dual a un espacio normado, puede ser provisto de una estructura natural de espacio normado. La topología en E^*, correspondiente a la norma introducida, se llama *topología fuerte en* E^*. Si es deseable subrayar que E^* se considera como un espacio normado, en lugar de E^* escribiremos $(E^*, \|\cdot\|)$.

Establezcamos la siguiente propiedad importante del espacio dual a un espacio normado.

TEOREMA 1. *El espacio dual* $(E^*, \|\cdot\|)$ *es completo.*

DEMOSTRACION. Sea $\{f_n\}$ una sucesión fundamental de funcionales lineales. Entonces, para cada $\varepsilon > 0$ existe un N tal que $\|f_n - f_m\| \leqslant \varepsilon$ para todo $n, m \geqslant N$. De aquí obtenemos para cualquier $x \in E$

$$|f_n(x) - f_m(x)| \leqslant \|f_n - f_m\| \cdot \|x\| < \varepsilon \|x\|,$$

es decir, para cualquier $x \in E$ la sucesión numérica $\{f_n(x)\}$ converge. Pongamos

$$f(x) = \lim_{n \to \infty} f_n(x).$$

Comprobemos que f representa una funcional lineal continua. Comprobemos primero la linealidad:

$$f(\alpha x + \beta y) = \lim_{n \to \infty} f_n(\alpha x + \beta y) =$$
$$= \lim_{n \to \infty} [\alpha f_n(x) + \beta f_n(y)] = \alpha f(x) + \beta f(y).$$

Escojamos ahora N de manera que $\|f_n - f_{n+p}\| < 1$ para todo $n \geqslant N$ y todo $p \geqslant 0$. Entonces, $\|f_{n+p}\| < \|f_n\| + 1$ para todo $p \geqslant 0$. Por consiguiente,

$$|f_{n+p}(x)| \leqslant (\|f_n\| + 1) \cdot \|x\|.$$

Pasando al límite para $p \longrightarrow \infty$, obtenemos

$$\lim_{p \to \infty} |f_{n+p}(x)| = |f(x)| \leqslant (\|f_n\| + 1)\|x\|,$$

es decir, la funcional $f(x)$ es acotada. Demostremos ahora que la funcional f es el límite de la sucesión $f_1, f_2, \ldots, f_n, \ldots$ Sea $\varepsilon > 0$. Escojamos n tan grande que $\|f_n - f_{n+p}\| < \frac{\varepsilon}{3}$ para todo $p \geqslant 0$. Por definición de la norma existe un elemento $x_{n,\varepsilon}$ tal que

$$\|f_n - f\| \leqslant \frac{|f_n(x_{n,\varepsilon}) - f(x_{n,\varepsilon})|}{\|x_{n,\varepsilon}\|} + \frac{\varepsilon}{3} =$$
$$= \left| f_n\left(\frac{x_{n,\varepsilon}}{\|x_{n,\varepsilon}\|}\right) - f\left(\frac{x_{n,\varepsilon}}{\|x_{n,\varepsilon}\|}\right) \right| + \frac{\varepsilon}{3}.$$

Entonces,
$$\|f_{n+p} - f\| \leqslant \left| f_{n+p}\left(\frac{x_{n,\varepsilon}}{\|x_{n,\varepsilon}\|}\right) - f\left(\frac{x_{n,\varepsilon}}{\|x_{n,\varepsilon}\|}\right) \right| + \frac{2\varepsilon}{3}.$$

Pero el primer sumando del miembro derecho tiende a cero para $p \longrightarrow \infty$. Por consiguiente, para todo p suficientemente grande se cumplirá la desigualdad

$$\|f_{n+p} - f\| < \varepsilon.$$

El teorema queda demostrado.

Subrayemos una vez más que este teorema es válido independientemente de si es completo o no el espacio inicial.

Observación. *Si el espacio normado E no es completo y $\bar{E}$ es su completación, los espacios E^* y $(\bar{E})^*$ son isomorfos.* En efecto, podemos aceptar que E está sumergido en $\bar{E}$ como un subconjunto siempre denso; por eso, toda funcional $f \in E^*$ lineal continua sobre E tiene la única prolongación por continuidad $\bar{f}$ de E a todo el $\bar{E}$. Está claro que $\bar{f} \in (\bar{E})^*$, que $\|\bar{f}\| = \|f\|$ y que toda funcional de $(\bar{E})^*$ es una prolongación de alguna funcional de E^* (a saber, de su restricción sobre E). Por consiguiente, la aplicación $f \longrightarrow \bar{f}$ representa una aplicación isomórfica del espacio E^* a todo el espacio $(\bar{E})^*$.

3°. Ejemplos de espacios duales. 1. Sea E el espacio lineal de n dimensiones (real o complejo). Escojamos en E una base $e_1, \ldots, e_n$; entonces, todo vector $x \in E$ se representa de manera única en la forma $x = \sum_{i=1}^{n} x_i e_i$. Si f es una funcional lineal sobre E, está claro que

$$f(x) = \sum_{i=1}^{n} f(e_i) x_i; \tag{1}$$

por consiguiente, una funcional lineal se determina unívocamente por los valores que toma en los vectores de la base $e_1, \ldots, e_n$ con la particularidad de que estos valores pueden escogerse arbitrariamente. Definamos las funcionales lineales $f_1, \ldots, f_n$, tomando

$$f_j(e_i) = \begin{cases} 1, & \text{cuando } i = j, \\ 0, & \text{cuando } i \neq j. \end{cases}$$

Es evidente que estas funcionales son linealmente independientes. Está claro que

$$f_j(x) = x_j;$$

por eso, la fórmula (1) se puede escribir en la forma

$$f(x) = \sum_{i=1}^{n} f(e_i) f_i(x).$$

Por consiguiente, las funcionales $f_1, \ldots, f_n$ forman una base en el espacio E^*, es decir, E^* es un espacio lineal de n dimensiones. La base $f_1, \ldots, f_n$ de E^* se llama *dual* respecto a la base

$e_1, \ldots, e_n$ de E. Diferentes normas en el espacio E inducen distintas normas en E^*. Señalemos algunos ejemplos de pares de normas correspondientes en E y en E^*; en las fórmulas (a), (b), (c) y (d) $x_1, \ldots, x_n$ son las coordenadas del vector $x \in E$ en la base $e_1, \ldots, e_n$, mientras que $f^1, \ldots, f^n$ son las coordenadas de la funcional $f \in E^*$ en la base dual $f_1, \ldots, f_n$.

(a) $\|x\| = \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}}, \quad \|f\| = \left(\sum_{i=1}^{n} |f^i|^2\right)^{\frac{1}{2}};$

(b) $\|x\| = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}, \quad \|f\| = \left(\sum_{i=1}^{n} |f^i|^q\right)^{\frac{1}{q}}; \quad \frac{1}{p} + \frac{1}{q} = 1;$

(c) $\|x\| = \sup_{1 \leqslant i \leqslant n} |x_i|, \quad \|f\| = \sum_{i=1}^{n} |f^i|;$

(d) $\|x\| = \sum_{i=1}^{n} |x_i|, \quad \|f\| = \sup_{1 \leqslant i \leqslant n} |f^i|.$

2. Consideremos el espacio c_0 de las sucesiones $x = (x_1, \ldots, x_n, \ldots)$ convergentes a cero con la norma $\|x\| = \sup_n |x_n|$ y demostremos que *el espacio* $(c_0^*, \|\cdot\|)$ *dual a él es isomorfo al espacio* l_1 *de todas las sucesiones absolutamente sumables* $f = \{f_1, \ldots, f_n, \ldots\}$ *con la norma* $\|f\| = \sum_{n=1}^{\infty} |f_n|$. Cualquier sucesión $f \in l_1$ define en el espacio c_0 una funcional lineal acotada $\bar{f}$ mediante la fórmula

$$\bar{f}(x) = \sum_{n=1}^{\infty} f_n x_n; \tag{2}$$

está claro que $|\bar{f}(x)| \leqslant \|x\| \sum_{n=1}^{\infty} |f_n|$, de manera que $\|\bar{f}\| \leqslant \|f\|$. Consideremos en c_0 los vectores

$$e_1 = (1, 0, \ldots, 0, \ldots),$$
$$\cdots\cdots\cdots\cdots$$
$$e_n = (0, 0, \ldots 1, 0, \ldots),$$
$$\cdots\cdots\cdots\cdots$$

y pongamos $x^{(N)} = \sum_{n=1}^{N} \frac{f_n}{|f_n|} e_n$ $\left(\text{si } f_n = 0, \text{ aceptamos que } \frac{f_n}{|f_n|} = 0\right)$.

Entonces, $x^{(N)} \in c_0$, $\|x^{(N)}\| \leqslant 1$ y

$$\tilde{f}(x^{(N)}) = \sum_{n=1}^{N} \frac{f_n}{|f_n|} \tilde{f}(e_n) = \sum_{n=1}^{N} |f_n|,$$

de manera que $\lim\limits_{N \to \infty} \tilde{f}(x^{(N)}) = \sum\limits_{n=1}^{\infty} |f_n| = \|f\|$. Por consiguiente, $\|\tilde{f}\| \geqslant \geqslant \|f\|$; comparando esto con la desigualdad de sentido opuesto, obtenida anteriormente, llegamos a la conclusión de que $\|\tilde{f}\| = \|f\|$. Hemos construido, de esta forma, una aplicación lineal isométrica $f \longrightarrow \tilde{f}$ del espacio l_1 en el espacio c_0^*; queda por comprobar que la imagen del espacio l_1 mediante esta aplicación coincide con todo el c_0^*, es decir, que toda funcional $\tilde{f} \in c_0^*$ se puede representar en la forma (2), donde $f = \{f_n\} \in l_1$. Para todo $x = \{x_n\} \in c_0$, tenemos $x = \sum\limits_{n=1}^{\infty} x_n e_n$; la serie que figura en el miembro derecho converge en c_0 al elemento x, ya que $\left\| x - \sum\limits_{n=1}^{N} x_n e_n \right\| = \sup\limits_{n > N} |x_n| \longrightarrow 0$ para $N \longrightarrow \infty$. Como la funcional $\tilde{f} \in c_0^*$ es continua, tenemos $\tilde{f}(x) = \sum\limits_{n=1}^{\infty} x_n \tilde{f}(e_n)$; por eso, es suficiente comprobar que $\sum\limits_{n=1}^{\infty} |\tilde{f}(e_n)| < \infty$. Tomando $x^{(N)} = \sum\limits_{n=1}^{N} \frac{\tilde{f}(e_n)}{|\tilde{f}(e_n)|} e_n$ y observando que $x^{(N)} \in c_0$, $\|x^{(N)}\| \leqslant 1$, tenemos

$$\sum_{n=1}^{N} \|\tilde{f}(e_n)| = \sum_{n=1}^{N} \frac{\tilde{f}(e_n)}{|\tilde{f}(e_n)|} \tilde{f}(e_n) = \tilde{f}(x^{(N)}) \leqslant \|\tilde{f}\|,$$

de donde, debido a la arbitrariedad de N, sacamos la conclusión de que $\sum\limits_{n=1}^{\infty} |\tilde{f}(e_n)| < \infty$.

3. No es difícil demostrar que el espacio l_1^* dual al espacio l_1 es isomorfo al espacio m compuesto por todas las sucesiones acotadas $x = \{x_n\}$ con la norma $\|x\| = \sup\limits_n |x_n|$.

4. Sean $p > 1$ y l_p el espacio de todas las sucesiones $x = \{x_n\}$ tales que

$$\|x\| = \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{\frac{1}{p}} < \infty;$$

se puede demostrar que el espacio l_p^* dual a él es isomorfo al espacio l_q, $\frac{1}{p}+\frac{1}{q}=1$. La forma general de una funcional lineal continua sobre l_p es:

$$f(x)=\sum_{n=1}^{\infty} f_n x_n; \quad x=\{x_n\}\in l_p, \quad f=\{f_n\}\in l_q.$$

La demostración se basa en la desigualdad de Hölder.

Pongamos en claro la estructura del espacio dual al de Hilbert.

TEOREMA 2. *Sea H un espacio real de Hilbert. Para toda funcional lineal f continua sobre H existe el único elemento $x_0 \in H$ tal que*

$$f(x)=(x, x_0), \quad x\in H; \tag{3}$$

además, resulta que $\|f\|=\|x_0\|$. Viceversa, si $x_0\in H$, la fórmula (3) *define una funcional lineal continua f tal que $\|f\|=$ $=\|x_0\|$. Por consiguiente, los espacios H^* y H son isomorfos.*

DEMOSTRACION. Es evidente que la fórmula (3) define para todo $x_0\in H$ una funcional lineal sobre H. Como $|f(x)|=|(x, x_0)|\leqslant$ $\leqslant\|x\|\cdot\|x_0\|$, esta funcional es continua y como $f(x_0)=\|x_0\|^2$, tenemos además $\|f\|=\|x_0\|$. Comprobemos que toda funcional lineal continua f sobre H se puede representar en la forma (3). Si $f=0$, tomamos $x_0=0$. Sea ahora $f\neq 0$ y sea $H_0=\{x: f(x)=0\}$ el subespacio nulo de la funcional f; como f es continua, H_0 es un subespacio lineal c e r r a d o de H. En el punto 6 del § 1 del cap. III ha sido demostrado que la codimensión del subespacio de ceros de cualquier funcional lineal es igual a 1. Por eso y tomando en consideración el corolario 3 del teorema 7 del § 4 del cap. III, llegamos a la conclusión de que el complemento ortogonal $H_0^{\perp}$ del espacio H_0 es unidimensional, esto es, existe un elemento y_0 (no nulo) ortogonal a H_0 y tal que todo vector $x\in H$ se puede representar de manera única en la forma $x=y+$ $+\lambda y_0$, donde $y\in H_0$. Podemos aceptar, evidentemente, que $\|y_0\|=1$; pongamos $x_0=f(y_0)\,y_0$. Entonces, para cualquier $x\in H$ tenemos

$$x=y+\lambda y_0, \quad y\in H_0,$$
$$f(x)=\lambda f(y_0),$$
$$(x, x_0)=\lambda(y_0, x_0)=\lambda f(y_0)(y_0, y_0)=\lambda f(y_0).$$

De manera que $f(x)=(x, x_0)$ para todo $x\in H$. Si $f(x)=(x, x_0')$, $x\in H$, tenemos $(x, x_0-x_0')=0$, de donde, tomando $x=x_0-x_0'$, encontramos que x_0-x_0'. El teorema queda demostrado.

Observaciones. 1. Sea E un espacio euclídeo n o c o m p l e t o y sea H el espacio de Hilbert que representa su completación. Como los espacios E^* y H^* son isomorfos (véase la observación de la pág. 196) y H^* es isomorfo a H, es válida la siguiente proposición: *el espacio E^* dual a un espacio euclídeo no completo E es isomorfo a la completación H del espacio E.*

2. El teorema 2 es válido también para un espacio de Hilbert complejo (la demostración es la misma, solamente hay que sustituir $x_0=f(y_0)y_0$ por $x_0=\overline{f(y_0)}\,y_0$). La única diferencia entre el caso complejo y real consiste en que, en el caso complejo, la aplicación del espacio H en el espacio H^*, que pone en correspondencia al elemento $x_0\in H$ la funcional $f(x)=(x, x_0)$, es un isomorfismo lineal *conjugado*, esto es, pone en correspondencia al elemento λx_0 la funcional $\bar{\lambda}f$.

4°. Estructura del espacio dual a un espacio normado numerable. Sea E un espacio normado numerable con las normas

$$\|x\|_1\leqslant\|x\|_2\leqslant\ldots\leqslant\|x\|_n\leqslant\ldots,\quad x\in E.$$

De lo demostrado en el punto 5 del § 1 se deduce que el espacio E^*, dual a E, se representa en forma de la unión

$$E^*=\bigcup_{n=1}^{\infty}E_n^*$$

de una cadena creciente $E_1^*\subset E_2^*\subset\ldots\subset E_n^*\subset\ldots$ de espacios normados E_n^* (en otras palabras, E_n^* es el conjunto de todas las funcionales lineales sobre E continuas respecto a la norma $\|\cdot\|_n$). De esta forma, si se conocen los espacios E_n^*, se conoce también el espacio E^* (en el sentido de que se conocen los elementos que lo componen y las operaciones lineales en él). En cuanto a la topología de E^*, ésta se puede introducir de diferentes modos; algunos de ellos serán examinados en lo sucesivo. Señalemos además que con frecuencia resulta cómodo considerar el espacio E_n^* como el dual a la completación del espacio E respecto a la norma $\|\cdot\|_n$; esto es posible en virtud de la observación que hemos hecho después de la demostración del teorema 1.

Ejemplo. Sea Φ un espacio real de Hilbert numerable compuesto por todas las sucesiones $x=\{x_n\}$ tales que

$$\|x\|_k=\left(\sum_{n=1}^{\infty}n^k x_n^2\right)^{\frac{1}{2}}<\infty,\quad k=1,\ 2,\ \ldots;$$

los productos escalares en Φ vienen dados por

$$(x,\ y)_k=\sum_{n=1}^{\infty}n^k x_n y_n,\quad k=1,\ 2,\ \ldots$$

El espacio Φ con el producto escalar $(\cdot,\ \cdot)_k$ es un espacio euclídeo; sea Φ_k su completación. Es fácil ver que Φ_k se puede identificar con el espacio de Hilbert de todas las sucesiones $x=\{x_n\}$ tales que $\|x\|_k<\infty$. De acuerdo con el teorema 2, el espacio Φ_k^*, dual al espacio Φ_k, es isomorfo al espacio Φ_k; este isomorfismo consiste en que a cada funcional lineal continua $f\in\Phi_k^*$ se pone en correspondencia una sucesión $\bar{f}=\{f_n\}$ tal que

$$\|f\|=\left(\sum_{n=1}^{\infty} n^k |f_n|^2\right)^{\frac{1}{2}}<\infty,$$

$$f(x)=(x,\ \bar{f})_k=\sum_{n=1}^{\infty} n^k x_n f_n,\quad x=\{x_n\}\in\Phi_k,$$

y, viceversa, cada sucesión de este tipo determina un elemento de Φ_k^*. Definamos ahora la funcional $f\in\Phi_k^*$ no mediante la sucesión $\{f_n\}$ sino a partir de la sucesión $\{g_n\}$, donde $g_n=n^k f_n$. Entonces,

$$f(x)=\sum_{n=1}^{\infty} x_n g_n \ \text{ y } \ \|f\|=\left(\sum_{n=1}^{\infty} n^{-k} g_n^2\right)^{\frac{1}{2}}.$$

Por consiguiente, Φ_k^* se puede identificar con el espacio de Hilbert de las sucesiones $\{g_n\}$, que verifican la condición

$$\sum_{n=1}^{\infty} n^{-k} g_n^2<\infty,$$

y con el producto escalar

$$(g^{(1)},\ g^{(2)})=\sum_{n=1}^{\infty} n^{-k} g_n^{(1)} g_n^{(2)}.$$

Como $\Phi^*=\bigcup_{k=1}^{\infty}\Phi_k^*$, Φ^* es el espacio de todas las sucesiones $\{g_n\}$ que satisfacen la condición: existe un número entero positivo k tal que

$$\sum_{n=1}^{\infty} n^{-k} g_n^2<\infty.$$

Además, cada una de estas funcionales toma en el elemento $x=\{x_n\}\in\Phi$ un valor determinado igual a $\sum_{n=1}^{\infty} x_n g_n$. Es decir, si el espacio Φ es la intersección de una cadena decreciente de espacios de Hilbert

$$\Phi=\bigcap\Phi_k,\quad \Phi_1\supset\Phi_2\supset\ldots\supset\Phi_k\supset\ldots,$$

el espacio Φ^* es la unión de una cadena creciente de espacios de Hilbert

$$\Phi^* = \bigcup \Phi_k^*, \quad \Phi_1^* \subset \Phi_2^* \subset \ldots \subset \Phi_k^* \subset \ldots$$

Conviene introducir la denotación $\Phi_k^* = \Phi_{-k}$. Si además denotamos l_2 mediante Φ_0, obtenemos la siguiente cadena infinita en ambos sentidos de espacios de Hilbert

$$\ldots \subset \Phi_k \subset \ldots \subset \Phi_1 \subset \Phi_0 \subset \Phi_{-1} \subset \ldots \subset \Phi_{-k} \subset \ldots,$$

donde $\Phi_k^* = \Phi_{-k}$ para cada $k = 0, \pm 1, \pm 2, \ldots$

5°. Topología en el espacio dual. Siendo E un espacio normado, hemos definido una vecindad del cero en E^* como el conjunto de funcionales que verifican la condición

$$\|f\| < \varepsilon.$$

En otras palabras, se toma por sistema de vecindades del cero en el espacio E^*, dual a un espacio normado, el conjunto de funcionales tales que $|f(x)| < \varepsilon$, cuando x recorre la bola unitaria $\|x\| \leqslant 1$ del espacio E. En el caso en que E no es un espacio normado, sino un espacio topológico lineal, es natural, para definir la topología en E^*, considerar en E en lugar de una bola unitaria un conjunto arbitrario acotado A y definir una vecindad del cero en E^* como el conjunto de funcionales lineales que verifican la condición

$$|f(x)| < \varepsilon \text{ para todos los } x \in A.$$

De esta forma en E^* queda determinado el conjunto de vecindades del cero, cada una de las cuales se define mediante un número positivo ε y un conjunto acotado $A \subset E$. No vamos a comprobar aquí, aunque no es difícil de hacerlo, que el sistema de vecindades, definido de esta manera, satisface efectivamente las condiciones que debe verificar un sistema de conjuntos para que pueda ser tomado por un sistema determinante de vecindades del cero en un espacio topológico lineal.

Está claro que si el espacio E es normado, la topología de E^*, que acabamos de describir, coincide con la que se define en E^* mediante la norma. La topología descrita del espacio, dual a un espacio topológico lineal (a un espacio normado, en particular), se llamará topología *fuerte* en E^* (a diferencia de la topología débil de E^* de la cual hablaremos en el § 3).

Subrayemos que *la topología fuerte del espacio E^* es necesariamente separable y localmente convexa* (independientemente de la topología que tiene E). En efecto, si $f_0 \in E^*$ y $f_0 \neq 0$, existe un elemento $x_0 \in E$ tal que $f_0(x_0) \neq 0$, pongamos $\varepsilon = \frac{1}{2}|f_0(x_0)|$ y

$A=\{x_0\}$; entonces, está claro que $f_0 \overline{\in} U_{\varepsilon, A}$, es decir, que E^* es separable. Para demostrar la convexidad local de la topología fuerte de E^*, es suficiente observar que para cualquier $\varepsilon > 0$ y cualquier conjunto acotado $A \subset E$ la vecindad $U_{\varepsilon, A}$ es un conjunto convexo en E^*. Denotaremos la topología fuerte de E^* mediante el símbolo b; si es preciso señalar que E^* se considera con la topología fuerte, escribiremos a veces (E^*, b) (en lugar de E^*).

6°. Segundo espacio dual. Puesto que las funcionales lineales continuas sobre un espacio topológico lineal E forman por sí mismas un espacio topológico lineal (el espacio (E^*, b) dual a E), se puede hablar del espacio E^{**} de funcionales lineales continuas sobre E^*, es decir, del *segundo espacio dual* a E, etc.

Señalemos que todo elemento x_0 de E define en E^* una funcional lineal. En efecto, tomemos

$$\psi_{x_0}(f) = f(x_0), \tag{4}$$

donde x_0 es un elemento fijo de E, mientras que f recorre todo el E^*. La igualdad (4) pone en correspondencia a cada f un número $\psi_{x_0}(f)$, esto es, define una funcional sobre E^*. Como además

$$\psi_{x_0}(\alpha f_1 + \beta f_2) = \alpha f_1(x_0) + \beta f_2(x_0) = \alpha\psi_{x_0}(f_1) + \beta\psi_{x_0}(f_2),$$

esta funcional es lineal.

Es más, toda funcional de este tipo es continua sobre E^*. En efecto, sea $\varepsilon > 0$ y sea A un conjunto acotado de E que contiene x_0. Consideremos en E^* la vecindad $U(\varepsilon, A)$ del cero. De acuerdo con la definición de $U(\varepsilon, A)$, tenemos

$$|\psi_{x_0}(f)| = |f(x_0)| < \varepsilon \quad \text{para} \quad f \in U(\varepsilon, A)$$

Pero esto significa que la funcional ψ_{x_0} es continua en el punto 0 y, por consiguiente, en todo el espacio E^*.

Hemos obtenido de esta forma una aplicación de todo el espacio E en un subconjunto del espacio E^{**}. Esta aplicación es, evidentemente, lineal. Si sobre E existe un número suficientemente grande de funcionales lineales (por ejemplo, si E es normado o localmente convexo y separable), esta aplicación resulta biunívoca, ya que para cualesquiera dos diferentes x', $x'' \in E$ existe una funcional $f \in E^*$ tal que $f(x') \neq f(x'')$, es decir, $\psi_{x'}$ y $\psi_{x''}$ son diferentes funcionales sobre E^*. Esta aplicación de E en E^{**} se llama *aplicación natural del espacio E en el segundo espacio dual*. Denotémosla con π. Si $\pi(E) = E^{**}$, el espacio E (separable y localmente convexo) se llama *semirreflexivo* En el espacio E^{**} (considerado como dual a (E^*, b)) se puede introducir la topología fuerte que denotaremos con b^*; esta topología induce en el espacio E la topología $\pi^{-1}(b^*)$ (tal que un conjunto $Q \subset E$ se considera abierto si su imagen $\pi(Q)$ es la intersección de $\pi(E)$ con un subconjunto abierto del espacio (E^{**}, b^*)). Se puede demostrar que la topología $\pi^{-1}(b^*)$ es no más débil que la topología inicial del espacio E (es decir, que todo conjunto abierto en esta topología inicial es también abierto en la topología $\pi^{-1}(b^*)$); esto significa que la aplicación π^{-1}, que transforma $\pi(E)$ en E, es continua. Siendo el espacio E semirreflexivo y la aplicación π: $E \longrightarrow E^{**}$ continua, se dice que E es un espacio *reflexivo*. De lo expuesto se deduce: si E es reflexivo, la aplicación natural $\pi: E \longrightarrow E^{**}$ representa un isomorfismo entre los espacios topológicos lineales E y $E^{**} = (E^{**}, b^*)$ (es decir, es biunívoca y bicontinua).

Puesto que podemos ahora considerar todo elemento de E también como un elemento del espacio E^{**}, conviene para los valores de la funcional lineal $f \in E^*$ emplear, en lugar de la denotación $f(x)$, una denotación más simétrica

$$f(x) = (f, x). \tag{5}$$

Siendo fijo $f \in E^*$, podemos considerar (f, x) como una funcional sobre E, y siendo fijo x, como una funcional sobre E^* (y en este caso x aparece ya como un elemento de E^{**}).

Si E es un espacio normado (y, por consiguiente, son también normados los espacios E^*, E^{**}, etc.), *la aplicación natural del espacio E en E^{**} es una isometría.*

En efecto, sea x un elemento de E. Designemos mediante $\|x\|$ su norma en E y mediante $\|x\|_2$ la norma de su imagen en E^{**}. Demostremos que $\|x\| = \|x\|_2$. Sea f un elemento arbitrario de E^*. Entonces,

$$|(f, x)| \leqslant \|f\| \cdot \|x\|, \quad \text{es decir,} \quad \|x\| \geqslant \frac{|(f, x)|}{|f|},$$

y, puesto que el miembro izquierdo de la última desigualdad no depende de f,

$$\|x\| \geqslant \sup \frac{|(f, x)|}{\|f\|} = \|x\|_2.$$

Por otro lado, según el teorema de Hahn—Banach, para todo $x_0 \in E$ existe una funcional lineal f_0 tal que

$$|(f_0, x_0)| = \|f_0\| \cdot \|x_0\| \tag{6}$$

(para construir esta funcional es suficiente tomar $f_0(x) = \alpha$ para los elementos de tipo $x = \alpha x_0$ y extender después esta funcional, conservando su norma, a todo el E). De (6) se desprende que

$$\|x\|_2 = \sup_{f \in E^*} \frac{|(f, x)|}{\|f\|} \geqslant \|x\|,$$

es decir, $\|x\| = \|x\|_2$, que es lo que queríamos demostrar. Por consiguiente, el espacio normado E es isométrico a la variedad lineal (en general, no cerrada) $\pi(E)$ de E^{**}; identificando E con $\pi(E)$, podemos aceptar que $E \tilde{\subset} E^{**}$.

Puesto que para los espacios normados la aplicación natural $\pi: E \longrightarrow E^{**}$ es isométrica, resulta que los conceptos de semirreflexividad y reflexividad coinciden en el caso de espacios normados.

Como el espacio, dual a un espacio normado, es completo, todo espacio normado reflexivo E es completo.

Los espacios euclídeos de dimensión finita y el espacio de Hilbert representan los ejemplos más sencillos de espacios reflexivos (para ellos se tiene incluso que $E = E^*$).

El espacio c_0 de sucesiones convergentes a cero ofrece un ejemplo de un espacio no reflexivo completo. En efecto, como hemos demostrado anteriormente (ejemplo 2 del § 2), el espacio dual a c_0 es el espacio l_1 de todas las series numéricas absolutamente convergentes y el dual de este último coincide con el espacio m de todas las sucesiones acotadas.

El espacio $C_{[a, b]}$ de funciones continuas sobre un segmento $[a, b]$ es también no reflexivo. Sin embargo, no daremos aquí la demostración de esta proposición[1].

Como ejemplo de un espacio reflexivo, que no coincide con su dual, puede servir el espacio l_p para $1 < p \neq 2$ (puesto que $l_p^* = l_q$, donde $\frac{1}{p} + \frac{1}{q} = 1$, tenemos $l_p^{**} = l_q^* = l_p$).

EJERCICIO. Demuéstrese que un subespacio cerrado de un espacio reflexivo es también reflexivo.

§ 3. TOPOLOGIA DEBIL Y CONVERGENCIA DEBIL

1°. Topología débil en un espacio topológico lineal. Consideremos un espacio topológico lineal E y el conjunto de todas las funcionales continuas sobre él. Si $f_1, \ldots, f_n$ es un sistema finito arbitrario de estas funcionales y ε es un número positivo, el conjunto

$$\{x : |f_i(x)| < \varepsilon; \quad i = 1, 2, \ldots, n\} \tag{1}$$

---

[1] Se puede demostrar incluso una proposición más fuerte: no existe *ningún* espacio normado para el cual $C_{[a, b]}$ sea el espacio dual.

es abierto en E y contiene el punto 0, esto es, representa una vecindad del cero. La intersección de dos vecindades de este tipo siempre contiene un conjunto de tipo (1) y, por consiguiente, en E se puede introducir una topología para la cual la totalidad de los conjuntos de tipo (1) constituirá el sistema determinante de vecindades del cero. Ella se denomina *topología débil* del espacio E. En otras palabras, la topología débil en E es la topología más débil de este espacio lineal respecto a la cual son continuas todas las funcionales lineales que son continuas respecto a la topología inicial de este espacio.

Está claro que todo conjunto de E abierto en el sentido de la topología débil es también abierto en la topología inicial del espacio E, pero la recíproca no es, en general, válida (los conjuntos de tipo (1) no forman necesariamente un sistema determinante de vecindades del cero en la topología inicial). De acuerdo con la terminología, aceptada en el § 5 del cap. II, esto significa que la topología débil del espacio E es más débil que la topología inicial. Esto justifica su denominación.

Si en E existe un número suficientemente grande de funcionales lineales continuas (por ejemplo, si E es normado), la topología débil de E verifica el axioma de separabilidad de Hausdorff. Es fácil comprobar asimismo que las operaciones de adición y multiplicación por números, definidas en E, son continuas respecto a la topología débil de este espacio.

2°. Convergencia débil. Incluso en el caso de espacios normados, la topología débil de E puede no satisfacer el primer axioma de numerabilidad. Por consiguiente, esta topología no puede describirse, en general, en términos de sucesiones convergentes. No obstante, la convergencia en E determinada por esta topología representa un concepto importante. Se llama *convergencia débil*. A diferencia de ésta, la convergencia definida por la topología inicial del espacio E (por la norma, si E es normado) se llama *convergencia fuerte*.

Es obvio que el concepto de convergencia débil se puede enunciar de la siguiente manera: la sucesión $\{x_n\}$ de elementos de E se llama *débilmente convergente* a $x_0 \in E$, cuando para cualquier funcional $\varphi(x)$ lineal continua sobre E la sucesión numérica $\{\varphi(x_n)\}$ converge a $\varphi(x_0)$. En efecto, admitiendo, para simplificar, que $x_0=0$, supongamos que $\varphi(x_n) \longrightarrow 0$ para toda $\varphi \in E^*$. Entonces, cualquiera que sea la vecindad débil

$$U=\{x: |\varphi_i(x)| < \varepsilon, \quad i=1,\ 2,\ \ldots,\ k\}$$

del punto 0, existe un N tal que $x_n \in U$ para todo $n \geqslant N$ (para ello es suficiente escoger N_i de manera que $|\varphi_i(x_n)| < \varepsilon$, cuando

$n \geqslant N_i$, y tomar después $N = \max N_i$). Viceversa, si para cada vecindad U del cero existe un N tal que $x_n \in U$ para todo $n \geqslant N$, la condición $\varphi(x_n) \to 0$ para $n \to \infty$ se cumple, evidentemente, para cada funcional fija $\varphi \in E^*$.

Consideremos más detalladamente el concepto de convergencia débil en el caso de espacios normados.

TEOREMA 1. *Si $\{x_n\}$ es una sucesión débilmente convergente en un espacio normado, existe una constante C tal que*

$$\|x_n\| \leqslant C.$$

En otras palabras, *toda sucesión débilmente convergente de un espacio normado es acotada.*

DEMOSTRACION. Siendo la sucesión $\{x_n\}$ no acotada, cualquiera que sea la bola cerrada $S\,[f_0,\ \varepsilon] = \{f : \|f - f_0\| \leqslant \varepsilon\}$ de E^*, el conjunto numérico

$$\{(f,\ x_n)\},$$

donde $n = 1,\ 2,\ \ldots$ y f recorre esta bola, no es acotado. En efecto, si la sucesión $\{x_n\}$ es acotada sobre la bola $S\,[f_0,\ \varepsilon]$ es también acotada sobre la bola $S\,[0,\ \varepsilon] = \{g : \|g\| \leqslant \varepsilon\}$, ya que siendo $g \in S\,[0,\ \varepsilon]$, tenemos $f_0 + g \in S\,[f_0,\ \varepsilon]$ y $(g,\ x_n) = (f_0 + g,\ x_n) - (f_0,\ x_n)$ y los números $(f_0,\ x_n)$ son acotados debido a la convergencia débil de la sucesión $\{x_n\}$. Pero si $|(g,\ x_n)| \leqslant C$ para todo $g \in S\,[0,\ \varepsilon]$, tenemos, debido a que la aplicación natural de E en E^{**} es isométrica,

$$\|x_n\| \leqslant \frac{C}{\varepsilon},$$

esto es, las normas $\|x_n\|$ están acotadas en su conjunto. Por consiguiente, si la sucesión $\{x_n\}$ es no acotada, tampoco será acotada sobre cualquier bola de E^*. Tomemos una bola $B_0 \subset E^*$. Existen un número n_1 y un elemento $f \in B_0$ tales que $|(f,\ x_{n_1})| > 1$. Puesto que $(f,\ x)$ depende continuamente de f, la desigualdad $|(f, x_{n_1})| > 1$ se verificará para todos los f pertenecientes a una bola cerrada $B_1 \subset B_0$. Razonando de la misma manera, encontraremos en B_1 una bola cerrada B_2 y un número n_2 tales que para todo $f \in B_2$ se cumple la desigualdad $|(f,\ x_{n_2})| > 2$, etc.; en general, para cada k encontraremos un número n_k y una bola $B_k \subset B_{k-1}$ tales que

$$|(f,\ x_{n_k})| > k \text{ para } f \in B_k.$$

Se puede admitir, además, que los radios de las bolas B_k tienden a cero cuando $k \to \infty$. Puesto que el espacio E^* es completo, existe un elemento $\bar{f}$ perteneciente a todas las bolas B_k (¡principio de bolas encajadas!). Pero en tal caso

$$|(\tilde{f}, x_{n_k})| > k$$

para todo k y esto contradice a la convergencia débil de la sucesión $\{x_n\}$.

Observación. Al demostrar que la sucesión $\{x_n\}$ es acotada respecto a la norma nos hemos valido solamente de que la sucesión numérica (f, x_n) es acotada para todo $f \in E^*$. Por eso, si la sucesión $\{x_n\}$ de E es tal que la sucesión numérica (f, x_n) es acotada para toda $f \in E^*$, existe una constante C tal que $\|x_n\| \leqslant C$. Este resultado admite la siguiente generalización: *todo subconjunto débilmente acotado* (esto es, acotado en la topología débil) *Q de un espacio normado E es fuertemente acotado (es decir, está contenido en una bola)*. En efecto, supongamos que existe una sucesión $\{x_n\} \subset Q$ tal que $\|x_n\| \longrightarrow \infty$ $(n \longrightarrow \infty)$. Como Q es débilmente acotado, el conjunto $\{x_n\}$ también es débilmente acotado, es decir, es absorbido por cualquier vecindad débil del cero; en particular, para cualquier $f \in E^*$ existe un N tal que $\{x_n\} \subset N\,\{x: |(f, x)| < 1\}$, de donde se sigue que $|(f, x_n)| < N$ para todo n. Pero esto, de acuerdo con la observación hecha anteriormente, contradice a que $\|x_n\| \longrightarrow \infty$. Si tenemos en cuenta que la acotación débil de un conjunto Q significa que cualquier funcional lineal continua sobre él es acotada, llegamos al siguiente resultado importante: *para que un subconjunto Q de un espacio normado sea acotado es necesario y suficiente que cualquier funcional $f \in E^*$ sea acotada sobre Q.*

El teorema que sigue permite frecuentemente determinar la convergencia débil de una u otra sucesión.

TEOREMA 2. *La sucesión $\{x_n\}$ de elementos de un espacio normado E converge débilmente al elemento $x \in E$, si*

1) $\|x_n\|$ están acotadas en su conjunto por una constante M;
2) $f(x_n) \longrightarrow f(x)$ para toda $f \in \Delta$, donde Δ es un conjunto cuya cápsula lineal es siempre densa en E^*.

DEMOSTRACION. Se desprende de las condiciones del teorema y de la definición de una funcional lineal que si φ es una combinación lineal de elementos de Δ, entonces,

$$\varphi(x_n) \longrightarrow \varphi(x).$$

Sea ahora φ un elemento arbitrario de E y sea $\{\varphi_k\}$ una sucesión de combinaciones lineales de elementos de Δ convergente a φ. Demostremos que $\varphi(x_n) \longrightarrow \varphi(x)$. Sea M tal que

$$\|x_n\| \leqslant M \quad (n = 1, 2, \ldots) \text{ y } \|x\| \leqslant M.$$

Estimemos la diferencia $|\varphi(x_n)| - \varphi(x)$. Como $\varphi_k \longrightarrow \varphi$, para cualquier $\varepsilon > 0$ existe un K tal que $\|\varphi - \varphi_k\| < \varepsilon$ para todos los $k \geqslant K$. Por eso,

$$|\varphi(x_n) - \varphi(x)| \leqslant |\varphi(x_n) - \varphi_k(x_n)| + \\ + |\varphi_k(x_n) - \varphi_k(x)| + |\varphi_k(x) - \varphi(x)| \leqslant \varepsilon M + \\ + \varepsilon M + |\varphi_k(x_n) - \varphi_k(x)|.$$

Pero, $\varphi_k(x_n) \longrightarrow \varphi_k(x)$ para $n \longrightarrow \infty$, por hipótesis. Luego, $\varphi(x_n) - \varphi(x) \longrightarrow 0$, cuando $n \longrightarrow \infty$, para cualquier $\varphi \in E^*$.

Ejemplos. Veamos el sentido que tiene el concepto de convergencia débil en algunos espacios concretos.

1. *En el espacio euclídeo de dimensión finita* R^n la convergencia débil coincide con la fuerte. En efecto, sea $e_1, \ldots, e_n$ una base ortonormal de R^n y sea $\{x_k\}$ una sucesión de R^n convergente débilmente al elemento x. Entonces,

$$(x_k, e_i) = x_k^i \longrightarrow (x, e_i), \quad i = 1, 2, \ldots, n,$$

es decir, las primeras coordenadas de los vectores x_k tienden a la primera coordenada del vector límite x, sus segundas coordenadas convergen a la segunda coordenada del vector x, etc. Pero entonces,

$$\rho(x_k, x) = \left(\sum_{i=1}^{n} (x_k^i - x^i)^2\right)^{\frac{1}{2}} \longrightarrow 0,$$

es decir, $\{x_n\}$ converge fuertemente a x. Puesto que la convergencia fuerte implica la débil, queda demostrada la equivalencia en R^n de estas convergencias.

2. *Convergencia débil en* l_2. Para que la sucesión acotada $\{x^{(k)}\}$ converja débilmente a x, es suficiente que se cumplan las condiciones

$$(x^{(k)}, e_i) = x_k^i \longrightarrow x^i = (x, e_i), \quad i = 1, 2, \ldots,$$

donde

$$e_1 = (1, 0, 0, \ldots), \quad e_2 = (0, 1, 0, \ldots), \ldots$$

En efecto, las combinaciones lineales de los elementos e_i son siempre densas en el espacio l_2 (que coincide, como hemos visto, con su dual). Por eso nuestra proposición se desprende del teorema 2.

Por consiguiente, la convergencia débil de una sucesión acotada $\{x^{(k)}\}$ de l_2 significa que la sucesión numérica de las coordenadas x_k^i de estos vectores converge para cada $k = 1, 2, \ldots$ No es difícil ver que la convergencia débil de l_2 no coincide con la fuerte. En efecto, demostremos que la sucesión $e_1, e_2, \ldots, e_n, \ldots$ converge débilmente al 0 en l_2. Toda funcional lineal

f en l_2 puede ser representada como el producto escalar $f(x)=(x, a)$ del vector $x \in l_2$ por un vector fijo $a=(a_1, a_2, \ldots)$. Por eso,

$$f(e_n)=a_n,$$

y como $a_n \longrightarrow 0$ para $n \longrightarrow \infty$ cualquiera que sea $a \in l_2$, obtenemos

$$\lim f(e_n)=0$$

para cada funcional lineal en l_2.

Al mismo tiempo, la sucesión $\{e_n\}$ no converge, en el sentido fuerte, a ningún límite.

EJERCICIOS. 1. Supongamos que la sucesión $\{x_n\}$ de elementos de un espacio de Hilbert H converge débilmente al elemento $x \in H$ de manera que $\|x_n\| \longrightarrow \|x\|$ para $n \longrightarrow \infty$. Demuéstrese que en este caso la sucesión $\{x_n\}$ converge fuertemente a x, es decir $\|x_n - x\| \longrightarrow 0$.

1. Demuéstrese que la proposición del ejercicio 1 sigue siendo válida si se sustituye la condición $\|x_n\| \longrightarrow \|x\|$ por la condición $\|x_n\| \leqslant \|x\|$ para todo n o por la condición $\overline{\lim\limits_{n\to\infty}} \|x_n\| \leqslant \|x\|$.

3. Sea H un espacio (separable) de Hilbert y sea Q un subconjunto suyo acotado. Entonces, la topología en Q inducida por la topología débil del espacio H se puede definir mediante una métrica.

4. Demuéstrese que todo subconjunto cerrado convexo de un espacio de Hilbert es cerrado respecto a la topología débil (en particular, todo subespacio lineal cerrado de un espacio de Hilbert es débilmente cerrado). Dése un ejemplo de un conjunto cerrado de un espacio de Hilbert que no sea débilmente cerrado.

3. *Convergencia débil en el espacio $C_{[a, b]}$ de funciones continuas.* Sea $\{x_n(t)\}$ una sucesión acotada de funciones de $C_{[a, b]}$ convergente débilmente a la función $x(t)$. Entre las funcionales definidas sobre $C_{[a, b]}$ se encuentran, en particular, las funcionales δ_{t_0} cada una de las cuales representa el valor de la función en un punto fijo t_0 (véase el ejemplo 4 del punto 3 del § 1). Para cada una de estas funcionales δ_{t_0} la condición

$$\delta_{t_0} x_n(t) \longrightarrow \delta_{t_0} x(t)$$

significa que

$$x_n(t_0) \longrightarrow x(t_0).$$

Por consiguiente, si la sucesión $\{x_n(t)\}$ converge débilmente, ella

1) es equiacotada, esto es, $|x_n(t)| \leqslant C$ para todo $n=1, 2, \ldots$ y cualquier $a \leqslant t \leqslant b$;

2) converge en cada punto.

Se puede demostrar que el conjunto de estas dos condiciones no es sólo necesario sino también suficiente para la convergencia débil de la sucesión $\{x_n(t)\}$ en $C_{[a, b]}$.

Está claro que esta convergencia no coincide con la convergencia respecto a la norma de $C_{[a, b]}$, es decir, con la convergencia uniforme de las funciones continuas. (Dése un ejemplo correspondiente.)

3°. Topología débil y convergencia débil en el espacio dual. En el punto 4 del parágrafo anterior hemos introducido en el espacio dual E^* la topología que hemos llamado fuerte y que se define de la siguiente manera: como sistema de vecindades del cero se toma la totalidad de los conjuntos de tipo

$$\{f: |f(x)| < \varepsilon,\ x \in A\},$$

donde A es un conjunto acotado cualquiera de E y ε es un número positivo arbitrario. Si ahora consideramos en lugar de conjuntos acotados todos los subconjuntos f i n i t o s $A \subset E$, obtendremos la así llamada *topología débil en el espacio dual* E^*. Como todo conjunto finito $A \subset E$ es acotado (pero no viceversa, en general), está claro que la topología débil del espacio E^* es más débil que la topología fuerte de este espacio. En general, estas dos topologías no coinciden.

La topología débil, introducida en E^*, define en este espacio una convergencia que se llama *convergencia débil de funcionales*. La convergencia débil de funcionales lineales representa un concepto importante que desempeña un papel esencial en diversas cuestiones del Análisis Funcional, en particular en la teoría de las así llamadas funciones generalizadas, de las cuales hablaremos en el parágrafo siguiente.

La convergencia débil de la sucesión $\{\varphi_n\}$ de funcionales lineales significa, evidentemente, la convergencia de esta sucesión en todo elemento fijo de E. En otras palabras, la sucesión $\{\varphi_n\}$ se llama débilmente convergente a $\varphi \in E^*$, cuando para cada $x \in E$ se cumple la relación

$$\varphi_n(x) \longrightarrow \varphi(x).$$

Sea E (y, por consiguiente, E^*) un espacio de Banach. Tiene lugar el siguiente teorema análogo al teorema 1.

TEOREMA 1*. *Si $\{f_n\}$ es una sucesión débilmente convergente de funcionales lineales sobre un espacio de Banach, existe un número constante C tal que*

$$\|f_n\| \leqslant C, \quad n = 1, 2, \ldots$$

En otras palabras, toda sucesion débilmente convergente de elementos del espacio, dual a un espacio de Banach, es acotada respecto a la norma.

La demostración de este teorema no difiere en nada de la demostracion del teorema 1.

El siguiente teorema es completamente análogo al teorema 2.

TEOREMA 2*. *Una sucesión de funcionales lineales* $\{\varphi_n\}$ *de* E^* *converge débilmente a* $\varphi \in E^*$, *cuando*
1) *esta sucesión es acotada*, esto es

$$\|\varphi_n\| \leqslant C, \quad n=1, 2, \ldots;$$

2) *la relación* $(\varphi_n, x) \longrightarrow (\varphi, x)$ *se cumple para todos los* x, *pertenecientes a un conjunto tal que las combinaciones lineales de sus elementos son siempre densas en* E.

La demostración es la misma que para el teorema 2.

Veamos un e j e m p l o. Sea E el espacio $C_{[a, b]}$ de funciones continuas y sea[1)]

$$\varphi(x) = x(0),$$

es decir, sea φ la δ-función (véase el § 1, ejemplo 4). Sea, además, $\{\varphi_n(t)\}$ una sucesión de funciones continuas que verifican las siguientes condiciones:

1) $\varphi_n(t) = 0$ para $|t| > \frac{1}{n}$, $\varphi_n(t) \geqslant 0$ para $|t| \leqslant \frac{1}{n}$;

2) $\int_a^b \varphi_n(t)\,dt = 1$.

Entonces, para cualquier función $x(t)$ continua sobre $[a, b]$, tenemos, empleando el teorema del valor medio,

$$\int_a^b \varphi_n(t)\,x(t)\,dt = \int_{-\frac{1}{n}}^{\frac{1}{n}} \varphi_n(t)\,x(t)\,dt \longrightarrow x(0) \quad \text{para } n \longrightarrow \infty.$$

La expresión

$$\int_a^b \varphi_n(t)\,x(t)\,dt$$

representa una funcional lineal sobre $C_{[a, b]}$. De esta forma la δ-función se puede representar como límite, en el sentido de convergencia débil de funcionales lineales sobre $C_{[a, b]}$, de una sucesión de funciones «corrientes».

Observación. El espacio E^* de las funcionales lineales sobre un espacio E se puede considerar desde dos puntos de vista: como espacio dual al espacio inicial E o como espacio principal E^*, relacionando con él su espacio dual E^{**}. De acuerdo

[1)] Consideramos que $0 \in [a, b]$. Se podría, claro está, tomar en lugar del punto $t=0$ otro punto cualquiera.

con esto, la topología débil puede introducirse en E^* de dos modos: o bien como en el espacio de funcionales, definiendo las vecindades en E^* mediante todos los sistemas finitos de elementos de E, o bien como en el espacio principal mediante el espacio E^{**}. En el caso de un espacio reflexivo esto, por supuesto, es lo mismo. En cambio, si E no es reflexivo, éstas serán dos topologías diferentes en E^*. Para evitar la confusión que puede surgir aquí, la topología débil definida en el espacio principal (esto es, la topología en E^*, definida mediante E^{**}) se llamará *topología débil*, mientras que la topología débil del espacio de las funcionales (esto es, la topología en E^*, definida mediante E) se llamará *topología* $*$-débil. Es evidente que la topología $*$-débil en E^* es más débil que la topología débil del espacio E^* (es decir, la topología débil tiene no menos conjuntos abiertos que la topología $*$-débil).

4°. Topología $*$-débil en conjuntos acotados. En diferentes aplicaciones del concepto de convergencia débil de funcionales lineales desempeña un papel importante el siguiente teorema.

TEOREMA 3. *Si E es un espacio normado lineal separable, cualquier sucesión acotada de funcionales lineales continuas sobre E contiene una subsucesión débilmente convergente.*

DEMOSTRACION. Escojamos en E un conjunto numerable siempre denso $(x_1, x_2, \ldots, x_n, \ldots)$. Si $\{\varphi_n\}$ es una sucesión acotada (respecto a la norma) de funcionales lineales sobre E, la sucesión numérica

$$\varphi_1(x_1), \varphi_2(x_1), \ldots, \varphi_n(x_1), \ldots$$

es acotada. Por eso, se puede extraer de $\{\varphi_n\}$ una subsucesión

$$\varphi_1^{(1)}, \varphi_2^{(1)}, \ldots, \varphi_n^{(1)}, \ldots,$$

de manera que la sucesión numérica

$$\varphi_1^{(1)}(x_1), \varphi_2^{(1)}(x_1), \ldots, \varphi_n^{(1)}(x_1), \ldots$$

converja. Asimismo se puede extraer de la subsucesión $\{\varphi_n^{(1)}\}$ una subsucesión

$$\varphi_1^{(2)}, \varphi_2^{(2)}, \ldots, \varphi_n^{(2)}, \ldots,$$

de manera que converja la sucesión

$$\varphi_1^{(2)}(x_2), \varphi_2^{(2)}(x_2), \ldots, \varphi_n^{(2)}(x_2), \ldots$$

Continuando este proceso, obtendremos un sistema de sucesiones

$$\begin{array}{l} \varphi_1^{(1)}, \varphi_2^{(1)}, \ldots, \varphi_n^{(1)}, \ldots, \\ \varphi_1^{(2)}, \varphi_2^{(2)}, \ldots, \varphi_n^{(2)}, \ldots, \\ \cdots\cdots\cdots\cdots \end{array}$$

(donde cada una es subsucesion de la anterior), tal que $\{\varphi_n^{(k)}\}$ converge en los puntos $x_1, x_2, \ldots, x_k$. Entonces, tomando la subsucesión «diagonal»

$$\varphi_1^{(1)}; \varphi_2^{(2)}, \ldots, \varphi_n^{(n)}, \ldots,$$

obtendremos una sucesión de funcionales lineales tal que

$$\varphi_1^{(1)}(x_n), \varphi_2^{(2)}(x_n), \ldots$$

converge para todos los n. Pero esto significa (en virtud del teorema 2*) que la sucesión $\varphi_1^{(1)}(x), \varphi_2^{(2)}(x), \ldots$ converge también para todo $x \in E$. El teorema queda demostrado.

Este teorema, junto con el teorema 1*, indica que en el espacio E^*, dual a un espacio separable de Banach, los subconjuntos acotados, y sólo ellos, son relativamente numerables compactos en la topología $*$-débil. Probemos que, de hecho, tiene lugar aquí la compacidad y no sólo la compacidad numerable.

Demostremos, ante todo, el siguiente teorema.

TEOREMA 4. *La topología inducida en la bola cerrada unitaria S del espacio E^*, dual a un espacio normado separable E, por la topología $*$-débil de este espacio se puede definir mediante la métrica*

$$\rho(f, g) = \sum 2^{-n} |(f-g, x_n)|,$$

donde $\{x_n\}$ es un conjunto fijo numerable y siempre denso en la bola unitaria del espacio E.

DEMOSTRACION. Está claro, que la función $\rho(f, g)$ tiene todas las propiedades de la distancia; además, es invariante respecto a las traslaciones:

$$\rho(f+h, g+h) = \rho(f, g).$$

Por eso, basta comprobar que a) cualquier «bola»

$$Q_\varepsilon = \{f : \rho(f, 0) < \varepsilon\}$$

contiene una intersección de S con alguna vecindad débil del cero en E^* y que b) toda vecindad débil del cero en E^* contiene una intersección de S con alguna Q_ε.

Escojamos N de manera que $2^{-N} < \frac{\varepsilon}{2}$ y consideremos la vecindad débil del cero

$$V = V_{x_1, \ldots x_N; \varepsilon/2} = \left\{f : |(f, x_k)| < \frac{\varepsilon}{2}, k = 1, 2, \ldots, N\right\}.$$

Entonces, si $f \in S \cap V$, tenemos

$$\rho(f, 0) = \sum_{n=1}^{N} 2^{-n} |(f, x_n)| + \sum_{n=N+1}^{\infty} 2^{-n} |(f, x_n)| \leqslant$$

$$\leqslant \frac{\varepsilon}{2} \sum_{n=1}^{N} 2^{-n} + \sum_{n=N+1}^{\infty} 2^{-n} < \varepsilon,$$

es decir, $S \cap V \subset Q_\varepsilon$. Con esto queda demostrada la proposición a). Demostremos la proposición b). Sea

$$U = U_{y_1, \ldots, y_m; \delta} = \{f: |(f, y_k)| < \delta, \ k = 1, 2. \ldots, m\}$$

una vecindad $*$-débil del cero en E^*. Podemos admitir que $\|y_k\| \leqslant 1$, $k = 1, 2, \ldots, m$; puesto que el conjunto $\{x_n\}$ es siempre denso en S, existen unos números $n_1, \ldots, n_m$ tales que $\|y_k - x_{n_k}\| < \frac{\delta}{2}$, $k = 1, 2, \ldots, m$. Sea $N = \max(n_1, \ldots, n_m)$ y sea $\varepsilon = 2^{-(N+1)}\delta$. Entonces, para $f \in S \cap Q_\varepsilon$ de la desigualdad

$$\sum_{n=1}^{\infty} 2^{-n} |(f, x_n)| < \varepsilon$$

obtenemos que $|(f, x_n)| < 2^n \varepsilon$; en particular,

$$|(f, x_{n_k})| < 2^{n_k} \varepsilon \leqslant 2^N \varepsilon = \frac{\delta}{2}.$$

Por consiguiente, para todo $k = 1, 2, \ldots, m$ tenemos

$$|(f, y_k)| \leqslant |(f, x_{n_k})| + |(f, y_k - x_{n_k})| < \frac{\delta}{2} + \|f\| \cdot \|y_k - x_{n_k}\| < \delta.$$

Es decir, $S \subset Q_\varepsilon \subset U$. El teorema queda demostrado. Está claro, que este resultado se extiende automáticamente a cualquier bola y, por consiguiente, a cualquier subconjunto acotado $M \subset E^*$.

Hemos demostrado (teorema 3) que de toda sucesión acotada de E^* se puede extraer una subsucesion $*$-débilmente convergente. En otras palabras, en el espacio E^*, dual a un espacio normado lineal separable y provisto de la topología $*$-débil, todo subconjunto acotado M es relativamente numerable compacto. Pero, de acuerdo con el último teorema, cada conjunto de este tipo es un espacio topológico metrizable y en el caso de espacios métricos la compacidad y la compacidad numerable coinciden. Por consiguiente, obtenemos el siguiente resultado.

TEOREMA 3*. *Todo conjunto acotado M del espacio E^*, dual a un espacio de Banach separable, es relativamente compacto en el sentido de la topología $*$-débil del espacio E^*.*

Si E es un espacio normado lineal separable, toda bola cerrada del espacio (E^*, b) es cerrada en la topología $*$-débil del espacio E^*.

Como una traslación en el espacio E^* transforma toda colección de conjuntos cerrados (en la topología $*$-débil) en sí misma, basta demostrar que en la topología $*$-débil es cerrada toda bola de tipo $S_c = \{f: \|f\| \leqslant c\}$. Sea $f_0 \bar{\in} S_c$. Según la definición de la norma de una funcional, existe un vector $x \in E$ tal que $\|x\| = 1$, $f_0(x) = \alpha > c$. Entonces, el conjunto $u_f = \left\{f: f(x) > \frac{\alpha + c}{2}\right\}$ es una vecindad $*$-débil de la funcional f_0 que no contiene ningún elemento de la bola S_c; por consiguiente, la bola S_c es cerrada en la topología $*$-débil.

De la proposición demostrada y del teorema 3* se deduce el siguiente teorema.

TEOREMA 5. *Toda bola cerrada del espacio, dual a un espacio normado separable, es compacta en la topología $*$-débil.*

El teorema 3* representa un caso particular del siguiente resultado general: *si E es un espacio topológico lineal localmente convexo, todo subconjunto acotado de E^* es relativamente compacto en el sentido de la topología $*$-débil.*

No daremos aquí la demostración de esta proposición general.

§ 4. FUNCIONES GENERALIZADAS

1°. Ampliación del concepto de función. En diferentes cuestiones del Análisis resulta necesario interpretar el término «función» con diferente grado de generalidad. A veces se consideran funciones continuas, en otros casos es preciso suponer que se trata de funciones diferenciables una o varias veces, etc. Sin embargo, en muchos casos el concepto clásico de función resulta insuficiente, aun cuando sea interpretado en el sentido más general, esto es, como una regla cualquiera que a todo valor de x del campo de definición de esta función pone en correspondencia un número $y = f(x)$. He aquí dos ejemplos importantes.

1) Es cómodo determinar la distribución de masas a lo largo de una recta mediante la densidad de esta distribución. Sin embargo, si la recta tiene puntos que llevan masas positivas, está claro que la densidad de esta distribución no se puede describir de antemano con ninguna función «corriente».

2) Aplicando el aparato del Análisis Matemático a unos u otros problemas, tropezamos con situaciones, cuando no se pueden efectuar unas u otras operaciones del Análisis; por ejemplo, una función que no tenga derivada (en varios o incluso en todos los pun-

tos) no se puede derivar, si por derivada se entiende una función «corriente». Claro está que las dificultades de este orden se podrían evitar limitándose a considerar solamente funciones, digamos, analíticas. Sin embargo, tal restricción del conjunto de funciones admisibles no es, en muchos casos, deseable.

Por suerte, resulta, sin embargo, que estas dificultades y otras semejantes pueden ser superadas con no menos éxito no restringiendo sino ampliando sustancialmente el concepto de función, introduciendo las así llamadas funciones generalizadas. Como base para introducir las definiciones correspondientes, nos servirá el concepto de espacio dual, considerado anteriormente.

Subrayamos una vez más, que la introducción de funciones generalizadas se debió no al deseo de ampliar lo más posible los conceptos del Análisis sino a problemas absolutamente concretos. De hecho, en la Física estos conceptos se empleaban ya desde hace mucho tiempo, en todo caso antes de que atrajeron la atención seria de los matemáticos. Antes de pasar a las definiciones exactas, expongamos la idea principal.

Sea f una función fija definida sobre la recta e integrable en cada intervalo finito y sea φ una función continua que se anula fuera de un intervalo finito (tales funciones se llamarán en lo sucesivo *terminales*). A cada función φ de este tipo se puede poner en correspondencia mediante la función fija f el número

$$(f, \varphi,) = \int_{-\infty}^{\infty} f(x)\,\varphi(x)\,dx \tag{1}$$

(debido a la terminalidad de $\varphi(x)$, la integral se toma, de hecho, respecto a un intervalo finito). En otras palabras, la función f se puede considerar como una funcional (lineal, debido a las propiedades principales de la integral) definida sobre un espacio de funciones terminales. Sin embargo, las funcionales de tipo (1) no son las únicas que se pueden definir en este espacio; por ejemplo, haciendo corresponder a cada función φ su valor en el punto $x=0$, obtendremos una funcional lineal que no se puede representar en la forma (1). De manera que las funciones $f(x)$ se incluyen de un modo natural en un conjunto más amplio, el conjunto de todas las funcionales lineales sobre funciones terminales.

El conjunto de funciones φ se puede escoger de diversas maneras; podrían tomarse, por ejemplo, todas las funciones terminales conjuntas. Sin embargo, como veremos más adelante, conviene sujetar las funciones admisibles φ no sólo a las condiciones de continuidad terminalidad sino también a unas condiciones suficientemente rígidas de diferenciabilidad.

2°. Espacio de funciones básicas. Pasemos ahora a las definiciones precisas. Consideremos sobre la recta el conjunto K de todas las funciones terminales φ que tienen derivadas continuas de todos los órdenes[1]. Las funciones, pertenecientes a K, constituyen un espacio lineal (con las operaciones habituales de adición de funciones y multiplicación de éstas por números). En este espacio no se puede introducir una norma que sea conveniente desde el punto de vista de la teoría que se expone a continuación; sin embargo, resulta natural definir en él del siguiente modo el concepto de convergencia.

La sucesión $\{\varphi_n\}$ de elementos de K se llama convergente a la función $\varphi \in K$, cuando: 1) existe un intervalo fuera del cual se anulan todas las φ_n; 2) la sucesión $\{\varphi_n^{(k)}\}$ de derivadas de orden[2] k ($k = 1, 2, \ldots$) converge uniformemente en este intervalo a $\varphi^{(k)}$.

El espacio lineal K con la convergencia que hemos definido en él se llamará espacio básico y sus elementos, funciones básicas.

No es difícil describir la topología de K que induce la convergencia definida en K. Esta topología es generada por un sistema de vecindades del cero tal que cada vecindad se determina por un sistema finito $\gamma_0, \ldots, \gamma_m$ de funciones continuas positivas y consta de todas aquellas funciones de K que verifican para todo x las desigualdades

$$|\varphi(x)| < \gamma_0(x), \ldots, |\varphi^{(m)}(x)| < \gamma_m(x).$$

Dejamos a cargo del lector la demostración de que esta topología genera efectivamente la convergencia en K descrita anteriormente. Señalemos, que en K existen también otras topologías que generan esta convergencia.

EJERCICIO. Sea K_m el subespacio del espacio K compuesto por todas las funciones $\varphi \in K$ que son iguales a 0 fuera del segmento $[-m, m]$. En el espacio K_m se puede definir una estructura de espacio normado numerable, si tomamos

$$\|\varphi\|_n = \sup_{\substack{0 \leqslant k \leqslant n \\ |x| \leqslant m}} |\varphi^{(k)}(x)|, \quad n = 0, 1, 2, \ldots$$

Compruébese que la topología (convergencia de sucesiones, respectivamente) del espacio K_m generada por este sistema de normas coincide con la topología (convergencia, respectivamente) inducida en K_m por la topología (convergencia) del espacio K descrita anteriormente. Está claro que $K_1 \subset K_2 \subset \ldots$ $\ldots \subset K_m \subset \ldots$ y que $K = \bigcup_{m=1}^{\infty} K_m$. Demuéstrese que un conjunto $Q \subset K$ es acotado respecto a la topología definida en K, cuando, y sólo cuando, existe un m tal que Q es un subconjunto acotado del espacio normado numerable K_m. Sea T una funcional lineal sobre el espacio K; demuéstrese que

[1] El intervalo, fuera del cual la función φ es igual a 0, puede ser diferente para distintas $\varphi \in K$.

[2] Por derivada de orden cero se comprende, como de costumbre, la propia función.

las cuatro siguientes condiciones son equivalentes: (a) la funcional T es continua respecto a la topología del espacio K; (b) la funcional T es acotada en todo conjunto acotado $Q \subset K$; (c) si $\varphi_n \in K$ y $\varphi_n \longrightarrow 0$ (en el caso de la convergencia de sucesiones definida en K), $T(\varphi_n) \longrightarrow 0$; (d) para todo m la restricción T_m de la funcional T al subespacio $K_m \subset K$ es una funcional continua sobre K_m.

3°. Funciones generalizadas.

DEFINICION 1. Se llama *función generalizada* (definida en la recta $-\infty < x < \infty$) a toda funcional lineal continua $T(\varphi)$ sobre el espacio básico K. La continuidad de la funcional se entiende aquí en el sentido de que $T(\varphi_n) \rightarrow T(\varphi)$ cuando la sucesión φ_n converge a φ en el espacio básico K.

Observemos, ante todo, que toda función $f(x)$ integrable en un intervalo finito cualquiera genera una función generalizada. En efecto, la expresión

$$T_f(\varphi) = \int_{-\infty}^{\infty} f(x)\,\varphi(x)\,dx \tag{2}$$

es una funcional lineal continua en K. Estas funciones generalizadas se llamarán en lo sucesivo *regulares*, mientras que todas las demás, esto es, las que no pueden representarse en la forma (2), se llamarán *singulares*.

Indiquemos algunos ejemplos de funciones generalizadas singulares.

1. «δ-función»:

$$T(\varphi) = \varphi(0).$$

Ella es una funcional lineal en K, es decir, de acuerdo a la terminología convenida más arriba, una función generalizada. Esta funcional se representa generalmente en la forma

$$\int_{-\infty}^{\infty} \delta(x)\,\varphi(x)\,dx, \tag{3}$$

donde por $\delta(x)$ se entiende una «función» que es igual a cero para todo $x \neq 0$ y se convierte en el infinito en el punto $x=0$ de manera que

$$\int_{-\infty}^{\infty} \delta(x)\,dx = 1.$$

En el § 1 hemos considerado ya la δ-función como una funcional en el espacio de todas las funciones continuas definidas en un segmento. Veremos, sin embargo, que la consideración de la δ-función como una funcional sobre K tiene determinadas ventajas.

2. «δ-función desplazada». Sea

$$T(\varphi)=\varphi(a).$$

De acuerdo con la denotación (3), es natural representar esta funcional en la forma

$$\int_{-\infty}^{\infty}\delta(x-a)\,\varphi(x)\,dx. \tag{4}$$

3. «Derivada de la δ-función». A cada $\varphi\in K$ se pone en correspondencia el número $-\varphi'(0)$. Más adelante explicaremos por qué es natural considerar esta funcional como derivada de la funcional señalada en el ejemplo 1.

4. Consideremos la función $\frac{1}{x}$. Ella no es integrable en ningún intervalo que contenga el punto cero. Sin embargo, para cada $\varphi\in K$ la integral

$$\int_{-\infty}^{\infty}\varphi(x)\,\frac{1}{x}\,dx$$

existe y es finita en el sentido del valor principal. En efecto,

$$\int_{-\infty}^{\infty}\varphi(x)\,\frac{1}{x}\,dx=\int_{a}^{b}\varphi(x)\,\frac{1}{x}\,dx=\int_{a}^{b}\frac{\varphi(x)-\varphi(0)}{x}\,dx+\int_{a}^{b}\frac{\varphi(0)}{x}\,dx.$$

Aquí (a, b) es el intervalo fuera del cual φ se anula. En la primera de estas integrales bajo el signo de la integral aparece una función acotada, mientras que la segunda integral, comprendida en el sentido del valor principal, es finita. De esta forma, $\frac{1}{x}$ es una funcional sobre K, es decir, una función generalizada. Se puede demostrar que ninguna de las funciones generalizadas señaladas en los ejemplos 1, 2, 3 y 4 es regular (es decir, no puede representarse en la forma (2) cualquiera que sea la función f localmente integrable).

4°. **Operaciones con funciones generalizadas.** El conjunto de funciones generalizadas es el espacio lineal dual a K. Por consiguiente, en este conjunto están definidas las operaciones de adición y multiplicación por números. Es evidente, además, que para las funciones generalizadas regulares, es decir, para las funciones «corrientes» sobre la recta, su adición como funciones generalizadas (esto es, como funcionales lineales) coincide con la operación habitual de adición de funciones. Lo mismo se refiere a la multiplicación por números.

Introduzcamos en el espacio de funciones generalizadas la operación de paso al límite. Diremos que la sucesión de funciones generalizadas $\{f_n\}$ converge a f, cuando para cada $\varphi \in K$ se cumple la relación

$$(f_n, \varphi) \longrightarrow (f, \varphi).$$

En otras palabras, definimos la convergencia de funciones generalizadas como la convergencia $*$-débil de funcionales lineales sobre K. El espacio de funciones generalizadas provisto de esta convergencia se denotará mediante K^*.

Si α es una función indefinidamente diferenciable, es natural definir el producto de α por la función generalizada f mediante la fórmula

$$(\alpha f, \varphi) = (f, \alpha\varphi)$$

(la expresión del miembro derecho tiene sentido, ya que $\alpha\varphi \in K$). Todas estas operaciones, adición y multiplicación por números y funciones indefinidamente diferenciables, son continuas.

No introducimos el concepto de producto de dos funciones generalizadas. Se puede demostrar que es imposible definir este producto si se quiere que esta operación sea continua y que, además, coincida para las funciones generalizadas regulares con la multiplicación corriente de funciones.

Definamos ahora la operación de diferenciación de funciones generalizadas y examinemos sus propiedades.

Sea primero T una funcional de K definida mediante una función f diferenciable (en el sentido corriente):

$$T(\varphi) = \int_{-\infty}^{\infty} f(x)\,\varphi(x)\,dx.$$

Es natural definir su derivada como la funcional $\frac{dT}{dx}$ dada mediante la fórmula

$$\frac{dT}{dx}(\varphi) = \int_{-\infty}^{\infty} f'(x)\,\varphi(x)\,dx.$$

Empleando la fórmula de integración por partes y teniendo en cuenta que toda función básica φ se anula fuera de un intervalo finito, obtenemos

$$\frac{dT}{dx}(\varphi) = \int_{-\infty}^{\infty} f'(x)\,\varphi(x)\,dx = -\int_{-\infty}^{\infty} f(x)\,\varphi'(x)\,dx;$$

de esta forma, hemos obtenido para $\frac{dT}{dx}$ una expresión en la que no figura la derivada de f. Estas consideraciones sugieren la siguiente definición.

DEFINICION 2. Se llama *derivada* $\frac{dT}{dx}$ de la función generalizada T a la funcional definida mediante la fórmula

$$\frac{dT}{dx}(\varphi) = -T(\varphi').$$

Está claro que la funcional definida mediante esta fórmula es lineal y continua, es decir, representa una función generalizada. Análogamente se definen la segunda, la tercera, etc. derivadas.

Denotando la función generalizada mediante el símbolo f, denotaremos su derivada (comprendida en el sentido que acabamos de definir) mediante el símbolo corriente f'.

Directamente de la definición de la derivada de una función generalizada se deduce la validez de las proposiciones siguientes:

1. *Toda función generalizada tiene derivadas de todos los órdenes.*

2. *Si una sucesión* $\{f_n\}$ *de funciones generalizadas converge a una función generalizada* f (en el sentido de la definición de convergencia de funciones generalizadas), *la sucesión* $\{f_n'\}$ *de derivadas converge a la derivada* f' *de la función límite.* Esto equivale a que toda serie convergente compuesta por funciones generalizadas se puede derivar término a término cualquier número de veces.

Veamos algunos ejemplos.

1. De lo expuesto anteriormente se ve que siendo f una función regular (es decir, «verdadera») cuya derivada existe y es continua (o continua a trozos), la derivada de ella, como de función generalizada, coincide con su derivada en el sentido corriente.

2. Sea

$$f(x) = \begin{cases} 1 \text{ para } x \geqslant 0, \\ 0 \text{ para } x < 0. \end{cases} \tag{5}$$

Esta función define la funcional lineal

$$(f, \varphi) = \int_0^\infty \varphi(x)\, dx.$$

De acuerdo con la definición de la derivada de una función generalizada, tenemos

$$(f', \varphi) = -(f, \varphi') = -\int_0^\infty \varphi'(x)\, dx = \varphi(0)$$

(puesto que φ es igual a 0 en el infinito). Por consiguiente, la derivada de la función (5) (de la «función salto unidad») es la δ-función.

3. De los ejemplos 1 y 2 se ve claramente que si f es una función que en los puntos $x_1, x_2, \ldots$ tiene saltos iguales a $h_1, h_2, \ldots$ respectivamente y es diferenciable en los demás puntos (en el sentido corriente), su derivada (como de una función generalizada) representa la suma de la derivada corriente f' (en los puntos donde ésta existe) más la suma de tipo

$$\sum_i h_i \delta(x-x_i).$$

4. Aplicando la definición de derivada de una función generalizada a la δ-función, obtendremos que esta derivada representa la funcional que en cada función $\varphi \in K$ toma el valor $-\varphi'(0)$. Es precisamente la funcional que hemos denominado ya «derivada de la δ-función».

5. Consideremos la serie

$$\sum_{n=1}^{\infty} \frac{\operatorname{sen} nx}{n}. \tag{6}$$

Su suma representa una función de período 2π que en el intervalo $(-\pi, \pi)$ se define mediante las fórmulas

$$f(x) = \begin{cases} \dfrac{\pi - x}{2} & \text{para } 0 < x \leqslant \pi \\ -\dfrac{\pi + x}{2} & \text{para } -\pi \leqslant x < 0. \end{cases}$$

La derivada generalizada de esta función es igual a

$$-\frac{1}{2} + \pi \sum_{k=-\infty}^{\infty} \delta(x - 2k\pi). \tag{7}$$

Representa una función generalizada (al aplicarla a cualquier función terminal $\varphi(x)$ siempre obtendremos solamente un número finito de sumandos diferentes de cero). Por otro lado, derivando término a término la serie $\sum\limits_{n=1}^{\infty} \frac{\operatorname{sen} nx}{n}$ obtenemos una serie divergente

$$\sum_{n=1}^{\infty} \cos nx.$$

Sin embargo, en el sentido de convergencia de funciones generalizadas, esta serie converge (y precisamente a la expresión (7)). Por consiguiente, el concepto de función generalizada permite dar un sentido completamente determinado a la suma de una serie que en el sentido corriente diverge. Lo mismo se refiere a muchas integrales divergentes. Con situaciones de este tipo se tropieza frecuentemente en la teoría cuántica del campo y en otras varias ramas de la Física Teórica.

5°. **Suficiencia del conjunto de funciones básicas.** Hemos definido las funciones generalizadas como funcionales lineales sobre un espacio, el espacio K de funciones terminales indefinidamente diferenciables. En general, podíamos escoger el espacio básico de alguna otra manera. Veamos las consideraciones que pueden determinar la selección de uno u otro espacio para el espacio de funciones básicas. Sujetando los elementos de K a las exigencias rígidas de terminalidad y diferenciabilidad indefinida, hemos obtenido con ello, en primer lugar, un conjunto suficientemente amplio de funciones generalizadas (toda restricción del espacio básico lleva, evidentemente, a la ampliación del espacio dual) y, en segundo lugar, una gran libertad para aplicar a las funciones generalizadas las operaciones principales del Análisis (paso al límite, diferenciación). Pero al mismo tiempo el espacio de funciones básicas no puede ser muy restringido. Debe tener un número suficiente de elementos para que mediante ellos se puedan distinguir dos funciones diferentes. Hablando con más rigor, si f_1 y f_2 son dos funciones integrables (es decir, dos funciones generalizadas regulares), debe existir un elemento φ del espacio básico tal que

$$\int_{-\infty}^{\infty} f_1(x)\,\varphi(x)\,dx \neq \int_{-\infty}^{\infty} f_2(x)\,\varphi(x)\,dx. \tag{8}$$

Comprobemos que el espacio K satisface esta condición. Con más precisión, *sean f_1 y f_2 dos distintas funciones continuas* (y, por consiguiente, localmente integrables) *sobre la recta. Entonces, existe una función $\varphi \in K$ tal que se cumple la condición* (8).

En efecto, pongamos $f(x) = f_1(x) - f_2(x)$. Si $f(x) \not\equiv 0$, existe un punto x_0 tal que $f(x_0) \neq 0$. Entonces, $f(x) \neq 0$ y conserva su signo en un intervalo (α, β) que contiene el punto x_0. Consideremos la función

$$\varphi(x) = \begin{cases} e^{-\frac{1}{(x-\alpha)^2}} \cdot e^{-\frac{1}{(x-\beta)^2}} & \text{para } \alpha < x < \beta, \\ 0 & \text{para los demás } x. \end{cases}$$

Esta función se anula fuera de (α, β) y es positiva dentro de este intervalo; además, tiene derivadas de todos los órdenes, de manera que $\varphi \in K$. Es evidente que

$$\int_{-\infty}^{\infty} f(x)\varphi(x)\,dx = \int_{\alpha}^{\beta} f(x)\varphi(x)\,dx \neq 0.$$

Luego, hemos demostrado que el espacio K contiene un número suficientemente grande de funciones para que sea posible distinguir dos funciones continuas cualesquiera [1].

6°. Reconstrucción de una función por su derivada. Ecuaciones diferenciales en la clase de funciones generalizadas. Las ecuaciones diferenciales son una de las regiones principales de aplicación de la teoría de funciones generalizadas. Fueron precisamente las ecuaciones diferenciales que estimularon, en gran medida, el desarrollo de la teoría de funciones generalizadas. Estas aplicaciones están relacionadas, en general, con ecuaciones en derivadas parciales y su estudio detallado sale de los márgenes de este libro. Sin embargo, consideraremos aquí algunas cuestiones más sencillas, relacionadas con la solución de ecuaciones diferenciales (ordinarias) con funciones generalizadas. Comencemos por la ecuación elemental de tipo

$$y' = f(x)$$

($f(x)$ es una función o bien generalizada o bien «corriente»), es decir, por el problema de reconstrucción de una función a partir de su derivada.

TEOREMA 1. *Sólo las constantes son soluciones (en la clase de funciones generalizadas) de la ecuación*

$$y' = 0. \tag{9}$$

DEMOSTRACION. La ecuación (9) significa que

$$(y', \varphi) = (y, -\varphi') = 0 \tag{10}$$

para toda función básica $\varphi \in K$. La igualdad (10) define la funcional y sobre el conjunto $K^{(1)}$ de funciones básicas, cada una de las cuales puede representarse como derivada de cierta función básica.

Una función básica φ_0 puede representarse como derivada de otra función básica si, y sólo si,

$$\int_{-\infty}^{\infty} \varphi_0(x)\,dx = 0. \tag{11}$$

[1] Esta proposición se puede extender también a funciones substancialmente más generales que las continuas, pero para ello habría que recurrir al concepto de integrabilidad según Lebesgue, del cual hablaremos en el capítulo VI.

En efecto, si $\varphi_0(x)=\varphi_1'(x)$, tenemos

$$\int_{-\infty}^{\infty} \varphi_0(x)\,dx=\varphi_1(x)\Big|_{-\infty}^{\infty}=0. \tag{12}$$

Viceversa, si se cumple la igualdad (11), la expresión

$$\varphi_1(x)=\int_{-\infty}^{x} \varphi_0(t)\,dt \tag{13}$$

es una función indefinidamente diferenciable y, además, terminal debido a (12). Su derivada es $\varphi_0(x)$. Observemos ahora que toda función básica φ se puede representar en la forma

$$\varphi=\varphi_0+c\varphi_1,$$

donde φ_1 es una función básica fija que verifica la condición

$$\int_{-\infty}^{\infty} \varphi_1(x)\,dx=1,$$

c es una constante y $\varphi_0(x)$ es tal que $\int\limits_{-\infty}^{\infty} \varphi_0(x)\,dx=0$. En efecto, basta tomar

$$c=\int_{-\infty}^{\infty} \varphi(x)\,dx \text{ y } \varphi_0(x)=\varphi(x)-\varphi_1(x)\int_{-\infty}^{\infty} \varphi(x)\,dx.$$

Por consiguiente, al definir el valor de la funcional (= función generalizada) y para la función básica $\varphi_1(x)$, la funcional resulta unívocamente determinada en todo el K. Poniendo $(y, \varphi_1)=\alpha$, obtenemos

$$(y, \varphi)=\alpha\int_{-\infty}^{\infty} \varphi(x)\,dx=\int_{-\infty}^{\infty} \alpha\varphi(x)\,dx,$$

es decir, la función generalizada y es la constante α, que es lo que necesitábamos demostrar. De aquí se deduce que si para dos funciones generalizadas f y g se cumple la igualdad $f'=g'$, se tiene $f-g=\text{const}$.

Consideremos ahora la ecuación

$$y'=f(x) \tag{14}$$

donde $f(x)$ es una función generalizada. Probemos que para todo $f\in K^*$ la ecuación (14) tiene una solución perteneciente a K^*. Es natural llamar esta solución *primitiva* de la función generalizada f. La ecuación (14) significa que

$$(y, \ -\varphi') = (y', \ \varphi) = (f, \ \varphi) = \left(f, \ \int_{-\infty}^{x} \varphi'(\xi)\, d\xi\right) \tag{15}$$

para cualquier función básica $\varphi \in K$. Esta igualdad define el valor de la funcional y para todas las funciones básicas que pueden ser representadas como φ', donde $\varphi \in K$. Como hemos visto al demostrar el teorema 1, cualquier elemento $\varphi \in K$ puede representarse en la forma

$$\varphi = \varphi_0 + \varphi_1(x) \int_{-\infty}^{\infty} \varphi(x)\, dx,$$

donde φ_0 es la derivada de una función $\varphi^* \in K$ y $\varphi_1(x)$ es un elemento fijo de K que verifica la condición $\int_{-\infty}^{\infty} \varphi_1(x)\, dx = 1$. Si tomamos, además, $(y, \ \varphi_1) = 0$, la funcional y quedará definida en todo el K:

$$(y, \ \varphi) = (y, \ \varphi_0) = \left(f, \ \int_{-\infty}^{x} \varphi_0(\xi)\, d\xi\right).$$

Es fácil comprobar que esta funcional es lineal y continua. Además, satisface la condición (14). En efecto, para todo $\varphi \in K$

$$(y', \ \varphi) = (y, \ -\varphi') = \left(f, \ \int_{-\infty}^{x} \varphi'(\xi)\, d\xi\right) = (f, \ \varphi).$$

Luego, para cada función generalizada $f(x)$ existe una solución de la ecuación

$$y' = f(x),$$

es decir, cada función generalizada tiene primitiva. Según el teorema 1, esta primitiva se define unívocamente, a menos de una constante aditiva arbitraria, por la función $f(x)$.

Los resultados obtenidos se pueden extender fácilmente a sistemas de ecuaciones lineales. Nos limitaremos aquí a dar los enunciados correspondientes, omitiendo las demostraciones.

Consideremos un sistema homogéneo de n ecuaciones lineales con n funciones incógnitas

$$y_i' = \sum_{k=1}^{n} a_{ik}(x)\, y_k, \quad i = 1, \ 2, \ \ldots, \ n, \tag{16}$$

donde a_{ik} son funciones indefinidamente diferenciables. Este sistema tiene una determinada cantidad de soluciones «clásicas» (es decir, soluciones que representan funciones «corrientes» y, además, indefinidamente diferenciables). Se puede demostrar que, en la clase de funciones generalizadas, el sistema (16) no tiene ninguna otra solución.

Para un sistema no homogéneo de tipo

$$y_i' = \sum_{k=1}^{n} a_{ik} y_k + f_i, \tag{17}$$

donde f_i son funciones generalizadas y a_{ik} funciones corrientes indefinidamente diferenciables, existe en la clase de funciones generalizadas una solución que, salvo una solución arbitraria del sistema homogéneo (16), es única.

Si en el sistema (17) no sólo a_{ik} sino también f_i son funciones «corrientes», todas las soluciones que este sistema tiene en K' también resultan ser funciones corrientes.

7°. **Algunas generalizaciones.** Hemos considerado más arriba funciones generalizadas «de una variable real», esto es, funciones generalizadas sobre la recta. Es posible, basándose en las mismas ideas, introducir funciones generalizadas sobre un conjunto acotado, digamos, sobre un segmento o una circunferencia, así como funciones generalizadas de varias variables, funciones generalizadas de argumento complejo, etc. Además, incluso para las funciones generalizadas sobre la recta, la definición dada anteriormente no es la única posible. Veamos brevemente algunas de estas generalizaciones y modificaciones del concepto de función generalizada.

a) *Funciones de varias variables.* Consideremos en el espacio n-dimensional el conjunto K^n de funciones $\varphi(x_1, x_2, \ldots, x_n)$ que tienen derivadas parciales de todos los órdenes respecto a todos sus argumentos y tales que cada una de estas funciones se anula fuera de un paralelepípedo

$$a_i \leqslant x_i \leqslant b_i, \quad i = 1, 2, \ldots, n.$$

El conjunto K^n representa un espacio lineal (con las habituales operaciones de adición de funciones y multiplicación de las mismas por números), en el cual se puede definir una convergencia del siguiente modo: $\varphi_k \to \varphi$, cuando existe un paralelepípedo

$$a_i \leqslant x_i \leqslant b_i, \quad i = 1, 2, \ldots, n,$$

fuera del cual se anula cada una de las funciones φ_k y cuando en este paralelepípedo se tiene uniformemente

$$\frac{\partial^r \varphi_k}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}} \to \frac{\partial^r \varphi}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}} \left(\sum_{i=1}^{n} \alpha_i = r \right),$$

cualesquiera que sean r, α_1, ..., α_n. Una función generalizada de n variables es cualquier funcional lineal continua sobre K^n. Toda función $f(x)$ «corriente» de n variables, integrable en cualquier recinto acotado del espacio n-dimensional, es, al mismo tiempo, una función generalizada. Los valores de la funcional que le corresponde vienen dados por la fórmula

$$(f, \varphi) = \int f(x)\varphi(x)\,dx \quad (x = (x_1, \ldots, x_n),\ dx = dx_1 \ldots dx_n).$$

Al igual que en el caso de $n = 1$, diferentes funciones continuas determinan diferentes funcionales (es decir, representan diferentes funciones generalizadas).

Para las funciones generalizadas de n variables los conceptos de paso al límite, de derivada, etc. se introducen con los mismos métodos que en el caso de una variable. Por ejemplo, las derivadas parciales de una función generalizada se introducen mediante las fórmulas

$$\left(\frac{\partial^r f(x)}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}}, \varphi(x) \right) = (-1)^r \left(f(x), \frac{\partial^r \varphi(x)}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}} \right).$$

De aquí se desprende que cada función generalizada de n variables tiene derivadas parciales de todos los órdenes.

b) *Funciones generalizadas complejas.* Escojamos ahora para las funciones básicas el conjunto de funciones terminales, indefinidamente diferenciables sobre la recta, que toman valores complejos. Las funcionales lineales sobre el espacio $\tilde{K}$ de estas funciones es natural llamarlas funciones generalizadas complejas. Recordemos que en un espacio lineal complejo existen funcionales lineales de primera y segunda especie. Las primeras satisfacen la condición (α es un número)

$$(f, d\varphi) = \alpha(f, \varphi)$$

y las segundas, la condición

$$(f, \alpha\varphi) = \bar{\alpha}(f, \varphi).$$

Si $f(x)$ es una función corriente de valores complejos sobre la recta, se le puede asignar una funcional lineal de primera especie sobre $\tilde{K}$ de dos modos

$$(f, \varphi)_1 = \int_{-\infty}^{\infty} f(x)\varphi(x)\,dx \tag{18$_1$}$$

y

$$(f, \varphi)_2 = \int_{-\infty}^{\infty} \overline{f(x)}\, \varphi(x)\, dx. \tag{18$_2$}$$

A esta misma función $f(x)$ se le pueden asignar dos funcionales lineales de segunda especie, a saber:

$$_1(f, \varphi) = \int_{-\infty}^{\infty} f(x)\, \overline{\varphi(x)}\, dx \tag{18$_3$}$$

y

$$_2(f, \varphi) = \int_{-\infty}^{\infty} \overline{f(x)}\, \overline{\varphi(x)}\, dx. \tag{18$_4$}$$

La selección de una de estas cuatro posibilidades significa una manera determinada de inmersión del espacio de funciones «corrientes» en el espacio de funciones generalizadas. Las operaciones con las funciones generalizadas complejas se definen análogamente al caso de funciones reales.

c) *Funciones generalizadas sobre una circunferencia.* A veces conviene considerar funciones generalizadas definidas sobre un conjunto acotado. A título de ejemplo más sencillo, consideremos las funciones sobre una circunferencia. Por espacio de funciones básicas tomaremos el conjunto de todas las funciones indefinidamente diferenciables sobre una cincunferencia definiendo para ellas las operaciones de adición y multiplicación por números del modo habitual. La sucesión $\{\varphi_n(x)\}$ de funciones de este espacio se llama convergente, cuando para cada $k=0, 1, 2, \ldots$ converge uniformemente sobre toda la circunferencia la sucesión de derivadas $\{\varphi_n^{(k)}(x)\}$. Puesto que el conjunto de argumentos (la cincunferencia) es, en este caso, acotado, la condición de terminalidad de las funciones básicas desaparece automáticamente. Las funcionales lineales sobre este espacio se llamarán funciones generalizadas sobre la circunferencia. Toda función corriente sobre la circunferencia se puede considerar como una función periódica definida sobre la recta. Extendiendo esta idea a las funciones generalizadas, se puede identificar las funciones generalizadas sobre la circunferencia con las funciones generalizadas periódicas. Es natural entender por función generalizada periódica (de período a) una funcional f que verifica la condición

$$(f(x), \varphi(x-a) = (f(x), \varphi(x))$$

para toda función básica φ. Como ejemplo de función generali zada periódica puede servir la función

$$\sum_{n=1}^{\infty} \cos nx = -\frac{1}{2} + \pi \sum_{k=-\infty}^{\infty} \delta(x-2k\pi),$$

que ha sido mencionada anteriormente.

b) *Otros espacios básicos.* Hemos definido más arriba las funciones generalizadas sobre la recta como las funcionales lineales sobre el espacio K de funciones terminales indefinidamente diferenciables. Sin embargo, esta selección del espacio básico no es la única posible. Por ejemplo, en lugar del espacio K de funciones terminales, se podría haber tomado el espacio, algo más amplio, de todas las funciones $\varphi(x)$ indefinidamente diferenciables sobre la recta que decrecen, junto con sus derivadas, más rápido que cualquier potencia de $\frac{1}{|x|}$. Hablando con más precisión, admitiremos que $\varphi(x)$ pertenece al espacio básico, que denotaremos S, cuando para cualesquiera fijos p, $q=0, 1, 2, \ldots$ existe una constante $C_{p,q}$ (que depende de p, q y φ) tal que

$$|x^p \varphi^{(q)}(x)| < C_{pq}, \quad -\infty < x < \infty. \tag{19}$$

La convergencia en S se define del siguiente modo: la sucesión $\{\varphi_n(x)\}$ se llama convergente hacia $\varphi(x)$, cuando para cada $q=0, 1, 2, \ldots$ la sucesión $\{\varphi_n^{(q)}(x)\}$ converge uniformemente en cualquier intervalo finito y cuando en las desigualdades

$$|x^p \varphi_n^{(q)}(x)| < C_{pq}$$

las constantes C_{pq} se pueden escoger de manera que no dependan de n.

De esta forma se obtiene un conjunto de funciones generalizadas algo más reducido que en el caso del espacio K. Por ejemplo, la función

$$f(x) = e^{x^2}$$

es una funcional lineal continua sobre K, pero no sobre S. Es cómodo tomar S por espacio básico al considerar, por ejemplo, la transformación de Fourier de funciones generalizadas. En general, como ha demostrado el desarrollo de la teoría de funciones generalizadas, no es necesario sujetarse a una elección determinada del espacio básico, sino que conviene variarla según los problemas que se consideran. Sin embargo, es necesario requerir de manera esencial que, por un lado, haya «un número suficientemente grande» de funciones básicas (para que se pueda mediante ellas distinguir las funciones «corrientes», más rigurosamente, las funcionales regulares) y que, por otro lado, estas funciones básicas sean diferenciables suficiente número de veces.

EJERCICIO. Compruébese que en el espacio S se puede introducir una estructura de espacio normado numerable, tomando, por ejemplo,

$$\|\varphi_n\| \sum_{p+q=n} \sup_{\substack{-\infty < x < \infty \\ 0 \leqslant i \leqslant p \\ 0 \leqslant j \leqslant q}} = |(1+|x|^i)\varphi^{(j)}(x)|,$$

y que la convergencia de sucesiones en este espacio normado numerable equivale a la convergencia definida más arriba.

§ 5. OPERADORES LINEALES

1°. Definición y ejemplos de operadores lineales. Sean E y E_1 dos espacios topológicos lineales. Se llama *operador lineal* que actúa de E en E_1 a toda aplicación

$$y = Ax \quad (x \in E,\ y \in E_1),$$

que verifica la condición

$$A(\alpha x_1 + \beta x_2) = \alpha A x_1 + \beta A x_2.$$

El conjunto D_A de todos los $x \in E$ para los cuales está definida la aplicación A se llama *campo de definición* del operador A; en general, no se supone que $D_A = E$; sin embargo, siempre admitiremos que D_A es una variedad lineal, esto es, que si $x,\ y \in D_A$, también $\alpha x + \beta y \in D_A$ para todos los $\alpha,\ \beta$.

Un operador A se llama *continuo en el punto* $x_0 \in D_A$, cuando para cualquier vecindad V del punto $y_0 = Ax_0$ existe una vecindad U del punto x_0 tal que $Ax \in V$ siempre que $x \in U \cap D_A$. El operador A se llama *continuo*, cuando es continuo en cada punto $x \in D_A$.

No es difícil probar que, cuando E y E_1 son espacios normados, esta definición equivale a la siguiente: un operador A se llama continuo, cuando para cualquier $\varepsilon > 0$ existe un $\delta > 0$ tal que de la desigualdad

$$\|x' - x''\| < \delta \quad (x',\ x'' \in D_A)$$

se deduce

$$\|Ax' - Ax''\| < \varepsilon.$$

Está claro que el concepto de funcional lineal, introducido al principio de este capítulo, es un caso particular de operador lineal. Es decir, una funcional lineal es un operador lineal que transforma el espacio dado E en la recta numérica E_1. En esta suposición, las definiciones de linealidad y continuidad, que hemos enunciado para un operador lineal, se transforman en las correspondientes definiciones que hemos introducido con anterioridad para las funcionales.

Del mismo modo, una serie de conceptos y resultados que se exponen a continuación para los operadores lineales, representan una generalización suficiente automática de los resultados expuestos ya en el § 1 de este capítulo para el caso de funcionales lineales.

Ejemplos de operadores lineales.

1. Sea E un espacio lineal topológico. Pongamos

$$Ax = x \text{ para todo } x \in E.$$

Este operador, que transforma cada elemento del espacio en sí mismo, se llama *operador unidad.*

2. Si E y E_1 son dos espacios topológicos lineales arbitrarios y

$$Ox = 0 \text{ para todo } x \in E$$

(aquí 0 es el elemento cero del espacio E_1), se dice que O es el *operador nulo.*

3. (Forma general de un operador lineal que transforma un espacio de dimensión finita en otro de dimensión finita). Sea A un operador lineal que transforma el espacio R^n de n dimensiones con la base $e_1, \ldots, e_n$ en el espacio R^m de m dimensiones con la base $f_1, \ldots, f_m$. Si x es un vector arbitrario de R^n, tenemos

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

y, debido a la linealidad del operador A.

$$Ax = \sum_{i=1}^{n} x_i A e_i.$$

De manera que el operador A queda definido si se conoce en qué transforma los vectores $e_1, \ldots, e_n$ de la base. Consideremos el desarrollo del vector Ae_i según la base $f_1, \ldots, f_m$. Tenemos

$$Ae_i = \sum_{j=1}^{m} a_{ij} f_j.$$

De aquí se ve que el operador A se define por la matriz de los coeficientes a_{ij}. La imagen en R^m del espacio R^n representa un subespacio lineal cuya dimensión es igual, evidentemente, al rango de la matriz $\|a_{ij}\|$, esto es, no es mayor, en todo caso, que n. Señalemos que en un espacio de dimensión finita todo operador lineal es automáticamente continuo.

4. Consideremos el espacio de Hilbert H y un subespacio H_1 suyo. Desarrollando H en la suma directa del subespacio H_1 y su complemento ortogonal, es decir, representando cada elemento

$h \in H$ en la forma

$$h = h_1 + h_2 \quad (h_1 \in H_1, \ h_2 \perp H_1)$$

podemos tomar

$$Ph = h_1.$$

Es natural llamar este operador P operador proyectivo, que proyecta todo el H sobre H_1. La linealidad y continuidad se comprueban sin dificultad.

5. Consideremos en el espacio de funciones continuas sobre el segmento $[a, b]$ el operador dado mediante la fórmula

$$\psi(t) = \int_a^b K(s, t)\, \varphi(s)\, ds, \tag{1}$$

donde $K(s, t)$ es una función fija continua de dos variables. La función $\psi(t)$ es continua cualquiera que sea la función continua $\varphi(s)$, de manera que el operador (1) transforma el espacio de funciones continuas en sí mismo. Su linealidad es obvia. Para poder hablar de su continuidad, es necesario señalar previamente la topología que se ha tomado en nuestro espacio de funciones continuas. Proponemos al lector demostrar la continuidad de este operador en los casos en que a) se considera el espacio $C_{[a, b]}$, esto es, el espacio de funciones continuas con la norma $\|\varphi\| = \max |\varphi(t)|$; b) se considera el espacio $C^2_{[a, b]}$, es decir, $\|\varphi\| = \left(\int_a^b \varphi^2(t)\, dt \right)^{\frac{1}{2}}$.

6. Consideremos en el mismo espacio de funciones continuas el operador

$$\psi(t) = \varphi_0(t)\, \varphi(t),$$

donde $\varphi_0(t)$ es una función continua fija. La linealidad de este operador es evidente. Demuéstrese su continuidad en el caso de las normas señaladas en el ejemplo anterior.

7. Uno de los ejemplos de operador lineal más importantes para el Análisis, es el operador de diferenciación. Puede ser considerado en diferentes espacios.

a) Consideremos el espacio de funciones continuas $C_{[a, b]}$ y el operador

$$Df(t) = f'(t)$$

que actúa en él. Este operador (que consideramos que actúa de $C_{[a, b]}$ en $C_{[a, b]}$ otra vez) está definido, evidentemente, no en todo el espacio de funciones continuas, sino en la variedad lineal de funciones que tienen derivada continua. El operador D es lineal, pero no es continuo. Esto se ve, por ejemplo, de que la sucesión

$$\varphi_n(t) = \frac{\operatorname{sen} nt}{n}$$

converge a 0 (en la métrica de $C_{[a, b]}$), mientras que la sucesión

$$D\varphi_n(t) = \cos nt$$

no converge.

b) El operador de diferenciación puede ser considerado como un operador que actúa del espacio D_1 de funciones con derivada continua sobre $[a, b]$ con la norma

$$\|\varphi\|_1 = \text{máx}\,|\varphi(t)| + \max|\varphi'(t)|$$

en el espacio $C_{[a, b]}$. En este caso el operador D es lineal y continuo y transforma todo el D_1 en todo el $C_{[a, b]}$.

c) No resulta muy conveniente considerar el operador de diferenciación como un operador que actúa de D_1 en $C_{[a, b]}$, ya que, a pesar de obtener en este caso un operador continuo definido sobre todo el espacio, este operador no se puede aplicar dos veces a cualquier función de D_1. Es más cómodo considerar el operador de diferenciación en un espacio aún más reducido que D_1, en el espacio D_∞ de funciones indefinidamente diferenciables sobre $[a, b]$, cuya topología se define mediante un sistema numerable de normas

$$\|\varphi\|_n = \sup_{n,\, t}(|\varphi(t)|,\ |\varphi'(t)|,\ \ldots,\ |\varphi^{(n)}(t)|).$$

El operador de diferenciación transforma este espacio en sí mismo y, como es fácil de comprobar, es continuo sobre este espacio.

d) Las funciones indefinidamente diferenciables constituyen una clase muy reducida. Las funciones generalizadas ofrecen la posibilidad de considerar el operador de diferenciación como un operador definido en un espacio substancialmente más amplio y, al mismo tiempo, como un operador continuo. En el parágrafo anterior hemos hablado ya de cómo se define la operación de diferenciación para las funciones generalizadas. De lo allí expuesto se desprende que la diferenciación es un operador lineal en el espacio de funciones generalizadas y que, además, es continuo en el sentido de que de la convergencia de la sucesión $\{f_n(t)\}$ de funciones generalizadas hacia $f(t)$ se deduce la convergencia de la sucesión de sus derivadas hacia la derivada de la función generalizada $f(t)$.

2°. Continuidad y acotación. Un operador lineal que actúa de E en E_1 se llama *acotado*, cuando está definido sobre todo el E y transforma cada conjunto acotado en un conjunto también acotado. Entre la acotación y la continuidad de un operador lineal existe una relación estrecha y tienen lugar las siguientes proposiciones.

I. *Todo operador continuo es acotado.*

En efecto, sea $M \subset E$ un conjunto acotado y sea $AM \subset E_1$ un conjunto no acotado. Entonces, existe en E_1 una vecindad V del cero tal que ninguno de los conjuntos $\frac{1}{n} AM$ está contenido en V. Pero en este caso, existe una sucesión $\{x_n\} \subset M$ tal que ninguno de los elementos $\frac{1}{n} Ax_n$ pertenece a V y obtenemos[1] que $\frac{1}{n} x_n \to 0$ en E, mientras que la sucesión $\left\{\frac{1}{n} Ax_n\right\}$ no converge hacia el 0 en E_1; esto contradice a la continuidad del operador A.

II. *Si A es un operador acotado que actúa de E en E_1 y si en el espacio E se cumple el primer axioma de numerabilidad, el operador A es continuo.*

En efecto, si A no es continuo, existen una vecindad V del cero en E_1 y un sistema determinante $\{U_n\}$ de vecindades del cero en E, tales que para cada n existe un elemento $x_n \in \frac{1}{n} U_n$ tal que $Ax_n \overline{\in} V$. La sucesión $\{nx_n\}$ es acotada en E (e incluso converge hacia 0), mientras que la sucesión $\{nAx_n\}$ no es acotada en E_1 (ya que no está contenida en ninguno de los conjuntos V). De manera que si el operador A no es continuo y en E se cumple el primer axioma de numerabilidad, el operador A tampoco es acotado. Nuestra proposición queda demostrada.

Por consiguiente, en los espacios con el primer axioma de numerabilidad (a los que pertenecen, en particular, todos los espacios normados y normados numerables) la acotación de un operador lineal equivale a su continuidad.

Todos los operadores mencionados en los ejemplos 1, 2, 3, 4, 5 y 6 del punto anterior, son continuos. Como en todos los espacios allí considerados se cumple el primer axioma de numerabilidad, todos los operadores mencionados son acotados.

Si E y E_1 son espacios normados, la condición de acotación de un operador que actúa de E en E_1, se puede enunciar así: un operador A se llama acotado si transforma toda bola en un conjunto acotado. Debido a la linealidad de A, esta condición

[1] Véase el ejercicio del punto 3 del § 5.

se puede enunciar así: A es acotado, cuando existe una constante C tal que para todo $f \in E$

$$\| Af \| \leqslant C \| f \|.$$

El menor de los números C que satisfacen esta desigualdad se llama *norma* del operador A y se denota mediante $\| A \|$.

TEOREMA 1. *Para cualquier operador acotado A que actúa de un espacio normado en otro espacio normado, se tiene*

$$\| A \| = \sup_{\| x \| \leqslant 1} \| Ax \| = \sup_{x \neq 0} \frac{\| Ax \|}{\| x \|}. \tag{2}$$

DEMOSTRACION. Denotemos [1]

$$\alpha = \sup_{\| x \| \leqslant 1} \| Ax \| = \sup_{x \neq 0} \frac{\| Ax \|}{\| x \|}.$$

Probemos primero que $\| A \| \geqslant \alpha$.

Puesto que $\alpha = \sup \frac{\| Ax \|}{\| x \|}$, para cualquier $\varepsilon > 0$ existe un elemento x_ε tal que

$$\frac{\| Ax_\varepsilon \|}{\| x_\varepsilon \|} > \alpha - \varepsilon,$$

es decir,

$$\| Ax_\varepsilon \| > (\alpha - \varepsilon) \| x_\varepsilon \|,$$

de donde se deduce que

$$\alpha - \varepsilon < \| A \|;$$

como ε es un número positivo arbitrario, obtenemos $\alpha \leqslant \| A \|$.

Probemos ahora que la desigualdad $\alpha < \| A \|$ es imposible. Sea

$$\| A \| - \alpha = \varepsilon > 0.$$

Entonces,

$$\alpha < \| A \| - \frac{\varepsilon}{2}.$$

Pero de aquí se deduce que cualquiera que sea el punto x se cumplen las desigualdades

$$\frac{\| Ax \|}{\| x \|} \leqslant \alpha < \| A \| - \frac{\varepsilon}{2}$$

o bien

$$\| Ax \| \leqslant \left(\| A \| - \frac{\varepsilon}{2} \right) \| x \|,$$

---

[1] La igualdad $\sup\limits_{\| x \| \leqslant 1} \| Ax \| = \sup\limits_{x \neq 0} \frac{\| Ax \|}{\| x \|}$ es evidente debido a la linealidad de A.

es decir, $\|A\|$ no es la cota inferior de aquellos M para los cuales

$$\|Ax\| \leqslant M\|x\|.$$

Por consiguiente,

$$\|A\| = \alpha.$$

3°. Suma y producto de operadores.

DEFINICION 1. Sean A y B dos operadores lineales que actúan del espacio topológico lineal E en el espacio E_1. La *suma* $A+B$ de ellos es el operador C que pone en correspondencia al elemento $x \in E$ el elemento

$$y = Ax + Bx \in E_1.$$

Es fácil comprobar que $C = A+B$ es un operador lineal, continuo, si son continuos A y B. El campo de definición D_C del operador C es la intersección $D_A \cap C_B$ de los campos de difinición de los operadores A y B.

Si E y E_1 son espacios normados y si los operadores A y B son acotados, el operador C es también acotado y, además,

$$\|C\| \leqslant \|A\| + \|B\|. \tag{3}$$

En efecto, para todo x

$$\|Cx\| = \|Ax + Bx\| \leqslant \|Ax\| + \|Bx\| \leqslant (\|A\| + \|B\|)\|x\|,$$

de donde se deduce (3).

DEFINICION 2. Sean A y B dos operadores lineales, A actuando del espacio E en E_1 y B actuando de E_1 en E_2. Se llama *producto* BA de los operadores A y B al operador C que pone en correspondencia al elemento $x \in E$ el elemento

$$z = B(Ax)$$

de E_2. El campo de definición D_C del operador $C = BA$ se compone de todos los $x \in D_A$ tales que $Ax \in D_B$. Está claro que el operador C es lineal. Es continuo, si son continuos A y B.

Si A y B son operadores acotados que actúan en espacios normados, el operador $C = BA$ también es acotado y, además,

$$\|C\| \leqslant \|B\| \cdot \|A\|. \tag{4}$$

En efecto,

$$\|Cx\| = \|B(Ax)\| \leqslant \|B\| \cdot \|Ax\| \leqslant \|B\| \cdot \|A\| \cdot \|x\|, \tag{5}$$

de donde se sigue (4).

La suma y el producto de tres y más operadores se definen sucesivamente. Ambas operaciones son asociativas.

El producto de un operador A por un número k (se denota con kA) se define como el operador que pone en correspondencia al elemento x el elemento kAx.

El conjunto $\mathcal{L}(E, E_1)$ de todos los operadores lineales continuos, definidos en todo el espacio E, que transforman E en E_1 (donde E y E_1 son dos espacios topológicos lineales fijos), forma un espacio lineal respecto a las operaciones de adición y multiplicación por números, definidas más arriba. Si E y E_1 son espacios normados, $\mathcal{L}(E, E_1)$ es también un espacio normado (con la norma de operador que se ha definido anteriormente).

EJERCICIO. Sean E un espacio normado y E_1 un espacio normado completo. Entonces, el espacio normado $\mathcal{L}(E, E_1)$ es completo. Si $A_k \in \mathcal{L}(E, E_1)$ y $\sum\limits_{1=k}^{\infty} \|A_k\| < \infty$, la serie $\sum\limits_{k=1}^{\infty} A_k$ converge a un operador $A \in \mathcal{L}(E, E_1)$ y

$$\|A\| = \left\| \sum_{k=1}^{\infty} A_k \right\| \leqslant \sum_{k=1}^{\infty} \|A_k\|. \tag{6}$$

4°. Operador inverso, inversibilidad. Sea A un operador que actúa de E en E_1 y sean D_A el campo de definición y R_A el campo de valores de este operador.

DEFINICION. Un operador A se llama *inversible*, cuando para cualquier $y \in R_A$ la ecuación

$$Ax = y$$

tiene solución única.

Si A es inversible, a cada $y \in R_A$ se puede poner en correspondencia un elemento único $x \in D_A$ que es la solución de la ecuación $Ax = y$. El operador que realiza esta correspondencia se llama *inverso* de A y se denota mediante A^{-1}.

TEOREMA 2. *Un operador A^{-1}, inverso de un operador lineal A, es también lineal.*

DEMOSTRACION. Basta comprobar que se cumple la igualdad

$$A^{-1}(\alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 A^{-1} y_1 + \alpha_2 A^{-1} y_2. \tag{7}$$

Pongamos $Ax_1 = y_1$ y $Ax_2 = y_2$. Debido a la linealidad de A, tenemos

$$A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 y_1 + \alpha_2 y_2. \tag{8}$$

De acuerdo con la definición de operador inverso,

$$A^{-1} y_1 = x_1, \quad A^{-1} y_2 = x_2,$$

de donde, multiplicando estas igualdades por α_1 y α_2, respectivamente, y sumando, encontramos

$$\alpha_1 A^{-1} y_1 + \alpha_2 A^{-1} y_2 = \alpha_1 x_1 + \alpha_2 x_2.$$

Por otro lado, de (8) y de la definición de operador inverso se deduce que

$$\alpha_1 x_1 + \alpha_2 x_2 = A^{-1}(\alpha_1 y_1 + \alpha_2 y_2)$$

que junto con la igualdad anterior da

$$A^{-1}(\alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 A^{-1} y_1 + \alpha_2 A^{-1} y_2.$$

TEOREMA. 3 (teorema de Banach sobre el operador inverso). *Sea A un operador lineal acotado, que efectúa una transformación biunívoca del espacio de Banach E sobre el espacio de Banach E_1. Entonces, el operador inverso A^{-1} es acotado.*

Para demostrar este teorema es necesario el siguiente lema.

LEMA. *Sea M un conjunto siempre denso de un espacio de Banach E. Entonces, todo elemento no nulo $y \in E$ se puede desarrollar en una serie*

$$y = y_1 + y_2 + \ldots + y_n + \ldots,$$

donde $y_k \in M$ e $\|y_k\| \leqslant \frac{3\|y\|}{2^k}$.

DEMOSTRACION. Construimos la sucesión de elementos y_k del siguiente modo: escogemos y_1 de manera que sea

$$\|y - y_1\| < \frac{\|y\|}{2}, \tag{9}$$

lo cual es posible, ya que la desigualdad (9) define una esfera de radio $\frac{\|y\|}{2}$ y centro en el punto y, dentro de la cual debe existir un elemento de M (M es siempre denso en E). Escojamos $y_2 \in M$ de manera que $\|y - y_1 - y_2\| < \frac{\|y\|}{4}$, y_3 de manera que $\|y - y_1 - y_2 - y_3\| \leqslant \frac{\|y\|}{8}$ y, en general, escojamos y_n de manera que $\|y - y_1 - \ldots - y_n\| \leqslant \frac{\|y\|}{2^n}$. Tal selección es siempre posible, ya que M es siempre denso en E. De acuerdo con la elección de los elementos y_k:

$$\left\| y - \sum_{k=1}^{n} y_k \right\| \to 0 \text{ para } n \to \infty,$$

esto es, la serie $\sum_{k=1}^{\infty} y_k$ converge hacia y. Estimemos las normas de los elementos y_k:

$$\|y_1\| = \|y_1 - y + y\| \leqslant \|y_1 - y\| + \|y\| = \frac{3\|y\|}{2},$$

$$\|y_2\| = \|y_2 + y_1 - y + y - y_1\| \leqslant \|y - y_1 - y_2\| + \|y - y_1\| \leqslant \frac{3\|y\|}{4}.$$

Finalmente,

$$\|y_n\| = \|y_n + y_{n-1} + \ldots + y_1 - y + y - y_1 - \ldots - y_{n-1}\| \leqslant$$
$$\leqslant \|y - y_1 - \ldots - y_n\| + \|y - y_1 - \ldots - y_{n-1}\| \leqslant \frac{3\|y\|}{2^n}.$$

El lema queda demostrado.

DEMOSTRACION DEL TEOREMA 3. Consideremos en el espacio E_1 los conjuntos M_k, donde M_k es la totalidad de los y y que satisfacen la desigualdad $\|A^{-1}y\| \leqslant k\|y\|$. Todo elemento del espacio E_1 se encuentra en cierto M_k, es decir, $E_1 = \bigcup_{k=1}^{\infty} M_k$. De acuerdo al teorema de Baire sobre las categorías (cap. II, § 3, punto 3°), al menos uno de los conjuntos M_k, digamos M_n, es denso en una bola B_0. Consideremos dentro de la bola B_0 una capa esférica P, esto es, el conjunto de puntos z, para los cuales se cumple la desigualdad

$$\beta < \|z - y_0\| < \alpha,$$

donde

$$0 < \beta < \alpha, \quad y_0 \in M_n.$$

Trasladando paralelamente la capa esférica P de manera que su centro coincida con el origen de coordenadas, obtendremos la capa esférica P_0.

Probemos que en P_0 es denso un conjunto M_N. Sea $z \in P \cap M_n$; entonces, $z - y_0 \in P_0$ y

$$\|A^{-1}(z - y_0)\| \leqslant \|A^{-1}z\| + \|A^{-1}y_0\| \leqslant n(\|z\| + \|y_0\|) \leqslant$$
$$\leqslant n(\|z - y_0\| + 2\|y_0\|) = n\|z - y_0\| \left(1 + \frac{2\|y_0\|}{\|z - y_0\|}\right) \leqslant$$
$$\leqslant n\|z - y_0\| \left(1 + \frac{2\|y_0\|}{\beta}\right). \quad (10)$$

La magnitud $n\left[1 + \frac{2\|y_0\|}{\beta}\right]$ no depende de z. Tomemos $N = 1 + \left[1 + \frac{2\|y_0\|}{\beta}\right]$. Entonces, debido a (10), $z - y_0 \in M_N$ y como M_n es denso en P, M_N será denso en P_0. Consideremos un elemento y arbitrario no nulo de E_1. Siempre se puede escoger λ de manera que sea $\beta < \|\lambda y\| < \alpha$, es decir, que $\lambda y \in P_0$. Puesto que M_N es denso en P_0, se puede construir una sucesión $y_k \in M_N$ convergente hacia λy. Entonces, la sucesión $\frac{1}{\lambda} y_k$ converge hacia y. Es evidente que si $y_k \in M_N$, también $\frac{1}{\lambda} y_k \in M_N$ para cualquier número real $\lambda \neq 0$; por consiguiente, M_N es denso en $E_1 \setminus \{0\}$ y, por esto, también en E_1.

Consideremos un elemento no nulo $y \in E_1$; de acuerdo con el lema, puede ser desarrollado en una serie de elementos de M_N.

$$y = y_1 + y_2 + \ldots + y_k + \ldots,$$

siendo $\|y_k\| < \frac{3\|y\|}{2^k}$.

Consideremos en el espacio E la serie formada por las imágenes recíprocas de los elementos y_k, es decir, por elementos $x_k = A^{-1} y_k$.

Esta serie converge a un elemento x, ya que tiene lugar la desigualdad $\| x_k \| = \| A^{-1} y_k \| \leqslant N \| y_k \| < N \frac{3\|y\|}{2^k}$; además,

$$\| x \| \leqslant \sum_{k=1}^{\infty} \| x_k \| \leqslant 3N \|y\| \sum_{k=1}^{\infty} \frac{1}{2^k} = 3N \|y\|.$$

Como la serie $\sum\limits_{n=1}^{\infty} x_n$ converge y el operador A es continuo, se puede aplicar A a esta serie término por término. Obtendremos

$$Ax = Ax_1 + Ax_2 + \ldots = y_1 + y_2 + \ldots = y,$$

de donde $x = A^{-1} y$. Además,

$$\| A^{-1} y \| = \| x \| \leqslant 3N \| y \|$$

y como esta estimación es válida para cualquier $y \neq 0$, el operador A^{-1} es acotado. El teorema queda demostrado.

EJERCICIOS. 1. Sean E, E_1 dos espacios normados. Un operador lineal A, que actúa de E en E_1, con el campo de definición $D_A \in E$, se llama *cerrado*, cuando de las condiciones $x_n \in D_A$, $x_n \to x$, $Ax_n \to y$ se deduce que $x \in D_A$ y $Ax = y$. Compruébese que todo operador acotado es cerrado. Consideremos el producto directo $E \times E_1$ de los espacios E y E_1, esto es, el espacio lineal normado, compuesto de todos los pares posibles $[x, y]$, donde $x \in E$, $y \in E_1$, con la norma $\| [x, y] \| = \| x \| + \| y \|_1$ (mediante $\| \cdot \|$ designamos la norma en E y mediante $\| \cdot \|_1$, la norma en E_1). Se puede poner en correspondencia al operador A el conjunto $G_A = \{[x, y] : x \in D_A, \ y = Ax\} \subset E \times E_1$, que se denomina gráfico del operador. Compruébese que G_A es un conjunto lineal de $E \times E_1$ que es cerrado cuando, y sólo cuando, el operador A es cerrado.

Demuéstrese que *si E y E_1 son espacios de Banach y si el operador A está definido en todo E y es cerrado, es acotado (teorema de Banach sobre el gráfico cerrado).*

Sugerencia. Aplíquese el teorema 3 al operador $P:\{x, Ax\} \to x$ que actúa de G_A en E.

2. Demuéstrese que siendo A un operador lineal continuo que transforma biunívocamente un espacio completo normado numerable E sobre un espacio completo normado numerable E_1, el operador inverso A^{-1} es continuo. Enúnciese y demuéstrese el teorema sobre el gráfico cerrado para el caso de espacios normados numerables.

TEOREMA 4. *Sea A_0 un operador lineal acotado que transforma un espacio de Banach E sobre otro espacio de Banach E_1 y que tiene el inverso acotado A_0^{-1} y sea ΔA un operador lineal acotado que transforma E en E_1 y tal que $\|\Delta A\| < \frac{1}{\|A_0^{-1}\|}$. Entonces, el operador $A = A_0 + \Delta A$ transforma E sobre E_1 y tiene inverso acotado.*

DEMOSTRACION. Fijemos un elemento cualquiera $y \in E_1$ y consideremos la aplicación B del espacio E en sí mismo, definida mediante la fórmula

$$Bx = A_0^{-1}y - A_0^{-1}\Delta Ax.$$

De la condición $\|\Delta A\| < \|A_0^{-1}\|^{-1}$, se deduce que la aplicación B es contraída. Puesto que E es completo, existe un único punto fijo x de la aplicación B

$$x = Bx = A_0^{-1}\, y - A_0^{-1}\Delta\, Ax,$$

de donde

$$Ax = A_0 x + \Delta\, Ax = y.$$

Si $Ax' = y$, también x' será un punto fijo de la aplicación B, de manera que $x' = x$. Por consiguiente, para todo $y \in E_1$ la ecuación $Ax = y$ tiene en E una solución única, esto es, el operador A tiene inverso A^{-1}. En vista del teorema 3, el operador A^{-1} es acotado, que es lo que se quería demostrar.

TEOREMA 5. *Sean E un espacio de Banach, I el operador unidad de E y A un operador lineal acotado que aplica E en sí mismo y tal que $\|A\| < 1$. Entonces, el operador $(I - A)^{-1}$ existe, es acotado y represéntase en la forma*

$$(I - A)^{-1} = \sum_{k=0}^{\infty} A^k. \tag{11}$$

DEMOSTRACION. La existencia y acotación del operador $(I - A)^{-1}$ se desprende del teorema 4 (además, esto se desprende también del razonamiento que sigue). Como $\|A\| < 1$, tenemos $\sum_{k=0}^{\infty} \|A^k\| \leqslant$

$\leqslant \sum_{k=0}^{\infty} \| A \|^k < \infty$. El espacio E es completo y por eso de la convergencia de la serie $\sum_{k=0}^{\infty} \| A^k \|$ se deduce que la suma de la serie $\sum_{k=0}^{\infty} A^k$ representa un operador lineal acotado. Para un n cualquiera tenemos

$$(I-A)\sum_{k=0}^{n} A^k = \sum_{k=0}^{n} A^k (I-A) = I - A^{n+1};$$

pasando al límite para $n \longrightarrow \infty$ y tomando en consideración que $\| A^{n+1} \| \leqslant \| A \|^{n+1} \longrightarrow 0$, encontramos

$$(I-A)\sum_{k=0}^{\infty} A^k = \sum_{k=0}^{\infty} A^k (I-A) = I,$$

de donde

$$(I-A)^{-1} = \sum_{k=0}^{\infty} A^k,$$

que es lo que queríamos demostrar.

EJERCICIO. Sea A un operador lineal acotado que aplica un espacio de Banach E sobre un espacio de Banach E_1. Demuéstrese que existe una constante $\alpha > 0$ tal que si $B \in \mathscr{L}(E, E_1)$ y $\| A - B \| < \alpha$, el operador B aplica E sobre todo el espacio E_1 (Banach).

5°. **Operadores conjugados.** Consideremos un operador lineal continuo $y = Ax$ que aplica un espacio lineal topológico E en otro espacio E_1 del mismo tipo. Sea g una funcional lineal continua definida sobre E_1, esto es, $g \in E_1^*$. Apliquemos la funcional g al elemento $y = Ax$; es fácil comprobar que $g(Ax)$ es una funcional lineal continua definida en E; denotémosla con f. La funcional f es, de esta forma, un elemento del espacio E^*. Hemos asignado a cada funcional $g \in E_1^*$ una funcional $f \in E^*$, es decir, hemos obtenido un operador que aplica E_1^* en E^*. Este operador se llama *conjugado* del operador A y se designa mediante A^*.

Denotando la funcional f mediante (f, x), tendremos $(g, Ax) = (f, x)$ o

$$(g, Ax) = (A^* g, x).$$

Esta relación se puede tomar por definición del operador conjugado.

Ejemplo. *El operador conjugado en un espacio de dimensión finita.* Supongamos que un operador A aplica un espacio E^n (n-dimensional) en un espacio E^m (m-dimensional) y sea (a_{ij}) la matriz de este operador.

La aplicación $y = Ax$ se puede representar mediante el sistema de igualdades

$$y_i = \sum_{j=1}^{n} a_{ij}x_j, \quad i = 1, 2, \ldots, m$$

y la funcional f, (x) en la forma

$$f(x) = \sum_{j=1}^{n} f_j x_j.$$

De la igualdad

$$f(x) = g(Ax) = \sum_{i=1}^{m} g_i y_i = \sum_{i=1}^{m} \sum_{j=1}^{n} g_i a_{ij} x_j = \sum_{j=1}^{n} x_j \sum_{i=1}^{m} g_i a_{ij},$$

tendremos $f_j = \sum_{i=1}^{m} g_i a_{ij}$. Como $f = A^*g$, de aquí se deduce que el operador A^* se define mediante la matriz traspuesta de la matriz del operador A.

Las siguientes propiedades de operadores conjugados siguen directamente de la definición.

1. El operador A^* es lineal.
2. $(A+B)^* = A^* + B^*$.
3. Si k es un número complejo, $(kA)^* = kA^*$.

Menos evidente es el siguiente resultado.

TEOREMA 6. *El operador A^*, conjugado a un operador lineal acotado A que aplica un B-espacio E en un B-espacio E_1, es también acotado y*

$$\| A^* \| = \| A \|.$$

DEMOSTRACION. En virtud de las propiedades de la norma de un operador, tenemos

$$|(A^*g, x)| = |(g, Ax)| \leqslant \| g \| \cdot \| A \| \cdot \| x \|,$$

de donde $\| A^*g \| \leqslant \| A \| \cdot \| g \|$; por consiguiente,

$$\| A^* \| \leqslant \| A \|. \tag{12}$$

Sea $x \in E$ y $Ax \neq 0$; tomemos $y_0 = \frac{Ax}{\| Ax \|} \in E_1$; es obvio que $\| y_0 \| = 1$. De acuerdo con el corolario del teorema de Hahn — Banach, existe una funcional g, tal que $\| g \| = 1$ y $(g_0, y_0) = 1$, es decir, tal que $(g, Ax) = \| Ax \|$. De las relaciones

$$\| Ax \| = |(g, Ax)| = |(A^*g, x)| \leqslant \| A^*g \| \cdot \| x \| \leqslant$$
$$\leqslant \| A^* \| \cdot \| g \| \cdot \| x \| = \| A^* \| \cdot \| x \|$$

obtenemos $\| A \| \leqslant \| A^* \|$, lo que, junto con la desigualdad (12), da

$$\|A^*\| = \|A\|.$$

El teorema queda demostrado.

6°. Operador conjugado en un espacio euclídeo. Operadores autoconjugados. Consideremos el caso, cuando A es un operador acotado que actúa en un espacio de Hilbert H (real o complejo). De acuerdo con el teorema sobre la expresión general de una funcional lineal continua en un espacio de Hilbert, la aplicación τ que a cada $y \in H$ asigna la funcional lineal

$$(\tau y)(x) = (x, y)$$

es un isomorfismo (o un isomorfismo conjugado, cuando H es complejo) del espacio H sobre todo el espacio dual H^*. Sea A^* el operador conjugado del operador A. Está claro que la aplicación $\tilde{A}^* = \tau^{-1}A^*\tau$ representa un operador lineal acotado que actúa en H; es fácil ver que para cualesquiera $y \in H$

$$(Ax, y) = (x, \tilde{A}^*y).$$

Como $\|A^*\| = \|A\|$ y las aplicaciones τ y τ^{-1} son isométricas, tenemos $\|\tilde{A}^*\| = \|A\|$.

Todo lo expuesto anteriormente para un espacio de Hilbert, es válido también, por supuesto, para un espacio euclídeo (real o complejo) de dimensión finita.

Tomemos el siguiente acuerdo. Siendo R un espacio euclídeo (de dimensión finita o infinita), llamaremos operador conjugado del operador A, que actúa en R, al operador $\tilde{A}^*$, definido más arriba, que actúa en el mismo espacio R.

Es preciso subrayar que esta definición d i f i e r e de la definición de operador conjugado en un espacio arbitrario de Banach E, de acuerdo con la cual el operador conjugado A^* actúa en el espacio dual E^*. A veces, el operador $\tilde{A}^*$, en diferencia del operador A^*, se denomina *conjugado de Hermite*. Para no complicar la terminología ni las denotaciones, en lo que sigue escribiremos A^* en lugar de $\tilde{A}^*$ y llamaremos este operador *conjugado*, teniendo en cuenta, sin embargo, que en el caso de un espacio euclídeo el concepto de operador conjugado se comprenderá siempre tal como ha sido enunciado en esta sección.

Está claro, que en un espacio euclídeo R el operador conjugado de A se puede definir como un operador que para todos $x, y \in R$ verifica la igualdad

$$(Ax, y) = (x, A^*y).$$

Puesto que en el caso de un espacio euclídeo los operadores A y

A^* actúan en un mismo espacio, puede tener lugar la igualdad $A=A^*$. Hagamos la siguiente definición que destaca una clase importante de operadores en un espacio euclídeo (en particular, de Hilbert).

DEFINICION. Un operador lineal acotado que actúa en un espacio euclídeo R se llama *autoconjugado*, cuando $A=A^*$, es decir, cuando

$$(Ax, y)=(x, Ay)$$

para todos los $x, y \in R$.

Señalemos la siguiente propiedad importante del operador A^* conjugado a un operador A que actúa en un espacio euclídeo R. Un subespacio R_1 del espacio R se llama *invariante* respecto al operador A, cuando de $x \in R_1$ se deduce que $Ax \in R_1$. *Si el subespacio R_1 es invariante respecto de A, su complemento ortogonal $R_1^\perp$ es invariante respecto de A^*.* En efecto, si $y \in R_1^\perp$, tenemos para todo $x \in R_1$

$$(x, A^*y)=(Ax, y)=0,$$

ya que $Ax \in R_1$. En particular, si A es un operador autoconjugado, el complemento ortogonal de cualquier subespacio invariante suyo es invariante respecto de A.

EJERCICIO. Demuéstrese que siendo A y B dos operadores lineales acotados en un espacio euclídeo, se verifican las igualdades

$$(\alpha A+\beta B)^*=\bar{\alpha}A^*+\bar{\beta}B^*,$$
$$(AB)^*=(B^*A^*),$$
$$(A^*)^*=A,$$
$$I^*=I \quad (I \text{ es el operador unidad}).$$

7°. Espectro de un operador. Resolvente [1]. En la teoría de los operadores y en sus aplicaciones desempeña un papel primordial el concepto de espectro de un operador. Recordemos primero este concepto para el caso de operadores en un espacio de dimensión finita.

Sea A un operador lineal en el espacio n-dimensional E^n. El número λ se llama *valor propio* del operador A, cuando la ecuación

$$Ax=\lambda x$$

tiene soluciones no nulas. El conjunto de todos los valores propios se denomina *espectro* del operador A y todos los demás valores de λ se llaman *regulares*. En otras palabras, λ es un punto

[1] Siempre que se hable del espectro de un operador, consideramos que el operador actúa en un espacio complejo.

regular, cuando el operador $(A-\lambda I)$ es invertible. Además, el operador $(A-\lambda I)^{-1}$, como todo operador en un espacio de dimensión finita, es acotado. Por consiguiente, en un espacio de dimensión finita existen dos posibilidades:

1) la ecuación $Ax=\lambda x$ tiene solución no nula, es decir, λ es un valor propio de A; en este caso el operador $(A-\lambda I)^{-1}$ no existe;

2) existe el operador acotado $(A-\lambda I)^{-1}$, esto es, λ es un punto regular.

En el caso de dimensión infinita puede darse una tercera posibilidad, a saber:

3) el operador $(A-\lambda I)^{-1}$ existe, es decir, la ecuación $Ax=\lambda x$ tiene solamente solución nula, pero este operador no es acotado.

Introduzcamos la terminología siguiente. El número λ se llama *regular* para el operador A, que actúa en un espacio E (complejo) topológico lineal, cuando el operador $(A-\lambda I)^{-1}$, llamado *resolvente* del operador A, *está definido en todo el E y es continuo*. El conjunto de todos los demás valores de λ se llama *espectro* del operador A. Al espectro pertenecen todos los valores propios del operador A, ya que, si $(A-\lambda I)x=0$ para un $x\neq 0$, no existe $(A-\lambda I)^{-1}$. El conjunto de ellos se llama *espectro puntual*. La parte restante del espectro, es decir, el conjunto de valores de λ para los cuales $(A-\lambda I)^{-1}$ existe, pero no es continuo, se llama *espectro continuo*. De manera que cada valor de λ es para el operador A o bien regular, o bien valor propio, o bien punto del espectro continuo. La posibilidad de que un operador tenga espectro continuo es lo que distingue de un modo substancial la teoría de operadores en un espacio de dimensión infinita del caso de dimensión finita.

Sea A un operador que actúa en un B-espacio. Si el punto λ es regular, es decir, el operador $(A-\lambda I)^{-1}$ existe y es acotado, para un δ suficientemente pequeño el operador $(A-(\lambda+\delta)I)^{-1}$ también existe y es acotado (teorema 4), es decir, el punto $\lambda+\delta$ es también regular. Por consiguiente, *los puntos regulares constituyen un conjunto abierto*. Es decir, *el espectro*, esto es, el complemento de este conjunto, *es un conjunto cerrado*.

TEOREMA 7. *Si A es un operador lineal acotado en un espacio de Banach y si $|\lambda|>\|A\|$, λ es un punto regular.*

DEMOSTRACION. Como, evidentemente,

$$(A-\lambda I)=-\lambda\left(I-\frac{1}{\lambda}A\right)$$

tenemos

$$R_\lambda=(A-\lambda I)^{-1}=-\frac{1}{\lambda}\left(I-\frac{A}{\lambda}\right)^{-1}=-\frac{1}{\lambda}\sum_{0=k}^{\infty}\frac{A^k}{\lambda^k}.$$

Para $\|A\| < |\lambda|$, esta serie converge (véase el teorema 5), es decir, el operador $A-\lambda I$ tiene inverso acotado. En otras palabras, el espectro del operador A está contenido en el círculo de radio $\|A\|$ con centro en el cero.

Ejemplo. Consideremos en el espacio $C_{[0,1]}$ el operador A definido mediante la fórmula

$$Ax(t) = \mu(t)\,x(t),$$

donde $\mu(t)$ es una función continua fija. Tenemos

$$(A-\lambda I)\,x(t) = (\mu(t)-\lambda)\,x(t),$$

de donde

$$(A-\lambda I)^{-1}\,x(t) = \frac{1}{\mu(t)-\lambda}\,x(t).$$

El espectro del operador considerado A se compone de todos los λ tales que $\mu(t)-\lambda$ se anula para cierto t, comprendido entre 0 y 1, es decir, el espectro coincide con el conjunto de todos los valores de la función $\mu(t)$ sobre el segmento $[0, 1]$. Por ejemplo, si $\mu(t)=t$, el espectro representa el segmento $[0, 1]$ y no hay valores propios, es decir, el operador de multiplicación por t representa un ejemplo de operador con espectro puramente continuo.

Observaciones. (1) Todo operador lineal acotado, definido en un espacio de Banach completo, que tiene al menos un elemento diferente de cero, tiene espectro no vacío. Existen operadores, cuyo espectro se compone de un solo punto (por ejemplo, el operador de multiplicación por un número).

(2). Se puede precisar el teorema 7 del siguiente modo. Sea

$$r = \lim_{n\to\infty} \sqrt[n]{\|A^n\|}$$

(se puede demostrar que este límite existe cualquiera que sea el operador acotado A); entonces, el espectro del operador A se encuentra íntegramente dentro del círculo de radio r y centro en el cero. El número r se llama *radio espectral* del operador A.

(3). Los operadores resolventes R_μ y R_λ, correspondientes a los puntos μ y λ, conmutan y satisfacen la relación

$$R_\mu - R_\lambda = (\mu-\lambda)\,R_\mu R_\lambda$$

que se puede comprobar fácilmente multiplicando ambos miembros de esta igualdad por

$$(A-\lambda I)(A-\mu I).$$

De aquí se deduce que, siendo λ_0 un punto regular de A, la derivada de R_λ respecto a λ en $\lambda=\lambda_0$, es decir, el límite

$$\lim_{\Delta\lambda\to 0} \frac{R_{\lambda_0+\Delta\lambda} - R_{\lambda_0}}{\Delta\lambda}$$

existe (en el sentido de convergencia según la norma de operador) y es igual a $R_{\lambda_0}^2$.

EJERCICIO. Sea A un operador autoconjugado y acotado en un espacio complejo de Hilbert H. Demuéstrese que su espectro es un subconjunto cerrado y acotado del eje real.

§ 6. OPERADORES TOTALMENTE CONTINUOS

1°. Definición y ejemplos de operadores totalmente continuos. Una clase de operadores, que se aproxima por sus propiedades a la clase de operadores que actúan en espacios de dimensión finita y que, al mismo tiempo, es muy importante desde el punto de vista de las aplicaciones, es la clase formada por los así llamados operadores totalmente continuos. Estudiaremos ahora las propiedades principales de estos operadores, limitándonos al caso de espacio de Banach.

DEFINICION. Un operador A que aplica el espacio de Banach E en sí mismo se llama *totalmente continuo*, cuando transforma cada conjunto acotado en uno relativamente compacto.

En un espacio normado de dimensión finita, todo operador lineal es totalmente continuo, ya que transforma cualquier conjunto acotado en otro, acotado también, y en un espacio de dimensión finita todo conjunto acotado es relativamente compacto.

En un espacio de dimensión infinita la continuidad total de un operador es un requerimiento más fuerte que su continuidad simple (es decir, que la acotación). Por ejemplo, el operador unidad en el espacio de Hilbert es continuo, pero no totalmente continuo. (Demuéstrese esto independientemente del ejemplo 1 que se considera a continuación).

Veamos algunos ejemplos.

1. Sea I el operador unidad en un espacio de Banach E. Probemos que siendo E de dimensión infinita, el operador I no es totalmente continuo. Para ello bastará, evidentemente, demostrar que la bola unitaria de E (que, por supuesto, se transforma mediante el operador I en sí misma) no es compacta. Esto, a su vez, se desprende del siguiente lema que nos hará falta también en lo sucesivo.

LEMA. Sean $x_1, x_2, \ldots$ *vectores linealmente independientes de un espacio normado* E *y sea* E_n *el subespacio generado por los vectores* $x_1, \ldots, x_n$. *Entonces, existe una sucesión de vectores* $y_1, \ldots, y_n$ *que satisface las siguientes condiciones:*

$$1)\ \|y_n\|=1;\quad 2)\ y_n \in E_n;\quad 3)\ \rho(y_n, E_{n-1}) > \frac{1}{2}$$

(aquí $\rho(y_n E_{n-1})$ es la distancia del vector y_n hasta E_{n-1}, es decir, $\inf_{x \in E_{n-1}} \|y_n - x\|$).

DEMOSTRACION. En efecto, como los vectores $x_1, x_2, \ldots$ son linealmente independientes, tenemos $x_n \overline{\in} E_{n-1}$ y $\rho(x_n, E_{n-1}) = \alpha > 0$. Sea x^* un vector de E_{n-1} tal que $\|x_n - x^*\| < 2\alpha$. Entonces, $\rho(x_n - x^*, E_{n-1}) = \alpha$ y el vector

$$y_n = \frac{x_n - x^*}{\|x_n - x^*\|}$$

satisface todas las condiciones 1), 2) y 3). Por y_1 se puede tomar $\frac{x_1}{\|x_1\|}$. El lema queda demostrado.

Empleando este lema, se puede construir en la bola unitaria de cualquier espacio normado de dimensión infinita una sucesión de vectores $\{y_n\}$ para la cual $\rho(y_{n-1}, y_n) > \frac{1}{2}$. Está claro que esta sucesión no puede contener ninguna subsucesión convergente. Esto significa precisamente que no hay compacidad.

2. Sea A un operador lineal continuo que transforma un espacio de Banach E en un subespacio suyo de dimensión finita. Este operador es totalmente continuo, ya que transforma todo subconjunto acotado $M \subset E$ en un subconjunto acotado de un espacio de dimensión finita, es decir, en un conjunto relativamente compacto.

En particular, en un espacio de Hilbert el operador de proyección ortogonal sobre un subespacio es totalmente continuo cuando, y sólo cuando, este subespacio tiene dimensión finita.

Un operador que transforma un espacio de Banach E en un subespacio de dimensión finita se llama *degenerado*.

3. Consideremos en el espacio l_2 el operador A definido del siguiente modo: si $x = (x_1, x_2, \ldots, x_n, \ldots)$, entonces,

$$Ax = \left(x_1, \frac{1}{2}x_2, \ldots, \frac{1}{n}x_n, \ldots\right). \tag{1}$$

Este operador es totalmente continuo. En efecto, como todo conjunto acotado de l_2 está contenido en alguna bola de este espacio, basta demostrar que las imágenes de las bolas son relativamente compactas; debido a la linealidad del operador, basta comprobar esto para la bola unitaria. Pero el operador (1) transforma la bola unitaria del espacio l_2 en el conjunto de puntos que satisfacen la condición

$$\sum n^2 x_n^2 \leqslant 1$$

y la compacidad de este conjunto ha sido demostrada ya en el cap. II, § 7.

EJERCICIO. Sea $Ax=(a_1x_1, a_2x_2, \ldots, a_nx_n, \ldots)$; ¿qué condiciones debe cumplir la sucesión de números $a_1, a_2, \ldots$ para que este operador sea totalmente continuo en l_2?

4. En el espacio de funciones continuas $C_{[a,\,b]}$, forman una clase importante de operadores totalmente continuos los operadores que pueden representarse en la forma

$$Ax=y(s)=\int_a^b K(s,\ t)\,x(t)\,dt. \tag{2}$$

Probemos la validez de la siguiente proposición: *si la función $K(s,\ t)$ es acotada sobre el cuadrado $a\leqslant s\leqslant b$, $a\leqslant t\leqslant b$ y todos sus puntos de discontinuidad se encuentran en un número finito de curvas*

$$t=\varphi_k(s), \quad k=1,\ 2,\ \ldots,\ n,$$

donde φ_k son funciones continuas, la fórmula (2) *define en el espacio $C_{[a,\,b]}$ un operador totalmente continuo.*

En efecto, observemos, ante todo, que en las condiciones señaladas la integral (2) existe para cualquier s del segmento $[a,\ b]$, es decir, la función $y(s)$ está definida. Sea ahora

$$M=\sup_{a\leqslant s,\ t\leqslant b}|K(s,\ t)|$$

y sea G el conjunto de puntos $(s,\ t)$ en los cuales se cumple, al menos para un $k=1,\ 2,\ \ldots,\ n$, la desigualdad

$$|t-\varphi_k(s)|<\frac{\varepsilon}{12\,Mn}.$$

Sea F el complemento del conjunto G respecto al cuadrado $a\leqslant s$, $t\leqslant b$. Como F es compacto y la función $K(s,\ t)$ es continua sobre F, existe un $\delta>0$ tal que

$$|K(s',\ t)-K(s'',\ t)|<\frac{\varepsilon}{3\,(b-a)}$$

para cualesquiera dos puntos $(s',\ t)$, $(s'',\ t)$ de F que satisfacen la condición

$$|s'-s''|<\delta. \tag{3}$$

Estimemos la diferencia $y(s')-y(s'')$ admitiendo que s' y s'' satisfacen la condición (3). Tenemos

$$|y(s')-y(s'')|\leqslant\int_a^b|K(s',\ t)-K(s'',\ t)|\,|x(t)|\,dt\,|;$$

para evaluar la integral que figura en el miembro derecho, dividamos el segmento de integración $[a, b]$ en dos partes: la unión de intervalos

$$\bigcup_{k=1}^{n}\left[\left\{t:\ |t-\varphi_k(s')|<\frac{\varepsilon}{12Mn}\right\}\cup\left\{t:\ |t-\varphi_k(s'')|<\frac{\varepsilon}{12Mn}\right\}\right],$$

que denotaremos mediante P, y la parte restante del segmento $[a, b]$ que denotaremos mediante Q. Observando que P es la unión de intervalos tales que la suma de sus longitudes no pasa de $\frac{\varepsilon}{3M}$, obtenemos

$$\int_P |K(s', t)-K(s'', t)|\ |x(t)|\,dt<\frac{2\varepsilon}{3}\|x\|.$$

La integral sobre Q admite, obviamente, la estimación

$$\int_Q |K(s', t)-K(s'', t)|\ |x(t)|\,dt<\frac{\varepsilon}{3}\|x\|.$$

Por consiguiente,

$$|y(s')-y(s'')|<\varepsilon\|x\|. \tag{4}$$

La desigualdad (4) muestra que la función $y(s)$ es continua, es decir, la fórmula (2) define, en efecto, un operador que transforma el espacio $C_{[a, b]}$ en sí mismo. Además, se ve de la misma desigualdad que, siendo $\{x(t)\}$ un conjunto acotado de $C_{[a, b]}$, el conjunto correspondiente $\{y(s)\}$ es equicontinuo. Finalmente, si $\|x\|\leqslant C$, tenemos

$$\|y\|=\sup|y(s)|\leqslant\sup\int_a^b|K(s, t)|\ |x(t)|\,dt\leqslant M(b-a)\|x\|.$$

De manera que el operador (2) transforma todo conjunto acotado de $C_{[a, b]}$ en un conjunto de funciones equiacotado y equicontinuo, es decir, relativamente compacto.

4a. La exigencia de que los puntos de discontinuidad de la función $K(s, t)$ se encuentren sobre un número finito de curvas que intersectan las rectas $s=\text{const}$ en un solo punto, es substancial. Sea, por ejemplo,

$$K(s, t)=\begin{cases}1 \text{ para } s<\frac{1}{2},\\ 0 \text{ para } s\geqslant\frac{1}{2};\end{cases}$$

el operador (2) con este núcleo, que está definido sobre el cuadrado $0 \leqslant s,\ t \leqslant 1$ y que tiene por puntos de discontinuidad todo el segmento $s = \frac{1}{2}$, $0 \leqslant t \leqslant 1$, transforma la función $x(t) \equiv 0$ en una función discontinua.

4b. Si se toma $K(s, t) = 0$ para $t > s$, el operador (2) obtiene la forma

$$y(s) = \int_a^s K(s, t)\, x(t)\, dt. \tag{5}$$

Admitiremos que la función $K(s, t)$ es continua para $t < s$; entonces, de lo dicho en el ejemplo 4 se deduce que el operador (5) es totalmente continuo en $C_{[a, b]}$.

Este operador se llama *operador de Volterra*[1].

Observación. Con la definición que hemos adoptado de operador totalmente continuo puede resultar que la imagen de la bola unitaria cerrada no sea compacta (aunque es relativamente compacta). En efecto, consideremos en el espacio $C_{[-1, 1]}$ el operador de integración

$$(Jx)(s) = \int_1^s x(t)\, dt;$$

según hemos demostrado más arriba, J es un operador totalmente continuo en $C_{[-1, 1]}$. Tomemos

$$x_n(t) = \begin{cases} 0 & \text{para } -1 \leqslant t \leqslant 0, \\ nt & \text{para } \quad 0 < t \leqslant \frac{1}{n}, \\ 1 & \text{para } \quad \frac{1}{n} < t \leqslant 1. \end{cases}$$

En este caso, $x_n \in G_{[-1, 1]}$, $\|x_n\| = 1$ para todos los n e

$$y_n(t) = (Jx_n)(t) = \begin{cases} 0 & \text{para } -1 \leqslant t \leqslant 0, \\ \frac{1}{2} nt^2 & \text{para } \quad 0 < t \leqslant \frac{1}{n}, \\ t - \frac{1}{2n} & \text{para } \quad \frac{1}{n} < t \leqslant 1. \end{cases}$$

Está claro que la sucesión $\{y_n\}$ converge en $C_{[-1, 1]}$ a la función

$$y(t) = \begin{cases} 0 & \text{para } -1 \leqslant t \leqslant 0, \\ t & \text{para } \quad 0 \leqslant t \leqslant 1, \end{cases}$$

[1] Vitto Volterra, matemático italiano, autor de varias obras sobre Análisis Funcional y Ecuaciones Integrales.

que no es imagen (en la aplicación J) de ninguna función de $C_{[-1, 1]}$, ya que la función $y'(t)$ es discontinua.

Sin embargo, se puede demostrar que si el espacio es reflexivo (por ejemplo, de Hilbert) la imagen de la bola unitaria cerrada por una aplicación lineal totalmente continua es un compacto (para la demostración hay que valerse del resultado del ejercicio del punto 6 del § 2).

2°. Propiedades principales de operadores totalmente continuos.

TEOREMA 1. *Si $\{A_n\}$ es una sucesión de operadores totalmente continuos en un espacio de Banach E que converge, según la norma, a un operador A, el operador A es también totalmente continuo.*

DEMOSTRACION. Para probar la continuidad total del operador A bastará probar que cualquiera que sea la sucesión acotada $x_1, x_2, \ldots, x_n, \ldots$ de elementos de E ($\|x_n\| \leqslant C$), se puede extraer de la sucesión $\{Ax_n\}$ una subsucesión convergente.

Como el operador A_1 es totalmente continuo, de la sucesión $\{A_1x_n\}$ se puede extraer una subsucesión convergente. Sea

$$x_1^{(1)},\ x_2^{(1)},\ \ldots,\ x_n^{(1)},\ \ldots \tag{6}$$

una sucesión tal que $\{A_1x_n^{(1)}\}$ converge. Consideremos ahora la sucesión $\{A_2x_n^{(1)}\}$. De ella podemos también extraer una subsucesión convergente. Sea

$$x_1^{(2)},\ x_2^{(2)},\ \ldots,\ x_n^{(2)},\ \ldots$$

tal subsucesión de la sucesión (6) que $\{A_2x_n^{(2)}\}$ converge. Es evidente, entonces, que $\{A_1x_n^{(2)}\}$ también converge. Razonando de un modo análogo, escojamos de la sucesión $\{x_n^{(2)}\}$ una subsucesión

$$x_1^{(3)},\ x_2^{(3)},\ \ldots,\ x_n^{(3)},\ \ldots$$

tal que $\{A_3x_n^{(3)}\}$ converge, etc. Tomemos después la sucesión diagonal

$$x_1^{(1)},\ x_2^{(2)},\ \ldots,\ x_n^{(n)},\ \ldots$$

Cada uno de los operadores $A_1, A_2, \ldots, A_n, \ldots$ transforma esta sucesión en una convergente. Probemos que la sucesión $\{Ax_n^n\}$ también converge. Con ello quedará demostrada la continuidad total de A. Como el espacio E es completo, bastará demostrar que $\{Ax_n^{(n)}\}$ es una sucesión fundamental. Tenemos

$$\|Ax_n^{(n)} - Ax_m^{(m)}\| \leqslant \|Ax_n^{(n)} - A_kx_n^{(n)}\| + \\ + \|A_kx_n^{(n)} - A_kx_m^{(m)}\| + \|A_kx_m^{(m)} - Ax_m^{(m)}\|. \tag{7}$$

Escojamos primero k de manera que $\|A - A_k\| < \frac{\varepsilon}{3C}$ y después busquemos un N tal que para todos los $n > N$ y $m > N$ se cumpla la relación

$$\| A_k x_n^{(n)} - A_k x_m^{(m)} \| < \frac{\varepsilon}{3}$$

(esto es posible, ya que la sucesión $\{A_k x_n^{(n)}\}$ converge). En estas condiciones, obtenemos de (7) que

$$\| A x_n^{(n)} - A x_m^{(m)} \| < \varepsilon$$

para todos los n y m suficientemente grandes. El teorema queda demostrado.

Es fácil comprobar que una combinación lineal de operadores totalmente continuos es de nuevo un operador totalmente continuo. Por consiguiente, los operadores totalmente continuos forman, en el espacio $\mathscr{L}(E, E)$ de todos los operadores lineales acotados, definidos en E, un subespacio lineal cerrado.

Veamos ahora si el conjunto de operadores totalmente continuos está cerrado respecto a la operación de multiplicación de operadores. Resulta que en este orden es válida una afirmación substancialmente más profunda.

TEOREMA 2. *Si A es un operador totalmente continuo y B un operador acotado, los operadores AB y BA son totalmente continuos.*

DEMOSTRACION. Si el conjunto $M \subset E$ es acotado, BM también es acotado. Por consiguiente, ABM es relativamente compacto y esto significa precisamente que el operador AB es totalmente continuo. Además, si M es acotado, tenemos que AM es relativamente compacto y, debido a la continuidad de B, el conjunto BAM resulta también relativamente compacto, es decir, el operador BA es totalmente continuo. El teorema queda demostrado.

COROLARIO. *En un espacio E de dimensión infinita un operador totalmente continuo no puede tener un inverso acotado.*

En efecto, en el caso contrario, el operador unidad $I = A^{-1}A$ sería totalmente continuo en E, lo cual es imposible (véase el ejemplo 1).

Observación. El teorema 2 indica que los operadores totalmente continuos forman en el anillo de todos los operadores acotados $\mathscr{L}(E, E)$ un ideal bilateral [1)].

TEOREMA 3. *El operador conjugado a un operador totalmente continuo es totalmente continuo.*

DEMOSTRACION. Sea A un operador totalmente continuo en un espacio de Banach E. Probemos que el operador conjugado A^*,

---

1) Un *ideal* (bilateral) de un anillo R es un subanillo $\mathfrak{A}$ tal que, si $a \in \mathfrak{A}$ y $r \in R$, se tiene $ar \in \mathfrak{A}$ y $ra \in \mathfrak{A}$.

que actúa en E^*, transforma cada subconjunto acotado de E^* en uno relativamente compacto. Como todo subconjunto acotado de un espacio normado se encuentra en una bola, bastará demostrar que A^* transforma cada bola en un conjunto relativamente compacto. Debido a la linealidad del operador A^*, es suficiente demostrar que la imagen A^*S^* de la bola unitaria cerrada $S^* \subset E^*$ es relativamente compacta.

Consideremos los elementos de E^* como funciones definidas no sobre todo el espacio E, sino solamente sobre el compacto AS, que es la adherencia de la imagen de la bola unitaria por la aplicación A. Entonces, el conjunto Φ de funciones, correspondientes a las funcionales que pertenecen a S^*, será equiacotado y equicontinuo. En efecto, si $\|\varphi\| \leqslant 1$; tenemos

$$\sup_{x \in \overline{AS}} |\varphi(x)| = \sup_{x \in AS} |\varphi(x)| \leqslant \|\varphi\| \sup_{x \in S} \|Ax\| \leqslant \|A\|$$

y

$$|\varphi(x') - \varphi(x'')| \leqslant \|\varphi\| \cdot \|x' - x''\| \leqslant \|x' - x''\|.$$

Por consiguiente, el conjunto Φ es relativamente compacto en el espacio $C(\overline{AS})$ (en virtud del teorema de Arzelá). Pero el conjunto Φ, considerado con la métrica inducida por la métrica habitual del espacio de funciones continuas $C(\overline{AS})$, es isométrico al conjunto A^*S^* (con la métrica inducida por la norma del espacio E^*). En efecto, si $g_1, g_2 \in S^*$, se tiene

$$\|A^*g_1 - A^*g_2\| = \sup_{x \in S} |(A^*g_1 - A^*g_2, x)| =$$
$$= \sup_{x \in S} |(g_1 - g_2, Ax)| = \sup_{z \in AS} |(g_1 - g_2 z)| =$$
$$= \sup_{z \in \overline{AS}} |(g_1 - g_2, z)| = \rho(g_1, g_2).$$

Como Φ es relativamente compacto, es totalmente acotado; luego es también totalmente acotado el conjunto A^*S^* isométrico a él. Por esto A^*S^* es relativamente compacto en E^*. El teorema queda demostrado.

Observación. No es difícil comprobar que el conjunto Φ es cerrado en $C(\overline{AS})$, de manera que es compacto; por eso también es compacto el conjunto A^*S^*, aunque (como se deduce de la observación hecha en la pág. 254) la imagen por una aplicación totalmente continua arbitraria de la bola cerrada unitaria puede no ser un compacto. La situación en el teorema que acabamos de demostrar difiere de la general en que la bola cerrada unitaria S^* de E^* es compacta en la topología $*$-débil del espacio E^* (véase el teorema 3 del § 3). De aquí se deduce precisamente la compacidad (según la métrica del espacio E^*) de la imagen del conjunto S^* para cualquier operador totalmente continuo.

17—3427

EJERCICIOS. 1. Sea A un operador lineal acotado en un espacio de Banach. Demuéstrese que siendo el operador A^* totalmente continuo, el operador A es también totalmente continuo.

2. Para que un operador lineal A en un espacio de Hilbert H sea totalmente continuo es necesario y suficiente que su operador conjugado (de Hermite) A^* sea totalmente continuo.

3°. Valores propios de un operador totalmente continuo.

TEOREMA 4. *Todo operador totalmente continuo A en un espacio de Banach E tiene para cualquier $\rho > 0$ sólo un número finito de vectores propios linealmente independientes, correspondientes a valores propios, cuyos valores absolutos no son mayores que ρ.*

DEMOSTRACION. Observemos, ante todo, que el subespacio invariante E_λ, compuesto por todos los vectores propios del operador A que corresponden a un valor propio λ no nulo, es de dimensión finita. En efecto, si fuese E_λ de dimensión infinita, el operador A no sería totalmente continuo en el subespacio E_λ y, por consiguiente, en todo el E también. Por eso, para terminar la demostración del teorema bastará probar que, si $\{\lambda_n\}$ es una sucesión arbitraria de valores propios, diferentes dos a dos, de un operador totalmente continuo A, se tiene $\lambda_n \to 0$ para $n \to \infty$. A su vez, para ello es suficiente probar que no existe una sucesión infinita de valores propios $\{\lambda_n\}$, diferentes dos a dos, tal que la sucesión $\left\{\frac{1}{\lambda_n}\right\}$ sea acotada. Supongamos que existe tal sucesión y sea x_n el vector propio correspondiente al valor propio λ_n. Los vectores $x_1, x_2, \ldots$ son linealmente independientes[1]. Sea E_n $(n = 1, 2, \ldots)$ el subespacio generado por los vectores $x_1, \ldots, x_n$, es decir, sea E_n el conjunto de todos los elementos de tipo

$$y = \sum_{k=1}^{n} \alpha_k x_k.$$

Para cada $y \in E_n$ tenemos

$$y - \frac{1}{\lambda_n} Ay = \sum_{k=1}^{n} \alpha_k x_k - \sum_{k=1}^{n} \frac{\alpha_k \lambda_k}{\lambda_n} x_k = \sum_{k=1}^{n-1} \alpha_k \left(1 - \frac{\lambda_k}{\lambda_n}\right) x_k,$$

de donde se ve que

$$y - \frac{1}{\lambda_n} Ay \in E_{n-1}.$$

[1] La independencia lineal de los vectores correspondientes a diferentes valores propios de un operador, que actúa en un espacio normado, se demuestra igual que para los operadores en un espacio de dimensión finita (véase, por ejemplo, *Kurosch A. G., Curso de Álgebra Superior*. Editorial MIR, Moscú, 1968, pág. 213.)

Escojamos una sucesión $\{y_n\}$ de manera que

$$1)\ y_n \in E_n;\ 2)\ \|y_n\| = 1;\ 3)\ \rho(y_n, E_{n-1}) = \inf_{x \in E_{n-1}} \|y_n - x\| > \frac{1}{2}$$

(la existencia de una sucesión de este tipo ha sido demostrada en el lema de la pág. 250). Si la sucesión $\left\{\frac{1}{\lambda_n}\right\}$ es acotada, entonces, $\left\{\frac{y_n}{\lambda_n}\right\}$ es una sucesión acotada en E. Pero, al mismo tiempo, la sucesión $\left\{A\left(\frac{y_n}{\lambda_n}\right)\right\}$ no contiene ninguna subsucesión convergente, ya que para cualesquiera $p > q$

$$\left\|A\left(\frac{y_p}{\lambda_p}\right) - A\left(\frac{y_q}{\lambda_q}\right)\right\| = \left\|y_p - \left[y_p - \frac{1}{\lambda_p} Ay_p + A\left(\frac{y_q}{\lambda_q}\right)\right\| > \frac{1}{2},$$

puesto que $y_p - \frac{1}{\lambda} Ay_p + A\left(\frac{y_q}{\lambda_q}\right) \in E_{p-1}$. La contradicción obtenida demuestra el teorema.

4°. Operadores totalmente continuos en un espacio de Hilbert. En lo que precede hemos tratado de operadores totalmente continuos en un espacio de Banach arbitrario. Ahora completaremos nuestra exposición con algunos resultados referentes a operadores totalmente continuos en un espacio de Hilbert.

Hemos llamado un operador A totalmente continuo, cuando transforma todo conjunto acotado en uno relativamente compacto. Como $H = H^*$, es decir, H es el espacio dual a uno separable, todos los conjuntos acotados de él (y solamente ellos) son débilmente compactos. Por consiguiente, un operador totalmente continuo en un espacio de Hilbert puede definirse como un operador que transforma un conjunto débilmente compacto en un conjunto relativamente compacto según la topología fuerte. Finalmente, resulta cómoda, a veces, la siguiente definición de un operador totalmente continuo en un espacio de Hilbert: un operador A se llama totalmente continuo en H, cuando transforma toda sucesión débilmente convergente en una convergente fuertemente.

En efecto, supongamos que esta condición se cumple y sea M un conjunto acotado de H. Cada subconjunto infinito del conjunto M contiene una sucesión débilmente convergente. Si ésta se transforma en una sucesión fuertemente convergente, AM es compacto. Viceversa, sean A un operador totalmente continuo, $\{x_n\}$ una sucesión de convergencia débil y x su límite débil. Entonces, $\{Ax_n\}$ contiene una subsucesión que converge fuerte. Al mismo tiempo, debido a la continuidad de A, $\{Ax_n\}$ converge débilmente hacia Ax, de donde se sigue que $\{Ax_n\}$ no puede tener más de un punto de acumulación. Por consiguiente, $\{Ax_n\}$ es una sucesión convergente.

5°. Operadores autoconjugados y totalmente continuos en H. Para el caso de operadores lineales autoconjugados, que actúan en un espacio euclídeo de dimensión finita, se conoce el teorema sobre la reducción de la matriz de una transformación lineal de este tipo a la forma diagonal respecto a una base ortonormal. En este punto demostraremos un teorema que representa la generalización de este resultado al caso de operadores autoconjugados y totalmente continuos en un espacio de Hilbert. Los resultados de este punto son válidos tanto para el espacio de Hilbert real, como complejo. Para concretar, admitiremos que H es complejo.

Demostremos, ante todo, algunas propiedades de los vectores y valores propios de operadores autoconjugados en H, que son además, totalmente análogas a las propiedades correspondientes de operadores autoconjugados de dimensión finita.

I. Todos los valores propios de un operador A, autoconjugado y acotado en H, son reales.

En efecto, sea $Ax=\lambda x$, $\|x\|\neq 0$; entonces,

$$\lambda(x, x)=(Ax, x)=(x, Ax)=(x, \lambda x)=\bar{\lambda}(x, x),$$

de donde $\lambda=\bar{\lambda}$.

II. Los vectores propios de un operador autoconjugado y acotado, correspondientes a diferentes valores propios, son ortogonales.

Efectivamente, si $Ax=\lambda x$, $Ay=\mu y$ y $\lambda\neq\mu$, tenemos

$$\lambda(x, y)=(Ax, y)=(x, Ay)=(y, \mu y)=\mu(x, y),$$

de donde $(x, y)=0$.

Demostremos ahora el siguiente teorema fundamental.

TEOREMA 5 (Hilbert—Schmidt). *Para cualquier operador lineal A autoconjugado y totalmente continuo en un espacio de Hilbert H existe un sistema ortonormal $\{\varphi_n\}$ de vectores propios, correspondientes a los valores propios $\{\lambda_n\}$ tal que cada elemento $\xi\in H$ se puede escribir de manera única en la forma*

$$\xi=\sum c_k\varphi_k+\xi',$$

donde el vector ξ' verifica la condición $A\xi'=0$; además,

$$A\xi=\sum\lambda_k c_k\varphi_k$$

y

$$\lim\lambda_n=0 \quad (n\to\infty).$$

Para demostrar este teorema principal necesitaremos de las siguientes proposiciones auxiliares.

LEMA 1. *Si $\{\xi_n\}$ converge débilmente hacia ξ y el operador A es totalmente continuo, se tiene*

$$Q(\xi_n) = (A\xi_n, \xi_n) \to (A\xi, \xi) = Q(\xi).$$

DEMOSTRACION. Para cualquier n

$$|(A\xi_n, \xi_n) - (A\xi, \xi)| \leqslant |(A\xi_n, \xi_n) - (A\xi, \xi_n)| + |(A\xi, \xi_n) - (A\xi, \xi)|.$$

Pero

$$|(A\xi_n, \xi_n) - (A\xi, \xi_n)| \leqslant \|\xi_n\| \cdot \|A(\xi_n - \xi)\|$$

y

$$|(A\xi_n, \xi_n) - (A\xi, \xi)| = |(\xi, A(\xi_n - \xi))| \leqslant \|\xi\| \cdot \|A(\xi_n - \xi)\|$$

y, como los números $\|\xi_n\|$ son acotados y $\|A(\xi_n - \xi)\| \longrightarrow 0$, tenemos

$$|(A\xi_n, \xi_n) - (A\xi, \xi)| \longrightarrow 0,$$

que es lo que se quería demostrar.

LEMA 2. *Si una funcional*

$$|Q(\xi)| = |(A\xi, \xi)|,$$

donde A es un operador lineal autoconjugado y acotado, alcanza un máximo en el punto ξ_0 de la bola unitaria, entonces

$$(\xi_0, \eta) = 0$$

implica que

$$(A\xi_0, \eta) = (\xi_0, A\eta) = 0.$$

DEMOSTRACION. Es obvio que $\|\xi_0\| = 1$. Tomemos

$$\xi = \frac{\xi_0 + a\eta}{\sqrt{1 + |a|^2 \|\eta\|^2}},$$

donde a es un número complejo arbitrario. De $\|\xi_0\| = 1$ se sigue que

$$\|\xi\| = 1.$$

Como

$$Q(\xi) = \frac{1}{1 + |a|^2 \|\eta\|^2} [Q(\xi_0) + 2a(A\xi_0, \eta) + a^2 Q(\eta)],$$

para valores pequeños de a tenemos

$$Q(\xi) = Q(\xi_0) + 2a(A\xi_0, \eta) + O(a^2).$$

De la última igualdad se ve claramente que si $(A\xi_0, \eta) \neq 0$, se puede escoger a de manera que $|Q(\xi)| > |Q(\xi_0)|$ y esto contradice a la condición del lema.

Del lema 2 se deduce inmediatamente que si $|Q(\xi)|$ alcanza un máximo para $\xi=\xi_0$, entonces, ξ_0 es un vector propio del operador.

DEMOSTRACION DEL TEOREMA. Construiremos los vectores φ_k por inducción en el orden de decrecimiento de los valores absolutos de sus correspondientes valores propios

$$|\lambda_1| \geqslant |\lambda_2| \geqslant \ldots \geqslant |\lambda_n| \geqslant \ldots$$

Para construir el elemento φ_1 consideremos la expresión $|Q(\xi)| = |(A\xi, \xi)|$ y demostremos que alcanza un máximo sobre la bola unitaria. Sea

$$S = \sup_{\|\xi\| \leqslant 1} |(A\xi, \xi)|$$

y sea $\xi_1, \xi_2, \ldots$ una sucesión tal que $\|\xi_n\| = 1$ y

$$|(A\xi_n, \xi_n)| \longrightarrow S \text{ para } n \longrightarrow \infty.$$

Como la bola unitaria es débilmente compacta en H, se puede escoger de $\{\xi_n\}$ una subsucesión que converge débilmente hacia un elemento η. En este caso $\|\eta\| \leqslant 1$ y, en virtud del lema 1

$$|(A\eta, \eta)| = S.$$

Tomaremos por φ_1 el elemento η. Está claro que $\|\eta\|$ es exactamente igual a 1. (En efecto, sea $\eta = \eta_1$ y $\|\eta_1\| < 1$. Tomemos $\eta = \frac{\eta_1}{\|\eta_1\|}$; entonces, $\|\eta\| = 1$ y $|(A\eta, \eta)| > S$, lo que contradice a la definición de S.) Además,

$$A\varphi_1 = \lambda_1 \varphi_1,$$

de donde

$$|\lambda_1| = \frac{|(A\varphi_1, \varphi_1)|}{(\varphi_1, \varphi_1)} = |(A\varphi_1, \varphi_1)| = S.$$

Supongamos ahora que se han construido ya los vectores propios

$$\varphi_1, \varphi_2, \ldots, \varphi_n,$$

correspondientes a los valores propios

$$\lambda_1, \lambda_2, \ldots, \lambda_n.$$

Consideremos la funcional

$$|(A\xi, \xi)|$$

sobre el conjunto de elementos pertenecientes a

$$M_n' = H \ominus M(\varphi_1, \varphi_2, \ldots, \varphi_n)$$

(es decir, ortogonales a $\varphi_1, \varphi_2, \ldots, \varphi_n$), tales que $\|\xi\| < 1$. M_n' representa un subespacio invariante respecto de A (ya que

$M(\varphi_1, \varphi_2, \ldots, \varphi_n)$ es invariante y A es autoconjugado). Aplicando a M'_n los razonamientos anteriores, encontraremos en M'_n un vector propio del operador A; denotémoslo mediante φ_{n+1}.

Se puede dar dos casos: 1) después de un número finito de pasos obtendremos un subespacio M'_{n_0} en el cual $(A\xi, \xi) \equiv 0$; 2) $(A\xi, \xi) \not\equiv 0$ sobre M'_n para todo n.

En el primer caso, el lema 2 implica que el operador A transforma M'_{n_0} en cero, esto es, que M'_{n_0} consta solamente de los vectores propios correspondientes a $\lambda = 0$. El sistema construido de vectores $\{\varphi_n\}$ consta de un número finito de elementos.

En el segundo caso, obtendremos una sucesión $\{\varphi_n\}$ de vectores propios para cada uno de los cuales $\lambda_n \neq 0$. Probemos que $\lambda_n \to 0$. La sucesión $\{\varphi_n\}$ (como cualquier sucesión ortonormal) converge débilmente hacia el cero y por esto los elementos $A\varphi_n = \lambda_n\varphi_n$ deben converger, según la norma, hacia el cero, de donde $\lambda_n = \|A\varphi_n\| \to 0$.

Sea

$$M' = H \ominus M\{\varphi_n\} = \bigcap_n M'_n \neq 0.$$

Si $\xi \in M'$ y $\xi \neq 0$, tenemos

$$(A\xi, \xi) \leqslant \lambda_n \|\xi\|^2 \text{ para todo } n,$$

es decir,

$$(A\xi, \xi) = 0,$$

de donde, en virtud del lema 2 (para $\max|(A\xi, \xi)| = 0$) aplicado a M', obtenemos $A\xi = 0$, es decir, el operador A transforma el subespacio M' en cero.

De la construcción del conjunto $\{\varphi_n\}$, está claro que todo vector se puede representar en la forma

$$\xi = \sum c_k\varphi_k + \xi', \text{ donde } A\xi' = 0,$$

de donde se desprende que

$$A\xi = \sum \lambda_k c_k \varphi_k.$$

El teorema queda demostrado. Este teorema desempeña un papel fundamental en la teoría de Ecuaciones Integrales, de las cuales hablaremos en el capítulo X.

Observación. El teorema demostrado significa que para todo operador autoconjugado y totalmente continuo A de H existe una base ortogonal del espacio H compuesta por los vectores propios de este operador. En efecto, para obtener una base de este tipo bastará completar el sistema de vectores propios $\{\varphi_n\}$ construido en la demostración del teorema con una base ortogonal arbitraria del subespacio M' que es transformado por el operador A en el cero. En otras palabras, obtenemos aquí un resultado completamente análogo al teorema sobre la reducción de la matriz de un operador autoconjugado de dimensión finita a la forma diagonal en una base ortogonal.

Para los operadores no autoconjugados de un espacio n-dimensional esta reducción es, en general, imposible, sin embargo, es válido el siguiente teorema: *toda transformación lineal en un espacio n-dimensional tiene al menos un vector propio.* Es fácil ver que esta proposición no es extensible a operadores totalmente continuos en H. He aquí un ejemplo correspondiente. Consideremos en l_2 el siguiente operador A:

$$Ax = A(x_1, x_2, \ldots, x_n, \ldots) = \left(0, x_1, \frac{x_2}{2}, \ldots, \frac{x_{n-1}}{n-1}, \ldots\right).$$

Este operador no tiene ningún vector propio. En efecto, si

$$Ax = \lambda x,$$

se tiene

$$\lambda x_1 = 0, \ \lambda x_2 = x_1, \ \ldots, \ \lambda x_n = \frac{x_{n-1}}{n-1}, \ \ldots,$$

de donde $x_1 = x_2 = \ldots = x_n = \ldots = 0.$

CAPITULO V

ELEMENTOS DEL CALCULO DIFERENCIAL EN ESPACIOS LINEALES

En las cuestiones del Análisis Funcional que hemos tocado en los capítulos anteriores el papel principal correspondió a los conceptos de funcional lineal y operador lineal. Sin embargo, algunos problemas que surgen en el Análisis Funcional tienen un carácter sustancialmente no lineal e imponen la necesidad de desarrollar, junto al Análisis Funcional «lineal», el Análisis Funcional «no lineal», es decir, estudiar funcionales no lineales y operadores no lineales en espacios de dimensión infinita. Al Análisis Funcional no lineal pertenece, de hecho, una rama clásica de las Matemáticas que es el Cálculo de Variaciones, cuyos fundamentos fueron dados ya en los siglos XVII y XVIII en las obras de Bernoulli, Euler, Legendre y Jacobi. No obstante, el Análisis Funcional no lineal representa, en su conjunto, una rama relativamente moderna de las Matemáticas, aún muy lejos de su culminación. En este capítulo expondremos algunos conceptos primarios referentes al Análisis Funcional no lineal, principalmente, a la teoría de diferenciación, así como algunas aplicaciones de estos conceptos.

§ 1. DIFERENCIACION EN ESPACIOS LINEALES

1°. Diferencial fuerte (diferencial de Fréchet). Sean X e Y dos espacios normados y F una aplicación que actúa de X en Y y está definida sobre un subconjunto abierto O del espacio X. Diremos que esta aplicación es *diferenciable* en un punto dado $x \in O$, cuando existe un operador lineal acotado $L_x \in \mathscr{L}(X, Y)$ tal que

$$F(x+h)-F(x)=L_x(h)+\alpha(x, h), \tag{1}$$

donde
$$\frac{\|\alpha(x, h)\|}{\|h\|} \longrightarrow 0 \text{ para } \|h\| \longrightarrow 0. \tag{2}$$

La expresión $L_x(h)$ (que para cada $h \in X$ representa, evidentemente, un elemento del espacio Y) se llama *diferencial fuerte* (o *diferencial de Fréchet*) de la aplicación F en el punto x. El propio operador lineal L_x se llama *derivada*, más precisamente, *derivada fuerte* de la aplicación F en el punto x. Denotaremos esta derivada mediante el símbolo $F'(x)$.

Si la aplicación F es diferenciable en el punto x, la derivada correspondiente se determina de manera única. En efecto sea

$$F(x+h)-F(x) = L_x^{(1)}(h)+\alpha_1(x, h) = L_x^{(2)}(h)+\alpha_2(x, h);$$

entonces,
$$L_x^{(1)}(h) = L_x^{(2)}(h) = \alpha_2(x, h)-\alpha_1(x, h)$$

y, en virtud de (2),
$$\frac{\|L_x^{(1)}(h)-L_x^{(2)}(h)\|}{h} \longrightarrow 0 \text{ para } \|h\| \longrightarrow 0. \tag{3}$$

Pero, si para algún h se tiene
$$\frac{\|L_x^{(1)}(h)-L_x^{(2)}(h)\|}{h} = \lambda \neq 0,$$

tendremos para cualquier $\varepsilon \neq 0$
$$\frac{\|L_x^{(1)}(\varepsilon h)-L_x^{(2)}(\varepsilon h)\|}{h} = \lambda,$$

y la relación (3) no se cumple.

Señalemos ahora algunos resultados elementales que se deducen directamente de la definición de la derivada.

1. *Si* $F(x) = y_0 =$ const, *se tiene* $F'(x) \equiv 0$ *(es decir,* $F'(x)$ *es, en este caso, el operador nulo).*

2. *La derivada de una aplicación lineal continua* L *es esta misma aplicación.*

En efecto, tenemos, por definición,
$$L(x+h)-L(x) = L(h).$$

Menos obvio es el siguiente resultado importante.

3. *(Derivada de una función compuesta). Sean* X, Y *y* Z *tres espacios normados,* $U(x_0)$ *una vecindad del punto* $x_0 \in X$, F *una aplicación continua de esta vecindad en* Y, $y_0 = F(x_0)$, $V(y_0)$ *una vecindad del punto* $y_0 \in Y$ *y* G *una aplicación continua de esta vecindad en* Z. *Entonces, si la aplicación* F *es diferenciable en el punto* x_0 *y* G *es diferenciable en el punto* y_0, *la aplicación* $H = GF$

(que está definida y es continua en una vecindad del punto x_0) *es diferenciable en el punto* x_0 y

$$H'(x_0)=G'(y_0)F'(x_0). \tag{4}$$

Efectivamente, de acuerdo con las suposiciones hechas

$$F(x_0+\xi)=F(x_0)+F'(x_0)\xi+o_1(\xi)$$

y

$$G(y_0+\eta)=G(y_0)+G'(y_0)\eta+o_2(\eta).$$

Pero, $F'(x_0)$ y $G'(y_0)$ son operadores lineales acotados. Por eso

$$H(x_0+\xi)=G(y_0+F'(x_0)\xi+o_1(\xi))=$$
$$=G(y_0)+G'(y_0)(F'(x_0)\xi+o_1(\xi))+o_2(F'(x_0)\xi+o_1(\xi))=$$
$$=G(y_0)+G'(y_0)F'(x_0)\xi+o_3(\xi).$$

Siendo F, G y H funciones numéricas, la fórmula (4) se convierte en la conocida regla de diferenciación de una función compuesta.

4. *Sean F y G dos aplicaciones continuas que actúan de X en Y. Si F y G son diferenciables en el punto x_0, las aplicaciones $F+G$ y aF* (a es un número) *son también diferenciables en este punto y*

$$(F+G)'(x_0)=F'(x_0)+G'(x_0) \tag{5}$$

y

$$(aF)'(x_0)=aF'(x_0). \tag{6}$$

En efecto, de las definiciones de suma de operadores y de producto de un operador por un número, obtenemos inmediatamente que

$$(F+G)(x_0+h)=F(x_0+h)+G(x_0+h)=$$
$$=F(x_0)+G(x_0)+F'(x_0)h+G'(x_0)h+o_1(h)$$

y

$$aF(x_0+h)=aF(x_0)+aF'(x_0)h+o_2(h)$$

de donde se deducen las igualdades (5) y (6).

2°. Diferencial débil (diferencial de Gato). Sea de nuevo F una aplicación que actúa de X en Y. Se llama *diferencial débil*, o *diferencial de Gato*, de la aplicación F en el punto x al límite

$$DF(x,\ h)=\frac{d}{dt}F(x+th)\Big|_{t=0}=\lim_{t\to 0}\frac{F(x+th)-F(x)}{t},$$

donde la convergencia se entiende como la convergencia según la norma del espacio Y.

La diferencial débil $DF(x, h)$ puede no ser lineal respecto a h. Si esta linealidad tiene lugar, es decir, si

$$DF(x,\ h)=F'_c(x)h,$$

donde $F'_c(x)$ es un operador lineal, este operador se llama *derivada débil* (o *derivada de Gato*).

Señalemos que para las derivadas débiles no se cumple, como regla general, el teorema sobre la diferenciación de una función compuesta. (Dése un ejemplo).

3°. Fórmula de incremento finito. Supongamos que O es un conjunto abierto de X y que el segmento $[x_0, x]$ está contenido íntegramente en O. Sea, además, F una aplicación de X en Y, definida sobre O, que tiene derivada débil F'_c en cada punto del segmento $[x_0, x]$. Poniendo $\Delta x = x - x_0$ y tomando una funcional arbitraria $\varphi \in Y^*$, consideremos la función numérica

$$f(t) = \varphi(F_c(x_0 + t\,\Delta x)),$$

definida para $0 \leqslant t \leqslant 1$. Esta función es diferenciable respecto a t. Efectivamente, en la expresión

$$\frac{f(t+\Delta t) - f(t)}{\Delta t} = \varphi\,\frac{F(x_0 + t\,\Delta x + \Delta t\,\Delta x) - F(x_0 + t\,\Delta x)}{\Delta t}$$

se puede pasar al límite bajo el signo de la funcional lineal continua φ. Tendremos, entonces,

$$f'(t) = \varphi(F'_c(x_0 + t\,\Delta x)(\Delta x)).$$

Aplicando en el segmento $[0, 1]$ a la funcion f la fórmula de incremento finito, encontraremos

$$f(1) - f(0) = f'(\theta), \text{ donde } 0 \leqslant \theta \leqslant 1,$$

es decir,

$$\varphi(F(x) - F(x_0)) = \varphi(F'_c(x_0 + \theta\Delta x)(\Delta x)). \tag{7}$$

Esta relación tiene lugar para cualquier funcional $\varphi \in Y^*$ (el valor θ depende, claro está, de φ). De (7) obtenemos

$$|\varphi(F(x) - F(x_0))| \leqslant \|\varphi\| \sup_{0 \leqslant \theta \leqslant 1} \|F'_c(x_0 + \theta\Delta x)\| \cdot \|\Delta x\|. \tag{8}$$

Escojamos ahora una funcional no nula φ de manera que

$$\varphi(F(x) - F(x_0)) = \|\varphi\| \cdot \|F(x) - F(x_0)\|$$

(tal funcional φ existe en virtud del teorema de Hahn—Banach). Entonces, obtenemos de (8)

$$\|F(x) - F(x_0)\| \leqslant \sup_{0 \leqslant \theta \leqslant 1} \|F'_c(x_0 + \theta\Delta x)\| \cdot \|\Delta x\| \quad (\Delta x = x - x_0). \tag{9}$$

Esta desigual dad puede ser considerada como un análogo del teorema del valor medio para las funciones numéricas

Aplicando la desigualdad (9) a la aplicación

$$x \longrightarrow F(x) - F'_c(x_0)(\Delta x),$$

obtendremos la desigualdad siguiente:

$$\| F(x)-F(x_0)-F'_c(x_0)(\Delta x)\| \leqslant$$
$$\leqslant \sup_{0<\theta<1} \| F'_c(x_0+\theta\Delta x)-F'_c(x_0)\| \cdot \|\Delta x\|. \quad (10)$$

4°. Relación entre las diferenciabilidades débil y fuerte. Las diferenciabilidades débil y fuerte constituyen conceptos diferentes incluso en el caso de espacios de dimensión finita. Efectivamente, es bien conocido del Análisis que para una función numérica

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

la existencia de la derivada

$$\frac{d}{dt} f(x+th)$$

para cualquier $h=(h_1, \ldots, h_n)$ fijo no implica aún, en el caso de $n \geqslant 2$, la diferenciabilidad de esta función, es decir, la posibilidad de representar su incremento $f(x+h)-f(x)$ como la suma de una parte lineal (respecto a h) y un miembro infinitésimo de orden superior al primero respecto a $|h|$.

Como ejemplo elemental, puede servir aquí la función de dos variables

$$f(x_1, x_2)=\begin{cases} x_1+x_2+\dfrac{x_1^3x_2}{x_1^4+x_2^2}, & \text{cuando } x_1^2+x_2^2\neq 0, \\ 0, \text{ cuando } x_1=x_2=0. \end{cases} \quad (11)$$

Esta función es continua en todo el plano, incluido el punto (0, 0). Tiene diferencial débil en el punto (0, 0), ya que

$$\lim_{t\to 0} \frac{f(0+th)-f0}{t}=\lim_{t\to 0}\left(h_1+h_2+\frac{t^4h_1^3h_2}{t^4h_1^4+t^2h_2^2}\right)=h_1+h_2.$$

Sin embargo, esta diferencial no constituye la parte lineal principal del incremento de la función (11) en el punto (0, 0). En efecto, sea

$$\omega(0, h)=f(0+h)-f(0)-(h_1+h_2)=\frac{h_1^3h_2}{h_1^4+h_2^2}.$$

Entonces, tomando $h_2=h_1^2$, tendremos

$$\lim_{\|h\|\to 0}\frac{\omega(0, h)}{\|h\|}=\lim_{h_1\to 0}\frac{h_1^5}{2h_1^4\sqrt{h_1^2+h_1^4}}=\frac{1}{2}\neq 0.$$

Al mismo tiempo, si una aplicación F es diferenciable en el sentido fuerte, es diferenciable también débilmente y, además, las diferenciales fuerte y débil coinciden. Efectivamente, para una aplicación fuertemente diferenciable, tenemos

$$F(x+th)-F(x)=F'(x)(th)+o(th)=tF'(x)(h)+o(th)$$

y

$$\frac{F(x+th)-F(x)}{t}=F'(x)(h)+\frac{o(th)}{t}\longrightarrow F'(x)(h).$$

Busquemos las condiciones en las cuales la diferenciabilidad débil de una aplicación F implica su diferenciabilidad fuerte.

TEOREMA 1. *Si la derivada débil $F'_c(x)$ de la aplicación F existe en una vecindad $U(x_0)$ del punto x_0 y representa en esta vecindad una función continua (operadora) de x, la derivada fuerte $F'(x_0)$ existe en el punto x_0 y coincide con la débil.*

DEMOSTRACION. Por hipótesis, la aplicación F tiene derivada débil, esto es, $DF(x_0, h)=F'_c(x_0)h$. Escojamos h de manera que $x_0+h\in U(x_0)$ y consideremos la expresión

$$\omega(x_0, h)=F(x_0+h)-F(x_0)-F'_c(x_0)h. \tag{12}$$

Si e es ahora un elemento arbitrario del espacio y^* dual a Y, obtenemos de (12)

$$(\omega(x_0, h), e)=(F(x_0+h)-F(x_0), e)-(F'_c(x_0)h, e). \tag{13}$$

Consideremos la función $f(t)=(F(x_0+th), e)$ de argumento numérico t. Esta función es diferenciable respecto a t y para ella

$$\frac{df}{dt}=\lim_{\Delta t\to 0}\left(\frac{F(x_0+th+\Delta th)-F(x_0+th)}{\Delta t}, e\right)=(F'_c(x_0+th)h, e).$$

Por eso, aplicando a f la fórmula de incremento finito, podemos escribir la igualdad (13) en la forma

$$(\omega(x_0, h), e)=([F'_c(x_0+\tau h)-F'_c(x_0)]h, e), \tag{13'}$$

donde $0\leqslant\tau\leqslant 1$. Para un h fijo, el elemento $e\in Y^*$ se puede escoger de manera que $\|e\|=1$ y que se cumpla la desigualdad

$$|(\omega(x_0, h), e|\geqslant\frac{1}{2}\|\omega(x_0, h)\|\cdot\|e\|=\frac{1}{2}\|\omega(x_0, h)\|.$$

De aquí y de la igualdad (13') encontramos que

$$\|\omega(x_0, h)\|\leqslant 2\|F'_c(x_0+\tau h)-F'_c(x_0)\|\cdot\|h\|.$$

Pero $F'_c(x)$ es, por hipótesis, una función operadora continua de x; por eso,

$$\lim_{h\to 0}\|F'_c(x_0+\tau h)-F'_c(x_0)\|=0,$$

de manera que $\|\omega(x_0, h)\|$ es una infinitésima de orden superior al primero respecto a $\|h\|$, es decir, $F'_c(x_0)h$ constituye la parte principal de la diferencia $F(x_0+h)-F(x_0)$. Con esto queda demostrado tanto la existencia de la derivada fuerte $F'(x_0)$ como su coincidencia con la derivada débil. En lo sucesivo consideraremos, siempre que no se diga lo contrario, aplicaciones diferenciables en el sentido fuerte y, por consiguiente, también en el débil.

5°. Funcionales diferenciables. Hemos introducido el concepto de diferencial de una aplicación F que actúa de un espacio normado X en otro espacio normado Y. La derivada $F'(x)$ de esta aplicación representa para cada x un operador lineal de X en Y, esto es, un elemento del espacio $\mathscr{L}(X, Y)$. En particular, si Y es la recta numérica, F es una función sobre X que toma valores numéricos, es decir, una funcional. En este caso, la derivada de la funcional F en el punto x_0 es una funcional lineal (que depende de x_0), es decir, un elemento del espacio X^*.

Ejemplo. Consideremos en el espacio de Hilbert real H la funcional $F(x)=\|x\|^2$. Entonces,

$$\|x+h\|^2-\|x\|^2=2(x, h)+\|h\|^2;$$

la expresión $2(x, y)$ constituye la parte lineal principal de esta expresión y, por consiguiente[1],

$$F'(x)=F'_c(x)=2x.$$

EJERCICIO. Calcúlese la derivada de la funcional $\|x\|$. $\left(\text{Respuesta: } \frac{x}{\|x\|}\right.$ para $x \neq 0$; para $x=0$ no existe$\left.\right)$.

6°. Funciones abstractas. Supongamos ahora que el espacio de argumentos X coincide con la recta numérica. La aplicación $F(x)$ que pone en correspondencia al número x un elemento de un espacio de Banach Y se llama *función abstracta*. La derivada de una función abstracta $F'(x)$ (si es que existe) representa (para cada x) un elemento del espacio Y. Para una función abstracta (que representa una función de un argumento numérico) la diferenciabilidad débil coincide con la fuerte.

7°. Integral. Sea F una función abstracta de argumento real t con valores en un espacio de Banach Y. Si F está definida sobre un segmento $[a, b]$, se puede definir la integral de la

[1] Basándonos en el teorema sobre la expresión general de una funcional lineal continua en un espacio de Hilbert, identificamos aquí las funcionales de H^* con los elementos correspondientes de H.

función F en el segmento $[a, b]$. Esta integral se comprende como el límite de las sumas integrales

$$\sum_{k=0}^{n-1} F(\xi_k)(t_{k+1}-t_k)$$

correspondientes a las particiones

$$a=t_0<t_1<\ldots<t_n=b, \quad \xi_k \in [t_k, t_{k+1}]$$

cuando $\max |t_{k+1}-t_k| \to 0$. Esta integral (que representa, evidentemente, un elemento de Y) se denota mediante el símbolo

$$\int_a^b F(t)\,dt.$$

Razonamientos, análogos a los empleados para funciones que toman valores numéricos, demuestran que la integral de una función continua sobre un segmento existe; además, ella tiene propiedades análogas a las propiedades de la integral corriente de Riemann. Señalemos entre estas propiedades las siguientes.

1. Si U es una aplicación lineal continua fija del espacio Y en un espacio Z, se tiene

$$\int_a^b UF(t)\,dt = U\int_a^b F(t)\,dt.$$

2. Si $F(t)$ es de la forma $f(t)y_0$, donde f es una función numérica e y_0 un elemento fijo de Y, se tiene

$$\int_a^b F(t)\,dt = y_0\int_a^b f(t)\,dt.$$

3. $$\left\| \int_a^b F(t)\,dt \right\| \leqslant \int_a^b \| F(t) \|\,dt.$$

Sean X e Y de nuevo espacios normados y sea $BC(X, Y)$ el espacio lineal de todas las aplicaciones continuas acotadas[1] de X en Y. En el espacio $BC(X, Y)$ se puede introducir una topología tomando por vecindades del cero los conjuntos

$$U_{n,\,\varepsilon} = \{F: \sup_{\|x\| \leqslant n} \|F(x)\| < \varepsilon\}.$$

Esta topología coincide en el subespacio $\mathscr{L}(X, Y) \subset BC(X, Y)$

[1] Una aplicación $F: X \to Y$ se llama *acotada*, cuando para todo conjunto acotado $Q \subset X$ el conjunto $F(Q)$ es acotado en Y. Una aplicación continua no lineal no es necesariamente acotada.

de todas las aplicaciones lineales continuas de X en Y con la topología corriente de $\mathscr{L}(X, Y)$ definida por la norma de operador. Sea $J=[x_0, x_0+\Delta x]$ un segmento rectilíneo de X. Supongamos dada una aplicación continua de este segmento en el espacio $BC(X, Y)$, es decir, supongamos que a cada punto $x \in j$ se ha asignado una aplicación $F(x) \in BC(X, Y)$ que depende continuamente del parámetro vectorial $x \in J$. Entonces, se puede definir la integral de $F(x)$ en el segmento J, tomando

$$\int_{x_0}^{x_0+\Delta} F(x)\,dx = \int_0^1 F(x_0+t\Delta x)\,\Delta x\,dt \tag{14}$$

(aquí $F(x_0+t\Delta x)\,\Delta x$ es para cada $t \in [0, 1]$ un elemento del espacio Y, precisamente la imagen del elemento $\Delta x \in X$ mediante la aplicación $F(x_0+t\Delta x)$. Está claro que la integral que figura en el miembro derecho de la fórmula (14) existe y representa un elemento del espacio Y.

Apliquemos estas ideas al problema de reconstrucción de una aplicación a partir de su derivada.

Consideremos una aplicación F que actúa de X en Y y que tiene en el segmento $[x_0, x_0+\Delta x]$ derivada fuerte *continua* $F'(x) \in \mathscr{L}(X, Y)$. Entonces, existe la integral $\int_{x_0}^{x_0+\Delta x} F'(x)\,dx$. Demostremos que tiene lugar la igualdad

$$\int_{x_0}^{x_0+\Delta x} F'(x)\,dx = F(x_0+\Delta x) - F(x_0) \tag{15}$$

que generaliza la fórmula de Newton-Leibniz. En efecto, por definición,

$$\int_{x_0}^{x_0+\Delta x} F'(x)\,dx = \lim_{\lambda \to 0} \sum_{k=0}^{n-1} F'(x_0+t_k\Delta x)(\Delta x)(t_{k+1}-t_k) =$$
$$= \lim_{\lambda \to 0} \sum_{k=0}^{n-1} F'(\tilde{x}_k)(\Delta x_k),$$

donde $\tilde{x}_k = x_0+t_k\Delta x$, $\Delta x_k = (t_{k+1}-t_k)\,\Delta x$ y $\lambda = \max\limits_k (t_{k+1}-t_k)$. Pero, al mismo tiempo, para cualquier partición del segmento $0 \leqslant t \leqslant 1$ tenemos

$$F(x_0+\Delta x) - F(x_0) = \sum_{k=0}^{n-1} [F(x_0+t_{k+1}\Delta x) - F(x_0+t_k\Delta x] =$$
$$= \sum_{k=0}^{n-1} [F(x_{k+1}) - F(x_k)].$$

De la fórmula (10) de incrementos finitos, encontramos

$$\left\| \sum_{k=0}^{n-1} [F(x_{k+1}) - F(x_k) - F'(\tilde{x}_k)\,\Delta x_k] \right\| \leqslant$$

$$\leqslant \| \Delta x \| \sum_{k=0}^{n-1} (t_{k+1} - T_k) \sup \| F'(x_k + \theta \Delta x_k) - F'(\tilde{x}_k) \|. \quad (16)$$

Como la derivada F' es continua y, por consiguiente, también uniformemente continua sobre el segmento $[x_0, \ x_0 + \Delta x]$, el miembro derecho de la desigualdad (16) tiende a cero, cuando disminuyen indefinidamente las longitudes de los elementos de la partición del segmento $[x_0, \ x_0 + \Delta x]$, y de aqui se sigue la igualdad (15).

8°. **Derivadas de órdenes superiores.** Sea F una aplicación diferenciable que actúa de X en Y. Su derivada $F'(x)$ es, para cada $x \in X$, un elemento de $\mathscr{L}(X, Y)$, es decir, F' es una aplicación del espacio X en el espacio de operadores lineales $\mathscr{L}(X, Y)$. Si esta aplicación es diferenciable, la derivada correspondiente a ella se llama *segunda derivada* de la aplicación F y se denota mediante el símbolo F''. De manera que $F''(x)$ es un elemento del espacio $\mathscr{L}(X, \mathscr{L}(X, Y))$ de operadores lineales que actúan de X en $\mathscr{L}(X, Y)$. Probemos que los elementos de este espacio admiten una interpretación más cómoda y más clara a partir de las así llamadas aplicaciones bilineales.

Decimos que se tiene una *aplicación bilineal* B del espacio X en el espacio Y, cuando a cada par ordenado de elementos x, x' de X corresponde un elemento $y = B(x, x') \in Y$ de manera que se cumplen las siguientes condiciones:

1) para cualesquiera x_1, x_2, x_1', x_2' de X y cualesquiera números α, β se verifican las igualdades:

$$B(\alpha x_1 + \beta x_2, \ x') = \alpha B(x_1, \ x') + \beta B(x_2, \ x'),$$
$$B(x, \ \alpha x_1' + \beta x_2') = \alpha B(x, \ x_1') + \beta B(x, \ x_2');$$

2) existe un número positivo M tal que

$$\| B(x, \ x') \| \leqslant M \| x \| \cdot \| x' \| \quad (17)$$

para todos los $x, x' \in X$.

En otras palabras, la primera de estas condiciones significa que la aplicación B es lineal respecto a cada uno de sus dos argumentos; no es difícil comprobar que la segunda condición equivale a la continuidad de B respecto al conjunto de argumentos. El menor de los números M que satisface la condición (17) se llama *norma* de la aplicación bilineal B y se denota con $\| B \|$. De una manera evidente se definen las operaciones lineales para aplicaciones bilineales que tienen las propiedades habituales.

De esta forma, las aplicaciones bilineales del espacio X en el espacio Y constituyen un espacio lineal normado que denotaremos con $B(X^2, Y)$.

A cada elemento A del espacio $\mathscr{L}(X, \mathscr{L}(X, Y))$ se puede poner en correspondencia un elemento de $B(X^2, Y)$, tomando

$$B(x, x') = (Ax)x'. \tag{18}$$

Es obvio que esta correspondencia es lineal. Probemos que es, además, isométrica y transforma el espacio $\mathscr{L}(X, \mathscr{L}(X, Y))$ en todo el espacio $B(X^2, Y)$. En efecto, si $y = B(x, x') = (Ax)x'$, tenemos

$$\| y \| \leqslant \| Ax \| \cdot \| x' \| \leqslant \| A \| \cdot \| x \| \cdot \| x' \|,$$

de donde

$$\| B \| \leqslant \| A \|. \tag{19}$$

Por otro lado, dada una aplicación bilineal B, la aplicación $x' \to (Ax)x' = B(x, x')$ es, para un $x \in X$ fijo, una aplicación lineal del espacio X en Y.

Por consiguiente, a cada $x \in X$ se pone en correspondencia un elemento Ax del espacio $\mathscr{L}(X, Y)$; es obvio que Ax depende linealmente de x, es decir, que la aplicación bilineal B define un elemento A del espacio $\mathscr{L}(X, \mathscr{L}(X, Y))$. Además, está claro que la aplicación B se reconstruye a partir de A mediante la fórmula (18) y que

$$\| Ax \| = \sup_{\| x' \| \leqslant 1} \| (Ax)x' \| = \sup_{\| x' \| \leqslant 1} \| B(x, x') \| \leqslant \| B \| \cdot \| x \|,$$

de donde

$$\| A \| \leqslant \| B \|. \tag{20}$$

Comparando (19) y (20), obtenemos $\| A \| = \| B \|$. De modo que la correspondencia entre $B(X^2, Y)$ y $\mathscr{L}(X, \mathscr{L}(X, Y))$ definida por la igualdad (18) es lineal e isométrica y, por consiguiente, biunívoca. Además, la imagen del espacio $\mathscr{L}(X, \mathscr{L}(X, Y))$ es todo el espacio $B(X^2, Y)$.

Hemos visto que la segunda derivada $F''(x)$ es un elemento del espacio $\mathscr{L}(X, \mathscr{L}(X, Y))$. De acuerdo con lo expuesto podemos considerar que $F''(x)$ es un elemento del espacio $B(X^2, Y)$.

Veamos algunos ejemplos. Sean X e Y espacios euclídeos de dimensión finita, m y n respectivamente. Entonces, toda aplicación lineal de X en Y se puede definir mediante una $(m \times n) =$ $=$ matriz. De manera que la derivada $F'(x)$ de la aplicación F, que actúa de X en Y, es una matriz (dependiente de $x \in X$). Si en X e Y se escogen unas bases, digamos,

$$e_1, \ldots, e_m \text{ en } X \text{ y } f_1, \ldots, f_n \text{ en } Y,$$

18*

tendremos

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_m e_m,$$
$$y = y_1 f_1 + y_2 f_2 + \ldots + y_n f_n$$

y en este caso

$$F'(x) = \left\{ \begin{matrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_1} \\ \cdot & \cdot & \cdot & \cdot \\ \frac{\partial y_1}{\partial x_m} & \frac{\partial y_2}{\partial x_m} & \cdots & \frac{\partial y_n}{\partial x_m} \end{matrix} \right\}.$$

La segunda derivada $F''(x)$ se determina por un conjunto de $m \times m \times n$ valores $a_{ij}^k = \frac{\partial^2 y_k}{\partial x_i \partial x_j}$. Este conjunto de valores a_{ij}^k puede ser considerado o bien como una aplicación lineal del espacio X en el espacio $\mathscr{L}(X, Y)$ definida por

$$b_j^k = \sum_{i=1}^{m} a_{ij}^k x_i$$

o bien como una aplicación bilineal del espacio X en Y definida mediante la fórmula

$$c^k = \sum_{i, j=1}^{m} a_{ij}^k x_i x_j'.$$

De una manera análoga se puede introducir el concepto de tercera, cuarta y, en general, n-ésima derivada de la aplicación F, que actúa de X en Y, definiendo la n-ésima derivada como la derivada de la derivada de orden $(n-1)$. Es obvio que la n-ésima derivada constituye un elemento del espacio $\mathscr{L}(X, \mathscr{L}(X, \ldots, \mathscr{L}(X, Y)))$. Repitiendo los razonamientos, empleados para la segunda derivada, se puede asignar de un modo natural a cada elemento de este espacio un elemento del espacio $N(X^n, Y)$ de las aplicaciones n-lineales de X en Y. Por una *aplicación n-lineal* se entiende aquí una correspondencia $y = N(x', x'', \ldots, x^{(n)})$ entre los sistemas ordenados $(x', x'', \ldots, x^{(n)})$ de elementos de X y los elementos del espacio Y que es lineal respecto a cada x^i, cuando son fijos los elementos $x', \ldots, x^{(i-1)}, x^{(i+1)}, \ldots, x^{(n)}$, y que verifica para un $M > 0$ determinado la condición

$$\| N(x', x'', \ldots, x^{(n)}) \| \leqslant M \| x' \| \cdot \| x'' \| \ldots \| x^{(n)} \|.$$

Por consiguiente, la n-ésima derivada de la aplicación F se puede considerar como un elemento del espacio $N(X^n, Y)$.

9°. Diferenciales de orden superior. Hemos definido la diferencial (fuerte) de una aplicación F como el resultado de aplicar al elemento $h \in X$ el operador lineal $F'(x)$: $dF = F'(x)(h)$. La diferencial de segundo orden se define como $d^2F = F''(x)(h, h)$, es decir, como una expresión cuadrática correspondiente a la aplicación $F''(x) \in B(X^2, Y)$. De un modo análogo, la diferencial de orden n se define mediante $d^nF = F^{(n)}(x)(h, h, \ldots, h)$, esto es, como aquel elemento del espacio Y en el cual se transforma por la aplicación $F^{(n)}(x)$ el elemento $(h, h, \ldots, h) \in X \times X \times \ldots \times X = X^n$.

10°. **Fórmula de Taylor.** La diferenciabilidad fuerte de la aplicación F significa que la diferencia $F(x+h) - F(x)$ se puede representar como la suma de un miembro lineal y un sumando de orden superior al primero respecto a $\|h\|$. Este resultado se generaliza en una fórmula análoga a la fórmula de Taylor para las funciones numéricas, conocida del Análisis.

TEOREMA 2. *Sea F una aplicación que actúa de X en Y, que está definida en una región $O \subset X$ y tal que $F^{(n)}(x)$ existe y representa una función uniformemente continua de x en O. Entonces, tiene lugar la igualdad*

$$F(x+h) - F(x) = F'(x)(h) + \frac{1}{2!}F''(x)(h, h) + \ldots$$
$$\ldots + \frac{1}{n!}F^{(n)}(x)(h, \ldots, h) + \omega(x, h) \qquad (21)$$

donde $\|\omega(x, h)\| = o(\|h\|^n)$.

La DEMOSTRACION se realiza por inducción. Para $n=1$ la igualdad (21) es trivial Supongamos que ella es válida para $n-1$ cualquiera que sea la aplicación que satisface las condiciones del teorema. Entonces, para la aplicación F' tenemos

$$F'(x+h) = F'(x) + F''(x)(h) + \frac{1}{2!}F'''(x)(h, h) +$$
$$+ \ldots + \frac{1}{(n-1)!}F^{(n)}(x)(h, h, \ldots, h) + \omega_1(x, h), \qquad (22)$$

donde $\|\omega_1(x, h)\| = o(\|h\|^{n-1})$. Integrando en el segmento $[x, x+h]$ ambos miembros de la igualdad (22) y empleando la fórmula (15) de Newton—Leibniz, encontraremos

$$F(x+h) - F(x) = \int_0^1 F'(x+th)h\,dt = \int_0^1 \Big\{F'(x) + tF''(x)(h) +$$
$$+ \frac{1}{2!}t^2F'''(x)(h, h) + \ldots$$
$$\ldots + \frac{1}{(n-1)!}t^{n-1}F^{(n)}(x)(h, \ldots, h)\Big\} h\,dt + R_n, \qquad (23)$$

donde $R_n = \int_0^1 \omega_1(x, th)h\,dt$.

De (23) obtenemos

$$F(x+h)-F(x)=F'(x)(h)+\frac{1}{2!}F''(x)(h,\ h)+\ldots$$
$$\ldots+\frac{1}{n!}F^{(n)}(x)(h,\ \ldots,\ h)+R_n,$$

siendo

$$\| R_n\| \leqslant \int_0^1 \|\omega_1(x,\ th)\|\cdot\|h\|\,dt=o(\|h\|^n).$$

Con esto nuestra proposición queda demostrada.

La fórmula (21) se llama fórmula de Taylor para aplicaciones.

§ 2. PROBLEMAS EXTREMALES

Una de las secciones más antiguas y más elaboradas del Análisis Funcional no lineal es la búsqueda de extremos de funcionales. El estudio de estos problemas constituye el contenido del así llamado Cálculo de variaciones. Los métodos que se utilizan en el Cálculo de variaciones están sujetos, en su mayor parte a la forma especial de aquellas funcionales cuyos valores extremales se buscan. Sin embargo, se puede enunciar algunos resultados y métodos generales para funcionales más o menos arbitrarios. Sin plantearnos la tarea de dar una exposición un tanto detallada de los métodos variacionales, nos limitaremos a dar un examen breve de aquellos elementos de la teoría general de problemas para funcionales que constituyen el fundamento del Cálculo de variaciones.

1°. Condición necesaria de extremo. Sea F una funcional que toma valores reales, definida en un espacio de Banach X. Se dice que la funcional F alcanza un mínimo en el punto x_0, cuando para todos los x, suficientemente próximos a x_0 y tales que $F(x)$ está definido, se cumple la desigualdad $F(x)-F(x_0)\geqslant 0$. De manera análoga se define un máximo de una funcional. Si en un punto dado x_0 la funcional F alcanza mínimo o máximo, diremos que la funcional tiene en este punto extremo.

Diferentes problemas mecánicos y físicos pueden ser reducidos a la búsqueda del extremo de unas u otras funcionales.

Para las funciones de n variables es bien conocida la siguiente condición necesaria de extremo: si la función f es diferenciable en el punto $x_0=(x_1^0,\ x_2^0,\ \ldots,\ x_n^0)$ y tiene extremo en este punto, en este punto $df=0$ ó, lo que es equivalente,

$$\frac{\partial f}{\partial x_1}=\frac{\partial f}{\partial x_2}=\ldots=\frac{\partial f}{\partial x_n}=0.$$

Esta condición se extiende fácilmente a las funcionales.

TEOREMA 1. *Para que una funcional diferenciable F alcance extremo en el punto x_0 es necesario que su diferencial en este punto sea igual a cero para todo h:*

$$F'(x_0)(h) \equiv 0.$$

DEMOSTRACION. Por definición de la diferenciabilidad, tenemos

$$F(x_0+h) - F(x_0) = F'(x_0)(h) + o(\|h\|). \qquad (1)$$

Si $F'(x_0)(h) \neq 0$ para algún h, entonces, para valores reales suficientemente pequeños de λ, el signo de toda la expresión $F'(x_0)(\lambda h) + o(\|h\|)$ coincide con el signo de su término principal $F'(x_0)(\lambda h)$. Pero $F'(x_0)$ es una funcional lineal y por eso $F'(x_0)(\lambda h) = \lambda F'(x_0)(h)$. De manera que siendo $F'(x_0)(h) \neq 0$, la expresión (1) puede tomar, para h arbitrariamente pequeños, tanto valores positivos, como negativos, es decir, no puede haber extremo en el punto x_0.

Veamos algunos ejemplos.

1. Sea

$$F(x) = \int_a^b f(t, x(t))\,dt, \qquad (2)$$

donde f es una función continuamente diferenciable. Esta funcional, considerada en el espacio $C_{[a, b]}$ de funciones continuas, es diferenciable. En efecto,

$$F(x+h) - F(x) = \int_a^b [f(t, x+h) - f(t, x)]\,dt = \int_a^b f'_x = (t, x)\,h(t)\,dt + o(\|h\|),$$

de donde

$$dF = \int_a^b f'_x(t, x(t))\,h(t)\,dt.$$

La igualdad a cero de esta funcional lineal para todos los $h \in C_{[a, b]}$ significa que $f'_x(t, x) = 0$. Efectivamente, para todo $x(t) \in C_{[a, b]}$ la derivada $f'_x(t, x)$ es una función continua de t. Si ella es diferente de cero en algún punto t_0, digamos, $f'_x(t_0, x(t_0)) > 0$, esta igualdad tendrá lugar también en una vecindad (α, β) del punto t_0. Entonces, tomando

$$h(t) = \begin{cases} (t-\alpha)(\beta-t) & \text{para } \alpha \leqslant t \leqslant \beta, \\ 0 & \text{para los demás } t, \end{cases}$$

obtendremos que

$$\int_a^b f'_x(t, x)\, h(t)\, dt > 0.$$

La contradicción obtenida demuestra nuestra proposición. La ecuación $f'_x(t, x) = 0$ determina, en general, una curva en la cual la funcional (2) puede alcanzar un extremo.

2. Consideremos en el mismo espacio $C_{[a, b]}$ la funcional

$$F(x) = \int_a^b \int_a^b K(\xi_1, \xi_2)\, x(\xi_1)\, x(\xi_2)\, d\xi_1\, d\xi_2, \tag{3}$$

donde $K(\xi_1, \xi_2)$ es una función continua que satisface la condición $K(\xi_1, \xi_2) = K(\xi_2, \xi_1)$. Es fácil calcular que la diferencial de esta funcional es igual a

$$dF = 2 \int_a^b \int_a^b K(\xi_1, \xi_2)\, x(\xi_1)\, h(\xi_2)\, d\xi_1\, d\xi_2.$$

Si esta expresión es igual a cero para todo $h \in C_{[a, b]}$, tenemos, por los mismos razonamientos que en el ejemplo 1,

$$\int_a^b K(\xi_1, \xi_2)\, x(\xi_1)\, d\xi_1 = 0 \quad \text{para todo} \quad \xi_2,\ a \leqslant \xi_2 \leqslant b.$$

Una de las soluciones de esta ecuación es la función $x \equiv 0$. La respuesta a la pregunta de si existe extremo en este punto y si existen otros puntos en los que es posible un extremo, depende de la forma de la función $K(\xi_1, \xi_2)$ y exige un estudio complementario.

3. Consideremos la funcional

$$F(x) = \int_a^b f(t, x(t), x'(t))\, dt, \tag{4}$$

definida en el espacio $C^1_{[a, b]}$ de funciones continuamente diferenciables sobre el segmento $[a, b]$. Aquí $x'(t) = \frac{dx(t)}{dt}$, y $f(t, x, x')$ es una función dos veces diferenciable de sus argumentos. La funcional (4) desempeña un papel principal en varias cuestiones del Cálculo de variaciones. Busquemos su diferencial. Utilizando la fórmula de Taylor, encontramos

$$F(x+h)-F(x)=\int_a^b [f(t,\ x+h,\ x'+h')-f(t,\ x,\ x')]\,dt=$$

$$=\int_a^b (f'_x h+f'_{x'}h')\,dt+o(\|h\|),$$

donde $\|h\|$ es la norma de la función h como elemento del espacio $C^1_{[a,\ b]}$. Por consiguiente, la condición necesaria de extremo de la funcional (4) es

$$dF=\int_a^b (f'_x h+f'_{x'}h')\,dt=0. \tag{5}$$

En su forma integral esta condición es poco útil para buscar la función x en la que se alcanza el extremo. Démosle una forma más cómoda, integrando por partes en (5) el término $f'_{x'}h'$. Tendremos

$$\int_a^b f'_{x'}h'\,dt=f'_{x'}h\Big|_a^b-\int_a^b h\frac{d}{dt}f'_{x'}\,dt.$$

De manera que

$$dF=\int_a^b \left(f'_x=-\frac{d}{dt}f'_{x'}\right)h\,dt+f'_{x'}h\Big|_a^b=0. \tag{6}$$

Esta igualdad debe verificarse para todo h, en particular, también cuando $h(a)=h(b)=0$. Por consiguiente,

$$\int_a^b \left(f'_x-\frac{d}{dt}f'_{x'}\right)h\,dt=0$$

para todos los h tales que $h(a)=h(b)=0$, de donde, con razonamientos análogos a los empleados en el ejemplo 1, encontramos

$$f'_x-\frac{d}{dt}f'_{x'}=0. \tag{7}$$

Por eso, la igualdad (6) se reduce a

$$f'_{x'}h\,\Big|_a^b=0. \tag{8}$$

Si la funcional (4) se considera para todas las funciones x continuamente diferenciables definidas sobre $[a,\ b]$, podemos escoger h de modo que $h(a)=0$, $h(b)\neq 0$ y entonces obtenemos de la igualdad (8)

$$f'_{x'}\big|_{t=b}=0; \tag{9}$$

por otro lado, tomando $h(b)=0$, $h(a)\neq 0$, obtenemos

$$f'_{x'}\big|_{t=a}=0. \tag{10}$$

Por consiguiente, de la condición (6) de igualdad a cero de la diferencial de la funcional (4) hemos obtenido que la función x, que ofrece extremo a la funcional (4) debe verificar la ecuación diferencial (7) y las condiciones de contorno (9) y (10) en los extremos del segmento $[a, b]$. Como la solución general de una ecuación diferencial de segundo orden contiene dos constantes arbitrarias, tenemos a nuestra disposición un número de condiciones de contorno necesario precisamente para encontrar estas constantes.

2°. **Segunda diferencial. Condiciones suficientes de extremo de una funcional.** Volvamos de nuevo al problema sobre la búsqueda del extremo de una función de n variables. Supongamos que para la función $f(x_1, \ldots, x_n)$ se cumple en el punto $(x_1^0, \ldots, x_n^0)$ la condición $df=0$. Entonces, como se sabe, para resolver el problema de si hay o no hay efectivamente en este punto un extremo, debe considerarse la segunda diferencial. Tienen lugar las siguientes proposiciones.

1. *Si una función $f(x_1, \ldots, x_n)$ tiene en un punto $(x_1^0, \ldots, x_n^0)$ mínimo, en ese punto $d^2f \geqslant 0$.* (Análogamente, si en un punto $(x_1^0, \ldots, x_n^0)$ hay máximo, en ese punto $d^2f \leqslant 0$).

2. *Si en un punto $(x_1^0, \ldots, x_n^0)$ se cumplen las condiciones*

$$df=0 \quad \text{y} \quad d^2f=\sum_{i,\,k=1}^{n}\frac{\partial^2 f}{\partial x_i \partial x_k}dx_i\,dx_k>0$$

(cuando no todo $dx_i=0$), *la función $f(x)$ tiene en ese punto mínimo* (análogamente, máximo, si $d^2f<0$).

Veamos en qué medida subsisten estos resultados para funcionales definidas en un espacio de Banach.

TEOREMA 2. *Sea F una funcional real, definida en un espacio de Banach X, con segunda derivada continua en una vecindad del punto x_0. Si esta funcional alcanza un mínimo en el punto x_0, entonces, $d^2F(x_0)\geqslant 0$* [1].

DEMOSTRACION. Empleando la fórmula de Taylor, tenemos

$$F(x_0+h)-F(x_0)=F'(x_0)(h)+\frac{1}{2}F''(x_0)(h, h)+o(\|h\|^2).$$

[1] Esta desigualdad significa que $F''(x_0)(h, h)\geqslant 0$ para todo h.

Si la funcional F tiene mínimo en el punto x_0, entonces, $F'(x_0)=0$ y queda la igualdad

$$F(x_0+h)-F(x_0)=\frac{1}{2}F''(x_0)(h,\ h)+o(\|h\|^2). \qquad (11)$$

Si para algún h admisible tiene lugar la desigualdad

$$F''(x_0)(h,\ h)<0 \qquad (12)$$

veremos, teniendo en cuenta que $F''(x_0)(\varepsilon h,\ \varepsilon h)=\varepsilon^2 F''(x_0)(h,\ h)$, que existen elementos h de norma tan pequeña como se quiera para los cuales también se cumple (12). Pero, el signo de toda la expresión (11) depende, para $\|h\|$ suficientemente pequeño, del signo de su término principal $\frac{1}{2}F(x_0)(h,\ h)$ y obtenemos que

$$F(x_0+h)-F(x_0)=\frac{1}{2}F''(x_0)(h,\ h)+o(\|h\|^2)<0,$$

es decir, que no hay mínimo en el punto x_0. Análogamente se considera el caso de máximo.

El teorema demostrado es una generalización directa del teorema correspondiente para las funciones de un número finito de variables. La situación es distinta en el caso de la condición suficiente. La condición mencionada más arriba $d\ F''(x_0)(h,\ h)>0$, que es suficiente para el mínimo en el caso de funciones de n variables, no resulta suficiente para funcionales definidas en un espacio de Banach de dimensión infinita. Veamos un ejemplo sencillo. Consideremos en el espacio de Hilbert la funcional

$$F(x)=\sum_{n=1}^{\infty}\frac{x_n^2}{n^3}-\sum_{n=1}^{\infty}x_n^4.$$

En el punto 0, la primera diferencial de esta funcional es igual a 0 y la segunda es igual a $\sum_{n=1}^{\infty}\frac{h_n^2}{n^3}$, es decir, representa una funcional definida positiva. Sin embargo, en el punto 0 no hay mínimo, ya que

$$F(0)=0 \text{ y } F\left(0,\ \ldots,\ 0,\ \frac{1}{n},\ 0,\ \ldots\right)=\frac{1}{n^5}-\frac{1}{n^4}<0.$$

Por consiguiente, en cualquier vecindad del punto 0 existen puntos en los cuales $F(x)<F(0)$.

Introduzcamos el siguiente concepto. Una funcional cuadrática B se llama *fuertemente positiva*, cuando existe un número $c>0$ tal que $B(x,\ x)\geqslant c\|x\|^2$ para todo x.

TEOREMA 3. *Si una funcional F, definida en un espacio de Banach X, verifica las condiciones*

1) $F'(x_0)=0$,

2) $F''(x_0)$ es una funcional cuadrática fuertemente positiva, F tiene mínimo en el punto x_0.

DEMOSTRACION. Escojamos $\varepsilon>0$ tan pequeño que para $\|h\|<\varepsilon$ la magnitud $o(\|h\|^2)$ en la igualdad (11) verifique la condición $|o(\|h\|^2)|<\frac{c}{4}\|h\|^2$. Entonces,

$$F(x_0+h)-F(x_0)=\frac{1}{2}F''(x_0)(h,h)+o(\|h\|^2)>\frac{c}{4}\|h\|^2>0$$
$$\text{para } \|h\|<\varepsilon.$$

En un espacio de dimensión finita la positividad fuerte de una forma cuadrática equivale a que sea definida positiva y por eso (siendo igual a cero la primera diferencial) es una condición suficiente de mínimo de una función de un número finito de variables el que la segunda diferencial sea definida positiva. En el caso de dimensión infinita (como muestra el ejemplo dado más arriba), la positividad fuerte es una condición más fuerte que la de definida positiva.

La condición de positividad fuerte de la segunda diferencial que garantiza la existencia de mínimo es cómoda porque se puede aplicar a cualquier funcional (independientemente de su forma concreta) dos veces diferenciable en cualquier espacio de Banach. Al mismo tiempo, esta condición resulta demasiado tosca y difícilmente comprobable en casos prácticos importantes. En el Cálculo de variaciones se establecen unas condiciones suficientes de extremo más finas (que emplean la forma concreta de las funcionales que se consideran en los problemas variacionales); sin embargo, la exposición de estos temas no entra en la tarea de nuestro libro.

§ 3. METODO DE NEWTON

Uno de los métodos bien conocidos de resolución de ecuaciones de tipo

$$f(x)=0 \tag{1}$$

(f es una función numérica de argumento numérico, definida en un segmento $[a, b]$) es el así llamado *método de Newton* o *método de tangentes*. Consiste en que para resolver la ecuación (1) se buscan las aproximaciones sucesivas de acuerdo con la fórmula recurrente

$$x_{n+1}=x_n-\frac{f(x_n)}{f'(x_n)} \tag{2}$$

(por aproximación nula x_0 se toma aquí un punto arbitrario del segmento donde está definida f). La interpretación geométrica de este método viene dada en la fig. 19. Se puede demostrar que si x^* es la única raíz de la ecuación (1) en el segmento $[a, b]$ y si la función f tiene en este segmento la primera derivada diferente de cero y la segunda derivada acotada, existe una vecin-

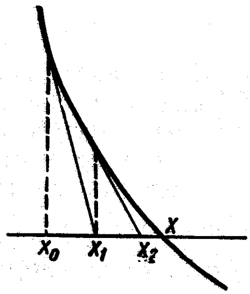


FIG. 19

dad de la raíz x^* tal que si el punto x_0 se toma en esta vecindad, la sucesión (2) converge hacia x^*.

El método de Newton se puede extender a las ecuaciones en operadores. Expondremos aquí este método para el caso de ecuaciones en operadores en espacios de Banach.

Consideremos la ecuación

$$F(x)=0, \tag{3}$$

donde F es una aplicación de un espacio de Banach X en otro espacio de Banach Y. Supongamos que la aplicación F es fuertemente diferenciable en una bola $B(x_0, r)$ de radio r (cuyo centro x_0 tomaremos como la aproximación nula de la solución que buscamos) y que su derivada F' satisface en esta bola la condición de Lipschitz, es decir,

$$\|F'(x_1)-F'(x_2)\| \leqslant L\|x_1-x_2\| \quad (L=\text{const}). \tag{4}$$

Sustituyendo, al igual que en el caso unidimensional, la expresión $F(x_0)-F(x)$ por su parte lineal principal, esto es, por el elemento $F'(x_0)(x_0-x)$, obtendremos de (3) una ecuación lineal $F'(x_0)(x_0-x)=F(x_0)$, cuya solución $x_1=x_0-[F'(x_0)]^{-1}F(x_0)$ es natural tomar por la siguiente aproximación de la solución x de la ecuación $F(x)=0$ (aquí se presupone, claro está, la existencia del operador $[F'(x_0)]^{-1}$). Repitiendo estos razonamientos, obtendremos una sucesión

$$x_{n+1}=x_n-[F'(x_n)]^{-1}(F(x_n)) \tag{5}$$

de soluciones aproximadas de la ecuación (3). En el caso de dimensión infinita, la búsqueda del operador inverso $[F'(x_n)]^{-1}$ puede resultar una tarea suficientemente compleja. Por eso, conviene, a veces, emplear aquí el así llamado *método modificado de Newton*. La modificación consiste en que, en lugar de la sucesión (5), se considera la sucesión definida por la fórmula

$$x_{n+1}=x_n-[F'(x_0)]^{-1}(F(x_n)), \tag{6}$$

es decir, en cada paso el operador inverso $[F'(x_0)]^{-1}$ se toma para un mismo valor del argumento $x=x_0$. Aunque esta modificación reduce la velocidad de convergencia, resulta con frecuencia conveniente desde el punto de vista de cálculo. Pasemos ahora al enunciado y a la demostración de la proposición exacta.

TEOREMA 1. *Sean $M=\|[F'(x_0)]^{-1}\|$, $k=\|[F'(x_0)]^{-1}F(x_0)\|$ y sea L la constante que figura en la desigualdad* (4). *Entonces, si $h=MkL<\frac{1}{4}$* y *t_0 es la menor de las raíces de la ecuación $ht^2-t+1=0$, la ecuación $F(x)=0$ tiene en la bola $\|x-x_0\|\leqslant t_0k$* una solución única *x^* y la sucesión $\{x_n\}$ definida por la fórmula recurrente* (6), *converge a esta solución.*

DEMOSTRACION. Consideremos en el espacio X la aplicación $Ax=x-[F'(x_0)]^{-1}F(x)$. Esta aplicación transforma la bola $\|x-x_0\|\leqslant t_0k$ en sí misma. En efecto,

$$Ax-x_0=x-x_0-[F'(x_0)]^{-1}F(x)=$$
$$=[F'(x_0)]^{-1}\{F'(x_0)(x-x_0)-F(x)+F(x_0)\}-[F'(x_0)]^{-1}F(x_0).$$

Por eso,

$$\|Ax-x_0\|\leqslant\|[F'(x_0)]^{-1}\|\cdot\|F'(x_0)(x-x_0)-F(x)+$$
$$+F(x_0)\|+\|[F'(x_0)]^{-1}F(x_0)\|,$$

es decir,

$$\|Ax-x_0\|\leqslant M\|F'(x_0)(x-x_0)-F(x)+F(x_0)\|+k. \tag{7}$$

Consideremos la aplicación auxiliar

$$\Phi(x)=F(x)-F(x_0)-F'(x_0)(x-x_0).$$

Es diferenciable y su derivada es igual a

$$\Phi'(x)=F'(x)-F'(x_0).$$

Si $\|x-x_0\|<t_0k$, tiene lugar la estimación

$$\|\Phi'(x)\|=\|F'(x)-F'(x_0)\|\leqslant L\|x-x_0\|\leqslant Lt_0k.$$

De aquí, según el teorema del valor medio, obtenemos

$$\|\Phi(x)\|=\|\Phi(x)-\Phi(x_0)\|\leqslant Lt_0k\|x-x_0\|\leqslant Lt_0^2k^2. \qquad (8)$$

De manera que siendo $\|x-x_0\|\leqslant t_0k$, tenemos de (7) y (8)

$$\|Ax-x_0\|\leqslant MLt_0^2k^2+k=k(MLt_0^2k+1)=k(ht_0^2+1)=kt_0,$$

y esto significa que la aplicación A transforma la bola $\|x-x_0\|\leqslant kt_0$ en sí misma. Probemos ahora que A es una aplicación contraída de esta bola. Para $\|x-x_0\|\leqslant kt_0$ tenemos

$$A'(x)=I-[F'(x_0)]^{-1}F'(x)=[F'(x_0)]^{-1}(F'(x_0)-F'(x)),$$

de donde

$$\|A'(x)\|\leqslant M\|F'(x_0)-F'(x)\|\leqslant ML\|x-x_0\|\leqslant MLkt_0.$$

Pero t_0 es la menor de las raíces de la ecuación

$$ht^2-t+1=0, \text{ es decir,}$$

$$t_0=\frac{1-\sqrt{1-4h}}{2h}.$$

Por esto,

$$A'(x)\|\leqslant MLkt_0=ht_0=$$
$$=h\frac{1-\sqrt{1-4h}}{2h}=\frac{1-\sqrt{1-4h}}{2}=q<\frac{1}{2}, \qquad (9)$$

de donde

$$\|Ax_1-Ax_2\|\leqslant\frac{1}{2}\|x_1-x_2\|,$$

es decir, A es una aplicación contraída.

Por consiguiente, la aplicación A tiene en la bola $\|x-x_0\|\leqslant\leqslant kt_0$ un punto fijo x^*, y sólo uno. Para este punto

$$x^*=x^*-[F'(x_0)]^{-1}F(x^*), \text{ es decir, } F(x^*)=0.$$

Al mismo tiempo, $Ax_n=x_n-[F'(x_0)]^{-1}F(x_n)=x_{n+1}$ y, en virtud del teorema sobre las aplicaciones contraídas, la sucesión $\{x_n\}$ converge hacia x^*.

De la desigualdad (9) se desprende inmediatamente la siguiente estimación para la velocidad de convergencia del método modificado de Newton:

$$\|x_n-x^*\|\leqslant\frac{q^n}{1-q}\|[F'(x_0)]^{-1}F(x_0)\|,$$

es decir, el error del método modificado de Newton disminuye como los términos de una serie geométrica. Para comparar, indiquemos que el método corriente de Newton (en el que las aproximaciones se definen mediante la fórmula (5) en lugar de la fórmula (6)) converge más rápidamente que una serie geométrica: para este método

$$\| x_n - x^* \| \leqslant \frac{1}{2^{n-1}} (2h)^{2^n-1} k.$$

Ejemplo. Consideremos la ecuación integral no lineal

$$x(s) = \int_a^b K(s, t, x(t))\, dt, \tag{10}$$

donde $K(s, t, u)$ es una función continua y continuamente diferenciable de sus argumentos. Introduciendo la aplicación $y = F(x)$, definida por la igualdad

$$y(s) = x(s) - \int_a^b K(s, t, x(t))\, dt,$$

podemos escribir la ecuación (10) en la forma

$$F(x) = 0.$$

Sea x_0 la aproximación nula para la solución de esta ecuación. Entonces, la primera rectificación $\Delta x(s) = x_1 - x_0$ se encuentra de la ecuación

$$F'(x_0)\, \Delta x = -F(x_0). \tag{11}$$

Si la función $K(s, t, u)$ y el espacio funcional, en el que se considera la ecuación (10), son tales que la derivada $F'(x)$ de la aplicación F se puede calcular «diferenciando bajo el símbolo de la integral», es decir, si

$$z = F'(x_0)(x)$$

significa que

$$z(s) = x(s) - \int_a^b K'_u(s, t, x_0(t))\, x(t)\, dt,$$

la ecuación (11) se representa en la forma

$$\Delta x(s) = \int_a^b K'_u(s, t, x_0(t))\, \Delta x(t)\, dt + \varphi_0(s), \tag{12}$$

donde

$$\varphi_0(s) = \int_a^b K(s, t, x_0(t))\, dt - x_0(s).$$

Análogamente se buscan las rectificaciones siguientes.

De manera que para buscar cada aproximación siguiente de la solución de la ecuación (10) hay que resolver una ecuación integral lineal. Cuando se emplea el método modificado de Newton, resulta que en cada uno de estos pasos hay que resolver una ecuación lineal con el mismo núcleo.

CAPITULO VI

MEDIDA, FUNCIONES MEDIBLES, INTEGRAL

El concepto de medida $\mu(A)$ de un conjunto A constituye una generalización natural de los siguientes conceptos:

1) de la longitud $l(\Delta)$ de un segmento Δ,
2) del área $S(F)$ de una figura plana F,
3) del volumen $V(G)$ de una figura G del espacio,
4) del incremento $\varphi(b)-\varphi(a)$ de una función no decreciente $\varphi(t)$ en el semisegmento $[a, b)$,
5) de la integral de una función no negativa en una región lineal, plana o del espacio, etc.

Este concepto, surgido inicialmente en la Teoría de funciones de variable real, encontró más tarde múltiples aplicaciones en la Teoría de Probabilidades, la Teoría de Sistemas Dinámicos, el Análisis Funcional y en otras ramas de las Matemáticas.

En el § 1 de este capítulo exponemos la teoría de medida para el caso de conjuntos planos, partiendo del concepto del área de un rectángulo. La teoría general de medida es explicada en los §§ 2 y 3. El lector podrá notar que todos los razonamientos que se realizan en el § 1 tienen un carácter general y se repiten, sin modificaciones sustanciales, en la teoría abstracta.

§ 1. MEDIDA DE CONJUNTOS PLANOS

1°. Medida de conjuntos elementales. Consideremos el sistema $\mathfrak{S}$ de conjuntos del plano (x, y), cada uno de los cuales se determina por una desigualdad de tipo

$$a \leqslant x \leqslant b,$$
$$a < x \leqslant b,$$
$$a \leqslant x < b,$$
$$a < x < b$$

y por una desigualdad de tipo

$$c \leqslant y \leqslant d,$$
$$c < y \leqslant d,$$
$$c \leqslant y < d,$$
$$c < y < d,$$

donde a, b, c, y d son números arbitrarios. Los conjuntos pertenecientes a este sistema se llamarán rectángulos. Un rectángulo cerrado definido por las desigualdades

$$a \leqslant x \leqslant b, \quad c \leqslant y \leqslant d$$

es un rectángulo en el sentido corriente (con su frontera) cuando $a < b$ y $c < d$; es un segmento (cuando $a = b$ y $c < d$ ó $a < b$ y $c = d$); es un punto (para $a = b$ y $c = d$) o, finalmente, el conjunto vacío (cuando $a > b$ ó $c > d$). Un rectángulo abierto

$$a < x < b, \quad c < y < d$$

representa en función de la relación entre a, b, c y d, o bien un rectángulo sin frontera, o bien el conjunto vacío. Cada rectángulo de los tipos restantes (que llamaremos rectángulos semiabiertos) constituye o bien un rectángulo sin uno, dos o tres lados, o bien un intervalo, o bien un semisegmento o bien, finalmente, un conjunto vacío.

Partiendo del concepto de área, conocido de la Geometría Elemental, definiremos la medida de cada rectángulo de la siguiente forma:

a) la medida del conjunto vacío es igual a 0;

b) la medida de un rectángulo no vacío (cerrado, abierto o semiabierto) determinado por los números a, b, c y d es igual a

$$(b-a)(d-c).$$

Luego, hemos asignado a todo rectángulo P un número $m(P)$, la medida de este rectángulo, de manera que se cumplen, evidentemente, las siguientes condiciones:

1) la medida $m(P)$ toma valores reales no negativos;

2) la medida $m(P)$ es aditiva, esto es, si $P = \bigcup_{k=1}^{n} P_k$ y $P_i \cap P_k = \varnothing$ para $i \neq k$, entonces,

$$m(P) = \sum_{k=1}^{n} m(P_k).$$

Nuestra tarea es extender la medida $m(P)$, definida por ahora para rectángulos, a una clase más general de conjuntos, conservando las condiciones 1) y 2).

El primer paso en esta dirección consiste en extender el concepto de medida a los así llamados conjuntos elementales. Un conjunto plano se llamará *elemental* cuando puede ser representado, al menos de una forma, como la unión de un número finito de rectángulos disjuntos dos a dos.

En lo que sigue necesitaremos el siguiente teorema.

TEOREMA 1. *La unión, la intersección, la diferencia y la diferencia simétrica de dos conjuntos elementales son también conjuntos elementales.*

DEMOSTRACION. Está claro que la intersección de dos rectángulos es de nuevo un rectángulo. Por eso, si

$$A=\bigcup_k P_k \quad \text{y} \quad B=\bigcup_j Q_j$$

son dos conjuntos elementales, también

$$A\cap B=\bigcup_{k,\,j}(P_k\cap Q_j)$$

es un conjunto elemental.

Es fácil ver que la diferencia de dos rectángulos es un conjunto elemental. Consecuentemente, sustrayendo de un rectángulo un conjunto elemental, obtenemos de nuevo un conjunto elemental (como intersección de conjuntos elementales). Sean ahora A y B dos conjuntos elementales. Es obvio que existe un rectángulo P que contiene a ambos. Entonces,

$$A\cup B=P\setminus[(P\setminus A)\cap(P\setminus B)]$$

es, de acuerdo con lo señalado anteriormente, un conjunto elemental. De las igualdades

$$A\setminus B=A\cap(P\setminus B)$$

y

$$A\triangle B=(A\cup B)\setminus(A\cap B)$$

se deduce entonces que la diferencia y la diferencia simétrica de conjuntos elementales son conjuntos elementales. El teorema queda demostrado. ■

Definamos ahora la medida $m'(A)$ de conjuntos elementales del siguiente modo: si

$$A=\bigcup_k P_k,$$

donde P_k son rectángulos disjuntos dos a dos, tomamos

$$m'(A)=\sum_k m(P_k).$$

Probemos que $m'(A)$ no depende de la forma de representar al conjunto A como unión de rectángulos. Sea

$$A=\bigcup_k P_k=\bigcup_j Q_j,$$

donde P_k y Q_j son rectángulos y $P_i\cap P_k=\varnothing$, $Q_i\cap Q_k=\varnothing$ para $i\neq k$. Como la intersección $P_k\cap Q_j$ de dos rectángulos es un rectángulo, tenemos, en virtud de la aditividad de la medida de rectángulos,

$$\sum_k m(P_k)=\sum_{k,j} m(P_k\cap Q_j)=\sum_j m(Q_j).$$

Es fácil ver que la medida de conjuntos elementales definida de esta forma es no negativa y aditiva.

Establezcamos la siguiente propiedad de la medida de conjuntos elementales importante para lo sucesivo.

TEOREMA 2. *Si A es un conjunto elemental y $[A_n]$ es un sistema finito o numerable de conjuntos elementales tal que*

$$A\subset\bigcup_n A_n,$$

entonces,

$$m'(A)\leqslant\sum_n m'(A_n). \tag{1}$$

DEMOSTRACION. Para cualquier $\varepsilon>0$ y un conjunto A dado es posible, evidentemente, encontrar un conjunto elemental cerrado $\bar{A}$ contenido en A que verifique la condición

$$m'(\bar{A})\geqslant m'(A)-\frac{\varepsilon}{2}.$$

(Para ello es suficiente sustituir cada uno de los k rectángulos P_l que componen A por un rectángulo cerrado contenido en él de área mayor que $m(P_l)-\frac{\varepsilon}{2k}$).

Además, para cada A_n se puede encontrar un conjunto elemental abierto $\tilde{A}_n$ que contiene A y verifica la condición

$$m'(\tilde{A}_n)\leqslant m'(A_n)+\frac{\varepsilon}{2^{n+1}}.$$

Está claro que

$$\overline{A} \subset \bigcup_n \tilde{A}_n.$$

De acuerdo con el lema de Heine—Borel, se puede extraer de $\{\tilde{A}_n\}$ un sistema finito $\tilde{A}_{n_1}, \ldots, \tilde{A}_{n_s}$ que cubre $\overline{A}$. Es obvio que

$$m'(\overline{A} \leqslant \sum_{i=1}^{s} m'(\tilde{A}_{n_i})$$

(de lo contrario $\overline{A}$ resultaría cubierto por un número finito de rectángulos de un área total menor que $m'(\overline{A})$, lo cual es imposible). Por eso,

$$m'(A) \leqslant m'(\overline{A}) + \frac{\varepsilon}{2} \leqslant \sum_{i=1}^{s} m'(\tilde{A}_{n_i}) + \frac{\varepsilon}{2} \leqslant \sum_n m'(\tilde{A}_n) + \frac{\varepsilon}{2} \leqslant$$
$$\leqslant \sum_n m'(A_n) + \sum_n \frac{\varepsilon}{2^{n+1}} + \frac{\varepsilon}{2} = \sum m'(A_n) + \varepsilon,$$

de donde, debido a la arbitrariedad de $\varepsilon > 0$, se desprende (1). □

2°. **Medida de Lebesgue de conjuntos planos.** La clase de conjuntos elementales no agota todos los conjuntos que se consideran en la Geometría elemental y en el Análisis clásico. Resulta natural, por eso, plantear el problema de la extensión del concepto de medida, conservando sus propiedades principales, a una clase de conjuntos más amplia que la compuesta por uniones finitas de rectángulos de lados paralelos a los ejes de coordenadas.

Este problema fue resuelto, en cierto sentido de un modo definitivo, por H. Lebesgue a principos del siglo XX.

Al presentar la teoría de medida de Lebesgue tendremos que considerar no sólo uniones finitas sino también uniones infinitas de rectángulos.

Para evitar que aparezcan en este caso conjuntos de «medida infinita», nos limitaremos a considerar en lo sucesivo conjuntos contenidos íntegramente en el cuadrado $E = \{0 \leqslant x \leqslant 1;\ 0 \leqslant y \leqslant 1\}$.

En la clase de estos conjuntos definiremos como sigue dos funciones $\mu^*(A)$ y $\mu_*(A)$.

DEFINICION 1. Se llama *medida superior* $\mu^*(A)$ del conjunto A al número

$$\inf_{A \subset \bigcup P_k} \sum m(P_k),$$

donde la cota inferior se toma respecto a todos los cubrimientos del conjunto A por medio de sistemas finitos o numerables de rectángulos.

DEFINICION 2. Se llama *medida inferior* $\mu_*(A)$ del conjunto A al número

$$1-\mu^*(E \setminus A).$$

Es fácil ver que siempre

$$\mu_*(A) \leqslant \mu^*(A).$$

En efecto, supongamos que para un conjunto $A \subset E$ se tiene

$$\mu_*(A) > \mu^*(A),$$

es decir,

$$\mu^*(A) + \mu^*(E \setminus A) < 1.$$

De acuerdo con la definición de la cota inferior máxima, existirán entonces dos sistemas de rectángulos $\{P_i\}$ y $\{Q_i\}$, que cubren A y $E \setminus A$, respectivamente, tales que

$$\sum_i m(P_i) + \sum_k m(Q_k) < 1.$$

Sea $\{R_j\}$ la unión de los sistemas $\{P_i\}$ y $\{Q_i\}$; tenemos

$$E \subset \bigcup_j R_j \quad \text{y} \quad m(E) > \sum_j m(R_j),$$

lo que contradice al teorema 2.

DEFINICION 3. Un conjunto A se llama *medible* (en el sentido de Lebesgue) cuandc

$$\mu_*(A) = \mu^*(A).$$

El valor común $\mu(A)$ de las medidas superior e inferior de un conjunto medible A es su *medida de Lebesgue.*

3°. **Propiedades principales de la medida de Lebesgue y de los conjuntos medibles.** Demostremos primero la siguiente propiedad de la medida superior.

TEOREMA 3. *Si*

$$A \subset \bigcup_n A_n,$$

donde $\{A_n\}$ *es un sistema finito o numerable de conjuntos, se tiene*

$$\mu^*(A) \leqslant \sum_n \mu^*(A_n).$$

DEMOSTRACION. De acuerdo con la definición de medida superior, para cada A_n existe un sistema de rectángulos $\{P_{nk}\}$, finito o nu.

merable, tal que $A_n \subset \bigcup_k P_{nk}$ y

$$\sum_k m(P_{nk}) \leqslant \mu^*(A_n) + \frac{\varepsilon}{2^n},$$

donde $\varepsilon > 0$ es escogido arbitrariamente. En este caso

$$A \subset \bigcup_n \bigcup_k P_{nk}$$

$$\mu^*(A) \leqslant \sum_n \sum_k m(P_{nk}) \leqslant \sum_n \mu^*(A_n) + \varepsilon.$$

Como $\varepsilon > 0$ es arbitrario, de aquí se deduce la afirmación del teorema.

Más arriba hemos introducido ya el concepto de medida para conjuntos que hemos llamado elementales. El teorema que sigue muestra que en el caso de conjuntos elementales la definición 3 lleva al mismo resultado.

TEOREMA 4. *Los conjuntos elementales son medibles y para ellos la medida de Lebesgue coincide con la medida $m'(A)$ construida anteriormente.*

DEMOSTRACION. Si A es un conjunto elemental y $P_1, P_2, \ldots, P_k$ son los rectángulos disjuntos dos a dos que lo componen, tenemos por definición

$$m'(A) = \sum_{i=1}^{k} m(P_i).$$

Como los rectángulos P_i cubren todo el A, tenemos

$$\mu^*(A) \leqslant \sum_i m(P_i) = m'(A).$$

Pero si $\{Q_j\}$ es un sistema arbitrario finito o numerable de rectángulos que cubre A, entonces, de acuerdo con el teorema 2, $m'(A) \leqslant \sum_j m(Q_j)$; de manera que $m'(A) = \mu^*(A)$.

Como $E \setminus A$ es también un conjunto elemental, tenemos $m'(E \setminus A) = \mu^*(E \setminus A)$. Pero

$$m'(E \setminus A) = 1 - m'(A) \quad \text{y} \quad \mu^*(E \setminus A) = 1 - \mu_*(A),$$

de donde

$$m'(A) = \mu_*(A).$$

Por consiguiente,

$$\mu^*(A)=\mu_*(A)=m'(A).$$

Del resultado obtenido se desprende que el teorema 2 es un caso particular del teorema 3.

TEOREMA 5. *Para que un conjunto A sea medible es necesario y suficiente que se cumpla la siguiente condición: cualquiera que sea $\varepsilon>0$ existe un conjunto elemental B, tal que*

$$\mu^*(A\triangle B)<\varepsilon.$$

De esta forma, son medibles aquellos conjuntos, y sólo aquéllos, que pueden ser «aproximados» con cualquier grado de presición por conjuntos elementales. Para demostrar el teorema 5 necesitaremos el siguiente lema.

LEMA. *Para dos cualesquiera conjuntos A y B se tiene*

$$|\mu^*(A)-\mu^*(B)|\leqslant\mu^*(A\triangle B).$$

DEMOSTRACION DEL LEMA. Como

$$A\subset B\cup(A\triangle B),$$

tendremos, en virtud del teorema 3,

$$\mu^*(A)\leqslant\mu^*(B)+\mu^*(A\triangle B).$$

De aquí se desprende la proposición del lema para el caso en que $\mu^*(A)\geqslant\mu^*(B)$. En cambio, si $\mu^*(A)\leqslant\mu^*(B)$, la afirmación del lema se desprende de la desigualdad

$$\mu^*(B)\leqslant\mu^*(A)+\mu^*(A\triangle B),$$

que se demuestra de una manera análoga.

DEMOSTRACION DEL TEOREMA 5. SUFICIENCIA. Supongamos que para cualquier $\varepsilon>0$ existe un conjunto elemental B tal que

$$\mu^*(A\triangle B)<\varepsilon.$$

Entonces, de acuerdo al lema,

$$|\mu^*(A)-m'(B)|=|\mu^*(A)-\mu^*(B)|<\varepsilon \tag{2}$$

y como

$$(E\setminus A)\triangle(E\setminus B)=A\triangle B,$$

de la misma forma obtenemos que

$$|\mu^*(E\setminus A)-m'(E\setminus B)|<\varepsilon. \tag{3}$$

Teniendo en cuenta que

$$m'(B)+m'(E\setminus B)=m'(E)=1,$$

encontramos de las desigualdades (2) y (3)

$$|\mu^*(A)+\mu^*(E\setminus A)-1|<2\varepsilon$$

y como $\varepsilon>0$ es arbitrario, tenemos

$$\mu^*(A)+\mu^*(E\setminus A)=1,$$

es decir, el conjunto A es medible.

NECESIDAD. Sea A medible, esto es,

$$\mu^*(A)+\mu^*(E\setminus A)=1.$$

Para un $\varepsilon>0$ arbitrario busquemos unos cubrimientos de los conjuntos A y $E\setminus A$ mediante sistemas de rectángulos $\{B_n\}$ y $\{C_n\}$ tales que

$$\sum_n m(B_n)\leqslant\mu^*(A)+\frac{\varepsilon}{3}\ \text{ y }\ \sum_n m(C_n)\leqslant\mu^*(E\setminus A)+\frac{\varepsilon}{3}.$$

Como $\sum\limits_n m(B_n)<\infty$, existirá un N tal que

$$\sum_{n>N} m(B_n)<\frac{\varepsilon}{3};$$

tomemos

$$B=\bigcup_{n=1}^{N} B_n.$$

Demostremos que el conjunto elemental B satisface la condición $\mu^*(A\triangle B)<\varepsilon$. Está claro que el conjunto

$$P=\bigcup_{n>N} B_n$$

contiene $A\setminus B$, que el conjunto

$$Q=\bigcup_n (B\cap C_n)$$

contiene $B\setminus A$ y que, por consiguiente, $A\triangle B\subset P\cup Q$. Además,

$$\mu^*(P)\leqslant\sum_{n>N} m(B_n)<\frac{\varepsilon}{3}.$$

Estimemos $\mu^*(Q)$. Para ello observemos que

$$\left(\bigcup_n B_n\right)\cup\left(\bigcup_n (C_n\setminus B)\right)=E$$

de manera que

$$\sum_n m(B_n) + \sum_n m'(C_n \setminus B) \geqslant 1. \tag{4}$$

Pero, por hipótesis,

$$\sum_n m(B_n) + \sum_n m(C_n) \leqslant \mu^*(A) + \mu^*(E \setminus A) + \frac{2\varepsilon}{3} = 1 + \frac{2\varepsilon}{3}. \tag{5}$$

Sustrayendo (4) de (5), obtenemos

$$\sum_n m(C_n) - \sum_n m'(C_n \setminus B) = \sum_n m'(C_n \cap B) < \frac{2}{3}\varepsilon,$$

es decir,

$$\mu^*(Q) < \frac{2}{3}\varepsilon.$$

Por eso,

$$\mu^*(A \triangle B) \leqslant \mu^*(P) + \mu^*(Q) < \varepsilon.$$

Luego, si A es medible, cualquiera que sea $\varepsilon > 0$ existe un conjunto elemental B tal que $\mu^*(A \triangle B) < \varepsilon$. El teorema queda demostrado. □

TEOREMA 6. *La unión y la intersección de un número finito de conjuntos medibles son conjuntos medibles.*

DEMOSTRACION. Es obvio que basta realizar la demostración para el caso de dos conjuntos. Sean A_1 y A_2 conjuntos medibles. Entonces, para cualquier $\varepsilon > 0$ existen conjuntos elementales B_1 y B_2 tales que

$$\mu^*(A_1 \triangle B_1) < \frac{\varepsilon}{3}, \quad \mu^*(A_2 \triangle B_2) < \frac{\varepsilon}{2}.$$

Como

$$(A_1 \cup A_2) \triangle (B_1 \cup B_2) \subset (A_1 \triangle B_1) \cup (A_2 \triangle B_2),$$

tenemos

$$\mu^*[(A_1 \cup A_2) \triangle (B_1 \cup B_2)] \leqslant \mu^*(A_1 \triangle B_1) + \mu^*(A_2 \triangle B_2) < \varepsilon.$$

Pero, $B_1 \cup B_2$ es un conjunto elemental; luego, en virtud del teorema 4, el conjunto $A_1 \cup A_2$ es medible.

Por definición de conjunto medible, siendo A medible, también $E \setminus A$ es medible; por esto, la intersección de dos conjuntos medibles es medible en vista de la relación

$$A_1 \cap A_2 = E \setminus [(E \setminus A_1) \cup (E \setminus A_2)].$$

COROLARIO. *La diferencia y la diferencia simétrica de dos conjuntos medibles son medibles.*

Esto se deduce del teorema 6 y de las igualdades

$$A_1 \setminus A_2 = A_1 \cap (E \setminus A_2), \quad A_1 \triangle A_2 = (A_1 \setminus A_2) \cup (A_2 \setminus A_1).$$

TEOREMA 7. *Si $A_1, \ldots, A_n$ son conjuntos medibles disjuntos dos a dos, entonces,*

$$\mu\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mu(A_k). \tag{6}$$

DEMOSTRACION. Al igual que en el teorema 6 es suficiente considerar el caso $n=2$. Escojamos arbitrariamente un $\varepsilon > 0$ y sean B_1 y B_2 conjuntos elementales tales que

$$\mu^*(A_1 \triangle B_1) < \varepsilon, \tag{7}$$
$$\mu^*(A_2 \triangle B_2) < \varepsilon. \tag{8}$$

Pongamos $A = A_1 \cup A_2$ y $B = B_1 \cup B_2$. De acuerdo con el teorema 6, el conjunto A es medible. Como los conjuntos A_1 y A_2 no se intersecan,

$$B_1 \cap B_2 \subset (A_1 \triangle B_1) \cup (A_2 \triangle B_2)$$

y, por consiguiente,

$$m'(B_1 \cap B_2) \leqslant 2\varepsilon. \tag{9}$$

De (7) y (8) resulta, en virtud del lema del teorema 5, que

$$|m'(B_1) - \mu^*(A_1)| < \varepsilon, \tag{10}$$
$$|m'(B_2) - \mu^*(A_2)| < \varepsilon. \tag{11}$$

Puesto que la medida es aditiva en la clase de conjuntos elementales, obtenemos de (9), (10) y (11)

$$m'(B) = m'(B_1) + m'(B_2) - m'(B_1 \cap B_2) \geqslant \mu^*(A_1) + \mu^*(A_2) - 4\varepsilon.$$

Observando, además, que $A \triangle B \subset (A_1 \triangle B_1) \cup (A_2 \triangle B_2)$, encontramos finalmente

$$\mu^*(A) \geqslant m'(B) - \mu^*(A \triangle B) \geqslant m'(B) - 2\varepsilon \geqslant \mu^*(A_1) + \mu^*(A_2) - 6\varepsilon.$$

Como $\varepsilon > 0$ se puede escoger tan pequeño como se quiera, tenemos

$$\mu^*(A) \geqslant \mu^*(A_1) + \mu^*(A_2).$$

Por ser siempre válida (en virtud del teorema (3)) la desigualdad opuesta

$$\mu^*(A) \leqslant \mu^*(A_1) + \mu^*(A_2)$$

para $A=A_1\cup A_2$, obtenemos en conclusión

$$\mu^*(A)=\mu^*(A_1)+\mu^*(A_2);$$

como A_1, A_2 y A son medibles, se puede sustituir aquí μ^* por μ. El teorema queda demostrado.

TEOREMA 8. *La unión e intersección de un número numerable de conjuntos medibles son conjuntos medibles.*

DEMOSTRACION. Sea

$$A_1, A_2, \ldots, A_n, \ldots$$

un sistema numerable de conjuntos medibles y sea $A=\bigcup_{n=1}^{\infty} A_n$. Tomemos $A_n'=A_n\setminus\bigcup_{k=1}^{n-1} A_k$. Está claro que $A=\bigcup_{n=1}^{\infty} A_n'$ y que los conjuntos A_n' son disjuntos dos a dos. En virtud del teorema 6 y de su corolario, todos los conjuntos A_n' son medibles. En virtud del teorema 7 y de la definición de la medida superior, para cualquier n finito

$$\sum_{k=1}^{n} \mu(A_k')=\mu\left(\bigcup_{k=1}^{n} A_k'\right)\leqslant \mu^*(A),$$

por lo que la serie

$$\sum_{n=1}^{\infty} \mu(A_n')$$

converge; de manera que para todo $\varepsilon>0$ existe un N tal que

$$\sum_{n>N} \mu(A_n')<\frac{\varepsilon}{2}. \tag{12}$$

Por ser medible el conjunto $C=\bigcup_{n=1}^{N} A_n'$ (como unión de un número finito de conjuntos medibles), existe un conjunto elemental B tal que

$$\mu^*(C\triangle B)<\frac{\varepsilon}{2}. \tag{13}$$

Como

$$A\triangle B\subset(C\triangle B)\cup\left(\bigcup_{n>N} A_n'\right),$$

de (12) y (13) se deduce que

$$\mu^*(A\triangle B) < \varepsilon.$$

En virtud del teorema 5, esto significa que el conjunto A es medible.

Puesto que los complementos de conjuntos medibles son medibles, la parte del teorema correspondiente a las intersecciones se desprende de la igualdad

$$\bigcap_n A_n = E\setminus\bigcup_n (E\setminus A_n).$$

El teorema 8 es una generalización del teorema 6. El teorema que sigue constituye una generalización correspondiente del teorema 7.

TEOREMA 9. *Si $\{A_n\}$ es una sucesión de conjuntos medibles disjuntos dos a dos y si $A=\bigcup\limits_n A_n$, se tiene*

$$\mu(A)=\sum_n \mu(A_n).$$

DEMOSTRACION. Según el teorema 7, para cualquier N

$$\mu\left(\bigcup_{n=1}^{N} A_n\right)=\sum_{n=1}^{N}\mu(A_n) < \mu(A).$$

Pasando al límite para $N\to\infty$, obtenemos

$$\mu(A)\geqslant\sum_{n=1}^{\infty}\mu(A_n). \tag{14}$$

Por otro lado, según el teorema 3,

$$\mu(A)\leqslant\sum_{n=1}^{\infty}\mu(A_n). \tag{15}$$

De (14) y (15) se desprende la afirmación del teorema.

La propiedad de la medida establecida en el teorema 9 es llamada *aditividad numerable* o *σ-aditividad*. De la σ-aditividad se deduce la siguiente propiedad de la medida llamada *continuidad*.

TEOREMA 10. *Si $A_1\supset A_2\supset\ldots$ es una sucesión de conjuntos medibles sumergidos unos en otros y si $A=\bigcap\limits_n A_n$, se tiene*

$$\mu(A)=\lim_{n\to\infty}\mu(A_n).$$

Obviamente bastará considerar el caso $A=\varnothing$, ya que el caso general se reduce a éste sustituyendo A_n por $A_n \setminus A$. Entonces,

$$A_1=(A_1\setminus A_2)\cup(A_2\setminus A_3)\cup\ldots$$

y

$$A_n=(A_n\setminus A_{n+1})\cup(A_{n+1}\setminus A_{n+2})\cup\ldots$$

Por consiguiente,

$$\mu(A_1)=\sum_{k=1}^{\infty}\mu(A_k\setminus A_{k+1}) \tag{16}$$

y

$$\mu(A_n)=\sum_{k=n}^{\infty}\mu(A_k\setminus A_{k+1}); \tag{17}$$

como la serie (16) converge, su resto (17) tiende a cero para $n\to\infty$. De manera que

$$\mu(A_n)\to 0 \text{ para } n\to\infty,$$

que es lo que necesitábamos demostrar.

COROLARIO. *Si* $A_1\subset A_2\subset\ldots$ *es una sucesión creciente de conjuntos medibles y si*

$$A=\bigcup_n A_n$$

se tiene

$$\mu(A)=\lim_{n\to\infty}\mu(A_n)$$

Para demostrarlo es suficiente pasar de los conjuntos A_n a sus complementos y recurrir al teorema 10.

De esta forma hemos extendido la medida de conjuntos elementales a una clase más amplia de conjuntos, llamados medibles, cerrada respecto a las operaciones de unión e intersección numerables. La medida construida es σ-aditiva en esta clase de conjuntos. Los teoremas demostrados permiten hacerse una idea de la clase de todos los conjuntos medibles según Lebesgue.

Como todo conjunto cerrado, contenido en E, se puede representar como la unión de un número finito o numerable de rectángulos abiertos, esto es, de conjuntos medibles, todos los conjuntos abiertos son, en virtud del teorema 8, medibles. Los conjuntos cerrados son complementos de los abiertos y, consecuentemente, son también medibles. Según el teorema 8, serán también medibles todos aquellos conjuntos que se puedan obtener a partir de conjuntos abiertos y cerrados mediante un número finito o numerable de operaciones consistentes en considerar uniones o intersecciones numerables. Se puede demostrar, sin embargo, que con estos conjuntos no se agota la clase de todos los conjuntos medibles según Lebesgue.

4°. Algunos suplementos y generalizaciones. Hemos considerado anteriormente sólo aquellos conjuntos del plano que son subconjuntos del cuadrado unidad $E=\{0\leqslant x,\ y\leqslant 1\}$. No es difícil librarse de esta restricción, por ejemplo, del siguiente modo. Considerando todo el plano como la unión de los cuadrados $E_{nm}=\{n\leqslant x\leqslant n+1,\ m\leqslant y\leqslant m+1\}$ (n, m son números enteros), diremos que un conjunto plano A es medible cuando es medible su intersección $A_{nm}=A_n\cap E_{nm}$ con cada uno de estos cuadrados y cuando la serie $\sum\limits_{n,m}\mu(A_{nm})$ converge, tomando por definición

$$\mu(A)=\sum_{n,m}\mu(A_{nm}).$$

Todas las propiedades de la medida que hemos establecido anteriormente se extienden de manera obvia a este caso.

En este parágrafo hemos expuesto la construcción de la medida de Lebesgue para los conjuntos planos. De manera análoga se puede construir la medida de Lebesgue en la recta, en el espacio de tres dimensiones y, en general, en un espacio euclídeo de cualquier dimensión n. En todos estos casos la medida es construida siguiendo las mismas ideas: a partir de la medida definida de antemano para un sistema de conjuntos elementales (rectángulos en el caso del plano; intervalos (a, b), segmentos $[a, b]$ y semisegmentos $(a, b]$ y $[a, b)$ en el caso de la recta; etc.) definimos primero la medida para uniones finitas de estos conjuntos, extendiéndola después a una clase mucho más amplia de conjuntos, a la clase de conjuntos medibles según Lebesgue. Para conjuntos de un espacio de cualquier dimensión la propia definición de conjunto medible se conserva textualmente.

Al introducir el concepto de medida de Lebesgue hemos partido de la definición habitual del área. En el caso unidimensional la construcción análoga se basa en el concepto de la longitud de un intervalo (de un segmento, de un semisegmento). No obstante, es posible introducir en este caso el concepto de la medida de otra forma, algo más general (que frecuentemente aparece en la práctica).

Sea $F(t)$ una función no decreciente y continua a la izquierda definida en la recta. Pongamos

$$\begin{aligned} m(a,\ b)&=F(b)-F(a+0),\\ m[a,\ b]&=F(b+0)-F(a),\\ m(a,\ b]&=F(b+0)-F(a+0),\\ m[a,\ b)&=F(b)-F(a). \end{aligned}$$

Es fácil ver que la función de intervalo m, definida de esta forma, es no negativa y aditiva. Aplicando a ella razonamientos análogos a los realizados en este parágrafo, podemos construir una «medida» $\mu_F(A)$. La clase de conjuntos medibles según esta medida será cerrada respecto a las uniones e intersecciones numerables, mientras que la medida μ_F será σ-aditiva. La clase de conjuntos medibles según μ_F dependerá, en general, de la selección de la función F. Sin embargo, cualquiera que sea F, los conjuntos abiertos y cerrados y, por consiguiente, todas las uniones e intersecciones numerables de los mismos serán medibles. Las medidas que se obtienen a partir de una u otra función F se llaman *medidas de Lebesgue — Stieltjes*. En particular, a la función $F(t) \equiv t$ corresponde la medida corriente de Lebesgue en la recta.

Una medida μ_F que se anula en cualquier conjunto, cuya medida corriente de Lebesgue es igual a 0, se llama *absolutamente continua*. Una medida μ_F concentrada totalmente en un conjunto finito o numerable de puntos (esto ocurrirá cada vez que el conjunto de valores de la función $F(t)$ sea finito o numerable) se llama *discreta*. Una medida μ_F se llama *singular* cuando es igual a cero para cualquier conjunto compuesto de un punto y cuando existe un conjunto M de medida de Lebesgue igual a 0 tal que la medida μ_F de su complemento es igual a 0.

Se puede demostrar que toda medida μ_F es suma de una medida absolutamente continua, una medida discreta y una medida singular. A las medidas de Lebesgue—Stieltjes volveremos en el capítulo siguiente.

Existencia de conjuntos no medibles. Como se ha demostrado, la clase de conjuntos medibles según Lebesgue es muy amplia. Surge la pregunta natural de si existen, en general, conjuntos no medibles. Vamos a demostrar que este problema se resuelve positivamente. Lo más sencillo es construir conjuntos no medibles en la circunferencia.

Sea C una circunferencia de longitud 1 y sea α un número irracional. Asignemos a una misma clase aquellos puntos de la circunferencia C que se transforman unos en otros mediante una rotación de la circunferencia C de valor angular $n\alpha\pi$ (n es un número entero). Cada una de estas clases quedará compuesta, obviamente, por un conjunto numerable de puntos. Escojamos ahora un punto en cada una de estas clases. Probemos que el conjunto obtenido de esta forma (denotémoslo con Φ_0) no es medible. Sea Φ_n el conjunto que se obtiene de Φ_0 por una rotación de valor angular $n\alpha\pi$. Es fácil ver que los conjuntos Φ_n son disjuntos dos a dos y que la unión de ellos es la circunferencia C. Si el conjunto Φ_0 fuese medible, también serían medibles los conjuntos Φ_n congruentes a él. Como

$$C = \bigcup_{n=-\infty}^{\infty} \Phi_n, \quad \Phi_n \cap \Phi_m = \varnothing \text{ para } n \neq m,$$

podríamos concluir de aoui, debido a la σ-aditividad de la medida, que

$$1 = \sum_{n=-\infty}^{\infty} \mu(\Phi_n). \tag{18}$$

Pero los conjuntos congruentes tienen la misma medida; luego, si Φ_0 es medible, tenemos

$$\mu(\Phi_n) = \mu(\Phi).$$

Esto demuestra que la igualdad (18) es imposible, ya que la suma de la serie, que figura en el miembro derecho de la igualdad (18), es igual a 0 cuando $\mu(\Phi_0) = 0$ y es infinita cuando $\mu(\Phi_0) > 0$. De manera que el conjunto Φ_0 (y, consecuentemente, cualquier conjunto Φ_n) no es medible.

§ 2. CONCEPTO GENERAL DE MEDIDA. PROLONGACION DE UNA MEDIDA DE UN SEMIANILLO A UN ANILLO. ADITIVIDAD Y σ-ADITIVIDAD [1]

1°. Definición de medida. Al construir la medida de conjuntos planos hemos partido de la medida (el área) de un rectángulo, extendiendo después el concepto de medida a una clase más amplia de conjuntos. Lo esencial en la construcción, expuesta en el parágrafo anterior, no es de ninguna manera la expresión concreta del área de un rectángulo; son esenciales para esta construcción dos hechos generales: 1) el área de un rectángulo es una función de conjunto no negativa que satisface la condición de aditividad, esto es,

$$m(P_1 \cup P_2) = m(P_1) + m(P_2) \text{ cuando } P_1 \cap P_2 = \varnothing$$

y 2) el conjunto de rectángulos constituye un semianillo de conjuntos. Por esto, a la construcción expuesta en el § 1 para el caso de conjuntos planos se le puede dar una forma totalmente abstracta y general. Con ello se ampliará sustancialmente la posibilidad de aplicar nuestras construcciones. A esto están dedicados los dos parágrafos que siguen.

Introduzcamos, ante todo, la siguiente definición fundamental.

DEFINICION 1. Una función $\mu(A)$ de conjunto se llama *medida* cuando

---

[1] En este parágrafo en lo que sigue emplearemos sistemáticamente los conceptos y resultados expuestos en el § 5 del cap. I.

1) el campo de definición $\mathfrak{S}_\mu$ de la función $\mu(A)$ es un semianillo de conjuntos;

2) los valores de la función $\mu(A)$ son reales y no negativos;

3) $\mu(A)$ es aditiva, esto es, para cualquier descomposición finita

$$A = A_1 \cup \ldots \cup A_n$$

de un conjunto $A \in \mathfrak{S}_\mu$ en conjuntos $A_k \in \mathfrak{S}_\mu$ se verifica la igualdad

$$\mu(A) = \sum_{k=1}^{n} \mu(A_k).$$

Observación. De la descomposición $\varnothing = \varnothing \cup \varnothing$ se deduce que $\mu(\varnothing = 2\mu(\varnothing)$, es decir, $\mu(\varnothing) = 0$.

Los dos teoremas que vienen a continuación sobre medidas en semianillos se emplearán frecuentemente en lo sucesivo.

TEOREMA 1. *Sea μ una medida definida en un semianillo $\mathfrak{S}_\mu$. Si los conjuntos $A_1, \ldots, A_n$, A pertenecen a $\mathfrak{S}_\mu$ y A_k son subconjuntos disjuntos dos a dos de A, se tiene*

$$\sum_{k=1}^{n} \mu(A_k) \leqslant \mu(A).$$

DEMOSTRACION. Por ser $\mathfrak{S}_\mu$ un semianillo existe, en virtud del lema 1 del § 5 del capítulo I, la descomposición

$$A = \bigcup_{k=1}^{s} A_k, \quad s \geqslant n, \quad A_k \in \mathfrak{S}_\mu,$$

donde los n primeros conjuntos coinciden con los conjuntos dados $A_1, \ldots, A_n$. Como la medida de cualquier conjunto es no negativa, tenemos

$$\sum_{k=1}^{n} \mu(A_k) \leqslant \sum_{k=1}^{s} \mu(A_k) = \mu(A).$$

TEOREMA 2. *Si $A_1, \ldots, A_n$, A pertenecen a $\mathfrak{S}_\mu$ y $A \subset \bigcup_{k=1}^{n} A_k$, se tiene*

$$\mu(A) \leqslant \sum_{k=1}^{n} \mu(A_k).$$

DEMOSTRACION. En virtud del lema 2 del § 5 del capítulo I, existe un sistema de conjuntos disjuntos dos a dos $B_1, \ldots, B_t$

de $\mathfrak{S}_\mu$ tal que cualquier conjunto $A_1, A_2, \ldots, A_n, A$ se puede representar como la unión de determinados conjuntos B_s:

$$A = \bigcup_{s \in M_0} B_s, \quad A_k = \bigcup_{s \in M_k} B_s, \quad k = 1, 2, \ldots, n,$$

donde cada índice $s \in M_0$ pertenece también a un M_k. Luego, cada término de la suma

$$\sum_{s \in M_0} \mu(B_s) = \mu(A)$$

figura una o varias veces en la suma doble

$$\sum_{k=1}^{n} \sum_{s \in M_k} \mu(B_s) = \sum_{k=1}^{n} \mu(A_k).$$

De aquí se desprende precisamente que

$$\mu(A) \leqslant \sum_{k=1}^{n} \mu(A_k).$$

En particular para $n = 1$ tenemos el resultado siguiente.

COROLARIO. *Si* $A \subset A'$, *se tiene* $\mu(A) \leqslant \mu(A')$.

2°. Prolongación de una medida en un semianillo al anillo generado. El primer paso en construir la medida de conjuntos planos consistió en extender el concepto de medida de un rectángulo a los conjuntos elementales, es decir, a las uniones finitas de rectángulos disjuntos dos a dos. Veamos ahora el análogo abstracto de este problema. Enunciemos, ante todo, la siguiente definición.

DEFINICION 2. Una medida μ se llama *prolongación* de una medida m cuando $\mathfrak{S}_m \subset \mathfrak{S}_\mu$ y cuando para todo $A \in \mathfrak{S}_m$ se cumple la igualdad

$$\mu(A) = m(A).$$

El objeto de este punto es demostrar la proposición que sigue.

TEOREMA 3. *Para cada* $m(A)$, *definida en un semianillo* $\mathfrak{S}_m$, *existe una prolongación* $\mu(A)$, *y sólo una, que tiene como campo de definición el anillo* $\mathfrak{R}(\mathfrak{S}_m)$ *(esto es, el anillo minimal sobre* $\mathfrak{S}_m$*).*

DEMOSTRACION. Para todo conjunto $A \in \mathfrak{R}(\mathfrak{S}_m)$ existe la descomposición

$$A = \bigcup_{k=1}^{n} B_k \quad (B_k \in \mathfrak{S}_m) \tag{1}$$

(teorema 3, § 5, capítulo I). Tomemos por definición

$$\mu(A)=\sum_{k=1}^{n} m(B_k). \tag{2}$$

Es fácil ver que la magnitud $\mu(A)$, definida por la igualdad (2), no depende de cómo se escoja la descomposición (1).
En efecto, consideremos dos descomposiciones

$$A=\bigcup_{i=1}^{r} B_i=\bigcup_{j=1}^{n} C_j,\quad B_i\in\mathfrak{S}_m,\quad C_j\in\mathfrak{S}_m.$$

Como todas las intersecciones $B_i \cap C_j$ pertenecen a $\mathfrak{S}_m$, tenemos, en vista de la aditividad de la medida m,

$$\sum_{i=1}^{r} m(B_i)=\sum_{i=1}^{r}\sum_{j=1}^{n} m(B_i\cap C_j)=\sum_{j=1}^{n} m(C_j),$$

que es lo que queríamos demostrar. Es evidente, que la función $\mu(A)$, definida por la igualdad (2), es no negativa y aditiva. Luego, hemos demostrado la existencia de la prolongación μ de la medida m al anillo $\mathfrak{R}(\mathfrak{S}_m)$. Para demostrar su unicidad observemos que, de acuerdo a la definición de prolongación, si $A=\bigcup_{k=1}^{n} B_k$, donde B_k son conjuntos disjuntos de $\mathfrak{S}_m$, tenemos para cualquier prolongación $\tilde{\mu}$ de la medida m al anillo $\mathfrak{R}(\mathfrak{S}_m)$

$$\tilde{\mu}(A)=\sum\tilde{\mu}(B_k)=\sum m(B_k)=\mu(A),$$

es decir, la medida $\tilde{\mu}$ coincide con la medida μ definida por la igualdad (2). El teorema queda demostrado.

Este teorema constituye, de hecho, la repitición, en términos abstractos, de la construcción realizada en el § 1 al prolongar la medida de rectángulos a la clase de conjuntos elementales que representa precisamente el anillo minimal sobre el semianillo de rectángulos.

3°. Aditividad numerable. En diferentes cuestiones del Análisis es preciso considerar, además de uniones finitas, uniones de un número numerable de conjuntos. En este orden la condición de aditividad, a la que hemos sometido las medidas (definición 1) resulta insuficiente y es natural sustituirla por una condición más fuerte de la así llamada aditividad numerable.

DEFINICION 3. Una medida μ se llama *aditiva numerable* (o σ-aditiva) cuando para cualesquiera conjuntos $A, A_1, A_2, \ldots, A_n, \ldots,$

que pertenecen a su campo de definición $\mathfrak{S}_\mu$ y que verifican las condiciones

$$A = \bigcup_{n=1}^{\infty} A_n, \quad A_i \cap A_j = \varnothing \text{ para } i \neq j,$$

tiene lugar la igualdad

$$\mu(A) = \sum_{n=1}^{\infty} \mu(A_n).$$

La medida plana de Lebesgue, construida en el § 1, es σ-aditiva (teorema 9). Un ejemplo de una medida σ-aditiva de una naturaleza totalmente distinta se puede obtener del siguiente modo. Sea

$$X = \{x_1, x_2, \ldots\}$$

un conjunto numerable arbitrario y sean los números $p_n > 0$ tales que

$$\sum_{n=1}^{\infty} p_n = 1.$$

El campo $\mathfrak{S}_m$ se compone de todos los subconjuntos del conjunto X. Tomemos para cada $A \subset X$

$$\mu(A) = \sum_{x_n \in A} p_n.$$

Es fácil ver que $\mu(A)$ será una medida σ-aditiva y que $\mu(X) = 1$. Este ejemplo surge de un modo natural en diferentes cuestiones de la Teoría de probabilidades.

Señalemos un ejemplo de una medida aditiva que no es σ-aditiva. Sea X el conjunto de todos los puntos racionales del segmento $[0, 1]$ y sea $\mathfrak{S}_\mu$ el conjunto formado por las intersecciones del conjunto X con intervalos (a, b), segmentos $[a, b]$ o semisegmentos $(a, b]$ y $[a, b)$ arbitrarios. Es fácil ver que $\mathfrak{S}_\mu$ forma un semianillo. Tomemos para cada conjunto de este tipo

$$\mu(A_{ab}) = b - a.$$

Esta medida es aditiva, pero no es σ-aditiva, ya que $\mu(X) = 1$ y al mismo tiempo X es la unión de un número numerable de puntos cada uno de los cuales tiene medida 0.

Las medidas que consideraremos ahora y en el parágrafo siguiente se suponen σ-aditivas.

TEOREMA 4. *Si una medida m definida en un semianillo $\mathfrak{S}_m$ es σ-aditiva, también es σ-aditiva la medida $\mu = r(m)$ que se obtiene prolongándola al anillo $\mathfrak{R}(\mathfrak{S}_m)$.*

DEMOSTRACION. Sea

$$A \in \mathfrak{R}(\mathfrak{S}_m), \quad B_n \in \mathfrak{R}(\mathfrak{S}_m), \quad n=1, 2, \ldots,$$

y sea

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

donde $B_s \cap B_r = \varnothing$ para $s \neq r$. Entonces, existen conjuntos A_j y B_{ni} de $\mathfrak{S}_m$ tales que

$$A = \bigcup_j A_j, \quad B_n = \bigcup_i B_{ni},$$

donde los conjuntos que figuran en los miembros derechos de cada una de estas igualdades son disjuntos dos a dos y las uniones respecto a i y j son finitas (teorema 3, § 5, capítulo I).

Sea $C_{nij} = B_{ni} \cap A_j$. Es fácil ver que los conjuntos C_{nij} son disjuntos dos a dos y que

$$A_j = \bigcup_n \bigcup_i C_{nij},$$

$$B_{ni} = \bigcup_j C_{nij}.$$

Luego, debido a la σ-aditividad de la medida m sobre $\mathfrak{S}_m$, tenemos

$$m(A_j) = \sum_n \sum_i m(C_{nij}), \tag{3}$$

$$m(B_{ni}) = \sum_j m(C_{nij}) \tag{4}$$

y de acuerdo a la definición de la medida $\mu = r(m)$ sobre $\mathfrak{R}(\mathfrak{S})$, tenemos

$$\mu(A) = \sum_j m(A_j), \tag{5}$$

$$\mu(B_n) = \sum_i m(B_{ni}). \tag{6}$$

De (3), (4), (5) y (6) se desprende que $\mu(A) = \sum_n \mu(B_n)$. (Las sumas respecto a i y j son finitas y las series respecto a n convergen.)

Demostremos ahora las siguientes propiedades fundamentales de medidas σ-aditivas que constituyen una generalización de los teoremas 1 y 2 al caso de uniones numerables de conjuntos.

TEOREMA 5. *Sea m una medida σ-aditiva y sean $A_1, A_2, \ldots, A_n, \ldots$ conjuntos pertenecientes a $\mathfrak{S}_m$.* Entonces, Iσ si $\bigcup_{k=1}^{\infty} A_k \subset A$ y $A_i \cap A_j = \varnothing$ para $i \neq j$, se tiene

$$\sum_{k=1}^{\infty} m(A_k) \leqslant m(A);$$

IIσ si $\bigcup_{k=1}^{\infty} A_k \supset A$, se tiene

$$\sum_{k=1}^{\infty} m(A_k) \geqslant m(A).$$

DEMOSTRACION. Si todos los A_k son disjuntos y están contenidos en A, tenemos para cualquier n

$$\sum_{k=1}^{n} m(A_k) \leqslant m(A),$$

en virtud del teorema 1. Pasando aquí al límite para $n \longrightarrow \infty$, obtenemos la primera afirmación del teorema.

En cuanto a la segunda afirmación, es suficiente demostrarla, de acuerdo con el teorema 4, para medidas definidas sobre un anillo, ya que de la validez de la proposición IIσ para $\mu = r(m)$ se deduce directamente su validez para la medida m. Siendo $\mathfrak{S}_m$ un anillo, los conjuntos

$$B_n = (A \cap A_n) \setminus \bigcup_{k=1}^{n-1} A_k$$

pertenecen a $\mathfrak{S}_m$. Como

$$A = \bigcup_{n=1}^{\infty} B_n, \quad B_n \subset A$$

y los conjuntos B_n son disjuntos dos a dos, tenemos

$$m(A) = \sum_{n=1}^{\infty} m(B_n) \leqslant \sum_{n=1}^{\infty} m(A_n).$$

Observación. La afirmación Iσ del teorema demostrado no emplea, obviamente, la σ-aditividad de la medida considerada y sigue siendo válida para cualesquiera medidas aditivas. La afirmación IIσ, al contrario, se basa de un modo sustancial, en la σ-aditividad de la medida. Efectivamente, en el ejemplo dado anteriormente de una medida aditiva, pero no σ-aditiva, todo el espacio X, de medida 1, es cubierto por una unión numerable de conjuntos, compuestos de un solo punto, que tienen medida 0. Es más, no es difícil demostrar que la condición IIσ es, en realidad, equivalente a la σ-aditividad. En efecto, sea μ una medida y sean A, A_1, ..., A_n, ... conjuntos de $\mathfrak{S}_\mu$ tales que todos los A_k son disjuntos dos a dos y $A = \bigcup A_k$. Entonces, en virtud de la condición Iσ (que es válida, como hemos visto, para cualquier medida), se tiene

$$\sum_{k=1}^{\infty} \mu(A_k) \leqslant \mu(A).$$

Si μ verifica además la condición IIσ, tendremos (ya que los conjuntos A_k cubren A)

$$\sum_{k=1}^{\infty} \mu(A_k) \geqslant \mu(A)$$

de manera que

$$\sum_{k=1}^{\infty} \mu(A_k) = \mu(A).$$

En la práctica resulta con frecuencia más fácil comprobar que una medida verifica la condición IIσ que demostrar su σ-aditividad.

§ 3. PROLONGACIÓN DE LEBESGUE DE UNA MEDIDA

1°. Prolongación de Lebesgue de una medida definida en un semianillo con unidad. Si la medida m definida en un semianillo $\mathfrak{S}_m$ verifica sólo la condición de aditividad (pero no es σ-aditiva), su prolongación a $\mathfrak{R}(\mathfrak{S}_m)$, obtenida por el procedimiento descrito en el parágrafo anterior, agota en gran medida, las posibilidades de extender la medida del semianillo inicial a una clase más amplia de conjuntos. En cambio, si la medida considerada es σ-aditiva, puede ser extendida de $\mathfrak{S}_m$ a un sistema de conjuntos mucho más amplio que el anillo $\mathfrak{R}(\mathfrak{S}_m)$. La prolongación de una medida σ-aditiva, definida en un semianillo, a una clase de conjuntos, en cierto sentido maximal, se puede realizar mediante la así llamada prolongación de Lebesgue. Consideremos primero la prolongación de Lebesgue de una medida, definida en un semianillo con unidad. El caso general será estudiado en el punto siguiente.

Supongamos que en un semianillo de conjuntos $\mathfrak{S}_m$ con unidad E está definida una medida m σ-aditiva. Definamos en el sistema $\mathfrak{A}$ de todos los subconjuntos del conjunto E las funciones $\mu^*(A)$ y $\mu_*(A)$ del siguiente modo.

DEFINICION 1. Se llama *medida superior* del conjunto $A \subset E$ al número

$$\mu^*(A) = \inf_{A \subset \bigcup B_n} \sum_n m(B_n),$$

donde la cota inferior se toma respecto a todos los cubrimientos del conjunto A mediante sistemas finitos o numerables de conjuntos $B_n \in \mathfrak{S}_m$.

DEFINICION 2. Se llama *medida inferior* de un conjunto $A \subset E$ al número

$$\mu_*(A) = m(E) - \mu^*(E \setminus A).$$

El teorema 5 del § 2 implica que siempre $\mu_*(A) \leqslant \mu^*(A)$.

DEFINICION 3. Un conjunto $A \subset E$ se llama *medible* (según Lebesgue) cuando

$$\mu_*(A) = \mu^*(A).$$

Siendo A medible, el valor común $\mu_*(A) = \mu^*(A)$ es denotado mediante $\mu(A)$ y llamado medida (de Lebesgue) del conjunto A.

Si A es medible, también será, evidentemente, medible su complemento.

En vista del teorema 5 del § 2, para cualquier prolongación σ-aditiva $\tilde{\mu}$ de una medida m, definida en un semianillo, tiene lugar la desigualdad

$$\mu_*(A) \leqslant \tilde{\mu}(A) \leqslant \mu^*(A).$$

Consecuentemente, para un conjunto medible A toda prolongación σ-aditiva $\tilde{\mu}$ de una medida m (si esta prolongación está definida en A) toma necesariamente el valor $\mu_*(A) = \mu^*(A)$. La medida de Lebesgue no es otra cosa que la prolongación σ-aditiva de la medida m a la clase de todos los conjuntos medibles en el sentido de la definición 3. Es obvio que la definición de conjunto medible se puede enunciar también así:

DEFINICION 3'. Un conjunto $A \subset E$ se llama *medible* cuando

$$\mu^*(A) + \mu^*(E \setminus A) = m(E).$$

Conviene emplear, además de la medida inicial m, su prolongación $m' = r(m)$ al anillo $\mathfrak{R}(\mathfrak{S}_m)$ considerada anteriormente (§ 2). Está claro que la definición 1 es equivalente a la siguiente.

DEFINICION 1'. Se llama *medida superior* de un conjunto A al número

$$\mu^*(A) = \inf_{A \subset \bigcup_n B'_n} \sum_n m'(B'_n), \quad B'_n \subset \mathfrak{R}(\mathfrak{S}_m).$$

En efecto, como la medida m' es σ-aditiva (teorema 4 del § 2), cualquier suma $\sum_n m'(B'_n)$, donde $B'_n \in \mathfrak{R}(\mathfrak{S}_m)$, puede ser sustituida por la suma equivalente

$$\sum_{n,\,k} m(B_{nk}),\quad B_{nk}\in\mathfrak{S}_m,$$

donde $B'_n=\bigcup_k B_{nk}$ y $B_{ni}\cap B_{nj}=\varnothing$ para $i\neq j$.

Los resultados que siguen son fundamentales para la exposición ulterior.

TEOREMA 1. *Si*

$$A\subset\bigcup_n A_n,$$

donde $\{A_n\}$ es un sistema finito o numerable de conjuntos, se tiene

$$\mu^*(A)\leqslant\sum_n\mu^*(A_n).$$

TEOREMA 2. *Si $A\in\mathfrak{R}(\mathfrak{S}_m)$, se tiene $\mu_*(A)=m'(A)=\mu^*(A)$, es decir, todos los conjuntos de $\mathfrak{R}(\mathfrak{S}_m)$ son medibles y las medidas superior e inferior de los mismos coinciden con m'.*

TEOREMA 3. *Para que un conjunto A sea medible es necesaria y suficiente la siguiente condición:*

para cualquier $\varepsilon>0$ existe un $B\in\mathfrak{R}(\mathfrak{S}_m)$ tal que

$$\mu^*(A\triangle B)<\varepsilon.$$

En el § 1 estas proposiciones han sido demostradas para la medida plana de Lebesgue (teoremas 3, 4 y 5 del § 1). Las demostraciones dadas allí siguen siendo válidas en el caso general que estamos considerando y por eso no las repetimos.

TEOREMA 4. *El sistema $\mathfrak{M}$ de todos los conjuntos medibles es un anillo.*

DEMOSTRACION. Como

$$A_1\cap A_2=A_1\setminus(A_1\setminus A_2)$$

y

$$A_1\cup A_2=E\setminus[(E\setminus A_1)\cap(E\setminus A_2)],$$

basta demostrar que si $A_1\in\mathfrak{M}$ y $A_2\in\mathfrak{M}$, también

$$A=A_1\setminus A_2\in\mathfrak{M}.$$

Supongamos que A_1 y A_2 son medibles; en este caso existen $B_1\in\mathfrak{R}(\mathfrak{S}_m)$ y $B_2\in\mathfrak{R}(\mathfrak{S}_m)$ tales que

$$\mu^*(A_1\triangle B_1)<\frac{\varepsilon}{2}\ \text{y}\ \mu^*(A_2\triangle B_2)<\frac{\varepsilon}{2}.$$

Tomando $B = B_1 \setminus B_2 \in \mathfrak{R}(\mathfrak{S}_m)$ y empleando la relación

$$(A_1 \setminus A_2) \triangle (B_1 \setminus B_2) \subset (A_1 \triangle B_1) \cup (A_2 \triangle B_2)$$

encontramos

$$\mu^*(A \triangle B) < \varepsilon.$$

Como $\varepsilon > 0$ es arbitrario, de aquí se desprende que A es medible.

Observación. Es evidente que E constituye la unidad del anillo $\mathfrak{M}$ que de esta forma resulta ser un álgebra de conjuntos.

TEOREMA 5. *La función $\mu(A)$ es aditiva en el sistema $\mathfrak{M}$ de los conjuntos medibles.*

La demostración de este teorema es una repetición verbal de la demostración del teorema 7 del § 1.

TEOREMA 6. *La función $\mu(A)$ es σ-aditiva en el sistema $\mathfrak{M}$ de los conjuntos medibles.*

DEMOSTRACION. Sea

$$A = \bigcup_{n=1}^{\infty} A_n, \quad A, \quad A_i \in \mathfrak{M}; \quad A_i \cap A_j = \varnothing \quad \text{para} \quad i \neq j.$$

En virtud del teorema 1,

$$\mu^*(A) \leqslant \sum_n \mu(A_n), \tag{1}$$

y, de acuerdo al teorema 5,

$$\mu^*(A) \geqslant \mu^*\left(\bigcup_{n=1}^{N} A_n\right) = \sum_{n=1}^{N} \mu^*(A_n)$$

para cualquier N, de donde

$$\mu^*(A) \geqslant \sum_n \mu(A_n). \tag{2}$$

De (1) y (2) se deduce la afirmación del teorema.

Hemos demostrado de esta forma que la función $\mu(A)$, definida en el sistema $\mathfrak{M}$, posee todas las propiedades de una medida σ-aditiva.

Ello justifica la siguiente definición.

DEFINICION 4. Se llama *prolongación de Lebesgue* $\mu = L(m)$ de una medida m a la función $\mu(A)$, definida en el sistema $\mathfrak{S}_\mu = \mathfrak{M}$ de los conjuntos medibles y coincidente en este sistema con la medida superior $\mu^*(A)$.

En el § 1 hemos demostrado, al considerar la medida plana de Lebesgue, que son medibles no sólo las uniones e intersecciones finitas de conjuntos medibles, sino también las uniones e intersecciones numerables de conjuntos medibles. Esto sigue siendo válido también en el caso general, es decir, tiene lugar el siguiente teorema.

TEOREMA 7. *El sistema $\mathfrak{M}$ de conjuntos medibles según Lebesgue constituye un álgebra de Borel con unidad E.*

DEMOSTRACION. Como

$$\bigcap_n A_n = E \setminus \bigcup_n (E \setminus A_n)$$

y puesto que el complemento de un conjunto medible es medible, basta demostrar lo siguiente: si $A_1, A_2, \ldots, A_n, \ldots$ pertenecen a $\mathfrak{M}$, también $A = \bigcup_n A_n$ pertenece a $\mathfrak{M}$. La demostración de esta proposición dada en el teorema 8 del § 1 para los conjuntos planos se conserva textualmente en el caso general.

Al igual que en el caso de medida plana de Lebesgue, la σ-aditividad de la medida implica su continuidad, esto es, siendo μ una medida σ-aditiva definida en una B-álgebra y siendo $A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots$ una cadena decreciente de conjuntos medibles tal que

$$A = \bigcap_n A_n,$$

se tiene

$$\mu(A) = \lim_{n \to \infty} \mu(A_n)$$

y siendo $A_1 \subset A_2 \subset \ldots \subset A_n \subset \ldots$ una cadena creciente de conjuntos medibles tal que

$$A = \bigcup_n A_n,$$

se tiene

$$\mu(A) = \lim_{n \to \infty} \mu(A_n).$$

La demostración, dada para la medida plana en el § 1 (teorema 10), se extiende textualmente al caso general.

2°. Prolongación de una medida definida en un semianillo sin unidad. Si el semianillo $\mathfrak{S}_m$, en el cual está definida la medida inicial m, no tiene unidad, la construcción de la prolongación de Lebesgue, expuesta en el punto anterior, debe ser modificada. La definición 1 de la medida superior se conserva, pero la medida superior μ^* estará definida sólo en el sistema S_{μ^*} de aquellos conjuntos A para cada uno de los cuales existe un cubrimiento $\bigcup_n B_n$ mediante conjuntos de $\mathfrak{S}_m$ de suma finita

$$\sum_n m(B_n).$$

La definición 2 pierde su sentido. La medida inferior puede ser definida (de una manera algo distinta) también en el caso general; pero no lo haremos. Conviene definir ahora el concepto de conjunto medible a partir de la propiedad de conjuntos medibles señalada en el teorema 3.

DEFINICION 5. Un conjunto A se llama *medible* cuando para cualquier $\varepsilon > 0$ existe un conjunto $B \in \mathfrak{R}(\mathfrak{S}_m)$ tal que $\mu^*(A \triangle B) < \varepsilon$.

Los teoremas 4, 5 y 6 y la definición 4 subsisten. La existencia de la unidad ha sido empleada sólo durante la demostración del teorema 4. Para demostrar el teorema 4 en el caso general, debemos probar de una manera independiente que $A_1 \in \mathfrak{M}$ y $A_2 \in \mathfrak{M}$ implican que $A_1 \cup A_2 \in \mathfrak{M}$. Pero esto se desprende de la inclusión

$$A_1 \cup A_2 \triangle (B_1 \cup B_2) \subset (A_1 \triangle B_1) \cup (A_2 \triangle B_2).$$

En el caso en que $\mathfrak{S}$ no tiene unidad, el teorema 7 es sustituido por el teorema siguiente.

TEOREMA 8. *Cualquiera que sea la medida inicial m, el sistema $\mathfrak{M} = \mathfrak{S}_{L(m)}$ de conjuntos medibles según Lebesgue es un σ-anillo; siendo A_n medibles el conjunto $A = \bigcup_N A_n$ es medible cuando, y sólo cuando, las medidas $\mu\left(\bigcup_{n=1} A_n\right)$ están acotadas por una constante que no depende de N.*

Dejamos la demostración de este teorema a cargo del lector.

Observación. En nuestra exposición la medida es siempre finita, de manera que la necesidad de la última condición es obvia.

Del teorema 8 resulta:

COROLARIO. *El sistema $\mathfrak{M}_A$ de todos los conjuntos $B \in \mathfrak{M}$, que son subconjuntos de un conjunto fijado $A \in \mathfrak{M}$, constituye un álgebra boreliana.*

Por ejemplo, el sistema de los subconjuntos de cualquier segmento $[a, b]$ medibles según Lebesgue (en el sentido de la medida lebesguiana habitual $\mu^{(1)}$ en la recta) es un álgebra boreliana de conjuntos.

Para concluir señalemos otra propiedad de las medidas de Lebesgue.

DEFINICION 6. Una medida μ se llama *completa* cuando de $\mu(A)=0$ y $A'\subset A$ se desprende $A'\in\mathfrak{S}_\mu$.

Evidentemente, en este caso $\mu(A')=0$. No es difícil demostrar que la prolongación lebesguiana de cualquier medida es completa. Esto se debe a que para $A'\subset A$ y $\mu(A)=0$ es necesariamente $\mu^*(A')=0$ y a que es medible cualquier conjunto C para el cual $\mu^*(C)=0$, ya que $\varnothing\in\mathfrak{R}$ y

$$\mu^*(C\,\triangle\,\varnothing)=\mu^*(C)=0.$$

Indiquemos la relación existente entre el proceso de prolongación de una medida según Lebesgue y el proceso de completación de un espacio métrico. Observemos en este orden que $m'(A\,\triangle\,B)$ puede ser considerado como la distancia entre los elementos A y B del anillo $\mathfrak{R}(\mathfrak{S}_m)$. Entonces, $\mathfrak{R}(\mathfrak{S}_m)$ se convierte en un espacio métrico (no completo, como regla general) y su completación, de acuerdo con el teorema 3 del § 2, se compone precisamente de todos los conjuntos medibles (aunque, sin embargo, los conjuntos A y B no se pueden distinguir desde el punto de vista métrico cuando $\mu(A\,\triangle\,B)=0$).

3°. Prolongación de una medida según Jordan. Al estudiar en el § 2 de este capítulo las medidas que verifican solamente la condición de aditividad, hemos demostrado que cada una de estas medidas m puede ser extendida del semianillo $\mathfrak{S}_m$ al anillo minimal $\mathfrak{R}(\mathfrak{S}_m)$ generado por este semianillo. No obstante, existe también la posibilidad de extender la medida a un anillo más amplio que $\mathfrak{R}(\mathfrak{S}_m)$. La construcción correspondiente se llama *prolongación de una medida según Jordan* [1]. La idea de esta construcción, empleada en varios casos particulares ya por los matemáticos de la Grecia antigua, consiste en aproximar el conjunto «a medir» A por conjuntos A' y A'' de medida prescrita, por dentro y por fuera, esto es, de manera que

$$A'\subset A\subset A''.$$

Sea m una medida definida en un anillo $\mathfrak{R}(\mathfrak{S}_m)$.

DEFINICION 7. Diremos que un conjunto A *es medible según Jordan* cuando para cualquier $\varepsilon>0$ existen en el anillo $\mathfrak{R}$ conjuntos A' y A'' que satisfacen las condiciones

$$A'\subset A\subset A'',\quad m(A''\setminus A')<\varepsilon.$$

Es válida la siguiente proposición.

TEOREMA 9. *El sistema $\mathfrak{R}^*$ de los conjuntos medibles según Jordan es un anillo.*

[1] Camille Jordan, matemático francés (1838—1922).

Sea $\mathfrak{R}$ un sistema de conjuntos A para los cuales existe un conjunto $B \supset A$ de $\mathfrak{R}$. Para cualquier A de $\mathfrak{R}$ tomemos por definición

$$\bar{\mu}(A) = \inf_{B \supset A} m(B),$$
$$\underline{\mu}(A) = \sup_{B \subset A} m(B).$$

Las funciones $\bar{\mu}(A)$ y $\underline{\mu}(A)$ se llaman medida «exterior» e «interior», respectivamente, del conjunto A.

Es evidente que siempre

$$\underline{\mu}(A) \leqslant \bar{\mu}(A).$$

TEOREMA 10. *El anillo $\mathfrak{R}^*$ coincide con el sistema de aquellos conjuntos $A \in \mathfrak{R}$ para los cuales $\underline{\mu}(A) = \bar{\mu}(A)$.*

Para los conjuntos de $\mathfrak{R}$ tienen lugar los siguientes teoremas:

TEOREMA 11. *Si $A \subset \bigcup A_k$, se tiene $\bar{\mu}(A) \leqslant \sum_{k=1}^{n} \bar{\mu}(A_k)$.*

TEOREMA 12. *Si $A_k \subset A$ $(k = 1, 2, \ldots, n)$ y $A_i \cap A_j = \varnothing$, se tiene*

$$\underline{\mu}(A) \geqslant \sum_{k=1}^{n} \underline{\mu}(A_k).$$

Definamos ahora la función μ con campo de definición

$$\mathfrak{S}_\mu = \mathfrak{R}^*,$$

como el valor común de las medidas exterior e interior:

$$\mu(A) = \underline{\mu}(A) = \bar{\mu}(A).$$

De los teoremas 11 y 12 y del hecho evidente de que para $A \in \mathfrak{R}$ se tiene

$$\bar{\mu}(A) = \underline{\mu}(A) = m(A),$$

se desprende la siguiente afirmación.

TEOREMA 13. *La función $\mu(A)$ es una medida y es una prolongación de la medida m.*

La construcción expuesta es aplicable a cualquier medida m definida en un anillo. En particular, se puede aplicarla a los conjuntos del plano. En este caso, se toma como anillo inicial el sistema de conjuntos elementales (es decir, las uniones finitas de rectángulos). El anillo de conjuntos elementales depende, obviamente, de la selección del sistema de coordenadas en el plano (se toman rectángulos de lados paralelos a los ejes de coordenadas). Al pasar a la medida plana de Jordan

$$J^{(2)} = j(m_2),$$

esta dependencia de la selección del sistema de coordenadas desaparece: partiendo de cualquier sistema de coordenadas $\{\bar{x}_1, \bar{x}_2\}$, relacionado con el sistema inicial $\{x_1, x_2\}$ mediante una transformación ortogonal

$$\overline{x}_1 = \cos\alpha\cdot x_1 + \operatorname{sen}\alpha\cdot x_2 + a_1,$$
$$\overline{x}_2 = -\operatorname{sen}\alpha\cdot x_1 + \cos\alpha\cdot x_2 + a_2,$$

obtendremos una misma medida de Jordan

$$J^{(2)} = j(m_2) = j(\overline{m}_2)$$

(aquí $\overline{m}_2$ es la medida construida a partir de los rectángulos de lados paralelos a los ejes $\overline{x}_1$, $\overline{x}_2$). Este resultado se desprende del siguiente teorema general.

TEOREMA 14. *Para que las prolongaciones de Jordan* $\mu_1 = j(m_1)$ y $\mu_2 = j(m_2)$ *de las medidas* m_1 *y* m_2, *definidas en los anillos* $\mathfrak{R}_1$ *y* $\mathfrak{R}_2$, *coincidan es necesario y suficiente que se cumplan las condiciones:*

$$\mathfrak{R}_1 \subset \mathfrak{S}_{\mu_1} \quad m_1(A) = \mu_2(A) \text{ en } \mathfrak{R}_1,$$
$$\mathfrak{R}_2 \subset \mathfrak{S}_{\mu_2} \quad m_2(A) = \mu_1(A) \text{ en } \mathfrak{R}_2.$$

Si la medida inicial m es definida en un semianillo en lugar de un anillo, es natural llamar prolongación de Jordan a la medida

$$j(m) = j(r(m))$$

que se obtiene mediante la extensión de m al anillo $\mathfrak{R}(\mathfrak{S}_m)$ y la prolongación ulterior según Jordan.

4°. Unicidad de prolongación de una medida. Si el conjunto A es medible según Jordan respecto a la medida μ, esto es, pertenece a $\mathfrak{R}^*(\mathfrak{S}_m)$, entonces, para cualquier medida $\tilde{\mu}$, que es prolongación de m y que está definida en $\mathfrak{R}^*(\mathfrak{S}_m)$, el valor $\tilde{\mu}(A)$ coincide con el valor $J(A)$ de la prolongación de Jordan $J = j(m)$. Se puede demostrar que la prolongación de la medida m a un sistema más amplio que $\mathfrak{S}_{j(m)}$ no será única. Con más precisión esto significa lo siguiente. Un conjunto A se llamará conjunto de unicidad de una medida m cuando

1) existe una medida que es prolongación de la medida m y que está definida en el conjunto A;

2) para cualesquiera dos medidas de este género μ_1 y μ_2

$$\mu_1(A) = \mu_2(A).$$

Tiene lugar el teorema: *el sistema de conjuntos de unicidad de una medida* m *coincide con el sistema de los conjuntos medibles según Jordan respecto a la medida* m, *es decir, coincide con el sistema* $\mathfrak{S}_{j(m)}$.

Sin embargo, si son consideradas solamente medidas σ-aditivas y sus prolongaciones (σ-aditivas), el sistema de los conjuntos de unicidad será, en general, más amplio.

Como el caso más importante es precisamente el de las medidas σ-aditivas, tomaremos la siguiente definición.

DEFINICION 8. Un conjunto A se llama *conjunto de σ-unicidad* de una medida σ-aditiva m cuando

1) existe una prolongación σ-aditiva λ de la medida m definida en A (es decir, tal que $A \in \mathfrak{S}_\lambda$);

2) para cualesquiera dos extensiones σ-aditivas λ_1 y λ_2 de este género es válida la igualdad

$$\lambda_1(A) = \lambda_2(A).$$

Si A es un conjunto de σ-unicidad de una medida σ-aditiva μ, existe de acuerdo con nuestra definición, el único valor posible $\lambda(A)$ para la prolongación σ-aditiva de la medida μ definida en A.

Es fácil ver que cada conjunto A medible según Jordan es medible también según Lebesgue (¡pero no viceversa!; dése un ejemplo) y que sus medidas de Jordan y de Lebesgue coinciden. De aquí se desprende inmediatamente que la prolongación de Jordan de una medida σ-aditiva es σ-aditiva.

Cada conjunto A medible según Lebesgue es un conjunto de σ-unicidad para la medida inicial m. En efecto, cualquiera que sea $\varepsilon > 0$ existe para A un $B \in \mathfrak{R}$ tal que $\mu^*(A \triangle B) < \varepsilon$. Cualquiera que sea la prolongación λ, definida en A, de la medida m, tenemos

$$\lambda(B) = m'(B),$$

ya que la prolongación de la medida m a $\mathfrak{R} = \mathfrak{R}(\mathfrak{S}_m)$ es única. Además,

$$\lambda(A \triangle B) \leqslant \mu^*(A \triangle B) < \varepsilon$$

y, consecuentemente,

$$|\lambda(A) - m'(B)| < \varepsilon.$$

Luego, para dos cualesquiera prolongaciones σ-aditivas $\lambda_1(A)$ y $\lambda_2(A)$ de la medida m, tenemos

$$|\lambda_1(A) - \lambda_2(A)| < 2\varepsilon,$$

de donde, debido a la arbitrariedad de $\varepsilon > 0$,

$$\lambda_1(A) = \lambda_2(A).$$

Se puede probar que los conjuntos medibles según Lebesgue agotan todo el sistema de los conjuntos de σ-unicidad de la medida inicial m.

Sea m una medida σ-aditiva con campo de definición $\mathfrak{S}$ y sea $\mathfrak{M} = L(\mathfrak{S})$ el campo de definición de su prolongación de Lebesgue. Del teorema 3 de este parágrafo se desprende fácilmente que cualquiera que sea el semianillo $\mathfrak{S}_1$ tal que

$$\mathfrak{S} \subset \mathfrak{S}_1 \subset \mathfrak{M},$$

siempre

$$L(\mathfrak{S}_1) = L(\mathfrak{S}).$$

§ 4. FUNCIONES MEDIBLES

1°. Definición y propiedades principales de funciones medibles. Sean X e Y dos conjuntos arbitrarios en los que se han escogido dos sistemas de subconjuntos $\mathfrak{S}$ y $\mathfrak{S}'$, respectivamente. Una función abstracta $y = f(x)$ con campo de definición X y con valores en Y se llama $(\mathfrak{S}, \mathfrak{S}')$-*medible* cuando de $A \in \mathfrak{S}'$ se deduce que $f^{-1}(A) \in \mathfrak{S}$.

Por ejemplo, si X e Y son la recta numérica R^1 (es decir, si se consideran funciones reales de variable real) y si $\mathfrak{S}$ y $\mathfrak{S}'$ son el sistema de todos los subconjuntos abiertos (o todos los subconjuntos cerrados) de R^1, la definición dada de función medible coincide con la de continuidad. Si tomamos para $\mathfrak{S}$ y $\mathfrak{S}'$ el sistema de todos los conjuntos borelianos, obtendremos las así llamadas funciones B-medibles (o medibles según Borel).

En lo sucesivo, el concepto de función medible nos interesará fundamentalmente desde el punto de vista de la teoría de integración. En este plano, el papel principal corresponde al concepto de μ-medibilidad de las funciones reales definidas en un conjunto X, coincidiendo $\mathfrak{S}$ con el sistema de todos los subconjuntos μ-medibles del conjunto X y $\mathfrak{S}'$ con la colección de todos los B-conjuntos de la recta. Para simplificar, aceptaremos que X es la unidad del campo de definición $\mathfrak{S}_\mu$ de la medida μ. Como toda medida σ-aditiva puede ser prolongada, de acuerdo con los resultados del § 3, a un álgebra boreliano, es natural admitir desde el principio que $\mathfrak{S}_\mu$ es una B-álgebra. Por lo tanto, para las funciones reales daremos la siguiente definición de medibilidad:

DEFINICION 1. Una función real $f(x)$ definida en un conjunto X se llama *μ-medible* cuando

$$f^{-1}(A) \in \mathfrak{S}_\mu$$

cualquiera que sea el conjunto boreliano A de la recta numérica.

TEOREMA 1. *Para que una función $f(x)$ sea μ-medible es necesario y suficiente que para cualquier c real el conjunto $\{x: f(x) < c\}$ sea μ-medible (es decir, pertenezca a $\mathfrak{S}_\mu$).*

DEMOSTRACION. La necesidad de la condición es obvia ya que la semirrecta $(-\infty, c)$ es un conjunto boreliano. Para demostrar la suficiencia, observemos, ante todo, que la adherencia boreliana $B(\Sigma)$ del sistema Σ de todas las semirrectas $(-\infty, c)$ coincide con el sistema B^1 de todos los conjuntos borelianos de la recta numérica. Por hipótesis, $f^{-1}(\Sigma) \subset \mathfrak{S}_\mu$. Luego,

$$f^{-1}(B(\Sigma)) = B(f^{-1}(\Sigma)) \subset B(\mathfrak{S}_\mu).$$

Pero, $B(\mathfrak{S}_\mu) = \mathfrak{S}_\mu$, ya que $\mathfrak{S}_\mu$ es una B-álgebra, por hipótesis. El teorema queda demostrado.

TEOREMA 2. *El límite de una sucesión convergente para cada $x \in X$ de funciones μ-medibles es μ-medible.*

DEMOSTRACION. Sea $f_n(x) \to f(x)$; entonces,

$$\{x: f(x) < c\} = \bigcup_k \bigcup_n \bigcap_{m>n} \left\{x: f_m(x) < c - \frac{1}{k}\right\}. \qquad (1)$$

En efecto, si $f(x) < c$, existe un k tal que $f(x) < c - \frac{2}{k}$; además, para este k se puede escoger n tan grande que para $m \geqslant n$ se cumpla la desigualdad

$$f_m(x) < c - \frac{1}{k}$$

y esto significará precisamente que x figura en el miembro derecho de (1).

Viceversa, si x pertenece al miembro derecho de la igualdad (1), existe un k tal que para todos los m suficientemente grandes

$$f_m(x) < c - \frac{1}{k};$$

luego, $f(x) < c$, es decir, x figura en el miembro izquierdo de la igualdad (1).

Si las funciones $f_n(x)$ son medibles, los conjuntos

$$\left\{x : f_m(x) < c - \frac{1}{k}\right\}$$

pertenecen a $\mathfrak{S}_\mu$. Como $\mathfrak{S}_\mu$ es un álgebra boreliana, los conjuntos

$$\{x : f(x) < c\}$$

pertenecen también, en virtud de (1), a $\mathfrak{S}_\mu$ y esto demuestra que $f(x)$ es medible.

TEOREMA 3. *Una función B-medible de una función μ-medible es μ-medible.*

DEMOSTRACION 2. Sea $f(x) = \varphi[\psi(x)]$, donde φ es medible según Borel y ψ es μ-medible. Si $A \subset D'$ es un conjunto μ-medible arbitrario, su imagen recíproca $A' = \varphi^{-1}(A)$ es B-medible y la imagen recíproca $A'' = \psi^{-1}(A')$ del conjunto A' es B-medible. Como $f^{-1}(A) = A''$, de aquí sigue la medibilidad de la función f.

El teorema demostrado es aplicable, en particular, al caso de funciones continuas φ (que son siempre B-medibles).

2°. Funciones simples. En vista del estudio ulterior de funciones medibles conviene representar cada una de ellas como límite de una sucesión de así llamadas funciones simples.

DEFINICION 2. Una función $f(x)$ se llama *simple* cuando es μ-medible y toma a lo sumo un número numerable de valores.

Está claro que el concepto de función simple depende de la selección de la medida μ.

La estructura de las funciones simples es caracterizada por el siguiente teorema.

TEOREMA 4. *Una función $f(x)$ que toma a lo sumo un número numerable de valores distintos*

$$y_1, y_2, \ldots, y_n, \ldots$$

es μ-medible cuando, y sólo cuando, todos los conjuntos

$$A_n = \{x : f(x) = y_n\}$$

son μ-medibles.

DEMOSTRACION. La necesidad de la condición está clara ya que cada A_n es la imagen recíproca del conjunto compuesto por un solo punto $\{y_n\}$ y cualquier conjunto compuesto de un solo punto es boreliano. La suficiencia se deduce de que la imagen recíproca $f^{-1}(B)$ de cualquier conjunto $B \subset D^1$ es, por hipótesis, la unión $\bigcup_{y_n \in B} A_n$ de una cantidad a lo sumo numerable de conjuntos medibles A_n, es decir, es medible.

El empleo ulterior de las funciones simples se basa en el siguiente teorema.

TEOREMA 5. *Para que la función $f(x)$ sea μ-medible es necesario y suficiente que pueda ser representada como límite de una sucesión uniformemente convergente de funciones medibles simples.*

DEMOSTRACION. La suficiencia se desprende del teorema 2. Para demostrar la necesidad, consideremos una función medible arbitraria $f(x)$ y tomemos $f_n(x) = \frac{m}{n}$ cuando $\frac{m}{n} \leqslant f(x) < \frac{m+1}{n}$ (aquí m son enteros y n son enteros positivos). Está claro que $f_n(x)$ son funciones simples; para $n \to \infty$ ellas convergen uniformemente hacia $f(x)$ ya que $|f(x) - f_n(x)| \leqslant \frac{1}{n}$.

3°. Operaciones aritméticas con funciones medibles.

TEOREMA 6. *La suma, la diferencia y el producto de dos funciones medibles son funciones medibles. El cociente de dos funciones medibles es también medible si el denominador no se anula.*

Realicemos la demostración de este teorema en varios pasos.

a) La suma de dos funciones medibles es medible.

Sean primero $f(x)$ y $g(x)$ dos funciones μ-medibles simples que toman los valores

$$f_1, f_2, \ldots, f_n, \ldots$$

y

$$g_1, g_2, \ldots, g_n, \ldots$$

respectivamente. La suma $h(x) = f(x) + g(x)$ puede tomar solamente los valores $h=f_i+g_j$, con la particularidad de que estos valores se toman en los conjuntos

$$\{x:h(x)=h\}=\bigcup_{f_i+g_j=h}(\{x:f(x)=f_i\}\bigcap\{x:g(x)=g_j\}). \qquad (2)$$

El número de los valores posibles h es finito o numerable y los conjuntos $\{x:h(x)=h\}$ correspondientes a estos valores son medibles, ya que el miembro derecho de la igualdad (2) es, obviamente, un conjunto medible. Luego, $h(x)=f(x)+g(x)$ es una función medible simple.

Sean ahora $f(x)$ y $g(x)$ dos funciones medibles arbitrarias; consideremos las sucesiones $\{f_n(x)\}$ y $\{g_n(x)\}$ de funciones simples convergentes hacia $f(x)$ y $g(x)$ respectivamente. Entonces, las funciones simples $f_n(x)+g_n(x)$ convergen uniformemente hacia la función $f(x)+g(x)$ que, en virtud del teorema 5, es medible.

b) El producto de una función μ-medible por un número constante es μ-medible. Esta afirmación es obvia.

De a) y b) resulta:

c) La diferencia de dos funciones μ-medibles es μ-medible.

d) El producto de funciones μ-medibles es μ-medible. En efecto, consideremos la identidad $fg=\frac{1}{4}[(f+g)^2-(f-g)^2]$. La expresión del miembro derecho es una función μ-medible. Esto se desprende de a), b) y c) y de que el cuadrado de una función medible es, en virtud del teorema 3, una función medible.

e) Si $f(x)$ es medible y $f(x)\neq 0$, también $\frac{1}{f(x)}$ es medible. En efecto, tenemos

$$\left\{x:\frac{1}{f(x)}<c\right\}=\left\{x:f(x)>\frac{1}{c}\right\}\bigcup\{x:f(x)<0\}$$

para $c>0$,

$$\left\{x:\frac{1}{f(x)}<c\right\}=\left\{x:0>f(x)>\frac{1}{c}\right\}$$

para $c<0$ y

$$\left\{x:\frac{1}{f(x)}<c\right\}=\{x:f(x)<c\}$$

para $c=0$.

En el miembro derecho figura cada una de las veces un conjunto medible. De d) y e) se deduce la medibilidad del cociente $\frac{f(x)}{g(x)}$ (con la condición de que $g(x)\neq 0$).

En resumen, hemos probado que las operaciones aritméticas con funciones medibles llevan de nuevo a funciones medibles.

4°. **Equivalencia.** En el estudio de funciones medibles pueden ser despreciados frecuentemente los valores de una función en un conjunto de medida nula. En este orden introduciremos la siguiente definición.

DEFINICION 3. Dos funciones f y g definidas en un mismo conjunto medible E se llaman *equivalentes* (en símbolo $f \sim g$) cuando

$$\mu\{x: f(x) \neq g(x)\} = 0.$$

Diremos que una propiedad se verifica en *casi todo* el E cuando se verifica en todos los puntos de E con la excepción de puntos que forman un conjunto de medida nula. De esta forma, se puede decir que dos funciones se llaman equivalentes cuando coinciden en casi todos los puntos.

TEOREMA 7. *Si dos funciones f y g continuas en un segmento E son equivalentes (respecto a la medida de Lebesgue), ellas coinciden.*

DEMOSTRACION. Supongamos que en un punto x_0 se tiene $f(x_0) \neq g(x_0)$, es decir, $f(x_0) - g(x_0) \neq 0$. Como $f-g$ es una función continua, existirá una vecindad del punto x_0 tal que en todos sus puntos la función $f-g$ es diferente de cero. Esta vecindad tiene medida positiva; de manera que

$$\mu\{x: f(x) \neq g(x)\} > 0,$$

es decir, las funciones continuas f y g no pueden ser equivalentes, si toman diferentes valores aunque sea en un punto.

Evidentemente, para funciones medibles arbitrarias (esto es, discontinuas, en general) la equivalencia de dos funciones de ninguna manera implica su coincidencia; por ejemplo, la función igual a la unidad en los puntos racionales y al cero en los puntos irracionales es equivalente a la función igual idénticamente a cero.

TEOREMA 8. *Una función $f(x)$, definida en un conjunto medible E y equivalente en este conjunto a una función medible $g(x)$, es también medible.*

En efecto, de la definición de equivalencia se desprende que los conjuntos

$$\{x: f(x) < a\} \text{ y } \{x: g(x) < a\}$$

pueden diferir uno de otro solamente en un conjunto de medida nula; por consiguiente, si es medible el segundo de ellos, es medible también el primero.

5°. Convergencia en casi todos los puntos. Puesto que en muchos casos el comportamiento de las funciones medibles en uno u otro conjunto de medida nula no tendrá importancia para nosotros, resulta natural generalizar del siguiente modo el concepto habitual de convergencia de una sucesión de funciones.

DEFINICION 4. Una sucesión de funciones $f_n(x)$ definidas en un espacio X se llama *convergente en casi todos los puntos* hacia $F(x)$ cuando

$$\lim_{n\to\infty} f_n(x) = F(x) \tag{3}$$

para casi todo $x \in X$ (es decir, el conjunto de aquellos puntos x en los que no se verifica (3) es de medida nula).

Ejemplo. La sucesión de funciones $f_n(x) = (-x)^n$, definidas en el segmento $[0, 1]$, converge para $n \to \infty$ hacia la función $F(x) \equiv 0$ en casi todos los puntos (a saber, en todos puntos a excepción del punto $x = 1$).

El teorema 2 admite esta generalización.

TEOREMA 2'. *Si una sucesión de funciones μ-medibles $f_n(x)$ converge hacia una función $F(x)$ en casi todo el espacio X, $F(x)$ es también medible.*

DEMOSTRACION. Sea A el conjunto, donde

$$\lim_{n\to\infty} f_n(x) = F(x).$$

Por hipótesis, $\mu(X \setminus A) = 0$. La función $F(x)$ es medible en A y como cualquier función es medible, evidentemente, en un conjunto de medida nula, $F(x)$ es medible en $X \setminus A$; luego, es medible también en el conjunto X.

EJERCICIO. Supongamos que una sucesión de funciones medibles $f_n(x)$ converge en casi todos los puntos hacia una función límite $f(x)$. Demuéstrese que la sucesión $\{f_n(x)\}$ converge en casi todo punto hacia $g(x)$ cuando, y sólo cuando, $g(x)$ es equivalente a $f(x)$.

6°. Teorema de Egórov. D. F. Egórov demostró en 1913 el siguiente teorema importante que establece la relación entre la convergencia en casi todos los puntos y la convergencia uniforme.

TEOREMA 9. *Supongamos que una sucesión de funciones medibles $\{f_n(x)\}$ converge hacia $f(x)$ en casi todo el conjunto E. Entonces, para cualquier $\delta > 0$ existe un conjunto medible $E_\delta \subset E$ tal que*
1) $\mu(E_\delta) > \mu(E) - \delta$;
2) la sucesión $\{f_n(x)\}$ converge hacia $f(x)$ uniformemente en el conjunto E_δ.

DEMOSTRACION. De acuerdo con el teorema 2′ la función $f(x)$ es medible. Pongamos

$$E_n^m = \bigcap_{i \geqslant n} \left\{ x : |f_i(x) - f(x)| < \frac{1}{m} \right\}.$$

De esta forma, E_n^m representa, para m y n fijos, el conjunto de todos los puntos x para los cuales

$$\left| f_i(x) - f(x) < \frac{1}{m} \right|$$

cualquiera que sea $i \geqslant n$. Sea

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

Está claro, de la definición de los conjuntos E_n^m, que para un m fijo

$$E_1^m \subset E_2^m \subset \ldots \subset E_n^m \subset \ldots$$

Debido a que una medida σ-aditiva es continua, para cualquier m y cualquier $\delta > 0$ existirá un $n_0(m)$ tal que

$$\mu(E^m \setminus E^m_{n_0(m)}) < \frac{\delta}{2^m}.$$

Tomemos

$$E_\delta = \bigcap_{m=1}^{\infty} E^m_{n_0(m)}$$

y probemos que el conjunto E_δ construido de esta forma verifica las condiciones del teorema.

Demostremos primero que la sucesión $\{f_i(x)\}$ converge uniformemente en E_δ hacia la función $f(x)$. Esto se desprende inmediatamente de que para $x \in E_\delta$ y cualquier m

$$|f_i(x) - f(x)| < \frac{1}{m} \text{ cuando } i > n_0(m).$$

Estimemos ahora la medida del conjunto $E \setminus E_\delta$. Observemos para ello que $\mu(E \setminus E^m) = 0$ cualquiera que sea m. En efecto, si $x_0 \in E \setminus E^m$, existen valores tan grandes como se quiera de i para los cuales

$$|f_i(x_0) - f(x_0)| \geqslant \frac{1}{m},$$

es decir, la sucesión $\{f_n(x)\}$ no converge hacia $f(x)$ en el punto x_0. Como, por hipótesis, $\{f_n(x)\}$ converge hacia $f(x)$ en casi todos los puntos, tenemos

$$\mu(E\setminus E^m)=0.$$

De aquí se sigue que

$$\mu(E\setminus E^m_{n_0(m)})=\mu(E^m\setminus E^m_{n_0(m)})<\frac{\delta}{2^m}.$$

Luego,

$$\mu(E\setminus E_\delta)=\mu\left(E\setminus\bigcap_{m=1}^{\infty}E^m_{n_0(m)}\right)=$$

$$=\mu\left(\bigcup_{m=1}^{\infty}(E\setminus E^m_{n_0(m)})\right)\leqslant\sum_{m=1}^{\infty}\mu(E\setminus E^m_{n_0(m)})<\sum_{m=1}^{\infty}\frac{\delta}{2^m}=\delta.$$

Hemos demostrado el teorema.

7°. Convergencia en medida.

DEFINICION 5. Se dice que una sucesión de funciones medibles $f_n(x)$ *converge en medida* hacia una función $F(x)$ cuando para cualquier $\sigma>0$

$$\lim_{n\to\infty}\mu\{x:|f_n(x)-F(x)|\geqslant\sigma\}=0.$$

Los teoremas 10 y 11 que siguen establecen la relación entre los conceptos de convergencia en casi todos los puntos y convergencia en medida.

TEOREMA 10. *Si una sucesión de funciones medibles $f_n(x)$ converge en casi todos los puntos hacia una función $F(x)$, converge en medida hacia la misma función límite $F(x)$.*

DEMOSTRACION. Del teorema 2' se desprende que la función límite $F(x)$ es medible. Sea A el conjunto (de medida nula) en el que $f_n(x)$ no tiende a $F(x)$. Sean, además,

$$E_k(\sigma)=\{x:|f_k(x)-F(x)|\geqslant\sigma\},$$

$$R_n(\sigma)=\bigcup_{k=n}^{\infty}E_k(\sigma),$$

$$M=\bigcap_{n=1}^{\infty}R_n(\sigma).$$

Está claro que todos estos conjuntos son medibles. Como

$$R_1(\sigma)\supset R_2(\sigma)\supset\ldots.$$

tenemos, debido a la continuidad de la medida,

$$\mu(R_n(\sigma))\longrightarrow\mu(M)\text{ para }n\longrightarrow\infty.$$

Comprobemos ahora que

$$M\subset A. \qquad (4)$$

En efecto, si $x_0 \overline{\in} A$, es decir, si

$$\lim_{n \to \infty} f_n(x_0) = F(x_0),$$

para un $\sigma > 0$ dado existe un n, tal que

$$|f_n(x_0) - F(x_0)| < \sigma,$$

es decir, $x_0 \overline{\in} E_n(\sigma)$ y, con más razón, $x_0 \overline{\in} M$.

Pero, $\mu(A) = 0$; de manera que (4) implica $\mu(M) = 0$ y, por consiguiente,

$$\mu(R_n(\sigma)) \longrightarrow 0 \text{ para } n \longrightarrow \infty;$$

como $E_n(\sigma) \subset R_n(\sigma)$, esto demuestra el teorema.

Es fácil ver en un ejemplo que la convergencia en medida de una sucesión de funciones no implica, en general, su convergencia en casi todos los puntos. Efectivamente, definamos para cada k natural en el semisegmento $(0, 1]$ k funciones

$$f_1^{(k)}, f_2^{(k)}, \ldots, f_k^{(k)}$$

del siguiente modo

$$f_i^{(k)}(x) = \begin{cases} 1 \text{ para } \dfrac{i-1}{k} < x \leqslant \dfrac{i}{k}, \\ 0 \text{ para los demás valores de } x. \end{cases}$$

Enumerando una tras otra todas estas funciones obtendremos una sucesión que, como es fácil de comprobar, converge en medida hacia cero y al mismo tiempo no converge en ningún punto (¡demuéstrese esto!).

EJERCICIO. Supongamos que una sucesión de funciones medibles $f_n(x)$ converge en medida hacia una función límite $f(x)$. Demuéstrese que la sucesión $f_n(x)$ converge en medida hacia una función $g(x)$ cuando, y sólo cuando, $g(x)$ es equivalente a $f(x)$.

Aunque el ejemplo dado demuestra que el teorema 10 no puede invertirse completamente, tiene lugar el siguiente resultado.

TEOREMA 11. *Supongamos que una sucesión de funciones medibles $f_n(x)$ converge en medida hacia $f(x)$. Entonces, de esta sucesión se puede extraer una sucesión parcial $\{f_{n_k}(x)\}$ que converge en casi todos los puntos hacia $f(x)$.*

DEMOSTRACION. Sea $\varepsilon_1, \varepsilon_2, \ldots$ una sucesión de números positivos tal que

$$\lim_{n \to \infty} \varepsilon_n = 0$$

y sean $\eta_1, \eta_2, \ldots, \eta_n, \ldots$ unos números positivos tales que la serie

$$\eta_1 + \eta_2 + \ldots$$

converge. Construyamos una sucesión de índices

$$n_1 < n_2 < \ldots$$

del modo siguiente: n_1 es número natural tal que

$$\mu\{x: |f_{n_1}(x) - f(x)| \geqslant \varepsilon_1\} < \eta_1$$

(un número n_1 de este tipo existe obligatoriamente); n_2 es un número tal que

$$\mu\{x:|f_{n_2}(x)-f(x)|\geqslant\varepsilon_2\}<\eta_2.$$

En general, n_k es un número tal que

$$\mu\{x:|f_{n_k}(x)-f(x)|\geqslant\varepsilon_k\}<\eta_k$$
$$(n_k>n_{k-1}).$$

Mostremos que la sucesión construida converge hacia $f(x)$ en casi todos los puntos. Efectivamente, sean

$$R_i=\bigcup_{k=i}^{\infty}\{x:|f_{n_k}(x)-f(x)|\geqslant\varepsilon_k,$$

$$Q=\bigcap_{i=1}^{\infty}R_i.$$

Como

$$R_1\supset R_2\supset R_3\supset\ldots\supset R_n\supset\ldots$$

tendremos, debido a la continuidad de la medida, $\mu(R_i)\longrightarrow\mu(Q)$.

Por otro lado, está claro que $\mu(R_i)<\sum\limits_{k=i}^{\infty}\eta_k$, de donde se desprende que $\mu(R_i)\longrightarrow 0$ para $i\longrightarrow\infty$, es decir, $\mu(Q)=0$. Resta probar que en todos los puntos del conjunto $E\diagdown Q$ tiene lugar la relación

$$f_{n_k}(x)\longrightarrow f(x).$$

Sea $x_0\in E\diagdown Q$. Entonces, existirá un i_0 tal que $x_0\,\overline{\in}\,R_{i_0}$. Ello significa que para todos los $k\geqslant i_0$

$$x_0\,\overline{\in}\,\{x:|f_{n_k}(x)-f(x)|\geqslant\varepsilon_k\},$$

es decir,

$$|f_{n_k}(x_0)-f(x_0)|<\varepsilon_k.$$

Como, por hipótesis, $\varepsilon_k\longrightarrow 0$, se tiene

$$\lim_{k\to\infty}f_{n_k}(x_0)=f(x_0).$$

El teorema queda demostrado.

8°. Teorema de Luzin. *C*-propiedad. La definición de función medible, dada al principio de este parágrafo, se refiere a funciones sobre conjuntos arbitrarios y no está relacionada, en el caso general, de manera alguna con el concepto de continuidad de una función. Sin embargo, si se trata de funciones definidas en un segmento, tiene lugar el siguiente teorema importante demostrado por N. N. Luzin en 1913:

TEOREMA 12. *Para que una función $f(x)$ definida en un segmento $[a, b]$ sea medible es necesario y suficiente que para cualquier $\varepsilon>0$ exista una función $\varphi(x)$ continua en $[a, b]$ tal que*

$$\mu\{x:f(x)\neq\varphi(x)\}<\varepsilon.$$

En otras palabras, una función medible se puede convertir en una continua en $[a, b]$ variando sus valores en un conjunto de medida tan pequeña como se quiera. De una función en un segmento, que mediante una «deformación pequeña» de este tipo puede ser hecha continua, se dice que verifica la C-propiedad (término de N. N. Luzin). El teorema de Luzin muestra que para funciones de argumento numérico la C-propiedad puede ser tomada como base de la propia definición de medibilidad. Es fácil obtener la demostración del teorema de Luzin valiéndose del teorema de Egórov (¡realícese esta demostración!).

§ 5. INTEGRAL DE LEBESGUE

El concepto de la integral de Riemann, conocido del curso elemental del Análisis, es aplicable sólo a aquellas funciones que o bien son continuas o bien no tienen «demasiados» puntos de discontinuidad. Para funciones medibles que pueden ser discontinuas en todo punto donde estén definidas (o incluso pueden estar definidas en un conjunto abstracto de manera que el concepto de continuidad carece de sentido para ellas), la construcción de Riemann de la integral no es válida. Al mismo tiempo, para estas funciones existe un concepto perfecto y flexible de la integral introducido por Lebesgue.

La idea principal de la integral de Lebesgue consiste en que, a diferencia de la integral de Riemann, los puntos x se agrupan no de acuerdo a su proximidad en el eje x sino de acuerdo a la proximidad de los valores de la función en estos puntos. Esto ofrece inmediatamente la posibilidad de extender el concepto de integral a una clase muy amplia de funciones.

Además, la integral de Lebesgue se define de un mismo modo para funciones determinadas en cualesquiera espacios provistos de medida, mientras que la integral de Riemann se introduce primero para funciones de una variable y solamente después se extiende, con las modificaciones correspondientes, al caso de varias variables. Para funciones en espacios abstractos provistos de medida la integral de Riemann simplemente no tiene sentido.

En lo que sigue se considerará, siempre que no se diga lo contrario, una medida σ-aditiva $\mu(A)$ definida en un álgebra boreliana de conjuntos con unidad X. Todos los conjuntos considerados $A \subset X$ se supondrán μ-medibles y las funciones $f(x)$ estarán definidas para $x \in X$ y serán μ-medibles.

1°. Integral de Lebesgue para funciones simples. Introduciremos primero el concepto de la integral de Lebesgue para las funciones que hemos llamado con anterioridad simples, es decir, para funciones medibles con un número finito o numerable de valores.

Sea f una función simple que toma los valores

$$y_1, y_2, \ldots, y_n, \ldots; \; y_i \neq y_j \text{ para } i \neq j.$$

Es natural definir la integral de la función f en el conjunto A mediante la igualdad

$$\int_A f(x)\,d\mu=\sum_n y_n\,\mu(A_n), \text{ donde } A_n=\{x: x\in A,\quad f(x)=y_n\}, \qquad (1)$$

si la serie que figura en el miembro derecho converge. De esta forma llegamos a la siguiente definición (en la que, por razones obvias, se postula de antemano la convergencia *absoluta* de esta serie).

DEFINICION 1. Una función simple f se llama *integrable* o *sumable* (respecto a la medida μ) en un conjunto A cuando la serie (1) converge absolutamente. Si f es integrable, la suma de la serie (1) se llama integral de f en el conjunto A.

En esta definición se supone que todos los y_n son diferentes. Sin embargo, el valor de la integral de una función simple se puede representar como la suma de productos de tipo $c_k\mu(B_k)$ sin suponer que todos los c_k son distintos. Esto es posible hacerlo gracias al siguiente lema.

LEMA. *Supongamos que* $A=\bigcup\limits_k B_k$, $B_i\cap B_j=\varnothing$ *para* $i=j$, *y que la función* f *toma en cada conjunto* B_k *un valor único* c_k; *entonces,*

$$\int_A f(x)\,d\mu=\sum_k c_k\mu(B_k) \qquad (2)$$

y la función f *es integrable en* A *cuando, y sólo cuando, la serie* (2) *converge absolutamente.*

DEMOSTRACION. Es fácil ver que cada conjunto

$$A_n=\{x: x\in A,\quad f(x)=y_n\}$$

es la unión de aquellos B_k para los cuales $c_k=y_n$. Luego,

$$\sum_n y_n\,\mu(A_n)=\sum_n y_n\sum_{c_k=y_n}\mu(B_k)=\sum_k c_k\mu(B_k).$$

Como la medida es no negativa, tenemos

$$\sum_n |y_n|\,\mu(A_n)=\sum_n |y_n|\sum_{c_k=y_n}\mu(B_k)=\sum_k |c_k|\,\mu(B_k),$$

es decir, las series $\sum\limits_n y_n\mu(A_n)$ y $\sum\limits_k c_k\mu(B_k)$ son ambas absolutamente convergentes o ambas divergentes. El lema queda demostrado.

Indiquemos algunas propiedades de la integral de Lebesgue de funciones simples:

A) $$\int_A f(x)\,d\mu + \int_A g(x)\,d\mu = \int_A [f(x)+g(x)]\,d\mu,$$

con la particularidad de que la existencia de las integrales del miembro izquierdo implica la existencia de la integral del miembro derecho.

Supongamos, para demostrar esta propiedad, que f toma los valores f_i en los conjuntos $F_i \subset A$ y g toma los valores g_j en los conjuntos $G_j \subset A$, de manera que

$$J_1 = \int_A f(x)\,d\mu = \sum_i f_i \mu(F_i), \tag{3}$$

$$J_2 = \int_A g(x)\,d\mu = \sum_j g_j \mu(G_j). \tag{4}$$

Entonces, de acuerdo con el lema,

$$J = \int_A [f(x)+g(x)]\,d\mu = \sum_i \sum_j (f_i+g_j)\,\mu(F_i \cap G_j), \tag{5}$$

pero,

$$\mu(F_i) = \sum_j \mu(F_i \cap G_j), \quad \mu(G_j) = \sum_i \mu(F_i \cap G_j),$$

de manera que la convergencia absoluta de las series (3) y (4) implica la convergencia absoluta de la serie (5); además,

$$J = J_1 + J_2.$$

B) Para cualquier constante k

$$k\int_A f(x)\,d\mu = \int_A \{kf(x)\}\,d\mu,$$

donde la existencia de la integral de miembro izquierdo implica la existencia de la integral del miembro derecho. (La comprobación es inmediata).

C) Una función simple f acotada en un conjunto A es integrable en A y, además, si $|f(x)| \leqslant M$ en A, se tiene

$$\left|\int_A f(x)\,d\mu\right| \leqslant M\mu(A).$$

(La comprobación es inmediata).

2°. Integral de Lebesgue en conjuntos de medida finita.

DEFINICION 2. Una función $f(x)$ se llamará *integrable (sumable) en un conjunto* A cuando exista una sucesión de funciones simples f_n integrables en A convergente uniformemente hacia f.

El límite

$$I = \lim_{n \to \infty} \int_A f_n(x)\, d\mu \tag{6}$$

será denotado mediante

$$\int_A f(x)\, d\mu$$

y llamado *integral* de la función f *en el conjunto* A.

Esta definición es correcta si se verifican las siguientes condiciones:

1. El límite (6) existe cualquiera que sea la sucesión uniformemente convergente de funciones simples integrables en A.

2. Este límite no depende, para una función fijada f, de la selección de la sucesión $\{f_n\}$.

3. Para las funciones simples las definiciones de la integrabilidad y de la integral coinciden con las dadas en el punto 1.

Todas estas condiciones quedan, realmente, satisfechas.

Para demostrar la primera, basta observar que, debido a las propiedades A), B) y C) de la integral de funciones simples,

$$\left| \int_A f_n(x)\, d\mu - \int_A f_m(x)\, d\mu \right| \leqslant \mu(A) \sup_{x \in A} |f_n(x) - f_m(x)|. \tag{7}$$

Para demostrar la segunda condición es preciso considerar dos sucesiones $\{f_n\}$ y $\{f_n^*\}$ convergentes hacia f. Si el límite (6) tomara valores distintos para cada una de estas dos sucesiones, no existiría el límite (6) para la sucesión obtenida como unión de estas dos, lo que estaría en contradicción con la primera condición. Finalmente, para probar la validez de la tercera condición es suficiente considerar la sucesión $f_n = f$.

Establezcamos las propiedades fundamentales de la integral de Lebesgue. Una consecuencia directa de la definición es que

I. $$\int_A 1 \cdot d\mu = \mu(A). \tag{8}$$

II. Para cualquier constante k

$$\int_A \{kf(x)\}\, d\mu = k \int_A f(x)\, d\mu, \tag{9}$$

donde la existencia de la integral del miembro izquierdo implica la existencia de la integral del miembro derecho.

Esta propiedad se deduce, mediante el paso al límite, de la propiedad B) para las integrales de funciones simples.

III. $$\int_A f(x)\,d\mu + \int_A g(x)\,d\mu = \int_A \{f(x)+g(x)\}\,d\mu, \tag{10}$$

donde la existencia de las integrales del miembro izquierdo implica la existencia de la integral del miembro derecho.

La demostración se obtiene, mediante el paso al límite, de la propiedad A) de la integral para funciones simples.

IV. Una función f acotada en un conjunto A es integrable en A.

La demostración se obtiene, mediante el paso al límite, de la propiedad B) de la integral de funciones simples.

V. Si $f(x) \geqslant 0$, se tiene

$$\int_A f(x)\,d\mu \geqslant 0 \tag{11}$$

(suponiendo que la integral existe).

Para las funciones simples esta propiedad se deduce directamente de la definición y en el caso general la demostración se basa en la posibilidad de aproximar una función no negativa mediante funciones simples no negativas.

De la última propiedad se desprende inmediatamente que para $f(x) \geqslant g(x)$ se tiene

$$\int_A f(x)\,d\mu \geqslant \int_A g(x)\,d\mu, \tag{12}$$

de manera que siendo $m \leqslant f(x) \leqslant M$ para todos (o casi todos) los $x \in A$, tendremos

$$m\mu(A) \leqslant \int_A f(x)\,d\mu \leqslant M\mu(A). \tag{13}$$

VI. Si $\mu(A)=0$, se tiene $\int_A f(x)\,d\mu = 0$.

Esta afirmación se deduce directamente de la definición de la integral de Lebesgue.

VII. Si una función φ es integrable en A y $|f(x)| \leqslant \varphi(x)$ en casi todo el A, entonces f es integrable en A.

En efecto, si f y φ son funciones simples, omitiendo del conjunto A un conjunto de medida nula, podemos representar el conjunto A' que queda como la unión de una cantidad finita o numerable de conjuntos, en cada uno de los cuales f y φ son constantes

$$f(x) = a_n, \quad \varphi(x) = b_n \text{ y, además, } |a_n| \leqslant b_n.$$

De la integrabilidad de φ se desprende que

$$\sum_n |a_n| \mu(A_n) \leqslant \sum_n b_n \mu(A_n) = \int_{A'} \varphi(x)\, d\mu = \int_A \varphi(x)\, d\mu.$$

De manera que f es también integrable y

$$\left|\int_A f(x)\, d\mu\right| = \left|\int_{A'} f(x)\, d\mu\right| = \left|\sum_n a_n \mu(A_n)\right| \leqslant$$
$$\leqslant \sum_n |a_n| \mu(A_n) = \int_{A'} |f(x)|\, d\mu \leqslant \int_A \varphi(x)\, d\mu.$$

En el caso general esta afirmación se demuestra pasando al límite.

VIII. Las integrales

$$I_1 = \int_A f(x)\, d\mu, \quad I_2 = \int_A |f(x)|\, d\mu \tag{14}$$

o bien existen ambas o bien ambas no existen.

En efecto, la existencia de la integral I_2 implica la existencia de la integral I_1.

Lo recíproco se desprende, en el caso de una función simple, de la definición de la integral y en el caso general se demuestra mediante el paso al límite y valiéndose de la desigualdad

$$||a|-|b|| \leqslant |a-b|.$$

3°. σ-aditividad y continuidad absoluta de la integral de Lebesgue. En el punto anterior hemos enunciado las propiedades de la integral de Lebesgue en un conjunto fijado. Ahora estableceremos algunas propiedades de la integral de Lebesgue considerando la expresión

$$F(A) = \int_A f(x)\, d\mu$$

como función de conjunto, definida en la clase de conjuntos medibles. Probemos, ante todo, la propiedad siguiente.

TEOREMA 1. Si $A = \bigcup\limits_n A_n$; $A_i \cap A_j = \varnothing$ para $i \neq j$, *entonces,*

$$\int_A f(x)\, d\mu = \sum \int_{A_n} f(x)\, d\mu, \tag{15}$$

donde la existencia de la integral del miembro izquierdo implica la existencia de las integrales y la convergencia absoluta de la serie del miembro derecho.

DEMOSTRACION. Probemos primero la afirmación del teorema para el caso de una función simple f que toma los valores

$$y_1, y_2, \ldots, y_n, \ldots$$

Sean

$$B_k = \{x : x \in A, \quad f(x) = y_k\},$$
$$B_{nk} = \{x : x \in A_n, \quad f(x) = y_k\}.$$

Entonces,

$$\int_A f(x)\, d\mu = \sum_k y_k \mu(B_k) = \sum_k y_k \sum_n \mu(B_{nk}) =$$
$$= \sum_n \sum_k y_k \mu(B_{nk}) = \sum_n \int_{A_n} f(x)\, d\mu. \quad (16)$$

Como la serie $\sum y_k \mu(B_k)$ converge absolutamente, suponiendo que f es integrable en A, y como las medidas de todos los conjuntos son no negativas, también convergen absolutamente todas las demás series de la cadena de igualdades (16).

En el caso de una función f arbitraria, se desprende, de su integrabilidad en A, que para cualquier $\varepsilon > 0$ existe una función simple g integrable en A que verifica la condición

$$|f(x) - g(x)| < \varepsilon. \quad (17)$$

Para g tenemos

$$\int_A g(x)\, d\mu = \sum_n \int_{A_n} g(x)\, d\mu, \quad (18)$$

donde g es integrable en cada conjunto A_n y la serie (18) converge absolutamente. De este último hecho y de la estimación (17) se desprende que f es también integrable en cada A_n y que

$$\sum_n \left| \int_{A_n} f(x)\, d\mu - \int_{A_n} g(x)\, d\mu \right| \leqslant \sum_n \varepsilon\mu(A_n) = \varepsilon\mu(A),$$
$$\left| \int_A f(x)\, d\mu - \int_A g(x)\, d\mu \right| \leqslant \varepsilon\mu(A),$$

esto y (18) demuestra la convergencia absoluta de la serie $\sum_n \int_{A_n} f(x)\, d\mu$ y lleva a la estimación

$$\left| \sum_n \int_{A_n} f(x)\, d\mu - \int_A f(x)\, d\mu \right| \leqslant 2\varepsilon\mu(A).$$

Como $\varepsilon > 0$ es arbitrario, obtenemos

$$\sum_n \int_{A_n} f(x)\, d\mu = \int_A f(x)\, d\mu.$$

COROLARIO. *Si f es integrable en A, f es integrable también en cualquier conjunto medible $A' \subset A$.*

Hemos demostrado que la integrabilidad en un conjunto A de una funcion $f(x)$ implica, en el caso en que $A = \bigcup A_n$ y $A_i \cap A_j \neq \varnothing$, que $f(x)$ sea integrable en cada A_n y que la integral en A sea igual a la suma de las integrales en los conjuntos A_n.

Esta afirmación puede ser invertida en el sentido siguiente.

TEOREMA 2. *Si $A = \bigcup\limits_n A_n$, $A_i \cap A_j = \varnothing$ para $i \neq j$ y la serie*

$$\sum_n \int_{A_n} |f(x)|\, d\mu \tag{19}$$

converge, la función f es integrable en A y

$$\int_A f(x)\, d\mu = \sum_n \int_{A_n} f(x)\, d\mu.$$

Lo nuevo aquí, en comparación con el teorema anterior, es la afirmación de que la convergencia de la serie (19) implica la integrabilidad de f en A.

Realicemos la demostración primero para el caso de una función simple f que toma los valores f_i en los conjuntos B_i. Tomando

$$A_{ni} = A_n \cap B_i,$$

tenemos

$$\int_{A_n} |f(x)|\, d\mu = \sum_i |f_i|\, \mu(A_{ni}).$$

La convergencia de la serie (19) implica la convergencia de las series

$$\sum_n \sum_i |f_i|\, \mu(A_{ni}) = \sum_i |f_i|\, \mu(B_i \cap A).$$

La convergencia de la última serie significa la existencia de la integral

$$\int_A f(x)\, d\mu = \sum_i f_i \mu(B_i \cap A).$$

En el caso general, aproximaremos la función f mediante una función simple $\tilde{f}$ de manera que

$$|f(x) - \tilde{f}(x)| < \varepsilon. \tag{20}$$

Entonces,

$$\int_{A_n} |\tilde{f}(x)|\, d\mu \leqslant \int_{A_n} |f(x)|\, d\mu + \varepsilon\mu(A_n)$$

y como la serie

$$\sum_n \mu(A_n) = \mu(A)$$

converge, la convergencia de la serie (19) implica la convergencia de la serie

$$\sum_n \int_{A_n} |\tilde{f}(x)|\, d\mu,$$

es decir, implica, de acuerdo con lo demostrado, la integrabilidad en A de la función simple $\tilde{f}$. Pero, entonces, la función inicial f será también integrable en A, debido a (20). El teorema queda demostrado.

Desigualdad de Chébishev. *Si* $\varphi(x) \geqslant 0$ *en* A, *entonces*

$$\mu\{x : x \in A,\ \varphi(x) \geqslant c\} \geqslant \frac{1}{c} \int_A \varphi(x)\, d\mu. \tag{21}$$

En efecto, sea

$$A' = \{x : x \in A,\ \varphi(x) \geqslant c\}.$$

Entonces,

$$\int_A \varphi(x)\, d\mu = \int_{A'} \varphi(x)\, d\mu + \int_{A \setminus A'} \varphi(x)\, d\mu \geqslant \int_{A'} \varphi(x) d\mu \geqslant c\mu(A').$$

COROLARIO. *Si*

$$\int_A |f(x)|\, d\mu = 0,$$

es $f(x) = 0$ *en casi todos los puntos.*

Efectivamente, tenemos, en virtud de la desigualdad de Chébishev,

$$\mu\left\{x : x \in A,\ |f(x)| \geqslant \frac{1}{n}\right\} \leqslant n \int_A |f(x)|\, d\mu = 0$$

para todos los n. Por la tanto,

$$\mu\{x : x \in A,\ f(x) \neq 0\} \leqslant \sum_{n=1}^{\infty} \mu\left\{x : x \in A,\ \ |f(x)| \geqslant \frac{1}{n}\right\} = 0.$$

En el punto anterior hemos indicado que la integral de Lebesgue en un conjunto de medida nula es igual a cero cualquiera que sea la función f.

Esta afirmación puede ser considerada como el caso límite del siguiente teorema importante.

TEOREMA 3 (continuidad absoluta de la integral de Lebesgue). *Si $f(x)$ es una función sumable en un conjunto A, para cada $\varepsilon > 0$ existe un $\delta > 0$ tal que*

$$\left| \int_e f(x)\, d\mu \right| < \varepsilon$$

para cualquier conjunto medible $e \subset A$ tal que $\mu(e) < \delta$.

DEMOSTRACION. Observemos, ante todo, que nuestra afirmación es evidente cuando f es acotada. Sea ahora f una función arbitraria sumable en A. Tomemos

$$A_n = \{x : x \in A,\ n \leqslant |f(x)| \leqslant n+1\}$$

y

$$B_N = \bigcup_{n=0}^{N} A_n,\ C_N = A \setminus B_N.$$

Entonces, en virtud del teorema 1,

$$\int_A |f(x)|\, d\mu = \sum_{n=0}^{\infty} \int_A |f(x)|\, d\mu.$$

Escojamos N de manera que

$$\sum_{n=N+1}^{\infty} |f(x)|\, d\mu = \int_{C_N} |f(x)|\, d\mu < \frac{\varepsilon}{2},$$

y sea

$$0 < \delta < \frac{\varepsilon}{2(N+1)}.$$

Si ahora $\mu(e) < \delta$, tendremos

$$\left| \int_e f(x)\, d\mu \right| \leqslant \int_e |f(x)|\, d\mu = \int_{e \cap B_N} |f(x)|\, d\mu + \int_{e \cap C_N} |f(x)|\, d\mu.$$

La primera de las integrales del miembro derecho no pasa de $\frac{\varepsilon}{2}$ (propiedad V), mientras que la segunda no es mayor que la integral referida a todo el conjunto C_N, es decir, tampoco pasa de $\frac{\varepsilon}{2}$; luego, tenemos

$$\int_e |f(x)|\, d\mu < \varepsilon.$$

Las propiedades establecidas de la integral como función de conjunto llevan al resultado siguiente. Sea f una función no negativa, sumable en un espacio X respecto a una medida μ. Entonces, la función

$$F(A)=\int_A f(x)\,d\mu$$

está definida para todos los conjuntos medibles $A\subset X$ y es no negativa y σ-aditiva, es decir, satisface la condición: si $A=\bigcup_n A_n$ y $A_i\cap A_j=\varnothing$, se tiene $F(A)=\sum_n F(A_n)$. En otras palabras la integral de una función no negativa posee, considerándola como función de conjunto, todas las propiedades de una medida σ-aditiva. Esta medida está definida en la misma σ-álgebra en la que está definida la medida inicial μ y, además, está relacionada con μ mediante la condición: si $\mu(A)=0$, también $F(A)=0$.

4°. Paso al límite bajo el signo de la integral de Lebesgue. La cuestión sobre el paso al límite bajo el signo de la integral o, que es lo mismo, sobre la posibilidad de integrar término por término una serie convergente, surge con frecuencia en diferentes problemas.

En el Análisis clásico se demuestra que una condición suficiente para poder realizar este paso al límite es la convergencia uniforme de la sucesión (serie) correspondiente.

Demostraremos ahora ciertos teoremas acerca del paso al límite bajo el signo de la integral de Lebesgue que constituyen unas generalizaciones de largo alcance de los teoremas correspondientes del Análisis clásico.

TEOREMA 4 (Lebesgue). *Si una sucesión $\{f_n\}$ converge en A hacia f y para todo n*

$$|f_n(x)|\leqslant\varphi(x),$$

donde φ es integrable en A, la función límite f es también integrable en A y

$$\int_A f_n(x)\,d\mu\longrightarrow\int_A f(x)\,d\mu.$$

DEMOSTRACION. Se desprende fácilmente de las condiciones del teorema que $|f(x)|\leqslant\varphi(x)$. Por eso, como hemos señalado en el punto anterior (propiedad VII), $f(x)$ es integrable. Sean

$$A_k=\{x:k-1\leqslant\varphi(x)\leqslant k\};\quad B_m=\bigcup_{k\geqslant m}A_k=\{x:\varphi(x)\geqslant m\}.$$

Por el teorema 1

$$\int_A \varphi(x)\,d\mu = \sum_k \int_{A_k} \varphi(x)\,d\mu \tag{22}$$

y la serie (22) converge absolutamente. Además,

$$\int_{B_m} \varphi(x)\,d\mu = \sum_{k \geqslant m} \int_{A_k} \varphi(x)\,d\mu.$$

De la convergencia de la serie (22) se desprende la existencia de un m tal que

$$\int_{B_m} \varphi(x)\,d\mu < \frac{\varepsilon}{5}.$$

En $A \setminus B_m$ se cumple la desigualdad $\varphi(x) < m$. En virtud del teorema de Egórov, se puede representar $A \setminus B_m$ en la forma $A \setminus B_m = C \cup D$, donde $\mu(D) < \frac{\varepsilon}{5m}$ y en el conjunto C la sucesión $\{f_n\}$ converge uniformemente hacia f.

Escojamos N de manera que para $n > N$ se tenga

$$|f_n(x) - f(x)| < \frac{\varepsilon}{5\mu(C)}$$

en el conjunto C. Entonces,

$$\int_A (f_n(x) - f(x))\,d\mu = \int_{B_m} f_n(x)\,d\mu - \int_{B_m} f(x)\,d\mu + \\ + \int_D f_n(x)\,d\mu - \int_D f(x)\,d\mu + \int_C [f_n(x) - f(x)]\,d\mu.$$

De aquí

$$\int_A |f_n(x) - f(x)|\,d\mu < 5 \cdot \frac{\varepsilon}{5} = \varepsilon.$$

COROLARIO. *Si* $|f_n(x)| \leqslant M = \text{const}$ y $f_n \to f$, *tenemos*

$$\int_A f_n(x)\,d\mu \to \int_A f(x)\,d\mu.$$

Observación. Como quiera que los valores que asume una función en un conjunto de medida 0 no influyen en el valor de la integral, bastaría suponer en el teorema 4 que $\{f_n\}$ converge hacia f en casi todos los puntos y que las desigualdades $|f_n(x)| \leqslant \varphi(x)$ se cumplen también en casi todos los puntos.

TEOREMA 5 (Beppo Levi). *Supongamos que*

$$f_1(x) \leqslant f_2(x) \leqslant \ldots \leqslant f_n(x) \leqslant \ldots$$

en un conjunto A, en el que las funciones f_n son integrables, y que sus integrales son acotadas en su conjunto

$$\int_A f_n(x)\,d\mu \leqslant K.$$

Entonces, existe en casi todo el A el límite (finito)

$$f(x) = \lim_{n\to\infty} f_n(x), \tag{23}$$

la función f es integrable en A y

$$\int_A f_n(x)\,d\mu \longrightarrow \int_A f(x)\,d\mu.$$

Además, en el conjunto, en el que no existe el límite (23), la función f puede estar definida arbitrariamente, tomándose, por ejemplo, $f(x)=0$ en este conjunto.

DEMOSTRACION. Vamos a suponer que $f_1(x) \geqslant 0$, ya que el caso general se reduce a éste pasando a las funciones

$$\bar{f}_n = f_n - f_1.$$

Consideremos el conjunto

$$\Omega = \{x : x \in A,\ f_n(x) \longrightarrow \infty\}.$$

Es fácil ver que $\Omega = \bigcap_r \bigcup_n \Omega_n^{(r)}$, donde

$$\Omega_n^{(r)} = \{x : x \in A,\ \ f_n(x) > r\}.$$

En virtud de la desigualdad de Chébishev (21),

$$\mu(\Omega_n^{(r)}) \leqslant \frac{K}{r}.$$

Como $\Omega_1^{(r)} \subset \Omega_2^{(r)} \subset \ldots \subset \Omega_n^{(r)} \subset \ldots$, tenemos $\mu\left(\bigcup_n \Omega_n^{(r)}\right) \leqslant K/r$, y como para cualquier r

$$\Omega \subset \bigcup_n \Omega_n^{(r)},$$

encontramos de aquí que $\mu(\Omega) \leqslant K/r$. Por ser r arbitrario,

$$\mu(\Omega) = 0.$$

Con esto queda demostrado que la sucesión monótona $\{f_n(x)\}$ tiene en casi todo el A un límite finito $f(x)$.

Sea ahora $\varphi(x) = r$ para aquellos x en los cuales

$$r-1 \leqslant f(x) < r, \quad r = 1, 2, \ldots$$

Si demostramos la integrabilidad de $\varphi(x)$ en A, la afirmación de nuestro teorema se desprenderá inmediatamente del teorema 4.

Denotemos por A_r el conjunto de aquellos puntos $x \in A$ para los cuales $\varphi(x) = r$ y pongamos

$$B_s = \bigcup_{r=1}^{s} A_r.$$

Como las funciones f_n y f son acotadas en B_s y siempre $\varphi(x) \leqslant f(x) + 1$, tenemos

$$\int_{B_s} \varphi(x)\,d\mu \leqslant \int_{B_s} f(x)\,d\mu + \mu(A) =$$

$$= \lim_{n \to \infty} \int_{B_s} f_n(x)\,d\mu + \mu(A) \leqslant K + \mu(A).$$

Pero,

$$\int_{B_s} \varphi(x)\,d\mu = \sum_{r=1}^{s} r\mu(A_r).$$

La acotación de estas sumas implica la convergencia de la serie

$$\sum_{r=1}^{\infty} r\mu(A_r) = \int_A \varphi(x)\,d\mu.$$

Luego, hemos demostrado que φ es integrable en A.

COROLARIO. *Si* $\psi_n(x) \geqslant 0$ *y*

$$\sum_{n=1}^{\infty} \int_A \psi_n(x)\,d\mu < \infty,$$

entonces, la serie $\sum \psi_n(x)$ *converge en casi todo el* A *y*

$$\int_A \left(\sum_{n=1}^{\infty} \psi_n(x) \right) d\mu = \sum_{n=1}^{\infty} \int_A \psi_n(x)\,d\mu.$$

TEOREMA 6. (Fatou). *Si una sucesión* $\{f_n\}$ *de funciones medibles no negativas converge en casi todo el* A *hacia* f *y*

$$\int_A f_n(x)\,d\mu \leqslant K,$$

f *es integrable en* A *y*

$$\int_A f(x)\,d\mu \leqslant K.$$

DEMOSTRACION. Tomemos

$$\varphi_n(x) = \inf_{k \geqslant n} f_k(x);$$

φ_n es medible, ya que

$$\{x : \varphi_n(x) < c\} = \bigcup_{k \geqslant n} \{x : f_k(x) < c\}.$$

Además, como $0 \leqslant \varphi_n(x) \leqslant f_n(x)$, φ_n son medibles y

$$\int_A \varphi_n(x)\, d\mu \leqslant \int_A f_n(x)\, d\mu \leqslant K;$$

finalmente,

$$\varphi_1(x) \leqslant \varphi_2(x) \leqslant \ldots \leqslant \varphi_n(x) \leqslant \ldots$$

y

$$\lim_{n \to \infty} \varphi_n(x) = f(x)$$

en casi todos los puntos. Por lo tanto, aplicando a $\{\varphi_n\}$ el teorema anterior, obtenemos el resultado necesario.

5°. **Integral de Lebesgue en un conjunto de medida infinita.** Al hablar de la integral y de sus propiedades, hemos aceptado hasta este momento que se consideran funciones definidas en uno u otro conjunto medible de medida finita. Sin embargo, tropezamos frecuentemente con funciones definidas en un conjunto de medida infinita, por ejemplo, con funciones definidas en la recta con su medida lebesguiana. Por esto es importante extender el concepto de la integral también a este caso. Nos limitaremos al caso prácticamente más esencial en el que el conjunto X puede ser representado como la unión de una cantidad a lo sumo numerable de conjuntos de medida finita[1]:

$$X = \bigcup_n X_n, \quad \mu(X_n) < \infty. \tag{24}$$

Llamaremos sucesión *exhaustiva* a toda sucesión $\{X_n\}$ de subconjuntos medibles del conjunto X que verifica la condición (24). Introduzcamos la definición siguiente.

[1] Si un espacio X, en el que está definida una medida μ, puede ser representado como la unión de una cantidad numerable de conjuntos de medida finita, la medida en X se llama σ-finita. Como ejemplos de medidas σ-finitas pueden servir las medidas de Lebesgue en la recta, el plano y el espacio n-dimensional. Medidas que no son σ-finitas se pueden obtener, por ejemplo, asignando a cada punto de la recta el peso 1 y llamando medibles a todos los subconjuntos finitos de la recta (que constituyen un anillo).

DEFINICION 3. Una función medible f, definida en un conjunto X de medida σ-finita X_μ, se llama sumable en X cuando es sumable en cada subconjunto medible $A \subset X$ de medida finita y cuando para cada sucesión exhaustiva $\{X_n\}$ de conjuntos medibles existe el límite

$$\lim_{n \to \infty} \int_{X_n} f(x)\, d\mu(x). \tag{25}$$

Este límite se llama integral de f en el conjunto X y se denota con

$$\int_X f(x)\, d\mu(x).$$

Está claro que el límite (25) no depende de la selección de la sucesión exhaustiva $\{X_n\}$, ya que, de lo contrario, uniendo dos sucesiones exhaustivas podríamos construir una sucesión exhaustiva para la cual el límite (25) no existiría. Está claro también que si la función f se anula fuera de un conjunto de medida finita, la definición de la integral que acabamos de enunciar coincide con la que ha sido dada en el punto 2.

Observación. La definición de la integral de una función simple, dada en el punto 1, puede ser conservada textualmente también en el caso de medida infinita. Está claro que para que una función simple sea sumable es necesario entonces que asuma cada valor diferente del cero solamente en un conjunto de medida finita. La definición de integrabilidad, dada en el punto 2, está relacionada estrechamente con la suposición de que la medida del conjunto X sea finita. En efecto, si $\mu(X) = \infty$, la convergencia uniforme de una sucesión de funciones simples sumables $\{\varphi_n\}$ no implica, en el caso general, la convergencia de la sucesión de sus integrales (¡dése un ejemplo!).

Los resultados expuestos en los puntos 2 y 3 para el caso de medida finita subsisten, en lo fundamental, para las integrales en conjuntos de medida infinita.

La diferencia substancial consiste en que, en el caso de $\mu(X) = \infty$, una función medible acotada en X no es necesariamente sumable. En particular, si $\mu(X) = \infty$, ninguna constante diferente del cero es integrable en X.

El lector comprobará fácilmente que los teoremas de Lebesgue, de Beppo Levi y de Fatou subsisten en el caso de medida σ-finita.

6°. Comparación de la integral de Lebesgue con la integral de Riemann. Veamos la relación que existe entre la integral de Lebesque y la integral de Riemann, limitándonos al caso más sencillo de la medida lineal de Lebesgue en la recta.

TEOREMA 7. *Si existe la integral de Riemann*

$$I=(R)\int_a^b f(x)\,dx,$$

f es integrable en $[a, b]$ *según Lebesgue y*

$$\int_{[a,\,b]} f(x)\,d\mu=I.$$

DEMOSTRACION. Consideremos la partición del segmento $[a, b]$ en 2^n intervalos mediante los puntos

$$x_k=a+\frac{k}{2^n}(b-a)$$

y consideremos las sumas de Darboux correspondientes a esta partición

$$\Omega_n=\frac{b-a}{2^n}\sum_{k=1}^{2^n} M_{nk},$$

$$\omega_n=\frac{b-a}{2^n}\sum_{k=1}^{2^n} m_{nk},$$

donde M_{nk} es la cota superior de f en el segmento

$$x_{k-1}\leqslant r\leqslant x_k,$$

y m_{nk} la cota inferior de f en el mismo segmento. Por definición de la integral de Riemann,

$$I=\lim_{n\to\infty}\Omega_n=\lim_{n\to\infty}\omega_n.$$

Tomemos

$$\bar{f}_n(x)=M_{nk}\ \text{para}\ x_{k-1}\leqslant x\leqslant x_k,$$
$$\underline{f}_n(x)=m_{nk}\ \text{para}\ x_{k-1}\leqslant x\leqslant x_k.$$

En el punto $x=b$ las funciones $\bar{f}_n$ y $\underline{f}_n$ pueden ser definidas arbitrariamente. Es fácil probar que

$$\int_{[a,\,b]}\bar{f}_n(x)\,d\mu=\Omega_n,$$

$$\int_{[a,\,b]}\underline{f}_n(x)\,d\mu=\omega_n.$$

Como la sucesión $\{\bar{f}_n\}$ es no creciente y la sucesión $\{\underline{f}_n\}$ no decreciente, tenemos en casi todos los puntos

$$\bar{f}_n(x) \longrightarrow \bar{f}(x) \geqslant f(x), \quad \underline{f}_n(x) \longrightarrow \underline{f}(x) \leqslant f(x).$$

En virtud del teorema 7,

$$\int_{[a, b]} \bar{f}(x)\, d\mu = \lim_{n \to \infty} \Omega_n = I = \lim_{n \to \infty} \omega_n = \int_{[a, b]} \underline{f}(x)\, d\mu.$$

De manera que

$$\int_{[a, b]} |\bar{f}(x) - \underline{f}(x)|\, d\mu = \int_{[a, b]} \{\bar{f}(x) - \underline{f}(x)\}\, d\mu = 0$$

y, por consiguiente, en casi todos los puntos

$$\bar{f}(x) - \underline{f}(x) = 0,$$

es decir,

$$\bar{f}(x) = \underline{f}(x) = f(x)$$

y

$$\int_{[a, b]} f(x)\, d\mu = I.$$

El teorema queda demostrado.

Es fácil señalar ejemplos de funciones acotadas integrables según Lebesgue, pero no integrables según Riemann (por ejemplo, la función de Dirichlet, mencionada anteriormente, que es igual a 1 para los x racionales y al 0 para los x irracionales).

Las funciones no acotadas no son integrables según Riemann, pero muchas de ellas son integrables según Lebesgue. En particular, cualquier función $f(x) \geqslant 0$ para la cual la integral de Riemann

$$\int_{a+\varepsilon}^{b} f(x)\, dx$$

existe para cada $\varepsilon > 0$ y tiene un límite finito I para $\varepsilon \longrightarrow 0$, es integrable en $[a, b]$ según Lebesgue y

$$\int_{[a, b]} f(x)\, d\mu = \lim_{\varepsilon \to 0} \int_{a+\varepsilon}^{b} f(x)\, dx.$$

Señalemos que las integrables impropias

$$\int_{a}^{b} f(x)\, dx = \lim_{\varepsilon \to 0} \int_{a+\varepsilon}^{b} f(x)\, dx,$$

en el caso en que

$$\lim_{\varepsilon \to 0} \int_{a+\varepsilon}^{b} |f(x)|\, dx = \infty,$$

no existen en el sentido lebesguiano ya que, de acuerdo a la propiedad VIII del punto 2, la sumabilidad de la función $f(x)$ implica que la función $|f(x)|$ sea también sumable. Por ejemplo, la integral

$$\int_{0}^{1} \frac{1}{x} \operatorname{sen} \frac{1}{x}\, dx$$

existe como integral impropia de Riemann (convencionalmente convergente), pero no existe como integral de Lebesgue.

Si una función se considera en toda la recta (o en una semirrecta), su integral de Riemann puede existir solamente en el sentido impropio. Si esta integral converge absolutamente, también existirá en este caso la correspondiente integral de Lebesgue teniendo el mismo valor. En cambio, si esta integral converge convencionalmente, la función no será integrable en el sentido de Lebesgue. Por ejemplo, la función

$$\frac{\operatorname{sen} x}{x}$$

no es integrable según Lebesgue en toda la recta ya que

$$\int_{-\infty}^{\infty} \left| \frac{\operatorname{sen} x}{x} \right| dx = \infty.$$

Sin embargo, como se sabe, la integral impropia

$$\int_{-\infty}^{\infty} \frac{\operatorname{sen} x}{x}\, dx$$

existe y es igual a π.

§ 6. PRODUCTOS DIRECTOS DE SISTEMAS DE CONJUNTOS Y DE MEDIDAS. TEOREMA DE FUBINI

En el Análisis desempeñan un papel importante los teoremas sobre la reducción de una integral doble (o, en general, múltiple) a la integral reiterada. El resultado fundamental de la teoría de las integrales múltiples de Lebesgue es el teorema de Fubini que demostraremos al final de este parágrafo. Previamente expondremos algunos conceptos y resultados auxiliares que tienen, además, interés por sí mismos.

1°. **Productos de sistemas de conjuntos.** Un conjunto Z de pares ordenados (x, y), donde $x \in X$ e $y \in Y$, se llama *producto directo* de los conjuntos X e Y y se denota con $Z = X \times Y$. Del mismo modo, el conjunto Z de sucesiones finitas ordenadas $(x_1, x_2, \ldots, x_n)$, donde $x_k \in X_k$, se llama producto directo de los conjuntos $X_1, X_2, \ldots, X_n$ y se denota con

$$Z = X_1 \times X_2 \times \ldots \times X_n = |\overline{\times}| X_k.$$

En particular, cuando

$$X_1 = X_2 = \ldots = X_n = X,$$

el conjunto Z es la n-ésima potencia del conjunto X:

$$Z = X^n.$$

Por ejemplo, el espacio de coordenadas n-dimensional R^n es la n-ésima potencia de la recta numérica R^1. El cubo unidad I^n, esto es, el conjunto de elementos de R^n con coordenadas que verifican las desigualdades

$$0 \leqslant x_k \leqslant 1, \quad k = 1, 2, \ldots, n$$

constituye la n-ésima potencia del segmento unidad $I^1 = [0, 1]$.

Si $\mathfrak{S}_1, \mathfrak{S}_2, \ldots, \mathfrak{S}_n$ son sistemas de subconjuntos de los conjuntos $X_1, X_2, \ldots, X_n$, entonces,

$$\mathfrak{R} = \mathfrak{S}_1 \times \mathfrak{S}_2 \times \ldots \times \mathfrak{S}_n$$

representa el sistema de subconjuntos del conjunto $X = |\overline{\times}| X_k$ que se pueden escribir en la forma

$$A = A_1 \times A_2 \times \ldots \times A_n,$$

donde $A_k \in \mathfrak{S}_k$.

Si $\mathfrak{S}_1 = \mathfrak{S}_2 = \ldots = \mathfrak{S}_n = \mathfrak{S}$, es $\mathfrak{R}$ la n-ésima potencia de $\mathfrak{S}$:

$$\mathfrak{R} = \mathfrak{S}^n.$$

Por ejemplo, el sistema de paralelepípedos de R^n es la n-ésima potencia del sistema de segmentos de R^1.

TEOREMA 1. *Si $\mathfrak{S}_1, \mathfrak{S}_2, \ldots, \mathfrak{S}_n$ son semianillos, también $\mathfrak{R} = |\overline{\times}| \mathfrak{S}_k$ es un semianillo.*

DEMOSTRACION. De acuerdo con la definición de semianillo debemos probar que si $A, B \in \mathfrak{R}$, entonces $A \cap B \in \mathfrak{R}$ y que, además, para $B \subset A$

se tiene $A=\bigcup_{i=1}^{m} C_i$, donde $C_1=B$, $C_i \cap C_j=\varnothing$ para $i \neq j$ y $C_i \in \mathfrak{R}$ $(i=1, 2, \ldots, n)$.

Realicemos la demostración para el caso en que $n=2$.

I. Sea $A \in \mathfrak{S}_1 \times \mathfrak{S}_2$, $B \in \mathfrak{S}_1 \times \mathfrak{S}_2$; esto significa que

$$A=A_1 \times A_2, \quad A_1 \in \mathfrak{S}_1, \quad A_2 \in \mathfrak{S}_2,$$
$$B=B_1 \times B_2, \quad B_1 \in \mathfrak{S}_1, \quad B_2 \in \mathfrak{S}_2.$$

Entonces,

$$A \cap B=(A_1 \cap B_1) \times (A_2 \cap B_2)$$

y como

$$A_1 \cap B_1 \in \mathfrak{S}_1, \quad A_2 \cap B_2 \in \mathfrak{S}_2,$$

tenemos

$$A \cap B \in \mathfrak{S}_1 \times \mathfrak{S}_2.$$

II. Supongamos ahora adicionalmente que $B \subset A$. Entonces,

$$B_1 \subset A_1, \quad B_2 \subset A_2$$

y, como $\mathfrak{S}_1$ y $\mathfrak{S}_2$ son semianillos, tienen lugar las descomposiciones

$$A_1=B_1 \cup B_1^{(1)} \cup \ldots \cup B_1^{(k)},$$
$$A_2=B_2 \cup B_2^{(1)} \cup \ldots \cup B_2^{(l)},$$
$$A=A_1 \times A_2=(B_1 \times B_2) \cup (B_1 \times B_2^{(1)}) \cup \ldots \cup (B_1 \times B_2^{(l)}) \cup$$
$$\cup (B_1^{(1)} \times B_2) \cup (B_1^{(1)} \times B_2^{(1)}) \cup \ldots \cup (B_1^{(1)} \times B_1^{(l)}) \cup$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\cup (B_1^{(k)} \times B_2) \cup (B_1^{(k)} \times B_2^{(1)}) \cup \ldots \cup (B_1^{(k)} \times B_2^{(l)}).$$

En la última descomposición el primer término es $B_1 \times B_2=B$ y todos los términos pertenecen al sistema $\mathfrak{S}_1 \times \mathfrak{S}_2$. Esto demuestra el teorema.

Sin embargo, si los sistemas $\mathfrak{S}_k$ son anillos o anillos borelianos, ello no implica aún, en el caso general, que el producto $\left|\overline{\times}_k\right| \mathfrak{S}_k$ sea un anillo o, respectivamente, un anillo boreliano.

2°. Productos de medidas. Supongamos que en los semianillos $\mathfrak{S}_1, \mathfrak{S}_2, \ldots, \mathfrak{S}_n$ se han definido las medidas

$$\mu_1(A_1), \ \mu_2(A_2), \ \ldots, \ \mu_n(A_n) \ \ A_k \in \mathfrak{S}_k.$$

Definamos en

$$\mathfrak{R}=\mathfrak{S}_1 \times \mathfrak{S}_2 \times \ldots \times \mathfrak{S}_n$$

la medida

$$\mu=\mu_1 \times \mu_2 \times \ldots \times \mu_n$$

de la siguiente forma: si

$$A = A_1 \times A_2 \times \ldots \times A_n,$$

entonces,

$$\mu(A) = \mu_1(A_1)\mu_2(A_2) \ \ldots \ \mu_n(A_n).$$

Es preciso demostrar que $\mu(A)$ es una medida, esto es, que $\mu(A)$ es aditiva. Haremos esto para el caso en que $n=2$. Supongamos que se tiene la descomposición

$$A = A_1 \times A_2 = \bigcup_k B^{(k)}, \quad B^{(i)} \cap B^{(j)} = \varnothing \text{ para } i \neq j,$$

$$B^{(k)} = B_1^{(k)} \times B_2^{(k)}.$$

En el capítulo I (lema 2 del § 5) ha sido demostrada la existencia de descomposiciones

$$A_1 = \bigcup_m C_1^{(m)}, \quad A_2 = \bigcup_n C_2^{(n)}$$

tales que los conjuntos $B_1^{(k)}$ son uniones de ciertos $C_1^{(m)}$ y los conjuntos $B_2^{(k)}$ son uniones de determinados $C_2^{(n)}$. Es obvio que

$$\mu(A) = \mu_1(A_1)\mu_2(A_2) = \sum_n \sum_m \mu_1(C_1^{(m)})\mu_2(C_2^{(n)}), \tag{1}$$

$$\mu(B^{(k)}) = \mu_1(B_1^{(k)})\mu_2(B_2^{(k)}) = \sum_m \sum_n \mu_1(C_1^{(m)})\mu_2(C_2^{(n)}), \tag{2}$$

donde en el miembro derecho de la igualdad (1) aparecen justamente una vez todos los términos que figuran en los miembros derechos de las igualdades (2). Por lo tanto,

$$\mu(A) = \sum_k \mu(B_k),$$

que es lo que debíamos demostrar.

En particular, la aditividad de las medidas elementales en un espacio euclídeo n-dimensional se desprende de la aditividad de la medida lineal en la recta.

TEOREMA 2. *Si las medidas $\mu_1, \mu_2, \ldots, \mu_n$ son σ-aditivas, también es σ-aditiva la medida $\mu = \mu_1 \times \mu_2 \times \ldots \times \mu_n$.*

Daremos la demostración para el caso en que $n=2$. Supongamos que λ_1 es la prolongación lebesguiana de la medida μ_1. Sea $C = \bigcup_{n=1}^{\infty} C_n$, donde los conjuntos C y C_n pertenecen a $\mathfrak{S}_1 \times \mathfrak{S}_2$,

es decir,

$$C = A \times B, \quad A \in \mathfrak{S}_1, \quad B \in \mathfrak{S}_2,$$
$$C_n = A_n \times B_n, \quad A_n \in \mathfrak{S}_1, \quad B_n \in \mathfrak{S}_2.$$

Tomemos para $x \in X$

$$f_n(x) = \begin{cases} \mu_2(B_n) \text{ cuando } x \in A_n, \\ 0 \text{ cuando } x \overline{\in} A_n. \end{cases}$$

Es fácil ver que para $x \in A$

$$\sum_n f_n(x) = \mu_2(B)$$

y por esto, de acuerdo con el teorema de Beppo Levi (teorema 5 del § 5)

$$\sum_n \int_A f_n(x)\, d\lambda_1 = \int_A \mu_2(B)\, d\lambda_1 = \lambda_1(A) \cdot \mu_2(B) = \mu_1(A) \cdot \mu_2(B).$$

Pero,

$$\int_A f_n(x)\, d\lambda_1 = \mu_2(B_n) \cdot \mu_1(A_n) = \mu(C_n)$$

de modo que

$$\sum_n \mu(C_n) = \mu(C).$$

La prolongación lebesguiana de la medida $\mu_1 \times \mu_2 \times \ldots \times \mu_n$ se llamará producto de las medidas μ_k y se denotará con

$$\mu_1 \otimes \mu_2 \otimes \ldots \otimes \mu_n = \boxed{\otimes}\, \mu_k.$$

En particular, para

$$\mu_1 = \mu_2 = \ldots = \mu_n = \mu$$

obtenemos la n-ésima potencia de la medida μ:

$$\mu^n = \boxed{\bigotimes_{k=1}^{n}}\, \mu_k, \quad \mu_k = \mu.$$

Por ejemplo, la medida n-dimensional de Lebesgue μ^n es la n-ésima potencia de la medida lineal de Lebesgue μ.

3°. Representación de la medida plana en términos de la integral de la medida lineal de secciones y definición geométrica de la integral de Lebesgue. Sea G una región del plano (x, y) limitada por las verticales $x = a$, $x = b$ y las curvas $y = \varphi(x)$, $y = \psi(x)$.

Como es conocido, el área de la región G es igual a la integral

$$V(G) = \int_a^b \{\varphi(x) - \psi(x)\}\, dx.$$

La diferencia $\varphi(x_0)-\psi(x_0)$ representa aquí la longitud de la sección de la región G mediante la vertical $x=x_0$. Nuestro objetivo es extender este modo de medir áreas al caso de medidas-productos arbitrarios

$$\mu=\mu_x \otimes \mu_y.$$

Vamos a suponer en lo sucesivo que las medidas μ_x y μ_y están definidas en álgebras borelianas, son σ-aditivas y verifican la condición de plenitud (si $B\subset A$ y $\mu(A)=0$, entonces B es medible), condición que, como hemos visto con anterioridad, verifican todas las prolongaciones lebesguianas.

Introduciremos las denotaciones siguientes

$$A_x=\{y\colon (x, y)\in A\} \qquad (x \text{ fijo}),$$
$$A_y=\{x\colon (x, y)\in A\} \qquad (y \text{ fijo}).$$

Si X e Y son rectas numéricas (de modo que $X\times Y$ es el plano), A_{x_0} es la proyección sobre el eje Y de la sección del conjunto A mediante la recta vertical $x=x_0$.

TEOREMA 3. *En las suposiciones señaladas se tiene*

$$\mu(A)=\int_X \mu_y(A_x)\,d\mu_x=\int_Y \mu_x(A_y)\,d\mu_y$$

para cualquier conjunto μ-medible A [1].

DEMOSTRACION. Es obvio que basta demostrar la igualdad

$$\mu(A)=\int_X \varphi_A(x)\,d\mu_x, \quad \text{donde } \varphi_A(x)=\mu_y(A_x), \tag{3}$$

ya que la segunda parte del teorema es totalmente análoga a la primera. Observemos que el teorema incluye automáticamente la aseveración de que para casi todos los x (en el sentido de la medida μ_x) los conjuntos A_x son medibles según la medida μ_y y de que la función $\varphi_A(x)$ es medible respecto de la medida μ_x. De lo contrario, la fórmula (3) no tendría sentido.

---

[1] Observemos que la integración en X se reduce, de hecho, a la integración en el conjunto $\bigcup_y A_y \subset X$ fuera del cual el integrando es igual a cero.

Análogamente, $\int_Y = \int_{\bigcup_x A_x}$.

La medida μ es la extensión lebesguiana de la medida
$$m=\mu_x \times \mu_y$$
definida en el sistema $\mathfrak{S}_m$ de conjuntos de tipo
$$A=A_{y_0} \times A_{x_0}.$$

Para estos conjuntos la igualdad (3) es evidente, ya que en este caso
$$\varphi_A(x)=\begin{cases} \mu_y(A_{x_0}) & \text{para } x \in A_{y_0}, \\ 0 & \text{para } x \overline{\in} A_{y_0}. \end{cases}$$

La igualdad (3) se extiende sin dificultad también a los conjuntos de $\mathfrak{R}(\mathfrak{S}_m)$ que se descomponen en uniones finitas de conjuntos disjuntos dos a dos de $\mathfrak{S}_m$.

En el caso general, la demostración de la igualdad (3) se basa en el lema siguiente que tiene interés independiente en la teoría de prolongaciones lebesguianas.

LEMA. *Para cualquier conjunto μ-medible A existe un conjunto B tal que*
$$B=\bigcap_n B_n, \quad B_1 \supset B_2 \supset \ldots \supset B_n \supset \ldots$$
$$B_n=\bigcup_k B_{nk}, \quad B_{n1} \subset B_{n2} \subset \ldots \subset B_{nk} \subset \ldots,$$
donde los conjuntos B_{nk} son elementos de $\mathfrak{R}(\mathfrak{S}_m)$ y $A \subset B$ y
$$\mu(A)=\mu(B).$$

La DEMOSTRACION del lema se basa en el hecho de que, cualquiera que sea n, el conjunto A puede ser incluido, por definición de medibilidad, en la unión $C_n=\bigcup_r \Delta_{nr}$ de conjuntos Δ_{nr} de $\mathfrak{S}_m$ de manera que
$$\mu(C_n)<\mu(A)+\frac{1}{n}.$$

Tomando $B_n=\bigcap_{k=1}^{n} C_k$, veremos sin dificultad que los conjuntos B_n tienen la forma $B_n=\bigcup_s \delta_{ns}$ donde δ_{ns} son elementos de $\mathfrak{S}_m$. Tomando, finalmente, $B_{nk}=\bigcup_{s=1}^{k} \delta_{ns}$, obtendremos un sistema de conjuntos B_{nk} con las propiedades requeridas. Esto demuestra el lema.

Empleando el teorema de Beppo Levi (teorema 5 del § 5), es fácil extender la igualdad (3) de los conjuntos $B_{nk} \in \mathfrak{R}(\mathfrak{S}_m)$ a los conjuntos B_n y B, ya que

$$\varphi_{B_n}(x) = \lim_{k \to \infty} \varphi_{B_{nk}}(x), \quad \varphi_{B_{n1}} \leqslant \varphi_{B_{n2}} \leqslant \ldots,$$

$$\varphi_B(x) = \lim_{n \to \infty} \varphi_{B_n}(x), \quad \varphi_{B_1} \geqslant \varphi_{B_2} \geqslant \ldots .$$

Si $\mu(A) = 0$, entonces $\mu(B) = 0$ y en casi todos los puntos

$$\varphi_B(x) = \mu_y(B_x) = 0.$$

Como $A_x \subset B_x$, para casi todos los x el conjunto A_x es medible y

$$\varphi_A(x) = \mu_y(A_x) = 0,$$

$$\int \varphi_A(x)\, d\mu_x = 0 = \mu(A).$$

Por consiguiente, la fórmula (3) es válida para conjuntos A tales que $\mu(A) = 0$. En el caso general, representaremos A en la forma $B \setminus C$, donde, en virtud de (4),

$$\mu(C) = 0.$$

Como la fórmula (3) es válida para los conjuntos B y C, es fácil probar que se cumple también para el conjunto A. Hemos terminado la demostración del teorema (3).

Consideremos ahora el caso en que Y es la recta numérica, μ_y es la medida lineal de Lebesgue y el conjunto A es el conjunto formado por puntos (x, y) de tipo

$$\{x \in M,\ 0 \leqslant y \leqslant f(x)\}, \tag{5}$$

donde M es un conjunto μ_x-medible y $f(x)$ es una función integrable no negativa. En este caso,

$$\mu_y(A_x) = \begin{cases} f(x) & \text{para } x \in M, \\ 0 & \text{para } x \overline{\in} M \end{cases}$$

y

$$\mu(A) = \int_M f(x)\, d\mu_x.$$

Hemos demostrado con esto el teorema siguiente.

TEOREMA 4. *La integral de Lebesgue de una función no negativa $f(x)$ es igual a la medida $\mu = \mu_x \times \mu_y$ del conjunto A definido por las relaciones* (5).

En el caso en que X es también la recta numérica, el conjunto M es un segmento y la función $f(x)$ es integrable según Riemann, este teorema se reduce a la expresión conocida de la integral como el área comprendida debajo del gráfico de la función.

4°. Teorema de Fubini. Consideremos el producto triple $U=X\times Y\times Z$; si en X, Y y Z están definidas las medidas μ_x, μ_y y μ_z, la medida

$$\mu_u=\mu_x\otimes\mu_y\otimes\mu_z$$

puede ser definida o bien como

$$\mu_u=(\mu_x\otimes\mu_y)\otimes\mu_z$$

o bien como

$$\mu_u=\mu_x\otimes(\mu_y\otimes\mu_z).$$

Es fácil probar que, de hecho, estas definiciones son equivalentes.

El teorema siguiente constituye el resultado principal en la teoría de integrales múltiples.

TEOREMA DE FUBINI. *Supongamos que las medidas μ_x y μ_y están definidas en anillos borelianos, son σ-aditivas y completas; supongamos, además, que*

$$\mu=\mu_x\otimes\mu_y$$

y que la función $f(x, y)$ es integrable respecto a la medida μ en un conjunto

$$A\subset X\times Y. \tag{6}$$

Entonces[1],

$$\int_A f(x, y)\,d\mu=\int_X\left(\int_{A_x} f(x, y)\,d\mu_y\right)d\mu_x=$$

$$=\int_Y\left(\int_{A_y} f(x, y)\,d\mu_x\right)d\mu_y. \tag{7}$$

La afirmación del teorema incluye la existencia de las integrales en los paréntesis para casi todos los valores de la variable respecto a la cual se toman estas integrales.

DEMOSTRACION. Realicemos primero la demostración para el caso en que $f(x, y)\geqslant 0$. Consideremos con este fin el producto triple

$$U=X\times Y\times Z,$$

[1] Véase la llamada al pie de la página 356.

donde el tercer factor es la recta numérica, y el producto de medidas

$$\lambda = \mu_x \otimes \mu_y \otimes \mu^1 = \mu \otimes \mu^1,$$

donde μ^1 es la medida lebesguiana lineal.

Definamos en U un subconjunto W tomando

$$(x, y, z) \in W$$

cuando

$$x \in A, \quad y \in A_x,$$
$$0 \leqslant z \leqslant f(x, y).$$

De acuerdo con el teorema 4,

$$\lambda(W) = \int_A f(x, y)\, d\mu. \tag{8}$$

Por otra parte, en vista del teorema 3,

$$\lambda(W) = \int_X \xi(W_x)\, d\mu_x, \tag{9}$$

donde $\xi = \mu_y \times \mu^1$ y W_x es el conjunto de pares (y, z) tales que $(x, y, z) \in W$. Además, de acuerdo con el teorema 3,

$$\xi(W_x) = \int_{A_x} f(x, y)\, d\mu_y. \tag{10}$$

Comparando (8), (9) y (10), encontramos

$$\int_A f(x, y)\, d\mu = \int_X \left(\int_{A_x} f(x, y)\, d\mu_y \right) d\mu_x,$$

que es los que queríamos demostrar.

El caso general se reduce al estudiado mediante las relaciones

$$f(x, y) = f^+(x, y) - f^-(x, y),$$
$$f^+(x, y) = \frac{|f(x, y)| + f(x, y)}{2}, \quad f^-(x, y), = \frac{|f(x, y)| - f(x, y)}{2}.$$

Observación. Como veremos en los ejemplos que indicamos más abajo, la existencia de las integrales reiteradas

$$\int_X \left(\int_{A_x} f\, d\mu_y \right) d\mu_x \quad \text{y} \quad \int_Y \left(\int_{A_y} f\, d\mu_x \right) d\mu_y \tag{11}$$

no implica, en el caso general, ni la igualdad (7) ni la integrabilidad de la función $f(x, y)$ en A. Sin embargo, *si existe al menos una de las integrales*

$$\int_X \left(\int_{A_x} |f(x, y)|\, d\mu_y \right) d\mu_x \quad \text{ó} \quad \int_Y \left(\int_{A_y} |f(x, y)|\, d\mu_x \right) d\mu_y, \tag{12}$$

$f(x, y)$ es integrable en A y tiene lugar la igualdad (7).

En efecto, supongamos, por ejemplo, que la primera de las integrales (12) existe y es igual a M. La función $f_n(x, y) = \min \times \times \{|f(x, y)|, n\}$ es medible, acotada y, consecuentemente, sumable en A. Por el teorema de Fubini

$$\int_A f_n(x, y)\, d\mu = \int_X \left(\int_{A_x} f_n(x, y)\, d\mu_y \right) d\mu_x \leqslant M. \tag{13}$$

Las funciones f_n forman una sucesión monótona no decreciente que converge en casi todos los puntos hacia $|f(x, y)|$. En virtud del teorema de Beppo Levi, de aquí y de la desigualdad (13) se desprende que la función $|f(x, y)|$ es sumable en A. Pero, entonces, también $f(x, y)$ es sumable y para ella es válido el teorema de Fubini. De aquí se deduce nuestra afirmación.

Hemos demostrado el teorema de Fubini suponiendo que las medidas μ_x y μ_y (y, por consiguiente, también μ) son finitas. Sin embargo, es fácil probar que este teorema subsiste también en el caso de medidas σ-finitas.

Veamos unos ejemplos de funciones, para las cuales existen las integrales reiteradas (11), pero no tiene lugar la igualdad (7).

1. Sea

$$A = [-1, 1]^2$$

y

$$f(x, y) = \frac{xy}{(x^2 + y^2)^2};$$

entonces,

$$\int_{-1}^{1} f(x, y)\, dx = 0$$

para $y \neq 0$ y

$$\int_{-1}^{1} f(x, y)\, dy = 0$$

para $x \neq 0$. Por lo tanto,

$$\int_{-1}^{1} \left(\int_{-1}^{1} f(x, y)\, dx \right) dy = \int_{-1}^{1} \left(\int_{-1}^{1} f(x, y)\, dy \right) dx = 0;$$

pero la integral, en el sentido de integral doble de Lebesgue, en el cuadrado no existe, ya que

$$\int_{-1}^{1} \int_{-1}^{1} |f(x, y)|\, dx\, dy \geqslant \int_0^1 dr \int_0^{2\pi} \frac{|\operatorname{sen} \varphi \cos \varphi|}{r}\, d\varphi = 2 \int_0^1 \frac{dr}{r} = \infty.$$

2. $A=[0, 1]^2$

$$f(x, y)=\begin{cases} 2^n \text{ para } \frac{1}{2^n}\leqslant x\leqslant\frac{1}{2^{n-1}}; \ \frac{1}{2^n}\leqslant y<\frac{1}{2^{n-1}}, \\ -2^{2n+1} \text{ para } \frac{1}{2^{n+1}}\leqslant x<\frac{1}{2^n}; \ \frac{1}{2^n}\leqslant y<\frac{1}{2^{n-1}}, \\ 0 \text{ en los demás casos.} \end{cases}$$

Se puede calcular que

$$\int_0^1\left(\int_0^1 f(x, y)\,dx\right)dy=0$$

mientras que

$$\int_0^1\left(\int_0^1 f(x, y)\,dy\right)dx=\frac{1}{4}.$$

CAPITULO VII

INTEGRAL INDEFINIDA DE LEBESGUE. TEORIA DE DIFERENCIACION

En este capítulo continuaremos el estudio de la integral de Lebesgue, limitándonos fundamentalmente al caso de funciones en la recta y aceptando que la medida, respecto a la cual se toma esta integral, es la medida habitual lineal de Lebesgue.

Si f es una función sumable definida en un conjunto medible X de medida μ, la integral

$$\int_A f(x)\, d\mu \tag{1}$$

existe para cada $A \subset X$ medible y, siendo f fija, representa una función de conjunto definida para todos los subconjuntos medibles $A \subset X$. Si la función f está definida en un segmento de la recta numérica y el conjunto A, respecto al cual se toma la integral (1), también es un segmento, esta integral es función de par de puntos, esto es, de los extremos del segmento A. Fijando uno de los extremos del segmento de integración, digamos el izquierdo, podemos considerar la integral referida al segmento $[a, x]$ como función de una variable x. Estudiaremos las propiedades de la integral

$$\int_a^x f(t)\, dt,$$

tomada en el segmento $[a, x]$ con la medida habitual lineal de Lebesgue en este segmento, como función del extremo superior de integración x. Este problema nos llevará a considerar algunas clases importantes de funciones en la recta. El estudio general de la integral de Lebesgue de una función fija f como función de conjunto se realiza en el § 5.

Son conocidas del curso elemental del Análisis las siguientes igualdades fundamentales que establecen la relación entre las operaciones de diferenciación e integración: si f es una función continua y F una función de derivada continua, entonces,

1) $$\frac{d}{dx}\int_a^x f(t)\,dt = f(x),$$

2) $$\int_a^b F'(t)\,dt = F(b) - F(a).$$

Surgen las preguntas: a) ¿es válida la igualdad 1) para funciones sumables en el sentido de Lebesgue? y b) ¿cuál es la clase (la más amplia posible) de funciones para la que se verifica la igualdad 2)?

Estas cuestiones son estudiadas en los parágrafos próximos del presente capítulo.

§ 1. FUNCIONES MONÓTONAS. DIFERENCIABILIDAD DE LA INTEGRAL RESPECTO AL EXTREMO SUPERIOR

1°. Propiedades fundamentales de funciones monótonas. Comenzaremos el estudio de las propiedades de la integral de Lebesgue

$$\Phi(x) = \int_a^x f(t)\,dt \tag{1}$$

como función del extremo superior con la siguiente observación obvia, pero importante: si la función f es no negativa, $\Phi(x) = \int_a^x f(t)\,dt$ es una función monótona no decreciente; además, como toda función sumable es diferencia de dos funciones sumables no negativas:

$$f(t) = f^+(t) - f^-(t), \tag{2}$$

la integral (1) es la diferencia de dos funciones monótonas no decrecientes. Consecuentemente, el estudio de la integral de Lebesgue como función del extremo superior puede ser reducido al estudio de funciones monótonas del mismo género. Las funciones monótonas (independientemente de su origen) poseen una serie de propiedades simples e importantes que pasamos a exponer.

Recordemos algunos conceptos necesarios para lo sucesivo. Siempre que no se diga lo contrario, se considerarán funciones definidas en un segmento.

Una función f se llama *monótona no decreciente* cuando $x_1 \leqslant x_2$ implica

$$f(x_1) \leqslant f(x_2);$$

análogamente, f se llama *monótona no creciente* cuando $x_1 \leqslant x_2$ implica

$$f(x_1) \geqslant f(x_2).$$

Sea f una función arbitraria en la recta. El límite

$$\lim_{\substack{h \to 0 \\ h > 0}} f(x_0 + h)$$

(si es que existe) se llama *límite a la derecha* de la función f en el punto x_0 y se denota con $f(x_0 + 0)$. Análogamente, *el límite a la izquierda* $f(x_0 - 0)$ de la función f en el punto x_0 se define como

$$\lim_{\substack{h \to 0 \\ h > 0}} f(x_0 - h).$$

La igualdad $f(x_0 + 0) = f(x_0 - 0)$ significa, obviamente, que la función f es continua en el punto x_0 o tiene en él una discontinuidad evitable. Un punto en el que $f(x_0 + 0)$ y $f(x_0 - 0)$ existen, pero no coinciden, se llama *punto de discontinuidad de primera especie* y la diferencia $f(x_0 + 0) - f(x_0 - 0)$ se llama *salto* de la función f en este punto.

Una función f se llama *continua a la izquierda* en el punto x_0 cuando $f(x_0) = f(x_0 - 0)$ y *continua a la derecha* en este punto cuando $f(x_0) = f(x_0 + 0)$.

Demostremos las propiedades fundamentales de las funciones monótonas. Para concretar, hablaremos de funciones monótonas no decrecientes aunque todo lo que se dice a continuación se extiende automáticamente a las funciones monótonas no crecientes.

1. *Toda función f monótona no decreciente en $[a, b]$ es medible y acotada y, por consiguiente, sumable.*

En efecto, debido a la monotonía,

$$f(x) \leqslant f(b) \text{ en } [a, b].$$

Además, para cualquier constante c el conjunto

$$A_c = \{x : f(x) < c\}]$$

es o bien un segmento o bien un semisegmento (o el conjunto vacío). Efectivamente, si existen puntos en los que $f(x) < c$, designemos mediante d la cota superior mínima de todos estos x. Entonces, A_c es o bien el segmento $[a, d]$ o bien el semisegmento $[a, d)$.

2. *Una función monótona puede tener solamente discontinuidad de primera especie.*

En efecto, supongamos que x es un punto arbitrario de $[a, b]$ y que $x_n \longrightarrow x_0$ siendo $x_n < x_0$. En este caso, $\{f(x_n)\}$ es una sucesión monótona no decreciente acotada por arriba (por el valor $f(x_0)$, por ejemplo). Luego, para cualquier sucesión de este tipo existe $\lim\limits_{n \to \infty} f(x_n)$, es decir, existe $f(x_0-0)$. De manera análoga se demuestra la existencia de $f(x_0+0)$.

Es evidente que una función monótona no es necesariamente continua. No obstante, es válida la proposición siguiente.

3. *El conjunto de los puntos de discontinuidad de una función monótona es a lo sumo numerable.*

Efectivamente, la suma de los saltos de una función f monótona en $[a, b]$ no pasa de $f(b)-f(a)$. Consecuentemente, para cada n el número de saltos de magnitud mayor que $\frac{1}{n}$ es finito. Sumando respecto a todos los $n=1, 2, \ldots$, obtenemos que el número total de saltos es finito o numerable.

Entre las funciones monótonas las más sencillas son las así llamadas *funciones de saltos*. Son funciones que se obtienen del siguiente modo. Supongamos que en un segmento $[a, b]$ se ha escogido una cantidad finita o numerable de puntos

$$x_1, x_2, \ldots, x_n, \ldots$$

y supongamos que a cada uno de estos puntos se ha asignado un número positivo h_n de manera que $\sum\limits_n h_n < \infty$. Definamos en $[a, b]$ una función f tomando

$$f(x) = \sum_{x_n < x} h_n. \tag{3}$$

Esta función es, obviamente, monótona no decreciente. Además, es continua *a la izquierda* en cada punto, el conjunto de sus puntos de discontinuidad coincide con el conjunto $\{x_n\}$ y el salto en el punto x_n es igual a h_n. En efecto,

$$f(x-0) = \lim_{\substack{\varepsilon \to 0 \\ \varepsilon > 0}} f(x-\varepsilon) = \lim_{\substack{\varepsilon \to 0 \\ \varepsilon > 0}} \sum_{x_n < x-\varepsilon} h_n$$

y, como cada x_n que verifica la condición $x_n < x$ también verifica la condición $x_n < x-\varepsilon$ para ε suficientemente pequeño, el último límite es igual a $\sum\limits_{x_n<x} h_n = f(x)$. Luego[1],

$$f(x-0)=f(x).$$

Si el punto x coincide con uno de los puntos x_n, digamos $x=x_{n_0}$, se tiene

$$f(x_{n_0}+0)=\lim_{\varepsilon\to 0} f(x_{n_0}+\varepsilon)=\lim_{\varepsilon\to 0}\sum_{x_n<x_{n_0}+\varepsilon} h_n=\sum_{x_n\leqslant x_{n_0}} h_n,$$

es decir, $f(x_{n_0}+0)-f(x_{n_0}-0)=h_{n_0}$.

En lo sucesivo entendremos por función de saltos cualquier función que puede ser obtenida mediante la construcción descrita. El tipo más sencillo de funciones de saltos son las funciones escalonadas, cuyos puntos de discontinuidad pueden ser representados mediante una sucesión monótona

$$x_1<x_2<\ldots<x_n<\ldots$$

En el caso general, una función de saltos puede tener una estructura más compleja; por ejemplo, si $\{x_n\}$ es el conjunto de todos los puntos racionales del segmento $[a, b]$ y $h_n=\frac{1}{2^n}$, la fórmula (3) determina una función de saltos discontinua en los puntos racionales y continua en los irracionales.

Otra clase de funciones monótonas, en cierto sentido opuesta a la de las funciones de saltos, es la formada por las funciones monótonas continuas. Tiene lugar la afirmación siguiente.

4. *Toda función monótona puede representarse como suma de una función monótona continua y de una función de saltos.*

En efecto, sea f una función no decreciente cualquiera, sean $x_1, x_2, \ldots$ sus puntos de discontinuidad y sean $h_1, h_2, \ldots$ sus saltos en estos puntos. Tomemos

$$H(x)=\sum_{x_n<x} h_n.$$

La diferencia

$$\varphi=f-H$$

es una función no decreciente continua. Consideremos, para demostrar esto, la diferencia

$$\varphi(x'')-\varphi(x')=[f(x'')-f(x')]-[H(x'')-H(x')].$$

---

[1] Si hubiésemos definido f mediante la fórmula

$$f(x)=\sum_{x_n\leqslant x} h_n,$$

obtendríamos una función continua *a la derecha*.

La expresión que figura en el miembro derecho es la diferencia entre el incremento total de la función f en el segmento $[x', x'']$ y la suma de sus saltos en este segmento, es decir, coincide con la medida del conjunto de valores que toma esta función en sus puntos de continuidad pertenecientes a $[x', x'']$. Evidentemente, esta magnitud es no negativa y, por lo tanto, φ es una función no decreciente. Además, para un punto x^* arbitrario tenemos

$$\varphi(x^*-0) = \lim_{x \to x^*-0} f(x) - \lim_{x \to x^*-0} H(x) = f(x^*-0) - \sum_{x_n < x^*} h_n,$$

$$\varphi(x^*+0) = f(x^*+0) - H(x^*+0) = f(x^*+0) - \sum_{x_n \leqslant x^*} h_n,$$

de donde

$$\varphi(x^*+0) - \varphi(x^*-0) = f(x^*+0) - f(x^*-0) - h^* = 0$$

(aquí h^* es el salto de la función H en el punto x^*).

En consecuencia, φ es efectivamente continua.

2°. Diferenciabilidad de una función monótona. Después de haber expuesto estas propiedades elementales de las funciones monótonas, pasemos a estudiar el problema sobre la existencia de la derivada de una función monótona.

TEOREMA 1 (Lebesgue). *Una función monótona f definida en un segmento $[a, b]$ tiene derivada finita en casi todos los puntos de este segmento.*

Introduciremos, ante todo, algunos conceptos que emplearemos en la demostración de este teorema.

Como es conocido, la derivada de una función f en un punto x_0 es el límite del cociente

$$\frac{f(x)-f(x_0)}{x-x_0} \tag{4}$$

para $x \to x_0$. Este límite puede, por supuesto, no existir, pero siempre tienen sentido las cuatro magnitudes siguientes (que pueden tomar también valores infinitos):

Λ_d que es el límite superior del cociente (4) cuando x tiende a x_0 por la derecha (es decir, de manera que $x - x_0 > 0$). Esta magnitud se llama número derivado superior derecho.

λ_d (número derivado inferior derecho) que es el límite inferior del cociente (4) cuando $x \to x_0$ por la derecha.

Λ_i (número derivado superior izquierdo) que es el límite superior del cociente (4) cuando $x \to x_0$ por la izquierda.

λ_i (número derivado inferior izquierdo) que es el límite inferior del cociente (4) cuando $x \to x_0$ por la izquierda.

La fig. 20 aclara el contenido geométrico de estas magnitudes. Está claro que siempre

$$\lambda_d \leqslant \Lambda_d \quad \text{y} \quad \lambda_i \leqslant \Lambda_i.$$

Si Λ_d y λ_d son finitos y coinciden, su valor común es la derivada de la función $f(x)$ en el punto x_0 a la derecha. Análo-

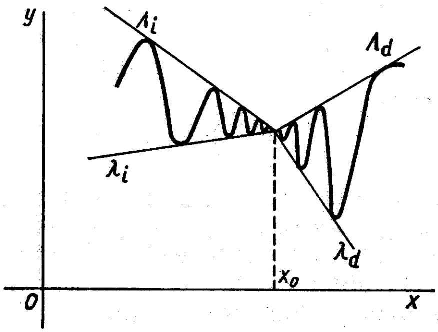


FIG. 20

gamente, si $\Lambda_i = \lambda_i$, su valor común es la derivada a la izquierda. La existencia de la derivada finita de f en el punto x_0 equivale a que en este punto son finitos y coinciden todos los números derivados de la función f. Por lo tanto, el teorema de Lebesgue puede ser enunciado del siguiente modo: *para una función monótona en* $[a, b]$ *las relaciones*

$$-\infty < \lambda_i = \lambda_d = \Lambda_i = \Lambda_d < \infty$$

se cumplen en casi todos los puntos de $[a, b]$.

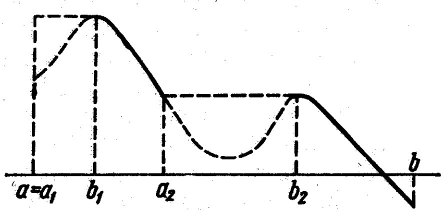


FIG. 21

La demostración del teorema de Lebesgue se basa en el lema que damos más abajo y que será empleado también en lo sucesivo.

Introduzcamos la siguiente definición. Sea $g(x)$ una función continua definida en un segmento $a \leqslant x \leqslant b$. Un punto x_0 de este segmento se llamará *punto invisible por la derecha* para la función g cuando exista un punto ξ, $x_0 < \xi \leqslant b$, tal que $g(x_0) < g(\xi)$ (véase la fig. 21).

LEMA (Riesz). *Cualquiera que sea la función continua g el conjunto de puntos invisibles por la derecha es abierto en el segmento $[a, b]$ y, por consiguiente, puede ser representado como la unión de un número finito o numerable de intervalos (a_k, b_k) disjuntos dos a dos. En los puntos extremos de estos intervalos se cumplen las desigualdades*

$$g(a_k) \leqslant g(b_k). \tag{5}$$

DEMOSTRACION DEL LEMA. Si x_0 es un punto invisible por la derecha para g, la misma propiedad la tendrá, debido a la continuidad de g, cualquier punto suficientemente próximo a x_0. Luego, el conjunto de estos puntos es abierto. Sea (a_k, b_k) uno de los intervalos que lo componen. Supongamos que

$$g(a_k) > g(b_k); \tag{6}$$

entonces, existe en el intervalo (a_k, b_k) un punto interior x_0 en el cual $g(x_0) > g(b_k)$. Sea x^* el punto más a la derecha de aquellos puntos x de (a_k, b_k) en los que $g(x) = g(x_0)$ (véase la fig. 21).

Como $x^* \in (a_k, b_k)$, existe un punto $\xi > x^*$ tal que $g(\xi) > g(x^*)$. El punto ξ no puede pertenecer al intervalo (a_k, b_k), ya que x^* es el punto *más a la derecha* de este intervalo en el que $g(x) = g(x_0)$, mientras que $g(b_k) < g(x_0)$. Por otro lado, la desigualdad $\xi > b_k$ también es imposible, puesto que tendríamos $g(b_k) < g(x_0) < g(\xi)$ *no* siendo b_k un punto invisible por la derecha. La contradicción obtenida prueba que la desigualdad (6) no tiene lugar, es decir, que $g(a_k) \leqslant g(b_k)$ y esto demuestra el lema. El lector podrá probar fácilmente que, de hecho, $g(a_k) = g(b_k)$ siempre que $a_k \neq a$.

Observación. Un punto x_0 se llama *invisible por la izquierda* para una función $g(x)$ cuando existe un $\xi < x_0$ tal que $g(\xi) > g(x_0)$. Los mismos razonamientos permiten establecer que el conjunto de estos puntos es la suma de un número finito o numerable de intervalos (a_k, b_k) disjuntos dos a dos para cada uno de los cuales se cumple la desigualdad

$$g(a_k) \geqslant g(b_k).$$

Pasemos ahora directamente a la demostración del teorema de Lebesgue. Demostrémoslo primero suponiendo que f es una función monótona continua no decreciente.

Para demostrar el teorema basta probar que en casi todos los puntos

$$1)\ \Lambda_d < \infty \quad \text{y} \quad 2)\ \lambda_i \geqslant \Lambda_d.$$

En efecto, si tomamos $f^*(x) = -f(-x)$, f^* será también una función monótona continua no decreciente. Si Λ_d^* y λ_i^* son los números derivados superior derecho e inferior izquierdo para f^*, es fácil probar que

$$\Lambda_d^* = \Lambda_i, \quad \lambda_i^* = \lambda_d.$$

Por esto, aplicando a $f^*(x)$ la desigualdad (2), tendremos

$$\lambda_d \geqslant \Lambda_i. \tag{7}$$

Uniendo en una cadena las desigualdades obtenidas y valiéndonos de la definición de los números derivados, tendremos

$$\Lambda_d \leqslant \lambda_i \leqslant \Lambda_i \leqslant \lambda_d \leqslant \Lambda_d$$

y esto significa que

$$\lambda_i = \lambda_d = \Lambda_i = \Lambda_d.$$

Probemos primero que $\Lambda_d < \infty$ en casi todos los puntos. Si $\Lambda_d = \infty$ en un punto x_0, cualquiera que sea $C = \text{const}$ existirá a la derecha del punto x_0 un punto ξ tal que

$$\frac{f(\xi) - f(x_0)}{\xi - x_0} > C,$$

es decir,

$$f(\xi) - f(x_0) > C(\xi - x_0)$$

o

$$f(\xi) - C\xi > f(x_0) - Cx_0.$$

En otras palabras, el punto x_0 es un punto invisible por la derecha para la función

$$g(x) = f(x) - Cx.$$

En vista del lema de Riesz, el conjunto formado por estos puntos es abierto y en los extremos de los intervalos (a_k, b_k) que lo componen se cumplen las desigualdades

$$f(a_k) - Ca_k \leqslant f(b_k) - Cb_k,$$

es decir,

$$f(b_k) - f(a_k) \geqslant C(b_k - a_k).$$

Dividiendo por C y sumando las desigualdades obtenidas en todos los intervalos (a_k, b_k), encontramos

$$\sum_k (b_k - a_k) \leqslant \sum_k \frac{f(b_k)-f(a_k)}{C} \leqslant \frac{f(b)-f(a)}{C}.$$

Aquí C se puede escoger tan grande como se quiera. De manera que el conjunto de aquellos puntos en los que $\Lambda_d = \infty$ puede ser cubierto por intervalos tales que la suma de sus longitudes sea tan pequeña como se quiera. Consecuentemente, la medida de este conjunto es igual a 0.

El mismo procedimiento, ligado al lema de Riesz, permite demostrar que en casi todos los puntos $\lambda_i \geqslant \Lambda_d$, sólo que este procedimiento debe ser empleado ahora dos veces. Consideremos un par de números c y C tales que $0 < c < C < \infty$ y tomemos $\rho = \frac{c}{C}$. Sea E_ρ el conjunto de aquellos x para los cuales $\Lambda_d > C$ y $\lambda_i < c$. Si logramos demostrar que $\mu E_\rho = 0$, ello implicará que $\lambda_i \geqslant \Lambda_d$ en casi todos los puntos ya que el conjunto de puntos, donde $\lambda_i < \Lambda_d$, puede ser representado, evidentemente, como la unión de una cantidad a lo sumo numerable de conjuntos de tipo E_ρ.

Conviene destacar el lema siguiente.

LEMA. *Para cualquier intervalo* $(\alpha, \beta) \subset [a, b]$ *tenemos*

$$\mu\{x : x \in E_\rho \cap (\alpha, \beta)\} \leqslant \rho(\beta - \alpha).$$

DEMOSTRACION. Consideremos primero el conjunto formado por aquellos $x \in (\alpha, \beta)$ para los cuales $\lambda_i < c$. Cualquiera que sea el punto x, existe un $\xi < x$ tal que $\frac{f(\xi)-f(x)}{\xi - x} < c$, es decir, $f(\xi) - c\xi > f(x) - cx$. Por lo tanto, x es invisible por la izquierda para la función $f(x) - cx$ y, en vista del lema de Riesz (véase la observación en la pág. 370), el conjunto de estos x puede ser representado como la unión, a lo sumo numerable, de intervalos $(\alpha_k, \beta_k) \subset (\alpha, \beta)$ tales que $f(\alpha_k) - c\alpha_k \geqslant f(\beta_k) - c\beta_k$, es decir,

$$f(\beta_k) - f(\alpha_k) \leqslant c(\beta_k - \alpha_k). \tag{8}$$

Consideremos en cada uno de los intervalos (α_k, β_k) el conjunto G_k formado por aquellos x en los que $\Lambda_d > C$. Aplicando una vez más el lema de Riesz (ahora para los puntos invisibles por la derecha, lo mismo que al demostrar la desigualdad $\Lambda_d < \infty$), veremos que G_k se puede representar como la unión, a lo sumo numerable, de intervalos disjuntos dos a dos $(\alpha_{k_j}, \beta_{k_j})$ y que

$$\beta_{k_j} - \alpha_{k_j} \leqslant \frac{1}{C}[f(\beta_{k_j}) - f(\alpha_{k_j})]. \tag{9}$$

Está claro que el sistema de intervalos $(\alpha_{kj}, \beta_{kj})$ cubre el conjunto $E_\rho \cap (\alpha, \beta)$ y que, además, tenemos, de acuerdo con (8) y con (9),

$$\sum_{k,j} (\beta_{kj} - \alpha_{kj}) \leqslant \frac{1}{C} \sum_{k,j} [f(\beta_{kj}) - f(\alpha_{kj})] \leqslant$$
$$\leqslant \frac{1}{C} \sum_k [f(\beta_k) - f(\alpha_k)] \leqslant \frac{c}{C} \sum_k (\beta_k - \alpha_k) \leqslant \rho(\beta - \alpha);$$

esto demuestra el lema.

Ahora es fácil probar que efectivamente $\mu E_\rho = 0$.

Para ello es suficiente emplear sólo aquella propiedad del conjunto E_ρ que ha sido demostrada en el último lema. Sea $\mu E_\rho = t$. Cualquiera que sea $\varepsilon > 0$, existe un conjunto abierto G, igual a la unión numerable de intervalos (a_m, b_m), tal que $E_\rho \subset G$ y $\sum\limits_m (b_m - a_m) < t + \varepsilon$. Tomemos $t_m = \mu [E_\rho \cap (a_m, b_m)]$. Es obvio que $t = \sum\limits_m t_m$. Por el lema, $t_m \leqslant \rho(b_m - a_m)$. Luego, $t \leqslant \rho \sum\limits_m \times$ $\times (b_m - a_m) < \rho(t + \varepsilon)$ y como $\varepsilon > 0$ es arbitrario, tenemos $t \leqslant \rho t$. Pero, $0 < \rho < 1$; por lo tanto $t = 0$.

De manera que hemos demostrado el teorema 1 para el caso en que f es una función continua. Los mismos razonamientos sirven para el caso de una función monótona discontinua, si se recurre a la siguiente generalización del lema de Riesz al caso de funciones con discontinuidades de primera especie solamente.

Sea g una función en un segmento $[a, b]$ que tiene solamente discontinuidades de primera especie. Diremos que un punto $x_0 \in [a, b]$ *es invisible por la derecha* para $g(x)$ cuando exista un $\xi > x_0$ tal que

$$\max [g(x_0 - 0), g(x_0), g(x_0 + 0)] < g(\xi).$$

Entonces, lo mismo que en el caso en que g es continua, el conjunto de puntos invisibles por la derecha para g es abierto y en los extremos de los intervalos (a_k, b_k) que lo componen se cumplen las desigualdades

$$g(a_k) \leqslant g(b_k).$$

Aunque la demostración del teorema 1 es extensa, tiene una interpretación simple y obvia. Expliquemos, por ejemplo, por qué Λ_d (y Λ_i) deben ser finitos en casi todos los puntos. El cociente $\frac{\Delta f}{\Delta x}$ es el «coeficiente de dilatación» del segmento $[a, b]$ en el punto dado x bajo la aplicación f. Como esta aplicación transforma el segmento finito $[a, b]$ en un segmento finito $[f(a), f(b)]$, la «dilatación» no puede ser infinita en un conjunto de medida positiva. Tampoco resulta difícil interpretar el último razonamiento que se basa en el lema demostrado en la pág. 372. Significa simplemente que si un subconjunto medible A en cualquier intervalo (α, β) es parte de este intervalo de medida no superior a $\rho(\beta-\alpha)$, donde $\rho < 1$ está fijado, entonces A no puede ser de medida positiva.

3°. Derivada de la integral respecto al extremo superior. Como la integral

$$\int_a^x \varphi(t)\,dt$$

de cualquier función sumable puede ser representada como diferencia de dos funciones monótonas, del teorema 1 se obtiene inmediatamente el resultado siguiente.

TEOREMA 2. *Cualquiera que sea la función sumable* φ, *la derivada*

$$\frac{d}{dx}\int_a^x \varphi(t)\,dt \tag{10}$$

existe para casi todos los puntos x.

Es preciso subrayar que, aunque hemos demostrado la existencia de la derivada (10) en casi todos los puntos, el problema sobre la igualdad

$$\frac{d}{dx}\int_a^x \varphi(t)\,dt = \varphi(x)$$

no se ha discutido aún. Resulta (véase el § 3) que esta igualdad es válida en casi todos los puntos cualquiera que sea la función sumable φ.

§ 2. FUNCJONES DE VARIACIÓN ACOTADA

El problema sobre la derivada de la integral de Lebesgue respecto al extremo superior nos ha llevado a considerar la clase de funciones que pueden ser representadas como diferencia de funciones monótonas. En este parágrafo daremos una descripción distinta de estas funciones, que no se basa en el concepto de monotonía, y estudiaremos sus propiedades principales. Comencemos por las definiciones necesarias.

DEFINICION 1. Una función f definida en un segmento se llama *función de variación acotada* cuando existe una constante C tal que, cualquiera que sea la partición del segmento $[a, b]$ por puntos

$$a=x_0<x_1<\ldots<x_n=b,$$

se cumple la desigualdad

$$\sum_{k=1}^{n}|f(x_k)-f(x_{k-1})|\leqslant C. \tag{1}$$

Toda función monótona es de variación acotada, ya que para ella la suma que figura en el miembro izquierdo de (1) no depende de la partición y es siempre igual a $|f(b)-f(a)|$.

DEFINICION 2. Sea f una función de variación acotada. La cota superior mínima de las sumas (1) correspondientes a todas las particiones finitas del segmento $[a, b]$ se denomina *variación total* de la función f en el segmento $[a, b]$ y se designa mediante $V_a^b[f]$. De manera que

$$V_a^b[f]=\sup\sum_{k=1}^{n}|f(x_k)-f(x_{k-1})|.$$

Observación. Una función f definida en toda la recta se llama función de variación acotada cuando las magnitudes

$$V_a^b[f]$$

están acotadas en su conjunto. En este caso,

$$\lim_{\substack{b\to\infty\\ a\to-\infty}} V_a^b[f]$$

se llama variación total de la función f en la recta $-\infty<x<\infty$ y se designa con $V_{-\infty}^{\infty}[f]$.

Veamos las propiedades fundamentales de la variación total de una función.

1. *Si α es un número constante, se tiene*

$$V_a^b[\alpha f]=|\alpha|V_a^b[f].$$

Esto se desprende inmediatamente de la definición de $V_a^b[f]$.

2. *Si f y g son funciones de variación acotada, también $f+g$ es de variación acotada y*

$$V_a^b[f+g]\leqslant V_a^b[f]+V_a^b[g]. \tag{2}$$

En efecto, cualquiera que sea la partición del segmento $[a, b]$, tenemos

$$\sum_k|f(x_k)+g(x_k)-f(x_{k-1})-g(x_{k-1})|\leqslant$$
$$\leqslant\sum_k|f(x_k)-f(x_{k-1})|+\sum_k|g(x_k)-g(x_{k-1})|$$

y, como siempre

$$\sup (A+B) \leqslant \sup A + \sup B,$$

obtenemos de aquí la desigualdad necesaria.

Las propiedades 1 y 2 significan que una combinación lineal de funciones de variación acotada (definidas en un segmento dado $[a, b]$) es de nuevo una función de variación acotada. En otras palabras, las funciones de variación acotada constituyen un espacio lineal (a diferencia del conjunto de funciones monótonas que no forma un espacio lineal).

3. *Si* $a < b < c$, *se tiene*

$$V_a^b[f] + V_b^c[f] = V_a^c[f]. \tag{3}$$

En efecto, consideremos primero una partición del segmento $[a, c]$ tal que b es uno de los puntos de la partición, digamos $x_r = b$. En este caso,

$$\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})| = \sum_{k=1}^{r} |f(x_k) - f(x_{k-1})| +$$
$$+ \sum_{k=r+1}^{n} |f(x_k) - f(x_{k-1})| \leqslant V_a^b[f] + V_b^c[f]. \tag{4}$$

Consideremos ahora una partición arbitraria del segmento $[a, c]$. Está claro que, si agregamos a los puntos de partición uno más, a saber, el punto b, la suma

$$\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})|$$

en todo caso no disminuirá. Consecuentemente, la desigualdad (4) se cumple para *cualquier* partición del segmento $[a, c]$, es decir,

$$V_a^c[f] \leqslant V_a^b[f] + V_b^c[f].$$

Por otro lado, cualquiera que sea $\varepsilon > 0$, existen unas particiones de los segmentos $[a, b]$ y $[b, c]$ tales que

$$\sum_i |f(x_i') - f(x_{i-1}')| > V_a^b[f] - \frac{\varepsilon}{2}$$

y

$$\sum_j |f(x_j'') - f(x_{j-1}'')| > V_b^c[f] - \frac{\varepsilon}{2}.$$

Uniendo estas dos particiones, encontraremos una partición del segmento $[a, c]$ tal que

$$\sum_k |f(x_k)-f(x_{k-1})| = \sum_i |f(x_i')-f(x_{i-1}')| + \\ + \sum_j |f(x_j'')-f(x_{j-1}'')| > V_a^b[f]+V_b^c[f]-\varepsilon.$$

Como $\varepsilon > 0$ es arbitrario, de aquí se desprende que

$$V_a^c[f] \geqslant V_a^b[f]+V_b^c[f]. \tag{5}$$

De (4) y (5) se deduce (3).

Como la variación de cualquier función en un segmento cualquiera es no negativa, obtenemos inmediatamente de la propiedad 3 que:

4. *La función*

$$v(x)=V_a^x[f]$$

es monótona no decreciente.

5. *Si f es continua a la izquierda en el punto x^*, también v es continua a la izquierda en este punto.*

En efecto, sea dado $\varepsilon > 0$. Escojamos $\delta > 0$ de manera que $|f(x^*)-f(x)| < \frac{\varepsilon}{2}$ siempre que $|x^*-x| < \delta$. Escojamos, además, una partición

$$a=x_0 < x_1 < \ldots < x_n = x^*$$

tal que

$$V_a^{x^*}[f]-\sum_{k=1}^{n} |f(x_k)-f(x_{k-1})| < \frac{\varepsilon}{2}. \tag{6}$$

Podemos suponer que

$$|x^*-x_{n-1}| < \delta$$

(de lo contrario, podríamos agregar otro punto de partición ya que con esto la diferencia que figura en el miembro izquierdo de (6) sólo podría disminuir); entonces,

$$|f(x^*)-f(x_{n-1})| < \frac{\varepsilon}{2}$$

y, por consiguiente,

$$V_a^{x^*}[f]-\sum_{k=1}^{n-1} |f(x_k)-f(x_{k-1})| < \varepsilon.$$

Pero, entonces, con mayor razón

$$V_a^{x^*}[f]-V_a^{x_{n-1}}[f] < \varepsilon, \text{ es decir, } v(x^*)-v(x_{n-1}) < \varepsilon.$$

Como v es una función monótona no decreciente, de aquí se deduce que $v(x^*)-v(x)<\varepsilon$ para todo x tal que $x_{n-1}\leqslant x\leqslant x^*$. Esto significa precisamente que la función v es continua a la izquierda en el punto x^*.

Razonamientos análogos demuestran que si f es continua a la derecha en el punto x^*, también v es continua a la derecha en este punto. Por consiguiente, si f es continua en un punto (o en todo el segmento $[a, b]$), también v lo es.

Sea f una función arbitraria de variación acotada en $[a, b]$ y sea v su variación total en $[a, x]$. Consideremos la diferencia

$$\varphi=v-f.$$

Esta diferencia es una función monótona no decreciente. En efecto, sea $x'\leqslant x''$. Entonces,

$$\varphi(x'')-\varphi(x')=[v(x'')-v(x')]-[f(x'')-f(x')]. \qquad (7)$$

Pero, siempre

$$|f(x'')-f(x')|\leqslant v(x'')-v(x'),$$

de manera que el miembro derecho y, consecuentemente, también el miembro izquierdo de la igualdad (7) son no negativos.

Como

$$f=v-\varphi$$

hemos obtenido de esta forma el resultado siguiente.

TEOREMA 1. *Toda función de variación acotada puede ser representada como la diferencia de dos funciones monótonas no decrecientes.*

Es decir, el conjunto de funciones que pueden ser representadas como diferencia de funciones monótonas y que ha sido considerado en el parágrafo anterior, es precisamente el conjunto de funciones de variación acotada.

Del teorema 1 y del teorema de Lebesgue sobre la existencia de la derivada de una función monótona, demostrado en el parágrafo anterior, se desprende inmediatamente que *toda función de variación acotada posee derivada finita en casi todos los puntos.*

EJERCICIOS. 1. Si f tiene derivada acotada en $[a, b]$ (es decir, $f'(x)$ existe en todo punto y $|f'(x)|<C$), la función f es de variación acotada y

$$V_a^b[f]\leqslant C(b-a).$$

2. Sea $f(x)=x \operatorname{sen}\frac{1}{x}$. Demuéstrese que f no es una función de variación acotada.

Las constantes, y sólo ellas, son las funciones cuya variación total es igual a 0. Tomemos

$$\|f\| = V_a^b[f].$$

La magnitud $V_a^b[f]$ posee las propiedades 2) y 3) de la norma (véase la pág. 149), pero no verifica la propiedad 1). Si consideramos solamente las funciones sujetas a la condición adicional $f(a)=0$, ellas también formarán un espacio lineal en el que la magnitud $V_a^b[f]$ tendrá ya todas las propiedades de la norma. El espacio $V_{[a,b]}^0$ de funciones de variación acotada en $[a, b]$ que verifican la condición $f(a)=0$, con las habituales operaciones de adición y multiplicación por números y con la norma

$$\|f\| = V_a^b[f]$$

se llama espacio de funciones de variación acotada.

§ 3. DERIVADA DE LA INTEGRAL INDEFINIDA DE LEBESGUE

En el § 1 hemos demostrado que la integral de Lebesgue

$$\int_a^x f(t)\,dt$$

tiene, como función de x, derivada finita en casi todos los puntos. Sin embargo, no hemos revelado aún cómo está relacionada esta derivada con el integrando. Ahora demostraremos el resultado que hemos mencionado al final del § 1.

TEOREMA 1. *Cualquiera que sea la función sumable f, en casi todos los puntos se cumple la igualdad*

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x).$$

DEMOSTRACION. Tomemos

$$\Phi(x) = \int_a^x f(t)\,dt.$$

Probemos primero que en casi todos los puntos

$$f(x) \geqslant \Phi'(x).$$

Si $f(x) < \Phi'(x)$, existirán números racionales α y β tales que

$$f(x) < \alpha < \beta < \Phi'(x). \tag{1}$$

Sea $E_{\alpha\beta}$ el conjunto formado por los puntos en los que se verifica la desigualdad (1). Demostremos que la medida de cada uno de estos conjuntos $E_{\alpha\beta}$ es igual a cero. Como el número de estos conjuntos es numerable, de aquí se desprenderá que

$$\mu\{x : f(x) < \Phi'(x)\} = 0.$$

Sea $\varepsilon > 0$ arbitrario y sea $\delta > 0$ tal que

$$\left| \int_e f(t)\,dt \right| < \varepsilon$$

siempre que $\mu(e) < \delta$ (un tal δ existe cualquiera que sea ε debido a la continuidad absoluta de la integral). Escojamos ahora un conjunto abierto $G \subset [a, b]$ de manera que

$$G \supset E_{\alpha\beta} \text{ y } \mu(G) < \mu(E_{\alpha\beta}) + \delta.$$

Si $x \in E_{\alpha\beta}$, tenemos

$$\frac{\Phi(\xi) - \Phi(x)}{\xi - x} > \beta \qquad (2)$$

para todos los $\xi > x$ suficientemente próximos a x. Escribiendo la desigualdad (2) en la forma

$$\Phi(\xi) - \beta\xi > \Phi(x) - \beta x,$$

obtenemos que el punto x es un punto invisible por la derecha para la función $\Phi(x) - \beta x$ en cualquiera de los intervalos que componen el conjunto G. Valiéndonos del lema de Riesz, podemos por eso indicar un conjunto $S = \bigcup_k (a_k, b_k)$ tal que $E_{\alpha\beta} \subset S \subset G$ y

$$\Phi(b_k) - \beta b_k \geqslant \Phi(a_k) - \beta a_k,$$

es decir,

$$\Phi(b_k) - \Phi(a_k) \geqslant \beta(b_k - a_k)$$

o

$$\int_{a_k}^{b_k} f(t)\,dt \geqslant \beta(b_k - a_k).$$

Sumando estas desigualdades correspondientes a todos los intervalos (a_k, b_k) que componen S, obtenemos

$$\int_S f(t)\,dt \geqslant \beta\mu(S). \qquad (3)$$

Al mismo tiempo,

$$\int_S f(t)\,dt = \int_{E_{\alpha\beta}} f(t)\,dt + \int_{S \setminus E_{\alpha\beta}} f(t)\,dt <$$

$$< \alpha\mu(E_{\alpha\beta}) + \varepsilon \leqslant \alpha\mu(S) + \varepsilon + |\alpha|\delta. \qquad (4)$$

Comparando (3) y (4), encontramos

$$\alpha\mu(S)+\varepsilon+|\alpha|\delta\geqslant\beta\mu(S),$$

es decir,

$$\mu(S)\leqslant\frac{\varepsilon+|\alpha|\delta}{\beta-\alpha}.$$

De modo que el conjunto $E_{\alpha\beta}$ se puede sumergir en un conjunto abierto de medida tan pequeña como se quiera (podemos admitir que $|\alpha|\delta\leqslant\varepsilon$) y esto significa precisamente que $\mu(E_{\alpha\beta})=0$. Hemos demostrado, pues, que

$$f(x)\geqslant\Phi'(x)$$

en casi todos los puntos. Sustituyendo $f(x)$ por $-f(x)$, encontraremos de la misma forma que en casi todos los puntos

$$-f(x)\geqslant-\Phi'(x),$$

es decir,

$$f(x)\leqslant\Phi'(x)$$

y, consecuentemente, en casi todos los puntos

$$f(x)=\Phi'(x)=\frac{d}{dx}\int_a^x f(t)\,dt.$$

El teorema queda demostrado.

§ 4. RECONSTRUCCIÓN DE UNA FUNCIÓN A PARTIR DE SU DERIVADA. FUNCIONES ABSOLUTAMENTE CONTINUAS

Hemos resuelto, pues, el primero de los problemas planteados en la introducción a este capítulo demostrando que para una función f sumable en $[a, b]$

$$\frac{d}{dx}\int_a^x f(t)\,dt=f(x)$$

en casi todos los puntos. Consideremos ahora el segundo de los problemas antes planteados, es decir, estudiemos cómo se generaliza al caso de la integral de Lebesgue la fórmula de Newton — Leibniz

$$F(x)=F(a)+\int_a^x F'(t)\,dt \tag{1}$$

que es bien conocida, en el caso de funciones de derivada continua, del Análisis elemental.

Es natural limitarse desde el principio a considerar aquellas funciones F que son de antemano diferenciables en casi todos los puntos (de lo contrario, la igualdad (1) no tiene simplemente sentido). Como sabemos ya, son de este tipo las funciones de variación acotada.

Por otro lado, la integral que figura en el miembro derecho de (1) es una función de variación acotada. Luego, la igualdad (1) no puede ser válida para una clase más amplia de funciones. Puesto que toda función de variación acotada es diferencia de dos monótonas no decrecientes, son precisamente las funciones monótonas las que deben ser consideradas en primer término.

Sin embargo, para funciones monótonas arbitrarias la igualdad (1) no tiene lugar, en general. Pero, es válida la afirmación siguiente.

TEOREMA 1. *La derivada f' de una función monótona no decreciente f es sumable y*

$$\int_a^b f'(x)\,dx \leqslant f(b)-f(a).$$

DEMOSTRACION. Por definición, la derivada de una función f en un punto x es el límite del cociente[1]

$$\varphi_h(x)=\frac{f(x+h)-f(x)}{h} \qquad (2)$$

para $h\to 0$. La sumabilidad de f implica que cada una de las funciones φ_h sea también sumable y, por lo tanto, la igualdad (2) puede ser integrada. Esto nos da

$$\int_a^b \varphi_h(x)\,dx=\frac{1}{h}\int_a^b f(x+h)\,dx-\frac{1}{h}\int_a^b f(x)\,dx=$$

$$=-\frac{1}{h}\int_b^{b+h} f(x)\,dx-\frac{1}{h}\int_a^{a+h} f(x)\,dx.$$

La expresión que figura en el miembro derecho tiende para $h\to+0$ hacia $f(b)-f(a+0)$. Luego, aplicando el teorema de Fatou,

[1] Para que la expresión $f(x+h)$ tenga sentido cualquiera que sea $x\in[a, b]$, podemos aceptar que $f(x)=f(b)$ para $x>b$ y $f(x)=f(a)$ para $x<a$.

encontramos

$$\int_a^b f'(x)\,dx \leqslant \lim_{h\to 0} \int_a^b \varphi_h(x)\,dx = f(b) - f(a+0) \leqslant f(b) - f(a)$$

(el teorema de Fatou garantiza asimismo la existencia de la integral de f'). Hemos demostrado el teorema.

Es fácil dar ejemplos de funciones monótonas para las cuales tiene lugar la desigualdad estricta

$$\int_a^b f'(x)\,dx < f(b) - f(a).$$

Es suficiente, por ejemplo, tomar

$$f(x) = \begin{cases} 0 \text{ para } 0 \leqslant x < \frac{1}{2}, \\ 1 \text{ para } \frac{1}{2} \leqslant x \leqslant 1. \end{cases}$$

Es de mayor interés, sin embargo, el hecho de que existen funciones monótonas continuas para las cuales se cumple la desigualdad estricta $\int_a^x f'(t)\,dt < f(x) - f(a)$ para todos los $x > a$. He aquí uno de los ejemplos más sencillos. Consideremos en el segmento $[0, 1]$ el conjunto perfecto de Cantor y definamos primero f en los intervalos contiguos tomando

$$f(t) = \frac{k}{2^n}, \quad k = 1, 3, 5, \ldots, 2n-1$$

para el k-ésimo intervalo contiguo, contando de izquierda a la derecha, de n-ésimo rango (incluyendo sus extremos). Es decir,

$$f(t) = \frac{1}{2} \text{ para } \frac{1}{3} \leqslant t \leqslant \frac{2}{3}, \; f(t) = \frac{1}{4} \text{ para } \frac{1}{9} \leqslant t \leqslant \frac{2}{9},$$

$$f(t) = \frac{3}{4} \text{ para } \frac{7}{9} \leqslant t \leqslant \frac{8}{9},$$

etc. (fig. 22). De esta forma f está definida en todo el segmento $[0, 1]$, excepto los puntos de segunda especie del conjunto de Cantor (es decir, los puntos que no pertenecen ni a los intervalos contiguos ni al conjunto de sus extremos). Completemos la definición de f en los puntos restantes del siguiente modo. Sea t^* uno de estos puntos y sea $\{t_n\}$ una sucesión creciente, convergente hacia este punto, de puntos de primera especie (esto es, de extremos de los intervalos contiguos). Entonces, existe el límite

$$\lim_{n\to\infty} f(t_n); \tag{3}$$

análogamente existe también el límite

$$\lim_{n\to\infty} f(t'_n) \tag{4}$$

donde $\{t'_n\}$ es una sucesión decreciente de puntos de primera especie que converge hacia t^*; además, los límites (3) y (4)

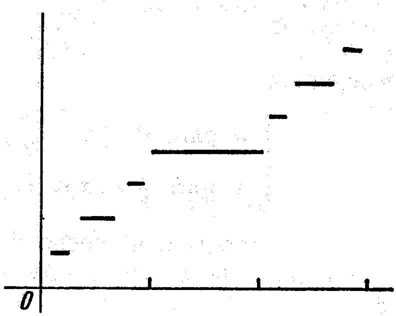


FIG. 22

coinciden. Tomando este valor común igual a $f(t^*)$, obtendremos una función monótona definida y continua en todo el segmento [0, 1]. La derivada de esta función, llamada «escalera de Cantor», es igual, evidentemente, a cero en cada punto de cualquier intervalo contiguo, esto es, en casi todos los puntos. Consecuentemente, tenemos para esta función

$$0=\int_0^1 f(t)\,dt < f(1)-f(0)=1.$$

Para poder describir la clase de funciones para las cuales tiene lugar la igualdad

$$\int_a^b f'(x)\,dx = f(b)-f(a)$$

introduciremos la definición siguiente.

DEFINICION 1. Una función f definida en un segmento $[a, b]$ se llama *absolutamente continua* en este segmento cuando para cualquier $\varepsilon > 0$ existe un $\delta > 0$ tal que, cualquiera que sea el sistema finito de intervalos disjuntos dos a dos

$$(a_k, b_k), \quad k = 1, 2, \ldots, n$$

pertenecientes a $[a, b]$ y tal que

$$\sum_{k=1}^{n} (b_k - a_k) < \delta,$$

se cumple la desigualdad

$$\sum_{k=1}^{n} |f(b_k) - f(a_k)| < \varepsilon.$$

Está claro que toda función absolutamente continua es uniformemente continua. Lo recíproco, en general, no tiene lugar: por ejemplo, la «escalera de Cantor» descrita más arriba es continua (y, por consiguiente, uniformemente continua) en el segmento $[0, 1]$ y, sin embargo, no es absolutamente continua. En efecto, el conjunto de Cantor puede ser cubierto por un sistema finito de intervalos (a_k, b_k) cuya suma de longitudes es tan pequeña como se quiera. Al mismo tiempo, para cada uno de estos sistemas de intervalos se cumple, evidentemente, la igualdad

$$\sum_{k=1}^{n} | f(b_k) - f(a_k) | = 1.$$

Indiquemos las propiedades fundamentales de las funciones absolutamente continuas.

1. Observemos ante todo que en la definición se puede tomar, en lugar de cualquier sistema finito de intervalos de longitud total $< \delta$, cualquier sistema finito o numerable de intervalos de longitud total $< \delta$. En efecto, supongamos que para un $\varepsilon > 0$ dado hemos escogido $\delta > 0$ de manera que

$$\sum_{k=1}^{n} | f(b_k) - f(a_k) | < \varepsilon$$

para cualquier sistema finito de intervalos (a_k, b_k) que verifica la condición

$$\sum_{k=1}^{n} (b_k - a_k) < \delta$$

y sea (α_k, β_k) un sistema numerable de intervalos de longitud total no mayor que δ. Entonces, para cualquier n tenemos

$$\sum_{k=1}^{n} |f(\beta_k - f(\alpha_k)| < \varepsilon;$$

pasando aquí al límite para $n \to \infty$, obtenemos que

$$\sum_{k=1}^{\infty} |f(\beta_k) - f(\alpha_k)| \leqslant \varepsilon.$$

2. *Toda función absolutamente continua es de variación acotada.* En efecto, la continuidad absoluta de una función f en un segmento $[a, b]$ significa que para cualquier $\varepsilon > 0$ se puede escoger $\delta > 0$ de manera que la variación total de la función f en un segmento de longitud $< \delta$ no será mayor que ε. Puesto que el segmento $[a, b]$ puede ser dividido en un número finito de segmentos de longitud $< \delta$, la variación total de la función f en $[a, b]$ también será finita.

3. *La suma, la diferencia y el producto por un número de funciones absolutamente continuas son funciones absolutamente continuas.*

Esto se desprende inmediatamente de la definición de continuidad absoluta y de las propiedades del valor absoluto de suma, diferencia y producto.

Las propiedades 2 y 3 muestran que las funciones absolutamente continuas constituyen una variedad lineal en el espacio de todas las funciones de variación acotada.

4. *Toda función absolutamente continua puede ser representada como diferencia de dos funciones absolutamente continuas no decrecientes.*

En efecto, una función absolutamente continua, como toda función de variación acotada, puede ser representada en la forma

$$f = v - g,$$

donde

$$v(x) = V_a^x[f] \text{ y } g(x) = v(x) - f(x)$$

son funciones no decrecientes. Probemos que cada una de estas funciones es absolutamente continua. Demostremos esto para v. Sea dado un $\varepsilon > 0$; escojamos $\varepsilon > 0$ a partir de este ε de acuerdo con la continuidad absoluta de la función f. Tomemos un sistema de intervalos (a_k, b_k) de longitud total menor que δ y consideremos la suma

$$\sum_{k=1}^{n} [v(b_k) - v(a_k)]. \tag{5}$$

Esta suma es la cota superior mínima de los números

$$\sum_{k=1}^{n}\sum_{l=1}^{m_k}|f(x_{k,l+1})-f(x_{k,l})| \tag{6}$$

correspondientes a todas las particiones finitas posibles

$$\begin{array}{l} a_1 < x_{1,1} < x_{1,2} < \ldots < x_{1,m_1} = b_1, \\ a_2 < x_{2,1} < x_{2,2} < \ldots < x_{2,m_2} = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_n < x_{n,1} < x_{n,2} < \ldots < x_{n,m_n} = b_n \end{array}$$

de los intervalos $(a_1, b_1), \ldots, (a_n, b_n)$. Como la suma de longitudes de todos los intervalos $(x_{k,l}, x_{k,l+1})$ correspondientes a la suma (6), no pasa de $\delta > 0$, cada una de las sumas (6) es no mayor que ε. Consecuentemente, la suma (5), que es la cota superior mínima de estas sumas, tampoco pasa de ε.

Los dos teoremas que vienen a continuación muestran la relación estrecha que existe entre los conceptos de la continuidad absoluta y de la integral indefinida de Lebesgue.

TEOREMA 2. *La función*

$$F(x) = \int_a^x f(t)\,dt,$$

que representa la integral indefinida de una función sumable, es absolutamente continua.

DEMOSTRACION. Si $\{(a_k, b_k)\}$ es un sistema de intervalos disjuntos dos a dos, tenemos

$$\sum_{k=1}^{n}|F(b_k)-F(a_k)| = \sum_{k=1}^{n}\left|\int_{a_k}^{b_k} f(t)\,dt\right| \leqslant$$

$$\leqslant \sum_{k=1}^{n}\int_{a_k}^{b_k}|f(t)|\,dt = \int_{\bigcup_k (a_k,\, b_k)} |f(t)|\,dt;$$

debido a la continuidad absoluta de la integral de Lebesgue, la última expresión tiende a 0 cuando la longitud total de los intervalos (a_k, b_k) tiende a cero.

TEOREMA 3 (Lebesgue). *La derivada $f = F'$ de una función absolutamente continua, definida en un segmento $[a, b]$, es sumable en este segmento y para todo $x\,(a \leqslant x \leqslant b)$*

$$\int_a^x f(t)\,dt = F(x) - F(a).$$

Los teoremas 2 y 3 indican que las funciones absolutamente continuas, y sólo ellas, pueden ser reconstruidas (salvo un sumando constante) a partir de su derivada mediante la operación de integración.

Para demostrar el teorema 3 nos hará falta el lema siguiente.

LEMA. *Si la derivada de una función absolutamente continua monótona no decreciente f es igual a 0 en casi todos los puntos, esta función es una constante.*

DEMOSTRACION DEL LEMA. Como f es una función monótona continua, su campo de valores es el segmento $[f(a), f(b)]$. Probemos que la longitud de este segmento es igual a cero cuando $f'(x)=0$ en casi todos los puntos. Con esto quedará demostrado el lema. Dividamos el conjunto de puntos del segmento $[a, b]$ en dos clases: el conjunto E de aquellos puntos en los que $f'(x)=0$ y el conjunto Z, complemento de E. Por hipótesis, $\mu(Z)=0$. Tomando un $\varepsilon>0$, busquemos aquel $\delta>0$ que corresponde a este ε en virtud de la continuidad absoluta de la función f e incluyamos Z en un conjunto abierto, cuya medida es inferior a δ (esto es posible ya que $\mu(Z)=0$). En otras palabras, cubramos Z por un sistema finito o numerable de intervalos (a_k, b_k) de longitud total menor que δ. De acuerdo con la selección de δ, tenemos

$$\sum_k |f(b_k)-f(a_k)|<\varepsilon.$$

Luego, todo el sistema de intervalos (a_k, b_k) (y, con mayor razón, el conjunto Z perteneciente a la unión de ellos) es transformado por la función en un conjunto de medida inferior a ε. Es decir, $\mu(f(Z))=0$.

Consideremos ahora el conjunto $E=[a, b]\setminus Z$. Sea $x_0\in E$. Entonces, como $f'(x_0)=0$, para todos los x suficientemente próximos a x_0 se cumple la desigualdad

$$\frac{f(x)-f(x_0)}{x-x_0}<\varepsilon,$$

es decir, para concretar aceptamos que $x>x_0$

$$f(x)-f(x_0)<\varepsilon(x-x_0)$$

o bien

$$\varepsilon x_0-f(x_0)<\varepsilon x-f(x);$$

luego, x_0 es un punto invisible por la derecha para la función $g(x)=\varepsilon x-f(x)$. Entonces, de acuerdo con el lema de Riesz, el conjunto E está contenido en un sistema finito o numerable de intervalos (α_k, β_k), en cuyos extremos se cumplen las condiciones

$$\varepsilon\beta_k - f(\beta_k) \geqslant \varepsilon\alpha_k - f(\alpha_k),$$

es decir,

$$f(\beta_k) - f(\alpha_k) \leqslant \varepsilon(\beta_k - \alpha_k),$$

de donde

$$\sum_k (f(\beta_k) - f(\alpha_k)) \leqslant \varepsilon \sum_k (\beta_k - \alpha_k) \leqslant \varepsilon(b-a).$$

En otras palabras, la función f transforma el conjunto E en un conjunto que puede ser cubierto por un sistema de intervalos de longitud sumaria menor que $\varepsilon(b-a)$. Debido a la arbitrariedad de ε, de aquí se desprende que $\mu(f(E)) = 0$.

Luego, tanto $f(E)$ como $f(Z)$ son de medida nula. Pero, la unión de estos dos conjuntos es el segmento $[f(a), f(b)]$. Con esto queda demostrado que la longitud de este segmento es cero, es decir, que $f(x) = \text{const}$.

Ahora es fácil ya demostrar el teorema 3. Basta, evidentemente, limitarse al caso en que la función $F(x)$ es no decreciente. Entonces,

$$\Phi(x) = F(x) - \int_a^x f(t)\,dt \tag{7}$$

será también una función monótona no decreciente. En efecto, si $x'' > x'$, tenemos, en virtud de (7),

$$\Phi(x'') - \Phi(x') = F(x'') - F(x') - \int_{x'}^{x''} f(t)\,dt \geqslant 0.$$

Además, Φ es absolutamente continua (como diferencia de dos funciones absolutamente continuas) y $\Phi'(x) = 0$ en casi todos los puntos (en vista del teorema 1). Por lo tanto, Φ es una constante, de acuerdo con el lema. Tomando en (7) $x = a$, encontramos que esta constante es igual a $F(a)$. El teorema queda demostrado.

Anteriormente, al considerar las funciones de variación acotada, hemos probado que toda función f de este tipo puede ser representada como suma de una función de saltos H y de una función continua φ de variación acotada

$$f = H + \varphi.$$

Consideremos ahora una función continua φ, pero no absolutamente continua, de variación acotada y tomemos

$$\psi(x) = \int_a^x \varphi'(t)\,dt.$$

La diferencia

$$\chi = \varphi - \psi$$

es una función continua de variación acotada. Además,

$$\frac{d}{dx}\chi(x) = \varphi'(x) - \frac{d}{dx}\int_a^x \varphi'(t)\,dt = 0$$

(en casi todos los puntos).

Diremos que una función continua de variación acotada es *singular* cuando su derivada es nula en casi todos los puntos. Podemos enunciar entonces el resultado siguiente:

toda función de variación acotada puede ser representada como suma de tres componentes:

$$f = H + \psi + \chi, \tag{8}$$

es decir, de una función de saltos, de una función absolutamente continua y de una función singular.

Es fácil probar que cada uno de los tres sumandos de la descomposición (8) queda determinado unívocamente, salvo una constante, por la función $f(x)$. Si normamos todas las funciones que figuran en la igualdad (8), exigiendo, por ejemplo, que cada una de ellas sea nula en el punto $x = a$, la descomposición (8) será única. Derivando la igualdad (8), encontramos que en casi todos los puntos

$$f'(x) = \psi'(x).$$

Luego, al integrar la derivada de una función de variación acotada, se reconstruye no esta función sino solamente su componente absolutamente continua. Las otras dos componentes (la función de saltos y la singular) «desaparecen sin dejar huella».

Es aleccionador comparar los resultados de este parágrafo con lo que da la teoría de funciones generalizadas. Al igual que en el capítulo IV, entenderemos por función generalizada una funcional lineal continua sobre el espacio K de funciones terminales indefinidamente diferenciables. A cada función localmente sumable f se asigna una funcional que opera en los elementos $\varphi \in K$ de acuerdo con la fórmula $(f, \varphi) = \int_{-\infty}^{\infty} f(x)\varphi(x)\,dx$. La derivada generalizada de esta funcional es la funcional que pone en correspondencia al elemento $\varphi \in K$ el número $(f', \varphi) = \int_{-\infty}^{\infty} f(x)\varphi'(x)\,dx$. Como en la clase de funciones generalizadas la ecuación $y' = 0$

tiene solamente soluciones corrientes (constantes), toda función generalizada se reconstruye a partir de su derivada unívocamente, salvo una constante. En particular, toda función localmente sumable f puede ser reconstruida, salvo una constante, a partir de su *derivada generalizada* f' en casi todos los puntos. Supongamos ahora que la función f tiene derivada en casi todos los puntos, por ejemplo, supongamos que f es una función monótona. Sea $\frac{df}{dx}$ la derivada habitual de la función f. $\Big($Hemos visto ya que $\frac{df}{dx}$ puede ser igual a 0 en casi todos los puntos¡ a pesar de que $f(x) \not\equiv$ const!$\Big)$ La función $\frac{df}{dx}$ es localmente sumable (suponemos que f es monótona) y, consecuentemente, podemos asignar a esta función una funcional (función generalizada) $(f_1, \varphi) = \int_{-\infty}^{\infty} \frac{df}{dx} \varphi(x)\, dx$. El hecho sustancial consiste en que *la función generalizada f_1 no coincide, en el caso general, con la función generalizada f'*. Por ejemplo, si

$$f(x) = \begin{cases} 1 & \text{para } x > 0, \\ 0 & \text{para } x \leqslant 0, \end{cases}$$

tenemos $f_1 = 0$ y $f' = \delta$ (véase el ejemplo 2 de la pág. 222).

El teorema 3 significa precisamente que entre todas las funciones de variación acotada la derivada, comprendida en el sentido habitual, de las funciones absolutamente continuas (¡y sólo de ellas!) coincide con la derivada generalizada de estas funciones.

Aquí tropezamos una vez más con la situación de la cual hemos hablado ya en el § 4 del capítulo IV: para poder efectuar las operaciones principales del Análisis (en este caso se trata de la reconstrucción de una función a partir de su derivada) resulta necesario o bien, manteniéndose en los márgenes de las definiciones clásicas, limitarse a una clase suficientemente estrecha de funciones (la de funciones absolutamente continuas) o bien, al contrario, ampliar sustancialmente el concepto de función (generalizando al mismo tiempo la definición de la derivada).

EJERCICIOS. 1. Demuéstrese que la definición de continuidad absoluta, enunciada anteriormente, equivale a la siguiente: f es absolutamente continua en $[a, b]$ cuando transforma cada subconjunto de medida nula de este segmento en un conjunto de medida nula también.

2. Calcúlese la derivada generalizada de la «escalera de Cantor».

3. Sean f una función de variación acotada, f' su derivada generalizada y f_1 la funcional (función generalizada) determinada por la derivada «habitual» $\frac{df}{dx}$ de la función f. Demuéstrese que:

a) si f es absolutamente continua, entonces $f' = f_1$;
b) si $f' = f_1$, entonces $f(x)$ es equivalente a una función absolutamente continua, esto es, coincide con una función de este tipo en casi todos los puntos. En particular, si $f' = f_1$ y f es continua, es f absolutamente continua.

§ 5. INTEGRAL DE LEBESGUE COMO FUNCIÓN DE CONJUNTO. TEOREMA DE RADON — NIKODYM

1°. Cargas. Descomposición de Hahn y descomposición de Jordan. Los conceptos y resultados, expuestos en los parágrafos anteriores para funciones sobre la recta, son extensibles, en gran medida, a las funciones definidas en un espacio arbitrario provisto de medida.

Sea X un espacio provisto de medida μ y sea f una función en X sumable respecto a μ. En este caso, f será sumable en cada subconjunto medible A del conjunto X y, consecuentemente, la integral

$$\Phi(A) = \int_A f(x)\, d\mu \tag{1}$$

(donde f es una función fijada) representa una función de conjunto definida en la colección γ_μ de todos los subconjuntos medibles del conjunto X; además, esta función es σ-aditiva, esto es, cualquiera que sea la descomposición

$$A = \bigcup_k A_k$$

del conjunto medible A en una unión, finita o numerable, de conjuntos medibles disjuntos dos a dos, se cumple la igualdad

$$\Phi(A) = \sum_k \Phi(A_k).$$

En otras palabras, la función Φ, definida por la igualdad (1), posee todas las propiedades de una medida σ-aditiva, excepto, es posible, la de no negatividad (si f toma valores negativos).

DEFINICION. Una función arbitraria σ-aditiva Φ de conjuntos, definida en un σ-anillo de subconjuntos de un espacio X dado, se llama *medida de signo alterno* o, simplemento, *carga.*

El concepto de carga es una generalización natural del concepto de medida σ-aditiva y, como veremos más abajo, se reduce, en cierto sentido, al concepto de medida (esto es, de carga de signo determinado).

EJERCICIOS. Demuéstrese que para cualquier carga Φ, definida en una σ-álgebra de conjuntos $\mathfrak{S}$, existe una constante c tal que $|\Phi(A)| \leqslant c$ para todos los $A \in \mathfrak{S}$.

Si consideramos una carga eléctrica real distribuida, digamos, en una superficie, esta superficie puede ser dividida en dos partes: la que lleva carga positiva (es decir, tal que cualquier parte suya lleva una carga positiva) y la que lleva carga negativa. El equivalente matemático de este resultado es el teorema 1 que damos a continuación.

Introduzcamos primero la terminología siguiente. Sea Φ una carga definida en una σ-álgebra $\mathfrak{S}$ de subconjuntos del espacio X. Un conjunto E se llama *negativo* respecto a Φ cuando para cualquier $F \in \mathfrak{S}$ el conjunto $E \cap F$ pertenece a $\mathfrak{S}$ y $\Phi(E \cap F) \leqslant 0$; de un modo análogo, E se llama *positivo* cuando $E \cap F \in \mathfrak{S}$ y $\Phi(E \cap F) \geqslant 0$ para todos los $F \in \mathfrak{S}$.

TEOREMA 1. *Si Φ es una carga definida en X, existe un subconjunto medible $A^- \subset X$ tal que A^- es negativo y $B^+ = X \setminus A^-$ es positivo (respecto a Φ).*

DEMOSTRACION. Pongamos

$$a = \inf \Phi(A),$$

donde la cota inferior se toma respecto a todos los conjuntos negativos medibles A. Sea $\{A_n\}$ una sucesión de conjuntos medibles negativos tal que

$$\lim_{n \to \infty} \Phi(A_n) = a.$$

Entonces, $A^- = \bigcup\limits_n A_n$ es un conjunto medible negativo tal que

$$\Phi(A^-) = a.$$

Probemos que A^- es el conjunto deseado, esto es, demostremos que

$$B^+ = X \setminus A^-$$

es positivo. Supongamos lo contrario, es decir, admitamos que B^+ contiene un subconjunto medible C_0 tal que $\Phi(C_0) < 0$. El conjunto C_0 no puede ser negativo, ya que entonces lo agregaríamos a A^- obteniendo así un conjunto negativo $\tilde{A}$ tal que

$$\Phi(\tilde{A}) < a$$

y esto es imposible. Luego, existe un número entero mínimo k para el que se puede encontrar en C_0 un subconjunto C_1 que verifique la condición

$$\Phi(C_1) \geqslant \frac{1}{k_1}.$$

Está claro que $C_1 \neq C_0$. Podemos repetir para el conjunto $C_0 \setminus C_1$ el razonamiento aplicado a C_0; obtendremos un conjunto C_2 que verifica la condición

$$\Phi(C_2) \geqslant \frac{1}{k_2} \quad (k_2 > k_1),$$

etc. Tomemos finalmente

$$F_0 = C_0 \setminus \bigcup_{i=1}^{\infty} C_i.$$

El conjunto F_0 no es vacío ya que $\Phi(C_0) < 0$ y $\Phi(C_i) > 0$ para $i \geqslant 1$. De la construcción se desprende que F_0 es negativo. Por lo tanto, agregándolo a A^-, llegaremos de nuevo a una contradicción con la definición de a. Luego, para todos los $E \subset X \setminus A^-$ tenemos

$$\Phi(E) \geqslant 0,$$

es decir, $X \setminus A^-$ es positivo. El teorema queda demostrado.

La partición del espacio X en la parte negativa A^- y en la positiva B^+ se llama *descomposición de Hahn*.

En general, la descomposición de Hahn no es única; sin embargo, si

$$X = A_1^- \cup B_1^+ \text{ y } X = A_2^- \cup B_2^+$$

son dos descomposiciones de este tipo, entonces, para cualquier $E \in \mathfrak{S}$ se tiene

$$\Phi(E \cap A_1^-) = \Phi(E \cap A_2^-) \text{ y } \Phi(E \cap B_1^+) = \Phi(E \cap B_2^+). \tag{2}$$

En efecto,

$$E \cap (A_1^- \setminus A_2^-) \subset E \cap A_1^- \tag{3}$$

y al mismo tiempo

$$E \cap (A_1^- \setminus A_2^-) \subset E \cap B_2^+; \tag{4}$$

de (3) se desprende que

$$\Phi(E \cap (A_1^- \setminus A_2^-)) \leqslant 0$$

y de (4) que

$$\Phi(E \cap (A_1^- \setminus A_2^-)) \geqslant 0.$$

Luego,

$$\Phi(E \cap (A_1^- \setminus A_2^-)) = 0;$$

análogamente encontramos que

$$\Phi(E \cap (A_2^- \setminus A_1^-)) = 0.$$

De aquí se deduce que

$$\Phi(E \cap A_1^-) = \Phi(E \cap A_2^-).$$

Del mismo modo se demuestra la segunda de las igualdades (2).

Consecuentemente, una carga Φ en $\mathfrak{S}$ determina unívocamente dos funciones no negativas de conjunto, a saber:

$$\Phi^+(E) = \Phi(E \cap B^+) \text{ y } \Phi^-(E) = -\Phi(E \cap A^-),$$

que son llamadas *variación superior* y *variación inferior*, respectivamente, de la carga Φ. Además, es obvio que

1) $\Phi = \Phi^+ - \Phi^-$,

2) Φ^+ y Φ^- representan funciones de conjunto no negativas y σ-aditivas, esto es, son medidas.

También será medida, evidentemente, la función $|\Phi| = \Phi^+ + \Phi^-$; ella se llama *variación total* de la carga Φ y la representación de Φ como diferencia de las variaciones superior e inferior se llama *descomposición de Jordan* de esta carga Φ.

Observación. Hemos considerado aquí cargas finitas, esto es, funciones Φ cuyos valores están acotados tanto superiormente, como inferiormente. Además, Φ^+ y Φ^- son, en este caso, medidas finitas. Lo expuesto puede ser generalizado a cargas acotadas solamente por un lado, esto es, a cargas, para las cuales al menos una de las funciones Φ^+ o Φ^- es una medida finita.

2°. Principales tipos de cargas. Sea μ una medida σ-aditiva definida en un σ-anillo de conjuntos $\mathfrak{S}$ del espacio X que llamaremos medibles. Introduzcamos los conceptos siguientes.

Diremos que una carga Φ definida en conjuntos $E \in \mathfrak{S}$ *está concentrada en un conjunto medible* A_0 cuando $\Phi(E) = 0$ para cada $E \subset X \setminus A_0$.

Una carga Φ se llama *continua* cuando $\Phi(E) = 0$ para cualquier conjunto E compuesto por un solo punto. Una carga Φ se llama *discreta* cuando está concentrada en un conjunto finito o numerable. En otras palabras, el hecho de que una carga sea discreta significa que existe un conjunto finito o numerable de puntos $c_1, c_2, \ldots, c_n, \ldots$ tal que para todo $E \subset X$ se tiene

$$\Phi(E) = \sum_{c_k \in E} \Phi(c_k).$$

Una carga Φ se llama *absolutamente continua* (respecto a la medida dada μ) cuando $\Phi(A) = 0$ para todo A medible tal que $\mu(A) = 0$.

Una carga Φ se llama *singular* (respecto a la medida μ) cuando está concentrada en un conjunto de μ-medida nula. Está claro que una carga absolutamente continua y singular a la vez es nula.

3°. Cargas absolutamente continuas. Teorema de Radon—Nikodym. Como ejemplo de carga absolutamente continua respecto a la medida dada μ puede servir la integral de Lebesgue

$$\Phi(A)=\int_A f(x)\,d\mu$$

de una función sumable fija f considerada como función de conjunto. Resulta que con esto se agotan todas las cargas absolutamente continuas. En otras palabras, tiene lugar el teorema siguiente.

TEOREMA 2. (Radon—Nikodym). *Sea μ una medida σ-aditiva definida en una σ-álgebra de subconjuntos de X y sea Φ una carga definida en conjuntos μ-medibles. Entonces, existe en X una función f sumable respecto a μ tal que*

$$\Phi(A)=\int_A f(x)\,d\mu$$

para cada A medible. Esta función, llamada derivada de la carga Φ respecto a la medida μ, se determina unívocamente, salvo una equivalencia.

DEMOSTRACION. Toda carga puede ser representada como diferencia de dos cargas no negativas (véase el punto 2), con la particularidad de que una carga absolutamente continua puede ser representada como diferencia de cargas absolutamente continuas. Por lo tanto, basta demostrar el teorema para el caso de cargas no negativas, esto es, para medidas. Sea, pues, Φ una medida absolutamente continua respecto a la medida dada μ. Demostremos el lema siguiente.

LEMA. *Sea Φ una medida absolutamente continua respecto a μ y distinta de cero idéntico. Entonces existen un n y un conjunto medible B tales que $\mu(B)>0$ y B es positivo respecto a la carga $\Phi-\frac{1}{n}\mu$.*

DEMOSTRACION DEL LEMA. Sea $X=A_n^-\cup B_n^+$ la descomposición de Hahn correspondiente a la carga $\Phi-\frac{1}{n}\mu$, $n=1, 2, \ldots$, y sea

$$A_0=\bigcap A_n^-, \quad B_0=\bigcup B_n^+.$$

Entonces,

$$\Phi(A_0)\leqslant\frac{1}{n}\mu(A_0) \text{ para todo } n,$$

es decir, $\Phi(A_0)=0$ y, consecuentemente, $\Phi(B_0)>0$, de manera que también $\mu(B_0) > 0$ (debido a la continuidad absoluta de Φ respecto a μ). Luego, existe un n tal que $\mu(B_n^+) > 0$. Este n y el conjunto $B = B_n^+$ verifican las condiciones del lema.

Pasemos ahora a la demostración directa del teorema. Sea K un conjunto de funciones f en X que verifican las condiciones siguientes: f son no negativas, integrables respecto a μ y $\int\limits_A f(x)\,d\mu \leqslant \Phi(A)$ para todo A medible. Sea

$$M = \sup\left\{\int\limits_X f(x)\,d\mu \text{ respecto a todos los } f \in K\right\}.$$

Tomemos en K una sucesión $\{f_n\}$ de funciones tal que

$$\lim_{n\to\infty} \int\limits_X f_n(x)\,d\mu = M.$$

Pongamos ahora

$$g_n(x) = \max(f_1(x),\ f_2(x),\ \ldots,\ f_n(x)).$$

Probemos que para todo E medible es

$$\int\limits_E g_n(x)\,d\mu \leqslant \Phi(E).$$

En efecto, podemos representar E en la forma $\bigcup\limits_{k=1}^{n} E_k$, donde E_k no se intersecan y $g_n(x) = f_k(x)$ en E_k; luego,

$$\int\limits_E g_n(x)\,d\mu = \sum_{k=1}^{n} \int\limits_{E_k} f_k(x)\,d\mu \leqslant \sum_{k=1}^{n} \Phi(E_k) = \Phi(E).$$

Sea

$$f(x) = \sup\{f_n(x)\}.$$

Está claro que $f(x) = \lim\limits_{n\to\infty} g_n(x)$ y, consecuentemente, en virtud del teorema de Beppo Levi,

$$\int\limits_X f(x)\,d\mu = \lim_{n\to\infty} \int\limits_X g_n(x)\,d\mu = M.$$

Probemos ahora que

$$\Phi(E) - \int\limits_E f(x)\,d\mu \equiv 0.$$

Por la forma de construirla, la función de conjunto

$$\lambda(E)=\Phi(E)-\int_E f(x)\,d\mu$$

es no negativa y posee todas las propiedades de una medida. Si $\lambda \not\equiv 0$, existen, de acuerdo con el lema, un $\varepsilon>0$ y un B, $\mu(B)>0$, tales que

$$\varepsilon\mu(E\cap B)\leqslant\lambda(E\cap B)$$

para cualquier E medible. Tomando entonces $h(x)=f(x)+\varepsilon\chi_B(x)$, donde χ_B es la función característica del conjunto E, tendríamos para cualquier conjunto E medible

$$\int_E h(x)\,d\mu=\int_E f(x)\,d\mu+\varepsilon\mu(E\cap B)\leqslant\int_{E\setminus B} f(x)\,d\mu+\Phi(E\cap B)\leqslant\Phi(E).$$

Esto significaría que la función h pertenece al conjunto K definido anteriormente. Pero, al mismo tiempo,

$$\int_X h(x)\,d\mu=\int_X f(x)\,d\mu+\varepsilon\mu(B)>M,$$

lo que contradice a la definición de M. Luego, hemos demostrado la existencia de una función f tal que

$$\Phi(A)=\int_A f(x)\,d\mu.$$

Probemos su unicidad. Si para todo $A\in\mathfrak{S}_\mu$

$$\Phi(A)=\int_A f_1(x)\,d\mu=\int_A f_2(x)\,d\mu,$$

entonces cualquiera que sea n para los conjuntos

$$A_n=\left\{x:\ f_2(x)-f_1(x)>\frac{1}{n}\right\}$$

tenemos

$$\mu(A_n)\leqslant n\int_{A_n} f_1(x)-f_2(x)\,d\mu=0.$$

De la misma forma, para $B_m=\left\{x:\ f_1(x)-f_2(x)>\frac{1}{m}\right\}$ tenemos

$$\mu=(B_m)=0.$$

Como

$$\{x\colon f_1(x) \neq f_2(x)\} = \bigcup_n A_n \cup \bigcup_m B_m,$$

tenemos

$$\mu\{x\colon f_1(x) \neq f_2(x)\} = 0,$$

es decir, $f_1(x) = f_2(x)$ en casi todos los puntos. Hemos terminado la demostración.

Observación. El teorema de Radon—Nikodym constituye, evidentemente, una generalización natural del teorema de Lebesgue de que toda función absolutamente continua es integral de su derivada. Sin embargo, al considerar las funciones en la recta, tenemos en nuestro poder un método efectivo para buscar la derivada, como es el cálculo del límite del cociente $\frac{\Delta f}{\Delta x}$, mientras que el teorema de Radon—Nikodym sólo afirma la existencia de la derivada $\frac{d\Phi}{d\mu}$ de una carga absolutamente continua Φ respecto a la medida μ; pero, no ofrece método alguno para calcularla. Se puede indicar este método; pero, aquí no vamos a detenernos en ello. En líneas generales, este método consiste en calcular el límite del cociente $\frac{\Phi(A)}{\mu(A)}$ según un sistema de conjuntos que «se contraen», en cierto sentido, alrededor del punto dado. Estas cuestiones son estudiadas detalladamente, por ejemplo, en el libro de G. E. Shílov y B. L. Gurévich *«Integral, medida, derivada»* [13].

§ 6. INTEGRAL DE STIELTJES

1°. **Medidas de Stieltjes.** Al hablar, en el § 1 del capítulo precedente, de la construcción de la medida de Lebesgue, hemos mencionado ya la construcción siguiente. Supongamos definida en un segmento $[a, b]$ una función monótona no decreciente F; aceptaremos, para concretar, que es continua a la izquierda. Definiendo las medidas de todos los segmentos, los intervalos y los semisegmentos, pertenecientes al segmento básico $[a, b]$, mediante las igualdades

$$\begin{aligned} m(\alpha, \beta) &= F(\beta) - F(\alpha+0), \\ m[\alpha, \beta] &= F(\beta+0) - F(\alpha), \\ m(\alpha, \beta] &= F(\beta+0) - F(\alpha+0), \\ m[\alpha, \beta) &= F(\beta) - F(\alpha), \end{aligned}$$

podemos extender después esta medida, empleando el procedimiento de Lesbesque de prolongación de medida, a un σ-anillo que contiene todos los subconjuntos abiertos y cerrados (y, consecuentemente, todos los subconjuntos borelianos) del segmento $[a, b]$. La medida μ_F que se obtiene a partir de esta construcción se llama *medida de Lebesgue—Stieltjes* correspondiente a la función F, mientras que la propia función F se llama *función generatriz* de esta medida.

Consideremos algunos casos particulares de medidas de Lebesgue—Stieltjes.

1. Sean F una función de saltos, $x_1, x_2, \ldots$ sus puntos de discontinuidad y $h_1, h_2, \ldots$ sus saltos en estos puntos. Entonces, la medida μ_F correspondiente a esta función generatriz está construida del siguiente modo: todos los subconjuntos del segmento $[a, b]$ son medibles y la medida de un conjunto A es

$$\mu_F(A) = \sum_{x_i \in A} h_i. \tag{1}$$

En efecto, de la definición de la medida de Lebesgue—Stieljes se ve inmediatamente que la medida de cada punto x_i es igual a h_i, mientras que la medida del complemento del conjunto $\{x_i\}_{i=1}^{\infty}$ es igual a cero. De aquí, debido a la σ-aditividad de la medida μ_F, se deduce la igualdad (1) para cualquier $A \subset [a, b]$. Una medida μ_F construida a partir de una función de saltos se llama medida *discreta*.

2. Sea F una función no decreciente absolutamente continua en $[a, b]$ y sea $f = F'$ su derivada. En este caso, la medida correspondiente μ_F está definida en todos los subconjuntos de $[a, b]$ medibles según Lebesgue y, además, para cada conjunto A de este tipo

$$\mu_F(A) = \int_A f(x)\,dx. \tag{2}$$

Efectivamente, en virtud del teorema de Lebesgue, tenemos para cada intervalo (α, β)

$$\mu_F(\alpha, \beta) = F(\beta) - F(\alpha) = \int_\alpha^\beta f(x)\,dx.$$

Como la extensión lebesguiana de toda medida σ-aditiva se determina unívocamente por sus valores en el semianillo inicial, de aquí se desprende la validez de (2) para todo $A \subset [a, b]$ medible según Lebesgue. Una medida μ_F correspondiente a una función absolutamente continua F se llama medida *absolutamente continua*.

3. Si F es una función continua singular, su medida correspondiente μ_F está concentrada íntegramente en aquel conjunto de medida lebesguiana nula en el que F' es diferente de cero o no existe. La propia medida μ_F se llama en este caso medida *singular*.

Está claro que, si $F=F_1+F_2$, se tiene $\mu_F=\mu_{F_1}+\mu_{F_2}$ y, puesto que toda función monótona puede ser representada como suma de una función de saltos y de los componentes absolutamente continuo y singular, de aquí se desprende que *toda medida de Lebesgue—Stieltjes puede ser representada como suma de una medida discreta, absolutamente continua y singular*. Una función monótona se descompone, salvo un sumando constante, en una función de saltos, absolutamente continua y singular. Luego, la descomposición de toda medida de Lebesgue—Stieltjes en las componentes discreta, absolutamente continua y singular es única.

Lo expuesto se refiere a medidas de Lebesgue—Stieltjes en un segmento. Si ahora F es una función monótona no decreciente acotada (superior e inferiormente) en toda la recta, entonces, definiendo la medida de cualquier segmento, intervalo y semisegmento de la recta mediante fórmulas análogas a (1) y (2), obtendremos una medida finita en toda la recta que llamaremos medida de Lebesgue—Stieltjes (en la recta). En particular, la medida de toda la recta será, en este caso, igual a

$$F(\infty)-F(-\infty),$$

donde

$$F(\infty)=\lim_{x\to\infty} F(x), \quad F(-\infty)=\lim_{x\to\infty} F(x)$$

(la existencia de estos límites se debe a que F es monótona y acotada).

El concepto de medida de Lebesgue—Stieltjes abarca, de hecho, todas las medidas (esto es, todas las funciones de conjunto finitas no negativas y σ-aditivas) en la recta. En efecto, sea μ cualquier medida de este tipo. Tomando

$$F(x)=\mu(-\infty, x),$$

obtendremos una función monótona tal que su correspondiente medida de Lebesgue—Stieltjes coincide con la medida inicial μ. Es decir, el término de «medidas de Lebesgue—Stieltjes» no significa, de hecho, una clase especial de medidas en la recta, sino que indica simplemente el método concreto de construir esta medida a partir de una función generatriz.

2°. Integral de Lebesgue — Stieltjes. Sea μ_F una medida en el segmento $[a, b]$ generada por una función monótona F. Para esta medida se define de manera habitual la clase de funciones sumables y se introduce el concepto de la integral de Lebesgue

$$\int_a^b f(x)\, d\mu_F(x).$$

Una integral de este tipo tomada respecto a una medida μ_F, correspondiente a una función generatriz F, se llama *integral de Lebesgue—Stieltjes* y se designa mediante

$$\int_a^b f(x)\, dF(x).$$

Consideremos algunos casos particulares.

1. Si F es una función de saltos (esto es, μ_F es una medida discreta), la integral

$$\int_a^b f(x)\, dF(x)$$

se reduce, evidentemente, a la suma

$$\sum_i f(x_i)\, h_i,$$

donde x_i son los puntos de discontinuidad de la función F y h_i, los saltos de F en los puntos x_i.

2. Si F es una función absolutamente continua, su integral correspondiente de Lebesgue—Stieltjes

$$\int_a^b f(x)\, dF(x)$$

es igual a

$$\int_a^b f(x)\, F'(x)\, dx$$

es decir, a la integral de $f(x)\,F'(x)$ respecto a la medida lebesguiana habitual. En efecto, si $f(x) = \text{const}$, la igualdad

$$\int_a^b f(x)\, dF(x) = \int_a^b f(x)\, F'(x)\, dx \tag{3}$$

se deduce de (2). Debido a la σ-aditividad de las integrales, la igualdad (3) es extensible también a las funciones simples sumables según la medida μ_F. Sea ahora $\{f_n\}$ una sucesión de funciones simples convergente uniformemente hacia f. Sin perder generalidad, podemos aceptar que $\{f_n\}$ es una sucesión no decreciente. Entonces, $\{f_n(x)F'(x)\}$ es una sucesión no decreciente convergente en casi todos los puntos hacia $f(x)F'(x)$ y, en virtud del teorema de Beppo Levi, podemos pasar al límite en la igualdad

$$\int_a^b f_n(x)\,dF(x) = \int_a^b f_n(x)F'(x)\,dx$$

para $n \to \infty$.

De lo expuesto de deduce que, siendo F una suma de una función de saltos y otra absolutamente continua, la integral de Lebesgue—Stieltjes respecto a la medida μ_F se reduce a una serie (o una suma finita) más una integral respecto a la medida habitual de Lebesgue. En cambio, si F contiene también una componente singular, esta reducción resulta imposible.

El concepto de la integral de Lebesgue—Stieltjes puede ser ampliado de un modo natural, pasando de las funciones monótonas a las funciones de variación acotada. Sea Φ una función de este tipo. Representémosla como diferencia de dos funciones monótonas

$$\Phi = v - g,$$

donde v es la variación total de la función φ en el segmento $[a, x]$. Introduzcamos ahora la integral de Lebesgue—Stieltjes respecto a Φ, tomando, por definición,

$$\int_a^b f(x)\,d\Phi(x) = \int_a^b f(x)\,dv(x) - \int_a^b f(x)\,dg(x).$$

Es fácil probar que si se tiene otra representación de Φ como diferencia de dos funciones monótonas, digamos

$$\Phi = v^* - g^*$$

entonces,

$$\int_a^b f(x)dv(x) - \int_a^b f(x)\,dg(x) = \int_a^b f(x)\,dv^*(x) - \int_a^b f(x)\,dg^*(x),$$

es decir, para calcular la integral de Lebesgue—Stieltjes respecto a una función Φ dada puede ser empleada cualquier representación de esta función por medio de la diferencia de dos funciones monótonas.

3°. Algunas aplicaciones de la integral de Lebesgue — Stieltjes en la teoría de probabilidades. La integral de Lebesgue —Stieltjes encuentra aplicación tanto en el Análisis, como en otras muchas cuestiones aplicadas. En particular, este concepto se emplea ampliamente en la teoría de probabilidades. Recordemos que se llama función de distribución de una variable aleatoria ξ a una función F (no decreciente, obviamente) definida para cada x por la igualdad

$$F(x)=P(\xi<x),$$

es decir, $F(x)$ es la probabilidad de que la variable aleatoria ξ tome un valor menor que x. Es evidente que cada función de distribución es monótona no decreciente, continua a la derecha y verifica las condiciones

$$F(-\infty)=0, \quad F(+\infty)=1.$$

Viceversa, toda función de este tipo puede ser considerada como función de distribución de una variable aleatoria.

Son características sustanciales de una variable aleatoria su valor medio (o esperanza matemática)

$$M\xi=\int_{-\infty}^{\infty} x\,dF(x) \tag{4}$$

y varianza

$$V\xi=\int_{-\infty}^{\infty}(x-M\xi)^2\,dF(x). \tag{5}$$

Entre las variables aleatorias suelen destacarse las variables aleatorias discretas y continuas. Una variable aleatoria se llama *discreta* cuando puede tomar sólo un número finito o numerable de valores

$$x_1, x_2, \ldots, x_n, \ldots$$

(por ejemplo, el número de llamadas telefónicas que se reciben en una central durante un intervalo de tiempo es una variable aleatoria discreta).

Si $p_1, p_2, \ldots, p_n, \ldots$ son las probabilidades con que la variable ξ toma los valores $x_1, x_2, \ldots, x_n, \ldots$, la función de distribución para ξ es, evidentemente, una función de saltos. Para ella las integrales (4) y (5) se reducen respectivamente a las sumas

$$M\xi=\sum_i x_i p_i$$

y

$$V\xi=\sum (x_i-a)^2 p_i \quad (a=M\xi).$$

Una variable aleatoria ξ se llama *continua* cuando su función de distribución F es absolutamente continua. La derivada F' de esta función de distribución se llama *densidad de distribución de probabilidades* de la variable aleatoria ξ. De acuerdo con lo expuesto en el punto anterior, las integrales de Stieltjes que expresan el valor medio y la varianza de una variable aleatoria continua se reducen a integrales respecto a la medida lebesguiana habitual:

$$M\xi = \int_{-\infty}^{\infty} xp(x)\,dx,$$

$$V\xi = \int_{-\infty}^{\infty} (x-a)^2\,p(x)\,dx,$$

donde $p = F'$ es la densidad de distribución de probabilidades para ξ y $a = M\xi$.

Los cursos elementales de teoría de probabilidades se limitan generalmente al estudio de variables aleatorias discretas o continuas que en lo fundamental son las únicas que aparecen en cuestiones aplicadas. Sin embargo, una función de distribución de una variable aleatoria puede tener, en el caso general, una componente singular, de modo que no toda variable aleatoria puede ser representada como una combinación de variables aleatorias discreta y continua.

Sean ξ una variable aleatoria, F su función de distribución y $\eta = \varphi(\xi)$ otra variable aleatoria que representa una función de ξ. El valor medio $M\eta$ de la variable se puede representar, por definición, como

$$\int_{-\infty}^{\infty} x\,d\Phi(x),$$

donde Φ es la función de distribución para η. Es sustancial, sin embargo, que si φ es sumable respecto a la medida generada en la recta por la función F, el valor medio de la variable η se puede representar también a través de la función de distribución F de la variable ξ, a saber:

$$M\eta = M\varphi(\xi) = \int_{-\infty}^{\infty} \varphi(x)\,dF(x).$$

En efecto, la función $y = \varphi(x)$ determina una aplicación de la recta $(-\infty < x < \infty)$ con la medida μ_F (generada por F) en la recta $(-\infty < y < \infty)$ con la medida μ_Φ, en la que se transforma por la aplicación $y = \varphi(x)$ la medida μ_F. Pero, de los resultados del capítulo VI se desprende que si (X, μ) y (Y, ν) son dos espacios provistos de medidas, φ es una aplicación que conserva la medida y transforma (X, μ) en (Y, ν) y f es una función sumable en (Y, ν), entonces

$$\int_Y f(y)\,d\nu = \int_X f(\varphi(x))\,d\mu$$

(sustitución de variables en la integral de Lebesgue). Tomando aquí $f(y)=y$ y $\mu=\mu_F$, $\nu=\mu_\Phi$, obtendremos la igualdad necesaria. Luego, para calcular el valor medio (y, claro está, la varianza también) de una función de una variable aleatoria ξ es suficiente conocer solamente la función de distribución de la propia variable ξ.

4°. Integral de Riemann — Stieltjes. Además de la integral de Lebesgue — Stieltjes, considerada en el punto anterior, que representa de hecho, la diferencia de las integrales lebesguianas de una función dada f respecto a dos medidas, definidas en la recta, se puede definir también la así llamada integral de Riemann — Stieltjes. Ella se introduce como límite de sumas integrales, análogas a las sumas integrales habituales de Riemann.

Sea de nuevo Φ una función continua a la izquierda de variación acotada, definida en un segmento $[a, b]$, y sea f una función arbitraria en este segmento. Consideremos una partición

$$a=x_0<x_1<x_2<\ldots<x_n=b$$

del segmento $[a, b]$ y, escogiendo en cada elemento $[x_{i-1}, x_i]$ de esta partición un punto arbitrario ξ_i, formemos la suma

$$\sum_{i=1}^{n} f(\xi_i)\,[\Phi(x_i)-\Phi(x_{i-1})]. \tag{6}$$

Si para $\max(x_i-x_{i-1})\to 0$, estas sumas tienden a un límite determinado (que no depende de cómo se ha dividido el segmento $[a, b]$ ni de cómo se han escogido los puntos ξ_i en cada elemento de la partición), este límite se llama integral de Riemann — Stieltjes de la función f respecto a la función Φ y se designa con

$$\int_a^b f(x)\,d\Phi(x). \tag{7}$$

TEOREMA 1. *Si la función f es continua en $[a, b]$, su integral de Riemann — Stieltjes respecto a Φ existe y coincide con la integral de Lebesgue — Stieltjes correspondiente.*

DEMOSTRACION. La suma (6) puede ser considerada como la integral de Lebesgue — Stieltjes de la función escalonada

$$f_n(x)=\xi_i \quad \text{para} \quad x_{i-1}\leqslant x<x_i.$$

Al subdividir la partición del segmento $[a, b]$, la sucesión de estas funciones converge uniformemente hacia f. Por lo tanto, el límite de estas sumas existe y representa la integral de Lebesgue — Stieltjes de la función límite f (teorema sobre el paso al límite bajo el signo de la integral). Al mismo tiempo, este límite es precisamente la integral de Riemann—Stieltjes (7). El teorema queda demostrado.

Demostremos algunas propiedades elementales de la integral de Riemann—Stieltjes. Siempre suponemos que f es continua en $[a, b]$.

1. *Tiene lugar la estimación* (teorema del valor medio)

$$\left|\int_a^b f(x)\,d\Phi(x)\right| \leqslant \max|f(x)|\,V_a^b[\Phi] \tag{8}$$

($V_a^b[\Phi]$ es la variación total de la función Φ en $[a, b]$).

En efecto, cualquiera que sea la partición del segmento $[a, b]$ se cumple la desigualdad

$$\left|\sum_{i=1}^{n} f(\xi_i)(\Phi(x_i)-\Phi(x_{i-1}))\right| \leqslant \sum_{i=1}^{n} |f(\xi_i)|\cdot|\Phi(x_i)-\Phi(x_{i-1})| \leqslant$$

$$\leqslant \max|f(x)|\cdot\sum_{i=1}^{n}|\Phi(x_i)-\Phi(x_{i-1})| \leqslant \max|f(x)|\,V_a^b[\Phi].$$

Pasando en esta desigualdad de las sumas integrales al límite de las mismas, obtenemos la estimación (8). Para $\Phi(x)=x$ ella se convierte en la estimación conocida

$$\left|\int_a^b f(x)\,dx\right| \leqslant (b-a)\max|f(x)|$$

para la integral de Riemann.

2. *Si* $\Phi=\Phi_1+\Phi_2$, *se tiene*

$$\int_a^b f(x)\,d\Phi(x) = \int_a^b f(x)\,d\Phi_1(x) = \int_a^b f(x)\,d\Phi_2(x).$$

En efecto, para toda partición del segmento $[a, b]$ se cumple la desigualdad correspondiente para las sumas integrales; por consiguiente, ella se conserva también cuando se pasa al límite, es decir, para las integrales.

3. *Si* ψ *es una función de variación acotada distinta de cero solamente en un conjunto finito o numerable de puntos,*

$$\int_a^b f(x)\,d\psi(x)=0$$

para cualquier función f *continua en* $[a, b]$.

En efecto, esto es evidente para una función distinta de cero en un unico punto x_0 (ya que al tomar particiones tan pequeñas como se quiera del segmento $[a, b]$ sin que el punto x_0 sea un punto de división, obtendremos sumas integrales iguales a cero); luego, debido a la aditividad, esto es válido también para cualquier función diferente de cero en un número finito de puntos. Supongamos ahora que ψ es distinta de cero en los puntos

$$r_1, r_2, \ldots, r_n \ldots$$

y sean

$$y_1, y_2, \ldots, y_n, \ldots$$

sus valores en estos puntos. Como ψ es de variación acotada, tenemos $\sum |y_n| < \infty$. Escojamos ahora N de manera que $\sum\limits_{n>N} |y_n| < \varepsilon$ y representemos ψ como suma

$$\psi = \psi_N + \tilde{\psi}$$

donde ψ_N toma los valores $y_1, \ldots, y_N$ en los puntos $r_1, \ldots, r_N$ y es igual a 0 en los demás, mientras que $\tilde{\psi}$ es distinta de 0 solamente en los puntos $r_{N+1}, r_{N+2}, \ldots$. Las sumas integrales correspondientes a $\tilde{\psi}$ verifican la desigualdad

$$\left| \sum_{i=1}^{m} f(\xi_i)(\tilde{\psi}(x_i) - \tilde{\psi}(x_{i-1})) \right| \leqslant 2M \cdot \sum_{n>N} |y_n'| \leqslant 2M\varepsilon,$$

donde $M = \max |f(x)|$. Por eso,

$$\left| \int_a^b f(x)\, d\psi(x) \right| \leqslant \left| \int f(x)\, d\psi_N(x) \right| + \left| \int f(x)\, d\tilde{\psi}(x) \right| \leqslant 2M\varepsilon;$$

de aquí se deduce, debido a la arbitrariedad de ε, nuestra afirmación.

4. *Si f es una función continua, la integral de Riemann — Stieltjes $\int_a^b f(x)\, d\Phi(x)$ no depende de los valores que toma Φ en sus puntos de discontinuidad.*

En efecto, sean Φ_1 y Φ_2 dos funciones de variación acotada coincidentes en todos sus puntos de continuidad. Entonces, la diferencia

$$\psi = \Phi_1 - \Phi_2$$

representa una función distinta de cero solamente en un conjunto a lo sumo numerable de puntos. Lo demás se desprende de las propiedades 2 y 3.

Puesto que la integral de Riemann—Stieltjes de una función continua coincide con la correspondiente integral de Lebesgue—Stieltjes, para la integral de Riemann—Stieltjes son válidas las igualdades:

$$\int_a^b f(x)\,d\Phi(x) = \sum_i f(x_i)\,h_i,$$

donde Φ es una función de saltos, y

$$\int_a^b f(x)\,d\Phi(x) = \int_a^b f(x)\,\Phi'(x)\,dx, \tag{9}$$

donde Φ es una función absolutamente continua. Además, si Φ' es integrable según Riemann, la integral que figura en el miembro derecho de (9) puede ser comprendida en el sentido de Riemann.

Todo lo expuesto para la integral de Riemann—Stieltjes en el caso de un segmento finito se puede extender fácilmente al caso en que la integral se considera en toda la recta o en una semirrecta.

Observación. En el caso de la integral de Stieltjes, a diferencia de la integral habitual de Riemann, los valores de la integral en el intervalo (a, b), el segmento $[a, b]$ y los semisegmentos $(a, b]$ y $[a, b)$ no coinciden, en general (los puntos aislados tienen medida de Stieltjes positiva, si la función que genera la medida es en ellos discontinua). El símbolo

$$\int_a^b f(x)\,d\Phi(x)$$

se interpreta comúnmente, si no se dice lo contrario, como la integral referida al semisegmento $[a, b)$.

5°. **Paso al límite bajo el signo de la integral de Stieltjes.** En el capítulo VI hemos demostrado varios teoremas acerca del paso al límite bajo el signo de la integral de Lebesgue. El problema se planteaba allí del modo siguiente: dadas una sucesión $\{f_n\}$ de funciones y las integrales de estas funciones respecto a una medida determinada, nos interesaba la posibilidad de pasar al límite de esta sucesión bajo el signo de la integral. Sin embargo, en el caso de la integral de Stieltjes tiene también interés plantear el problema del modo siguiente: sea dada una sucesión $\{\Phi_n\}$ de funciones de variación acotada; ¿bajo qué condiciones se puede pasar al límite bajo el signo de la integral

$$\int_a^b f(x)\, d\Phi_n(x)$$

para una función fija f?

En este orden tiene lugar el teorema siguiente.

TEOREMA 2 (primer teorema de Helly). *Supongamos que las funciones Φ_n de variación acotada en un segmento $[a, b]$ convergen en cada punto de este segmento hacia una función Φ y que las variaciones totales de las funciones Φ_n están acotadas en conjunto*

$$V_a^b[\Phi_n] \leqslant C \quad (n=1,\ 2,\ \ldots).$$

Entonces la función límite Φ es también de variación acotada y cualquiera que sea la función continua f tiene lugar la igualdad

$$\lim_{n\to\infty} \int_a^b f(x)\, d\Phi_n(x) = \int_a^b f(x)\, d\Phi(x). \tag{10}$$

DEMOSTRACION. Probemos ante todo que la variación total de la función límite Φ no es mayor que la constante C con la que están acotadas todas las $V_a^b[\Phi_n]$. En efecto, para cualquier partición del segmento $[a, b]$ por los puntos

$$a = x_0 < x_1 < \ldots < x_m = b$$

tenemos

$$\sum_{k=1}^{m} |\Phi(x_k) - \Phi(x_{k-1})| = \lim_{n\to\infty} \sum_{k=1}^{m} |\Phi_n(x_k) - \Phi_n(x_{k-1})| \leqslant C;$$

luego,

$$V_a^b[\Phi] \leqslant C.$$

Probemos ahora que la relación (10) es válida en el caso en que f es una función escalonada. Supongamos que f toma los valores h_k en los intervalos (x_{k-1}, x_k). Entonces,

$$\int_a^b f(x)\, d\Phi_n(x) = \sum_k h_k [\Phi_n(x_k) - \Phi_n(x_{k-1})] \tag{11}$$

y

$$\int_a^b f(x)\, d\Phi(x) = \sum_k h_k [\Phi(x_k) - \Phi(x_{k-1})]. \tag{12}$$

Está claro que la primera de estas expresiones se convierte en la segunda cuando $n \longrightarrow \infty$.

Sea ahora f una función continua y sea ε un número positivo arbitrario. Escojamos una función escalonada f_ε de manera que

$$|f(x)-f_\varepsilon(x)| < \frac{\varepsilon}{3C}.$$

Entonces,

$$\left|\int_a^b f(x)\,d\Phi(x) - \int_a^b f(x)\,d\Phi_n(x)\right| \leqslant \left|\int_a^b f(x)\,d\Phi(x) - \int_a^b f_\varepsilon(x)\,d\Phi(x)\right| +$$
$$+\left|\int_a^b f_\varepsilon(x)\,d\Phi(x) - \int_a^b f_\varepsilon(x)\,d\Phi_n(x)\right| +$$
$$+\left|\int_a^b f_\varepsilon(x)\,d\Phi_n(x) - \int_a^b f(x)\,d\Phi_n(x)\right|.$$

En virtud del teorema del valor medio para la integral de Stieltjes, el primero y tercer sumandos son menores que $\frac{\varepsilon}{3}$, mientras que el segundo sumando es menor que $\frac{\varepsilon}{3}$ para n suficientemente grandes. Como $\varepsilon > 0$ es arbitrario, de aquí se deduce la afirmación del teorema.

Observación. Este teorema subsiste también en el caso en que uno o ambos extremos de la integral

$$\int_a^b f(x)\,d\Phi_n(x)$$

son infinitos. Sin embargo, la función f en este caso debe tender en el infinito a un límite finito (esto permite aproximarla uniformemente en el intervalo infinito mediante funciones escalonadas que toman solamente un número finito de valores).

El primer teorema de Helly ofrece las condiciones en las que se puede pasar al límite respecto a una sucesión $\{\Phi_n\}$ de funciones de variación acotada en la integral de Stieltjes. El teorema que sigue explica cuándo puede ser garantizada la existencia misma de la sucesión que satisface las condiciones del teorema anterior.

TEOREMA 3 (segundo teorema de Helly). *De cualquier conjunto infinito Φ de funciones φ que están definidas en un segmento $[a, b]$ y que verifican las condiciones*

$$\max|\varphi(x)| \leqslant C, \quad V_a^b[\varphi] \leqslant K, \tag{13}$$

donde C y K son constantes (las mismas para toda $\varphi \in \Phi$), *se puede extraer una sucesión parcial convergente en cada punto del segmento $[a, b]$.*

DEMOSTRACION. Basta demostrar este teorema para las funciones no decrecientes. En efecto, sea

$$\varphi = v - g,$$

donde $v(x)$ es la variación total de la función φ en el segmento $[a, x]$. Entonces, las funciones v correspondientes a todas las $\varphi \in \Phi$ satisfacen las desigualdades

$$\max|v(x)| \leqslant C, \quad V_a^b[v] = V_a^b[\varphi] \leqslant K,$$

esto es, verifican las condiciones del teorema, y son monótonas. Suponiendo que el teorema es válido para las funciones monótonas, escojamos en Φ una sucesión $\{\varphi_n\}$ de manera que las v_n correspondientes converjan hacia un límite v. Las funciones

$$g_n = v_n - \varphi_n$$

serán también monótonas y verificarán las condiciones del teorema. Luego, se puede extraer de $\{\varphi_n\}$ una sucesión parcial $\{\varphi_{n_k}\}$ tal que g_{n_k} convergen hacia un límite g. Pero, en este caso,

$$\varphi_{n_k}(x) \to \varphi(x) = v(x) - g(x).$$

Demostremos, pues, el teorema para una familia Φ de funciones monótonas. Sean

$$r_1, r_2, \ldots, r_n, \ldots$$

todos los puntos racionales del segmento $[a, b]$. Debido a (13), los números $\varphi(r_1)$ $(\varphi \in \Phi)$ forman un conjunto acotado y, por eso, existe una sucesión $\{\varphi_n^{(1)}\}$ convergente en el punto r_1. Escojamos ahora en ella una sucesión parcial $\{\varphi_n^{(2)}\}$ que converja en r_2 (y, por supuesto, en r_1). Escojamos después en $\{\varphi_n^{(2)}\}$ una sucesión parcial $\{\varphi_n^{(3)}\}$ convergente en el punto r_3, etc. La sucesión diagonal $\{\varphi_n^{(n)}\}$ convergerá evidentemente, en todos los puntos racionales del segmento $[a, b]$. El límite de esta sucesión será una función no decreciente φ definida, por ahora, solamente en los puntos $r_1, r_2, \ldots, r_n, \ldots$ Definámosla en los demás puntos del segmento $[a, b]$, tomando para las x irracionales

$$\varphi(x) = \lim_{r \to x-0} \varphi(r) \quad (r \text{ son racionales}).$$

Probemos que la función no decreciente φ, obtenida de esta forma, es, en todos los puntos de continuidad, el límite de la sucesión $\{\varphi_n^{(n)}\}$. Sea x^* uno de estos puntos. Entonces, para un $\varepsilon > 0$ dado se puede escoger un $\delta > 0$ tal que

$$|\varphi(x^*) - \varphi(x)| < \frac{\varepsilon}{6} \text{ siempre que } |x^* - x| < \delta. \tag{14}$$

Escojamos unos puntos racionales r' y r'' de manera que $r' < x^* < r''$, $r' > x^* - \delta$ y $r'' < x^* + \delta$. Sea ahora n tan grande que para $n > n_0$ se cumplen las desigualdades

$$|\varphi_n(r') - \varphi(r')| < \frac{\varepsilon}{6} \text{ y } |\varphi_n(r'') - \varphi(r'')| < \frac{\varepsilon}{6}. \tag{15}$$

De (14) y (15) se desprende que

$$|\varphi_n(r')| - \varphi_n(r'')| < \frac{2}{3}\varepsilon.$$

Como la función φ_n es no decreciente, tenemos $\varphi_n(r') \leqslant \varphi_n(x^*) \leqslant \leqslant \varphi_n(r'')$. Luego,

$$|\varphi(x^*) - \varphi_n(x^*)| \leqslant |\varphi(x^*) - \varphi(r')| + |\varphi(r') - \varphi_n(r')| + \\ + |\varphi_n(r') - \varphi_n(x^*)| \leqslant \frac{\varepsilon}{6} + \frac{4\varepsilon}{6} + \frac{\varepsilon}{6} = \varepsilon$$

y esto significa precisamente que $\lim\limits_{n\to\infty} \varphi_n(x^*) = \varphi(x^*)$.

Hemos logrado construir una sucesión de funciones de Φ que converge hacia la función límite φ en todo punto, excepto, posiblemente, los puntos de discontinuidad de la función φ. Como el conjunto de estos puntos es a lo sumo numerable, aplicando de nuevo el proceso diagonal, podemos extraer de la sucesión $\{\varphi_n\}$ una sucesión parcial que converja hacia φ también en estos puntos, es decir, que converja en todo $[a, b]$.

6°. Representación general de funcionales lineales continuas en el espacio de funciones continuas. Hemos señalado ya algunas aplicaciones de la integral de Stieltjes. Ahora estudiaremos otro problema, relacionado con este concepto, determinando la forma general de una funcional lineal en el espacio $C_{[a, b]}$.

TEOREMA 4 (Riesz). *Toda funcional lineal continua F en el espacio $C_{[a, b]}$ puede ser representada en la forma*

$$F(f) = \int_a^b f(x)\, d\varphi(x) \tag{16}$$

donde φ es una función de variación acotada. Además,

$$\|F\| = V_a^b[\varphi].$$

DEMOSTRACIÓN. El espacio $C_{[a,\,b]}$ puede ser considerado como subespacio del espacio $M_{[a,\,b]}$ de todas las funciones acotadas en este segmento con la misma norma

$$\|f\| = \sup |f(x)|$$

que existe en $C_{[a,\,b]}$. Sea F una funcional lineal continua en $C_{[a,\,b]}$. En virtud del teorema de Hahn — Banach, puede ser prolongada, conservando su norma, de $C_{[a,\,b]}$ a todo el $M_{[a,\,b]}$. Esta funcional prolongada estará definida, en particular, para todas las funciones de tipo

$$f_{\tau}(x) = \begin{cases} 1 & \text{para } x \leqslant \tau, \\ 0 & \text{para } x > \tau. \end{cases}$$

Tomemos

$$\varphi(\tau) = F(f_{\tau}) \tag{17}$$

y probemos que la función φ es de variación acotada en el segmento $[a, b]$. En efecto, tomemos una partición arbitraria

$$a = x_0 < x_1 < \ldots < x_n = b$$

de este segmento y pongamos

$$\alpha_k = \operatorname{sgn}[\varphi(x_k) - \varphi(x_{k-1})], \quad (k = 1, 2, \ldots, n).$$

Entonces,

$$\sum_{k=1}^{n} |\varphi(x_k) - \varphi(x_{k-1})| = \sum_{k=1}^{n} \alpha_k [\varphi(x_k) - \varphi(x_{k-1})] =$$

$$= \sum_{k=1}^{n} \alpha_k F(f_{x_k} - f_{x_{k-1}}) = F\left[\sum_{k=1}^{n} \alpha_k (f_{x_k} - f_{x_{k-1}})\right] \leqslant$$

$$\leqslant \|F\| \cdot \left\| \sum_{k=1}^{n} \alpha_k (f_{x_k} - f_{x_{k-1}}) \right\|.$$

Pero, la función $\sum_{k=1}^{n} \alpha_k (f_{x_k} - f_{x_{k-1}})$ toma sólo los valores ± 1 y 0. Por eso, su norma es igual a 1. Por lo tanto,

$$\sum_{k=1}^{n} |\varphi(x_k - \varphi(x_{k-1})| \leqslant \|F\|.$$

Como esto es válido para cualquier partición del segmento $[a, b]$, tenemos

$$V_a^b[\varphi] \leqslant \|F\|.$$

Es decir, hemos construido a partir de la funcional F una función φ de variación acotada. Probemos que es precisamente esta función mediante la cual la funcional F puede ser representada a través de la integral de Stieltjes (16).

Sea f una función continua cualquiera en $[a, b]$. Tomando arbitrariamente un ε positivo, escojamos $\delta > 0$ de manera que $|f(x'') - f(x')| < \varepsilon$ siempre que $|x'' - x'| < \delta$. Dividamos ahora el segmento $[a, b]$ mediante los puntos t_k en partes, de longitud cada una menor que δ, y consideremos la función escalonada

$$f^{(\varepsilon)}(x) = f(x_k) \text{ para } x_{k-1} \leqslant x \leqslant x_k, \quad k = 1, 2, \ldots, n.$$

Ella puede ser representada, evidentemente, en la forma

$$f^{(\varepsilon)}(x) = \sum_{k=1}^{n} f(x_k) [f_{x_k}(x) - f_{x_{k-1}}(x)],$$

donde f_τ es la función definida por la igualdad (17). Está claro que $|f(x) - f^{(\varepsilon)}(x)| < \varepsilon$ para todas las x, $a \leqslant x \leqslant b$, es decir,

$$\| f(x) - f^{(\varepsilon)}(x) \| < \varepsilon.$$

Calculemos el valor de la funcional F en el elemento $f^{(\varepsilon)}$. Debido a la linealidad de esta funcional y de acuerdo con la definición de la función f_τ, este valor es igual a

$$F(f^{(\varepsilon)}) = \sum_{k=1}^{n} f(x_k) [F(f_{x_k}) - F(f_{x_{k-1}})] = \sum_{k=1}^{n} f(x_k) [\varphi(x_k) - \varphi(x_{k-1})],$$

es decir, representa la suma integral para la integral

$$\int_a^b f(x)\, d\varphi(x).$$

Luego, para una partición suficientemente pequeña de $[a, b]$, tenemos

$$\left| F(f^{(\varepsilon)}) - \int_a^b f(x)\, d\varphi(x) \right| < \varepsilon.$$

Pero, al mismo tiempo,

$$|F(f) - F(f^{(\varepsilon)})| \leqslant \|F\| \cdot \|f - f_\varepsilon\| \leqslant \|F\| \cdot \varepsilon.$$

Por lo tanto,

$$\left| F(f) - \int_a^b f(x)\, d\varphi(x) \right| < \varepsilon(1 + \|F\|),$$

de donde, debido a la arbitrariedad de ε, obtenemos la igualdad

$$F(f) = \int_a^b f(x)\, d\varphi(x).$$

Hemos visto que la variación total de la función φ, dada por (17), satisface la desigualdad siguiente:

$$V_a^b[\varphi] \leqslant \|F\|. \tag{18}$$

Por otro lado, del teorema del valor medio para la integral de Stieltjes, se desprende inmediatamente que

$$\|F\| \leqslant V_a^b[\varphi]. \tag{19}$$

Comparando (18) y (19), obtenemos la igualdad

$$\|F\| = V_a^b[\varphi].$$

El teorema queda demostrado completamente.

Observación. Está claro que siendo φ una función arbitraria de variación acotada φ en el segmento $[a, b]$, la relación

$$F(f) = \int_a^b f(x)\, d\varphi(x)$$

determina una funcional lineal en el espacio $C_{[a, b]}$. Además, dos funciones φ_1 y φ_2, coincidentes en todos los puntos excepto, posiblemente, los de un conjunto a lo sumo numerable, determinan una misma funcional lineal; viceversa, si φ_1 y φ_2 determinan una misma funcional en $C_{[a, b]}$, esto es, si

$$\int_a^b f(x)\, d\varphi_1(x) = \int_a^b f(x)\, d\varphi_2(x)$$

para toda función continua, entonces φ_1 y φ_2 coinciden en sus puntos de continuidad, es decir, en todos los puntos excepto, posiblemente, los de un conjunto finito o numerable de puntos.

Luego, existe una aplicación biunívoca entre las funcionales lineales en $C_{[a, b]}$ y las clases de funciones de variación acotada en $[a, b]$ que verifican la condición $\varphi(a) = 0$, perteneciendo dos funciones a una misma clase cuando coinciden en sus puntos de continuidad. Para una función arbitraria φ de la clase correspondiente a la funcional dada F se cumple la desigualdad

$$\|F\| \leqslant V_a^b[\varphi];$$

la igualdad puede no tener lugar; pero, como se desprende de la demostración del teorema, en cada una de estas clases existe al menos una función para la cual esta igualdad se alcanza.

CAPITULO VIII

ESPACIOS DE FUNCIONES SUMABLES

Entre las diferentes clases de espacios normados que se emplean en el Análisis una de las más importantes es la de los espacios de funciones medibles de cierta potencia sumable y, en primero término, el espacio L_1 de todas las funciones sumables y el espacio L_2 de funciones de cuadrado sumable. Estudiaremos ahora las propiedades fundamentales de estos espacios. El contenido de este capítulo se basa, por un lado, en las propiedades generales de los espacios métricos y los espacios normados lineales, expuestas en los capítulos II, III y IV, y, por otro lado, en el concepto de la integral de Lebesgue introducido en el capítulo VI.

§ 1. ESPACIO L_1

1°. Definición y propiedades fundamentales del espacio L_1. Sea X un espacio provisto de una medida μ; la medida del propio X puede ser finita o infinita. Consideremos el conjunto de todas las funciones sumables en X. Como una combinación lineal de funciones sumables es de nuevo una función sumable, este conjunto, con las operaciones habituales de adición de funciones y multiplicación de las mismas por números, constituye un espacio lineal. Denotaremos este espacio mediante $L_1(X, \mu)$ o simplemente L_1. Introduzcamos en L_1 una norma tomando[1)]

$$\|f\| = \int |f(x)|\, d\mu. \tag{1}$$

1) Aquí y en lo sucesivo el símbolo $\int$ representa la integración en todo el espacio X.

Está claro que

$$\|\alpha f\| = |\alpha| \cdot \|f\|$$

y

$$\|f_1 + f_2\| \leqslant \|f_1\| + \|f_2\|.$$

Sin embargo, para que se cumpla también la última condición de la norma, a saber

$$\|f\| > 0 \text{ cuando } f \neq 0,$$

es preciso aceptar que las funciones equivalentes en X no se distinguen y representan un mismo elemento del espacio L_1. En particular, el elemento nulo de L_1 es el conjunto de todas las funciones iguales a cero en casi todos los puntos. En este caso, la expresión (1) poseerá todas las propiedades de la norma. Llegamos así a la definición siguiente.

DEFINICION 1. Se llama espacio L_1 al espacio normado cuyos elementos son las clases de funciones sumables equivalentes; la adición de elementos de L_1 y la multiplicación de los mismos por números se definen como las operaciones habituales de adición y multiplicación de funciones[1] y la norma se define mediante la fórmula

$$\|f\| = \int |f(x)|\, d\mu.$$

En L_1, al igual que en cualquier espacio normado, la distancia se introduce mediante la fórmula

$$\rho(f, g) = \|f - g\|.$$

La convergencia de una sucesión de funciones sumables, comprendida en el sentido de esta distancia, se llama *convergencia media.* El espacio L_1 puede ser considerado como compuesto de funciones complejas (espacio complejo L_1) o solamente de funciones reales (espacio real L_1). Los resultados de este parágrafo son válidos en ambos casos.

Para muchas cuestiones del Análisis tiene gran importancia el resultado siguiente.

TEOREMA 1. *El espacio L_1 es completo.*

DEMOSTRACION. Sea $\{f_n\}$ una sucesión fundamental en L_1, esto es,

$$\|f_n - f_m\| \longrightarrow 0 \text{ para } n, m \longrightarrow \infty.$$

---

[1] Con más precisión: puesto que cada elemento de L_1 es una clase de funciones sumables equivalentes, para sumar dos clases de este tipo tomamos un representante en cada una de ellas y llamamos suma de estas clases a la clase que contiene la suma de los representantes elegidos. Está claro que el resultado no depende de la selección de los representantes en las clases dadas.

Entonces, se puede encontrar una sucesión $\{n_k\}$ de índices tal que

$$\|f_{n_k}-f_{n_{k+1}}\|=\int |f_{n_k}(x)-f_{n_{k+1}}(x)|\,d\mu<\frac{1}{2^k}.$$

De esta desigualdad y del teorema de Beppo Levi se desprende que la serie

$$|f_{n_1}|+|f_{n_2}-f_{n_1}|+\dots$$

converge en casi todo X. Pero, entonces, también la serie

$$f_{n_1}+f_{n_2}-f_{n_1}+\dots$$

converge en casi todo X hacia una función

$$f(x)=\lim_{n\to\infty} f_{n_k}(x).$$

Luego, hemos probado que una sucesión fundamental de L_1 contiene sucesión parcial convergente en casi todos los puntos.

Probemos ahora que la sucesión parcial $\{f_{n_k}\}$ converge en la media hacia la misma función f. Como la sucesión $\{f_n\}$ es fundamental, para cualquier $\varepsilon>0$ fijo y k y l suficientemente grandes tenemos

$$\int |f_{n_k}(x)-f_{n_l}(x)|\,d\mu<\varepsilon.$$

De acuerdo con el teorema de Fatou, podemos pasar en esta desigualdad al límite para $l\longrightarrow\infty$ bajo el signo de la integral. Tendremos

$$\int |f_{n_k}(x)-f(x)|\,d\mu\leqslant\varepsilon,$$

de donde se deduce que $f\in L_1$ y $f_{n_k}\longrightarrow f$. Pero, una sucesión fundamental que contiene una sucesión parcial convergente hacia un límite converge hacia el mismo límite. El teorema queda demostrado.

2°. Conjuntos siempre densos en L_1. Por definición de la integral de Lebesgue, cualquiera que sea la función f sumable en X y cualquiera que sea $\varepsilon>0$ existe una función siempre sumable $\varphi(x)$ tal que

$$\int |f(x)-\varphi(x)|\,d\mu<\varepsilon.$$

Además, puesto que para una función sumable simple que toma los valores $y_1, y_2, \dots$ en los conjuntos $E_1, E_2, \dots$ la integral se define como la suma de la serie

$$\sum_{n=1}^{\infty} y_n\mu(E_n)$$

(si es que ella converge absolutamente), está claro que toda función sumable simple puede ser representada como límite (en media) de una sucesión de funciones sumables simples que toman solamente un número finito de valores. Luego, *en el espacio* L_1 *son siempre densas las funciones cada una de las cuales toma solamente un número finito de valores* (es decir, representa una combinación lineal de funciones características).

Sea R un espacio métrico provisto de una medida que verifica la condición siguiente (que se cumple para la medida de Lebesgue en un espacio euclídeo y en otros muchos casos de interés práctico): *todos los conjuntos abiertos y todos los conjuntos cerrados de* R *son medibles y para cualquier* $M \subset R$

$$\mu(M) = \inf_{M \subset G} \mu(G), \tag{2}$$

donde la cota inferior se toma respecto a todos los conjuntos abiertos G que contienen M. Entonces tiene lugar el teorema siguiente.

TEOREMA 2. *El conjunto de todas las funciones continuas es siempre denso en* $L_1(R, \mu)$.

DEMOSTRACION. En vista de lo explicado anteriormente, basta demostrar que toda función simple que toma un número finito de valores es límite, en el sentido de la convergencia media, de funciones continuas. Además, como toda función simple que toma un número finito de valores es una combinación lineal de las funciones características $\chi_M(x)$ de los conjuntos medibles, basta realizar la demostración para estas últimas. Sea M un conjunto medible del espacio métrico R. De la condición (2) se desprende inmediatamente que para cualquier $\varepsilon > 0$ existen un conjunto cerrado F_M y un conjunto abierto G_M tales que

$$F_M \subset M \subset G_M \text{ y } \mu(G_M) - \mu(F_M) < \varepsilon.$$

Definamos ahora la función $\varphi_\varepsilon(x)$, tomando[1])

$$\varphi_\varepsilon(x) = \frac{\rho(x, R \setminus G_M)}{\rho(x, R \setminus G_M) + \rho(x, F_M)}.$$

Esta función es igual a 0 cuando $x \in R \setminus G_M$ y es igual a 1 cuando $x \in F_M$. Es continua, ya que cada una de las funciones $\rho(x, F_M)$ y $\rho(x, R \setminus G_M)$ es continua y la suma de ellas nunca se anula. La función $\chi_M - \varphi_\varepsilon$ no pasa de 1 en $G_M \setminus F_M$ y es igual a 0 fuera de este conjunto. Luego,

$$\int |\chi_M(x) - \varphi_\varepsilon(x)| \, d\mu < \varepsilon,$$

de donde se desprende la afirmación del teorema.

---

[1]) $\rho(x, A)$ representa la distancia entre el punto x y el conjunto A.

Está claro que el espacio $L_1(X, \mu)$ depende tanto de la selección del espacio X como de la medida en éste. Por ejemplo, si la medida μ está concentrada en un número finito de puntos, $L_1(X, \mu)$ será un espacio de dimensión finita. En el Análisis desempeñan un papel fundamental los espacios L_1 de dimensión infinita, pero provistos de un subconjunto numerable siempre denso. Para poder describir estos espacios L_1, introduciremos un concepto más correspondiente, en realidad, a la teoría general de la medida.

DEFINICION 2. Se dice que una medida μ *tiene una base numerable* cuando existe un sistema numerable

$$\mathcal{A}=\{A_n\} \quad (n=1, 2, \ldots)$$

de subconjuntos medibles del espacio X (la base numerable de la medida μ) tal que para cualquier medible $M\subset X$ y cualquier $\varepsilon>0$ existe un $A_k\in\mathcal{A}$ tal que

$$\mu(M\triangle A_k)<\varepsilon.$$

En particular, una medida μ tiene, evidentemente, una base numerable, si puede ser representada como prolongación lebesguiana de una medida definida inicialmente en un semianillo numerable $\mathfrak{S}_m$. Efectivamente, el anillo $\mathfrak{R}(\mathfrak{S}_m)$ (obviamente numerable) representa en este caso la base necesaria. De aquí se ve, por ejemplo, que tiene base numerable la medida de Lebesgue de un segmento, ya que para ella se puede tomar como sistema inicial de conjuntos elementales la totalidad de intervalos, segmentos y semisegmentos con extremos racionales.

El producto $\mu=\mu_1\times\mu_2$ de dos medidas de base numerable tiene también base numerable, ya que las sumas finitas de productos de dos en dos de elementos de la base de la medida μ_1 por los elementos de la base de medida μ_2 forman, como se comprueba fácilmente, una base de la medida $\mu=\mu_1\times\mu_2$. Luego, la medida de Lebesgue en el plano (y en un espacio n-dimensional también) tiene base numerable.

Sea

$$A_1^*,\ A_2^*,\ \ldots,\ A_n^*,\ \ldots \tag{3}$$

una base numerable de la medida μ. Es fácil ver que ampliando el sistema de conjuntos (3) se puede formar una base numerable de la medida μ

$$A_1,\ A_2,\ \ldots,\ A_n, \tag{4}$$

cerrada respecto a la sustracción y las uniones e intersecciones finitas.

TEOREMA 3. *Si la medida μ tiene base numerable, existe en $L_1(X, \mu)$ un conjunto numerable de funciones siempre denso*

$$f_1, f_2, \ldots, f_n.$$

DEMOSTRACION. Probemos que las sumas finitas

$$\sum_{k=1}^{n} c_k f_k(x), \tag{5}$$

donde c_k son números racionales y f_k son las funciones características de los elementos de la base numerable de la medida μ, forman un conjunto numerable siempre denso en $L_1(X, \mu)$.

La numerabilidad de este conjunto es evidente; probemos que es siempre denso en $L_1(X, \mu)$. Como hemos visto, el conjunto de funciones escalonadas, que toman sólo un número finito de valores distintos, es siempre denso en L_1. Es obvio que cualquier función de este conjunto puede ser aproximada tanto como se quiera por una función del mismo tipo, pero de valores racionales; por lo tanto, basta demostrar que cualquier función escalonada f, que toma los valores

$$y_1, y_2, \ldots, y_n \quad \text{(todos los } y_i \text{ racionales)}$$

en los conjuntos

$$E_1, E_2, \ldots, E_n \quad \left(\bigcup_i E_i = X, \; E_i \cap E_j = \varnothing \text{ para } i \neq j\right),$$

puede ser aproximada tanto como se quiera, en el sentido de la métrica de L_1, por funciones de tipo (5). Teniendo en cuenta la observación hecha, podemos aceptar, sin perder generalidad, que la base de la medida μ está cerrada respecto a las operaciones de sustracción y uniones e intersecciones finitas.

Por definición de una base numerable de una medida μ, para cualquier $\varepsilon > 0$ existen en ella unos conjuntos $A_1, A_2, \ldots, A_n$ tales que

$$\mu\,[(E_k \setminus A_k) \cup (A_k \setminus E_k)] < \varepsilon.$$

Tomemos

$$A_k' = A_k \setminus \bigcup_{i<k} A_i \quad (k = 1, 2, \ldots, n)$$

y definamos f^* mediante

$$f^*(x) = \begin{cases} y_k & \text{para } x \in A_k' \\ 0 & \text{para } x \in R \setminus \bigcup_{i=1}^{n} A_i. \end{cases}$$

Es fácil ver que para ε suficientemente pequeño la magnitud

$$\mu\{x: f(x) \neq f^*(x)\}$$

es tan pequeña como se quiera y, consecuentemente, la integral

$$\int |f(x) - f^*(x)|\, d\mu \leqslant (2 \max |y_n|)\, \mu\{x: f(x) \neq f^*(x)\}$$

es tan pequeña como se quiera para ε suficientemente pequeño.

En vista de las suposiciones hechas respecto a la base de la medida μ, la función f^* es una función de tipo (5). El teorema queda demostrado.

Para el caso particular en que X es un segmento de la recta numérica y μ es la medida de Lebesgue, la base numerable de $L_1(X, \mu)$ puede ser obtenida también de un modo más clásico: puede ser considerado como base de este tipo, por ejemplo, el conjunto de todos los polinomios de coeficientes racionales. Es siempre denso (incluso en el sentido de convergencia uniforme) en el conjunto de funciones continuas y estas últimas forman un conjunto siempre denso en $L_1(X, \mu)$.

§ 2. ESPACIO L_2

1°. Definición y propiedades fundamentales. El espacio L_1 es, como hemos visto, un espacio lineal normado completo (esto es, un espacio de Banach). Sin embargo, no es euclídeo: la norma definida en él no se puede introducir mediante ningún producto escalar. Esto se deduce del «teorema del paralelogramo» demostrado al final del § 4 del cap. III. En efecto, tomando, por ejemplo, en el segmento $[0, 2\pi]$ las funciones integrables $f \equiv 1$ y $g = \operatorname{sen} x$ vemos que la relación

$$\|f+g\|^2 + \|f-g\|^2 = 2(\|f\|)^2 + \|g\|^2)$$

no se cumple en L_1 para ellas.

Un espacio funcional que, además de ser normado, es euclídeo se puede obtener considerando el conjunto de funciones de cuadrado integrable. Introduzcamos las definiciones correspondientes. Supongamos que se consideran funciones reales f definidas en un espacio X provisto de una medida μ tal que $\mu(X) < \infty$. Todas las funciones se suponen medibles y definidas en casi todo X. Las funciones equivalentes no se distinguen.

DEFINICION 1. Una función f se llama *función de cuadrado integrable* en X cuando la integral

$$\int f^2(x)\, d\mu$$

existe (es finita). El conjunto de todas las funciones de cuadrado integrable en X se designa con $L_2(X, \mu)$ o, brevemente, L_2.

Veamos las propiedades fundamentales de las funciones de cuadrado integrable.

1. *El producto de dos funciones de cuadrado integrable es una función integrable.*

Esto se desprende directamente de la desigualdad

$$|f(x)g(x)| \leqslant \frac{1}{2}[f^2(x)+g^2(x)]$$

y de las propiedades de la integral de Lebesgue.

COROLARIO. *Toda función de cuadrado integrable f es integrable.*
En efecto, basta tomar $g(x) \equiv 1$ y emplear la propiedad 1.

2. *La suma de dos funciones de L_2 es también de L_2.* En efecto,

$$[f(x)+g(x)]^2 \leqslant f^2(x)+2|f(x)g(x)|+g^2(x)$$

y, de acuerdo con la propiedad 1, cada una de las tres funciones que figuran en el miembro derecho es integrable.

3. *Si $f \in L_2$ y α es un número arbitrario, entonces $\alpha f \in L_2$.*
En efecto, si $f \in L_2$ tenemos,

$$\int [\alpha f(x)]^2 d\mu = \alpha^2 \int f^2(x)\, d\mu < \infty.$$

Las propiedades 2 y 3 muestran que las combinaciones lineales de funciones de L_2 son de nuevo elementos de L_2; además, es evidente que la adición de funciones de L_2 y la multiplicación de las mismas por números verifican todas las condiciones de la definición de un espacio lineal (cap. III, § 1). Luego, *el conjunto L_2 de funciones de cuadrado integrable constituye un espacio lineal.*

Definamos ahora el producto escalar en L_2 tomando

$$(f, g) = \int f(x) g(x)\, d\mu.$$

Está claro que todas las condiciones de la definición de un producto escalar (véase el cap. III, § 4), a saber:

1) $(f, g) = (g, f)$,

2) $(f_1+f_2, g) = (f_1, g)+(f_2, g)$,

3) $(\alpha f, g) = \alpha (f, g)$,

4) $(f, f) > 0$ cuando $f \neq 0$,

se cumplen en este caso. En particular, la condición 4) se cumple debido a que hemos convenido no distinguir las funciones equivalentes (por elemento nulo se toma, de esta forma, el conjunto de todas las funciones equivalentes a $f \equiv 0$ en X).

De esta forma, después de definir para las funciones de cuadrado integrable las operaciones de adición y de multiplicación por números y de introducir el producto escalar, llegamos, en conclusión, a la definición siguiente.

DEFINICION 2. Se llama espacio L_2 al espacio euclídeo, cuyos elementos son las clases de funciones equivalentes de cuadrado integrable, en el que las operaciones de adición y multiplicación por números se definen como las operaciones habituales de adición y multiplicación y el producto escalar se define mediante la fórmula

$$(f, g) = \int f(x) g(x) \, d\mu.$$

En L_2, al igual que en cualquier espacio euclídeo, tiene lugar la desigualdad de Cauchy—Buniakovski, que toma en este caso la forma

$$\left(\int f(x) g(x) \, d\mu\right)^2 \leqslant \int f^2(x) \, d\mu \int g^2(x) \, d\mu,$$

y la desigualdad triangular, que toma la forma

$$\sqrt{\int [f(x) + g(x)]^2 \, d\mu} \leqslant \sqrt{\int f^2(x) \, d\mu} + \sqrt{\int g^2(x) \, d\mu}.$$

En particular, para $g(x) \equiv 1$ la desigualdad de Cauchy—Buniakovski se reduce a la siguiente desigualdad útil:

$$\left(\int f(x) \, d\mu\right)^2 \leqslant \mu(X) \int f^2(x) \, d\mu. \tag{1}$$

La norma en L_2 se define por la fórmula

$$\|f\| = \sqrt{(f, f)} = \sqrt{\int f^2(x) \, d\mu}$$

y la distancia entre los elementos f y g, por la fórmula

$$\rho(f, g) = \|f - g\| = \sqrt{\int [f(x) - g(x)]^2 \, d\mu}.$$

La magnitud

$$\int [f(x) - g(x)]^2 \, d\mu = \|f - g\|^2$$

se llama también *desviación cuadrática* entre las funciones f y g.

La convergencia de una sucesión funcional en el sentido de la métrica del espacio L_2 se llama convergencia cuadrática. En el caso en que no haya peligro de confundir este concepto con el de convergencia en L_1, introducido en el parágrafo anterior, emplearemos el término más breve «convergencia media».

TEOREMA 1. *El espacio L_2 es completo.*

DEMOSTRACION. Sea $\{f_n\}$ una sucesión fundamental de L_2, es decir,

$$\|f_n - f_m\| \to 0 \text{ para } n, m \to \infty.$$

De acuerdo con la desigualdad (1), tenemos entonces

$$\int |f_n(x) - f_m(x)|\, d\mu \leqslant [\mu(X)]^{1/2} \left\{ \int [f_n(x) - f_m(x)]^2\, d\mu \right\}^{1/2} \leqslant \varepsilon [\mu(X)]^{1/2}, \qquad (2)$$

esto es, la sucesión $\{f_n\}$ es fundamental también respecto a la métrica del espacio L_1. Repitiendo los razonamientos que hemos empleado al demostrar la complitud del espacio L_1, podemos escoger de $\{f_n\}$ una sucesión parcial $\{f_{n_k}\}$ que converge en casi todos los puntos hacia una función f. En la desigualdad

$$\int [f_{n_k}(x) - f_{n_l}(x)]^2\, d\mu < \varepsilon$$

que es válida para elementos de esta sucesión parcial con k y l suficientemente grandes, podemos pasar, en virtud del teorema de Fatou, al límite para $l \longrightarrow \infty$. Tendremos

$$\int [f_{n_k}(x) - f(x)]^2\, d\mu \leqslant \varepsilon$$

de donde se deduce que $f \in L_2$ y que $f_{n_k} \longrightarrow f$. Para terminar la demostración basta, al igual que en el teorema 1 del § 1, señalar que toda sucesión fundamental que contiene una sucesión parcial convergente converge hacia el mismo límite.

EJERCICIO. Definamos $L_p(X, \mu)$ como el conjunto de las clases de funciones equivalentes para las cuales $\int |f|^p\, d\mu < \infty$, donde $1 \leqslant p < \infty$. Demuéstrese que $L_p(X, \mu)$ es un espacio de Banach respecto a la norma $\|f\| = \left(\int |f|^p\, d\mu\right)^{1/p}$.

2°. Caso de medida infinita. En el punto anterior hemos considerado funciones de cuadrado integrable definidas en un espacio X de medida finita. La condición $\mu(X) < \infty$ ha sido empleada de un modo sustancial. La hemos empleado, primero, al demostrar que toda función de cuadrado sumable es sumable en primer grado y, después, al deducir la desigualdad (2) en la que se basa la demostración de la complitud del espacio L_2. Si se consideran funciones en un conjunto de medida infinita (por ejemplo, en toda la recta con la medida lebesguiana en ella), no toda función de L_2 será elemento de L_1. Por ejemplo, la función $\frac{1}{\sqrt{1+x^2}}$ no es integrable en toda la recta, pero su cuadrado es integrable. Además, en el caso $\mu(X)<\infty$ tiene lugar la desigualdad (1) según la cual la convergencia de una sucesión de funciones en L_2 implica su convergencia en L_1. Cuando $\mu(X)=\infty$, esto ya no tiene lugar: por ejemplo, la sucesión de funciones en la recta

$$\varphi_n(x)=\begin{cases}\frac{1}{n} & \text{para } |x|\leqslant n,\\ 0 & \text{para } |x|>n\end{cases}$$

converge hacia el 0 en el espacio $L_2(-\infty, \infty)$ de funciones de cuadrado integrable en la recta, pero no converge hacia ningún límite en $L_1(-\infty, \infty)$. Sin embargo, *el teorema sobre la complitud del espacio* L_2 *sigue siendo válido también para* $\mu(X)=\infty$ [1].

Demostremos esta afirmación. Vamos a suponer, lo mismo que en el § 5 del cap. VI, donde hemos introducido el concepto de integral en un conjunto de medida infinita, que todo el espacio X puede ser representado como la unión numerable de conjuntos de medida finita. Sea

$$X=\bigcup_{n=1}^{\infty} X_n, \quad \mu(X_n)<\infty, \quad X_n\cap X_m=\varnothing \text{ para } n\neq m$$

una representación de este tipo y sea $\{f_n\}$ una sucesión fundamental en $L_2(X, \mu)$. Entonces, para cada $\varepsilon>0$ existe un N tal que

$$\int [f_k(x)-f_l(x)]^2\,d\mu<\varepsilon \text{ para todos los } k,\ l\geqslant N.$$

Tomemos

$$\varphi^{(n)}(x)=\begin{cases}\varphi(x) & \text{para } x\in X_n\\ 0 & \text{para los demás } x.\end{cases}$$

Entonces, debido a la σ-aditividad de la integral de Lebesgue, tenemos

$$\int [f_k(x)-f_l(x)]^2\,d\mu=\sum_{n=1}^{\infty}\int_{X_n}[f_k^{(n)}(x)-f_l^{(n)}(x)]\,d\mu<\varepsilon.$$

Consecuentemente, para cada M finito es, con mayor razón,

$$\sum_{n=1}^{M}\int_{X_n}[f_k^{(n)}(x)-f_l^{(n)}(x)]^2\,d\mu<\varepsilon. \tag{3}$$

---

[1] La demostración de la complitud del espacio L_1, expuesta en el § 1, no depende, evidentemente, de si es o no finita la medida del espacio X.

El conjunto de funciones de cuadrado integrable en cada X_n es un espacio completo. Tomando

$$f^{(n)}(x) = \lim_{l \to \infty} f_l^{(n)}(x)$$

(donde la convergencia se entiende como convergencia en el espacio $L_2(X_n, \mu)$), podemos pasar al límite para $l \to \infty$ en la desigualdad (3). Tendremos

$$\sum_{n=1}^{M} \int_{X_n} [f_k^{(n)}(x) - f^{(n)}(x)]^2 \, d\mu \leqslant \varepsilon.$$

Puesto que esta desigualdad se cumple para todos los M, podemos pasar en ella al límite para $M \to \infty$. De este modo, tenemos

$$\sum_{n=1}^{\infty} \int_{X_n} [f_k^{(n)}(x) - f^{(n)}(x)]^2 \, d\mu \leqslant \varepsilon.$$

Tomando

$$f(x) = f^{(n)}(x) \quad \text{para} \quad x \in X_n,$$

podemos dar a esta última desigualdad la forma

$$\int [f_k(x) - f(x)]^2 \, d\mu \leqslant \varepsilon.$$

De aquí se deduce tanto que f es un elemento de $L_2(X, \mu)$, como la convergencia de la sucesión $\{f_k\}$ hacia f.

3°. Conjuntos siempre densos en L_2. Teorema sobre el isomorfismo. Así pues, el espacio $L_2(X, \mu)$ de funciones de cuadrado integrable es un espacio euclídeo completo. Excepto casos triviales, la dimensión de este espacio es infinita. Desde el punto de vista de diversas aplicaciones en el Análisis, es importante conocer cuándo el espacio $L_2(X, \mu)$ contiene un conjunto numerable siempre denso. Hemos visto en el § 1 que en el caso del espacio $L_1(X, \mu)$ la existencia de un conjunto numerable siempre denso se desprende de la existencia de una base numerable de la medida μ. No es difícil comprobar que esta misma condición garantiza también la existencia de un conjunto numerable siempre denso en $L_2(X, \mu)$. En efecto, toda función de $L_2(X, \mu)$ puede ser aproximada con precisión necesaria mediante funciones cada una de las cuales es igual al 0 fuera de un conjunto de medida finita[1]. Después, los mismos razonamientos que han sido empleados al demostrar el teorema 3 del § 1 muestran que

[1] Si es $\mu(X) < \infty$, este paso sobra.

en el conjunto de funciones de este tipo se puede escoger un conjunto numerable siempre denso.

Luego, si la medida μ tiene base numerable, el espacio $L_2(X, \mu)$ es un espacio euclídeo completo provisto de un conjunto numerable siempre denso. En otras palabras, dejando a un lado el caso en que $L_2(X, \mu)$ es de medida finita, obtenemos el resultado siguiente: *si la medida μ es de base numerable, el espacio $L_2(X, \mu)$ es de Hilbert.*

En virtud del teorema sobre el isomorfismo de los espacios de Hilbert, esto significa que todos los espacios $L_2(X, \mu)$ de este tipo son isomorfos. En particular, todo espacio de este tipo $L_2(X, \mu)$ es isomorfo al espacio l_2 de sucesiones numéricas con la suma de cuadrados convergente (que puede ser considerado como el espacio $L_2(X, \mu)$ correspondiente a la medida μ definida en una sucesión numerable de puntos). *En lo que sigue sólo se considerarán espacios $L_2(X, \mu)$ correspondientes a medidas de base numerable.* En los casos en que no puedan surgir confusiones cada espacio de este tipo se designará simplemente mediante L_2.

Como el espacio L_2 representa, según lo explicado, una realización del espacio de Hilbert, se pueden extender a L_2 todos los conceptos y resultados dados en el § 4 del cap. III para un espacio de Hilbert abstracto.

En particular, como, de acuerdo con el teorema de Riesz, toda funcional lineal en el espacio de Hilbert H puede ser representada mediante el producto escalar

$$F(h) = (h, a),$$

donde a es un vector fijo de H, *toda funcional lineal en L_2 es de la forma*

$$F(f) = \int f(x)\, g(x)\, d\mu$$

donde g es una función fijada de cuadrado integrable en X.

4°. Espacio complejo L_2. Hemos considerado hasta aquí el espacio real L_2. Los resultados expuestos se extienden sin dificultad al caso complejo. Una función compleja f definida en un espacio X provisto de una medida μ, se llama función de cuadrado integrable cuando la integral

$$\int_X |f(x)|^2\, d\mu$$

es finita. Definiendo del modo habitual la adición de estas funciones y la multiplicación de las mismas por números e introduciendo el producto escalar mediante la fórmula

$$(f, g) = \int_X f(x)\overline{g(x)}\,d\mu,$$

obtenemos un espacio euclídeo llamado espacio complejo L_2. (Aquí, lo mismo que en el caso real, consideramos las funciones equivalentes como un mismo elemento del espacio). Este espacio es completo y, además, si la medida μ es de base numerable, es también separable. Luego (omitiendo el caso en que este espacio es de medida finita), obtenemos que el espacio complejo L_2, correspondiente a una medida de base numerable, es el espacio complejo de Hilbert. Todos los espacios de este tipo son isomorfos y para ellos son válidos los resultados expuestos en el § 4 del cap. III.

5°. **Convergencia cuadrática y su relación con otros tipos de convergencia de sucesiones funcionales.** Al introducir en el espacio L_2 la norma, hemos definido con ello para las funciones de cuadrado integrable el siguiente concepto de convergencia:

$$f_n \to f,$$

cuando

$$\lim_{n\to\infty} \int [f_n(x) - f(x)]^2\,d\mu = 0.$$

Esta convergencia la hemos llamado convergencia cuadrática. Veamos cómo está relacionada con otros tipos de convergencia de sucesiones funcionales. Supongamos primero que la medida del espacio X en el que están definidas las funciones es finita.

1. *Si la sucesión $\{f_n\}$ de funciones de $L_2(X, \mu)$ converge en la métrica de $L_2(X, \mu)$, también converge en la métrica de $L_1(X, \mu)$.*

En efecto, debido a la desigualdad (1), tenemos

$$\int |f_n(x) - f(x)|\,d\mu \leqslant \left[\mu(X)\int (f_n(x) - f(x))^2\,d\mu\right]^{\frac{1}{2}}$$

y de aquí se desprende nuestra afirmación.

2. *Si la sucesión $\{f_n\}$ converge uniformemente, también converge cuadráticamente.*

En efecto, cualquiera que sea $\varepsilon > 0$, tenemos para n suficientemente grandes

$$|f_n(x) - f(x)| < \varepsilon$$

y, consecuentemente,

$$\int |f_n(x) - f(x)|^2\,d\mu < \varepsilon^2\mu(X),$$

de donde se deduce nuestra afirmación.

3. *Si una sucesión* $\{f_n\}$ *de funciones sumables converge en media, también converge en medida en* X.

Esta afirmación se desprende directamente de la desigualdad de Chébishev.

De aquí y del teorema 11 del § 4 del cap. VI, se sigue:

4. *Si una sucesión* $\{f_n\}$ *converge en media, se puede extraer de ella una sucesión parcial* $\{f_{n_k}\}$ *convergente en casi todos los puntos.*

Observemos que al demostrar el teorema sobre la complitud del espacio L_1 hemos encontrado este resultado sin basarnos en el teorema 11 del § 4 del cap. VI.

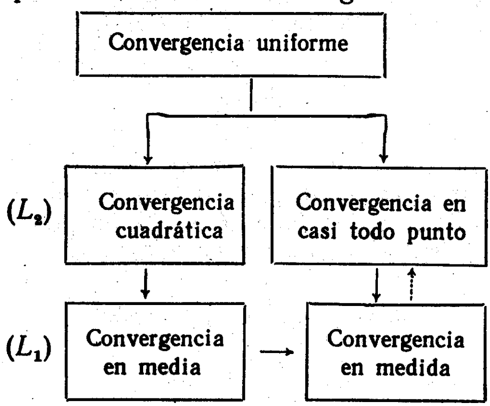
Es fácil ver que la convergencia media (e incluso cuadrática) de una sucesión no implica, en general, la convergencia de esta sucesión en casi todos los puntos. Efectivamente, la sucesión $\{f_{n_k}\}$, construida en el § 4 del cap. VI, converge en media y cuadráticamente hacia $f\equiv 0$; pero, al mismo tiempo, como ha sido allí demostrado, no tiende hacia 0 en nungún punto. Viceversa, una sucesión $\{f_n\}$ puede converger en casi todos los puntos (e incluso en todo punto) y no converger en media. Consideremos, por ejemplo, en el segmento $[0, 1]$ la sucesión de funciones

$$f_n(x)=\begin{cases} n \text{ para } x\in\left(0,\ \frac{1}{n}\right) \\ 0 \text{ para los demás } x, \end{cases}$$

tal que $f_n(x)\longrightarrow 0$ para todo $x\in[0, 1]$. Al mismo tiempo,

$$\int_0^1 |f_n(x)|\,dx=1 \text{ para todo } n.$$

Las relaciones que existen entre diferentes tipos de convergencia de funciones, definidas en un espacio de medida finita, pueden ser esquematizadas del modo siguiente:



donde la flecha de puntos significa la posibilidad de escoger de una sucesión convergente en medida una sucesión parcial convergente en casi todo punto.

En el caso $\mu(X)=\infty$ (por ejemplo, para funciones en toda la recta numérica con la medida de Lebesgue en ella) las relaciones encontradas no tienen ya lugar. Por ejemplo, la sucesión

$$f_n(x)=\begin{cases} \frac{1}{\sqrt{n}} & \text{para } |x|\leqslant n, \\ 0 & \text{para } |x|>n, \end{cases}$$

converge uniformemente en toda la recta hacia la función $f\equiv 0$ y, sin embargo, no converge ni en media ni cuadráticamente. Además, para $\mu(X)=\infty$ la convergencia cuadrática (esto es, en L_2) no implica, según hemos señalado ya, la convergencia media (esto es, en L_1) de la misma sucesión.

A su vez, la convergencia media no implica, en general, la convergencia cuadrática (esta última observación es válida tanto cuando $\mu(X)<\infty$, como cuando $\mu(X)=\infty$).

§ 3. SISTEMAS ORTOGONALES DE FUNCIONES EN L_2. SERIES RESPECTO A SISTEMAS ÓRTOGONALES

De los teoremas generales, demostrados en el § 4 del cap. III para los espacios euclídeos, se desprende que en L_2 existen sistemas completos ortogonales (en particular, ortogonales y normales) de funciones. Estos sistemas pueden ser obtenidos, por ejemplo, aplicando el proceso de ortogonalización, descrito en el § 4 del cap. III, a uno u otro sistema completo. Si en L_2 se ha escogido un sistema $\{\varphi_n\}$ completo ortogonal, todo elemento $f\in L_2$ puede ser representado, en vista también de los resultados del § 4 del cap. III, como la suma de la serie

$$f=\sum_{n=1}^{\infty} c_n\varphi_n$$

esto es, como la suma de la serie de Fourier de la función f respecto al sistema ortogonal $\{\varphi_n\}$. Además, los coeficientes c_n, es decir, los coeficientes de Fourier de la función f respecto al sistema $\{\varphi_n\}$, se definen mediante las fórmulas

$$c_n=\frac{1}{\|\varphi_n\|^2}\int f(x)\,\varphi_n(x)\,d\mu$$

$$\left(\|\varphi_n\|^2=\int \varphi_n^2(x)\,d\mu\right).$$

En este parágrafo consideraremos algunos ejemplos de mayor importancia de sistemas ortogonales en el espacio L_2 y los desarrollos que les corresponden.

1°. Sistema trigonométrico. Serie trigonométrica de Fourier. Consideremos el espacio $L_2(-\pi, \pi)$ de funciones de cuadrado integrable en el segmento $[-\pi, \pi]$ con la medida de Lebesgue habitual en este segmento. En este espacio las funciones

$$\{\cos nx, \operatorname{sen} nx\} \quad (n=0, 1, 2, \ldots) \tag{1}$$

forman un sistema completo ortogonal, llamado *sistema trigonométrico*. La ortogonalidad de este sistema se comprueba fácilmente mediante el cálculo directo; por ejemplo, para $n \neq m$

$$\int_{-\pi}^{\pi} \cos nx \cos mx\, dx = \frac{1}{2} \int_{-\pi}^{\pi} \left[\cos \frac{n+m}{2} x + \cos \frac{n-m}{2} x\right] dx = 0,$$

etc. La complitud del sistema (1) se desprende del teorema de Weierstrass sobre la aproximación de cualquier función periódica continua mediante polinomios trigonométricos[1]. El sistema (1) no es normal. El sistema normal correspondiente está compuesto por las funciones

$$\frac{1}{\sqrt{2\pi}}, \quad \frac{\cos nx}{\sqrt{\pi}}, \quad \frac{\operatorname{sen} nx}{\sqrt{\pi}} \quad (n=1, 2, \ldots).$$

Sea f una función de $L_2(-\pi, \pi)$; sus coeficientes de Fourier, correspondientes a las funciones 1, $\cos nx$ y $\operatorname{sen} nx$, se acostumbra designar con $\frac{a_0}{2}$, a_n y b_n. Por lo tanto, de acuerdo con las fórmulas generales para los coeficientes de Fourier, tenemos

$$\frac{a_0}{2} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\, dx, \text{ es decir, } a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)\, dx$$

y

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx\, dx, \quad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \operatorname{sen} nx\, dx.$$

[1] En el § 2 del cap. IX demostraremos el teorema de Fejér que constituye una generalización del teorema de Weierstrass. Con ello daremos una demostración de la complitud del sistema trigonométrico (demostración que no se basa, claro está, en los resultados que aquí exponemos).

La serie correspondiente de Fourier es

$$\frac{a_0}{2}+\sum_{n=1}^{\infty} a_n \cos nx + b_n \operatorname{sen} nx$$

y cualquiera que sea $f \in L_2$ converge cuadráticamente hacia esta función. Si

$$S_n=\frac{a_0}{2}+\sum_{k=1}^{n} a_k \cos kx + b_k \operatorname{sen} kx$$

es la suma parcial de la serie de Fourier, la desviación cuadrática entre S_n y f puede ser encontrada mediante la fórmula

$$\| f(x)-S_n(x)\|^2=\| f\|^2-\pi\left(\frac{a_0^2}{2}+\sum_{k=1}^{n} a_k^2+b_k^2\right).$$

Entre todos los polinomios trigonométricos

$$T_n(x)=\frac{\alpha_0}{2}+\sum_{k=1}^{n} \alpha_k \cos kx + \beta_k \operatorname{sen} kx$$

con n fijo la suma parcial S_n de la serie de Fourier es la que mejor aproxima (en la métrica de L_2) la función f. Para el sistema trigonométrico la desigualdad de Bessel da

$$\frac{a_0}{2}+\sum_{n=1}^{\infty} a_n^2+b_n^2 \leqslant \frac{1}{\pi}\int_{-\pi}^{\pi} f^2(x)\,dx.$$

Como el sistema trigonométrico es completo, para cualquier función de L_2 tiene lugar, de hecho, la igualdad de Parseval

$$\frac{a_0}{2}+\sum_{n=1}^{\infty} a_n^2+b_n^2 = \frac{1}{\pi}\int_{-\pi}^{\pi} f^2(x)\,dx.$$

Para cualquier función $f \in L_2$ los cuadrados de sus coeficientes de Fourier forman una serie convergente. Viceversa, si los números a_0, a_n, b_n $(n=1, 2, \ldots)$ son tales que la serie $\sum a_n^2+b_n^2$ converge, la serie

$$\frac{a_0}{2}+\sum_{n=1}^{\infty} a_n \cos nx + b_n \operatorname{sen} nx$$

también converge (en L_2) y su suma es una función para la cual a_0, a_n y b_n son sus coeficientes de Fourier.

Todo lo que se acaba de exponer para funciones definidas en el segmento $[-\pi, \pi]$, se extiende fácilmente a funciones definidas en un segmento de longitud arbitraria, digamos en $[-l, l]$. Si f es una función de cuadrado integrable en $[-l, l]$, la sustitución $x=\frac{\pi t}{l}$, es decir, $t=\frac{lx}{\pi}$, convierte $f(t)$ en la función $f^*(x)=f\frac{lx}{\pi}$ en el segmento $[-\pi, \pi]$.

De acuerdo con esto

$$a_n=\frac{1}{l}\int_{-l}^{l} f(t)\cos\frac{n\pi t}{l}\,dt$$

y

$$b_n=\frac{1}{l}\int_{-l}^{l} f(t)\operatorname{sen}\frac{n\pi t}{l}\,dt.$$

La serie de Fourier para una función f definida en un segmento de longitud $2l$ es

$$\frac{a_0}{2}+\sum_{n=1}^{\infty} a_n\cos\frac{n\pi t}{l}+b_n\operatorname{sen}\frac{n\pi t}{l}.$$

Observaciones 1. La teoría de series trigonométricas fue elaborada, en gran medida, en las obras de J. Fourier, matemático francés, relacionadas con sus investigaciones en la Física Matemática y, en primer término, en la teoría de propagación del calor. No obstante, las fórmulas para los coeficientes a_n y b_n aparecen ya en los trabajos de Euler. Inicialmente los términos «serie de Fourier», «coeficientes de Fourier», etc., se relacionaban precisamente con el sistema trigonométrico ortogonal y solamente mucho más tarde empezaron a usarse en ese sentido general en el que los hemos empleado en el § 4 del cap. III (esto es, para un sistema ortogonal arbitrario en un espacio euclídeo cualquiera).

2. De la complitud del sistema trigonométrico y de los teoremas generales demostrados en el § 4 del cap. III se desprende que cualquiera que sea $f\in L_2$ su serie de Fourier

$$\frac{a_0}{2}+\sum_{n=1}^{\infty} a_n\cos nx+b_n\operatorname{sen} nx$$

converge en media hacia la función dada f. Pero, desde el punto de vista de los problemas concretos del Análisis es importante encontrar las condiciones en que esta serie converge hacia f en otros sentidos, digamos en cada punto o uniformemente. Estas cuestiones serán consideradas en el capítulo siguiente.

2°. Sistemas trigonométricos en el segmento $[0, \pi]$. Las funciones

$$1, \cos x, \cos 2x, \ldots \tag{2}$$

y

$$\operatorname{sen} x, \operatorname{sen} 2x, \ldots \tag{3}$$

forman en su conjunto un sistema ortogonal completo en el segmento $[-\pi, \pi]$. Probemos que cada uno de los sistemas (2) y (3) es ortogonal y completo en el segmento $[0, \pi]$. La ortogonalidad se comprueba mediante el cálculo directo. Demostremos la complitud del sistema (2). Sea f una función de cuadrado integrable en $[0, \pi]$. Definámosla en el segmento $[-\pi, 0]$, tomando

$$f(-x) = f(x)$$

y desarrollémosla en serie de Fourier respecto al sistema

$$1, \cos nx, \operatorname{sen} nx \qquad (n = 1, 2, \ldots).$$

Como la función f, definida ahora en $[-\pi, \pi]$, es par, todos sus coeficientes en los senos son iguales a cero; esto se ve inmediatamente de las fórmulas para los coeficientes: para una función par f y $n \geqslant 1$ tenemos

$$\int_{-\pi}^{\pi} f(x) \operatorname{sen} nx\, dx = \int_{-\pi}^{0} f(x) \operatorname{sen} nx\, dx + \int_{0}^{\pi} f(x) \operatorname{sen} nx\, dx =$$

$$= -\int_{0}^{\pi} f(x) \operatorname{sen} nx\, dx + \int_{0}^{\pi} f(x) \operatorname{sen} nx\, dx = 0.$$

En otras palabras, esta función puede ser aproximada en media en $[-\pi, \pi]$ (y con mayor razón en $[0, \pi]$), con precisión arbitraria, mediante combinaciones lineales de los elementos del sistema (2). De aquí se deduce la complitud del sistema (2). La complitud en $[0, \pi]$ del sistema (3) se demuestra análogamente, prolongando al semisegmento $[-\pi, 0]$ la función $f(x)$, definida en $[0, \pi]$, mediante la fórmula

$$f(-x) = -f(x).$$

La función obtenida de esta forma es impar en $[-\pi, \pi]$ y se desarrolla en este segmento en serie solamente respecto a los senos.

3°. Forma compleja de la serie de Fourier. La serie trigonométrica de Fourier de una función f en el segmento $[-\pi, \pi]$ puede ser representada en una forma más compacta, si se emplean las fórmulas de Euler

$$\cos nx = \frac{e^{inx} + e^{-inx}}{2}, \quad \operatorname{sen} nx = \frac{e^{inx} - e^{-inx}}{2i}.$$

Colocando estas expresiones en la serie de Fourier, tenemos

$$\frac{a_0}{2}+\sum_{n=1}^{\infty} a_n \cos nx + b_n \operatorname{sen} nx = \frac{a_0}{2} + \\ + \sum_{n=1}^{\infty}\left(a_n \frac{e^{inx}+e^{-inx}}{2} - ib_n \frac{e^{inx}-e^{-inx}}{2}\right) = \\ = \frac{a_0}{2} + \sum_{n=1}^{\infty} \frac{a_n - ib_n}{2} e^{inx} + \sum_{n=1}^{\infty} \frac{a_n+ib_n}{2} e^{-inx} = \sum_{n=-\infty}^{\infty} c_n e^{inx},$$

donde $c_0 = \frac{a_0}{2}$ y para $n \geqslant 1$

$$\left.\begin{aligned} c_n &= \frac{a_n - ib_n}{2}, \\ c_{-n} &= \frac{a_n + ib_n}{2}. \end{aligned}\right\} \tag{4}$$

La expresión

$$\sum_{n=-\infty}^{\infty} c_n e^{inx}$$

se llama serie trigonométrica de Fourier en forma compleja. Los coeficientes c_n de esta serie se expresan por las fórmulas (4) a través de a_n y b_n; sin embargo, es fácil obtener unas fórmulas para poder calcularlos directamente. En efecto, el cálculo inmediato da

$$\int_{-\pi}^{\pi} e^{inx} \cdot e^{-imx}\, dx = \begin{cases} 0 & \text{para } n \neq m, \\ 2\pi & \text{para } n = m. \end{cases}$$

Luego, multiplicando la igualdad

$$f(x) = \sum_{n=-\infty} c_n e^{inx} \tag{5}$$

por e^{-imx} $(m = 0, \pm 1, \pm 2, \ldots)$ e integrándola, obtenemos

$$\int_{-\pi}^{\pi} f(x)\, e^{-imx}\, dx = 2\pi c_m,$$

es decir,

$$c_m = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\, e^{-imx}\, dx \quad (m = 0, \pm 1, \pm 2, \ldots). \tag{6}$$

El desarrollo (5) subsiste para las funciones complejas de cuadrado integrable en el segmento $[-\pi, \pi]$. En otras palabras, las funciones e^{inx} constituyen una base del espacio $L_2[-\pi, \pi]$ de funciones complejas con el cuadrado integrable del valor absoluto en $[-\pi, \pi]$. Además, las expresiones (6) representan los productos escalares de f por e^{imx} en este espacio complejo.

Está claro que, sustituyendo e^{inx} por $e^{i\frac{n\pi}{l}x}$, todo lo expuesto puede ser extendido al espacio $L_2(-l, l)$ de funciones complejas en un segmento de longitud arbitraria $2l$.

4°. Polinomios de Legendre. Las combinaciones lineales de las funciones

$$1, \ x, \ x^2, \ \ldots \tag{7}$$

constituyen el conjunto de polinomios. Luego, el sistema (7) es completo en el espacio L_2 de funciones en un segmento[1]. Ortogonalizando el sistema (7) en el segmento $[-1, 1]$, es decir, respecto al producto escalar

$$(f, g) = \int_{-1}^{1} f(x)\, g(x)\, dx,$$

obtenemos un sistema ortogonal completo

$$Q_0(x), \ Q_1(x), \ Q_2(x), \ \ldots,$$

donde Q_n es un polinomio de grado n. Probemos que cada uno de los polinomios que se obtienen al ortogonalizar el sistema (7) coincide, salvo un factor constante, con el polinomio

$$P_n(x) = \frac{d^n}{dx^n}(x^2-1)^n.$$

Probemos que el sistema $\{P_n\}$ es ortogonal. Sea $n \geqslant m$. Como

$$\left.\frac{d^k}{dx^k}(x^2-1)^n\right|_{x=-1} = \left.\frac{d^k}{dx^k}(x^2-1)^n\right|_{x=1} = 0$$

para todos los $k=0, 1, \ldots, n-1$, obtenemos, integrando por partes,

$$\int_{-1}^{1} P_m(x)\, P_n(x)\, dx = -\int_{-1}^{1} \frac{d^{m+1}}{dx^{m+1}}(x^2-1)^m \frac{d^{n-1}}{dx^{n-1}}(x^2-1)^n = \ldots$$

$$\ldots = (-1)^n \int_{-1}^{1} \frac{d^{m+n}}{dx^{m+n}}(x^2-1)^m (x^2-1)^n\, dx. \tag{8}$$

---

[1] La complitud del sistema de polinomios en el espacio $L_2[a, b]$ de funciones de cuadrado integrable en un segmento cualquiera $[a, b]$ se desprende del teorema de Weierstrass sobre la aproximación uniforme de cualquier función continua en un segmento mediante polinomios. Véase el final del § 2 del capítulo IX.

Si $m < n$, bajo el signo de la última integral figura el cero idéntico y de aquí se desprende la ortogonalidad del sistema $\{P_n\}$. Además, está claro que el polinomio P_n es de grado n, es decir, todo P_n se encuentra en el subespacio generado por los $n+1$ primeros elementos del sistema (7). Luego, tanto el sistema $\{P_n\}$, como el sistema $\{Q_n\}$, poseen las propiedades siguientes:

1) ortogonalidad,

2) el n-ésimo elemento del sistema pertenece al subespacio generado por los elementos $1, x, \ldots, x^{n-1}$.

Pero cada elemento del sistema queda determinado por estas propiedades unívocamente, salvo un factor constante (teorema 1 del § 4 del cap. III).

En el caso $m=n$ la igualdad (8) lleva al resultado siguiente

$$\int_{-1}^{1} P_n^2(x)\,dx = \int_{-1}^{1} \frac{d^{2n}}{dx^{2n}}(x^2-1)^n (x^2-1)^n\,dx =$$

$$= (2n!) \int_{-1}^{1} (x^2-1)^n\,dx = \frac{(n!)^2 \cdot 2^{2n+1}}{2n+1}.$$

En otras palabras, la norma del polinomio P_n es igual a

$$\frac{n!\,2^n\sqrt{2}}{\sqrt{2n+1}}.$$

Por lo tanto, el sistema de polinomios

$$\frac{\sqrt{2n+1}}{n!\,2^n\sqrt{2}}\,P_n,$$

además de ser ortogonal, es también normal.

En lugar de estos polinomios normalizados suelen considerarse los polinomios definidos por la fórmula

$$L_n(x) = \frac{1}{n!\,2^n}\frac{d^n}{dx^n}(x^2-1)^n$$

llamados *polinomios de Legendre*. De los cálculos realizados se deduce que

$$\int_{-1}^{1} L_n(x) L_m(x)\,dx = \begin{cases} 0 & \text{para } n \neq m, \\ \dfrac{2}{2n+1} & \text{para } n = m. \end{cases}$$

Señalemos en forma explícita los cinco primeros polinomios de Legendre:

$$L_0(x) = 1, \quad L_1(x) = x, \quad L_2(x) = \frac{3}{2}x^2 - \frac{1}{2},$$
$$L_3(x) = \frac{5}{2}x^3 - \frac{3}{2}x, \quad L_4(x) = \frac{35}{8}x^4 - \frac{15}{4}x^2 + \frac{3}{8}.$$

El desarrollo de una función f en el segmento $[-1, 1]$ respecto a los polinomios de Legendre es de la forma

$$f(x) = \sum_{n=0}^{\infty} c_n L_n(x),$$

donde

$$c_n = \frac{2n+1}{2} \int_{-1}^{1} f(x) L_n(x)\, dx.$$

5°. Sistemas ortogonales en productos. Series múltiples de Fourier. Sean definidas en los conjuntos X' y X'' las medidas μ' y μ''. Designemos mediante L_2' y L_2'' los espacios correspondientes de funciones de cuadrado integrable. Consideremos en el producto

$$X = X' \times X''$$

la medida

$$\mu = \mu' \times \mu''.$$

Designemos mediante L_2 el espacio correspondiente de funciones de cuadrado integrable. Interpretaremos las funciones de L_2 como funciones de dos variables.

TEOREMA 1. *Si $\{\varphi_m\}$ y $\{\psi_n\}$ son sistemas ortonormales de L_2' y L_2'', respectivamente, el sistema de todos los productos*

$$f_{mn}(x, y) = \varphi_m(x)\psi_n(y)$$

es un sistema ortonormal completo de L_2.

La demostración de ortonormalidad es muy sencilla:

1. $\displaystyle\int_X f_{mn}^2(x, y)\, d\mu = \int_{X'} \varphi_m^2(x) \left(\int_{X''} \psi_n^2(y)\, d\mu'' \right) d\mu' = 1.$

2. Si $m \neq m_1$, tenemos

$$\int_X f_{mn}(x, y) f_{m_1 n_1}(x, y)\, d\mu =$$
$$= \int_{X''} \psi_n(y)\psi_{n_1}(y) \left(\int_{X'} \varphi_m(x)\varphi_{m_1}(x)\, d\mu' \right) d\mu'' = 0.$$

3. Si $m=m_1$, pero $n\neq n_1$, tenemos

$$\int_X f_{mn}(x, y) f_{m_1n_1}(x, y)\,d\mu=$$

$$=\int_{X'} \varphi_m^2(x) \left(\int_{X''} \psi_n(y)\psi_{n_1}(y)\,d\mu''\right) d\mu'=0.$$

Probemos la complitud del sistema $\{f_{mn}\}$. Supongamos que en L_2 existe una función f ortogonal a todas las funciones f_{mn}. Tomemos

$$F_m(y)=\int_{X'} f(x, y)\varphi_m(x)\,d\mu'.$$

En este caso, cualquiera que sea n,

$$\int_{X''} F_m(y)\psi_n(y)\,d\mu''=\int_X f(x, y) f_{mn}(x, y)\,d\mu=0.$$

Debido a la complitud del sistema $\{\psi_n\}$, de aquí se desprende que

$$F_m(y)=0$$

para casi todo y. Luego, para casi todo y tienen lugar las igualdades

$$\int_{X'} f(x, y)\varphi_m(x)\,d\mu'=0$$

cualquiera que sea m. Debido a la complitud del sistema $\{\varphi_m\}$, obtenemos de aquí que para casi todo y el conjunto de aquellos x en los que

$$f(x, y)\neq 0$$

es de medida nula. En virtud del teorema de Fubini, esto significa que la función $f(x, y)$ es igual a 0 en casi todo el X. El teorema queda demostrado.

Apliquemos este teorema a algunos sistemas ortogonales concretos. En el espacio de funciones de cuadrado integrable de dos variables

$$f(x, y) \quad (-\pi\leqslant x\leqslant\pi, \ -\pi\leqslant y\leqslant\pi)$$

constituyen un sistema ortogonal completo los productos de dos en dos de los elementos de los sistemas:

$$1, \cos mx, \operatorname{sen} mx \quad (m=1, 2, \ldots)$$

y

$$1, \cos ny, \operatorname{sen} ny \quad (n=1, 2, \ldots),$$

es decir, las funciones

$$1,\ \cos mx,\ \operatorname{sen} mx,\ \cos ny,\ \operatorname{sen} ny,\ \cos mx\ \operatorname{sen} ny,$$
$$\cos mx \cos ny,\ \operatorname{sen} mx \operatorname{sen} ny,\ \operatorname{sen} mx \cos ny.$$

La expresión de la serie correspondiente de Fourier es, en cierto grado, voluminosa y por eso conviene recurrir aquí a las funciones exponenciales de argumento imaginario, esto es, a las funciones

$$e^{imx}\cdot e^{iny} = e^{i(mx+ny)} \quad (n,\ m = 0,\ \pm 1,\ \pm 2,\ \ldots).$$

A esta base corresponde la serie de Fourier

$$f(x,\ y) = \sum_{n,\ m=-\infty}^{\infty} C_{mn} e^{i(mx+ny)},$$

donde

$$C_{mn} = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f(x,\ y)\, e^{-i(mx+ny)}\, dx\, dy.$$

Empleando los polinomios de Legendre, podemos obtener en el espacio de funciones, definidas en el cuadrado

$$-1 \leqslant x \leqslant 1, \quad -1 \leqslant y \leqslant 1,$$

un sistema ortonormal completo compuesto por los polinomios

$$Q_{mn}(x,\ y) = \frac{\sqrt{(2m+1)(2n+1)}}{m!\, n!\, 2^{m+n+1}} \frac{d^m}{dx^m}(x^2-1)^m \frac{d^n}{dy^n}(y^2-1)^n.$$

Todo lo expuesto se extiende, de manera obvia, a las funciones de varias variables. En particular, la serie trigonométrica de Fourier para una función de k variables es

$$f(x_1,\ \ldots,\ x_k) = \sum_{n_1,\ \ldots,\ n_k=-\infty}^{\infty} C_{n_1 n_2 \ldots n_k} e^{i(n_1x_1+\ldots+n_kx_k)},$$

donde

$$C_{n_1\ldots n_k} = \frac{1}{(2\pi)^k} \int_{-\pi}^{\pi} \ldots \int_{-\pi}^{\pi} f(x_1,\ \ldots,\ x_k)\, e^{-i(n_1x_1+\ldots+n_kx_k)}\, dx_1 \ldots dx_k.$$

6°. Polinomios ortogonales respecto a un núcleo dado. Hemos llegado a los polinomios de Legendre ortogonalizando las funciones

$$1,\ x,\ x^2,\ \ldots,\ x^n,\ \ldots \tag{9}$$

respecto al producto escalar

$$\int_{-1}^{1} f(x)\, g(x)\, dx$$

correspondiente a la medida habitual de Lebesgue en el segmento $[-1, 1]$. Si se define en este segmento otra medida μ, satisfaciendo la condición de que las funciones (9) sean linealmente independientes en el espacio correspondiente, obtendremos, aplicando al sistema (9) el proceso de ortogonalización, un sistema de polinomios $\{P_n\}$ que depende, en general, de la selección de la medida μ. Supongamos que la medida $\mu(E)$ está definida para los subconjuntos medibles del segmento $[-1, 1]$ mediante la fórmula

$$\mu(E) = \int_E g(x)\,dx, \tag{10}$$

donde g es una función sumable no negativa fija. En este caso, la condición de ortonormalidad

$$(P_m, P_n) = \begin{cases} 1 \text{ para } m = n, \\ 0 \text{ para } m \neq n, \end{cases}$$

es de la forma

$$\int_{-1}^{1} P_m(x)\,P_n(x)\,g(x)\,dx = \begin{cases} 1 \text{ para } m = n, \\ 0 \text{ para } m \neq n. \end{cases} \tag{11}$$

La función g, que define la medida (10), es llamada núcleo o función de peso. Por esto, los polinomios que verifican la condición (11) se dicen ortogonales respecto al núcleo g. La selección de uno u otro núcleo lleva a diferentes sistemas de polinomios. En particular, tomando

$$g(x) = \frac{1}{\sqrt{1-x^2}},$$

obtendremos unos polinomios que coinciden, salvo un coeficiente constante, con los así llamados *polinomios de Chébishev* que se definen mediante la fórmula

$$T_n(x) = \cos n \arccos x \quad (n = 1, 2, \ldots)$$

y desempeñan un papel importante en diferentes problemas de interpolación.

La ortogonalidad de estos polinomios respecto al núcleo $\frac{1}{\sqrt{1-x^2}}$ se comprueba fácilmente. En efecto, tomando

$$x = \cos\theta, \quad dx = -\operatorname{sen}\theta\,d\theta, \quad \sqrt{1-x^2} = \operatorname{sen}\theta,$$

encontramos

$$\int_{-1}^{1} \frac{T_m(x)\,T_n(x)}{\sqrt{1-x^2}}\,dx = \int_{0}^{\pi} \cos m\theta \cos n\theta\, d\theta = \delta_{mn} = \begin{cases} 1 & \text{para } n=m \\ 0 & \text{para } n\neq m. \end{cases}$$

7°. Base ortogonal en el espacio $L_2(-\infty, \infty)$. Funciones de Hermite. En lo que precede hemos considerado diferentes sistemas ortogonales completos en un segmento, esto es, en un conjunto de medida finita. Consideremos ahora el caso de medida infinita y concretamente el espacio $L_2(-\infty, \infty)$ de funciones de cuadrado integrable en toda la recta numérica. Un sistema ortogonal de funciones de este espacio no se puede construir ni a partir de polinomios ni a partir de funciones trigonométricas, ya que todas estas funciones no pertenecen al espacio $L_2(-\infty, \infty)$. Los «materiales» para construir una base en $L_2(-\infty, \infty)$ hay que buscarlos entre las funciones que decrecen suficientemente rápido en el infinito. Probemos que se puede obtener un sistema ortogonal completo en $L_2(-\infty, \infty)$ ortogonalizando la sucesión

$$x^n e^{-\frac{x^2}{2}} \quad (n=0, 1, 2, \ldots)$$

En efecto, toda función de tipo $P(x)e^{-x^2/2}$, donde P es un polinomio, pertenece, evidentemente, a $L_2(-\infty, \infty)$. Además, el conjunto de estas funciones es siempre denso en $L_2(-\infty, \infty)$ (esto será demostrado en el § 4 del capítulo IX).

Aplicando a las funciones $x^n e^{-x^2/2}$ el proceso de ortogonalización, obtenemos el sistema de funciones de tipo

$$\varphi_n(x) = H_n(x)\,e^{-x^2/2} \quad (n=0, 1, 2, \ldots),$$

donde H_n es un polinomio de grado n. Estos polinomios se llaman *polinomios de Hermite* y las propias funciones φ_n se llaman *funciones de Hermite.* Es fácil mostrar que los polinomios de Hermite coinciden, salvo un coeficiente constante, con los polinomios

$$H_n^*(x) = (-1)^n e^{x^2} \frac{d^n e^{-x^2}}{dx^n}.$$

En efecto, el polinomio H_n^* es, evidentemente, de grado n. La relación de ortogonalidad

$$\int_{-\infty}^{\infty} H_n^*(x)\,H_m^*(x)\,e^{-x^2}\,dx = 0 \quad (n \neq m)$$

puede ser comprobada directamente integrando por partes. Pero, debido al corolario del teorema sobre la ortogonalización, existe, salvo un coeficiente constante, solamente un sistema de funciones ortogonales de tipo $P_n(x)e^{-x^2/2}$, donde P_n es un polinomio de grado n.

El resultado obtenido puede ser interpretado también de la siguiente forma. Consideremos en la recta la medida μ de densidad e^{-x^2}, esto es, tal que

$$d\mu = e^{-x^2}\,dx.$$

Esta es una medida finita en la recta. En el espacio de funciones de cuadrado integrable respecto a esta medida en la recta el producto escalar tiene la forma

$$(f, g) = \int_{-\infty}^{\infty} f(x)\,g(x)\,e^{-x^2}\,dx$$

y los polinomios de Hermite forman en este espacio un sistema ortogonal.

Consideremos también el espacio $L_2(0, \infty)$ de funciones de cuadrado integrable en la semirrecta. Tomando en este espacio el sistema de funciones

$$x^n e^{-x}$$

y aplicando a éste el proceso de ortogonalización, obtenemos el sistema de funciones

$$L_n(x)\,e^{-x}$$

llamadas *funciones de Laguerre.*

Los polinomios correspondientes L_n se llaman *polinomios de Laguerre*. Los polinomios de Laguerre pueden ser considerados como una base ortogonal en el espacio de funciones de cuadrado integrable en la semirrecta $(0, \infty)$ con la medida

$$d\mu = e^{-x}\,dx$$

en esta semirrecta. En el § 4 del cap. IX demostraremos que el sistema de funciones de Laguerre es completo en $L_2(0, \infty)$.

8°. Polinomios ortogonales respecto a un núcleo discreto. Supongamos que a $n+1$ puntos diferentes $x_0, x_1, \ldots, x_n$ de la recta real se les han prescrito, a título de núcleo, unos números positivos $p_0, p_1, \ldots, p_n$ y que la medida μ está definida por

$$\mu(E) = \sum_{x_k \in E} p_k,$$

es decir, la medida $\mu(E)$ es igual a la suma de los núcleos de los puntos x_k pertenecientes a E. En esta «medida degenerada» todo conjunto E que no contiene puntos $x_k (k=0, 1, \ldots, n)$ es de medida 0. Luego, la integral de una función f en toda la recta real es igual a

$$\int_{-\infty}^{\infty} f(x)\, d\mu = \sum_{k=0}^{n} p_k f(x_k).$$

Es obvio que las funciones f y g serán equivalentes respecto a la medida μ cuando

$$f(x_k) = g(x_k)$$

en todos los puntos $x_0, x_1, \ldots, x_n$ y solamente en este caso. Para este caso degenerado el problema sobre la mejor aproximación en el sentido de la distancia de L_2 se reduce a buscar las sumas

$$c_0\varphi_0 + c_1\varphi_1 + c_2\varphi_2 + \ldots + c_m\varphi_m$$

que ofrecen el mínimo a la expresión

$$\sum_{k=0}^{n} p_k \left\{ f(x_k) - \sum_{i=0}^{m} c_i \varphi_i(x_k) \right\}^2,$$

es decir, al problema de «interpolación por el método de cuadrados mínimos».

Chébishev fue el primero en desarrollar, partiendo del problema de interpolación por el método de cuadrados mínimos mediante polinomios

$$c_0 + c_1 x + c_2 x^2 + \ldots + c_m x^m$$

de un grado dado m, la teoría de polinomios ortogonales.

Para exponer los resultados de Chébishev, relacionados a este problema, observemos que respecto a nuestra medida μ el sistema

$$1,\ x,\ x^2,\ \ldots,\ x^n$$

es linealmente independiente ya que el producto escalar (x^r, x^s) viene dado por la fórmula

$$(x^r, x^s) = \sum_{k=0}^{n} p_k x_k^{r+s} \tag{12}$$

y el determinante de Gram del sistema (12) es (las sumas son respecto a k desde 0 hasta n):

$$\begin{vmatrix} \sum p_k & \sum p_k x_k & \sum p_k x_k^2 & \dots & \sum p_k x_k^n \\ \sum p_k x_k & \sum p_k x_k^2 & \sum p_k x_k^3 & \dots & \sum p_k x_k^{n+1} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \sum p_k x_k^n & \sum p_k x_k^{n+1} & \sum p_k x_k^{n+2} & \dots & \sum p_k x_k^{2n} \end{vmatrix} =$$

$$= \begin{vmatrix} \sqrt{p_0} & \sqrt{p_1} & \dots & \sqrt{p_n} \\ \sqrt{p_0}x_0 & \sqrt{p_1}x_1 & \dots & \sqrt{p_n}x_n \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \sqrt{p_0}x_0^n & \sqrt{p_1}x_1^n & \dots & \sqrt{p_n}x_n^n \end{vmatrix}^2 =$$

$$= (p_0 p_1 \dots p_n) \begin{vmatrix} 1 & 1 & \dots & 1 \\ x_0 & x_1 & \dots & x_n \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ x_0^n & x_1^n & \dots & x_n^n \end{vmatrix}^2 \neq 0.$$

Al contrario, para $r > n$ las potencias x^r dependen linealmente de las funciones del sistema (12) ya que L_2 es, en nuestro caso, de dimensión $n+1$. Por lo tanto, el proceso de ortogonalización llevará a un sistema finito de polinomios

$$P_0, \ P_1, \ \dots, \ P_n,$$

ortonormales en el sentido de que

$$\sum_{k=0}^{n} p_k P_r(x_k) P_s(x_k) = \delta_{rs},$$

y cada función f se desarrollará en una serie finita

$$f \sim \sum_{r=0}^{n} c_r P_r,$$

donde

$$c_r = \sum_{k=0}^{n} p_k P_r(x_k) f(x_k).$$

En los puntos x_k se cumplen las desigualdades

$$f(x_k) = \sum_{r=0}^{n} c_r P_r(x_k) \quad (k = 0, 1, \dots, n),$$

es decir, la suma de la serie es simplemente el polinomio de interpolación de Lagrange. Las sumas incompletas

$$Q_m = \sum_{r=0}^{m} c_r P_r \ (m < n)$$

son polinomios de grado m que mejor aproximan f en los puntos x_k en el sentido de que la expresión

$$\sum_{k=0}^{n} p_k \{f(x_k) - Q_m(x_k)\}^2$$

es menor para Q_m que para cualquier otro polinomio del mismo grado m.

CAPITULO IX

SERIES TRIGONOMETRICAS. TRANSFORMACION DE FOURIER

§ 1. CONDICIONES DE CONVERGENCIA DE LA SERIE DE FOURIER

1°. Condiciones suficientes de convergencia de la serie de Fourier en un punto. Consideremos de nuevo el espacio $L_2[-\pi, \pi]$ de funciones de cuadrado sumable en el segmento $[-\pi, \pi]$. Como se ha demostrado en el cap. VIII, es un espacio euclídeo completo de dimensiones infinitas, esto es, un espacio de Hilbert. Las funciones

$$1, \cos nx, \operatorname{sen} nx \quad (n = 1, 2, \ldots) \tag{1}$$

forman en él un sistema ortogonal completo, de manera que para toda función $f \in L_2[-\pi, \pi]$ la serie de Fourier

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + b_n \operatorname{sen} nx, \tag{2}$$

donde

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx, \qquad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \operatorname{sen} nx \, dx, \tag{3}$$

converge hacia f cuadráticamente, es decir, en la métrica del espacio $L_2[-\pi, \pi]$. Sin embargo, desde el punto de vista de aplicación de las series de Fourier a los problemas de la Física Matemática y a otras cuestiones, es importante determinar las condiciones en las que se puede garantizar no sólo la convergencia media de la serie de Fourier hacia f, sino también la convergencia en un punto dado, en todos los puntos o, incluso, uniforme. Determinaremos ahora las condiciones suficientes para la convergencia de la serie trigonométrica en un punto dado. Hagamos unas observaciones preliminares.

Ante todo, en vez de hablar de funciones definidas en el segmento $[-\pi, \pi]$, podemos hablar de funciones periódicas de período 2π en toda la recta, ya que cualquier función definida en un segmento puede ser prolongada periódicamente. Observemos, además, que, por ser acotadas las funciones que constituyen el sistema trigonométrico, las fórmulas (3), que definen los coeficientes de Fourier respecto a este sistema, tienen sentido para cualquier función sumable[1] (y no sólo para las funciones de cuadrado sumable). Luego, a toda función $f \in L_1[-\pi, \pi]$ corresponde el conjunto de sus coeficientes de Fourier y su serie de Fourier

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + b_n \operatorname{sen} nx.$$

Pasemos ahora a examinar el problema acerca de la convergencia de esta serie en un punto dado x hacia el valor de la función f en ese punto. Tomemos

$$S_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos kx + b_k \operatorname{sen} kx. \tag{4}$$

Transformemos primeramente $S_n(x)$ sustituyendo en (4) los coeficientes a_n y b_n por sus expresiones integrales (3). Designando la variable de integración con t obtendremos

$$S_n(x) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \left\{ \frac{1}{2} + \sum_{k=1}^{n} \cos kx \cos kt + \operatorname{sen} kx \operatorname{sen} kt \right\} dt =$$

$$= \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \left\{ \frac{1}{2} + \sum_{k=1}^{n} \cos k(t-x) \right\} dt$$

Empleando la fórmula bien conocida[2]

$$\frac{1}{2} + \cos u + \cos 2u + \ldots + \cos nu = \frac{\operatorname{sen} \frac{2n+1}{2} u}{2 \operatorname{sen} \frac{u}{2}}, \tag{5}$$

---

[1] En este caso, claro está, no hacemos ninguna afirmación acerca de la convergencia de la serie (2) para una función sumable cualquiera.

[2] Para obtener esta fórmula basta sumar las igualdades

$$\operatorname{sen} \frac{u}{2} = \frac{1}{2} \cdot 2 \operatorname{sen} \frac{u}{2},$$

$$\operatorname{sen} \frac{3u}{2} - \operatorname{sen} \frac{u}{2} = \cos u \cdot 2 \operatorname{sen} \frac{u}{2},$$

$$\cdots\cdots\cdots\cdots\cdots$$

$$\operatorname{sen} \frac{2n+1}{2} u - \operatorname{sen} \frac{2n-1}{2} u = \cos nu \cdot 2 \operatorname{sen} \frac{u}{2}.$$

encontramos

$$S_n(x) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \frac{\operatorname{sen} \frac{2n+1}{2}(t-x)}{2 \operatorname{sen} \frac{t-x}{2}} dt. \tag{6}$$

Tomemos $t-x=z$. Teniendo en cuenta que bajo el signo de la integral (6) figura una función periódica de período 2π, de manera que su integral en cualquier segmento de lontitud 2π tiene el mismo valor, podemos conservar, al realizar la integración respecto a z, los mismos extremos $-\pi$ y π. Tendremos

$$S_n(x) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x+z) \frac{\operatorname{sen} \frac{2n+1}{2} z}{2 \operatorname{sen} \frac{z}{2}} dz.$$

La función

$$D_n(z) = \frac{1}{2\pi} \frac{\operatorname{sen} \frac{2n+1}{2} z}{\operatorname{sen} \frac{z}{2}}$$

es llamada *núcleo de Dirichlet*. De la igualdad (5) se desprende inmediatamente que cualquiera que sea n

$$\int_{-\pi}^{\pi} D_n(z)\, dz = 1.$$

Empleando este resultado, podemos representar la diferencia $S_n(x)-f(x)$ en la forma

$$S_n(x)-f(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} [f(x+z)-f(x)] \frac{\operatorname{sen} \frac{2n+1}{2} z}{\operatorname{sen} \frac{z}{2}} dz. \tag{7}$$

De esta forma, hemos reducido el problema de la convergencia de $S_n(x)$ hacia $f(x)$ al problema de la convergencia de la integral (7) hacia el cero. El estudio de esta integral se basa en el lema siguiente.

LEMA. *Si φ es una función sumable en el segmento $[a, b]$, entonces*

$$\lim_{p \to \infty} \int_a^b \varphi(x) \operatorname{sen} px\, dx = 0.$$

DEMOSTRACION. Si φ es una función continua diferenciable, tenemos, integrando por partes, (para $p \longrightarrow \infty$)

$$\int_a^b \varphi(x) \operatorname{sen} px\, dx = -\varphi(x) \frac{\cos px}{p} \Big|_a^b + \int_a^b \varphi'(x) \frac{\cos px}{p} dx \longrightarrow 0. \tag{8}$$

Sea ahora φ una función sumable arbitraria en $[a, b]$. Como las funciones continuamente diferenciables (incluso solamente los polinomios) son siempre densas en $L_1[a, b]$, cualquiera que sea $\varepsilon > 0$ existe una función continuamente diferenciable φ_ε tal que

$$\int_a^b |\varphi(x) - \varphi_\varepsilon(x)|\, dx < \frac{\varepsilon}{2}. \tag{9}$$

Tenemos

$$\left| \int_a^b \varphi(x) \operatorname{sen} px\, dx \right| \leqslant \left| \int_a^b [\varphi(x) - \varphi_\varepsilon(x)] \operatorname{sen} px\, dx \right| + \\ + \left| \int_a^b \varphi_\varepsilon(x) \operatorname{sen} px\, dx \right|.$$

Aquí el primer sumando del miembro derecho es menor que $\frac{\varepsilon}{2}$, en virtud de (9), y el segundo sumando tiende a cero para $p \longrightarrow \infty$, debido a (8). El lema queda demostrado.

Es fácil demostrar ahora el siguiente criterio suficiente de convergencia de la serie de Fourier.

TEOREMA 1. *Si f es una función sumable y para x fijo existe la integral*

$$\int_{-\delta}^{\delta} \frac{f(x+t) - f(x)}{t} dt \tag{10}$$

para algún valor de $\delta > 0$, las sumas parciales S_n de la serie de Fourier de la función f convergen en este punto x hacia $f(x)$.

DEMOSTRACION. La integral (7) se puede escribir en la forma

$$\frac{1}{\pi} \int_{-\pi}^{\pi} \frac{f(x+z) - f(x)}{z} \frac{z}{2 \operatorname{sen} \frac{z}{2}} \operatorname{sen} \frac{2n+1}{2} z\, dz. \tag{11}$$

Si la función

$$\frac{f(x+z)-f(x)}{z}$$

es integrable (respecto a z) entre $-\delta$ y δ, también es integrable en todo el segmento $[-\pi, \pi]$ (ya que $f \in L_1$ $[-\pi, \pi]$). Luego, es integrable también la función

$$\frac{f(x+z)-f(x)}{z} \cdot \frac{z}{2 \operatorname{sen} \frac{z}{2}};$$

por lo tanto, se puede aplicar a la integral (11) el lema demostrado, obteniendo así que esta integral converge a cero para $n \to \infty$. El teorema queda demostrado.

Observación 1. La condición de la existencia de la integral (10) es conocida como *condición de Dini*. Ella se cumple, en particular, cuando la función f tiene en el punto dado x derivada finita o, por lo menos, derivada a la izquierda o a la derecha.

Los razonamientos realizados al demostrar el teorema 1 son válidos asimismo, si, en vez de la condición de Dini, se exige la existencia de las dos integrales siguientes:

$$\int_{-\delta}^{0} \frac{f(x+z)-f(x-0)}{z}\,dz \quad \text{y} \quad \int_{0}^{\delta} \frac{f(x+z)-f(x+0)}{z}\,dz, \tag{12}$$

donde $f(x-0)$ y $f(x+0)$ son los límites a la izquierda y a la derecha, respectivamente, de la función f en el punto x (se supone que x es un punto de discontinuidad de primera especie para f). En efecto, considerando la diferencia

$$S_n(x) - \frac{f(x+0)+f(x-0)}{2},$$

que puede ser representada en la forma

$$\frac{1}{\pi}\int_{-\pi}^{0} [f(x+z)-f(x-0)] \frac{\operatorname{sen} \frac{2n+1}{2} z}{2 \operatorname{sen} \frac{z}{2}}\,dz +$$

$$+\frac{1}{\pi}\int_{0}^{\pi} [f(x+z)-f(x+0)] \frac{\operatorname{sen} \frac{2n+1}{2} z}{2 \operatorname{sen} \frac{z}{2}}\,dz,$$

vemos que si existen las integrales (12), esta diferencia tiende a cero para $n \to \infty$. De aquí se obtiene el teorema siguiente sobre las condiciones suficientes de convergencia de la serie de Fourier, que suele darse en los cursos de Análisis.

Supongamos que f es una función acotada de período 2π que tiene solamente discontinuidades de primera especie y supongamos que f posee en cada punto derivadas a la izquierda y a la derecha. Entonces, su serie de Fourier converge en todo punto y su suma es igual a $f(x)$ en los puntos de continuidad y a $\frac{f(x+0)+f(x-0)}{2}$ en los puntos de discontinuidad.

Observación 2. El núcleo de Dirichlet $D_n(z)$, que ha desempeñado un papel fundamental en nuestros razonamientos, es una función que toma en el punto $z=0$ el valor $\frac{2n+1}{2\pi}$ y que oscila rápidamente para valores grandes de n (fig. 23). El sentido del teorema demostrado más arriba consiste en que para una función f,

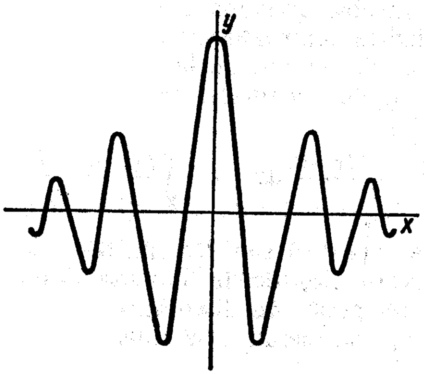


FIG. 23

que satisface en un punto dado x la condición de Dini, el aporte fundamental para grandes n en la integral

$$\int_{-\pi}^{\pi} f(x+z)\, D_n(z)\, dz$$

corresponde sólo a una pequeña vecindad del punto x, cuyas dimensiones tienden a cero para $n \to \infty$. Se puede decir que los núcleos de Dirichlet D_n forman una sucesión de funcionales que converge débilmente hacia la δ-función en el conjunto de funciones f que pueden ser desarrolladas en serie de Fourier.

Está claro que la sucesión $\{D_n\}$ no converge hacia ningún límite en el sentido habitual de convergencia de funciones y es por esto que no hemos podido aplicar al estudio de la integral (7) los teoremas típicos sobre el paso al límite bajo el signo de la integral.

Observación 3. La condición de Dini que garantiza la convergencia de la serie de Fourier puede ser sustituida por otras condiciones; pero, no puede ser simplemente omitida en el teorema 1. En efecto, incluso entre las funciones continuas existen funciones tales que su serie de Fourier diverge en algunos puntos. Entre las funciones sumables existen tales que la serie correspondiente de Fourier diverge en todo punto (A. N. Kolmogórov). Ya en 1915 N. N. Luzin planteó el problema siguiente: ¿existen en L_2 funciones cuya serie de Fourier diverge en un conjunto de medida positiva? Como lo ha demostrado Carleson (en 1966) tales funciones no existen.

La existencia de funciones continuas para las cuales la serie correspondiente de Fourier no converge en todo punto se desprende fácilmente de los teoremas generales sobre la convergencia débil de funcionales. Observemos, ante todo, que

$$\int_{-\pi}^{\pi} |D_n(z)|\, dz \longrightarrow \infty \text{ para } n \longrightarrow \infty. \tag{13}$$

Efectivamente, el numerador de la fracción

$$|D_n(z)| = \frac{\left|\operatorname{sen} \frac{2n+1}{2} z\right|}{2\pi \left|\operatorname{sen} \frac{z}{2}\right|}$$

es igual a 1 en los puntos en los que

$$\frac{2n+1}{2} z = \left(k + \frac{1}{2}\right)\pi, \quad k = 0, 1, \ldots, n. \tag{14}$$

Construyamos alrededor de cada uno de los puntos z, definidos por la condición (14), el intervalo

$$\left|\frac{2n+1}{2} z - \frac{2k+1}{2}\pi\right| < \frac{\pi}{3}. \tag{15}$$

La longitud de cada uno de estos intervalos es, evidentemente, igual a $\frac{2\pi}{3(2n+1)}$. En cada uno de estos intervalos, $\left|\operatorname{sen} \frac{2n+1}{2} z\right|$ es no menor que $\frac{1}{2}$. Estimemos la magnitud de $\operatorname{sen} \frac{z}{2}$ en estos mismos intervalos. Tenemos

$$\operatorname{sen} \frac{z}{2} < \frac{z}{2} < \frac{1}{2}\left(\frac{2k+1}{2}\pi + \frac{\pi}{3}\right)\left(\frac{2n+1}{2}\right)^{-1} < \frac{k+1}{2n+1}\pi.$$

Luego, la integral de $|D_n(z)|$, tomada sólo respecto a los intervalos definidos por la condición (15), es mayor que la suma

$$\frac{1}{2\pi}\sum_{k=0}^{n}\frac{1}{2}\,\frac{1}{\frac{k+1}{2n+1}\pi}\;\frac{2\pi}{3(2n+1)}=\frac{1}{6\pi}\sum_{k=0}^{n}\frac{1}{k+1}.$$

Esta suma tiende a ∞ para $n\to\infty$. De aquí se deduce (13). La relación (13) significa que las normas de las funciones D_n no están acotadas en su conjunto en el espacio de funciones continuas. Pero, entonces, en virtud del teorema sobre la convergencia débil de funcionales, esta sucesión no puede converger débilmente en el espacio de funciones continuas, esto es, existe una función continua f para la cual no existe

$$\lim_{n\to\infty}\int_{-\pi}^{\pi}D_n(x)f(x)\,dx.$$

2°. **Condiciones de convergencia uniforme de la serie de Fourier.** Hemos encontrado las condiciones suficientes para que la serie de Fourier de una función f converge en todo punto. La clase de funciones que satisfacen estas condiciones es muy amplia: incluso la condición de continuidad de la función no es necesaria para poderla representar como la suma de una serie trigonométrica que converja en todo punto. La situación resulta distinta, si nos interesamos por las condiciones de la convergencia uniforme de la serie de Fourier. Está claro que si la función $f(x)$ tiene al menos una discontinuidad, su serie de Fourier no puede converger uniformemente hacia ella ya que la suma de una serie uniformemente convergente de funciones continuas es siempre una función continua. Luego, la continuidad es una condición necesaria (pero, no suficiente, por supuesto) para la convergencia uniforme de la serie de Fourier.

El teorema siguiente ofrece una condición suficiente sencilla para que la serie de Fourier converja uniformemente.

TEOREMA 2. *Si una función f de período 2π es absolutamente continua y la derivada de f pertenece a $L_2[-\pi, \pi]$, la serie de Fourier de la función f converge hacia f uniformemente en toda la recta.*

DEMOSTRACION. Designemos mediante a_n' y b_n' los coeficientes de Fourier de la función f'. Como f es absolutamente continua, se puede aplicar a la integral

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx$$

la fórmula de integración por partes. Obtenemos

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx =$$

$$= \frac{1}{\pi} f(x) \frac{\operatorname{sen} nx}{n} \Big|_{-\pi}^{\pi} - \frac{1}{n\pi} \int_{-\pi}^{\pi} f'(x) \operatorname{sen} nx \, dx = -\frac{b_n'}{n}$$

y análogamente

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \operatorname{sen} nx \, dx = \frac{a_n'}{n}.$$

Luego,

$$\sum_{n=1}^{\infty} |a_n| + |b_n| = \sum_{n=1}^{\infty} \frac{|b_n'|}{n} + \frac{|a_n'|}{n}. \tag{16}$$

La serie numérica

$$\left|\frac{a_0}{2}\right| + \sum_{n=1}^{\infty} |a_n| + |b_n| \tag{17}$$

es, evidentemente, una mayorante para la serie de Fourier de la función f. De (16) se desprende que la serie (17) converge[1]. Entonces, por el criterio de Weierstrass, la serie de Fourier de la función f converge uniformemente. Queda por demostrar que la suma de esta serie es f. Sea φ la suma de la serie de Fourier para la función f. En tal caso φ tiene los mismos coeficientes de Fourier que tiene f. Como ambas funciones son continuas, de aquí obtenemos que $f = \varphi$.

La condición de convergencia uniforme de la serie de Fourier de una función f puede ser enunciada en una forma análoga a la condición de Dini, a saber:

---

[1] En efecto, $\frac{|b_n'|}{n} \leqslant \frac{1}{2}\left(b_n'^2 + \frac{1}{n^2}\right)$, donde $\sum_{n=1}^{\infty} b_n'^2 < \infty$ debido a la desigualdad de Bessel. Análogamente para $\frac{|a_n'|}{n}$.

Si f es una función sumable acotada en un conjunto $E \subset [-\pi, \pi]$ y la condición de Dini se cumple en E uniformemente, esto es, para todo $\varepsilon > 0$ existe un $\delta > 0$ tal que

$$\int_{-\delta}^{\delta} \frac{|f(x+z)-f(x)|}{|z|}\, dz < \varepsilon$$

simultáneamente para todos los $x \in E$, la serie de Fourier de la función f converge uniformemente en E hacia esta función.

La demostración de este teorema se basa en el lema siguiente que constituye una acentuación del lema demostrado en la pág. 451.

Si B es un conjunto de funciones sumables compacto en la métrica de $L_1[-\pi, \pi]$, entonces, cualquiera que sea $\varepsilon > 0$ existe un $N = N(\varepsilon)$ tal que

$$\left| \int_a^b f(t) \operatorname{sen} \lambda t\, dt \right| < \varepsilon$$

para $\lambda \geqslant N(\varepsilon)$ y para todas las $f \in B$ a la vez.

Para demostrar el lema tomemos en B una $\frac{\varepsilon}{2}$-red finita $\varphi_1, \ldots, \varphi_k$ y escojamos N de manera que

$$\left| \int_a^b \varphi_i(t) \operatorname{sen} \lambda t\, dt \right| < \frac{\varepsilon}{2}, \quad i = 1, 2, \ldots, k, \text{ para } \lambda \geqslant N.$$

Si ahora f es una función cualquiera de B, tenemos para cierto i

$$\| f - \varphi_i \| < \frac{\varepsilon}{2}$$

y, consecuentemente,

$$\left| \int_a^b f(t) \operatorname{sen} \lambda t\, dt \right| \leqslant \left| \int_a^b \varphi_i(t) \operatorname{sen} \lambda t\, dt \right| + \left| \int_a^b (f - \varphi_i) \operatorname{sen} \lambda t\, dt \right| < \varepsilon.$$

Esto demuestra el lema.

La aplicación de este lema a la demostración del teorema 2 se basa en que, como es fácil de demostrar, el conjunto de funciones

$$\varphi_x(t) = \frac{f(x+t)-f(x)}{t}$$

es compacto.

Hasta aquí hemos hablado de funciones definidas en el segmento $[-\pi, \pi]$. Está claro que todo lo expuesto puede ser extendido automáticamente a las funciones definidas en un segmento de una longitud arbitraria $2l$.

Además, también en el caso de varias variables independientes se pueden enunciar tanto las condiciones suficientes para la convergencia de la serie de Fourier en todo punto como las condiciones de la convergencia uniforme de la serie de Fourier. No vamos a detenernos en estas condiciones.

§ 2. TEOREMA DE FEJÉR

1°. Teorema de Fejér. Sea f una función continua en la recta de período 2π. Esta función se determina unívocamente por su serie de Fourier

$$\frac{a_0}{2}+\sum_{n=1}^{\infty} a_n \cos nx + b_n \operatorname{sen} nx. \tag{1}$$

En efecto, si f_1 y f_2 son dos funciones continuas que tienen los mismos coeficientes de Fourier, entonces f_1-f_2 es una función continua igual a cero en casi todo punto, es decir, idénticamente igual a cero. Sin embargo, como la serie de Fourier de una función continua no es necesariamente convergente, no podemos obtener la función f sumando directamente su serie de Fourier. Un método de reconstrucción de una función continua a partir de su serie de Fourier ofrece el teorema que damos a continuación, demostrado por Fejér en 1905.

Sea

$$S_k(x)=\frac{a_0}{2}+\sum_{j=1}^{k} a_j \cos jx + b_j \operatorname{sen} jx \tag{2}$$

la suma parcial de la serie de Fourier de la función f. Tomemos

$$\sigma_n(x)=\frac{S_0(x)+S_1(x)+\ldots+S_{n-1}(x)}{n}. \tag{3}$$

Las medias aritméticas σ_n de las sumas S_k se llaman *sumas de Fejér* de la función f.

TEOREMA 1 (Fejér). *Si f es una función continua de período 2π, la sucesión $\{\sigma_n\}$ de sus sumas de Fejér converge hacia f uniformemente en toda la recta numérica.*

DEMOSTRACION. Empleemos la representación integral (6), obtenida en el parágrafo anterior para las sumas parciales de la serie de

Fourier,

$$S_k(x)=\frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\frac{\operatorname{sen}\frac{2k+1}{2}(t-x)}{2\operatorname{sen}\frac{t-x}{2}}\,dt$$

Introduciendo estas expresiones en la igualdad (3), obtenemos la siguiente expresión para $\sigma_n(x)$:

$$\sigma_n(x)=\frac{1}{2n\pi}\int_{-\pi}^{\pi}\left\{\sum_{k=0}^{n-1}\frac{\operatorname{sen}\frac{2k+1}{2}(t-x)}{\operatorname{sen}\frac{t-x}{2}}\right\}f(t)\,dt,$$

que mediante la fórmula[1)]

$$\sum_{k=0}^{n-1}\operatorname{sen}(2k+1)u=\frac{\operatorname{sen}^2 nu}{\operatorname{sen} u}$$

puede ser reducida a la forma

$$\sigma_n(x)=\frac{1}{2n\pi}\int_{-\pi}^{\pi}\left(\frac{\operatorname{sen} n\frac{t-x}{2}}{\operatorname{sen}\frac{t-x}{2}}\right)^2 f(t)\,dt. \tag{4}$$

La expresión

$$\Phi_n(z)=\frac{1}{2n\pi}\left(\frac{\operatorname{sen} n\frac{z}{2}}{\operatorname{sen}\frac{z}{2}}\right)^2 \tag{5}$$

es el así llamado *núcleo de Fejér*. Tomando $t-x=z$ y teniendo en cuenta que el valor de la integral de una función periódica en un segmento de longitud igual al período es siempre el mismo, podemos escribir (4) en la forma

$$\sigma_n(x)=\int_{-\pi}^{\pi} f(x+z)\,\Phi_n(z)\,dz. \tag{6}$$

Debemos demostrar que para $n\to\infty$ esta expresión converge uniformemente hacia $f(x)$. Señalemos primeramente las propiedades

---

1) Esta fórmula se obtiene fácilmente sumando respecto a k las igualdades $2\operatorname{sen}(2k+1)u\cdot\operatorname{sen}u=\cos 2ku-\cos 2(k+1)u$ y realizando transformaciones trigonométricas elementales.

siguientes del núcleo de Fejér:

1) $\Phi_n(z) \geqslant 0$,

2) $\int\limits_{-\pi}^{\pi} \Phi_n(z)\,dz = 1$,

3) para todo $\sigma > 0$ fijo y $n \to \infty$ tenemos

$$\int\limits_{-\pi}^{-\delta} \Phi_n(z)\,dz = \int\limits_{\delta}^{\pi} \Phi_n(z)\,dz = \eta_n(\delta) \to 0.$$

La primera de estas propiedades es evidente; la segunda se obtiene directamente de la igualdad (6), si tomamos $f(x) \equiv 1$ y observamos que para esta función $\sigma_n(x) \equiv 1$ para todo n; finalmente, la tercera propiedad se desprende inmediatamente de que para $\sigma < z \leqslant \pi$ se tiene $\operatorname{sen} \frac{z}{2} \geqslant \frac{2\delta}{\pi}$ y, consecuentemente,

$$\left(\frac{\operatorname{sen} n \frac{z}{2}}{\operatorname{sen} \frac{z}{2}} \right)^2 \leqslant \left(\frac{\pi}{2\delta} \right)^2.$$

Tomando en consideración estas propiedades del núcleo de Fejér no es difícil demostrar el teorema. Como f es una función continua y periódica, es acotada y uniformemente continua en toda la recta. En otras palabras, existe una constante M tal que para todo x

$$|f(x)| \leqslant M \tag{7}$$

y, además, para todo $\varepsilon > 0$ existe un $\delta > 0$ tal que

$$|f(x'') - f(x')| < \frac{\varepsilon}{2} \tag{8}$$

siempre que

$$|x'' - x'| < \delta.$$

Para demostrar el teorema debemos estimar la diferencia

$$f(x) - \sigma_n(x) = \int\limits_{-\pi}^{\pi} [f(x) - f(x+z)]\,\Phi_n(z)\,dz,$$

que puede ser representada como suma de las tres integrales siguientes:

$$J_1 = \int_{-\pi}^{-\delta} \{f(x) - f(x+z)\} \Phi_n(z)\, dz,$$

$$J_2 = \int_{-\delta}^{\delta} \{f(x) - f(x+z)\} \Phi_n(z)\, dz,$$

$$J_3 = \int_{\delta}^{\pi} \{f(x) - f(x+z)\} \Phi_n(z)\, dz.$$

De (7) y (8) se obtienen directamente las estimaciones siguientes:

$$|J_1| \leqslant 2M\eta_n(\delta),$$
$$|J_3| \leqslant 2M\eta_n(\delta),$$
$$|J_2| \leqslant \frac{\varepsilon}{2} \int_{-\delta}^{\delta} \Phi_n(z)\, dz < \frac{\varepsilon}{2}.$$

Escojamos ahora n_0 tan grande que para $n \geqslant n_0$ y un δ dado se cumpla la desigualdad

$$2M\eta_n(\delta) < \frac{\varepsilon}{4}.$$

Entonces,

$$|f(x) - \sigma_n(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \varepsilon$$

y de aquí, debido a la arbitrariedad de ε, se desprende la afirmación del teorema.

Observemos que en la demostración han sido empleadas solamente las propiedades 1), 2) y 3) del núcleo de Fejér. Esto permite obtener diferentes generalizaciones del teorema 1 (véase el punto 3 de este parágrafo).

2°. Complitud del sistema trigonométrico. Teorema de Weierstrass. Del teorema de Fejér se desprende la complitud del sistema de funciones trigonométricas en el espacio $L_2[-\pi, \pi]$. Efectivamente, de acuerdo con este teorema cualquier función continua es el límite de una sucesión de polinomios trigonométricos σ_n convergente uniformemente (y, por lo tanto, también en media). Como las funciones continuas son siempre densas en L_2, de aquí se deduce la complitud del sistema trigonométrico.

El teorema de Fejér puede ser considerado como un complemento del teorema de Weierstrass sobre la aproximación de funciones continuas mediante polinomios trigonométricos: este último afirma que toda función continua periódica es límite uniforme de *cierta* sucesión de polinomios trigonométricos, mientras que el teorema de Fejér indica una sucesión concreta con esta propiedad, la sucesión de sumas de Fejér (3). Del teorema de Weierstrass sobre la aproximación uniforme de una función continua periódica mediante polinomios trigonométricos se deduce fácilmente el segundo teorema de Weierstrass, esto es, el teorema sobre la aproximación de cualquier función continua en un segmento $[a, b]$ mediante polinomios algebraicos. En efecto, si $f(x)$ es una función de este tipo, tomando $t=\frac{x-a}{b-a}\pi$, es decir, $x=\frac{t(b-a)}{\pi}+a$, obtenemos una función $\varphi(t)$ de t definida en el segmento $[0, \pi]$. Prolonguémosla primero al semisegmento $[-\pi, 0]$, tomando $\varphi(-t)=\varphi(t)$, y después, por periodicidad, a toda la recta. Construyamos ahora un polinomio trigonométrico T_n que verifique la condición

$$|T_n(t)-\varphi(t)|<\frac{\varepsilon}{2} \text{ para todo } t.$$

Ahora bien, todo polinomio trigonométrico puede ser desarrollado en una serie de Taylor que converge uniformemente en cualquier intervalo finito. Sea P_m la suma parcial de la serie de Taylor para T_n tal que

$$|T_n(t)-P_m(t)|<\frac{\varepsilon}{2} \quad \text{para } 0\leqslant t\leqslant\pi.$$

Entonces,

$$|\varphi(t)-P_m(t)|<\varepsilon \text{ para } 0\leqslant t\leqslant\pi.$$

Efectuando en $P_m(t)$ la sustitución $t=\frac{x-a}{b-a}\pi$, obtenemos un polinomio $Q_m(x)$ que satisface, evidentemente, la condición

$$|f(x)-Q_m(x)|<\varepsilon \text{ para } a\leqslant x\leqslant b.$$

3°. Teorema de Fejér en el caso del espacio L_1. En el teorema de Fejér se ha logrado alcanzar cierta simetría entre la hipótesis y la tesis del teorema. El hecho de que una función f pertenezca al espacio $C_{[-\pi, \pi]}$ de funciones continuas implica que las sumas de Fejér, correspondientes a ella, converjan hacia f en la métrica de ese mismo espacio $C_{[-\pi, \pi]}$. Los teoremas análogos se pueden obtener también para otros espacios funcionales, en particular, para el espacio $L_1[-\pi, \pi]$. Hablando con más rigor, tiene lugar el siguiente teorema que es natural llamarlo teorema de Fejér para funciones sumables:

Si f es una función sumable en el segmento $[-\pi, \pi]$, sus sumas de Fejér convergen hacia f respecto a la norma del espacio $L_1[-\pi, \pi]$.

La demostración de este teorema se puede obtener mediante razonamientos próximos a los expuestos en el punto 1. No los daremos aquí pero si indicaremos un resultado importante que se desprende del teorema de Fejér para funciones sumables:

Toda función sumable se determina unívocamente (salvo una equivalencia) *por sus coeficientes de Fourier.*

En efecto, sean f y g dos funciones sumables que tienen coeficientes de Fourier iguales. Entonces, todos los coeficientes de Fourier de la función $f-g$ son iguales a 0. Luego, son iguales idénticamente a cero también todas las sumas de Fejér para $f-g$. Consecuentemente, el límite de ellas, L_1, esto es, la función $f-g$, es 0 en casi todo punto.

§ 3. INTEGRAL DE FOURIER

1°. Teorema fundamental. En el § 1 se han encontrado las condiciones en que una función definida en un segmento finito (o, que es lo mismo, una función periódica en toda la recta) puede ser desarrollada en la serie convergente de Fourier, es decir, puede ser representada como superposición de ondas armónicas. Tratemos ahora de extender este resultado a funciones no periódicas. Como veremos, bajo unas condiciones adicionales bastante generales, es posible obtener esta representación sólo no mediante una serie, sino mediante una integral, la así llamada integral de Fourier.

Comencemos por unas consideraciones sugestivas formales. Sea f una función que en cada intervalo finito satisface las condiciones que garantizan su desarrollo en la serie de Fourier. En otras palabras, supongamos que f es sumable en cualquier intervalo finito y que verifica en todo punto la condición de Dini. Considerando f, digamos, en el intervalo $(-l, l)$, podemos escribir el desarrollo de esta función en la serie de Fourier:

$$f(x)=\frac{a_0}{2}+\sum_{k=1}^{\infty} a_k \cos\frac{k\pi}{l}x+b_k \operatorname{sen}\frac{k\pi}{l}x. \tag{1}$$

Sustituyendo aquí a_k y b_k por sus expresiones

$$a_0=\frac{1}{l}\int_{-l}^{l} f(t)\,dt,\quad a_n=\frac{1}{l}\int_{-l}^{l} f(t)\cos\frac{k\pi}{l}t\,dt,$$

$$b_n=\frac{1}{l}\int_{-l}^{l} f(t)\operatorname{sen}\frac{k\pi}{l}t\,dt.$$

obtenemos

$$f(x)=\frac{1}{2l}\int_{-l}^{l} f(t)\,dt+\sum_{k=1}^{\infty}\frac{1}{l}\int_{-l}^{l} f(t)\cos\frac{k\pi}{l}x\cos\frac{k\pi}{l}t\,dt+$$

$$+\frac{1}{l}\int_{-l}^{l} f(t)\operatorname{sen}\frac{k\pi}{l}x\operatorname{sen}\frac{k\pi}{l}t\,dt,$$

es decir,

$$f(x)=\frac{1}{2l}\int_{-l}^{l} f(t)\,dt+$$

$$+\frac{1}{l}\sum_{k=1}^{\infty}\int_{-l}^{l} f(t)\left[\cos\frac{k\pi}{l}x\cos\frac{k\pi}{l}t+\operatorname{sen}\frac{k\pi}{l}x\operatorname{sen}\frac{k\pi}{l}t\right]dt=$$

$$=\frac{1}{2l}\int_{-l}^{l} f(t)\,dt+\frac{1}{\pi}\sum_{k=1}^{\infty}\frac{\pi}{l}\int_{-l}^{l} f(t)\cos\frac{k\pi}{l}(t-x)\,dt. \tag{2}$$

Completemos las suposiciones hechas respecto a la función f con la siguiente: esta función es absolutamente integrable en toda la recta, es decir,

$$\int_{-\infty}^{\infty} |f(t)|\,dt<\infty. \tag{3}$$

Pasemos en la igualdad (2) al límite para $l\to\infty$ (por ahora de una manera formal). Debido a (3), el primer sumando del miembro derecho de la igualdad (2) tiende a cero para $l\to\infty$. El segundo sumando puede ser considerado como la suma integral (referida a todo el intervalo infinito) de la integral

$$\int_{0}^{\infty} F(\lambda)\,d\lambda$$

de la función

$$F(\lambda)=\int_{-l}^{l} f(t)\cos\lambda(t-x)\,dt,$$

si tomamos $\lambda_k=\frac{k\pi}{l}$ y $\Delta\lambda=\frac{\pi}{l}$. Luego, el paso formal al límite para $l\to\infty$ en (2) lleva a la igualdad

$$f(x)=\frac{1}{\pi}\int_{0}^{\infty} d\lambda\int_{-\infty}^{\infty} f(t)\cos\lambda(t-x)\,dt. \tag{4}$$

Esta es precisamente la representación deseada. Designando

$$a_\lambda=\frac{1}{\pi}\int_{-\infty}^{\infty} f(t)\cos\lambda t\,dt,\quad b_\lambda=\frac{1}{\pi}\int_{-\infty}^{\infty} f(t)\operatorname{sen}\lambda t\,dt,$$

la igualdad (4) puede ser representada en la siguiente forma analoga a la serie de Fourier:

$$f(x)=\int_0^{\infty}(a_\lambda \cos\lambda x+b_\lambda \operatorname{sen}\lambda x)\,d\lambda. \tag{5}$$

Por ahora hemos obtenido la igualdad (4), llamada *fórmula de Fourier*, mediante el paso formal al límite. La validez de este paso (con las suposiciones hechas respecto a la función f) puede ser argumentada; sin embargo, es más simple demostrar la igualdad (4) directamente. Demostremos, pues, el teorema siguiente.

TEOREMA 1. *Si una función f es absolutamente integrable en toda la recta y verifica la condición de Dini en un punto x, tiene lugar la igualdad*

$$f(x)=\frac{1}{\pi}\int_0^{\infty}d\lambda\int_{-\infty}^{\infty}f(t)\cos\lambda(t-x)\,dt.$$

DEMOSTRACION. Designemos

$$J(A)=\frac{1}{\pi}\int_0^{A}d\lambda\int_{-\infty}^{\infty}f(t)\cos\lambda(t-x)\,dt. \tag{6}$$

Debemos demostrar que $\lim\limits_{A\to\infty}J(A)$ existe y es igual a $f(x)$. Como f es absolutamente integrable, la integral interior en (6) converge y, además, uniformemente, para todos los valores de λ. Por lo tanto, aplicando el teorema de Fubini, podemos cambiar el orden de integración en la integral reiterada (6). Esto nos da

$$J(A)=\frac{1}{\pi}\int_{-\infty}^{\infty}dt\int_0^{A}f(t)\cos\lambda(t-x)\,d\lambda=\frac{1}{\pi}\int_{-\infty}^{\infty}f(t)\frac{\operatorname{sen}A(t-x)}{t-x}\,dt.$$

Poniendo $t-x=z$, podemos representar esta integral en la forma

$$J(A)=\frac{1}{\pi}\int_{-\infty}^{\infty}f(x+z)\frac{\operatorname{sen}Az}{z}\,dz. \tag{7}$$

Valiéndonos ahora de la conocida igualdad

$$\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\operatorname{sen}Az}{z}\,dz=1\quad(A>0),$$

podemos escribir la diferencia $J(A)-f(x)$ en la forma

$$J(A)-f(x)=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{f(x+z)-f(x)}{z}\operatorname{sen} Az\,dz. \tag{8}$$

Consideremos la integral del miembro derecho de esta relación como suma de tres componentes

$$J(A)-f(x)=\frac{1}{\pi}\int_{-N}^{N}\frac{f(x+z)-f(x)}{z}\operatorname{sen} Az\,dz+$$
$$+\frac{1}{\pi}\int_{|z|\geqslant N}\frac{f(x+z)}{z}\operatorname{sen} Az\,dz-\frac{f(x)}{\pi}\int_{|z|\geqslant N}\frac{\operatorname{sen} Az}{z}\,dz.$$

Como el segundo y el tercer términos del miembro derecho son integrales convergentes, cada uno de ellos puede ser hecho menor que $\frac{\varepsilon}{3}$ si se escoge suficientemente grande el número N. Finalmente, el primer sumando del miembro derecho tiende (para N fijo) a cero cuando $A\to\infty$ (en virtud del lema del § 1 y de la condición de Dini). Consecuentemente, obtenemos

$$\lim_{A\to\infty}(J(A)-f(x))=0,$$

que es lo que se quería demostrar.

2°. **Forma compleja de la integral de Fourier.** Puesto que en la fórmula integral de Fourier (4) la integral interior es una función par respecto a λ, podemos escribir esta fórmula así:

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}d\lambda\int_{-\infty}^{\infty}f(t)\cos\lambda(t-x)\,dt. \tag{9}$$

Ahora bien, la integrabilidad absoluta de la función f implica que la integral $\int_{-\infty}^{\infty}f(t)\operatorname{sen}\lambda(t-x)\,dt$ exista y sea una función impar respecto a λ. Por lo tanto,

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}d\lambda\int_{-\infty}^{\infty}f(t)\operatorname{sen}\lambda(t-x)\,dt=0, \tag{10}$$

si la integral respecto a λ se comprende aquí en el sentido del valor principal, esto es, como $\lim\limits_{N\to\infty}\int\limits_{-N}^{N}$. Sumando a (9) la igualdad (10) multiplicada por $-i$, obtenemos

$$f(x)=\frac{1}{2\pi}\int\limits_{-\infty}^{\infty}d\lambda\int\limits_{-\infty}^{\infty}f(t)\,e^{-i\lambda(t-x)}\,dt.$$

Esta igualdad será llamada *fórmula compleja de Fourier.*

§ 4. TRANSFORMACIÓN DE FOURIER, SUS PROPIEDADES Y SUS APLICACIONES

1°. Transformación de Fourier y fórmula de inversión. La fórmula integral de Fourier puede ser transformada del modo siguiente. Tomemos

$$g(\lambda)=\int\limits_{-\infty}^{\infty}f(t)\,e^{-i\lambda t}\,dt. \tag{1}$$

Entonces

$$f(x)=\frac{1}{2\pi}\int\limits_{-\infty}^{\infty}g(\lambda)\,e^{i\lambda x}\,d\lambda. \tag{2}$$

Notemos que la fórmula (1) tiene sentido para cualquier función f absolutamente integrable. Es decir, mediante la fórmula (1) a toda $f\in L_1(-\infty,\infty)$ ponemos en correspondencia una función determinada g definida en toda la recta numérica. Esta última se llama *transformación de Fourier* de la función inicial f. La fórmula (2) que expresa f a través de su transformación de Fourier se llama *fórmula de inversión* de la transformación de Fourier. Conviene prestar atención a la semejanza que existe entre las fórmulas (1) y (2). La segunda difiere de la primera sólo en el signo de la exponente y en el coeficiente $\frac{1}{2\pi}$ delante de la integral. Podríamos alcanzar una simetría aún mayor, si hubiésemos tomado

$$g(\lambda)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}f(x)\,e^{-i\lambda x}\,dx. \tag{1'}$$

La fórmula de inversión tendría entonces la forma

$$f(x)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}g(\lambda)\,e^{i\lambda x}\,d\lambda, \tag{2'}$$

es decir, la diferencia consistiría sólo en el signo de la exponente.

Sin embargo, a pesar de la semejanza exterior de las fórmulas (1) y (2), son sustancialmente distintas: en la primera, la integral existe en el sentido habitual (ya que $f \in L_1(-\infty, \infty)$, mientras que en la segunda, sólo en el sentido del valor principal. Además, la igualdad (1) es la *definición* de la función g, mientras que la igualdad (2), que constituye una forma de la fórmula integral de Fourier, contiene la *afirmación* de que la integral que figura en su miembro derecho es igual a la función inicial f. Como hemos visto más arriba, para que esta igualdad sea válida, la función f, además de ser integrable, debe verificar unas condiciones adicionales, digamos la condición de Dini.

Observación. Hemos definido la transformación de Fourier g para toda función f de $L_1(-\infty, \infty)$ y hemos demostrado que una función f, que en todo punto cumple la condición de Dini, puede ser expresada mediante g a través de la fórmula de inversión. La situación aquí es totalmente análoga a la que tiene lugar para las series de Fourier. En efecto, los coeficientes de Fourier

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx}\, dx$$

están definidos para toda $f \in L_1[-\pi, \pi]$; sin embargo, la convergencia de la serie de Fourier

$$\sum_{n=-\infty}^{\infty} c_n e^{inx}$$

(que desempeña aquí el papel de la fórmula de inversión) puede ser garantizada solamente bajo ciertas condiciones adicionales (condición de Dini). Al mismo tiempo, para la transformación de Fourier (lo mismo que para la serie; véase el final del § 2) tiene lugar el resultado siguiente: *si para una función $f \in L_1(-\infty, \infty)$ se tiene*

$$\int_{-\infty}^{\infty} f(x) e^{-i\lambda x}\, dx \equiv 0,$$

es $f(x) = 0$ en casi todo punto. En efecto, observemos, en primer lugar, que de la igualdad anterior se desprende que para todos los t y λ reales es

$$\int_{-\infty}^{\infty} f(x+t) e^{-i\lambda x}\, dx = 0.$$

Pongamos ahora

$$\varphi(x)=\int_0^{\xi} f(x+t)\,dt,$$

donde ξ es un número real fijado. Empleando el teorema de Fubini y la condición a la que se ha sometido la función f, es fácil ver que la función φ (que también pertenece a $L_1(-\infty, \infty)$) satisface la misma condición, esto es,

$$\int_{-\infty}^{\infty} \varphi(x)\,e^{-i\lambda x}\,dx=0$$

para todo λ real. Pero, la función φ es, como se comprueba fácilmente, una función absolutamente continua en cada segmento finito y, por lo tanto, tiene derivada finita en casi todo punto. En particular, esta función satisface en casi todo punto la condición de Dini. De manera que, en virtud del teorema 1 del § 3, se anula en casi todo punto ya que su transformación de Fourier es el 0 idéntico. Pero, φ es continua; luego, $\varphi(x)\equiv 0$. De aquí se desprende, en particular, que para todo ξ real

$$\int_0^{\xi} f(t)\,dt=0$$

y, por consiguiente, $f(x)=0$ en casi todo punto.

Consideremos algunos ejemplos.

1. Sea $f(x)=e^{-\gamma|x|}$, $\gamma>0$. Busquemos la transformación de Fourier de esta función. Tenemos

$$g(\lambda)=\int_{-\infty}^{\infty} e^{-\gamma|x|}e^{-i\lambda x}\,dx=$$

$$=\int_{-\infty}^{\infty} e^{-\gamma|x|}(\cos\lambda x-i\operatorname{sen}\lambda x)\,dx=2\int_0^{\infty} e^{-\gamma x}\cos\lambda x\,dx.$$

Integrando dos veces por partes, encontramos

$$g(\lambda)=\frac{2}{\lambda^2+\gamma^2}.$$

2. Sea

$$f(x)=\begin{cases} 1 & \text{para } |x|\leqslant a,\\ 0 & \text{para } |x|>a.\end{cases}$$

Entonces,

$$g(\lambda)=\int_{-\infty}^{\infty} f(x)\, e^{-i\lambda x}\, dx=\int_{-a}^{a} e^{-i\lambda x}\, dx=\frac{e^{i\lambda a}-e^{-i\lambda a}}{i\lambda}=\frac{2 \operatorname{sen} \lambda a}{\lambda}.$$

(Conviene prestar atención a que la función g no pertenece, en este caso, a $L_1(-\infty, \infty)$).

3. Sea $f(x)=\frac{1}{x^2+a^2}$. Entonces,

$$g(\lambda)=\int_{-\infty}^{\infty} e^{-i\lambda x}\frac{dx}{x^2+a^2}. \tag{3}$$

Lo más sencillo para calcular esta integral es emplear la teoría de residuos. Supongamos primero que $\lambda>0$. Completando el eje real, respecto al que se toma la integral (2), mediante una semicircunferencia de radio indefinidamente grande, situada en el semiplano inferior (esto es, en el semiplano donde la exponente $e^{-i\lambda x}$ tiende a cero), obtenemos que la integral (3) es igual a la suma de residuos del integrando en el plano inferior dividida por $2\pi i$. La función $\frac{e^{-i\lambda x}}{x^2+a^2}$ tiene en el semiplano inferior un polo de orden 1 en el punto $x=-ai$. El residuo en este punto se calcula de acuerdo con la siguiente regla conocida: si $f(z)=\frac{\varphi(z)}{\psi(z)}$, $\varphi(a)\neq 0$ y $\psi(z)$ *tiene en el punto* $z=a$ *un cero de primer orden, el residuo de la función* f *en el punto* a *es igual a* $\frac{\varphi(a)}{\psi'(a)}$. Luego, obtenemos en nuestro caso que

$$g(\lambda)=\frac{1}{2\pi i}\frac{e^{-a\lambda}}{-2ai}=\frac{e^{-a\lambda}}{4\pi a} \quad \text{para } \lambda>0.$$

Para $\lambda<0$ obtenemos análogamente (considerando solamente el semiplano superior en vez del inferior)

$$g(\lambda)=\frac{e^{a\lambda}}{4\pi a}.$$

Es decir, resumiendo,

$$g(\lambda)=\frac{e^{-a|\lambda|}}{4\pi a} \quad (-\infty<\lambda<\infty).$$

4. Pongamos $f(x)=e^{-ax^2}$. Tenemos

$$g(\lambda)=\int_{-\infty}^{\infty} e^{-ax^2}\, e^{-i\lambda x}\, dx. \tag{4}$$

El integrando representa aquí una función analítica que no tiene singularidades en ninguna parte finita del plano. Por lo tanto, en virtud del teorema de Cauchy, la integral (4) no cambiará de valor, si, en lugar de tomarla respecto al eje real, se toma respecto a cualquier recta $z=x+iy$ ($y=\text{const}$) paralela a este eje. Es decir,

$$g(\lambda)=\int_{-\infty}^{\infty} e^{-a(x+iy)^2}\cdot e^{-i\lambda(x+iy)}\,dx=$$
$$=e^{ay^2+\lambda y}\int e^{-ax^2-2aixy-i\lambda x}\,dx=e^{ay^2+\lambda y}\int e^{-ax^2-ix(2ay+\lambda)}\,dx.$$

Escojamos ahora la constante y de manera que se anule la parte imaginaria en la exponente del integrando, esto es, tomemos $y=\frac{-\lambda}{2a}$. Entonces

$$g(\lambda)=e^{a\frac{\lambda^2}{4a^2}-\frac{\lambda^2}{2a}}\int_{-\infty}^{\infty} e^{-ax^2}\,dx=e^{-\frac{\lambda^2}{4a}}\sqrt{\frac{\pi}{a}},$$

ya que $\int_{-\infty}^{\infty} e^{-ax^2}\,dx=\sqrt{\frac{\pi}{a}}$.

En particular, tenemos para $a=\frac{1}{2}$

$$f(x)=e^{-\frac{x^2}{2}},\quad g(\lambda)=\sqrt{2\pi}\,e^{-\frac{\lambda^2}{2}},$$

es decir, la función $e^{-\frac{x^2}{2}}$ corresponde a sí misma (salvo un coeficiente constante) en la transformación de Fourier.

2°. Propiedades fundamentales de la transformación de Fourier. De la fórmula (1) que define la transformación de Fourier se desprenden varias propiedades de esta transformación. Consideremos estas propiedades. Para abreviar, designaremos mediante el símbolo $F[f]$ la transformación de Fourier de la función f. En otras palabras, designamos mediante F el operador lineal, definido en el espacio $L_1(-\infty,\infty)$, que pone en correspondencia a toda función de este espacio su transformación de Fourier[1)].

1. *Si una sucesión* $\{f_n\}$ *de funciones de* $L_1(-\infty,\infty)$ *converge en la métrica del espacio* $L_1(-\infty,\infty)$, *la sucesión de las transformaciones de Fourier* $g_n=F[f_n]$ *converge uniformemente en toda la recta.*

---

1) No perteneciente, en general, a L_1.

Esta afirmación se desprende inmediatamente de la estimación evidente

$$|g_n(\lambda)-g_m(\lambda)| \leqslant \int_{-\infty}^{\infty} |f_n(x)-f_m(x)|\,dx.$$

2. *La transformación de Fourier g de una función absolutamente integrable f es una función continua acotada; además, $g(\lambda)$ tiende a cero para $|\lambda| \to \infty$.*

En efecto, de la estimación

$$|g(\lambda)| \leqslant \int_{-\infty}^{\infty} |f(x)|\,dx$$

se ve inmediatamente que la función $g=F[f]$ es acotada. Ahora bien, si f es la función característica del intervalo (a, b), tenemos para ella

$$g(\lambda)=\int_a^b e^{-i\lambda x}\,dx=\frac{e^{-i\lambda a}-e^{-i\lambda b}}{i\lambda}.$$

Esta función es, evidentemente, continua y converge hacia cero para $|\lambda| \to \infty$. Como la operación F consistente en pasar de f a g es lineal, de aquí se deduce que la transformación de Fourier de cualquier función escalonada (esto es, de una combinación lineal de funciones características de intervalos) es también una función continua que tiende a cero para $\lambda \to \pm\infty$. Finalmente, las funciones escalonadas son siempre densas en $L_1(-\infty, \infty)$ y, por lo tanto, si $f \in L_1$, existe una sucesión $\{f_n\}$ de funciones escalonadas convergente hacia f en $L_1(-\infty, \infty)$. Entonces, en virtud de la propiedad 1, la sucesión de funciones $g_n=F[f_n]$ converge uniformemente en la recta hacia la función $g=F[f]$. Pero, en tal caso, la función límite g es también continua y tiende a cero para $|\lambda| \to \infty$.

3. *Si f es absolutamente continua en todo intervalo finito y $f' \in L_1(-\infty, \infty)$, tiene lugar la igualdad*

$$F[f']=i\lambda F[f].$$

Es decir, a la diferenciación de una función le corresponde (en las condiciones señaladas) la multiplicación de su transformación de Fourier por $i\lambda$.

En efecto, una función absolutamente continua en todo intervalo finito puede ser representada en la forma

$$f(x)=f(0)+\int_0^x f'(t)\,dt.$$

La integrabilidad absoluta de f' implica que la expresión que figura aquí en el miembro derecho tenga un límite para $x \to \infty$ y $x \to -\infty$. Este límite puede ser solamente el cero ya que, de lo contrario, la función f no sería integrable en toda la recta. Teniendo esto en cuenta, obtenemos, integrando por partes,

$$F[f'](\lambda) = \int_{-\infty}^{\infty} f'(x) e^{-i\lambda x} dx = f(x) e^{-i\lambda x} \Big|_{-\infty}^{\infty} + i\lambda \int_{-\infty}^{\infty} f(x) e^{-i\lambda x} dx =$$

$$= i\lambda F[f](\lambda),$$

que es lo que se quería demostrar.

Si la función f es tal que $f^{(k-1)}$ es absolutamente continua en todo intervalo y $f, \ldots, f^{(k)} \in L_1(-\infty, \infty)$, obtenemos, por los mismos razonamientos, que

$$F[f^{(k)}] = (i\lambda)^k F[f]. \qquad (5)$$

4. *Relación entre el grado de diferenciabilidad de una función y la velocidad de decrecimiento en el infinito de su transformación de Fourier.* Dividiendo la igualdad (5) por $(i\lambda)^k$ y recordando que la transformación de Fourier es siempre una función que tiende a cero en el infinito (propiedad 2), obtenemos que *si $f^{(k)}$ es absolutamente integrable, se tiene*

$$|F[f]| = \frac{|F[f^{(k)}]|}{|\lambda|^k} \to 0,$$

es decir, bajo estas condiciones, $F[f]$ decrece en el infinito más rápido que $\frac{1}{|\lambda|^k}$. Luego, cuanto más derivadas tenga f en L_1 tanto más rápido decrece en el infinito su transformación de Fourier.

5. *Si f'' existe y pertenece a $L_1(-\infty, \infty)$, es absolutamente integrable $F[f]$.*

En efecto, en estas condiciones, $F[f]$ es acotada y decrece en el infinito más rápido que $\frac{1}{\lambda^2}$. De aquí se desprende la integrabilidad.

Hemos demostrado anteriormente (propiedad 4) que cuanto más derivadas tenga una función f, tanto más rápido decrece en el infinito su transformación de Fourier. Es válida también la afirmación dual, es decir, cuanto más rápido decrece f, tanto mayor grado de diferenciabilidad tiene su transformación de Fourier. Hablando con más precisión:

6. *Supongamos que tanto la función $f(x)$ como $xf(x)$ son absolutamente integrables. Entonces, la función $g = F[f]$ es diferen-*

ciable y

$$g'(\lambda)=F[-ixf(x)]. \tag{6}$$

En efecto, derivando respecto a λ la integral

$$\int_{-\infty}^{\infty} f(x)\, e^{-i\lambda x}\, dx$$

que define g, obtenemos la integral

$$-i\int_{-\infty}^{\infty} x\, f(x)\, e^{-i\lambda x}\, dx$$

que (debido a la integrabilidad de la función $xf(x)$) converge uniformemente respecto a λ. Luego, la derivada de la función g existe y tiene lugar (6).

Cuando f es tal que son absolutamente integrables las funciones $f(x)$, $xf(x)$, ..., $x^p f(x)$, se puede demostrar con razonamientos análogos que la función g tiene derivadas hasta el orden p inclusive y, además,

$$g^{(k)}(\lambda)=F[(-ix)^k f(x)] \quad (k=0,\ 1,\ \ldots,\ p).$$

7. Si exigimos que la función f decrezca en el infinito aún más rápido, g tendrá un grado mayor de diferenciabilidad. Si $x^p f(x)\in L_1(-\infty,\ \infty)$ para todo p, la función g es indefinidamente diferenciable. Supongamos ahora que $e^{\delta|x|} f(x)\in L_1(-\infty,\infty)$ para cierto $\delta>0$. En este caso, $g(\lambda)$ puede ser prolongada, como una función analítica, del eje real λ a una franja del plano de la variable compleja $\zeta=\lambda+i\mu$, con la particularidad de que la anchura de esta franja es tanto mayor cuanto más grande sea δ. En todo caso, se puede afirmar que g será una función analítica para $|\mu|<\delta$. En efecto, la integral

$$\int_{-\infty}^{\infty} f(x)\, e^{-ix\zeta}\, dx$$

converge, evidentemente, para $|\mu|<\delta$ definiendo una función continua que coincide en el eje real con la transformación de Fourier de la función f. El hecho de que esta función es diferenciable para $|\mu|<\delta$ en el sentido de la teoría de funciones analíticas se demuestra igual que la propiedad 6.

3°. Complitud de las funciones de Hermite y de Laguerre. Empleando las ideas expuestas en el párrafo anterior se puede demostrar que *si una función medible f es diferente de* 0 *en casi todo punto de un intervalo* $(a,\ b)$, *donde* $-\infty\leqslant a<b\leqslant\infty$

y satisface la condición $|f(x)| \leqslant Ce^{-\delta|x|}$, *donde* $\delta > 0$, *entonces el sistema de funciones* $x^n f(x)$, $n = 0, 1, 2, \ldots$, *es completo en* $L_2(a, b)$.

De aquí se desprenderá, en particular, que las funciones de Hermite constituyen un sistema completo en $L_2(-\infty, \infty)$ y las funciones de Laguerre, en $L_2(0, \infty)$ (véase el punto 7 del § 3 del cap. VIII).

Demostremos la proposición sobre la complitud enunciada más arriba. Supongamos que el sistema $\{x^n f(x)\}$ no es completo. Entonces, de acuerdo con el teorema de Hahn—Banach, existe una función no nula $h \in L_2(-\infty, \infty)$ tal que

$$\int_{-\infty}^{\infty} x^n f(x) h(x)\, dx = 0 \quad (n = 0, 1, 2, \ldots).$$

(Hemos empleado aquí el teorema sobre la forma general de una funcional lineal continua en el espacio de Hilbert; si se considera el espacio complejo $L_2(a, b)$ habrá que escribir $\overline{h(x)}$ en lugar de $h(x)$.). Está claro que $fh \in L_1(a, b)$; es más, tenemos $e^{\delta_1|x|} fh \in L_1(a, b)$ para cualquier $\delta_1 < \delta$. En lo sucesivo conviene aceptar que $a = -\infty$ y $b = \infty$ tomando, si es necesario, iguales a cero las funciones f y h fuera de (a, b). Sea g la transformación de Fourier de la función fh, es decir,

$$g(\lambda) = \int_{-\infty}^{\infty} f(x) h(x) e^{-i\lambda x}\, dx.$$

De lo explicado anteriormente se desprende que la función g puede ser prolongada, como una función analítica, a la franja $|\operatorname{Im} \zeta| < \delta$. Por otro lado, en virtud de la propiedad 6, todas las derivadas de esta función son iguales a 0 para $\lambda = 0$, de manera que $g(\lambda) \equiv 0$. Por la propiedad de unicidad demostrada en el punto 1, de aquí se deduce que $f(x) h(x) = 0$ en casi todo punto y, consecuentemente, $h(x) = 0$ en casi todo punto ya que $f(x)$ es diferente de 0 en casi todo punto. Pero, esto contradice a nuestra suposición de que h es una función no nula. La contradición obtenida demuestra la complitud del sistema $\{x^n f(x)\}$.

4°. Transformación de Fourier de funciones indefinidamente diferenciables y rápidamente decrecientes. Teniendo en cuenta que al pasar de la función f a su transformación de Fourier g las propiedades de diferenciabilidad y de decrecimiento en el infinito corresponden una a otra, es fácil indicar clases naturales de funciones que se aplican en sí mismas por la transformación de Fourier.

Sea S_∞ el conjunto de funciones indefinidamente diferenciables en la recta, para cada una de las cuales existe una constante C_{pq} (que depende de la función f y de los números p y q) tal que

$$|x^p f^{(q)}(x)| < C_{pq}. \tag{7}$$

Probemos que, si $f \in S_\infty$, también $g = F[f] \in S_\infty$. Observemos, ante todo, que de acuerdo con (7) cada una de las funciones

$$x^p f^{(q)}(x)$$

es absolutamente integrable. En efecto, como (7) se cumple para todos los p y q, la función

$$x^{p-2} f^{(q)}(x)$$

es acotada, esto es, $x^p f^{(q)}(x)$ decrece no más lento que $\frac{1}{x^2}$; consecuentemente, la función $F[f]$ tiene derivadas de todos los órdenes. Finalmente, de acuerdo con el punto 2, la sumabilidad de $f^{(q)}(x)$, $q = 1, 2, \ldots$ implica que $g = F[f]$ decrezca en el infinito más rápido que $\frac{1}{|\lambda|^q}$. Consideremos ahora las funciones

$$(i\lambda)^q g^{(p)}(\lambda) = (-i)^q F[(x^p f(x))^{(q)}];$$

cada una de las cuales es acotada por una constante D_{pq}, como la transformación de Fourier de una función integrable. De modo que, si $f \in S_\infty$, también $g = F[f] \in S_\infty$. Viceversa, sea $g \in S_\infty$; entonces, de acuerdo con lo demostrado, la función

$$f^*(x) = \int_{-\infty}^{\infty} g(\lambda) e^{-i\lambda x} dx$$

pertenece a S_∞. Pongamos $f(x) = \frac{1}{2\pi} f^*(-x)$. Está claro que $f \in S_\infty$. Al mismo tiempo, por la fórmula de inversión

$$g(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f^*(x) e^{i\lambda x} dx,$$

es decir, g es la transformación de Fourier de la función $f \in S_\infty$. Luego, la transformación de Fourier aplica la clase S_∞ en la misma clase S_∞. Está claro que esta aplicación es biunívoca.

EJERCICIO. Sea $f \in S_\infty$ y $\int_{-\infty}^{\infty} x^p f(x)\,dx = 0$ para todo $p \geqslant 0$. ¿Se desprende de aquí que $f(x) \equiv 0$?

5°. Transformación de Fourier y convolución de funciones. Sean f_1 y f_2 funciones integrables en toda la recta. La función

$$f(x)=\int_{-\infty}^{\infty} f_1(\xi) f_2(x-\xi)\,d\xi$$

se llama *convolución* de estas funciones. La función $f(x)$ está definida para casi todo x y es integrable. En efecto, la integral doble

$$\iint f_1(\xi) f_2(x-\xi)\,d\xi\,dx$$

existe, ya que existe la integral

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |f_1(\xi) f_2(\eta)|\,d\xi\,d\eta$$

(véase la observación al teorema de Fubini en el cap. VI). Consecuentemente, existe también la integral

$$\int_{-\infty}^{\infty} f(x)\,dx=\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} f_1(\xi) f_2(x-\xi)d\xi.$$

La función f se designa mediante el símbolo f_1*f_2. Calculemos la transformación de Fourier de la convolución de dos funciones de L_1. Aplicando el teorema de Fubini y tomando $x-\xi=\eta$, obtenemos

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)e^{-i\lambda x}\,dx &= \int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty} f_1(\xi) f_2(x-\xi)d\xi\right\} e^{-i\lambda x}\,dx=\\ &=\int_{-\infty}^{\infty} f_1(\xi)\left\{\int_{-\infty}^{\infty} f_2(x-\xi)e^{-i\lambda x}\,dx\right\}d\xi=\\ &=\int_{-\infty}^{\infty} f_1(\xi)\left\{\int_{-\infty}^{\infty} f_2(\eta)e^{-i\lambda\eta}e^{-i\lambda\xi}\,d\eta\right\}d\xi=\\ &=\int_{-\infty}^{\infty} f_2(\eta)e^{-i\lambda\eta}d\eta\cdot\int_{-\infty}^{\infty} f_1(\xi)e^{-i\lambda\xi}\,d\xi,\end{aligned}$$

es decir,

$$F[f_1*f_2]=F[f_1]\cdot F[f_2].$$

Luego, la transformación de Fourier convierte la operación de convolución en una operación más simple: la multiplicación de funciones. Este hecho desempeña un papel importante en varias aplicaciones de la transformación de Fourier.

6°. Aplicación de la transformación de Fourier a la solución de la ecuación de conducción del calor. La aplicación de la transformación de Fourier a ecuaciones diferenciales se basa en que, según hemos demostrado en el punto 3, esta transformación convierte la operación de diferenciación en la de multiplicación por la variable independiente. Luego, si tenemos una ecuación diferencial lineal de coeficientes constantes

$$y^{(n)}+a_1y^{(n-1)}+\ldots+a_{n-1}y'+a_ny=\varphi(x), \tag{8}$$

ella es convertida por la transformación de Fourier en una ecuación algebraica de tipo

$$(i\lambda)^nz+a_1(i\lambda)^{n-1}z+\ldots+a_{n-1}i\lambda z+a_nz=\psi(\lambda), \text{ donde } \psi=F[\varphi]. \tag{9}$$

Sin embargo, este método es poco fecundo en el caso de ecuaciones diferenciales ordinarias ya que la solución de ecuaciones lineales de coeficientes constantes, a las que puede ser aplicado, no ofrece, por sí misma, grandes dificultades. Además, el paso de (8) a (9) es posible, si la función incógnita $y=y(x)$ es integrable en toda la recta, mientras que para las soluciones de ecuaciones lineales de coeficientes constantes esto, como regla general, no tiene lugar.

Es de mayor importancia la aplicación de la transformación de Fourier a ecuaciones en derivadas parciales, donde permite, en ciertas condiciones, reducir la solución de una ecuación de este tipo a la solución de una ecuación diferencial ordinaria. Ilustremos esto resolviendo el problema de Cauchy para la ecuación de conducción del calor. Busquemos para $-\infty<x<\infty$ y $t\geqslant 0$ la solución de la ecuación

$$\frac{\partial u(x, t)}{\partial t}=\frac{\partial^2u(x, t)}{\partial x^2} \tag{10}$$

que se convierte para $t=0$ en una función prefijada $u_0(x)$. El contenido físico de este problema consiste en determinar la temperatura de una varilla termoconductiva infinita para cualquier momento $t>0$, si en el momento inicial $t=0$ su temperatura en cada punto es $u_0(x)$.

Suponiendo que $u_0(x)$, $u_0'(x)$ y $u_0''(x)$ pertenecen a $L_1(-\infty, \infty)$, buscaremos la solución del problema planteado en la clase de funciones $u(x, t)$ que satisfacen las condiciones siguientes:

1) Las funciones $u(x, t)$, $u_x(x, t)$ y $u_{xx}(x, t)$ son absolutamente integrables en todo el eje x para cualquier $t\geqslant 0$ fijado.

2) La función $u_t(x, t)$ tiene en todo intervalo finito $0 \leqslant t \leqslant T$ una mayorante integrable $f(x)$ (que no depende de t):

$$|u_t(x, t)| \leqslant f(x), \quad \int_{-\infty}^{\infty} f(x)\,dx < \infty.$$

Realicemos en la ecuación (10) la transformación de Fourier respecto a x. Entonces, en el miembro derecho tendremos

$$F[u_{xx}(x, t)] = -\lambda^2 v(\lambda, t), \text{ donde } v(\lambda, t) = F[u(x, t)],$$

mientras que en el miembro izquierdo, en virtud de 2), tendremos

$$F[u_t] = \int_{-\infty}^{\infty} u_t(x, t)\, e^{-i\lambda x}\,dx = \frac{\partial}{\partial t}\int_{-\infty}^{\infty} u(x, t)\, e^{-i\lambda x}\,dx = v_t(\lambda, t).$$

De esta forma la transformación de Fourier convierte la ecuación (10) en una ecuación diferencial ordinaria

$$v_t(\lambda, t) = -\lambda^2 v(\lambda, t)$$

y debemos ahora buscar la solución de esta ecuación que para $t=0$ da

$$v_0(\lambda) = F[u_0(x)] = \int_{-\infty}^{\infty} u_0(x)\, e^{-i\lambda x}\,dx.$$

Esta solución es, evidentemente,

$$v(\lambda, t) = e^{-\lambda^2 t} v_0(\lambda).$$

Ahora, para obtener la solución del problema inicial, resta encontrar aquella función $u(x, t)$ cuya transformación de Fourier es la función encontrada $v(\lambda, t)$.

Recurriendo al ejemplo 4 del punto 1, tenemos

$$e^{-\lambda^2 t} = F\left[\frac{1}{2\sqrt{\pi t}}\, e^{-\frac{x^2}{4t}}\right].$$

Por lo tanto,

$$v(\lambda, t) = F\left[\frac{1}{2\sqrt{\pi t}}\, e^{-\frac{x^2}{4t}}\right] \cdot F[u_0(x)] = F\left[\frac{1}{2\sqrt{\pi t}}\, e^{-\frac{x^2}{4t}} * u_0(x)\right],$$

es decir,

$$u(x, t) = \frac{1}{2\sqrt{\pi t}} \int_{-\infty}^{\infty} e^{-\frac{\xi^2}{4t}}\, u_0(x-\xi)\,d\xi.$$

Hemos obtenido la así llamada *fórmula de Poisson* para la solución de la ecuación de conducción del calor.

7°. Transformación de Fourier de funciones de varias variables. El concepto de la transformación de Fourier, considerado anteriormente para funciones de una variable, puede ser extendido fácilmente al caso de funciones de varias variables.

Sea $f(x_1, x_2, \ldots, x_n)$ una función integrable en todo el espacio n-dimensional R^n. Su transformación de Fourier es la función

$$g(\lambda_1, \lambda_2, \ldots, \lambda_n)=$$
$$=\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_n)e^{-i(x_1\lambda_1+x_2\lambda_2+\ldots+x_n\lambda_n)}\,dx_1\ldots dx_n.$$

Esta integral n-ple, existente de antemano ya que $f(x_1, x_2, \ldots$ $\ldots, x_n)$ es integrable, puede ser representada, de acuerdo con el teorema de Fubini, mediante la siguiente integral reiterada:

$$g(\lambda_1, \lambda_2, \ldots, \lambda_n)=\int_{-\infty}^{\infty}\left\{\ldots\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty} f(x_1, x_2, \ldots\right.\right.$$
$$\left.\left.\ldots, x_n)e^{-ix_1\lambda_1}dx_1\right\}e^{-ix_2\lambda_2}dx_2\ldots\right\}e^{-ix_n\lambda_n}dx_n. \qquad (11)$$

En otras palabras, el paso de una función de n variables a su transformación de Fourier puede ser realizado aplicando sucesivamente esta transformación a cada variable (en un orden cualquiera). Invirtiendo cada una de las n operaciones sucesivas que figuran en el miembro derecho de (11), obtenemos la fórmula

$$f(x_1, x_2, \ldots, x_n)=\frac{1}{(2\pi)^n}\int_{-\infty}^{\infty}\left\{\ldots\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty} g(\lambda_1, \lambda_2, \ldots\right.\right.$$
$$\left.\left.\ldots, \lambda_n)e^{ix_n\lambda_n}d\lambda_n\right\}e^{ix_{n-1}\lambda_{n-1}}d\lambda_{n-1}\ldots\right\}e^{ix_1\lambda_1}d\lambda_1.$$

que puede ser representada mediante una integral n-ple

$$f(x_1, x_2, \ldots, x_n)=\frac{1}{(2\pi)^n}\int\ldots$$
$$\ldots\int g(\lambda_1, \lambda_2, \ldots, \lambda_n)e^{i(x_1\lambda_1+x_2\lambda_2+\ldots+x_n\lambda_n)}\,d\lambda_1\ldots d\lambda_n; \qquad (12)$$

sin embargo, puesto que la función $g(\lambda_1, \lambda_2, \ldots, \lambda_n)$ puede no ser, en general, sumable en todo el R^n, es preciso indicar en qué sentido debe comprenderse esta integral y las condiciones que deben imponerse a $f(x_1, x_2, \ldots, x_n)$ para que sea posible representar $f(x_1, x_2, \ldots, x_n)$ mediante la integral (12).

Una de las respuestas posibles a estos problemas la da el teorema siguiente.

TEOREMA 1. *Supongamos que la función $f(x_1, x_2, \ldots, x_n)$ es integrable en todo el espacio R^n y satisface las condiciones siguientes:*

$$\left.\begin{aligned}
&|f(x_1+t_1, x_2, \ldots, x_n)-f(x_1, x_2, \ldots, x_n)| \leqslant C|t_1|^\alpha,\\
&|f(x_1, x_2+t_2, \ldots, x_n)-f(x_1, x_2, \ldots, x_n)| \leqslant\\
&\qquad\qquad\qquad\qquad \leqslant C(x_1)|t_2|^\alpha,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&|f(x_1, x_2, \ldots, x_n+t_n)-f(x_1, x_2, \ldots, x_n)| \leqslant\\
&\qquad\qquad \leqslant C(x_1, x_2, \ldots, x_{n-1})|t_n|^\alpha,\\
&\text{donde } 0\leqslant\alpha\leqslant 1, \quad \int_{-\infty}^{\infty} C(x_1)\,dx_1 < \infty, \ \ldots,\\
&\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} C(x_1, x_2, \ldots, x_{n-1})\,dx_1dx_2\ldots dx_{n-1} < \infty.
\end{aligned}\right\} \qquad (13)$$

Entonces, la fórmula de inversión (12) *es válida, si la integral en ella se entiende como*

$$\frac{1}{(2\pi)^n}\cdot\lim_{N_1\to\infty}\int_{-N_1}^{N_1}\left\{\ldots\lim_{N_{n-1}\to\infty}\int_{-N_{n-1}}^{N_{n-1}}\left\{\lim_{N_n\to\infty}\int_{-N_n}^{N_n} g(\lambda_1, \lambda_2, \ldots, \lambda_n)\times\right.\right.$$

$$\left.\left.\times e^{ix_n\lambda_n}\,d\lambda_n\right\} e^{ix_{n-1}\lambda_{n-1}}\,d\lambda_{n-1}\ \ldots\right\} e^{ix_1\lambda_1}\,d\lambda_1.$$

En efecto, como $f(x_1, x_2, \ldots, x_n)$ es sumable en R^n, es sumable también, de acuerdo con el teorema de Fubini, respecto a x_1 para casi todos los $x_2, \ldots, x_n$. Luego, existe la función

$$f_1(\lambda_1, x_2, \ldots, x_n) = \int_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_n)\,e^{-ix_1\lambda_1}dx_1,$$

De (13) se desprende que $f(x_1, x_2, \ldots, x_n)$ como función de x_1 verifica la condición del teorema 1 del § 3; por lo tanto, $f(x_1, x_2, \ldots, x_n)$ puede ser expresada a través de f_1 mediante la fórmula de inversión

$$f(x_1, x_2, \ldots, x_n) = \lim_{N_1 \to \infty} \frac{1}{2\pi} \int_{-N_1}^{N_1} f_1(\lambda_1, x_2, \ldots, x_n) e^{ix_1\lambda_1} d\lambda_1.$$

Ahora bien, si tomamos

$$f_2(\lambda_1, \lambda_2, \ldots, x_n) = \int_{-\infty}^{\infty} f_1(\lambda_1, x_2, \ldots, x_n) e^{-ix_2\lambda_2} dx_2,$$

de la condición (13) se deducirá que para f_1 es válida la fórmula de inversión

$$f_1(\lambda_1, x_2, \ldots, x_n) = \lim_{N_2 \to \infty} \frac{1}{2\pi} \int_{-N_2}^{N_2} f_2(\lambda_1, \lambda_2, \ldots, x_n) e^{ix_2\lambda_2} d\lambda_2,$$

es decir,

$$f(x_1, x_2, \ldots, x_n) = \lim_{N_1 \to \infty} \frac{1}{2\pi} \int_{-N_1}^{N_1} \left\{ \lim_{N_2 \to \infty} \frac{1}{2\pi} \int_{-N_2}^{N_2} f_2(\lambda_1, \lambda_2, \ldots, x_n) \times e^{ix_2\lambda_2} d\lambda_2 \right\} e^{ix_1\lambda_1} d\lambda_1.$$

Definiendo de un modo análogo $f_3(\lambda_1, \lambda_2, \ldots, x_n)$, etc., obtendremos la fórmula (12).

La transformación de Fourier de funciones de varias variables encuentra amplia aplicación en la teoría de ecuaciones en derivadas parciales. Consideremos, por ejemplo, la ecuación

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \tag{14}$$

que describe la conducción del calor en un plano. Sea dada la temperatura en el momento $t = 0$:

$$u(0, x, y) = u_0(x, y).$$

Sometiendo la solución que buscamos de la ecuación (14) a condiciones análogas a las indicadas en el punto 5, podemos efectuar en la ecuación (14) la transformación de Fourier respecto a las variables x e y. Obtendremos como resultado la ecuación ordinaria

$$\frac{dv}{dt} = -(\lambda^2 + \sigma^2) v,$$

donde

$$v(t, \lambda\sigma)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u(t, x, y)e^{-i(\lambda x+\sigma y)}\,dx\,dy. \tag{15}$$

Resolviendo la ecuación (15), podemos después encontrar la solución de la ecuación inicial (14) por medio de la fórmula de inversión.

§ 5. TRANSFORMACION DE FOURIER EN EL ESPACIO $L_2(-\infty, \infty)$

1°. Teorema de Plancherel. Volvamos primero a los resultados obtenidos para las series de Fourier. Para una mayor analogía con la transformación de Fourier, consideraremos la forma compleja de la serie de Fourier, esto es, tomaremos en el segmento $[-\pi, \pi]$ el sistema completo ortogonal de funciones e^{inx}, $n=0, \pm 1, \ldots$, y pondremos en correspondencia a toda función f sumable en el segmento $[-\pi, \pi]$ su sucesión de coeficientes de Fourier

$$c_n=\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx \quad (n=0, \pm 1, \pm 2, \ldots).$$

Si la funcion f, además de ser sumable, es de cuadrado sumable, sus coeficientes de Fourier verifican la condición

$$\sum_{n=-\infty}^{\infty} |c_n|^2<\infty.$$

En otras palabras, el paso de una función de cuadrado sumable al conjunto de sus coeficientes de Fourier constituye una aplicación del espacio euclídeo L_2 sobre el espacio euclídeo l_2; además, esta aplicación es lineal y verifica la igualdad de Parseval:

$$2\pi\sum_{n=-\infty}^{\infty} |c_n|^2=\int_{-\pi}^{\pi} |f(x)|^2\,dx \tag{1}$$

(es decir, este paso difiere sólo en un coeficiente constante de una aplicación que conserva la norma).

Consideremos ahora la transformación de Fourier para funciones definidas en toda la recta y, veamos, si es posible interpretar esta transformación como un operador lineal en $L_2(-\infty, \infty)$. La dificultad principal consiste aquí en que una función de cuadrado integrable no pertenece necesariamente a $L_1(-\infty, \infty)$, es decir, en que su transformación de Fourier puede no existir en el sentido definido en el § 4. Sin embargo, es posible definir, en cierto sentido, la transformación de Fourier para toda $f \in L_2(-\infty, \infty)$. Entonces, se obtiene el siguiente teorema que puede ser considerado como un análogo de la igualdad de Parseval (1).

TEOREMA. (Plancherel, 1910). *Para cualquier función $f \in L_2(-\infty, \infty)$ la integral*

$$g_N(\lambda) = \int_{-N}^{N} f(x) e^{-i\lambda x}\, dx$$

es, cualquiera que sea N, una función de λ perteneciente a $L_2(-\infty, \infty)$. Para $N \to \infty$ las funciones g_N convergen en la métrica del espacio L_2 hacia un límite g tal que

$$\int_{-\infty}^{\infty} |g(\lambda)|^2 d\lambda = 2\pi \int_{-\infty}^{\infty} |f(x)|^2 dx. \tag{2}$$

Esta función g es llamada transformación de Fourier de la función $f \in L_2$. Si f pertenece también a $L_1(-\infty, \infty)$, la función correspondiente g coincide con la transformación habitual de Fourier de la función f.

DEMOSTRACION. La idea principal de la demostración consiste en que la igualdad (2) es comprobada primero para todas las funciones que pertenecen a la clase S_∞ de funciones indefinidamente diferenciables y rápidamente decrecientes, que son siempre densas en $L_2(-\infty, \infty)$, después es extendida, por continuidad, a todo el $L_2(-\infty, \infty)$. Realicemos ahora detalladamente esta idea.

1) Sean $f_1, f_2 \in S_\infty$. Designemos con g_1 y g_2, respectivamente, sus transformaciones de Fourier. Tenemos

$$\int_{-\infty}^{\infty} f_1(x)\overline{f_2(x)}\, dx = \int_{-\infty}^{\infty} \frac{1}{2\pi} \int_{-\infty}^{\infty} [g_1(\lambda) e^{i\lambda x} d\lambda]\, \overline{f_2(x)}\, dx =$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[g_1(\lambda) \overline{\int_{-\infty}^{\infty} f_2(x) e^{-i\lambda x} dx} \right] d\lambda = \frac{1}{2\pi} \int_{-\infty}^{\infty} g_1(\lambda)\overline{g_2(\lambda)}\, d\lambda,$$

donde el cambio del orden de integración está justificado ya que la función

$$g_1(\lambda)\overline{f_2(x)}\, e^{i\lambda x}$$

es absolutamente integrable en el plano (x, λ). Tomando $f_1 = f_2 = f$

y $g_1 = g_2 = g$ en la igualdad obtenida, encontramos que la fórmula (2) es válida para cualquier función $f \in S_\infty$.

2) Sea ahora f una función arbitraria de $L_2(-\infty, \infty)$ que se anula fuera de un intervalo $(-a, a)$. Entonces, f es integrable en el intervalo $(-a, a)$ (es decir, pertenece a $L_1(-a, a)$) y, consecuentemente, en toda la recta. Luego, para ella está definida la transformación de Fourier

$$g(\lambda) = \int_{-\infty}^{\infty} f(x) e^{-i\lambda x} dx.$$

Sea ahora $\{f_n\}$ una sucesión de funciones de S_∞, nulas fuera de un intervalo $(-a, a)$, que converge hacia f respecto a la norma del espacio $L_2(-\infty, \infty)$. Como f y toda f_n son diferentes de cero sólo en un intervalo finito, la sucesión $\{f_n\}$ converge hacia f también respecto a la norma del espacio $L_1(-\infty, \infty)$. Por lo tanto, la sucesión $\{g_n\}$ converge hacia g uniformemente en toda la recta (véase el punto 2 del § 3). Además, la sucesión $\{g_n\}$ es fundamental en $L_2(-\infty, \infty)$. En efecto, $g_n - g_m \in S_\infty$; luego, en virtud de lo ya demostrado, tenemos

$$\int_{-\infty}^{\infty} |g_n(\lambda) - g_m(\lambda)|^2 d\lambda = 2\pi \int_{-\infty}^{\infty} |f_n(x) - f_m(x)|^2 dx,$$

de donde se desprende que $\{g_n\}$ es una sucesión fundamental. Ello significa que esta sucesión converge en L_2 y, además, hacia la misma función g, hacia la cual converge uniformemente. Por esto, podemos pasar al límite para $n \to \infty$ en la igualdad

$$\|f_n\|^2 = \frac{1}{2\pi} \|g_n\|^2.$$

Luego, la igualdad (2) es válida para toda $f \in L_2$ que se anula fuera de un intervalo.

3) Sea, finalmente, f una función arbitaria de L_2. Tomemos

$$f_N(x) = \begin{cases} f(x) & \text{para} \quad |x| \leqslant N, \\ 0 & \text{para} \quad |x| > N. \end{cases}$$

Está claro que

$$\|f - f_N\| \to 0 \quad \text{para} \quad N \to \infty.$$

La función f_N pertenece a $L_1(-\infty, \infty)$ y, consecuentemente, para ella existe la transformación de Fourier que es igual a

$$g_N(\lambda) = \int_{-\infty}^{\infty} f_N(x) e^{-i\lambda x} dx = \int_{-N}^{N} f(x) e^{i\lambda x} dx.$$

Pero, en virtud del punto 2) de nuestros razonamientos,

$$\|f_N - f_M\|^2 = \frac{1}{2\pi}\|g_N - g_M\|^2,$$

es decir, las funciones g_N convergen en L_2 hacia un límite que designaremos con g. Por lo tanto, en la igualdad

$$\|f_N\|^2 = \frac{1}{2\pi}\|g_N\|^2$$

se puede pasar al límite para $N \to \infty$, de donde se obtiene la relación (2) para cualquier $f \in L_2(-\infty, \infty)$. Si ahora f pertenece tanto a $L_2(-\infty, \infty)$ como a $L_1(-\infty, \infty)$, existe para ella la transformación de Fourier

$$\tilde{g}(\lambda) = \int_{-\infty}^{\infty} f(x) e^{-i\lambda x}\, dx,$$

comprendida en el sentido corriente. Como las funciones f_N convergen en $L_1(-\infty, \infty)$ hacia f, sus transformaciones de Fourier g_N convergen uniformemente hacia $\tilde{g}$. Pero, hemos demostrado, además, que las funciones g_N convergen respecto a la métrica de $L_2(-\infty, \infty)$ hacia un límite que hemos designado con g. De aquí se desprende que $\tilde{g}$ coincide con g. Hemos terminado la demostración.

De la relación (2) se deduce inmediatamente que para cualesquiera $f_1, f_2 \in L_2(-\infty, \infty)$ se cumple la igualdad

$$\int_{-\infty}^{\infty} f_1(x)\overline{f_2(x)}\, dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} g_1(\lambda)\overline{g_2(\lambda)}\, d\lambda.$$

Para demostrarla basta escribir la igualdad (2) para la función $f_1 + f_2$ y comparar después las expresiones en los miembros derecho e izquierdo.

2°. **Funciones de Hermite.** El teorema de Plancherel, expuesto en el punto anterior, señala que la transformación de Fourier puede ser considerada como un operador lineal acotado F que transforma el espacio $L_2(-\infty, \infty)$ en sí mismo. Escogiendo en este espacio un sistema ortonormal completo, podemos definir este operador F (al igual que cualquier otro operador lineal) mediante la matriz infinita correspondiente. La forma de esta matriz depende, claro está, de cómo se escoge la base. Una matriz, correspondiente a uno u otro operador, adquiere su forma más sencilla si la base correspondiente está compuesta por las funciones propias del operador dado: en este caso, la matriz es de la forma diagonal. Veamos si existe una base de este tipo para la transformación de Fourier F. En otras palabras, veamos qué funciones de $L_2(-\infty, \infty)$ son propias para la transformación de Fourier F. Observemos, con este fin, que la ecuación

$$\frac{d^2f}{dx^2}-x^2f=\mu f \tag{3}$$

se convierte mediante la transformación de Fourier en una ecuación del mismo tipo $\left(\text{ya que la operación } \frac{d^2}{dx^2} \text{ corresponde a la multiplicación por } -\lambda^2 \text{ y la multiplicación por } -x^2 \text{ corresponde a la operación } \frac{d^2}{d\lambda^2}\right)$[1]. Por esto resulta natural buscar las funciones propias del operador F como soluciones de la ecuación (3). Busquemos las soluciones de esta ecuación que tienen la forma

$$f=we^{-\frac{x^2}{2}},$$

donde w es un polinomio. Introduciendo en (3) esta expresión, obtenemos para w la ecuación

$$w''-2xw'=(\mu+1)\,w.$$

Tomando

$$w=a_0+a_1x+\ldots+a_nx^n, \tag{4}$$

encontramos la igualdad

$$(2a_2+3\cdot 2\cdot a_3x+\ldots+n(n-1)\,a_nx^{n-2})-$$
$$-2x\,(a_1+2a_2x+\ldots+na_nx^{n-1})=(\mu+1)\,(a_0+a_1x+\ldots+a_nx^n).$$

Comparando en ella los coeficientes de potencias iguales de x en los miembros izquierdo y derecho, encontramos que

$$-2na_n=(\mu+1)\,a_n,\quad -2\,(n-1)\,a_{n-1}=(\mu+1)\,a_{n-1},$$

etc.; en general,

$$k\,(k-1)\,a_k-2\,(k-2)\,a_{k-2}=(\mu+1)\,a_{k-2}. \tag{5}$$

Como el coeficiente a_n se supone distinto de cero, tenemos

$$\mu=(2n+1) \quad \text{y} \quad a_{n-1}=0,$$

es decir, μ debe ser un número entero negativo impar. Todos los coeficientes del polinomio w se determinan por las relaciones (5) unívocamente, salvo un coeficiente constante. Además,

---

[1] Suponiendo, claro está, que la función incógnita f verifica las condiciones correspondientes de diferenciabilidad y de decrecimiento en el infinito.

son iguales a cero todos aquellos coeficientes cuyos subíndices tienen paridad opuesta a la del número n, es decir, de la potencia del polinomio w. Al contrario, los coeficientes cuyos subíndices tienen la misma paridad que n son distintos de cero y se calculan por la fórmula recurrente

$$a_{k-2}=\frac{k(k-1)}{2k-2n-4}a_k$$

(para el valor dado de a_n). De esta forma obtenemos para w la expresión siguiente

$$w_n(x)=a_n\left(x^n-\frac{n-(n-1)}{4}x^{n-2}+\frac{n(n-1)(n-2)(n-3)}{4\cdot 8}x^{n-4}-\ldots\right).$$

Hemos construido, pues, el sistema de funciones de tipo

$$\varphi_n(x)=w_n(x)e^{-\frac{x^2}{2}}\ (n=0,\ 1,\ 2,\ \ldots).$$

Es obvio que cada una de estas funciones pertenece a $L_2(-\infty, \infty)$ $\left(\text{ya que decrece en el infinito más rápido que cualquier potencia de } \frac{1}{|x|}\right)$. Además, estas funciones son dos a dos ortogonales. En efecto, de acuerdo con (3), tenemos

$$\varphi_n''(x)-x^2\varphi_n(x)=-(2n+1)\varphi_n(x),$$
$$\varphi_m''(x)-x^2\varphi_m(x)=-(2m+1)\varphi_m(x).$$

Multiplicando la primera de estas igualdades por φ_m y la segunda por φ_n y restándolas una de otra, obtenemos

$$\varphi_n''\varphi_m-\varphi_m''\varphi_n=2(n-m)\varphi_n\varphi_m.$$

Teniendo en cuenta esta igualdad, encontramos para $n\neq m$, integrando por partes,

$$\int_{-\infty}^{\infty}\varphi_n(x)\varphi_m(x)\,dx=\frac{1}{2(n-m)}\int_{-\infty}^{\infty}[\varphi_n''\varphi_m-\varphi_m''\varphi_n]\,dx=$$
$$=\frac{1}{2(n-m)}\left\{[\varphi_n'\varphi_m-\varphi_m'\varphi_n]_{-\infty}^{\infty}-\int_{-\infty}^{\infty}[\varphi_n'\varphi_m'-\varphi_m'\varphi_n']\,dx=0.\right.$$

Con esto queda demostrada la ortogonalidad.

De esta forma cada uno de los elementos φ_n del sistema ortogonal obtenido es un polinomio de grado n multiplicado por $e^{-\frac{x^2}{2}}$. Luego, los elementos de este sistema deben coincidir, salvo unos factores numéricos, con las funciones de Hermite que hemos construido en el § 3 del cap. VII mediante la ortogonalización de la sucesión

$$e^{-\frac{x^2}{2}},\quad xe^{-\frac{x^2}{2}},\ \ldots,\ x^n e^{-\frac{x^2}{2}},\ \ldots$$

en el espacio $L_2(-\infty, \infty)$.

Probemos ahora que las funciones φ_n son funciones propias de la transformación de Fourier:

$$F\varphi_n = c_n \varphi_n. \tag{6}$$

Esto se deduce de los siguientes hechos.

1) La ecuación (3) es invariante respecto a la transformación F.

2) Para todo n la ecuación (3) tiene una solución única, salvo un coeficiente constante, de tipo $P_n(x)e^{-\frac{x^2}{2}}$, donde P_n es un polinomio de grado n.

3) La transformación de Fourier convierte $x^n e^{-\frac{x^2}{2}}$ en $\left(i\frac{d}{dx}\right)^n e^{-\frac{x^2}{2}} = Q_n(x)e^{-\frac{x^2}{2}}$, donde Q_n es un polinomio de grado n (la última afirmación se comprueba fácilmente por inducción). De la igualdad (6) se deduce que para todo k entero

$$F^k \varphi_n = c_n^k \varphi_n.$$

Pero, la transformación de Fourier, aplicada cuatro veces consecutivas, convierte toda función en sí misma multiplicada por $4\pi^2$ Por lo tanto,

$$c_n^4 = 4\pi^2$$

es decir, c_n puede tomar solamente los valores $\pm\sqrt{2\pi}$ y $\pm i\sqrt{2\pi}$. Luego, en $L_2(-\infty, \infty)$ la transformación de Fourier es un operador lineal que en la base compuesta por las funciones de Hermite se define mediante una matriz diagonal, en la que los elementos diagonales toman los valores $\pm\sqrt{2\pi}$ y $\pm i\sqrt{2\pi}$ [1].

---

[1] Si la transformación de Fourier se define mediante la fórmula

$$F[f] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-i\lambda x}\, dx$$

(esto es, mediante la fórmula (1') del § 4 y no mediante la fórmula (1)), su cuarta potencia será el operador unidad y obtenemos para F, en la base compuesta por las funciones de Hermite, la matriz diagonal con elementos ± 1 y $\pm i$.

§ 6. TRANSFORMACION DE LAPLACE

1°. Definición y propiedades fundamentales de la transformación de Laplace. Las posibilidades de aplicar la transformación de Fourier a ecuaciones diferenciales y a otras varias cuestiones están considerablemente limitadas debido a que la transformación de Fourier está definida sólo para funciones sumables en toda la recta. En particular, la transformación de Fourier no existe para funciones que tienden al infinito para $x \to -\infty$ o $x \to \infty$. Al mismo tiempo, al resolver ecuaciones diferenciales aparecen frecuentemente funciones de este tipo. Uno de los caminos posibles que permiten superar esta dificultad es extender el concepto de transformación de Fourier a las funciones generalizadas y acerca de él hablaremos brevemente en el § 8 de este capítulo. Otro camino posible, que nos permite quedarnos en los márgenes del concepto clásico de función y de los métodos clásicos del Análisis, consiste en sustituir la transformación de Fourier por la así llamada transformación de Laplace.

Supongamos que una función f (no integrable, en general, en toda la recta) resulta ser integrable si se multiplica por $e^{-\gamma x}$, donde γ es un número real. Entonces, la integral

$$g(s) = \int_{-\infty}^{\infty} f(x) e^{-isx} dx = \int_{-\infty}^{\infty} f(x) e^{-ix\lambda} e^{x\mu} dx$$

converge para determinados valores complejos de $s = \lambda + i\mu$, en particular, converge en la recta $\mu = -\gamma$. En esta recta representa la transformación de Fourier de la función $f(x) e^{x\mu}$.

El caso más importante para las aplicaciones, en el que se cumplen nuestras suposiciones sobre la integrabilidad de la función $f(x) e^{-\gamma x}$, es el caso en que f verifica las condiciones siguientes:

$$\left.\begin{array}{l} f(x) < Ce^{\gamma_0 x} \text{ para } x \geqslant 0, \\ f(x) = 0 \text{ para } x < 0 \end{array}\right\} \tag{1}$$

(γ_0 y C son constantes). Entonces, la integral

$$g(s) = \int_{-\infty}^{\infty} f(x) e^{-isx} dx = \int_{0}^{\infty} f(x) e^{-isx} dx \tag{2}$$

existe para todos los $s = \lambda + i\mu$ tales que $\mu < \gamma_0$, es decir, en el semiplano limitado por la recta Im $s = -\gamma_0$, y representa la transformación de Fourier de la función

$$f(x) e^{\mu x}.$$

Esta última puede ser obtenida a partir de g por medio de la fórmula de invérsión

$$f(x)e^{\mu x}=\frac{1}{2\pi}\int_{-\infty}^{\infty}g(s)e^{i\lambda x}\,d\lambda,$$

de donde

$$f(x)=\frac{1}{2\pi}\int_{i\mu-\infty}^{i\mu+\infty}g(s)e^{isx}\,ds\quad(s=\lambda+i\mu). \tag{3}$$

Como la función $f(x)e^{\mu x}$ decrece para $\mu<-\gamma_0$ como una función exponencial (en virtud de (1)), su transformación de Fourier q y, consecuentemente, también $g(s)e^{isx}$ son funciones analíticas en el semiplano $\operatorname{Im} s<-\gamma_0$.

Por eso, en la fórmula de inversión (3) la integral puede ser tomada respecto a cualquier recta paralela al eje real y perteneciente a este semiplano.

Efectuemos ahora en las fórmulas (2) y (3) un cambio de variable, tomando $p=is$ y designando $g(s)$ mediante $\Phi(p)$. Tendremos

$$\Phi(p)=\int_{0}^{\infty}f(x)e^{-px}\,dx$$

y

$$f(x)=\frac{1}{2\pi}\int_{-\mu-i\infty}^{-\mu+i\infty}\Phi(p)e^{px}\frac{dp}{i}=\frac{1}{2\pi i}\int_{-\mu-i\infty}^{-\mu+i\infty}\Phi(p)e^{px}\,dp.$$

La función Φ está definida y es analítica en el semiplano $\operatorname{Re} p>\mu_0$; se llama *transformación de Laplace* de la función f (que verifica las condiciones (1)).

De lo expuesto se observa que la transformación de Laplace difiere poco, en general, de la transformación de Fourier. Sin embargo, esta pequeña diferencia lleva a que la clase de funciones, para las cuales está definida la transformación de Laplace, se distinga de un modo sustancial de la clase $L_1(-\infty,\infty)$ de funciones, para las cuales existe la transformación de Fourier.

2°. Aplicación de la transformación de Laplace a la solución de ecuaciones diferenciales (método operacional). La transformación de Laplace puede ser empleada para resolver ecuaciones cuando se buscan soluciones determinadas por ciertas condiciones iniciales. Supongamos que está dada una ecuación diferencial lineal de coeficientes constantes

$$y^{(n)}+a_1y^{(n-1)}+\ldots+a_ny=b(x) \tag{4}$$

y que se busca su solución que satisface las condiciones iniciales

$$y(0)=y_0,\ y'(0)=y_1,\ \ldots,\ y^{(n-1)}(0)=y_{n-1}. \tag{5}$$

Apliquemos la transformación de Laplace a la ecuación (4), es decir, multipliquémosla por e^{-px} e integrémosla entre 0 y ∞. Sea

$$Y(p)=\int_0^\infty y(x)e^{-px}\,dx$$

la transformación de Laplace de la función y. Integrando por partes, encontramos la transformación de Laplace de las derivadas $y', \ldots, y^{(n)}$:

$$\int_0^\infty y'(x)e^{-px}\,dx=y(x)e^{-px}\Big|_0^\infty+p\int_0^\infty y(x)e^{-px}\,dx=-y_0+pY(p);$$

$$\int_0^\infty y''(x)e^{-px}\,dx=y'(x)e^{-px}\Big|_0^\infty+p\int_0^\infty y'(x)e^{-px}\,dx=$$
$$=-y_1+p(-y_0+pY(p))=-y_1-py_0+p^2Y(p);$$

. . . .

$$\int_0^\infty y^{(n)}(x)e^{-px}\,dx=y^{(n-1)}(x)e^{-px}\Big|_0^\infty+p\int_0^\infty y^{(n-1)}(x)e^{-px}\,dx=$$
$$=-y_{n-1}+p(-y_{n-2}-py_{n-3}-\ldots+p^{n-1}Y(p))=$$
$$=y_{n-1}-py_{n-2}-\ldots+p^nY(p).$$

Sea, finalmente,

$$B(p)=\int_0^\infty b(x)e^{-px}\,dx.$$

Obtenemos que la transformación de Laplace convierte la ecuación diferencial (4) (teniendo en cuenta las condiciones iniciales (5)) en la ecuación

$$Q(p)+R(p)Y(p)=B(p),$$

donde B es la transformación de Laplace de la función b, Q es un polinomio respecto a p de grado $n-1$ y R es un polinomio en p de grado $\leqslant n$. De esta ecuación tenemos

$$Y(p)=\frac{B(p)-Q(p)}{R(p)}.$$

La solución y de la ecuación (4) se obtiene de aquí mediante la fórmula de inversión

$$y(x)=\frac{1}{2\pi i}\int_{-\mu-i\infty}^{-\mu+i\infty}\frac{B(p)-Q(p)}{R(p)}e^{px}\,dp.$$

Esta integral suele calcularse mediante los residuos.

Para resolver ecuaciones diferenciales lineales de coeficientes constantes se emplea ampliamente el así llamado método operacional. Consiste en que el miembro izquierdo de una ecuación de este tipo

$$y^{(n)}+a_1y^{(n-1)}+\ldots+a_ny=b(x)$$

se considera como resultado de aplicar a la función incógnita y el operador

$$\left(\frac{d^n}{dx^n}+a_1\frac{d^{n-1}}{dx^{n-1}}+\ldots+a_n\right), \tag{6}$$

y la solución de la ecuación se considera como el resultado de aplicar al miembro derecho de la ecuación el operador inverso del operador (6). Es fácil, mediante cálculos directos, encontrar el resultado que se obtiene al aplicar este operador a algunas funciones elementales: trigonométricas, exponenciales, potenciales y sus combinaciones. Esto permite escribir automáticamente la solución de una ecuación lineal de coeficientes constantes, siempre que su miembro derecho sea una combinación de estas funciones elementales. Está claro que el método operacional constituye, de hecho, la aplicación, en forma indirecta, de la transformación de Laplace; esta última puede servir precisamente para argumentar este método que frecuentemente figura en textos técnicos en forma de una «receta».

§ 7. TRANSFORMACION DE FOURIER — STIELTJES

1°. Definición de la transformación de Fourier — Stieltjes. Volvamos de nuevo a la transformación de Fourier en el espacio $L_1(-\infty, \infty)$:

$$g(\lambda)=\int_{-\infty}^{\infty}e^{-i\lambda x}f(x)\,dx.$$

Esta fórmula puede ser escrita en forma de la integral de Stieltjes

$$g(\lambda)=\int_{-\infty}^{\infty}e^{-i\lambda x}\,dF(x), \tag{1}$$

donde

$$F(x) = \int_{-\infty}^{x} f(t)\, dt \tag{2}$$

es una función absolutamente continua de variación acotada $\left(\text{igual a } \int_{-\infty}^{\infty} |f(x)|\, dx\right)$ en todo eje numérico. Sin embargo, la igualdad (1) tiene sentido no sólo para funciones de tipo (2) sino para cualesquiera funciones de variación acotada en toda la recta. La integral

$$g(\lambda) = \int_{-\infty}^{\infty} e^{-i\lambda x}\, dF(x),$$

donde F es una función arbitraria de variación acotada en la recta se llama *transformación de Fourier—Stieltjes de la función F.* Para la transformación de Fourier—Stieltjes se conservan muchas de las propiedades que hemos demostrado anteriormente para la transformación habitual de Fourier como, por ejemplo, la siguiente: la función g definida por la integral (1) es continua y acotada en toda la recta.

En efecto,

$$|g(\lambda_1) - g(\lambda_2)| = \int_{-N}^{N} [e^{-i\lambda_1 x} - e^{-i\lambda_2 x}]\, dF(x) + \\ + \int_{|x| > N} [e^{-i\lambda_1 x} - e^{-i\lambda_2 x}]\, dF(x).$$

El segundo sumando del miembro derecho puede ser hecho tan pequeño como se quiera para cualesquiera λ_1 y λ_2, si se escoge N suficiente grande, mientras que el primer sumando tiende a cero para N fijo cuando $\lambda_1 - \lambda_2 \to 0$.

Al mismo tiempo, no todas las propiedades de la transformación de Fourier subsisten para la transformación de Fourier—Stieltjes. Por ejemplo, la transformación de Fourier—Stieltjes no tiende, en general, a cero para $|\lambda| \to \infty$. Sea, por ejemplo,

$$F(x) = \begin{cases} 0 \text{ para } x < 0, \\ 1 \text{ para } x \geqslant 0. \end{cases}$$

Entonces

$$g(\lambda) = \int_{-\infty}^{\infty} d^{-i\lambda x}\, dF(x) = 1.$$

Análogamente la transformación de Fourier—Stieltjes de la función igual a 0 para $x < x_0$ y a 1 para $x \geqslant x_0$ es $e^{ix_0\lambda}$, es decir, una función periódica de λ.

Si F es una función de saltos tal que los puntos

$$n = 0, \ \pm 1, \ \pm 2, \ \ldots$$

son sus puntos de discontinuidad y

$$\ldots, a_{-1}, a_0, a_1, \ldots, a_n, \ldots \ (\text{donde } \sum |a_n| < \infty)$$

son los valores de sus saltos en estos puntos, tenemos que

$$\int_{-\infty}^{\infty} e^{-i\lambda x}\, dF(x) = \sum_n a_n e^{-in\lambda}$$

es una función de período 2π. En cambio, si F tiene saltos a_n en los puntos x_n, que forman una sucesión arbitraria en el eje numérico, su transformación de Fourier—Stieltjes es de la forma

$$\sum_n a_n e^{-ix_n\lambda}.$$

Las funciones de este tipo pertenecen a las así llamadas funciones *casi periódicas*.

2°. Aplicación de la transformación de Fourier — Stieltjes a la teoría de probabilidades. Para funciones sumables en $(-\infty, \infty)$ hemos introducido en el § 4 el concepto de convolución

$$f(x) = f_1(x) * f_2(x) = \int_{-\infty}^{\infty} f_1(x-\xi) f_2(\xi)\, d\xi. \tag{3}$$

Pongamos ahora

$$F(x) = \int_{-\infty}^{x} f(t)\, dt, \quad F_1(x) = \int_{-\infty}^{x} f_1(t)\, dt \ \text{ y } \ F_2(x) = \int_{-\infty}^{x} f_2(t)\, dt.$$

Entonces, integrando la igualdad (3), podemos escribirla en la forma siguiente:

$$F(x) = \int_{-\infty}^{x} f(t)\, dt = \int_{-\infty}^{x} dt \int_{-\infty}^{\infty} f_1(t-\xi) f_2(\xi)\, d\xi =$$

$$= \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{x} f_1(t-\xi)\, dt \right\} f_2(\xi)\, d\xi = \int_{-\infty}^{\infty} F_1(x-\xi)\, dF_2(\xi)$$

(el cambio del orden de integración está justificado aquí en virtud del teorema de Fubini y de la integrabilidad absoluta de la función f). La relación que hemos obtenido

$$F(x) = \int_{-\infty}^{\infty} F_1(x-\xi)\, dF_2(\xi)$$

pone en correspondencia a las funciones F_1 y F_2 la función F. Sin embargo, la integral que figura aquí en el miembro derecho tiene sentido no sólo para funciones absolutamente continuas de tipo (2) sino también para dos funciones cualesquiera de variación acotada en toda la recta. Llamaremos a la expresión

$$F(x) = \int_{-\infty}^{\infty} F_1(x-\xi)\, dF_2(\xi), \tag{4}$$

donde F_1 y F_2 son funciones arbitrarias de variación acotada en la recta, *convolución de estas funciones*. Probemos que la expresión (4) representa una función definida para todos los valores de x y de variación acotada en toda la recta.

En efecto, F_1 es una función medible acotada y, consecuentemente, la integral (4) existe para todo x. Además,

$$|F(x_1)-F(x_2)| = \left| \int_{-\infty}^{\infty} (F_1(x_1-\xi)-F_1(x_2-\xi))\, dF_2(\xi) \right| \leqslant$$

$$\leqslant \int_{-\infty}^{\infty} |F_1(x_1-\xi)-F_1(x_2-\xi)|\, d(\operatorname{var} F_2(\xi)),$$

de donde

$$V[F] \leqslant V[F_1]\, V[F_2].$$

TEOREMA 1. *Si F es la convolución de dos funciones F_1 y F_2 de variación acotada y g, g_1 y g_2 son sus transformaciones de Fourier—Stieltjes, se tiene*

$$g(\lambda) = g_1(\lambda)\, g_2(\lambda).$$

DEMOSTRACION. Sea $F = F_1 * F_2$ y sea

$$a = x_0 < x_1 < \ldots < x_n = b$$

una partición del segmento $[a, b]$. Entonces, para todo λ

$$\int_a^b e^{i\lambda x}\, dF(x) = \lim_{\max \Delta x_k \to 0} \sum_{k=1}^{n} e^{i\lambda x_k} (F(x_k) - F(x_{k-1})) =$$

$$= \lim_{\max \Delta x_k \to 0} \int_{-\infty}^{\infty} \sum_{k=1}^{n} e^{-i\lambda(x_k-\xi)} (F_1(x_k-\xi) - F_1(x_{k-1}-\xi))\, e^{-i\lambda\xi}\, dF_2(\xi),$$

es decir,

$$\int_a^b e^{-i\lambda x}\,dF(x) = \int_{-\infty}^{\infty}\left\{\int_{a-\xi}^{b-\xi} e^{-i\lambda x}\,dF_1(x)\right\} e^{-i\lambda\xi}\,dF_2(\xi).$$

Pasando aquí al límite para $a \to -\infty$ y $b \to \infty$, obtenemos

$$\int_{-\infty}^{\infty} e^{-i\lambda x}\,dF(x) = \int_{-\infty}^{\infty} e^{-i\lambda x}\,dF_1(x)\int_{-\infty}^{\infty} e^{-i\lambda\xi}\,dF_2(\xi),$$

esto es,

$$g(\lambda) = g_1(\lambda)\,g_2(\lambda).$$

El teorema de que la transformación de Fourier—Stieltjes transforma la convolución de funciones en producto se emplea ampliamente en la teoría de probabilidades (método de funciones características). Si ξ y η son dos variables aleatorias independientes y F_1 y F_2 son sus funciones de distribución, a la variable $\xi+\eta$ le corresponde la función de distribución

$$F = F_1 * F_2.$$

Frecuentemente resulta necesario considerar en la teoría de probabilidades la suma de variables aleatorias. El paso de las funciones de distribución a sus transformaciones de Fourier—Stieltjes, a las así llamadas funciones características, permite sustituir la operación de convolución por la operación de multiplicación, más simple y más cómoda. De esto hemos hablado ya anteriormente en relación con el concepto de convolución de dos funciones absolutamente integrables. Sin embargo, en aquel momento no disponíamos aún del concepto de la transformación de Fourier—Stieltjes y hemos tenido que limitarnos a variables aleatorias continuas (a la suma de las cuales, siendo ellas independientes, corresponde la convolución de sus densidades de distribución). El concepto de la transformación de Fourier—Stieltjes permite aplicar este mismo método a sumas de variables aleatorias arbitrarias.

EJERCICIOS. 1. Demuéstrese que la transformación de Fourier—Stieltjes verifica la propiedad de unicidad: si la función f es continua a la izquierda y su transformación de Fourier—Stieltjes es idénticamente nula, entonces $F(x) = \text{const}$.

2. Demuéstrese que la operación de convolución de funciones de variación acotada es conmutativa y asociativa.

§ 8. TRANSFORMACION DE FOURIER DE FUNCIONES GENERALIZADAS

Las posibilidades de aplicar el método de la transformación de Fourier, comprendida en el sentido habitual, a diferentes problemas, digamos, a ecuaciones diferenciales, resultan considerablemente restringidas debido a que esta transformación está definida sólo para funciones absolutamente integrales en toda la recta. Se puede obtener una ampliación sustancial del campo de aplicación de la transformación de Fourier introduciendo el concepto de transformación de Fourier para funciones generalizadas. Expongamos las ideas fundamentales de esta construcción.

Consideremos primero en la recta el espacio S_∞ de funciones indefinidamente diferenciables y decrecientes en el infinito junto con sus derivadas más rápido que cualquier potencia de $\frac{1}{|x|}$ (véase el § 4 del cap. IV).

Tomando S_∞ por el espacio de funciones básicas, consideremos el espacio correspondiente de funciones generalizadas S_∞^*.

Definamos ahora la transformación de Fourier en el espacio S_∞^*. Para ello recordemos, ante todo, que la transformación de Fourier (comprendida en el sentido habitual) aplica el espacio S_∞ en sí mismo: si $\varphi \in S_\infty$, también $F[\varphi] \in S_\infty$ y, ademas F es una aplicación biunívoca de S_∞ sobre todo el espacio S_∞. Apoyándonos en esto, daremos la definición siguiente. *Se llama transformación de Fourier de una función generalizada $f \in S_\infty^*$ a la funcional lineal $g \in S_\infty^*$ definida mediante la fórmula*

$$(g, \psi) = (f, \varphi), \text{ donde } \psi = F[\varphi]. \tag{1}$$

Cuando φ recorre todo el S_∞, también $\psi = F[\varphi]$ recorre todo el S_∞; luego, la igualdad (1) define efectivamente una funcional sobre S_∞. Es inmediato comprobar la linealidad y continuidad de esta funcional.

Para aquellos elementos de S_∞^* que representan funciones absolutamente integrables, la definición que acabamos de enunciar de la transformación de Fourier coincide con la habitual. En efecto, si $f \in L_1(-\infty, \infty)$, $\varphi \in S_\infty$ y $g = F[f]$, $\psi = F[\varphi]$, entonces del teorema de Plancherel se desprende la igualdad

$$(f, \varphi) = (g, \psi);$$

además, para f prefijada existe solamente una función g salvo una equivalencia, que satisface esta igualdad para toda $\varphi \in S_\infty$. De manera que la definición de la transformación de Fourier para funciones generalizadas, enunciada más arriba, constituye una extensión de la definición clásica a una clase más amplia de objetos.

Ejemplos. 1. Sea $f(x)=c=\text{const}$. Entonces,

$$(f,\ \varphi)=\int_{-\infty}^{\infty} c\varphi(x)\,dx=c\psi(0)\quad (\psi=F[\varphi]),$$

es decir, la transformación de Fourier de una constante es igual a esta constante multiplicada por la δ-función.

2. Sea $f(x)=e^{iax}$. Entonces,

$$(f,\ \varphi)=\int_{-\infty}^{\infty} e^{iax}\varphi(x)\,dx=\psi(-a),$$

es decir, la transformación de Fourier de la función e^{iax} es δ-función desplazada $\delta(x-a)$.

3. Sea $f(x)=x^2$. Entonces, de la igualdad

$$\psi''(\lambda)=-\int_{-\infty}^{\infty} x^2\varphi(x)\,e^{-i\lambda x}\,dx,$$

obtenemos, tomando en ella $\lambda=0$,

$$(x^2,\ \varphi(x))=-\psi''(0)$$

es decir, la transformación de Fourier de la función x^2 es la segunda derivada de la δ-función tomada con el signo menos.

Hemos definido la transformación de Fourier para las funciones generalizadas en S_∞. Pero, podríamos tomar cualquier otro espacio básico, por ejemplo, el espacio K de funciones terminales indefinidamente diferenciables. Para toda función $\varphi\in K$ la transformación de Fourier (en el sentio habitual) existe y se puede comprobar que es una función analítica entera de orden de crecimiento exponencial. Hablando con más precisión, la transformación de Fourier es un operador lineal que aplica el espacio K en el espacio Z, cuyos elementos son funciones analíticas enteras ψ, para cada una de las cuales se cumplen las desigualdades.

$$|s|^q\,|\psi(s)|\leqslant C_q e^{a|\tau|}\quad (q=1,\ 2,\ \ldots),$$

donde C_q y a son constantes, que dependen de ψ y $\tau=\operatorname{Im} s$. Puesto que en el espacio K se ha introducido anteriormente un concepto de convergencia, la aplicación F que transforma K en Z induce cierto concepto de convergencia en Z; una sucesión $\{\psi_n\}$ converge en Z hacia ψ cuando para las imágenes recíprocas se cumple la relación $\varphi_n\to\varphi$. Además, es fácil enunciar este concepto de convergencia sin recurrir el espacio K[1].

Sea ahora f un elemento arbitrario de K^*. Asignémosle una funcional lineal g sobre Z, tomando:

$$(g, \psi) = (f, \varphi), \text{ donde } \psi = F[\varphi].$$

Esta funcional g se llamará transformación de Fourier de la funcional f. De esta forma la transformación de Fourier de una función generalizada f sobre el espacio básico K es una función generalizada sobre Z, es decir, sobre aquel espacio en el que se aplica K por la transformación de Fourier comprendida en el sentido habitual.

Esta misma construcción puede ser repetida también para funciones generalizadas definidas en otros espacios de funciones básicas. Cada vez surgirá un esquema que incluye cuatro espacios: un espacio inicial de funciones básicas, el conjunto de las transformaciones de Fourier de estas funciones básicas (es decir, el segundo espacio de funciones básicas) y dos espacios duales. Este esquema se reduce a dos espacios cuando por funciones básicas se toma el espacio S_∞ ya que la transformación de Fourier lo aplica en sí mismo.

El concepto de la transformación de Fourier para funciones generalizadas ha encontrado amplia aplicación en la teoría de ecuaciones diferenciales en derivadas parciales. El lector podrá encontrar un tratamiento de estos problemas en el libro de G. E. Shílov [12]

---

[1] A saber: $\psi_n \longrightarrow 0$ en Z cuando se cumplen las desigualdades

$$|s^q \psi_n(s)| \leqslant c_q e^{a|\tau|},$$

y $\psi_n \longrightarrow 0$ uniformemente en todo intervalo finito del eje real.

CAPITULO X

ECUACIONES INTEGRALES LINEALES

§ 1. DEFINICIONES FUNDAMENTALES. ALGUNOS PROBLEMAS QUE LLEVAN A ECUACIONES INTEGRALES

1°. Tipos de ecuaciones integrales. Se llama ecuación integral a una ecuación que contiene la función incógnita bajo el signo de integral. Tal es, por ejemplo, la ecuación

$$\varphi(s)=\int_b^a K(s,\ t)\,\varphi(t)\,dt+f(s), \tag{1}$$

donde f y K son funciones dadas y φ la función que debemos encontrar. Las variables s y t recorren aquí un segmento prefijado $[a, b]$.

La particularidad característica de la ecuación (1) es su linealidad: la función incógnita φ entra en ella de un modo lineal. Varios problemas conducen a la necesidad de considerar también ecuaciones integrales no lineales, por ejemplo, la ecuación de tipo

$$\varphi(s)=\int_a^b K(s,\ t)\,g(\varphi(t),\ t)\,dt,$$

donde K y g son funciones dadas. Sin embargo, nos limitaremos en lo sucesivo a ecuaciones integrales lineales.

Algunas ecuaciones integrales fueron estudiadas ya al principio del siglo pasado. Por ejemplo, Abel consideró en 1823 la ecuación que lleva ahora su nombre

$$f(s)=\int_0^s \frac{\varphi(t)}{(s-t)^\alpha}dt \quad (0<\alpha<1,\ f(0)=0),$$

donde f es una función dada y φ es la función incógnita, y demostró que la solución de esta ecuación es de la forma

$$\varphi(t)=\frac{\operatorname{sen}\pi\alpha}{\pi}\int_0^t \frac{f'(s)}{(t-s)^{1-\alpha}}ds.$$

Sin embargo, la teoría general de ecuaciones integrales lineales fue elaborada sólo en el límite de los siglos XIX y XX en las obras, fundamentalmente, de Volterra, de Fredholm y de Hilbert.

La ecuación (1) se llama *ecuación de Fredholm de segunda especie* (véase el § 7 del capítulo IV) mientras que la ecuación

$$\int_a^b K(s,\ t)\,\varphi(t)\,dt+f(s)=0 \tag{2}$$

(donde la función incógnita figura sólo bajo el signo de integral) es llamada *ecuación de Fredholm de primera especie.*

La ecuación de Abel mencionada anteriormente pertenece a las así llamadas *ecuaciones de Volterra*; la forma general de estas ecuaciones es:

$$\int_a^s K(s,\ t)\,\varphi(t)\,dt=f(s) \tag{3}$$

(ecuación de Volterra de primera especie) o

$$\varphi(s)=\int_a^s K(s,\ t)\,\varphi(t)\,dt+f(s) \tag{4}$$

(ecuación de Volterra de segunda especie). Está claro que la ecuación de Volterra puede ser considerada como una ecuación de Fredholm en la que la función K verifica la condición

$$K(s,\ t)=0 \text{ para } t>s.$$

Sin embargo, conviene destacar las ecuaciones de tipo Volterra en una clase especial ya que ellas poseen una serie de propiedades que no tienen lugar para ecuaciones arbitrarias de Fredholm.

Si en las ecuaciones (1), (2) o (3) la función f es igual a cero, esta ecuación se llama homogénea. En el caso contrario la ecuación se llama no homogénea.

2°. Ejemplos de problemas que llevan a ecuaciones integrales. En los parágrafos posteriores de este capítulo estudiaremos las propiedades fundamentales de las ecuaciones integrales lineales. Sin embargo, indicaremos previamente algunos problemas típicos que conducen a ellas.

1. *Equilibrio de una cuerda cargada.* Consideremos una cuerda, esto es, un hilo material de longitud l, que flexiona libremente, pero ofrece una resistencia a la dilatación, proporcional a la magnitud de ésta. Supongamos fijados los extremos de la cuerda en los puntos $x=0$ y $x=l$. Entonces en la posición de equilibrio la cuerda coincide con el segmento $0 \leqslant x \leqslant l$ del eje x. Supongamos ahora que en el punto $x=\xi$ se ha aplicado una fuerza vertical $P=P_\xi$. Bajo el efecto de esta fuerza la cuerda tomará, evidentemente, la forma de quebrada indicada en la fig. 24.

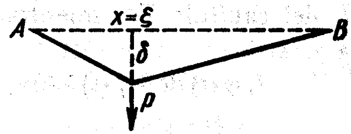


FIG. 24

Busquemos la magnitud δ de la flecha de la cuerda en el punto ξ de su posición de equilibrio bajo la acción de la fuerza P_ξ aplicada en este punto. Si la magnitud de la fuerza P_ξ es pequeña en comparación con la tensión T_0 de la cuerda sin carga, podemos aceptar que la tensión de la cuerda cargada sigue siendo T_0. Entonces, de la condición de equilibrio de la cuerda encontramos la igualdad siguiente:

$$T_0 \frac{\delta}{\xi} + T_0 \frac{\delta}{l-\xi} = P_\xi,$$

de donde

$$\delta = \frac{P_\xi (l-\xi)\xi}{T_0 l}.$$

Sea ahora $u(x)$ la flecha de la cuerda en un punto x bajo la acción de la fuerza P_ξ. Tenemos

$$u(x) = P_\xi G(x, \xi),$$

donde

$$G(x, \xi) = \begin{cases} \dfrac{x(l-\xi)}{T_0 l} & \text{para} \quad 0 \leqslant x \leqslant \xi, \\[2mm] \dfrac{(l-x)\xi}{T_0 l} & \text{para} \quad \xi \leqslant x \leqslant l. \end{cases}$$

En particular, de estas fórmulas se ve inmediatamente que $G(x, \xi)=G(\xi, x)$. Supongamos ahora que sobre la cuerda actúa una fuerza distribuida continuamente a lo largo suyo con la densidad $p(\xi)$. Si esta fuerza es pequeña, la deformación otra vez dependerá linealmente de la fuerza y la forma de la cuerda cargada de este modo será descrita mediante la función

$$u(x)=\int_0^l G(x, \xi)\, p(\xi)\, d\xi. \tag{5}$$

Luego, si está dada la carga que actúa sobre la cuerda, la fórmula (5) permite encontrar la forma que toma la cuerda bajo la acción de esta carga.

Consideremos ahora el problema recíproco: hallar la distribucíón de la carga p bajo la cual la cuerda toma la forma prefijada u. Para encontrar la función p a partir de la función dada u obtenemos una ecuación que coincide, salvo las denotaciones, con la ecuación (2), es decir, una ecuación integral de Fredholm de primera especie.

2. *Oscilaciones libres y forzadas de una cuerda.* Supongamos ahora que la cuerda no se encuentra en reposo y realiza ciertas oscilaciones. Sea $u(x, t)$ la posición en el momento t de aquel punto de la cuerda cuya abscisa es x y sea ρ la densidad lineal de la cuerda[1]. Entonces, sobre un elemento de la cuerda de longitud dx actúa una fuerza de inercia igual a

$$-\frac{\partial^2 u(x, t)}{\partial t^2}\rho dx, \text{ de donde}$$

$$p(\xi)=-\frac{\partial^2 u(\xi, t)}{\partial t^2}\rho.$$

Tomando en (5) esta expresión en lugar de $p(\xi)$, obtenemos

$$u(x, t)=-\int_0^l G(x, \xi)\,\rho\frac{\partial^2 u(\xi, t)}{\partial t^2}\, d\xi. \tag{6}$$

Supongamos que la cuerda realiza oscilaciones armónicas de una frecuencia prefijada ω y de una amplitud $u(x)$ que depende de x. En otras palabras sea

$$u(x, t)=u(x)\operatorname{sen}\omega t.$$

Introduciendo esta expresión en (6) y dividiendo ambos miembros de la igualdad por $\operatorname{sen}\omega t$, obtenemos la siguiente

[1] Aceptamos que ρ=const aunque esto no es sustancial para lo sucesivo.

ecuación integral para u:

$$u(x) = \rho\omega^2 \int_0^l G(x, \xi) u(\xi) d\xi. \qquad (7)$$

Si la cuerda no oscila libremente, sino, bajo la acción de una fuerza exterior, realiza oscilaciones forzadas, es fácil comprobar que la correspondiente ecuación de las oscilaciones armónicas de la cuerda es de la forma

$$u(x) = \rho\omega^2 \int_0^l G(x, \xi) u(\xi) d\xi + f(x),$$

es decir, representa una ecuación no homogénea de Fredholm de segunda especie.

3. *Reducción de ecuaciones diferenciales a ecuaciones integrales.* La resolución de una u otra ecuación diferencial conviene reducirla en varios casos a la resolución de una ecuación integral. Por ejemplo, en el cap. II hemos visto que para demostrar la existencia y la unicidad de la solución de la ecuación diferencial

$$y' = f(x, y)$$

con la condición inicial $y(x_0) = y_0$, conviene reducirla a la ecuación integral (no lineal)

$$y = y_0 + \int_{x_0}^{x} f(\xi, y) d\xi.$$

Las ecuaciones de orden superior al primero también pueden ser reducidas a una ecuación integral. Consideremos, por ejemplo, la ecuación de segundo orden

$$y'' + f(x) y = 0.$$

Tomando $f(x) = \rho^2 - \sigma(x)$, donde $\rho = \text{const}$, podemos escribirla en la forma

$$y'' + \rho^2 y = \sigma(x) y. \qquad (8)$$

Como se sabe, la solución de la ecuación

$$y'' + \rho^2 y = g(x)$$

puede ser representada en la forma

$$y(x) = \cos\rho(x - a) + \frac{1}{\rho} \int_a^x \operatorname{sen}\rho(x - \xi) g(\xi) d\xi.$$

Luego, la resolución de la ecuación (8) se reduce a la resolución de la ecuación integral

$$y(x)-\frac{1}{\rho}\int_a^x \sigma(\xi)\operatorname{sen}\rho(x-\xi)\,y(\xi)\,d\xi=\cos\rho(x-a).$$

§ 2. ECUACIONES INTEGRALES DE FREDHOLM

1°. Operador integral de Fredholm. En este parágrafo estudiaremos las ecuaciones de Fredholm de segunda especie, esto es, las ecuaciones de tipo

$$\varphi(s)=\int_a^b K(s,\ t)\,\varphi(t)\,dt+f(s). \tag{1}$$

Respecto a la función K, llamada *núcleo* de esta ecuación, supondremos que es medible y verifica la condición

$$\int_a^b\int_a^b |K^2(s,\ t)|\,ds\,dt<\infty. \tag{2}$$

El término independiente f de esta ecuación es una función dada de $L_2[a,\ b]$ y φ es la función incógnita perteneciente a $L_2[a,\ b]$.

Pongamos en correspondencia a la ecuación (1) el operador A definido del modo siguiente:

$$A\varphi=\psi$$

significa que

$$\psi(s)=\int_a^b K(s,\ t)\,\varphi(t)\,dt. \tag{3}$$

El estudio de la ecuación (1) se reduce, por supuesto, al estudio de las propiedades de este operador, llamado operador de Fredholm de núcleo K.

TEOREMA 1. *La igualdad* (3), *donde* $K(s,\ t)$ *es una función de cuadrado integrable, define en el espacio* $L_2[a,\ b]$ *un operador lineal totalmente continuo* A *cuya norma satisface la desigualdad*

$$\|A\|\leqslant\sqrt{\int_a^b\int_a^b |K^2(s,\ t)|\,ds\,dt}. \tag{4}$$

DEMOSTRACION. Observemos, ante todo, que la integral

$$\int_a^b |K^2(s,\ t)|\,dt$$

existe, debido al teorema de Fubini y a la condición (2), para casi todo s. En otras palabras, $K(s, t)$ pertenece como función de t a $L_2[a, b]$ para casi todo s. Como el producto de dos funciones de cuadrado sumables es sumable, la integral que figura en el miembro derecho de (3) existe para casi todo s, es decir, la función ψ está definida en casi todos los puntos. Probemos que $\psi \in L_2[a, b]$. En virtud de la desigualdad de Cauchy—Buniakovski, tenemos para casi todo s

$$|\psi^2(s)| = \left| \int_a^b K(s, t)\varphi(t)\,dt \right|^2 \leqslant$$

$$\leqslant \int_a^b |k^2(s, t)|\,dt \cdot \int_a^b |\varphi^2(t)|\,dt = \|\varphi\|^2 \cdot \int_a^b |k^2(s, t)|\,dt.$$

Integrando respecto a s y sustituyendo la integral reiterada de $|K^2(s, t)|$ por una doble, obtenemos la desigualdad

$$\|A\varphi\|^2 = \int_a^b |\psi^2(s)|\,ds \leqslant \|\varphi\|^2 \int_a^b \int_a^b |K^2(s, t)|\,ds\,dt$$

que además de probar la integrabilidad de $|\psi^2(s)|$ demuestra la estimación (4) para la norma del operador A. Resta probar que el operador A es totalmente continuo. Sea $\{\psi_n\}$ un sistema completo ortogonal de $L_2[a, b]$. Entonces, todos los productos pares de tipo $\psi_m(s)\psi_n(t)$ forman un sistema completo en el espacio $L_2([a, b]\times[a, b])$ y, consecuentemente,

$$K(s, t) = \sum_m \sum_n a_{mn}\psi_m(s)\psi_n(t).$$

Pongamos ahora

$$K_N(s, t) = \sum_{m,\, n=1}^{N} a_{mn}\psi_m(s)\psi_n(t)$$

y sea A_N el operador correspondiente al núcleo K_N. Este operador es totalmente continuo ya que transforma todo el espacio $L_2[a, b]$ en un subespacio de dimensión finita (los operadores de este tipo han sido llamados en el cap. IV degenerados). En efecto, si $\varphi \in L_2[a, b]$, se tiene

$$A_N\varphi = \int_a^b K_N(s, t)\varphi(t)\,dt = \sum_{m,\, n=1}^{N} a_{mn}\psi_m(s) \int_a^b \varphi(t)\psi_n(t)\,dt =$$

$$= \sum_{m=1}^{N} \psi_m(s) \sum_{n=1}^{N} a_{mn}b_n,$$

donde

$$b_n = \int_a^b \varphi(t)\,\psi_n(t)\,dt,$$

es decir, todo elemento $\varphi \in L_2[a, b]$ es transformado por el operador A_N en un elemento del subespacio de dimensión finita generado por los vectores $\psi_1, \ldots, \psi_N$. Ahora bien, como K_N es la suma parcial de la serie de Fourier de la función K, tenemos

$$\int_a^b \int_a^b (K(s, t) - K_N(s, t))^2\,ds\,dt \longrightarrow 0 \text{ para } N \to \infty.$$

Aplicando la estimación (4) al operador $A - A_N$, encontramos de aquí

$$\| A - A_N \| \longrightarrow 0 \quad \text{para} \quad N \longrightarrow \infty.$$

Empleando ahora el teorema de que el límite de una sucesión convergente de operadores totalmente continuos es un operador totalmente continuo, obtenemos la continuidad total del operador A.

El teorema queda demostrado

Observaciones. 1. Al demostrar el teorema 1 hemos probado que todo operador de Fredholm puede ser representado como límite (en el sentido de la convergencia según la norma) de una sucesión de operadores integrales degenerados.

2. Sean A_1 y A_2 dos operadores de tipo (3) y sean K_1 y K_2 sus núcleos correspondientes. Si los operadores A_1 y A_2 son iguales, es decir, $A_1\varphi = A_2\varphi$ para toda $\varphi \in L_2[a, b]$, entonces $K_1(s, t) = K_2(s, t)$ casi en todos los puntos. En efecto, si

$$A_1\varphi - A_2\varphi = \int_a^b (K_1(s, t) - K_2(s, t)\,\varphi(t)\,dt = 0$$

para toda $\varphi \in L_2[a, b]$, entonces para casi todo $s \in [a, b]$ se tiene

$$\int_a^b | K_1(s, t) - K_2(s, t) |^2\,dt = 0,$$

es decir,

$$\int_a^b \int_a^b | K_1(s, t) - K_2(s, t)|^2\,ds\,dt = 0,$$

de donde se desprende nuestra afirmación. Luego, si convenimos, como siempre, en no distinguir las funciones sumables equivalentes, podemos decir que la correspondencia entre los operadores integrales y los núcleos es biunívoca.

TEOREMA 2. *Sea A un operador de Fredholm correspondiente a un núcleo $K(s, t)$. Entonces, el operador conjugado A^* se define por el núcleo «conjugado» $\overline{K(t, s)}$.*

DEMOSTRACION. Empleando el teorema de Fubini, tenemos

$$(Af, g)=\int_a^b\left\{\int_a^b K(s, t)f(t)\,dt\right\}\overline{g(s)}\,ds=\int_a^b\int_a^b K(s, t)f(t)\overline{g(s)}\,dt\,ds=$$

$$=\int_a^b\left\{\int_a^b K(s, t)\overline{g(s)}\,ds\right\}f(t)\,dt=\int_a^b f(t)\,\overline{\left\{\int_a^b \overline{K(s, t)}\,g(s)\,ds\right\}}\,dt$$

y de aquí se deduce la afirmación del teorema.

En particular, un operador A de tipo (3) es autoconjugado en $L_2[a, b]$, es decir, $A^*=A$, cuando, y sólo cuando, $\overline{K(s, t)}= =K(t, s)$. En el caso en que se considera el espacio de Hilbert real (y, por lo tanto, núcleos reales K) la condición de autoconjugación es la igualdad $K(s, t)=K(t, s)$.

Observación. Hemos considerado operadores integrales que actúan en el espacio $L_2[a, b]$ en un segmento. No obstante, todo lo expuesto se extiende sin modificaciones al caso en que se considera, en lugar del segmento $[a, b]$, un espacio cualquiera provisto de medida.

2°. Ecuaciones de núcleo simétrico. Consideremos una ecuación integral de Fredholm de segunda especie

$$\varphi(s)=\int_a^b K(s, t)\,\varphi(t)\,dt+f(s), \tag{5}$$

cuyo núcleo verifica las condiciones

1) $$\int_a^b\int_a^b |K^2(s, t)|\,ds\,dt<\infty,$$

2) $$K(s, t)=\overline{K(t, s)}.$$

Estas ecuaciones serán llamadas ecuaciones de *núcleo simétrico.* En virtud de los teoremas 1 y 2 del punto anterior, el operador de Fredholm correspondiente

$$A\varphi=\int_a^b K(s, t)\,\varphi(t)\,dt \tag{6}$$

es totalmente continuo y autoconjugado. Luego, es válido para él el teorema de Hilbert—Schmidt (punto 5, § 6, cap. IV). Apliquemos este teorema a la resolución de la ecuación (5). Como lo que importa no es la forma integral del operador (6) sino el hecho de que este operador es totalmente continuo y autoconjugado, es natural escribir la ecuación (5) en forma simbólica

$$\varphi = A\varphi + f. \tag{7}$$

De acuerdo con el teorema de Hilbert—Schmidt, existe para A un sistema ortonormal $\{\psi_n\}$ de funciones propias, correspondientes a los valores propios $\{\lambda_n\}$, tal que todo elemento ξ de L_2 puede ser representado en la forma

$$\xi = \sum_n a_n \psi_n + \xi', \text{ donde } A\xi' = 0.$$

Tomemos

$$f = \sum_n b_n \psi_n + f' \quad (Af' = 0) \tag{8}$$

y busquemos la solución φ de la ecuación (7) en la forma

$$\varphi = \sum_n x_n \psi_n + \varphi' \quad (A\varphi' = 0). \tag{9}$$

Introduciendo los desarrollos (8) y (9) en (7), obtenemos

$$\sum_n x_n \psi_n + \varphi' = \sum_n x_n \lambda_n \psi_n + \sum b_n \psi_n + f'.$$

Esta igualdad se cumple cuando, y sólo cuando,

$$f' = \varphi'$$

y

$$x_n (1 - \lambda_n) = b_n \quad (n = 1, 2, \ldots\ldots),$$

es decir, cuando

$$f' = \varphi',$$

$$x_n = \frac{b_n}{1 - \lambda_n} \text{ para } \lambda_n \neq 1,$$

$$b_n = 0 \text{ para } \lambda_n = 1.$$

La última igualdad es una condición necesaria y suficiente para que la ecuación (7) tenga solución. Obtenemos de esta forma el resultado siguiente: *si* 1 *no es valor propio del operador* A, *la ecuación* (7) *tiene una solución, y sólo una, cualquiera que sea* f; *en cambio, si* 1 *es valor propio del operador* A *la ecuación* (7) *tiene solución cuando, y sólo cuando, el término independiente* f *es ortogonal a todas las funciones propias del operador* A *correspondientes al valor propio* 1; *si esta última condición se cumple, la ecuación* (7) *tiene un conjunto infinito de soluciones.*

3°. Teoremas de Fredholm. Caso de núcleo degenerado. Pasemos ahora a estudiar las ecuaciones de Fredholm de segunda especie con núcleos que verifican la condición

$$\int_a^b \int_a^b |K^2(s, t)|\, ds\, dt < \infty.$$

(que garantiza la continuidad total del operador correspondiente), pero que no son simétricos.

Supongamos primero que se considera la ecuación

$$\varphi(s) = \int_a^b K(s, t)\, \varphi(t)\, dt + f(s), \tag{10}$$

cuyo núcleo es degenerado, es decir, de la forma

$$K(s, t) = \sum_{i=1}^{n} P_i(s)\, Q_i(t), \tag{11}$$

donde P_i, Q_i son funciones de L_2. El operador con núcleo de tipo (11) transforma toda función $\varphi \in L_2$ en la suma

$$\sum_{i=1}^{n} P_i(s) \int_a^b Q_i(t)\, \varphi(t)\, dt,$$

es decir, en un elemento del subespacio de dimensión finita generado por las funciones P_i, $i = 1, 2, \ldots, n$. Notemos que en la expresión (11) las funciones $P_1, \ldots, P_n$ pueden ser consideradas linealmente independientes. En efecto, si esto no fuese así, podríamos, expresando cada una de las funciones P_i como combinación lineal de las independientes, representar este mismo núcleo $K(s, t)$ como suma de un número menor de sumandos de tipo $\tilde{P}_j(s)\, \tilde{Q}_j(t)$ de manera que las funciones $\tilde{P}_j$ sean linealmente independientes.

Busquemos, pues, la solución de la ecuación (10) con núcleo degenerado (11) en el que las funciones $P_1, \ldots, P_n$ son linealmente independientes. Tomando en la ecuación (10) en lugar de $K(s, t)$ la suma correspondiente, obtenemos

$$\varphi(s) = \sum_{i=1}^{n} P_i(s) \int_a^b Q_i(t)\, \varphi(t)\, dt + f(s). \tag{12}$$

Designando

$$\int_a^b Q_i(t)\, \varphi(t)\, dt = q_i,$$

podemos escribir la ecuación (12) en la forma

$$\varphi(s)=\sum_{i=1}^{n} q_i P_i(s)+f(s).$$

Tomando en la ecuación (10) esta expresión para φ, obtenemos

$$\sum_{i=1}^{n} q_i P_i(s)+f(s)=$$
$$=\sum_{i=1}^{n} P_i(s)\int_a^b Q_i(t)\left[\sum_{j=1}^{n} q_j P_j(t)+f(t)\right]dt+f(s). \quad (13)$$

Poniendo

$$\int_a^b Q_i(t)\,P_j(t)\,dt=a_{ij}, \qquad \int_a^b Q_i(t)\,f(t)\,dt=b_i,$$

la igualdad (13) resulta:

$$\sum_{i=1}^{n} q_i P_i(s)=\sum_{i=1}^{n} P_i(s)\left[\sum_{j=1}^{n} a_{ij}q_j+b_i\right].$$

Como las funciones P_i son, por suposición, linealmente independientes, esta igualdad implica la igualdad de los respectivos coeficientes de $P_i(s)$:

$$q_i=\sum_{j=1}^{n} a_{ij}q_j+b_i, \quad i=1,\ 2,\ \ldots,\ n. \quad (14)$$

Hemos obtenido para los coeficientes q_i un sistema de ecuaciones lineales. Resolviéndolo obtenemos la función

$$\varphi(s)=\sum_{i=1}^{n} q_i P_i(s)+f(s).$$

Esta función satisface la ecuación integral (10) ya que todos los razonamientos, mediante los cuales hemos pasado de la ecuación (10) al sistema (14), pueden ser realizados en orden contrario.

Luego, *la resolución de una ecuación integral de núcleo degenerado se reduce a la resolución del correspondiente sistema* (14) *de ecuaciones algebraicas lineales.*

Para sistemas de ecuaciones lineales son bien conocidas las condiciones de existencia y de unicidad de la solución.

1. Un sistema de ecuaciones algebraicas lineales

$$Tx=y\ (T=\|a_{ik}\|,\ x=(x_1,\ \ldots,\ x_n),\ y=(y_1,\ \ldots,\ y_n))$$

tiene solución cuando, y sólo cuando, el vector y es ortogonal a toda solución del sistema homogéneo conjugado

$$T^*z=0 \quad (T^*=\|a_{kl}\|).$$

II. Si el determinante de la matriz T es diferente de cero, la ecuación $Tx=y$ tiene solución única cualquiera que sea y. En cambio, si el determinante de la matriz T es igual a cero, la ecuación homogénea $Tx=0$ tiene soluciones no nulas.

III. Como la matriz T y la matriz conjugada T^* son del mismo rango, los sistemas homogéneos $Tx=0$ y $T^*z=0$ tienen el mismo número de soluciones linealmente independientes.

Debido a la relación que, como hemos visto, existe entre ecuaciones integrales de núcleo degenerado y sistemas de ecuaciones algebraicas lineales, estas proposiciones pueden ser consideradas como teoremas referentes a las soluciones de ecuaciones integrales degeneradas. En el punto siguiente demostraremos que, de hecho, estos mismos teoremas tienen lugar también para ecuaciones de núcleo arbitrario (no degenerado). Sin embargo, puesto que para operadores integrales no degenerados no tienen sentido conceptos como rango y determinante de una matriz, los teoremas correspondientes deben ser enunciados de manera que en ellos no figuran estos conceptos.

4°. Teoremas de Fredholm para ecuaciones de núcleo no degenerado. Volvamos a considerar la ecuación

$$\varphi(s)=\int_a^b K(s,\ t)\,\varphi(t)\,dt+f(s), \tag{15}$$

suponiendo ahora que su núcleo verifica sólo la condición

$$\int_a^b\int_a^b |K^2(s,\ t)|\,ds\,dt<\infty$$

(que garantiza la continuidad total del operador correspondiente), es decir, no suponemos ahora el núcleo ni degenerado ni simétrico. Nos interesan las condiciones en las que la ecuación (15) tiene solución y las propiedades de sus soluciones. Además, para nosotros será esencial sólo la propiedad de continuidad total del operador correspondiente a la ecuación (15) y no su forma integral. Por lo tanto, realizaremos todas las consideraciones sucesivas para la ecuación en operadores

$$\varphi=A\varphi+f, \tag{16}$$

donde A es un operador arbitrario totalmente continuo definido en el espacio de Hilbert H.

Tomando $T = I - A$ (donde I es el operador unidad), escribiremos la ecuación (16) en la forma

$$T\varphi = f. \tag{17}$$

Además de esta ecuación, consideraremos la ecuación homogénea

$$T\varphi_0 = 0 \tag{18}$$

y las ecuaciones conjugadas

$$T^*\psi = g, \tag{19}$$
$$T^*\psi_0 = 0 \tag{20}$$

($T^* = I - A^*$). La relación existente entre las soluciones de estas cuatro ecuaciones viene expresada en los siguientes teoremas de Fredholm.

I. *La ecuación no homogénea* $T\varphi = f$ *tiene solución para aquellas* f, *y sólo aquéllas, que son ortogonales a toda solución de la ecuación homogénea conjugada* $T^*\psi_0 = 0$.

II. *(Alternativa de Fredholm.) O bien la ecuación* $T\varphi = f$ *tiene una solución, y sólo una, cualquiera que sea* $f \in H$ *o bien la ecuación homogénea* $T\varphi_0 = 0$ *tiene solución no nula.*

III. *Las ecuaciones homogéneas* (17) *y* (19) *tienen el mismo número, además finito, de soluciones linealmente independientes.*

Antes de pasar a demostrar estos teoremas, observemos que son válidos (en virtud de lo dicho en el punto 2) para ecuaciones de núcleo simétrico. Además, como A y A^* coinciden en este caso, el teorema III resulta trivial.

Por otro lado, si A es un operador integral degenerado, las ecuaciones correspondientes se reducen, como hemos visto, a sistemas de ecuaciones algebraicas lineales; los teoremas de Fredholm se convierten, evidentemente, en este caso en los teoremas sobre sistemas lineales que hemos enunciado en el punto anterior.

Aprovechando que todo operador integral es límite de una sucesión convergente de operadores degenerados, podríamos demostrar los teoremas de Fredholm mediante el correspondiente paso al límite (de núcleos degenerados a núcleos no degenerados). Sin embargo, escogemos otro camino y daremos una demostración de estos teoremas que no está relacionada con la consideración de ecuaciones degeneradas.

DEMOSTRACION DE LOS TEOREMAS DE FREDHOLM. Sea $N(B)$ el conjunto de los ceros de un operador lineal continuo B, es decir, el conjunto de todos aquellos $x \in H$ para los que $Bx = 0$. Está claro que $N(B)$ es siempre un subespacio lineal cerrado. Sea $R(B)$ el campo de valores del operador B, es decir, el conjunto de vectores de tipo $y = Bx$. El conjunto $R(B)$ constituye también una variedad lineal, pero, en general, no cerrada. Demostraremos ahora que para el operador $T=I-A$ esta variedad es cerrada.

LEMA 1. *La variedad $R(T)$ es cerrada.*

DEMOSTRACION. Sean $y_n \in R(T)$ y sea $y_n \to y$. Por hipótesis, existen vectores $x_n \in H$ tales que

$$y_n = Tx_n = x_n - Ax_n. \tag{21}$$

Restando, si hace falta, de x_n su proyección sobre $N(T)$, podemos aceptar que estos vectores son ortogonales a $N(T)$. Además, podemos aceptar que $\|x_n\|$ están acotadas en conjunto. En efecto, de lo contrario, pasando a una sucesión, tendríamos $\|x_n\| \to \infty$ y, dividiendo por $\|x_n\|$, deduciríamos de (21) que $\frac{x_n}{\|x_n\|} - A\frac{x_n}{\|x_n\|} \to 0$. Pero, como el operador A es totalmente continuo, podemos aceptar, pasando de nuevo a una subsucesión, que la sucesión $\left\{A\frac{x_n}{\|x_n\|}\right\}$ converge. Por lo tanto, también $\frac{x_n}{\|x_n\|}$ convergerá, digamos a un vector $z \in H$. Está claro que $\|z\| = 1$ y $T(z) = 0$, es decir, $z \in N(T)$. Suponemos, sin embargo, que los vectores x_n son ortogonales a $N(T)$; luego, también el vector z debe ser ortogonal a $N(T)$. La contradicción obtenida permite suponer que $\|x_n\|$ están acotadas en conjunto. Al mismo tiempo, la sucesión $\{Ax_n\}$ puede ser supuesta en este caso convergente; entonces, como se desprende de (21), también será convergente la sucesión $\{x_n\}$. Si x es el límite de esta sucesión, de (21) se desprende que $y = Tx$. El lema queda demostrado.

LEMA 2. *El espacio H es la suma directa ortogonal de los subespacios cerrados $N(T)$ y $R(T)$, es decir,*

$$N(T) \oplus R(T^*) = H \tag{22}$$

y análogamente

$$N(T^*) \oplus R(T) = H. \tag{23}$$

DEMOSTRACION. Sabemos ya que los dos subespacios que figuran en el miembro izquierdo de (22) son cerrados. Además, son ortogonales ya que, si $h \in N(T)$, se tiene $(h, T^*x) = (Th, x) = 0$ para todo $x \in H$. Falta por demostrar que no existe ningún vector no nulo ortogonal simultáneamente a $R(T^*)$ y $N(T)$. Pero, si el vector z es ortogonal a $R(T^*)$, entonces para cualquier $x \in H$ tenemos $(Tz, x) = (z, T^*x) = 0$, es decir, $z \in N(T)$. El lema queda demostrado.

Del lema 2 se desprende inmediatamente el primer teorema de Fredholm. En efecto, $f \perp N(T^*)$ cuando, y sólo cuando, $f \in R(T)$, es decir, cuando existe un φ tal que $T\varphi = f$.

Pongamos ahora $H^k = R(T^k)$ para cada k entero, de manera que, en particular, $H^1 = R(T)$. Está claro que los subespacios H^k forman una cadena de subespacios encajados

$$H \supset H^1 \supset H^2 \supset \ldots \tag{24}$$

y que, en virtud del lema 1, todos estos subespacios son cerrados. Además, $T(H^k) = H^{k+1}$.

LEMA 3. *Existe un j tal que $H^{k+1} = H^k$ para todo $k \geqslant j$.*

DEMOSTRACION. Si no existe un tal j, es evidente que todos los H^k son distintos. En este caso podemos construir una sucesión ortonormal $\{x_k\}$ tal que $x_k \in H^k$ y son ortogonales a H^{k+1}. Sea $l > k$. Entonces,

$$Ax_l - Ax_k = -x_k + (x_l + Tx_k - Tx_l)$$

y, consecuentemente, $\|Ax_l - Ax_k\| \geqslant 1$ ya que $x_l + Tx_k - Tx_l \in H^{k+1}$. Luego, de la sucesión $\{Ax_k\}$ no se puede extraer ninguna subsucesión convergente, lo que contradice a la continuidad total del operador A. Con esto queda demostrado el lema.

LEMA 4. *Si $N(T) = \{0\}$, se tiene $R(T) = H$.*

DEMOSTRACION. Si $N(T) = \{0\}$, el operador T es biunívoco; de manera que, siendo $R(T) \neq H$, la cadena (24) consta de diferentes subespacios y esto contradice al lema 3. Luego, $R(T) = H$. Análogamente, $R(T^*) = H$, si se tiene $N(T^*) = \{0\}$.

LEMA 5. *Si $R(T) = H$, se tiene $N(T) = \{0\}$.*

DEMOSTRACION. Como $R(T) = H$, tenemos, en virtud del lema 2, $N(T^*) = \{0\}$; pero, entonces, en virtud del lema 4, $R(T) = H$ y, consecuentemente, $N(T) = \{0\}$ por el lema 2.

Los lemas 4 y 5 constituyen en su conjunto el contenido del segundo teorema (la alternativa) de Fredholm. Con esto queda demostrado este teorema.

Demostremos, finalmente, el tercer teorema de Fredholm.

Supongamos que el subespacio $N(T)$ es de dimensión infinita. Entonces, existe en este subespacio un sistema ortonormal infinito $\{x_k\}$. Además, $Ax_k = x_k$ de manera que para $k \neq l$ tenemos $\|Ax_k - Ax_l\| = \sqrt{2}$. Pero, esto significa que de la sucesión $\{Ax_k\}$ no se puede extraer ninguna subsucesión convergente, lo que contradice a la continuidad total del operador A.

Sea μ la dimensión de $N(T)$ y sea ν la dimensión de $N(T^*)$. Supongamos que $\mu < \nu$. Sea $\{\varphi_1, \ldots, \varphi_\mu\}$ una base ortonormal en $N(T)$ y sea $\psi_1, \ldots, \psi_\nu$ una base ortonormal de $N(T^*)$. Tomemos

$$Sx = Tx + \sum_{j=1}^{\mu} (x, \varphi_j) \psi_j.$$

Como el operador S se obtiene del operador T agregándole un operador degenerado, todos los resultados demostrados más arriba para el operador T son válidos también para el operador S.

Probemos que la ecuación $Sx = 0$ tiene solamente solución trivial. En efecto, supongamos que

$$Tx + \sum_{j=1}^{\mu} (x, \varphi_j) \psi_j = 0. \tag{25}$$

Como los vectores ψ_j son ortogonales, en virtud del lema 2, a todos los vectores de tipo Tx, de (25) se deduce que

$$Tx = 0$$

y

$$(x, \varphi_j) = 0 \text{ para } 1 \leqslant j \leqslant \mu.$$

Luego, el vector x debe ser, por un lado, una combinación lineal de los vectores φ_j y, por otro lado, debe ser ortogonal a ellos. Consecuentemente, $x = 0$. De modo que la ecuación $Sx = 0$ tiene solamente solución trivial. Existe entonces, de acuerdo con el teorema 2, un vector y tal que

$$Ty + \sum_{j=1}^{\mu} (y, \varphi_j) \psi_j = \psi_{\mu+1}.$$

Está claro que multiplicando esta igualdad escalarmente por $\psi_{\mu+1}$, obtendremos 0 en el miembro izquierdo y 1 en el miembro derecho. Esta contradicción ha surgido porque hemos supuesto que $\mu < \nu$. Luego, $\mu \geqslant \nu$. Sustituyendo ahora el operador T por T^*, encontraremos que $\mu \geqslant \nu$ y, consecuentemente, $\mu = \nu$. El teorema III queda demostrado completamente.

Observación 1. Los teoremas de Fredholm tratan, de hecho, sobre la posibilidad de invertir el operador $A - I$ y significan que $\lambda = 1$ es o bien un punto regular para A o bien un valor propio de multiplicidad finita. Por supuesto, todo lo que se afirma en estos teoremas sigue siendo válido también para los operadores $A - \lambda I$, donde $\lambda \neq 0$. Luego, *todo punto distinto de* 0 *del espectro de un operador totalmente continuo es un valor propio suyo de multiplicidad finita.* Además, como sabemos, el conjunto de estos valores propios es, a lo sumo, numerable. Recordemos, de paso, que el punto 0 siempre pertenece al espectro de un operador totalmente continuo en un espacio de dimensión infinita; pero, en general, no es necesariamente valor propio. Los operadores totalmente continuos, para los cuales 0 es el *único* punto del espectro, son llamados *operadores* (abstractos) *de Volterra*.

Observación 2. Hemos demostrado los teoremas de Fredholm para la ecuación de tipo $\varphi = A\varphi + f$, donde A es un operador totalmente continuo en el espacio de Hilbert. Estos teoremas pueden ser extendidos, sin modificaciones sustanciales, al caso de un espacio de Banach arbitrario E. En este caso, claro está, la ecuación conjugada $\psi = A^*\psi + g$ será una ecuación en el espacio E^*, la condición de ortogonalidad $(f, \psi_0) = 0$ debe comprenderse en el sentido de que toda funcional del subespacio $N^* \subset E^*$ de soluciones de la ecuación $A^*\psi_0 = 0$ se anula en el elemento $f \in E$, etc. Una exposición de los teoremas de Fredholm para el caso de ecuaciones en espacios de Banach se puede ver, por ejemplo, en el libro de L. A. Lusternik y V. I. Sóbolev «Elementos del Análisis Funcional».

5°. **Ecuaciones de Volterra.** Se llama *ecuación de Volterra* (de segunda especie) a la ecuación integral

$$\varphi(s) = \int_a^s K(s, t)\, \varphi(t)\, dt + f(s), \tag{26}$$

donde $K(s, t)$ es una función medible acotada: $|K(s, t)| \leqslant M$. Puesto que esta ecuación puede ser considerada como un caso particular de la ecuación de Fredholm (con núcleo igual a cero para $t > s$), los teoremas de Fredholm son válidos también para la ecuación (26). No obstante, para las ecuaciones de Volterra estos teoremas pueden ser precisados del modo siguiente. *La ecuación de Volterra* (26) *tiene una solución, y sólo una, cualquiera que sea la función* $f \in L_2$.

En efecto, repitiendo textualmente los razonamientos del punto 4 del § 4 del cap. II, veremos que cierta potencia del operador

$$A\varphi = \int_a^s K(s, t)\, \varphi(t)\, dt$$

es un operador contraído y, por lo tanto, la ecuación homogénea tiene solución única (trivial). De aquí se desprende en virtud de los teoremas de Fredholm, nuestra afirmación.

EJERCICIO. Consideremos en un segmento una ecuación integral de Fredholm de segunda especie con núcleo continuo. Demuéstrense para esta ecuación los teoremas de Fredholm *en el espacio de funciones continuas*. En este caso, el papel de «ecuación conjugada» lo desempeña la ecuación integral de núcleo transpuesto y la ortogonalidad se comprende en el sentido de L_2.

6°. Ecuaciones integrales de primera especie. Se llama *ecuación abstracta de Fredholm de primera especie* a la ecuación de tipo

$$A\varphi = f, \tag{27}$$

es decir, a una ecuación que contiene la función incógnita sólo bajo el signo de operador totalmente continuo.

La resolución de una ecuación de este tipo constituye un problema más complejo, en general, que la resolución de una ecuación de segunda especie y para una función arbitraria $f \in L_2$ la ecuación (27) puede no tener solución.

Consideremos primero, a título de ejemplo elemental, la ecuación

$$f(s) = \int_a^s \varphi(t)\, dt,$$

es decir, una ecuación de núcleo

$$K(s, t) = \begin{cases} 1 & \text{para } t \leqslant s, \\ 0 & \text{para } t > s. \end{cases}$$

Ella tiene solución obvia $\varphi(s) = f'(s)$ cuando f es absolutamente continua y pertenece a L_2; no tiene solución en el caso contrario.

Probemos que también en el caso general la ecuación (27) puede no tener solución para una $f \in H$ arbitraria. En efecto, si la ecuación $A\varphi = f$ tiene solución para cualquier $f \in H$, ello significa que este operador transforma H en todo el H. Probemos que esto es imposible. Todo el espacio H puede ser representado como la unión de una cantidad numerable de bolas S_n (por ejemplo, de las bolas de radio 1, 2, ..., n... y centro en el cero). El operador totalmente continuo A transforma cada una de estas bolas en un conjunto compacto. De manera que AH es la unión de una cantidad numerable de compactos. Pero, cualquier compacto es nunca denso en H; al mismo tiempo, H, al igual que cualquier espacio métrico completo, no puede ser representado como la unión de una cantidad numerable de conjuntos nunca densos. Luego, $AH \neq H$; en otras palabras, cualquiera que sea el operador A totalmente continuo en H la ecuación $A\varphi = f$ no puede tener solución para toda $f \in H$.

Otro momento sustancial en la solución de ecuaciones de primera especie consiste en que en H un operador, inverso de un operador totalmente continuo, no es acotado. Por lo tanto, si f_1 y f_2 son dos elementos próximos de H y ambas ecuaciones

$$A\varphi_1 = f_1 \quad \text{y} \quad A\varphi_2 = f_2$$

tienen solución, las soluciones correspondientes $\varphi_1 = A^{-1}f_1$ y $\varphi_2 = A^{-1}f_2$ pueden distinguirse considerablemente una de otra. En otras palabras, un error tan pequeño como se quiera en el término independiente de la ecuación puede conducir a un error tan grande como se quiera en la solución. Los problemas, en los que una pequeña variación en los datos iniciales lleva a una pequeña variación en la solución (la palabra «pequeña» puede ser comprendida en diferentes problemas de modo distinto), se llaman *correctos*. La solución de una ecuación integral de primera especie (a diferencia de una ecuación de segunda especie) es un problema *no correcto*. En los últimos tiempos se han difundido mucho los problemas no correctos y han obtenido un gran desarrollo los métodos de su regularización (es decir, métodos que permiten reducirlos a problemas correctos en uno u otro sentido). Sin embargo, la exposición de estas cuestiones sale de los márgenes de este libro.

§ 3. ECUACIONES INTEGRALES CON PARAMETRO. MÉTODO DE FREDHOLM

1°. Espectro de un operador totalmente continuo en H. Consideremos la ecuación

$$\varphi = \lambda A\varphi + f$$

o, que es lo mismo,

$$(I - \lambda A)\varphi = f, \tag{1}$$

donde A es un operador totalmente continuo en el espacio H de Hilbert y λ es un parámetro numérico.

En virtud de la alternativa de Fredholm, pueden darse dos y sólo dos casos:

1) La ecuación (1) tiene para λ prefijado una solución, y sólo una, cualquiera que sea $f \in H$.

2) La ecuación homogénea $\varphi = \lambda A\varphi$ tiene solución no nula.

En el primer caso el operador $I - \lambda A$ aplica, además biunívocamente, H en todo el H. De aquí se desprende la existencia del operador inverso acotado $(I - \lambda A)^{-1}$. Está claro que esto equivale a que el operador $\left(A - \frac{1}{\lambda} I\right)^{-1}$ existe y es acotado; en otras palabras, $\frac{1}{\lambda}$ no pertenece, en este caso, al espectro del operador A.

Supongamos ahora que tiene lugar la segunda posibilidad, esto es, que existe un elemento $\varphi_\lambda \in H$ diferente de cero tal que

$$\varphi_\lambda = \lambda A\varphi_\lambda \text{ ó } A\varphi_\lambda = \frac{1}{\lambda}\varphi_\lambda,$$

es decir, $\frac{1}{\lambda}$ es valor propio del operador A.

Obtenemos el resultado siguiente: *todo número* $\mu = \frac{1}{\lambda}$ *distinto de cero o bien es un valor propio del operador totalmente continuo A o bien es un valor regular*. En otras palabras, el espectro continuo de un operador totalmente continuo o bien no existe o o bien consta solamente del punto $\mu = 0$.

Uniendo lo que acabamos de decir con el teorema 4 del § 6 del cap. IV, obtenemos que el espectro de un operador totalmente continuo en H puede ser descrito del modo siguiente. El espectro de cualquier operador A totalmente continuo en H consta de un número finito o numerable de valores propios distintos de cero $\mu_1, \mu_2, \ldots, \mu_n \ldots$, cada uno de los cuales es de multiplicidad finita[1], y del punto cero; éste es el único punto posible de acumulación de la sucesión $\{\mu_n\}$. El propio punto $\mu = 0$ puede ser o bien un valor propio de multiplicidad finita o infinita o bien un punto de acumulación del conjunto de valores propios. Como hemos demostrado en el punto 5, para la ecuación

$$\varphi = \lambda B\varphi + f,$$

donde B es un operador de Volterra, siempre tiene lugar el primer caso de la alternativa de Fredholm (existe la solución para cualquier $f \in L_2$). En otras palabras, el espectro de un operador de tipo de Volterra consta sólo del punto $\mu = 0$.

2° **Representación de la solución en forma de una serie de potencias de λ. Determinantes de Fredholm.** La solución de la ecuación

$$(I - \lambda A)\varphi = f$$

puede ser escrita formalmente como

$$\varphi = (I - \lambda A)^{-1} f. \tag{2}$$

Esta fórmula define, efectivamente, la solución cuando $\|\lambda A\| < 1$, es decir, cuando $|\lambda| < \frac{1}{\|A\|}$, ya que en este caso el operador

[1] $\mu = 0$ pertenece necesariamente al espectro del operador A ya que A^{-1} no puede ser acotada en H.

$(I-\lambda A)^{-1}$ existe, está definido en todo H y es acotado (véase el punto 7 del § 5 del cap. IV). Además, el operador $(I-\lambda A)^{-1}$ puede ser considerado, en este caso, como la suma de la serie de potencias

$$(I-\lambda A)^{-1}=I+\lambda A+\lambda^2A^2+\ldots+\lambda^nA^n+\ldots,$$

cuya convergencia (respecto a la norma) está garantizada por la condición $|\lambda|<\frac{1}{\|A\|}$. Luego, la solución (2) de nuestra ecuación (1) puede ser representada en la forma

$$\varphi=f+\lambda Af+\lambda^2A^2f+\ldots+\lambda^nA^nf+\ldots \tag{3}$$

Este mismo resultado se obtiene, si la solución de la ecuación (1) se busca en forma de la serie de potencias

$$\varphi_\lambda=\varphi_0+\lambda\varphi_1+\ldots+\lambda^n\varphi_n+\ldots$$

(donde φ_n ya no dependen de λ). Tomando esta serie en lugar de φ en los miembros derecho e izquierdo de la ecuación $\varphi=\lambda A\varphi+f$ e igualando después los coeficientes de potencias iguales de λ en ambos miembros de la igualdad, obtendremos

$$\varphi_0=f,\ \varphi_1=Af,\ \ldots,\ \varphi_n=A\varphi_{n-1}=A^nf,\ \ldots,$$

es decir, la serie (3).

Probemos que si A es un operador integral definido por un núcleo K de cuadrado integrable, el operador $(I-\lambda A)^{-1}$ puede ser representado, para valores suficientemente pequeños de λ, como la suma $I+\Gamma_\lambda$ del operador unidad I y de un operador integral Γ_λ con núcleo de cuadrado integrable que depende de λ. Veamos primero la forma que tienen en este caso los operadores A^2, A^3, etc. Consideremos, con este fin, un problema más general: sean dados dos operadores integrales

$$A\varphi=\int_a^b K(s,\ t)\,\varphi(t)\,dt,\quad B\varphi=\int_a^b Q(s,\ t)\,\varphi(t)\,dt,$$

donde

$$\int_a^b\int_a^b|K^2(s,\ t)|\,ds\,dt=k^2<\infty,\quad \int_a^b\int_a^b|Q^2(s,\ t)|\,ds\,dt=q^2<\infty.$$

Busquemos la forma del operador AB. Tenemos

$$AB\varphi=\int_a^b\left\{K(s,u)\int_a^b Q(u,t)\,\varphi(t)\,dt\right\}du=\int_a^b\left\{\int_a^b K(s,u)\,Q(u,t)\,du\right\}\varphi(t)\,dt.$$

La posibilidad de cambiar aquí el orden de integración se desprende del teorema de Fubini ya que el integrando

$$K(s,\ u)\,Q(u,\ t)\,\varphi(t)$$

es sumable respecto al conjunto de variables u y t por ser producto de dos funciones

$$K(s,\ u)\,\varphi(t) \text{ y } Q(u,\ t)$$

de cuadrado integrable cada una.

Tomemos

$$R(s,\ t)=\int_a^b K(s,\ u)\,Q(u,\ t)\,du; \tag{4}$$

en virtud de la desigualdad de Cauchy—Buniakovski, tenemos

$$|R^2(s,t)|\leqslant\int_a^b |K^2(s,\ u)|\,du\int_a^b |Q^2(u,\ t)|\,du,$$

de donde

$$\int_a^b\int_a^b R^2(s,\ t)\ ds\,dt\leqslant k^2q^2.$$

Luego, el producto de los operadores integrales de tipo de Fredholm es un operador del mismo tipo cuyo núcleo viene dado por la fórmula (4). En particular, tomando $A=B$, encontramos que A^2 es un operador integral de núcleo

$$K_2(s,\ t)=\int_a^b K(s,\ u)\,K(u,\ t)\,du$$

que verifica la condición

$$\int_a^b\int_a^b |K_2^2(s,\ t)|\,ds\,dt\leqslant\left[\int_a^b\int_a^b |K^2(s,\ t)|\,ds\,dt\right]^2=k^4,$$

de donde $\|A^2\|\leqslant k^2$.

Análogamente obtenemos que cada uno de los operadores A^n está definido por el núcleo

$$K_n(s,\ t)=\int_a^b K_{n-1}(s,\ u)\,K(u,\ t)\,du\quad (n=2,\ 3,\ \ldots)$$

que satisface la condición

$$\int_a^b \int_a^b |K_n^2(s, t)|\, ds\, dt \leqslant k^{2n}, \tag{5}$$

donde, igual que antes, $k^2 = \int_a^b \int_a^b |K^2(s, t)|\, ds\, dt$.

En virtud de la estimación (5), la serie

$$\lambda K(s, t) + \lambda^2 K_2(s, t) + \ldots + \lambda^n K_n(s, t) + \ldots$$

converge para $|\lambda| < \frac{1}{k}$ en el espacio $L_2([a, b] \times [a, b])$ hacia una función $\Gamma(s, t, \lambda)$, cuyo cuadrado es sumable respecto a s y t para todo $|\lambda| < \frac{1}{k}$. El operador integral Γ_λ, para el que la función $\Gamma(s, t, \lambda)$ sirve de núcleo, es la suma de la serie

$$\lambda A + \lambda^2 A^2 + \ldots + \lambda^n A^n + \ldots \tag{6}$$

Pero, agregando a esta suma el operador unidad I, obtendremos precisamente el operador $(I - \lambda A)^{-1}$. Luego, para $|\lambda| < \frac{1}{k}$ el operador $(I - \lambda A)^{-1}$ es, efectivamente, la suma del operador unidad I y del operador totalmente continuo Γ_λ de núcleo

$$\Gamma(s, t, \lambda) = \sum_{n=1}^{\infty} \lambda^n K_n(s, t).$$

La condición $|\lambda| < \frac{1}{k}$ es suficiente para la convergencia de la serie (6), pero no necesaria. En algunos casos puede ocurrir que esta serie converja para todos los valores de λ. Por ejemplo, si A es un operador de tipo de Volterra con un núcleo que satisface la condición

$$|K(s, t)| \leqslant M$$

se puede probar mediante cálculo directo que para los correspondientes núcleos $K_n(s, t)$ es válida la estimación

$$|K_n(s, t)| \leqslant \frac{M^n (b-a)^{n-1}}{(n-1)!},$$

de donde se deduce la convergencia de la serie (6) para cualquier λ.

Sin embargo, la serie de potencias (6) tiene, en general, un radio finito de convergencia. Al mismo tiempo, la ecuación $\varphi = \lambda A \varphi + f$ tiene solución para todos los λ, excepto un número finito o numerable de valores, a saber, excepto aquellos valores para los cuales $\frac{1}{\lambda}$ es valor propio del operador A. Fredholm demostró que para un operador integral A definido por un núcleo acotado y *continuo* $K(s, t)$ la solución de la ecuación $\varphi = \lambda A\varphi + f$ puede ser encontrada del modo siguiente. Designemos

$$K\begin{pmatrix} s_1 & \dots & s_n \\ t_1 & \dots & t_n \end{pmatrix} = \begin{vmatrix} K(s_1, t_1) & \dots & K(s_n, t_n) \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ K(s_n, t_1) & \dots & K(s_n, t_n) \end{vmatrix}$$

y definamos las funciones $D(\lambda)$ y $D(s, t, \lambda)$ mediante las fórmulas

$$D(\lambda) = 1 - \lambda \int_a^b K\begin{pmatrix} \xi_1 \\ \xi_1 \end{pmatrix} d\xi_1 + \frac{\lambda^2}{2} \int_a^b \int_a^b K\begin{pmatrix} \xi_1 & \xi_2 \\ \xi_1 & \xi_2 \end{pmatrix} d\xi_1\, d\xi_2 + \dots$$

$$\dots + (-1)^n \frac{\lambda^n}{n!} \int_a^b \dots \int_a^a K\begin{pmatrix} \xi_1 & \dots & \xi_n \\ \xi_1 & \dots & \xi_n \end{pmatrix} d\xi_1 \dots d\xi_n + \dots \qquad (7)$$

$$D(s, t, \lambda) = K\begin{pmatrix} s \\ t \end{pmatrix} - \lambda \int_a^b K\begin{pmatrix} s & \xi_1 \\ t & \xi_1 \end{pmatrix} d\xi_1 +$$

$$+ \frac{\lambda^2}{2} \int_a^b \int_a^b K\begin{pmatrix} s & \xi_1 & \xi_2 \\ t & \xi_1 & \xi_2 \end{pmatrix} d\xi_1\, d\xi_2 + \dots$$

$$\dots + (-1)^n \frac{\lambda^n}{n!} \int_a^b \dots \int_b^b K\begin{pmatrix} s & \xi_1 & \dots & \xi_n \\ t & \xi_1 & \dots & \xi_n \end{pmatrix} d\xi_1 \dots d\xi_n + \dots \qquad (8)$$

Entonces, según Fredholm, la solución de la ecuación integral

$$\varphi(t) = \lambda \int_a^b K(s, t)\, \varphi(s)\, ds + f(t)$$

viene dada por la fórmula

$$\varphi(t) = f(t) + \int_a^b \frac{D(s, t, \lambda)}{D(\lambda)} f(s)\, ds \qquad (9)$$

para todos los valores de λ tales que $\frac{1}{\lambda}$ no es valor propio del operador integral A correspondiente al núcleo $K(s, t)$. Además, las funciones $D(\lambda)$ y $D(s, t, \lambda)$ son funciones analíticas enteras del parámetro λ y $D(\lambda)=0$ cuando, y sólo cuando, $\frac{1}{\lambda}$ es un valor propio del operador integral A. Como ha demostrado T. Carleman en 1921, las fórmulas (7), (8) y (9), obtenidas por Fredholm para el caso de un núcleo continuo $K(s, t)$, tienen lugar también para cualquier núcleo de cuadrado integrable. No daremos aquí la deducción de la fórmula (9) y de las fórmulas (7) y (8).

BIBLIOGRAFIA

1. Ajiezer N. I., Glazman I. M., *Teoría de los operadores lineales* (Ахиезер Н. И., Глазман И. М., *Теория линейных операторов*, „Наука“, 1966.)
2. Guelfand I. M., Ráikov D. A., Shílov G. E., *Anillos normados conmutativos.* (Гельфанд И. М., Райков Д. А., Шилов Г. Е., *Коммутативные нормированные кольца*, Физматгиз, 1960)
3. Guelfand I. M., Shílov G. E., *Funciones generalizadas, vol. 1*; *Funciones generalizadas y operaciones sobre ellas.* (Гельфанд И. М., Шилов Г. Е., *Обобщенные функции, вып. 1*; *Обобщенные функции и действия над ними*, Физматгиз, 1959)
4. Guelfand I. M., Shílov G. E., *Funciones generalizadas, vol 2*; *Espacios de funciones fundamentales y generalizados.* (Гельфанд И. М., Шилов Г. Д., *Обобщенные функции, вып. 2*; *Пространства основных и обобщенных*, Физматгиз, 1958)
5. Guelfand I. M., Shílov G. E., *Funciones generalizadas, vol. 3.* Algunos problemas de la teoría de las ecuaciones diferenciales. (Гельфанд И. М., Шилов Г. Е., *Обобщенные функции, вып. 3*; *Некоторые вопросы теории дифференциальных уравнений*, Физматгиз, 1958)
6. Guelfand I. M., Vilenkin N. Ya., *Funciones generalizadas, vol. 4*; *Algunas aplicaciones del análisis armónico. Espacios de Hilbert.* (Гельфанд И. М., Виленкин Н. Я., *Обобщенные функции, вып. 4*; *Некоторые применения гармонического анализа. Оснащенные гильбертовы пространства*, Физматгиз, 1961)
7. Kantoróvich L. V., Akílov G. P., Análisis funcional en los espacios normados. (Канторович Л. В., Акилов Г. П., Функциональный анализ в нормированных пространствах, Физматгиз, 1959)
8. Natanson I. P., *Theory of functions of a real variable*, Ungar, Nueva York, 1955.
9. Petrovski I. G., *Vorlesungen über die Theorie der Intgralgleichungen*, Physica—Verlag, Würzburg, 1953; trad. de la 2ª ed. rusa, 1951. Hay también traducción inglesa: *Lectures on the theory of integral equations*, Graylock Press, Rochester, 1957.
10. J. Rey Pastor, *Elementos de la teoría de funciones*, 3ª ed., Iberoamericana, Madrid, 1953.
11. S. Ríos, *Teoría de la integral*, Rev. Acad. Ciencias, vol. 36, Madrid, 1942

12. Shílov G. E. *Análisis matemático*. Segundo curso especial. (Шилов Г. Е., Математический анализ. Второй специальный курс, Физматгиз. 1965).
13. Shílov G. E., Gurévich B. L., *Integral, medida y derivada. Teoría general*. (Шилов Г. Е., Гуревич Б. Л., *Интеграл, мера и производная. Общая теория*, „Наука", 1967).
14. Shílov G. E., Fan Dik Tin, *Integral, medida y derivada en los espacios lineales*
(Шилов Г. Е., Фан Дык Тинь, *Интеграл, мера и производная на линейных пространствах*, „Наука", 1967).
15. Vilenkin N. Ya. y otros, *Análisis funcional* (de la serie "Libros de consulta de Matemáticas").
(Виленкин Н. Я. и др., *Функциональный анализ* (в серии „Справочная математическая библиотека"), „Наука", 1964).

INDICE ALFABETICO