CAPÍTULO 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,

$$1) \qquad \frac{d}{dx}\int_a^x f(t)\,dt = f(x),$$

$$2) \qquad \int_a^b F'(t)\,dt = F(b) - F(a).$$

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 $x_n \to x_0$ 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 $\lim_{n\to\infty} f(x_n)$, 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 $\sum_n h_n < \infty$. 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 entenderemos 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-

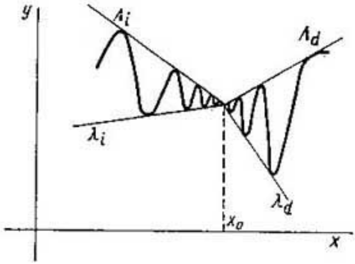


FIG. 20

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 $[a, b]$.

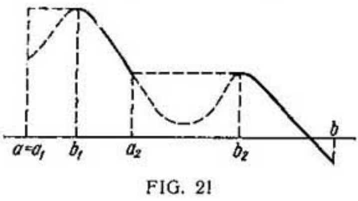


FIG. 21

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 $(\alpha_{k_j}, \beta_{k_j})$ cubre el conjunto $E_\rho \cap (\alpha, \beta)$ y que, además, tenemos, de acuerdo con (8) y con (9),

$$\sum_{k,j} (\beta_{k_j} - \alpha_{k_j}) \leqslant \frac{1}{C} \sum_{k,j} [f(\beta_{k_j}) - f(\alpha_{k_j})] \leqslant$$
$$\leqslant \frac{1}{C} \sum_k [f(\beta_k) - f(\alpha_k)] \leqslant \frac{c}{C} \sum_k (\beta_k - \alpha_k) \leqslant \rho(\beta - \alpha);$$

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 $\sum_m (b_m - a_m) < t + \varepsilon$. Tomemos $t_m = \mu [E_\rho \cap (a_m, b_m)]$. Es obvio que $t = \sum_m t_m$. Por el lema, $t_m \leqslant \rho (b_m - a_m)$. Luego, $t \leqslant \rho \sum_m \times$ $\times (b_m - a_m) < \rho(t + \varepsilon)$ 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. FUNCIONES 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,

$$\varphi(x'')-\varphi(x')=[v(x'')-v(x')]-[f(x'')-f(x')]. \tag{7}$$

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

$$V_b^a[f]\leqslant C(b-a).$$

2. Sea $f(x)=x \operatorname{sen}\dfrac{1}{x}$. 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

$$\frac{\Phi(\xi) - \Phi(x)}{\xi - x} > \beta \tag{2}$$

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

$$\int_S f(t)\,dt \geqslant \beta\mu(S). \tag{3}$$

Al mismo tiempo,

$$\int_S f(t)\,dt = \int_{E_{\alpha\beta}} f(t)\,dt + \int_{S \setminus E_{\alpha\beta}} f(t)\,dt <$$

$$< \alpha\mu(E_{\alpha\beta}) + \varepsilon \leqslant \alpha\mu(S) + \varepsilon + |\alpha|\delta. \tag{4}$$

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

$$F(x)=F(a)+\int_a^x F'(t)\,dt \qquad (1)$$

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]

$$\varphi_h(x)=\frac{f(x+h)-f(x)}{h} \tag{2}$$

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)

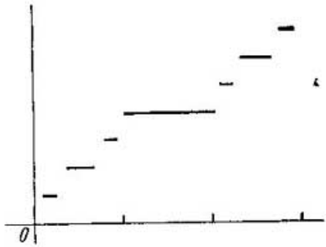


FIG. 22

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

$$\begin{aligned} a_1 < x_{1,1} < x_{1,2} < \ldots < x_{1,m_1} = b_1, \\ a_2 < x_{2,1} < x_{2,2} < \ldots < x_{2,m_2} = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_n < x_{n,1} < x_{n,2} < \ldots < x_{n,m_n} = b_n \end{aligned}$$

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 $f(x) \not\equiv \text{const}!\Big)$ La función $\frac{df}{dx}$ es localmente sumable (suponemos que f es monótona) y, consecuentemente, podemos asignar a esta función una funcional (función generalizada) $(f_1, \varphi) = \int\limits_{-\infty}^{\infty} \frac{df}{dx} \varphi(x)\, dx$. 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 FUNCION 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, $A^- = \bigcup_n A_n$ 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 $\int_A f(x)\,d\mu \leqslant \Phi(A)$ para todo A medible. Sea

$$M = \sup\left\{\int_X f(x)\,d\mu \text{ respecto a todos los } f \in K\right\}.$$

Tomemos en K una sucesión $\{f_n\}$ de funciones tal que

$$\lim_{n\to\infty} \int_X f_n(x)\,d\mu = M.$$

Pongamos ahora

$$g_n(x) = \max(f_1(x),\ f_2(x),\ \ldots,\ f_n(x)).$$

Probemos que para todo E medible es

$$\int_E g_n(x)\,d\mu \leqslant \Phi(E).$$

En efecto, podemos representar E en la forma $\bigcup_{k=1}^{n} E_k$, donde E_k no se intersecan y $g_n(x) = f_k(x)$ en E_k; luego,

$$\int_E g_n(x)\,d\mu = \sum_{k=1}^{n} \int_{E_k} f_k(x)\,d\mu \leqslant \sum_{k=1}^{n} \Phi(E_k) = \Phi(E).$$

Sea

$$f(x) = \sup\{f_n(x)\}.$$

Está claro que $f(x) = \lim_{n\to\infty} g_n(x)$ y, consecuentemente, en virtud del teorema de Beppo Levi,

$$\int_X f(x)\,d\mu = \lim_{n\to\infty} \int_X g_n(x)\,d\mu = M.$$

Probemos ahora que

$$\Phi(E) - \int_E f(x)\,d\mu \equiv 0.$$

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

$$\{x:\ f_1(x) \neq f_2(x)\} = \bigcup_n A_n \cup \bigcup_m B_m,$$

tenemos

$$\mu\{x:\ f_1(x) \neq f_2(x)\} = 0,$$

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 único 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,

$$\left| \int_a^b f(x)\, d\psi(x) \right| \leqslant \left| \int f(x)\, d\psi_N(x) \right| + \left| \int^b f(x)\, d\tilde{\psi}(x) \right| \leqslant 2M\varepsilon;$$

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 $n \to \infty$.

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,

$$\|F\| = V_a^b[\varphi].$$

DEMOSTRACION. 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] Aqui 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,

$$\|f_n-f_m\|\to 0 \text{ para } n,\ m\to\infty.$$

[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

$$|f_{n_1}| + |f_{n_2} - f_{n_1}| + \ldots$$

converge en casi todo X. Pero, entonces, también la serie

$$f_{n_1} + f_{n_2} - f_{n_1} + \ldots$$

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 $y_1, y_2, \ldots$ en los conjuntos $E_1, E_2, \ldots$ 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

$$\mathcal{A} = \{A_n\} \qquad (n = 1, 2, \ldots)$$

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

$$\mu\left[(E_k \setminus A_k) \cup (A_k \setminus E_k)\right] < \varepsilon.$$

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

$$\int |f_n(x) - f_m(x)|\, d\mu \leqslant [\mu(X)]^{1/2} \left\{ \int [f_n(x) - f_m(x)]^2\, d\mu \right\}^{1/2} \leqslant \varepsilon [\mu(X)]^{1/2}, \tag{2}$$

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 $l \to \infty$. Tendremos

$$\int [f_{n_k}(x) - f(x)]^2\, d\mu \leqslant \varepsilon$$

de donde se deduce que $f \in L_2$ y que $f_{n_k} \to f$. 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

$$\varphi_n(x) = \begin{cases} \frac{1}{\sqrt{n}} & \text{para } |x| \leqslant n, \\ 0 & \text{para } |x| > n \end{cases}$$

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

$$f(x) = f^{(n)}(x) \text{ para } x \in X_n,$$

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

$$\int |f_n(x)-f(x)|\,d\mu\leqslant[\mu(X)\int (f_n(x)-f(x))^2\,d\mu]^{\frac{1}{2}}$$

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.

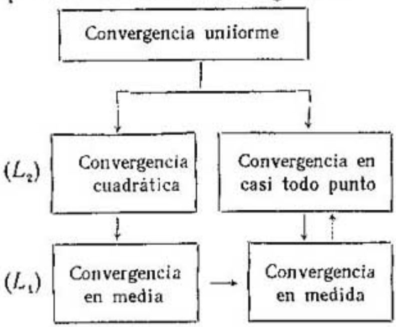
Es fácil ver que la convergencia media (e incluso cuadrática) de una sucesión no implica, en general, la convergencia de esta sucesión en casi todos los puntos. Efectivamente, la sucesión $\{f_{n_k}\}$, construida en el § 4 del cap. VI, converge en media y cuadráticamente hacia $f\equiv 0$; pero, al 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

$$f_n(x)=\begin{cases} \dfrac{1}{\sqrt{n}} & \text{para } |x|\leqslant n, \\ 0 & \text{para } |x|>n, \end{cases}$$

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 ORTOGONALES

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

$$\int_X f_{mn}(x, y)\, f_{m_1n_1}(x, y)\, d\mu=$$
$$=\int_{X''}\psi_n(y)\,\psi_{n_1}(y)\left(\int_{X'}\varphi_m(x)\,\varphi_{m_1}(x)\, d\mu'\right)d\mu''=0.$$

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

$$f(x_1,\ \ldots,\ x_k)=\sum_{n_1,\ \ldots,\ n_k=-\infty}^{\infty} C_{n_1n_2\ldots n_k}e^{i(n_1x_1+\ldots+n_kx_k)},$$

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

$$x=\cos\theta,\ dx=-\operatorname{sen}\theta\,d\theta,\quad \sqrt{1-x^2}=\operatorname{sen}\theta,$$

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

$$Q_m = \sum_{r=0}^{m} c_r P_r \quad (m < n)$$

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

$$\operatorname{sen} \frac{u}{2} = \frac{1}{2} \cdot 2 \operatorname{sen} \frac{u}{2},$$

$$\operatorname{sen} \frac{3u}{2} - \operatorname{sen} \frac{u}{2} = \cos u \cdot 2 \operatorname{sen} \frac{u}{2},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\operatorname{sen} \frac{2n+1}{2} u - \operatorname{sen} \frac{2n-1}{2} u = \cos nu \cdot 2 \operatorname{sen} \frac{u}{2}.$$

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 $p \to \infty$)

$$\int_a^b \varphi(x) \operatorname{sen} px\, dx = -\varphi(x) \frac{\cos px}{p} \Big|_a^b + \int_a^b \varphi'(x) \frac{\cos px}{p} dx \to 0. \tag{8}$$

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 $p \to \infty$, 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,

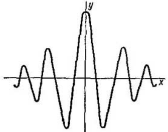


FIG. 23

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

$$\int_{-\pi}^{\pi} |D_n(z)|\, dz \to \infty \text{ para } n \to \infty. \tag{13}$$

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

$$\frac{1}{2\pi}\sum_{k=0}^{n}\frac{1}{2}\frac{1}{\frac{k+l}{2n+1}\pi}\quad\frac{2\pi}{3(2n+1)}=\frac{1}{6\pi}\sum_{k=0}^{n}\frac{1}{k+1}.$$

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.*

DEMOSTRACIÓN. 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 sí 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,

$$f(x)=\frac{1}{2l}\int_{-l}^{l}f(t)\,dt+$$

$$+\frac{1}{l}\sum_{k=1}^{\infty}\int_{-l}^{l}f(t)\left[\cos\frac{k\pi}{l}x\cos\frac{k\pi}{l}t+\operatorname{sen}\frac{k\pi}{l}x\operatorname{sen}\frac{k\pi}{l}t\right]dt=$$

$$=\frac{1}{2l}\int_{-l}^{l}f(t)\,dt+\frac{1}{\pi}\sum_{k=1}^{\infty}\frac{\pi}{l}\int_{-l}^{l}f(t)\cos\frac{k\pi}{l}(t-x)\,dt. \qquad (2)$$

Completemos las suposiciones hechas respecto a la función f con la siguiente: esta función es absolutamente integrable en toda la recta, es decir,

$$\int_{-\infty}^{\infty}|f(t)|\,dt<\infty. \qquad (3)$$

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

$$f(x)=\frac{1}{\pi}\int_{0}^{\infty}d\lambda\int_{-\infty}^{\infty}f(t)\cos\lambda(t-x)\,dt. \qquad (4)$$

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 análoga 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 $\int\limits_{-\infty}^{\infty} e^{-ax^2}\,dx=\sqrt{\frac{\pi}{a}}$.

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

$$F[f^{(k)}] = (i\lambda)^k F[f]. \tag{5}$$

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 $f(x)$, $xf(x)$, $\ldots$, $x^p f(x)$, 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

$$\int_{-\infty}^{\infty} f(x) e^{-i\lambda x}\, dx = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} f_1(\xi) f_2(x-\xi) d\xi \right\} e^{-i\lambda x}\, dx =$$

$$= \int_{-\infty}^{\infty} f_1(\xi) \left\{ \int_{-\infty}^{\infty} f_2(x-\xi) e^{-i\lambda x}\, dx \right\} d\xi =$$

$$= \int_{-\infty}^{\infty} f_1(\xi) \left\{ \int_{-\infty}^{\infty} f_2(\eta) e^{-i\lambda\eta} e^{-i\lambda\xi}\, d\eta \right\} d\xi =$$

$$= \int_{-\infty}^{\infty} f_2(\eta) e^{-i\lambda\eta} d\eta \cdot \int_{-\infty}^{\infty} f_1(\xi) e^{-i\lambda\xi}\, d\xi,$$

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

$$\frac{\partial u(x,\ t)}{\partial t}=\frac{\partial^2 u(x,\ t)}{\partial x^2} \tag{10}$$

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 $f(x_1, x_2, \ldots, x_n)$ es integrable, puede ser representada, de acuerdo con el teorema de Fubini, mediante la siguiente integral reiterada:

$$g(\lambda_1, \lambda_2, \ldots, \lambda_n) = \int_{-\infty}^{\infty} \left\{ \ldots \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_n)\, e^{-ix_1\lambda_1}\, dx_1 \right\} e^{-ix_2\lambda_2}\, dx_2 \ldots \right\} e^{-ix_n\lambda_n}\, dx_n. \quad (11)$$

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

$$f(x_1, x_2, \ldots, x_n) = \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} \left\{ \ldots \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} g(\lambda_1, \lambda_2, \ldots, \lambda_n) e^{ix_n\lambda_n}\, d\lambda_n \right\} e^{ix_{n-1}\lambda_{n-1}}\, d\lambda_{n-1} \ldots \right\} e^{ix_1\lambda_1} d\lambda_1,$$

que puede ser representada mediante una integral n-ple

$$f(x_1, x_2, \ldots, x_n) = \frac{1}{(2\pi)^n} \int \ldots \int g(\lambda_1, \lambda_2, \ldots, \lambda_n)\, e^{i(x_1\lambda_1 + x_2\lambda_2 + \ldots + x_n\lambda_n)}\, d\lambda_1 \ldots d\lambda_n; \quad (12)$$

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 à $f(x_1, x_2, \ldots, x_n)$ 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:*

$$\left.\begin{aligned}
&|f(x_1+t_1, x_2, \ldots, x_n)-f(x_1, x_2, \ldots, x_n)| \leqslant C|t_1|^\alpha,\\
&|f(x_1, x_2+t_2, \ldots, x_n)-f(x_1, x_2, \ldots, x_n)| \leqslant\\
&\qquad\qquad\qquad\qquad\leqslant C(x_1)|t_2|^\alpha,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&|f(x_1, x_2, \ldots, x_n+t_n)-f(x_1, x_2, \ldots, x_n)| \leqslant\\
&\qquad\qquad\qquad\leqslant C(x_1, x_2, \ldots, x_{n-1})|t_n|^\alpha,\\
&\text{donde } 0\leqslant\alpha\leqslant 1, \quad \int_{-\infty}^{\infty} C(x_1)\,dx_1<\infty, \ \ldots,\\
&\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} C(x_1, x_2, \ldots, x_{n-1})\,dx_1dx_2\ldots dx_{n-1}<\infty.
\end{aligned}\right\} \quad (13)$$

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

$$v(t, \lambda\sigma) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u(t, x, y)\, e^{-i(\lambda x+\sigma y)}\, dx\, dy. \qquad (15)$$

Resolviendo la ecuación (15), podemos después encontrar la solución de la ecuación inicial (14) por medio de la fórmula de inversión.

§ 5. TRANSFORMACIÓN 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

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x) e^{-inx} \quad dx \, (n=0, \pm 1, \pm 2, \ldots).$$

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:

$$2\pi \sum_{n=-\infty}^{\infty} |c_n|^2 = \int_{-\pi}^{\pi} |f(x)|^2\, dx \qquad (1)$$

(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

$$f_N(x)=\begin{cases} f(x) & \text{para } |x|\leqslant N,\\ 0 & \text{para } |x|>N.\end{cases}$$

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

$$\frac{d^2 f}{dx^2} - x^2 f = \mu f \tag{3}$$

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

$$w = a_0 + a_1 x + \ldots + a_n x^n, \tag{4}$$

encontramos la igualdad

$$(2a_2 + 3 \cdot 2 \cdot a_3 x + \ldots + n(n-1) a_n x^{n-2}) -$$
$$-2x(a_1 + 2a_2 x + \ldots + na_n x^{n-1}) = (\mu + 1)(a_0 + a_1 x + \ldots + a_n x^n).$$

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

$$e^{-\frac{x^2}{2}},\quad xe^{-\frac{x^2}{2}},\ \dots,\ x^n e^{-\frac{x^2}{2}},\ \dots$$

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:

$$\left.\begin{aligned} f(x) &< Ce^{\gamma_0 x} \text{ para } x \geqslant 0, \\ f(x) &= 0 \text{ para } x < 0 \end{aligned}\right\} \tag{1}$$

(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 $\operatorname{Im} s = -\gamma_0$, 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 inversión

$$f(x)e^{\mu x}=\frac{1}{2\pi}\int_{-\infty}^{\infty} g(s)e^{i\lambda x}\,d\lambda,$$

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 periodo 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

$$F(x) = \int_{-\infty}^{x} f(t)\, dt, \quad F_1(x) = \int_{-\infty}^{x} f_1(t)\, dt \quad \text{y} \quad F_2(x) = \int_{-\infty}^{x} f_2(t)\, dt.$$

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, además 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. Shilov [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.

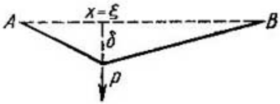


FIG. 24

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 distribución 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 $\rho=\text{const}$ aunque esto no es sustancial para lo sucesivo.

ecuación integral para u:

$$u(x)=\rho\omega^2\int_0^l G(x,\ \xi)\,u(\xi)\,d\xi. \tag{7}$$

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

$$y''+\rho^2y=\sigma(x)\,y. \tag{8}$$

Como se sabe, la solución de la ecuación

$$y''+\rho^2y=g(x)$$

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

$$\int_a^b \int_a^b (K(s, t) - K_N(s, t))^2\,ds\,dt \to 0 \text{ para } N \to \infty.$$

Aplicando la estimación (4) al operador $A - A_N$, encontramos de aquí

$$\| A - A_N \| \to 0 \quad \text{para} \quad N \rightrightarrows \infty.$$

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

$$T^*z = 0 \quad (T^* = \| a_{ki} \|).$$

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

$$Tx + \sum_{j=1}^{\mu} (x, \varphi_j)\psi_j = 0. \qquad (25)$$

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 PARÁMETRO. 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.

$(1-\lambda A)^{-1}$ 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

$$(I-\lambda A)^{-1} = I + \lambda A + \lambda^2 A^2 + \ldots + \lambda^n A^n + \ldots,$$

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

$$\varphi = f + \lambda A f + \lambda^2 A^2 f + \ldots + \lambda^n A^n f + \ldots \tag{3}$$

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

$$\varphi_0 = f,\ \varphi_1 = Af,\ \ldots,\ \varphi_n = A\varphi_{n-1} = A^n f,\ \ldots,$$

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

$$K\begin{pmatrix} s_1 & \dots & s_n \\ t_1 & \dots & t_n \end{pmatrix} = \begin{vmatrix} K(s_1, t_1) & \dots & K(s_n, t_n) \\ \dots & \dots & \dots \\ K(s_n, t_1) & \dots & K(s_n, t_n) \end{vmatrix}$$

y definamos las funciones $D(\lambda)$ y $D(s, t, \lambda)$ mediante las fórmulas

$$D(\lambda) = 1 - \lambda \int_a^b K\begin{pmatrix} \xi_1 \\ \xi_1 \end{pmatrix} d\xi_1 + \frac{\lambda^2}{2} \int_a^b \int_a^b K\begin{pmatrix} \xi_1 & \xi_2 \\ \xi_1 & \xi_2 \end{pmatrix} d\xi_1\, d\xi_2 + \dots$$

$$\dots + (-1)^n \frac{\lambda^n}{n!} \int_a^b \dots \int_a^a K\begin{pmatrix} \xi_1 & \dots & \xi_n \\ \xi_1 & \dots & \xi_n \end{pmatrix} d\xi_1 \dots d\xi_n + \dots \quad (7)$$

$$D(s, t, \lambda) = K\begin{pmatrix} s \\ t \end{pmatrix} - \lambda \int_a^b K\begin{pmatrix} s & \xi_1 \\ t & \xi_1 \end{pmatrix} d\xi_1 +$$

$$+ \frac{\lambda^2}{2} \int_a^b \int_a^b K\begin{pmatrix} s & \xi_1 & \xi_2 \\ t & \xi_1 & \xi_2 \end{pmatrix} d\xi_1\, d\xi_2 + \dots$$

$$\dots + (-1)^n \frac{\lambda^n}{n!} \int_a^b \dots \int_b^b K\begin{pmatrix} s & \xi_1 & \dots & \xi_n \\ t & \xi_1 & \dots & \xi_n \end{pmatrix} d\xi_1 \dots d\xi_n + \dots \quad (8)$$

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

$$\varphi(t) = f(t) + \int_a^b \frac{D(s, t, \lambda)}{D(\lambda)} f(s)\, ds \quad (9)$$

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 Dík 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

ESTIMADO LECTOR:

La Editorial le quedará muy agradecida, si Ud. nos manda su opinión acerca del libro que le ofrecemos, así como de la traducción y presentación del mismo.

Le agradeceremos también cualquier otra sugerencia respecto a la edición de libros que le interesan.

Dirija, por favor, su opinión y sugerencias a la Editorial Mir:

Editorial Mir, 1 Rizhski per, 2, Moscú, 129820, I-110, URSS.