

ANÁLISIS FUNCIONAL
Notas de curso

M.A. RODRÍGUEZ
Departamento de Física Teórica II
Universidad Complutense de Madrid

21 de septiembre de 2012

Índice general

1. La integral de Lebesgue	3
1.1. Introducción	3
1.2. Medidas	4
1.2.1. Medidas en intervalos	4
1.2.2. Medidas en anillos	7
1.2.3. Medidas exteriores	10
1.2.4. Medidas interiores	12
1.2.5. La medida de Lebesgue	12
1.2.6. Propiedades de la medida de Lebesgue	13
1.3. Funciones medibles	15
1.4. Integración	16
1.4.1. Conceptos generales	17
1.4.2. La integral de Lebesgue	18
1.4.3. La integral de Riemann	21
1.4.4. Integrales impropias	22
1.5. Productos de medidas	26
2. Topología, distancia y norma	29
2.1. Introducción	29
2.2. Espacios topológicos	29
2.3. Espacio métricos	34
2.4. Espacios normados	36
3. Espacios de Banach	43
3.1. Espacios normados de dimensión finita	43
3.2. Espacios de sucesiones l^p	46
3.3. Espacios L^p	48
4. Espacios de Hilbert	55
4.1. Introducción	55
4.2. Definiciones y propiedades	55
4.3. Ortogonalidad	58
4.4. Polinomios ortogonales	64
4.5. Series de Fourier	65
5. Operadores en espacios de Hilbert	67
5.1. Introducción	67
5.2. Operadores acotados	67
5.3. Funcionales lineales continuos	69
5.4. Funcionales bilineales hermiticos	71
5.5. Operadores autoadjuntos y unitarios	72
5.6. Operadores positivos	77
5.7. Operadores compactos	78

5.8. Operadores de clase de traza y Hilbert-Schmidt	78
5.9. Ecuaciones integrales	79
6. Espectros de operadores	81
6.1. Introducción	81
6.2. El espectro puntual y el espectro continuo	82
6.3. El resolvente	83
6.4. El espectro de un operador y de su adjunto	83
6.5. Espectro de operadores acotados normales	84
6.6. Espectro de operadores compactos	84
6.7. Ecuaciones integrales	84
7. Distribuciones	87
7.1. Ideas básicas sobre distribuciones	87
7.1.1. El espacio de las funciones prueba de soporte compacto	87
7.1.2. El espacio de distribuciones	87
7.1.3. Operaciones con distribuciones	88
7.1.4. Convergencia en el espacio de distribuciones	88
7.2. Topología	88
7.2.1. Espacios localmente convexos	88
7.2.2. Funciones prueba de soporte compacto	89
7.2.3. Distribuciones	90
7.2.4. Topología en \mathcal{D}'	90
7.3. Propiedades y operaciones	92
7.3.1. Derivación	92
7.3.2. Multiplicación por funciones	93
7.3.3. El producto de convolución	94
7.4. Ejemplos de distribuciones	97
7.4.1. La delta de Dirac	97
7.4.2. Distribuciones asociadas a capas simples y dobles	97
7.4.3. Las fórmulas de Sochozki	98
7.4.4. Derivadas de distribuciones	99
7.4.5. Soluciones fundamentales	101
7.5. Transformada de Fourier	103
7.5.1. El espacio \mathcal{S}_n	103
7.5.2. La transformación de Fourier en $L^1(\mathbf{R}^n)$	103
7.5.3. La transformación de Fourier en \mathcal{S}_n	104
7.5.4. La fórmula de inversión	105
7.5.5. La transformación de Fourier en $L^2(\mathbf{R}^n)$	106
7.5.6. Distribuciones temperadas	106
7.5.7. Transformación de Fourier en \mathcal{S}'_n	107
7.5.8. Convolución en \mathcal{S}_n y \mathcal{S}'_n	108
7.6. Ejemplos de transformadas de Fourier	108
7.7. Soluciones fundamentales	113
7.7.1. Propiedades y cálculo de soluciones fundamentales	113
7.7.2. Soluciones fundamentales de ecuaciones diferenciales ordinarias	114
7.7.3. Soluciones fundamentales de las ecuaciones del calor y ondas	117

Capítulo 1

La integral de Lebesgue

1.1. Introducción

El objetivo de este capítulo es un estudio de la integral de Lebesgue. Los espacios vectoriales topológicos más utilizados en Física son los espacios de funciones, y en particular (al menos en Mecánica Cuántica) los espacios de funciones integrables Lebesgue (o cuya potencia p sea integrable Lebesgue). Se hace necesario por tanto presentar, aunque sea someramente, el concepto de integral de Lebesgue y previamente a éste (por razones además de aplicación a ciertos elementos de cálculo funcional de operadores) una introducción al concepto de medida. La pregunta es, si queremos espacios de funciones de cuadrado integrable (por ejemplo en Mecánica Cuántica), ¿por qué no usamos la integral de Riemann? Una de las respuestas es que esta integral es insuficiente, en particular en lo que respecta a su comportamiento en relación con sucesiones. El ejemplo más inmediato es el siguiente.

Ejemplo 1.1 Sea $\{r_1, r_2, \dots, r_m, \dots\}$ el conjunto de números racionales del intervalo $(0, 1)$. Consideremos la siguiente sucesión de funciones:

$$f_n(x) = \begin{cases} 1, & x \in \{r_1, \dots, r_n\} \\ 0, & x \in [0, 1] \setminus \{r_1, \dots, r_n\} \end{cases} \quad (1.1)$$

Las funciones f_n son integrables Riemann (tienen un número finito de discontinuidades, son continuas a trozos). Además la integral en $[0, 1]$ es cero. Cuando se toma el límite (puntual) de la sucesión f_n se tiene la función:

$$f(x) = \begin{cases} 1, & x \in \mathbf{Q} \\ 0, & x \in [0, 1] \setminus \mathbf{Q} \end{cases} \quad (1.2)$$

Pero la función f no es integrable Riemann. En cualquier partición del intervalo $[0, 1]$, los subintervalos contienen puntos racionales e irracionales, por lo que la función oscila entre 0 y 1 y las sumas inferiores y superiores son constantemente iguales a 0 y 1.

Otro ejemplo que también, aunque no de manera tan clara, obliga a extender el concepto de integral es el siguiente relativo a la completación de ciertos espacios de funciones.

Ejemplo 1.2 Sea la sucesión de funciones en $[0, 1]$ definidas por ($n \geq 2$):

$$f_n(x) = \begin{cases} 0, & x < \frac{1}{2} - \frac{1}{n} \\ nx - \frac{n}{2} + 1, & \frac{1}{2} - \frac{1}{n} \leq x \leq \frac{1}{2} \\ 1, & x > \frac{1}{2} \end{cases} \quad (1.3)$$

Es evidente que se trata de una sucesión de funciones continuas. Pero su límite (de nuevo puntual) es una función discontinua:

$$f(x) = \begin{cases} 0, & x < \frac{1}{2} \\ 1, & x \geq \frac{1}{2} \end{cases} \quad (1.4)$$

Consideremos el espacio $C[0, 1]$ de las funciones continuas en el intervalo $[0, 1]$ (continuidad en 0 y 1 quiere decir que existen los límites por la derecha e izquierda respectivamente). En este espacio definimos una distancia por medio de la integral de Riemann:

$$d(f, g) = \int_0^1 |f - g| dx \quad (1.5)$$

bien definida al ser las funciones continuas en $[0, 1]$ integrables Riemann (aunque no entraremos en detalles ahora, se trata de una distancia en el sentido usual). De acuerdo con ella, la sucesión de funciones es de Cauchy, es decir:

$$d(f_n, f_m) = \int_0^1 |f_n(x) - f_m(x)| dx = \frac{1}{2} \left| \frac{1}{n} - \frac{1}{m} \right| \rightarrow 0 \quad (1.6)$$

cuando n, m tienden a infinito. Sin embargo su límite no es una función continua y no está en $C[0, 1]$. La cuestión es si existe una función continua a la que tienda esta función en la distancia definida anteriormente (aunque no tienda puntualmente a ella). La respuesta es que no. En efecto, sea $f(x)$ una función en $C[0, 1]$, tal que:

$$d(f_n, f) = \int_0^1 |f_n(x) - f(x)| dx \quad (1.7)$$

tienda a cero cuando $n \rightarrow \infty$. Se tiene:

$$\begin{aligned} \int_0^1 |f_n(x) - f(x)| dx &= \int_0^{\frac{1}{2} - \frac{1}{n}} |f(x)| dx + \\ &\int_{\frac{1}{2} - \frac{1}{n}}^{\frac{1}{2}} \left| nx - \frac{n}{2} + 1 - f(x) \right| dx + \int_{\frac{1}{2}}^1 |1 - f(x)| dx \end{aligned}$$

Los tres sumandos son positivos luego todos deben hacerse cero cuando $n \rightarrow \infty$. El segundo es de esa forma, pues f es continua y la longitud del intervalo tiende a cero. El tercero no depende de n luego el integrando debe ser cero, y por tanto $f(x) = 1, x > 1/2$. En cuanto al primero, cuando $n \rightarrow \infty$, el extremo superior tiende a $1/2$, y por lo tanto $f(x) = 0, x < 1/2$. Pero entonces f no puede ser continua.

Se dice que el espacio $C[0, 1]$ no es completo con esta distancia. Como veremos la completación de este espacio con esta distancia, es decir qué funciones hay que añadir para conseguir un espacio completo, se consigue al incluir las funciones integrables Lebesgue (no basta con las integrables Riemann).

1.2. Medidas

Nuestro objetivo es como hemos dicho definir la integral de Lebesgue y estudiar sus propiedades. Para ello necesitamos unas nociones básicas de teoría de la medida que iremos introduciendo a medida que se necesiten, teniendo siempre como guía la mencionada integral de Lebesgue.

1.2.1. Medidas en intervalos

Consideremos una clase de conjuntos en la recta real,

$$\mathcal{P} = \{ [a, b) : -\infty < a \leq b < \infty \}$$

es decir, los intervalos semicerrados acotados de la recta real. En esta clase de conjuntos definimos una función (de conjuntos) con valores en la recta real:

$$\lambda([a, b)) = b - a \quad (1.8)$$

que tiene unas propiedades claras:

Proposición 1.2.1 *La función λ verifica ($I \in \mathcal{P}$):*

1. $\lambda(I) \geq 0$
2. $\lambda(\emptyset) = 0$
3. $\lambda(I) < \infty$

Nótese que la clase \mathcal{P} no es cerrada bajo la unión de conjuntos (la unión de dos intervalos no es en general un intervalo) ni bajo la diferencia de conjuntos.

Veamos algunas propiedades más de esta función λ y la clase \mathcal{P} .

Proposición 1.2.2 *Sea $I_0 = [a_0, b_0] \in \mathcal{P}$, $I_j = [a_j, b_j] \in \mathcal{P}$, $j = 1, \dots, n$, una clase de conjuntos disjuntos, contenidos en I_0 . Entonces,*

$$\sum_{j=1}^n \lambda(I_j) \leq \lambda(I_0)$$

La demostración es muy sencilla y responde a una idea intuitiva de la medida de un intervalo. Puesto que los intervalos I_j son disjuntos podemos considerarlos ordenados:

$$a_0 \leq a_1 \leq b_1 \leq a_2 \leq \dots \leq b_n \leq b_0$$

Por definición de λ :

$$\sum_{j=1}^n \lambda(I_j) = \sum_{j=1}^n (b_j - a_j) \leq \sum_{j=1}^n (b_j - a_j) + \sum_{j=1}^{n-1} (a_{j+1} - b_j) = b_n - a_1 \leq \lambda(I_0)$$

Nótese que en este caso se tiene una unión de conjuntos de \mathcal{P} que están dentro de otro (también en \mathcal{P}) y se comparan las medidas. Veamos ahora la situación en la que un conjunto está contenido en la unión de otros. En primer lugar consideraremos el caso de un conjunto cerrado contenido en una unión de abiertos.

Proposición 1.2.3 *Sea $F = [a_0, b_0]$ un intervalo cerrado contenido en la unión de una familia finita de intervalos abiertos acotados, $A_j = (a_j, b_j)$, $j = 1, \dots, n$. Entonces:*

$$b_0 - a_0 < \sum_{j=1}^n (b_j - a_j)$$

Nótese que los conjuntos de los que se está hablando ahora no están en \mathcal{P} , por lo que no se menciona a λ que, en principio, solo se aplica a la clase \mathcal{P} .

Puesto que F está en la unión de los intervalos abiertos, a_0 está en alguno de ellos. Sea A_{k_1} dicho intervalo. Si $b_{k_1} \leq b_0$, b_{k_1} es un punto de F y por tanto estará en algún otro intervalo, A_{k_2} . De esta forma podemos continuar hasta que $b_{k_m} > b_0$. Para simplificar, reordenamos los intervalos y eliminamos los posteriores al A_{k_m} pues no juegan ningún papel en la proposición (es decir, sea $m = n$). Tenemos entonces:

$$a_1 < a_0 < b_1, \quad a_n < b_0 < b_n$$

y para los intermedios (si es que existen):

$$a_{j+1} < b_j < b_{j+1}, \quad j = 1, \dots, n-1$$

Ahora es muy sencillo obtener la desigualdad del enunciado:

$$b_0 - a_0 < b_n - a_1 = b_1 - a_1 + \sum_{j=1}^{n-1} (b_{j+1} - b_j) \leq \sum_{j=1}^n (b_j - a_j)$$

Con este resultado podemos pasar ahora a uno similar en la clase \mathcal{P} .

Proposición 1.2.4 Sea $I_0 = [a_0, b_0) \in \mathcal{P}$, $I_j = [a_j, b_j) \in \mathcal{P}$, $j = 1, 2, \dots$ una sucesión de conjuntos, tales que $I_0 \subset \bigcup_{j=1}^{\infty} I_j$. Entonces,

$$\lambda(I_0) \leq \sum_{j=1}^{\infty} \lambda(I_j)$$

Supongamos que $b_0 > a_0$ (en otro caso el resultado es trivial). Entonces, existe ϵ con $0 < \epsilon < b_0 - a_0$. Sea $F_0 = [a_0, b_0 - \epsilon]$ y para cualquier $\delta > 0$ definimos:

$$A_j = \left(a_j - \frac{\delta}{2^j}, b_j \right)$$

Entonces $F_0 \subset I_0$, $I_j \subset A_j$, y se tiene

$$F_0 \subset \bigcup_{j=1}^{\infty} A_j$$

Pero F_0 es un compacto, y por tanto de todo recubrimiento por abiertos podemos extraer un subrecubrimiento finito:

$$F_0 \subset \bigcup_{i=1}^n A_i$$

Ahora estamos en las condiciones del teorema anterior, y entonces:

$$b_0 - a_0 - \epsilon < \sum_{j=1}^n \left(b_j - a_j + \frac{\delta}{2^j} \right) = \sum_{j=1}^n (b_j - a_j) + \delta \sum_{j=1}^n \frac{1}{2^j}$$

es decir,

$$\lambda(I_0) - \epsilon < \sum_{j=1}^{\infty} \lambda(I_j) + \delta$$

Pero ϵ y δ son arbitrarios (suficientemente pequeños) y se tiene la desigualdad buscada.

La unión de intervalos semicerrados no es en general un intervalo semicerrado. Lo que es lo mismo, \mathcal{P} no es cerrado bajo uniones de conjuntos. Sin embargo, en algunos casos se da esta situación.

Proposición 1.2.5 Dada una clase numerable de conjuntos de \mathcal{P} , I_j , $j = 1, 2, \dots$, disjuntos dos a dos, tal que $\bigcup_{j=1}^{\infty} I_j \in \mathcal{P}$, se tiene:

$$\lambda \left(\bigcup_{j=1}^{\infty} I_j \right) = \sum_{j=1}^{\infty} \lambda(I_j)$$

Sea $I = \bigcup_{j=1}^{\infty} I_j$, con I, I_j , ($j = 1, 2, \dots$) $\in \mathcal{P}$. Por tanto, $I_j \subset I$ para todo j . Estamos en las hipótesis de la proposición 1.2.2. Entonces

$$\sum_{j=1}^n \lambda(I_j) \leq \lambda(I)$$

para cualquier valor de n . Pero eso quiere decir que:

$$\sum_{j=1}^{\infty} \lambda(I_j) \leq \lambda(I)$$

Aplicando la desigualdad en sentido contrario que se deduce de la proposición 1.2.4 se llega a la conclusión buscada.

Como hemos visto en las anteriores proposiciones, la función λ se aplica a conjuntos de una determinada clase y tiene una serie de propiedades que la hacen muy interesante para nuestros propósitos. Por supuesto hay casos más generales que la definición dada en (1.8) y conjuntos más generales a los que aplicarla que \mathcal{P} .

1.2.2. Medidas en anillos

Definiremos a continuación una serie de conceptos que forman la base de la teoría de la medida.

La idea fundamental es asignar a un conjunto un número (real) que será la medida de ese conjunto. Pero antes necesitamos especificar a que tipo de conjuntos les vamos a asignar una medida.

Sea X un conjunto.

Definición 1.2.1 *Se dice que una colección \mathcal{R} de subconjuntos de X es un anillo si se verifican las dos propiedades siguientes:*

1. Si $A, B \in \mathcal{R}$, entonces $A \cup B \in \mathcal{R}$
2. Si $A, B \in \mathcal{R}$, entonces $A - B \in \mathcal{R}$,

La diferencia de dos subconjuntos de X se define por:

$$A, B \in X, \quad A - B = A \cap B'$$

donde B' es el complementario de B en X .

La diferencia simétrica de dos conjuntos se define como:

$$A \Delta B = (A - B) \cup (B - A)$$

Un anillo es cerrado bajo esta operación.

Por ejemplo, la clase \mathcal{P} no es un anillo, pero la colección formada por las uniones (finitas) de los intervalos semicerrados, $\mathcal{R}_{\mathcal{P}}$, es un anillo.

Dada una colección, \mathcal{E} , de subconjuntos de X existe un anillo minimal que los contiene, $\mathcal{R}(\mathcal{E})$ (es la intersección de todos los anillos que contienen a \mathcal{E}). Se dice que $\mathcal{R}(\mathcal{E})$ está generado por \mathcal{E} .

Si sustituimos la condición de unión finita por la de unión numerable, tendremos un σ -anillo:

$$A_i \in \mathcal{S}, \quad i \in I \text{ (numerable)}, \quad \bigcup_{i \in I} A_i \in \mathcal{S}$$

Definición 1.2.2 *Una σ -álgebra es un σ -anillo que contiene al conjunto X .*

Por tanto, si un conjunto pertenece a una σ -álgebra también su complementario está en ella.

Sobre estas estructuras definimos una medida. Sea $\bar{\mathbf{R}}$ la recta real extendida: $\bar{\mathbf{R}} = \mathbf{R} \cup \{-\infty, \infty\}$

Definición 1.2.3 *Una medida es una función definida en un anillo con valores en la recta extendida, tal que:*

1. $\forall A \in \mathcal{R} \quad \mu(A) \geq 0$
2. $\mu(\emptyset) = 0$
3. Si $A_n, n = 1, 2, \dots$ es una clase numerable de elementos de \mathcal{R} disjuntos dos a dos, tales que $\bigcup_{n=1}^{\infty} A_n \in \mathcal{R}$, entonces

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

Se dice que μ es σ -aditiva.

Volvamos a la función λ . Tenemos una función de conjunto λ (1.8) definida en \mathcal{P} . Pero \mathcal{P} no es un anillo. Como $\mathcal{P} \subset \mathcal{R}_{\mathcal{P}}$ y $\mathcal{R}_{\mathcal{P}}$ sí es un anillo, extendemos la función λ a una medida en $\mathcal{R}_{\mathcal{P}}$.

Proposición 1.2.6 *Existe una única medida $\bar{\lambda}$ (que además es finita) en $\mathcal{R}_{\mathcal{P}}$ que es la extensión de λ al anillo $\mathcal{R}_{\mathcal{P}}$ (es decir, si $A \in \mathcal{P}$, entonces $\lambda(A) = \bar{\lambda}(A)$).*

Probaremos primeramente que la medida de un conjunto de $\mathcal{R}_{\mathcal{P}}$, que se puede expresar como un unión finita de elementos de \mathcal{P} , no depende de la elección de esos elementos. Sea $A \in \mathcal{R}_{\mathcal{P}}$ y

$$A = \bigcup_{k=1}^n I_k = \bigcup_{l=1}^m J_l, \quad I_k, J_l \in \mathcal{P}, \text{ disjuntos}$$

Descomponemos cada I_k de la forma siguiente:

$$I_k = \bigcup_{l=1}^m (I_k \cap J_l)$$

que es una clase de conjuntos disjuntos en \mathcal{P} . Aplicando los resultados anteriores:

$$\lambda(I_k) = \sum_{l=1}^m \lambda(I_k \cap J_l)$$

y sumando en k

$$\sum_{k=1}^n \lambda(I_k) = \sum_{k=1}^n \sum_{l=1}^m \lambda(I_k \cap J_l)$$

De la misma forma:

$$\sum_{l=1}^m \lambda(J_l) = \sum_{l=1}^m \sum_{k=1}^n \lambda(J_l \cap I_k)$$

Como los términos de la derecha son iguales los de la izquierda también:

$$\bar{\lambda}(A) = \sum_{k=1}^n \lambda(I_k) = \sum_{l=1}^m \lambda(J_l)$$

Cualquier otra función sobre $\mathcal{R}_{\mathcal{P}}$ que sea finitamente aditiva y que sobre \mathcal{P} sea igual a λ debe coincidir con $\bar{\lambda}$ (pues los elementos de $\mathcal{R}_{\mathcal{P}}$ son uniones finitas de elementos disjuntos de \mathcal{P} y la función debe ser finitamente aditiva).

Veamos ahora que, además, $\bar{\lambda}$ es σ -aditiva. Consideremos una sucesión de elementos disjuntos de $\mathcal{R}_{\mathcal{P}}$, A_k , $k = 1, 2, \dots$, con unión, A , en $\mathcal{R}_{\mathcal{P}}$ (recuérdese que $\mathcal{R}_{\mathcal{P}}$ no es un σ -anillo, por lo que siempre tenemos que imponer esta condición). Por estar A_k en $\mathcal{R}_{\mathcal{P}}$, es una unión de elementos disjuntos de \mathcal{P} :

$$A_k = \bigcup_{l=1}^{n_k} I_{kl}$$

Por definición de $\bar{\lambda}$,

$$\bar{\lambda}(A_k) = \sum_{l=1}^{n_k} \lambda(I_{kl})$$

Consideremos en primer lugar el caso en que A es un elemento de \mathcal{P} . Entonces:

$$\bar{\lambda}(A) = \lambda(A) = \sum_{k=1}^{\infty} \sum_{l=1}^{n_k} \lambda(I_{kl}) = \sum_{k=1}^{\infty} \bar{\lambda}(A_k)$$

Si $A \notin \mathcal{P}$, entonces:

$$A = \bigcup_{r=1}^m J_r, \quad J_r \in \mathcal{P}, \quad \text{disjuntos}$$

y para cada J_r usamos el resultado anterior:

$$\bar{\lambda}(A) = \sum_{r=1}^m \bar{\lambda}(J_r) = \sum_{r=1}^m \sum_{k=1}^{\infty} \bar{\lambda}(A_k \cap J_r) = \sum_{k=1}^{\infty} \sum_{r=1}^m \bar{\lambda}(A_k \cap J_r)$$

pero

$$\sum_{r=1}^m \bar{\lambda}(A_k \cap J_r) = \bar{\lambda}(A_k)$$

luego $\bar{\lambda}$ es σ -aditiva.

En resumen, partiendo de una función de conjuntos en la clase de intervalos semicerrados acotados de \mathbf{R} , se ha definido una medida σ -finita (es decir, todo conjunto medible está contenido en la unión numerable de una sucesión de conjuntos medibles de medida finita) en un anillo formado por uniones finitas disjuntas de elementos de \mathcal{P} y esa medida restringida a la clase \mathcal{P} reproduce la función inicial.

Las medidas (y en particular la definida anteriormente) tienen unas propiedades interesantes en cuanto a su relación con las sucesiones de conjuntos. Veamos primeramente que algunas de las propiedades establecidas anteriormente son generalizables a cualquier medida en un anillo.

Una medida en un anillo es monótona (si $A \subset B$ entonces $\mu(A) \leq \mu(B)$) y substractiva (si $A \subset B$, $\mu(A) < \infty$, entonces $\mu(A - B) = \mu(A) - \mu(B)$). Las mismas propiedades que encontramos en el caso particular de intervalos de la recta real se tienen ahora en general.

Proposición 1.2.7 *Sea μ una medida en un anillo \mathcal{R} , $A \in \mathcal{R}$, $A_n \in \mathcal{R}$, $n = 1, 2, \dots$, tales que $A \subset \cup_n A_n$. Entonces $\mu(A) \leq \sum_n \mu(A_n)$*

Proposición 1.2.8 *Sea μ una medida en un anillo \mathcal{R} , $A \in \mathcal{R}$, $A_n \in \mathcal{R}$, $n = 1, 2, \dots$, disjuntos, tales que $\cup_n A_n \subset A$. Entonces $\sum_n \mu(A_n) \leq \mu(A)$*

Dada una sucesión de subconjuntos $\{A_n\}$ de un conjunto X , se dice que A es el límite superior de la sucesión si todo elemento de A está en un número infinito de elementos de la sucesión. Se dice que A es el límite inferior de la sucesión si todo elemento de A está en todos excepto en un número finito de elementos de la sucesión. Si ambos límites coinciden se dice que A es el límite de la sucesión. Las sucesiones monótonas, $A_n \subset A_{n+1}$ (crecientes) o $A_n \supset A_{n+1}$ (decrecientes) tienen límite. En el primer caso (crecientes) el límite es la unión de todos los elementos de la sucesión. En el segundo (decrecientes) el límite es la intersección de todos los elementos de la sucesión.

Tenemos las dos siguientes propiedades de las medidas:

Proposición 1.2.9 *Sea μ una medida en un anillo \mathcal{R} y A_n una sucesión creciente de elementos del anillo, con límite en \mathcal{R} . Entonces*

$$\mu(\lim_n A_n) = \lim_n \mu(A_n)$$

Si A_n es una sucesión decreciente y $\mu(A_{n_0}) < \infty$ para algún n_0 , se tiene también la igualdad anterior.

En el primer caso (sea $A_0 = \emptyset$), se tiene (el único problema es que la sucesión no es en general de conjuntos disjuntos):

$$\mu(\lim_n A_n) = \mu\left(\bigcup_n A_n\right) = \mu\left(\bigcup_n (A_n - A_{n-1})\right)$$

Ahora ya tenemos una colección de elementos disjuntos y como μ es σ -aditiva,

$$\begin{aligned} \mu(\lim_n A_n) &= \sum_n \mu(A_n - A_{n-1}) = \lim_n \sum_{k=1}^n \mu(A_k - A_{k-1}) \\ &= \lim_n \mu\left(\bigcup_{k=1}^n (A_k - A_{k-1})\right) = \lim_n \mu(A_n) \end{aligned}$$

En el segundo caso, al ser $\mu(A_{n_0}) < \infty$, $\mu(A_n) < \infty$ cuando $n \geq n_0$. Por lo tanto el límite también tiene medida menor que ∞ . Además la sucesión $A_{n_0} - A_n$ es creciente y se puede aplicar el resultado anterior:

$$\begin{aligned} \mu(A_{n_0}) - \mu(\lim_n A_n) &= \mu(\lim_n (A_{n_0} - A_n)) = \lim_n (\mu(A_{n_0}) - \mu(A_n)) \\ &= \mu(A_{n_0}) - \lim_n \mu(A_n) \end{aligned}$$

1.2.3. Medidas exteriores

El que un conjunto esté en un anillo, no implica, obviamente, que sus subconjuntos lo estén. Debido a esto, es posible que un subconjunto de medida cero contenga subconjuntos que ni siquiera sean medibles (estén en el anillo correspondiente).

Consideremos un σ -anillo \mathcal{H} , en el que todo subconjunto de un conjunto de \mathcal{H} esté también en \mathcal{H} (se llaman anillos hereditarios).

Definición 1.2.4 *Se dice que una función μ^* sobre \mathcal{H} con valores en la recta extendida es una medida exterior si:*

1. $A \in \mathcal{H}, \mu^*(A) \geq 0$, (no negativa)
2. $A, B \in \mathcal{H}, A \subset B \Rightarrow \mu^*(A) \leq \mu^*(B)$ (monótona)

3.

$$A_n \in \mathcal{H}, n \in \mathbf{N} \Rightarrow \mu^* \left(\bigcup_{n=1}^{\infty} A_n \right) \leq \sum_{n=1}^{\infty} \mu^*(A_n) \quad (\text{subaditiva})$$

4. $\mu^*(\emptyset) = 0$

Pues bien, toda medida en un anillo \mathcal{R} se puede extender a una medida exterior en el σ -anillo hereditario generado por \mathcal{R} de la forma siguiente:

$$\mu^*(A) = \inf \left\{ \sum_{n=1}^{\infty} \mu(A_n) : A_n \in \mathcal{R}, A \subset \bigcup_{n=1}^{\infty} A_n \right\}$$

El problema con las medidas exteriores es que no son σ -aditivas. Para solucionar este problema, definimos lo que es un conjunto μ^* -medible (Caratheorody).

Definición 1.2.5 *Dado un σ -anillo hereditario, \mathcal{H} , se dirá que $A \in \mathcal{H}$ es μ^* -medible si*

$$\mu^*(B) = \mu^*(B \cap A) + \mu^*(B \cap A'), \quad \forall B \in \mathcal{H}$$

Se tiene entonces que la clase de los conjuntos μ^* -medibles ($\bar{\mathcal{S}}$) de un anillo hereditario es un anillo. Si el anillo hereditario de partida es un σ -anillo, también la clase de los conjuntos μ^* -medibles es un σ -anillo. Y lo que resulta es que μ^* es σ -aditiva en el siguiente sentido:

$$\mu^*(A \cap B) = \sum_{n=1}^{\infty} \mu^*(A \cap B_n)$$

donde A es un elemento del σ -anillo hereditario de partida, y B_n es una sucesión de conjuntos disjuntos μ^* -medibles cuya unión es igual a B . La demostración de estas cuestiones es como sigue. Sean B y C dos conjuntos en $\bar{\mathcal{S}}$ y $A \in \mathcal{H}$. Al ser B, C μ^* -medibles se tiene:

$$\begin{aligned} \mu^*(A) &= \mu^*(A \cap B) + \mu^*(A \cap B') \\ &= \mu^*(A \cap B \cap C) + \mu^*(A \cap B \cap C') + \\ &\quad \mu^*(A \cap B' \cap C) + \mu^*(A \cap B' \cap C') \end{aligned}$$

Ahora sustituimos A por $A \cap (B \cup C)$ y se obtiene:

$$\mu^*(A) = \mu^*(A \cap (B \cup C)) + \mu^*(A \cap (B \cup C)')$$

relación que prueba que $B \cup C \in \bar{\mathcal{S}}$. De forma similar se prueba que $B - C \in \bar{\mathcal{S}}$. Como $\emptyset \in \bar{\mathcal{S}}$ se tiene que este conjunto es un anillo. Para probar que es un σ -anillo (en el caso de que el anillo hereditario de partida lo sea) sustituimos B, C por B_1, B_2 en las expresiones anteriores:

$$\mu^*(A \cap (B_1 \cup B_2)) = \mu^*(A \cap B_1) + \mu^*(A \cap B_2)$$

Sea $F_n = \cup_{i=1}^n A_i$. Entonces para todo n :

$$\mu^*(A \cap F_n) = \sum_{i=1}^n \mu^*(A \cap B_i)$$

Podemos de esta forma acotar $\mu^*(A)$:

$$\mu^*(A) \geq \sum_{i=1}^n \mu^*(A \cap B_i) + \mu^*(A \cap B')$$

y pasar (puesto que la desigualdad es cierta para cualquier n) a:

$$\mu^*(A) \geq \sum_{i=1}^{\infty} \mu^*(A \cap B_i) + \mu^*(A \cap B') \geq \mu^*(A \cup B) + \mu^*(A \cap B')$$

Esta desigualdad es suficiente para probar que tenemos un σ -anillo. Además se tiene:

$$\sum_{i=1}^{\infty} \mu^*(A \cap B_i) + \mu^*(A \cap B') \geq \mu^*(A \cup B) + \mu^*(A \cap B')$$

Simplemente hay que sustituir A por $A \cup B$ para obtener el resultado final.

Ahora restringimos la medida exterior a la clase de los conjuntos μ^* -medibles y allí se tiene el siguiente resultado fundamental.

Teorema 1.2.1 *Sea μ^* una medida exterior en un σ -anillo hereditario \mathcal{H} y $\bar{\mathcal{S}}$ la clase de conjuntos μ^* -medibles. Todo conjunto de medida exterior cero es μ^* -medible y la función $\bar{\mu}$ definida como la restricción de μ^* a $\bar{\mathcal{S}}$ es una medida completa en el σ -anillo $\bar{\mathcal{S}}$ (una medida en un anillo es completa si todos los subconjuntos de un conjunto de medida cero están en el anillo). Se llama la medida inducida por la medida exterior.*

Sea A un conjunto de medida exterior cero en \mathcal{H} . Para todo conjunto $B \in \mathcal{H}$ se tiene:

$$\mu^*(B) = \mu^*(A) + \mu^*(B) \geq \mu^*(B \cap A) + \mu^*(B \cap A')$$

lo que implica que A es μ^* -medible (la desigualdad en sentido contrario se da siempre). Hay que demostrar que $\bar{\mu}$ es una medida completa. La σ -aditividad se sigue de la demostración del resultado anterior y la completitud se demuestra como sigue. Sea $A \in \bar{\mathcal{S}}$ y B un subconjunto suyo. Si $\bar{\mu}(A) = 0$ se sigue que su medida exterior es también cero. Por ser $B \subset A$, la medida exterior de B también es cero. Pero hemos visto que esto implica que B es μ^* -medible y por tanto $B \in \bar{\mathcal{S}}$. Podemos ahora establecer un resultado que permite la extensión de una medida definida en un anillo al σ -anillo generado por él.

Proposición 1.2.10 *Si μ es una medida en un anillo \mathcal{R} y μ^* la correspondiente medida exterior en $\mathcal{H}(\mathcal{R})$, entonces todo conjunto del σ -anillo generado por \mathcal{R} , $\mathcal{S}(\mathcal{R})$, es μ^* -medible (es decir, $\mathcal{S}(\mathcal{R}) \subset \bar{\mathcal{S}}$).*

Supongamos que $A \in \mathcal{R}$ y $B \in \mathcal{H}(\mathcal{R})$. Dado $\epsilon > 0$ se tiene una sucesión de conjuntos de \mathcal{R} , A_n , con $B \subset \cup_{n=1}^{\infty} A_n$ tal que:

$$\mu^*(A) + \epsilon \geq \sum_{n=1}^{\infty} \mu(A_n) = \sum_{n=1}^{\infty} (\mu(A_n \cap A) + \mu(A_n \cap A')) \geq \mu^*(A_n \cap A) + \mu^*(A_n \cap A')$$

Tenemos entonces que $A \in \bar{\mathcal{S}}$. Como $\mathcal{S}(\mathcal{R})$ es el menor de los σ -anillos que contienen a \mathcal{R} , se tiene que $\mathcal{S}(\mathcal{R}) \subset \bar{\mathcal{S}}$.

El resultado general, sin introducir medidas exteriores, se resume en los dos teoremas siguientes:

Teorema 1.2.2 *Supongamos que en un anillo \mathcal{R} se tiene definida una medida μ que es σ -finita. Entonces existe una única medida (que también es σ -finita) definida en el σ -anillo $\mathcal{S}(\mathcal{R})$ que extiende a μ .*

Sin embargo, la medida extendida no tiene por qué ser completa. La extensión a una medida completa se puede hacer en base al siguiente resultado.

Teorema 1.2.3 *Dada una medida en un σ -anillo \mathcal{S} , la clase de los conjuntos de la forma*

$$\bar{\mathcal{S}} = \{ A \Delta N : A \in \mathcal{S}, N \subset E \in \mathcal{S}, \mu(E) = 0 \}$$

es también un σ -anillo y la función definida por:

$$\bar{\mu}(A \Delta N) = \mu(A)$$

es una medida completa en la clase extendida, que se llama la completación de μ .

Como usando medidas exteriores también hemos obtenido completaciones de medidas, necesitamos un resultado que las relacione.

Teorema 1.2.4 *Si μ es una medida σ -finita en un anillo y μ^* es la medida exterior asociada, la completación de la extensión de μ (en el sentido del teorema anterior) es igual a μ^* en la clase de los conjuntos μ^* -medibles.*

1.2.4. Medidas interiores

Sea \mathcal{S} un σ -anillo y $\mathcal{H}(\mathcal{S})$ el σ -anillo hereditario que genera.

Definición 1.2.6 *Si $A \in \mathcal{H}(\mathcal{S})$, se define la medida interior $\mu_*(A)$ por:*

$$\mu_*(A) = \sup\{ \mu(B) : A \supset B \in \mathcal{S} \}$$

Se pueden estudiar propiedades similares a las de las medidas exteriores. Y se llega a un resultado fundamental:

Teorema 1.2.5 *Sea \mathcal{S} un σ -anillo y $\mathcal{H}(\mathcal{S})$ el σ -anillo hereditario construido sobre él. Sea $\bar{\mathcal{S}}$ el anillo construido en el teorema 1.2.3. Entonces,*

$$\mu^*(A) = \mu_*(A) = \bar{\mu}(A), \quad \forall A \in \bar{\mathcal{S}}$$

Además,

$$A \in \mathcal{H}(\mathcal{S}), \mu^*(A) = \mu_*(A) < \infty \Rightarrow A \in \bar{\mathcal{S}}$$

Y se puede llegar al teorema de extensión anterior definiendo los conjuntos μ^* -medibles como aquellos para los que la medida exterior e interior coinciden.

1.2.5. La medida de Lebesgue

Introduciremos ahora la medida de Lebesgue. Ya hemos definido la medida λ y demostrado algunas de sus propiedades más importantes.

Dada la clase \mathcal{P} se puede definir el σ -anillo generado por ella, \mathcal{B} , que es realmente una σ -álgebra, pues \mathbf{R} es una unión numerable de intervalos de la clase \mathcal{P} . Los elementos de \mathcal{B} se llaman **conjuntos de Borel**. Recuérdese que en el apartado 1.2.1 habíamos introducido la clase \mathcal{R} que era solo un anillo.

Como consecuencia de los resultados vistos hasta ahora, se tiene que la función λ se puede extender a \mathcal{B} . Llamaremos también λ a la medida extendida. De acuerdo con las propiedades de medidas exteriores, se construye ahora el conjunto $\bar{\mathcal{B}}$, formado por los conjuntos que son diferencias simétricas de elementos de \mathcal{B} y subconjuntos de conjuntos de medida nula en \mathcal{B} . Como sabemos, este conjunto es un σ -anillo (en realidad una σ -álgebra) y en él se puede definir la completación de λ (que seguiremos llamando λ). Ésta es la **medida de Lebesgue** en la recta real. Los conjuntos de $\bar{\mathcal{B}}$ se llaman conjuntos medibles Lebesgue. La medida λ en \mathcal{B} no es finita, pues $\lambda(\mathbf{R}) = \infty$, pero sí lo es en \mathcal{P} . Por lo tanto, las correspondientes medidas en \mathcal{B} y $\bar{\mathcal{B}}$ son σ -finitas. En realidad son totalmente σ -finitas pues la recta \mathbf{R} también es σ -finita.

1.2.6. Propiedades de la medida de Lebesgue

En esta sección estudiaremos algunas propiedades de la medida de Lebesgue.

Proposición 1.2.11 *Todo conjunto numerable es un conjunto de Borel de medida nula.*

Demostraremos que la medida de un conjunto que contiene un solo punto es cero. La proposición se sigue inmediatamente debido a que la medida es σ -aditiva. Para ello basta considerar la siguiente sucesión de intervalos:

$$I_n = \left[a, a + \frac{1}{n} \right)$$

Está claro que su intersección (es decir el límite) es $\{a\}$. Además:

$$\lambda(I_n) = \frac{1}{n}$$

y la conclusión es inmediata.

A pesar de lo curioso de la introducción de los conjuntos de Borel, resulta que el álgebra de Borel coincide con el σ -anillo generado por los abiertos de \mathbf{R} . Para demostrarlo, se prueba primero que todo intervalo abierto es un conjunto de Borel. Pero esto es sencillo, pues basta restar de un intervalo semicerrado (que es de Borel) el extremo inferior del intervalo (que también es de Borel). Como todo abierto de \mathbf{R} es una unión numerable de intervalos abiertos, se concluye que el σ -anillo generado por los abiertos está contenido en el de Borel. Para demostrar la inclusión opuesta, basta escribir el conjunto con un solo punto como una intersección numerable de intervalos abiertos:

$$\{a\} = \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, a + \frac{1}{n} \right)$$

Esto hace que estos conjuntos de un solo punto estén en el σ -anillo generado por los abiertos. Como los intervalos semicerrados son uniones de intervalos abiertos con un conjunto formado por un solo punto, se tiene que están en el σ -anillo generado por los abiertos y por lo tanto también está el σ -anillo de Borel.

La medida exterior usada en la construcción de la medida de Borel se puede definir ahora como:

$$\lambda^*(A) = \inf \{ \lambda(B) : A \subset B, B \text{ abierto} \}$$

La medida de Lebesgue es invariante bajo traslaciones y se comporta como era de esperar bajo dilataciones.

Como hemos visto, la medida de un conjunto numerable es igual a cero. Podría pensarse que todo conjunto de medida cero es numerable, pero no es el caso.

El conjunto de Cantor

Consideremos el intervalo de la recta real $[0, 1]$. Cualquier número de este intervalo se puede escribir como:

$$x = \sum_{n=1}^{\infty} \frac{x_n}{3^n}$$

donde los coeficientes x_n valen 0, 1 o 2 (es la expresión en base 3 de ese número). Tomemos ahora los números que no tienen ningún coeficiente x_n igual a 1 y llamemos \mathcal{C} a ese conjunto (hay que prestar atención a como se desarrollan los números en base 3. Por ejemplo, si se escriben como en la notación decimal:

$$x = \sum_{n=1}^{\infty} \frac{x_n}{3^n} = 0, x_1 x_2 x_3 \dots$$

el número $1/3$ es claramente:

$$\frac{1}{3} = 0,100\dots$$

Pero también es igual a

$$\frac{1}{3} = 0,0222\dots = 2 \sum_{n=2}^{\infty} \frac{1}{3^n} = 2 \frac{1/9}{1 - 1/3}$$

Así que $1/3$ está en \mathcal{C} .

Desde un punto de vista geométrico, consideremos ahora el intervalo $I = [0, 1]$ que dividimos en tres partes. Eliminamos la parte central, es decir nos quedamos con:

$$I - I_1, \quad I_1 = \left(\frac{1}{3}, \frac{2}{3}\right)$$

En cada una de estas dos partes hacemos lo mismo, dividimos en tres partes y eliminamos la parte central

$$I_2 = \left(\frac{1}{9}, \frac{2}{9}\right), \quad I_3 = \left(\frac{7}{9}, \frac{8}{9}\right)$$

El conjunto que queda ahora es:

$$I - (I_1 \cup I_2 \cup I_3)$$

Se repite el proceso indefinidamente. Pues bien el conjunto \mathcal{C} coincide con el definido anteriormente:

$$\mathcal{C} = I - \bigcup_{n=1}^{\infty} I_n$$

Para demostrarlo, consideremos los desarrollos ternarios (si hay más de una posibilidad se elige según el criterio anterior). Sea $x \in I$. Si en su desarrollo ternario, $x_1 = 1$, entonces $x \in I_1$ (y viceversa). Pues:

$$x = \frac{1}{3} + \dots > \frac{1}{3}$$

lo que es evidente. Pero además:

$$\frac{1}{3} + \dots < \frac{2}{3}$$

pues en el caso más desfavorable se tiene el número:

$$\frac{1}{3} + \sum_{n=2}^{\infty} \frac{2}{3^n} = \frac{1}{3} + \frac{2/9}{1 - 1/3} = 2/3 = 0,2000\dots$$

que no es de los que estamos considerando. De forma similar se prueba que si $x_1 \neq 1$ y $x_2 = 1$ entonces $x \in I_2 \cup I_3$ (y viceversa). Así se prueba que ambos conjuntos coinciden. Ocorre que \mathcal{C} es un conjunto medible y que su medida es cero. El que sea medible proviene de la manera de obtenerlo, restando de un conjunto medible otro, la unión de una familia numerable de conjuntos medibles. El que la medida sea cero es también fácil de obtener (los I_n son disjuntos, y en el paso n hay 2^{n-1} intervalos y cada uno de ellos mide $1/3^n$):

$$\lambda\left(\bigcup_{n=1}^{\infty} I_n\right) = \sum_{n=1}^{\infty} \lambda(I_n) = \frac{1}{3} + 2\frac{1}{9} + 4\frac{1}{27} + \dots = \sum_{n=1}^{\infty} \frac{2^{n-1}}{3^n} = 1$$

Como la medida de $[0, 1]$ es también 1, se concluye que la medida del conjunto de Cantor es 0. Pero el conjunto de Cantor no es numerable. Tiene el mismo número de elementos que el conjunto de los números reales (una forma de hacerlo es establecer una correspondencia entre los irracionales del conjunto de Cantor y los irracionales del intervalo $[0, 1]$, mediante el paso de la expresión ternaria a una binaria). El cardinal de la clase de conjuntos de Borel es el continuo. Pero todos los subconjuntos del conjunto de Cantor son medibles (y de medida cero). Por tanto, hay conjuntos medibles Lebesgue que no son de Borel.

Pero también se puede demostrar que hay conjuntos en la recta real que no son medibles Lebesgue. Para ello, consideremos (Vitali) el intervalo $[0, 1]$ y la relación de equivalencia $x \sim y$ si su diferencia es un

número racional. Usando el axioma de elección, se puede elegir un representante en cada clase y formar un conjunto V . Se tiene que este conjunto no es medible. Para demostrarlo, supongamos que V es medible Lebesgue y consideremos el conjunto de los números racionales en $[-1, 1]$, $\{r_1, r_2, \dots\}$. Definamos los conjuntos

$$V_n = \{r_n + x : x \in V\}$$

que son medibles pues son traslaciones de V . Los conjuntos V_n son disjuntos y miden todos lo mismo que V . Finalmente V verifica:

$$[0, 1] \subset \bigcup_{n=1}^{\infty} V_n \subset [-1, 2]$$

Apliquemos ahora que λ es una medida:

$$\lambda\left(\bigcup_{n=1}^{\infty} V_n\right) = \sum_{n=1}^{\infty} \lambda(V_n) = \lim_{n \rightarrow \infty} (n\lambda(V)) \leq \lambda([-1, 2]) = 3$$

La conclusión es que $\lambda(V) = 0$. Pero

$$\lambda\left(\bigcup_{n=1}^{\infty} V_n\right) \geq 1$$

lo que es una contradicción.

Es posible probar que no se puede extender la medida de Lebesgue a la clase de todos los conjuntos de la recta real de forma que se obtenga una medida invariante bajo traslaciones.

La medida de Lebesgue se puede generalizar en el siguiente sentido. Sea g una función finita, creciente y continua; existe una única medida completa definida en un σ -anillo $\bar{\mathcal{S}}_g$ (que contiene a los conjuntos de Borel), tal que:

$$\bar{\mu}(g([a, b])) = g(b) - g(a)$$

Se le llama la medida de *Lebesgue-Stieltjes* asociada a g . Para cada elemento de $\bar{\mathcal{S}}_g$ existe un boreliano cuya diferencia simétrica con el primer conjunto tiene medida de Lebesgue-Stieltjes igual a cero.

1.3. Funciones medibles

Diremos que una cierta propiedad se verifica casi doquiera si es cierta salvo en un conjunto medible de medida nula. Un espacio de medida es un conjunto X , con un σ -anillo \mathcal{S} y una medida μ definida en él. Exigiremos que la unión de todos los conjuntos de \mathcal{S} sea el espacio total (todo punto está en algún conjunto de \mathcal{S}) y nos limitaremos al caso en que \mathcal{S} es una σ -álgebra. Se suele decir que un subconjunto de X es medible si pertenece a \mathcal{S} (como hemos usado anteriormente). Sea f una función definida en un espacio de medida con valores en la recta real:

$$f: X \rightarrow \mathbf{R}$$

Definición 1.3.1 *Se dice que una función f es medible si dado B , conjunto de Borel de la recta real, el conjunto $f^{-1}(B)$ es un conjunto medible de X*

Recuérdese que los conjuntos de Borel son los de la σ -álgebra generada por los conjuntos abiertos.

Los mismos conceptos se pueden aplicar también a la medida Lebesgue (y se habla de funciones medibles Borel o medibles Lebesgue).

Además de funciones con valores reales, es necesario considerar funciones con valores en la recta real extendida. Basta considerar que los conjuntos $\{-\infty\}$ y $\{+\infty\}$ son borelianos. Una función es entonces medible si se cumple la definición anterior y los conjuntos $f^{-1}(\{+\infty\})$ y $f^{-1}(\{-\infty\})$ son medibles.

Para que una función sea medible, basta aplicar la definición a borelianos del tipo $(-\infty, c)$.

Un ejemplo sencillo de funciones medibles son las funciones características de conjuntos medibles. Si el conjunto X es medible, las funciones constantes son medibles.

No es difícil probar que el valor absoluto de una función medible es medible (y cualquier potencia del valor absoluto también lo es). Además la suma y el producto de funciones medibles son medibles. También son medibles la parte positiva y negativa de una función:

$$f^+ = \frac{1}{2}(|f| + f), \quad f^- = \frac{1}{2}(|f| - f)$$

El concepto de sucesiones de funciones medibles será fundamental en el estudio de la integral.

Teorema 1.3.1 *Dada una sucesión de funciones medibles, f_n , y una función f tal que $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ casi doquiera, entonces f es medible.*

Una generalización de las funciones características de conjuntos son las funciones simples. Sea X un espacio de medida. Se dice que la función $s(x)$ es una función simple si existe una colección finita de conjuntos medibles y disjuntos $\{A_1, \dots, A_n\}$ y de números reales $\{c_1, \dots, c_n\}$ tal que:

$$s(x) = \begin{cases} c_i, & x \in A_i \\ 0, & x \notin A_1 \cup \dots \cup A_n \end{cases}$$

Las funciones simples son medibles y se pueden escribir como combinaciones lineales de funciones características:

$$s(x) = \sum_{i=1}^n c_i \chi_{A_i}(x)$$

La importancia de las funciones simples viene dada por el siguiente resultado.

Teorema 1.3.2 *Sea f una función medible no negativa en un espacio de medida. Entonces, existe una sucesión creciente de funciones simples no negativas que tienden a f .*

La demostración es como sigue: definimos para cada n el conjunto.

$$A_n^i = \left\{ x \in X : \frac{i-1}{2^n} \leq f(x) < \frac{i}{2^n} \right\}, \quad i = 1, \dots, n2^n$$

Los conjuntos A_n^i son disjuntos si $i \neq j$ y son medibles, al serlo f . Se definen las funciones simples:

$$s_n(x) = \sum_{i=1}^{n2^n} \frac{i-1}{2^n} \chi_{A_n^i}$$

Se trata de una sucesión de funciones simples no negativas crecientes:

$$0 \leq s_n(x) \leq s_{n+1}(x) \leq f(x)$$

La convergencia en cada x es consecuencia de la desigualdad (x fijo):

$$0 \leq f(x) - s_n(x) \leq \frac{1}{2^n}$$

cuando n es suficientemente grande.

1.4. Integración

El último concepto que se introducirá en este capítulo es el de integral de una función. La definición es muy sencilla después de todo lo anterior, y a continuación estudiaremos algunas de sus propiedades en especial las referidas a su comportamiento con las sucesiones de funciones.

1.4.1. Conceptos generales

Definición 1.4.1 Se dice que una función simple es una función escalón si los conjuntos A_n que definen a la función simple son de medida finita.

Definimos primero lo que es la integral de una función escalón.

Definición 1.4.2 Sea (X, \mathcal{S}, μ) un espacio de medida y

$$\varphi(x) = \sum_{i=1}^n c_i \chi_{A_i}(x)$$

una función escalón. Se define la integral de s sobre el espacio X como:

$$\int_X \varphi \, d\mu = \sum_{i=1}^n c_i \mu(A_i)$$

Una función escalón se puede escribir de diversas formas en términos de funciones características, pero el valor de la integral no depende de esas descomposiciones. A veces interesa definir la integral no sobre todo el espacio X sino sobre una parte de él, $A \subset X$:

$$\int_A \varphi \, d\mu = \int_X \chi_A \varphi \, d\mu$$

La integral así definida verifica una serie de propiedades, que resultan naturales después de haberlas estudiado en la integral de Riemann.

Proposición 1.4.1 Todas las funciones que aparecen a continuación son funciones escalón.

1. La integral es lineal

$$\int (a\varphi + b\psi) \, d\mu = a \int \varphi \, d\mu + b \int \psi \, d\mu$$

2. Si φ es no negativa, la integral también lo es.

3.

$$\int |\varphi + \psi| \, d\mu \leq \int |\varphi| \, d\mu + \int |\psi| \, d\mu$$

4.

$$\left| \int \varphi \, d\mu \right| \leq \int |\varphi| \, d\mu$$

5. Si A es un conjunto medible y φ una función escalón, tal que $a \leq \varphi(x) \leq b$ para $x \in A$, se tiene:

$$a\mu(A) \leq \int_A \varphi \, d\mu \leq b\mu(A)$$

Es posible definir la integral indefinida de una función escalón. Si A es un conjunto medible:

$$F(A) = \int_A \varphi \, d\mu$$

La integral indefinida de una función escalón es absolutamente continua (una función de conjunto, g es absolutamente continua, con respecto a una medida μ , si para todo $\epsilon > 0$ existe $\delta > 0$, tal que para todo conjunto medible A , con $\mu(A) < \delta$, entonces $g(A) < \epsilon$).

El objetivo es definir la integral para funciones más generales que las funciones simples. Para ello estudiaremos las sucesiones de funciones escalón.

Proposición 1.4.2 Sea φ_n una sucesión decreciente de funciones escalón que tiende a 0 (casi doquiera). Entonces la sucesión de las integrales de φ_n es decreciente y tiende a 0. Y por lo tanto, si una sucesión creciente de funciones escalón tiende a una función escalón φ , también la sucesión de las integrales es creciente y tiende a la integral de esa función.

1.4.2. La integral de Lebesgue

Definición 1.4.3 Una función se llama superior si existe una sucesión creciente de funciones escalón $\varphi_n(x)$, con integral finita, que converge a ella (casi doquiera).

Llamaremos sucesión generadora de una función superior u a la sucesión de funciones escalón $\{\varphi_n\}$. Recordemos que una función escalón se escribe como:

$$\varphi(x) = \sum_{i=1}^n c_i \chi_{A_i}$$

donde a_1, \dots, a_n es una colección de números reales y los conjuntos A_i son disjuntos y medibles, de medida finita.

Una función superior tiene distintas sucesiones generadoras. Pero el límite $\int \varphi_n d\mu$ (que existe) es el mismo para todas ellas. Debido a esto, se puede definir:

Definición 1.4.4 La integral de una función superior u , con una sucesión generadora φ_n se define como

$$\int u d\mu = \lim_{n \rightarrow \infty} \int \varphi_n d\mu$$

La integral así definida es lineal:

$$\int (u + v) d\mu = \int u d\mu + \int v d\mu$$

pero solo es homogénea cuando el coeficiente es no negativo:

Proposición 1.4.3 Si u es una función superior y $a \in \mathbf{R}$, au es una función superior si y solo si $a \geq 0$. Entonces:

$$\int af d\mu = a \int f d\mu, \quad a \geq 0$$

El problema con $a < 0$ es que la sucesión pasa de ser creciente a ser decreciente, con lo que no se cumple la definición de función superior. Se tiene también que el máximo y el mínimo de dos funciones superiores es una función superior.

La integral de funciones superiores es monótona:

$$u \geq v, \text{ c.d.} \Rightarrow \int u d\mu \geq \int v d\mu$$

Para demostrarlo basta considerar dos sucesiones generadoras φ_n y ϕ_n de u y g respectivamente. Entonces, al ser g menor o igual que f (casi doquiera, por supuesto), la sucesión de funciones escalón $\min\{\varphi_n, \phi_n\}$ es también una sucesión generadora de v . Pero

$$\varphi_n \geq \min\{\varphi_n, \phi_n\}$$

de donde se concluye el resultado. En particular la integral de una función superior mayor o igual a cero es también mayor o igual a cero.

Las sucesiones crecientes de funciones superiores (con integrales finitas) son funciones superiores.

Sin embargo, las funciones superiores no forman un espacio vectorial. El concepto de integral se extiende a funciones más generales (que sí forman un espacio vectorial).

Definición 1.4.5 Se dice que una función es integrable si es la diferencia de dos funciones superiores (c.d.). La integral se define como:

$$\int f d\mu = \int u d\mu - \int v d\mu$$

La integral de Lebesgue-Stieltjes se define como la integral de Lebesgue, pero asociada a la medida de Lebesgue-Stieltjes.

El valor de la integral es independiente de las funciones superiores que se usen para calcularla. Una función integrable es medible (las funciones superiores son medibles como límites de funciones escalón).

Las funciones integrables forman un espacio vectorial. Si una función es integrable, su valor absoluto es integrable, así como sus partes positiva y negativa. En particular una función es integrable si y solo si lo son sus partes positiva y negativa.

La integral es finita. Al igual que con las funciones escalón, la suma de funciones integrables es integrable, el valor absoluto y la parte positiva y negativa de una función son integrables. Se puede definir también la integral sobre un conjunto tal y como se ha hecho en el caso de la integral de funciones escalón.

Proposición 1.4.4 *La integral es lineal:*

$$\int (af + bg) d\mu = a \int f d\mu + b \int g d\mu$$

Se tiene el siguiente resultado sobre funciones positivas (que se puede usar para introducir la integral):

Teorema 1.4.1 *Si una función integrable es no negativa (c.d.), entonces es una función superior.*

Como f es integrable, existen dos funciones superiores tales que $f = u - v$. La diferencia de las sucesiones generadoras de u y v tiende a f (aunque no sea en principio un función generadora de f). Como F es no negativa, tenemos una sucesión de funciones escalón no negativas que tiende a f :

$$\varphi^+ \rightarrow f, \quad \text{c.d.}$$

Un resultado previo 1.3.2 asegura que toda función medible es el límite de una sucesión creciente de funciones simples. En nuestro caso:

$$0 \leq s_n \uparrow f, \quad \text{c.d.}$$

Tomemos la sucesión

$$\phi_n = \min\{s_n, \max\{\varphi_i^+, i = \dots, n\}\}$$

Ahora solo queda probar que las integrales $\int \phi_n d\mu$ están acotadas. Se tiene la siguiente cadena de desigualdades (usando la monotonía de la integral):

$$\begin{aligned} \phi_n + v &\leq f + v = u \\ \int \phi_n d\mu + \int v d\mu &\leq \int u d\mu \\ \int \phi_n d\mu &\leq \int u d\mu - \int v d\mu < \infty \end{aligned}$$

Proposición 1.4.5 *La integral tiene las siguientes propiedades:*

1. *Dada una función integrable no negativa casi doquiera, la integral es cero si y solo si la función es igual a cero (c.d.)*
2. *La integral de una función integrable sobre un conjunto de medida cero es cero.*
3. *Si una función integrable y positiva c.d. tiene integral cero sobre un conjunto, la medida de este conjunto es cero.*
4. *Si una función tiene integral cero sobre cualquier conjunto medible, entonces es cero c.d.*
5. *Si una función medible está acotada entre dos funciones integrables (c.d.) es integrable.*
6. *Sean f y g integrables. Entonces, $f \geq g \Rightarrow \int f d\mu \geq \int g d\mu$*

7. Si f es integrable se tiene: $|\int f \, d\mu| \leq \int |f| \, d\mu$
8. Sean f medible y g integrable. Si $|f| \leq |g|$ c.d., entonces f es integrable.
9. Una función medible es integrable si y solo si su valor absoluto es integrable.
10. Si f es esencialmente acotada (acotada c.d.) y el conjunto A es medible, entonces f es integrable sobre A .

Si f es integrable, su parte positiva y negativa son funciones superiores y por tanto para calcular la integral se pueden usar como una descomposición de f en una diferencia de funciones superiores:

$$\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu$$

Veamos a continuación algunos resultados concernientes al comportamiento de la integral frente a sucesiones. Nos interesa especialmente el comportamiento de las sucesiones de funciones integrables. Como se recordará ésta fue una de las razones para justificar la introducción de la integral de Lebesgue.

Teorema 1.4.2 Sea $\{f_n\}$ una sucesión creciente de funciones integrables, tal que $\lim \int f_n \, d\mu < \infty$. Entonces, existe una función integrable que es el límite de la sucesión $f_n \uparrow f$ y se tiene:

$$\int f_n \, d\mu \uparrow \int f \, d\mu$$

Teorema 1.4.3 (Lema de Fatou) Sea f_n una sucesión de funciones integrables no negativas (como siempre, c.d.) y tal que $\liminf_n \int f_n \, d\mu < \infty$. Entonces, $\liminf f_n$ es una función integrable y

$$\int \liminf_n f_n \, d\mu \leq \liminf_n \int f_n \, d\mu$$

El límite inferior de una sucesión a_n es el límite de los ínfimos de los conjuntos $\{a_n, a_{n+1}, \dots\}$. El resultado más importante es el teorema de convergencia dominada de Lebesgue:

Teorema 1.4.4 Sea $\{f_n\}$ una sucesión de funciones integrables que verifican $|f_n| \leq g$, $n = 1, 2, \dots$, casi doquiera, siendo g una función integrable. Si $f_n \rightarrow f$ c.d., entonces f es integrable y se tiene:

$$\lim \int f_n \, d\mu = \int \lim f_n \, d\mu = \int f \, d\mu$$

Demostración. Puesto que todas las funciones están acotadas por g , el límite también lo está. Pero g es integrable, luego $|f|$ y por tanto f , también lo es. Veamos ahora la cuestión del límite. A la sucesión $g - f_n$ se le aplica el lema de Fatou (pues es positiva y $\liminf_n \int (g - f_n) \, d\mu < \infty$). Además

$$\liminf_n (g - f_n) = g - f, \quad \text{c.d.}$$

Por tanto:

$$\begin{aligned} \int g \, d\mu - \int f \, d\mu &= \int (g - f) \, d\mu = \int \liminf_n (g - f_n) \, d\mu \\ &\leq \liminf_n \int (g - f_n) \, d\mu = \int g \, d\mu - \limsup_n \int f_n \, d\mu \end{aligned}$$

es decir

$$\limsup_n \int f_n \, d\mu \leq \int f \, d\mu$$

Los mismos razonamientos se aplican a la sucesión $g + f_n$:

$$\begin{aligned} \int g \, d\mu + \int f \, d\mu &= \int (g + f) \, d\mu = \int \liminf_n (g + f_n) \, d\mu \\ &\leq \liminf_n \int (g + f_n) \, d\mu = \int g \, d\mu + \liminf_n \int f_n \, d\mu \end{aligned}$$

y se tiene:

$$\int f \, d\mu \leq \liminf_n \int f_n \, d\mu$$

Pero eso implica que los límites inferior y superior son iguales y por tanto el límite existe y es el que afirma el teorema.

Todas las integrales definidas hasta ahora son finitas. Pero la integral de una función simple con coeficientes positivos se puede definir, aunque posiblemente sea infinita (depende de la medida de los conjuntos sobre los que no es cero). Se puede decir que estas funciones tienen una integral que es igual a infinito. Entonces, si una función es el límite de una sucesión creciente de funciones simples positivas, el límite de las integrales de esa sucesión existe aunque posiblemente sea infinito. Se dice entonces que la integral de la función límite es infinito (aunque no es usual llamarla integrable). De acuerdo con esta idea, toda función medible positiva tiene una integral (finita o no).

1.4.3. La integral de Riemann

La integral de Riemann se construye introduciendo unas particiones del intervalo (cerrado y acotado) donde se define la función (que es acotada). A continuación se definen las sumas superiores (e inferiores), sumando los productos del valor máximo (mínimo para la suma inferior) de la función en cada subintervalo de la partición por la longitud del intervalo. Si la partición se refina (se introducen más puntos), las sumas superiores no crecen y las inferiores no decrecen. Para dos particiones arbitrarias, la suma inferior de una de ellas es menor o igual que la suma superior de la otra. De esta forma se construye la integral de Riemann inferior como el supremo de las sumas inferiores sobre todas las particiones y la integral superior como el ínfimo de las sumas superiores sobre todas las particiones. La integral inferior siempre es menor o igual que la superior. Con estos resultados se puede hacer la siguiente definición.

Definición 1.4.6 Una función f acotada en un intervalo $[a, b]$ es integrable (Riemann) si la integral superior e inferior, coinciden. Al valor común se le llama integral de Riemann de esa función.

El valor de la integral se denota por:

$$\int_a^b f(x) \, dx$$

El criterio de Riemann proporciona una caracterización de las funciones integrables en un intervalo.

Proposición 1.4.6 (Riemann) Una función acotada en un intervalo $[a, b]$ es integrable (Riemann) si para todo $\epsilon > 0$ existe una partición del intervalo $[a, b]$ tal que la diferencia entre la suma superior e inferior asociada a ese intervalo es menor que ϵ .

Para una función integrable (Riemann), si en las sumas superiores e inferiores se toma un punto del subintervalo y el valor de la función en ese punto (y no el supremo o el ínfimo de la función en el subintervalo), se tiene que el límite de la suma de los valores de la función en ese punto por las longitudes de los subintervalos tiende a la integral Riemann cuando la longitud del mayor de los subintervalos tiende a cero.

Pues bien, las funciones integrables Riemann son integrables Lebesgue.

Teorema 1.4.5 Una función definida en un intervalo $[a, b]$ integrable Riemann es integrable Lebesgue y los valores de las integrales coinciden:

$$\int f \, d\lambda = \int_a^b f(x) \, dx$$

Consideremos una partición del intervalo $[a, b]$ en 2^n subintervalos de la misma longitud:

$$P_n = \{x_0, x_1, \dots, x_n\}, \quad x_i = a + i(b-a)2^{-n}$$

Definimos dos familias de funciones escalón por:

$$\varphi_n = \sum_{i=1}^{2^n} m_i \chi_{[x_{i-1}, x_i]}, \quad \phi_n = \sum_{i=1}^{2^n} M_i \chi_{[x_{i-1}, x_i]}$$

donde

$$m_i = \inf\{f(x) : x \in [x_{i-1}, x_i]\}, \quad M_i = \sup\{f(x) : x \in [x_{i-1}, x_i]\}$$

De esta forma φ_n (ϕ_n) es una sucesión creciente (decreciente) de funciones escalón que acotan a la función f :

$$\varphi_n(x) \leq f(x) \leq \phi_n(x), \quad x \in [a, b]$$

Supongamos que $\varphi_n(x) \uparrow g$ y $\phi_n(x) \downarrow h$. Entonces g y h son integrables Lebesgue y acotan a f :

$$g(x) \leq f(x) \leq h(x), \quad x \in [a, b]$$

Pero tal y como se han definido las sucesiones de funciones escalón está claro que coinciden con las sumas inferior y superior para la partición considerada. Se tiene:

$$\begin{aligned} 0 &\leq \int (h - g) \, d\lambda = \lim \int (\phi_n - \varphi_n) \, d\lambda = \lim \int \phi_n \, d\lambda - \lim \int \varphi_n \, d\lambda = \\ &\lim S^*(f, P_n) - \lim S_*(f, P_n) = 0 \end{aligned}$$

donde $S^*(f, P_n)$ ($S_*(f, P_n)$) es la suma superior (inferior) de la función f en la partición considerada. Pero, por las propiedades de la integral Lebesgue, esto implica que $h = g = f$ casi doquiera, así que f es integrable Lebesgue (es una función superior). Además:

$$\int f \, d\lambda = \lim \int \varphi_n \, d\lambda = \lim S_*(f, P_n) = \int_a^b f(x) \, dx$$

Lebesgue dio un criterio que permite establecer bajo que condiciones una función acotada es integrable (Riemann) en un intervalo:

Teorema 1.4.6 *Una función acotada en el intervalo $[a, b]$ es integrable Riemann si y solo si es una función continua casi doquiera (es decir, salvo en un conjunto de medida nula).*

Las funciones continuas son entonces integrables (Riemann) y su cálculo se hace mediante el teorema fundamental del cálculo infinitesimal.

1.4.4. Integrales impropias

Para una función definida en un intervalo no acotado, es posible en ocasiones definir la integral Riemann como:

$$\int_a^\infty f(x) \, dx = \lim_{b \rightarrow \infty} \int_a^b f(x) \, dx$$

cuando este límite existe. Se habla de una integral impropia.

Si f (definida en $[a, \infty)$) es integrable en todo subintervalo cerrado y acotado de $[a, \infty)$, la integral (impropia) en $[a, \infty)$ existe si para todo $\epsilon > 0$ existe $M > 0$ tal que

$$\left| \int_{a_1}^{b_1} f(x) \, dx \right| < \epsilon, \quad a_1, b_1 \geq M$$

Es fácil ver entonces que si f es integrable en todo subintervalo cerrado y acotado de $[a, \infty)$, y la integral

$$\int_a^\infty |f(x)| \, dx$$

existe, entonces también existe $\int_a^\infty f(x) \, dx$.

El resultado inverso es falso (hay funciones que son integrables Riemann, en sentido impropio, pero su valor absoluto no lo es). En el caso de funciones integrables Lebesgue se tiene:

Teorema 1.4.7 *Supongamos que una función definida en el intervalo $[a, \infty)$ es integrable Riemann en todo subintervalo cerrado. Entonces, f es integrable Lebesgue si y solo si la integral de Riemann (impropia) $\int |f(x)| dx$ existe y se tiene en este caso:*

$$\int f d\lambda = \int_a^\infty f(x) dx$$

El comportamiento de la integral frente a límites y derivaciones (paramétricas) se refleja en los dos siguientes resultados.

Teorema 1.4.8 *Sea (X, \mathcal{S}, μ) un espacio de medida e I un intervalo de \mathbf{R} . Consideremos una función $f : X \times I \rightarrow \mathbf{R}$, tal que $f(x, t)$ es una función medible en su primera variable para todo $t \in I$. Supongamos que existe una función g , integrable, que acota a f : $|f(x, t)| \leq g(x)$ para casi todo x y para todo t y sea t_0 un punto en el cierre de I , que puede ser igual a $\pm\infty$. Entonces, si para algún punto t_0 existe el límite de $f(x, t)$ cuando $t \rightarrow t_0$, siendo $h(x) = \lim_{t \rightarrow t_0} f(x, t)$ para casi todo x , se tiene que h es integrable y además:*

$$\lim_{t \rightarrow t_0} \int f(x, t) d\mu(x) = \int \lim_{t \rightarrow t_0} f(x, t) d\mu(x)$$

Teorema 1.4.9 *Sea (X, \mathcal{S}, μ) un espacio de medida e I un intervalo abierto acotado de \mathbf{R} . Consideremos una función $f : X \times I \rightarrow \mathbf{R}$, tal que $f(x, t)$ es una función integrable Lebesgue en su primera variable para todo $t \in I$. Supongamos que para algún $t_0 \in I$ existe la derivada parcial $\partial_t f(x, t_0)$, para casi todo x . Supongamos que existe una función integrable g , y un entorno de t_0 , U , tal que se da la siguiente acotación:*

$$\left| \frac{f(x, t) - f(x, t_0)}{t - t_0} \right| \leq g(x)$$

para casi todo x y para todo $t \in U$. En este caso, $\partial_t f(t_0, x)$ es una función integrable en x y la función $F : I \rightarrow \mathbf{R}$, dada por la integral:

$$F(t) = \int f(x, t) d\mu(x)$$

es diferenciable en t_0 y su derivada es:

$$F'(t_0) = \int \frac{\partial f}{\partial t}(x, t_0) d\mu(x)$$

Con estos resultados es posible evaluar la integral de ciertas integrales impropias.

Ejemplo 1.3 (Euler) Queremos calcular (si existe) el valor de la integral

$$\int_0^\infty e^{-x^2} dx$$

La integral de Riemann existe por la acotación $0 \leq e^{-x^2} \leq e^{-x}$, teniendo en cuenta que esta última función es integrable en la semirrecta positiva. La integral es también una integral Lebesgue, pues la función e^{-x^2} es positiva en toda la semirrecta. Consideremos ahora dos funciones definidas por:

$$f(t) = \left(\int_0^t e^{-x^2} dx \right)^2, \quad g(t) = \int_0^1 \frac{e^{-t^2(x^2+1)}}{x^2+1} dx$$

La derivada de la primera función no ofrece dificultades. La dependencia en t está en el extremo superior de la integral y por tanto (teorema fundamental del cálculo):

$$f'(t) = 2e^{-t^2} \int_0^t e^{-x^2} dx, \quad t > 0$$

Para la otra función, $g(t)$, se puede ver que se satisfacen las condiciones para poder aplicar la derivación parcial bajo el signo integral y se tiene:

$$g'(t) = -2e^{-t^2} \int_0^1 te^{-x^2 t^2} = -2e^{-t^2} \int_0^t e^{-x^2}, \quad t > 0$$

Por lo tanto, $f'(t) = -g'(t)$, luego $f(t)$ y $-g(t)$ difieren en una constante para $t \geq 0$ (pues son continuas). Pero la constante se puede calcular muy fácilmente:

$$f(t) = -g(t) + \frac{\pi}{4}$$

Usando las propiedades de cálculo del límite bajo el signo integral, se tiene:

$$\frac{\pi}{4} = \lim_{t \rightarrow \infty} f(t) + \lim_{t \rightarrow \infty} g(t) = \left(\int_0^\infty e^{-x^2} dx \right)^2$$

y por tanto

$$\int_0^\infty e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$$

Ejemplo 1.4 La integral que nos proponemos evaluar a continuación es:

$$F(t) = \int_0^\infty e^{-x^2} \cos(2xt) dx$$

La integral de Riemann existe (en sentido impropio), pues el coseno está acotado por 1 en valor absoluto y hemos probado que la integral de e^{-x^2} existe. Por la misma razón que en el ejemplo anterior, existe la integral de Lebesgue. Derivemos la integral con respecto a t bajo el signo integral (se cumplen las condiciones adecuadas para poder hacerlo):

$$F'(t) = -2 \int_0^\infty xe^{-x^2} \sin(2xt) dx = -2t \int_0^\infty e^{-x^2} \cos(2xt) dx$$

Se tiene entonces:

$$F'(t) = -2tF(t)$$

Esta ecuación diferencial es inmediata de resolver y como $F(0) = \sqrt{\pi}/2$ se obtiene:

$$F(t) = \frac{\sqrt{\pi}}{2} e^{-t^2}$$

Ejemplo 1.5 Calculemos ahora para $t \geq 0$ la siguiente integral:

$$F(t) = \int_0^\infty e^{-xt} \frac{\operatorname{sen} x}{x} dx$$

Supongamos que $t > 0$. Acotando el valor absoluto del integrando por la función e^{-xt} se ve que la integral de Riemann en sentido impropio existe y también la integral Lebesgue. La función $F(t)$ verifica las dos propiedades siguientes:

$$\lim_{t \rightarrow \infty} F(t) = 0, \quad F'(t) = -\frac{1}{1+t^2}$$

Para la primera, basta ver que se cumplen las hipótesis para poder calcular el límite bajo el signo integral. Para la segunda, hay que verificar que se cumplen las correspondientes a poder derivar bajo la integral. Para poder llegar a ese resultado hay que integrar por partes:

$$\int_0^r e^{-xt} \operatorname{sen} x dx = -\frac{e^{-rt}}{1+t^2} (t \operatorname{sen} r + \cos r) + \frac{1}{1+t^2}$$

y tomar el límite cuando $r \rightarrow \infty$. Ahora basta con integrar y obtener:

$$F(t) = -\arctan t + C$$

Cuando $t \rightarrow \infty$, $F(t) \rightarrow 0$. Por lo tanto $C = \pi/2$.

Veamos ahora que pasa cuando $t = 0$. La integral es:

$$\int_0^{\infty} \frac{\operatorname{sen} x}{x} dx$$

y no es integrable Lebesgue. La integral impropia del valor absoluto del integrando no existe. En un intervalo $[(k-1)\pi, k\pi]$ se tiene

$$\int_{(k-1)\pi}^{k\pi} \frac{|\operatorname{sen} x|}{x} dx \geq \frac{1}{k\pi} \int_{(k-1)\pi}^{k\pi} |\operatorname{sen} x| dx = \frac{2}{k\pi}$$

En un intervalo $[0, n\pi]$ la integral se obtiene como una suma de de las anteriores:

$$\int_0^{n\pi} \frac{|\operatorname{sen} x|}{x} dx = \sum_{k=1}^n \int_{(k-1)\pi}^{k\pi} \frac{|\operatorname{sen} x|}{x} dx \geq \frac{2}{\pi} \sum_{k=1}^n \frac{1}{k}$$

y la suma no es convergente.

Sin embargo es integrable Riemann en sentido impropio, debido a las cancelaciones entre las partes negativas y positivas. En efecto:

$$\int_a^b \frac{\operatorname{sen} x}{x} dx = \frac{\cos a}{a} - \frac{\cos b}{b} - \int_a^b \frac{\cos x}{x^2} dx$$

y acotando:

$$\left| \int_a^b \frac{\operatorname{sen} x}{x} dx \right| \leq \frac{1}{b} + \frac{1}{a} + \int_a^b \frac{1}{x^2} dx = \frac{2}{a}$$

Según hemos visto anteriormente esto implica que la función es integrable Riemann. Si ponemos $a = n$, $b = n+1$, tenemos:

$$\left| \int_n^{n+1} \frac{\operatorname{sen} x}{x} dx \right| \leq \frac{2}{n}$$

y por tanto,

$$\int_n^{n+1} \frac{\operatorname{sen} x}{x} dx \rightarrow 0, \quad n \rightarrow \infty$$

Vamos a calcular ahora el valor de esa integral (en sentido Riemann). Definamos la sucesión de funciones:

$$f_n(t) = \int_0^n e^{-xt} \frac{\operatorname{sen} x}{x} dx, \quad t \geq 0$$

que verifica:

$$\lim_{n \rightarrow \infty} f_n(n) = 0$$

ya que:

$$|f_n(n)| \leq \int_0^n e^{-nx} dx = \frac{1}{n}(1 - e^{-n^2}) \leq \frac{1}{n}$$

Téngase en cuenta que

$$f_n(0) = \int_0^n \frac{\operatorname{sen} x}{x} dx$$

Se cumplen las condiciones necesarias para poder calcular la derivada bajo el signo integral:

$$f'_n(t) = - \int_0^n e^{-xt} \operatorname{sen} x dx = \frac{e^{-nt}(t \operatorname{sen} n + \cos n) - 1}{1 + t^2}$$

y esta sucesión de funciones se puede acotar por una función que es integrable Lebesgue en $[0, \infty)$:

$$|f'_n(t)| \leq \frac{1}{1+t^2} (1 + (1+t)e^{-t})$$

Además:

$$\lim_{n \rightarrow \infty} f'_n(t) = -\frac{1}{1+t^2}$$

Consideremos la sucesión f'_n restringida a los intervalos $[0, n]$:

$$g_n(x) = f'_n(x)\chi_{[0, n]}$$

Ahora tenemos una sucesión de funciones integrables Lebesgue, definidas en toda la semirrecta positiva. La sucesión $g_n(t)$ está acotada en valor absoluto por $f'_n(t)$ y tiene el mismo límite que ésta. Entonces, aplicando el teorema de convergencia dominada de Lebesgue:

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_0^n f_n(t) dt &= \lim_{n \rightarrow \infty} \int_0^\infty g_n(t) dt = \int_0^\infty \lim_{n \rightarrow \infty} g_n(t) dt = \\ &= \int_0^\infty \lim_{n \rightarrow \infty} f'_n(t) dt = - \int_0^\infty \frac{1}{1+t^2} dt = -\frac{\pi}{2} \end{aligned}$$

Finalmente en la relación:

$$\int_0^{n+1} \frac{\operatorname{sen} x}{x} dx = \int_0^n \frac{\operatorname{sen} x}{x} dx + \int_n^{n+1} \frac{\operatorname{sen} x}{x} dx = f_n(0) + \int_n^{n+1} \frac{\operatorname{sen} x}{x} dx$$

se toma el límite cuando $n \rightarrow \infty$ y se sustituye

$$\lim_{n \rightarrow \infty} f_n(0) = \frac{\pi}{2}$$

igualdad que se obtiene del teorema fundamental del cálculo:

$$\int_0^n f'_n(t) dt = f_n(n) - f_n(0)$$

y tomando el límite cuando $n \rightarrow \infty$.

Tenemos por tanto:

$$\int_0^\infty \frac{\operatorname{sen} x}{x} dx = \frac{\pi}{2}$$

que coincide con el valor de $F(t)$ cuando $t \rightarrow 0$.

1.5. Productos de medidas

Nos interesa en esta sección el estudio de medidas en espacios que son producto cartesiano de dos (o más) espacios de medida. El primer resultado se refiere a estructuras de anillos en espacios producto.

Proposición 1.5.1 *Si \mathcal{S} y \mathcal{T} son dos anillos de subconjuntos en dos conjuntos X e Y respectivamente, en el espacio $X \times Y$ la clase formada por las uniones finitas disjuntas de rectángulos, es decir conjuntos de la formas $A \times B$, $A \subset X$, $B \subset Y$ es un anillo, que se llama $\mathcal{S} \times \mathcal{T}$.*

Entonces si (X, \mathcal{S}) e (Y, \mathcal{T}) son dos espacios medibles, el espacio $(X \times Y, \mathcal{S} \times \mathcal{T})$ es un espacio medible. Recuérdese que X e Y son uniones de conjuntos medibles en cada medida.

Para conjuntos que no sean rectángulos (es decir, de la forma $A \times B$) necesitamos relacionarlos de alguna forma con conjuntos en los espacios que forman el producto. Introducimos la siguiente definición:

Definición 1.5.1 *Sean (X, \mathcal{S}) e (Y, \mathcal{T}) dos espacios medibles, con espacio producto $(X \times Y, \mathcal{S} \times \mathcal{T})$. Sea $U \subset X \times Y$. Dado $x \in X$, el conjunto de X , $U_x = \{y : (x, y) \in U\}$ se llama la sección de U determinada por x . De igual forma se determinan secciones en el espacio Y : $U^y = \{x : (x, y) \in U\}$*

De forma similar se define lo que es una sección de una función:

Definición 1.5.2 Sean (X, \mathcal{S}) e (Y, \mathcal{T}) dos espacios medibles, con espacio producto $(X \times Y, \mathcal{S} \times \mathcal{T})$. Sea f una función definida en un conjunto U de $X \times Y$. Se llama sección de f por x a la función $f_x(y) = f(x, y)$, $y \in U_x$. De forma similar se definen las secciones f_y .

Los resultados que relacionan los conjuntos y funciones medibles en el espacio producto con los espacios que lo forman son:

Teorema 1.5.1 Toda sección de un conjunto medible es un conjunto medible.

Teorema 1.5.2 Toda sección de una función medible es una función medible.

Pasemos ahora a resultados sobre las medidas μ y ν .

Teorema 1.5.3 Sean (X, \mathcal{S}, μ) e (Y, \mathcal{T}, ν) dos espacios de medida, con espacio producto $(X \times Y, \mathcal{S} \times \mathcal{T})$. Si U es un conjunto medible de $X \times Y$, las funciones $f(x) = \nu(U_x)$, $g(y) = \mu(U^y)$ son funciones medibles no negativas y

$$\int f \, d\mu = \int g \, d\nu$$

En el espacio $(X \times Y, \mathcal{S} \times \mathcal{T})$ se define la medida producto $\mu \times \nu$ de la forma siguiente:

Teorema 1.5.4 Sean (X, \mathcal{S}, μ) e (Y, \mathcal{T}, ν) dos espacios de medida, con medidas σ -finitas y espacio producto $(X \times Y, \mathcal{S} \times \mathcal{T})$. La función λ definida por:

$$\lambda(U) = \int \nu(U_x) \, d\mu(x) = \int \mu(U^y) \, d\nu(y)$$

es una medida σ -finita y para todo rectángulo medible:

$$\lambda(A \times B) = \mu(A)\nu(B)$$

Se llama la medida producto cartesiano de las medidas μ y ν , $\lambda = \mu \times \nu$.

La última condición determina a λ de forma única.

El teorema más importante relativo a la integración con medidas producto es el teorema de Fubini, que relaciona la integral en el espacio producto con las integrales en los espacios que lo forman. Supongamos que $h(x, y)$ es una función integrable definida en el espacio $X \times Y$, con la medida $\lambda = \mu \times \nu$ (como siempre se tienen dos espacios de medida (X, \mathcal{S}, μ) e (Y, \mathcal{T}, ν) y el espacio producto $(X \times Y, \mathcal{S} \times \mathcal{T}, \mu \times \nu)$). La integral (la integral doble) es:

$$\int h(x, y) \, d\lambda(x, y) \equiv \int h(x, y) (\, d\mu(x) \times d\nu(y))$$

Supongamos que $h_x(y)$ es la sección de h en $x \in X$, y sea

$$f(x) = \int h_x(y) \, d\nu(y)$$

en el caso en que esté bien definida. Si $f(x)$ es integrable, se tiene:

$$\int f(x) \, d\mu(x) = \int \left(\int h(x, y) \, d\nu(y) \right) d\mu(x) \equiv \int d\mu(x) \int h(x, y) \, d\nu(y)$$

De manera similar se puede repetir el argumento con la otra sección de $h(x, y)$, $h_y(x)$, y

$$g(y) = \int h_y(x) \, d\mu(x)$$

$$\int g(y) d\nu(y) = \int \left(\int h(x, y) d\mu(x) \right) d\nu(y) \equiv \int d\nu(y) \int h(x, y) d\mu(x)$$

Estas integrales se llaman iteradas. Tenemos tres tipos de integrales definidas para una función: la doble (asociada al espacio de medida producto) y las dos iteradas. El problema es si tienen alguna relación. Este es el contenido del teorema de Fubini, que establece la igualdad de estas tres integrales bajo ciertas condiciones.

Proposición 1.5.2 *Un subconjunto de $X \times Y$ es de medida cero si y solo si casi toda X -sección (o bien casi toda Y -sección) tiene medida cero.*

La demostración es una simple consecuencia de la definición de λ , la medida producto:

$$\lambda(A) = \int \nu(A_x) d\mu(x) = \int \mu(A^y) d\nu(y)$$

Proposición 1.5.3 *Para toda función medible y no negativa en el producto $X \times Y$ se tiene:*

$$\int h d(\mu \times \nu) = \int \left(\int h d\mu \right) d\nu = \int \left(\int h d\nu \right) d\mu$$

y como conclusión

Teorema 1.5.5 (Fubini) *Sea $h(x, y)$ una función integrable en $X \times Y$, tal que toda sección de h es integrable. Entonces, las funciones:*

$$f(x) = \int h(x, y) d\nu(y), \quad g(y) = \int h(x, y) d\mu(x)$$

son integrables y se tiene:

$$\int h d(\mu \times \nu) = \int \left(\int h d\mu \right) d\nu = \int \left(\int h d\nu \right) d\mu$$

Capítulo 2

Espacios topológicos, métricos y normados

2.1. Introducción

En este capítulo se estudian conjuntos de puntos con ciertas propiedades derivadas de la existencia de una topología. Como lo que nos preocupa esencialmente son los espacios de Hilbert, introduciremos unas nociones elementales de los espacios topológicos y métricos para detenernos con más detalle en los espacios normados.

Está claro que no pretendemos hacer una introducción pormenorizada sino simplemente esbozar las ideas fundamentales y las relaciones que existen entre estos conceptos

2.2. Espacios topológicos

Sea X un conjunto. Llamaremos puntos a sus elementos. Como se ha visto en \mathbf{R} , la noción de valor absoluto de un número real da lugar a la de distancia entre dos números reales (como el valor absoluto de la diferencia) y esto a una serie de relaciones de proximidad de los puntos, a nociones de convergencia de sucesiones, de intervalos abiertos y cerrados, de conjuntos acotados, etc. Sin embargo, la noción de distancia no es esencial para definir estos conceptos. A continuación introducimos la definición de una topología y veremos como aparecen en cualquier espacio dotado de ella muchos de los conceptos que se han visto en \mathbf{R} .

Definición 2.2.1 *Sea X un conjunto de puntos. Una topología \mathcal{T} en X es una clase de subconjuntos de X que verifican las siguientes propiedades:*

1. $\emptyset, X \in \mathcal{T}$
2. $A_1, \dots, A_n \in \mathcal{T} \Rightarrow \bigcap_{i=1}^n A_i \in \mathcal{T}$
3. $A_i \in \mathcal{T}, i \in I \Rightarrow \bigcup_{i \in I} A_i \in \mathcal{T}$

Es decir las intersecciones finitas de elementos de \mathcal{T} están en \mathcal{T} y las uniones arbitrarias de elementos de \mathcal{T} están en \mathcal{T} . Se dice que (X, \mathcal{T}) es un espacio topológico. A los elementos de \mathcal{T} se les llama abiertos.

Hay dos topologías triviales en todo conjunto. Una, en la que son abiertos todos los subconjuntos de X . Otra, en la que los únicos conjuntos abiertos son el total y el vacío: $\{\emptyset, X\}$. Pero hay topologías más interesantes. Por ejemplo en \mathbf{R} , con la distancia euclidiana, donde un conjunto es abierto si para cada uno de sus puntos existe un intervalo abierto que lo contiene y que está contenido en el conjunto. Es fácil ver que se cumplen las propiedades que deben verificar los conjuntos abiertos.

Definición 2.2.2 *Un subconjunto de X es cerrado si su complementario es abierto.*

Para los conjuntos cerrados se tienen las siguientes propiedades

Proposición 2.2.1 *En un espacio topológico se verifica:*

1. \emptyset, X son cerrados
2. F_1, \dots, F_n cerrados, entonces $\bigcup_{i=1}^n F_i$ es cerrado
3. $F_i, i \in I$ cerrados, entonces $\bigcap_{i \in I} F_i$ es cerrado.

La demostración es muy sencilla pues basta aplicar las leyes de complementariedad (el complementario de la unión es la intersección de los complementarios).

La noción de proximidad que en \mathbf{R} (y en todo espacio métrico, ver más adelante) da la distancia, se realiza aquí mediante la noción de entorno.

Definición 2.2.3 *Sea (X, \mathcal{T}) un espacio topológico. Se dice que el conjunto V es un entorno de un punto $x \in X$ si existe un abierto de X tal que*

$$x \in A \subset V$$

Nótese que el entorno no tiene por qué ser abierto. En particular un abierto es un entorno de cada uno de sus puntos.

Dado un punto $x \in X$ se puede definir lo que se entiende por una base de entornos de x .

Definición 2.2.4 *Sea x un punto de un espacio topológico. Se dice que la clase de conjuntos $V_i, i \in I$ es una base de entornos de x si dado cualquier entorno de x, U , existe $i \in I$ tal que $x \in V_i \subset U$.*

Se dice que el espacio verifica el primer axioma de numerabilidad (ANI) si para todo punto del espacio existe una base de entornos numerable. Una definición similar se puede hacer para todo el espacio.

Definición 2.2.5 *Sea (X, \mathcal{T}) un espacio topológico. Se dice que la clase de conjuntos $V_i, i \in I$ es una base de la topología si son abiertos y para todo punto $x \in X$ y para todo entorno de x, U , existe $i \in I$ tal que $x \in V_i \subset U$.*

Se dice que el espacio verifica el segundo axioma de numerabilidad (ANII) si la topología tiene una base numerable. Si un espacio es ANII también es ANI, pues la clase de conjuntos de base de la topología que contiene a un punto x es una base de entornos de ese punto. Sin embargo hay espacios que son ANI pero no ANII.

Clasificaremos ahora los puntos en relación con sus propiedades respecto a conjuntos.

Definición 2.2.6 *Sea B un subconjunto de X . Se dice que $x \in X$ es un punto interior de B si existe un entorno de x contenido en B .*

Esto implica evidentemente que $x \in B$. Todos los puntos de un abierto son interiores (inmediato de la definición). Pero un cerrado puede no tener puntos interiores (por ejemplo en \mathbf{R} con la topología anteriormente definida, los puntos son cerrados y, obviamente, no contienen puntos interiores). El conjunto de los puntos interiores de un conjunto B es un abierto, que se llama el interior de ese conjunto, B° . Es el mayor abierto contenido en ese conjunto. En efecto, si $A \subset B$, A abierto, para todo punto x de A existe un entorno de x que está contenido en A y por tanto en B . De acuerdo con la definición $x \in B^\circ$. Luego $A \subset B^\circ$.

Un conjunto es abierto si y solo si coincide con el conjunto de sus puntos interiores (una fácil consecuencia de lo dicho anteriormente).

Definición 2.2.7 *Sea $B \subset X$. Se dice que el punto $x \in X$ es un punto de adherencia (o de cierre) del conjunto B si para todo entorno V de x se tiene $V \cap B \neq \emptyset$*

El conjunto de los puntos de adherencia de un conjunto B se llama el cierre de ese conjunto \bar{B} . El cierre de un conjunto es un conjunto cerrado. Para demostrarlo hay que probar que su complementario es abierto. Si $x \in X - \bar{B}$, entonces $x \notin \bar{B}$ lo que quiere decir que hay un entorno de x que está contenido en $X - B$, pues es evidente que $B \subset \bar{B}$. Es el menor cerrado que contiene al conjunto (la demostración es similar a la que hemos hecho con el interior). Un conjunto es cerrado si y solo si coincide con su cierre.

Definición 2.2.8 *Se dice que un conjunto B es denso en X si su cierre coincide con X .*

Por ejemplo, los racionales y los reales. En todo intervalo abierto que contenga un número real hay un número racional. Luego los reales son el cierre de los racionales.

Definición 2.2.9 *Se dice que un espacio topológico es separable si tiene un subconjunto numerable denso.*

El ejemplo anterior muestra que los racionales constituyen un conjunto denso en los reales y son numerables, luego \mathbf{R} es separable.

Definición 2.2.10 *Se dice que un punto x es de acumulación de un conjunto B si todo entorno de x excluido el propio punto x tiene intersección no vacía con el conjunto:*

$$(V - \{x\}) \cap B \neq \emptyset$$

Un ejemplo sencillo es la sucesión $\{1/n\}$ en los reales. El 0 es un punto de acumulación. El conjunto de los puntos de acumulación es un conjunto cerrado y se llama el conjunto derivado, B' . El cierre de un conjunto es la unión del conjunto derivado con el propio conjunto.

Definición 2.2.11 *Se dice que un punto x es frontera de un conjunto B si todo entorno de x tiene intersección no vacía con el conjunto y su complementario:*

$$V \cap B \neq \emptyset, \quad V \cap (X - B) \neq \emptyset$$

El intervalo $(0, 1]$ tiene dos puntos frontera en \mathbf{R} , el 0 que no está en el intervalo y el 1 que sí lo está. El conjunto de puntos frontera de un conjunto se llama la frontera de ese conjunto. Un conjunto es cerrado si contiene a su frontera y es abierto si es disjunto con ella. El intervalo anterior no es abierto ni cerrado: ni contiene a su frontera ni es disjunto con ella.

Aunque en un espacio topológico las sucesiones no juegan el papel fundamental que realizan en un espacio métrico, no por ello dejan de ser importantes.

Definición 2.2.12 *Sea una sucesión x_n en X . Se dice que x_n converge a x_0 cuando n tiende a infinito, si para todo entorno V de x_0 existe n_0 tal que si $n \geq n_0$ entonces $x_n \in V$.*

En un espacio topológico se pueden introducir lo que se llama los axiomas de separación. Nos restringiremos a la separación de tipo Hausdorff. Los espacios de tipo Hausdorff son los más utilizados en las aplicaciones.

Definición 2.2.13 *Un espacio topológico se dice que es de Hausdorff si dados dos puntos $x \neq y$ en X , existen dos entornos V y W de x e y respectivamente tal que $V \cap W = \emptyset$.*

Está claro que \mathbf{R} es un espacio de Hausdorff. Pero un conjunto X con la topología que solo tiene como abiertos el vacío y el conjunto total no puede ser Hausdorff (si tiene dos puntos o más). Definamos finalmente, en lo que a conjuntos se refiere, lo que es un conjunto compacto. Primero definimos qué es un recubrimiento.

Definición 2.2.14 *Sea X un espacio topológico y $B \subset X$. Se dice que la colección de conjuntos abiertos A_i , $i \in I$ es un recubrimiento de B si $B \subset \cup A_i \in I$.*

y a continuación lo que es un subrecubrimiento.

Definición 2.2.15 Sea X un espacio topológico, $B \subset X$ y $A_i, i \in I$ un recubrimiento de B . Un subrecubrimiento de $A_i, i \in I$ es una familia $A_i, i \in J$, con $J \subset I$, tal que $B \subset \cup_{i \in J} A_i$.

para acabar con la definición de un compacto.

Definición 2.2.16 Se dice que $K \subset X$ es un conjunto compacto si todo recubrimiento abierto admite un subrecubrimiento finito.

Para ser precisos, la noción de compacto se establece sobre espacios topológicos y no sobre subconjuntos. La definición sobre subconjuntos se hace estudiando la topología relativa inducida sobre el subconjunto (en ella un conjunto es abierto si es la intersección de un abierto del espacio ambiente con el subconjunto).

El intervalo $(0, 1) \subset \mathbf{R}$ no es un compacto. Se pueden construir recubrimientos abiertos que no contengan ningún subrecubrimiento finito. El problema es que los extremos del intervalo no están en él. Como veremos en la siguiente proposición, este ejemplo se resuelve de manera inmediata. Pero vamos a discutir un ejemplo de recubrimiento abierto de $(0, 1)$ que no admite subrecubrimientos finitos. Sea $A_n = (1/(n+1), 1)$, $n = 1, 2, \dots$ una familia de intervalos abiertos que recubren a $(0, 1)$, es decir:

$$(0, 1) \subset \bigcup_{n=1}^{\infty} \left(\frac{1}{n+1}, 1 \right)$$

Supongamos que tenemos una subfamilia finita. Entonces existe un n_0 máximo, tal que A_{n_0} está en la subfamilia pero A_{n_0+1} no. Está claro que hay puntos de $(0, 1)$ que no son cubiertos por el subrecubrimiento (los menores que $1/(n_0+2)$).

Los conjuntos cerrados no son en general compactos. Pero el inverso es cierto.

Proposición 2.2.2 Sea K un compacto de un espacio topológico (X, \mathcal{T}) de Hausdorff. Entonces K es cerrado. Además los subconjuntos cerrados de un conjunto compacto son compactos.

Demostremos que $X - K$ es abierto. Sea $y \in X - K$. Para cada $x \in K$, al ser X un espacio de Hausdorff, existe un entorno abierto de x a cuyo cierre no pertenece y . La colección de esos abiertos es un recubrimiento abierto de K , y al ser K compacto, podemos extraer un subrecubrimiento finito. El cierre de la unión (que es la unión de los cierres al ser finita) contiene a K y no contiene a y . Luego su complementario es un abierto que contiene a y y está contenido en $X - K$ que es un abierto. En cuanto a la segunda parte, sea K un compacto y $F \subset K$ cerrado. Dado un recubrimiento abierto de F , consideramos el recubrimiento de K añadiendo al anterior el abierto $K - F$. Pero K es compacto, luego podemos extraer un subrecubrimiento finito. Eliminando $X - F$ de ese subrecubrimiento obtenemos un recubrimiento de F que es un subrecubrimiento finito del recubrimiento inicial. Luego F es compacto.

Como hemos dicho antes, esta propiedad implica que $(0, 1)$ no puede ser compacto, al no ser cerrado. La recta real (que es cerrada) no es compacta (con la distancia usual no está acotada). Un caso más interesante desde el punto de vista de las aplicaciones a espacios vectoriales normados (en particular a espacios de Hilbert) es que la bola unidad cerrada, es decir, el conjunto de los vectores de ese espacio cuya norma es menor o igual que 1, es cerrada y acotada pero no es compacta cuando la dimensión es infinita (como veremos más adelante, hay sucesiones infinitas que no tienen puntos de acumulación).

El siguiente resultado es fundamental en topología.

Teorema 2.2.1 (Tychonoff) El producto de espacios compactos es un espacio compacto.

Como consecuencia, el conjunto $[0, 1]^I$ es un compacto (I es un conjunto cualquiera). Sin embargo la esfera unidad en un espacio normado (ver sección siguiente) de dimensión infinita no es un compacto.

Aunque un espacio no sea compacto puede tener esa propiedad desde un punto de vista local.

Definición 2.2.17 Sea (X, \mathcal{T}) un espacio topológico. Se dice que es localmente compacto si todo punto posee un entorno cuyo cierre es compacto.

Aunque \mathbf{R} no es compacto, es localmente compacto. Todo punto tiene un entorno cuyo cierre es compacto (porque en \mathbf{R} los cerrados y acotados son compactos). Como veremos más adelante, los espacios vectoriales de dimensión finita (como \mathbf{R}), con la topología que les hace espacios vectoriales topológicos, son localmente compactos, pero este resultado no es cierto cuando se trata de espacios vectoriales de dimensión infinita.

Introducimos ahora la noción de función continua entre espacios topológicos.

Definición 2.2.18 Sea $f : (X_1, \mathcal{T}_1) \rightarrow (X_2, \mathcal{T}_2)$ una función. Se dice que f es continua en un punto $x \in X_1$ si para todo entorno V de $f(x)$ existe un entorno U de x tal que $f(U) \subset V$.

Una función continua entre dos espacios topológicos se caracteriza de la forma siguiente:

Definición 2.2.19 Sea $f : (X_1, \mathcal{T}_1) \rightarrow (X_2, \mathcal{T}_2)$. La función f es continua si y solo si la imagen inversa de un abierto es un abierto.

Está claro que si f es continua en X_1 lo es en cada punto de X_1 . Dado $x \in X_1$, y un entorno cualquiera de $f(x)$, tenemos entonces un abierto de X_2 , A_2 tal que, $f(x) \in A_2 \subset V$. Por ser f continua, $f^{-1}(A_2)$ es un abierto de X_1 y por tanto un entorno de x . Entonces, $f(f^{-1}(A_2)) \subset V$. Y si f es continua en todos los puntos del espacio X_1 , se tiene que f es continua en X_1 . En efecto, sea V un abierto en el segundo espacio, X_2 . Hay que probar que su imagen inversa mediante f^{-1} es un abierto de X_1 . Sea $x \in f^{-1}(V)$. Por ser f continua en x , y ser $f(x) \in V$, que es un abierto y por tanto un entorno de $f(x)$, se tiene que existe U , entorno de x tal que $f(U) \subset V$. Pero $U \subset f^{-1}(f(U)) \subset f^{-1}(V)$, que por lo tanto es abierto.

Se tiene el siguiente resultado en cuanto a los conjuntos compactos:

Proposición 2.2.3 Sea $f : (X_1, \mathcal{T}_1) \rightarrow (X_2, \mathcal{T}_2)$ continua. Entonces si K es un compacto de X_1 , se tiene que $f(K)$ es un compacto de X_2 .

La demostración es muy sencilla. Se toma un recubrimiento abierto arbitrario de $f(K)$. La imagen inversa de cada elemento del recubrimiento es un abierto del primer espacio, pues f es continua. Pero la unión de esas imágenes inversas es ciertamente un recubrimiento abierto de K . Como K es compacto se puede extraer un subrecubrimiento finito. Los conjuntos del recubrimiento de $f(K)$ correspondientes al subrecubrimiento (de K) escogido son también un subrecubrimiento finito de $f(K)$, que por lo anterior es compacto.

Nótese que la imagen de un abierto bajo una aplicación continua no tiene por qué ser abierta. Sin embargo se puede definir el siguiente concepto.

Definición 2.2.20 Sea f una aplicación entre dos espacios topológicos. Se dice que es abierta si la imagen de un abierto es abierta.

Si entre dos espacios topológicos se puede establecer una aplicación biyectiva y tanto ella como su inversa son continuas, se dice que son homeomorfos y la aplicación se llama un homeomorfismo. Un homeomorfismo es claramente una aplicación abierta.

Finalmente sean (X_1, \mathcal{T}_1) y (X_2, \mathcal{T}_2) . En el conjunto producto cartesiano $X_1 \times X_2$ se puede definir una topología, la topología producto. Una base de esta topología está formada por los productos cartesianos de abiertos de X_1 y X_2 . Es decir un conjunto $A \subset X_1 \times X_2$ es un abierto en la topología producto si para todo punto $(x_1, x_2) \in A$ se tienen dos entornos U_1 y U_2 de x_1 y x_2 respectivamente tales que:

$$(x_1, x_2) \in U_1 \times U_2 \subset A$$

Proposición 2.2.4 Las proyecciones $p_i : X_1 \times X_2 \rightarrow X_i$ son continuas cuando en el espacio producto se usa la topología definida anteriormente.

La demostración es una consecuencia directa de la definición de la topología producto. Consideremos un abierto A_1 de X_1 . Su imagen inversa mediante la proyección, $p_1^{-1}(A_1) = A_1 \times X_2$ es un abierto en la topología producto.

2.3. Espacio métricos

Consideremos nuevamente un espacio X y en él una distancia, es decir una aplicación de $X \times X$ en \mathbf{R} verificando las propiedades que se detallarán a continuación. Estudiaremos entonces que es posible definir una topología asociada a esta métrica y determinaremos sus propiedades más importantes.

Definición 2.3.1 Sea X un conjunto y d una aplicación:

$$d : X \times X \rightarrow \mathbf{R}$$

que verifica:

1. $d(x, y) \geq 0$, $d(x, y) = 0 \Leftrightarrow x = y$
2. $d(x, y) = d(y, x)$
3. $d(x, y) \leq d(x, z) + d(z, y)$ (desigualdad triangular)

Un espacio métrico es un conjunto de puntos con una distancia. El conjunto \mathbf{R} con la distancia definida a partir del valor absoluto es un espacio métrico:

$$d(x, y) = |x - y|, \quad x, y \in \mathbf{R}$$

La aplicación

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

definida en un conjunto X lo convierte en un espacio métrico.

Si $d(x, y)$ es una distancia, la aplicación:

$$\rho(x, y) = \frac{d(x, y)}{1 + d(x, y)}$$

es también una distancia, con valores entre 0 y 1.

Veamos como definir una topología a partir de la métrica. Primero definimos lo que se entiende por bola abierta:

Definición 2.3.2 Sea (X, d) un espacio métrico. Se define la bola abierta de centro $a \in X$ y radio $r > 0$ por:

$$B(a, r) = \{x \in X : d(x, a) < r\}$$

y a continuación lo que entendemos por un abierto en (X, d) .

Definición 2.3.3 Sea (X, d) un espacio métrico. Se dice que $A \subset X$ es un abierto (en la topología de la métrica) si para todo $x \in A$ existe una bola de centro x y radio r contenida en A .

Lo que hay que hacer ahora es comprobar que esta colección de abiertos así definidos es una topología en X . Es claro que \emptyset y X son abiertos (trivialmente de la definición de abierto). Veamos ahora que la intersección de dos abiertos es un abierto. Sean A_1 y A_2 abiertos en X . Sea $x \in A_1 \cap A_2$. Existen dos bolas de centro x y radios r_1 y r_2 tales que:

$$B(x, r_1) \subset A_1, \quad B(x, r_2) \subset A_2$$

Sea $r = \min\{r_1, r_2\}$. Es claro que

$$B(x, r) \subset A_1 \cap A_2$$

En cuanto a la unión, si $x \in \cup_{i \in I} A_i$ entonces existe $i_0 \in I$ tal que $x \in A_{i_0}$ que, por ser abierto, contiene a una bola de centro x y radio r , que también estará contenida en la unión.

Luego efectivamente estos abiertos constituyen una topología en X . El conjunto de bolas de centro en cualquier punto de X y radios arbitrarios es una base de la topología, así como el conjunto de bolas

de centro x fijado y radio arbitrario es una base de entornos de ese punto. Lo que se tiene además es que la familia

$$B\left(x, \frac{1}{n}\right), \quad n = 1, 2, \dots$$

es también una base de entornos del punto x , y es numerable, por lo que todo espacio métrico verifica el primer axioma de numerabilidad.

De forma similar a las bolas abiertas, se definen las cerradas

$$\bar{B}(a, r) = \{x \in X : d(x, a) \leq r\}$$

Toda bola cerrada es un cerrado como es fácil de ver. Podemos ahora dar una caracterización de la convergencia de una sucesión en términos de la métrica.

Proposición 2.3.1 *Sea (X, d) un espacio métrico y $\{x_n\}$ una sucesión en ese espacio. Se dice que $\{x_n\}$ converge a x_0 (en la topología de la métrica) si para todo $\epsilon > 0$ existe n_0 tal que si $n \geq n_0$ se tiene $d(x_n, x_0) < \epsilon$. Dicho de otra forma, $x_n \rightarrow x_0$ cuando $n \rightarrow \infty$ si $d(x_n, x_0) \rightarrow 0$ en \mathbf{R} .*

La demostración se basa simplemente en un uso adecuado de bolas abiertas y la propiedad de ser una base de entornos del punto.

Proposición 2.3.2 *Sea (X, d) un espacio métrico. Entonces, $x \in X$ es un punto de acumulación de un conjunto $A \subset X$ si y solo si existe una sucesión de puntos en A (distintos de x) que converge a x .*

Consideremos las bolas $B(x, 1/n)$. Por ser x un punto de acumulación, en todas las bolas hay puntos de A distintos de x . Sea $x_n \in B(x, 1/n) \cap A$. La sucesión x_n está en A y converge a x . En sentido contrario es inmediato por la definición de punto de acumulación. Nótese la importancia de tener una base de entornos de x formada por las bolas de centro x y radios $1/n$.

Sobre un mismo conjunto es posible definir en general distintas métricas. La pregunta es como son las topologías que determinan esas métricas. Se dice que dos métricas son equivalentes si las topologías que generan son iguales (es decir los abiertos en una de ellas son abiertos en la otra y viceversa).

Por ejemplo, las métricas definidas anteriormente por $d(x, y)$, $\rho(x, y)$ son equivalentes.

Como hemos visto, los puntos de adherencia de un conjunto se caracterizan por el hecho de que existe una sucesión contenida en el conjunto cuyo límite es el punto en cuestión. Si es de acumulación la sucesión se puede elegir de puntos distintos. Esta caracterización está asociada a las propiedades de ser un espacio métrico (en particular a ser ANI) y no se verifica en un espacio topológico arbitrario.

Usando la noción de distancia es posible definir lo que quiere decir acotación en un espacio métrico.

Definición 2.3.4 *Sea (X, d) un espacio métrico. Se dice que un conjunto es acotado si existe un número real positivo k tal que la distancia entre dos puntos cualquiera del conjunto es menor que k .*

Proposición 2.3.3 *De toda sucesión acotada en un espacio métrico se puede extraer una subsucesión convergente.*

Un concepto importante en espacios métricos es el de sucesiones de Cauchy.

Definición 2.3.5 *Sea (X, d) un espacio métrico y x_n una sucesión en X . Se dice que x_n es una sucesión de Cauchy (fundamental) si para todo $\epsilon > 0$ existe n_0 tal que si $n, m \geq n_0$ se tiene que $d(x_n, x_m) < \epsilon$.*

Una sucesión de Cauchy está acotada. Es fácil demostrar que toda sucesión convergente es de Cauchy, pero el inverso no siempre es cierto (el ejemplo más sencillo es el de los racionales con la topología de la métrica heredada de \mathbf{R}). Esto nos lleva a la introducción del concepto de espacio métrico completo.

Definición 2.3.6 *Se dice que un espacio métrico es completo si toda sucesión de Cauchy es convergente.*

Por ejemplo, \mathbf{R} es completo.

Definición 2.3.7 Sean (X_1, d_1) , (X_2, d_2) dos espacios métricos y $f : X_1 \rightarrow X_2$ una aplicación. Se dice que f es una isometría si

$$d_2(f(x), f(y)) = d_1(x, y), \quad x, y \in X_1$$

Es obvio que una isometría es continua. Las isometrías son siempre inyectivas (si $f(x) = f(y)$ entonces $d_2(f(x), f(y)) = 0$ y por ser isometría, $d_1(x, y) = 0$, luego $x = y$). Si la aplicación es sobreyectiva se dice que los espacios son isométricos (la isometría es un homeomorfismo, es decir, dos espacios isométricos son homeomorfos).

Es posible demostrar que todo espacio métrico no completo admite un completado (es decir existe un espacio métrico completo que lo contiene y del que es un subespacio denso, y cuya métrica restringida al primer espacio es la métrica de partida). El completado es único (salvo isometría). Se dice que un espacio métrico es precompacto si su completado es compacto.

Para los espacios métricos la caracterización de compactos puede hacerse de otras formas. Se tiene el siguiente resultado.

Teorema 2.3.1 (Bolzano-Weierstrass) Sea (X, d) un espacio métrico. Un conjunto K es compacto si y solo si todo subconjunto infinito de K tiene un punto de acumulación en K .

La condición es necesaria en cualquier espacio topológico. La parte suficiente (para la que se necesita la estructura métrica) no es fácil de probar.

Como consecuencia se tiene:

Proposición 2.3.4 Sea (X, d) un espacio métrico. Un conjunto K es compacto si y solo si toda sucesión en K tiene una subsucesión convergente a un punto de K .

Los conjuntos compactos en \mathbf{R}^n tienen una caracterización muy sencilla. Como veremos en la sección siguiente, todo espacio vectorial topológico de dimensión finita es isomorfo (en el sentido algebraico y en el topológico) a \mathbf{K}^n donde \mathbf{K} es el cuerpo sobre el que está construido (en nuestro caso \mathbf{R} o \mathbf{C}). Se tiene el resultado siguiente.

Teorema 2.3.2 (Heine-Borel) Un subconjunto de un espacio \mathbf{K}^n es compacto si y solo si es cerrado y acotado (en la métrica usual).

Supongamos que $K \subset \mathbf{K}^n$ es compacto. De acuerdo con la proposición 2.2.2 es cerrado. Veamos que también es acotado. Si no lo estuviera se podría construir una sucesión de puntos x_n tales que $\|x_n\| > n$ para todo $n \in \mathbf{N}$ ($\|x\|$ es la norma usual en \mathbf{K}^n , ver sección siguiente). Esta sucesión (que tiene infinitos puntos distintos) no tiene subsucesiones convergentes, lo que es una contradicción con el hecho de que K sea compacto.

Supongamos ahora que K es cerrado y acotado. Al ser acotado, toda sucesión contenida en él está acotada. Pero de toda sucesión acotada es posible extraer una subsucesión convergente (ver 2.3.3). Como K es cerrado el límite está en K (ver 2.3.1).

2.4. Espacios normados

Supongamos que el conjunto de puntos que estudiamos ahora es un espacio vectorial L (real o complejo, $\mathbf{K} = \mathbf{R}$ o \mathbf{C}). Este es el caso fundamental a tratar en estas notas (en particular cuando además tenemos un producto escalar). En el espacio vectorial se puede definir una topología. Cuando esta topología verifica ciertas propiedades en relación con las operaciones de L se dice que tenemos un espacio vectorial topológico (e.v.t.). Los espacios normados son un caso particular de los e.v.t. (y los espacios de Hilbert un caso particular de los normados como veremos más tarde).

Definición 2.4.1 Sea L un espacio vectorial y \mathcal{T} una topología en él. Se dice que (L, \mathcal{T}) es un espacio vectorial topológico si se verifica:

1. La aplicación $(x, y) \rightarrow x + y$ de $L \times L$ en L es continua.

2. La aplicación $(\lambda, x) \rightarrow \lambda x$ de $\mathbf{K} \times L$ en L es continua.

La topología en el cuerpo es la usual (dada por la distancia asociada al módulo). Y en los casos de un producto, la topología es la producto.

Los entornos de un punto en un e.v. t. están directamente relacionados con los entornos del vector 0. En efecto, sea U un entorno de 0. Entonces, el conjunto

$$x + U = \{x + y : y \in U\}$$

es un entorno de x . Y viceversa, todos los entornos de x son de esta forma: sea V un entorno de x . Entonces, $U = -x + V$ es un entorno de 0 y por lo tanto $V = x + U$.

Al disponer de una estructura adicional (la de espacio vectorial), hay conjuntos importantes desde este punto de vista, como son los subespacios. Estos conjuntos no tienen por qué ser cerrados (lo son si la dimensión es finita), pero se verifica:

Proposición 2.4.1 *Si (L, \mathcal{T}) es un espacio vectorial topológico y M un subespacio, el cierre de M en la topología de L es también un subespacio.*

La suma es una aplicación continua, luego $\bar{M} + \bar{M}$ está contenido en \bar{M} . Nótese que lo que se está aplicando es que para una aplicación continua, si $f(A) \subset B$ entonces $f(\bar{A}) \subset \bar{B}$. Por razones semejantes, $\mathbf{K}\bar{M} \subset \bar{M}$.

Dado un subespacio, el espacio vectorial cociente es también un e.v.t. (aunque no insistiremos en como determinar la topología cociente de este espacio). Pero, para ver la importancia que tiene el que los subespacios sean cerrados para satisfacer ciertas propiedades, enunciaremos la siguiente proposición.

Proposición 2.4.2 *Sea L un espacio vectorial topológico y M un subespacio de L . Entonces el espacio vectorial cociente L/M con la topología cociente es Hausdorff si y solo si M es cerrado.*

La dimensión de un espacio vectorial es el cardinal de una de sus bases (algebraicas), que se puede demostrar que existen (se llaman bases de Hamel) y tienen todas el mismo cardinal. Los espacios vectoriales de dimensión finita juegan un papel muy particular. Todo e.v.t. de Hausdorff de dimensión n sobre \mathbf{K} (que es \mathbf{R} o \mathbf{C}) es isomorfo a \mathbf{K}^n (isomorfo quiere decir que existe un isomorfismo algebraico que es también homeomorfismo). La demostración se basa en el isomorfismo algebraico que se puede establecer entre ambos espacios. Fijada una base de L , $\{u_1, \dots, u_n\}$, sea f la siguiente aplicación:

$$f : L \rightarrow \mathbf{K}^n, \quad f(x) = (a_1, \dots, a_n)$$

donde $x = a_1 u_1 + \dots + a_n u_n$. Esta aplicación es biyectiva. Es fácil ver que es continua y abierta.

Además, en este caso las aplicaciones lineales tienen la siguiente propiedad que no es cierta en el caso de dimensión infinita, donde existen aplicaciones lineales que no son continuas.

Proposición 2.4.3 *Sea M un e.v.t. de dimensión finita y L un e.v.t. de dimensión arbitraria. Entonces, toda aplicación lineal de M en L es continua.*

La demostración se puede hacer usando el isomorfismo anterior (de esa forma, una aplicación lineal viene dada por una matriz, y en el espacio de matrices se puede definir una norma). Veremos más tarde como la acotación de una aplicación lineal usando una norma permite demostrar la continuidad.

La propiedad de compacidad local está directamente relacionada con la dimensión finita.

Teorema 2.4.1 *Si L es un e.v.t. de Hausdorff localmente compacto, entonces tiene dimensión finita.*

Otros conjuntos importantes en la teoría de e.v.t. son las variedades lineales. Dado un e.v. L y un subespacio M , una variedad lineal es un conjunto $x + M$ donde $x \in L$, definido por:

$$x + M = \{x + y : y \in M\}$$

La dimensión de una variedad lineal es la del subespacio que la define. Un hiperplano es una variedad lineal propia (distinta de L) de dimensión maximal. Los hiperplanos están relacionados con las formas lineales en L (es decir las aplicaciones lineales de L en \mathbf{K}).

Proposición 2.4.4 *Un subconjunto H de L , e.v., es un hiperplano si y solo si existe una forma lineal f y un escalar $a \in \mathbf{K}$ tales que:*

$$H = \{x : f(x) = a\}$$

La forma f y el escalar a determinan H salvo un factor.

En particular si L es un espacio vectorial topológico, un hiperplano es cerrado o denso en L . Es cerrado si y solo si la forma f que lo define es continua.

A pesar de no tener definida una distancia es posible definir un concepto de acotación en un e.v.t.

Definición 2.4.2 *Se dice que un subconjunto A de un e.v.t. está acotado si para todo entorno de 0 , U , existe $c \in \mathbf{K}$ tal que:*

$$A \subset cU$$

Se puede probar que toda sucesión de Cauchy está acotada en el sentido anterior. Resulta muy interesante ver como propiedades de acotación, tan elementales como la que sigue, implican que la topología de un e.v.t se puede asociar a una métrica

Teorema 2.4.2 *Todo e.v.t. de Hausdorff localmente acotado (es decir, tiene un entorno de 0 acotado) es metrizable (es decir, existe una métrica tal que las bolas abiertas son una base de la topología de L).*

En la teoría de e.v.t. los conjuntos convexos juegan un papel fundamental.

Definición 2.4.3 *Se dice que un subconjunto de un espacio vectorial es convexo si dados dos puntos cualquiera del conjunto, el segmento que los une también está en él.*

Es decir: si $x, y \in A \subset L$, entonces $tx + (1-t)y \in A$ para todo $t \in (0, 1)$.

Definición 2.4.4 *Se dice que un espacio vectorial topológico es localmente convexo si es Hausdorff y todo entorno de un punto x contiene un entorno convexo de x .*

Los espacios localmente convexos están asociados a seminormas.

Definición 2.4.5 *Sea L un e.v. y p una aplicación de L en \mathbf{R} . Se dice que p es una seminorma si se verifica:*

1. $p(ax) = |a|p(x)$, $a \in \mathbf{K}$, $x \in L$
2. $p(x+y) \leq p(x) + p(y)$, $x, y \in L$

Una familia de seminormas determina en L una topología que es localmente convexa. Pero no entraremos en estos detalles ahora.

Lo que más nos interesará en estas notas es una clase particular de seminormas. Una norma en un espacio vectorial L se define por:

Definición 2.4.6 *Sea L un e.v. Una aplicación $\|\cdot\| : L \rightarrow \mathbf{R}$ se dice que es una norma si se verifica:*

1. $\|x\| = 0 \Rightarrow x = 0$
2. $\|ax\| = |a| \|x\|$, $a \in \mathbf{K}$, $x \in L$
3. $\|x+y\| \leq \|x\| + \|y\|$, $x, y \in L$

Se tiene entonces que dada una norma es posible definir una distancia

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

Se dice que un espacio vectorial topológico es un espacio normado, si está dotado de una norma y la topología es la derivada de la norma (de la distancia asociada a la norma). Un espacio normado completo se llama un espacio de Banach.

Entre otras caracterizaciones de un espacio normable (es decir un e.v.t. cuya topología proviene de una norma) se tiene la siguiente

Proposición 2.4.5 *Un e.v.t. de Hausdorff es normable si y solo si tiene un entorno de 0 que es convexo y acotado.*

La topología de un espacio normado puede venir definida por diferentes normas. Dos normas son equivalentes si dan lugar a la misma topología.

Proposición 2.4.6 *Dos normas $\|\cdot\|_1$ y $\|\cdot\|_2$ en un espacio vectorial L son equivalentes si y solo si existen dos constantes $k_1, k_2 > 0$ que verifican para todo $x \in X$:*

$$k_1\|x\|_1 \leq \|x\|_2 \leq k_2\|x\|_1$$

Basta probar que toda bola abierta en una norma contiene una bola abierta de la otra norma. Sea $B_i(0, r)$ una bola abierta en la norma $\|\cdot\|_i$ (para otros puntos el resultado se obtiene por traslación). Entonces, debido a las desigualdades de la proposición, se tienen los contenidos:

$$B_2(0, k_1r) \subset B_1(0, r), \quad B_1\left(0, \frac{r}{k_2}\right) \subset B_2(0, r)$$

lo que implica la igualdad de las topologías. Supongamos ahora que las normas son equivalentes. Entonces la bola $B_2(0, 1)$ contiene una bola $B_1(0, r)$ para algún r . Es decir si $\|x\|_2 < 1$ entonces $\|x\|_1 < r$. Por las propiedades de las normas frente a homotecias, si $\|x\|_2 < k$ entonces $\|x\|_1 < kr$ para todo $k > 0$. Eligiendo $k = \|x\|_2$ se tiene:

$$\|x\|_1 < r\|x\|_2$$

La otra desigualdad se demuestra de modo similar (nótese que se ha empleado en la demostración la propiedad de homogeneidad de la norma).

Como se ha visto, todo e.v.t. de Hausdorff de dimensión finita es isomorfo a \mathbf{K}^n . Además se tiene el siguiente resultado.

Proposición 2.4.7 *Todas las normas en un espacio normado de dimensión finita son equivalentes*

La demostración en un sentido es muy sencilla. Lo haremos en \mathbf{K}^n . Tomemos la base canónica de \mathbf{K}^n , $\{u_1, \dots, u_n\}$. Sea $\|\cdot\|$ una norma arbitraria y $\|\cdot\|_\infty$ la norma del supremo. Se tiene:

$$\|x\| = \left\| \sum_{i=1}^n x_i u_i \right\| \leq \sum_{i=1}^n |x_i| \|u_i\| \leq \|x\|_\infty \sum_{i=1}^n \|u_i\| = k_1 \|x\|_\infty$$

con $k_1 = \sum_{i=1}^n \|u_i\|$. Sin embargo, en el otro sentido la situación es bastante más complicada. Se hace por inducción en la dimensión del espacio vectorial. Para $n = 1$ es trivial. Supongamos que es cierta para $n - 1$. Consideremos \mathbf{K}^n y supongamos que no existe una constante k_2 tal que

$$\|x\|_\infty \leq k_2 \|x\|, \quad \forall x \in \mathbf{K}^n$$

Es decir, dado un número natural m , es posible hallar $x \in \mathbf{K}^n$ distinto de cero tal que:

$$\|x\|_\infty > m \|x\|, \quad \forall x \in \mathbf{K}^n$$

Sea $j \in \{1, \dots, n\}$, tal que $|x_j| = \|x\|_\infty > 0$. Entonces, si

$$x_m = \frac{1}{\|x\|_\infty} x$$

se tiene:

$$\|x_m\| = 1, \quad \|x_m\| < \frac{1}{m}$$

Consideremos ahora el conjunto infinito S de números naturales m para los que se tienen estas condiciones (en el cual podemos suponer que j está fijado, ya que solo puede tomar un número finito de valores y

por tanto uno de ellos se repite un número infinito de veces). El conjunto de vectores de \mathbf{K}^n que tienen la componente j igual a cero, W , es un espacio vectorial de dimensión $n - 1$. En este espacio tenemos:

$$\|y\|_\infty \leq C\|y\|, \quad \forall y \in W$$

Si $m \in S$ se define

$$y_m = x_m - u_j$$

que está en W (pues su componente j es cero). Demostremos que $\{y_m\}$, $m \in S$ es una sucesión de Cauchy. Sea $\epsilon > 0$ y un número natural n_0 tal que $\epsilon > \frac{2}{n_0}$. para $n, m \in S$, $n, m \geq n_0$ se tiene:

$$\|y_n - y_m\| \leq \|x_n - x_m\| \leq \|x_n\| + \|x_m\| \leq \frac{1}{n} + \frac{1}{m} \leq \frac{2}{n_0} \leq \epsilon$$

Por lo que $\{y_m\}$, $m \in S$ es de Cauchy también en la norma del supremo. Al ser el espacio W completo, su límite (cuando $m \rightarrow \infty$, $m \in S$) es un vector $y \in W$. Pero

$$\|u_j + y\| \leq \|u_j + y + y_m - y_m\| \leq \|u_j + y_m\| + \|y - y_m\| \leq \frac{1}{m} + C\|y - y_m\|_\infty$$

Cuando $m \rightarrow \infty$, el segundo miembro tiende a cero y por tanto $y = -u_j$, lo que es imposible, pues $y \in W$ pero $u_j \notin W$.

Los conjuntos acotados definidos anteriormente adquieren ahora, en un espacio normado, un aspecto más cercano. Un conjunto de un espacio normado es acotado si la norma está acotada en ese conjunto.

Las aplicaciones continuas en un e.v.t. tienen la siguiente propiedad.

Proposición 2.4.8 Sean L, M dos espacios vectoriales topológicos. Una aplicación lineal $T : L \rightarrow M$ es continua si y solo si es continua en 0.

Obviamente si es continua es continua en 0. Supongamos que es continua en $x = 0$. Para todo entorno de $0 \in M$, U , existe un entorno de $0 \in L$, V , tal que $T(V) \subset U$. Sea W un entorno de $Tx \in M$. Entonces $-Tx + W$ es un entorno de $0 \in M$, por lo que existe un entorno V' de $0 \in L$ tal que $T(V') \subset -Tx + W$. Pero entonces $x + V'$ es un entorno de x y $T(x + V') \subset W$.

Las aplicaciones lineales continuas entre espacios normados pueden caracterizarse en términos de la norma.

Proposición 2.4.9 Sean L, M dos espacios vectoriales normados. Una aplicación lineal $T : L \rightarrow M$ es continua si y solo si existe $c > 0$ tal que:

$$\|Tx\| \leq c\|x\|, \quad x \in L$$

Si T es continua, lo es en $x = 0$. Por tanto existen bolas en L y M de centro 0, tal que $T(B_L(0, r)) \subset B_M(0, r')$, es decir $\|T(x)\| \leq r'$ si $\|x\| \leq r$. Pero, para todo $x \in L$ ($x \neq 0$), $\frac{r}{2\|x\|}x$ está en $B_L(0, r)$, luego

$$\left\| T \left(\frac{r}{2\|x\|} x \right) \right\| \leq r'$$

es decir:

$$\frac{r}{2\|x\|} \|T(x)\| \leq r' \Rightarrow \|T(x)\| \leq \frac{2r'}{r} \|x\|$$

Es inmediato probar que la aplicación T es continua en 0 si se cumple esa acotación. Se dice que la aplicación lineal está acotada.

La proposición anterior permite definir la norma de una aplicación lineal continua entre espacios normados.

Definición 2.4.7 Sean L, M dos espacios normados y $u : L \rightarrow M$ una aplicación lineal continua. Entonces, el número real definido por:

$$\|u\| = \sup\{\|u(x)\| : \|x\| \leq 1\}$$

es finito y se tiene una norma en el espacio de las aplicaciones lineales continuas de L en N , $\mathcal{L}(L, N)$

Si M es un espacio de Banach, entonces $\mathcal{L}(L, M)$ es también un espacio de Banach. Para el caso en que $M = \mathbf{K}$, se tiene el espacio de las formas lineales continuas de L , es decir, lo que se llama el dual topológico de L . En este caso, $L' = \mathcal{L}(L, \mathbf{K})$ es un espacio de Banach (aunque L sea solo un espacio normado). Téngase en cuenta que en el caso de dimensión infinita hay formas lineales que no son continuas, por lo que el dual algebraico no coincide con el espacio de formas lineales continuas (dual topológico).

En el espacio de las aplicaciones continuas entre dos espacios de Banach se pueden definir diferentes topologías. Además de la de la norma (que se suele llamar topología de operadores uniforme), hay dos topologías particularmente importantes, relacionadas con la acción de la aplicación. Una de ellas se llama la topología de operadores fuerte. Se define como la más débil de las topologías que hace que las evaluaciones (es decir, las aplicaciones $\mathbf{T}_x : \mathcal{L}(L, M) \rightarrow M$, $\mathbf{T}_x(T) = Tx$) sean continuas. Lo que implica, una de sus propiedades más inmediatas, que una sucesión T_n de operadores converge a un operador T en esta topología si y solo si las sucesiones de vectores de M , $\{T_n(x)\}$ convergen a Tx para todo x de L (por supuesto en la norma de M).

La última topología que vamos a mencionar es la topología de operadores débil. Como anteriormente, se define como la más débil de las que hacen que las aplicaciones $\mathbf{T}_{x,f}$ de $\mathcal{L}(L, M)$ en \mathbf{C} , definidas para cada $x \in L$ y $f \in \mathcal{L}(M)$ por $\mathbf{T}_{x,f}(T) = f(Tx)$, sean continuas. Como antes, esto implica que una sucesión de operadores $\{T_n\}$ converge en esta topología a un operador T si y solo si $f(T_n x)$ converge a $f(Tx)$ (en \mathbf{C}) para cada $x \in L$ y $f \in \mathcal{L}(M)$.

En todo espacio de Banach se puede definir una topología (no la de la norma), como la más débil que hace que los funcionales lineales acotados sean continuos (es decir, todos los funcionales continuos en la topología de la norma son continuos en esta topología). Una sucesión en el espacio $\{x_n\}$ converge en la topología débil si las sucesiones $f(x_n)$ lo hacen en \mathbf{C} para todo funcional acotado. Se puede probar que un funcional lineal es continuo débilmente si y solo si lo es en norma. La topología débil no es metrizable si el espacio es de dimensión infinita.

Como ejemplo de la potencia de los resultados que se obtienen con estos conceptos, enunciaremos a continuación el teorema de Hahn-Banach, uno de los pilares del análisis funcional. Este teorema establece que, bajo ciertas condiciones, existen formas lineales continuas no triviales sobre un espacio vectorial. Veamos primeramente un lema.

Proposición 2.4.10 *Si L es un e.v.t. real de Hausdorff de dimensión mayor o igual que 2 y B es un conjunto abierto de L que no contiene al vector 0, entonces existe un subespacio de dimensión 1 que no corta a B .*

Una formulación geométrica de este teorema es la siguiente:

Teorema 2.4.3 (Mazur) *Sea L un e.v.t. y M una variedad lineal en M . Sea A un subconjunto de L , convexo, abierto y no vacío, tal que su intersección con M sea el vacío. Entonces existe un hiperplano cerrado de L que contiene a M y que no corta a A .*

La demostración no es trivial en absoluto, pero no requiere más conocimientos que los expuestos en estas notas (aparte de algún resultado de la teoría de conjuntos). Supongamos primero que el espacio es real. Podemos tomar un subespacio en vez de una variedad lineal (basta hacer una traslación). Consideremos los subespacios cerrados de L que contienen al subespacio M y que no cortan al abierto A . Se trata de una clase de conjuntos, \mathcal{M} , no vacía, pues el cierre de M (que es un subespacio cerrado que contiene a M) está en A . Esta clase se puede ordenar por inclusión. Consideremos un subconjunto del conjunto \mathcal{M} que esté totalmente ordenado. Tiene un supremo (que es el cierre de la unión) y por el Lema de Zorn existe un elemento maximal. Sea H_0 . Al ser cerrado, el espacio cociente es de Hausdorff, de dimensión mayor o igual que 1 (porque A es no vacío). Supongamos que fuera de dimensión mayor o igual que 2. La proyección canónica de L sobre L/H_0 es abierta, por lo que $\phi^{-1}(B)$ es un abierto y convexo, que no contiene al vector 0. H_0 no corta a A . Entonces, existe un subespacio unidimensional que no corta a B , con lo que la imagen inversa de este subespacio mediante ϕ es un espacio vectorial de L que no corta a A y contiene a H_0 (siendo distinto de él). Esto es contradictorio con el hecho de que H_0 sea el supremo y por tanto la dimensión de L/H_0 es igual a 1. Por tanto H_0 es un hiperplano cerrado que cumple el teorema.

En el caso complejo se usa simplemente la complejización de un espacio real de L/H_0

Como resultado de este teorema, se tiene que en un e.v.t. que contenga un subconjunto abierto convexo y no vacío existen formas lineales continuas no triviales (y viceversa).

La forma de este teorema para espacios normados es la siguiente.

Teorema 2.4.4 (Hahn-Banach) *Sea $(L, \|\cdot\|)$ un espacio normado y M un subespacio. Sea f una forma lineal que verifique $|f(x)| \leq \|x\|$ para todo $x \in M$ (por tanto continua). Entonces f admite una extensión (lineal), f_1 , a todo L con la misma acotación $|f_1(x)| \leq \|x\|$ para todo x en L (por tanto, continua con la misma norma).*

Veamos a continuación otro de los teoremas fundamentales del análisis funcional, el teorema de Banach-Steinhaus. Sean L, M dos espacio normados y $A \subset \mathcal{L}(L, M)$. Se dice que A es un conjunto puntualmente acotado si para cada $x \in A$ el conjunto

$$\{T(x) : T \in A\}$$

es un conjunto acotado. Todo conjunto acotado en norma es puntualmente acotado. El problema, que resuelve el principio de la acotación uniforme o teorema de Banach-Steinhaus, es el inverso.

Teorema 2.4.5 (Banach-Steinhaus) *Sea L un espacio de Banach y M un espacio normado. Un subconjunto de $\mathcal{L}(L, M)$ es acotado en norma si y solo si es puntualmente acotado.*

Para acabar esta introducción enunciaremos dos teoremas que cierran el conjunto de teoremas fundamentales del análisis funcional.

Teorema 2.4.6 (Aplicación abierta) *Sean L, M dos espacios de Banach, y $T \in \mathcal{L}(L, M)$. Si T es sobreyectiva, entonces T es abierta (es decir transforma abiertos en abiertos). Por tanto, si T es también inyectiva es un isomorfismo (topológico).*

Teorema 2.4.7 (Grafo cerrado) *Sean L, M dos espacios de Banach, y $T \in \mathcal{L}(L, M)$. Si el grafo de T , $(G_T = \{(x, T(x)) : x \in L\} \subset L \times M)$, es cerrado en $L \times M$ entonces T es acotado.*

Capítulo 3

Espacios de Banach

Estudiaremos en este capítulo espacios vectoriales reales o complejos, de dimensión finita o infinita dotados de una norma y en especial espacios normados completos (espacios de Banach).

3.1. Espacios normados de dimensión finita

Sea $\mathbf{K} = \mathbf{R}, \mathbf{C}$. Consideremos los espacios vectoriales \mathbf{K}^n y en ellos la siguiente norma:

$$\|x\|_2 = \left(\sum_{i=1}^n |x_i|^2 \right)^{\frac{1}{2}}, \quad x \in \mathbf{K}^n$$

Demostraremos que se trata de una norma. La única propiedad difícil es la tercera, la desigualdad triangular (Minkovski). Sean $x, y \in \mathbf{K}^n$.

$$\begin{aligned} \|x + y\|^2 &= \sum_{i=1}^n |x_i + y_i|^2 = \sum_{i=1}^n (\bar{x}_i + \bar{y}_i)(x_i + y_i) \\ &= \|x\|^2 + \|y\|^2 + 2 \sum_{i=1}^n \operatorname{Re}(\bar{x}_i y_i) \end{aligned}$$

pero está claro que

$$\operatorname{Re}(\bar{x}_i y_i) \leq |\bar{x}_i y_i|$$

así que solo tenemos que usar la desigualdad de Cauchy-Schwarz (Hölder):

$$\sum_{i=1}^n |\bar{x}_i y_i| \leq \left(\sum_{i=1}^n |x_i|^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^n |y_i|^2 \right)^{\frac{1}{2}}$$

pues entonces

$$\|x + y\|^2 \leq \|x\|^2 + \|y\|^2 + 2\|x\| \|y\| = (\|x\| + \|y\|)^2$$

Pero hay otras normas en \mathbf{K}^n que (como sabemos) llevan a la misma topología.

$$\|x\|_1 = \sum_{i=1}^n |x_i|$$

es una norma en \mathbf{K}^n . Como antes, la única dificultad se encuentra en la desigualdad triangular. Pero en este caso es una consecuencia inmediata de la desigualdad triangular para el módulo de un número real o complejo. La expresión

$$\|x\|_\infty = \max\{|x_i| : i = 1, 2, \dots, n\}$$

es también una norma. La desigualdad triangular es inmediata de las propiedades del máximo. Que estas tres normas son equivalentes proviene de las siguientes desigualdades:

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty$$

$$\|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty$$

En general, la siguiente aplicación es una norma:

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}$$

donde $p \geq 1$. Como en el caso $p = 2$, la dificultad mayor está en la desigualdad triangular. Para demostrarla, utilizaremos la desigualdad de Hölder (que para $p = q = 2$ es la de Cauchy-Schwarz).

Proposición 3.1.1 (Desigualdad de Hölder) *Sea $x, y \in \mathbf{K}^n$ y $p, q > 1$ tales que:*

$$\frac{1}{p} + \frac{1}{q} = 1$$

(se dice que p y q son conjugados). Entonces

$$\sum_{i=1}^n |x_i| |y_i| \leq \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^n |y_i|^q \right)^{\frac{1}{q}}$$

Podemos demostrarlo de la forma siguiente. Está claro que basta probarlo para dos vectores a, b con coordenadas mayores que 0. Consideramos los vectores:

$$\tilde{a} = \frac{a}{\left(\sum_{i=1}^n a_i^p \right)^{\frac{1}{p}}}, \quad \tilde{b} = \frac{b}{\left(\sum_{i=1}^n b_i^q \right)^{\frac{1}{q}}}$$

que verifican:

$$\sum_{i=1}^n \tilde{a}_i^p = 1, \quad \sum_{i=1}^n \tilde{b}_i^q = 1$$

Como las componentes de \tilde{a} e \tilde{b} son positivas, se tiene:

$$\tilde{a}_i = e^{r_i/p}, \quad \tilde{b}_i = e^{s_i/q}, \quad i = 1, \dots, n$$

para ciertos números reales r_i, s_i . Pero la función exponencial es convexa, es decir, verifica:

$$f(\alpha r + (1 - \alpha)s) \leq \alpha f(r) + (1 - \alpha)f(s), \quad 0 \leq \alpha \leq 1$$

(ver en la sección 3.3 algunos resultados elementales sobre funciones convexas). Sea $\alpha = 1/p$, lo que implica $1 - \alpha = 1/q$. Entonces:

$$e^{\frac{r_i}{p} + \frac{s_i}{q}} \leq \frac{1}{p} e^{r_i} + \frac{1}{q} e^{s_i}$$

y sustituyendo los valores de r_i y s_i :

$$\tilde{a}_i \tilde{b}_i \leq \frac{1}{p} \tilde{a}_i^p + \frac{1}{q} \tilde{b}_i^q$$

Si ahora sumamos en i , se tiene:

$$\sum_{i=1}^n \tilde{a}_i \tilde{b}_i \leq \frac{1}{p} \sum_{i=1}^n \tilde{a}_i^p + \frac{1}{q} \sum_{i=1}^n \tilde{b}_i^q = \frac{1}{p} + \frac{1}{q} = 1$$

Sustituyendo los valores de \tilde{a}_i y \tilde{b}_i :

$$\sum_{i=1}^n \frac{a_i}{\left(\sum_{i=1}^n a_i^p\right)^{\frac{1}{p}}} \frac{b_i}{\left(\sum_{i=1}^n b_i^q\right)^{\frac{1}{q}}} \leq 1$$

que es la desigualdad buscada. Evidentemente, cuando $p = q = 1/2$ se obtiene la desigualdad de Cauchy-Schwarz. En realidad, la desigualdad de Cauchy-Schwarz es la siguiente:

$$\left| \sum_{i=1}^n \tilde{x}_i y_i \right| \leq \|x\|_2 \|y\|_2$$

pero es una consecuencia inmediata de la anterior:

$$\left| \sum_{i=1}^n \tilde{x}_i y_i \right| \leq \sum_{i=1}^n |x_i| |y_i|$$

En el caso general ($p \geq 1$), la desigualdad triangular es:

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p$$

Para demostrarlo (para $p > 1$, si $p = 1$ es inmediata) aplicamos la desigualdad de Hölder:

$$\begin{aligned} \sum_{i=1}^n |x_i| |x_i + y_i|^{p-1} &\leq \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^n |x_i + y_i|^{(p-1)q} \right)^{\frac{1}{q}} \\ \sum_{i=1}^n |y_i| |x_i + y_i|^{p-1} &\leq \left(\sum_{i=1}^n |y_i|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^n |x_i + y_i|^{(p-1)q} \right)^{\frac{1}{q}} \end{aligned}$$

Si ahora sumamos las desigualdades, se obtiene ($(p-1)q = p$):

$$\begin{aligned} \sum_{i=1}^n (|x_i| + |y_i|) |x_i + y_i|^{p-1} &\leq \\ \left[\left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^n |y_i|^p \right)^{\frac{1}{p}} \right] &\left(\sum_{i=1}^n |x_i + y_i|^p \right)^{\frac{1}{q}} \end{aligned}$$

y por lo tanto:

$$\begin{aligned} \sum_{i=1}^n |x_i + y_i|^p &\leq \sum_{i=1}^n (|x_i| + |y_i|) |x_i + y_i|^{p-1} \leq \\ \left[\left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^n |y_i|^p \right)^{\frac{1}{p}} \right] &\left(\sum_{i=1}^n |x_i + y_i|^p \right)^{\frac{1}{q}} \end{aligned}$$

Finalmente

$$\left(\sum_{i=1}^n |x_i + y_i|^p \right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^n |y_i|^p \right)^{\frac{1}{p}}$$

Para $0 < p < 1$ la desigualdad no se cumple. Lo que ahora se tiene es la siguiente desigualdad. El conjugado de p es ahora negativo, $q < 0$. Entonces (si $a_i \geq 0$ y $b_i > 0$):

$$\sum_{i=1}^n a_i b_i \geq \left(\sum_{i=1}^n a_i^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^n b_i^q \right)^{\frac{1}{q}}$$

Sea

$$u_i = (a_i b_i)^p, \quad v_i = b_i^{-p}$$

Aplicamos la desigualdad de Hölder a u, v con coeficientes conjugados $1/p$ y $-q/p$ (que son mayores que 1):

$$\sum_{i=1}^n u_i v_i \leq \left(\sum_{i=1}^n u_i^{1/p} \right)^p \left(\sum_{i=1}^n v_i^{-q/p} \right)^{-\frac{p}{q}}$$

Sustituyendo las expresiones de u_i y v_i , tenemos:

$$\sum_{i=1}^n a_i^p \leq \left(\sum_{i=1}^n a_i b_i \right)^p \left(\sum_{i=1}^n b_i^q \right)^{-\frac{p}{q}}$$

Elevándolo todo a la potencia $1/p$ y despejando adecuadamente:

$$\left(\sum_{i=1}^n a_i^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^n b_i^q \right)^{\frac{1}{q}} \leq \sum_{i=1}^n a_i b_i$$

como queríamos probar. Comprobemos con un simple ejemplo que la desigualdad triangular no se cumple cuando $p = 1/2$. Sea $x = (1, 0, \dots, 0)$, $y = (0, 1, 0, \dots, 0)$. Se tiene:

$$\|x\|_{1/2} = 1, \quad \|y\|_{1/2} = 1$$

Pero, $x + y = (1, 1, 0, \dots, 0)$ y se tiene

$$\|x + y\|_{1/2} = (1^{1/2} + 1^{1/2})^2 = 4$$

En lugar de la desigualdad triangular lo que se verifica es la siguiente relación:

$$\left(\sum_{i=1}^n (a_i + b_i)^p \right)^{\frac{1}{p}} \geq \left(\sum_{i=1}^n a_i^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^n b_i^p \right)^{\frac{1}{p}}$$

La demostración es similar a la que se ha hecho en el caso $p > 1$. Lo que ocurre ahora es que la desigualdad de Hölder se ve sustituida por la desigualdad que acabamos de demostrar.

Sin embargo, si $0 < p < 1$

$$d(x, y) = \|x - y\|_p^p$$

es una distancia aunque no proceda de una norma.

3.2. Espacios de sucesiones l^p

Pasemos ahora al caso de dimensión infinita. Consideremos el espacio de las sucesiones para las que la suma de los módulos a la potencia $p > 0$ de sus términos es convergente:

$$l^p = \left\{ \{x_n\} : x_n \in \mathbf{K}, \sum_{n=1}^{\infty} |x_n|^p < \infty \right\}$$

Consideremos en primer lugar el caso $p \geq 1$. Demostraremos en primer lugar que se trata de un espacio vectorial. El único problema es probar que la suma de dos elementos de este conjunto está también en el conjunto. Para ello, comprobemos que la siguiente expresión es una norma en l^p .

$$\|x\|_p = \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{\frac{1}{p}}$$

Como siempre la desigualdad triangular es la única complicación. Pero podemos repetir los pasos que dimos en el caso de dimensión finita y llegar a

$$\left(\sum_{i=1}^n |x_i + y_i|^p \right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^n |y_i|^p \right)^{\frac{1}{p}}$$

Pero

$$\left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{\frac{1}{p}} < \infty$$

Por lo tanto el miembro de la izquierda de la desigualdad está acotado para todo n y la desigualdad es válida cuando $n \rightarrow \infty$. De esta forma se prueba que l^p es un espacio vectorial y además $\|x\|_p$ es una norma.

Los espacios l^p con $p \geq 1$ son espacios normados de dimensión infinita (no numerable). Si $0 < p < 1$, l^p no es un espacio normado, aunque sí un espacio vectorial y además un espacio métrico (con una distancia que no proviene de una norma).

En efecto,

$$d_p(x, y) = \sum_{n=1}^{\infty} |x_n - y_n|^p, \quad 0 < p < 1$$

es una distancia, pues la desigualdad de Minkovski se cumple:

$$d_p(x, y) \leq d_p(x, z) + d_p(z, y)$$

Vamos a demostrar ahora que l^p es un espacio de Banach. Sea $\{x^n\}$ una sucesión en l^p , con $1 \leq p$. Es decir:

$$x^n = (x_1^n, \dots, x_k^n, \dots)$$

Supongamos que es de Cauchy. Para todo $\epsilon > 0$ existe n_0 tal que

$$\|x^n - x^m\|_p < \epsilon, \quad n, m \geq n_0$$

Esto implica que cada una de las sucesiones $\{x_k^n\}_{n=1}^{\infty}$ es de Cauchy en \mathbf{K} , pues

$$|x_k^n - x_k^m| \leq \|x^n - x^m\|_p < \epsilon$$

luego convergen al ser \mathbf{K} completo. Sea $x = (x_1, x_2, \dots)$ el vector de los límites. Hay que probar que está en l^p y que es el límite de la sucesión. Al ser $\{x^n\}$ de Cauchy, para todo ϵ , por ejemplo $\epsilon = 1$, existe n_0 tal que para todo $n, m \geq n_0$ y para todo N , se tiene:

$$\sum_{k=1}^N |x_k^n - x_k^m|^p \leq \|x^n - x^m\|_p^p < 1$$

En la suma finita podemos hacer $n \rightarrow \infty$ y se tiene una acotación para todo $m \geq n_0$ y para todo N :

$$\sum_{k=1}^N |x_k - x_k^m|^p \leq 1$$

y por lo tanto es cierta cuando $N \rightarrow \infty$:

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

De esta forma hemos demostrado que $x - x^m$ está en l^p cuando $m \geq n_0$. Pero $x^m \in l^p$, luego $x \in l^p$. Probemos ahora que es el límite de la sucesión. Hacemos lo mismo que antes con ϵ arbitrario:

$$\sum_{k=1}^{\infty} |x_k - x_k^m|^p \leq \epsilon, \quad m \geq n_0$$

luego

$$\|x - x^m\|_p^p < \epsilon$$

La demostración es válida para cualquier p . Por lo tanto todo l^p , $0 < p$, es completo (si es normado es de Banach).

Los espacios l^p están relacionados unos con otros, de acuerdo con las inclusiones:

$$l^r \subset l^s, \quad 0 < r < s$$

En efecto, sea $x = \{x_n\} \in l^r$, $x \neq 0$. Entonces

$$\frac{|x_n|}{\|x\|_r} \leq 1 \Rightarrow \frac{|x_n|^r}{\|x\|_r^r} \geq \frac{|x_n|^s}{\|x\|_r^s}$$

Sumando:

$$\frac{\sum_{n=1}^{\infty} |x_n|^r}{\|x\|_r^r} \geq \frac{\sum_{n=1}^{\infty} |x_n|^s}{\|x\|_r^s}$$

y por lo tanto $x \in l^s$. Como, además:

$$\frac{\sum_{n=1}^{\infty} |x_n|^r}{\|x\|_r^r} = 1$$

se tiene la desigualdad de Jensen:

$$\|x\|_s \leq \|x\|_r$$

El conjunto c_0 de las sucesiones que tienden a cero es también un espacio de Banach (que contiene a cualquier l^p y que está contenido en l^∞ , el espacio de las sucesiones acotadas).

El conjunto de las sucesiones de tipo finito (es decir con un número finito de elementos distintos de cero, que es lo mismo que el conjunto de las funciones continuas con soporte compacto de \mathbf{N} en \mathbf{K}) está contenido en cualquier l^p y contiene al espacio \mathbf{K}^n para cualquier n . Pero no es cerrado en ningún l^p . Para demostrarlo basta ver que es denso y distinto del espacio total. Para ver que es distinto, basta considerar sucesiones como $1/n$ en l^p si $p > 1$ y $1/n^2$ si $p \geq 1$. Para ver que es denso, se hace con $1 \leq p < \infty$. Pero además este espacio tiene un cierre en la norma infinito que está contenido en c_0 y entonces no puede ser denso en l^∞ . La dimensión de este espacio es \aleph_0 . Los espacios l^p , con $1 \leq p < \infty$ y c_0 son separables (porque el espacio de las sucesiones de tipo finito es denso en l^p (con su norma) y en c_0 (con la norma del supremo)). Sin embargo l^∞ no es separable.

Los espacios l^p guardan unas relaciones de dualidad. El dual de l^p es l^q , con p, q conjugados. El dual de l^1 es l^∞ , pero l^1 no es el dual de l^∞ , sino de c_0 .

3.3. Espacios L^p

Dado cualquier conjunto X , se puede considerar las funciones de X en \mathbf{K} acotadas. Llamaremos $B(X)$ a este conjunto de funciones:

$$B(X) = \{f : X \rightarrow \mathbf{K} : f \text{ acotada}\}$$

El conjunto $B(X)$ es un espacio vectorial sobre \mathbf{K} en el que se puede definir una norma:

$$\|f\|_\infty = \sup\{|f(x)| : x \in X\}$$

Se llama la norma de la convergencia uniforme. Es además un espacio de Banach. Para probarlo, consideremos una sucesión $\{f_n\}$ de funciones que sea de Cauchy. Por tanto

$$\sup_{x \in X} |f_n(x) - f_m(x)| \rightarrow 0, \quad n \rightarrow \infty$$

Como en el caso de las sucesiones, es sencillo ver que $\{f_n(x)\}$ es también una sucesión de Cauchy en \mathbf{K} . Al ser \mathbf{K} completo, tiene límite. Sea $f(x)$ el límite para cada $x \in X$. Veamos que f es una función

acotada y que es el límite de la sucesión en el sentido de la convergencia uniforme. El razonamiento es similar al usado en las sucesiones. Al ser $\{f_n\}$ de Cauchy,

$$\|f_n - f_m\|_\infty \leq 1, \quad n, m \geq n_0$$

y por lo tanto, para cualquier $x \in X$:

$$|f_n(x) - f_m(x)| \leq 1, \quad n, m \geq n_0$$

Cuando $n \rightarrow \infty$, se tiene:

$$|f(x) - f_m(x)| \leq 1, \quad m \geq n_0$$

Por lo tanto, la función $f - f_m$ está acotada y en $B(X)$ que es un espacio vectorial, luego $f \in B(X)$. Veamos que es el límite. Basta tomar un ϵ arbitrario y repetir el razonamiento:

$$|f(x) - f_m(x)| \leq \epsilon, \quad m \geq n_0$$

implica que

$$\|f - f_m\|_\infty \leq \epsilon, \quad m \geq n_0$$

Podemos demostrar ahora que l^∞ es un espacio de Banach. Basta tomar $X = \mathbf{N}$ y la medida discreta. Además se puede probar que

$$\lim_{p \rightarrow \infty} \|x\|_p = \|x\|_\infty$$

Como sabemos, cualquier subespacio cerrado de un espacio de Banach es un espacio de Banach. Sea X un espacio topológico compacto. Entonces, el conjunto de funciones continuas de X en \mathbf{K} es un subespacio del conjunto de funciones acotadas (pues X es compacto). Además es un subespacio cerrado (en la norma del supremo de $B(X)$), pues la norma es la de la convergencia uniforme y por tanto el límite de una sucesión de funciones continuas es una función continua. Luego

$$C(X) = \{f : X \rightarrow \mathbf{K} : f \text{ continua}\}$$

es un espacio de Banach (en la norma del supremo).

Consideremos un espacio de medida, (X, \mathcal{S}, μ) , (que será en lo que respecta a las aplicaciones la recta real y la medida de Lebesgue). Si f es una función (con valores en \mathbf{R} o \mathbf{C}) medible de este espacio, se tiene que $|f|^p$, $p > 0$ es también medible. \mathcal{L}^p es el conjunto de las funciones medibles tales que $|f|^p$ es integrable, con $0 < p < \infty$.

El conjunto \mathcal{L}^p es un espacio vectorial. Para ver que la suma de dos funciones de \mathcal{L}^p está en este espacio, se puede usar la siguiente desigualdad (ver la Nota más adelante):

$$|a + b|^p \leq 2^p(|a|^p + |b|^p)$$

Entonces

$$\int |f + g|^p d\mu \leq \int 2^p(|f|^p + |g|^p) d\mu < \infty$$

Nota. Funciones convexas

Las funciones convexas dan lugar a desigualdades interesantes, por lo que daremos aquí una breve noticia de sus propiedades. Se dice que una función real de variable real es convexa en un intervalo (a, b) si

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y), \quad x, y \in (a, b), \quad 0 < \alpha < 1$$

Las funciones convexas verifican la desigualdad de Jensen, que no es más que la generalización de la definición a n puntos:

$$f\left(\sum_{i=1}^n \alpha_i x_i\right) \leq \sum_{i=1}^n \alpha_i f(x_i), \quad x_i \in (a, b), \quad \sum_{i=1}^n \alpha_i = 1$$

Se puede demostrar la siguiente propiedad, relativa a la relación entre funciones convexas y derivadas:

Si existe la derivada primera de una función f y es no decreciente en un intervalo abierto, entonces f es convexa en ese intervalo.

Para probarlo basta usar el teorema del valor medio. Sea $y > x$, $0 \leq \alpha \leq 1$.

$$\begin{aligned} & \alpha f(x) + (1 - \alpha)f(y) - f(\alpha x + (1 - \alpha)y) \\ &= \alpha[f(x) - f(\alpha x + (1 - \alpha)y)] + (1 - \alpha)[f(y) - f(\alpha x + (1 - \alpha)y)] \end{aligned}$$

Pero, por el teorema del valor medio, al ser f derivable:

$$\begin{aligned} f(\alpha x + (1 - \alpha)y) - f(x) &= (1 - \alpha)(x - y)f'(c), \quad c \in (x, \alpha x + (1 - \alpha)y) \\ f(y) - f(\alpha x + (1 - \alpha)y) &= \alpha(x - y)f'(d), \quad d \in (\alpha x + (1 - \alpha)y, y) \end{aligned}$$

Por tanto,

$$\begin{aligned} & \alpha f(x) + (1 - \alpha)f(y) - f(\alpha x + (1 - \alpha)y) = \\ & -\alpha(1 - \alpha)(x - y)f'(c) + (1 - \alpha)\alpha(x - y)f'(d) = \\ & \alpha(1 - \alpha)(x - y)(f'(d) - f'(c)) \geq 0 \end{aligned}$$

Como consecuencia, si existe la derivada segunda y es no negativa en un intervalo abierto, la función es convexa en ese intervalo.

Como aplicación, si $a, b \geq 0$, $p > 1$ se tiene:

$$(a + b)^p \leq 2^{p-1}(a^p + b^p)$$

Basta usar el hecho de que $h(x) = x^p$ es convexa (su derivada segunda $p(p-1)x^{p-2}$ en el caso en que $p > 1$ es claramente positiva en cualquier intervalo abierto de la semirrecta positiva). Entonces:

$$(a + b)^p = 2^p \left(\frac{1}{2}a + \frac{1}{2}b \right)^p \leq 2^p \left(\frac{1}{2}a^p + \frac{1}{2}b^p \right) = 2^{p-1}(a^p + b^p)$$

De manera evidente:

$$(a + b)^p \leq 2^p(a^p + b^p)$$

pues $2^{p-1} \leq 2^p$. Pero esta desigualdad es también correcta cuando $0 < p < 1$. Téngase en cuenta que el anterior razonamiento no puede aplicarse cuando $0 < p < 1$, pues entonces x^p no es convexa. Por ejemplo:

$$(1 + 4)^{\frac{1}{2}} \leq 2^{-\frac{1}{2}}(1^{\frac{1}{2}} + 4^{\frac{1}{2}}) \Rightarrow 10^{1/2} \leq 3$$

lo que no es correcto. Por tanto si esta segunda desigualdad se cumple para $0 < p < 1$ habrá que demostrarlo de otra forma.

Suponiendo que $a \neq 0$ (si $a = 0$ es trivialmente cierto), si $t = b/a$, lo que tenemos que probar es que:

$$\frac{1+t}{2} \leq (1+t^p)^{\frac{1}{p}}, \quad t \geq 0, \quad 0 < p < 1$$

La desigualdad es correcta (estrictamente) para $t = 0$. Si probamos que la derivada de la función de la derecha es mayor que $1/2$ (que es la derivada de la función de la izquierda), habremos probado la desigualdad. Derivando:

$$\frac{1}{p}(1+t^p)^{\frac{1}{p}-1} p t^{p-1} = (1+t^p)^{\frac{1-p}{p}} t^{p-1} = \left(1 + \frac{1}{t^p}\right)^{\frac{1-p}{p}} > 1$$

En este espacio vectorial de funciones \mathcal{L}^p , se verifica la desigualdad de Hölder. Si $\frac{1}{p} + \frac{1}{q} = 1$, $1 < p, q < \infty$, entonces

$$\int |fg| d\mu \leq \left(\int |f|^p d\mu \right)^{\frac{1}{p}} \left(\int |g|^q d\mu \right)^{\frac{1}{q}}, \quad f \in \mathcal{L}^p, \quad g \in \mathcal{L}^q$$

Para probarlo, veamos en primer lugar que si a, b son dos números reales no negativos, entonces, para todo $0 < \alpha < 1$ se tiene:

$$a^\alpha b^{1-\alpha} \leq \alpha a + (1-\alpha)b$$

Supongamos que $a, b > 0$. Entonces, la función:

$$f(x) = 1 - \alpha + \alpha x - x^\alpha$$

es positiva para todo $x > 0$. La función tiene solo un mínimo en la semirrecta $x \geq 0$ ($f(0) = 1 - \alpha > 0$), pues su derivada es:

$$f'(x) = \alpha(1 - x^{\alpha-1}) = 0 \Rightarrow x = 1$$

Pero en $x = 1$, $f(1) = 0$ y por tanto se tiene la desigualdad buscada.

Para demostrar la desigualdad de Hölder, supongamos que f y g no son nulas casi doquiera (en caso contrario la desigualdad es trivial). Sea

$$a = \frac{|f(x)|^p}{\int |f|^p d\mu}, \quad b = \frac{|g(x)|^q}{\int |g|^q d\mu}$$

Aplicando la desigualdad anterior con $\alpha = 1/p$, tenemos:

$$\frac{|f(x)g(x)|}{\left(\int |f|^p d\mu\right)^{\frac{1}{p}} \left(\int |g|^q d\mu\right)^{\frac{1}{q}}} \leq \frac{1}{p} \frac{|f(x)|^p}{\int |f|^p d\mu} + \frac{1}{q} \frac{|g(x)|^q}{\int |g|^q d\mu}$$

Debido a esta acotación, $fg \in \mathcal{L}^1$, con lo que podemos integrar y obtener:

$$\frac{\int |f(x)g(x)| d\mu}{\left(\int |f|^p d\mu\right)^{\frac{1}{p}} \left(\int |g|^q d\mu\right)^{\frac{1}{q}}} \leq \frac{1}{p} + \frac{1}{q} = 1$$

y la desigualdad de Hölder queda probada. Usando esta desigualdad podemos probar la de Minkovski:

$$\left(\int |f(x) + g(x)| d\mu\right)^{\frac{1}{p}} \leq \left(\int |f|^p d\mu\right)^{\frac{1}{p}} + \left(\int |g|^p d\mu\right)^{\frac{1}{p}}$$

Supongamos que $p > 1$ (pues es inmediato en el caso $p = 1$). Sea q el coeficiente conjugado de p . Debido a la relación entre p y q se tiene que $|f + g|^{p-1} \in \mathcal{L}^q$ (pues $(p-1)q = p$). Aplicando la desigualdad de Hölder a las funciones $|f||f + g|^{p-1}$ y $|g||f + g|^{p-1}$, que están en \mathcal{L}^1 , se tiene:

$$\begin{aligned} \int |f||f + g|^{p-1} d\mu &\leq \left(\int |f|^p d\mu\right)^{\frac{1}{p}} \left(\int |f + g|^{(p-1)q} d\mu\right)^{\frac{1}{q}} \\ &= \left(\int |f|^p d\mu\right)^{\frac{1}{p}} \left(\int |f + g|^p d\mu\right)^{\frac{1}{q}} \end{aligned}$$

$$\begin{aligned} \int |g||f + g|^{p-1} d\mu &\leq \left(\int |g|^p d\mu\right)^{\frac{1}{p}} \left(\int |f + g|^{(p-1)q} d\mu\right)^{\frac{1}{q}} \\ &= \left(\int |g|^p d\mu\right)^{\frac{1}{p}} \left(\int |f + g|^p d\mu\right)^{\frac{1}{q}} \end{aligned}$$

y sumando

$$\int |f + g|^p d\mu \leq \left[\left(\int |f|^p d\mu\right)^{\frac{1}{p}} + \left(\int |g|^p d\mu\right)^{\frac{1}{p}} \right] \left(\int |f + g|^p d\mu\right)^{\frac{1}{q}}$$

que lleva a la desigualdad de Minkovski. La aplicación

$$f \rightarrow \left(\int |f|^p d\mu \right)^{\frac{1}{p}}$$

no es una norma, porque el hecho de que la integral sea cero no implica que la función lo sea, ya que puede ser cero casi doquiera. La idea es introducir una relación de equivalencia en el conjunto \mathcal{L}^p de forma que dos funciones están en la misma clase si son iguales casi doquiera. De esta forma se construye el espacio de clases L^p y la aplicación anterior es una norma en este espacio. Hay que tener en cuenta esta circunstancia en las aplicaciones de esta teoría. En la práctica seguiremos llamando a estas clases de funciones con el mismo nombre que a una función, aunque prestando atención a que todo lo que se diga de una de estas funciones será siempre salvo equivalencia, es decir, salvo una función que es cero casi doquiera. Por tanto, los espacios L^p son espacios normados para $1 \leq p < \infty$. Se puede demostrar que estos espacios son de Banach.

Teorema 3.3.1 *Sea $1 \leq p < \infty$. El espacio L^p es un espacio de Banach.*

Para demostrarlo, consideremos una sucesión de funciones en L^p que sea de Cauchy, y supondremos que:

$$\|f_{n+1} - f_n\|_p \leq \frac{1}{2^n}$$

(lo que siempre se puede conseguir considerando subsucesiones si es necesario). A partir de esta sucesión construimos otra sucesión por:

$$g_1 = 0, g_2 = |f_1|, g_3 = |f_1| + |f_2 - f_1|, \dots, g_n = |f_1| + |f_2 - f_1| + \dots + |f_n - f_{n-1}|, \dots$$

Esta sucesión es de funciones no negativas y creciente. Además:

$$\int |g_n|^p d\mu \leq \left(\|f_1\|_p + \sum_{i=2}^{\infty} \|f_i - f_{i-1}\|_p \right)^p \leq (\|f_1\|_p + 1)^p$$

Esto quiere decir (por el teorema de Levi, ver bibliografía) que la sucesión g_n converge casi doquiera a un función g cuya potencia p es integrable ($g \in L^p$). Además:

$$\begin{aligned} |f_{n+k}(x) - f_n(x)| &= \left| \sum_{i=n+1}^{n+k} (f_i(x) - f_{i-1}(x)) \right| \leq \sum_{i=n+1}^{n+k} |f_i(x) - f_{i-1}(x)| \\ &= g_{n+k}(x) - g_n(x) \end{aligned}$$

Por lo tanto, la sucesión $f_n(x)$ es de Cauchy y f_n converge puntualmente casi doquiera a una función f . Solo falta probar que esta función está en L^p y la sucesión f_n converge a ella en la norma de L^p . Pero

$$|f_n(x)| = |f_1(x) + \sum_{i=2}^n (f_i(x) - f_{i-1}(x))| \leq g_n(x) \leq g(x)$$

casi doquiera para todo n , luego $|f(x)| \leq g(x)$ y por tanto $f \in L^p$. Además:

$$|f(x) - f_n(x)| \leq 2g(x) \quad \text{c.d.}$$

Como $\lim_{n \rightarrow \infty} |f(x) - f_n(x)|^p = 0$, el límite de la norma de la diferencia es cero por el teorema de la convergencia dominada:

$$\lim_{n \rightarrow \infty} \|f - f_n\|_p = 0$$

Hay que tener en cuenta que el que una sucesión de funciones de L^p converja en norma a una función, no quiere decir que la convergencia sea puntual (incluso casi doquiera). También es posible construir sucesiones que convergen casi doquiera a una función en L^p pero que no convergen en norma.

La desigualdad (cierta para $1 \leq p < \infty$ y $a, b \geq 0$) $(a + b)^p \leq 2^{p-1}(a^p + b^p)$ y el lema de Fatou permiten demostrar la siguiente propiedad:

Proposición 3.3.1 *Sea f_n una sucesión de funciones en el espacio de Banach L^p y $f \in L^p$. Si $f_n(x)$ converge a $f(x)$ casi doquiera y la sucesión de normas $\|f_n\|_p$ converge a $\|f\|_p$, entonces f_n converge en norma a f ($\|f_n - f\| \rightarrow 0$).*

Hasta ahora nos hemos limitado en la construcción de estos espacios de funciones al intervalo $1 \leq p < \infty$. Pero es posible definir un espacio que llamaremos L^∞ . Para ello definimos primero el supremo de una función f . Se dice que K es una cota (esencial) de una función (de valores reales o complejos) f si $|f(x)| \leq K$ casi doquiera. De esta forma, una función está esencialmente acotada si tiene una cota (esencial). Llamaremos el supremo (esencial) al ínfimo de las cotas (esenciales), cuando la función está esencialmente acotada:

$$\|f\|_\infty = \inf\{K : |f(x)| \leq K, \text{ c.d.}\}$$

El conjunto de las funciones medibles esencialmente acotadas se llama \mathcal{L}^∞ y es un espacio vectorial. Como antes es necesario establecer una relación de equivalencia para poder tener una norma. El espacio de clases se llamará L^∞ y es un espacio vectorial normado con la norma del supremo (en lo que sigue prescindiremos de la palabra esencial). Este espacio es un espacio de Banach. La demostración es inmediata, basándose como es usual en el hecho de que las sucesiones de números reales (o complejos) $f_n(x)$ son de Cauchy si f_n lo es en L^∞ .

Entre las propiedades interesantes de estos espacio L^p , $1 \leq p < \infty$, podemos citar que las funciones escalón están contenidos en ellos y que incluso son densas en la norma correspondiente. Igual les pasa a las funciones continuas de soporte compacto (con algunas restricciones sobre el espacio y la medida, que obviamente cumplen \mathbf{R} y la medida de Lebesgue). Con estas condiciones la completación de los espacios de funciones continuas con soporte compacto son los espacios L^p (cuando se considera la norma $\|\cdot\|_p$). El conjunto $C([0, 1])$ con la norma p no es, de acuerdo con lo anterior, completo. La convergencia es más débil que la uniforme y al completarlo se obtienen los espacios L^p . Sin embargo, el conjunto de las funciones que tienden a cero en el infinito $C_0(X)$ (con X un espacio localmente compacto) es un subconjunto cerrado del conjunto de funciones acotadas (L^∞) y por lo tanto es un espacio de Banach.

Se tiene también un resultado que relaciona unos espacios con otros:

Proposición 3.3.2 *Si el espacio tiene medida finita y $1 \leq p < q \leq \infty$, entonces $L^q \subset L^p$ y $\|f\|_p \leq \|f\|_q$.*

La demostración usa el hecho de que la función constante igual a 1 está en todos los espacios L^p en este caso (porque la medida del espacio total es finita). Es evidente que $L^\infty \subset L^p$ para todo p con $1 \leq p$ (de nuevo porque el espacio total tiene medida finita). Si $q < \infty$, sea $r = q/p > 1$, y sea s su coeficiente conjugado. Toda función en L^q tiene su módulo a la potencia p en L^r . Por la desigualdad de Hölder el producto de esta función por la función 1 (que está en L^s) está en L^1 . Por lo tanto $f \in L^p$.

Pero si el espacio no tiene medida finita la situación no es tan simple. El siguiente ejemplo muestra una función que está en L^p pero en ningún otro L^q con $p \neq q$. Sea

$$f(x) = \frac{1}{x(1 + |\log x|)^2}$$

Probemos en primer lugar que esta función está en $L^1(0, \infty)$, aunque la función diverge cuando $x \rightarrow 0$ y el intervalo es infinito. La integral del valor absoluto (la función es positiva) es (haciendo $t = \log x$):

$$\begin{aligned} \int_0^\infty \frac{1}{x(1 + |\log x|)^2} dx &= \int_{-\infty}^\infty \frac{1}{(1 + |t|)^2} dt = \\ &= \int_{-\infty}^0 \frac{1}{(1 - t)^2} dt + \int_0^\infty \frac{1}{(1 + t)^2} dt = 1 + 1 = 2 \end{aligned}$$

Por lo tanto, la función $g(x) = (f(x))^{\frac{1}{p}}$ está en L^p , pues $|g(x)|^p = f(x)$ es integrable. Sin embargo, consideremos $g^q = (f(x))^{\frac{q}{p}}$ con $q \neq p$ y veamos si la integral ($\alpha = q/p$)

$$\int_0^\infty \frac{1}{x^\alpha(1 + |\log x|)^{2\alpha}} dx = \int_0^1 \frac{1}{x^\alpha(1 - \log x)^{2\alpha}} dx + \int_1^\infty \frac{1}{x^\alpha(1 + \log x)^{2\alpha}} dx$$

converge (si lo hiciera, $g \in L^q$). El integrando de la primera integral es singular en $x = 0$. De los criterios usuales de convergencia se tiene que si $\alpha > 1$ esa integral diverge (la singularidad que en $x = 0$ tiene $x^{-\alpha}$ es demasiado fuerte para la convergencia y el logaritmo no puede compensarla, aunque esté elevado a 2α). El punto singular para la segunda integral es ∞ y ahora el criterio es que la integral diverge si $\alpha < 1$ (en este caso, $x^{-\alpha}$ no tiende a cero suficientemente deprisa y la contribución del logaritmo tampoco es suficiente). En cualquier caso, para $\alpha \neq 1$ (es decir $p \neq q$) la integral diverge en uno u otro extremo y por tanto $g \notin L^q$.

Los espacios L^p y L^q correspondientes a coeficientes conjugados tienen unas relaciones interesantes cuando se estudian sus funcionales. Existe un teorema de Riesz (que estudiaremos en el caso en que $p = q = 2$) que asegura que cualquier funcional lineal acotado de L^p se puede representar por una función de L^q en el siguiente sentido:

$$F(g) = \int fg \, d\mu, \quad g \in L^p, f \in L^q$$

donde f viene fijada unívocamente por el funcional F . Sin embargo este no es el caso (en general) para los espacios L^1 y L^∞ .

Los espacios $L^p(\mathbf{R})$ y $L^q(\mathbf{R})$ son duales (con p, q conjugados, mayores que 1). Sin embargo, de forma similar a como ocurría con los espacios l^p , el dual de $L^1(\mathbf{R})$ es $L^\infty(\mathbf{R})$, pero el dual de $L^\infty(\mathbf{R})$ no es $L^1(\mathbf{R})$ (hay funcionales lineales continuos sobre $L^\infty(\mathbf{R})$ que no están en $L^1(\mathbf{R})$). Se puede probar que este último espacio no es el dual de ningún espacio de Banach.

Los espacios de sucesiones l^p que hemos estudiado previamente se pueden considerar como un caso particular de la teoría de los espacios L^p cuando la medida es la medida discreta en el espacio \mathbf{N} .

Capítulo 4

Espacios de Hilbert

4.1. Introducción

Los espacios de Hilbert constituyen el punto central de este curso. Aunque en los tratados elementales no ocupan demasiado lugar (reservado a los espacios vectoriales topológicos y a los espacios de Banach) desde el punto de vista de las aplicaciones físicas constituyen una herramienta fundamental. En lo que sigue desarrollaremos la teoría elemental, con especial énfasis en los aspectos geométricos, para luego, en capítulos posteriores estudiar la teoría de operadores sobre estos espacios.

4.2. Definiciones y propiedades

Definición 4.2.1 Sea L un espacio vectorial sobre $\mathbf{K} = \mathbf{R}, \mathbf{C}$. Consideremos la siguiente aplicación:

$$\begin{aligned} L \times L &\rightarrow \mathbf{K} \\ x, y &\mapsto (x, y) \end{aligned}$$

que verifica las siguientes propiedades:

1. $(x, x) \geq 0$, $x \in \mathcal{H}$; $(x, x) = 0 \Leftrightarrow x = 0$
2. $(x, y + z) = (x, y) + (x, z)$, $x, y, z \in L$
3. $(x, \lambda y) = \lambda(x, y)$, $\lambda \in \mathbf{K}$, $x \in L$
4. $(x, y) = \overline{(y, x)}$ $x, y \in L$

Las propiedades 2,3,4 equivalen a decir que $(,)$ es una forma sesquilineal hermítica (simétrica si $\mathbf{K} = \mathbf{R}$). La propiedad 1 quiere decir que la forma es definida positiva. Se dice que el espacio L es un espacio pre-Hilbert con el producto escalar $(,)$.

Proposición 4.2.1 (Desigualdad de Cauchy-Schwarz) Sea L un espacio pre-Hilbert. Para todo par de vectores $x, y \in \mathcal{H}$ se tiene:

$$|(x, y)| \leq \sqrt{(x, x)(y, y)}$$

Para todo $\lambda \in \mathbf{C}$, $x, y \in L$ (supongamos $y \neq 0$) se tiene:

$$0 \leq (x + \lambda y, x + \lambda y) = (x, x) + 2\operatorname{Re}(\lambda(x, y)) + |\lambda|^2(y, y)$$

Sea $(x, y) = re^{i\theta}$ y escojamos $\lambda = te^{-i\theta}$, donde $t \in \mathbf{R}$. Entonces:

$$0 \leq (x, x) + 2t|(x, y)| + t^2(y, y)$$

En particular, para $t = -|(x, y)|/(y, y)$

$$0 \leq (x, x) - \frac{|(x, y)|^2}{(y, y)}$$

es decir:

$$|(x, y)| \leq \sqrt{(x, x)(y, y)}$$

Proposición 4.2.2 *Sea \mathcal{H} un espacio pre-Hilbert. La aplicación*

$$x \in \mathcal{H} \rightarrow \|x\| = (x, x)^{1/2}$$

es una norma en \mathcal{H} .

La demostración es la siguiente. Por las propiedades de producto escalar, $\|x\| \geq 0$ y $\|x\| = 0$ si y solo si $x = 0$. Además, $\|\lambda x\| = (\lambda x, \lambda x)^{1/2} = |\lambda|(x, x)^{1/2}$. Finalmente,

$$\begin{aligned} \|x + y\|^2 &= (x + y, x + y) = \|x\|^2 + \|y\|^2 + 2\operatorname{Re}(x, y) \leq \\ &\|x\|^2 + \|y\|^2 + 2|(x, y)| \leq \|x\|^2 + \|y\|^2 + 2\|x\| \|y\| = (\|x\| + \|y\|)^2 \end{aligned}$$

Por lo tanto, un espacio pre-Hilbert es un espacio normado y como consecuencia un espacio métrico. Los espacios que más nos van a interesar son los espacios pre-Hilbert que son completos (y por consiguiente espacios de Banach). El producto escalar es una aplicación continua en la topología de la métrica que deriva de él.

Definición 4.2.2 *Un espacio de Hilbert es un espacio pre-Hilbert completo, con la topología de la métrica derivada del producto escalar.*

Como hemos visto, de todo producto escalar deriva un norma, pero el inverso no es cierto (hay espacios de Banach que no son espacios de Hilbert).

Proposición 4.2.3 (Identidad de polarización) *En un espacio pre-Hilbert L se tiene la siguiente igualdad*

$$(x, y) = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2) - \frac{i}{4}(\|x + iy\|^2 - \|x - iy\|^2), \quad x, y \in L$$

En el caso de un espacio real la identidad de polarización se escribe como:

$$(x, y) = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2) = \frac{1}{2}(\|x + y\|^2 - \|x\|^2 - \|y\|^2), \quad x, y \in L$$

Para comprobarlo basta desarrollar los segundos miembros. Esto implica en particular que si una norma proviene de un producto escalar, éste es único. Se cumple también la siguiente identidad (de clara interpretación geométrica)

Proposición 4.2.4 (Ley del paralelogramo) *Sea L un espacio pre-Hilbert. Entonces,*

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2, \quad x, y \in L$$

que se demuestra fácilmente aplicando la definición de la norma en función del producto escalar. Pero lo que es más interesante es que esta propiedad caracteriza las normas que se pueden definir a través de un producto escalar.

Proposición 4.2.5 *Una norma en un espacio vectorial proviene de un producto escalar (es decir, existe un producto escalar tal que $\|x\| = \sqrt{(x, x)}$) si y solo si la norma verifica la identidad del paralelogramo.*

La demostración consiste en definir un producto escalar a través de la identidad de polarización y comprobar que efectivamente se dan las propiedades del producto escalar. Hagámoslo en el caso real. El producto escalar tiene que venir definido por la identidad de polarización

$$(x, y) = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2)$$

Obviamente, $(x, x) \geq 0$ y solo es cero cuando $x = 0$. Además $(x, y) = (y, x)$. Veamos la linealidad en la primera variable:

$$(x, y + z) = (x, y) + (x, z)$$

es equivalente a

$$\|x + y + z\|^2 - \|y + z\|^2 - \|x + y\|^2 - \|x + z\|^2 + \|x\|^2 + \|y\|^2 + \|z\|^2 = 0$$

Para probar esta identidad, pongamos, usando la ley del paralelogramo:

$$\begin{aligned} \|x + y + z\|^2 &= -\|x - z\|^2 + 2\|x + \frac{1}{2}y\|^2 + 2\|z + \frac{1}{2}y\|^2 \\ \|x - z\|^2 &= -\|x + z\|^2 - 2\|x\|^2 - 2\|z\|^2 \\ \|x + \frac{1}{2}y\|^2 &= \frac{1}{4}\|2x + y\|^2 = -\frac{1}{4}\|y\|^2 + \frac{1}{2}\|x + y\|^2 + \frac{1}{2}\|x\|^2 \\ \|z + \frac{1}{2}y\|^2 &= \frac{1}{4}\|2z + y\|^2 = -\frac{1}{4}\|y\|^2 + \frac{1}{2}\|z + y\|^2 + \frac{1}{2}\|z\|^2 \end{aligned}$$

y sumando obtenemos el resultado deseado. Para probar que $(\lambda x, y) = \lambda(x, y)$ se prueba en primer lugar para naturales (usando la propiedad anterior). Resulta inmediato probar que $(-x, y) = -(x, y)$. Y para racionales se tiene:

$$\begin{aligned} \left(\frac{1}{q}x, y\right) &= \frac{1}{4}\left(\left\|\frac{1}{q}x + y\right\|^2 - \left\|\frac{1}{q}x - y\right\|^2\right) = \\ &= \frac{1}{4q^2}\left(\|x + qy\|^2 - \|x - qy\|^2\right) = \frac{1}{q^2}(x, qy) = \frac{1}{q}(x, y) \end{aligned}$$

y por continuidad se demuestra en \mathbf{R} . De forma similar se hace para \mathbf{C} .

El espacio l^2 de las sucesiones complejas cuya suma de cuadrados es convergente es un espacio de Hilbert. Sin embargo, el conjunto de funciones continuas en el intervalo $[0, 1]$ (de cuadrado integrable), no es un espacio de Hilbert sino solo un espacio pre-Hilbert. Es claro que

$$(f, g) = \int_0^1 \overline{f(t)}g(t) dt$$

es un producto escalar en ese espacio. Sin embargo no es completo. Sea la sucesión de funciones:

$$f_n(t) = \inf\{n, t^{-1/3}\}$$

Se trata de una sucesión de funciones continuas en el intervalo $[0, 1]$, que además es de Cauchy. Sin embargo, su límite en la norma

$$\|f\|_2 = \left(\int_0^1 |f(t)|^2 dt\right)^{1/2}$$

no es una función continua. Se puede probar que el completado es el conjunto de funciones de cuadrado integrable Lebesgue.

Un subespacio vectorial de un espacio de Hilbert no tiene porqué ser cerrado. Sin embargo su cierre también es un subespacio. Como estamos en un espacio métrico completo, los conjuntos cerrados son completos (y viceversa). Por lo tanto, todo subespacio cerrado de un espacio de Hilbert es un espacio de Hilbert. Si un subespacio es de dimensión finita, es cerrado (y completo, también en el caso en que el espacio total es solo pre-Hilbert).

4.3. Ortogonalidad

Uno de los conceptos más importantes en un espacio de Hilbert es el de ortogonalidad. Aunque muchas de las propiedades siguientes son ciertas en un espacio pre-Hilbert, consideremos para abreviar que estamos en un espacio de Hilbert.

Definición 4.3.1 *Sea \mathcal{H} un espacio de Hilbert. Se dice que dos vectores son ortogonales si su producto escalar es cero.*

Dos vectores ortogonales (no nulos) son linealmente independientes (l.i.).

Definición 4.3.2 *Un subconjunto S de un espacio de Hilbert \mathcal{H} es ortogonal si $x \neq 0, \forall x \in S$ y se tiene $(x, y) = 0, \forall x, y \in S$.*

El sistema se llama ortonormal si además de lo anterior, $\|x\| = 1, x \in S$.

Teorema 4.3.1 (Pitágoras) *Sea $\{x_1, \dots, x_n\}$ un conjunto de vectores en un espacio de Hilbert, que forman un conjunto ortogonal. Entonces:*

$$\left\| \sum_{i=1}^n x_i \right\|^2 = \sum_{i=1}^n \|x_i\|^2$$

La demostración es inmediata. Sin embargo, el recíproco no es cierto (en \mathbf{C}). Un sistema ortogonal es linealmente independiente. Basta usar el teorema de Pitágoras para probarlo.

Definición 4.3.3 *Se dice que un subconjunto S de un espacio de Hilbert, \mathcal{H} , es total si el único vector de \mathcal{H} que es ortogonal a todos los vectores de S es el vector 0.*

Proposición 4.3.1 *El conjunto de vectores ortogonales a un subconjunto $S \subset \mathcal{H}$ es un subespacio vectorial cerrado de \mathcal{H} (se le llama complemento ortogonal, S^\perp).*

La demostración es inmediata de las propiedades de linealidad y continuidad del producto escalar.

La ortogonalidad cumple las siguientes propiedades.

Proposición 4.3.2 *Sea \mathcal{H} un espacio de Hilbert. Entonces:*

1. $A \subset B \Rightarrow B^\perp \subset A^\perp$
2. $A \subset (A^\perp)^\perp$
3. $(\cup_i A_i)^\perp = \cap_i A_i^\perp$
4. $A^\perp = (\bar{A})^\perp = (\text{lin } A)^\perp$

La demostración de las tres primeras es muy sencilla. En cuanto a la cuarta, $A \subset \bar{A}$ y por tanto $(\bar{A})^\perp \subset A^\perp$. Además $A \subset \text{lin } A$ y por tanto $(\text{lin } A)^\perp \subset A^\perp$. En sentido contrario, si $x \in A^\perp$ y x es ortogonal a una sucesión convergente de vectores en A , también lo es a su límite (que está en \bar{A}), luego $x \in (\bar{A})^\perp$. Para la envolvente lineal el argumento es similar.

El cierre de un conjunto siempre está contenido en el ortogonal del ortogonal: $\bar{A} \subset (A^\perp)^\perp$. Y en el caso de que se trate de un subespacio vectorial cerrado, el ortogonal del ortogonal coincide con el conjunto de partida. Nótese que si un subespacio L es denso en \mathcal{H} , su ortogonal es igual a $\{0\}$, pues,

$$L^\perp = (\bar{L})^\perp = \mathcal{H}^\perp = \{0\}$$

La suma directa (ortogonal) de dos subespacios cerrados ortogonales, es un subespacio cerrado. Pero en general, la suma de subespacios cerrados no tiene por qué ser cerrada. Veamos un ejemplo de esta situación. Supongamos que en un espacio de Hilbert tenemos dos sucesiones ortonormales, $\{x_n\}$ e $\{y_n\}$ (que supondremos infinitas, ya que en un espacio de Hilbert de dimensión finita todos los subespacios

son cerrados). Supongamos también que los vectores de la primera sucesión son ortogonales a los de la segunda y construimos una tercera sucesión como combinación de las dos primeras:

$$z_n = \left(\cos \frac{1}{n} \right) x_n + \left(\sin \frac{1}{n} \right) y_n$$

Debido a las propiedades de las sucesiones $\{x_n\}$ e $\{y_n\}$ y a la elección de los coeficientes, la sucesión $\{z_n\}$ es ortonormal. Sean

$$L_x = \overline{(\text{lin}\{x_n\})}, \quad L_z = \overline{(\text{lin}\{z_n\})}$$

Es evidente que la suma no es ortogonal:

$$(x_n, z_n) = \cos \frac{1}{n}$$

Para cada $N \in \mathbf{N}$, la suma de vectores

$$\sum_{n=1}^N \left(\sin \frac{1}{n} \right) y_n$$

está en $L_x + L_z$:

$$\left(\sin \frac{1}{n} \right) y_n = - \left(\cos \frac{1}{n} \right) x_n + z_n \in L_x + L_z$$

La sucesión de esos vectores es una serie convergente pues:

$$\sum_{n=1}^{\infty} \sin^2 \frac{1}{n} < \infty$$

Por tanto,

$$y \in \overline{L_x + L_z}$$

Pero $y \notin L_x + L_z$. En efecto, si fuera así:

$$y = x + z, \quad x \in L_x, \quad z \in L_z$$

y se tendría

$$\begin{aligned} \sin \frac{1}{m} &= (y, y_m) = (x + z, y_m) = (z, y_m) = \left(\sum_n (z_n, z) z_n, y_m \right) \\ &= \sum_n (z, z_n) (z_n, y_m) = (z, z_m) \sin \frac{1}{m} \end{aligned}$$

lo que implica que los coeficientes de Fourier de z respecto a la base de L_z son todos iguales a 1. Pero como el conjunto es infinito esto es imposible (ver más adelante las condiciones sobre bases ortonormales).

La noción de proyección sobre un conjunto puede establecerse en un espacio métrico (y puede no existir para un subconjunto arbitrario): se plantea en términos de distancia mínima. Sin embargo, al limitarnos a espacios de Hilbert, la introduciremos en términos de ortogonalidad, con lo que las cuestiones de aproximación serán una consecuencia.

Teorema 4.3.2 (Teorema de la proyección ortogonal) *Sea \mathcal{H} un espacio de Hilbert y L un subespacio lineal cerrado. El espacio \mathcal{H} se puede escribir como la suma directa (ortogonal) de L y su complemento ortogonal L^\perp . Es decir todo vector u se puede poner de forma única como la suma de dos vectores ortogonales, $v \in L$, $w \in L^\perp$. Escribiremos:*

$$\mathcal{H} = L \oplus L^\perp$$

y diremos que v es la proyección ortogonal de u sobre L .

Nótese que se ha exigido que el subespacio sea cerrado. En primer lugar probaremos que existe un vector v de L que minimiza la distancia $\|u - x\|$, para $u \in \mathcal{H}$ fijado, y es único. Sea $a = \inf\{\|u - x\| : x \in L\}$. Construimos una sucesión (x_n) en L tal que $\lim_{n \rightarrow \infty} \|u - x_n\| = a$ (que existe por ser a el ínfimo). Veamos que x_n es de Cauchy. Utilizando la ley del paralelogramo,

$$\|(x_n - u) - (x_m - u)\|^2 + \|x_n + x_m - 2u\|^2 = 2\|u - x_n\|^2 + 2\|u - x_m\|^2$$

es decir:

$$\|x_n - x_m\|^2 = 2\|u - x_n\|^2 + 2\|u - x_m\|^2 - 4\|\frac{1}{2}(x_n + x_m) - u\|^2$$

Pero $\frac{1}{2}(x_n + x_m) \in L$ (nótese que bastaría con exigir que L fuera convexo, aunque no fuera un subespacio, para tener esta propiedad). Por tanto:

$$\|x_n - x_m\|^2 \leq 2(\|u - x_n\|^2 - a^2) + 2(\|u - x_m\|^2 - a^2)$$

Si $n, m \rightarrow \infty$, el miembro de la derecha se hace tan pequeño como queramos y por tanto la sucesión (x_n) es de Cauchy. Pero el subespacio L es cerrado y por tanto el límite de esta sucesión, v , está en L (como antes, bastaría que el subespacio L fuera un subconjunto completo de un espacio pre-Hilbert). Además, $\|u - v\| = a$, pues $\lim_{n \rightarrow \infty} (u - x_n) = u - v$ y por tanto $\lim_{n \rightarrow \infty} \|u - x_n\| = \|u - v\| = a$. Finalmente para acabar esta primera parte, probaremos que v es único. Sea $v' \in L$ tal que $\|u - v'\| = a$. Usaremos nuevamente la ley del paralelogramo.

$$\begin{aligned} \|v - v'\|^2 &= 2\|u - v\|^2 + 2\|u - v'\|^2 - \|v + v' - 2u\|^2 = \\ &2\|u - v\|^2 + 2\|u - v'\|^2 - 4\|\frac{1}{2}(v + v') - u\|^2 \leq 2a^2 + 2a^2 - 4a^2 = 0 \end{aligned}$$

Luego $v' = v$. Pasemos ahora a la cuestión de la suma directa ortogonal. Téngase en cuenta que la suma de subespacios puede ser directa sin ser ortogonal. Tenemos que probar que $u - v$ es un vector ortogonal a todos los vectores de L . Sea $y \in L$, y supongamos que $\|y\| = 1$ (obviamente, no se pierde generalidad con esta hipótesis). Se tiene ($w = u - v$):

$$\|w - \mu y\|^2 = (w, w - \mu y) - (\mu y, w - \mu y) = \|w\|^2 - |(w, y)|^2 + |(w, y) - \mu|^2$$

Tomando $\mu_0 = (w, y)$ obtenemos:

$$\|w - \mu_0 y\|^2 = \|w\|^2 - |(w, y)|^2$$

Pero $w - \mu_0 y = u - (v + \mu_0 y)$ y el vector $v + \mu_0 y$ está en L . Por lo tanto, la distancia de u a este vector es mayor o igual que la de u a v : $\|w\| = \|u - v\| \leq \|w - \mu_0 y\|$. Entonces

$$\|w\|^2 \leq \|w\|^2 - |(w, y)|^2$$

y no puede ocurrir la desigualdad estricta, luego $(w, y) = 0$.

La aplicación $P_L : \mathcal{H} \rightarrow \mathcal{H}$ que asigna a cada vector u el vector v definido en la proposición anterior, se llama el proyector ortogonal asociado a L . Podemos definir también el proyector correspondiente a L^\perp . Estas aplicaciones son lineales (como se puede comprobar fácilmente) y verifican las siguientes propiedades:

Proposición 4.3.3 *Sea L un subespacio vectorial cerrado de un espacio de Hilbert \mathcal{H} . Sea P_L la proyección ortogonal sobre L . Entonces:*

1. $(P_L x, y) = (x, P_L y)$, $\forall x, y \in \mathcal{H}$
2. $P_L x = x$, $\forall x \in L$
3. $P_L y = 0$, $\forall y \in L^\perp$
4. $P_L^2 = P$

$$5. (P_L x, x) = \|P_L x\|^2 \leq \|x\|^2$$

$$6. P_L + P_{L^\perp} = I$$

$$7. P_L P_{L^\perp} = P_{L^\perp} P_L = 0$$

La primera propiedad implica la siguiente igualdad. Si $x = x_L + x_{L^\perp}$, $y = y_L + y_{L^\perp}$, se tiene:

$$(P_L x, y) = (x, P_L y) = (x_L, y_L)$$

La segunda, tercera y cuarta son evidentes. En cuanto a la quinta,

$$(P_L x, x) = \|x_L\|^2 \leq \|x\|^2$$

Las dos últimas son también inmediatas.

Dada una sucesión ortogonal (un conjunto ortogonal numerable) en un espacio de Hilbert, se tiene la siguiente desigualdad que relaciona los productos escalares del vector x con los vectores del conjunto ortogonal y la norma del vector.

Proposición 4.3.4 (Desigualdad de Bessel) *Sea $\{e_n\}$, $n \in \mathbf{N}$ un sistema ortonormal. Entonces, para todo $x \in \mathcal{H}$ se tiene que la serie $\sum_{n=1}^{\infty} |(e_n, x)|^2$ es convergente y :*

$$\sum_{n=1}^{\infty} |(e_n, x)|^2 \leq \|x\|^2$$

(La proposición puede extenderse a familias arbitrarias de vectores ortonormales) Para demostrarla, probemos que las sumas parciales de la serie están acotadas:

$$\begin{aligned} 0 &\leq \left(x - \sum_{n=1}^N (e_n, x)e_n, x - \sum_{n=1}^N (e_n, x)e_n\right) = \\ &\|x\|^2 - \sum_{n=1}^N (e_n, x)(x, e_n) - \sum_{n=1}^N (x, e_n)(e_n, x) + \\ &\sum_{n,m=1}^N \overline{(e_n, x)}(e_m, x)(e_n, e_m) = \|x\|^2 - \sum_{n=1}^N |(e_n, x)|^2 \end{aligned}$$

Al estar las sumas parciales acotadas, la serie es convergente y su suma está acotada por $\|x\|^2$. Los escalares (e_n, x) se llaman los coeficientes de Fourier de x respecto al sistema ortonormal $\{e_n\}$. Como consecuencia de la desigualdad de Bessel se tiene el siguiente teorema.

Teorema 4.3.3 *Sea \mathcal{H} un espacio de Hilbert y $\{e_n\}$ un sistema ortonormal.*

1. *Para todo $x \in \mathcal{H}$, la serie $\sum_{n=1}^{\infty} (e_n, x)e_n$ es convergente.*
2. *Para todo $x, y \in \mathcal{H}$, la serie de escalares $\sum_{n=1}^{\infty} (e_n, x)(e_n, y)$ converge a $(\sum_{n=1}^{\infty} (e_n, x)e_n, y)$*
3. *Si M es el subespacio cerrado generado por el sistema ortonormal, y P_M la proyección sobre ese espacio, se tiene:*

$$P_M(x) = \sum_{n=1}^{\infty} (e_n, x)e_n, \quad \forall x \in \mathcal{H}$$

Probaremos que la sucesión de sumas parciales es de Cauchy.

$$\left\| \sum_{k=n}^m (e_k, x)e_k \right\|^2 = \sum_{k=n}^m |(e_k, x)|^2$$

Pero la serie de la derecha es convergente por la desigualdad de Bessel, luego para todo $\epsilon > 0$ existe n_0 tal que si $n, m \geq n_0$ se tiene:

$$\sum_{k=n}^m |(e_k, x)|^2 \leq \epsilon^2$$

luego,

$$\left\| \sum_{k=n}^m (e_k, x) e_k \right\| < \epsilon$$

y la sucesión de sumas parciales es de Cauchy, con lo que la serie es convergente (al ser \mathcal{H} un espacio completo por ser de Hilbert). Para demostrar la segunda parte, se puede usar el hecho de que la aplicación $x \rightarrow (y, x)$ (para un $y \in \mathcal{H}$ fijado) es lineal y continua, y por lo tanto, la imagen de una sucesión convergente es convergente. Al haber probado que la serie $\sum_{n=1}^{\infty} (e_n, x) e_n$ es convergente, su imagen $\sum_{n=1}^{\infty} (e_n, x)(y, e_n)$ mediante la aplicación anterior es convergente, y su suma es la imagen de la suma, es decir: $(\sum_{n=1}^{\infty} (e_n, x) e_n, y)$. Finalmente,

$$(e_m, x - \sum_{n=1}^{\infty} (e_n, x) e_n) = (e_m, x) - (e_m, \sum_{n=1}^{\infty} (e_n, x) e_n) = 0$$

luego $x - \sum_{n=1}^{\infty} (e_n, x) e_n$ es ortogonal a la familia $\{e_n\}$ y por tanto al subespacio que generan, luego pertenece al ortogonal. Si P_M es la proyección sobre el cierre del subespacio generado por $\{e_n\}$, M , entonces

$$x - P_M(x) \in M^\perp$$

y por tanto

$$P_M(x) = \sum_{n=1}^{\infty} (e_n, x) e_n$$

Este vector es además el que hace mínima la distancia de x a M :

$$\left\| x - \sum_{n=1}^{\infty} (e_n, x) e_n \right\| \leq \|x - y\|, \quad \forall y \in M$$

Definición 4.3.4 Una base ortonormal en un espacio de Hilbert es un sistema ortonormal maximal (en el sentido de que no está contenido propiamente en otro sistema ortonormal).

Proposición 4.3.5 Sea \mathcal{H} un espacio de Hilbert y S un sistema ortonormal. Entonces, existe un sistema ortonormal maximal que contiene a S .

La demostración se basa en el lema de Zorn. Como corolario se tiene:

Corolario 4.3.1 Todo espacio de Hilbert posee una base ortonormal.

A partir de una sucesión de vectores l.i. de un espacio de Hilbert es posible construir un sistema ortogonal.

Proposición 4.3.6 (Ortogonalización de Gram-Schmidt) Sea $\{v_n : n \in \mathbf{N}\}$ una sucesión de vectores linealmente independientes en un espacio de Hilbert \mathcal{H} . Entonces existe una sucesión ortogonal de vectores $\{u_n : n \in \mathbf{N}\}$ que verifica que el subespacio generado por $\{v_k : k = 1, \dots, n\}$ es igual al generado por $\{u_k : k = 1, \dots, n\}$ para todo $n \in \mathbf{N}$. El sistema $\{u_n : n \in \mathbf{N}\}$ está determinado de manera única salvo constantes.

Sea $M_n = \text{lin}\{v_k : k = 1, \dots, n\}$. Al ser un espacio de dimensión finita es cerrado (y por tanto completo) y existe la proyección P_{M_n} . Sea $u_1 = v_1$ y supongamos que el proceso de construcción de los u_k se ha completado hasta n . Veamos como construir u_{n+1} . Sea

$$u_{n+1} = v_{n+1} - P_{M_n}(v_{n+1})$$

El vector u_{n+1} es distinto de cero, pues los vectores que generan M_n son l.i. con v_{n+1} . Por construcción,

$$u_{n+1} \in M_n^\perp$$

lo que implica que se cumple la primera condición (ser ortogonales). Además, $u_{n+1} \in \text{lin}(M_n \cup \{v_{n+1}\}) = M_{n+1}$ y $v_{n+1} \in \text{lin}(M_n \cup \{u_{n+1}\}) \subset M_{n+1}$, por lo que se concluye que

$$\text{lin}\{v_1, \dots, v_n\} = \text{lin}\{u_1, \dots, u_n\}$$

En cuanto a la unicidad, si hubiera otro sistema $\{u'_n : n \in \mathbf{N}\}$ con las mismas propiedades que $\{u_n : n \in \mathbf{N}\}$, se tendría:

$$u'_n \in \text{lin}\{u'_1, \dots, u'_{n-1}\}^\perp = \text{lin}\{u_1, \dots, u_{n-1}\}^\perp$$

Además $u'_n \in M_n$ por lo tanto:

$$u'_n = \sum_{k=1}^n \lambda_k u_k$$

de donde $u'_n = \lambda_n u_n$. En la práctica, la construcción de la familia ortogonal es:

$$u_{n+1} = v_{n+1} + \sum_{k=1}^n \lambda_k u_k, \quad \lambda_k = -\frac{(u_k, v_{n+1})}{\|u_k\|^2}$$

Como consecuencia de esto, se puede demostrar:

Proposición 4.3.7 *Un espacio de Hilbert es separable si y solo si existe una base hilbertiana (ortogonal) numerable*

La demostración es muy sencilla. Si el espacio tiene una base hilbertiana numerable, el subespacio cerrado que genera es separable y coincide con todo el espacio (basta entonces considerar combinaciones lineales con coeficientes racionales para acabar la demostración). Y viceversa, si el espacio es separable existe una sucesión que es densa en el espacio. Usando el proceso de Gram-Schmidt podemos conseguir un sistema ortogonal que es total. Para ello, en primer lugar hay que seleccionar una subsucesión que sea l.i. Por inducción se prueba que esta subsucesión genera todo el espacio y ahora se aplica Gram-Schmidt.

Definición 4.3.5 *Se llama dimensión hilbertiana de un espacio de Hilbert al cardinal de una de sus bases ortonormales.*

La definición está justificada por la siguiente proposición:

Proposición 4.3.8 1. *Dos bases ortonormales de un espacio de Hilbert tienen el mismo cardinal*

2. *Dos espacios de Hilbert son isomorfos si y solo si su dimensión hilbertiana es la misma.*

Una base ortonormal se puede caracterizar por cualquiera de las siguientes propiedades (el resultado es válido en espacios no separables, pero lo enunciaremos solo en el caso en que se tenga una base ortonormal infinita numerable. El caso de dimensión finita es trivial).

Proposición 4.3.9 *Sea $S = \{e_n : n \in \mathbf{N}\}$ un sistema ortogonal de un espacio de Hilbert \mathcal{H} . Son equivalentes:*

1. *S es una base ortonormal.*
2. *El cierre de la envolvente lineal de S es igual a \mathcal{H} (la envolvente lineal es densa en \mathcal{H}).*
3. *El único vector ortogonal a S es el vector 0.*
4. *Para todo $x \in \mathcal{H}$ la serie $\sum_{n=1}^{\infty} (e_n, x)e_n$ converge al vector x .*

Veamos la equivalencia de estas propiedades.

1) \Rightarrow 2): Si S es una base ortonormal y $\text{lin}(S)$ no fuera denso en \mathcal{H} , su complemento ortogonal sería distinto de $\{0\}$ y habría un vector ortogonal a todo S , que no sería maximal.

2) \Rightarrow 3): Por ser $\text{lin}(S)$ denso en \mathcal{H} , $\text{lin}(S)^\perp = \{0\}$, luego $S^\perp = \{0\}$.

3) \Rightarrow 4): Consideremos la serie en $\text{lin}(S)$

$$\sum_{n=1}^{\infty} (e_n, x) e_n$$

Por la desigualdad de Bessel sabemos que esta serie es convergente a un vector x' . Veamos que $x' = x$. En efecto,

$$(e_m, x - x') = (e_m, x - \sum_{n=1}^{\infty} (e_n, x) e_n) = (e_m, x) - \sum_{n=1}^{\infty} (e_n, x) (e_m, e_n) = 0, \quad \forall m$$

y por 3), $x - x' = 0$.

4) \Rightarrow 1): Si $(x, e_n) = 0$ para todo n , se tiene por 4) que $x = 0$.

Como resumen de estas definiciones y propiedades se tiene la identidad de Parseval. Todo vector de un espacio de Hilbert separable se puede escribir como una serie en los elementos de una base ortonormal:

$$x = \sum (e_n, x) e_n$$

El producto escalar se puede escribir así:

$$(x, y) = \sum_{n=1}^{\infty} \overline{(e_n, x)} (e_n, y)$$

y se tiene la identidad de Parseval:

$$\sum_{n=1}^{\infty} |(e_n, x)|^2 = \|x\|^2$$

Como un espacio de Hilbert es un espacio vectorial, admite una base algebraica, es decir un conjunto de vectores linealmente independientes que generan (mediante combinaciones lineales, que son sumas finitas) el espacio. Una base ortonormal no es una base algebraica en un espacio de dimensión infinita. A estas últimas se les llama bases de Hamel.

4.4. Polinomios ortogonales

En el espacio de Hilbert $L^2[a, b]$ existen numerosas bases ortonormales que aparecen en las aplicaciones, pero entre las más importantes se encuentran las de polinomios ortogonales con un peso $p(x)$. Sea $I \subset \mathbf{R}$ un intervalo y $p : I \rightarrow \mathbf{R}$ una función positiva (en el interior de I) y que verifica:

$$\int_I |t|^n p(t) dt < \infty, \quad \forall n \in \mathbf{N}$$

Definimos el espacio:

$$E_p = \{ f \in C(I) : \int_I |f(t)|^2 p(t) dt < \infty \}$$

Se trata de un espacio vectorial que contiene a los polinomios (debido a las propiedades de la función $p(t)$). En este espacio podemos definir un producto escalar:

$$(f, g) = \int_I \overline{f(t)} g(t) p(t) dt$$

siendo $p(t)$ la función peso. Con este producto escalar, E_p es un espacio pre-Hilbert (si en vez de considerar las funciones continuas usamos las funciones de cuadrado integrable Lebesgue obtenemos un espacio de Hilbert) y la distancia entre dos funciones es la desviación cuadrática media ponderada. Por ejemplo, para diferentes intervalos y funciones peso se obtienen las familias de polinomios que se detallan a continuación.

I	$p(t)$	Polinomios
$[-1, 1]$	1	Legendre
$[-1, 1]$	$\frac{1}{\sqrt{1-t^2}}$	Tchebychef
\mathbf{R}	e^{-t^2}	Hermite
$[0, \infty)$	e^{-t}	Laguerre

El punto más delicado al establecer que se trata de bases de un espacio de Hilbert está en la cuestión de si son o no un sistema maximal. Se tiene el siguiente resultado:

Proposición 4.4.1 *Sea $\{p_n(t)\}$ una familia de polinomios ortonormales, que se obtiene de la familia $\{1, t, t^2, \dots\}$ utilizando Gram-Schmidt respecto al producto escalar:*

$$(f, g) = \int_{-\infty}^{\infty} f(t)g(t)m(t) dt$$

Si la función peso $m(t)$ es no negativa (y no idénticamente cero) y verifica:

$$\int_{-\infty}^{\infty} m(t) dt < \infty, \quad \int_{-\infty}^{\infty} e^{\alpha|t|}m(t) dt < \infty$$

para algún $\alpha > 0$, entonces $\{\sqrt{m(t)}p_n(t)\}$ son una base ortonormal en $L^2(A)$, donde $A = \text{sop } m$

4.5. Series de Fourier

Entre las bases ortonormales más importantes en las aplicaciones se encuentran las series trigonométricas de Fourier. Consideremos el espacio de Hilbert $L^2[0, 2\pi]$ con el producto escalar usual:

$$(f, g) = \int_0^{2\pi} \overline{f(x)}g(x) dx$$

En este espacio, las funciones

$$\frac{1}{\sqrt{2\pi}}e^{inx}, \quad n \in \mathbf{Z}$$

son una base ortonormal. El teorema de Weierstrass permite afirmar que estas funciones (su envolvente lineal con más precisión) son densas en el conjunto de funciones continuas con la norma del supremo, lo que implica que lo son con la norma que deriva del producto escalar. Pero las funciones continuas son densas en $L^2[0, 2\pi]$ en cualquier norma $\|\cdot\|_p$.

Los coeficientes de Fourier se calculan fácilmente:

$$f \in L^2[0, 2\pi], \quad f(x) = \sum_{n \in \mathbf{Z}} c_n e^{inx}$$

donde

$$c_n = \frac{1}{2\pi} \int_0^{2\pi} e^{-inx} f(x) dx$$

También se puede usar una base ortonormal formada por funciones reales:

$$\left\{ \frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos nx, \frac{1}{\sqrt{\pi}} \operatorname{sen} nx, n = 1, 2, \dots \right\}$$

En este caso, los coeficientes de Fourier son:

$$a_0 = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} f(x) dx,$$

$$a_n = \frac{1}{\sqrt{\pi}} \int_0^{2\pi} f(x) \cos nx dx, \quad b_n = \frac{1}{\sqrt{\pi}} \int_0^{2\pi} f(x) \operatorname{sen} nx dx$$

para el desarrollo:

$$f(x) = \frac{1}{\sqrt{2\pi}} a_0 + \frac{1}{\sqrt{\pi}} \sum_{n=1}^{\infty} (a_n \cos nx + b_n \operatorname{sen} nx)$$

Aunque se pone el signo de igualdad, las series trigonométricas de Fourier no convergen necesariamente de forma puntual (solo en la norma $\|\cdot\|_2$). Sin embargo, si la función f tiene primera derivada continua, la convergencia es no solo puntual sino incluso uniforme. Para funciones en $L^2[0, 2\pi]$, se sabe que la convergencia es puntual casi doquiera. No podemos entrar aquí en un estudio detallado de las series de Fourier.

Como conclusión de este capítulo, mencionemos que un espacio de Hilbert separable es isomorfo a l^2 (basta escoger una base ortonormal). Sin embargo, este isomorfismo no siempre permite estudiar de la mejor manera posible las propiedades del espacio de Hilbert y los operadores en él definidos.

Capítulo 5

Operadores en espacios de Hilbert

5.1. Introducción

La teoría de operadores en espacios de Hilbert tiene múltiples aplicaciones en Física, especialmente en Mecánica Cuántica. La teoría es sin embargo extraordinariamente compleja si uno pretende cubrir todos los tipos de operadores que se presentan en Física. En esta corta introducción estableceremos las propiedades más importantes de los operadores acotados. En cuanto a la teoría espectral, pospondremos para otro capítulo unas consideraciones elementales sobre este importante tema, fundamental en las aplicaciones.

5.2. Operadores acotados

Consideraremos como caso más importante las aplicaciones lineales (operadores, dado que todas las aplicaciones que aparecerán son lineales omitiremos a veces el adjetivo lineal para el término operador) continuas entre dos espacios de Hilbert.

Sean $\mathcal{H}_1, \mathcal{H}_2$ dos espacios de Hilbert. Como ya hemos dicho, un operador acotado es una aplicación lineal $T : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ que verifica:

$$\|Tx\| \leq k\|x\|, \forall x \in \mathcal{H}, k \geq 0$$

Ya demostramos en el capítulo 3 que esta condición equivale (en un espacio normado) a la continuidad del operador T . Además, permite definir la norma del operador:

$$\|T\| = \sup \left\{ \frac{\|Tx\|}{\|x\|} : x \in \mathcal{H}, x \neq 0 \right\} = \sup \{ \|Tx\| : x \in \mathcal{H}, \|x\| = 1 \}$$

Se tiene entonces:

$$\|Tx\| \leq \|T\| \|x\|, \forall x \in \mathcal{H}$$

Téngase en cuenta que esta definición implica que $\|T\|$ es el ínfimo de las constantes k que acotan al operador. Y que no tiene por qué existir un vector x que verifique $\|Tx\| = \|T\| \|x\|$ (aunque veremos casos como el de los operadores compactos, donde esto sí ocurre). La norma de un operador verifica:

$$\|AB\| \leq \|A\| \|B\|$$

No es difícil construir ejemplos sencillos de operadores acotados, aunque no sea tan sencillo calcular su norma. Sea $\mathcal{H}_1 = \mathcal{H}_2 = l^2$ y T la aplicación en l^2 que transforma el vector x en y con $y_1 = 0, y_n = x_{n-1}$ (operadores de desplazamiento). Se trata de un operador acotado y su norma es igual a 1.

Hay ocasiones en las que los operadores no aparecen definidos sobre todo el espacio de Hilbert sino sobre un subespacio (necesariamente se trata de un subespacio al ser el operador lineal). Se llama dominio del operador T , $D(T)$, al subespacio en el que está definido este operador. Se dice que un operador es acotado en su dominio si

$$\sup \left\{ \frac{\|Tx\|}{\|x\|} : x \in D(T), x \neq 0 \right\} < \infty$$

El resultado más interesante es que un operador acotado se puede extender de manera única al cierre de su dominio, manteniendo la norma, con lo cual siempre se puede suponer que el dominio de un operador acotado es cerrado (y por lo tanto un espacio de Hilbert). Si el dominio de un operador acotado es denso en el espacio de Hilbert, entonces dicho operador se puede extender a todo el espacio. Este resultado no es cierto para un operador no acotado. La demostración se basa en el uso de sucesiones de Cauchy. Dado $x \in \overline{D(T)}$, existe una sucesión en $D(T)$ que tiende a x en \mathcal{H} , que obviamente es de Cauchy en $D(T)$. Al ser T continuo, la imagen de esta sucesión de Cauchy es también de Cauchy y por lo tanto tiene límite en el espacio final. Se llama rango (o recorrido) de un operador a su imagen.

Los operadores no acotados son fundamentales en Mecánica Cuántica, aunque su teoría es mucho más complicada que la de los acotados. Por ejemplo, el operador momento, $P = -i\frac{d}{dx}$, no es acotado en el siguiente dominio de $L^2[0, 1]$

$$D(T) = \{f \in L^2[0, 1] : f \in C^1[0, 1]\}$$

Resulta necesario restringir el dominio, pues no todas las funciones de $L^2[0, 1]$ son derivables. Si se toma en este espacio el conjunto de funciones e^{inx} , que tienen norma 1, se observa como las normas de las imágenes no están acotadas. La restricción puede dar lugar a operadores con propiedades muy diferentes. Supongamos que P está definido en el siguiente dominio (un subespacio de $L^2[0, 1]$):

$$D(T) = \{f \in L^2[0, 1] : f \text{ absolutamente continua}\}$$

(absolutamente continua equivale esencialmente a ser (casi doquiera) la integral indefinida de una función continua). Pues bien, el espectro (concepto que discutiremos en el próximo capítulo) de T es todo \mathbf{C} . Pero si el dominio se define como:

$$D(T) = \{f \in L^2[0, 1] : f \text{ absolutamente continua, } f(0) = 0\}$$

entonces el espectro es vacío (aunque los dominios son diferentes, ambos son densos en $L^2[0, 1]$).

No todo operador admite un inverso. Aunque lo admita, es posible que dicho inverso no sea acotado. El problema es que, en general, el rango de un operador (aunque sea acotado), no es un subespacio cerrado (aunque obviamente sea un subespacio). Una condición suficiente para que el rango sea cerrado, es la siguiente:

Proposición 5.2.1 *Dado un operador acotado T en un espacio de Hilbert \mathcal{H} , si existe $k > 0$ tal que*

$$\|Tx\| \geq k\|x\|, \quad x \in \mathcal{H}$$

entonces el rango de T es cerrado.

La demostración es sencilla. Sea $\{x_n\}$ una sucesión de vectores en \mathcal{H} e $y_n = Tx_n$ sus correspondientes imágenes. Supongamos que la sucesión de imágenes converge a $y \in \mathcal{H}$. Pero entonces:

$$\|x_n - x_m\| \leq \frac{1}{k} \|T(x_n - x_m)\| = \frac{1}{k} \|y_n - y_m\|$$

luego $\{x_n\}$ es de Cauchy y converge a $x \in \mathcal{H}$. Como el operador es acotado,

$$Tx = \lim_{n \rightarrow \infty} Tx_n = y$$

luego y está en el rango de T .

Puede entonces demostrarse la siguiente proposición relativa al operador inverso.

Proposición 5.2.2 *Un operador acotado en un espacio de Hilbert, \mathcal{H} , es invertible (posee un operador inverso que es acotado) si y solo si su rango es denso en \mathcal{H} y existe $k > 0$ tal que*

$$\|Tx\| \geq k\|x\|, \quad \forall x \in \mathcal{H}$$

Si el operador es invertible, el rango coincide con todo \mathcal{H} . Además, al ser T^{-1} acotado,

$$\|T^{-1}y\| \leq \|T^{-1}\| \|y\|, \quad y \in \mathcal{H}$$

es decir, escribiendo $x = T^{-1}y$:

$$\|x\| \leq \|T^{-1}\| \|Tx\|, \quad x \in \mathcal{H}$$

luego

$$\|Tx\| \geq k\|x\|, \quad x \in \mathcal{H}$$

con $k = 1/\|T^{-1}\|$. Supongamos ahora que el rango es denso y se da la acotación anterior. Por la proposición demostrada en primer lugar, el rango es cerrado y por tanto coincide con \mathcal{H} . Veamos que el operador T es inyectivo:

$$\|Tx - Ty\| \geq k\|x - y\|$$

pero si $Tx = Ty$, entonces $\|x - y\| \leq 0$ y T es inyectivo. Luego existe T^{-1} (y es lineal). Para probar que es acotado basta usar nuevamente la acotación de la proposición.

Pero no es difícil encontrar ejemplos de operadores acotados que admiten un inverso, pero éste no es acotado. Por ejemplo, el operador (lineal acotado) en l^2 definido sobre la base ortonormal canónica,

$$Te_n = \frac{1}{n^2}e_n, \quad n = 1, 2, \dots$$

tiene un inverso (porque es inyectivo). El dominio del inverso no es sin embargo, todo l^2 , pero es denso en él (pues contiene a la base ortonormal $\{e_n\}$). Si T^{-1} fuera continuo, de acuerdo con lo anterior, se podría extender su acción a todo l^2 . Pero no lo es:

$$T^{-1}e_n = n^2e_n$$

y los elementos de la base tienen norma 1. Luego no es acotado. Es posible encontrar elementos de l^2 sobre los que no está definido T^{-1} . Por ejemplo $x = (1, 1/2, 1/3, \dots)$

Como se vio en el capítulo dedicado a los espacios de Banach, el espacio de aplicaciones lineales continuas entre dos espacios de Hilbert (que son de Banach) es un espacio de Banach, con la norma definida anteriormente. Pero en general no es un espacio de Hilbert (con más precisión solo lo es si el espacio final tiene dimensión 1, es decir, es el dual topológico). Sin embargo, en ocasiones es más conveniente usar otras topologías en este espacio de aplicaciones. Supongamos que el espacio final e inicial son el mismo y consideremos el espacio de operadores acotados $\mathcal{A}(\mathcal{H})$. La topología inducida por la norma se llama también topología uniforme. Una sucesión de operadores A_n converge en esta topología al operador A si y solo si $\|A - A_n\| \rightarrow 0$ cuando $n \rightarrow \infty$. Se puede definir también una topología fuerte, en la que la convergencia se describe como sigue. Una sucesión de operadores A_n converge en la topología fuerte a un operador A si y solo si para cada vector x del espacio \mathcal{H} , la sucesión de vectores de \mathcal{H} , $A_n x$, converge a Ax . Esta claro que si una sucesión converge en norma, también lo hace en la topología fuerte (no entraremos en detalles de como se define la topología fuerte. Téngase en cuenta que el hecho de saber cuando una sucesión es convergente o no, no basta para determinar completamente la topología). La topología débil reduce aún más el concepto de convergencia llevándolo al cuerpo: una sucesión de operadores acotados A_n es convergente débilmente a un operador A si y solo si la sucesión de escalares $(x, A_n y)$ converge a (x, Ay) para cualquier par de vectores $x, y \in \mathcal{H}$. La topología uniforme es más fina que la fuerte y ésta más fina que la débil. Es decir, hay sucesiones que convergen débilmente pero no lo hacen fuertemente y sucesiones que lo hacen fuertemente pero no en norma. Sin embargo, si una sucesión converge en norma también lo hace fuertemente y débilmente.

5.3. Funcionales lineales continuos

Como en el caso de operadores entre espacios de Banach, un caso muy importante de operadores aparece cuando el espacio final es el cuerpo \mathbf{K} sobre el que está construido el espacio de Hilbert. Se habla entonces de funcionales lineales. Los más interesantes son los continuos, y en el caso de espacios

de Hilbert, como se hacía en dimensión finita, es muy sencillo construirlos utilizando el producto escalar. Veamos dos ejemplos de funcionales sobre el espacio $l^2(\mathbf{C})$. Sea

$$F(x) = \sum_{n=1}^{\infty} \frac{x_n}{n}, \quad x = (x_n) \in l^2(\mathbf{C})$$

Se trata de un funcional lineal, pues si

$$\sum_{n=1}^{\infty} |x_n|^2 < \infty$$

se tiene también (por la desigualdad de Cauchy-Schwarz)

$$\left| \sum_{n=1}^{\infty} \frac{x_n}{n} \right| < \infty$$

Este funcional se puede escribir como:

$$\sum_{n=1}^{\infty} \frac{x_n}{n} = (y, x), \quad y = \left(\frac{1}{n} \right) \in l^2(\mathbf{C})$$

Es continuo, pues

$$|(y, x)| \leq \|y\| \|x\|, \quad \forall x \in l^2(\mathbf{C})$$

Se tiene además que la norma del funcional es:

$$\|F\| = \|y\| = \sqrt{\sum_{n=1}^{\infty} \frac{1}{n^2}}$$

De hecho todos los funcionales de esta forma son continuos, y además, como veremos a continuación por el teorema de Riesz-Fréchet, todos los funcionales lineales continuos son de esta forma. Sea ahora

$$G(x) = \sum_{n=1}^{\infty} \frac{x_n}{\sqrt{n}}, \quad x = (x_n) \in l^2(\mathbf{C})$$

definido cuando la serie de término general $\frac{x_n}{\sqrt{n}}$ es convergente. Este funcional es lineal, pero no es continuo. Si lo fuera, estaría representado por el vector $(1/\sqrt{n})$ (según el teorema de Riesz-Fréchet) que no pertenece a $l^2(\mathbf{C})$.

Como en el caso de dimensión finita, se verifica el teorema de Riesz-Fréchet.

Teorema 5.3.1 (Riesz-Fréchet) *Sea \mathcal{H} un espacio de Hilbert. Entonces, para todo funcional continuo $F : \mathcal{H} \rightarrow \mathbf{K}$ existe un único vector $x_F \in \mathcal{H}$ tal que:*

$$F(y) = (x_F, y), \quad \forall y \in \mathcal{H}$$

Demostremos este resultado. En primer lugar, el núcleo de un funcional continuo (y en general, de un operador acotado) es un subespacio cerrado. Si coincide con \mathcal{H} el funcional es el cero. Si no coincide, sea

$$\mathcal{H} = \ker F \oplus (\ker F)^\perp$$

la descomposición ortogonal de \mathcal{H} respecto al núcleo de F . Consideremos un vector no nulo $x \in (\ker F)^\perp$. Todo vector $y \in \mathcal{H}$ se puede escribir como:

$$y = y - ax + ax$$

donde $a \in \mathbf{K}$. Elijamos a de forma que $y - ax \in \ker F$:

$$F(y - ax) = F(y) - aF(x) = 0 \Rightarrow a = \frac{F(y)}{F(x)}$$

Entonces

$$y = \left(y - \frac{F(y)}{F(x)}x \right) + \frac{F(y)}{F(x)}x$$

y el vector x es ortogonal a $y - \frac{F(y)}{F(x)}x$, con lo que:

$$(x, y) = \frac{F(y)}{F(x)}\|x\|^2 \Rightarrow F(y) = \frac{F(x)}{\|x\|^2}(x, y) = \left(\frac{\overline{F(x)}}{\|x\|^2}x, y \right) = (x_F, y)$$

donde

$$x_F = \frac{\overline{F(x)}}{\|x\|^2}x$$

A pesar de que x se ha elegido arbitrariamente en $(\ker F)^\perp$, el vector x_F es único. Si $z \in \mathcal{H}$ verifica:

$$F(y) = (x_F, y) = (z, y), \quad \forall y \in \mathcal{H}$$

entonces $z = x_F$.

Una consecuencia del teorema de Riesz-Fréchet es que la dimensión del subespacio ortogonal al núcleo de un funcional lineal continuo es 1. Téngase en cuenta que estos resultados son ciertos para funcionales lineales continuos. Si no es continuo, el núcleo no es un subespacio cerrado, de hecho es denso en el espacio total (y distinto de él).

El espacio de funcionales lineales continuos (el dual, \mathcal{H}^*) es un espacio de Hilbert, con un producto escalar definido de forma natural como:

$$(F, G) = (x_F, x_G)$$

y existe un anti-isomorfismo (lineal conjugado) entre \mathcal{H} y su dual \mathcal{H}^* .

5.4. Funcionales bilineales hermíticos

Dado un espacio de Hilbert \mathcal{H} , se pueden definir aplicaciones de $\mathcal{H} \times \mathcal{H}$ en \mathbf{K} que verifiquen las siguientes propiedades:

$$\varphi(x, \alpha y + \beta z) = \alpha \varphi(x, y) + \beta \varphi(x, z)$$

$$\varphi(\alpha x + \beta y, z) = \bar{\alpha} \varphi(x, z) + \bar{\beta} \varphi(y, z)$$

Es decir lineales en la segunda variable y antilineales en la primera. Se llamarán funcionales (o formas) bilineales (por abuso de lenguaje ciertamente). En el caso particular en que se verifique además:

$$\varphi(x, y) = \overline{\varphi(y, x)}$$

se llamarán formas hermíticas (o sesquilineales).

Asociados a ellos aparecen las formas cuadráticas, dadas por

$$Q_\varphi(x) = \varphi(x, x)$$

(no entraremos en detalles de una definición independiente de formas cuadráticas ni la restringiremos al caso en que φ sea hermítica, aunque esto sea lo más natural).

Debido a las propiedades expuestas anteriormente, cuando φ es un funcional hermítico, $Q(x)$ es real (y viceversa). En este caso, se dice que la forma cuadrática es definida positiva si $Q(x) > 0$ para todo $x \in \mathcal{H}$ distinto de cero.

La forma bilineal se puede expresar en términos de su forma cuadrática asociada (identidades de polarización):

$$\varphi(x, y) = Q_\varphi\left(\frac{1}{2}(x+y)\right) - Q_\varphi\left(\frac{1}{2}(x-y)\right) + iQ_\varphi\left(\frac{1}{2}(x+iy)\right) - iQ_\varphi\left(\frac{1}{2}(x-iy)\right)$$

Los conceptos de acotación (continuidad) para funcionales bilineales y formas cuadráticas son similares a los utilizados para funcionales lineales. Se dice que una forma bilineal es acotada si existe $k > 0$ tal que

$$|\varphi(x, y)| \leq k\|x\| \|y\|, \quad \forall x, y \in \mathcal{H}$$

y una forma cuadrática es acotada si

$$|Q_\varphi(x)| \leq k\|x\|^2, \quad \forall x \in \mathcal{H}$$

Se comprueba fácilmente que es posible definir las normas de formas bilineales y cuadráticas acotadas por:

$$\|\varphi\| = \sup\{|\varphi(x, y)| : \|x\| = \|y\| = 1\}, \quad \|Q_\varphi\| = \sup\{|Q_\varphi(x)| : \|x\| = 1\}$$

y que una forma bilineal y su forma cuadrática asociada son acotadas o no simultáneamente.

Además se dan la siguientes desigualdades entre sus normas cuando son acotadas:

$$\|Q_\varphi\| \leq \|\varphi\| \leq 2\|Q_\varphi\|$$

La demostración es una consecuencia de la identidad de polarización y la ley del paralelogramo. Pero si φ es hermitica, entonces:

$$\|Q_\varphi\| = \|\varphi\|$$

5.5. Operadores autoadjuntos y unitarios

Consideremos ahora el caso de operadores lineales acotados (en principio se pueden hacer desarrollos similares para operadores no acotados, pero no podemos entrar en ello ahora) de un espacio de Hilbert en sí mismo. Sea pues $A : \mathcal{H} \rightarrow \mathcal{H}$ un operador lineal acotado.

Definición 5.5.1 *Se llama operador adjunto de A al operador A^+ que verifica:*

$$(x, Ay) = (A^+x, y), \quad \forall x, y \in \mathcal{H}$$

Se puede demostrar que el operador A^+ existe y es único, usando formas bilineales. Si A es un operador acotado,

$$\varphi_A(x, y) = (x, Ay)$$

es una forma bilineal (no necesariamente hermitica). Pero es también acotada y se puede probar fácilmente que

$$\|A\| = \|\varphi_A\|$$

Partiendo de φ_A se puede definir para cada $x \in \mathcal{H}$ fijado, un funcional lineal:

$$\varphi_x(y) = \varphi(x, y) = (x, Ay)$$

que es también acotado. Por el teorema de Riesz-Fréchet, existe $z_x \in \mathcal{H}$ tal que:

$$\varphi_x(y) = (z_x, y) = (x, Ay)$$

La correspondencia

$$x \rightarrow z_x$$

es lineal y define un operador lineal y acotado (como se puede probar fácilmente) A^+ , que cumple

$$(A^+x, y) = (x, Ay), \quad \|A\| = \|A^+\|$$

Proposición 5.5.1 *El operador adjunto verifica las siguientes propiedades:*

1. $(aA + bB)^+ = \bar{a}A^+ + \bar{b}B^+$
2. $(A^+)^+ = A$

3. $(AB)^+ = B^+A^+$
4. Si el operador A tiene un inverso que está acotado, entonces A^+ también lo tiene y se verifica $(A^{-1})^+ = (A^+)^{-1}$
5. $\|A\|^2 = \|AA^+\|$

La demostración de la última propiedad, se sigue de la acotación siguiente

$$\|AA^+\| \leq \|A\| \|A^+\| = \|A\|^2$$

y de:

$$\|Ax\|^2 = (Ax, Ax) = (A^+Ax, x) \leq \|A^+Ax\| \|x\| \leq \|A^+A\| \|x\|^2$$

lo que implica que

$$\|A^+A\| \geq \|A\|^2$$

Una clase importante entre los operadores acotados es la de los que coinciden con su adjunto.

Definición 5.5.2 Sea A un operador lineal acotado definido sobre un espacio de Hilbert \mathcal{H} . Se dice que A es autoadjunto si

$$A = A^+$$

Se tiene entonces que $A : \mathcal{H} \rightarrow \mathcal{H}$, lineal y acotado, es autoadjunto si y solo si:

$$(x, Ay) = (Ax, y), \quad x, y \in \mathcal{H}$$

o también si y solo si (Ax, x) es real para todo $x \in \mathcal{H}$. La norma de un operador autoadjunto se puede calcular mediante la expresión:

$$\|A\| = \sup\{|(x, Ax)| : x \in \mathcal{H}, \|x\| = 1\}$$

En efecto,

$$|(x, Ax)| \leq \|A\| \|x\|^2$$

y por tanto (si $x \neq 0$):

$$\frac{|(x, Ax)|}{\|x\|^2} \leq \|A\|$$

de donde

$$k = \sup \left\{ \frac{|(x, Ax)|}{\|x\|^2}, x \neq 0 \right\} \leq \|A\|$$

La desigualdad en sentido contrario se puede obtener de la forma siguiente. Aplicando A a una combinación con dos vectores $x, y \in \mathcal{H}$ y siendo $\lambda \in \mathbf{K}$ se tiene:

$$\begin{aligned} k\|x + \lambda y\|^2 &\geq |(x + \lambda y, A(x + \lambda y))| \\ &= |(x, Ax) + |\lambda|^2(y, Ay) + 2\operatorname{Re}(\lambda(x, Ay))| \end{aligned}$$

Cambiando λ por $-\lambda$

$$\begin{aligned} k\|x - \lambda y\|^2 &\geq |(x - \lambda y, A(x - \lambda y))| \\ &= |(x, Ax) + |\lambda|^2(y, Ay) - 2\operatorname{Re}(\lambda(x, Ay))| \end{aligned}$$

Es decir,

$$\begin{aligned} (x, Ax) + |\lambda|^2(y, Ay) + 2\operatorname{Re}(\lambda(x, Ay)) &\leq k\|x + \lambda y\|^2 \\ -(x, Ax) - |\lambda|^2(y, Ay) + 2\operatorname{Re}(\lambda(x, Ay)) &\leq k\|x - \lambda y\|^2 \end{aligned}$$

y sumando

$$4\operatorname{Re}(\lambda(x, Ay)) \leq k(\|x + \lambda y\|^2 + \|x - \lambda y\|^2) = 2k(\|x\|^2 + |\lambda|^2\|y\|^2)$$

Si ahora tomamos

$$\lambda = \frac{\overline{(x, Ay)}\|x\|}{(x, Ay)\|y\|}$$

se obtiene

$$\operatorname{Re} \overline{(x, Ay)} \leq k\|x\|\|y\|$$

para toda pareja de vectores $x, y \in \mathcal{H}$. Eligiendo $x = Ay$, se tiene:

$$\|Ay\|^2 = |(Ay, Ay)| \leq k\|Ay\|\|y\|$$

Luego $\|Ay\| \leq k\|y\|$ y por tanto $\|A\| \leq k$ lo que demuestra el resultado.

Un operador autoadjunto es evidentemente normal, es decir conmuta con su adjunto. Como veremos los operadores unitarios son también normales y comparten muchas propiedades con los autoadjuntos debido a este hecho. Los operadores autoadjuntos son fundamentales en las aplicaciones de la teoría en la Mecánica Cuántica, en la que los observables vienen representados por este tipo de operadores. Téngase en cuenta que cuando el dominio del operador no es todo \mathcal{H} , es posible que el dominio de su operador adjunto no coincida con el del operador de partida. Si se verifica $(x, Ay) = (Ax, y)$ para todo par de vectores $x, y \in D(A)$, se dice que el operador es simétrico (o a veces hermítico) (con $D(A^+) \subset D(A)$), reservando el término autoadjunto para el caso de operadores simétricos cuyos dominios coincidan. En ciertos casos es posible definir lo que se denomina extensiones autoadjuntas, pero no podemos entrar en este tema aquí.

Consideremos por ejemplo el operador posición en Mecánica Cuántica:

$$(Qf)(x) = xf(x)$$

Está claro que su dominio no es todo $L^2(\mathbf{R})$. Se puede elegir:

$$D(Q) = \{f \in L^2(\mathbf{R}) : xf \in L^2(\mathbf{R})\}$$

que no es todo $L^2(\mathbf{R})$, pero que es denso en él. Efectivamente, no es difícil comprobar que el dominio contiene a las funciones continuas en un compacto (en cualquiera) y por lo tanto a todas las funciones de soporte compacto. Pero sabemos que el conjunto de funciones con soporte compacto es denso en $L^2(\mathbf{R})$ (en cualquier $L^p(\mathbf{R})$). Por lo tanto tenemos un operador con dominio denso en $L^2(\mathbf{R})$. Sin embargo este operador no es acotado en su dominio. Consideremos las funciones $\chi_{[-n, +n]}(x)$ que están obviamente en $D(Q)$. Su norma L^2 es:

$$\int_{\mathbf{R}} |\chi_{[-n, +n]}(x)|^2 dx = \int_{[-n, +n]} dx = 2n$$

Sin embargo, la norma de las imágenes es:

$$\int_{\mathbf{R}} |x\chi_{[-n, +n]}(x)|^2 dx = \int_{[-n, +n]} x^2 dx = \frac{2n^3}{3}$$

y el cociente es:

$$\frac{\|Q\chi_{[-n, +n]}\|}{\|\chi_{[-n, +n]}\|} = \frac{n^2}{3}$$

que no está acotado.

El operador adjunto (que se puede definir, aunque Q no sea acotado) verifica:

$$(Q^+f, g) = (f, Qg), \quad f \in D(Q^+), \quad g \in D(Q)$$

En este caso, el operador Q es hermítico (simétrico), pues:

$$(Qf, g) = (f, Qg), \quad f, g \in D(Q)$$

y, por tanto (de la construcción del operador adjunto),

$$D(Q) \subset D(Q^+)$$

Se puede demostrar también que ambos dominios coinciden (pero no lo haremos aquí) y el operador es autoadjunto.

Otro tipo de operadores importantes, de los que ya hemos visto un ejemplo cuando estudiamos la descomposición ortogonal del espacio de Hilbert (en realidad, como veremos, todos son así), son los proyectores ortogonales.

Definición 5.5.3 *Se dice que un operador acotado definido en un espacio de Hilbert \mathcal{H} es un proyector ortogonal si es autoadjunto ($P = P^+$) e idempotente ($P^2 = P$).*

Al ser acotado, el núcleo de P es un subespacio cerrado y por tanto

$$\mathcal{H} = \ker P \oplus (\ker P)^\perp$$

(suma directa ortogonal). Pero se tiene que

$$(\ker P)^\perp = \text{Ran } P$$

pues si $x \in \text{Ran } P$, entonces $x = Px$ por ser P idempotente. De esta forma,

$$(x, y) = (Px, y) = (x, Py) = 0$$

cuando $y \in \ker P$. Luego $\text{Ran } P \subset (\ker P)^\perp$. Si $x \in (\ker P)^\perp$, entonces, para todo $y \in \mathcal{H}$:

$$(x - Px, y) = (x, y - Py) = 0$$

pues $y - Py \in \ker P$ y por tanto, $x - Px = 0$ y $x \in \text{Ran } P$.

Por tanto, todo proyector induce una descomposición ortogonal del espacio de Hilbert, en suma de su núcleo y su rango. Y viceversa, como hemos visto, dado un subespacio cerrado, M , se tiene un proyector del cuál es el rango. Veamos que es acotado.

$$\mathcal{H} = M \oplus M^\perp, \quad x = y + z, \quad y \in M, \quad z \in M^\perp$$

$$Px = y, \quad \|Px\|^2 = \|y\|^2 \leq \|y\|^2 + \|z\|^2 = \|x\|^2$$

luego $\|P\| \leq 1$ y si $x \in M$, $Px = x$. y la norma de un proyector ortogonal es igual a 1. El ser autoadjunto e idempotente es inmediato (como ya vimos en su momento)

Un operador en un espacio de Hilbert puede actuar de forma irreducible o reducible. En el primer caso, no existe ningún subespacio cerrado (nos limitaremos a este caso) invariante excepto el total y el $\{0\}$. En el segundo existe algún subespacio no trivial tal que $AM \subset M$. Pero en este segundo caso se puede uno preguntar si el subespacio complementario a M es también invariante. De entre los posibles subespacios complementarios (cerrados), el más interesante es el ortogonal. Se dice que un subespacio reduce a un operador (el operador es completamente reducible) si tanto M como M^\perp son invariantes bajo A . Como se puede probar, un subespacio reduce a un operador si y solo si es invariante bajo A y A^+ (consecuencia inmediata de que si M es un subespacio invariante para un operador A , su complementario ortogonal es invariante bajo A^+).

Es muy sencillo construir un operador que sea reducible pero no completamente reducible. Consideremos el espacio de Hilbert l^2 y en él el operador de desplazamiento:

$$A(x_1, x_2, \dots) = (0, x_1, x_2, \dots)$$

Su operador adjunto se calcula fácilmente:

$$A^+(x_1, x_2, \dots) = (x_2, x_3, \dots)$$

El operador A es reducible. Cualquier subespacio N con vectores de la forma:

$$(0, \dots, 0, x_m, x_{m+1}, \dots)$$

es invariante bajo A . Su ortogonal, formado por vectores:

$$(x_1, \dots, x_{m-1}, 0, \dots)$$

lo es bajo A^+ . Sin embargo, N no es invariante bajo A^+ , luego este tipo de subespacios no reduce a A .

Pero aún más. No hay ningún subespacio no trivial que reduzca a A . Supongamos que M fuera uno de tales subespacios. Sea

$$x = (0, \dots, 0, x_m, x_{m+1}, \dots)$$

un vector no nulo de M , con $x_m \neq 0$. Aplicando m veces A^+ se tiene:

$$(A^+)^m x = (x_{m+1}, x_{m+2}, \dots)$$

y m veces A :

$$A^m (A^+)^m x = (0, \dots, 0, x_{m+1}, x_{m+2}, \dots)$$

que, debido a la invariancia de M bajo A y A^+ , es un vector de M y por lo tanto:

$$x - A^m (A^+)^m x = (0, \dots, 0, x_m, 0, \dots) \in M$$

Es decir, los vectores de la base ortonormal canónica de l^2 están en M (al estar uno de ellos y ser invariante bajo A y A^+ es inmediato que están todos). Por lo tanto $M = l^2$.

Resulta entonces que las condiciones de invariancia de un subespacio se pueden escribir en términos de proyectores. Así es posible demostrar que un subespacio (cerrado) es invariante bajo un operador A (siempre acotado) si su proyección asociada P verifica

$$AP = PA$$

De forma similar, un subespacio reduce a un operador si su proyector conmuta con él.

Dados dos proyectores, P_1 y P_2 su suma no es en general un proyector. Sin embargo si $P_1 P_2 = P_2 P_1 = 0$, entonces su suma es un proyector sobre el espacio suma de los asociados a P_1 y P_2 (que son necesariamente ortogonales). Para poder estudiar en detalle la teoría espectral de operadores y los teoremas de descomposición espectral, es necesario introducir sumas (no necesariamente finitas ni siquiera numerables) de proyectores y una relación de orden (asociada a la inclusión de los correspondientes subespacios). Pero no podemos ir por ahora más lejos en este asunto.

El tercer tipo de operadores que vamos a estudiar es fundamental en Mecánica Cuántica: operadores unitarios (operadores de evolución).

Definición 5.5.4 *Se dice que un operador lineal acotado en un espacio de Hilbert es unitario si*

$$UU^+ = U^+U = 1_{\mathcal{H}}$$

Si un operador es unitario conserva el producto escalar:

$$(Ux, Uy) = (x, y), \quad \forall x, y \in \mathcal{H}$$

De hecho esta propiedad, unida a que el operador sea biyectivo, define a los operadores unitarios. La norma de un operador unitario es igual a 1. Los operadores que conservan el producto escalar (y por tanto de la norma) sin ser necesariamente biyectivos se llaman isometrías. Existe otra clase de operadores llamados isometrías parciales que son fundamentales, por ejemplo, en la teoría de colisiones. Una isometría parcial es un operador que restringido al ortogonal de su núcleo es una isometría.

Por ejemplo, el operador desplazamiento en l^2 :

$$T(x_1, x_2, \dots) = (0, x_1, x_2, \dots)$$

es isométrico, sin embargo no es unitario. No es biyectivo (su rango son los vectores con primera coordenada cero). Su adjunto es

$$T^+(x_1, x_2, \dots) = (x_2, x_3, \dots)$$

y se tiene:

$$T^+T(x_1, x_2, \dots) = (x_1, x_2, x_3, \dots)$$

luego $T^+T = I$, pero $TT^+ \neq I$.

La transformada de Fourier en $L^2(\mathbf{R})$

Un operador unitario muy importante en las aplicaciones a la Física es la transformada de Fourier. Su definición usual es

$$(\mathcal{F}f)(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} f(x) dx$$

Para definirla en el espacio $L^2(\mathbf{R})$ utilizaremos una base ortonormal. Sea la base formada por las funciones de Hermite:

$$\varphi_n(x) = \frac{1}{(n!2^n\sqrt{\pi})^{1/2}} e^{-x^2/2} H_n(x), \quad n = 0, 1, 2, \dots$$

Sobre esta base, la transformada de Fourier se define por:

$$(\mathcal{F}\varphi_n)(k) = (-i)^n \varphi_n(k), \quad n = 0, 1, \dots$$

(se puede comprobar que coincide con la definición anterior). A partir de aquí, por linealidad lo construimos para cualquier combinación lineal finita de las funciones de esta base y viene dada por la integral anterior. Se puede probar fácilmente que es biyectivo (sobre $\text{lin}\{\varphi_n\}$) e isométrico. Su inverso viene dado por:

$$(\mathcal{F}^{-1}g)(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} g(k) dk$$

Finalmente se prueba que existe un único operador lineal acotado en $L^2(\mathbf{R})$ que coincide con éste en el subespacio $\text{lin}\{\varphi_n\}$. La actuación sobre funciones arbitrarias de $L^2(\mathbf{R})$ se define por:

$$(\mathcal{F}f)(k) = \frac{i}{\sqrt{2\pi}} \frac{d}{dk} \int_{-\infty}^{\infty} \frac{e^{-ikx} - 1}{x} f(x) dx$$

$$(\mathcal{F}^{-1}g)(x) = -\frac{i}{\sqrt{2\pi}} \frac{d}{dx} \int_{-\infty}^{\infty} \frac{e^{ikx} - 1}{k} g(k) dk$$

En el caso en que la función f sea integrable en módulo en \mathbf{R} , es decir esté en $L^1(\mathbf{R})$ además de en $L^2(\mathbf{R})$, es sencillo probar que las fórmulas anteriores se reducen a:

$$(\mathcal{F})(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} f(x) dx$$

$$(\mathcal{F}^{-1}g)(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{iky} g(k) dk$$

5.6. Operadores positivos

En un espacio de Hilbert es posible introducir un concepto de operador positivo utilizando el producto escalar.

Definición 5.6.1 *Se dice que el operador acotado A en el espacio de Hilbert \mathcal{H} es positivo si $(Ax, x) \geq 0$ para todo x de \mathcal{H} .*

Para un operador positivo es posible definir su raíz cuadrada, como el único operador positivo que verifica que su cuadrado es el operador de partida. Con esta definición, el “módulo” de un operador arbitrario se define como la raíz cuadrada del producto de ese operador por su adjunto (producto que es positivo):

$$|A| = \sqrt{A^+A}$$

Usando estas definiciones es posible escribir la descomposición polar de un operador (similar a la de los números complejos).

Teorema 5.6.1 *Dado un operador lineal acotado en un espacio de Hilbert, existe una isometría parcial U tal que*

$$A = |A|U$$

La isometría U se determina de manera única si se exige que su núcleo sea igual al de A . El rango de U es el cierre del rango de A .

5.7. Operadores compactos

Los operadores compactos (o completamente continuos), $\mathcal{C}(\mathcal{H})$, forman una clase de operadores en espacios de Hilbert con numerosas aplicaciones. Su relación con los operadores integrales hace que les dediquemos unas breves líneas.

Definición 5.7.1 *Se dice que un operador lineal en un espacio de Hilbert es compacto si la imagen de un subconjunto compacto es precompacta (es decir su completación (o, en este caso, su cierre) es un compacto).*

La siguiente propiedad es equivalente a la definición: un operador es compacto si toda sucesión acotada tiene como imagen una sucesión que admite subsucesiones convergentes. Se puede probar que un operador compacto transforma sucesiones débilmente convergentes (en la topología débil, es decir, x_n converge débilmente a x_0 si (x_n, y) converge en \mathbf{C} para todo $y \in \mathcal{H}$ a (x_0, y)) en sucesiones convergentes en el sentido de la norma. Todo operador compacto es continuo (en dimensión finita coinciden con todos los operadores). La norma de un operador compacto se puede obtener para algún vector $x \in \mathcal{H}$:

$$\|Tx\| = \|T\|, \quad \|x\| = 1$$

Además si un operador es compacto, su adjunto también lo es.

Un caso sencillo de operadores compactos son los de rango finito (es decir, el rango de T es un subespacio vectorial de dimensión finita). Pero se puede decir aún más. Un operador compacto en un espacio de Hilbert separable es el límite (en norma) de una sucesión de operadores de rango finito. Esto hace que muchas de las propiedades de los espacios de dimensión finita puedan ser trasladadas a dimensión infinita. En particular los relacionados con las soluciones de la ecuación $Tx = \lambda x$. En el caso de dimensión finita esta ecuación tiene solución no trivial o bien el operador $(T - \lambda 1_{\mathcal{H}})^{-1}$ existe. En el caso infinito la situación es más complicada. Pero para operadores compactos se sigue dando esa propiedad. Lo veremos con más detalle cuando estudiemos los espectros de operadores.

5.8. Operadores de clase de traza y operadores Hilbert-Schmidt

La traza de un operador en un espacio de Hilbert se puede definir de la forma siguiente. Dada una base ortonormal $\{u_n\}$ y un operador acotado positivo, la traza de A es:

$$\operatorname{tr} A = \sum_{n=1}^{\infty} (u_n, Au_n)$$

La traza no depende de la base elegida. Se dice que un operador es de la clase de traza si la traza de su módulo (en el sentido definido en la sección 5.6) es finita. Los operadores de clase de traza forman un ideal en el espacio de los operadores acotados. En este espacio es posible definir una norma, $\|\cdot\|_1 = \operatorname{tr} |A|$. De esta forma, el espacio de los operadores de clase de traza es un espacio de Banach. Además

$$\|A\| \leq \|A\|_1$$

Los operadores de clase de traza son compactos. Para operadores no necesariamente positivos (pero en $\mathcal{C}(\mathcal{H})$), la traza se define por:

$$\operatorname{tr} A = \sum_{n=1}^{\infty} (u_n, Au_n)$$

pues se puede demostrar que esta suma es convergente y además no depende de la base ortonormal elegida (se usan conceptos de operadores de Hilbert-Schmidt que definiremos a continuación). Uno de los aspectos más interesantes de los operadores de clase de traza es que son el dual del espacio de los operadores compactos (usando la traza).

Los operadores Hilbert-Schmidt son una clase especial de operadores compactos, con interesantes propiedades algebraicas. Sea A un operador lineal acotado que verifica que para alguna base ortonormal

u_n se tiene $\sum_n \|Au_n\|^2 < \infty$. Es decir, tal que $\text{tr } AA^+$ es finita. Se escribe

$$\|A\|_2 = \sqrt{\sum_{n=1}^{\infty} \|Au_n\|^2}$$

Definición 5.8.1 *Se dice que un operador acotado es Hilbert-Schmidt si*

$$\|A\|_2 < \infty$$

La norma (usual) de un operador Hilbert-Schmidt es siempre menor o igual que la norma Hilbert-Schmidt ($\|\cdot\|_2$ es una norma en la clase de operadores Hilbert-Schmidt). El conjunto de los operadores de tipo Hilbert-Schmidt es un espacio de Banach con esta norma $\|\cdot\|_2$ (de hecho, presenta una estructura más rica, pero no entraremos ello). Más aún, el conjunto de operadores de tipo Hilbert-Schmidt es un espacio de Hilbert. Se le denota por $\mathcal{C}_2(\mathcal{H})$. Todo operador Hilbert-Schmidt es un operador de la clase de traza y por tanto un operador compacto

Una de las aplicaciones más importantes de los operadores Hilbert-Schmidt se basa en su relación con las ecuaciones integrales que estudiaremos a continuación.

5.9. Ecuaciones integrales

Comenzaremos por establecer la relación existente entre los operadores integrales y los operadores de Hilbert-Schmidt en una realización concreta.

Sea $\mathcal{H} = L^2(I)$, donde I es un intervalo de la recta real (en general, las funciones de cuadrado integrable en un cierto espacio de medida). Si T es un operador de Hilbert-Schmidt en \mathcal{H} , entonces existe una función de cuadrado integrable en $I \times I$, $k(x, y)$,

$$\int_{I \times I} |k(x, y)|^2 dx dy < \infty$$

tal que el operador puede definirse por:

$$(Tf)(x) = \int_I k(x, y)f(y) dy$$

Además, la norma Hilbert-Schmidt es:

$$\|A\|_2^2 = \int_{I \times I} |k(x, y)|^2 dx dy$$

El anterior es un operador integral aunque no todos los operadores integrales son Hilbert-Schmidt. Definiremos primeramente lo que es el núcleo de un operador integral.

Definición 5.9.1 *Sea $L^2(I)$ el espacio de Hilbert de las funciones de cuadrado integrable con valores en \mathbf{C} definidas en el intervalo I de la recta real. Un núcleo es una función $k(x, y)$ medible Borel en $I \times I$, tal que la integral*

$$\int_I |k(x, y)f(y)| dy$$

está en $L^2(I)$ para toda función $f \in L^2(I)$.

Con las funciones $k(x, y)$ es posible definir un operador en el espacio $L^2(I)$:

$$K : L^2(I) \rightarrow L^2(I), \quad (Kf)(x) = \int_I k(x, y)f(y) dy$$

Se trata de un operador lineal continuo (de núcleo $k(x, y)$). El núcleo define obviamente al operador, y viceversa, si dos operadores integrales son iguales, entonces sus núcleos son iguales casi doquiera. Dado

un operador integral (con núcleo $k(x, y)$), su adjunto es otro operador integral cuyo núcleo es $\overline{k(y, x)}$. Por tanto un operador integral es autoadjunto si y solo si el núcleo es simétrico, es decir:

$$k(x, y) = \overline{k(y, x)}$$

A partir de un operador integral K , se pueden construir las potencias K^n . Los núcleos de estos operadores se llaman núcleos iterados y verifican la ecuación de recurrencia:

$$k_n(x, y) = \int_I k(x, t)k_{n-1}(t, y) dt, \quad k_1(x, y) = k(x, y)$$

Un operador integral se llama de tipo compacto si K^n es un operador compacto para algún n .

Asociadas a este tipo de operadores se pueden considerar ecuaciones integrales que se presentan muy a menudo en Física. Se llaman ecuaciones integrales de Fredholm homogéneas, de primera y de segunda clase y corresponden al esquema (donde f es la función a calcular):

$$(K - \lambda)f = g$$

Si $g = 0$ se llaman homogéneas. Si $\lambda = 0$ se llaman de primera clase. Y en el caso general, de segunda clase. Si el núcleo $k(x, y)$ verifica

$$k(x, y) = 0, \quad x < y$$

se llaman ecuaciones de Volterra.

Como veremos en el siguiente tema, la resolución de las ecuaciones integrales es equivalente a obtener el espectro de los operadores integrales asociados.

Capítulo 6

Espectros de operadores

6.1. Introducción

El análisis del espectro de un operador en un espacio de Hilbert es fundamental en las aplicaciones de esta teoría, especialmente en Mecánica Cuántica. No podemos entrar aquí en todas las implicaciones que conlleva, especialmente en el caso de operadores no acotados, pero intentaremos al menos dar algunos resultados que pueden ser útiles para ciertas clases de operadores.

La situación es muy diferente en los casos de dimensión finita e infinita. Recuérdese que dado un operador en un espacio de Hilbert de dimensión finita, un escalar λ es un autovalor si existe un vector no nulo tal que

$$Ax = \lambda x$$

Se dice entonces que x es un autovector con autovalor λ . Y no hay ningún otro elemento esencial en la teoría.

Sea T un operador (no necesariamente acotado) en un espacio de Hilbert \mathcal{H} cuyo dominio es denso en \mathcal{H} (aunque no tiene por qué ser igual a \mathcal{H}). El operador fundamental a la hora de clasificar los puntos del plano complejo en relación a T es el operador

$$(T - \lambda 1_{\mathcal{H}})^{-1}$$

Este operador puede o no existir, ser acotado o no, y tener un dominio u otro. De acuerdo con estas características se tiene lo siguiente (el operador $(T - \lambda 1_{\mathcal{H}})^{-1}$ caso de existir se supone que está definido en el rango de $T - \lambda 1_{\mathcal{H}}$ y tiene como rango el dominio de ese operador).

Definición 6.1.1 *Clasificación de los puntos de \mathbf{C} atendiendo al operador T .*

1. Se dice que $\lambda \in \mathbf{C}$ está en el espectro puntual ($\lambda \in \sigma_p(T)$) del operador T si no existe el operador $(T - \lambda 1_{\mathcal{H}})^{-1}$.
2. Se dice que λ está en el espectro residual del operador T ($\lambda \in \sigma_r(T)$) si el operador $(T - \lambda 1_{\mathcal{H}})^{-1}$ existe, pero su dominio (es decir, el rango del operador $T - \lambda 1_{\mathcal{H}}$) no es denso en el espacio \mathcal{H} .
3. Se dice que λ está en el espectro continuo de T ($\lambda \in \sigma_c(T)$) si existe $(T - \lambda 1_{\mathcal{H}})^{-1}$ y su dominio es denso en \mathcal{H} , pero no es acotado en su dominio.
4. Se dice que λ está en el resolvente de T ($\lambda \in \rho(T)$) si existe $(T - \lambda 1_{\mathcal{H}})^{-1}$, su dominio es denso en \mathcal{H} y es acotado en su dominio.

Esta definición clasifica a todos los puntos del plano complejo en cuatro conjuntos disjuntos. En el caso de dimensión finita solo existe espectro puntual y conjunto resolvente.

El operador $R_{\lambda}(T) = (T - \lambda 1_{\mathcal{H}})^{-1}$ (que se llama operador resolvente) existe (en cuyo caso estamos en el conjunto resolvente) o no (en cuyo caso estamos en el espectro, que es solo puntual cuando la dimensión es finita).

Cuando el operador es acotado el espectro es no vacío.

6.2. El espectro puntual y el espectro continuo

El espectro puntual es en todos sus aspectos igual en dimensión finita e infinita. Si un punto pertenece a σ_p , existe un vector del espacio (autovector), que verifica $Tx = \lambda x$. Esto es debido a la no existencia del operador inverso. Evidentemente puede existir más de un vector l.i. con esta propiedad. El conjunto de los autovectores asociados a un autovalor es el espacio propio de ese autovalor y es un espacio invariante bajo el operador T . Pero puede ser de dimensión infinita. Si dos autovectores corresponden a autovalores distintos son linealmente independientes.

Supongamos el operador posición en $L^2(\mathbf{R})$. Como hemos demostrado es un operador con dominio denso en $L^2(\mathbf{R})$, pero no es acotado. Su espectro puntual es vacío pues de:

$$(\lambda - x)f(x) = 0$$

se tiene que $f(x) = 0$ casi doquiera. Al ser autoadjunto no existe espectro residual (como veremos más adelante). Su espectro continuo es toda la recta. Podemos usar el siguiente criterio para demostrarlo.

Proposición 6.2.1 *Sea T un operador en un espacio de Hilbert con dominio $D(T)$. El operador $(T - \lambda 1_{\mathcal{H}})^{-1}$ no existe o si existe no está acotado en su dominio si y solo si existe una sucesión de vectores en el dominio de T de norma 1 tal que $T - \lambda 1_{\mathcal{H}}$ aplicado a esa sucesión, nos da una sucesión que tiende a 0 cuando $n \rightarrow \infty$.*

Supongamos que existe dicha sucesión x_n y que existe el operador $(T - \lambda 1_{\mathcal{H}})^{-1}$. Construimos la sucesión:

$$y_n = \frac{1}{\|(T - \lambda 1_{\mathcal{H}})x_n\|} (T - \lambda 1_{\mathcal{H}})x_n$$

que es una sucesión de vectores de norma 1. Aplicando $(T - \lambda 1_{\mathcal{H}})^{-1}$

$$(T - \lambda 1_{\mathcal{H}})^{-1}y_n = \frac{1}{\|(T - \lambda 1_{\mathcal{H}})x_n\|} x_n$$

obtenemos una sucesión de vectores de norma no acotada. Luego $(T - \lambda 1_{\mathcal{H}})^{-1}$ no está acotado. Supongamos ahora que no existe el operador inverso $(T - \lambda 1_{\mathcal{H}})^{-1}$. Eso implica que existe $x \in \mathcal{H}$ (que podemos elegir de norma 1), tal que

$$(T - \lambda 1_{\mathcal{H}})x = 0$$

En este caso, podemos definir $x_n = x$ para todo n . La sucesión $(T - \lambda 1_{\mathcal{H}})x_n \rightarrow 0$ obviamente. La otra situación es que exista el inverso $(T - \lambda 1_{\mathcal{H}})^{-1}$ pero no esté acotado. Existe por tanto una sucesión, y_n , de vectores de norma 1 en el dominio de $(T - \lambda 1_{\mathcal{H}})^{-1}$, tales que:

$$\|(T - \lambda 1_{\mathcal{H}})^{-1}y_n\| \rightarrow \infty$$

Entonces la sucesión

$$x_n = \frac{1}{\|(T - \lambda 1_{\mathcal{H}})^{-1}y_n\|} (T - \lambda 1_{\mathcal{H}})^{-1}y_n$$

cumple las condiciones de la proposición. Esto acaba la demostración.

Si un punto $\lambda \in \mathbf{C}$ está en el espectro puntual, el operador $(T - \lambda 1_{\mathcal{H}})^{-1}$ no existe y la sucesión de la que habla la proposición se puede construir como hemos visto en la demostración (y de otras formas usando diferentes autovectores con el mismo autovalor) aunque posiblemente contenga infinitos términos iguales. Si λ está en el espectro residual, la sucesión no puede contener infinitos términos iguales.

Demostremos ahora que el espectro continuo del operador posición es toda la recta real (por ser autoadjunto, como veremos más adelante, el espectro es real). Consideremos la sucesión de funciones de $L^2(\mathbf{R})$:

$$f_n(x) = \left(\frac{n^2}{\pi}\right)^{1/4} e^{-n^2(x-\lambda)^2/2}$$

que tienen norma 1. La sucesión de sus imágenes mediante $Q - \lambda 1_{\mathcal{H}}$

$$(Q - \lambda 1_{\mathcal{H}})f_n = \left(\frac{n^2}{\pi}\right)^{1/4} (x - \lambda)e^{-n^2(x-\lambda)^2/2}$$

tienen norma

$$\|(Q - \lambda 1_{\mathcal{H}})f_n\| = \left(\frac{n^2}{\pi}\right)^{1/2} \int_{\mathbf{R}} (x - \lambda)^2 e^{-n^2(x-\lambda)^2} dx = \frac{1}{2n^2}$$

que tiende a 0 cuando $n \rightarrow \infty$.

6.3. El resolvente

En el caso de dimensión finita el resolvente es el conjunto de puntos que no son autovalores. Sin embargo, en dimensión infinita la situación es más compleja. Si nos limitamos a operadores acotados definidos en \mathcal{H} lo único que hay que probar es que para que un punto esté en el resolvente, el dominio de $(T - \lambda 1_{\mathcal{H}})^{-1}$ debe ser denso en \mathcal{H} (lo que por otra parte equivale a que este operador sea acotado en todo \mathcal{H}). Se puede demostrar que el conjunto resolvente es siempre un abierto del plano:

Teorema 6.3.1 *Sea \mathcal{H} un espacio de Hilbert y T un operador lineal acotado. Entonces, el conjunto resolvente es un abierto y el resolvente (como función de λ) es analítico en su dominio (en cada componente conexa para ser más precisos, con valores en el espacio de operadores acotados de \mathcal{H}). Para dos puntos cualquiera del conjunto resolvente, los operadores resolventes asociados conmutan y se tiene:*

$$R_{\lambda}(T) - R_{\mu}(T) = (\mu - \lambda)R_{\mu}(T)R_{\lambda}(T)$$

El resolvente se puede escribir usando la serie de Neumann. Desarrollando formalmente:

$$R_{\lambda}(T) = (T - \lambda 1_{\mathcal{H}})^{-1} = \frac{1}{T - \lambda 1_{\mathcal{H}}} = -\frac{1}{\lambda} \frac{1}{1 - T/\lambda} = -\frac{1}{\lambda} \left(1 + \sum_{n=1}^{\infty} \left(\frac{T}{\lambda}\right)^n\right)$$

que es una serie que converge cuando $\lambda > \|T\|$.

El espectro (la unión del puntual, el continuo y el residual) es un cerrado. Es más, si el operador está acotado, el espectro es no vacío, compacto y contenido en el disco de centro el origen y radio la norma del operador. Si el operador es autoadjunto, el radio espectral (el supremo de los puntos del espectro) es igual a la norma del operador.

6.4. El espectro de un operador y de su adjunto

Entre el espectro de un operador y el de su operador adjunto existen unas relaciones que generalizan las sencillas del caso de dimensión finita (donde si λ es un autovalor, su conjugado es un autovalor del operador adjunto). Así, es posible probar el siguiente resultado para operadores acotados con dominio en \mathcal{H} .

Proposición 6.4.1 *Los puntos del resolvente de un operador son los conjugados del resolvente del operador adjunto. De la misma forma los puntos del espectro continuo de un operador son conjugados a los del espectro continuo del operador adjunto. Sin embargo si un punto está en el espectro puntual de un operador su conjugado puede estar en el puntual del adjunto o en el residual. Pero si un punto está en el residual de un operador, su conjugado está en el puntual del adjunto.*

6.5. Espectro de operadores acotados normales

Un operador acotado es normal si conmuta con su adjunto. La propiedad más importante (en relación al espectro) que verifican los operadores normales es que si un punto del plano complejo es un autovalor, entonces su conjugado es un autovalor del operador adjunto (recuérdese que para operadores arbitrarios lo más que puede decirse es que es un autovalor o está en el espectro residual). Además si dos autovectores corresponden a autovalores distintos, entonces son ortogonales. Ambas propiedades son inmediatas de demostrar y se corresponden al caso de dimensión finita. El espectro residual es vacío para operadores normales, es decir no hay puntos del plano complejo para los que exista $(T - \lambda 1_{\mathcal{H}})^{-1}$ y el rango de $T - \lambda 1_{\mathcal{H}}$ no sea denso en \mathcal{H} .

Si U es un operadores unitario, entonces es normal y por lo tanto se le aplican los resultados anteriores. Pero además, el espectro (que consta de puntual y continuo) se encuentra sobre la circunferencia unidad.

Proposición 6.5.1 *Si U es un operador unitario, todo punto del espectro tiene módulo 1.*

De forma similar, para los operadores autoadjuntos se tiene la siguiente proposición.

Proposición 6.5.2 *Sea A un operador acotado autoadjunto en \mathcal{H} . Entonces, los puntos del espectro son reales y están contenidos en el intervalo $[\inf\{(v, Av) : \|v\| = 1\}, \sup\{(v, Av) : \|v\| = 1\}]$. Se tiene además que los extremos de este intervalo están en el espectro de A .*

De aquí se deduce, usando el resultado sobre la norma que presentamos cuando definimos los operadores autoadjuntos, que la norma de un operador autoadjunto es igual al máximo de los valores absolutos de los extremos del intervalo en el que está contenido el espectro puntual.

Para proyectores ortogonales el resultado es el siguiente.

Proposición 6.5.3 *Sea P un proyector ortogonal distinto de 0 y la identidad. Entonces, su espectro es solo puntual e igual a $\{0, 1\}$.*

6.6. Espectro de operadores compactos

Para los operadores compactos se tienen resultados más precisos sobre la estructura de su espectro. Se puede demostrar el siguiente resultado.

Proposición 6.6.1 *Sea A un operador compacto en un espacio de Hilbert. Se tiene entonces que la suma de la dimensiones de los subespacios propios correspondientes a autovalores $\lambda > k > 0$ para todo $k > 0$, es finita. Además, el espectro puntual es numerable (o finito) y el único punto de acumulación (en el caso de que tenga alguno) es el 0. Todos los puntos del plano complejo, salvo el cero, están en el espectro puntual o en el resolvente y el cero está en el espectro.*

6.7. Ecuaciones integrales

Volvemos ahora al tema de las ecuaciones integrales, centrándonos ahora en las soluciones de las ecuaciones de Fredholm. En primer lugar estudiaremos los núcleos de rango finito. Consideremos el siguiente tipo de núcleos:

$$k(x, y) = \sum_{i=1}^n a_i(x)b_i(y)$$

La ecuación integral homogénea es

$$\int_I \sum_{i=1}^n a_i(x)b_i(y)f(y) dy = \lambda f(x)$$

que se puede escribir como:

$$\sum_{i=1}^n \left(\int_I b_i(y) f(y) dy \right) a_i(x) = \lambda f(x)$$

es decir, si $\lambda \neq 0$, f es una combinación lineal de las funciones $a_i(x)$, $i = 1, \dots, n$, y desde el punto de vista de operadores, se trata de calcular el espectro puntual de un operador de rango finito. Las soluciones deben ser:

$$f(x) = \sum_{i=1}^n \alpha_i a_i(x)$$

y sustituyendo en la ecuación:

$$\sum_{i=1}^n \left(\int_I b_i(y) \sum_{j=1}^n \alpha_j a_j(y) dy \right) a_i(x) = \lambda \sum_{i=1}^n \alpha_i a_i(x)$$

es decir

$$\sum_{j=1}^n \beta_{ij} \alpha_j = \lambda \alpha_i, \quad i = 1, \dots, n$$

donde

$$\beta_{ij} = \int_I b_i(y) a_j(y) dy$$

Definiendo la matriz $n \times n$, A , de elementos β_{ij} , el problema queda reducido a encontrar los autovalores no nulos de la matriz A .

Capítulo 7

Distribuciones

Las distribuciones (o funciones generalizadas) permiten justificar de una manera adecuada muchas operaciones en Física que no estaban adecuadamente formalizadas y encontrar nuevas e importantes aplicaciones. En esta breve introducción nos centraremos en las distribuciones sobre funciones de soporte compacto y de decrecimiento rápido. Las primeras como introducción al tema y las segundas para poder definir la transformada de Fourier y desarrollar algunas aplicaciones importantes.

7.1. Ideas básicas sobre distribuciones

7.1.1. El espacio de las funciones prueba de soporte compacto

El espacio \mathcal{D} es el conjunto de funciones en n variables reales con valores complejos en C^∞ que son de soporte compacto. Por soporte de una función φ ($\text{sop } \varphi$) entendemos el cierre del conjunto de puntos donde la función es distinta de cero.

El conjunto \mathcal{D} es un espacio vectorial complejo de dimensión infinita. El espacio \mathcal{D} es no vacío. Por ejemplo:

$$\varphi(x) = \begin{cases} 0 & |x| \geq 1 \\ \exp\left(-\frac{1}{1-x^2}\right) & |x| < 1 \end{cases} \quad (7.1)$$

Pero hay muchas más funciones en este espacio. En particular, toda función continua de soporte compacto se puede aproximar de manera uniforme por una función que esté en \mathcal{D} .

En el espacio \mathcal{D} se introduce la siguiente definición para el estudio de la convergencia de sucesiones. Sea $\varphi_n \in \mathcal{D}$, $n \in \mathbf{N}$. Se dice que la sucesión $\{\varphi_n\}$ es convergente en \mathcal{D} a una función $\varphi \in \mathcal{D}$ si existe un conjunto $K \subset \mathbf{R}^n$, acotado, que contiene a $\text{sop } \varphi_n$ para todo n y si en cada orden de derivadas, se tiene la convergencia uniforme de las derivadas de φ_n a las derivadas de la función φ .

7.1.2. El espacio de distribuciones

Una distribución, $u \in \mathcal{D}'$, es un funcional lineal continuo sobre \mathcal{D} . La continuidad se entiende en el sentido de que si φ_n converge a φ en \mathcal{D} , de acuerdo con la definición anterior, entonces $u(\varphi_n)$ converge a $u(\varphi)$ en \mathbf{C} .

El espacio de las distribuciones \mathcal{D}' es un espacio vectorial complejo de dimensión infinita (no todos los funcionales sobre \mathcal{D} son continuos, aunque no se puede escribir uno que no lo sea).

El primer ejemplo de distribuciones es el de las asociadas a funciones que son localmente integrables (integrables sobre acotados). Sea $f \in L^1_{\text{loc}}$ una función localmente integrable. Definimos el funcional:

$$u_f(\varphi) = \int_{\mathbf{R}^n} f(x)\varphi(x) dx, \quad \varphi \in \mathcal{D} \quad (7.2)$$

Es inmediato probar que es continuo, por tanto $u_f \in \mathcal{D}'$. Dos funciones localmente integrables que coinciden salvo en un conjunto de medida nula (casi doquiera) definen la misma distribución (y a la inversa).

El segundo ejemplo, por su importancia, es la delta de Dirac, δ :

$$\delta(\varphi) = \varphi(0), \quad \varphi \in \mathcal{D} \quad (7.3)$$

es una distribución sobre \mathcal{D} .

El soporte de una distribución es el cerrado más pequeño fuera del cuál la distribución es nula, es decir, aplicada a una función prueba cuyo soporte esté fuera de ese cerrado, el resultado es cero.

7.1.3. Operaciones con distribuciones

Las distribuciones se pueden sumar y multiplicar por escalares (\mathcal{D}' es un espacio vectorial). Además es posible definir la derivada de una distribución. Sea $u \in \mathcal{D}'(\mathbf{R})$.

$$u'(\varphi) = -u(\varphi') \quad (7.4)$$

En general,

$$u \in \mathcal{D}'(\mathbf{R}^n), \quad D_J u(\varphi) = (-1)^{|J|} u(D_J \varphi) \quad (7.5)$$

donde

$$D_J = \frac{\partial^{|J|}}{\partial x_1^{j_1} \dots \partial x_n^{j_n}}, \quad J = (j_1, \dots, j_n) \in \mathbf{N}^n, \quad |J| = j_1 + \dots + j_n \quad (7.6)$$

Las distribuciones tienen derivadas de todos los órdenes (y no dependen del orden en el que se efectúa las derivadas).

Sin embargo, no siempre es posible definir el producto de dos distribuciones. Sin embargo, el producto de una función en C^∞ por una distribución nos da una nueva distribución:

$$f u(\varphi) = u(f \varphi), \quad f \in C^\infty, \quad \varphi \in \mathcal{D} \quad (7.7)$$

7.1.4. Convergencia en el espacio de distribuciones

En \mathcal{D}' se define la siguiente noción de convergencia. Sea u_n una sucesión de distribuciones en \mathcal{D}' . Se dice que $u_n \rightarrow u$ en \mathcal{D}' si para toda función prueba $\varphi \in \mathcal{D}$ se tiene $u_n(\varphi) \rightarrow u(\varphi)$ en \mathbf{C} .

Si una sucesión de distribuciones u_n verifica la propiedad anterior, es decir, $u_n(\varphi)$ es convergente para todo $\varphi \in \mathcal{D}$, entonces existe una distribución que es el límite de la sucesión u_n (hay que probar que el límite es continuo).

La derivación es continua en \mathcal{D}' .

7.2. Topología de los espacios de funciones prueba y distribuciones

7.2.1. Topologías localmente convexas en espacios de funciones

Sea Ω un abierto no vacío de \mathbf{R}^n . Se sabe que Ω se puede escribir como la unión de una cantidad numerable de conjuntos compactos K_n no vacíos elegidos de forma que K_n esté contenido en el interior de K_{n+1} .

El espacio vectorial $\mathcal{C}(\Omega)$ es el conjunto de las funciones definidas sobre Ω , con valores complejos, continuas. En él se define una familia de seminormas:

$$p_n(f) = \sup\{|f(x)| : x \in K_n\} \quad (7.8)$$

(No se puede definir una norma del supremo sobre Ω que es un abierto y por tanto no compacto. Esta familia de seminormas separa el espacio, en el sentido que para cada función $f \neq 0$ existe una seminorma p_n tal que $p_n(f) \neq 0$).

Se tiene entonces que la familia de seminormas define una topología con una base local de entornos convexos:

$$V_n = \left\{ f \in \mathcal{C}(\Omega) : p_n(f) < \frac{1}{n} \right\} \quad (7.9)$$

y el espacio $\mathcal{C}(\Omega)$ está dotado de una topología localmente convexa. Además, la métrica:

$$d(f, g) = \sum_{n=1}^{\infty} \frac{2^{-n} p_n(f - g)}{1 + p_n(f - g)} \quad (7.10)$$

es compatible con la topología anterior, es decir el espacio es metrizable (el ser metrizable es una consecuencia de que el espacio posea una base numerable de entornos (IAN). El que la métrica sea compatible con la topología se comprueba demostrando que las bolas abiertas de centro el origen son una base local del espacio). Finalmente, las sucesiones de Cauchy son convergentes y por tanto el espacio es completo. En efecto, sea f_m una sucesión de Cauchy en $\mathcal{C}(\Omega)$. Entonces para cada seminorma (para cada n) se tiene

$$p_n(f_m - f_{m'}) \rightarrow 0, \quad m, m' \rightarrow \infty \quad (7.11)$$

Como p_n es el supremo en el compacto K_n está claro que la sucesión f_n converge uniformemente sobre cada compacto K_n a una cierta función continua $f \in \mathcal{C}(\Omega)$. Solo queda probar que $d(f_m, f)$ tiende a cero cuando $m \rightarrow \infty$.

Un espacio localmente convexo y completo se llama un espacio de Fréchet.

Hay que tener en cuenta que $\mathcal{C}(\Omega)$ no es un espacio normable (no es localmente acotado). En particular los entornos anteriores V_n no están acotados.

De forma similar se definen los espacios $\mathcal{C}^\infty(\Omega)$. Una función f está en $\mathcal{C}^\infty(\Omega)$ si $D^J f \in \mathcal{C}(\Omega)$ para todo multi-índice J .

El soporte de una función f es el cierre del conjunto donde f es distinta de cero. Sea K un compacto de \mathbf{R}^n . Se define:

$$\mathcal{D}_K = \{f \in \mathcal{C}^\infty(\Omega) : \text{sop } f \subset K\} \quad (7.12)$$

Los conjuntos \mathcal{D}_K son (o se pueden identificar con) subespacios de $\mathcal{C}^\infty(\Omega)$ si $K \subset \Omega$. En $\mathcal{C}^\infty(\Omega)$ se introduce una topología de la forma siguiente. Como antes, se elige una sucesión de compactos cuya unión es Ω y K_n contenido en el interior de K_{n+1} . También se define una familia de seminormas, pero ahora tomamos el máximo de las derivadas de f hasta un orden dado:

$$p_N(f) = \max\{|D^J f(x)| : x \in K_N, |J| \leq N\} \quad (7.13)$$

Como antes, estas seminormas definen en $\mathcal{C}^\infty(\Omega)$ una topología metrizable y localmente convexa. Los subespacios \mathcal{D}_K son cerrados (\mathcal{D}_K es la intersección de los espacios nulos de los funcionales de evaluación ($f \rightarrow f(x)$, que son continuos) cuando $x \in \mathbf{R}^n - K$).

También como en el caso de las funciones continuas (lo cierto es que este es un proceso general de construcción de topologías localmente convexas) se tiene una base de entornos convexos:

$$V_N = \left\{ f \in \mathcal{C}^\infty(\Omega) : p_N(f) < \frac{1}{N} \right\} \quad (7.14)$$

Si se tiene una sucesión de Cauchy en $\mathcal{C}^\infty(\Omega)$, f_n , entonces la sucesión $D^J f_n$ converge uniformemente en cualquier compacto de Ω . Como consecuencia, $f_n \rightarrow f$ y $D^J f_n \rightarrow D^J f$ en el espacio $\mathcal{C}^\infty(\Omega)$. Al igual que $\mathcal{C}(\Omega)$, $\mathcal{C}^\infty(\Omega)$ es un espacio de Fréchet y sus subespacios cerrados \mathcal{D}_K también lo son.

Además, el espacio $\mathcal{C}^\infty(\Omega)$ tiene la propiedad de Heine-Borel, los conjuntos cerrados y acotados son compactos. Debido a esta propiedad, este espacio no es normable (no existe una norma que defina esta topología) ya que no es localmente acotado (si lo fuera sería de dimensión finita y no lo es). Lo mismo ocurre con los espacios \mathcal{D}_K cuando K tiene interior no vacío (si no, el espacio es trivial).

7.2.2. El espacio de las funciones prueba de soporte compacto

Sea Ω un abierto no vacío de \mathbf{R}^n . El espacio de funciones prueba es

$$\mathcal{D}(\Omega) = \bigcup_K \mathcal{D}_K \quad (7.15)$$

donde $K \subset \Omega$, compacto. Es decir el espacio de las funciones sobre $\Omega \subset \mathbf{R}^n$ con valores complejos que son infinitamente derivables y tienen soporte compacto.

Sobre este espacio se puede definir un conjunto de normas

$$\|\varphi\|_N = \max\{D^J\varphi(x) : x \in \Omega, |J| \leq N\} \quad (7.16)$$

que cuando se restringen a \mathcal{D}_K son las seminormas que hemos discutido anteriormente. El problema es que la topología en $\mathcal{D}(\Omega)$ no es completa (una sucesión de funciones que es de Cauchy, puede tener un límite que no es de soporte compacto).

Como esta propiedad (el ser completa) es muy importante en las aplicaciones, se introduce otra topología que sí es completa. El inconveniente es que no será metrizable.

Sea Ω un abierto no vacío de \mathbf{R}^n . Si K es un compacto en Ω , podemos construir \mathcal{D}_K , con la topología (τ_K) que se ha introducido en la sección anterior. Estos espacio son de Fréchet (localmente convexos y completos). Sea β la clase de los conjuntos convexos y equilibrados ($\lambda W \subset W$ si $|\lambda| \leq 1$) de $\mathcal{D}(\Omega)$, W , que verifican que su intersección con cualquier \mathcal{D}_K es un abierto en ese espacio: $W \cap \mathcal{D}_K \in \tau_K$. Finalmente, τ es la clase de los conjuntos de la forma $\varphi + W$ donde $\varphi \in \mathcal{D}(\Omega)$ y $W \in \beta$ (la descripción anterior equivale a construir el límite inductivo de las topologías τ_K).

Se tiene entonces que τ es una topología en $\mathcal{D}(\Omega)$, que β es una base de entornos y $(\mathcal{D}(\Omega), \tau)$ es un espacio vectorial localmente convexo, aunque no es metrizable (una consecuencia del teorema de Baire).

Esta topología induce sobre \mathcal{D}_K la topología τ_K (\mathcal{D}_K es un subespacio de $\mathcal{D}(\Omega)$). Además verifica la propiedad de Heine-Borel, todos los conjuntos cerrados y acotados son compactos.

Las sucesiones de Cauchy en $\mathcal{D}(\Omega)$ están contenidas en algún \mathcal{D}_K (para algún compacto K) y allí son de Cauchy en las normas introducidas anteriormente, $\|\cdot\|_N$. Si una sucesión tiende a 0 en $\mathcal{D}(\Omega)$ (con la topología τ), los soportes de todos los elementos de la sucesión están en algún compacto y todas las derivadas tienden a cero uniformemente. Este resultado es de gran importancia para entender la convergencia (que será uno de los conceptos más usados) en el espacio de las funciones prueba. Finalmente $\mathcal{D}(\Omega)$ es un espacio completo (una consecuencia de la completitud de \mathcal{D}_K).

7.2.3. Distribuciones

Una vez que tenemos definida una topología sobre el espacio de funciones prueba de soporte compacto $\mathcal{D}(\Omega)$, estudiaremos los funcionales lineales continuos sobre este espacio. Aunque $\mathcal{D}(\Omega)$ no es un espacio normado, los funcionales lineales continuos y acotados coinciden. Además, el hecho de que \mathcal{D}_K sea metrizable y la topología sea un límite inductivo, hace que la condición de continuidad pueda ser sustituida por la siguiente. Si φ_n es una sucesión en $\mathcal{D}(\Omega)$ que tiende a cero, entonces su imagen mediante el funcional lineal también tiende a cero. Finalmente, la condición de continuidad también es equivalente a la continuidad de las restricciones del funcional a los subespacios \mathcal{D}_K .

Una vez establecidas las principales propiedades de los funcionales lineales continuos pasamos a la definición de distribución que es justamente ésta, una distribución no es nada más que un funcional lineal continuo sobre el espacio $\mathcal{D}(\Omega)$. Aunque pueda parecer sencilla, téngase en cuenta la complejidad de la topología (al menos en su definición y el establecimiento de sus propiedades) del espacio $\mathcal{D}(\Omega)$. El espacio de distribuciones se llamará $\mathcal{D}'(\Omega)$.

El principal problema es caracterizar de forma sencilla cuándo un funcional lineal sobre $\mathcal{D}(\Omega)$ es continuo y por tanto define una distribución. Usando las propiedades dadas anteriormente, se tiene que $u \in \mathcal{D}'(\Omega)$ si y solo si la restricción a \mathcal{D}_K es continua, para todo $K \subset \Omega$ compacto. Dicho de forma más precisa, para todo $K \subset \Omega$, compacto, existe N (entero no negativo) y una constante $k \in \mathbf{R}$ tales que:

$$|u(\varphi)| \leq k\|\varphi\|_N, \quad \forall \varphi \in \mathcal{D}_K \quad (7.17)$$

Recuérdese que $\|\varphi\|_N = \max\{|D^J\varphi(x)| : x \in \Omega, |J| \leq N\}$. Se llama orden de u al menor de los N para los que se da la anterior acotación. El orden puede ser infinito.

7.2.4. Topología en \mathcal{D}'

El espacio de las distribuciones es un espacio de funcionales, aplicaciones lineales entre dos espacios vectoriales. La topología en este espacio se puede definir de varias formas, relacionadas siempre con las topologías de los espacios inicial y final. En particular al ser un espacio dual (el espacio final es \mathbf{C}) podemos

definir la llamada topología débil* (el asterisco proviene de que se trata de una topología introducida en el dual).

Dada una función prueba, $\varphi \in \mathcal{D}(\Omega)$, podemos definir un funcional lineal sobre $\mathcal{D}'(\Omega)$, por

$$F_\varphi(u) = u(\varphi), \quad u \in \mathcal{D}'(\Omega) \quad (7.18)$$

El funcional F_φ es claramente lineal. Además la familia $\{F_\varphi : \varphi \in \mathcal{D}(\Omega)\}$ separa puntos de $\mathcal{D}'(\Omega)$. Si $F_\varphi(u_1) = F_\varphi(u_2)$, para todo $\varphi \in \mathcal{D}(\Omega)$, se tiene:

$$u_1(\varphi) = u_2(\varphi), \quad \forall \varphi \in \mathcal{D}(\Omega) \quad (7.19)$$

luego $u_1 = u_2$.

La situación es la siguiente. Tenemos un espacio vectorial, $\mathcal{D}'(\Omega)$, y una familia de funcionales (los funcionales F_φ) que separan puntos en $\mathcal{D}'(\Omega)$. La topología débil en $\mathcal{D}'(\Omega)$ inducida por los funcionales F_φ es la más fina de las topologías que hacen que los funcionales F_φ sean continuos¹. Se tiene entonces que la topología débil es localmente convexa y el dual de $\mathcal{D}'(\Omega)$ es el espacio vectorial de los funcionales F_φ (que es isomorfo a $\mathcal{D}(\Omega)$). Además todo funcional que sea continuo en esta topología es de la forma F_φ .

En definitiva, dotado de la topología débil*, $\mathcal{D}'(\Omega)$ es un espacio localmente convexo, y la convergencia de sucesiones en este espacio es la siguiente

$$\lim_{n \rightarrow \infty} u_n = u \Leftrightarrow \lim_{n \rightarrow \infty} u_n(\varphi) = u(\varphi), \quad \forall \varphi \in \mathcal{D}(\Omega) \quad (7.20)$$

Veamos un resultado en cierto sentido contrario. Si $\lim_{n \rightarrow \infty} u_n(\varphi)$ existe para todo $\varphi \in \mathcal{D}(\Omega)$, entonces podemos definir una aplicación:

$$u(\varphi) = \lim_{n \rightarrow \infty} u_n(\varphi), \quad \forall \varphi \in \mathcal{D}(\Omega) \quad (7.21)$$

Pues bien, esta aplicación u es una distribución y la convergencia $u_n \rightarrow u$ se verifica también para las derivadas: $D^J u_n \rightarrow D^J u$. La demostración es muy sencilla, eso sí, una vez que uno conoce los teoremas fundamentales del análisis funcional (en este caso el de Banach-Steinhaus).

Dada una distribución $u \in \mathcal{D}'(\Omega)$ es posible definir su soporte. Se dice que una distribución se anula en un abierto A de Ω si su valor sobre cualquier función de $\mathcal{D}(\Omega)$ cuyo soporte esté en A , es cero. Si ahora construimos la unión de todos estos A , el soporte de la distribución es su complementario en Ω .

El soporte de una distribución tiene las siguientes importantes propiedades. Si $u \in \mathcal{D}'(\Omega)$ y $\varphi \in \mathcal{D}(\Omega)$ y sus soportes tienen intersección vacía, entonces $u(\varphi) = 0$. La única distribución que tiene soporte vacío es la cero. Si el soporte de una distribución es compacto, la distribución tiene orden finito y se extiende de manera única a un funcional lineal continuo sobre $\mathcal{C}^\infty(\Omega)$.

Para acabar esta sección sobre propiedades de las distribuciones y la topología, veamos como una distribución es (salvo detalles que ahora veremos) la derivada (en sentido de distribuciones) de una función.

El primer resultado es local y se refiere a cualquier distribución. Sea $u \in \mathcal{D}'(\Omega)$ y K un compacto de Ω . Existe una función continua f sobre Ω tal que, para algún J :

$$u(\varphi) = (-1)^{|J|} \int_{\Omega} f(x) D^J \varphi(x) dx, \quad \varphi \in \mathcal{D}_K \quad (7.22)$$

La demostración no es muy complicada pero requiere entre otras cosas el uso del teorema de Hahn-Banach sobre extensión de funcionales. La expresión anterior establece que u es la derivada de una función continua cuando actúa sobre funciones en \mathcal{D}_K .

Si u es una distribución de soporte compacto, entonces el resultado es global. Sea A un abierto contenido en Ω y $u \in \mathcal{D}'(\Omega)$, de soporte compacto K contenido en A . Sea N el orden (finito) de u . Existe

¹Los abiertos de esta topología son las uniones de intersecciones finitas de los conjuntos $F_\varphi^{-1}(V)$, siendo V un abierto de \mathbb{C} .

entonces un número finito de funciones continuas en Ω , f_J , (J es un multi-índice tal que $J_i \leq N + 2$), que tienen sus soportes en A y que verifican $u = \sum_J D^J f_J$. Es decir, como antes,

$$u(\varphi) = \sum_J (-1)^{|J|} \int_{\Omega} f_J(x) D^J \varphi(x) dx, \quad \varphi \in \mathcal{D}(\Omega) \quad (7.23)$$

Finalmente, el resultado más general es el siguiente. Si $u \in \mathcal{D}'(\Omega)$, existen funciones f_J continuas en Ω , tales que cada subconjunto compacto de Ω corta al soporte de un número finito de esas funciones y $u = \sum_J D^J f_J$. En el caso de que el orden de u sea finito las funciones f_J se pueden escoger de forma que solo un número finito de ellas sea distinto de cero.

7.3. Propiedades y operaciones en el espacio de distribuciones

Una función localmente integrable en Ω define una distribución en $\mathcal{D}'(\Omega)$. Localmente integrable (L^1_{loc}), significa que se trata de una función medible (Lebesgue) y que es integrable (su integral es finita) sobre cada compacto contenido en Ω . Entonces

$$u_f(\varphi) = \int_{\Omega} f(x) \varphi(x) dx, \quad \varphi \in \mathcal{D}(\Omega) \quad (7.24)$$

es una distribución. Obviamente es lineal así que solo habría que probar que es continua. Pero es acotada sobre cada \mathcal{D}_K :

$$\left| \int_{\Omega} f(x) \varphi(x) dx \right| \leq \left(\int_K |f(x)| dx \right) \|\varphi\|_0 \quad (7.25)$$

Se tiene entonces una distribución de orden 0. Una medida (una medida de Borel, es decir una medida sobre la σ -álgebra de Borel, generada por los abiertos, o una medida positiva finita sobre compactos) define de la misma manera una distribución.

7.3.1. Derivación

Siendo uno de los objetivos de la introducción de distribuciones derivar funciones que no son derivables, veamos como efectivamente es posible definir derivadas de todos los órdenes para una distribución.

Sea f una función en $L^1_{\text{loc}}(\mathbf{R})$, derivable con continuidad. Se tiene:

$$\int_{\Omega} f'(x) \varphi(x) dx = - \int_{\Omega} f(x) \varphi'(x) dx, \quad \varphi \in \mathcal{D}(\Omega) \quad (7.26)$$

pues φ se anula fuera de un compacto contenido (estrictamente) en Ω . Por lo tanto resulta razonable definir (al menos para este tipo de funciones f):

$$u'_f(\varphi) \equiv u_{f'}(\varphi) = -u_f(\varphi') \quad (7.27)$$

Extendiendo esta definición a cualquier distribución, tenemos, si $u \in \mathcal{D}'(\Omega)$:

$$D^J u(\varphi) = (-1)^{|J|} u(D^J \varphi), \quad \varphi \in \mathcal{D}(\Omega) \quad (7.28)$$

Veamos que el funcional así definido es una distribución. Como siempre debemos probar la acotación.

$$|D^J u(\varphi)| = |u(D^J \varphi)| \leq C \|D^J \varphi\|_N \leq C \|\varphi\|_{N+|J|} \quad (7.29)$$

Luego se trata de una distribución.

Dada una función $f \in L^1_{\text{loc}}$ sabemos que existe una distribución asociada a ella. De acuerdo con lo anterior, esta distribución es derivable indefinidamente. Pero si f es también derivable y la derivada es localmente integrable, la pregunta es si

$$D^J u_f = u_{D^J f} \quad (7.30)$$

La respuesta es que esto no es correcto en general. Supongamos que $\Omega = (a, b)$ es un intervalo de \mathbf{R} y que f es una función definida en Ω continua por la izquierda y de variación acotada². Entonces f es derivable casi doquiera³ y su derivada es una función localmente integrable.

Calculemos la derivada de u_f en el sentido de distribuciones. Para ello, apliquemos el teorema de Fubini a la siguiente integral. Sea

$$C = \{(x, y) : x \in \Omega, y \in \Omega, x < y\} \quad (7.31)$$

y μ la medida definida por $\mu([c, d]) = f(d) - f(c)$. Entonces

$$\int_C \varphi'(x) dx d\mu(y) = \int_\Omega \left(\int_a^y \varphi'(x) dx \right) d\mu(y) = \int_\Omega (\varphi(y) - \varphi(a)) d\mu(y) = u_\mu(\varphi) \quad (7.32)$$

pues $\varphi(a) = 0$. También:

$$\int_C \varphi'(x) dx d\mu(y) = \int_\Omega \varphi'(x) \left(\int_x^b d\mu(y) \right) dx = \int_\Omega (f(b) - f(x)) \varphi'(x) dx \quad (7.33)$$

Pero

$$\int_\Omega (f(b) - f(x)) \varphi'(x) dx = f(b) \int_\Omega \varphi'(x) dx - \int_\Omega f(x) \varphi'(x) dx = - \int_\Omega f(x) \varphi'(x) dx \quad (7.34)$$

es decir (por Fubini):

$$u_\mu(\varphi) = - \int_\Omega f(x) \varphi'(x) dx = -u_f(\varphi') \equiv u'_f(\varphi) \quad (7.35)$$

y por tanto $u_\mu = u'_f$

De esta forma, la derivada de u_f en el sentido de distribuciones es una distribución asociada a una medida $Du_f = u_\mu$. Para que $Du_f = u_{Df}$ se tendría que cumplir $\mu = Df$, es decir,

$$u_\mu(\varphi) = \int_\Omega \varphi(x) f'(x) dx \quad (7.36)$$

lo que es cierto si y solo si f es absolutamente continua⁴ pues en este caso,

$$d\mu = f'(x) dx \quad (7.37)$$

y

$$f(x) - f(a) = \int_a^x f'(y) dy \quad (7.38)$$

lo que equivale a que f sea absolutamente continua⁵.

7.3.2. Multiplicación por funciones

Las distribuciones forman un espacio vectorial, así que la suma de distribuciones o el producto de una distribución por un escalar son distribuciones. Sin embargo, el producto de distribuciones no está bien definido en general.

²Se dice que una función f es de variación acotada en el intervalo $[a, b]$ si existe una constante positiva c tal que para cualquier partición del intervalo $a = x_0 < x_1 < \dots < x_n = b$ se tiene $\sum_{i=1}^n |f(x_i) - f(x_{i-1})| < c$. Toda función de variación acotada es la diferencia de dos funciones no decrecientes

³Teorema de Lebesgue: una función monótona definida en un intervalo $[a, b]$ tiene derivada finita casi doquiera. Como consecuencia toda función de variación acotada tiene derivada casi doquiera.

⁴Una función es absolutamente continua en el intervalo $[a, b]$ si, dado $\epsilon > 0$, existe $\delta > 0$ tal que $\sum_{i=1}^n |f(b_i) - f(a_i)| < \epsilon$, si $\sum_i (b_i - a_i) < \delta$ para cualquier elección (finita) de intervalos $(a_i, b_i) \subset [a, b]$ disjuntos dos a dos. Toda función absolutamente continua es de variación acotada.

⁵Si F es absolutamente continua en un intervalo su derivada es sumable. F es absolutamente continua en un intervalo si y solo si se cumple la ecuación (7.38).

Lo que sí podemos definir es el producto de una distribución por una función \mathcal{C}^∞ :

$$fu(\varphi) = u(f\varphi), \quad u \in \mathcal{D}'(\Omega), \quad f \in \mathcal{C}^\infty(\Omega), \quad \varphi \in \mathcal{D}(\Omega) \quad (7.39)$$

La definición es correcta pues $f\varphi \in \mathcal{D}(\Omega)$. Claramente es un funcional lineal. Debemos demostrar que es continuo. La prueba se basa en la fórmula de Leibniz:

$$D^J(fu) = \sum_{I \leq J} c_{JI} (D^{J-I}f) D^I u \quad (7.40)$$

Sea K un compacto de Ω . Sabemos que al ser u una distribución, existen N y k tales que:

$$|u(\varphi)| \leq k \|\varphi\|_N \quad (7.41)$$

Usando la fórmula de Leibniz,

$$\|f\varphi\|_N \leq k' \|\varphi\|_N \quad (7.42)$$

donde la constante k' depende de f , N y K . Finalmente

$$|fu(\varphi)| = |u(f\varphi)| \leq k \|f\varphi\|_N \leq kk' \|\varphi\|_N \quad (7.43)$$

y por tanto fu es continua.

7.3.3. El producto de convolución

Como hemos dicho antes, el producto de distribuciones puede no estar bien definido. Sin embargo, es posible definir la convolución de una función y una distribución y en algunos caso la de dos distribuciones.

Consideremos que $\Omega = \mathbf{R}^n$. Sean f, g funciones con valores complejos definidas sobre \mathbf{R}^n . Se define la convolución de f y g como la función

$$(f * g)(x) = \int_{\mathbf{R}^n} f(y)g(x-y) dy \quad (7.44)$$

si la integral (Lebesgue) existe. Podemos generalizarla en el siguiente sentido

$$(u * \varphi)(x) = u(\tau_x \check{\varphi}), \quad u \in \mathcal{D}'(\mathbf{R}^n), \quad \varphi \in \mathcal{D}(\mathbf{R}^n) \quad (7.45)$$

es decir, la convolución de una distribución con una función prueba es otra función, definida como arriba. En esta definición τ_x es la traslación en x :

$$(\tau_x f)(y) = f(y-x) \quad (7.46)$$

y si f es una función, \check{f} es la función definida por:

$$\check{f}(x) = f(-x) \quad (7.47)$$

La traslación τ_x se puede aplicar a distribuciones (basta ver qué ocurre cuando la distribución es la asociada a una función localmente integrable):

$$(\tau_x u)(\varphi) = u(\tau_{-x} \varphi), \quad u \in \mathcal{D}'(\mathbf{R}^n), \quad \varphi \in \mathcal{D}(\mathbf{R}^n) \quad (7.48)$$

La definición anterior se reduce al caso expuesto al principio si u_f es la distribución asociada a una función localmente integrable, f :

$$\begin{aligned} (u_f * \varphi)(x) &= u_f(\tau_x \check{\varphi}) = \int_{\mathbf{R}^n} f(y)(\tau_x \check{\varphi})(x-y) dy = \int_{\mathbf{R}^n} f(y)\check{\varphi}(x-y-x) dy \\ &= \int_{\mathbf{R}^n} f(y)\check{\varphi}(-y) dy = \int_{\mathbf{R}^n} f(y)\varphi(y) dy = (f * \varphi)(x) \end{aligned} \quad (7.49)$$

Veamos a continuación las propiedades más interesantes de la convolución.

La traslación, τ_x , en una convolución se puede hacer sobre cualquiera de los dos factores:

$$\tau_x(u * \varphi) = (\tau_x u) * \varphi = u * (\tau_x \varphi), \quad u \in \mathcal{D}(\mathbf{R}^n), \varphi \in \mathcal{D}(\mathbf{R}^n), x \in \mathbf{R}^n \quad (7.50)$$

En efecto,

$$\tau_x(u * \varphi)(y) = (u * \varphi)(y - x) = u(\tau_{y-x} \check{\varphi}) \quad (7.51)$$

$$((\tau_x u) * \varphi)(y) = (\tau_x u)(\tau_y \check{\varphi}) = u(\tau_{-x} \tau_y \check{\varphi}) = u(\tau_{y-x} \check{\varphi}) \quad (7.52)$$

$$(u * (\tau_x \varphi))(y) = u(\tau_y (\tau_x \check{\varphi})) = u(\tau_y \tau_{-x} \check{\varphi}) = u(\tau_{y-x} \check{\varphi}) \quad (7.53)$$

De la misma forma, las derivadas se pueden efectuar sobre cualquiera de los factores.

$$D^J(u * \varphi) = (D^J u) * \varphi = u * (D^J \varphi), \quad u \in \mathcal{D}(\mathbf{R}^n), \varphi \in \mathcal{D}(\mathbf{R}^n), J \in \mathbf{N}^n \quad (7.54)$$

lo que, por otra parte indica que $u * \varphi \in \mathcal{C}^\infty(\mathbf{R}^n)$

Para demostrarlo, usemos la identidad

$$\tau_x((D^J \varphi)) = (-1)^{|J|} D^J(\tau_x \check{\varphi}) \quad (7.55)$$

que puede ser comprobada fácilmente, por ejemplo en una dimensión:

$$(\tau_x(\partial_y \varphi))(y) = (\partial_y \varphi)(y - x) = (\partial_y \varphi)(x - y) \quad (7.56)$$

pero

$$(\partial_y \varphi)(x - y) = -\partial_y(\tau_x \check{\varphi})(y) \quad (7.57)$$

resultado que se generaliza al caso de n variables y orden de derivación cualquiera.

Apliquemos ahora la distribución u :

$$u(\tau_x((D^J \varphi))) = (u * D^J \varphi)(x) \quad (7.58)$$

$$u((-1)^{|J|} D^J(\tau_x \check{\varphi})) = (-1)^{|J|} u(D^J(\tau_x \check{\varphi})) = D^J u(\tau_x \check{\varphi}) = ((D^J u) * \varphi)(x) \quad (7.59)$$

La otra igualdad de (7.54) se demuestra primero para una derivada direccional, en la dirección de un vector unitario \mathbf{e} de \mathbf{R}^n .

Definimos la traslación y dilatación: $\sigma_r = \frac{1}{r}(\tau_0 - \tau_{r\mathbf{e}})$ y la aplicamos a una convolución. Usando (7.50) y la definición de convolución:

$$\sigma_r(u * \varphi) = u * (\sigma_r \varphi), \quad (u * (\sigma_r \varphi))(x) = u(\tau_x(\sigma_r \varphi)) \quad (7.60)$$

Teniendo en cuenta que

$$(\sigma_r \varphi)(x) = \frac{1}{r}((\tau_0 \varphi)(x) - (\tau_{r\mathbf{e}} \varphi)(x)) = \frac{\varphi(x) - \varphi(x + r\mathbf{e})}{r} \rightarrow D_{\mathbf{e}} \varphi \quad (7.61)$$

cuando $r \rightarrow 0$ en $\mathcal{D}(\mathbf{R}^n)$, llegamos a

$$\tau_x((\sigma_r \varphi)) \rightarrow \tau_x(D_{\mathbf{e}} \varphi), \quad u(\tau_x((\sigma_r \varphi))) \rightarrow u(\tau_x(D_{\mathbf{e}} \varphi)) \quad (7.62)$$

cuando $r \rightarrow 0$, es decir:

$$u * (\sigma_r \varphi)(x) \rightarrow u * (D_{\mathbf{e}} \varphi)(x), \quad \forall x \in \mathbf{R}^n \quad (7.63)$$

Como

$$\sigma_r(u * \varphi) \rightarrow D_{\mathbf{e}}(u * \varphi) \quad (7.64)$$

se tiene:

$$u * (D_{\mathbf{e}} \varphi) = D_{\mathbf{e}}(u * \varphi) \quad (7.65)$$

como queríamos probar (la generalización es inmediata).

La convolución es asociativa en el siguiente sentido:

$$u * (\varphi_1 * \varphi_2) = (u * \varphi_1) * \varphi_2 \quad (7.66)$$

Para demostrarlo, veamos en primer lugar que

$$(\varphi_1 * \varphi_2)\checkmark(t) = \int_{\mathbf{R}^n} \check{\varphi}_2(s)(\tau_s \check{\varphi}_1)(t) ds \quad (7.67)$$

lo que es inmediato

$$\begin{aligned} (\varphi_1 * \varphi_2)\checkmark(t) &= (\varphi_1 * \varphi_2)(-t) = (\varphi_2 * \varphi_1)(-t) = \int_{\mathbf{R}^n} \varphi_2(s)\varphi_1(-t-s) ds \\ &= \int_{\mathbf{R}^n} \varphi_2(-s)\varphi_1(-t+s) ds \\ &= \int_{\mathbf{R}^n} \check{\varphi}_2(s)\check{\varphi}_1(t-s) ds = \int_{\mathbf{R}^n} \check{\varphi}_2(s)(\tau_s \check{\varphi}_1)(t) ds \\ &= (\check{\varphi}_2 * \check{\varphi}_1)(t) \end{aligned} \quad (7.68)$$

Las funciones prueba φ_i tienen soporte (compacto) igual a K_i , $i = 1, 2$. Sea $K = K_1 + K_2$ (suma en \mathbf{R}^n). Consideremos la función $s \rightarrow \check{\varphi}_1(s)\tau_s \check{\varphi}_2$, que pasa de valores en \mathbf{R}^n a funciones en $\mathcal{D}(\mathbf{R}^n)$, más concretamente funciones de \mathcal{D}_K , que además se anula fuera de K_2 . Se tiene

$$(\varphi_1 * \varphi_2)\checkmark = \int_{K_2} \check{\varphi}_2(s)(\tau_s \check{\varphi}_1) ds \quad (7.69)$$

y por tanto,

$$\begin{aligned} (u * (\varphi_1 * \varphi_2))(0) &= u((\varphi_1 * \varphi_2)\checkmark) = \int_{K_2} \check{\varphi}_2(s)u(\tau_s \check{\varphi}_1) ds \\ &= \int_{\mathbf{R}^n} (u * \varphi_1)(s)\varphi_2(-s) ds \\ &= ((u * \varphi_1) * \varphi_2)(0) \end{aligned} \quad (7.70)$$

Sustituyamos ahora φ_2 por $\tau_{-x}\varphi_2$

$$(u * (\varphi_1 * \tau_{-x}\varphi_2))(0) = ((u * \varphi_1) * \tau_{-x}\varphi_2)(0) \quad (7.71)$$

Pero

$$\varphi_1 * \tau_{-x}\varphi_2 = \tau_{-x}(\varphi_1 * \varphi_2) \quad (7.72)$$

$$u * (\varphi_1 * \tau_{-x}\varphi_2) = u * (\tau_{-x}(\varphi_1 * \varphi_2)) = \tau_{-x}(u * (\varphi_1 * \varphi_2)) \quad (7.73)$$

y por consiguiente

$$(\tau_{-x}(u * (\varphi_1 * \varphi_2)))(0) = (\tau_{-x}((u * \varphi_1) * \varphi_2))(0) \quad (7.74)$$

es decir

$$(u * (\varphi_1 * \varphi_2))(x) = ((u * \varphi_1) * \varphi_2)(x) \quad (7.75)$$

Si u es una distribución de soporte compacto, es posible definir (igual que antes) su convolución con una función en $\mathcal{C}^\infty(\mathbf{R}^n)$. Las propiedades son como anteriormente.

Finalmente, si $u, v \in \mathcal{D}'(\Omega)$ y al menos una de las dos tiene soporte compacto entonces es posible definir su convolución:

$$(u * v)(\varphi) = (u * (v * \varphi))(0) \quad (7.76)$$

y se tiene:

$$(u * v) * \varphi = u * (v * \varphi) \quad (7.77)$$

7.4. Ejemplos de distribuciones

7.4.1. La delta de Dirac

La distribución delta de Dirac, definida por

$$\delta(\varphi) = \varphi(0) \quad (7.78)$$

no proviene de ninguna función localmente integrable aunque está asociada a una medida (discreta). La manera más clásica de introducirla es considerar la densidad de una esfera homogénea de masa constante (1) y radio ϵ tendiendo a cero:

$$\rho(x) = \begin{cases} \frac{3}{4\pi\epsilon^3} & |x| < \epsilon \\ 0 & |x| > \epsilon \end{cases} \quad (7.79)$$

Se tiene

$$\int_{\mathbf{R}^3} \rho(x) dx = 1 \quad (7.80)$$

independientemente de ϵ . Demostremos que para cualquier función continua $f(x)$ se verifica:

$$\lim_{\epsilon \rightarrow 0} \int_{\mathbf{R}^n} \rho(x) f(x) dx = f(0) \quad (7.81)$$

En efecto,

$$\left| \int_{\mathbf{R}^n} \rho(x) f(x) dx - f(0) \right| = \frac{3}{4\pi\epsilon^3} \int_{|x| < \epsilon} |f(x) - f(0)| \quad (7.82)$$

Como f es continua, dado $\eta > 0$ existe $\epsilon_0 > 0$, tal que si $|x| < \epsilon_0$, entonces $|f(x) - f(0)| < \eta$. Por tanto, para todo $\epsilon < \epsilon_0$ se tiene:

$$\left| \int_{\mathbf{R}^n} \rho(x) f(x) dx - f(0) \right| \leq \eta \quad (7.83)$$

Haciendo tender η a cero se obtiene el resultado deseado. Vemos como la distribución δ aparece como un límite de distribuciones asociada a funciones que son localmente integrables (esencialmente con una discontinuidad de salto). Es posible definir el trasladado de la distribución δ :

$$(\tau_a \delta)(\varphi) = \delta(\tau_{-a} \varphi) = \varphi(a) \quad (7.84)$$

Usualmente escribiremos

$$\tau_a \delta = \delta_a \quad (7.85)$$

Las distribuciones que tiene un punto como soporte se pueden expresar como combinación lineal de la distribución δ y sus derivadas. Sea $u \in \mathcal{D}(\Omega)$, tal que $\text{sop } u = \{p\} \subset \Omega$ y su orden sea N . En estas circunstancias, existen constantes c_J que verifican

$$u = \sum_{|J| \leq N} c_J D^J \delta_p \quad (7.86)$$

El inverso es claramente cierto. La demostración no es muy complicada pero no la haremos aquí (ver por ejemplo Rudin, Teorema 6.25).

7.4.2. Distribuciones asociadas a capas simples y dobles

La delta de Dirac aparece asociada según lo anterior a distribuciones (en el sentido estadístico) puntuales de masa (o de carga). A veces nos interesa estudiar distribuciones superficiales de masa, los llamados problemas de capa simple.

Sea S una superficie en \mathbf{R}^3 suave a trozos. Definimos δ_S como una distribución en $\mathcal{D}'(\mathbf{R}^3)$, dada por:

$$\delta_S(\varphi) = \int_S \varphi(x) d\sigma, \quad \varphi \in \mathcal{D}(\mathbf{R}^3) \quad (7.87)$$

donde $d\sigma$ es el elemento de superficie en S . Esta distribución tiene su soporte en la superficie S .

Si en vez de una carga puntual tenemos dos de distinto signo muy próximas, la estructura correspondiente (un dipolo) tiene carga cero, pero presenta un momento dipolar. Veamos como calcular una distribución asociada a esta situación. La densidad de carga del dipolo es:

$$\frac{1}{\epsilon}(\delta_\epsilon - \delta), \quad \epsilon > 0 \quad (7.88)$$

Calculemos el límite en el espacio $\mathcal{D}'(\mathbf{R})$

$$\frac{1}{\epsilon}(\delta_\epsilon - \delta)(\varphi) = \frac{1}{\epsilon}(\varphi(\epsilon) - \varphi(0)) \rightarrow \varphi'(0), \quad \epsilon \rightarrow 0 \quad (7.89)$$

Por lo tanto, la distribución en el límite es:

$$\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon}(\delta_\epsilon - \delta) = -\delta' \quad (7.90)$$

Esta es pues la densidad de carga asociada a un dipolo. Por supuesto la carga total es cero (en realidad $1 \notin \mathcal{D}(\mathbf{R})$, pero se puede regularizar adecuadamente, en el sentido de obtener una función en $\mathcal{D}(\mathbf{R})$ que valga 1 en un entorno de 0, tenga soporte compacto y sea diferenciable infinitas veces).

$$-\delta'(1) = 0 \quad (7.91)$$

y el momento dipolar:

$$-\delta'(x) = 1 \quad (7.92)$$

De forma similar podemos estudiar una capa doble en \mathbf{R}^3 . Sea S una superficie suave a trozos con normal \mathbf{n} y $f(x)$ una función continua sobre la superficie (que es la densidad de carga superficial). Se define la distribución

$$\partial_{\mathbf{n}}(\nu\delta_S) \quad (7.93)$$

por

$$\partial_{\mathbf{n}}(\nu\delta_S)(\varphi) = \int_S \nu(x)\partial_{\mathbf{n}}\varphi(x) d\sigma \quad (7.94)$$

Esta distribución tiene su soporte en la superficie S . Se llama una distribución de capa doble en S con densidad $f(x)$.

7.4.3. Las fórmulas de Sochozki

Introduciremos en esta sección otras distribuciones y las relaciones que existen entre ellas. Se define la “pseudofunción”

$$\mathcal{P}\left(\frac{1}{x}\right) \quad (7.95)$$

como una distribución, a través de:

$$\mathcal{P}\left(\frac{1}{x}\right)(\varphi) = \text{VP} \int_{\mathbf{R}} \frac{\varphi(x)}{x} dx = \lim_{\epsilon \rightarrow 0^+} \int_{|x| \geq \epsilon} \frac{\varphi(x)}{x} dx \quad (7.96)$$

Claramente se trata de una aplicación lineal. Veamos que además es continua. Para demostrarlo, consideremos una sucesión de funciones φ_n en $\mathcal{D}(\mathbf{R})$ que convergen a 0 y tales que sus soportes están contenidos en el intervalo $[-a, a]$

$$\mathcal{P}\left(\frac{1}{x}\right)(\varphi_n) = \lim_{\epsilon \rightarrow 0^+} \int_{-a}^a \frac{\varphi_n(x)}{x} dx \quad (7.97)$$

Sumando y restando $\varphi(0)$:

$$\begin{aligned} \text{VP} \int_{-a}^a \frac{\varphi_n(x)}{x} dx &= \text{VP} \int_{-a}^a \frac{\varphi_n(x) - \varphi_n(0)}{x} dx + \text{VP} \int_{-a}^a \frac{\varphi_n(0)}{x} dx \\ &= \int_{-a}^a \frac{\varphi_n(x) - \varphi_n(0)}{x} dx \end{aligned} \quad (7.98)$$

donde no hace falta incluir el valor principal, al ser el integrando continuo en $x = 0$. Utilizando los teoremas elementales del cálculo,

$$\left| \frac{\varphi_n(x) - \varphi_n(0)}{x} \right| \leq \max\{|\varphi'_n(x)| : x \in [-a, a]\} \quad (7.99)$$

y por tanto

$$\left| \mathcal{P}\left(\frac{1}{x}\right)(\varphi_n) \right| \leq \int_{-a}^a \left| \frac{\varphi_n(x) - \varphi_n(0)}{x} \right| dx \leq 2a \max\{|\varphi'_n(x)| : x \in [-a, a]\} \quad (7.100)$$

que tiende a cero cuando $n \rightarrow \infty$.

Calculemos ahora el límite cuando $\epsilon \rightarrow 0_+$ de las distribuciones asociadas a las funciones localmente integrables siguientes:

$$\frac{1}{x + i\epsilon} \quad (7.101)$$

Aplicando la distribución a una función prueba,

$$\lim_{\epsilon \rightarrow 0_+} \int_{\mathbf{R}} \frac{\varphi(x)}{x + i\epsilon} dx = \lim_{\epsilon \rightarrow 0_+} \int_{-a}^a \varphi(x) \frac{x - i\epsilon}{x^2 + \epsilon^2} dx = \quad (7.102)$$

restando y sumando, como en el cálculo anterior $\varphi(0)$

$$\lim_{\epsilon \rightarrow 0_+} \left\{ \int_{-a}^a (\varphi(x) - \varphi(0)) \frac{x - i\epsilon}{x^2 + \epsilon^2} dx + \varphi(0) \int_{-a}^a \frac{x - i\epsilon}{x^2 + \epsilon^2} dx \right\} = \quad (7.103)$$

$$\int_{-a}^a \frac{\varphi(x) - \varphi(0)}{x} dx - \lim_{\epsilon \rightarrow 0_+} i\epsilon \varphi(0) \int_{-a}^a \frac{1}{x^2 + \epsilon^2} dx =$$

$$\mathcal{P}\left(\frac{1}{x}\right)(\varphi) - 2i\varphi(0) \lim_{\epsilon \rightarrow 0_+} \arctan \frac{a}{\epsilon} \quad (7.104)$$

y por tanto, si se escribe

$$\lim_{\epsilon \rightarrow 0_+} \frac{1}{x + i\epsilon} = \frac{1}{x + i0_+} \quad (7.105)$$

se tiene

$$\frac{1}{x + i0_+} = \mathcal{P}\left(\frac{1}{x}\right) - i\pi\delta \quad (7.106)$$

De forma similar, se puede probar que:

$$\frac{1}{x - i0_+} = \mathcal{P}\left(\frac{1}{x}\right) + i\pi\delta \quad (7.107)$$

Estas son las fórmulas de Sochozki.

7.4.4. Derivadas de distribuciones

Estudiaremos en esta sección algunas propiedades de las derivadas de distribuciones. La función

$$\log |x| \quad (7.108)$$

es localmente integrable, luego define una distribución. Veamos cuál es su derivada.

$$(\log |x|)'(\varphi) = -(\log |x|)(\varphi') = - \int_{\mathbf{R}} \varphi'(x) \log |x| dx \quad (7.109)$$

Descomponemos la integral en dos sumandos e integramos por partes

$$\begin{aligned} \int_0^\infty \varphi'(x) \log |x| dx &= \lim_{\epsilon \rightarrow 0} \int_\epsilon^\infty \varphi'(x) \log x dx \\ &= \lim_{\epsilon \rightarrow 0} \left(\varphi(x) \log x \Big|_\epsilon^\infty - \int_\epsilon^\infty \frac{\varphi(x)}{x} dx \right) \\ &= \lim_{\epsilon \rightarrow 0} \left(-\varphi(\epsilon) \log \epsilon - \int_\epsilon^\infty \frac{\varphi(x)}{x} dx \right) = \lim_{\epsilon \rightarrow 0} \left(-\varphi(0) \log \epsilon - \int_\epsilon^\infty \frac{\varphi(x)}{x} dx \right) \end{aligned} \quad (7.110)$$

De forma similar

$$\int_{-\infty}^0 \varphi'(x) \log |x| \, dx = \lim_{\epsilon \rightarrow 0} \left(\varphi(0) \log \epsilon - \int_{-\infty}^{-\epsilon} \frac{\varphi(x)}{x} \, dx \right) \quad (7.111)$$

Sumando ambas expresiones

$$-\int_{\mathbf{R}} \varphi'(x) \log |x| \, dx = -\lim_{\epsilon \rightarrow 0} \left(\int_{\epsilon}^{\infty} \frac{\varphi(x)}{x} \, dx + \int_{-\infty}^{-\epsilon} \frac{\varphi(x)}{x} \, dx \right) \quad (7.112)$$

y por tanto

$$(\log |x|)' = \mathcal{P}\left(\frac{1}{x}\right) \quad (7.113)$$

La derivada de una función continua a trozos (continua salvo en un número finito de puntos en los que tiene discontinuidades de salto finito), y derivable, salvo obviamente en esos puntos, no existe en sentido de funciones pero sí en el de distribuciones. Sea $f(x)$ derivable en los intervalos $x < x_0$ y $x > x_0$ y con una discontinuidad de salto en x_0 :

$$[f]_{x_0} = f(x_0 + 0) - f(x_0 - 0) \quad (7.114)$$

Calculemos la derivada en sentido de distribuciones:

$$f'(\varphi) = -\int_{\mathbf{R}} f(x) \varphi'(x) \, dx = -\int_{-\infty}^{x_0} f(x) \varphi'(x) \, dx - \int_{x_0}^{\infty} f(x) \varphi'(x) \, dx \quad (7.115)$$

Integrando por partes:

$$f'(\varphi) = (f(x_{0+}) - f(x_{0-}))\varphi(x_0) + \int_{-\infty}^{x_0} \{f'\}(x)\varphi(x) \, dx + \int_{x_0}^{\infty} \{f'\}(x)\varphi(x) \, dx \quad (7.116)$$

y por tanto,

$$f' = \{f'\} + (f(x_{0+}) - f(x_{0-}))\delta_{x_0} \quad (7.117)$$

El resultado se generaliza sin dificultad a cualquier número de discontinuidades de este tipo (de hecho, aunque sean infinitas). También es posible generalizar este tipo de expresiones a funciones de varias variables.

Supongamos que Ω es un abierto conexo de \mathbf{R}^n acotado por una frontera S que es suave a trozos. Sea f una función diferenciable con continuidad en el cierre de Ω y en el cierre de $\Omega' = \mathbf{R}^n \setminus \bar{\Omega}$. Supongamos que $[f]_S$ es el salto en la función f al atravesar la frontera:

$$[f]_S(x) = \lim_{x' \rightarrow x, x' \in \Omega'} f(x') - \lim_{x' \rightarrow x, x' \in \Omega} f(x') \quad (7.118)$$

Para hallar la derivada parcial de f en el sentido de distribuciones usamos una fórmula de Green ($\mathbf{n}x_i$ es el ángulo que forma la normal con la dirección x_i):

$$\frac{\partial f}{\partial x_i}(\varphi) = -\int_{\mathbf{R}^n} f(x) \frac{\partial \varphi}{\partial x_i} \, dx = \int_{\mathbf{R}^n} \left\{ \frac{\partial f}{\partial x_i} \right\} \varphi(x) \, dx + \int_S [f]_S(x) \cos(\mathbf{n}x_i) \varphi(x) \, dx \quad (7.119)$$

y por tanto

$$\frac{\partial f}{\partial x_i} = \left\{ \frac{\partial f}{\partial x_i} \right\} + [f]_S \cos(\mathbf{n}x_i) \delta_S \quad (7.120)$$

Si la función es dos veces derivable (en las condiciones anteriores) se tiene una fórmula similar con la derivada segunda:

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \left\{ \frac{\partial^2 f}{\partial x_i \partial x_j} \right\} + \frac{\partial}{\partial x_j} ([f]_S \cos(\mathbf{n}x_i) \delta_S) + \left\{ \frac{\partial}{\partial x_i} \right\} \cos(\mathbf{n}x_j) \delta_S \quad (7.121)$$

Particularizando esta expresión para el caso $i = j$ tenemos:

$$\frac{\partial^2 f}{\partial x_i^2} = \left\{ \frac{\partial^2 f}{\partial x_i^2} \right\} + \frac{\partial}{\partial x_i} ([f]_S \cos(\mathbf{n}x_i) \delta_S) + \left\{ \frac{\partial}{\partial x_i} \right\} \cos(\mathbf{n}x_i) \delta_S \quad (7.122)$$

y si sumamos en i :

$$\Delta f = \{\Delta f\} + \sum_i \frac{\partial}{\partial x_i} ([f]_S \cos(\mathbf{n}x_i) \delta_S) + \sum \left\{ \frac{\partial}{\partial x_i} \right\} \cos(\mathbf{n}x_i) \delta_S \quad (7.123)$$

pero:

$$\sum_i \frac{\partial}{\partial x_i} ([f]_S \cos(\mathbf{n}x_i) \delta_S) = \frac{\partial}{\partial n} ([f]_S \delta_S) \quad (7.124)$$

y

$$\sum \left\{ \frac{\partial}{\partial x_i} \right\} \cos(\mathbf{n}x_i) \delta_S = \left[\frac{\partial f}{\partial n} \right]_S \delta_S \quad (7.125)$$

lo que nos permite escribir el resultado final

$$\Delta f = \{\Delta f\} + \left[\frac{\partial f}{\partial n} \right]_S \delta_S + \frac{\partial}{\partial n} ([f]_S \delta_S) \quad (7.126)$$

7.4.5. Soluciones fundamentales

Una solución fundamental de un operador lineal de coeficientes constantes es una distribución \mathcal{E} que verifica

$$P\mathcal{E} = \delta \quad (7.127)$$

El cálculo de soluciones fundamentales lo abordaremos después de estudiar la transformada de Fourier de distribuciones. Aquí nos limitamos a comprobar que ciertas distribuciones satisfacen unas ecuaciones diferenciales, es decir, estudiamos como se derivan distribuciones.

Calculemos en $n = 2$ el laplaciano de la distribución $\log r$ (que es una función localmente integrable).

$$(\Delta \log r)(\varphi) = (\log r)(\Delta \varphi) = \lim_{\epsilon \rightarrow 0} \int_{\epsilon < r < R} (\log r) \Delta \varphi(x) dx \quad (7.128)$$

suponiendo que el soporte de φ está contenido en el disco de radio R y centro el origen. Usemos ahora una de las fórmulas de Green:

$$\int_{\Omega} (f \Delta g - g \Delta f) dx = \int_{\partial \Omega} \left(f \frac{\partial g}{\partial n} - g \frac{\partial f}{\partial n} \right) ds \quad (7.129)$$

siendo Ω la corona circular $\epsilon < r < R$ y n la normal exterior.

$$\begin{aligned} & \int_{\epsilon < r < R} (\varphi(x) \Delta \log r - (\log r) \Delta \varphi(x)) dx = \\ & \int_{C_{\epsilon}} \left(\varphi(x) \frac{\partial \log r}{\partial n} - (\log r) \frac{\partial \varphi}{\partial n} \right) ds + \int_{C_R} \left(\varphi \frac{\partial \log r}{\partial n} - (\log r) \frac{\partial \varphi}{\partial n} \right) ds \end{aligned} \quad (7.130)$$

Pero en esa corona

$$\Delta \log r = 0 \quad (7.131)$$

la normal exterior es el radio, $n = r$ para C_R y $n = -r$ para C_{ϵ} y el arco $ds = r d\theta$, $r = \epsilon$ o $r = R$. Además $\varphi(x)$ y sus derivadas se hacen cero en C_R . Por tanto:

$$\begin{aligned} & - \int_{\epsilon < r < R} (\log r) \Delta \varphi(x) dx = -\epsilon \int_0^{2\pi} \left(\frac{\varphi(x)}{r} - (\log r) \frac{\partial \varphi}{\partial r} \right) d\theta + \\ & \int_0^{2\pi} \left(\frac{\varphi(x)}{r} - (\log r) \frac{\partial \varphi}{\partial r} \right) R d\theta = - \int_0^{2\pi} \varphi(x) d\theta + (\epsilon \log \epsilon) \int_0^{2\pi} \frac{\partial \varphi}{\partial r} d\theta \end{aligned} \quad (7.132)$$

Tenemos entonces:

$$\begin{aligned} (\Delta \log r)(\varphi) &= \lim_{\epsilon \rightarrow 0} \int_{\epsilon < r < R} (\log r) \Delta \varphi(x) dx = \\ & - \lim_{\epsilon \rightarrow 0} (\epsilon \log \epsilon) \int_0^{2\pi} \frac{\partial \varphi}{\partial r} d\theta + \lim_{\epsilon \rightarrow 0} \int_0^{2\pi} \varphi(x) d\theta = 2\pi \varphi(0) \end{aligned} \quad (7.133)$$

es decir

$$\Delta \log r = 2\pi\delta \quad (7.134)$$

Este resultado tiene su correspondiente en cualquier número de variables:

$$\Delta \frac{1}{r^{n-2}} = -(n-2)\sigma_n\delta, \quad \sigma_n = \frac{2\pi^{n/2}}{\Gamma\left(\frac{n}{2}\right)} \quad (7.135)$$

siendo σ_n el área de la esfera unidad en dimensión n .

Veamos cuál es la situación en la ecuación de ondas, para $n = 2$. Demostraremos que

$$(\partial_t^2 - c^2\partial_x^2)\frac{1}{2c}\theta(ct - |x|) = \delta(x, t) \quad (7.136)$$

Aplicamos la distribución de la izquierda a una función prueba

$$\begin{aligned} (\partial_t^2 - c^2\partial_x^2)\frac{1}{2c}\theta(ct - |x|)(\varphi) &= \frac{1}{2c}\theta(ct - |x|)(\partial_t^2\varphi - c^2\partial_x^2\varphi) = \\ \frac{1}{2c}\int_{\mathbf{R}^2}\theta(ct - |x|)(\partial_t^2\varphi) dt dx - \frac{c}{2}\int_{\mathbf{R}^2}\theta(ct - |x|)(\partial_x^2\varphi) dx dt &= \\ \frac{1}{2c}\int_{-\infty}^{\infty}\left(\int_{|x|/c}^{\infty}(\partial_t^2\varphi) dt\right) dx - \frac{c}{2}\int_0^{\infty}\left(\int_{-ct}^{ct}(\partial_x^2\varphi) dx\right) dt &= \\ -\frac{1}{2c}\int_{-\infty}^{\infty}\partial_t\varphi\left(x, \frac{|x|}{c}\right) dx - \frac{c}{2}\int_0^{\infty}(\partial_x\varphi(ct, t) - \partial_x\varphi(-ct, t)) dt & \end{aligned} \quad (7.137)$$

Separamos la primera integral en dos intervalos:

$$\begin{aligned} &-\frac{1}{2c}\int_{-\infty}^0\partial_t\varphi\left(x, \frac{|x|}{c}\right) dx - \frac{1}{2c}\int_0^{\infty}\partial_t\varphi\left(x, \frac{|x|}{c}\right) dx - \\ &\frac{c}{2}\int_0^{\infty}(\partial_x\varphi(ct, t) - \partial_x\varphi(-ct, t)) dt = \\ &-\frac{1}{2c}\int_0^{\infty}\partial_t\varphi\left(-x, \frac{x}{c}\right) dx + \frac{c}{2}\int_0^{\infty}\partial_x\varphi(-ct, t) dt - \\ &\frac{1}{2c}\int_0^{\infty}\partial_t\varphi\left(x, \frac{x}{c}\right) dx - \frac{c}{2}\int_0^{\infty}\partial_x\varphi(ct, t) dt \\ &= -\frac{1}{2}\int_0^{\infty}(\partial_t\varphi(-ct, t) - c\partial_x\varphi(-ct, t)) dt - \\ &\frac{1}{2}\int_0^{\infty}(\partial_t\varphi(ct, t) + c\partial_x\varphi(ct, t)) dt \end{aligned} \quad (7.138)$$

Pero la derivada total de φ es

$$\frac{d}{dt}\varphi(-ct, t) = \partial_t\varphi(-ct, t) - c\partial_x\varphi(-ct, t) \quad (7.139)$$

$$\frac{d}{dt}\varphi(ct, t) = \partial_t\varphi(ct, t) + c\partial_x\varphi(ct, t) \quad (7.140)$$

y por tanto tenemos:

$$\begin{aligned} &= -\frac{1}{2}\int_0^{\infty}\frac{d}{dt}\varphi(-ct, t) dt - \frac{1}{2}\int_0^{\infty}\frac{d}{dt}\varphi(ct, t) dt = \frac{1}{2}\varphi(0, 0) + \frac{1}{2}\varphi(0, 0) \\ &= \varphi(0, 0) \end{aligned} \quad (7.141)$$

7.5. Distribuciones y transformación de Fourier

La transformación de Fourier es un elemento indispensable en la teoría de ecuaciones diferenciales (lineales de coeficientes constantes especialmente). Por esta razón, y vistas las aplicaciones que las distribuciones tienen en este campo, nos interesa establecer bajo que condiciones y que propiedades tiene la transformación de Fourier en los espacios de distribuciones. Como veremos es necesario limitar éstas, pues la transformada de Fourier de una función de soporte compacto no lo es y este problema se traslada a las distribuciones. En las secciones siguientes definiremos un nuevo espacio de distribuciones y en él la transformación de Fourier.

7.5.1. El espacio \mathcal{S}_n

Este espacio (el espacio de las funciones que decrecen rápidamente) contiene al espacio $\mathcal{D}(\mathbf{R}^n)$ y su topología se construye de una forma similar. Está formado por las funciones $f \in \mathcal{C}^\infty(\mathbf{R}^n)$ para las cuales:

$$\sup_{|J| \leq N} \sup_{x \in \mathbf{R}^n} (1 + |x|^2)^N |D^J f(x)| < \infty, \quad N = 0, 1, 2, \dots \quad (7.142)$$

Es decir, f y todas sus derivadas decrecen más deprisa que cualquier polinomio. Obviamente una función de soporte compacto está en \mathcal{S}_n , pero las funciones de \mathcal{S}_n no tienen por qué ser de soporte compacto. Estas funciones, sus derivadas y el producto de un polinomio por ellas están en $L^1(\mathbf{R}^n)$.

El espacio \mathcal{S}_n , con la topología definida por la familia de normas anteriores es un espacio localmente convexo completo (un espacio de Fréchet).

En el espacio \mathcal{S}_n la aplicación multiplicar por un polinomio y multiplicar por otra función de \mathcal{S}_n son continuas, así como la de derivar a cualquier orden.

7.5.2. La transformación de Fourier en $L^1(\mathbf{R}^n)$

Sea $f \in L^1(\mathbf{R}^n)$. Definimos su transformada de Fourier por:

$$\hat{f}(k) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} f(x) e^{-ik \cdot x} dx, \quad k \in \mathbf{R}^n \quad (7.143)$$

con $k \cdot x = \sum_{i=1}^n k_i x_i$. Está claro que esta función existe en las condiciones establecidas. Recordemos algunas de sus propiedades más importantes. La transformada de una función trasladada se puede obtener fácilmente de la transformada de la función original.

$$\begin{aligned} (\widehat{\tau_y f})(k) &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} f(x-y) e^{-ik \cdot x} dx = \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} f(x) e^{-ik \cdot (x+y)} dx \\ &= e^{-ik \cdot y} \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} f(x) e^{-ik \cdot x} dx = e^{-ik \cdot y} \hat{f}(k) \end{aligned} \quad (7.144)$$

De forma similar, la transformada de una exponencial por una función produce una traslación:

$$(\widehat{e^{ik' \cdot x} f})(k) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} f(x) e^{-i(k-k') \cdot x} dx = \tau_{k'} \hat{f}(k) \quad (7.145)$$

La transformada de Fourier de una convolución es el producto de las transformadas de Fourier (salvo un factor).

$$\begin{aligned} &\frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} \left(\int_{\mathbf{R}^n} f(x-y) g(y) dy \right) e^{-ik \cdot x} dx \\ &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} g(y) \left(\int_{\mathbf{R}^n} f(x-y) e^{-ik \cdot (x-y)} dx \right) e^{-ik \cdot y} dy \\ &= (2\pi)^{n/2} \hat{f}(k) \hat{g}(k) \end{aligned} \quad (7.146)$$

La transformación de Fourier es un operador unitario en $L^2(\mathbf{R}^n)$, pero su actuación es más complicada de introducir (pues no todas las funciones de $L^2(\mathbf{R}^n)$ están en $L^1(\mathbf{R}^n)$).

7.5.3. La transformación de Fourier en \mathcal{S}_n

La transformada de Fourier es una aplicación lineal continua de \mathcal{S}_n en \mathcal{S}_n . Para demostrar este resultado, veamos primeramente el comportamiento con respecto a las derivadas. Sea P un polinomio en las variables $(\lambda_1, \dots, \lambda_n)$ con coeficientes en \mathbf{C} . Se define $P(D)$, donde D es el operador de derivación, como:

$$P(\lambda) = \sum_J c_J \lambda^J, \quad P(D) = \sum_J c_J D^J \quad (7.147)$$

Entonces,

$$P(\widehat{D})f = P(ik)\hat{f} \quad (7.148)$$

En efecto⁶, si hacemos la convolución con $e^{ik \cdot x}$ y aplicamos $P(D)$:

$$(P(D)f) * e^{ik \cdot x} = f * (P(D)e^{ik \cdot x}) \quad (7.149)$$

pero

$$P(D)e^{ik \cdot x} = \sum_J c_J D^J e^{ik \cdot x} = \sum_J c_J (ik)^J e^{ik \cdot x} = P(ik)e^{ik \cdot x} \quad (7.150)$$

Por tanto,

$$(P(D)f) * e^{ik \cdot x} = f * (P(ik)e^{ik \cdot x}) = P(ik)(f * e^{ik \cdot x}) \quad (7.151)$$

Si escribimos explícitamente las relación anterior:

$$\frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} (P(D)f)(x-y)e^{ik \cdot y} dy = \frac{1}{(2\pi)^{n/2}} P(ik) \int_{\mathbf{R}^n} f(x-y)e^{ik \cdot y} dy \quad (7.152)$$

y sustituyendo en $x=0$ (y cambiando y por $-y$):

$$\frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} (P(D)f)(y)e^{-ik \cdot y} dy = \frac{1}{(2\pi)^{n/2}} P(ik) \int_{\mathbf{R}^n} f(y)e^{-ik \cdot y} dy \quad (7.153)$$

es decir

$$P(\widehat{D})f = P(ik)\hat{f} \quad (7.154)$$

como queríamos probar. También podemos probar una igualdad inversa en cierto sentido a la anterior:

$$P(\widehat{x})f = P(-iD)\hat{f} \quad (7.155)$$

Veámosla para la derivada parcial con respecto a x_1 .

$$\frac{\hat{f}(k_1 + \epsilon, k_2, \dots, k_n) - \hat{f}(k)}{\epsilon} = \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} f(x)e^{-ik \cdot x} \frac{e^{i\epsilon x_1} - 1}{\epsilon} dx \quad (7.156)$$

Si $\epsilon \rightarrow 0$ y aplicamos el teorema de la convergencia dominada,

$$-i \frac{\partial}{\partial k_1} \hat{f}(k) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} x_1 f(x)e^{-ik \cdot x} dx = x_1 \widehat{f}(x) \quad (7.157)$$

resultado que se generaliza a cualquier polinomio P .

Para finalizar demostremos que la transformación de Fourier es una aplicación continua del espacio \mathcal{S}_n en si mismo. Primeramente demostraremos que si $f \in \mathcal{S}_n$ entonces $\hat{f} \in \mathcal{S}_n$. Si $f \in \mathcal{S}_n$, la función $x^J f(x)$ también está en ese espacio. Pero hemos visto que la transformada de Fourier de esta función es (salvo constantes) la derivada $D^J \hat{f}$. Y el producto de un polinomio por esta transformada es la transformada de una derivada de $x^J f(x)$. Es decir, el producto de cualquier polinomio por \hat{f} o por $D^J \hat{f}$ está acotado, por lo que $\hat{f} \in \mathcal{S}_n$ (lo que hemos aplicado es que la función f , sus derivadas y polinomios por sus derivadas están en $L^1(\mathbf{R}^n)$). El problema de la continuidad es más sencillo (aunque mucho más complicado si se ataca desde el principio). Si una sucesión de funciones en \mathcal{S}_n tiende a una función en \mathcal{S}_n , también lo hace en $L^1(\mathbf{R}^n)$ y como consecuencia las transformadas de Fourier convergen puntualmente a la transformada de la función límite. Una aplicación del teorema del gráfico cerrado (el punto difícil) permite concluir la continuidad de la transformación de Fourier de \mathcal{S}_n en \mathcal{S}_n .

⁶En el caso de $L^1(\mathbf{R}^n)$ esta propiedad no se ha tratado, pues allí las funciones no tienen por qué ser diferenciables.

7.5.4. La fórmula de inversión

Una vez introducida la transformación de Fourier en \mathcal{S}_n y demostrado que se trata de una aplicación lineal y continua de \mathcal{S}_n en \mathcal{S}_n , vamos a ver ahora como es posible recuperar la función original a partir de su transformada. Se tiene la siguiente transformación inversa:

$$g(x) = \int_{\mathbf{R}^n} \hat{g}(k) e^{ik \cdot x} dk, \quad g \in \mathcal{S}_n \quad (7.158)$$

Pero, además, la transformación de Fourier es una aplicación lineal, biyectiva y continua, tanto ella como su inversa, de \mathcal{S}_n en \mathcal{S}_n . Su periodo es 4 (aplicada cuatro veces se obtiene la identidad).

Dadas dos funciones $f, g \in L^1(\mathbf{R}^n)$, se tiene, a partir de la siguiente integral:

$$\frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} \int_{\mathbf{R}^n} f(x) g(k) e^{-ik \cdot x} dx dk \quad (7.159)$$

y usando Fubini,

$$\int_{\mathbf{R}^n} \hat{f}(k) g(k) dk = \int_{\mathbf{R}^n} f(x) \hat{g}(x) dx \quad (7.160)$$

Supongamos ahora que $g \in \mathcal{S}_n$ y que $f(x) = h(x/\lambda)$, $h \in \mathcal{S}_n$, $\lambda > 0$. La transformada de Fourier de f es:

$$\begin{aligned} \hat{f}(k) &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} f(x) e^{-ik \cdot x} dx = \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} h\left(\frac{x}{\lambda}\right) e^{-ik \cdot x} dx \\ &= \frac{1}{(2\pi)^{n/2}} \lambda^n \int_{\mathbf{R}^n} h(y) e^{-i\lambda k \cdot y} dy = \lambda^n \hat{h}(\lambda k) \end{aligned} \quad (7.161)$$

y por tanto,

$$\int_{\mathbf{R}^n} \hat{h}(x) g\left(\frac{x}{\lambda}\right) dx = \int_{\mathbf{R}^n} \lambda^n \hat{h}(\lambda k) g(k) dk = \int_{\mathbf{R}^n} h\left(\frac{x}{\lambda}\right) \hat{g}(x) dx \quad (7.162)$$

Se puede aplicar el teorema de la convergencia dominada cuando $\lambda \rightarrow \infty$. En este caso,

$$g\left(\frac{x}{\lambda}\right) \rightarrow g(0), \quad h\left(\frac{x}{\lambda}\right) \rightarrow h(0) \quad (7.163)$$

lo que lleva a

$$g(0) \int_{\mathbf{R}^n} \hat{h}(x) dx = h(0) \int_{\mathbf{R}^n} \hat{g}(x) dx \quad (7.164)$$

Sea $h(x) = \exp(-\frac{1}{2}|x|^2)$. Su transformada de Fourier es

$$\hat{h}(k) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{1}{2}|x|^2} e^{-ik \cdot x} dx = \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{1}{2}|x|^2 - ik \cdot x} dx \quad (7.165)$$

Usando

$$\frac{1}{2}(|x|^2 + 2ik \cdot x - |k|^2) + \frac{1}{2}|k|^2 = \frac{1}{2}|x - ik|^2 + \frac{1}{2}|k|^2 \quad (7.166)$$

resulta

$$\hat{h}(k) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{1}{2}|x-ik|^2 - \frac{1}{2}|k|^2} dx = e^{-\frac{1}{2}|k|^2} \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{1}{2}|x-ik|^2} dx \quad (7.167)$$

La integral se hace fácilmente usando técnicas de variable compleja:

$$\int_{\mathbf{R}^n} e^{-\frac{1}{2}|x-ik|^2} dx = \prod_{i=1}^n \int_{\mathbf{R}} e^{-\frac{1}{2}(p-iq)^2} dp \quad (7.168)$$

Basta cerrar un circuito con el eje x (el eje p) la recta $y = -iq$ y dos segmentos perpendiculares que las unan. Cuando esos segmentos se llevan a infinito, la integral en el ciclo es cero (pues no hay singularidades)

y por tanto la integral en el eje y en la recta paralela son iguales (de sentido contrario en realidad). Pero sobre el eje es la integral gaussiana, y tenemos que:

$$\prod_{i=1}^n \int_{\mathbf{R}} e^{-\frac{1}{2}(p-iq)^2} dp = \prod_{i=1}^n (2\pi)^{1/2} = (2\pi)^{n/2} \quad (7.169)$$

con lo que

$$\hat{h}(k) = e^{-\frac{1}{2}|k|^2} \quad (7.170)$$

es decir, $\hat{h} = h$. Volvamos a la igualdad (7.164) (escribamos la variable como k)

$$g(0) \int_{\mathbf{R}^n} e^{-\frac{1}{2}|k|^2} dk = (2\pi)^{n/2} g(0) = \int_{\mathbf{R}^n} \hat{g}(k) dk \quad (7.171)$$

Despejando:

$$g(0) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} \hat{g}(k) dk \quad (7.172)$$

Para recuperar la fórmula de inversión, hagamos una traslación en x :

$$g(x) = (\tau_{-x}g)(0) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} (\widehat{\tau_{-x}g})(k) dk = \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} e^{ik \cdot x} \hat{g}(k) dk \quad (7.173)$$

De las fórmulas de las transformaciones directa e inversa se deduce que la transformación de Fourier es una aplicación inyectiva. Además al aplicar dos veces la transformación de Fourier se obtiene la función de partida pero en la variable $-x$. Aplicada cuatro veces se obtiene la función original. Entonces, la inversa de esta transformación es igual a componerla tres veces y por lo tanto si la transformación es continua, su inversa también lo es.

7.5.5. La transformación de Fourier en $L^2(\mathbf{R}^n)$

Como hemos dicho, la transformación de Fourier convierte funciones en $L^1(\mathbf{R})$ en funciones de $L^1(\mathbf{R})$. La transformación de Fourier en \mathcal{S}_n se extiende a una isometría lineal de $L^2(\mathbf{R}^n)$ en $L^2(\mathbf{R}^n)$.

En primer lugar, como \mathcal{S}_n es denso en $L^1(\mathbf{R}^n)$ y en $L^2(\mathbf{R}^n)$ (como se puede demostrar), la transformación se extiende sin dificultad a $L^1(\mathbf{R}^n) \cap L^2(\mathbf{R}^n)$. Posteriormente se pasa de $L^1(\mathbf{R}^n) \cap L^2(\mathbf{R}^n)$ a $L^2(\mathbf{R}^n)$. Téngase en cuenta que si $f \notin L^1(\mathbf{R}^n)$ no es obvio como definir la transformación de Fourier (el procedimiento pasa por definirlo en una base ortonormal y luego en sus combinaciones finitas y finalmente en todo el espacio). Al ser $L^2(\mathbf{R}^n)$ un espacio de Hilbert, la transformación de Fourier resulta ser un operador unitario en este espacio.

7.5.6. Distribuciones temperadas

Al igual como hicimos en el caso de las distribuciones sobre funciones prueba de soporte compacto, la definición de transformación de Fourier de distribuciones sobre \mathcal{S}_n se hace considerando su acción sobre estas funciones de decrecimiento rápido. Veamos antes algunas propiedades de \mathcal{S}_n .

Como hemos visto $\mathcal{D} \subset \mathcal{S}_n$, al menos como conjuntos. Pero desde un punto de vista topológico las relaciones son más estrechas. Se puede probar que el espacio \mathcal{D} es denso en \mathcal{S}_n (con la topología de \mathcal{S}_n) y que la inclusión es continua (con las topologías de \mathcal{D} y \mathcal{S}_n respectivamente).

Si F es un funcional lineal continuo sobre \mathcal{S}_n , $F : \mathcal{S}_n \rightarrow \mathbf{C}$, la inclusión $i : \mathcal{D}(\mathbf{R}^n) \rightarrow \mathcal{S}_n$ es continua, como hemos dicho, y por tanto la composición de ambas también lo es:

$$u_F : \mathcal{D}(\mathbf{R}^n) \rightarrow \mathbf{C}, \quad u_F = F \circ i \quad (7.174)$$

Además, como $\mathcal{D}(\mathbf{R}^n)$ es denso en \mathcal{S}_n , la correspondencia $F \rightarrow u_F$ es inyectiva. Tenemos un isomorfismo entre \mathcal{S}'_n (los funcionales F) y un espacio de funcionales lineales continuos sobre $\mathcal{D}(\mathbf{R}^n)$, es decir un subespacio del espacio de distribuciones $\mathcal{D}'(\mathbf{R}^n)$. A las distribuciones de este subespacio se les llama temperadas. En general se prescinde de u_F y se llama a la distribución F . De esta forma, las distribuciones

temperadas son los elementos de \mathcal{S}'_n y se caracterizan por el hecho de que su acción se puede extender de manera continua a \mathcal{S}_n .

Muchas de las distribuciones que se han estudiado previamente son temperadas. Pero no todas lo son. Por ejemplo e^x es una función localmente integrable y por tanto es una distribución. Pero no es temperada (crece demasiado deprisa). Si una distribución es de soporte compacto, es una distribución temperada. No todas las medidas de Borel dan lugar a distribuciones temperadas. Pero si

$$\int_{\mathbf{R}^n} (1 + |x|^2)^{-k} d\mu(x) < \infty \quad (7.175)$$

para algún $k > 0$, la distribución correspondiente es temperada.

De forma similar, si g es medible en \mathbf{R}^n y

$$\int_{\mathbf{R}^n} |(1 + |x|^2)^{-N} g(x)|^p dx = C < \infty \quad (7.176)$$

para algún $p \geq 1$ y $N > 0$, entonces g es una distribución temperada. Como consecuencia, las funciones de $L^p(\mathbf{R}^n)$ con $1 \leq p < \infty$, son distribuciones temperadas. Las demostraciones de todos estos ejemplos pasan por demostrar la continuidad del funcional cuando actúa sobre las funciones de \mathcal{S}_n .

Como en el caso de las distribuciones, si u es una distribución temperada, sus derivadas también lo son, y el producto de cualquier polinomio o cualquier función de \mathcal{S}_n por u también lo es.

7.5.7. Transformada de Fourier de una distribución temperada

Definimos ahora la transformada de Fourier de una distribución temperada:

$$\hat{u}(\phi) = u(\hat{\phi}), \quad u \in \mathcal{S}'_n, \phi \in \mathcal{S}_n \quad (7.177)$$

La transformación de Fourier es una aplicación continua de \mathcal{S}_n en \mathcal{S}_n . Como consecuencia $\hat{u} \in \mathcal{S}'_n$. Las propiedades de la transformación de Fourier en \mathcal{S}' son similares a las de la transformación en \mathcal{S} .

Las funciones de $L^p(\mathbf{R}^n)$ definen distribuciones temperadas como hemos visto. Pero las funciones en estos espacios para $p = 1, 2$ tienen transformadas de Fourier. Se puede probar que las definiciones son consistentes. Es decir, la transformada de Fourier de una de estas funciones da lugar a una distribución que es la transformada de Fourier de la distribución asociada a la función original. La demostración es consecuencia de una propiedad que se probó al hallar la fórmula de inversión:

$$\int_{\mathbf{R}^n} \hat{f}g = \int_{\mathbf{R}^n} f\hat{g} \quad (7.178)$$

Tal y como hemos dicho, las propiedades de la transformación de Fourier en \mathcal{S}'_n son similares a las de la correspondiente transformación en \mathcal{S}_n . En particular, la transformación de Fourier es lineal, continua y biyectiva y su inversa también lo es. Además tiene periodo 4. En cuanto a su comportamiento con respecto a la derivada se tiene:

$$(P(\widehat{D})u)(\phi) = (P(D)u)\hat{\phi} = u(\tilde{P}(D)\hat{\phi}) \quad (7.179)$$

donde \tilde{P} es el polinomio que se obtiene al sustituir D^J por $(-1)^{|J|}D^J$. Pero de las propiedades de la transformación de Fourier en \mathcal{S}_n sabemos que:

$$\widehat{P(x)f} = P(-iD)\hat{f} \quad (7.180)$$

De igual forma podemos probar

$$\widehat{P(x)u} = P(-iD)\hat{u} \quad (7.181)$$

7.5.8. Convolución en \mathcal{S}_n y \mathcal{S}'_n

Al igual que para las funciones en $\mathcal{D}(\mathbf{R}^n)$ y las distribuciones en $\mathcal{D}'(\mathbf{R}^n)$ es posible estudiar la convolución para las distribuciones temperadas. Sea $u \in \mathcal{S}'_n$, $\phi \in \mathcal{S}_n$. Se define la convolución:

$$(u * \phi)(x) = u(\tau_x \check{\phi}) \quad (7.182)$$

Las propiedades son muy similares a las estudiadas entonces. Por ejemplo,

$$u * \phi \in \mathcal{C}^\infty, \quad D^J(u * \phi) = (D^J u) * \phi = u * (D^J \phi), \quad u \in \mathcal{S}'_n, \quad \phi \in \mathcal{S}_n \quad (7.183)$$

La distribución $u * \phi$ es de crecimiento polinómico (y por tanto temperada). Además

$$(u * \phi) * \psi = u * (\phi * \psi), \quad u \in \mathcal{S}'_n, \quad \phi, \psi \in \mathcal{S}_n \quad (7.184)$$

En relación con la transformación de Fourier las propiedades más interesantes son:

$$\widehat{u * \phi} = \hat{\phi} \hat{u}, \quad \widehat{u * \phi} = \hat{\phi} \hat{u} \quad (7.185)$$

7.6. Ejemplos de transformadas de Fourier de distribuciones

Hemos calculado anteriormente la transformada de Fourier de la función de \mathcal{S}_n , $e^{-\frac{1}{2}|x|^2}$ (usaremos a menudo la notación $\mathcal{F}(f) = \hat{f}$):

$$\mathcal{F}(e^{-\frac{1}{2}|x|^2}) = e^{-\frac{1}{2}|k|^2} \quad (7.186)$$

Este resultado se generaliza al caso en que $|x|^2$ se sustituye por una forma cuadrática $x^T Q x$ definida positiva (Q simétrica):

$$\mathcal{F}(e^{-\frac{1}{2}x^T Q x}) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{1}{2} \sum_{i,j=1}^n q_{ij} x_i x_j} e^{-ik \cdot x} dx \quad (7.187)$$

Mediante un cambio de base es posible reducir la forma cuadrática a una suma de cuadrados:

$$x = Py, \quad \langle Qx, x \rangle = x^T Q x = y^T P^T Q P y = \sum_{i=1}^n y_i^2 \quad (7.188)$$

$$Q = (P P^T)^{-1}, \quad \det Q = \frac{1}{(\det P)^2} \quad (7.189)$$

Cambiando la variable de x a y

$$dx = |\det P| dy \quad (7.190)$$

$$\begin{aligned} \mathcal{F}(e^{-\frac{1}{2}x^T Q x}) &= \frac{|\det P|}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{1}{2} \sum_{i=1}^n y_i^2} e^{-ik \cdot (Py)} dy \\ &= \frac{|\det P|}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{1}{2} \sum_{i=1}^n y_i^2} e^{-i(P^T k) \cdot y} dy \\ &= |\det P| \prod_{i=1}^n \frac{1}{(2\pi)^{1/2}} \int_{\mathbf{R}} e^{-\frac{1}{2}(y_i^2 + 2ik'_i y_i)} dy \end{aligned} \quad (7.191)$$

con $k' = P^T k$. Como

$$y_i^2 + 2ik'_i y_i = (y_i + ik'_i)^2 + (k'_i)^2 \quad (7.192)$$

$$= |\det P| \prod_{i=1}^n \frac{e^{-\frac{1}{2}(k'_i)^2}}{(2\pi)^{1/2}} \int_{\mathbf{R}} e^{-\frac{1}{2}(y_i + ik'_i)^2} dy \quad (7.193)$$

$$= |\det P| e^{-\frac{1}{2}|k'|^2} = |\det P| e^{-\frac{1}{2}(P P^T k) \cdot k} \quad (7.194)$$

es decir:

$$\mathcal{F}(e^{-\frac{1}{2}(Qx) \cdot x}) = \frac{1}{\sqrt{\det Q}} e^{-\frac{1}{2}(Q^{-1}k) \cdot k} \quad (7.195)$$

Una propiedad que se aplicará en numerosas ecuaciones es la continuidad de la transformación de Fourier. En particular se tiene

$$\mathcal{F}(\phi)(k) = \lim_{a \rightarrow \infty} \int_{-a}^a \phi(x) e^{-ik \cdot x} dx, \quad \phi \in \mathcal{S}_n \quad (7.196)$$

donde la convergencia se entiende en \mathcal{S}'_n

Por ejemplo, calculemos la transformada de Fourier de e^{ix^2} . La integral de Fresnel es:

$$\int_{-\infty}^{\infty} e^{ix^2} dx = \sqrt{\pi} e^{i\frac{\pi}{4}} \quad (7.197)$$

Si queremos calcular la transformada de Fourier de e^{ix^2} en \mathcal{S}_1 ,

$$\begin{aligned} \int_{-a}^a e^{ix^2} e^{-ikx} dx &= \int_{-a}^a e^{i(x^2 - kx)} dx = e^{-\frac{1}{4}ik^2} \int_{-a}^a e^{i(x - \frac{1}{2}k)^2} dx \\ &= e^{-\frac{1}{4}ik^2} \int_{-a - \frac{1}{2}k}^{a + \frac{1}{2}k} e^{iy^2} dy \end{aligned} \quad (7.198)$$

Si se hace $a \rightarrow \infty$ la integral converge (uniformemente en k , en un intervalo acotado)

$$\lim_{a \rightarrow \infty} \int_{-a}^a e^{ix^2} e^{-ikx} dx = \sqrt{\pi} e^{-\frac{1}{4}(k^2 - \pi)} \quad (7.199)$$

y por tanto

$$\mathcal{F}(e^{ix^2}) = \frac{1}{\sqrt{2}} e^{-\frac{1}{4}(k^2 - \pi)} \quad (7.200)$$

Queda solo un problema y es que la ecuación anterior sería válida sobre las funciones de $\mathcal{D}(\mathbf{R})$ (que tienen el soporte compacto, por la cuestión de la convergencia uniforme sobre intervalos compactos). Pero como $\mathcal{D}(\mathbf{R})$ es denso en \mathcal{S}_1 la fórmula es cierta en \mathcal{S}'_1 .

Calculemos la transformada de Fourier de la función

$$\phi(x) = \frac{1}{r^2}, \quad r = |x|, \quad x \in \mathbf{R}^3 \quad (7.201)$$

Como antes, integremos en un intervalo $[-a, a]$ y hagamos tender a a infinito. Tomando coordenadas esféricas (con k en la dirección del eje z)

$$\int_{-a}^a \frac{e^{-ik \cdot x}}{r^2} dx = \int_{-a}^a dr \int_0^\pi d\theta e^{-i|k|r \cos \theta} \sin \theta \int_0^{2\pi} d\phi = \frac{4\pi}{|k|} \int_{-a}^a \frac{\text{sen } |k|r}{r} dr \quad (7.202)$$

El problema ahora es calcular la integral

$$\int_0^\infty \frac{\text{sen } |k|r}{r} dr \quad (7.203)$$

Pero éste es un problema clásico de la teoría de residuos.

$$\int_{-\infty}^{\infty} \frac{\text{sen } |k|r}{r} dr \quad (7.204)$$

$$\begin{aligned} &= \frac{1}{2i} \lim_{\epsilon \rightarrow 0} \left(\int_{-\infty}^{-\epsilon} \frac{e^{i|k|r}}{r} dr + \int_{\epsilon}^{\infty} \frac{e^{i|k|r}}{r} dr - \int_{-\infty}^{-\epsilon} \frac{e^{-i|k|r}}{r} dr - \int_{\epsilon}^{\infty} \frac{e^{-i|k|r}}{r} dr \right) \\ &= \frac{1}{i} \lim_{\epsilon \rightarrow 0} \left(\int_{-\infty}^{-\epsilon} \frac{e^{i|k|r}}{r} dr + \int_{\epsilon}^{\infty} \frac{e^{i|k|r}}{r} dr \right) \end{aligned} \quad (7.205)$$

El eje x , donde se está integrando se cierra con un semicírculo cuyo radio tenderá a infinito en el semiplano superior (de modo que allí la integral tiende a cero), y la singularidad del origen se salva con un semicírculo de radio ϵ , recorrido en sentido horario. Al ser el origen un polo simple el valor de la integral cuando ϵ tiende a cero es πi veces el residuo del integrando en ese punto. En nuestro caso el residuo es 1 y se tiene por tanto:

$$\int_{-\infty}^{\infty} \frac{\operatorname{sen} |k|r}{r} dr = \frac{1}{i} i\pi = \pi \quad (7.206)$$

Finalmente,

$$\mathcal{F}\left(\frac{1}{r^2}\right) = \sqrt{\frac{\pi}{2}} \frac{1}{|k|} \quad (7.207)$$

El cálculo de transformadas de Fourier es en general una tarea complicada y solo para algunas funciones o distribuciones es posible alcanzar un resultado que se pueda expresar de una forma compacta. Veamos por ejemplo el cálculo de la transformada de Fourier de la pseudofunción definida en $n = 2$ por

$$\mathcal{P}\left(\frac{1}{r^2}\right)(\phi) = \int_{r < 1} \frac{\phi(x) - \phi(0)}{r^2} dx + \int_{r > 1} \frac{\phi(x)}{r^2} dx \quad (7.208)$$

Hagamos el cálculo a partir de la definición:

$$\mathcal{F}\left(\mathcal{P}\left(\frac{1}{r^2}\right)\right)(\phi) = \mathcal{P}\left(\frac{1}{k^2}\right)\hat{\phi} = \int_{|k| < 1} \frac{\hat{\phi}(k) - \hat{\phi}(0)}{|k|^2} dk + \int_{|k| > 1} \frac{\hat{\phi}(k)}{|k|^2} dk \quad (7.209)$$

pero

$$\hat{\phi}(k) - \hat{\phi}(0) = \frac{1}{2\pi} \int_{\mathbf{R}^2} \phi(x)(e^{-ik \cdot x} - 1) dx \quad (7.210)$$

luego

$$\begin{aligned} \mathcal{F}\left(\mathcal{P}\left(\frac{1}{r^2}\right)\right)(\phi) &= \int_{|k| < 1} \frac{1}{|k|^2} \frac{1}{2\pi} \int_{\mathbf{R}^2} \phi(x)(e^{-ik \cdot x} - 1) dx dk \\ &+ \int_{|k| > 1} \frac{1}{|k|^2} \frac{1}{2\pi} \int_{\mathbf{R}^2} \phi(x)e^{-ik \cdot x} dx dk \end{aligned} \quad (7.211)$$

$$\begin{aligned} &= \frac{1}{2\pi} \int_{\mathbf{R}^2} \phi(x) \left(\int_{|k| < 1} \frac{e^{-ik \cdot x} - 1}{|k|^2} dk \right) dx \\ &+ \frac{1}{2\pi} \int_{\mathbf{R}^2} \phi(x) \left(\int_{|k| > 1} \frac{e^{-ik \cdot x}}{|k|^2} dk \right) dx \end{aligned} \quad (7.212)$$

$$\begin{aligned} \int_{|k| < 1} \frac{e^{-ik \cdot x} - 1}{|k|^2} dk &= \int_0^{2\pi} \int_0^1 \frac{e^{-i|k|r \cos \varphi} - 1}{|k|^2} |k| d|k| d\varphi \\ &= \int_0^1 \frac{d|k|}{|k|} \left(\int_0^{2\pi} (e^{-i|k|r \cos \varphi} - 1) d\varphi \right) \end{aligned} \quad (7.213)$$

Pero la integral en el ángulo φ se puede expresar en términos de la función de Bessel J_0 :

$$J_0(z) = \frac{1}{\pi} \int_0^\pi \cos(z \cos \varphi) d\varphi = \frac{1}{2\pi} \int_0^\pi (e^{iz \cos \varphi} + e^{-iz \cos \varphi}) d\varphi \quad (7.214)$$

$$\int_0^\pi e^{iz \cos \varphi} d\varphi = \int_\pi^{2\pi} e^{iz \cos(\varphi - \pi)} d\varphi = \int_\pi^{2\pi} e^{-iz \cos \varphi} d\varphi \quad (7.215)$$

luego

$$J_0(z) = \frac{1}{2\pi} \int_0^{2\pi} e^{-iz \cos \varphi} d\varphi \quad (7.216)$$

Sustituyendo,

$$\begin{aligned} \int_0^1 \frac{d|k|}{|k|} \left(\int_0^{2\pi} (e^{-i|k|r \cos \varphi} - 1) d\varphi \right) &= 2\pi \int_0^1 \frac{J_0(|k|r) - 1}{|k|} d|k| \\ &= 2\pi \int_0^r \frac{J_0(t) - 1}{t} dt \end{aligned} \quad (7.217)$$

luego

$$\frac{1}{2\pi} \int_{\mathbf{R}^2} \phi(x) \left(\int_{|k|<1} \frac{e^{-ik \cdot x} - 1}{|k|^2} dk \right) dx = \int_{\mathbf{R}^2} \phi(x) \left(\int_0^r \frac{J_0(t) - 1}{t} dt \right) dx \quad (7.218)$$

La otra integral es:

$$\frac{1}{2\pi} \int_{\mathbf{R}^2} \phi(x) \left(\int_{|k|>1} \frac{e^{-ik \cdot x}}{|k|^2} dk \right) dx = \int_{\mathbf{R}^2} \phi(x) \left(\int_r^\infty \frac{J_0(t)}{t} dt \right) dx \quad (7.219)$$

y la suma es:

$$\int_{\mathbf{R}^2} \phi(x) \left(\int_0^r \frac{J_0(t) - 1}{t} dt + \int_r^\infty \frac{J_0(t)}{t} dt \right) dx \quad (7.220)$$

Podemos escribir

$$\int_0^r \frac{J_0(t) - 1}{t} dt + \int_r^\infty \frac{J_0(t)}{t} dt = - \int_1^r \frac{dt}{t} + \int_0^1 \frac{J_0(t) - 1}{t} dt + \int_1^\infty \frac{J_0(t)}{t} dt \quad (7.221)$$

Como

$$\int_1^r \frac{dt}{t} = \log r \quad (7.222)$$

se tiene:

$$\mathcal{F} \left(\mathcal{P} \left(\frac{1}{r^2} \right) \right) (\phi) = \int_{\mathbf{R}^2} \phi(x) (c_0 - \log r) dx \quad (7.223)$$

donde

$$c_0 = \int_0^1 \frac{J_0(t) - 1}{t} dt + \int_1^\infty \frac{J_0(t)}{t} dt \quad (7.224)$$

y finalmente:

$$\mathcal{F} \left(\mathcal{P} \left(\frac{1}{r^2} \right) \right) = c_0 - \log |k| \quad (7.225)$$

La transformada de Fourier de la delta de Dirac es muy sencilla

$$\mathcal{F}(\delta)(\phi) = \delta(\hat{\phi}) = \hat{\phi}(0) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} \phi(x) dx \quad (7.226)$$

y por lo tanto,

$$\mathcal{F}(\delta) = \frac{1}{(2\pi)^{n/2}} \quad (7.227)$$

Aplicando la transformada inversa, (formalmente) se tiene

$$\delta(x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} \frac{1}{(2\pi)^{n/2}} e^{ik \cdot x} dk = \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{ik \cdot x} dk \quad (7.228)$$

Como consecuencia de lo anterior se tiene también

$$\mathcal{F}(1) = (2\pi)^{n/2} \delta \quad (7.229)$$

La transformada de Fourier de $\theta(R - |x|)$ se puede calcular directamente

$$\frac{1}{\sqrt{2\pi}} \int_{\mathbf{R}} \theta(R - |x|) e^{-ikx} dx = \frac{1}{(2\pi)^{1/2}} \int_{-R}^R e^{-ikx} dx \quad (7.230)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-R}^R e^{-ikx} dx = \frac{1}{\sqrt{2\pi}} \frac{-1}{ik} (e^{-ikR} - e^{ikR}) = \sqrt{\frac{2}{\pi}} \frac{\text{sen } kR}{k} \quad (7.231)$$

Para calcular la transformada de Fourier de la función paso (en \mathbf{R}) calculemos la de $\theta(x)e^{-ax}$ con $a > 0$:

$$\frac{1}{\sqrt{2\pi}} \int_{\mathbf{R}} \theta(x) e^{-ax} e^{-ikx} dx = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} e^{-(a+ik)x} dx = -\frac{i}{\sqrt{2\pi}} \frac{1}{k-ia} \quad (7.232)$$

es decir

$$\mathcal{F}(e^{-ax}\theta) = -\frac{i}{\sqrt{2\pi}} \frac{1}{k-ia} \quad (7.233)$$

Tomando límites cuando $a \rightarrow 0$:

$$\mathcal{F}(\theta) = \lim_{a \rightarrow 0} \mathcal{F}(e^{-ax}\theta) = -\frac{i}{\sqrt{2\pi}} \lim_{a \rightarrow 0} \frac{1}{k-ia} = -\frac{i}{\sqrt{2\pi}} \frac{1}{k-i0} \quad (7.234)$$

La transformada de $\theta(-x)$ se calcula de forma similar ($a > 0$):

$$\frac{1}{\sqrt{2\pi}} \int_{\mathbf{R}} \theta(-x) e^{ax} e^{-ikx} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 e^{(a-ik)x} dx = \frac{i}{\sqrt{2\pi}} \frac{1}{k+ia} \quad (7.235)$$

$$\mathcal{F}(\theta(-x)) = \frac{i}{\sqrt{2\pi}} \frac{1}{k+i0} \quad (7.236)$$

La transformada de la función signo se calcula ahora fácilmente:

$$\mathcal{F}(\text{sgn}(x)) = \mathcal{F}(\theta(x) - \theta(-x)) = -\frac{i}{\sqrt{2\pi}} \left(\frac{1}{k-i0} + \frac{1}{k+i0} \right) \quad (7.237)$$

y utilizando las fórmulas de Sochozki

$$\frac{1}{k \pm i0_+} = \mathcal{P}\left(\frac{1}{k}\right) \mp i\pi\delta \quad (7.238)$$

se tiene:

$$\mathcal{F}(\text{sgn}(x)) = -i\sqrt{\frac{2}{\pi}} \mathcal{P}\left(\frac{1}{k}\right) \quad (7.239)$$

La transformada de la pseudofunción $\mathcal{P}(1/x)$ es:

$$\mathcal{F}\left(\mathcal{P}\left(\frac{1}{x}\right)\right) = i\sqrt{\frac{\pi}{2}} \mathcal{F}^2(\text{sgn}(k)) = -i\sqrt{\frac{\pi}{2}} \text{sgn}(k) \quad (7.240)$$

Otra manera de verlo es, directamente de las fórmulas de Sochozki:

$$\frac{1}{k-i0_+} = \mathcal{P}\left(\frac{1}{k}\right) + i\pi\delta \quad (7.241)$$

$$\mathcal{F}^{-1}\left(\frac{1}{k-i0_+}\right) = \mathcal{F}^{-1}\left(\mathcal{P}\left(\frac{1}{k}\right)\right) + i\pi\mathcal{F}^{-1}(\delta) \quad (7.242)$$

$$i\sqrt{2\pi}\theta(x) = \mathcal{F}^{-1}\left(\mathcal{P}\left(\frac{1}{k}\right)\right) + \frac{i\pi}{\sqrt{2\pi}} \quad (7.243)$$

de donde

$$\mathcal{F}^{-1}\left(\mathcal{P}\left(\frac{1}{k}\right)\right) = i\sqrt{2\pi}\left(\theta(x) - \frac{1}{2}\right) = i\sqrt{\frac{\pi}{2}} \text{sgn}(x) \quad (7.244)$$

7.7. Soluciones fundamentales

Como ya hemos dicho, una solución fundamental de un operador diferencial es una distribución que verifica la ecuación:

$$P(D)\mathcal{E} = \delta \quad (7.245)$$

El propósito de este capítulo es presentar las propiedades más importantes de las soluciones fundamentales y calcularlas en un cierto número de casos que aparecen en las ecuaciones clásicas de la Física Matemática.

Si el operador del que se trata es un operador diferencial lineal de coeficientes constantes (aunque no necesariamente deben ser constantes para lo siguiente),

$$P(D) = \sum_J c_J D^J, \quad c_J \in \mathbf{R} \quad (7.246)$$

la acción sobre una distribución es

$$(P(D)u)(\varphi) = \left(\sum_J c_J D^J u\right)(\varphi) = \sum_J c_J (D^J u)(\varphi) = \sum_J (-1)^{|J|} c_J u(D^J \varphi) \quad (7.247)$$

Supongamos que la ecuación a resolver es

$$P(D)u = f \quad (7.248)$$

Si u es una distribución que verifica esa ecuación, se dice que es una solución del problema anterior en sentido de distribuciones o una solución generalizada del problema. Pero si f es una función continua en un cierto dominio y se tiene una solución generalizada suficientemente regular (con tantas derivadas como el orden del operador), entonces se trata de una solución en sentido clásico.

7.7.1. Propiedades y cálculo de soluciones fundamentales

Dado un operador diferencial lineal de coeficientes constantes, la solución fundamental, caso de existir, no será única. Pues si \mathcal{E}_0 es una solución en sentido de distribuciones de la ecuación homogénea $P(D)\mathcal{E}_0 = 0$, entonces la distribución $\mathcal{E} + \mathcal{E}_0$ es una solución fundamental si \mathcal{E} lo era.

Debido a las características del operador diferencial (lineal y de coeficientes constantes) la transformada de Fourier es una herramienta muy útil en el cálculo de soluciones fundamentales.

Sea la ecuación

$$P(D)\mathcal{E} = \delta \quad (7.249)$$

y apliquemos la transformación de Fourier a ambos miembros de la ecuación

$$\mathcal{F}(P(D)\mathcal{E}) = \mathcal{F}(\delta) \quad (7.250)$$

De acuerdo con lo desarrollado en el capítulo anterior,

$$P(ik)\hat{\mathcal{E}} = \frac{1}{(2\pi)^{n/2}} \quad (7.251)$$

En esta expresión $P(ik)$ es un polinomio en la variable vectorial k , de grado el del operador que estamos estudiando. La implicación es doble. La distribución \mathcal{E} es una solución fundamental de la ecuación diferencial si y solo si su transformada de Fourier satisface (7.251). Esta ecuación no se puede resolver simplemente despejando. Por ejemplo, las soluciones (en el sentido de distribuciones) de la ecuación

$$ku = 1 \quad (7.252)$$

no son obviamente

$$u = \frac{1}{k} \quad (7.253)$$

que ni siquiera está bien definida como distribución. Aunque sí que es verdad que la solución debe coincidir con esa función fuera de los ceros del polinomio (en esta caso fuera de $x = 0$). Veamos como

$$u = \mathcal{P}\left(\frac{1}{k}\right) \quad (7.254)$$

es una solución:

$$k\mathcal{P}\left(\frac{1}{k}\right)(\varphi) = \mathcal{P}\left(\frac{1}{k}\right)(k\varphi) = \text{VP} \int_{\mathbf{R}} \frac{1}{k} k\varphi(k) dk = \int_{\mathbf{R}} \varphi(k) dk = 1(\varphi) \quad (7.255)$$

Si ahora consideramos una distribución con soporte en $x = 0$ (la delta de Dirac) se tiene que

$$\mathcal{P}\left(\frac{1}{k}\right) + a\delta, \quad a \in \mathbf{R} \quad (7.256)$$

es también solución de la ecuación. En particular, usando las fórmulas de Sochozki,

$$\mathcal{P}\left(\frac{1}{k}\right) \mp i\delta = \frac{1}{k \pm i0} \quad (7.257)$$

son también soluciones de la ecuación. Desde otro punto de vista, la delta de Dirac es solución de la ecuación homogénea

$$ku = 0 \quad (7.258)$$

Volvamos a la ecuación (7.249). Un resultado de Malgrange y Ehrenpreis permite asegurar que siempre existe una solución fundamental.

Una vez que hemos calculado una solución fundamental de la ecuación diferencial, veamos como podemos usarla en el cálculo de soluciones de la ecuación homogénea. Se la ecuación

$$P(D)u = f \quad (7.259)$$

donde f es una distribución en \mathcal{D}' para la que existe la convolución $\mathcal{E} * f$. Entonces, $u = \mathcal{E} * f$ es una solución del problema no homogéneo. En efecto,

$$P(D)(\mathcal{E} * f) = (P(D)\mathcal{E}) * f = \delta * f = f \quad (7.260)$$

La solución de este problema es única (dentro de la clase de las distribuciones en \mathcal{S}'_n para las que existe la convolución con f). La razón es que en esa clase, la única solución del problema homogéneo es la trivial:

$$u = u * \delta = u * (P(D)\mathcal{E}) = P(D)(u) * \mathcal{E} = 0 \quad (7.261)$$

7.7.2. Soluciones fundamentales de ecuaciones diferenciales ordinarias

Sea L un operador lineal de coeficientes constante en una variable x

$$L = P(D) = a_n \frac{d^n}{dx^n} + a_{n-1} \frac{d^{n-1}}{dx^{n-1}} + \cdots + a_1 \frac{d}{dx} + a_0 \quad (7.262)$$

Queremos calcular una solución fundamental, $L\mathcal{E} = \delta$. Para ello calculamos la transformada de Fourier:

$$P(ik) = a_n (ik)^n + a_{n-1} (ik)^{n-1} + \cdots + a_1 ik + a_0 \quad (7.263)$$

y buscamos las soluciones de

$$P(ik)\hat{\mathcal{E}} = \frac{1}{\sqrt{2\pi}} \quad (7.264)$$

Sea z_0 un cero del polinomio $P(\lambda)$, supongamos que simple. En un entorno de ese punto (donde no hay otros ceros), la transformada de Fourier de la solución fundamental buscada es:

$$\hat{\mathcal{E}} \sim \frac{1}{\sqrt{2\pi}(\lambda - z_0)} = \frac{1}{\sqrt{2\pi}(ik - z_0)} \quad (7.265)$$

y por tanto, de acuerdo con lo ya visto,

$$\frac{1}{k - ia} = i\sqrt{2\pi}\mathcal{F}(e^{-ax}\theta) \quad (7.266)$$

se tiene

$$\frac{-i}{k + iz_0} = \sqrt{2\pi}\mathcal{F}(e^{z_0x}\theta), \quad \frac{1}{\sqrt{2\pi}}\mathcal{F}^{-1}\left(\frac{1}{ik - z_0}\right) = e^{z_0x}\theta(x) \quad (7.267)$$

lo que lleva a una solución fundamental dada por:

$$\mathcal{E} \sim e^{z_0x}\theta(x) \quad (7.268)$$

es decir, la función paso por una solución del problema homogéneo $Lu = 0$, que verifica $u(0) = 1$. Vemos que la solución del problema completo es

$$\mathcal{E}(x) = \theta(x)u(x) \quad (7.269)$$

donde $u(x)$ es una solución de $Lu = 0$ con las condiciones iniciales

$$u(0) = u'(0) = \dots = u^{(n-1)}(0) = 0, \quad u^{(n-1)}(0) = 1 \quad (7.270)$$

En efecto, las derivadas sucesivas de esta función son:

$$\mathcal{E}' = \theta'(x)u(x) + \theta(x)u'(x) = u\delta + \theta(x)u'(x) = u(0)\delta + \theta(x)u'(x) = \theta(x)u'(x) \quad (7.271)$$

y en general

$$\mathcal{E}^{(s)} = \theta(x)u^{(s)}(x), \quad n = 0, 1, \dots, n-1 \quad (7.272)$$

Finalmente

$$\mathcal{E}^{(n)} = \theta'(x)u^{(n-1)}(x) + \theta(x)u^{(n)}(x) = \delta + \theta(x)u^{(n)}(x) \quad (7.273)$$

Por tanto

$$L\mathcal{E} = L(u)\theta + \delta = \delta \quad (7.274)$$

La obtención de la solución fundamental puede hacerse directamente por la transformación inversa. Por ejemplo, calculemos una solución fundamental de la ecuación:

$$\mathcal{E}' + a\mathcal{E} = \delta \quad (7.275)$$

La ecuación transformada es:

$$ik\hat{\mathcal{E}} + a\hat{\mathcal{E}} = \frac{1}{\sqrt{2\pi}}, \quad \hat{\mathcal{E}} = -\frac{i}{\sqrt{2\pi}}\frac{1}{k - ia} \quad (7.276)$$

La transformada inversa es

$$\mathcal{E} = -\frac{i}{2\pi} \int_{-\infty}^{\infty} \frac{e^{ikx}}{k - ia} dk \quad (7.277)$$

La integral se hace en el plano complejo. Si $k > 0$, cerramos el eje real con un semicírculo en el semiplano superior. El comportamiento en el semicírculo viene dado por una exponencial decreciente. La única singularidad es el polo simple ia . Por tanto la transformada inversa es

$$-\frac{i}{2\pi} \int_C \frac{e^{ikx}}{k - ia} dk = -\frac{i}{2\pi} 2\pi i \operatorname{Res}\left(\frac{e^{ikx}}{k - ia}, ia\right) = e^{-ax} \quad (7.278)$$

Cuando $k < 0$ se cierra el circuito por el semiplano inferior donde no hay singularidades luego el resultado es 0. Finalmente

$$\mathcal{E} = \theta(x)e^{-ax} \quad (7.279)$$

Para la ecuación

$$\mathcal{E}'' + a^2\mathcal{E} = \delta \quad (7.280)$$

La ecuación transformada es:

$$-k^2 \hat{\mathcal{E}} + a^2 \hat{\mathcal{E}} = \frac{1}{\sqrt{2\pi}}, \quad \hat{\mathcal{E}} = -\frac{1}{\sqrt{2\pi}} \frac{1}{k^2 - a^2} \quad (7.281)$$

La transformada inversa es

$$\mathcal{E} = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{ikx}}{k^2 - a^2} dk \quad (7.282)$$

En este caso, el circuito de integración pasa por las singularidades ($\pm a$). Supongamos que las desplazamos sumando $i\epsilon$, ($\epsilon > 0$). Si $x > 0$ podemos cerrar por el semiplano superior. Los dos polos (simples) están en interior del circuito y por tanto

$$\begin{aligned} -\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{ikx}}{k^2 - a^2} dk &= -\lim_{\epsilon \rightarrow 0} \frac{1}{2\pi} 2\pi i \left(\frac{e^{i(a+i\epsilon)x}}{2(a+i\epsilon)} + \frac{e^{i(-a+i\epsilon)x}}{2(-a+i\epsilon)} \right) = \\ &= -\frac{i}{2a} (e^{iax} - e^{-iax}) = \frac{\text{sen } ax}{a} \end{aligned} \quad (7.283)$$

Si $x < 0$ cerramos por el semiplano inferior y el resultado es cero (no hay singularidades). Finalmente

$$\mathcal{E}_1 = \theta(x) \frac{\text{sen } ax}{a} \quad (7.284)$$

Pero podemos desplazar los polos al semiplano inferior. En este caso, la integral es cero cuando $x > 0$. Cuando $x < 0$

$$\begin{aligned} -\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{ikx}}{k^2 - a^2} dk &= \lim_{\epsilon \rightarrow 0} \frac{1}{2\pi} 2\pi i \left(\frac{e^{i(a-i\epsilon)x}}{2(a-i\epsilon)} + \frac{e^{i(-a-i\epsilon)x}}{2(-a-i\epsilon)} \right) = \\ &= \frac{i}{2a} (e^{iax} - e^{-iax}) = -\frac{\text{sen } ax}{a} \end{aligned} \quad (7.285)$$

y la solución es:

$$\mathcal{E}_2 = -\theta(-x) \frac{\text{sen } ax}{a} \quad (7.286)$$

Las soluciones fundamentales deben diferir en una solución de la ecuación homogénea

$$\mathcal{E}_1 - \mathcal{E}_2 = (\theta(x) + \theta(-x)) \frac{\text{sen } ax}{a} = \frac{\text{sen } ax}{a} \quad (7.287)$$

Otra posible solución es desplazar un polo al semiplano superior y otro al inferior. Ahora las dos integrales (para $x > 0$ y $x < 0$) son distintas de 0:

$$-\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{ikx}}{k^2 - a^2} dk = -i \lim_{\epsilon \rightarrow 0} \frac{e^{i(a+i\epsilon)x}}{2(a+i\epsilon)} = -\frac{i}{2a} e^{iax} \quad (7.288)$$

$$-\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{ikx}}{k^2 - a^2} dk = i \lim_{\epsilon \rightarrow 0} \frac{e^{i(-a-i\epsilon)x}}{2(-a-i\epsilon)} = -\frac{i}{2a} e^{-iax} \quad (7.289)$$

por tanto

$$\mathcal{E}_3 = -\frac{i}{2a} \left(\theta(x) e^{iax} + \theta(-x) e^{-iax} \right) \quad (7.290)$$

Y, por último, podemos desplazar los polos en sentido contrario.

$$-\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{ikx}}{k^2 - a^2} dk = -i \lim_{\epsilon \rightarrow 0} \frac{e^{i(-a+i\epsilon)x}}{2(-a+i\epsilon)} = \frac{i}{2a} e^{-iax} \quad (7.291)$$

$$-\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{ikx}}{k^2 - a^2} dk = i \lim_{\epsilon \rightarrow 0} \frac{e^{i(a-i\epsilon)x}}{2(a-i\epsilon)} = \frac{i}{2a} e^{iax} \quad (7.292)$$

por tanto

$$\mathcal{E}_4 = \frac{i}{2a} \left(\theta(x) e^{-iax} + \theta(-x) e^{iax} \right) \quad (7.293)$$

7.7.3. Soluciones fundamentales de las ecuaciones del calor y ondas

Queremos hallar ahora la solución del problema ($c > 0$):

$$(\partial_t - c\Delta)\mathcal{E} = \delta(x, t) \quad (7.294)$$

Para ello, aplicamos la transformada de Fourier a las coordenadas espaciales

$$\partial_t \hat{\mathcal{E}} + c|k|^2 \hat{\mathcal{E}} = \frac{1}{(2\pi)^{n/2}} \delta(t) \quad (7.295)$$

con lo que obtenemos una ecuación diferencial ordinaria en la variable t . Su solución es inmediata de lo que hemos visto antes.

$$\hat{\mathcal{E}} = \frac{1}{(2\pi)^{n/2}} \theta(t) e^{-ct|k|^2} \quad (7.296)$$

Haciendo la transformación inversa

$$\mathcal{E}(x, t) = \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} \theta(t) e^{-ct|k|^2} e^{ik \cdot x} dk \quad (7.297)$$

Como siempre, completando cuadrados

$$|k|^2 - ik \cdot \frac{x}{ct} = \left| k - \frac{ix}{2ct} \right|^2 + \frac{|x|^2}{4c^2t^2} \quad (7.298)$$

$$\mathcal{E}(x, t) = \frac{\theta(t)}{(4\pi ct)^{n/2}} e^{-|x|^2/4ct} \quad (7.299)$$

Escribamos la ecuación de ondas en $3 + 1$ dimensiones:

$$(\partial_t^2 - c^2 \Delta)\mathcal{E} = \delta(x, t) \quad (7.300)$$

y hagamos una transformación de Fourier en las variables espaciales

$$\partial_t^2 \hat{\mathcal{E}} + c^2 |k|^2 \hat{\mathcal{E}} = \frac{1}{(2\pi)^{3/2}} \delta(t) \quad (7.301)$$

Ya conocemos la solución de este problema

$$\hat{\mathcal{E}} = \frac{\theta(t)}{(2\pi)^{3/2}} \frac{\text{sen } ct|k|}{c|k|} \quad (7.302)$$

La transformada inversa de esta distribución es:

$$\mathcal{E}(x, t) = \frac{\theta(t)}{(2\pi)^3 c} \int_{\mathbf{R}^3} \frac{\text{sen } ct|k|}{|k|} e^{ik \cdot x} dk \quad (7.303)$$

En vez de calcular esta transformada inversa, vamos a calcular la directa de δ_S . Recordemos que, para una superficie S ,

$$\delta_S(\varphi) = \int_S \varphi(x) d\sigma \quad (7.304)$$

Supongamos ahora que S es la superficie esférica de radio R . La distribución δ_{S_R} es de soporte compacto (la superficie S_R). Para estas distribuciones de soporte compacto la transformada de Fourier pues calcularse como sigue.

$$\mathcal{F}(u)(k) = \frac{1}{(2\pi)^{n/2}} u(\phi(x)) e^{-ik \cdot x} \quad (7.305)$$

donde $\phi(x)$ es una función en \mathcal{D} que vale 1 sobre el soporte de u . Entonces, en nuestro caso

$$\begin{aligned}
 \mathcal{F}(\delta_{S_R})(k) &= \frac{1}{(2\pi)^{3/2}} \delta_{S_R}(\phi(x)e^{-ik \cdot x}) = \frac{1}{(2\pi)^{3/2}} \int_{S_R} \phi(x)e^{-ik \cdot x} d\sigma \\
 &= \frac{R^2}{(2\pi)^{3/2}} \int_{S_1} e^{-iRk \cdot x} d\sigma = \frac{2\pi R^2}{(2\pi)^{3/2}} \int_0^\pi e^{-iR|k| \cos \theta} \text{sen } \theta d\theta \\
 &= \frac{R^2}{(2\pi)^{1/2}} \int_0^\pi e^{-iR|k| \cos \theta} \text{sen } \theta d\theta = \frac{2R}{(2\pi)^{1/2}} \frac{\text{sen } R|k|}{|k|}
 \end{aligned} \tag{7.306}$$

Esta es justamente la función que hemos encontrado antes. Por tanto

$$\mathcal{E}_3(x, t) = \frac{\theta(t)}{4\pi c^2 t} \delta_{S_{at}} = \frac{\theta(t)}{2\pi c} \delta(c^2 t^2 - |x|^2) \tag{7.307}$$

Las soluciones fundamentales para las dimensiones espaciales iguales a 2 o a 1 se obtienen de forma similar o aplicando un método de descenso. Los resultados son:

$$\mathcal{E}_2(x, t) = \frac{\theta(at - |x|)}{2\pi c \sqrt{c^2 t^2 - |x|^2}} \tag{7.308}$$

$$\mathcal{E}_1(x, t) = \frac{1}{2c} \theta(at - |x|) \tag{7.309}$$

Bibliografía

- [1] L. Abellanas, A. Galindo, *Espacios de Hilbert (Geometría, Operadores, Espectros)*, Eudema, Madrid 1987.
- [2] L. Abellanas, L. Martínez Alonso, *Métodos Matemáticos de la Física: Problemas y Tests*, Tecnigraf, Talavera 1995.
- [3] C.D. Aliprantis, O. Burkinshaw, *Principles of real analysis*, Elsevier, New York 1981.
- [4] C.D. Aliprantis, K.C. Border, *Infinite dimensional analysis*, Springer, Berlin 1999.
- [5] S.K. Berberian, *Introducción al espacio de Hilbert*, Teide, Barcelona 1970.
- [6] S.K. Berberian, *Lectures in Functional Analysis and Operator Theory*, Springer, New York 1974.
- [7] N. Boccara, *Functional Analysis: An introduction for physicists*, Academic Press, San Diego 1990.
- [8] N. Bourbaki, *Éléments d'histoire des mathématiques*, Hermann, Paris 1969.
- [9] M. Carter, B. van Brunt, *The Lebesgue-Stieltjes Integral*, Springer, New York 2000.
- [10] B. Gelbaum, *Problems in Analysis*, Springer, New York 1982.
- [11] P.R. Halmos, *Measure Theory*, Van Nostrand, New York 1950.
- [12] P.R. Halmos, *Introduction to Hilbert space and the theory of spectral multiplicity*, Chelsea Publ. Co., New York 1951.
- [13] P.R. Halmos, *A Hilbert space problem book*, Van Nostrand, New York 1967.
- [14] T. Hawkins, *Lebesgue's Theory of Integration*, Chelsea, New York 1975.
- [15] J.M. Jauch, *Foundations of Quantum Mechanics*, Addison-Wesley, Reading 1968.
- [16] T.F. Jordan, *Linear Operators for Quantum Mechanics*, Wiley, New York 1969.
- [17] J.L. Kelley, *Topología general*, Eudeba, Buenos Aires 1975.
- [18] A.N. Kolmogorov, S.V. Fomin, *Introductory real analysis*, Dover, New York 1970.
- [19] E. Kreyszig, *Introductory Functional Analysis with Applications*, Wiley, New York 1978.
- [20] M.J. Lighthill, *Fourier analysis and generalized functions*, Cambridge Univ. Press, Cambridge 1970.
- [21] L. Lusternik, V. Sobolev, *Précis d'analyse fonctionnelle*, Mir, Moscou 1989.
- [22] C.C. Pugh, *Real Mathematical Analysis*, Springer, New York 2002.
- [23] M. Reed, B. Simon, *Methods of Modern Mathematical Physics. I: Functional Analysis*, Academic Press, New York 1972.
- [24] F. Riesz, B. Sz.-Nagy, *Functional Analysis*, Dover, Mineola 1955.

- [25] W. Rudin, *Functional Analysis*, Tata MacGraw-Hill, New Delhi 1973.
- [26] H.H. Schaefer, *Espacios vectoriales topológicos*, Teide, Barcelona 1974.
- [27] H.H. Schaefer, *Operator methods in Quantum Mechanics*, Dover, New York, 2002.
- [28] L. Schwartz, *Métodos Matemáticos para las Ciencias Físicas*, Selecciones Científicas, Madrid 1969.
- [29] L. Schwartz, *Analyse: Topologie générale et analyse fonctionnelle*, Hermann, Paris 1970.
- [30] L. Schwartz, *Théorie des distributions*, Hermann, Paris 1973.
- [31] V. Trénoquine, B. Pissarevski, T. Soboléva, *Problèmes et exercices d'analyse fonctionnelle*, Mir, Moscou 1987.
- [32] V.S. Vladimirov, *Generalized Functions in Mathematical Physics*, Mir, Moscow 1979.
- [33] V.S. Vladimirov, *Equations of Mathematical Physics*, Mir, Moscow 1984.