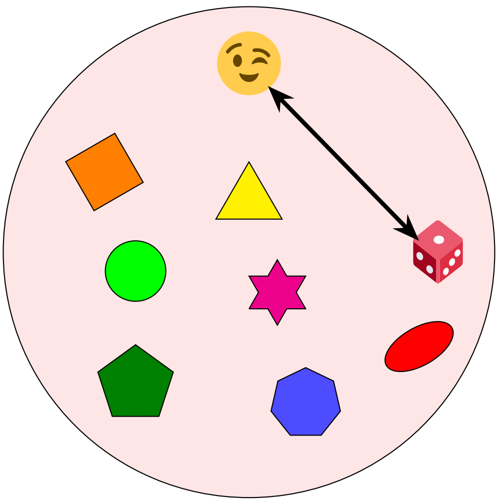
Introducción al Análisis Funcional

Prof. Renato Álvarez Nodarse

www: https://renato.ryn-fismat.es



$$\boxed{\|\alpha_1 x_1 + \cdots + \alpha_n x_n\| \geq c(|\alpha_1| + \cdots + |\alpha_n|)}$$

Versión 3.6 01/06/2026

Introducción al Análisis Funcional

Prof. Renato Álvarez-Nodarse

Índice general

Introducción al Análisis Funcional

Prof. Renato Álvarez-Nodarse

Prefacio

> Como la lectura, la clase es una obra en colaboración y quienes escuchan no son menos importantes que el que habla. En este libro está mi parte personal de aquellas sesiones. Espero que el lector las enriquezca, como las enriquecieron los oyentes.
>
> Jorge Luis Borges
>
> De *"El libro", Borges Oral* (1979)

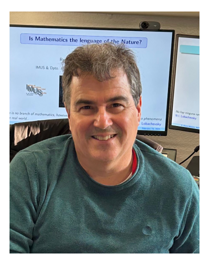
Estas notas contienen el contenido del curso *Análisis Funcional* impartido por el autor en la Facultad de Matemáticas de la Universidad de Sevilla. Dicho curso es una asignatura opcional del tercer curso del Grado en Matemáticas. Las mismas están divididas en ocho capítulos. El primero es un resumen de algunos de los conceptos del análisis matemático elemental que, por diversas razones, no siempre se dan de forma satisfactoria. El segundo es, de alguna forma, una motivación y está inspirada tras la lectura del primer capítulo de la magnífica monografía de Schechter [22].

En los capítulos 3, 4 y 5 se introducen los espacios métricos, normados (Banach) y euclídeos (Hilbert), así como se discuten algunas de sus principales propiedades. Son sin duda la base para entender los subsiguientes capítulos 6, 7 y 8. Cada uno de los capítulos 3, 4 y 5 se culmina enunciando y probando un teorema clásico de especial importancia en el contexto del mismo. Así en el el capítulo 3, dedicado a los espacios métricos, se culmina con el *Teorema de categorías de Baire*, en el 4, con el clásico *Teorema de Banach-Steinhaus* para los espacios de Banach y en el 5, con el *Teorema de representación de F. Riesz*. El capítulo 6 está dedicado a la teoría de operadores y culmina con la prueba del *Teorema espectral* para operadores compactos, cuando además son autoadjuntos o normales. El capítulo

7 contiene tres de los teoremas clásicos del Análisis funcional: el *Teorema de Hahn-Banach* sobre la extensión de funcionales lineales, el *Teorema de la aplicación abierta* y el *Teorema del grafo cerrado*. También en dicho capítulo 7 se discuten brevemente los conceptos de espacios duales y la convergencia débil. Finalmente, en el capítulo 8 se muestra, brevemente, la conexión de la teoría de operadores en espacios de Hilbert con la mecánica cuántica. Cada capítulo culmina con una colección de problemas (muchos de los cuales están resueltos con todo detalle) para que el lector interesado pueda ejercitar los conceptos y teoremas estudiados en ellos.

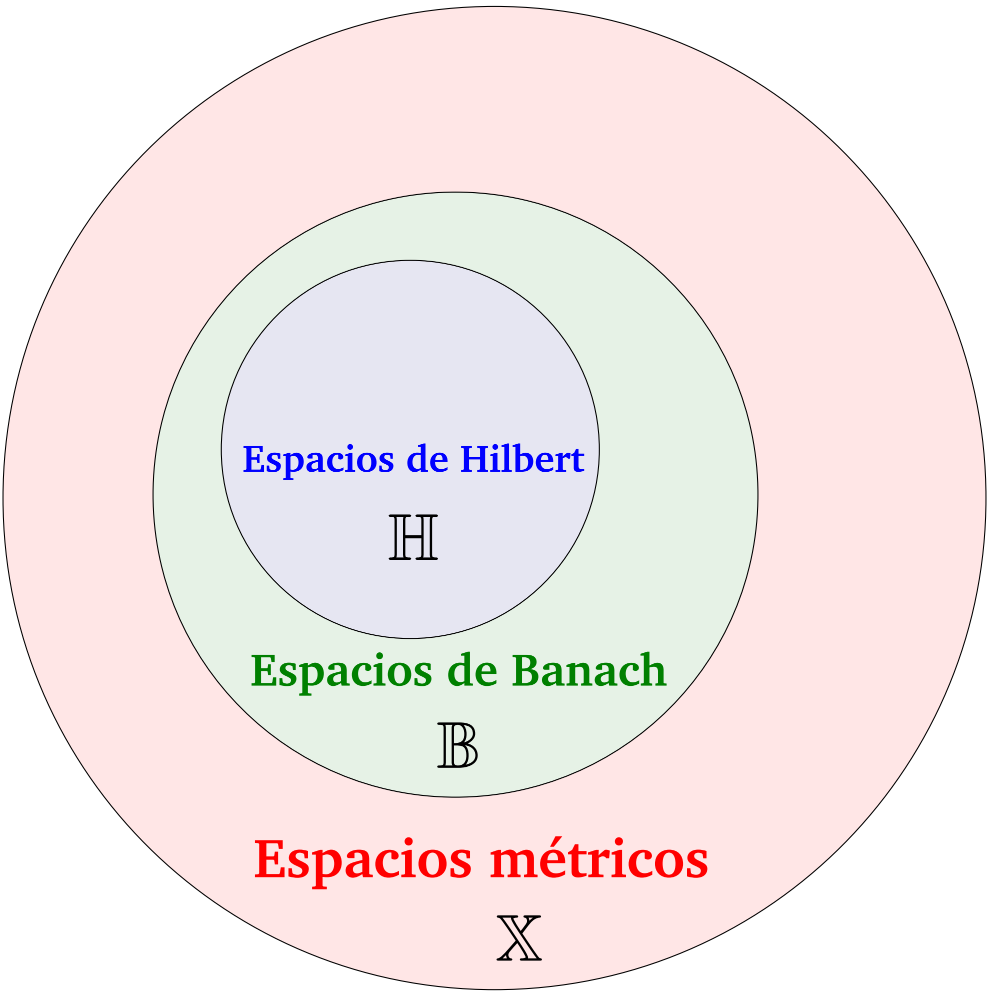
Quiero agradecer a todos los que de una forma u otra me han ayudado a que estas notas sean posibles. En primer lugar a mi familia, a los que les he robado un tiempo precioso. Sin duda también ocupan un lugar destacado los alumnos que participaron activamente en los tres primeros cursos durante el invierno/primavera de 2024, 2025 y 2026. En especial a Manuel Alanís, Felipe Garrido, Adán González, Mario Herrera, Antonio Medinilla, Jesús Mesa, David Ramos, Mateo Reino y Fernando Ruíz por sus muchas puntualizaciones, revisiones y críticas (siempre constructivas). Sin duda la presentación actual de estas notas son mucho mejor gracias a ellos y todos los errores responsabilidad exclusivamente mía. También soy deudor de los autores de las monografías mencionadas en la bibliografía. Aunque estas notas no coinciden en su totalidad con ninguna de ellas (si ese fuese el caso no hubiese sido necesario escribir este texto, hubiese bastado usar una de dichas monografías), de cada una de ellas he tomado prestado, entre otras cosas, demostraciones, ejemplos y problemas. Para ayudar al lector a la hora de consultar dichos textos en el capítulo dedicado a la bibliografía se menciona que texto se ha usado en cada uno de los grandes bloques (espacios métricos, normados y euclídeos), y, por tanto, muy recomendables para profundizar en los mismos.

Espero que el estudiante, y el lector en general, le encuentre utilidad a estas notas y disfrute de su lectura como yo he disfrutado al escribirlas.



Renato Álvarez Nodarse

Sevilla, 1 de junio de 2026



$$\boxed{\mathbb{H} \subseteq \mathbb{B} \subseteq \mathbb{X}}$$

Jerarquía de los espacios funcionales que se estudian en estas notas

Introducción al Análisis Funcional

Prof. Renato Álvarez-Nodarse

Capítulo 1

Preliminares

Estudia el pasado si quieres conocer el futuro

Atribuida a Confucio

En este capítulo recordaremos algunos conceptos básicos del análisis necesarios para entender el resto de los capítulos de estas notas. Asumiremos que el lector tiene unos conocimientos mínimos del cálculo infinitesimal, el álgebra lineal y los números complejos. Para más detalles consultar, por ejemplo, [24, 28, 29].

1.1. Sucesiones numéricas en $\mathbb{C}$

En adelante denotaremos por $\mathbb{N} = \{1, 2, 3, \ldots, n, n+1, \ldots\}$ al conjunto de los números naturales.

Definición 1.1.1 *La aplicación o función $f : \mathbb{N} \mapsto A$ que hace corresponder a cada número natural n un elemento a_n de cierto conjunto A se denomina sucesión de elementos de A y la denotaremos por $(a_n)_{n=1}^{\infty}$, o $(a_n)_n$.*

Por ejemplo, si $A = \mathbb{C}$, diremos que $f : \mathbb{N} \mapsto \mathbb{C}$ define una sucesión de números complejos. Como ejemplo tenemos las

$$a_n = (-1)^n = \{-1, 1, -1, 1, \ldots\}, \quad z_n = e^{in} = \{e^i, e^{2i}, e^{3i}, \ldots\}.$$

Definición 1.1.2 *Sea $\mathcal{N}$ un subconjunto de $\mathbb{N}$, $\mathcal{N} = \{n_1, n_2, ..., n_k, ..., \ \ k \in \mathbb{N}\}$ siendo $(n_k)_k$ una sucesión estrictamente creciente de números naturales. Sea $(a_n)_n$ una sucesión de números cualquiera. Construyamos una nueva sucesión cuyos elementos sean los elementos de a_n correspondientes a los valores n_k de $\mathcal{N}$. Es decir, construyamos un subconjunto del conjunto $\{a_n, \ n \in \mathbb{N}\}$. La nueva sucesión así obtenida la denotaremos $(a_{n_k})_k$ y la llamaremos subsucesión de $(a_n)_n$.*

Por ejemplo, sea $a_n = (-1)^n$. De la sucesión a_n podemos construir las subsucesiones de los elementos pares a_{2k} e impares a_{2k-1} correspondientes a los subconjuntos $\mathcal{N} = \{2, 4, ..., 2k, ..., \ \ k \in \mathbb{N}\}$ y $\mathcal{N} = \{1, 3, ..., 2k-1, ..., \ \ k \in \mathbb{N}\}$, respectivamente. Ellas son $a_{2k} = 1$ y $a_{2k-1} = -1$.

Definición 1.1.3 *Se dice que una sucesión $(z_n)_n$ está acotada, si para todo $n \in \mathbb{N}$, existe un $M \in \mathbb{R}$, $M > 0$ tal que $|z_n| \leq M$.*

Por ejemplo, las sucesiones (x_n) y $(z_n)_n$ definidas antes son claramente acotadas (basta escoger $M = 1$).

Definición 1.1.4 *Se dice que una sucesión $(z_n)_n$ es no acotada si para todo $M \in \mathbb{R}$, $M > 0$, existe un $n \in \mathbb{N}$ tal que $|z_n| > M$.*

Por ejemplo, la sucesión $b_n = (-1)^n n^2$ no está acotada.

Definiremos la distancia $d(z_1, z_2)$ entre dos números complejos $z_1 = x_1 + iy_1$ y $z_2 = x_2 + iy_2$ como

$$d(z_1, z_2) = |z_1 - z_2| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}. \tag{1.1.1}$$

Definiremos el ε-entorno o ε-vecindad de un número complejo z_0 a la “bola” $U_\varepsilon(z_0)$ definida por

$$U_\varepsilon(z_0) = \{z \in \mathbb{C}; \ |z - z_0| < \varepsilon\}.$$

Obviamente $U_\varepsilon(z_0)$ es un círculo del plano complejo de centro z_0 y radio ε excluyendo la frontera (es decir, la correspondiente circunferencia).

Sea $z = x + iy$, $z_0 = x_0 + iy_0$. Puesto que

$$\max\{|x-x_0|, |y-y_0|\} \leq |z-z_0| \leq |x-x_0|+|y-y_0|, \qquad |z+z_0| \leq |z|+|z_0|,$$

podemos construir la teoría de límites en $\mathbb{C}$ de la misma forma que se hace en $\mathbb{R}$. Así pues, tenemos la siguiente definición:

Definición 1.1.5 *Diremos una sucesión de números complejos $(z_n)_n$ tiene límite z, y lo denotaremos por $\lim_{n\to\infty} z_n = z$, si*[a]

$$\forall \varepsilon > 0,\ \varepsilon \in \mathbb{R}, \quad \exists N \in \mathbb{N};\ \forall n > N, \quad |z_n - z| < \varepsilon.$$

[a]En adelante al escribir $a > 0$ asumiremos que a es un número real, ya que los complejos no se pueden ordenar.

El significado geométrico del límite es claro: dentro de ε-entorno $U_\varepsilon(z)$ del límite z de la sucesión $(z_n)_n$ hay infinitos términos de la sucesión (de hecho todos a partir de cierto N) y fuera de él sólo hay un número finito de términos (a lo sumo los N primeros términos) de la misma.

Si escribimos $z_n = x_n + iy_n$ y $z = x + iy$, entonces, de las desigualdades

$$\max\{|x_n - x|, |y_n - y|\} \leq |z_n - z| \leq |x_n - x| + |y_n - y|,$$

se deduce que

$$\lim_{n\to\infty} z_n = z \quad \Longleftrightarrow \quad \lim_{n\to\infty} x_n = x, \quad \text{y} \quad \lim_{n\to\infty} y_n = y.$$

En particular, $z_n \stackrel{n\to\infty}{\longrightarrow} 0$ si y solo si $|z_n| \stackrel{n\to\infty}{\longrightarrow} 0$.

Nota 1.1.6 *De lo anterior se deduce que toda sucesión de números complejos es equivalente a dos sucesiones de números reales (la de sus partes reales e imaginarias).*

Así pues, por analogía con el caso real podemos definir las sucesiones de Cauchy, enunciar y probar el criterio de Cauchy y muchas otras propiedades de las las sucesiones[1]. No obstante al ser $\mathbb{C}$ un cuerpo **no** ordenado, se pierden todas las propiedades relacionadas con el orden (supremo, monotonía, etc). Gracias a la teoría de límites de sucesiones podemos definir el límite de funciones, continuidad de funciones, derivabilidad de funciones, etc.

Además, está claro que, al igual que en el caso real, si $(z_n)_n$ tiene límite, entonces $(z_n)_n$ es acotada, pero no al contrario.

[1]El lector no familiarizado con dichas definiciones puede consultar, por ejemplo, el capítulo 3 de la magnífica monografía [28]

Definición 1.1.7 *Una sucesión $(z_n)_n$ es de Cauchy (o fundamental) si para todo $\varepsilon > 0$ existe un $N \in \mathbb{N}$ tal que para todos $n, m > N$, se cumple $|z_n - z_m| < \varepsilon$. O, equivalentemente, z_n es de Cauchy si y solo si*

$$\forall\varepsilon > 0,\ \exists N \in \mathbb{N},\ \text{tal que}\ \forall n > N,\ \text{y}\ \forall p \in \mathbb{N}, \quad |z_{n+p} - z_n| < \varepsilon,$$

o, equivalentemente,

$$\forall\varepsilon > 0,\ \exists N \in \mathbb{N},\ \text{tal que}\ \forall n > m > N, \quad |z_n - z_m| < \varepsilon.$$

Teorema 1.1.8 (Criterio de Cauchy) *Para que una sucesión de números complejos $(z_n)_n$ sea de convergente es necesario y suficiente que sea de Cauchy.*

Demostración: La prueba se basa en la equivalencia entre $z_n = x_n + iy_n$ y las sucesiones reales $(x_n)_n$ y $(y_n)_n$. En efecto, basta notar que

$$\max\{|x_n - x_{n+p}|, |y_n - y_{n+p}|\} \le |z_n - z_{n+p}| \le |x_n - x_{n+p}| + |y_n - y_{n+p}|,$$

luego está claro que si z_n es de Cauchy, x_n e y_n también lo son, luego $x_n \overset{n\to\infty}{\longrightarrow} x$ y $y_n \overset{n\to\infty}{\longrightarrow} y$ pues $(x_n)_n$ e $(y_n)_n$ son ambas sucesiones reales para las cuales es cierto el criterio de Cauchy. Luego $z_n \overset{n\to\infty}{\longrightarrow} z = x + iy$. ■

De la equivalencia entre las sucesiones reales y complejas (ver Nota 1.1.6) se sigue inmediatamente el siguiente teorema:

Teorema 1.1.9 (Propiedades algebraicas de los límites) *Sean dos sucesiones convergentes $(a_n)_n$ y $(b_n)_n$ con $\lim\limits_{n\to\infty} a_n = a$, y $\lim\limits_{n\to\infty} b_n = b$. Entonces:*

1. $\lim\limits_{n\to\infty} a_n + b_n = a + b$.
2. $\lim\limits_{n\to\infty} a_n \cdot b_n = a \cdot b$. *En particular, para todo $\alpha \in \mathbb{C}$, $\lim\limits_{n\to\infty} \alpha\, a_n = \alpha\, a$.*
3. *Si $n \in \mathbb{N}$, $b_n \neq 0$, y $b \neq 0$, entonces, $\lim\limits_{n\to\infty} \dfrac{a_n}{b_n} = \dfrac{a}{b}$.*

Definición 1.1.10 *Un punto z de un conjunto $E \subset \mathbb{C}$ es interior si existe* *un entorno $U_\varepsilon(z)$ del mismo contenido por completo en E. Un conjunto es abierto si todos sus puntos son interiores.*

Evidentemente todo entorno $U_\varepsilon(z)$ de z es un abierto.

Definición 1.1.11 *Dado un conjunto $E \subset \mathbb{C}$ diremos que $z_0 \in \mathbb{C}$ es un punto de acumulación de E si en cualquier entorno $U_\varepsilon(z_0)$ existe al menos un punto de E distinto de z_0 (y, por tanto, infinitos).*

Por ejemplo, 0 es un punto de acumulación del conjunto $E = \{1, 1/2, 1/3, \dots, 1/n, \dots\}$, mientras que 10^{-k}, cualquiera sea $k \in \mathbb{N}$ no lo es.

De la Definición 1.1.11 se sigue inmediatamente el siguiente

Teorema 1.1.12 *z_0 es un punto de acumulación de E si y solo si existe al menos una sucesión de números $(z_n)_n$, con $z_n \neq z_{n+1}$, para todo $n \in \mathbb{N}$, tal que $\lim_{n\to\infty} z_n = z_0$.*

Definición 1.1.13 *Un conjunto E es cerrado si contiene todos sus puntos de acumulación.*

Por ejemplo, el conjunto $\overline{U}_\varepsilon(z_0) =: \{z \in \mathbb{C};\ |z - z_0| \leq \varepsilon\}$ es un cerrado.

Definición 1.1.14 *Un conjunto E es acotado si existe un $M > 0$ tal que, para todo $z \in E$, $|z| < M$.*

A partir del Lema de Bolzano-Weierstrass para los conjuntos acotados de $\mathbb{R}$ se sigue el siguiente resultado:

Lema 1.1.15 (Bolzano-Weierstrass) *Cualquier subconjunto infinito[a] acotado de $\mathbb{C}$ tiene por lo menos un punto de acumulación. En particular se tiene que de toda sucesión acotada se puede extraer una subsucesión convergente.*

[a]Es decir, con un número infinito de elementos.

1.2. Series

1.2.1. Series numéricas

Definición 1.2.1 *Dada una sucesión de números complejos $(a_n)_n$, la expresión*

$$\sum_{k=1}^{\infty} a_k = a_1 + a_2 + a_3 + \cdots + a_k + \cdots$$

se denomina serie infinita o serie y a los números a_1, a_2, ..., elementos de la serie. Las sumas

$$S_n = \sum_{k=1}^{n} a_k = a_1 + a_2 + a_3 + \cdots + a_n,$$

se denominan sumas parciales de la serie.

Definición 1.2.2 *Diremos que la serie $\sum_{k=1}^{\infty} a_k$ converge a S, si la sucesión de sumas parciales $(S_n)_n$ tiene límite finito S y a dicho número le denominaremos "suma" de la serie. Si por el contrario, la sucesión de sumas parciales no tiene límite, entonces diremos que la serie diverge.*

El Criterio de Cauchy para las sucesiones establece que una sucesión $(x_n)_n$ es convergente si y solo si $\forall \varepsilon > 0, \exists N \in \mathbb{N},$ tal que si $n > N,$ y $\forall p \in \mathbb{N}$ entonces $|x_{n+p} - x_n| < \varepsilon$. Una simple aplicación del mismo a la sucesión de sumas parciales nos conduce a nuestro primer criterio de convergencia para las series.

Teorema 1.2.3 (Criterio de Cauchy para las series) *La serie $\sum_{k=1}^{\infty} a_k$ es convergente si y solo si , para todo $\varepsilon > 0$, existe un $N \in \mathbb{N}$, tal que para todos $n > N$ y para todo $p \in \mathbb{N}$, $|a_{n+1} + \cdots + a_{n+p}| < \varepsilon$*

Este teorema tiene dos consecuencias inmediatas: la primera es que si alteramos un número finito de elementos de la serie, la convergencia de ésta no se afecta; y la segunda es el siguiente resultado:

Teorema 1.2.4 (Condición necesaria de convergencia de una serie) *Si la serie $\sum_{k=1}^{\infty} a_k$ es convergente, entonces $\lim\limits_{n\to\infty} a_n = 0$.*

Demostración: Basta tomar $p = 1$ en el Teorema 1.2.3. ■

Es *fácil* ver, a partir del teorema anterior que la serie $\sum_{n=1}^{\infty} \frac{n^2}{300n+n^2}$ es divergente. Con algo de más trabajo se puede probar que la serie $\sum_{n=1}^{\infty} \frac{1}{n}$, es divergente, lo cual ilustra que las condiciones del Teorema 1.2.4 son necesarias pero no suficientes.

Una propiedad de las series convergentes es que si $\sum_{k=1}^{\infty} a_k$ y $\sum_{k=1}^{\infty} b_k$ convergen, entonces la serie $\sum_{k=1}^{\infty} \alpha a_k + \beta b_k$ también es convergente, además

$$\sum_{k=1}^{\infty} \alpha a_k + \beta b_k = \alpha \sum_{k=1}^{\infty} a_k + \beta \sum_{k=1}^{\infty} b_k, \qquad \forall \alpha,\, \beta \in \mathbb{C}.$$

Nótese, además, que del carácter convergente de la serie $\sum_{k=1}^{\infty}(a_k + b_k)$ no se puede inferir el carácter convergente o divergente de las series $\sum_{k=1}^{\infty} a_k$ y $\sum_{k=1}^{\infty} b_k$ (¿por qué?).

Definición 1.2.5 *Diremos que una serie $\sum_{k=1}^{\infty} a_k$ es absolutamente convergente si la serie $\sum_{k=1}^{\infty} |a_k|$ es convergente.*

Teorema 1.2.6 (Weierstrass) *Una serie $\sum_{k=1}^{\infty} a_k$ es absolutamente convergente si y solo si la sucesión $S_n = \sum_{k=1}^{n} |a_k|$ está acotada.*

Demostración: La sucesión $(S_n)_n$ de sumas parciales $S_n = \sum_{k=1}^{n} |a_k|$ es una sucesión monótona, luego si está acotada, entonces existirá el límite que es precisamente la suma de la serie. Por el contrario, si la sucesión $S_n = \sum_{k=1}^{n} |a_k|$ es no acotada, entonces la serie será divergente. ■

A partir del criterio de Cauchy 1.2.3 se sigue el siguiente teorema:

Teorema 1.2.7 *Toda serie absolutamente convergente es convergente.*

Un ejemplo sencillo, y de gran importancia, de serie absolutamente convergente es la serie geométrica

$$\sum_{k=1}^{\infty} q^{k-1} = 1 + q + q^2 + \cdots = \frac{1}{1-q}, \quad |q| < 1.$$

Recordemos a continuación algunos criterios para estudiar la convergencia de series. El primero, y uno de los más sencillos, es el criterio de comparación. Su demostración es una simple consecuencia del Teorema de las sucesiones monótonas y acotadas similar a la prueba del Teorema 1.2.6.

Teorema 1.2.8 (Criterio de Comparación de Weierstrass) *Sean las series de números complejos $\sum_{k=1}^{\infty} a_k$ y $\sum_{k=1}^{\infty} b_k$. Si existe un $N \in \mathbb{N}$ tal que para todo $n > N$, $|a_n| \leq |b_n|$, entonces si la serie $\sum_{k=1}^{\infty} |b_k|$ converge, la serie $\sum_{k=1}^{\infty} |a_k|$ converge, y si $\sum_{k=1}^{\infty} |a_k|$ diverge, entonces $\sum_{k=1}^{\infty} |b_k|$ diverge.*

Ejercicio 1.2.9 Prueba, calculando el límite de sus sumas parciales, que la serie $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$ converge. Usando lo anterior y el criterio de comparación deduce la convergencia de la serie $\sum_{n=1}^{\infty} \frac{1}{n^2}$.

Teorema 1.2.10 (Criterio de comparación por paso al límite) *Sean las sucesiones $(a_n)_n$ y $(b_n)_n$ tales que $\lim_{n\to\infty} \frac{|a_n|}{|b_n|} = q > 0$. Entonces, $\sum_{k=1}^{\infty} |a_k|$ converge si y solo si $\sum_{k=1}^{\infty} |b_k|$ converge.*

Demostración: Este teorema es una consecuencia del Teorema de comparación 1.2.8. Como $\lim_{n\to\infty} a_n/b_n = q > 0$, ello significa que para todo $\varepsilon > 0$, existe un $N \in \mathbb{N}$, tal que para todo $n > N$, $|a_n/b_n - q| < \varepsilon$. Luego, si tomamos $\varepsilon = q/2$, a partir de cierto N tendremos que

$$\frac{1}{2} q b_n < a_n < \frac{1}{3} q b_n \,,$$

de donde, utilizando el criterio de comparación, se obtiene el resultado buscado. ■

Ejercicio 1.2.11 Estudiar la convergencia de las series

$$\sum_{n=1}^{\infty} \frac{1}{n^2}, \quad \sum_{n=1}^{\infty} \sin^2\left(\frac{1}{n}\right) \quad \text{y} \quad \sum_{n=1}^{\infty} \log\left(1 + \frac{1}{n}\right).$$

Nota 1.2.12 *De la prueba del Criterio de comparación por paso al límite se deduce que si $q = 0$, solo podemos afirmar lo mismo que en el Teorema de comparación 1.2.8 es decir, que si la serie $\sum_{k=1}^{\infty} |b_k|$ converge, la serie $\sum_{k=1}^{\infty} |a_k|$ converge, y si $\sum_{k=1}^{\infty} |a_k|$ diverge, $\sum_{k=1}^{\infty} |b_k|$ diverge.*

Veamos ahora dos criterios donde se involucra una única serie, es decir, dos criterios donde solo precisamos la información relacionada con la serie que queremos estudiar.

Teorema 1.2.13 (Criterio del cociente de D'Alembert) *Sea* $\sum_{k=1}^{\infty} a_k$ *una serie de números complejos tal que* $\lim_{n\to\infty} |a_{n+1}|/|a_n| = q$. *Entonces,*

1. *Si* $q < 1$, *la serie* $\sum_{k=1}^{\infty} a_k$ *es absolutamente convergente*
2. *Si* $q > 1$, *la serie* $\sum_{k=1}^{\infty} |a_k|$ *es divergente.*
3. *Si* $q = 1$, *la serie* $\sum_{k=1}^{\infty} |a_k|$ *puede ser convergente o divergente.*

Demostración: Si $\lim_{n\to\infty} |a_{n+1}/a_n| = q < 1$, entonces a partir de cierto $N \in \mathbb{N}$, existe un $\alpha \in (0,1)$, tal que $|a_{n+1}/a_n| \leq \alpha < 1$. Por tanto,

$$|a_{n+1}| \leq \alpha |a_n|, \quad |a_{n+1}| \leq \alpha^2 |a_{n-1}| \leq \alpha^2 |a_{n-2}| \Rightarrow |a_n| \leq \alpha^{n-N} |a_N| = C\alpha^n,$$

y el criterio de comparación, tomando $b_n = C\alpha^n$, nos conduce al resultado buscado. Análogamente se demuestra la divergencia en el caso $q > 1$. Como ejemplo del caso 3 tenemos las series cuando $a_n = 1/n$ y $a_n = 1/n^2$, para las cuales obviamente $\lim_{n\to\infty} |a_{n+1}/a_n| = 1$, siendo la primera divergente y la segunda convergente. ■

Este criterio no siempre funciona aún en casos sencillos como el siguiente: sea la serie $\sum_{k=1}^{\infty} \frac{2+(-1)^k}{2^k}$. Obviamente tenemos

$$\frac{a_{n+1}}{a_n} = \frac{1}{2}\frac{2+(-1)^{n+1}}{2+(-1)^n} = \begin{cases} 1/6 & n \text{ par} \\ 3/2 & n \text{ impar} \end{cases},$$

luego no existe el límite del cociente. No obstante,

$$\sum_{k=1}^{\infty} \frac{2+(-1)^k}{2^k} = \sum_{k=1}^{\infty} \frac{2}{2^k} + \frac{(-1)^k}{2^k} = \sum_{k=1}^{\infty} \frac{1}{2^{k-1}} + \sum_{k=1}^{\infty} \left(-\frac{1}{2}\right)^k,$$

y estas dos últimas obviamente son convergentes.

Teorema 1.2.14 (Criterio de la raíz de Cauchy) *Sea* $\sum_{k=1}^{\infty} a_k$ *una serie de números complejos tal que* $\lim_{n\to\infty} \sqrt[n]{|a_n|} = q$. *Entonces,*

1. *Si* $q < 1$, *la serie* $\sum_{k=1}^{\infty} a_k$ *es absolutamente convergente*
2. *Si* $q > 1$, *la serie* $\sum_{k=1}^{\infty} a_k$ *es divergente.*
3. *Si* $q = 1$, *la serie* $\sum_{k=1}^{\infty} a_k$ *puede ser convergente o divergente.*

Demostración: La demostración de este resultado se basa en si $\lim_{n\to\infty} \sqrt[n]{a_n} = q < 1$, entonces a partir de cierto $N \in \mathbb{N}$, $\sqrt[n]{|a_n|} \leq \alpha < 1$. Luego, si utilizamos el criterio de comparación tomando $b_n = \alpha^n$ el resultado es inmediato. Nótese que si $q = 1$, entonces pueden ocurrir ambas posibilidades. El mismo ejemplo, $a_n = 1/n$ y $a_n = 1/n^2$, teniendo en ambos casos $\lim_{n\to\infty} \sqrt[n]{a_n} = 1$, pero siendo la primera serie divergente y la segunda convergente. ∎

Si aplicamos este criterio a la serie $\sum_{k=1}^{\infty} \frac{2+(-1)^k}{2^k}$ tenemos

$$\frac{1}{2} = \sqrt[n]{\frac{-1+2}{2^n}} \leq \sqrt[n]{\frac{(-1)^n+2}{2^n}} \leq \sqrt[n]{\frac{1+2}{2^n}} = \frac{\sqrt[n]{3}}{2},$$

por tanto $\lim_{n\to\infty} \sqrt[n]{a_n} = \frac{1}{2} < 1$ y según el criterio de Cauchy 1.2.14 converge.

El ejemplo anterior nos indica que el criterio de Cauchy 1.2.14 es más general que el de D'Alembert 1.2.13 en el sentido que el criterio de criterio de D'Alembert implica el de Cauchy, pero no viceversa.

Nota 1.2.15 *Debemos destacar que tanto en el criterio de D'Alembert como el de Cauchy antes descritos, el resultado sigue siendo cierto si cambiamos los límites por los límites superiores.*

Ejercicio 1.2.16 Estudia el carácter de las series

$$\sum_{k=0}^{\infty} \frac{x^k}{k!}, \quad \sum_{k=1}^{\infty} \frac{(\sqrt{k})^k}{k!} \quad \text{y} \quad \sum_{k=0}^{\infty} \frac{k^2}{2^k}.$$

1.2.2. Más sobre series de números reales

Vamos a discutir ahora algunas propiedades de las series de números reales. Comenzaremos considerando serie de gran importancia en las matemáticas –y fuera de ellas–: las series del tipo $\sum_{n=1}^{\infty} \frac{1}{n^p}$ con $p \in \mathbb{R}$, $p > 0$. Estas series se conocen como series armónicas. Ya vimos ejemplos de ellas cuando $p = 1$, serie que divergía, y $p = 2$, serie que convergía. Una pregunta evidente es ¿Para cuáles $p > 0$ converge la serie armónica?

Es evidente del Teorema de comparación 1.2.8 que para todo $p > 2$ la serie converge y que para $p < 1$ diverge. Ahora bien, para $1 < p < 2$ no nos sirve el Teorema de comparación y un calculo sencillo nos muestra que tanto el criterio de Cauchy como el de D'Alembert nos dan 1 con lo cual no son concluyentes. ¿Cómo resolver este problema?

Teorema 1.2.17 (Criterio de McLaurin-Cauchy) *Sea $f(x)$ una función real no negativa definida en $[1, +\infty)$ y monótona decreciente en todo su dominio. Entonces la serie $\sum_{n=1}^{\infty} f(n)$ y la integral impropia $\int_1^{\infty} f(x)dx$ tienen el mismo carácter de convergencia.*

Demostración: Como f es decreciente y no negativa, tenemos que para todo $k = 0, 1, \dots$ y $x \in [k, k+1]$, $0 \le f(k+1) \le f(x) \le f(k)$, por tanto, usando la monotonía de la integral de Riemann –recuérdese que si f es monótona, entonces es integrable–, tenemos

$$f(k+1) \le \int_k^{k+1} f(x)dx \le f(k).$$

Luego, usando la aditividad de la integral,

$$\sum_{k=1}^{n} f(k+1) \le \int_1^{n+1} f(x)dx \le \sum_{k=1}^{n} f(k),$$

es decir, para las sumas parciales S_n de la serie $\sum_{n=1}^{\infty} f(n)$ y la función $F(t) := \int_1^t f(x)dx$, tenemos

$$S_{n+1} - f(1) \le F(n+1) \le S_n.$$

De la desigualdad anterior deducimos que si la serie $\sum_{n=1}^{\infty} f(n)$ converge, entonces sus sumas parciales S_n son acotadas y, por tanto, la función $F(t)$ es acotada, luego existe el límite $\lim_{t\to\infty} F(t)$, es decir, existe la integral impropia. Además, si existe la integral impropia, entonces la sucesión $F(n+1)$ es acotada, por tanto las sumas S_{n+1} también lo son, de donde deducimos que la serie converge (ver Teorema 1.2.6).

Si por el contrario la serie no converge, entonces las sumas parciales no están acotadas y, por tanto, F tampoco lo está, luego la integral impropia diverge. Análogamente, si la integral impropia no converge, entonces F es no acotada, luego S_n es no acotada y la serie diverge. ■

Ejercicio 1.2.18 Usa el criterio integral para estudiar el carácter de las series $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^p}$ y $\displaystyle\sum_{n=2}^{\infty} \frac{1}{n \log^p n}$ para $p > 0$.

Reordenación de términos en una serie

Dada una serie $\sum_{k=1}^{\infty} a_k$, construyamos una serie $\sum_{k=1}^{\infty} a'_k$ obtenida de la anterior mediante la reordenación de un número infinito de términos, es decir todos los términos de las sucesiones $(a_n)_n$ y $(a'_n)_n$ coinciden pero están, en general, situados en distintos sitios.

Teorema 1.2.19 *Si $\sum_{k=1}^{\infty} a_k$ es absolutamente convergente, entonces la serie $\sum_{k=1}^{\infty} a'_k$ también es absolutamente convergente y su suma coincide con la de la serie original.*

Demostración: Supondremos que la reordenación involucra a un número infinito de términos –si es un número finito sabemos que el resultado es cierto–. Denotemos por S la suma $S = \sum_{k=1}^{\infty} a_k$ y S' la suma $S' = \sum_{k=1}^{\infty} a'_k$.

Supongamos, que nuestra serie tiene todos sus términos positivos (o negativos). Sean $S'_n = \sum_{k=1}^{n} a'_k$ las sumas parciales de la serie reordenada. Escojamos ahora un $N \in \mathbb{N}$ tal que todos los elementos del conjunto $\{a'_1, a'_2, \dots a'_n\}$ estén contenidos en el conjunto $\{a_1, a_2, \dots a_N\}$. Entonces, obviamente $S'_n \leq S_N$. Pero la serie original es convergente, luego la serie S converge luego las S_n están acotadas y, por tanto, el conjunto de las sumas parciales de S' es acotado, luego S' converge. Además $S' \leq S$. Para probar que $S' = S$ procedemos al contrario, es decir, dado $S_n = \sum_{k=1}^{n} a_k$ escojamos un $N \in \mathbb{N}$ tal que todos los elementos del conjunto $\{a_1, a_2, \dots a_n\}$ estén contenidos en el conjunto $\{a'_1, a'_2, \dots a'_N\}$. Entonces, obviamente $S_n \leq S'_N$ y, por tanto, como hemos probado que S' converge tenemos $S \leq S'$ de donde deducimos que ambas series convergen y suman lo mismo.

Supongamos a continuación que la serie S tiene infinitos términos positivos e infinitos negativos –si alguno de ellos fuese un número finito el teorema estaría probado–. Entonces, si denotamos por a_n^+ y a_n^- a los términos positivos y negativos, respectivamente, tendríamos

$$\sum_{k=1}^{\infty} a_k = \sum_{k=1}^{\infty} a_k^+ + \sum_{k=1}^{\infty} a_k^-, \qquad \text{además} \qquad \sum_{k=1}^{\infty} |a_k| = \sum_{k=1}^{\infty} a_k^+ + \sum_{k=1}^{\infty} (-a_k^-),$$

de donde se deduce que ambas series $\sum_{k=1}^{\infty} a_k^+$ y $\sum_{k=1}^{\infty} a_k^-$ convergen. Sea, como antes, S' nuestra serie reordenada, entonces:

$$S' = \sum_{k=1}^{\infty} a'_k = \sum_{k=1}^{\infty} (a_k^+)' + \sum_{k=1}^{\infty} (a_k^-)' = \sum_{k=1}^{\infty} a_k^+ + \sum_{k=1}^{\infty} a_k^- = \sum_{k=1}^{\infty} a_k = S,$$

pues ambas series $\sum_{k=1}^{\infty} a_k^+$ y $\sum_{k=1}^{\infty} a_k^-$, convergen absolutamente y sus términos tienen todos el mismo signo. ∎

Este resultado solo es cierto para las series absolutamente convergentes. Veamos el siguiente ejemplo.

Sea la serie $S = \sum_{n=1}^{\infty} (-1)^{n+1}/n$. Esta serie es convergente como veremos a continuación, pero no lo hace absolutamente[2]. Además, como

$$\log(1+x) = \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k} x^k + \frac{(-1)^n}{1+\xi} \frac{x^{n+1}}{n+1}, \quad \text{con} \quad \xi \in (0,x),$$

tenemos, sustituyendo $x = 1$, que

$$\left| \log 2 - \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k} \right| \leq \frac{1}{n+1} \to 0, \quad \text{si} \quad n \to \infty,$$

es decir, nuestra serie S es, en efecto, convergente

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \cdots + \frac{1}{2k-1} - \frac{1}{2k} + \cdots = \log 2 \neq 0.$$

Reordenemos la serie

$$S = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \left(\frac{1}{5} - \frac{1}{6}\right) + \left(\frac{1}{7} - \frac{1}{8}\right) + \cdots + \left(\frac{1}{2k-1} - \frac{1}{2k}\right) + \cdots$$

de la siguiente forma:

$$\begin{aligned} S' &= \left(1 - \frac{1}{2} - \frac{1}{4}\right) + \left(\frac{1}{3} - \frac{1}{6} - \frac{1}{8}\right) + \left(\frac{1}{5} - \frac{1}{10} - \frac{1}{12}\right) + \cdots \\ &\qquad + \left(\frac{1}{2k-1} - \frac{1}{4k-2} - \frac{1}{4k}\right) + \cdots \\ &= \left(\frac{1}{2} - \frac{1}{4}\right) + \left(\frac{1}{6} - \frac{1}{8}\right) + \left(\frac{1}{10} - \frac{1}{12}\right) + \cdots + \left(\frac{1}{4k-2} - \frac{1}{4k}\right) + \cdots \\ &= \frac{1}{2}\left[\left(1 - \frac{1}{2}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \left(\frac{1}{5} - \frac{1}{6}\right) + \cdots + \left(\frac{1}{2k-1} - \frac{1}{2k}\right) + \cdots\right] \\ &= \frac{S}{2} = \frac{\log 2}{2}, \end{aligned}$$

[2]La serie $\sum_{n=1}^{\infty} 1/n$ diverge.

es decir ¡nuestra reordenación cambia el valor de la serie!

Definición 1.2.20 *Diremos que una serie es condicionalmente convergente si la serie* $\sum_{k=1}^{\infty} a_k$ *converge pero la serie* $\sum_{k=1}^{\infty} |a_k|$ *diverge.*

Teorema 1.2.21 (Teorema de reordenación de Riemann) *Las series condicionalmente convergentes se pueden reordenar de tal forma que sumen cualquier valor real prefijado de antemano.*

Demostración: Sea $S = \sum_{k=1}^{\infty} a_k$ nuestra serie. Obviamente S tiene infinitos términos positivos e infinitos negativos –pues si alguno de ellos fuese finito la serie seria absolutamente convergente–. Como antes denotaremos por a_n^+ y a_n^- a los términos positivos y negativos, respectivamente, y sean

$$S^+ = \sum_{k=1}^{\infty} a_k^+, \qquad \text{y} \qquad S^- = \sum_{k=1}^{\infty} a_k^-,$$

las correspondientes series de términos positivos y negativos y S_n^+ y S_n^- las sumas parciales de las mismas. Sean N^+ y N^- dos números naturales suficientemente grandes. Como

$$|S|_N := \sum_{k=1}^{N} |a_k| = \sum_{k=1}^{N^+} a_k^+ + \sum_{k=1}^{N^-}(-a_k^-) = S_{N^+} - S_{N^-},$$

entonces, para N^+ y N^- tendiendo a infinito se tiene que o S^+ tiende a $+\infty$ o S^- tiende a $-\infty$, pues si ambas convergieran, la serie S sería absolutamente convergente. Probemos que ambas series S^+ y S^- divergen. Supongamos que una de las dos series, digamos S^+, converge, entonces como

$$\sum_{k=1}^{N} a_k = \sum_{k=1}^{N^+} a_k^+ + \sum_{k=1}^{N^-} a_k^-,$$

y S^- es necesariamente divergente, S tendría que ser divergente, lo cual es una contradicción. Por tanto, y como ambas series S^+ y S^- son series de términos del mismo signo, tenemos

$$\lim_{n\to\infty} S_n^+ = +\infty, \qquad \text{y} \qquad \lim_{n\to\infty} S_n^- = -\infty.$$

Para probar el teorema vamos a fijar un número L cualquiera y vamos construir una reordenación de $(a_n)_n$ tal que la suma de nuestra serie S' sea L, es decir, $L = \sum_{k=1}^{\infty} a_k$. Por comodidad tomaremos $L > 0$

–el caso $L < 0$ es análogo–. Como $\lim S_n^+ = +\infty$, los n_1 primeros términos de nuestra nueva serie $\sum_{k=1}^{\infty} a'_k$ los escogemos, consecutivamente tomando los primeros términos positivos de la sucesión $(a_n)_n$ hasta que tengamos que $S_{n_1}^+ > L$. Así la suma parcial $S_{n_1}^+ = \sum_{k=1}^{n_1} a_k^+ > L$. Como $\lim S_n^- = -\infty$, vamos a seguir construyendo nuestra reordenación con los primeros m_1 términos negativos de la sucesión $(a_n)_n$ hasta que tengamos que $S_{n_1}^+ + S_{m_1}^- < L$. Y así sucesivamente construimos una sucesión de sumas parciales

$$S'_k = S_{n_1}^+ + S_{m_1}^- + S_{n_2}^+ + S_{m_2}^- + \cdots + S_{n_k}^+ + S_{m_k}^-,$$

tomando los términos positivos o negativos de $(a_n)_n$ de forma que en cada paso (o bloque de sumandos de un signo dado) estamos a la derecha o a la izquierda de L. Además, como nuestra serie $\sum_{k=1}^{\infty} a_k$ converge, entonces $a_n \to 0$ cuando $n \to \infty$, luego para todo $\varepsilon > 0$ siempre existe los valores $N = n_1 + \cdots + n_k$ y $M = m_1 + \cdots + m_k$ tales que si $n > K = N + M$, $a_n^+ < \varepsilon$ y $|a_n^-| < \varepsilon$ para todo $n > K$.

Probemos a continuación que la serie S', construida de esta forma, tiene límite L. Por construcción notemos que si escogemos $n > K$, siendo K el encontrado anteriormente, entonces $|S_n - L|$ se puede hacer tan pequeño como se quiera. En efecto, supongamos que n corresponde justo al último sumando, a'_k, de la suma S'_k, que no es más que el último sumando de $S_{n_k}^+$ o $S_{m_k}^-$, es decir, $a_{n_k}^+$ o $|a_{n_k}^-|$ entonces $|S_n - L| < |a'_k|$, y éste es en valor absoluto menor que ε como ya hemos visto. Es decir, si nos detenemos justo cuando en nuestro proceso saltamos por encima de L, entonces $|S_n - L| < \varepsilon$.

Si ahora $n > K$, que no sea el último sumando de $S_{n_k}^+$ o $S_{m_k}^-$, entonces como todos los a'_n son tales que $|a'_n| < \varepsilon$ si $n > K$, entonces vuelve a ser $|S_n - L| < |a'_k|$, siendo a'_k nuevamente el último sumando de la suma $S_{n_k}^+$ o $S_{n_k}^-$ anterior con lo que sigue siendo $|S_n - L| < \varepsilon$. Por tanto, existe un K suficientemente grande tal que la diferencia $|S_n - L|$ se puede hacer tan pequeña como se quiera para todo $n > K$, luego la serie S' tiene como suma a L. ∎

1.2.3. Series de potencias

Las definiciones y teoremas anteriores nos permiten definir y estudiar las series de potencias en $\mathbb{C}$.

Definición 1.2.22 *Dada una sucesión de números complejos $(a_n)_n$ y $z_0 \in \mathbb{C}$ definiremos serie*

$$\sum_{n=0}^{\infty} a_n (z - z_0)^n, \qquad z \in \mathbb{C}, \tag{1.2.1}$$

que denominaremos serie de potencias.

Como casos especiales de las series de potencias tenemos:

$$\sum_{k=0}^{\infty} x^k, \qquad \sum_{k=0}^{\infty} \frac{z^k}{k!}, \qquad \sum_{k=1}^{\infty} \frac{z^k}{k},$$

etc. Si la serie converge para todos los valores de z de cierto conjunto $A \subset \mathbb{C}$, entonces podemos definir la función suma $f(z)$ de la serie.

Propiedades de las series de potencias

Sin pérdida de generalidad vamos a considerar las series de potencias del tipo

$$\sum_{n=0}^{\infty} a_n z^n, \tag{1.2.2}$$

es decir, cuando $z_0 = 0$. Obviamente si tenemos una serie de potencias del tipo (1.2.1), la podemos reducir a la anterior (1.2.2) haciendo el cambio de variables $\zeta = z - z_0$ y viceversa.

Teorema 1.2.23 ***(Primer Teorema de Abel)*** *Sea la serie de potencias $\sum_{n=0}^{\infty} a_n z^n$, a_n, $z \in \mathbb{C}$. Si la serie converge para cierto[a] $w \in \mathbb{C} \setminus \{0\}$, entonces la serie converge absolutamente para todo $z \in \mathbb{C}$ con $|z| < |w|$.*

[a]Nótese que la serie siempre converge para $z = 0$.

Demostración: Como $w \neq 0$, entonces

$$\sum_{n=0}^{\infty} a_n z^n = \sum_{n=0}^{\infty} a_n w^n \left(\frac{z}{w}\right)^n.$$

Como $\sum_{n=0}^{\infty} a_n w^n$ es convergente, entonces por el Teorema 1.2.4 sabemos que $a_n w^n \stackrel{n\to\infty}{\longrightarrow} 0$, así pues, la sucesión $a_n\, w^n$ está acotada, luego existe un $M \in \mathbb{R}$, $M > 0$ tal que para todo $n \in \mathbb{N}$, $|a_n w^n| \leq M$, luego

$$\sum_{n=0}^{\infty} |a_n z^n| \leq \sum_{n=0}^{\infty} M \left|\frac{z}{w}\right|^n.$$

pero como $|z| < |w|$, $\left|\frac{z}{w}\right| < 1$, y la serie $\sum_{n=0}^{\infty} \left|\frac{z}{w}\right|$ es convergente, luego el criterio de comparación de Weierstrass 1.2.8 nos asegura la convergencia absoluta de la serie para $|z| < |w|$. ■

Corolario 1.2.24 *Si $\sum_{n=0}^{\infty} a_n z^n$ diverge para algún número complejo w, entonces diverge para todo z con $|z| > |w|$.*

Por tanto, el Primer Teorema de Abel nos asegura la existencia de regiones (círculos) de convergencia y divergencia, de una serie de potencias en $\mathbb{C}$. De hecho podemos afirmar que dada una serie de potencias cualquiera siempre existe un número no negativo $R \geq 0$ ó $R = +\infty$ tal que para todo z con $|z| < R$ la serie converge y si $|z| > R$ la serie diverge.

Definición 1.2.25 *El número $R \geq 0$ anterior se denomina radio de convergencia de una serie de potencias.*

Nota 1.2.26 *De lo anterior se deduce que la región de convergencia de una serie de potencias definida por la ecuación (1.2.1) es un círculo de radio R en el plano complejo está centrado en z_0 (ver la figura 1.1). Es sencillo[a] de ver que la región de convergencia de la serie definida por la ecuación (1.2.2) es entonces un círculo de radio R centrado en el origen. Nótese que se puede obtener una región de la otra simplemente mediante una traslación en el plano complejo (mediante el vector z_0).*

[a] Basta tomar $z - z_0 = \zeta$

Teorema 1.2.27 (Segundo Teorema de Abel) *Toda serie de potencias $\sum_{n=0}^{\infty} a_n z^n$, a_n, $z \in \mathbb{C}$, tiene radio de convergencia $R \geq 0$ ó $R = +\infty$. Además, la serie converge absolutamente para todo z tal que $|z| < R$.*

Demostración: Sea

$$A = \left\{\rho,\ \rho \geq 0;\ \sum_{n=0}^{\infty} a_n \rho^n \text{ converge}\right\}.$$

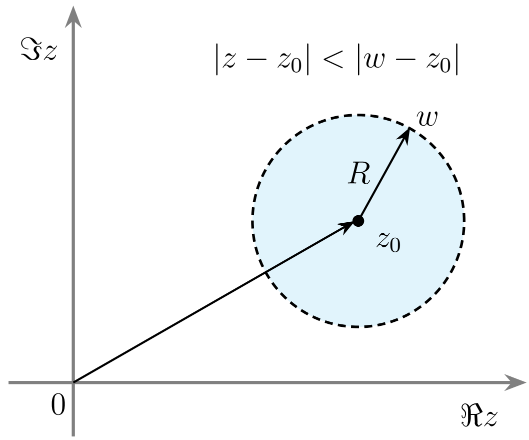


Figura 1.1: Región de convergencia de la serie $\sum_{n=0}^{\infty} a_n(z-z_0)^n$.

El conjunto A es no vacío pues para $\rho = 0$, la serie siempre converge. Entonces o bien $\sup A = +\infty$, o bien $\sup A < +\infty$. Sea $R = \sup A$. Si $R = 0$, entonces la serie sólo converge en $z = 0$. Si $R = +\infty$ la serie obviamente converge para todo z complejo y además por el Teorema de Abel lo hace absolutamente. Si $R < +\infty$, entonces por el Teorema de Abel la serie converge para todo $|z| < \rho < R$ y lo hace absolutamente. Además, por el corolario del primer Teorema de Abel la serie diverge para todo $|z| \geq \rho > R$, así pues R es el radio de convergencia. ∎

Nota 1.2.28 *Del teorema anterior se sigue que la mayor región de convergencia de la serie $\sum_{n=0}^{\infty} a_n z^n$ es el disco $U_R(0)$.*

Veamos ahora un criterio general para encontrar el radio de convergencia de una serie de potencias. Como corolario del criterio de Cauchy para las series tenemos el siguiente teorema:

Teorema 1.2.29 (Fórmula de Cauchy-Hadamard) *Dada una serie de potencias $\sum_{n=0}^{\infty} a_n z^n$, su radio de convergencia R viene dado por la fórmula*

$$R = \frac{1}{\overline{\lim_n} \sqrt[n]{|a_n|}}. \tag{1.2.3}$$

Demostración: Sea $z \in \mathbb{C}$. Sea $\limsup \sqrt[n]{|a_n z^n|} = q$. Como

$$\limsup \sqrt[n]{|a_n z^n|} = |z| \limsup \sqrt[n]{|a_n|},$$

entonces, usando el criterio generalizado de Cauchy 1.2.14 (ver la Nota 1.2.15), nuestra serie original converge si $|z| \limsup \sqrt[n]{|a_n|} < 1$ y además diverge si $|z| \limsup \sqrt[n]{|a_n|} > 1$. Obviamente si $\limsup \sqrt[n]{|a_n|} = 0$, entonces la convergencia es para todo z y, por tanto, en este caso $R = +\infty$. Si $\limsup \sqrt[n]{|a_n|} = +\infty$, entonces la serie converge solo para $z = 0$ y, por tanto, $R = 0$. Finalmente, si $0 < \limsup \sqrt[n]{|a_n|} < +\infty$, nuestra serie converge para $|z| < R$ y diverge para $|z| > R$ con $R = 1/\limsup \sqrt[n]{|a_n|}$, es decir, R es el radio de convergencia de la serie. ■

En particular, si existe el límite $\lim_{n\to\infty} \sqrt[n]{|a_n|} = l$, entonces $R = 1/\lim_{n\to\infty} \sqrt[n]{|a_n|}$, y si existe el límite $\lim_{n\to\infty} |a_n|/|a_{n+1}|$, entonces (ver el criterio de D'Alembert 1.2.13) $R = \lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|}$.

Ejercicio 1.2.30 *Estudiar la convergencia de las series*

$$\sum_{n=0}^{\infty} z^n, \quad \sum_{n=0}^{\infty} n! z^n, \quad \sum_{n=0}^{\infty} \frac{z^n}{n!}, \quad \sum_{n=1}^{\infty} \frac{z^n}{n}.$$

Ejercicio 1.2.31 *Calcular el radio de convergencia para las series de potencias*

$$\sum_{n=0}^{\infty} \frac{(-1)^n z^{2n+1}}{(2n+1)!}, \quad \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n}}{(2n)!}, \quad \sum_{n=1}^{\infty} \frac{z^{2n}}{n}.$$

Nótese que las series anteriores son tales que no todos los coeficientes a_n son distintos de cero. De hecho solo aparecen los pares o los impares.

Teorema 1.2.32 *Sea una serie de potencias $f(z) = \sum_{n=0}^{\infty} a_n z^n$ con radio de convergencia $R > 0$. Entonces*

1. *Para todo $k \geq 1$ la serie*

$$\sum_{n=0}^{\infty} n(n-1)\cdots(n-k+1) a_n z^{n-k}, \tag{1.2.4}$$

tiene radio de convergencia R.

2. *La función $f(z)$ es infinitamente diferenciable en $U_R(0)$ y la k-ésima derivada de f, $f^{(k)}(z)$, viene dada por la serie* (1.2.4).

3. *Para todo $n \geq 0$, $a_n = f^{(n)}(0)/n!$.*

Demostración: Asumiremos que $R > 0$. Para demostrar 1 es suficiente estudiar el caso $k = 1$ (para $k \geq 2$ el razonamiento es el mismo). Para $k = 1$ la serie (1.2.4) es de la forma $\sum_{n=0}^{\infty} n a_n z^{n-1} = 1/z \sum_{n=0}^{\infty} n a_n z^n$, y claramente $\limsup \sqrt[n]{n a_n} = \limsup \sqrt[n]{a_n}$ pues $\lim_{n\to\infty} \sqrt[n]{n} = 1$. Probemos 2. Sean

$$f(z) = \sum_{n=0}^{\infty} a_n z^n, \quad S_n(z) = \sum_{k=0}^{n} a_k z^k, \quad R_n(z) = \sum_{k=n+1}^{\infty} a_k z^k, \quad g(z) = \sum_{k=1}^{\infty} k a_k z^{k-1},$$

y escojamos z_0 tal que $|z_0| < r < R$. Entonces

$$\frac{f(z) - f(z_0)}{z - z_0} - g(z_0) = \left[\frac{S_n(z) - S_n(z_0)}{z - z_0} - S_n'(z_0)\right] + [S_n'(z_0) - g(z_0)] + \left[\frac{R_n(z) - R_n(z_0)}{z - z_0}\right].$$

Vamos a probar que cada sumando se puede hacer lo suficientemente pequeño. Comencemos por el tercero. Para ello nótese que

$$\frac{R_n(z) - R_n(z_0)}{z - z_0} = \sum_{k=n+1}^{\infty} a_k \frac{z^k - z_0^k}{z - z_0},$$

pero

$$\left|\frac{z^k - z_0^k}{z - z_0}\right| = |z^{k-1} + z^{k-2} z_0 + z^{k-3} z_0^2 + \cdots + z_0^{k-1}| \leq k r^{k-1},$$

luego entonces escogiendo $n > N_1$, suficientemente grande

$$\left|\frac{R_n(z) - R_n(z_0)}{z - z_0}\right| \leq \sum_{k=n+1}^{\infty} |a_k| k r^{k-1} \leq \frac{\varepsilon}{3},$$

pues la serie $\sum_{k=1}^{\infty} k a_k z^k$ converge absolutamente para todo z con $|z| < R$.

Para el segundo sumando notemos nuevamente que la serie $\sum_{k=1}^{\infty} k a_k z^{k-1}$ converge absolutamente para todo z con $|z| < R$, luego para dicho z_0

existirá un N_2 tal que para todo $n > N_2$, $|S'_n(z_0) - g(z_0)| \leq \varepsilon/3$. Sea $n > N = \max(N_1, N_2)$, entonces parta dicho n, en virtud de que $S_n(z)$ es diferenciable en z_0 (¿por qué?), existe un $\delta > 0$, tal que, para todo $z \in 0 < |z - z_0| < \delta$,

$$\left| \frac{S_n(z) - S_n(z_0)}{z - z_0} - S'_n(z_0) \right| < \frac{\varepsilon}{3}.$$

Juntando las tres desigualdades se sigue que para todo $\varepsilon > 0$, existe un $\delta > 0$, tal que, para todo $z \in 0 < |z - z_0| < \delta$,

$$\left| \frac{f(z) - f(z_0)}{z - z_0} - g(z_0) \right| < \varepsilon,$$

i.e., $f'(z_0) = g(z_0)$ que es lo que se quería probar.

Finalmente 3 se deduce de 1 y 2 sustituyendo en la fórmula

$$f^{(k)}(z) = \sum_{n=0}^{\infty} n(n-1)\cdots(n-k+1)a_n z^{n-k},$$

el valor $z = 0$. ■

Como consecuencia del teorema anterior se sigue que si $f(z)$ admite una serie de potencias, $f(z)$ es infinitamente diferenciable en $U_R(0)$, y cada una de las funciones $f^{(k)}(z)$, $k = 0, 1, 2, \ldots$, es continua en $U_R(0)$. Del tercer apartado se tiene además que la serie de potencias de una función dada coincide con la correspondiente serie de Taylor.

Ejercicio 1.2.33 *Prueba que la función $f(z) = \sum_{n=0}^{\infty} z^k/k!$ es infinitamente diferenciable y que $f'(z) = f(z)$, es decir, $f(z) = e^z$.*

1.2.4. Analiticidad de las funciones reales

Una de las principales diferencias entre las funciones complejas de variable compleja y las funciones reales de variable real está precisamente relacionada con la diferenciación y analiticidad. En el análisis real se dice que una función $f(x)$ es analítica en un entorno de un punto z_0 si en dicho entorno $f(x)$ se puede escribir como una serie de potencias

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n,$$

con radio de convergencia $R > 0$. Una propiedad muy importante de las series de potencias (reales o complejas) es que son infinitamente diferenciables. ¿Será cierto el recíproco? ¿Toda función C_A^∞ será analítica en algún entorno de $x_0 \in A$?

La respuesta a esta pregunta la dio Cauchy, en 1821. Definamos la función

$$f(x) = \begin{cases} e^{-1/x^2}, & x \neq 0, \\ \\ 0, & x = 0. \end{cases}$$

Esta función es infinitas veces derivable en $x = 0$ (de hecho en todo $\mathbb{R}$) además $f^{(k)}(0) = 0$ para todo $k \in \mathbb{N}$, luego su serie de Taylor en $x = 0$ es 0 y, por tanto, $f(x)$ no admite una serie de potencias en cero (o lo que es lo mismo, el radio de convergencia de la correspondiente serie es 0), i.e., la función de Cauchy no es analítica en $x_0 = 0$.

Teorema 1.2.34 (Condición necesaria y suficiente de analiticidad) *Una función $f(x)$ infinitamente derivable en todo un entorno de $x = x_0$ es analítica si y solo si el resto de Taylor*

$$R_n(x) = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x - x_0)^k,$$

tienda a cero para todo x de dicho entorno.

Demostración: Sea

$$P_n(x, x_0) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k. \qquad (1.2.5)$$

el polinomio de Taylor de f en el punto x_0 de una función f infinitas veces derivable en $x = x_0$ y sea $R_n(x, x_0) = f(x) - P_n(x, x_0)$ el resto o error de Taylor. Obviamente el polinomio de Taylor $P_n(x, x_0)$ coincide con las sumas parciales de la serie de potencias de f, luego $P_n(x, x_0)$ converge a $f(x)$ en $(x_0 + R, x_0 - R)$ si y solo si $R_n(x, x_0) \overset{n\to\infty}{\longrightarrow} 0$ para todo $x \in (x_0 + R, x_0 - R)$. ■

El caso complejo es mucho más *sencillo*. De hecho se tiene el siguiente sorprendente resultado[3] cuya demostración omitiremos:

[3] Ver, por ejemplo, A.I. Markushevich, *Teoría de las funciones analíticas*. Vol I. MIR, Moscú, 1970, pág 310.

Teorema 1.2.35 (Cauchy-Goursat) *Sea $\Omega \subset \mathbb{C}$ una región (abierto y conexo) y sea $f : \Omega \mapsto \mathbb{C}$ una función diferenciable (holomorfa). Entonces, para cada punto $z_0 \in \Omega$, $f(z)$ admite una serie de potencias $f(z) = \sum_{n=0}^{\infty} a_n(z-z_0)^n$ con radio de convergencia $R > 0$.*

1.3. Continuidad uniforme

1.3.1. Definición de continuidad uniforme

Recordemos la definición de continuidad de una función en un conjunto $A \in \mathcal{D}(f)$, donde $\mathcal{D}(f)$ denota al dominio de f.

Definición 1.3.1 *Se dice que una función $f : A \mapsto \mathbb{R}$ es continua en A si,*

$$\forall \varepsilon > 0,\ \forall a \in A,\ \exists \delta := \delta(a,\varepsilon) > 0;\ \textit{tal que } |x-a| < \delta \Rightarrow |f(x)-f(a)| < \varepsilon.$$

Hemos de destacar que en la definición anterior δ depende, en general, no sólo de ε, sino también del punto a. Veamos un ejemplo.

Sea $f(x) = 1/x$ en $(0,1]$. f es continua, de hecho lo será en cualquier intervalo que no contenga al cero. Para probarlo usamos la definición. Puesto que

$$|f(x)-f(a)| = \left|\frac{x-a}{x\,a}\right| = \frac{1}{|x\,a|}|x-a|.$$

Si $a > 0$ entonces podemos encontrar un $\delta_1 > 0$ tal que si $x \in (a-\delta_1, a+\delta_1) \subset (0,\infty)$, $1/|x\,a| < M$, de hecho, M es más grande a medida que a se acerque más a cero. En efecto, puesto que $1/x$ es decreciente tenemos que si $a \neq 0$ siempre podemos encontrar un $\delta_1 > 0$ (basta elegirlo menor que $a/2$) tal que $1/x \leq 1/(a/2) = 2/a$. Es decir,

$$|f(x)-f(a)| < \frac{2}{a^2}|x-a|, \qquad \forall x \in (a-\delta_1, a+\delta_1) \subset (0,\infty).$$

Sea $\varepsilon > 0$ cualquiera, entonces si x es tal que $0 < |x-a| < \delta = \varepsilon a^2/2$ tenemos que $|f(x)-f(a)| < \varepsilon$.

En la figura 1.2 se muestra lo que ocurre. Si prefijamos un $\varepsilon > 0$ tenemos los dos entornos $U(f(a))$ y $U(f(b))$ con $b < a$. Obviamente a medida que nos acercamos a cero el correspondiente entorno de a, $U(a)$ se hace

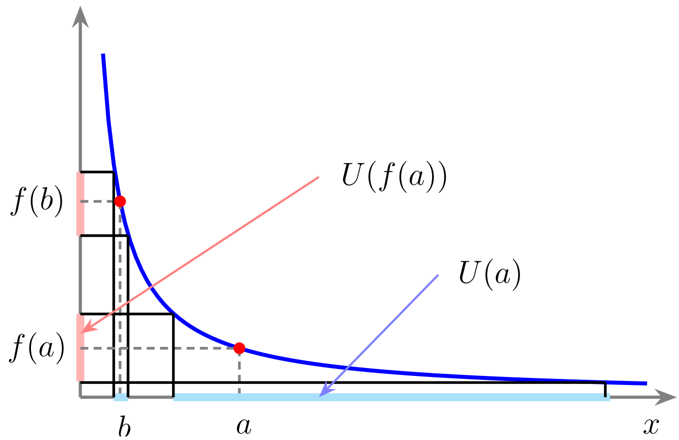


Figura 1.2: La función $f(x) = 1/x$, y la continuidad uniforme.

cada vez más pequeño además la longitud del mismo, como hemos visto es igual a $d = \varepsilon a^2/2$, así que si $a \to 0$, $d \to 0$.

No ocurre lo mismo si escogemos, por ejemplo, en intervalo $[\alpha, 1]$ cualquiera sea $\alpha \in (0, 1)$. En efecto, en este caso, nuestra desigualdad será

$$|f(x) - f(a)| < \frac{2}{a^2}|x - a| < \frac{2}{\alpha^2}|x - a|, \quad \forall x \in (a - \delta_1, a + \delta_1) \subset (0, \infty).$$

Es decir, que cualquiera sea $\varepsilon > 0$, entonces si x es tal que $0 < |x - a| < \delta = \varepsilon\alpha^2/2$ tenemos que $|f(x) - f(a)| < \varepsilon$, o sea, ¡nuestro δ en este caso no depende del punto a!, o sea, el mismo δ nos vale para todo el intervalo.

Otro tanto le ocurre a la función $f(x) = x^2$ en el intervalo $[0, \infty)$. Para $a > 0$ calculemos[4]

$$|f(x) - f(a)| = (x + a)|x - a| < (2a + \delta_1)|x - a|, \quad \text{si} \quad x \in (a - \delta_1, a + \delta_1).$$

Sea $\delta_1 = 1$, por ejemplo. Entonces, cualquiera sea $\varepsilon > 0$, si x es tal que $|x - a| < \delta = \varepsilon/(2a + 1)$ tenemos que $|f(x) - f(a)| < \varepsilon$, es decir, que a medida que a aumenta δ disminuye. Nuevamente no podemos obtener un δ válido para todo $a \geq 0$. Que ocurre si escogemos el intervalo $[0, b]$, $b < \infty$. Entonces,

$$|f(x) - f(a)| = (x + a)|x - a| < (2b)|x - a|, \qquad \forall x \in [0, b],$$

[4]Si $a = 0$ obviamente f es continua, además, las desigualdades también son válidas para $x \geq 0$.

luego, cualquiera sea $\varepsilon > 0$, si x es tal que $|x - a| < \delta = \varepsilon/(2b)$ tenemos que $|f(x) - f(a)| < \varepsilon$, es decir, δ no depende del punto a, por tanto el mismo δ nos vale para toda el intervalo.

Definición 1.3.2 *Se dice que una función $f : A \mapsto \mathbb{R}$ es uniformemente continua en un intervalo $[a, b]$ si para todo $\varepsilon > 0$ existe un $\delta > 0$ tal que para todos x e y de $[a, b]$, tales que $|x - y| < \delta$, entonces $|f(x) - f(y)| < \varepsilon$.*

Es decir, para cualquiera sea el punto x de $[a, b]$ que escojamos, $f(x)$ es continua y el δ en la definición anterior, a diferencia del de la definición de continuidad "estándar" 1.3.1 solo depende de ε y no del punto, o equivalentemente, que si escogemos un $\varepsilon > 0$ cualquiera y denotamos por $U(f)$ el ε-entorno de cualquier punto $y = f(x)$ de la imagen, entonces, existe un $\delta > 0$ tal que los correspondientes δ-entornos de x, $U(x)$ son tales que $f(U(x)) \subset U(f)$.

Es evidente que cualquier función uniformemente continua en $[a, b]$ es continua en dicho intervalo. Ahora bien, el recíproco, como hemos visto, no es cierto. Por tanto, es interesante tener teoremas que nos aseguren cuando una función es uniformemente continua y cuando no. Nótese que una función no es uniformemente continua en $A \subset \mathbb{R}$ si

$$\exists \varepsilon > 0, \quad \forall \delta > 0 \exists x, a \in A; \quad |x - a| < \delta, \quad \Rightarrow \quad |f(x) - f(a)| \geq \varepsilon.$$

Teorema 1.3.3 *Una función f no es uniformemente continua en A si existe un $\varepsilon > 0$ y dos sucesiones (a_n) y (b_n) con $|a_n - b_n| \overset{n\to\infty}{\longrightarrow} 0$ tales que $|f(a_n) - f(b_n)| \geq \varepsilon$.*

Demostración: Es inmediata a partir de la definición de convergencia uniforme (o lo que es lo mismo de su negación). ■

Ejercicio 1.3.4 *Prueba que $f(x) = 1/x$ y $g(x) = x^2$ no son uniformemente continuas en $(0, 1]$ y $[0, \infty)$, respectivamente, usando el teorema anterior.*

En el primer caso escogemos $a_n = 1/n$ y $b_n = 1/(n+1)$, $a_n - b_n = 1/(n(n+1)) \overset{n\to\infty}{\longrightarrow} 0$, pero $|f(a_n) - f(b_n)| = 1$. Para el segundo caso escogemos $a_n = \sqrt{n+1}$ y $b_n = \sqrt{n}$, de forma que $a_n - b_n = 1/(\sqrt{n+1} + \sqrt{n}) \overset{n\to\infty}{\longrightarrow} 0$, pero $|g(a_n) - g(b_n)| = 1$. ■

El siguiente teorema nos da un criterio muy útil para saber cuando una función es uniformemente continua en un intervalo cerrado y acotado $[a, b]$:

Teorema 1.3.5 *(Heine)*[a] *Si la función* $f : [a,b] \mapsto \mathbb{R}$ *es continua en todo* $[a,b]$ *entonces* f *es uniformemente continua en* $[a,b]$.

[a]Este teorema, a veces atribuido a Cantor, también es cierto para funciones de varias variables. Sea $A \subset \mathbb{R}^n$ un conjunto cerrado y acotado (compacto) y sea la función $f : A \mapsto \mathbb{R}$, continua en A. Entonces f es uniformemente continua en A.

Demostración: Supongamos que el teorema es falso. Entonces

$$\exists\varepsilon > 0, \quad \forall\delta > 0 \exists x, y \in A; \quad |x-y| < \delta, \quad \Rightarrow \quad |f(x) - f(y)| \geq \varepsilon.$$

Entonces para cada $n \in \mathbb{N}$, si escogemos $\delta = 1/n$, existen dos números a_n y b_n tales que $|a_n - b_n| < 1/n$ y $|f(a_n) - f(b_n)| \geq \varepsilon$, o sea, podemos construir sendas sucesiones (a_n) y (b_n). Ahora bien, como a_n es acotada, por el Teorema de Bolzano-Weierstrass para sucesiones existe una subsucesión $\{a_{k_n}\}$ convergente a cierto $x \in [a,b]$. Tomemos ahora la subsucesión $(b_{k_n})_n$ donde k_n es exactamente la misma subsucesión de números naturales que define a la subsucesión $(a_{k_n})_n$ de antes. Para dicha subsucesión tenemos que

$$|b_{k_n} - x| \leq \underbrace{|b_{k_n} - a_{k_n}|}_{<1/n} + \underbrace{|a_{k_n} - x|}_{\to 0} \stackrel{n\to\infty}{\longrightarrow} 0, \quad \Rightarrow b_{k_n} \stackrel{n\to\infty}{\longrightarrow} x,$$

es decir, la subsucesión $\{b_{k_n}\}$ de (b_n) también tiene límite x. Pero f es continua en $[a,b]$ así que

$$\lim_{n\to\infty} f(a_{k_n}) = f(x), \quad \lim_{n\to\infty} f(b_{k_n}) = f(x), \qquad \Rightarrow |f(a_{k_n}) - f(b_{k_n})| \stackrel{n\to\infty}{\longrightarrow} 0,$$

lo cual contradice que $|f(a_n) - f(b_n)| \geq \varepsilon$ (ya que como $\{a_{k_n}\}$ y $\{b_{k_n}\}$ son subsucesiones de $\{a_n\}$ y $\{b_n\}$, tendría que ser $|f(a_{k_n}) - f(b_{k_n})| \geq \varepsilon$). ∎

El teorema anterior nos indica que la función $f(x) = 1/x$ es uniformemente continua en cualquier cerrado y acotado que no contenga al cero y que la función $g(x) = x^2$ es uniformemente continua en cualquier cerrado y acotado de $\mathbb{R}$.

De la de la definición de continuidad uniforme se sigue (basta escoger $\delta < \varepsilon/K$) el siguiente:

Teorema 1.3.6 *(Condición suficiente de continuidad uniforme) Para que la función* f *sea uniformemente continua en* A*, es suficiente que exista una constante* $K > 0$ *tal que para todo* $x, y \in A$*,* $|f(x) - f(y)| \leq K|x-y|$.

Las funciones $f : A \mapsto \mathbb{R}$ que cumplen con que

$$\forall x, y \in A, \quad |f(x) - f(y)| \leq K|x - y|,$$

se llaman *funciones de Lipschitz*. Así que el teorema anterior se puede parafrasear diciendo que toda función de Lipschitz es uniformemente continua.

Obviamente toda función de Lipschitz es continua, pero no viceversa, no toda función continua es de Lipschitz. Por ejemplo, $f : [0,1] \mapsto \mathbb{R}$, $f(x) = \sqrt{x}$ es continua en $[0,1]$ (y por tanto uniformemente continua en $[0,1]$) pero no es de Lipschitz en $[0,1]$, pues (se supone que $x \neq y$ y $x, y \neq 0$)

$$|f(x) - f(y)| = |\sqrt{x} - \sqrt{y}| \leq \frac{1}{|\sqrt{x} + \sqrt{y}|}|x - y|,$$

y el factor $\frac{1}{|\sqrt{x}+\sqrt{y}|}$ no es acotado en $[0,1]$.

1.4. Sucesiones y series de funciones reales

Definición 1.4.1 *La sucesión de funciones $(f_n(x))_n$ tiene límite para cierto $x_0 \in \mathcal{D}(f_n)$ si la sucesión numérica $(f_n(x_0))_n$ converge para dicho x_0. En este caso diremos que la sucesión converge puntualmente en x_0,*

Definición 1.4.2 *El conjunto $\mathbb{X}$ de todos los puntos $x \in \mathcal{D}(f_n)$ donde la sucesión $(f_n(x))_n$ converge se denomina conjunto de convergencia de la sucesión de funciones.*

Evidentemente para todo $x \in \mathbb{X}$, existe un único valor del límite de sucesión numérica $(f_n(x))_n$, por tanto podemos definir una función $f(x)$ sobre $\mathbb{X}$ que es la función límite de $f_n(x)$. Así tenemos la siguiente

Definición 1.4.3 *En el conjunto $\mathbb{X}$ de convergencia de la sucesión de funciones $(f_n(x))_n$ la función $f(x)$ definida para cada $x \in \mathbb{X}$ de la forma $f(x) = \lim_{n\to\infty} f_n(x)$ se denomina función límite de $(f_n(x))_n$ y se dice que la sucesión $(f_n(x))_n$ converge puntualmente a $f(x)$ en $\mathbb{X}$ y lo denotaremos como $f_n(x) \to f(x)$.*

Ejemplo 1.4.4 *Consideremos los siguientes ejemplos representativos.*

1. Definamos en $[0,\infty)$ la sucesión $f_n(x) = x^n$, $n \in \mathbb{N}$. Es fácil comprobar que la sucesión de funciones converge si y solo si $x \in [0,1]$, además
$$\lim_{n\to\infty} x^n = f(x) = \begin{cases} 1, & x = 1 \\ 0 & 0 \le x < 1 \end{cases},$$
es la función límite. Nótese que si escogemos $\mathbb{X} = [0,q]$, $0 \le q < 1$, entonces $f(x) = 0$.

2. Sea $f_n(x) : \mathbb{R} \mapsto \mathbb{R}$, $f_n(x) = \dfrac{\operatorname{sen} n^2 x}{n}$. Es evidente que $\mathbb{X} = \mathbb{R}$ y $f(x) = 0$.

3. Sea $f_n(x) : \mathbb{R} \mapsto \mathbb{R}$, $f_n(x) = \dfrac{\operatorname{sen} nx}{n^2}$. Como en el ejemplo anterior $\mathbb{X} = \mathbb{R}$ y $f(x) = 0$.

4. Sea $f_n(x) : [0,1] \mapsto \mathbb{R}$, $f_n(x) = 2(n+1)x(1-x^2)^n$. Evidentemente si $x = 0$ o $x = 1$, $f_n(x) = 0$. Si $x \in (0,1)$ entonces $(1-x^2) < 1$, luego $f_n(x) \to 0$ cuando $n \to \infty$. Por tanto, $\mathbb{X} = [0,1]$ y $f(x) = 0$.

5. Sean $n, m \in \mathbb{N}$ y sea $f_m(x) = \lim_{n\to\infty} (\cos m!\pi x)^{2n}$. Si $m!x \in \mathbb{Z}$, entonces $f_m(x) = 1$, si $m!z \notin \mathbb{Z}$, entonces $f_m(x) = 0$. Veamos el límite cuando $m \to \infty$ de $f_m(x)$. Si $x \notin \mathbb{Q}$, entonces para todo $m \in \mathbb{N}$, $m!x \notin \mathbb{Z}$, por tanto $f_m(x) = 0$ y $f(x) = \lim_{m\to\infty} f_m(x) = 0$. Si $x \in \mathbb{Q}$, entonces a partir de cierto $m \in \mathbb{N}$, $m!x \in \mathbb{Z}$, por tanto, a partir de dicho m, $f_m(x) = 1$, luego $f(x) = \lim_{m\to\infty} f_m(x) = 1$. Así, tenemos que $\mathbb{X} = \mathbb{R}$ y $f(x)$ es la función de Dirichlet $D(x)$
$$D(x) = \begin{cases} 1, & x \in \mathbb{Q} \\ 0, & x \in \mathbb{I} \end{cases}.$$

Nuestro objetivo fundamental es estudiar, no cuando las sucesiones funcionales convergen o no –pues eso ya lo hemos visto en temas anteriores pues para cada x la sucesión $(f_n(x))_n$ es una sucesión numérica– sino decidir en que condiciones las propiedades de f_n se traspasan a la función límite. ¿Hereda $f(x)$ todas las propiedades de las $f_n(x)$? Los ejemplos anteriores sirven de ilustración a nuestro problema.

En el Ejemplo 1 las funciones x^n son continuas en $[0,1]$ pero la función

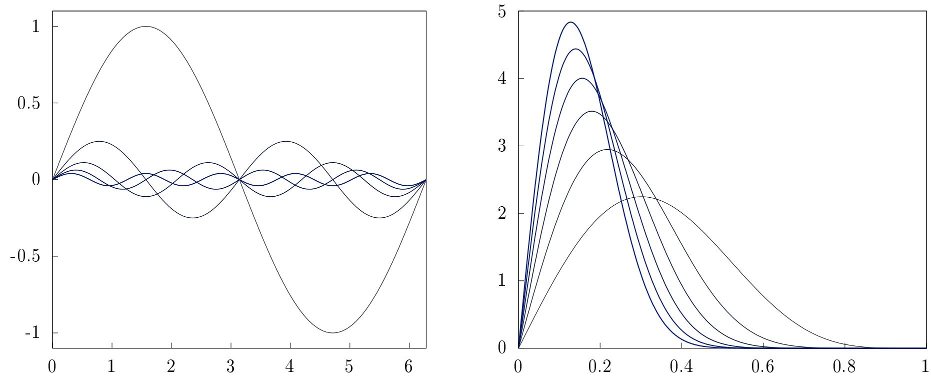


Figura 1.3: La sucesión $f_n(x) = \operatorname{sen} nx/n^2$ en $[0, 2\pi)$ (izquierda) y $f_n(x) = 2(n+1)x(1-x^2)^n$ en $[-1, 1]$ (derecha).

límite $f(x)$ no lo es. No ocurre lo mismo en el intervalo $\mathbb{X} = [0, q]$.

En el Ejemplo 2 la función $f_n(x) = (\operatorname{sen} n^2 x)/n$ es derivable en todo $\mathbb{R}$, y su función límite también, pero por ejemplo, la sucesión de sus derivadas $f_n'(x) = n \cos n^2 x$ no tiene límite excepto cuando $n^2 x = k\pi/2$, $k \in \mathbb{Z}$. Es decir, la sucesión de derivadas no converge a la derivada de la función límite.

En el Ejemplo 3, a diferencia del anterior la sucesión de sus derivadas $f_n'(x) = (\cos nx)/n$ converge a cero que justamente es la derivada de la función límite. Es decir, en este ejemplo sí que tenemos $f_n'(x) \to f'(x)$.

Analizando el Ejemplo 4 vemos que la función $f_n(x) = 2(n+1)x(1-x^2)^n$ es integrable en $[0, 1]$ y además

$$\int_0^1 f_n(x)dx = 2(n+1)\int_0^1 x(1-x^2)^n dx = 2(n+1)\left[-\frac{(1-x^2)^{n+1}}{n+1}\Bigg|_0^1 \right] = 1,$$

pero $\displaystyle\int f(x)dx = \int_0^1 0\,dx = 0$, luego $\displaystyle\int_0^1 f_n(x)dx \not\to \int f(x)dx$.

A diferencia del caso anterior, podemos comprobar que en $[0, 1]$ la función $f_m(x)$ del Ejemplo 5 vale cero excepto un número finito de puntos, luego f_m es integrable y su integral vale cero, pero su función límite es una función no integrable (Riemann).

Finalmente, en el Ejemplo 1 tenemos

$$\int_0^1 f_n(x)dx = \int_0^1 x^n\,dx = \frac{1}{n+1} \to 0 = \int_0^1 0\,dx = \int_0^1 f(x)dx.$$

Todo lo anterior nos dice que dada una sucesión $f_n(x)$ con determinadas propiedades como la continuidad, derivabilidad, integrabilidad, la función límite puede o no tenerlas. Por tanto nuestro objetivo es claro. Encontrar condiciones bajo las cuales la función límite "herede" las propiedades de las funciones de la sucesión.

Definición 1.4.5 *Una sucesión de funciones* $(f_n(x))_n$ *converge puntualmente en* $I \in \mathbb{X}$ *si para todo* $x_0 \in I$ *y todo* $\varepsilon > 0$ *existe un* $N \equiv N(x_0, \varepsilon) \in \mathbb{N}$ *tal que para todo* $n > N$, $|f_n(x_0) - f(x_0)| < \varepsilon$ *y lo denotaremos* $f_n \to f$.

Claramente vemos que, en general, nuestro N dependerá de x_0. Consideremos como ejemplo la sucesión $f_n(x) = x^n$ en $[0,1)$. Obviamente $f_n(x) \to 0$ para todo $x \in [0,1)$. Sea $0 < \varepsilon < 1$ cualquiera y $x_0 \in [0,1)$. Entonces, tendremos $|f_n(x_0) - 0| = x_0^n < \varepsilon$. Si escogemos N tal que $x_0^N < \varepsilon$, o equivalentemente $N \log x_0 < \log \varepsilon$, tendremos, como $\log x_0 < 0$ que, para todo $n > N = \log\varepsilon / \log x_0$, $x_0^n < \varepsilon$. Es decir, que nuestra N depende explícitamente de x_0 y lo que es peor, a medida que x_0 se va acercando a 1, N claramente va tomando valores más grandes. La siguiente figura 1.4 aclara lo anterior.

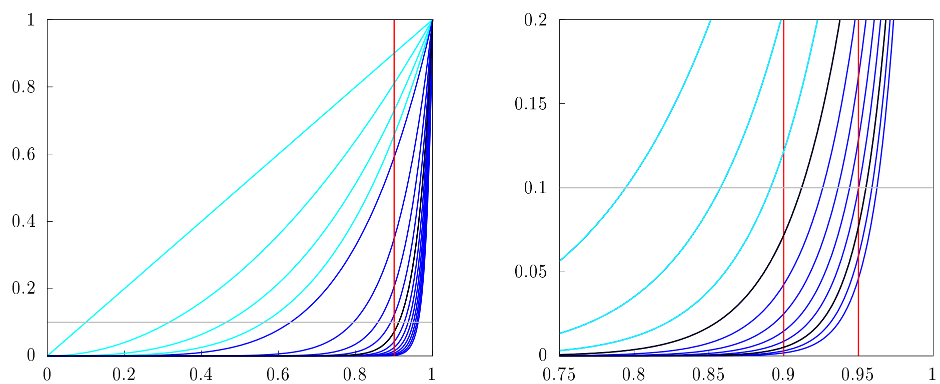


Figura 1.4: La sucesión $f_n(x) = x^n$ en $[0,1)$ (izquierda) y en $[q,1)$ (derecha).

En la figura 1.4 (izquierda) vemos la sucesión $f_n(x) = x^n$ dibujada en $[0, 1)$. Las curvas son, de arriba hacia abajo, x, x^2, x^3, x^5, $x^1 0 \ldots x^{60}$ (a partir de $n = 5$ se representan las x^{5n}). La linea horizontal representa a nuestro $\varepsilon > 0$ y la vertical un x_0 dado. Como vemos, dado el $\varepsilon > 0$ de la gráfica de la izquierda, a partir de cierto N tenemos que $x_0^N < \varepsilon$ es menor que el $\varepsilon > 0$ dado –a partir de la novena curva de arriba abajo–. Veamos que ocurre más cerca del 1, para ello quedémonos solo con un pequeño intervalo cercano al 1, digamos $[q, 1)$, estando q “suficientemente” cercano a 1. Aumentemos nuestro x_0 hasta hacerlo más próximo a 1 tal y como observamos en la figura de la derecha –la primera linea vertical era nuestro anterior x_0 y la segunda, más a la derecha es nuestro nuevo x_0–. Como vemos, ahora tenemos que aumentar nuestro N hasta llegar a la catorceava curva. Concretamente, en las gráficas de la figura 1.4 hemos escogido los siguientes valores numéricos: para $\varepsilon = 0{,}1$ y $x_0 = 0{,}9$ el valor de N es 25. En el segundo caso $\varepsilon = 0{,}1$, $x_0 = 0{,}95$ y el N necesario fue 45. Obviamente a medida que nos acercamos a 1 nuestro N crecerá tal y como hemos mostrado antes.

Una pregunta evidente es la siguiente. ¿Cualquiera sea una sucesión $(f_n(x))_n$ siempre ocurrirá lo mismo. Para responderla consideremos el siguiente ejemplo:

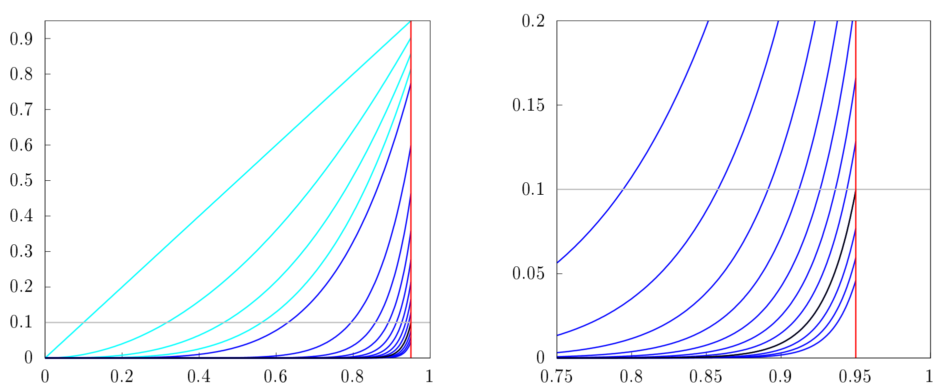


Figura 1.5: La sucesión $f_n(x) = x^n$, en $[0, q]$ (izq.) y una ampliación (der.)

Sea la misma sucesión $f_n(x) = x^n$ pero en $[0, q]$ con $q \in (0, 1)$. Hagamos un experimento numérico como el anterior usando como antes $\varepsilon = 0{,}1$. En la figura 1.5 representaremos nuevamente mediante linea horizontal a nuestro $\varepsilon > 0$, la primera linea vertical corresponderá a un x cualquiera en $[0,q]$ y la segunda linea vertical corresponderá a q, el extremo del intervalo. Dibujaremos la sucesión x^n para que se vea con más claridad el comportamiento de la misma. Obviamente $x \le q$, así que escogeremos $q = 0{,}95$, por ejemplo. Un simple vistazo a la figura 1.5 nos permite ver que cualquiera sea el $x \le q$ que escojamos, dado el $\varepsilon > 0$ tenemos un valor de N, que en nuestro caso corresponde a la séptima gráfica tal que $x^n < \varepsilon$ para todo $n > N$ y ¡cualquiera sea la $x \in [0,q]$! Es decir, podemos encontrar una N válida en todo punto de $[0,q]$. En realidad esto no es de extrañar pues para todo $\varepsilon > 0$ si escogemos $N = \log \varepsilon / \log q$, entonces tendremos $q^n < \varepsilon$, por tanto para todo $x \in [0,q]$, y para todo $n > N$ tendremos $x^n \le q^n < \varepsilon$, es decir, x^n tiende a 0 en todo $[0,q]$ y además lo hace *uniformemente* en todo el intervalo.

Lo anterior nos conduce de manera natural a la siguiente

Definición 1.4.6 *Dada una sucesión de funciones $(f_n(x))_n$ definidas en un intervalo $I \in \mathbb{X}$, diremos que $f_n(x)$ converge uniformemente a $f(x)$ en I si para todo $\varepsilon > 0$ existe un $N \equiv N(\varepsilon) \in \mathbb{N}$ tal que para todo $n > N$ y para todos los $x \in I$, $|f_n(x) - f(x)| < \varepsilon$, y lo denotaremos $f_n \rightrightarrows f$.*

Gráficamente la convergencia uniforme está representada en el gráfico de la izquierda de la figura 1.6.

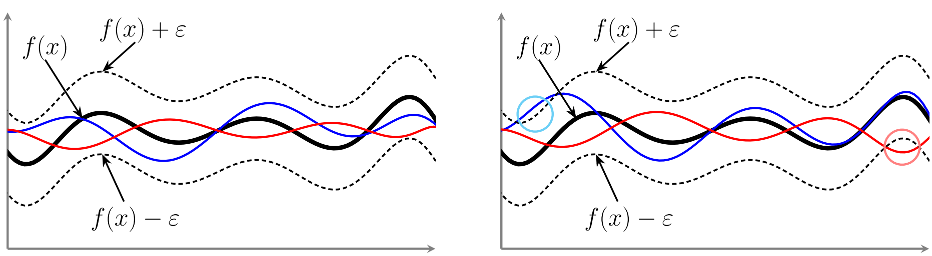


Figura 1.6: Convergencia uniforme (izquierda) y no uniforme (derecha) de una sucesión de funciones $(f_n(x))_n$.

La función f está al centro y está localizada en el interior de cierto entorno de longitud ε. Cualquiera sea la $x \in I$, $f_n(x)$ debe estar contenida en el entorno $(f(x)-\varepsilon, f(x)+\varepsilon)$ tal y como se muestra en dicha figura. En esta misma figura 1.6 podemos ver a la derecha lo que ocurre cuando la convergencia es no uniforme. Y es que para todo n siempre existe un x tal que $f_n(x)$ sale fuera del entorno $(f(x)-\varepsilon, f(x)+\varepsilon)$ (ver los círculos azul y rosa). Es evidente que si $f_n \rightrightarrows f$ entonces $f_n \to f$.

Nuestro próximo objetivo es dar condiciones necesarias y suficientes para que una sucesión de funciones converja uniformemente. Para ello vamos a considerar las funciones representadas en la figura 1.6. Vamos a representar gráficamente –ver gráfico 1.7– las funciones $f(x) - f_n(x)$ de la figura 1.6. Las líneas horizontales representan los valores de $\pm\varepsilon$ escogidos.

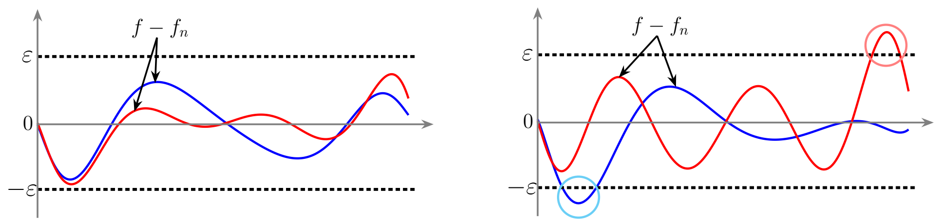


Figura 1.7: Gráficos de $f(x) - f_n(x)$ en el caso de la convergencia uniforme (izquierda) y no uniforme (derecha) de $(f_n(x))_n$.

Está claro de las gráficas de la figura 1.7 que para que la convergencia sea uniforme, todas las funciones $\phi_n(x) = |f(x) - f_n(x)|$ han de estar contenidos en el entorno $[0, \varepsilon)$ lo cual ocurrirá si los máximos o supremos de cada función $\phi_n(x)$ –que son números reales– están contenidos en dicho intervalo. No ocurre lo mismo en el caso de la convergencia no uniforme, donde siempre, cualquiera sea el ε que escojamos siempre habrá un valor de n tal que el valor de la función $\phi_n(x)$ en cierto punto x_0 del dominio sobrepasa a ε, pero entonces, si ese punto no corresponde al máximo, éste, si existe, también ha de sobrepasar a ε. Los anteriores razonamientos "geométricos" nos lleva a la siguiente condición necesaria y suficiente de convergencia uniforme para una sucesión de funciones:

Teorema 1.4.7 (Condición necesaria y suficiente de convergencia uniforme) *Dada una sucesión de funciones $(f_n(x))_n$ definidas en $I \in \mathbb{X}$, $f_n(x)$ converge uniformemente a $f(x)$ en I si y solo si*

$$\lim_{n\to\infty} \sup_{x\in I} |f_n(x) - f(x)| = 0.$$

Demostración: Supongamos que $f_n \rightrightarrows f$ en I. Entonces para todo $\varepsilon > 0$ existe un $N \in \mathbb{N}$ tal que para todo $n > N$ y para todos los $x \in I$, $|f_n(x) - f(x)| < \varepsilon/2$, luego para todo $\varepsilon > 0$ existe un N tal que para todo $n > N$ la sucesión $\sup_{x\in I} |f_n(x) - f(x)| < \varepsilon$, luego $\lim_{n\to\infty} \sup_{x\in I} |f_n(x) - f(x)| = 0$.

Por el contrario, si $\lim_{n\to\infty} \sup_{x\in I} |f_n(x) - f(x)| = 0$, entonces para todo $\varepsilon > 0$ existe un N tal que para todo $n > N$, $\sup_{x\in I} |f_n(x) - f(x)| < \varepsilon$, luego, por la propiedad del supremo tenemos que para todo $x \in I$, $|f_n(x) - f(x)| < \varepsilon$, es decir, para todo $\varepsilon > 0$ existe un $N \in \mathbb{N}$ tal que para todo $n > N$ y para todos los $x \in I$, $|f_n(x) - f(x)| < \varepsilon$, o sea, $f_n \rightrightarrows f$ en I.■

Corolario 1.4.8 *Para que una sucesión de funciones $(f_n(x))_n$ definidas en $I \in \mathbb{X}$, $f_n(x)$ converja uniformemente a $f(x)$ en I es necesario y suficiente que exista una sucesión $(a_n)_n$ de términos no negativos con $\lim_{n\to\infty} a_n = 0$, y un número $N \in \mathbb{N}$ tal que para todo $n > N$, $|f_n(x) - f(x)| \leq a_n$.*

Demostración: De las condiciones dadas

$$\sup_{x\in I} |f_n(x) - f(x)| \leq \sup_{x\in I} a_n = a_n \to 0,$$

de donde se sigue el resultado. ■

Ejercicio 1.4.9 *Probar que la sucesión $f_n(x) = x^n$ converge uniformemente a cero en $[0, q]$, $0 < q < 1$, pero no lo hace en $[0, 1)$.*

En efecto, en el primer caso tenemos

$$\sup_{x\in[0,q]} |x^n - 0| = q^n = a_n, \qquad \text{y} \qquad \lim_{n\to\infty} q^n = 0,$$

mientras que en el segundo $\sup_{x\in[0,1)} |x^n - 0| = 1$ que no tiende a 0 cuando $n \to \infty$. ■

Teorema 1.4.10 (Criterio de Cauchy de convergencia uniforme para una sucesión de funciones) *Dada una sucesión de funciones $(f_n(x))_n$ definidas en $I \in \mathbb{X}$, $f_n(x)$ converge uniformemente a $f(x)$ en I si y solo si para todo $\varepsilon > 0$, existe un $N \in \mathbb{N}$ tal que para todo $n > N$, para todo $p \in \mathbb{N}$ y para todo $x \in I$, $|f_{n+p}(x) - f_n(x)| < \varepsilon$.*

Demostración: Supongamos que $f_n \rightrightarrows f$ en I. Entonces para todo $\varepsilon > 0$ existe un $N \in \mathbb{N}$ tal que para todo $n > N$ y para todos los $x \in I$, $|f_n(x) - f(x)| < \varepsilon/2$. Entonces, para $n > N$, y todo $x \in I$, como $n + p \geq N$, tenemos

$$\begin{aligned} |f_{n+p}(x) - f_n(x)| &= |f_{n+p}(x) - f(x) + f(x) - f_n(x)| \\ &\leq |f_{n+p}(x) - f(x)| + |f_n(x) - f(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Por el contrario, si para todo $\varepsilon > 0$, existe un $N \in \mathbb{N}$ tal que para todo $n > N$, para todo $p \in \mathbb{N}$ y para todo $x \in I$, $|f_{n+p}(x) - f_n(x)| < \varepsilon/2$, entonces el criterio de Cauchy para las sucesiones nos asegura que existe el límite para cada $x \in I$. Pero además, tomando el límite $p \to \infty$ tenemos que para todo $x \in I$, $|f_n(x) - f(x)| \leq \varepsilon/2 < \varepsilon$, por tanto $f_n \rightrightarrows f$ en I. ■

Del teorema anterior se deduce el siguiente corolario.

Corolario 1.4.11 *Una sucesión de funciones $(f_n(x))_n$ definidas en $I \in \mathbb{X}$, converge uniformemente a $f(x)$ en I si y solo si para todo $\varepsilon > 0$, existe un $N \in \mathbb{N}$ tal que para todo $n > N$ y para todo $p \in \mathbb{N}$ $\sup_{x\in I} |f_{n+p}(x) - f_n(x)| < \varepsilon$.*

La demostración del siguiente resultado se deja como ejercicio al lector.

Proposición 1.4.12 *Si $(f_n(x))_n$ y $(g_n(x))_n$ son dos sucesiones de funciones uniformemente convergentes en I a las funciones $f(x)$ y $g(x)$, respectivamente, entonces, para todos $\alpha, \beta \in \mathbb{R}$ y toda función $h(x)$ acotada en I se tiene que, en I,*

$$\alpha f_n(x) + \beta g_n(x) \rightrightarrows \alpha f(x) + \beta g(x), \quad h(x) f_n(x) \rightrightarrows h(x) f(x).$$

Pasemos a continuación a estudiar lo que ocurre en el caso de las series de funciones. Por analogía con el caso de las series numéricas definiremos las series de funciones

Definición 1.4.13 *Dada una sucesión de funciones $(a_n(x))_n$, la expresión*

$$\sum_{k=1}^{\infty} a_k(x) = a_1(x) + a_2(x) + a_3(x) + \cdots + a_k(x) + \cdots$$

se denomina serie funcional infinita o serie de funciones. Las funciones

$$S_n(x) = \sum_{k=1}^{n} a_k(x) = a_1(x) + a_2(x) + a_3(x) + \cdots + a_n(x),$$

se denominan sumas parciales de la serie de funciones.

Obviamente las sumas parciales $S_n(x)$ son a su vez una sucesión de funciones por tanto podemos definir la convergencia puntual y uniforme de una serie de funciones.

Definición 1.4.14 *Diremos que la serie $\sum_{k=1}^{\infty} a_k(x)$ converge a $S(x)$, para cierto $x \in \mathbb{R}$ si la sucesión de sumas parciales $(S_n(x))_n$ tiene límite $S(x)$. Si por el contrario, la sucesión de sumas parciales no tiene límite, entonces diremos que la serie diverge.*

Definición 1.4.15 *Una serie de funciones $\sum_{k=1}^{\infty} a_k(x)$ converge puntualmente en $I \in \mathbb{X}$ si para todo $x \in I$ y todo $\varepsilon > 0$ existe un $N \equiv N(x, \varepsilon) \in \mathbb{N}$ tal que para todo $n > N$, $|S_n(x) - S(x)| < \varepsilon$ y lo denotaremos $S_n \to S$.*

Definición 1.4.16 *Una serie de funciones $\sum_{k=1}^{\infty} a_k(x)$ converge uniformemente a $S(x)$ en I si para todo $\varepsilon > 0$ existe un $N \equiv N(\varepsilon) \in \mathbb{N}$ tal que para todo $n > N$ y para todos los $x \in I$, $|S_n(x) - S(x)| < \varepsilon$ y lo denotaremos $S_n \rightrightarrows S$.*

Si definimos el resto de la serie

$$r_n(x) \equiv S(x) - S_n(x) = \sum_{k=n+1}^{\infty} a_k(x), \tag{1.4.1}$$

entonces, la serie converge uniformemente si y solo si $r_n(x) \rightrightarrows 0$, en caso contrario la convergencia no es uniforme. Una prueba sencilla es usar el Teorema 1.4.7 para la sucesión de funciones $S(x) - S_n(x)$. En efecto, puesto que $r_n(x) = S(x) - S_n(x)$, tenemos si $r_n(x) \rightrightarrows 0$, entonces existe una sucesión $(a_n)_n$ de términos positivos que tiende cero tal que $|r_n(x)| \leq a_n$, luego $|S(x) - S_n(x)| \leq a_n \to 0$. El recíproco es análogo.

Lo anterior nos conduce al siguiente

Teorema 1.4.17 *Si $S_n(x) \rightrightarrows S(x)$ en I, es decir, si la serie $\sum_{k=1}^{\infty} a_k(x)$ es uniformemente convergente en I, entonces, $a_n(x)$, el término general de la serie converge uniformemente a 0 en I, es decir, $a_n(x) \rightrightarrows 0$.*

Demostración: Como $S_n(x) \rightrightarrows S(x)$ en I, el criterio de Cauchy 1.4.10 nos dice que para todo $\varepsilon > 0$ existe un $N \in \mathbb{N}$ tal que, para todos $n > N$, $p \in \mathbb{N}$ y para todo $x \in I$ se tiene $|S_{n+1} - S_n| < \varepsilon$. Como $S_{n+1} - S_n = a_{n+1}(x)$, tenemos entonces que para todo $\varepsilon > 0$ existe un $N \in \mathbb{N}$ tal que, para todo $n > N$, y para todo $x \in I$, $|a_{n+1}| < \varepsilon$, es decir, $a_n \rightrightarrows 0$. ■

Esta condición es necesaria pero no suficiente.

Ejemplo 1.4.18 *Estudiar la convergencia de la serie* $\sum_{n=1}^{\infty} x^n$ *en* $I = [0,1)$.

Como $x^n \not\rightrightarrows 0$ en $[0,1)$, entonces la serie no converge uniformemente en $[0,1)$. ■

Un criterio muy útil para establecer la convergencia uniforme de una serie de funciones es el siguiente Criterio de Weierstrass:

Teorema 1.4.19 *Sea* $(a_n)_{n=0}^{\infty}$ *una sucesión de funciones definidas en* $I \subset \mathbb{R}$*, y* $(b_n)_{n=0}^{\infty}$ *una sucesión de números reales no negativos, tales que* $|a_n(x)| \leq b_n$ *para todo* $n \in \mathbb{N}$ *y* $x \in I$ *y cuya serie* $\sum_{n=1}^{\infty} b_n$ *es convergente. Entonces la serie de funciones* $\sum_{n=1}^{\infty} a_n(x)$ *es uniformemente convergente en* I*.*

Demostración: Usando en criterio de comparación de Weierstrass se sigue que la serie $\sum_{n=0}^{\infty} a_n(x)$ converge para cada $x \in I$. Para probar la convergencia uniforme basta probar que $r_n(x) \rightrightarrows 0$ en I. Como la serie $(b_n)_{n=0}^{\infty}$ converge, entonces $B_n = \sum_{k=1}^{n} b_k$ tiende a $B = \sum_{n=1}^{\infty} b_n$, si $n \to \infty$, o sea, la diferencia $B - B_n = \sum_{k=n+1}^{\infty} b_k$, tiende a cero si $n \to \infty$. Por tanto,

$$|S_n(x) - S(x)| = |r_n(x)| = \left|\sum_{k=n+1}^{\infty} a_k(x)\right| \leq \sum_{k=n+1}^{\infty} |a_k(x)| \leq \sum_{k=n+1}^{\infty} b_k,$$

de donde tenemos

$$\sup_{x\in I} |S_n(x) - S(x)| \leq \sum_{k=n+1}^{\infty} b_k = B - B_n \to 0, \text{ cuando } n \to \infty,$$

luego $S_n(x) \rightrightarrows S(x)$. ■

Ejercicio 1.4.20 *Sea la serie* $\sum_{n=0}^{\infty} x^n$*. La serie converge uniformemente en el intervalo* $x \in [-q,q]$*, para todo* $q \in (0,1)$*.*

Efectivamente, como $x^n \leq q^n$ y la serie $\sum_{n=0}^{\infty} q^n$ converge, entonces el criterio de Weierstrass nos asegura la convergencia. ■

1.4.1. Propiedades de las sucesiones y series de funciones

En este apartado vamos a estudiar en qué condiciones las propiedades de f_n se traspasan a la función límite. Es decir, responderemos a la pregunta formulada al principio: ¿hereda $f(x)$ todas las propiedades de las $f_n(x)$?

Teorema 1.4.21 (Sobre la conmutatividad del límite de una sucesión de funciones) *Sea $(f_n(x))_n$ una sucesión de funciones definidas en $I \subset \mathbb{R}$ tal que para cada $n \in \mathbb{N}$ existe el límite $\lim_{x\to x_0} f_n(x) = l_n$, $x_0 \in I$. Si $f_n(x) \rightrightarrows f(x)$ en I, entonces*

$$\lim_{n\to\infty}\left(\lim_{x\to x_0} f_n(x)\right) = \lim_{x\to x_0}\left(\lim_{n\to\infty} f_n(x)\right)$$

Demostración: Sea $f_n(x) \rightrightarrows f(x)$ en I. Entonces, para todo $\varepsilon > 0$ existe un $N \in \mathbb{N}$ tal que para todos $n > N$, $p \in \mathbb{N}$ y $x \in I$, $|f_{n+p}(x) - f_n(x)| < \varepsilon/3$. Como existe el $\lim_{x\to x_0} f_n(x) = l_n$, entonces, tomando límites $x \to x_0$ en la desigualdad $|f_{n+p}(x) - f_n(x)| < \varepsilon/3$, tenemos $|l_{n+p} - l_n| \le \varepsilon/3$, luego la sucesión $(l_n)_n$ es de Cauchy y, por tanto, existe el límite cuando $n \to \infty$ que denotaremos por l. Probemos entonces que $\lim_{x\to x_0} f(x) = l$. Para ello, usamos que

$$|f(x) - l| \le |f(x) - f_n(x)| + |f_n(x) - l_n| + |l_n - l|.$$

Comprobemos que en las condiciones del teorema todos los sumandos anteriores se pueden hacer tan pequeños como se quiera. Como $f_n(x) \rightrightarrows f(x)$, entonces, cualquiera sea el $x \in I$, escogiendo $n > N$ tenemos $|f(x) - f_n(x)| \le \varepsilon/3$. Ahora bien, como $\lim_{x\to x_0} f_n(x) = l_n$, entonces, si escogemos un $\delta > 0$ suficientemente pequeño, entonces para todo $x \in (x_0 - \delta, x_0 + \delta)$, $|f_n(x) - l_n| < \varepsilon/3$. Finalmente como $l_n \to l$ si $n \to \infty$, escogiendo n suficientemente grande, digamos mayor que cierto $M \in \mathbb{N}$, $|l_n - l| < \varepsilon/3$, luego, para todo $\varepsilon > 0$, escogiendo $n > \max(N, M)$ y $x \in (x_0 - \delta, x_0 + \delta)$, tenemos $|f(x) - l| < \varepsilon$. ■

Como corolario trivial tenemos el siguiente

Teorema 1.4.22 (Continuidad de una sucesión de funciones) *Sea $(f_n)_{n=0}^{\infty}$ una sucesión de funciones definidas en $I \subset \mathbb{R}$. Si f_n es continua en I para todo $n \in \mathbb{N}$, y f_n converge uniformemente a f en I entonces f es continua en I.*

Demostración: Como $f_n(x)$ es continua en $x = x_0$, entonces existe el límite $\lim\limits_{x\to x_0} f_n(x) = f_n(x_0)$ y el teorema anterior nos dice que

$$\lim_{n\to\infty}\left(\lim_{x\to x_0} f_n(x)\right) = \lim_{x\to x_0}\left(\lim_{n\to\infty} f_n(x)\right) \Rightarrow \lim_{x\to x_0} f(x) = \lim_{n\to\infty} f_n(x_0) = f(x_0).$$

Como lo anterior es cierto para todo $x_0 \in I$, entonces $f(x)$ es continua en I. ■

El inverso del teorema anterior es falso en general, es decir, que si $f_n(x) \to f(x)$ en I, y tanto $f_n(x)$ como $f(x)$ son continuas en I, ello no implica $f_n(x) \rightrightarrows f(x)$. Por ejemplo, sea la sucesión de funciones

$$f_n(x) = \begin{cases} \operatorname{sen} nx, & 0 \leq x \leq \frac{\pi}{n} \\ 0, & \frac{\pi}{n} \leq x \leq \pi \end{cases}, \quad \lim_{n\to\infty} f_n(x) = 0 = f(x), \quad \forall x \in [0,\pi].$$

Obviamente tanto $f_n(x)$ como $f(x)$ son continuas en $[0,\pi]$. Ahora bien,

$$\sup_{x\in[0,\pi]} |f_n(x) - f(x)| = \sup_{x\in[0,\pi]} |\operatorname{sen} nx| = 1 \not\to 0,$$

por tanto $f_n(x) \not\rightrightarrows 0$.

No obstante si imponemos que la sucesión $(f_n(x))_n$ sea monótona sí que se tiene el recíproco.

Teorema 1.4.23 (Dini) *Sea $(f_n(x))_n$ una sucesión de funciones continuas que tienden monótonamente (para cada x) a $f(x)$ en el intervalo $[a,b]$ (cerrado y acotado). Entonces $f(x)$ es continua si y solo si $f_n(x) \rightrightarrows f(x)$ en $[a,b]$.*

La demostración se deja como ejercicio.

Los teoremas anteriores 1.4.21 y 1.4.22 se pueden reescribir para las series de funciones. En efecto si tomamos como sucesión de funciones $(f_n(x))_n$ la sucesión de sumas parciales de una serie $\sum_{n=1}^{\infty} a_n(x)$, tenemos el siguiente teorema

Teorema 1.4.24 (Conmutatividad del límite y la suma para una serie de funciones) *Sea $\sum_{n=1}^{\infty} a_n(x)$ una serie de funciones uniformemente convergente en I tales que $\lim\limits_{x\to x_0} a_n(x) = a_n$, entonces*

$$\lim_{x\to x_0}\left(\sum_{n=1}^{\infty} a_n(x)\right) = \sum_{n=1}^{\infty}\left(\lim_{x\to x_0} a_n(x)\right) = \sum_{n=1}^{\infty} a_n.$$

Si además, las funciones $a_n(x)$ son continuas, entonces la suma $S(x)$ de la serie es una función continua.

Teorema 1.4.25 (Integrabilidad de una sucesión de funciones) *Sea $(f_n)_{n=0}^{\infty}$ una sucesión de funciones definidas en $[a,b] \subset \mathbb{R}$ e integrables en $[a,b]$ y sea f_n uniformemente convergente a f en $[a,b]$. Entonces*

$$\lim_{n\to\infty} \int_a^b f_n(x)dx = \int_a^b \left[\lim_{n\to\infty} f_n(x)\right] dx = \int_a^b f(x)dx.$$

Demostración: Ante todo hay que demostrar que $f(x)$ es integrable. Por sencillez asumiremos que las funciones f_n son continuas entonces por el Teorema 1.4.22 se sigue que la función límite f también es continua y, por tanto, integrable y se puede simplificar la demostración.[5] Como $f_n(x) \rightrightarrows f(x)$ en $[a,b]$, entonces para n suficientemente grande, y para todo $x \in [a,b]$

$$|f(x) - f_n(x)| \le \frac{\varepsilon}{2(b-a)}$$

de donde se sigue que

$$\left|\int_a^b f(x)dx - \int_a^b f_n(x)dx\right| \le \int_a^b |f(x)-f_n(x)|dx \le \int_a^b \frac{\varepsilon}{2(b-a)}dx = \frac{\varepsilon}{2} < \varepsilon,$$

y, por tanto,

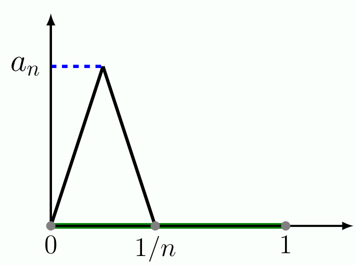
$$\lim_{n\to\infty} \int_a^b f_n(x)dx = \int_a^b f(x)dx = \int_a^b \left[\lim_{n\to\infty} f_n(x)\right] dx.$$

∎

El siguiente ejemplo pone de manifiesto la importancia de las condiciones del teorema que además son **suficientes**.

Ejemplo 1.4.26 Sea la sucesión de funciones $(f_n)_n$, $f_n : [0,1] \mapsto \mathbb{R}$, definida por

$$f_n(x) = \begin{cases} 2na_n x, & 0 \le x \le \frac{1}{2n}, \\ -2na_n(x - \frac{1}{n}), & \frac{1}{2n} \le x \le \frac{1}{n}, \\ 0, & \frac{1}{n} \le x \le 1. \end{cases}$$



[5] El caso general, para la integral de Riemann, esta desarrollado al final del capítulo.

Es evidente que las funciones $f_n(x)$ son continuas en $[0,1]$, y por tanto integrables en $[0,1]$ y que cualquiera sea la sucesión $(a_n)_n$ de sus máximos $\lim_{n\to\infty} f_n(x) = 0$. Como $\sup_{x\in[0,1]} |f_n(x)| = a_n$, si $\lim_{n\to\infty} a_n = 0$, entonces $f_n(x) \rightrightarrows 0$ en $[0,1]$ y

$$\lim_{n\to\infty} \int_0^1 f_n(x)dx = \int_0^1 f(x)dx = 0,$$

lo cual es fácil de comprobar mediante un cálculo directo pues el área de la figura $A_n = \frac{1}{2}\frac{1}{n}a_n \to 0$.

Ahora bien, podemos escoger a_n de forma que $A_n = \frac{1}{2}\frac{1}{n}a_n \to 0$, con $a_n \not\to 0$ –por ejemplo, $a_n = 1$, o incluso $a_n = \sqrt{n} \to \infty$–. En este caso también tenemos

$$\lim_{n\to\infty} \int_0^1 f_n(x)dx = \int_0^1 f(x)dx = 0.$$

Finalmente, si escogemos a_n de forma que el área $A_n = \frac{1}{2}\frac{1}{n}a_n \not\to 0$, –por ejemplo, $a_n = n$ o $a_n = n^2$, además en el primer caso el límite $\lim_{n\to\infty} \int_0^1 f_n(x)dx$ existe y en el segundo no (vale $+\infty$)– entonces $\lim_{n\to\infty} \int_0^1 f_n(x)dx \neq \int_0^1 f(x)dx = 0$.

En el caso de las series tenemos el siguiente resultado:

Teorema 1.4.27 ***(Integrabilidad de una serie de funciones)*** *Sea $(f_n)_{n=0}^\infty$ una sucesión de funciones definidas en $[a,b] \subset \mathbb{R}$ e integrables en $[a,b]$ tales que la serie $\sum_{n=1}^\infty a_n(x)$ es uniformemente convergente en I. Entonces su suma $S(x)$ es integrable y además*

$$\sum_{n=1}^\infty \int_a^b f_n(x)dx = \int_a^b \sum_{n=1}^\infty f_n(x)dx = \int_a^b S(x)dx.$$

Teorema 1.4.28 ***(Derivabilidad de una sucesión de funciones)*** *Sea $(f_n)_{n=0}^\infty$ una sucesión de funciones definidas en un intervalo acotado $I \subset \mathbb{R}$ y derivables en I tales que para cierto $x_0 \in I$ se tiene $\lim_{n\to\infty} f_n(x_0) = l$ y además la sucesión de funciones $(f_n')_{n=0}^\infty$ converja uniformemente a g en I. Entonces la sucesión $(f_n)_{n=0}^\infty$ converge uniformemente a cierta función f en I y además $f'(x) = g(x)$ para todo $x \in I$, es decir, la sucesión se puede* *derivar término a término, i.e.,*

$$\left(\lim_{n\to\infty} f_n(x)\right)' = \lim_{n\to\infty} f_n'(x).$$

Demostración: Definamos la sucesión de funciones $(\varphi_n(x))_n$, definidas por

$$\varphi_n(x) = \begin{cases} \frac{f_n(x)-f_n(x_0)}{x-x_0}, & x \neq x_0 \\ \\ f_n'(x_0), & x = x_0 \end{cases}$$

Las funciones $\varphi_n(x)$ son continuas en I –pues las funciones $f_n(x)$ son derivables en I por condición del teorema–.

Además $\varphi_n(x) \rightrightarrows \varphi(x)$ en I. En efecto, aplicando el criterio de Cauchy para la convergencia uniforme tenemos

$$\begin{aligned} |\varphi_{n+p}(x) - \varphi_n(x)| &= \left| \frac{(f_{n+p}(x) - f_n(x)) - (f_{n+p}(x_0) - f_n(x_0))}{x - x_0} \right| \\ &= \left| f_{n+p}'(\xi) - f_n'(\xi) \right|, \end{aligned}$$

donde $\xi \in (x, x_0)$. Para obtener la última igualdad hemos aplicado el Teorema del valor medio de Lagrange a la función $F(x) = f_{n+p}(x) - f_n(x)$, es decir, $F(x) - F(x_0) = F'(\xi)(x - x_0)$. Además, por definición tenemos $\varphi_{n+p}(x_0) - \varphi_n(x_0) = f_{n+p}'(x_0) - f_n'(x_0)$, luego $|\varphi_{n+p}(x) - \varphi_n(x)| = \left|f_{n+p}'(\xi) - f_n'(\xi)\right|$ para todo $x \in I$. Ahora bien, como $f_n'(x) \rightrightarrows g(x)$ en I, entonces, para todo $\varepsilon > 0$ existe un $N \in \mathbb{N}$ tal que para todo $n > N$, $p \in \mathbb{N}$ y para todo $x \in I$, $\left|f_{n+p}'(\xi) - f_n'(\xi)\right| < \varepsilon$, entonces

$$|\varphi_{n+p}(x) - \varphi_n(x)| = \left|f_{n+p}'(\xi) - f_n'(\xi)\right| < \varepsilon$$

de donde deducimos que $\varphi_n(x) \rightrightarrows \varphi(x)$ en I. Probemos ahora que $f_n(x) \rightrightarrows f(x)$ en I. Para ello usamos las identidades $f_{n+p}(x) = f_{n+p}(x_0) + (x - x_0)\varphi_{n+p}(x)$, y $f_n(x) = f_n(x_0) + (x - x_0)\varphi_n(x)$. Entonces,

$$\begin{aligned} |f_{n+p}(x) - f_n(x)| = &\ |f_{n+p}(x_0) - f_n(x_0) + (x - x_0)[\varphi_{n+p}(x) - \varphi_n(x)]| \\ \le &\ |f_{n+p}(x_0) - f_n(x_0)| + |x - x_0|\,|\varphi_{n+p}(x) - \varphi_n(x)| \end{aligned}$$

Ahora bien, como $f_n(x_0) \to f(x_0)$, entonces, para todo $\varepsilon > 0$ existe un $N_1 \in \mathbb{N}$ tal que para todo $n > N_1$, y $p \in \mathbb{N}$, $|f_{n+p}(x_0) - f_n(x_0)| < \varepsilon/2$. Además, $\varphi_n(x) \rightrightarrows \varphi(x)$ en I, por tanto, para todo $\varepsilon > 0$ existe un $N_2 \in \mathbb{N}$

tal que para todo $n > N$, $p \in \mathbb{N}$ y para todo $x \in I$, $|\varphi_{n+p}(x) - \varphi_n(x)| < \varepsilon/2(b-a)$, es decir, para todo $\varepsilon > 0$ existe un $N \in \mathbb{N}$, $N = \text{máx}(N_1, N_2)$, tal que para todo $n > N$, $p \in \mathbb{N}$ y para todo $x \in I$, $|f_{n+p}(x) - f_n(x)| < \varepsilon$, o sea, $f_n(x) \rightrightarrows f(x)$ en I.

Para culminar nuestra prueba nos resta demostrar que la función límite $f(x)$ es derivable en I y que $f'(x) = g(x)$ para todo $x \in I$, es decir, que $f'(x) = \text{lím}_{n\to\infty} f_n'(x)$.

Ante todo, notemos que puesto que $f_n(x) \rightrightarrows f(x)$ en I, entonces la condición $\text{lím}_{n\to\infty} f_n(x_0) = f(x_0)$ se cumple para cualquiera sea $x_0 \in I$ y no solo para el x_0 prefijado en el enunciado del teorema. Por tanto, si probamos que $f(x)$ es derivable en cierto $\zeta \in I$ –en particular x_0– y que $f_n'(\zeta) \to f'(\zeta) = g(\zeta)$, entonces $f'(x) = g(x)$ en todo I. Escojamos por tanto un ζ cualquiera y redefinamos las funciones

$$\varphi_n(x) = \begin{cases} \dfrac{f_n(x) - f_n(\zeta)}{x - \zeta}, & x \neq \zeta \\ f_n'(\zeta), & x = \zeta \end{cases}$$

Como antes, las funciones $\varphi_n(x)$ son continuas en I y $\varphi_n(x) \rightrightarrows \varphi(x)$ en I. Luego

$$\varphi(x) = \lim_{n\to\infty} \varphi_n(x) = \lim_{n\to\infty} \frac{f_n(x) - f_n(\zeta)}{x - \zeta} = \frac{f(x) - f(\zeta)}{x - \zeta}.$$

Lo anterior junto con el hecho de que todas las funciones $\varphi_n(x)$ son continuas en I, y que $\varphi_n(x) \rightrightarrows \varphi(x)$ en I, nos permiten utilizar el Teorema sobre la continuidad de una sucesión de funciones 1.4.22 para afirmar que $\varphi(x)$ es continua en I así como el Teorema sobre el límite 1.4.21 que nos da

$$\begin{aligned} \lim_{x\to\zeta} \frac{f(x) - f(\zeta)}{x - \zeta} &= \lim_{x\to\zeta} \left(\lim_{n\to\infty} \frac{f_n(x) - f_n(\zeta)}{x - \zeta} \right) = \lim_{n\to\infty} \left(\lim_{x\to\zeta} \varphi_n(x) \right) \\ &= \lim_{n\to\infty} \left(\lim_{x\to\zeta} \frac{f_n(x) - f_n(\zeta)}{x - \zeta} \right) = \lim_{n\to\infty} f_n'(\zeta). \end{aligned}$$

Esta igualdad nos asegura que existe la derivada de $f(x)$ para todo $x \in I$, y además que

$$f'(\zeta) = \lim_{x\to\zeta} \varphi(x) = \lim_{n\to\infty} f_n'(\zeta), \qquad \forall \zeta \in I,$$

que es que se quería demostrar. ■

Del teorema anterior se sigue *fácilmente* el siguiente teorema para las series de funciones:

Teorema 1.4.29 *Sea $S(x) = \sum_{n=1}^{\infty} a_n(x)$ una serie de funciones que converge al menos en un punto $x_0 \in I$, cuyos términos $a_n(x)$ son derivables en todo I, y cuya serie de las derivadas de $a_n(x)$, $\sum_{n=1}^{\infty} a_n'(x)$ converge uniformemente en I. Entonces la serie $S(x) = \sum_{n=1}^{\infty} a_n(x)$ converge uniformemente en I y además se cumple que*

$$S'(x) = \left(\sum_{n=1}^{\infty} a_n(x) \right)' = \sum_{n=1}^{\infty} \left(a_n'(x) \right),$$

o sea, es derivable término a término.

Ejemplo 1.4.30 Consideremos la serie $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$, que converge en el intervalo $[-R, R]$, cualquiera sea $R > 0$. Obviamente la serie $\sum_{n=0}^{\infty} \left(\frac{x^n}{n!}\right)$ $= \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!}$ converge uniformemente en $[-R, R]$. Entonces,

$$(e^x)' = \left(\sum_{n=0}^{\infty} \frac{x^n}{n!} \right)' dx = \sum_{n=0}^{\infty} \left(\frac{x^n}{n!} \right)' = \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} = \sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x,$$

Es decir, $(e^x)' = e^x$.

Ejemplo 1.4.31 Sea la función

$$\psi(x) = \begin{cases} x^2 \cos \dfrac{1}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

Dicha función es derivable en $\mathbb{R}$ y su derivada, definida por,

$$\psi'(x) = \begin{cases} 2x \cos \dfrac{1}{x} + \operatorname{sen} \dfrac{1}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

es continua para todo $x \in \mathbb{R}$ excepto el cero. Numeremos todos los puntos racionales del intervalo $[0, 1]$ mediante la sucesión $(x_n)_n$, lo cual es posible pues $\mathbb{Q}$ es un conjunto numerable (ser la sección 1.5.1). Definamos la sucesión $a_n(x) = \frac{1}{n^2}\psi(x - x_n)$. Entonces, $a_n'(x) = \frac{1}{n^2}\psi'(x - x_n)$ es discontinua en un único punto $x = x_n$, además

$$|a_n(x)| \le \frac{|x-x_n|^2}{n^2} \le \frac{1}{n^2}, \qquad |a_n'(x)| \le \frac{1+2|x-x_n|}{n^2} \le \frac{3}{n^2},$$

ambas series $\sum_{n=1}^\infty a_n(x)$ y $\sum_{n=1}^\infty a_n'(x)$ convergen uniformemente en $[0,1]$ pues ambas se pueden mayorar por la serie convergente $\sum_{n=1}^\infty 3/n^2$. Por tanto el teorema 1.4.29 nos dice que

$$S'(x) = \left(\sum_{n=1}^\infty a_n(x)\right)' = \sum_{n=1}^\infty (a_n'(x)),$$

es decir, que la función $S(x)$ es derivable en todo $[0,1]$ y su derivada es discontinua en cada uno de los puntos racionales del intervalo.

Anexo: demostración del Teorema 1.4.25 para funciones integrables

En el caso que las funciones f_n de la sucesión de funciones sean integrables Riemann podemos hacer lo siguiente: Como $f_n(x) \rightrightarrows f(x)$ en $[a,b]$, entonces para todo $\varepsilon > 0$ existe un $N \in \mathbb{N}$ tal que para todo $n > N$ y para todo $x \in [a,b]$, $|f_n(x) - f(x)| < \frac{\varepsilon}{4(b-a)}$. Por tanto, para todo $n > N$ tendremos

$$\begin{aligned} |f(x)-f(y)| = & \ |f(x)-f_n(x)+f_n(x)-f_n(y)+f_n(y)-f(y)| \\ \le & \ |f(x)-f_n(x)|+|f_n(x)-f_n(y)|+|f_n(y)-f(y)|. \end{aligned}$$

Obviamente, el primer y último sumando se pueden hacer más pequeños que $\frac{\varepsilon}{4(b-a)}$ pues $f_n(x) \rightrightarrows f(x)$ en $[a,b]$, i.e.

$$|f(x)-f(y)| \le |f_n(x)-f_n(y)| + \frac{\varepsilon}{2(b-a)}, \qquad \forall x,y \in [a,b].$$

Dividamos el intervalo $[a,b]$ en intervalos más pequeños $\Delta_i = [x_i, x_{i+1}]$ (una partición de $[a,b]$). Usando la última desigualdad tenemos

$$\sup_{x,y\in\Delta_i} |f(x)-f(y)| \le \sup_{x,y\in\Delta_i} |f_n(x)-f_n(y)| + \frac{\varepsilon}{2(b-a)}.$$

Entonces,

$$\sum_{i=1}^n \sup_{x,y\in\Delta_i} |f(x)-f(y)|\Delta x_i \le \sum_{i=1}^n \sup_{x,y\in\Delta_i} |f_n(x)-f_n(y)|\Delta x_i + \frac{\varepsilon}{2}.$$

Pero $f_n(x) \in \mathcal{R}_{[a,b]}$, luego, por el criterio de Riemann, existe una partición P_n de $[a,b]$, tal que $\sum_{i=1}^{n} \sup_{x,y\in\Delta_i} |f_n(x) - f_n(y)|\Delta x_i < \varepsilon/2$, por tanto, para dicha partición $S_f(P_n) - s_f(P_n) < \varepsilon$, luego $f(x) \in \mathcal{R}_{[a,b]}$. El resto de la prueba es análogo.

1.5. Sobre conjuntos infinitos

En este sección recordaremos algunas cuestiones sobre conjuntos infinitos haciendo hincapié en el caso de los números reales.

1.5.1. Conjuntos numerables

Sean dos conjuntos A y B cualesquiera. Por ejemplo, digamos que A es el conjunto de los números primos menores que un número dado (digamos 10), B el de los vértices de un dodecágono (12 lados). A y B son finitos, es decir, A y B están constituidos por un número finito de elementos. Luego, una forma natural para compararlos puede ser sencillamente usando el número de sus elementos. Obviamente A tiene cuatro elementos, $2, 3, 5, 7$ y B doce, así que B será más grande que A. Sea ahora C el conjunto de los lados de un cuadrado. Entonces A y C tienen el mismo número de elementos de forma que podemos hacer corresponder a cada uno de los elementos de C el correspondiente elemento de A y viceversa. Es decir, existe una correspondencia biunívoca entre los elementos de A y C. Obviamente tal correspondencia es imposible de encontrar entre los elementos de A y B o B y C (¿por qué?).

¿Qué ocurre si ahora A Y B tienen infinitos elementos, es decir, si son conjuntos infinitos? Ejemplo de tales conjuntos son conocidos: $\mathbb{N}$, $\mathbb{Q}$, $\mathbb{R}$, etc.

En este caso el primer método de comparar conjuntos no nos sirve pues no podemos "contar" los elementos ya que éstos son infinitos así que solo nos queda el segundo de ellos: intentar encontrar una correspondencia biunívoca entre conjuntos.

Recordemos que $\mathbb{N} = \{1, 2, 3, \dots, n, n+1, \dots\}$ es el conjunto de los números naturales que es el conjunto infinito más sencillo.

Definición 1.5.1 *Un conjunto A es numerable si*

1. *A es vacío.*

2. *A tiene un número finito de elementos.*

3. *Si A tiene un número infinito de elementos, estos se pueden poner en correspondencia biunívoca con $\mathbb{N}$.*

Veamos algunos ejemplos.

Por ejemplo, P, el conjunto de todos los números pares es numerable. En efecto, cualquiera sea $n \in \mathbb{N}$ existe un único $p \in P$ tal que $p = 2n$ y viceversa, cualquiera sea $p \in P$, existe un único $n \in \mathbb{N}$ tal que $n = p/2$, es decir existe una correspondencia biunívoca entre ambos conjuntos, o lo que es lo mismo *hay tantos números pares como naturales*.

Existe una forma muy sencilla de probar lo anterior. Escribamos todos los números naturales en una tabla

1	2	3	4	5	6	$\cdots$	$2n-1$	$2n$	$2n+1$	$2n+2$	$\cdots$

y eliminemos todos los impares y "contamos" los números resultantes. Así tenemos

$\cancel{1}$	2	$\cancel{3}$	4	$\cancel{5}$	6	$\cdots$	$2\cancel{n}-1$	$2n$	$2\cancel{n}+1$	$2n+2$	$\cdots$
	1		2		3	$\cdots$		n		$n+1$	$\cdots$

es decir, P es numerable, los números pares se pueden "contar", existe una correspondencia biyectiva entre los números pares y los naturales.

De forma similar se puede proceder para probar que el conjunto de los números impares I es numerable. Lo mismo ocurre con los enteros

0	1	-1	2	-2	3	-3	4	-4	5	-5	$\cdots$
1	2	3	4	5	6	7	8	9	10	11	$\cdots$

Veamos un ejemplo más complicado. Sea A el conjunto de puntos del plano de la forma (n, m), $n, m \in \mathbb{N}$. Entonces A es numerable.

Para comprobarlo vamos a "*contar*" todos los pares anteriores, es decir, vamos a construir una función biyectiva que a cada par le haga corresponder un único valor de $n \in \mathbb{N}$. Lo haremos siguiendo el siguiente esquema:

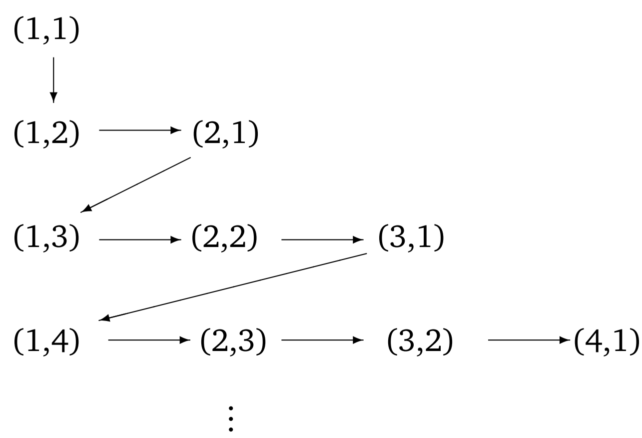


Figura 1.8: Contando el conjunto $A = \{(n, m),\ n, m \in \mathbb{N}\}$.

ordenamos los pares por filas según la suma $n + m = 2, 3, 4, \dots$ y cada fila la ordenamos de menor a mayor según el primer valor n y así tenemos una correspondencia biunívoca tal y como se ve en la figura 1.5.1

De forma análoga se puede probar la siguiente

Proposición 1.5.2 *El conjunto $B = \{(p, q)\,;\, p, q \in \mathbb{Z}\}$ es numerable, es decir, que se puede construir una "ordenación" de B similar a la del ejemplo anterior.*

Una consecuencia de la proposición anterior es la numerabilidad de $\mathbb{Q}$. En efecto, sea un número r racional cualquiera, $r \in \mathbb{Q}$. Entonces r se puede escribir como $r = p/q$, con $p \in \mathbb{Z}$ y $q \in \mathbb{N}$. Así que a cada número racional le podemos hacer corresponder el par (p, q) con $p \in \mathbb{Z}$ y $q \in \mathbb{N}$ que es un subconjunto de los pares B definidos en la Proposición 1.5.2 (¿por qué?), así que $\mathbb{Q}$ es numerable.

1.5.2. Propiedades de los conjuntos numerables

Veamos algunas propiedades de los conjuntos numerables. La siguiente proposición es sencilla de probar:

Proposición 1.5.3 1. *Si A es un conjunto numerable y B es un subconjunto de A, entonces B es numerable.*

2. *La unión de un conjunto numerable y uno finito es numerable.*

3. *La unión de dos conjuntos numerables es numerable, y por tanto la unión de un número finito de conjuntos numerables es numerable.*

4. *La unión de un número numerable de conjuntos numerables es numerable.*

Para probar el último apartado podemos escribir cada uno de los conjuntos numerables de la forma $A_1 = \{a_{1,1}, a_{1,2}, \ldots a_{1,n}, \ldots\}$, $A_2 = \{a_{2,1}, a_{2,2}, \ldots a_{2,n}, \ldots\}$,..., $A_m = \{a_{m,1}, a_{m,2}, \ldots a_{m,n}, \ldots\}$, y escribirlos como pares (m, n), $n, m \in \mathbb{N}$. Dichos pares son los representados en la figura 1.5.1 y, como ya vimos, son numerables. Luego tenemos 4. Está claro que 2 y 3 son casos particulares de 4. Para probar 1, basta usar la reducción al absurdo.

Veamos como usando los apartados anteriores podemos probar los ejercicios que antes hemos resuelto.

Sea $\mathbb{N}^-$ el conjunto de los números enteros negativos y sea $\mathbb{Z}_k$, $k \in \mathbb{N}$, los conjuntos de los números de la forma n/k, con $n \in \mathbb{Z}$, es decir, $\mathbb{Z}_k = \{n/k\,;\, n \in \mathbb{Z}\}$. Obviamente $\mathbb{Z} = \mathbb{N} \cup \{0\} \cup \mathbb{N}^-$, pero tanto $\mathbb{N}$ como $\mathbb{N}^-$ son numerables, así que $\mathbb{Z}$ es numerable. Además, por construcción $\mathbb{Z}_k$ es numerable. Además, $\mathbb{Q} = \mathbb{Z}_1 \cup \mathbb{Z}_2 \cup \cdots \cup \mathbb{Z}_k \cup \cdots$, o sea $\mathbb{Q}$ es la unión de una cantidad numerable de conjuntos numerables así que $\mathbb{Q}$ es numerable.

Definamos ahora el conjunto producto cartesiano de dos conjuntos. Sea A y B dos conjuntos cualesquiera. Definiremos el "producto cartesiano" de A y B que denotaremos por $A \otimes B$ al conjunto de todos los pares ordenados (a, b) con $a \in A$ y $b \in B$.

Está claro que si A y B son numerables el conjunto $A \otimes B$ es numerable. Para ello basta usar de forma inteligente las proposiciones 1.5.2 y 1.5.3.

Como consecuencia de lo anterior tenemos otra prueba de que $\mathbb{Q}$ es numerable pues, como hemos visto, a cada número racional le podemos hacer corresponder el par (p, q) con $p \in \mathbb{Z}$ y $q \in \mathbb{N}$, p, q primos entre si, o sea, $\mathbb{Q} = \mathbb{Z} \otimes \mathbb{N}$.

Usando la Proposición 1.5.3 se puede comprobar que el el conjunto $\mathbb{P}_n$ de todos los polinomios de grado a lo más n con coeficientes $a_0, \ldots a_n$ racionales:

$$\mathbb{P}_n = \{a_0 + a_1 x + a_2 x^2 + \cdots a_n x^n\,;\, a_0, a_1, a_2, \ldots a_n \in \mathbb{Q}\},$$

es numerable y, por tanto, el conjunto de todos los polinomios con coeficientes racionales

$$\mathbb{P} = \{a_0 + a_1 x + a_2 x^2 + \cdots a_n x^n\,;\, a_0, a_1, a_2, \ldots a_n \in \mathbb{Q},\ \forall n \in \mathbb{N}\},$$

también lo es pues $\mathbb{P} = \mathbb{P}_0 \cup \mathbb{P}_1 \cup \mathbb{P}_2 \cup \cdots$, y podemos usar el punto 4 de la Proposición 1.5.3.

Definición 1.5.4 *Sea $x \in \mathbb{R}$ tal que x es solución de la ecuación algebraica*

$$a_0 + a_1 x + a_2 x^2 + \cdots a_n x^n = 0, \quad , n \in \mathbb{N}, \quad a_0, a_1, a_2, \ldots a_n \in \mathbb{Q}.$$

Entonces se dice que x es un número algebraico. Si x no es solución de ninguna ecuación algebraica, entonces se dice trascendente.

Por ejemplo, 2 es un número algebraico ($x - 2 = 0$), $\sqrt{2}$ es un irracional algebraico pues es solución de $x^2 - 2 = 0$. Se puede probar[6] que e y π son trascendentes.

Nótese que, a diferencia de los números irracionales, ningún número racional puede ser trascendente (¿por qué?). Sea $\mathbb{A}$ el conjunto de todos los números algebraicos y $\mathbb{T}$ el de los trascendentes. Dado que un polinomio de grado n solo puede tener a lo más n raíces y que el conjunto $\mathbb{P}$ es numerable se sigue que el conjunto $\mathbb{A}$ es numerable.

Ejemplo 1.5.5 *Prueba que el conjunto de todos los subconjuntos finitos de un conjunto numerable A es numerable.*

Para comprobarlo usaremos la Proposición 1.5.3. Está claro que el conjunto A_1 de todos los subconjuntos de A con un solo elemento es numerable. En conjunto A_2 de todos los subconjuntos de A con dos elementos es numerable. En general, el conjunto A_n de todos los subconjuntos de A con $n \in \mathbb{N}$ elementos es numerable. De esta forma obtenemos un conjunto numerable $(A_1, A_2, \ldots)$ de conjuntos numerables. ■

Así, por ejemplo, el conjunto de todos los subconjuntos finitos de números racionales es numerable. Basta identificar $A = \mathbb{Q}$, A_1 con el conjuntos de todos los racionales q, A_2 con el conjunto de todos los pares de racionales (p, q), y así sucesivamente.

[6] El caso del número e fue probado por C. Hermite en 1873 y el de π por C.L. Lindemann en 1882.

Definición 1.5.6 *Dos conjuntos A y B se denominan equivalentes y se denota por $A \sim B$ si existe una correspondencia biunívoca entre sus elementos. En este caso el "número" de sus elementos es el mismo lo que se suele denotar por* $\text{card } A = \text{card } B$[a].

[a]En realidad el concepto de cardinal o potencia de un conjunto es ligeramente más complicado que el número de sus elementos, pero para nuestros objetivos esta definición es suficiente.

En fácil comprobar entonces que si $A \sim B$ entonces $B \sim A$ (¿por qué?) y que si $A \sim B$ y $B \sim C$, entonces $A \sim C$.

Por ejemplo, según hemos visto antes $\mathbb{N} \sim \mathbb{Z}$, $\mathbb{N} \sim \mathbb{Q}$, $\mathbb{N} \sim \mathbb{P}$, $\mathbb{N} \sim \mathbb{A}$, $\mathbb{Q} \sim \mathbb{A}$, etc.

Es evidente que el conjunto de los números racionales que pertenecen al conjunto $[0,1]$ es numerable pues son un subconjunto de $\mathbb{Q}$. La pregunta es ¿será también numerable el conjunto de los irracionales contenidos en $[0,1]$?, o, equivalentemente, ¿será numerable $[0,1]$?

Vamos a intentar responder a esta pregunta.

Supongamos que $[0,1] \sim \mathbb{N}$. Entonces ha de existir una sucesión de números reales $(x_n)_n$ tal que cualquiera sea $x \in [0,1]$, x coincidirá con al menos un miembro de la sucesión $(x_n)_n$ (en caso contrario $[0,1]$ no sería numerable, ¿por qué?).

Hagamos la siguiente construcción: sea $I_0 = [0,1]$. Escojamos un intervalo cerrado $I_1 \subset I_0$ que no contenga a x_1, o sea, $x_1 \notin I_1$. A continuación escojamos un intervalo cerrado $I_2 \subset I_1$ que no contenga a x_2, o sea, $x_2 \notin I_2$, y así sucesivamente. Entonces tenemos una sucesión de intervalos (cerrados) encajados tal que $I_{n+1} \subset I_n$ pero $x_{n+1} \notin I_{n+1}$. Pero, como $(I_n)_n$ es una sucesión de intervalos encajados, entonces, existe al menos un x que pertenece a todos los I_n (Teorema de los intervalos encajadas de Cantor), el cual, por construcción es distinto a todos los x_n, lo cual es una contradicción.

De lo anterior se deduce el Teorema de Cantor

Teorema 1.5.7 $\mathbb{R}$ *no es numerable. Es decir, hay más números reales que naturales, i.e.,* $\text{card } \mathbb{R} > \text{card } \mathbb{N}$.

Como corolario tenemos

Corolario 1.5.8 1. *$\mathbb{R} \neq \mathbb{Q}$, es decir, existe $\mathbb{I} = \mathbb{R} \setminus \mathbb{Q}$, el conjunto de los números irracionales, y además $\mathbb{I}$ no es numerable.*

2. *$\mathbb{A} \neq \mathbb{I}$, existe el conjunto de los números trascendentes $\mathbb{T}$, y $\mathbb{T}$ no es numerable.*

Resulta que la mayoría de los números irracionales que conocemos son algebraicos, por ejemplo, $\sqrt[n]{k}$ si $k \neq l^n$, $l \in \mathbb{N}$, $\sqrt{2} + \sqrt[3]{2}$ (prueba que este número es irracional algebraico), etc. Por el contrario se conocen pocos números trascendentes: π, e. No obstante resulta que éstos son la mayoría de todos los números ya que los números algebraicos (que incluyen, como hemos visto a los racionales) es un conjunto numerable, pero $\mathbb{R}$ no lo es y obviamente $\mathbb{R} = \mathbb{A} \cup \mathbb{T}$.

Automáticamente surge la pregunta. ¿Existirá algún conjunto infinito intermedio entre $\mathbb{N}$ y $\mathbb{R}$?, es decir, que tenga "más" elementos que $\mathbb{N}$ pero "menos" que $\mathbb{R}$?

Georg Cantor fue quien desarrolló la teoría moderna de conjuntos infinitos, a él debemos la notación y algunos de los resultados que hemos descrito aquí. Precisamente fue Cantor quien conjeturó que no existía dicho conjunto intermedio (hipótesis del continuo). Este fue el primero de los 23 famosos problemas que formuló Hilbert en 1900. La respuesta a este problema fue totalmente inesperada. Por un lado Kurt Gödel probó en 1940 que usando los axiomas de la teoría de conjuntos era imposible desmentir la hipótesis de Cantor. La respuesta definitiva la dio Paul Cohen en 1963 cuando demostró que bajo el mismo sistema de axiomas de la teoría de conjuntos era imposible probar la conjetura, o sea, la hipótesis del continuo se puede aceptar o no y eso no lleva a ninguna contradicción lógica dentro de la teoría de conjuntos.

1.6. Problemas

Problema 1.1 Prueba que si $\lim_{n\to\infty} z_n = 0$, entonces la sucesión de sus módulos converge y $\lim_{n\to\infty} |z_n| = 0$ pero la sucesión de sus argumentos ϕ_n puede no converger. ¿Qué ocurre en el caso cuando $\lim_{n\to\infty} z_n \neq 0$?

Problema 1.2 Calcula el $\displaystyle\lim_{n\to\infty} \frac{a^n}{1+a^n}$ en función de $a \in \mathbb{C}$.

Problema 1.3 *Suponiendo que* $\lim_{n\to\infty} z_n = z$*, prueba que* $\lim_{n\to\infty} \overline{z_n} = \overline{z}$*,* $\lim_{n\to\infty} \Re z_n = \Re z$*,* $\lim_{n\to\infty} \Im z_n = \Im z$ *y* $\lim_{n\to\infty} |z_n| = |z|$.

Problema 1.4 Prueba que toda sucesión $(z_n)_n$ de Cauchy es acotada.

Problema 1.5 Estudia el carácter de las siguientes series:

1. $\displaystyle\sum_{n=1}^{\infty} \frac{n^n}{3^n n!}, \quad \sum_{n=1}^{\infty} (\sqrt[n]{n}-1)^n, \quad \sum_{n=0}^{\infty} \frac{1}{\sqrt{2n^4+1}}, \quad \sum_{n=1}^{\infty} \log(1+4/n^2),$

2. $\displaystyle\sum_{n=1}^{\infty} \frac{|\operatorname{sen} n|}{n^2+n}, \quad \sum_{n=1}^{\infty} (-1)^n[\sqrt{n^2-1}-n], \quad \sum_{n=1}^{\infty} \frac{a^n+b^n}{a^n-b^n} x^n, \quad a,b,x \geq 0.$

Problema 1.6 Sea $\{u_n\}$ una sucesión de términos $u_n > 0$ tal que $\displaystyle\lim_{n\to\infty} u_n = 0$. Demuestra que las series $\sum_n u_n$ y $\sum_n \log(1+u_n)$ comparten el mismo carácter convergente o divergente. Aplica este resultado para determinar el carácter de la serie $\sum_{n=1}^{\infty} \log\left(1+\frac{2}{n^2+n+1}\right)$.

Problema 1.7 Determina el radio de convergencia de las series de potencias $\sum_{n=0}^{\infty} a_n z^n$ con a_n definido por:

a) $a_n = \log n$ b) $\displaystyle a_n = \frac{(-1)^n}{n}$, c) $\displaystyle a_n = \frac{[(-1)^n+3]^n}{n}$, d) $\displaystyle a_n = \frac{n^2}{a^2}$,

e) $\displaystyle a_n = \frac{n!}{n^n}$, f) $\displaystyle a_n = \frac{1}{n^n}$, g) $\displaystyle a_n = \frac{1}{n!}$, h) $\displaystyle a_n = \left(\operatorname{sen} \frac{n\pi}{4}\right)^n$.

Problema 1.8 Obtener las series de potencias de las siguientes funciones indicando la región de convergencia de las mismas:

a) $\displaystyle f(z) = \log\left(\frac{1+z}{1-z}\right)$, b) $f(z) = \arctan(z)$, c) $f(z) = \cos(\sqrt{z})$,

d) $\displaystyle f(z) = \frac{(z-2)^n}{3^n}$, e) $f(z) = \cosh(z)$, f) $f(z) = \log(a+bz)$, $(a \neq 0)$.

Problema 1.9 Estudia la convergencia uniforme de las sucesiones de funciones consideradas en el Ejemplo 1.4.4.

Problema 1.10 Estudia la convergencia puntual y uniforme de las siguientes sucesiones de funciones $\{f_n\}_{n\in\mathbb{N}}$:

$$f_n = \frac{1}{1+nx}, x \in [0,1], \quad f_n = xe^{-nx}, x \in [0,\infty), \quad f_n = \frac{2n^2x}{1+n^2x^2}, x \in \mathbb{R}.$$

Problema 1.11 Demuestra la Proposición 1.4.12

Problema 1.12 Demuestra el Teorema 1.4.17 que establece una condición necesaria de convergencia uniforme de una serie.

Problema 1.13 Sea $f_n(x) = \arctan(nx)$.

a) Estudiar la convergencia puntual de $\{f_n\}$. ¿Es uniformemente convergente en $\mathbb{R}$?.

b) Demostrar que converge uniformemente en $\{x \in \mathbb{R} : |x| \geq a > 0\}$.

c) Estudiar la convergencia uniforme de la serie de funciones

$$\sum_{n\geq 1} \frac{\arctan(nx)}{n^2+1}.$$

d) Demostrar que la función suma de la serie anterior es derivable en $\mathbb{R}\backslash\{0\}$.

Problema 1.14 Como vimos en el apartado 1.2.3, toda serie de potencias $\sum_{n=0}^{\infty} a_n z^n$, a_n, $z \in \mathbb{C}$, tiene un radio de convergencia $R \geq 0$ ó $R = +\infty$. Además, la serie converge absolutamente para todo z tal que $|z| < R$. Prueba que toda serie de potencias converge uniformemente en cualquier región Ω del plano complejo contenida en el círculo de radio $r < R$, i.e., $\Omega \subset \{z;\ |z| \leq r < R\}$.

Problema 1.15 Aplica los teoremas 1.4.24, 1.4.27 y 1.4.29 a las series de potencias y deduce los correspondientes teoremas. Compáralos con los del apartado §1.2.3. Discute cuando las series de potencias convergen uniformemente.

Capítulo 2

¿Qué es el análisis funcional?

> El arte de la Matemática consiste en encontrar ese caso especial que contiene todos los gérmenes de la generalidad.
>
> Atribuida a David Hilbert

¿Cómo comenzar un curso de Análisis funcional? Sin duda una posibilidad es con un poco de historia, pero para eso se puede consultar el interesante artículo de Berta Gamboa de Buen, *Historia del Análisis Funcional* publicado en la revista *Miscelánea Matemática de la Sociedad Matemática Mexicana*, volumen **28** (1999) páginas 17-39[1], Como bien cuenta su autora, el análisis funcional surge y se desarrolla en el siglo XX generado por la evolución del análisis matemático clásico, la física-matemática, las nuevas ideas del álgebra y la geometría y muy ligado a la topología. En otras palabras, en cierta forma el Análisis Funcional aúna las principales áreas de la matemática[2].

Nosotros, sin embargo, vamos a proceder de otra forma. Vamos a seguir la idea de Schechter en [22] y comenzar con un ejemplo relativamente sencillo que nos conducirá a redescubrir varios de los conceptos fundamentales del análisis clásico así como algunos otros del propio análisis funcional.

[1] https://miscelaneamatematica.org/ContenidoNumero/28

[2] Véase también el detallado estudio histórico de G. Birkhoff y E. Kreyszig titulado *The establishment of functional analysis* y publicado en *Historia Mathematica*, volumen **11** (1984) páginas 258-321, https://www.sciencedirect.com/science/article/pii/0315086084900363

2.1. Una ecuación diferencial lineal

Uno de los problemas típicos de un curso de ecuaciones diferenciales es la resolución del siguiente *problema de valores iniciales*

$$f''(x) + f(x) = g(x), \quad x \in [a,b], \quad f(a) = 1, \; f'(a) = 0, \tag{2.1.1}$$

donde por f' denotamos la derivada de f respecto a x. Además, por simplicidad asumiremos que g es una función continua en $[a,b]$ y que f, la solución buscada, es una función dos veces diferenciable con segunda derivada continua. Así, usaremos la notación $g \in C([a,b])$ y $f \in C^{(2)}([a,b])$, para denotar al conjuntos de funciones continuas en $[a,b]$ y las que son dos veces diferenciable con segunda derivada continua en $[a,b]$, respectivamente.

Se puede comprobar que la (una) solución de dicho problema de valores iniciales (2.1.1) viene dada por la expresión

$$f(x) = \cos(x-a) + \int_a^x \sin(x-t)g(t)dt, \qquad x \in [a,b]. \tag{2.1.2}$$

Comprobemos que, efectivamente, la función f anterior es solución de la ecuación (2.1.1) con los valores iniciales dados. Lo primero que es *obvio* es que $f(a) = 1$, pues la integral en (2.1.2) es cero. Calculemos ahora $f'(x)$. El primer sumando no tiene problemas, pero el segundo es algo más complicado pues la variable x aparece tanto en el límite de integración, como en la función a integrar. Este es un problema típico de la teoría de integrales dependientes de un parámetro. Así pues, para calcular la derivada de nuestra función f conviene recordar el siguiente teorema:[3]

Teorema: *Sea $h(t,x)$ una función continua tal que su derivada parcial respecto a x sea también continua en $D = \{(t,x) \in \mathbb{R}^2 \,|\, t \in [a,b], x \in [c,d]\}$. Entonces la función $H(x) = \int_a^b h(t,x)dt$ es continua y diferenciable en $[c,d]$ y*

$$H'(x) = \int_a^b \frac{\partial h}{\partial x}(t,x)dt.$$

Idea de la demostración: Para probar la continuidad se puede usar el Teorema de Heine 1.3.5 para funciones continuas en un intervalo cerrado

[3]Ver, por ejemplo, la sección §17.1.13 de [29].

y acotado que establece que si $f(x) := h(t,x)$ es continua en $[c,d]$ (t fijo), entonces f es uniformemente continua en $[c,d]$ es decir, para todo $\varepsilon > 0$ existe un $\delta > 0$ tal que para todos $x, x' \in [c,d]$, que satisfacen $|x-x'| < \delta$, se tiene $|h(t,x)-h(t,x')| < \varepsilon$. Entonces, para todos x, x' tales que $|x-x'| < \delta$,

$$|H(x) - H(x')| \le \int_a^b |h(t,x) - h(t,x')|dt \le \varepsilon(b-a),$$

i.e., $H(x) \to H(x')$ cuando $x \to x'$. Para probar la diferenciabilidad basta mostrar que para todo $x_0 \in [c,d]$

$$H(x_0+\delta) - H(x_0) - H'(x_0)\delta = o(\delta), \qquad \delta \in \mathbb{R},$$

donde $o(\delta)$ es tal que $o(\delta)/\delta \to 0$ cuando $\delta \to 0$.

Para ello supondremos, por simplicidad, que existe $\frac{\partial^2 h}{\partial x^2}(t,x)$ es continua en D y por tanto acotada. Entonces, usando el Teorema del valor medio en la variable x, y teniendo en cuenta que

$$\begin{aligned}|H(x_0+\delta) - H(x_0) - H'(x_0)\delta| &\le \int_a^b \left| h(t,x_0+\delta) - h(t,x_0) - \frac{\partial h}{\partial x}(t,x_0)\delta \right| dt \\ &= |\delta| \int_a^b \left| \frac{\partial h}{\partial x}(t,\xi) - \frac{\partial h}{\partial x}(t,x_0) \right| dt = |\delta|^2 \int_a^b \left| \frac{\partial^2 h}{\partial x^2}(t,\xi') \right| dt \le M|\delta|^2,\end{aligned}$$

donde $\xi, \xi' \in (c,d)$ y $M = (b-a)\max_{x\in[c,d]} \left|\frac{\partial^2 h}{\partial x^2}(t,x)\right|$.

El caso general lo dejamos como ejercicio al lector. ■

Sea $\alpha(x)$ una función diferenciable en $[c,d]$ cuya imagen esté contenida en $[a,b]$ y sea la función $\Phi(\alpha,x) = \int_a^\alpha h(t,x)dt$, $\alpha \in [a,b]$, $x \in [c,d]$. Usando el Teorema fundamental del cálculo así como el teorema anterior tenemos que

$$\frac{\partial \Phi}{\partial x}(\alpha,x) = \int_a^\alpha \frac{\partial h}{\partial x}(t,x)dt, \qquad \frac{\partial \Phi}{\partial \alpha}(\alpha,x) = h(\alpha,x).$$

Dado que ambas derivadas parciales son continuas, entonces $\Phi(\alpha,x)$ es diferenciable en $\alpha \in [a,b]$, $x \in [c,d]$. Definamos ahora la función compuesta $\Psi(x) = \Phi(\alpha(x),x)$, $x \in [c,d]$. Dicha función es diferenciable en $x \in [c,d]$ así que usando la regla de la cadena obtenemos

$$\frac{\partial \Psi}{\partial x} = \frac{\partial \Phi}{\partial \alpha}\frac{\partial \alpha}{\partial x} + \frac{\partial \Phi}{\partial x} = \alpha'(x)h(\alpha(x),x) + \int_a^{\alpha(x)} \frac{\partial h}{\partial x}(t,x)dt.$$

Aplicando la fórmula anterior a nuestra función f definida en (2.1.2), donde en el segundo sumando $\alpha(x) = x$ y $h(t,x) = \sin(x-t)g(t)$, obtenemos expresiones, r

$$f'(x) = -\sin(x-a) + \int_a^x \cos(x-t)g(t)dt, \qquad x \in [a,b]. \tag{2.1.3}$$

Análogamente,

$$f''(x) = -\cos(x-a) + g(x) - \int_a^x \sin(x-t)g(t)dt, \qquad x \in [a,b]. \tag{2.1.4}$$

Nótese que esta última expresión se puede escribir, usando (2.1.2), como

$$f''(x) = -f(x) + g(x) \quad \Rightarrow f''(x) + f(x) = g(x), \qquad x \in [a,b].$$

Además, de (2.1.3) se tiene que $f'(a) = 0$ y, por tanto, dado que $f(a) = 1$ como ya vimos, la función f definida por (2.1.2) es efectivamente la solución del problema (2.1.1).

El método anterior nos permite, al menos formalmente, plantearnos el siguiente problema: resolver el problema de valores iniciales

$$f''(x) + f(x) = \sigma(x)f(x), \quad x \in [a,b], \quad f(a) = 1, \; f'(a) = 0. \tag{2.1.5}$$

Está *claro* que la solución será

$$f(x) = \cos(x-a) + \int_a^x \sin(x-t)\sigma(t)f(t)dt, \qquad x \in [a,b]. \tag{2.1.6}$$

Lo anterior define una *ecuación integral*. Escribamos dicha ecuación (2.1.6) de una forma más conveniente.

Sea $D = \{(t,x) \in \mathbb{R}^2 \,|\, t \in [a,b], x \in [a,b]\}$ y sea $k(x,t)$ una función continua en D. Definamos el *operador* K, $K : C([a,b]) \mapsto C([a,b])$

$$Kf := K(f(x)) = \int_a^x k(x,t)f(t)dt. \tag{2.1.7}$$

Entonces (2.1.6) la podemos reescribir como

$$f = u + Kf, \quad u = \cos(x-a), \quad k(x,t) = \sin(x-t)\sigma(t).$$

¿Cómo resolver esta ecuación?

2.2. La ecuación integral de Volterra

Vamos a intentar resolver la ecuación

$$f(x) = u(x) + Kf(x), \quad x \in [a,b], \tag{2.2.1}$$

donde $u, f \in C([a,b])$ y K es el operador de Volterra definido en (2.1.7).

Nótese que K trasforma funciones continuas en funciones continuas. Una forma natural de resolverla es usar el *método iterativo* de Picard.

Así, elegimos una función $f_0 \in C([a,b])$ cualquiera y la sustituimos en el miembro derecho de (2.2.1) lo que nos dará, en principio, una nueva función continua f_1

$$f_1(x) = u(x) + Kf_0(x), \quad x \in [a,b]. \tag{2.2.2}$$

Si $f_1 = f_0$, lo cual es improbable, tendríamos la solución buscada. Si no, entonces sustituimos f_1 en el miembro derecho de (2.2.1) y obtenemos una nueva función f_2

$$f_2(x) = u(x) + Kf_1(x), \quad x \in [a,b]. \tag{2.2.3}$$

Si $f_2 = f_1$ tendríamos la solución, si no, sustituimos f_2 en en el miembro derecho de (2.2.1), y así, sucesivamente. De esta forma obtenemos una sucesión de ecuaciones

$$f_n(x) = u(x) + Kf_{n-1}(x), \quad x \in [a,b], \quad n = 1, 2, \ldots. \tag{2.2.4}$$

Como ya hemos dicho, es improbable que para algún n, $f_n = f_{n-1}$, así que la pregunta natural es ¿cuán cerca está f_n de f si n es muy grande? y ¿qué significa que f_n esté cerca de f, o que $f_n \to f$ en $[a,b]$ cuando $n \to \infty$?

Está claro que esta pregunta corresponde a saber cuando una sucesión de funciones continuas f_n converge a alguna función f, y para que f sea solución de (2.2.1) necesitamos además que f sea continua.

Por otro lado, nótese que si $f_n \to f$, otra pregunta natural es si $Kf_n \to Kf$. Si esto fuera así, entonces, tomando límites cuando $n \to \infty$ en (2.2.4) obtendríamos (2.2.1) y, por tanto, la solución de la ecuación. Intentemos formalizar lo anterior de una forma rigurosa.

Como las f_n son continuas y nos interesa que la función límite también lo sea, entonces podemos usar el concepto de convergencia uniforme de funciones (véase la sección 1.4). Allí vimos que una sucesión de funciones f_n converge uniformemente a f en $[a,b]$ si

$$\sup_{x\in[a,b]} |f_n(x) - f(x)| \to 0, \text{ cuando } n \to \infty.$$

Lo anterior nos da una idea de como definir la *distancia* ρ entre dos funciones f y g. Así, dada dos funciones $f, g \in C([a,b])$ definiremos la distancia $\rho(f,g)$ mediante la expresión

$$\rho(f,g) = \sup_{x\in[a,b]} |f(x) - g(x)|.$$

De esta definición se sigue que para todas $f, g, h \in C([a,b])$

1. $\rho(f,g) \geq 0$ y $\rho(f,g) = 0$ si y solo si $f = g$,
2. $\rho(f,g) = \rho(g,f)$, y
3. $\rho(f,g) \leq \rho(f,h) + \rho(h,g)$.

En muchas ocasiones se dice que la función ρ que satisface las propiedades anteriores define una *métrica*. Los conjuntos donde se puede definir una métrica se suelen denominar espacios métricos (ver la Definición 3.1.1).

Asociada a la definición de distancia se puede definir el *tamaño* o la *norma* de una función. Así, dada una función $f \in C([a,b])$ definiremos la norma de f, $\|f\|$ al número

$$\|f\| := \sup_{x\in[a,b]} |f(x)|. \tag{2.2.5}$$

Es fácil comprobar que la *norma* anterior satisface las siguientes propiedades: para todo $f, g \in C([a,b])$,

1. $\|f\| \geq 0$ y $\|f\| = 0$ si y solo si $f = 0$,
2. Para todo $\lambda \in \mathbb{R}$, $\|\lambda f\| = |\lambda| \|f\|$,
3. Para todos $f, g \in C([a,b])$, $\|f + g\| \leq \|f\| + \|g\|$.

Los conjuntos donde se puede definir una norma se denominan espacios normados (ver la Definición 4.2.1).

Lo siguiente que queremos hacer notar es que el espacio de las funciones continuas $C([a,b])$ satisface una serie de interesantes propiedades que pasamos a enumerar:

1. Para todos $f,g \in C([a,b])$, la suma, $h = f+g \in C([a,b])$ y para todos $f,g,h \in C([a,b])$ se cumple que:
   a) $f+g = g+f$
   b) $(f+g)+h = f+(g+h)$
   c) $f+\emptyset = \emptyset + f = f$, donde $\emptyset$ denota a la función $\emptyset : [a,b] \mapsto \mathbb{R}$, $\emptyset(x) = 0$.
   d) Cualquiera sea f existe una función $(-f)$ tal que $f+(-f) = (-f)+f = \emptyset$.
2. Para todo $f \in C([a,b])$, y $\alpha \in \mathbb{R}$, $g = \alpha \cdot f \in C([a,b])$, y para todos $f,g \in C([a,b])$, y $\alpha, \beta \in \mathbb{R}$ se cumple que:
   a) $\alpha \cdot (f+g) = \alpha \cdot f + \alpha \cdot g$,
   b) $(\alpha+\beta) \cdot f = \alpha \cdot f + \beta \cdot f$,
   c) $\alpha \cdot (\beta \cdot f) = (\alpha\beta) \cdot f$
   d) $1 \cdot f = f$

Los conjuntos que satisfacen las propiedades anteriores se denominan espacios vectoriales (ver la Definición 4.1.1).

De la propia definición del operador K (2.1.7), usando las propiedades de la integral, tenemos que, para todos $\alpha, \beta \in \mathbb{R}$ y $f,g \in C([a,b])$,

$$K(\alpha f + \beta g) = \alpha K f + \beta K g. \tag{2.2.6}$$

Un operador cumple con lo anterior se dice que es un *operador lineal*.

Además, para toda función $f \in C([a,b])$ se tiene, usando (2.1.7),

$$\begin{aligned} |Kf| &\leq \int_a^x |k(x,t)||f(t)|dt \leq \int_a^x \sup_{x,t\in[a,b]} |k(x,t)| \sup_{t\in[a,b]} |f(t)|dt \\ &\leq \kappa\|f\|(x-a) \leq \kappa\|f\|(b-a) = c\|f\|, \end{aligned} \tag{2.2.7}$$

donde $\kappa = \sup_{x,t\in[a,b]} |k(x,t)|$ y $c = \kappa(b-a)$. Es decir, para toda $f \in C([a,b])$, existe un $c \geq 0$ tal que

$$\|Kf\| = \sup_{t\in[a,b]} |Kf| \leq c\|f\|. \tag{2.2.8}$$

Los operadores que cumplen con dicha propiedad se denominan *operadores acotados*.

Armados de todo lo anterior estamos en condiciones de responder a nuestra pregunta sobre si la ecuación (2.2.1) tiene solución y si esta la podemos encontrar mediante el límite de la sucesión de funciones f_n definidas por (2.2.4).

Para ello vamos a estimar el valor de $\|f_n - f_m\|$ donde la norma es la definida en (2.2.5). Si resulta que su valor es tan pequeño como se quiera cuando n, m son lo suficientemente grandes (i.e., $(f_n)_n$ es una sucesión de Cauchy) entonces f_n converge uniformemente (ver Teorema 1.4.10) a una función f en $[a,b]$ que, por el Teorema 1.4.22, es continua en $[a,b]$. Esta propiedad de que toda sucesión de Cauchy sea convergente se conoce como propiedad de *completitud del espacio* (ver Definición 3.5.12).

Partiendo de (2.2.3) y usando (2.2.2) tenemos

$$f_2 = u + Kf_1 = u + K(u + Kf_0) = u + Ku + K(Kf_0) = u + Ku + K^2 f_0,$$

donde K^2 es el operador que se obtiene al componer, o multiplicar, K por si mismo: $K^2 : C([a,b]) \mapsto C([a,b])$, $K^2 f = (K \circ K)(f) = K(Kf)$. En general, definiremos el operador K^n inductivamente, i.e., $K^n = K \circ K^{n-1}$.

Continuando este proceso podemos comprobar por inducción que

$$f_n = u + Ku + K^2 u + \cdots + K^{n-1} u + K^n f_0,$$

por tanto, para todos $m < n$ obtenemos

$$f_n - f_m = K^m u + \cdots + K^{n-1} u + K^n f_0 - K^m f_0.$$

Entonces, usando las propiedades de la norma (concretamente la tercera)

$$\|f_n - f_m\| \leq \|K^m u\| + \cdots + \|K^{n-1} u\| + \|K^n f_0\| + \|K^m f_0\|. \tag{2.2.9}$$

Vamos a acotar cada uno de los sumandos. Nótese que estos se pueden escribir de la forma $\|K^n h\|$, $h \in C([a,b])$, $n = 1, 2, \ldots$.

Para $n = 1$ ya hemos visto en (2.2.8) que $\|Kh\| \leq c\|h\|$. Entonces,

$$\|K^2 h\| = \|K(Kh)\| \leq c\|Kh\| \leq c^2\|h\| \quad \Rightarrow \quad \|K^n h\| \leq c^n\|h\|. \quad (2.2.10)$$

De esta forma podríamos obtener una cota superior de la suma del miembro derecho de (2.2.9), pero dicha cota solo es útil si $c < 1$ pues en caso contrario cualquiera de los sumandos tendería a infinito cuando $n \to \infty$.

Vamos a *afinar* nuestra cota. Como vimos en (2.2.7) tenemos que

$$|Kh| \leq \kappa\|h\|(x-a), \quad x > a,$$

luego

$$\begin{aligned}
|K^2 h| = |K(Kh)| &\leq \int_a^x |k(x,t)||Kh|dt \leq \kappa \int_a^x |Kh|dt \\
&\leq \kappa^2\|h\| \int_a^x (x-a)dt = \frac{\kappa^2(x-a)^2}{2}\|h\|, \\
|K^3 h| = |K(K^2 h)| &\leq \int_a^x |k(x,t)||K^2 h|dt \leq \kappa \int_a^x |K^2 h|dt \\
&\leq \frac{\kappa^3\|h\|}{2} \int_a^x (x-a)^2 dt = \frac{\kappa^3(x-a)^3}{3!}\|h\|.
\end{aligned}$$

y, así, por inducción, obtenemos

$$|K^n h| \leq \frac{\kappa^n(x-a)^n}{n!}\|h\|,$$

que tomando supremos en $x \in [a,b]$ nos conduce a una cota mucho más fina que la obtenida en (2.2.10)

$$\|K^n h\| \leq \frac{\kappa^n(b-a)^n}{n!}\|h\|. \quad (2.2.11)$$

Por tanto, podemos reescribir (2.2.9), definiendo $\zeta = (b-a)\kappa$, como

$$\|f_n - f_m\| \leq \left(\frac{\zeta^m}{m!} + \cdots + \frac{\zeta^{n-1}}{(n-1)!}\right)\|u\| + \left(\frac{\zeta^m}{m!} + \frac{\zeta^n}{n!}\right)\|f_0\|. \quad (2.2.12)$$

Por un lado está claro que $\zeta^m/m! + \zeta^n/n! \to 0$, cuando $n, m \to \infty$. Por el otro,

$$\frac{\zeta^m}{m!} + \cdots + \frac{\zeta^{n-1}}{(n-1)!} \leq \sum_{k=m}^{\infty} \frac{\zeta^k}{k!} \to 0 \text{ cuando } m \to \infty$$

pues es la cola de la serie $\sum_{k=0}^{\infty} \zeta^k/k! = e^{\zeta}$. De esta forma hemos probado que f_n es uniformemente convergente y que límite f es una función continua en $[a, b]$.

Probemos ahora que $Kf_n \to Kf$ si $f_n \to f$. Para ello, usaremos la linealidad y la acotación de K (ver (2.2.6) y (2.2.8), respectivamente)), que nos conduce a

$$\|Kf_n - Kf\| = \|K(f_n - f)\| \leq c\|f_n - f\| \to 0 \text{ si } n \to \infty.$$

Por tanto, si ahora tomamos el límite cuando $n \to \infty$ en (2.2.4), obtenemos (2.2.1) lo que prueba que la sucesión f_n construida mediante el método iterativo de Picard aquí descrito tiende a la solución de la ecuación integral de Volterra (2.2.1).

Hemos encontrado una solución de (2.2.1) pero ¿es única? Probemos que sí. Supongamos que hay dos funciones continuas en $C([a, b])$, f y g, con $f \neq g$ tales que $f = u + Kf$ y $g = u + Kg$. Restando ambas tenemos $f - g = K(f - g)$. Sea $h = f - g$. Entonces

$$h = Kh \quad \Rightarrow \quad h = K(Kh) = K^2 h = K^2(Kh) = K^3 h = \cdots = K^n h,$$

y, por tanto, $\|h\| = \|K^n h\|$. Si tomamos el límite cuando $n \to \infty$ en la expresión anterior y usamos (2.2.11) obtenemos $\|K^n h\| \to 0$ de donde se deduce que $\|h\| = 0$, i.e. $f = g$, lo cual es una contradicción.

Recapitulemos el conjunto de propiedades que hemos usado:

1. El conjunto $C([a, b])$ las funciones continuas es un *espacio vectorial*.

2. En $C([a, b])$es un espacio *normado* y a su vez *métrico*.

3. El operador K es un operador *lineal* y *acotado*.

4. Toda sucesión $f_n \in C([a, b])$ tal que $f_n - f_m \to 0$ cuando $n, m \to \infty$, (es decir, $(f_n)_n$ es una sucesión de Cauchy) es convergente, i.e., $f_n \to f \in C([a, b])$ cuando $n \to \infty$. Es decir, el espacio $C([a, b])$ con la métrica del supremo es un espacio *completo*.

¿Existen otros espacios útiles para resolver este tipo de problemas? ¿Qué otras propiedades importantes son necesarias para lidiar con estos espacios? ¿Con qué otros tipos de operadores podemos encontrarnos en la práctica?

Nuestro objetivo será estudiar algunas de las propiedades de los espacios *funcionales* así como de cierta clase de operadores que generalizan el operador de Volterra discutido en este capítulo.

Capítulo 3

Espacios métricos

Mide lo que es medible y haz medible lo que no lo es.

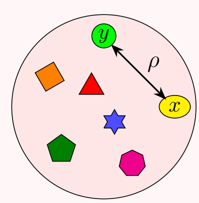
Atribuida a Galileo Galilei

Los espacios métricos son la *extensión natural* de los espacios $\mathbb{R}$ y $\mathbb{C}$. La idea consiste en, dado un conjunto cualquiera de elementos, poder medir la distancia entre ellos.

3.1. Definición de espacio métrico

Definición 3.1.1 *Un espacio métrico es un par* $(\mathbb{X}, \rho)$ *donde* $\mathbb{X}$ *es un conjunto y* $\rho := \rho(x, y)$ *es una función real (univaluada) no negativa definida para todos* $x, y, z \in \mathbb{X}$ *tal que*

1. $\rho(x, y) = 0 \Longleftrightarrow x = y$,
2. $\rho(x, y) = \rho(y, x)$,
3. $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$.



En adelante diremos simplemente que $\mathbb{X}$ es el espacio y ρ su métrica.

Veamos varios ejemplos representativos de espacios métricos. Nótese que muchos de ellos están definidos sobre el mismo conjunto $\mathbb{X}$ pero con diferentes funciones ρ, i.e., la métrica cambia.

Ejemplo 3.1.2 *Sea $\mathbb{X}$ un conjunto arbitrario y definamos la métrica* trivial $\rho(x,y)=1$ *si* $x\neq y$ *y* $\rho(x,y)=0$ *si* $x=y$. *Este espacio se suele denominar espacio métrico discreto.*

Ejemplo 3.1.3 *Si $\mathbb{X}$ es el conjunto de todos los números reales y $\rho(x,y)=|x-y|$, donde $|x|$ es el valor absoluto de x, obtenemos un espacio métrico normalmente denotado por $\mathbb{R}$.*

Ejemplo 3.1.4 *El espacio métrico $\mathbb{C}$ es el espacio definido por el par $\mathbb{X}=\mathbb{C}$ y la métrica $\rho(x,y)=|x-y|$, donde $|x|$ es el módulo de x.*

Ejemplo 3.1.5 $\mathbb{X}=\mathbb{R}^n$*, es decir, el espacio de las n-tuplas $x=(x_1,\dots,x_n)$ con la métrica*

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

Dicho espacio lo denotaremos por $\mathbb{R}^n_2$. Análogamente se puede definir para el caso complejo.

Ejemplo 3.1.6 *Si $\mathbb{X}$, es el espacio de las n-tuplas reales $x=(x_1,\dots,x_n)$ pero con la métrica*

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

obtenemos otro espacio métrico (distinto al anterior) que denotaremos por $\mathbb{R}^n_1$. Análogamente se puede definir para el caso complejo.

Una generalización de los dos ejemplos anteriores es el siguiente:

Ejemplo 3.1.7 $\mathbb{X}=\mathbb{R}^n$*, es decir, el espacio de las n-tuplas $x=(x_1,\dots,x_n)$ con la métrica*

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

Dicho espacio lo denotaremos por $\mathbb{R}^n_p$. Análogamente se puede definir para el caso complejo.

Ejemplo 3.1.8 $\mathbb{R}^n_\infty$. *Es decir,* $\mathbb{X} = \mathbb{R}^n$ *espacio de las* n*-tuplas* $x = (x_1, x_2, \ldots, x_n)$ *pero con la métrica:*

$$\rho(x, y) = \max_{k=1,\ldots,n} |x_k - y_k|.$$

Ejemplo 3.1.9 *Sea* $\mathbb{X} = C([a, b])$*, es decir, el espacio de las funciones continuas definidas sobre el segmento* $[a, b]$*. Definamos la función*

$$\rho(f, g) = \max_{x\in[a,b]} |f(x) - g(x)|.$$

El par obtenido $C^\infty_{[a,b]}$ *es un espacio métrico.*

Ejemplo 3.1.10 *Sea nuevamente* $\mathbb{X} = C([a, b])$*, es decir, el espacio de las funciones continuas definidas sobre el segmento* $[a, b]$ *y definamos la métrica*

$$\rho(f, g) = \left(\int_a^b |f(x) - g(x)|^p dx\right)^{1/p}, \qquad p \geq 1.$$

El par obtenido $C^p_{[a,b]}$ *es un espacio métrico. Como caso particular (de especial relevancia) tenemos el caso* $p = 2$*, i.e.,* $\mathbb{X} = C([a, b])$ *y* ρ *es la función*

$$\rho(f, g) = \sqrt{\int_a^b |f(x) - g(x)|^2 dx.}$$

Ejemplo 3.1.11 *Sea* $\mathbb{X}$ *el espacio de todas las sucesiones numéricas reales (o complejas)* $x = (x_n)_n$ *acotadas y definimos la métrica*

$$\rho(x, y) = \sup_{k\in\mathbb{N}} |x_k - y_k|.$$

El espacio obtenido es un espacio métrico que denotaremos por ℓ^∞*.*

Ejemplo 3.1.12 *Sea ahora $\mathbb{X}$ el espacio de todas las sucesiones reales (o complejas) $x = (x_n)_n$ tales que $\sum_{k=1}^{\infty} |x_k|^p < +\infty$ con la métrica*

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

Dicho espacio lo denotaremos por ℓ^p.

Un caso particular de los espacios ℓ^p es cuando $p = 2$, i.e., el espacio métrico ℓ^2 usualmente asociado al nombre de Hilbert. O sea, $\ell^2 = (\mathbb{X}, \rho)$ donde $\mathbb{X}$ es el espacio de todas las sucesiones $x = (x_1, x_2, \ldots, x_n, \ldots)$ tales que $\sum_{k=1}^{\infty} |x_k|^2 < +\infty$ con la métrica

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

Ejemplo 3.1.13 *Sea ahora $\mathbb{X}$ el espacio de todas las sucesiones reales $x = (x_1, x_2, \ldots, x_n, \ldots)$ y definamos la métrica por*

$$\rho(x, y) = \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{|x_j - y_j|}{1 + |x_j - y_j|}.$$

Este espacio también es un espacio métrico.

Ejemplo 3.1.14 *Sea el espacio métrico $(\mathbb{X}, \rho)$. Entonces si definimos sobre $\mathbb{X}$ una nueva métrica $\sigma(x, y) = \rho(x, y)/(1 + \rho(x, y))$, obtenemos un nuevo espacio métrico $(\mathbb{X}, \sigma)$.*

A continuación probaremos que los espacios considerados en los ejemplos anteriores son, en efecto, espacios métricos.

Ante todo notemos que en todos los casos es fácil comprobar los axiomas 1 y 2 de la definición 3.1.1. Probemos la desigualdad triangular. Para los ejemplos 3.1.2-3.1.4, 3.1.8, 3.1.9 y 3.1.11 las pruebas son sencillas y se basan en la desigualdad triangular en $\mathbb{R}$ (o $\mathbb{C}$) menos en el Ejemplo 3.1.2 donde es trivial. Nótese además que en todos estos casos la métrica (función ρ) está bien definida.

Para estudiar los ejemplos (3.1.5)-(3.1.7), (3.1.10) y (3.1.12) necesitaremos la desigualdad de Minkowski para n números, series o integrales. Así tenemos el siguiente:

Teorema 3.1.15 *(Desigualdad de Minkowski) Sea los números x_i e y_i, $i = 1, 2 \ldots, n$, cualesquiera. Entonces, para todo $p > 1$ se tiene*

$$\left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{\frac{1}{p}}, \tag{3.1.1}$$

donde la igualdad solo tiene lugar si $x_i = c\, y_i$ para todo $i = 1, 2, \ldots, n$ (es decir, si x_i y y_i son proporcionales).

Demostración: El primer paso consiste en probar que para todos $u, v \geq 0$ y $1 \leq p < +\infty$ se tiene

$$(u+v)^p = \inf_{t\in(0,1)} \left\{t^{1-p}u^p + (1-t)^{1-p}v^p; u, v \geq 0, 0 < t < 1\right\}. \tag{3.1.2}$$

Para ello basta comprobar que el mínimo absoluto de la función $f : (0,1) \mapsto \mathbb{R}$, $f(t) = t^{1-p}u^p + (1-t)^{1-p}v^p$ con $u, v > 0$ y $1 < p < +\infty$ (el caso $p = 1$ es inmediato) se alcanza en $t_* = u/(u+v)$ y $f(t_*) = (u+v)^p$. Entonces, para todos x_i e y_i, usando (3.1.2) tenemos

$$\begin{aligned} |x_i + y_i|^p \leq (|x_i| + |y_i|)^p &= \inf_{t\in(0,1)} (t^{1-p}|x_i|^p + (1-t)^{1-p}|y_i|^p) \\ &\leq t^{1-p}|x_i|^p + (1-t)^{1-p}|y_i|^p, \end{aligned}$$

que, sumando en i, nos conduce a

$$\sum_{i=1}^{n} |x_i + y_i|^p \leq t^{1-p}\sum_{i=1}^{n} |x_i|^p + (1-t)^{1-p}\sum_{i=1}^{n} |y_i|^p.$$

Llamando $u^p = \sum_{i=1}^{n} |x_i|^p$, $v^p = \sum_{i=1}^{n} |y_i|^p$ y tomando el ínfimo en $t \in (0,1)$ en la desigualdad anterior y usando nuevamente (3.1.2) se sigue que

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

de donde se obtiene el resultado[1] (3.1.1). ■

[1] Esta prueba se debe a Heinz König (A simple proof of the Minkowski inequality. In: Walter, W. (eds) *General Inequalities 6. International Series of Numerical Mathematics,* Vol. 103. Birkhäuser, Basel, 1992, pág. 469).

Del teorema anterior se sigue que los espacios $\mathbb{R}^n_p$ (o $\mathbb{C}^n_p$) del ejemplo (3.1.7) son espacios métricos.

La desigualdad de Minkowski (3.1.1) se puede extender fácilmente al caso de las series infinitas (ver Problema 3.5), en el caso de que las mismas converjan. Así se tiene, en particular para las sucesiones numéricas $(x_i)_i$ e $(y_i)_i$, que

$$\left(\sum_{i=1}^{\infty}|x_i+y_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^{\infty}|x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{\infty}|y_i|^p\right)^{\frac{1}{p}}. \tag{3.1.3}$$

De lo anterior se sigue que los espacios ℓ^p del Ejemplo 3.1.12 son espacios métricos.

Análogamente, para las funciones integrables correspondientes (i.e., donde estén definidas las integrales) se tienen, para todo $p \geq 1$, (ver Problema 3.6) la desigualdad de Minkowski para integrales,

$$\left(\int_a^b |f(x)+g(x)|^p dx\right)^{\frac{1}{p}} \leq \left(\int_a^b |f(x)|^p dx\right)^{\frac{1}{p}} + \left(\int_a^b |g(x)|^p dx\right)^{\frac{1}{p}}. \tag{3.1.4}$$

Luego los espacios $C^p_{[a,b]}$ del Ejemplo 3.1.10 son espacios métricos. Para los ejemplos 3.1.13 y 3.1.14 usamos que la función real $f : [0,\infty) \mapsto \mathbb{R}$, $f(x) = x/(1+x)$, es siempre creciente. En el primer caso tenemos que, para todo $a, b \in \mathbb{C}$, $|a+b| \leq |a|+|b| \Rightarrow f(|a+b|) \leq f(|a|+|b|)$, luego

$$\frac{|a+b|}{1+|a+b|} \leq \frac{|a|+|b|}{1+|a|+|b|} = \frac{|a|}{1+|a|+|b|} + \frac{|b|}{1+|a|+|b|} \leq \frac{|a|}{1+|a|} + \frac{|b|}{1+|b|},$$

de donde se sigue el resultado (ver Problema 3.17). En el segundo basta usar que $\rho(x,y) \leq \rho(x,z)+\rho(z,y)$, luego $f(\rho(x,y)) \leq f(\rho(x,z)+\rho(z,y))$.

3.2. Algunas definiciones topológicas

Definición 3.2.1 *Sea $\mathbb{X}$ un espacio métrico, $x_0 \in \mathbb{X}$ y $r > 0$. Definiremos la bola abierta $B(x_0,r)$ al conjunto*

$$B(x_0,r) = \{x \in \mathbb{X};\ \rho(x_0,x) < r\},$$

bola o esfera cerrada $S(x_0, r)$ *al conjunto*

$$S(x_0, r) = \{x \in \mathbb{X};\ \rho(x_0, x) \leq r\}.$$

Las bolas abiertas $B(x_0, \varepsilon)$ se suelen denominar ε-vecindades (o entornos) de x_0. Es evidente que toda ε-vecindad de x_0 contiene al propio x_0. Se suele decir que un entorno es pequeño si ε es pequeño.

Por ejemplo, sea $\mathbb{X} = \mathbb{R}$ con la métrica habitual[2]. Entonces $B(x_0, r) = (x_0 - r, x_0 + r)$ y $S(x_0, r) = [x_0 - r, x_0 + r]$.

Definición 3.2.2 *Un punto* x_0 *se denomina punto interior del conjunto* $M \subset \mathbb{X}$*,* $\mathbb{X}$ *espacio métrico, si existe un* $\varepsilon > 0$ *tal que* $B(x_0, \varepsilon) \subset M$*.*

Es decir, un punto x es un punto interior de M si existe un entorno de dicho punto contenido en M tal y como se muestra en la figura 3.2. Así, por ejemplo, 1 es un punto interior de $(0, 2]$ con la métrica de $\mathbb{R}$, sin embargo 2 no lo es.

Definición 3.2.3 *Se dice que un punto* $x_0 \in \mathbb{X}$ *(no necesariamente* $x_0 \in M \subset \mathbb{X}$*),* $\mathbb{X}$ *espacio métrico, es un punto de la frontera de* M *si en cualquier entorno de* x_0 *(tan pequeño como se quiera) hay al mismo tiempo elementos de* M *y de su complementario* $\mathbb{X}\backslash M$ *(pudiendo ser, en ambos casos, el propio* x_0*). El conjunto de todos los puntos frontera de* M *se denomina frontera de* M *y se suele denotar por* ∂M*.*

Por ejemplo, sea $\mathbb{X} = \mathbb{R}$ y $M = (0, 1]$. Todos los $x_0 \in (0, 1)$ son interiores y los puntos 0 y 1 constituyen la frontera.

Nótese que puesto que la definición de punto frontera es simétrica respecto al conjunto $M \subset \mathbb{X}$ y su complementario $\mathbb{X}\backslash M$, luego ambos tienen la misma frontera.

Definición 3.2.4 *Se dice que el conjunto* $M \subset \mathbb{X}$*,* $\mathbb{X}$ *espacio métrico, es abierto en* $\mathbb{X}$ *si todos sus puntos son interiores. Un conjunto* $M \subset \mathbb{X}$ *es cerrado en* $\mathbb{X}$ *si su complementario* $\mathbb{X}\backslash M$*, es abierto.*

Está claro que un conjunto $M \subset \mathbb{X}$ es abierto en $\mathbb{X}$ si y solo si todos sus puntos (elementos) se pueden encerrar en una bola abierta contenida

[2]En adelante, a no ser que se diga lo contrario, para $\mathbb{R}$ y los distintos subconjuntos de $\mathbb{R}$ usaremos la métrica habitual, es decir, la del Ejemplo 3.1.3.

completamente en M. Nótese también que, dado un entorno $B(x_0, \varepsilon)$ de x_0, y cualquiera sea $y \in B(x_0, \varepsilon)$ existe un entorno de y tal que $B(y, \delta) \subset B(x_0, \varepsilon)$. Es decir, cualquier punto dentro de una bola abierta (entorno) es el centro de un entorno contenido en dicha bola. Para probarlo notemos que todos los puntos z del entorno $B(x_0, \varepsilon)$ son tales que $\rho(x, z) < \epsilon$. Sea la bola $B(y, \delta)$ con $0 < \delta < \epsilon - \rho(x_0, y)$. Entonces, para todo $z \in B(y, \delta)$ tendremos que $\rho(x_0, z) \leq \rho(x_0, y) + \rho(y, z) \leq \rho(x_0, y) + \delta < \epsilon$. O sea $B(y, \delta) \subset B(x_0, \varepsilon)$. De lo anterior se deduce que las bolas son conjuntos abiertos.

Nótese que la definición de conjunto abierto es equivalente a decir que todos los puntos $x \in E$ están contenidos en un subconjunto abierto $F \subset E$ de E (si F es abierto todos los puntos de F son interiores, luego x es un punto interior de F y, por tanto, de E).

Por ejemplo, sea $\mathbb{X} = \mathbb{R}$. El conjunto $(0, 1)$ es abierto, $[0, 1]$ es cerrado y $[0, 1)$ no es ni abierto ni cerrado.

Proposición 3.2.5 *Un conjunto $M \subset \mathbb{X}$, $\mathbb{X}$ espacio métrico, es abierto si y solo si no contiene ningún punto frontera.*

Demostración: Si M es abierto todos sus puntos son interiores, luego cada punto $x \in M$ tiene un entorno de puntos que solo contiene puntos de M por lo que no cualquier entorno de x tiene puntos de su complementario $\mathbb{X} \backslash M$, luego ningún x puede ser un punto frontera. Por otro lado, si M no tiene ningún punto frontera, entonces cualquier $x \in M$ no es un punto frontera, luego tiene que haber al menos un entorno de x que solo contenga puntos de M, luego x es un punto interior. Como x es cualquiera M es abierto. ∎

Proposición 3.2.6 *Un conjunto $M \subset \mathbb{X}$, $\mathbb{X}$ espacio métrico, es cerrado si y solo si contiene todos sus puntos frontera.*

Demostración: Si M es cerrado, su complementario $\mathbb{X} \backslash M$ es abierto, luego por la Proposición 3.2.5 $\mathbb{X} \backslash M$ no contiene ningún punto frontera (recordemos que M y $\mathbb{X} \backslash M$ comparten la misma frontera), por tanto todos los puntos frontera de M pertenecen a M. Por otro lado, si M contiene a todos sus puntos frontera, entonces $\mathbb{X} \backslash M$ no contiene a ninguno y, por tanto, la Proposición 3.2.5 nos dice que $\mathbb{X} \backslash M$ es abierto, luego M es cerrado. ∎

Nótese que la frontera ∂M de cualquier conjunto M es un cerrado (es consecuencia de la Proposición 3.2.6), luego su complementario $\mathbb{X} \setminus \partial M$ es abierto.

La siguiente proposición es de gran interés en la práctica:

Proposición 3.2.7 *Sea Σ el conjunto de todos los subconjuntos abiertos de $\mathbb{X}$, $\mathbb{X}$ espacio métrico. Entonces*

1. $\emptyset \in \Sigma$, $\mathbb{X} \in \Sigma$,
2. *la unión (finita o infinita) de subconjuntos abiertos de $\mathbb{X}$ es abierto: Si U_k, $k = 1, 2, \dots$ son abiertos, $\bigcup_k U_k \in \Sigma$*
3. *La intersección de un número finito de abiertos es abierto: Si U_k, $k = 1, 2, \dots, n$ son abiertos, $\bigcap_{k=1}^n U_k \in \Sigma$.*

Demostración: El apartado 1 es evidente pues $\emptyset$ no contiene elementos y $\mathbb{X}$ es todo el espacio. Para probar 2 notemos que cualquiera sea $x \in U = \bigcup_k U_k \in \Sigma$, $x \in U_k$ para cierto k. Pero entonces existe una bola $B(x, \varepsilon) \subset U_k$ pues U_k es abierto y, por tanto, para todo $x \in U$, existe una bola $B(x, \varepsilon) \subset U$, de donde se sigue que U es abierto. Para probar 3 basta comprobar que es cierto para $n = 2$. Sean U_1 y U_2 dos abiertos. Si $U_1 \bigcap U_2 = \emptyset$ entonces es trivial (por 1). Asumamos $U = U_1 \bigcap U_2 \neq \emptyset$. Entonces para todo $x \in U$ existe una bola centrada en x tal que $B(x, \varepsilon_1) \subset U_1$ y otra tal que $B(x, \varepsilon_2) \subset U_2$. Pero entonces la bola $B(x, \varepsilon)$ con radio $\varepsilon = \min(\varepsilon_1, \varepsilon_2)$ está contenida en U. ■

Las tres propiedades anteriores son de extrema importancia. Tal es así que ellas definen un tipo de espacios muy generales: los *espacios topológicos*. Así, el par, dados un conjunto $\mathbb{X}$ y una colección Σ de subconjuntos de $\mathbb{X}$, $(\mathbb{X}, \Sigma)$ se denomina espacio topológico si Σ cumple con los *axiomas* (propiedades) 1, 2 y 3 de la proposición anterior. Al conjunto Σ se le denomina *topología* de $\mathbb{X}$. Así pues, todo espacio métrico es un espacio topológico.

Nota 3.2.8 *La intersección de infinitos conjuntos abiertos no tiene que ser un abierto:* $\bigcap_{k=1}^\infty (0, 1 + 1/k) = (0, 1]$. *La unión de infinitos conjuntos cerrados no tiene que ser cerrada:* $\bigcup_{k=1}^\infty [0, 1 - 1/k] = [0, 1)$.

Antes de continuar vamos a describir cómo es la *geometría* de algunos espacios métricos. Como ejemplo tomaremos $\mathbb{X} = \mathbb{R}^2$ y escogeremos

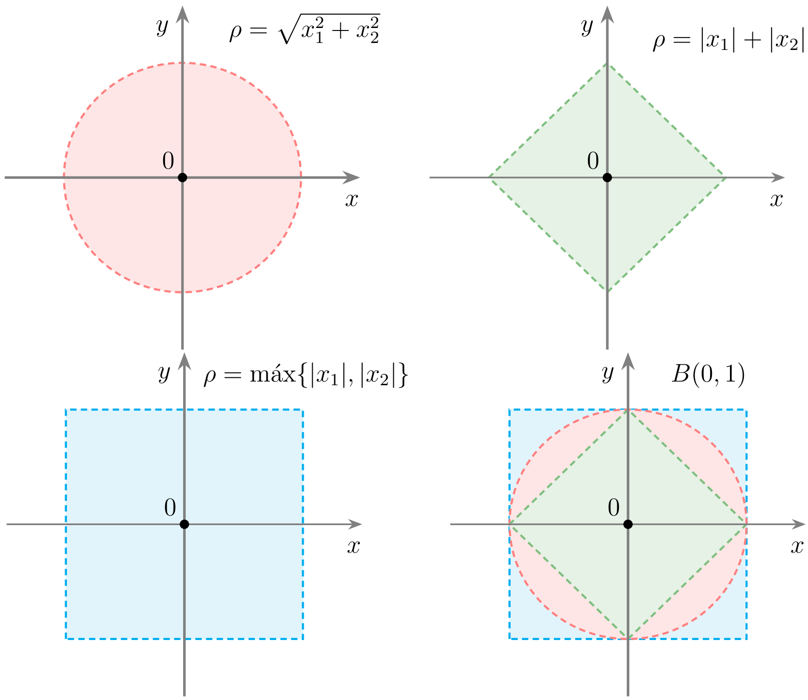


Figura 3.1: Entornos $B(0,1)$ según las métricas de los ejemplos 3.1.5 (encima a la izquierda) y 3.1.6 (encima a la derecha) y del Ejemplo 3.1.8 (debajo a la izquierda). Debajo a la derecha se muestran los tres entornos en una única figura. Las distancias ρ corresponden a la distancia entre los puntos $x = (x_1, x_2)$ y $x_0 = (0,0)$.

distintas métricas. En la figura 3.1 están representados las bolas $B(0,1)$ para las métricas de los ejemplos 3.1.5, 3.1.6 y 3.1.8, respectivamente, así como la comparación de las correspondientes bolas.

Definición 3.2.9 *Sea $M \subset \mathbb{X}$, $\mathbb{X}$ espacio métrico. Diremos que $x \in \mathbb{X}$ es un punto de contacto (o adherente) de M si en cualquier bola $B(x,\varepsilon)$, $\varepsilon > 0$ hay al menos un elemento de M. Así mismo, diremos que x es un punto de acumulación (o punto límite) de M si en cualquier bola $B(x,\varepsilon)$, $\varepsilon > 0$ hay al menos un elemento de M distinto de x, o equivalentemente, en cada bola $B(x,\varepsilon)$, $\varepsilon > 0$ hay infinitos elementos de M. Un punto x se denomina aislado de M si existe una bola $B(x,\varepsilon)$, $\varepsilon > 0$ que no contiene ningún elemento M excepto el propio x.*

En la figura 3.2 se muestras esquemáticamente un punto límite que se encuentra en la frontera, un punto interior (que también es un punto límite) y un punto aislado.

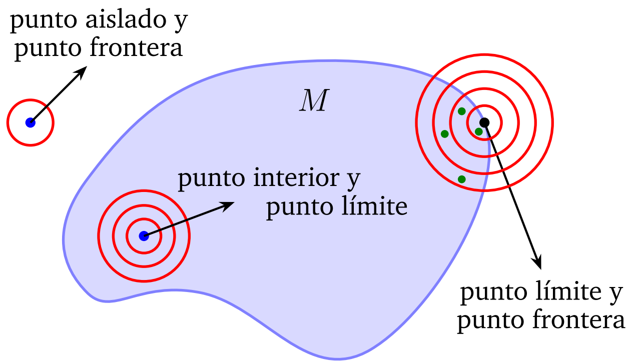


Figura 3.2: Esquema donde se muestra un punto interior, un punto límite y un punto aislado del conjunto M. Los círculos rojos representan los entornos (bolas) centrados en los diferentes.

Nótese que si $x \in M$, entonces obviamente x es un punto de contacto de M. En particular, todo punto interior de M es un punto de contacto. Además, todo punto interior es obviamente un punto de acumulación (ver figura 3.2). Por otro lado, x es un punto de acumulación de M si y solo si es un punto de contacto de $M \setminus \{x\}$. De lo anterior se deduce que los puntos de contacto de M o bien son puntos límites, o bien son aislados. De hecho, si $x \in M$ y no es un punto de acumulación, entonces es un punto aislado.

Nótese además que un punto límite x de M no puede ser un punto interior de $\mathbb{X} \setminus M$ (pues si lo fuese entonces algún entorno de x no contendría puntos de M), luego todo punto límite x o es un punto interior de M o es un punto frontera (ver figura 3.2). Obviamente todo conjunto contiene sus puntos interiores y si es cerrado, como ya vimos en la Proposición 3.2.6, también contiene sus puntos frontera. Luego un conjunto cerrado contiene todos sus puntos límites. De hecho se tiene el siguiente resultado:

Proposición 3.2.10 *Un conjunto $M \subset \mathbb{X}$, $\mathbb{X}$ espacio métrico, es cerrado si y solo si contiene a todos sus puntos límites.*

Demostración: Ya hemos visto que si M es cerrado, entonces contiene todos sus puntos límites. Supongamos ahora que M contiene todos sus puntos límites. Probemos que entonces M contiene todos sus puntos frontera y, por tanto, es cerrado (ver Proposición 3.2.6). Para ello notemos que si x es un punto de la frontera de M, entonces x o es un punto aislado, en cuyo caso $x \in M$, o bien es un punto límite. Probemos está última afirmación. Sea $x \in \partial M$, y supongamos que x no es un punto aislado, entonces en cualquier entorno de x tiene que haber elementos de M que no contengan a x (en caso contrario habría un entorno no contendría a ningún elemento de M a excepción de x y x sería aislado), lo que prueba que x es un punto límite. Pero como M contiene todos sus puntos límites, entonces contendrá todos sus puntos frontera que era lo que queríamos demostrar. ■

Definición 3.2.11 *Dado un subconjunto $M \subset \mathbb{X}$, $\mathbb{X}$ espacio métrico, se denomina clausura de M al conjunto $\overline{M}$ de los elementos de M y sus puntos de contacto.*

De la definición anterior se sigue que $\overline{M} = M \cup \{$conjunto de sus puntos límites$\}$ (¿por qué?). De hecho muchos textos definen la clausura de M al conjunto de de los elementos de M y sus puntos límites.

Por ejemplo, si $\mathbb{X} = \mathbb{Q}$, entonces $\overline{\mathbb{Q}} = \mathbb{R}$ pues todo $x \in \mathbb{R}$ es un punto límite de $\mathbb{Q}$ (¿por qué?).

Proposición 3.2.12 *Un subconjunto $M \subset \mathbb{X}$, $\mathbb{X}$ espacio métrico, es cerrado si y solo si $M = \overline{M}$. Como $M \subset \overline{M}$, entonces $\overline{M}$ es el menor conjunto cerrado que contiene a M.*

Demostración: El resultado se sigue del hecho que $\overline{M} = M \cup \{$conjunto de sus puntos límites$\}$ y la Proposición 3.2.10. No obstante, por completitud, daremos una prueba topológica del mismo.

Sea M cerrado. Como $M \subset \overline{M}$ basta probar que $M \supset \overline{M}$. Como M es cerrado, entonces $\mathbb{X} \setminus M$ es abierto así que para todo $x \in \mathbb{X} \setminus M$ existe una bola $B(x, r)$ completamente contenida en $\mathbb{X} \setminus M$, i.e., $B(x, r)$ no contiene puntos de M, luego $x \notin \overline{M}$ (pues si $x \in \overline{M}$ en cualquier bola $B(x, r)$

tendría que haber al menos un x' de M distinto de x lo cual contradice lo anterior), es decir, $x \in \mathbb{X} \setminus \overline{M}$, así que $\mathbb{X} \setminus M \subset \mathbb{X} \setminus \overline{M}$, i.e. $\overline{M} \subset M$.

Asumamos ahora que $M = \overline{M}$. Probaremos que $\mathbb{X} \setminus M$ es abierto. Sea $x \notin M$, entonces $x \notin \overline{M}$. Entonces $x \in \mathbb{X} \setminus M$. Como $x \notin \overline{M}$ entonces x no es un punto adherente de M así que debe existir al menos una bola $B(x, r)$ que no contiene a ningún elemento de M, es decir, $x \in \mathbb{X} \setminus M$ es un punto interior del conjunto $\mathbb{X} \setminus M$. Como x es arbitrario, $\mathbb{X} \setminus M$ es abierto luego M es cerrado. ∎

La siguiente propiedad de la clausura nos será de utilidad más adelante.

Ejercicio 3.2.13 *Prueba que dados $A, B \subset \mathbb{X}$, $\overline{A \cup B} = \overline{A} \cup \overline{B}$.*

En efecto,

$$A, B \subset A \cup B \Rightarrow \overline{A} \subset \overline{A \cup B} \text{ y } \overline{B} \subset \overline{A \cup B} \ \Rightarrow \ \overline{A} \cup \overline{B} \subset \overline{A \cup B}.$$

Está claro que $A \cup B \subset \overline{A} \cup \overline{B}$, pero $\overline{A \cup B}$ es el menor cerrado que contiene a $A \cup B$, luego cualquier cerrado que contenta a $A \cup B$ tiene que contener a $\overline{A \cup B}$. Así que el cerrado $\overline{A} \cup \overline{B}$ tiene que contener a $\overline{A \cup B}$, i.e., $\overline{A \cup B} \subset \overline{A} \cup \overline{B}$. ∎

Definición 3.2.14 *Un subconjunto $M \subset \mathbb{X}$, $\mathbb{X}$ espacio métrico, es acotado si su diámetro $d(M)$ definido por $d(M) = \sup\limits_{x,y \in M} \rho(x, y)$ es finito.*

Está claro que si $\mathbb{X} = \mathbb{R}$, los conjuntos $(0, 1)$ y $[-2, 1)$ son acotados.

3.3. Aplicaciones en espacios métricos

A lo largo de lo que queda del capítulo asumiremos que $\mathbb{X}$ es un espacio métrico.

Definición 3.3.1 *Por aplicación (operador) o función entenderemos una regla T que le hace corresponder a cada elemento del subconjunto $\mathcal{D}(T) \subset \mathbb{X}$ un **único** elemento del espacio métrico $\mathbb{Y}$. Así, $T : \mathbb{X} \mapsto \mathbb{Y}$, $y = Tx$ o $y = T(x)$, donde $x \in \mathcal{D}(T) \subset \mathbb{X}$ e $y \in \mathbb{Y}$. Al conjunto $\mathcal{D}(T) \subset \mathbb{X}$ se le denomina dominio de la aplicación.*

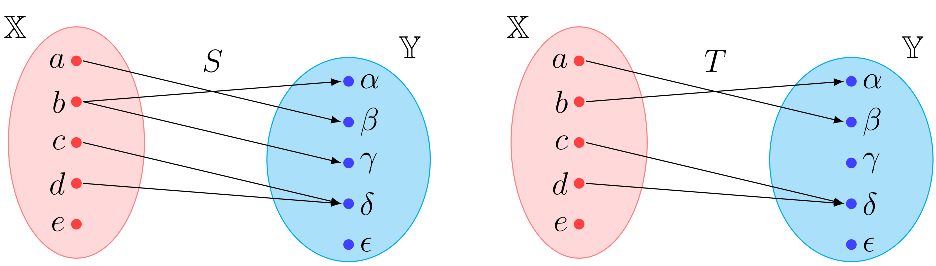


Figura 3.3: $T : \mathbb{X} \mapsto \mathbb{Y}$ (esquema de la derecha) es una aplicación pero S (a la izquierda) no lo es.

Definición 3.3.2 *Sea $x \in \mathcal{D}(T)$ cualquiera u sea $y = Tx \in \mathbb{Y}$. Diremos que Tx es la imagen de x según T. Al conjunto de todas las imágenes Tx le denominaremos imagen (o rango) de T y lo denotaremos por $\mathfrak{I}(T)$.*

Por ejemplo, el dominio de la aplicación T de la figura 3.3 (derecha) es el conjunto $\mathcal{D}(T) = \{a, b, c, d\}$ mientras que su imagen $\mathfrak{I}(T)$ es el conjunto $\{\alpha, \beta, \delta\}$.

Definición 3.3.3 *La imagen inversa de $y \in \mathbb{Y}$ es el conjunto de todas las $x \in \mathcal{D}(T)$ tales que $Tx = y$. La imagen inversa de un subconjunto $M \subset \mathbb{Y}$ es el conjunto de todas las $x \in \mathcal{D}(T)$ tales que $Tx = y$ para todos $y \in M$.*

La imagen inversa de un elemento $y \in \mathbb{Y}$ puede ser el conjunto vacío, un único punto (elemento) de $\mathcal{D}(T)$ o un subconjunto $M \subset \mathcal{D}(T)$.

Por ejemplo, para la aplicación T de la figura 3.3 (derecha) la imagen imagen inversa de δ es el conjunto $\{c, d\}$, la de β es $\{a\}$ y la de γ es el conjunto vacío $\emptyset$.

Definición 3.3.4 *Una aplicación $T : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathbb{Y}$ se llama sobreyectiva si todo elemento y de $\mathbb{Y}$ es imagen de algún elemento x del dominio, es decir, T es tal que*

$$\forall y \in \mathbb{Y}, \quad \exists x \in \mathcal{D}(T) \text{ tal que } Tx = y \quad \Longleftrightarrow \quad \mathfrak{I}(T) \equiv \mathbb{Y}.$$

Definición 3.3.5 *Una aplicación se llama inyectiva si todo elemento y de la imagen de T es imagen a lo sumo de uno y solo un elemento x del* *dominio. Es decir,* $T : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathbb{Y}$ *es tal que*

$$\forall y_1, y_2 \in \mathcal{I}(T), \text{ tales que } y_1 = Tx_1 = y_2 = Tx_2, \quad \Rightarrow \quad x_1 = x_2.$$

O, equivalentemente, si para todo $x_1, x_2 \in \mathcal{D}(T)$ *con* $x_1 \neq x_2$*, se tiene* $Tx_1 \neq Tx_2$*.*

Es decir, una aplicación inyectiva es tal que diferentes puntos tienen diferentes imágenes y, por tanto, la imagen inversa de cada $y \in \mathcal{I}(T)$ es un único elemento de $\mathcal{D}(T)$.

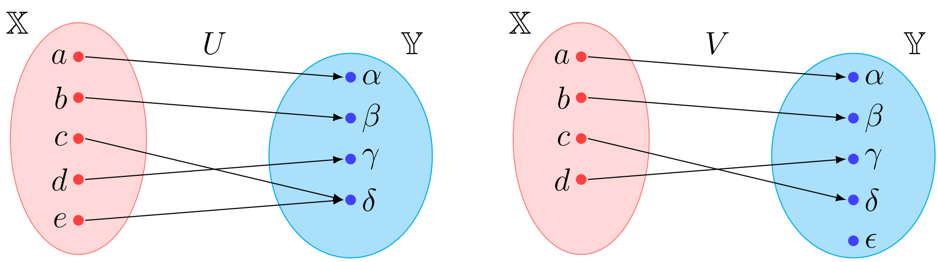


Figura 3.4: La aplicación $U : \mathbb{X} \mapsto \mathbb{Y}$ es sobreyectiva (pero no es inyectiva) y que la aplicación $V : \mathbb{X} \mapsto \mathbb{Y}$ es inyectiva (pero no es sobreyectiva).

Definición 3.3.6 *Una aplicación inyectiva y sobreyectiva se denomina biyectiva.*

Es decir, una aplicación es sobreyectiva si, para todo $y \in \mathbb{Y}$, la ecuación $Tx = y$ tiene al menos una solución, e inyectiva si la ecuación anterior tiene o bien una única solución, o bien no tiene solución. Así mismo, T es biyectiva si para todo $y \in \mathbb{Y}$, la ecuación $Tx = y$ tiene siempre una y solo una solución.

Si una aplicación es inyectiva, podemos definir si inversa.

Definición 3.3.7 *Sea* $T : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathbb{Y}$ *una aplicación inyectiva. Definiremos su aplicación inversa* T^{-1} *a la aplicación* $T^{-1} : \mathcal{I}(T) \subset \mathbb{Y} \mapsto \mathcal{D}(T) \subset \mathbb{X}$ *tal que a cada elemento* $y \in \mathcal{I}(T)$ *le hace corresponder un único* $x \in \mathcal{D}(T)$ *tal que* $Tx = y$*, i.e.,* $x = T^{-1}y$*.*

Para ilustrar los conceptos anteriores mostraremos algunos ejemplos con funciones reales.

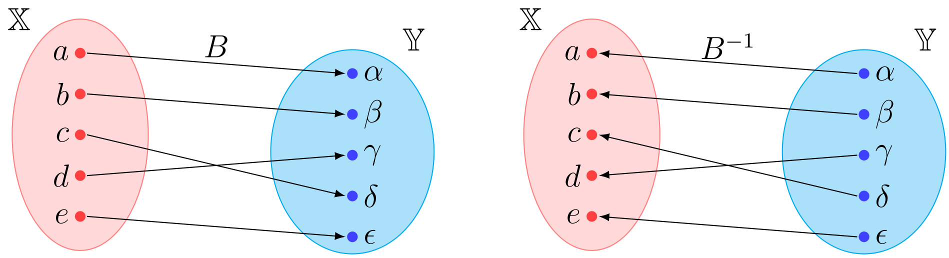


Figura 3.5: La aplicación $B : \mathbb{X} \mapsto \mathbb{Y}$ (izquierda) es biyectiva y $B^{-1} : \mathbb{Y} \mapsto \mathbb{X}$ (derecha) es su inversa.

Por ejemplo, la función $f : \mathbb{R} \mapsto \mathbb{R}$, $f(x) = x^2$ no es inyectiva pues para $f(x) = 4$, por ejemplo, existen dos valores de x del domino tales que $f(x) = 4$, ellos son $x = -2$ y $x = 2$. Esta función tampoco es sobreyectiva pues para $y = -1$ no existe ningún x del dominio tal que $f(x) = -1$ ($f(x) = -1 \iff x^2 = -1$).

Un ejemplo de función sobreyectiva es $h : \mathbb{R} \mapsto \mathbb{R}$, $h(x) = x^3$. La función $f : [0, +\infty) \mapsto [0, +\infty)$, $f(x) = x^2$ es inyectiva pues a cada $y \in f(A)$ le corresponde una $x \in [0, +\infty)$ tal que $f(x) = y$. Dicha x es $x = \sqrt{y}$. Por tanto la función $f : [0, +\infty) \mapsto [0, +\infty)$ tiene inversa y dicha inversa es $f^{-1} : [0, +\infty) \in [0, +\infty)$, $f^{-1}(x) = \sqrt{x}$.

Definición 3.3.8 (Composición de aplicaciones) *Sean $T : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathcal{I}(T) \subset \mathbb{Y}$ y $U : \mathcal{D}(U) \subset \mathbb{Y} \mapsto \mathcal{I}(U) \subset \mathbb{Z}$ dos aplicaciones tales que $\mathcal{I}(T) \subset \mathcal{D}(U)$. Entonces definiremos la aplicación $U \circ T : \mathbb{X} \mapsto \mathbb{Z}$ y la denominaremos aplicación compuesta de U y T a la aplicación que le hace corresponder a cada $x \in \mathcal{D}(T) \subset \mathbb{X}$ un elemento $z \in \mathbb{Z}$ tal que $z = (UT)x := (U \circ T)x = U(Tx)$.*

En general $UTx \neq TUx$, de hecho que exista $U \circ T$ no implica que exista $T \circ U$.

Por ejemplo, sea $f : \mathbb{R} \mapsto \mathbb{R}$, $f(x) = x^2$ y $g : \mathbb{R} \mapsto \mathbb{R}$, $g(x) = x + 2$. Es evidente que la imagen de f está contenida en el dominio de g, por tanto podemos definir la función $(g \circ f)(x) = g(f(x))$:

$$(g \circ f)(x) = g(f(x)) = g(x^2) = x^2 + 2.$$

Además, la imagen de g también está contenida en el dominio de f, por

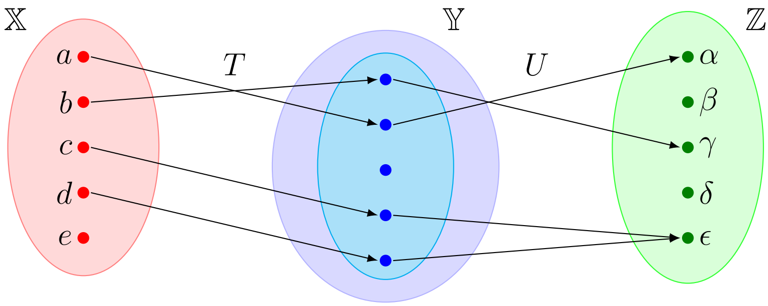


Figura 3.6: Composición de funciones $T \circ U$ de $T : \mathbb{X} \mapsto \mathcal{I}(T) \subset \mathbb{Y}$ y $U : \mathcal{D}(U) \subset \mathbb{Y} \mapsto \mathbb{Z}$, donde $\mathcal{I}(T) \subset \mathcal{D}(U)$.

lo que podemos definir la función compuesta $(f \circ g)(x) = f(g(x))$

$$(f \circ g)(x) = f(g(x)) = f(x+2) = (x+2)^2.$$

Nótese que $(f \circ g)(x) \neq (g \circ f)(x)$. De hecho que exista $(f \circ g)(x)$ no implica que exista $(g \circ f)(x)$. Por ejemplo, si $f : [-1, 1] \mapsto \mathbb{R}$, $f(x) = x$, y $g : [0, 1] \mapsto \mathbb{R}$, $g(x) = x^2$, entonces, por un lado, $(f \circ g)(x) : [0, 1] \mapsto \mathbb{R}$, $(f \circ g)(x) = x^2$, pero por el otro, $(g \circ f)(x)$ no existe.

Nota 3.3.9 *De las definiciones 3.3.7 y 3.3.8 se sigue que si una aplicación T es invertible, entonces $T \circ T^{-1} = I_{\mathbb{Y}}$ y $T^{-1} \circ T = I_{\mathbb{X}}$, donde $I_{\mathbb{Z}}$ es el operador identidad $I : \mathbb{Z} \mapsto \mathbb{Z}$, $Iz = z$ para todo $z \in \mathbb{Z}$ ($\mathbb{Z} = \mathbb{X}$ o $\mathbb{Y}$).*

Definición 3.3.10 *La restricción de una aplicación $T : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathbb{Y}$ a un subconjunto $B \subset \mathcal{D}(T)$ es la aplicación $T|_B : B \mapsto \mathbb{Y}$ que se obtiene de T al restringir el dominio de T al conjunto $B \subset \mathcal{D}(T)$, i.e., $T|_B x = Tx$ para todo $x \in B$. Una extensión de una aplicación $T : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathbb{Y}$ a un subconjunto C tal que $\mathbb{X} \supset C \supset \mathcal{D}(T)$ es una aplicación $\widetilde{T}$ tal que $\widetilde{T}|_{\mathcal{D}(T)} = T$, i.e., $\widetilde{T}x = Tx$ para todo $x \in \mathcal{D}(T)$.*

Definición 3.3.11 *Una aplicación $T : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathbb{Y}$ es continua en $x_0 \in \mathcal{D}(T)$ si para todo $\varepsilon > 0$, existe un $\delta > 0$ tal que para todo $x \in \mathcal{D}(T)$ con $\rho(x, x_0) < \delta$ es tal que[a] $\sigma(Tx, Tx_0) < \varepsilon$. Se dice que T es continua en* *todo $M \subset \mathcal{D}(T)$ si T es continua en todo $x \in M$.*

[a] Aquí ρ denota la métrica de $\mathbb{X}$ y σ la de $\mathbb{Y}$.

Nótese que la definición anterior es equivalente a decir que para cualquier bola $B(Tx_0, \varepsilon)$ de radio ε arbitrario, existe una bola de radio $\delta > 0$, $B(x_0, \delta)$, tal que $T(B(x_0, \delta)) \subset B(Tx_0, \varepsilon)$.

Concluiremos este apartado con una caracterización topológica de las funciones continuas que nos será de utilidad más adelante.

Proposición 3.3.12 *Una aplicación $T : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathbb{Y}$ es continua si y solo si la imagen inversa de cualquier subconjunto abierto (cerrado) de $\mathbb{Y}$ es un subconjunto abierto (cerrado) de $\mathbb{X}$.*

Demostración: Sea T continua y sea $S \subset \mathbb{Y}$ un abierto. Sea S_0 la imagen inversa de S. Si $S_0 = \emptyset$ la proposición es trivial así que asumiremos que $S_0 \neq \emptyset$. Sea $x_0 \in S_0$ cualquiera y sea $y_0 = Tx_0 \in S$ su imagen. Como S es abierto existe una bola $B(y_0, \varepsilon) \subset S$. Pero T es continua así que existe una bola $B(x_0, \delta)$ tal que $T(B(x_0, \delta)) \subset B(y_0, \varepsilon)$. Pero como $B(y_0, \varepsilon) \subset S$, entonces necesariamente $B(x_0, \delta) \subset S_0$, luego $x_0 \in S_0$ es interior y en virtud de que x_0 es arbitrario, S_0 es abierto.

Sea T una aplicación tal que imagen inversa de cualquier subconjunto abierto de $\mathbb{Y}$ es un subconjunto abierto de $\mathbb{X}$. Sea $x_0 \in \mathcal{D}(T)$ cualquiera, sea $S = B(Tx_0, \varepsilon)$ una bola abierta de radio ε arbitrario y sea S_0 la imagen inversa de dicha bola S que, por hipótesis, es un abierto. Como $Tx_0 \in S$, $x_0 \in S_0$. Pero entonces, como $T(S_0) = S$, y S_0 es un abierto, existe una una bola de radio $\delta > 0$, $B(x_0, \delta) \subset S_0$, tal que $T(B(x_0, \delta)) \subset B(Tx_0, \varepsilon) \subset S$, es decir, T es continua en x_0. Como x_0 era arbitrario, entonces T es continua en $\mathcal{D}(T)$. ∎

3.4. Espacios métricos separables

Como ya hemos comentado, un espacio métrico puede contener *muchos* o *pocos* elementos. Es fácil intuir que en un espacio con *pocos* elementos es en general más sencillo para trabajar que en uno con muchos. Un ejemplo de ello es $\mathbb{R}$. Cuando trabajamos con números reales estos pueden ser racionales $\mathbb{Q}$ o irracionales $\mathbb{I}$. En el primer caso es sencillo operar con ellos en la práctica, pero en el segundo es formalmente imposible pues para representar un irracional necesitamos un número infinito de decimales. Sin embargo en los cursos de análisis matemático elementales se prueba que dado un irracional (en general un real) cualquiera siempre se puede encontrar un racional que esté tan cerca como se quiera de dicho irracional. Esta propiedad se conoce como la densidad de $\mathbb{Q}$ en $\mathbb{R}$. Pero aparte de esta ventaja indudable de los racionales frente a los irracionales hay otra más: el conjunto $\mathbb{Q}$ de todos los racionales es numerable mientras que el de los irracionales (y, por tanto, $\mathbb{R}$) no lo es (ser la sección 1.5.1). En este apartado trataremos estas cuestiones para un espacio métrico general.

Definición 3.4.1 *Sea $M \subset \mathbb{X}$ un subconjunto de $\mathbb{X}$. Se dice que M es denso en $\mathbb{X}$ si su clausura coincide con $\mathbb{X}$, i.e., $\overline{M} = \mathbb{X}$.*

De la definición anterior, y teniendo en cuenta que $\overline{M} = M \cup \{$conjunto de sus puntos límites$\}$, se sigue que si M es denso $\mathbb{X}$ entonces cualquiera sea la bola $B(x, \varepsilon)$ (por pequeño que sea $\varepsilon > 0$) siempre contiene puntos de M. En otras palabras, si M es denso en $\mathbb{X}$ entonces siempre hay un elemento m de M tan cerca como se quiera a cualquier $x \in \mathbb{X}$, i.e., para todo $\varepsilon > 0$ existe un $m \in M$ tal que $\rho(x, m) < \varepsilon$. Por el contrario, si para todo $x \in \mathbb{X}$ existe un elemento de M tan cerca como se quiera, entonces o bien $x \in M$ (luego $x \in \overline{M}$), o bien x es punto límite de M, en cuyo caso $x \in \overline{M}$, luego $\mathbb{X} = \overline{M}$.

Por ejemplo, $\mathbb{Q}$ es denso en $\mathbb{R}$ pues, como ya hemos mencionado, para todo $x \in \mathbb{R}$ siempre existe un $q \in \mathbb{Q}$ tan cerca como se quiera de x, luego $\overline{\mathbb{Q}} = \mathbb{R}$.

Recordemos (ver el apartado 1.5.1) que un conjunto M cualquiera se denomina *numerable* si se puede poner en correspondencia biunívoca con $\mathbb{N} = \{1, 2, 3, \dots\}$. Es decir, existe una correspondencia biunívoca entre los elementos de M y los números naturales. Por ejemplo, $\mathbb{Q}$ es numerable, pero $\mathbb{R}$ no lo es.

Definición 3.4.2 *Un espacio métrico $\mathbb{X}$ es separable si contiene un subespacio numerable $M \subset \mathbb{X}$ denso en $\mathbb{X}$.*

Así pues, $\mathbb{R}$ es separable pues $\mathbb{Q}$ es numerable y denso en $\mathbb{R}$.

Ejercicio 3.4.3 *Prueba que $\mathbb{C}$ es separable.*

Basta usar la separabilidad de $\mathbb{R}$ y tener en cuenta la Nota 1.1.6 de la página 3. ∎

Ejemplo 3.4.4 *Estudia la separabilidad del espacio métrico trivial del Ejemplo 3.1.2.*

Es evidente que por la naturaleza de la métrica del espacio del ejemplo 3.1.2 este espacio no puede ser separable pues no en cualquier entorno de $x \in \mathbb{X}$ hay puntos de M (basta escoger $\varepsilon = 1/2$, por ejemplo), a no ser que M sea el propio $\mathbb{X}$. En otras palabras, el único posible subconjunto denso de $\mathbb{X}$ es el propio $\mathbb{X}$, así que $\mathbb{X}$ es separable si y solo si $\mathbb{X}$ es numerable. ∎

Introducción al Análisis Funcional

Prof. Renato Álvarez-Nodarse

Ejemplo 3.4.5 *Prueba que ℓ^2, el espacio del Ejemplo 3.1.12 con $p = 2$, es separable.*

Sea $Q \subset \ell^2$ el espacio de todas las sucesiones del tipo $y = (q_1, q_2, \ldots, q_n, 0, 0, \ldots)$, para todo $n \in \mathbb{N}$, donde para todo k, q_k es racional. Obviamente Q es un subespacio numerable de ℓ^2 (¿por qué?)[a]. Probemos que Q es denso en ℓ^2. Sea $(x_n)_n \in \ell^2$ una sucesión cualquiera de ℓ^2. Entonces para todo $\varepsilon > 0$ existe (¿por qué?) un $N \in \mathbb{N}$ tal que el resto r_n de la serie $\sum_{k=1}^{\infty} |x_k|^2$ satisface que

$$r_n = \sum_{k=n+1}^{\infty} |x_k|^2 < \frac{\varepsilon^2}{2}.$$

Como $\mathbb{Q}$ es denso en $\mathbb{R}$, entonces para todo x_k existe un racional q_k tan cerca como se quiera, luego existe una sucesión $q \in Q$ tal que (¿por qué?)

$$\sum_{k=1}^{n} |x_k - q_k|^2 < \frac{\varepsilon^2}{2}.$$

Entonces, para todo $x \in \ell^2$ podemos encontrar un $q \in Q$ tal que

$$\rho(x,q)^2 = \sum_{k=1}^{n} |x_k - q_k|^2 + \sum_{k=n+1}^{\infty} |x_k - \underbrace{q_k}_{=0}|^2 < \frac{\varepsilon^2}{2} + \frac{\varepsilon^2}{2} = \varepsilon^2,$$

i.e., para todo $x \in \ell^2$, existe un $q \in Q$ tal que $\rho(x,q) < \varepsilon$, luego Q es denso en ℓ^2. ■

[a]Ver Ejemplo 1.5.5.

Ejemplo 3.4.6 *Prueba que ℓ^∞ (ver Ejemplo 3.1.11) no es separable.*

Sea $y = (y_1, y_2, y_3, \ldots)$ una sucesión de ceros y unos. Obviamente $y \in \ell^\infty$. Si a cada y le asociamos un número real que en base dos tenga la forma $r = y_1/2 + y_2/2^2 + \cdots + y_n/2^n + \cdots$, entonces acabamos de establecer una relación biunívoca entre los números reales del intervalo $[0,1]$ y el conjunto de Y de todas las sucesiones de ceros y unos. Como $[0,1]$ no es numerable, Y tampoco lo es. Dadas dos sucesiones y e y' cualesquiera de Y, se tiene que $\rho(y,y') = 1$ si $y \neq y'$ y $\rho(y,y') = 0$ si $y = y'$ (¿por qué?). Sean $y \neq y'$ los centros de sendas bolas de radio 1/3 cada una. Obviamente dichas bolas no tienen puntos en común. Así pues, el conjunto de todas las bolas con centro en algún $y \in Y$ y radio $1/3$ no tiene puntos en común. Sea M cualquier conjunto denso en ℓ^∞. Entonces, en cada una de las bolas anteriores debe haber al menos un elemento de M, luego M no es numerable (el conjunto de las bolas no lo es) y como M era arbitrario, entonces ℓ^∞ no tiene ningún conjunto numerable denso, así que ℓ^∞ no es separable. ■

Ejercicio 3.4.7 *Sea $\mathbb{X}$ un espacio métrico separable y sea $M \subset \mathbb{X}$. Prueba que M también es separable.*

Como $\mathbb{X}$ es separable existe una sucesión $(x_n)_n \in \mathbb{X}$ densa en $\mathbb{X}$. Para cada $n, m \in \mathbb{N}$ elegimos $x \in M$ tal que $\rho(x, x_n) < 1/m$ (si existe, si no, no elegimos ninguno). Denotemos a estos elementos $x_{n,m}$, i.e., $\rho(x_{n,m}, x_n) < 1/m$. Está claro que el conjunto de todos los $x_{n,m}$ es numerable. Sea un $y \in M$ y un $\varepsilon > 0$ cualquiera. Elijamos $m \in \mathbb{N}$ tal que $1/m < \varepsilon/2$. Como $(x_n)_n$ es denso en $\mathbb{X}$, entonces existe un $n \in \mathbb{N}$ tal que $\rho(y, x_n) < 1/m < \varepsilon/2$ (nótese que de lo anterior se sigue que el conjunto de los $x_{n,m}$ es no vacío). Ahora

$$\rho(y, x_{n,m}) \leq \rho(y, x_n) + \rho(x_n, x_{n,m}) \leq \frac{1}{m} + \frac{1}{m} = \frac{2}{m} < \varepsilon.$$

Entonces, para todo $y \in M$ existe un elemento $x_{n,m}$ de un subconjunto numerable de M que está tan cerca como se quiera de él, luego M es separable. ■

Antes de discutir el concepto de convergencia en espacios métricos conviene hablar de otro tipo muy especial de conjuntos: los *conjuntos raros* o *densos en ninguna parte*.

En la Definición 3.4.1 vimos que un subconjunto $M \subset \mathbb{X}$ es denso (en todas partes) en $\mathbb{X}$ si su clausura $\overline{M} = \mathbb{X}$. Esto implica que si M no es denso en $\mathbb{X}$, entonces $\overline{M} \neq \mathbb{X}$, por lo que $\overline{M}$ dejará sin rellenar algún entorno[3] de $\mathbb{X}$. Sin embargo hay otros conjuntos que también es conveniente conocer. Así tenemos la siguiente:

Definición 3.4.8 *Un conjunto $M \subset \mathbb{X}$ es raro o denso en ninguna parte si su clausura $\overline{M}$ no contiene ningún entorno, i.e., el interior de $\overline{M}$ es el conjunto vacío $\emptyset$.*

En otras palabras, un conjunto M es raro si su clausura $\overline{M}$ no contiene puntos interiores. Ello implica que cualquier entorno de $\mathbb{X}$ contiene una bola que es disjunta a M. A diferencia de los conjuntos no densos en general, los no raros son tales que su clausura tiene que rellenar algún entorno de $\mathbb{X}$ pero no necesariamente todo el espacio.

Un conjunto cerrado M que no contiene ningún entorno es raro. Como M es cerrado, $M = \overline{M}$. Si M no fuese raro, entonces $\overline{M}$ contendría un entorno, pero como $M = \overline{M}$, M contendría un entorno lo que es una contradicción.

Proposición 3.4.9 *Sea M un abierto. Entonces $\overline{M} \setminus M$ es raro.*

Demostración: Como M es abierto, entonces no contiene sus puntos frontera (ver Proposición 3.2.5), luego todos los elementos de la frontera están en el cerrado $\overline{M}$ (ver Proposición 3.2.6) y, por tanto, $\overline{M} \setminus M = \partial M$. Pero la frontera ∂M de M es un cerrado que no tiene ningún entorno, luego por lo anterior $\overline{M} \setminus M$ es raro. ∎

Además, se tiene la siguiente caracterización:

Proposición 3.4.10 *Un conjunto $M \subset \mathbb{X}$ es raro si y solo si $\mathbb{X} \setminus \overline{M}$ es denso en $\mathbb{X}$, i.e., $\overline{\mathbb{X} \setminus \overline{M}} = \mathbb{X}$.*

Demostración: ($\Rightarrow$) Supongamos que M es raro pero $\overline{\mathbb{X} \setminus \overline{M}} \neq \mathbb{X}$. Entonces existen $x \in \mathbb{X}$ tales que ni pertenecen a $\mathbb{X} \setminus \overline{M}$ ni son puntos límites

[3] Recuérdese que un entorno de un punto $x \in \mathbb{X}$ es una bola centrada en x.

de $\mathbb{X} \setminus \overline{M}$ (¿por qué?), i.e., hay elementos $x \in \mathbb{X}$ tales que existe una bola (entorno) $U = B(x, \delta)$ que no contiene elementos de $\mathbb{X} \setminus \overline{M}$ y, por tanto, dicho entorno está contenido en $\overline{M}$ lo cual es una contradicción. ($\Leftarrow$) Supongamos ahora que $\overline{\mathbb{X} \setminus \overline{M}} = \mathbb{X}$ pero que M no es raro. Entonces existe un entorno (bola) $U = B(x, \delta) \subset \overline{M}$, luego $U \bigcap \mathbb{X} \setminus \overline{M} = \emptyset$. Pero entonces existen $x \in \mathbb{X}$ para los cuales existe un entorno que no contiene elementos de $\mathbb{X} \setminus \overline{M}$, es decir, existen $x \in \mathbb{X}$ que no son puntos límites de $\mathbb{X} \setminus \overline{M}$, luego $\overline{\mathbb{X} \setminus \overline{M}} \neq \mathbb{X}$, lo que es una contradicción. ■

Definición 3.4.11 *Un conjunto formado por la unión numerable de conjuntos raros se denomina de primera categoría o magro. Si un conjunto no es de primera categoría, entonces se dice que es de segunda categoría.*

Por ejemplo, si $\mathbb{X} = \mathbb{R}$, cualquier conjunto finito de puntos es raro (pues son cerrados que no contienen entornos) y, por tanto, de primera categoría. El conjunto de los números racionales $\mathbb{Q}$ es de primera categoría (es unión numerable de conjuntos raros: $U_n = \{q_n\}$, $q_n \in \mathbb{Q}$). Es conveniente aclarar que un conjunto o bien es de primera categoría o bien es de segunda. Eso implica que si $C = A \cup B$ y A es de primera categoría y C de segunda, entonces B tiene que ser de segunda pues claramente la unión de dos conjuntos de primera categoría es de primera categoría. Una cuestión que no suele ser *trivial* es saber cuando un conjunto es de segunda categoría. Por ejemplo, ¿es $\mathbb{R}$ de segunda categoría? ¿y el conjunto $(-1, 2]$? A esta cuestión regresaremos al final de este capítulo cuando probemos el Teorema de Baire 3.6.1.

Probemos ahora que el conjunto vacío $\emptyset$ es denso en ninguna parte (la palabra *raro* le viene muy bien), por tanto es de primera categoría. Para ello usamos que $\emptyset = \overline{\emptyset}$ ($\emptyset$ es cerrado) por lo que $\mathbb{X} \setminus \overline{\emptyset} = \mathbb{X} \setminus \emptyset = \mathbb{X}$, es decir, $\overline{\mathbb{X} \setminus \overline{\emptyset}} = \mathbb{X}$, luego por la Proposición 3.4.10, se sigue que $\emptyset$ es raro. Este resultado tiene una gran importancia pues de él se deduce que si un conjunto es de segunda categoría, entonces tiene que tener elementos, así que si probamos que un conjunto es de segunda categoría entonces sabemos que dicho conjunto contiene elementos luego no puede ser el conjunto vacío $\emptyset$.

Aunque no es nuestro objetivo tratar demasiado este tipo de cuestiones topológicas conviene mostrar un ejemplo de lo sutil que puede llegar a ser el concepto de conjuntos de primera y segunda categorías. Como mencionamos antes, un conjunto numerable de puntos de $\mathbb{R}$ (por ejemplo, los números naturales) es de primera categoría. El siguiente ejemplo muestra que no siempre es así.

Ejemplo 3.4.12 *Que un conjunto sea de primera o de segunda categoría puede depender del propio espacio métrico.*

Ante todo, recordemos que en $\mathbb{R}$, con la métrica habitual, cualquier conjunto numerable de puntos, en particular $\mathbb{N}$, es de primera categoría. Consideremos ahora el conjunto $\Omega = \{1, 2, 3, \dots\}$ de los números naturales y elijamos en Ω la métrica habitual de $\mathbb{R}$ de forma que Ω sea un espacio métrico. Es sencillo comprobar que los entornos (bolas) en Ω de radio menor que 1 están constituidos por puntos únicos. Sea $M \subset \Omega$ un subconjunto cualquiera no vacío de Ω. Está claro que M no tiene puntos límites (¿por qué?), luego $\overline{M} = M$. Análogamente se tiene que $\mathbb{X} \setminus M = \overline{\mathbb{X} \setminus M}$. Luego, $\overline{\mathbb{X} \setminus \overline{M}} = \overline{\mathbb{X} \setminus M} = \mathbb{X} \setminus M \neq \mathbb{X}$, y, por tanto, usando la Proposición 3.4.10 se concluye que ningún M no vacío es raro. Es decir, el único subconjunto de Ω que es denso en ninguna parte es el conjunto vacío y, por tanto, cualquier subconjunto no vacío de Ω es necesariamente de segunda categoría. ■

3.5. Convergencia en espacios métricos

En esta sección estudiaremos el concepto de convergencia en espacios métricos y alguna de sus consecuencias.

Definición 3.5.1 *Dada una sucesión $(x_n)_n$ de elementos de $\mathbb{X}$, diremos que $(x_n)_n$ es acotada si existe un subconjunto $M \subset \mathbb{X}$ acotado tal que $x_n \in M$ para todo $n \in \mathbb{N}$.*

Lo anterior es equivalente a que exista un $x \in \mathbb{X}$ y un número $K > 0$ tal que $\rho(x, x_n) < K$ para todo $n \in \mathbb{N}$.

Definición 3.5.2 *Una sucesión $(x_n)_n$ de elementos de $\mathbb{X}$ es convergente, y lo denotaremos por $\lim_{n\to\infty} x_n = x$, si existe un $x \in \mathbb{X}$ tal que para todo $\varepsilon > 0$ existe un $N \in \mathbb{N}$ tal que para todo $n > N$, $\rho(x, x_n) < \varepsilon$. En caso contrario diremos que $(x_n)_n$ es divergente.*

Nótese que en la propia definición de límite está explícito que el límite ha de ser un elemento de $\mathbb{X}$. Por ejemplo, sea $\mathbb{X}$ el intervalo abierto $(0, 1)$

con la métrica habitual de $\mathbb{R}$. La sucesión $x_n = 1/(n+1)$ no tiene límite en $\mathbb{X}$ ya que claramente $1/(n+1) \stackrel{n\to\infty}{\longrightarrow} 0$ pero $0 \notin (0,1)$.

Una consecuencia de la propia definición es que, si existe el límite, este es único (ver Problema 3.10).

Ejercicio 3.5.3 *Da una interpretación* geométrica (topológica) *del concepto de límite en un espacio métrico cualquiera.*

Gracias al concepto de convergencia podemos dar una caracterización analítica muy elegante de los conjuntos cerrados.

Proposición 3.5.4 Sea M un subespacio no vacío de un espacio métrico $\mathbb{X}$, y sea $\overline{M}$ su clausura. Entonces

a) $x \in \overline{M}$ si y solo si existe una sucesión $(x_n)_n$ de elementos de M tal que $\lim_{n\to\infty} x_n = x$.

b) M es cerrado si y solo si $\lim_{n\to\infty} x_n = x$ implica que $x \in M$.

Demostración: Probemos a). Si $x \in M$, tomamos la sucesión $x_n = x$. Si $x \notin M$, entonces es un punto de acumulación de M y es *fácil* construir la sucesión x_n (¿por qué?). Por otro lado, si $(x_n)_n \in M$ es tal que $x_n \to x$ entonces, o bien $x \in M$ o bien en todo entorno de x hay un $x_n \neq x$, i.e., x es un punto de acumulación, luego $x \in \overline{M}$. La propiedad b) se deduce de a) teniendo en cuenta que un conjunto es cerrado si y solo si $M = \overline{M}$ (ver Proposición 3.2.12). ■

Definición 3.5.5 *Un espacio métrico $\mathbb{X}$ se denomina (secuencialmente) compacto si cualquier sucesión $(x_n)_n$ de elementos de $\mathbb{X}$ tiene una subsucesión convergente.*

Entenderemos que $M \subset \mathbb{X}$ es compacto si M es compacto como subconjunto de $\mathbb{X}$, i.e., cualquier sucesión $(x_n)_n$ de elementos de M tiene una subsucesión convergente en M.

Ejercicio 3.5.6 *Prueba que el espacio métrico discreto del Ejemplo 3.1.2 constituido por infinitos puntos no es compacto. (**Ayuda:** Basta escoger una sucesión de elementos distintos.)*

Lema 3.5.7 *Si $M \subset \mathbb{X}$ es compacto, entonces M es cerrado y acotado.*

Demostración: Sea M compacto y sea $x \in \overline{M}$ cualquiera. Como $x \in \overline{M}$ entonces existe una sucesión $(x_n)_n$ en M tal que $x_n \stackrel{n\to\infty}{\longrightarrow} x \in \mathbb{X}$. Como M es compacto, entonces $x \in M$ (¿por qué?), luego $M = \overline{M}$ por lo que M es cerrado. Supongamos que M es compacto y no acotado. Entonces existe al menos una sucesión $(x_n)_n$ de elementos de M tal que, fijado un $b \in M$ arbitrario, se tiene que $\rho(x_n, b) > n$ (¿por qué?). Dicha sucesión obviamente no puede tener ninguna subsucesión convergente (pues en caso que la tuviera ésta sería acotada), por tanto, M no puede ser compacto. ■

El recíproco del resultado anterior es falso. Por ejemplo, consideremos el espacio ℓ^2 (ver Ejemplo 3.1.12) y sea M el conjunto formado por las sucesiones $e_k = \delta_{k,i}$, $k = 1, 2, \dots$, i.e., las sucesiones e_k cuyos términos son todos cero excepto el k-ésimo que vale 1. Está claro que para todo $k \in \mathbb{N}$, $\rho(0, e_k) = 1$, luego M es acotado. Además todos los puntos de M son aislados (¿por qué?), por tanto M es cerrado. Ahora bien, como M solo tiene puntos aislados, M no tiene ningún punto de acumulación por lo tanto ninguna sucesión que escojamos de elementos *distintos* de M contiene una subsucesión convergente.

Recordemos que una aplicación $T : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathbb{Y}$ es continua –ver la Definición 3.3.11– en $x_0 \in \mathcal{D}(T)$ si para todo $\varepsilon > 0$, existe un $\delta > 0$ tal que para todo $x \in \mathcal{D}(T)$ con $\rho(x, x_0) < \delta$ es tal que $\sigma(Tx, Tx_0) < \varepsilon$. T es continua en $M \subset \mathcal{D}(T)$ si es continua en todo $x \in M$.

Ejercicio 3.5.8 *Prueba que la Definición 3.3.11 de continuidad en un punto es equivalente a la siguiente definición por sucesiones: T es continua en x_0 si para cualquier sucesión $(x_n)_n$ con $x_n \stackrel{n\to\infty}{\longrightarrow} x_0$, se tiene que $Tx_n \stackrel{n\to\infty}{\longrightarrow} Tx_0$.*

En efecto, supongamos T es continua según la Definición 3.3.11 pero no lo es según la segunda. Entonces, ha de existir una sucesión $(x_n)_n$ tal que $x_n \stackrel{n\to\infty}{\longrightarrow} x_0$ pero $Tx_n \stackrel{n\to\infty}{\not\longrightarrow} Tx_0$, es decir, dada dicha sucesión x_n, existe un $\varepsilon > 0$ tal que cualquiera sea $N \in \mathbb{N}$, existe un $n > N$ tal que $\sigma(Tx_n, Tx_0) \geq \varepsilon$. Ahora bien, como $x_n \stackrel{n\to\infty}{\longrightarrow} x_0$, entonces cualquiera sea $\delta > 0$ existe un $N \in \mathbb{N}$ tal que para todo $n > N$, $\rho(x_n, x_0) < \delta$. Uniendo las dos condiciones anteriores deducimos que

$$\exists \varepsilon > 0, \quad \forall \delta > 0, \quad \exists x = x_n,\ \rho(x_n, x_0) < \delta \quad \Rightarrow \quad \sigma(Tx_n, Tx_0) \geq \varepsilon,$$

lo cual contradice nuestra hipótesis.[4]

Supongamos ahora que T es continua según la definición por sucesiones pero no lo es según la Definición 3.3.11. Entonces existe un $\varepsilon > 0$ tal que cualquiera sea $\delta > 0$ ha de existir un x tal que $\rho(x, x_0) < \delta$ que cumple con $\sigma(Tx, Tx_0) \geq \varepsilon$. Escojamos $\delta = 1/n$. Entonces, para cada $n \in \mathbb{N}$ existe un x_n tal que $\rho(x_n, x_0) < 1/n$, es decir, $x_n \to x_0$ pero $\sigma(Tx_n, Tx_0) \geq \varepsilon$. Luego $Tx_n \not\to Tx_0$, lo cual es una contradicción. ■

Teorema 3.5.9 *Sea $T : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathbb{Y}$ continua en el compacto $M \subset \mathcal{D}(T)$. Entonces la imagen de M, $T(M)$ también es un conjunto compacto. O sea, las aplicaciones continuas transforman compactos en compactos.*

Demostración: Sea una sucesión $(y_n)_n \in T(M) \subset \mathbb{Y}$ cualquiera. Probemos que existe una subsucesión $(y_{n_k})_k$ convergente en $\mathbb{Y}$. Para ello notemos que como para todo $n \in \mathbb{N}$, $y_n \in T(M) \subset \mathbb{Y}$, entonces existe x_n tal que $Tx_n = y_n$. Como M es compacto, entonces existe una subsucesión $(x_{n_k})_k$ convergente en M, pero entonces, por la continuidad de T, $Tx_{n_k} = y_{n_k} \to Tx \in T(M)$ cuando $k \to \infty$. Luego toda sucesión de $T(M)$ tiene una subsucesión convergente y, por tanto, $T(M)$ es compacto. ■

Corolario 3.5.10 *Sea una aplicación continua $T : M \subset \mathbb{X} \mapsto \mathbb{R}$ de un compacto M en los reales. Entonces T alcanza su máximo y su mínimo absolutos.*

Este corolario es una generalización del Teorema de Weierstrass para las funciones continuas.

Demostración: Del teorema anterior se sigue que $T(M) \subset \mathbb{R}$ es compacto. Entonces, por el Lema 3.5.7, $T(M)$ es cerrado y acotado y, por tanto, $\inf(T(M))$ y $\sup(T(M))$ están contenidos en $T(M)$, i.e., existen x_1 y x_2 de M tales que $T(x_1) = \inf(T(M))$ y $T(x_2) = \sup(T(M))$, de donde se sigue el resultado. ■

3.5.1. Espacios métricos completos

[4]Veamos una prueba directa de esta implicación. Sea $\varepsilon > 0$, cualquiera. Como que T es continua en x_0, existe $\delta > 0$ tal que si $\rho(x, x_0) \leq \delta$ entonces, $\sigma(Tx, Tx_0) \leq \varepsilon$. Sea $(x_n)_n$ cualquiera tal que $x_n \to x_0$. Entonces para todo $\delta > 0$ existe $N \in \mathbb{N}$ tal que para todo $n > N$, $\rho(x_n, x) < \delta$, pero entonces, por la continuidad de T se tiene que $\sigma(Tx_n, Tx_0) < \varepsilon$, de donde se sigue el resultado.

Definición 3.5.11 *Una sucesión $(x_n)_n$ de elementos de $\mathbb{X}$ se denomina de Cauchy o fundamental si existe para todo $\varepsilon > 0$ existe un $N \in \mathbb{N}$ tal que para todo $n > N$ y todo $p \in \mathbb{N}$, $\rho(x_n, x_{n+p}) < \varepsilon$.*

En $\mathbb{R}$ toda sucesión es convergente si y solo si es de Cauchy. Esta propiedad fundamental de $\mathbb{R}$ no es cierta para cualquier espacio métrico $\mathbb{X}$. Por ejemplo, si escogemos nuevamente $\mathbb{X}$ como el intervalo abierto $(0, 1)$ con la métrica habitual de $\mathbb{R}$, la sucesión $x_n = 1/(n+1)$, que es de Cauchy (¿por qué?) no tiene límite en $\mathbb{X}$.

Definición 3.5.12 *Un espacio métrico $\mathbb{X}$ se denomina completo si y solo si toda sucesión de Cauchy de elementos de $\mathbb{X}$ converge (a un elemento de $\mathbb{X}$).*

Por ejemplo, el espacio $\mathbb{X} = \mathbb{R}$ con la métrica usual de $\mathbb{R}$, es completo. También lo es $\mathbb{X} = \mathbb{C}$ con la métrica usual de $\mathbb{C}$. Sin embargo $\mathbb{Q}$, el conjunto de los números racionales, es incompleto (¿por qué?), y el conjunto $\mathbb{X} = (0, 1)$ de antes también lo es.

Proposición 3.5.13 *Sea $(x_n)_n$ una sucesión convergente de elementos de un espacio métrico $\mathbb{X}$. Entonces $(x_n)_n$ es de Cauchy.*

Demostración: Como $\lim_{n\to\infty} x_n = x$, entonces, para todo $\varepsilon > 0$ existe un $N \in \mathbb{N}$ tal que para todo $n > N$, $\rho(x, x_n) < \varepsilon/2$. Luego, para todo p

$$\rho(x_n, x_{n+p}) \leq \rho(x_n, x) + \rho(x, x_{n+p}) < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

con tal que $n > N$, i.e., $(x_n)_n$ es de Cauchy. ∎

Teorema 3.5.14 *Un subespacio M de un espacio métrico completo $\mathbb{X}$ es completo si y solo si es cerrado en $\mathbb{X}$.*

Demostración: Sea M completo. Probaremos que $M = \overline{M}$. Sea $x \in \overline{M}$ cualquiera, entonces existe una sucesión (ver Proposición 3.5.4) de elementos de M que converge a x. Entonces $(x_n)_n$ es de Cauchy (pues es convergente) pero M es completo, luego $x \in M$, i.e., $M = \overline{M}$.

Sea M cerrado (i.e., $M = \overline{M}$) y $(x_n)_n$ una sucesión de Cauchy en M. Entonces $\lim_{n\to\infty} x_n = x$ con $x \in \mathbb{X}$ ($\mathbb{X}$ es completo). Pero entonces $x \in \mathbb{X}$ es un punto de acumulación de M, i.e., $x \in \overline{M}$ y como M es cerrado $x \in M$, i.e., toda sucesión de Cauchy en M tiene límite en M, i.e., M es completo. ∎

Nota 3.5.15 *Nótese que en la prueba de la primera parte no se usa la completitud de $\mathbb{X}$. Es decir, se tiene que si un subespacio $M \subset \mathbb{X}$ es completo, entonces es cerrado en $\mathbb{X}$.*

Ejemplo 3.5.16 *El espacio métrico ℓ^2 del Ejemplo 3.1.12 ($p = 2$) es completo.*

Sea $x^{(n)} := (x_1^{(n)}, x_2^{(n)}, \dots, x_k^{(n)}, \dots)$ el n-ésimo término de una sucesión $(x^{(n)})_n$, de Cauchy de ℓ^2, i.e., $\sum_{k=1}^{\infty} |x_k^{(n)}|^2 < +\infty$. Entonces, para todo $\varepsilon > 0$

$$\sum_{k=1}^{\infty} |x_k^{(n)} - x_k^{(n+p)}|^2 < \varepsilon^2, \qquad \forall p \in \mathbb{N},$$

con tal que n sea lo suficientemente grande. Pero entonces, como todos los sumandos son positivos, podemos asegurar que, para todo $k \in \mathbb{N}$,

$$|x_k^{(n)} - x_k^{(n+p)}| < \varepsilon, \qquad \forall p \in \mathbb{N},$$

y n suficientemente grande. Es decir, cada una de las sucesiones reales (o complejas) $(x_k^{(n)})_n$, $k = 1, 2, 3, \dots$, es de Cauchy. Pero $\mathbb{R}$ (o $\mathbb{C}$) es un espacio completo por lo que existe $\lim_{n\to\infty} x_k^{(n)} = x_k$ para cada k. Sea $x = (x_1, x_2, \dots, x_k, \dots)$. Probemos que $x \in \ell^2$ y que $x^{(n)}$ converge a x en la métrica de ℓ^2. Como $(x^{(n)})_n$, de Cauchy, entonces para todo p si n es lo suficientemente grande las sumas parciales satisfacen que

$$\sum_{k=1}^{N} |x_k^{(n)} - x_k^{(n+p)}|^2 < \varepsilon^2, \qquad \forall N \in \mathbb{N}.$$

Como $x_k^{(n)}$ converge a x_k para cada k, tomando el límite cuando $p \to \infty$ tenemos que, para todo N,

$$\sum_{k=1}^{N} |x_k^{(n)} - x_k|^2 \leq \varepsilon^2 \quad \Rightarrow \quad \sum_{k=1}^{\infty} |x_k^{(n)} - x_k|^2 \leq \varepsilon^2$$

o, equivalentemente, si n es suficientemente grande, $\rho(x^{(n)}, x) \leq \varepsilon$. Luego hemos probado que $x^{(n)}$ converge a x en la métrica de ℓ^2. Finalmente, usando la desigualdad ($0 = (0, 0, 0, \dots)$)

$$\rho(x, 0) \le \rho(x, x^{(n)}) + \rho(x^{(n)}, 0),$$

y usando que $x^{(n)}$ es de ℓ^2, se tiene que $\rho(x, 0)$ está acotada, por lo que $x \in \ell^2$. Al ser $x^{(n)}$ una sucesión fundamental arbitraria, se sigue que ℓ^2 es completo. ■

De forma análoga se puede probar que ℓ^p, $p \ge 1$ es completo.

Ejemplo 3.5.17 *El espacio $C^2_{[a,b]}$ del Ejemplo 3.1.10 no es completo.*

Vamos a probar que existen sucesiones fundamentales de funciones continuas cuyo límite no puede ser, en la métrica de $C^2_{[a,b]}$, una función continua. Por sencillez tomaremos $[a, b] = [-1, 1]$, y construyamos la sucesión de funciones $(f_n)_n$ en $[-1, 1]$ definida por [a]

$$f_n(x) = \begin{cases} -1, & -1 \le x < -\frac{1}{n}, \\ nx, & -\frac{1}{n} \le x \le \frac{1}{n}, \\ 1, & \frac{1}{n} < x \le 1. \end{cases}$$

Para cada $x \in [-1, 1]$, f_n converge puntualmente a f definida por

$$\lim_{n\to\infty} f_n(x) = f(x) = \begin{cases} -1, & -1 \le x < 0, \\ 0, & x = 0, \\ 1, & 0 < x \le 1. \end{cases}$$

Por otro lado, como para todo $x \in [-1, 1]$, $|f_n(x) - f_{n+p}(x)| \le 1$, tenemos que $|f_n(x) - f_{n+p}(x)|^2 \le |f_n(x) - f_{n+p}(x)|$. Vamos a estimar la distancia entre dos funciones f_n y f_{n+p}. El área entre las funciones f_n y f_{n+p} está representada en la gráfica de la derecha de la figura 3.7. De dicha gráfica se deduce que dicha área está acotada por la de los cuadrados con vértices $(0, 0)$, $(0, 1)$, $(1/n, 1)$, $(1/n, 0)$ y $(0, 0)$, $(0, -1)$, $(-1/n, -1)$, $(-1/n, 0)$ que, junto con la desigualdad anterior, implica que

$$\int_{-1}^{1} |f_n(x) - f_{n+p}(x)|^2 dx \le \int_{-1}^{1} |f_n(x) - f_{n+p}(x)| dx < \frac{2}{n} \quad \Rightarrow$$

$$\rho(f_n, f_{n+p}) = \sqrt{\int_{-1}^{1} |f_n(x) - f_{n+p}(x)|^2 dx} < \sqrt{\frac{2}{n}} \overset{n\to\infty}{\longrightarrow} 0, \quad \forall p \in \mathbb{N},$$

es decir, $(f_n)_n$ es una sucesión fundamental. Además, se tiene que

$$\rho(f_n, f) = \sqrt{\int_{-1}^{1} |f_n(x) - f(x)|^2 dx} \leq \sqrt{\frac{2}{n}} \overset{n\to\infty}{\longrightarrow} 0.$$

Por otro lado, dada una función continua cualquiera $g \in C^2_{[a,b]}$ se cumple la desigualdad [b]

$$\rho(f, g) \leq \rho(g, f_n) + \rho(f_n, f). \tag{3.5.1}$$

Probemos que $\rho(f, g) \neq 0$. Como g es continua en $[-1, 1]$ entonces $g - f$ es continua en $I = [-1, 0) \cup (0, 1]$ (pues f lo es I) y no puede ser idénticamente nula en I pues si lo fuese entonces $\lim_{x\to 0-} g(x) = -1$ y $\lim_{x\to 0+} g(x) = 1$ lo cual es una contradicción (g es continua en 0). Entonces tiene que existir un entorno $U \subset I$ tal que $|f(x) - g(x)| > 0$, y en dicho entorno $\int_U |f(x) - g(x)| dx \neq 0$, pues $|f(x) - g(x)|$ es una función continua positiva de donde se sigue que existe un $d > 0$ tal que $\rho(f, g) \geq d > 0$.

Por otro lado, dado que $\rho(f_n, f)$ se puede hacer tan pequeño como se quiera, entonces de la desigualdad (3.5.1) se deduce que $\rho(f_n, g) \geq \rho(f, g) \geq d > 0$ por lo que no existe ninguna función $g \in C^2_{[a,b]}$ que sea límite de $(f_n)_n$ en la métrica de $C^2_{[a,b]}$. Luego $C^2_{[a,b]}$ no es completo. ■

[a] Un ejemplo análogo es el de la sucesión de funciones $f_n : \mathbb{R} \mapsto \mathbb{R}$, $f_n(x) = \arctan nx$. Otro más sencillo es definir $h_n(x) = 0$ para $x \in [-1, 0]$ y $h_n(x) = f_n(x)$ para $x \in (1/n, 1]$.

[b] Aunque $f \notin C^2_{[a,b]}$, la desigualdad es cierta como consecuencia de la desigualdad de Minkowski para integrales (3.1.4).

Ejemplo 3.5.18 *Probar que el espacio métrico $C^\infty_{[a,b]}$ del Ejemplo 3.1.9 es completo.*

Sea el espacio de las funciones continuas $C([a, b])$ con la métrica $\rho(f, g) = \max_{x\in[a,b]} |f(x) - g(x)|$ y sea $(f_n(x))_n$ una sucesión de cualquiera

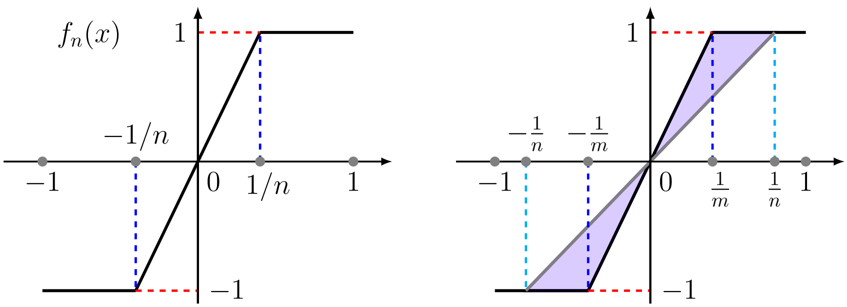


Figura 3.7: La función $f_n : [-1, 1] \mapsto [-1, 1]$, del Ejemplo 3.5.17 (izquierda) y el área entre las funciones f_n y f_{n+p} (derecha).

de Cauchy en dicho espacio. Entonces

$$\forall \varepsilon > 0, \quad \exists N \in \mathbb{N} \,|\, \forall n > N, \forall p \in \mathbb{N}, \quad \max_{x\in[a,b]} |f_n(x) - f_{n+p}(x)| < \varepsilon/2,$$

Luego por el Corolario 1.4.11 f_n es uniformemente convergente. Como $f_n \rightrightarrows f$ en $[a, b]$, entonces, por el Teorema 1.4.22, $f \in C([a, b])$. Así, tomando el límite $p \to \infty$ en la expresión anterior tenemos

$$\forall \varepsilon > 0, \quad \exists N \in \mathbb{N} \,|\, \forall n > N, \quad \rho(f_n, f) = \max_{x\in[a,b]} |f_n(x) - f(x)| < \varepsilon,$$

es decir, toda sucesión fundamental es convergente en $C^\infty_{[a,b]}$ y, por tanto, $C^\infty_{[a,b]}$ es completo. ∎

Definición 3.5.19 *La sucesión de esferas (bolas cerradas) $(S_n(x_n, r_n))_n$, $S_n(x_n, r_n) \subset \mathbb{X}$, tales que para todo $n \in \mathbb{N}$*

$$S_1(x_1, r_1) \supset S_2(x_2, r_2) \supset \cdots \supset S_n(x_n, r_n) \supset S_{n+1}(x_{n+1}, r_{n+1}) \supset \cdots ,$$

se denomina sucesión de esferas (cerradas) encajadas.

Teorema 3.5.20 (De las esferas encajadas) *Sea $\mathbb{X}$ un espacio métrico. Entonces, $\mathbb{X}$ es completo si y solo si, cualquier sucesión de esferas encajadas cuyos radios tiendan a cero ($r_n \overset{n\to\infty}{\longrightarrow} 0$) tiene intersección no vacía, i.e., $\bigcap_{n=1}^\infty S_n(x_n, r_n) \neq \emptyset$.*

Demostración: Sea $\mathbb{X}$ completo y sea $S_n(x_n, r_n)$, $n = 1, 2, 3, \ldots$, una sucesión de esferas encajadas con $r_n \overset{n\to\infty}{\longrightarrow} 0$. Entonces la sucesión $(x_n)_n$ de los centros de las esferas son una sucesión de Cauchy (¿por qué?). Como $\mathbb{X}$ es completo, entonces existe $\lim_{n\to\infty} x_n = x \in \mathbb{X}$. Probemos que $x \in \bigcap_{n=1}^{\infty} S_n(x_n, r_n)$. Como las esferas son encajadas, entonces cualquiera sea $n \in \mathbb{N}$, la esfera S_n contiene los elementos $x_n, x_{n+1}, x_{n+2}, \ldots$, así que x es un punto límite de S_n, para todo n, pero como S_n es cerrada $x \in S_n$ para todo n, luego $x \in \bigcap_{n=1}^{\infty} S_n(x_n, r_n)$.

Probemos ahora la implicación contraria. Sea $(x_n)_n$ una sucesión fundamental arbitraria de $\mathbb{X}$, i.e.,

$$\forall\varepsilon > 0,\ \exists N \in \mathbb{N},\ \mid \forall n > m > N, \quad \rho(x_n, x_m) < \varepsilon. \tag{*}$$

Sea $\varepsilon = 1/2$. Como $(x_n)_n$ es de Cauchy, existe $n_1 \in \mathbb{N}$ tal que para todo $n > n_1$, $\rho(x_n, x_{n_1}) < 1/2$ (tomamos $N = n_1 - 1$ en (*)). Sea la esfera $S_1 := S(x_{n_1}, 1)$. Está claro que $x_n \in S_1$ para todos los $n > n_1$.

Sea ahora $\varepsilon = 1/2^2$. Entonces existe $n_2 > n_1 \in \mathbb{N}$ tal que para todo $n > n_2$, $\rho(x_n, x_{n_2}) < 1/2^2$. Además, como $n_2 > n_1$, tenemos que $\rho(x_{n_2}, x_{n_1}) < 1/2$. Sea la esfera $S_2 := S(x_{n_2}, 1/2)$. Está claro que, por un lado $x_{n_2} \in S_1$ y que $x_n \in S_2$ para todos los $n > n_2$. Y así sucesivamente. Esta construcción está representada en la figura 3.8.

De esta forma construimos una sucesión de esferas $S_k := S(x_{n_k}, 1/2^{k-1})$ (representadas por los círculos rojos centrados en x_{n_k} de la figura 3.8) cuyos centros son los elementos de la subsucesión $(x_{n_k})_k$ de $(x_n)_n$ que cumplen que $\rho(x_{n_{k+1}}, x_{n_k}) < 1/2^k$. Nótese que los centros x_{n_m} con $m > k$ siempre están en el interior de las bolas abiertas $B(x_{n_k}, 1/2^k)$ (representadas por los círculos azules con líneas discontinuas centrados en x_{n_k} de la figura 3.8). Probemos que $S_{k+1} \subset S_k$. Para ello probaremos que para cualquiera sea $x \in S_{k+1}$, se tiene que $x \in S_k$. En efecto, si $x \in S_{k+1}$, entonces $\rho(x, x_{n_{k+1}}) \leq 1/2^k$, luego la distancia entre x y x_{n_k} cumple con

$$\rho(x, x_{n_k}) \leq \rho(x, x_{n_{k+1}}) + \rho(x_{n_{k+1}}, x_{n_k}) < \frac{1}{2^k} + \frac{1}{2^k} = \frac{1}{2^{k-1}} \quad \Rightarrow \quad x \in S_k,$$

i.e., $S_{k+1} \subset S_k$ y, por tanto, la sucesión $(S_k)_k$ es una sucesión de esferas encajadas cuyos radios, por construcción, son tales que $r_k = 1/2^{k-1} \to 0$. Pero como, por hipótesis, cualquier sucesión de esferas encajadas cuyos radios tiendan a cero tiene intersección no vacía, entonces existe $x \in \bigcap_{k=1}^{\infty} S_k$, y *obviamente* $\lim_{k\to\infty} x_{n_k} = x$ (¿por qué?). Pero $(x_n)_n$ es una

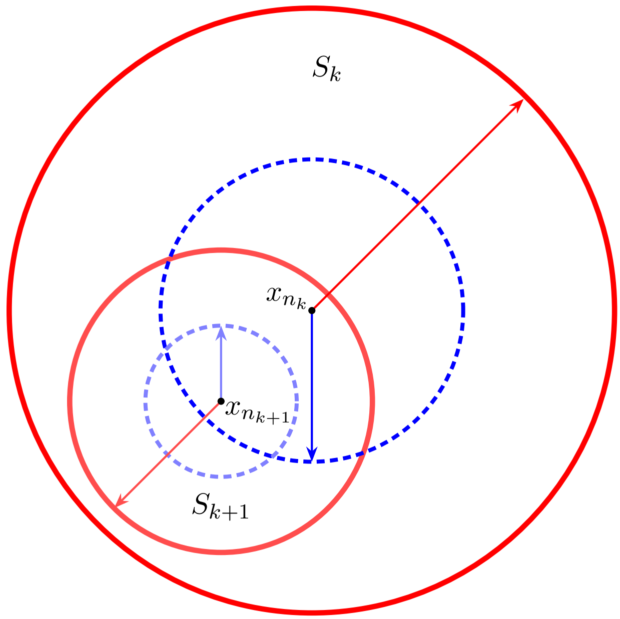


Figura 3.8: Esquema de la construcción de la sucesión de esferas encajadas $S_k := S(x_{n_k}, 1/2^{k-1})$.

sucesión fundamental y hemos probado que tiene una subsucesión convergente, luego la propia sucesión converge a x (ver el apartado 2 del Problema 3.13). Como $(x_n)_n$ era arbitraria, $\mathbb{X}$ es completo. ∎

Ejercicio 3.5.21 *Prueba que si $\mathbb{X}$ es completo, entonces $\bigcap_{n=1}^{\infty} S_n(x_n, r_n)$, con $r_n \to 0$, contiene un único punto. (**Ayuda:** Usa el teorema anterior y reducción al absurdo).*

Basta notar que, si para todo $n \in \mathbb{N}$, los elementos $x \neq y$ pertenecen a todas las esferas S_n, entonces $\rho(x, y) \leq 2r_n \to 0$, lo que implica que $x = y$. ∎

Apliquemos el teorema anterior al espacio métrico $\mathbb{R}$. Como sabemos $\mathbb{R}$ es completo, luego toda sucesión de esferas encajadas, en este caso los intervalos cerrados $[x_n - r_n, x_n + r_n]$, tales que $r_n \overset{n\to\infty}{\longrightarrow} 0$ tienen al menos un punto en común $\bigcap_{n=1}^{\infty}[x_n - r_n, x_n + r_n] \neq \emptyset$. De hecho tienen uno y solo uno. Lo anterior no es más que el Teorema de los intervalos encajados de Cantor en $\mathbb{R}$.

Nota 3.5.22 *Para la prueba del Teorema de las esferas encajadas 3.5.20 ha sido esencial el hecho de que la intersección infinita* $\bigcap_{n=1}^{\infty} S_n(x_n, r_n)$ *sea no vacía, lo cual no es cierto en general. Por ejemplo, si tomamos los conjuntos* $M_n = [n, \infty) \subset \mathbb{R}$*, entonces* $M_{n+1} \subset M_n$ *pero* $\bigcap_{n=1}^{\infty} M_n = \emptyset$*. En este caso el Teorema de las esferas encajadas no es aplicable (¿por qué?).*

Veamos, sin embargo, que si los conjuntos encajados son conjuntos compactos no vacíos, entonces su intersección infinita no es vacía. Así, probemos que una sucesión de conjuntos compactos no vacíos $(S_n)_n$ *son tales que, para todo* $n \in \mathbb{N}$*,* $S_{n+1} \subset S_n$*, entonces* $\bigcap_{n=1}^{\infty} S_n \neq \emptyset$*. Además, en el caso de que los diámetros* $d(S_n) \to 0$*, entonces dicha intersección consta de un único punto.*

Para demostrarlo vamos a elegir de cada S_n *un elemento* x_n *y vamos construir con ellos la sucesión* $(x_n)_n$*. Está claro que, por construcción,* $x_m \in S_k$ *para todos* $m \geq k$*. Como* S_1 *es compacto entonces existe una subsucesión* $(x_{n_k})_k$ *de* $(x_n)_n$ *que es convergente. Denotemos por* x *al límite de dicha subsucesión. Está claro que* $x \in S_1$*. Notemos además que, por construcción, todos los términos de* $(x_{n_k})_k$*, excepto quizá el primero pertenecen a* S_2 *(¿por qué?), luego* x *es un punto límite de* S_2*, y como* S_2 *es compacto, y por tanto cerrado,* $x \in S_2$*. Pero todos los términos de* $(x_{n_k})_k$*, excepto quizá los dos primeros pertenecen a* S_3*, luego luego* x *es un punto límite de* S_3*, y por tanto* $x \in S_3$*. Y así sucesivamente. De esta forma tenemos que* x *es punto límite de todas los* S_n *y por tanto* $x \in S_n$*, para todo* n*. Luego* $x \in \bigcap_{n=1}^{\infty} S_n$*, lo que prueba el resultado. Eso si, de la prueba se desprende que puede haber más elementos de* $\mathbb{X}$ *en la intersección de todos los* S_n*. Para poder asegurar que solo haya uno es necesario asumir que los diámetro* $d(S_n) \to 0$*, prueba que dejamos al lector.*

Para terminar este apartado vamos a discutir el concepto de completamiento de un espacio métrico. Para ello necesitamos la siguiente definición:

Definición 3.5.23 *Sea* $T : \mathbb{X} \mapsto \mathbb{Y}$ *una aplicación biyectiva del espacio métrico* $(\mathbb{X}, \rho)$ *al espacio métrico* $(\mathbb{Y}, \sigma)$*. Diremos que* T *es una isometría si*

$$\forall x_1, x_2 \in \mathbb{X}, \qquad \rho(x_1, x_2) = \sigma(Tx_1, Tx_2).$$

Dos espacios métricos $(\mathbb{X}, \rho)$ *e* $(\mathbb{Y}, \sigma)$ *son isométricos si existe una isometría entre* $(\mathbb{X}, \rho)$ *e* $(\mathbb{Y}, \sigma)$*.*

Esto tiene una implicación muy importante: si dos espacios son isométricos, entonces las distancias de los elementos originales y de sus imágenes según T son las mismas, es decir, los espacios solo difieren por la naturaleza del conjunto $\mathbb{X}$ pero son *idénticos* según la métrica. En particular los conceptos *topológicos* (entorno, cercanía, etc) son equivalentes en ambos espacios.

Definición 3.5.24 *Sea $\mathbb{X}$ un conjunto cualquiera y sea ρ una función tal que $(\mathbb{X}, \rho)$ sea un espacio métrico. Llamaremos completamiento de $(\mathbb{X}, \rho)$ al espacio métrico completo $(\mathbb{X}^*, \sigma)$ tal que $\mathbb{X} \subset \mathbb{X}^*$ y $\overline{\mathbb{X}} = \mathbb{X}^*$.*

Es decir, el conjunto $\mathbb{X}$ es un subconjunto denso de $\mathbb{X}^*$ en el espacio métrico $(\mathbb{X}^*, \sigma)$.

Por ejemplo, el conjunto de todos los reales $\mathbb{R}$ es el completamiento del conjunto de los racionales $\mathbb{Q}$ usando en ambos casos la métrica de $\mathbb{R}$ (ver Ejemplo 3.1.3).

Ejercicio 3.5.25 *¿Cuál es el completamiento del espacio $(0, 1)$ discutido anteriormente?*

Teorema 3.5.26 *Todo espacio métrico $(\mathbb{X}, \rho)$ tiene un completamiento. Dicho completamiento es único salvo isometrías. Es decir, si $\mathbb{X}^*$ y $\mathbb{X}^{**}$ son dos completamientos de $\mathbb{X}$, entonces existe una aplicación $T : \mathbb{X}^* \mapsto \mathbb{X}^{**}$, $x^{**} = Tx^*$ tal que $Tx = x$ para todo $x \in \mathbb{X}$ y $\rho^*(x^*, y^*) = \rho^{**}(Tx^*, Ty^*)$.*

La demostración de este Teorema se puede encontrar, por ejemplo, en [10, §3.4 pág. 76] y [11, §1.6, pág. 41].

Nota 3.5.27 *Una consecuencia del teorema anterior es que el espacio $C^2_{[a,b]}$ del ejemplo 3.5.17 se puede completar. El completamiento de dicho espacio se denota por $L^2_{[a,b]}$ y es el espacio de las funciones de cuadrado integrable, es decir, los espacios de las funciones tales que $\int_a^b |f(x)|^2 dx < +\infty$, donde se considera que dos funciones son iguales si y solo si difieren en un conjunto de medida nula. De forma análoga se pueden definir los espacios $L^p_{[a,b]}$ de las funciones tales que $\int_a^b |f(x)|^p dx < +\infty$.El estudio detallado de estos espacios requiere el uso de muchos resultados de la Teoría de la medida por lo que no lo consideraremos aquí. El lector interesado puede consular, por* *ejemplo [5, Capítulo 2], [14, Capítulo 13] y [23, Capítulo 10].*

3.5.2. El Teorema del punto fijo

Definición 3.5.28 *Sea $T : \mathbb{X} \mapsto \mathbb{X}$ una aplicación. Si existe un $\alpha \in (0,1)$ tal que*

$$\forall x, y \in \mathbb{X} \quad \Rightarrow \quad \rho(Tx, Ty) \leq \alpha\rho(x, y),$$

diremos que T es una aplicación de contracción.

Ejercicio 3.5.29 *Prueba que toda aplicación de contracción es continua.*

Usando la definición de continuidad mediante sucesiones (ver Ejercicio 3.5.8) tenemos que probar que si $x_n \to x$, entonces $Tx_n \to Tx$. Pero como T es de contracción $\rho(Tx_n, Tx) \leq \alpha\rho(x_n, x) \to 0$, de donde se tiene el resultado. Nótese que no es necesario que $\alpha \in (0,1)$ para probar el resultado. ■

Definición 3.5.30 *Sea $T : \mathbb{X} \mapsto \mathbb{X}$ una aplicación. El punto $x \in \mathbb{X}$ se denomina punto fijo de T si $Tx = x$.*

Teorema 3.5.31 (Del punto fijo) *Sea $\mathbb{X}$ un espacio métrico completo y $T : \mathbb{X} \mapsto \mathbb{X}$ una aplicación de contracción. Entonces T tiene un único punto fijo.*

Demostración: Sea $x_0 \in \mathbb{X}$ cualquiera. Construyamos la siguiente sucesión

$$x_0, \quad x_1 = Tx_0, \quad x_2 = Tx_1, \quad x_n = Tx_{n-1} \quad \Rightarrow x_n = T^n x_0.$$

Probemos que $(x_n)_n$ es fundamental. En efecto,

$$\begin{aligned}
\rho(x_n, x_{n+p}) &= \rho(T^n x_0, T^n(T^p x_0)) \leq \alpha\rho(T^{n-1}x_0, T^{n-1}(T^p x_0)) \leq \dots \\
&\leq \alpha^n \rho(x_0, T^p x_0) = \alpha^n \rho(x_0, x_p) \\
&\leq \alpha^n \left[\rho(x_0, x_1) + \rho(x_1, x_2) + \cdots + \rho(x_{p-1}, x_p)\right] \\
&\leq \alpha^n \left[\rho(x_0, x_1) + \rho(Tx_0, Tx_1) + \cdots + \rho(T^{p-1}x_0, T^{p-1}x_1)\right] \\
&\leq \alpha^n \left[\rho(x_0, x_1) + \alpha\rho(x_0, x_1) + \cdots + \alpha^{p-1}\rho(x_0, x_1)\right] \\
&= \alpha^n \rho(x_0, x_1) \frac{1-\alpha^p}{1-\alpha} \leq \frac{\alpha^n}{1-\alpha}\rho(x_0, x_1).
\end{aligned}$$

Luego, cualquiera sea $\varepsilon > 0$, existe un $N \in \mathbb{N}$, tal que para todo $n > N$ y todo $p \in \mathbb{N}$, $\rho(x_n, x_{n+p}) < \varepsilon$ (¿por qué?), i.e., $(x_n)_n$ es fundamental. Como $\mathbb{X}$ es completo entonces dicha sucesión tiene límite, i.e., $\lim_{n\to\infty} x_n = x$. Pero entonces, como T es continua $\lim_{n\to\infty} Tx_n = Tx$, pero $Tx_n = x_{n+1}$ y $\lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} x_n = x$, así que tomando límites en $Tx_n = x_{n+1}$ obtenemos $x = Tx$.

Para probar la unicidad asumamos que existen dos puntos fijos x e y con $x \neq y$. Así, $Tx = x$ y $Ty = y$, pero

$$\rho(x, y) = \rho(Tx, Ty) \leq \alpha\rho(x, y) \;\Rightarrow\; (1-\alpha)\rho(x, y) \leq 0 \;\Rightarrow\; \rho(x, y) = 0,$$

que implica $x = y$, lo cual es una contradicción. ■

Corolario 3.5.32 *Sea $\mathbb{X}$ un espacio métrico completo y sea $T : \mathbb{X} \mapsto \mathbb{X}$ una aplicación de contracción. Entonces cualquiera sea el elemento $x_0 \in \mathbb{X}$ la sucesión x_0, $x_1 = Tx_0$, . . . , $x_n = Tx_{n-1}$, . . . , tiene un único punto límite que coincide con el punto fijo de T. Además, $\rho(x_n, x) \leq \frac{\alpha^n}{1-\alpha}\rho(x_0, x_1)$.*

Demostración: La prueba es inmediata a partir de la desigualdad

$$\rho(x_n, x_{n+p}) \leq \frac{\alpha^n}{1-\alpha}\rho(x_0, x_1),$$

tomando el límite cuando $p \to \infty$. ■

Ejemplo 3.5.33 Como ejemplo *sencillo* consideremos las funciones reales en $f : [a, b] \mapsto \mathbb{R}$ tales que para todos x_1 e x_2 de $[a, b]$ se satisface la condición de Lipschitz con $K \in (0, 1)$, i.e.,

$$|f(x_1) - f(x_2)| \leq K|x_1 - x_2|, \qquad K \in (0, 1).$$

Entonces, f es una aplicación de contracción y por el Teorema del punto fijo la sucesión

$$x_0, \quad x_1 = f(x_0), \quad x_2 = f(x_1), \quad x_{n+1} = f(x_n), \quad \ldots$$

converge a un único límite x tal que $x = f(x)$. En particular, f satisface la condición de Lipschitz con $K \in (0, 1)$ si f es diferenciable y $|f'(x)| \leq K < 1$ en $[a, b]$. ■

Definición 3.5.34 *Sea $(T_n)_n$ una sucesión*[a] *de aplicaciones $T_n : \mathbb{X} \mapsto \mathbb{Y}$ y sea $M \subset \mathcal{D}(T)$. Diremos que T_n converge puntualmente en M a $T : \mathbb{X} \mapsto \mathbb{Y}$, si para todo $x \in M$ la sucesión $(T_n x)_n$ es convergente, i.e., para cada $x \in M$ $\lim_{n\to\infty} T_n x = Tx = y \in \mathbb{Y}$.*

[a]Se asume que todos los operadores T_n tienen el mismo dominio $\mathcal{D}(T_n)$.

En otras palabras

$$\forall \varepsilon > 0 \quad \text{y} \quad \forall x \in M \quad \exists N := N(\varepsilon, x) \in \mathbb{N}; \quad \forall n > N \;\Rightarrow\; \rho(T_n x, Tx) < \varepsilon.$$

Está claro de la definición anterior que el número N depende no solo del valor de ε sino también del elemento x. Para diferentes x tendremos en general diferentes N.

Definición 3.5.35 *Sea $(T_n)_n$ una sucesión de aplicaciones $T_n : \mathbb{X} \mapsto \mathbb{Y}$ y sea $M \subset \mathcal{D}(T)$. Diremos que T_n converge uniformemente en M a $T : \mathbb{X} \mapsto \mathbb{Y}$, si*

$$\forall \varepsilon > 0 \quad \exists N := N(\varepsilon) \in \mathbb{N}; \quad \forall n > N \text{ y } \forall x \in M \;\Rightarrow\; \rho(T_n x, Tx) < \varepsilon.$$

Es decir, fijado el $\varepsilon > 0$, podemos escoger un N tal que la desigualdad $\rho(T_n x, Tx) < \varepsilon$ es cierta en todo el subconjunto M.

Este concepto es bastante delicado incluso en el caso de las funciones reales como ya vimos en el capítulo 1 de estas notas.

Una aplicación del Teorema del punto fijo

Apliquemos el Teorema del punto fijo vamos a probar el siguiente resultado:

Teorema 3.5.36 (Picard) *Sea $f(x,y)$ una función continua en $\Omega \subset \mathbb{R}^2$, $\Omega = \{(x,y) |\ |x-x_0| \le a,\ |y-y_0| \le b\}$. Supongamos además que f satisface la condición de Lipschitz en y, en Ω, i.e.,*

$$\exists K > 0;\ \forall x, y, \tilde{y} \in \Omega,\ |f(x,y) - f(x,\tilde{y})| \le K|y - \tilde{y}|.$$

Entonces, existe un δ-entorno de x_0, donde el problema de valores iniciales

$$y'(x) = f(x,y), \quad y(x_0) = y_0,$$

tiene solución única.

Demostración: Ante todo notemos que el problema de valores iniciales es equivalente a la siguiente ecuación integral

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

La idea de la prueba es la siguiente. Vamos a probar que la aplicación

$$T : C^* \mapsto C^*, \quad \phi = Ty \quad \phi(x) = y_0 + \int_{x_0}^{x} f(t, y(t))dt,$$

es una aplicación de contracción en cierto espacio completo C^*.

Como f es continua en Ω que es cerrado y acotado, entonces (por el Teorema de Weierstrass para las funciones continuas) existe un $M > 0$ tal que para todos $(x, y) \in \Omega$, $|f(x, y)| \leq M$. Dado $\delta > 0$ tal que $K\delta < 1$, sea el conjunto cerrado y acotado $\Omega' \subset \Omega$, $\Omega' = \{(x, y)|\ |x - x_0| \leq \delta,\ |y - y_0| \leq M\delta\}$. Está claro que en Ω' también se tiene que $|f(x, y)| \leq M$.

Denotemos por C^* el espacio de las funciones continuas $y(x)$ en el intervalo I definido por $|x - x_0| \leq \delta$ y tales que $|y(x) - y_0| \leq M\delta$ (dicho intervalo I existe pues $y(x)$ es continua). Definamos en C^* la métrica $\rho(g, h) = \max_{x \in I} |g(x) - h(x)|$. Probemos que C^* es cerrado en el espacio C_I^∞ de las funciones continuas en I y, por tanto, completo (¿por qué?). Para ello tomemos una sucesión $(y_n(t))_n \in C^*$ cualquiera tal que $y_n(t) \stackrel{n\to\infty}{\longrightarrow} y(t)$. Como $(y_n(t))_n \in C^*$ entonces, para todo n y cualquiera sea $t \in I$ se cumple que $|y_n(t) - y_0| \leq M\delta$, de donde se sigue

$$|y(t) - y_0| \leq |y(t) - y_n(t)| + |y_n(t) - y_0| \leq \max_{x \in I} |y(t) - y_n(t)| + M\delta.$$

Tomando el límite cuando $n \to \infty$ en la expresión anterior y usando que $y_n(t) \stackrel{n\to\infty}{\longrightarrow} y(t)$, se tiene que $|y(t) - y_0| \leq M\delta$, i.e., $y(t) \in C^*$, luego, por el apartado b) de la Proposición 3.5.4 C^* es cerrado.

Definamos sobre C^* la siguiente aplicación T:

$$\phi = Ty \quad \phi(x) = y_0 + \int_{x_0}^{x} f(t, y(t))dt.$$

Probemos que T transforma C^* en C^*. En efecto, si $y(x) \in C^*$, entonces para todo $t \in I$, $|y(t) - y_0| \leq M\delta$ y $|f(t, y(t))| \leq M$, luego

$$|Ty - y_0| = |\phi(x) - y_0| \leq \int_{x_0}^{x} |f(t, y(t))|dt \leq M\delta \Rightarrow Ty \in C^*.$$

Además, cualesquiera sean y e $\tilde{y}$ de C^*, $\phi = Ty$, $\tilde{\phi} = T\tilde{y}$,

$$|\phi - \tilde{\phi}| \leq \int_{x_0}^{x} |f(t, y(t)) - f(t, \tilde{y}(t))| dt \leq K\delta \max_{x \in I} |y - \tilde{y}|,$$

de donde se sigue, tomando $\max$ en I que $\rho(\phi, \tilde{\phi}) \leq K\delta\rho(y, \tilde{y})$, con $K\delta < 1$. O sea, T es de contracción luego, por el Teorema 3.5.31, T tiene un único punto fijo, y por tanto la ecuación $Ty = y$ tiene solución única, i.e., el problema de valores iniciales tiene una única solución en el espacio C^* de las funciones continuas en $I = [x_0 - \delta, x_0 + \delta]$ tales que $|y(x) - y_0| \leq M\delta$. ■

Nótese que del corolario 3.5.32 se sigue que la solución del problema de valores iniciales puede ser aproximada mediante la sucesión

$$y_0(x) = y_0, \quad y_1(x) = y_0 + \int_{x_0}^{x} f(t, y_0(t)) dt \quad \cdots$$
$$y_n(x) = y_0 + \int_{x_0}^{x} f(t, y_{n-1}(t)) dt \quad \cdots$$

conocida como sucesión de Picard.[5]

3.6. El Teorema de las categorías de Baire

Vamos a terminar este capítulo demostrando un teorema importante relacionado con los espacios métricos completos que nos será de utilidad más adelante.

Teorema 3.6.1 (Baire) *Todo espacio métrico completo $\mathbb{X} \neq \emptyset$ es de segunda categoría.*

Demostración: Supongamos que $\mathbb{X}$ es magro (de primera categoría), entonces

$$\mathbb{X} = \bigcup_{k=1}^{\infty} M_k, \quad M_k \subset \mathbb{X} \text{ conjuntos raros.} \tag{3.6.1}$$

Sea M_1 raro. Entonces $\overline{M_1}$ no puede contener un entorno (abierto). Pero $\mathbb{X}$ si que puede contener un abierto (el propio $\mathbb{X}$), luego $\mathbb{X} \neq \overline{M_1} \neq \emptyset$, de

[5]Compárese con lo discutido en el apartado 2.2.

donde se sigue que $C(\overline{M_1}) := \mathbb{X} \setminus \overline{M_1}$ es no vacío. Sea $x_1 \in C(\overline{M_1})$. Como $x_1 \notin \overline{M_1}$, x_1 no puede ser ni punto de acumulación ni punto frontera de M_1 (¿por qué?) y, por tanto, ha de haber una bola[6] $B_1 := B(x_1, \varepsilon_1) \subset C(\overline{M_1})$, $\varepsilon_1 < 1/2$ ($\overline{M_1}$ es cerrado, luego $C(\overline{M_1})$ es abierto y x_1 es un punto interior de $C(\overline{M_1})$). Nótese que $B_1 \bigcap M_1 = \emptyset$.

Sea ahora M_2, que también es raro y, por tanto, $\overline{M_2}$ no puede contener abiertos. Entonces $\overline{M_2}$ no puede contener a la bola $B(x_1, \varepsilon_1/2)$, por tanto $C(\overline{M_2}) \bigcap B(x_1, \varepsilon_1/2) \neq \emptyset$ es no vacío y abierto, luego podemos elegir un x_2 y una bola $B_2 := B(x_2, \varepsilon_2)$, $\varepsilon_2 < \varepsilon_1/2 < 1/2^2$, tal que $B_2 \subset C(\overline{M_2}) \bigcap B(x_1, \varepsilon_1/2) \subset B_1$ (está claro que $B(x_1, \varepsilon_1/2) \subset B(x_1, \varepsilon_1) = B_1$) con $B_2 \bigcap M_2 = \emptyset$ (B_2 está en el complementario de $\overline{M_2}$). Y así sucesivamente.

De esta forma tendremos una sucesión de bolas abiertas[7] (entornos) $B_1, B_2, \ldots, B_n, \ldots$, con $\varepsilon_k < 1/2^k$ tales que $B_k \bigcap M_k = \emptyset$, y

$$B_{k+1} = B(x_{k+1}, \varepsilon_{k+1}) \subset B(x_k, \varepsilon_k/2) \subset B(x_k, \varepsilon_k) = B_k.$$

Es más, para todo $n > m$

$$B_n = B(x_n, \varepsilon_n) \subset B(x_m, \varepsilon_m/2) \subset B(x_m, \varepsilon_m) = B_m,$$

por tanto $\rho(x_n, x_m) < 1/2^m$, lo que significa que la sucesión $(x_n)_n$ de los centros de B_n es de Cauchy. Como $\mathbb{X}$ es completo, $x_n \to x$, $x \in \mathbb{X}$. Además, de la inclusión anterior $B_n \subset B(x_m, \varepsilon_m/2)$ se tiene que $x_n \in B(x_m, \varepsilon_m/2)$, y, por tanto, de la desigualdad triangular se sigue que para todo $n > m$

$$\rho(x_m, x) \leq \rho(x_m, x_n) + \rho(x_n, x) \leq \frac{\varepsilon_m}{2} + \rho(x_n, x) \to \frac{\varepsilon_m}{2}$$

cuando $n \to \infty$. Luego, $\rho(x_m, x) \leq \varepsilon_m/2 < \varepsilon_m$, i.e., $x \in B_m$ para todo $m \in \mathbb{N}$. Como $B_m \in C(\overline{M_m})$, entonces, para todo m, $x \notin M_m$, lo cual contradice (3.6.1). Luego $\mathbb{X}$ no puede ser de primera categoría lo cual prueba el teorema. ■

Corolario 3.6.2 *Sea un espacio métrico completo $\mathbb{X} \neq \emptyset$. Supongamos que*

$$\mathbb{X} = \bigcup_{k=1}^{\infty} M_k, \qquad M_k \subset \mathbb{X}, \quad M_k \textit{ conjuntos cerrados}.$$

[6]Recuérdese que las bolas son conjuntos abiertos.

[7]En la prueba se puede también razonar con cerrados y usar el Teorema de las esferas encajadas (ver Problema 3.22).

Entonces al menos un M_k contiene un abierto no vacío (un entorno).

Demostración: Si fuese falso todos los M_k serían conjuntos raros (pues al ser cerrados $\overline{M_k} = M_k$) y $\mathbb{X}$ sería un conjunto de primera categoría (magro) lo cual contradice el Teorema de Baire. ■

Nótese que dado un espacio métrico $\mathbb{X}$, si $\mathbb{X}$ es una unión numerable de conjuntos M_k, donde los subconjuntos $M_k \subset \mathbb{X}$, $k = 1, 2, \dots$, no son necesariamente cerrados, entonces

$$\mathbb{X} = \bigcup_{k=1}^{\infty} M_k = \bigcup_{k=1}^{\infty} \overline{M_k},$$

donde la última igualdad se sigue de que cambiar los M_k por sus clausuras no adiciona ningún nuevo elemento a $\mathbb{X}$. Si tenemos en cuenta el corolario 3.6.2 ello nos dice que si $\mathbb{X}$ es completo y es una unión numerable de conjuntos M_k, entonces al menos uno de los M_k es tal que su clausura $\overline{M_k}$ tiene interior no vacío.

Otro corolario del Teorema de Baire es que todo espacio métrico completo $\mathbb{X}$ no constituido por puntos aislados es necesariamente no numerable pues si lo fuese entonces $\mathbb{X} = \bigcup_{k=1}^{\infty} \{x_k\}$, pero cada conjunto $\{x_k\}$ constituido por un único punto x_k es raro y por tanto $\mathbb{X}$ sería de primera categoría, lo cual contradice el Teorema de Baire. De lo anterior se sigue, por ejemplo, que $\mathbb{R}$ no es numerable.

Dos aplicaciones "sencillas" al análisis clásico

El Teorema de Baire nos será de gran utilidad para probar algunos teoremas clásicos del análisis funcional como veremos más adelante. No obstante, en este apartado veremos dos ejemplos "*sencillos*" de aplicación al análisis clásico.

Ejemplo 1. No es complicado probar que existen funciones continuas en los irracionales y discontinuas en los racionales para distintos subconjuntos de $\mathbb{R}$ (ver, por ejemplo, el Problema 3.21). ¿Y al contrario? Es decir, ¿existen funciones continuas en los racionales y discontinuas en los irracionales? La respuesta es no.

Vamos a probar que no existen funciones definidas en $[0, 1]$ que sean continuas en cada racional de $(0, 1)$ y discontinuas en cada irracional de $(0, 1)$. Para ello necesitamos ciertas definiciones previas.

Sea $a \in I \subset \mathbb{R}$, I abierto, y sea $I_\delta(a) = (a-\delta, a+\delta) \subset I$. Definamos las cantidades

$$\omega(f, I) = \sup_{x,y \in I} |f(x) - f(y)|, \quad \omega(f, a) = \lim_{\delta \to 0} \omega(f, I_\delta(a)),$$

conocidas como *oscilaciones* de f en el abierto I y el punto $a \in I$, respectivamente. Es claro de lo anterior que para todo abierto $U \subset I$, $\omega(f, U) \leq \omega(f, I)$, luego $\omega(f, a) \leq \omega(f, I_\delta(a)) \leq \omega(f, I)$. Nótese que si $\omega(f, a) = 0$, entonces para todo $\varepsilon > 0$ existe un $\delta > 0$ tal que $\omega(f, I_\delta(a)) < \varepsilon$. Por tanto,

$$|f(x) - f(a)| \leq \sup_{x,y \in I_\delta(a)} |f(x) - f(y)| = \omega(f, I_\delta(a)) < \varepsilon,$$

de donde se sigue que f es continua en $x = a$. Además, de la definición de $\omega(f, a)$ se deduce que si f es continua en a, $\omega(f, a) = 0$. Así, tenemos la siguiente propiedad:

Propiedad I: Una función f es continua en $x = a$ si y solo si $\omega(f, a) = 0$.

Por otro lado, el conjunto $U(x, \varepsilon)$ de las $x \in I$ tales que $\omega(f, x) < \varepsilon$, para todo $\varepsilon > 0$ es abierto. Para ello tomemos $\varepsilon > 0$ y sea $x_0 \in U(x, \varepsilon)$. Sea $J \subset U(x, \varepsilon)$ un abierto que contenga a x_0 tal que $\omega(f, J) < \varepsilon$. Entonces, para todo $y \in J$ se tiene que $\omega(f, y) \leq \omega(f, J) < \varepsilon$, i.e., $y \in U(x, \varepsilon)$ y por tanto $J \subseteq U(x, \varepsilon)$, luego $U(x, \varepsilon)$ es abierto.

Propiedad II: Para cada $\varepsilon > 0$, el conjunto $U(x, \varepsilon) = \{x \in I \,|\, \omega(f, x) < \varepsilon\}$ es abierto.

Supongamos que tenemos una función continua en cada racional de $(0, 1)$ y discontinua en cada irracional de $(0, 1)$. Para ello usaremos que $\mathbb{X} = (0, 1)$ es un espacio métrico con la métrica habitual de $\mathbb{R}$. Sean los conjuntos

$$U_n = \{x \in (0, 1) \,|\, \omega(f, x) < 1/n\}, \quad n = 1, 2, \ldots.$$

De la Propiedad II se sigue que, para todo $n \in \mathbb{N}$, los U_n son abiertos. Como f es continua en los racionales $r_n \in (0, 1)$, la Propiedad I implica $\omega(f, r_n) = 0$, luego $\bigcap_{n=1}^\infty U_n = (0, 1) \bigcap \mathbb{Q}$, es decir, la intersección de todos los U_n solo pueden ser los racionales de $(0, 1)$ (pues f es discontinua en los irracionales, luego $\omega(f, i) \geq \varepsilon > 0$ para todo i irracional). Como los racionales son densos en $(0, 1)$, entonces cada U_n es denso en $(0, 1)$ (pues la intersección de todos los U_n son los racionales de $(0, 1)$). Sean ahora los conjuntos $V_n = (0, 1) \setminus U_n$, $n = 1, 2, \ldots$, que son cerrados (son los complementarios de los U_n que eran abiertos). Como los U_n son densos en $(0,1)$, entonces los V_n son raros en $(0,1)$ pues $(0,1) = \mathbb{X} = \overline{U_n} = \overline{\mathbb{X} \setminus V_n} = \mathbb{X} \setminus \overline{V_n}$ (ver la Proposición 3.4.10). Además[8], $(0,1) \setminus \mathbb{Q} = \bigcup_{n=1}^{\infty} V_n$. Como $\mathbb{Q}$ es numerable, y, por tanto, $\mathbb{Q} \bigcap (0,1)$ también, podemos construir una sucesión $(q_n)_n$ que enumera a los racionales de $(0,1)$. Entonces

$$(0,1) = V_1 \bigcup \{q_1\} \bigcup V_2 \bigcup \{q_2\} \bigcup \cdots$$

y, por tanto,

$$[0,1] = \{0\} \bigcup \{1\} \bigcup V_1 \bigcup \{q_1\} \bigcup V_2 \bigcup \{q_2\} \bigcup \cdots$$

que es una unión numerable de conjuntos raros y, por tanto, de primera categoría. Pero $[0,1]$ es completo (es un cerrado de $\mathbb{R}$), luego, por el Teorema de Baire es de segunda categoría, lo cual es una contradicción. ■

Ejemplo 2. Sea $C^{\infty}_{[0,1]}$ el espacio métrico de las funciones continuas en $[0,1]$ con la métrica del $\rho(f,g) = \max_{x \in [0,1]} |f(x) - g(x)|$. Como ya vimos en el Ejemplo 3.5.18, este espacio es completo. Sea A el espacio de las funciones continuas que tienen al menos en un punto una derivada lateral finita. Entonces $A \subset C^{\infty}_{[0,1]}$ es un conjunto de primera categoría.

Es decir, tenemos el siguiente teorema:

Teorema 3.6.3 *El conjunto de todas las funciones continuas que no tienen ninguna derivada finita en ningún punto del intervalo $[0,1]$ es de segunda categoría.*

Demostración: Para probar el teorema construiremos la siguiente familia de subconjuntos E_n de las funciones continuas $f \in C([0,1])$ tales que para cada $n \in \mathbb{N}$

$$E_n = \left\{ \frac{|f(x+h) - f(x)|}{h} \le n, \ \forall h \in (0, 1/n), \ \text{para algún } x \in [0, 1 - 1/n] \right\}.$$

Si una función continua es tal que en algún punto su derivada lateral por la derecha existe, entonces el cociente $\frac{|f(x+h)-f(x)|}{h}$, $h > 0$, es acotado (¿por qué?), de donde se deduce que todo elemento de A, el espacio de las funciones continuas que tienen al menos en un punto una derivada

[8] Ello se sigue de la ley de Morgan $(\mathbb{X} \setminus A) \bigcup (\mathbb{X} \setminus B) = \mathbb{X} \setminus (A \bigcap B)$.

lateral por la derecha finita, pertenece a algún E_n, y, por tanto, las que tienen derivada finita también. Está claro que el conjunto de los E_n es numerable.

Demostremos que E_n es cerrado para todo n. Para ello probaremos que si $f \in \overline{E_n}$, entonces $f \in E_n$. Como $f \in \overline{E_n}$, entonces existe una sucesión $(f_k)_k \in E_n$ tal que $f_k \to f$, con f continua (recordemos que en la métrica de $C^\infty_{[0,1]}$ la convergencia es uniforme). Como $f_k \in E_n$, existe para cada k un $x \in [0, 1-1/n]$, que denotaremos por x_k, tal que

$$\frac{|f_k(x_k+h) - f(x_k)|}{h} \leq n, \quad x_k \in [0, 1-1/n].$$

Como x_k es una sucesión acotada, podemos encontrar una subsucesión convergente $x_{k_j} \to x \in [0, 1-1/n]$. Para dicha subsucesión tenemos

$$\frac{|f_{k_j}(x_{k_j}+h) - f(x_{k_j})|}{h} \leq n, \quad x_{k_j} \in [0, 1-1/n].$$

Pero como $f_k \to f$, entonces $f_{k_j} \to f$, luego, tomando límite cuando $j \to \infty$ tenemos que

$$\frac{|f(x+h) - f(x)|}{h} \leq n,$$

para cierto $x \in [0, 1-1/n]$, i.e., $f \in E_n$.

Nota: Existe otra forma de probar que $f \in E_n$ que es la siguiente. Como $f \in E_n$, entonces existe $(f_k)_k \in E_n$ tal que $f_k \to f$, con f continua Ahora bien, como $f_k \to f$ tenemos que para todo $\varepsilon > 0$, existe un $N \in \mathbb{N}$, tal que para todo $k > N$,

$$\max_{x\in[0,1]} |f(x) - f_k(x)| < \frac{1}{2} h\varepsilon.$$

Por otro lado,

$$\frac{|f(x+h) - f(x)|}{h} \leq \frac{|f(x+h) - f_k(x+h)|}{h} + \frac{|f_k(x+h) - f_k(x)|}{h} + \frac{|f_k(x) - f(x)|}{h}.$$

Luego, luego para todo $\varepsilon > 0$ y todo $k > N$ tenemos que el primer y último sumandos son menores o iguales que $\varepsilon/2$ y, por tanto,

$$\frac{|f(x+h) - f(x)|}{h} \leq \frac{|f_k(x+h) - f_k(x)|}{h} + \varepsilon.$$

Pero, para cada $\varepsilon > 0$ y cada k existe un $x \in [0, 1-1/n]$ tal que

$$\frac{|f_k(x+h) - f_k(x)|}{h} \le n,$$

pues $f_k \in E_n$, de donde se sigue que

$$\frac{|f(x+h) - f(x)|}{h} \le n,$$

y, por tanto, $f \in E_n$, luego E_n es cerrado.

Probemos ahora que $\overline{C([0,1]) \setminus E_n} = C([0,1])$, entonces por la Proposición 3.4.10, E_n es raro y, por tanto, la unión de todos los E_n es de primera categoría, luego A también.

Notemos que $D_n = C([0,1]) \setminus E_n$ es el conjunto de las $f \in C([0,1])$ tales que para cada $n \in \mathbb{N}$

$$D_n = \left\{ \frac{|f(x+h) - f(x)|}{h} > n, \text{ para algún } h \in (0, 1/n), \text{ y } \forall x \in [0, 1-1/n] \right\}.$$

Para probar que $\overline{D_n} = C([0,1])$ lo que haremos en demostrar que en cualquier entorno de $C([0,1])$ hay un elemento de D_n, o lo que es lo mismo, que para cualquiera sea $\phi \in C([0,1])$ y cualquiera sea $\varepsilon > 0$, existe una $g \in D_n$ tal que $\rho(\phi, g) < \varepsilon$. Para ello necesitamos un teorema auxiliar cuya prueba[9] omitiremos:

Teorema 3.6.4 (de aproximación de Weierstrass) *Sea I un intervalo cerrado y acotado de $\mathbb{R}$ y sea $\phi(x)$ una función continua en I. Entonces, para todo $\varepsilon > 0$ existe un polinomio $p(x)$ definido en I tal que*

$$\max_{x \in I} |\phi(x) - p(x)| < \varepsilon.$$

Entonces, para toda función $\phi \in C([0,1])$ existe un polinomio $p(x)$ en $[0,1]$ tal que $\rho(\phi, p) < \varepsilon/4$. Para cada $n \in \mathbb{N}$ construyamos una función sierra (una recta creciente y una decreciente) $s_n(x)$ como la de la figura 3.9 cuya altura sea $\varepsilon/4$ y cuya pendiente m sea mayor, en valor absoluto, que $\max_{x\in[0,1]} |p'(x)| + n$, de forma que la función[10] $g(x) = p(x) + s_n(x)$

[9] Ver, e.g., [4, pág. 129–130].

[10] Nótese que para $s_n(x)$ se tiene que, para todo $x \in (0,1)$ siempre existe un $h > 0$ lo suficientemente chico tal que $|s_n(x+h) - s_n(x)|/h = m$, siendo m la pendiente de los dientes de s_n.

pertenece a cada D_n y es tal que

$$|\phi(x) - g(x)| \leq |\phi(x) - p(x)| + |s_n(x)| \leq \frac{\varepsilon}{2} \quad \Rightarrow \quad \rho(\phi, g) < \varepsilon.$$

Así E_n es raro pues su complementario es denso en $C([0,1])$. Ahora

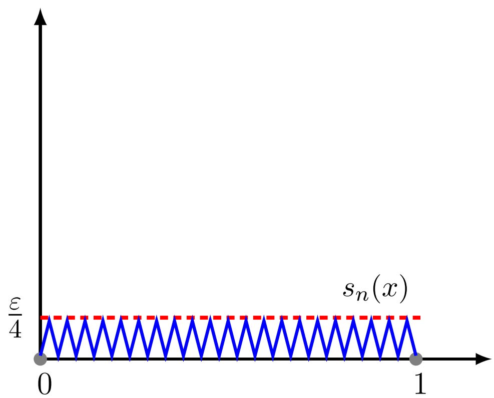


Figura 3.9: La función sierra $s_n(x)$ definida en cada E_n.

bien, como $A \subset \bigcup_{n=1}^{\infty} E_n$, y los E_n son raros, entonces A es de primera categoría, y como $C([0,1])$ es completo (ver Ejemplo 3.5.18) entonces, por el Teorema de Baire 3.6.1 es de segunda categoría. Ahora bien, $C([0,1]) = A \bigcup \{C([0,1]) \setminus A\}$, de donde se sigue que $C([0,1]) \setminus A$ es de segunda categoría. ■

Así pues, el conjunto de las funciones continuas en $[0,1]$ (que tienen una derivada direccional por la derecha en al menos un punto es un conjunto de primera categoría, y por tanto el conjunto de las que son derivables también. Sin embargo, el conjunto de las funciones continuas que no tienen derivada direccional por la derecha (y, por tanto, no son derivables) en ningún punto es de segunda categoría. Es decir, el conjunto de estas últimas funciones *monstruos* son las más comunes en el análisis, mientras que las funciones con las que estamos acostumbrados a trabajar constituyen un conjunto magro.

3.7. Problemas

Problema 3.1 Prueba la desigualdad de Young: dados dos números reales $a, b > 0$ y $p > 1$, entonces,

$$a^{\frac{1}{p}} b^{\frac{1}{q}} \leq \frac{1}{p} a + \frac{1}{q} b, \qquad \frac{1}{p} + \frac{1}{q} = 1. \tag{3.7.1}$$

Además la igualdad solo tiene lugar si $a = b$. **Ayuda**: Encuentra los extremos de la función $f : [0, \infty) \mapsto \mathbb{R}$, $f(x) = x^\alpha - \alpha x + \alpha - 1$, $\alpha \in (0, 1)$ y prueba que $f(x) \leq 0$ para todo $x \geq 0$. Escogiendo $x = a/b$, $a, b > 0$ y $\alpha = 1/p$, $p > 1$ se deduce el resultado.

Solución: La función f es continua y derivable en $(0, \infty)$.

$$f'(x) = \alpha(x^{\alpha-1} - 1), \quad \Rightarrow \quad f'(1) = 0.$$

Además $f''(x) = \alpha(\alpha - 1)x^{\alpha-2} < 0$ pues $\alpha \in (0, 1)$. Luego en $x = 1$ f tiene un máximo local. Pero como el único extremo local que tiene la función es éste, y además $f(0) = \alpha - 1 < 0$ y $\lim_{z\to+\infty} f(x) = -\infty$, entonces en $x = 1$ f tiene un máximo absoluto. Además $f(1) = 0$, por tanto, $f(x) \leq 0$ para todo $x > 0$, además la desigualdad es estricta para todo $x > 0$ distinto de 1.

Si elegimos $x = a/b$, $a, b > 0$ y $\alpha = 1/p$, $p > 1$, entonces, el resultado anterior nos conduce a (3.7.1), además la igualdad solo tiene lugar si $x = 1$, es decir, si $a = b$. ■

Problema 3.2 Prueba la desigualdad de Hölder: Sean los números x_i e y_i, $i = 1, 2 \ldots, n$ no negativos. Entonces, para todo $p > 1$ y $q > 1$, con $\frac{1}{p} + \frac{1}{q} = 1$, se tiene

$$\sum_{i=1}^{n} x_i \, y_i \leq \left(\sum_{i=1}^{n} x_i^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} y_i^q \right)^{\frac{1}{q}}, \quad \frac{1}{p} + \frac{1}{q} = 1, \tag{3.7.2}$$

donde la igualdad solo tiene lugar si $x_i^p = c\, y_i^q$ para todo $i = 1, 2, \ldots, n$ (es decir, si x_i^p y y_i^q son proporcionales). **Ayuda**: Usa la desigualdad de Young.

Solución: Si $x_i = y_i = 0$ para todo $i = 1, 2, \ldots, n$, la desigualdad es trivial. Supongamos entonces que al menos uno de los x_i y de los y_i son distintos de cero. Sea

$$X = \sum_{i=1}^{n} x_i^p > 0, \qquad Y = \sum_{i=1}^{n} y_i^q > 0, \quad \text{y} \quad a = \frac{x_i^p}{X}, \qquad b = \frac{y_i^q}{Y}, \qquad p > 1,$$

entonces usando la desigualdad de Young (3.7.1) obtenemos

$$\left(\frac{x_i^p}{X}\right)^{\frac{1}{p}}\left(\frac{y_i^p}{Y}\right)^{\frac{1}{q}} \leq \frac{1}{p}\frac{x_i^p}{X} + \frac{1}{q}\frac{y_i^q}{Y},$$

sumando en i desde $i = 1$ hasta $i = n$ obtenemos

$$\sum_{i=1}^{n} \frac{x_i}{X^{\frac{1}{p}}}\frac{y_i}{Y^{\frac{1}{q}}} \leq \frac{1}{p}\sum_{i=1}^{n}\frac{x_i^p}{X} + \frac{1}{q}\sum_{i=1}^{n}\frac{y_i^q}{Y} = \frac{1}{p} + \frac{1}{q} = 1,$$

de donde se deduce la desigualdad. Finalmente, la igualdad se obtiene si y solo si $a/b = 1$, de donde se deduce la proporcionalidad de los valores de x_i^p y y_i^q. ■

Problema 3.3 Prueba la desigualdad de Minkowski (3.1.1) a partir de la desigualdad de Hölder (3.7.2).

Solución: Si $x_i = y_i = 0$ para todo $i = 1, 2, \ldots, n$, la desigualdad es trivial. Supongamos entonces que al menos uno de los x_i y de los y_i son distintos de cero. Tenemos

$$\sum_{i=1}^{n}(x_i + y_i)^p = \sum_{i=1}^{n}(x_i + y_i)(x_i + y_i)^{p-1} = \sum_{i=1}^{n} x_i(x_i + y_i)^{p-1} + \sum_{i=1}^{n} y_i(x_i + y_i)^{p-1}.$$

A continuación aplicamos la desigualdad de Hölder a ambos sumandos

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

y

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

donde la igualdad tiene lugar si y solo si $x_i^p = c'(x_i + y_i)^{(p-1)q}$ y $y_i^p = c''(x_i + y_i)^{(p-1)q}$, es decir, si $x_i^p = c'/c'' y_i^p$, es decir si $x_i = cy_i$, para todo $i = 1, 2, \ldots, n$. Ahora bien, como $\frac{1}{p} + \frac{1}{q} = 1$, entonces $(p-1)q = p$, así pues,

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

de donde, dividiendo por $(\sum_{i=1}^{n}(x_i + y_i)^p)^{\frac{1}{q}}$ obtenemos el resultado. ■

Problema 3.4 Encontrar los extremos de la función $f : [0, \infty) \mapsto \mathbb{R}$, $f(x) = x^\alpha - \alpha x + \alpha - 1$, con $\alpha > 1$ o $\alpha < 0$ y probar que $f(x) \geq 0$ para todo $x \geq 0$. Deduce a partir de este resultado las desigualdades análogas a las desigualdades de Young (3.7.1), de Hölder (3.7.2) y Minkowski (3.1.1), respectivamente.

Solución: Este problema es análogo a los Problemas 3.1–3.3, simplemente las desigualdades son las contrarias.

Problema 3.5 Extiende a las series infinitas las desigualdades de Hölder

$$\sum_{i=1}^{\infty} x_i\, y_i \leq \left(\sum_{i=1}^{\infty} x_i^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{\infty} y_i^q\right)^{\frac{1}{q}}, \tag{3.7.3}$$

y Minkowski (3.1.3), respectivamente, donde $(x_i)_i$ y $(y_i)_i$ son sucesiones de números no-negativos. Se asume que las sucesiones son tales que las correspondientes series son convergentes.

Solución: Si las series $\sum_{i=1}^{\infty} x_i^p$ y $\sum_{i=1}^{\infty} y_i^q$ convergen, entonces, de la desigualdad (3.7.2) se tiene que las sumas parciales de la serie $\sum_{i=1}^{\infty} x_i\, y_i$ están acotadas y, al ser esta una serie de términos positivos, es convergente (Criterio de comparación de Weierstrass 1.2.8). Luego, tomando el límite cuando $n \to \infty$ en (3.7.2) se tiene (3.7.3). La desigualdad (3.1.3) se sigue de (3.1.1) tomando el límite cuando $n \to \infty$. ■

Problema 3.6 Sean f y g dos funciones continuas en $[a, b]$. Usando la desigualdad de Young (3.7.1) prueba las desigualdades de Hölder

$$\int_a^b |f(x)g(x)|dx \leq \left(\int_a^b |f(x)|^p dx\right)^{\frac{1}{p}} \left(\int_a^b |g(x)|^q dx\right)^{\frac{1}{q}}, \ p > 1, \tag{3.7.4}$$

donde q es tal que $\frac{1}{p} + \frac{1}{q} = 1$, y Minkowski

$$\left(\int_a^b |f(x) + g(x)|^p dx\right)^{\frac{1}{p}} \leq \left(\int_a^b |f(x)|^p dx\right)^{\frac{1}{p}} + \left(\int_a^b |g(x)|^p dx\right)^{\frac{1}{p}}, \ p > 1.$$

¿Son ciertas estas igualdades si f y g son integrables?

Solución: Sea

$$X=\int_a^b|f(x)|^p dx>0,\ Y=\int_a^b|g(x)|^q dx>0,\ a=\frac{|f(x)|^p}{X},\ b=\frac{|g(x)|^q}{Y}.$$

Aplicando la desigualdad de Young se sigue que

$$\left(\frac{|f(x)|^p}{X}\right)^{\frac{1}{p}}\left(\frac{|g(x)|^q}{Y}\right)^{\frac{1}{q}}\leq\frac{1}{p}\cdot\frac{|f(x)|^p}{X}+\frac{1}{q}\cdot\frac{|g(x)|^q}{Y}.$$

Luego, por la monotonía de la integral, tenemos

$$\int_a^b\left(\frac{|f(x)|}{X^{\frac{1}{p}}}\cdot\frac{|g(x)|}{Y^{\frac{1}{q}}}\right)dx\leq\frac{1}{p}\int_a^b\frac{|f(x)|^p}{X}dx+\frac{1}{q}\int_a^b\frac{|g(x)|^q}{Y}dx=\frac{1}{p}+\frac{1}{q}=1.$$

Multiplicando por $X^{1/p}$ e $Y^{1/q}$ obtenemos la desigualdad buscada.

Probemos ahora la desigualdad de Minkowski. Si f o g son idénticamente nulas, entonces la desigualdad es inmediata así que asumiremos que no lo son. Notemos que, para $p>1$,

$$\begin{aligned}\int_a^b|f(x)+g(x)|^p dx&=\int_a^b|f(x)+g(x)||f(x)+g(x)|^{p-1}dx\leq\\&\leq\int_a^b(|f(x)|+|g(x)|)|f(x)+g(x)|^{p-1}dx=\\&=\int_a^b|f(x)||f(x)+g(x)|^{p-1}dx+\int_a^b|g(x)||f(x)+g(x)|^{p-1}dx.\end{aligned}$$

Aplicando la desigualdad de Hölder (3.7.4) a los dos miembros de la última igualdad obtenemos

$$\begin{aligned}&\int_a^b|f(x)||f(x)+g(x)|^{p-1}dx+\int_a^b|g(x)||f(x)+g(x)|^{p-1}dx\leq\\&\leq\left(\left[\int_a^b|f(x)|^p dx\right]^{\frac{1}{p}}+\left[\int_a^b|g(x)|^p dx\right]^{\frac{1}{p}}\right)\left[\int_a^b|f(x)+g(x)|^{(p-1)q}dx\right]^{\frac{1}{q}}.\end{aligned}$$

Como $\frac{1}{p}+\frac{1}{q}=1$, se tiene que $(p-1)q=p$, entonces dividiendo la expresión anterior por $\left(\int_a^b|f(x)+g(x)|^p dx\right)^{\frac{1}{q}}$, y usando que

$$\frac{\displaystyle\int_a^b|f(x)+g(x)|^p dx}{\left(\int_a^b|f(x)+g(x)|^p dx\right)^{\frac{1}{q}}}=\left(\int_a^b|f(x)+g(x)|^p dx\right)^{1-\frac{1}{q}}=\left(\int_a^b|f(x)+g(x)|^p dx\right)^{p}$$

se obtiene el resultado.

Como se ve de la prueba la continuidad de f y g no es necesaria, lo que se necesita es que las funciones f y g y sus correspondientes potencias sean integrables. Está claro que si f y g son continuas son integrables. ■

Problema 3.7 Prueba que los espacios métricos definidos en los Ejemplos 3.1.5–3.1.7 y 3.1.12 son separables. **Ayuda:** Usa los subespacios definidos mediante los vectores $y = (q_1, q_2, \ldots, q_n)$, $q_k \in \mathbb{Q}$, $k = 1, 2, \ldots, n$ para los espacios de los ejemplos 3.1.5–3.1.7 y los subespacios definidos por los vectores $y = (q_1, q_2, \ldots, q_n, 0, 0, \ldots)$, para todo $n \in \mathbb{N}$, en el espacio del Ejemplo 3.1.12 ($p \geq 1$) y prueba que son numerables y densos en los correspondientes espacios métricos.

Solución: Como en el Problema 3.15 es suficiente con probarlo para $\mathbb{R}^n_p$, $p \geq 1$. Sean los vectores $y = (q_1, q_2, ..., q_n)$, $q_k \in \mathbb{Q}$, $k = 1, 2, ..., n$. Como la unión finita de conjuntos numerables es numerable, y $\mathbb{Q}$ es numerable, el conjunto Y de todos los y anteriores es numerable. Probemos que Y es denso en $\mathbb{R}^n_p$. Puesto que $\mathbb{Q}$ es denso en $\mathbb{R}$, entonces, para todo $x_i \in \mathbb{R}$ y $\varepsilon > 0$, existe un $q_i \in \mathbb{Q}$ tal que

$$|x_i - q_i| < \left(\frac{\varepsilon^p}{n}\right)^{1/p}.$$

Por tanto, para todo $x = (x_1, ..., x_n) \in \mathbb{R}^n_p$ existe un $q = (q_1, ..., q_n) \in Y$ tal que

$$\rho(x, q) = \left(\sum_{i=1}^{n} |x_i - q_i|^p\right)^{1/p} < (\varepsilon^p)^{1/p} \leq \varepsilon.$$

Es decir, Y es denso en $\mathbb{R}^n_p$, luego $\mathbb{R}^n_p$ es separable.

El caso de ℓ^p es totalmente análogo al caso especial de ℓ^2 discutido en el Ejemplo 3.4.5. ■

Problema 3.8 Prueba que el espacio $\mathbb{R}$ con la métrica discreta (ver Ejemplo 3.1.2) no es separable.

Solución: Supongamos que existe un conjunto $M \subset \mathbb{R}$ denso en $\mathbb{R}$. Como $\overline{M} = \mathbb{R}$, entonces todo $x \in \mathbb{R}$ es punto límite de M, luego existe una sucesión $(x_n)_n \in M$ tal que $x_n \to x$. Pero las únicas sucesiones que convergen en nuestro espacio son las que son constantes a partir de cierto $N \in \mathbb{N}$ (pues estamos tomando la métrica discreta), lo cual implica que $x \in M$. Luego, si M es denso, M no puede ser numerable y, por tanto, nuestro espacio no puede ser separable. ■

Problema 3.9 Sea $B(a,b)$ el conjunto de todas las funciones acotadas (no necesariamente continuas) en $[a,b]$. Prueba que $B(a,b)$ es un espacio métrico con la métrica $\rho(f,g) = \sup_{x\in[a,b]} |f(x) - g(x)|$. Prueba que dicho espacio no es separable. **Ayuda:** Sea el conjunto de todas las funciones $f_t(x) = 0$ para $0 \le x < t$ y $f_t(x) = 1$ para $t \le x \le b$, $t \in [a,b]$. Prueba que dicho conjunto no es numerable. Calcula la distancia entre dos funciones cualesquiera de dicho conjunto y razona como en el Ejemplo 3.4.6.

Solución: Probar que el espacio $B(a,b)$ es un espacio métrico es similar al caso del Ejemplo 3.1.10. Sea A el conjunto de todas las funciones $f_t(x) = 0$ para $0 \le x < t$ y $f_t(x) = 1$ para $t \le x \le b$, $t \in [a,b]$. Está claro que hay una función f_t para cada $t \in [a,b]$, luego A es no numerable pues $[a,b]$ no es numerable. Además, la distancia entre cualesquiera dos funciones distintas del conjunto A es 1, por tanto, razonando como en el Ejemplo 3.4.6 podemos construir un conjunto de bolas centradas en cada f_t y disjuntas entre sí, siendo no numerable dicho conjunto de bolas. Sea M cualquier conjunto denso en $B(a,b)$. Entonces, en cada una de las bolas anteriores debe haber al menos un elemento de M, luego M no es numerable (el conjunto de las bolas no lo es) y como M era arbitrario, entonces $B(a,b)$ no tiene ningún conjunto numerable denso, así que $B(a,b)$ tampoco lo es y, por tanto, $B(a,b)$ no puede ser separable. ■

Problema 3.10 Sea $\mathbb{X}$ un espacio métrico y sea $(x_n)_n$ una sucesión convergente. Prueba que

1. $(x_n)_n$ es acotada.
2. El límite de $(x_n)_n$ es único.
3. Si $x_n \overset{n\to\infty}{\longrightarrow} x$ e $y_n \overset{n\to\infty}{\longrightarrow} y$, entonces $\lim_{n\to\infty} \rho(x_n, y_n) = \rho(x,y)$, es decir, la aplicación $\rho : \mathbb{X} \mapsto \mathbb{R}$ es continua. **Ayuda:** Usa que $\rho(x_n, y_n) \le \rho(x_n, x) + \rho(x,y) + \rho(y, y_n)$.

Solución: 1. Por la definición de convergencia, sabemos que existe un $N \in \mathbb{N}$ talque $\rho(x_n, L) < 1$ para todo $n > N$. Además, existe el $\min_{n=1,\dots,N} \rho(x_n, L) = \delta$. Por tanto, si elegimos $r > \max\{1, \delta\} < +\infty$, tenemos que para todo n, $x_n \in B(L, r)$, i.e., está contenida en la bola centrada en L de radio r, luego $(x_n)_n$ es acotada.

2. Supongamos que la sucesión admite dos límites distintos $l_1 \ne l_2$. Entonces, por un lado, $d = \rho(l_1, l_2) > 0$, y por otro, como $x_n \to l_1$ y $x_n \to l_2$, entonces

para todo $\varepsilon > 0$ existe un $N \in \mathbb{N}$ tal que, para todo $n > N$ se cumple, al mismo tiempo, que $\rho(x_n, l_1) < \frac{\varepsilon}{2}$ y $\rho(x_n, l_2) < \frac{\varepsilon}{2}$, por tanto,

$$\rho(l_2, l_1) \leq \rho(l_2, x_n) + \rho(x_n, l_1) < \varepsilon \quad \Rightarrow \quad l_1 = l_2.$$

También se puede razonar como sigue: como $x_n \to l_1$ y $x_n \to l_2$, para todo $\varepsilon > 0$ existe (¿por qué?) un N tal que, para todo $n > N$, $x_n \in B(l_1; \frac{\varepsilon}{3})$ y $x_n \in B(l_2; \frac{\varepsilon}{3})$. Como $l_1 \neq l_2$ se puede siempre elegir $\varepsilon > 0$ tal que $B(l_1; \frac{\varepsilon}{3}) \bigcap B(l_2; \frac{\varepsilon}{3}) = \emptyset$. Pero entonces $x_n \in B(l_1; \frac{\varepsilon}{3}) \bigcap B(l_2; \frac{\varepsilon}{3}) = \emptyset$, lo cual es una contradicción.

3. Usando la desigualdad triangular tenemos

$$\rho(x_n, y_n) \leq \rho(x_n, y) + \rho(y, y_n) \leq \rho(x_n, x) + \rho(x, y) + \rho(y, y_n),$$

es decir, $\rho(x_n, y_n) - \rho(x, y) \leq \rho(x_n, x) + \rho(y, y_n)$. Análogamente

$$\rho(x, y) - \rho(x_n, y_n) \leq \rho(x_n, x) + \rho(y, y_n).$$

Combinando ambas desigualdades obtenemos

$$|\rho(x, y) - \rho(x_n, y_n)| \leq \rho(x_n, x) + \rho(y, y_n),$$

de donde se sigue el resultado pues $\rho(x_n, x) \to 0$, y $\rho(y_n, y) \to 0$. ■

Problema 3.11 Es conocido que en $\mathbb{R}$ de toda sucesión acotada se puede construir (extraer) una subsucesión convergente (Teorema de Bolzano-Weierstrass). Sea ahora $\mathbb{X}$ un espacio métrico arbitrario y sea $(x_n)_n$ una sucesión acotada de $\mathbb{X}$. ¿Se puede construir (extraer) de $(x_n)_n$ una subsucesión convergente? Justifica la respuesta.

Solución: No necesariamente. Por ejemplo, sea $\mathbb{X} = (0, 1)$ con la métrica habitual. La sucesión $x_n = 1/n$ es acotada pero no tiene límite en $\mathbb{X}$ y, por tanto, tampoco tiene límite ninguna de sus subsucesiones. De hecho es suficiente construir una sucesión acotada en $\mathbb{X}$ cuyo límite no pertenezca a $\mathbb{X}$ (como el ejemplo anterior). Puesto que todas las subsucesiones de una sucesión convergente tienen el mismo límite que la propia sucesión, ninguna subsucesión será convergente. Por tanto, la completitud de $\mathbb{X}$ es una condición necesaria. El próximo problema nos muestra que no es suficiente. ■

Problema 3.12 Sea $\mathbb{X}$ un espacio métrico completo. Sea $(x_n)_n$ una sucesión acotada en $\mathbb{X}$. ¿Se puede extraer de $(x_n)_n$ una subsucesión convergente (es decir, ¿es cierto el Teorema de Bolzano-Weierstrass para los espacios métricos completos?). Justifica tu respuesta. ¿Qué ocurre si $\mathbb{X}$ no es completo?

Solución: Si $\mathbb{X}$ no es completo, ver Ejercicio 3.11, donde vimos que la completitud era, en general, una condición necesaria. Supongamos que $\mathbb{X}$ es completo. Mostremos que tampoco es suficiente. Sea ℓ^∞ el espacio de las sucesiones acotadas con la métrica del supremo (ver Ejemplo 3.1.11). Este espacio es completo (ver Problema 3.16). Escojamos la sucesión $x^{(k)} = e_k$, donde e_k es la sucesión cuyos términos son todos nulos excepto en k-ésimo que tomamos igual a 1. Está claro que la sucesión $\|x^{(k)}\| = 1$ para todo k, luego es acotada. Además, para todos $n \neq m \in \mathbb{N}$, $\rho(x^{(k)}, x^{(m)}) = 1$, por tanto de dicha sucesión no podemos extraer ninguna subsucesión convergente. Este ejemplo también nos vale si en vez de usar ℓ^∞ usamos ℓ^2 como espacio métrico completo. ■

Problema 3.13 Sea $(x_n)_n$ una sucesión fundamental (de Cauchy) de elementos de un espacio métrico $\mathbb{X}$.

1. Prueba que $(x_n)_n$ es acotada.
2. Sea $(x_{n_k})_k$ una subsucesión de $(x_n)_n$. Prueba que si $(x_{n_k})_k$ converge a x, entonces $(x_n)_n$ también converge y $\lim_{n\to\infty} x_n = x$.

Solución: Sea una sucesión de Cauchy $(x_n)_n \in \mathbb{X}$. Entonces,

$$\forall\ \varepsilon > 0,\ \exists\, N \in \mathbb{N} : \forall\, n, m > N,\ \rho(x_n, x_m) < \varepsilon. \tag{3.7.5}$$

1. Tomando $\varepsilon = 1$ en (3.7.5) se tiene que para todo $n > N$, $\rho(x_n, x_N) \leq 1$. Razonando como en el punto uno del Problema 3.10 se tiene el resultado.

2. Sea $(x_{n_k})_k$ la subsucesión convergente a $x \in \mathbb{X}$. Entonces, por un lado que $x_{n_k} \to x$ si $k \to \infty$, luego existe $N > 0$ tal que $\rho(x_{n_k}, x) < \varepsilon/2$ si $n_k > N$, y, por el otro, como, para todo $k \in \mathbb{N}$, $n_k \geq k$, tenemos por (3.7.5) que para $k > N$ (que podemos tomar sin pérdida de generalidad igual que el N anterior) que $\rho(x_k, x_{n_k}) < \varepsilon/2$. Pero entonces, para $k > N$

$$\rho(x_k, x) \leq \rho(x_k, x_{n_k}) + \rho(x_{n_k}, x) < \varepsilon,$$

de donde se sigue el resultado. ■

Problema 3.14 Prueba que el espacio métrico trivial del Ejemplo 3.1.2 es completo.

Solución: Sea $(x_n)_n \subset \mathbb{X}$ una sucesión de Cauchy del espacio métrico trivial del Ejemplo 3.1.2. Entonces se tiene (3.7.5). Eligiendo $\varepsilon < 1$ cualquiera tenemos, por la definición de la métrica trivial, que para todos $n, m > N$, $\rho(x_n, x_m) = 0$, luego para todos $n, m > N$, $x_n = x_m$. Es decir, las únicas sucesiones de Cauchy de $\mathbb{X}$ son las que son constantes a partir de cierto N, es decir, $x_n = x_{N+1}$ $\forall n > N$. Luego, todas las sucesiones de Cauchy son convergentes y su límite es $x_{N+1} \in \mathbb{X}$ de donde se sigue la completitud del espacio métrico trivial. ■

Problema 3.15 Prueba que los espacios métricos de los ejemplos 3.1.5–3.1.7 son completos. ¿Y el del Ejemplo 3.1.8?

Solución: Vamos a probar que $\mathbb{R}^n_p$, $p \geq 1$ es completo de donde se sigue el resultado para los ejemplos 3.1.5-3.1.6. La prueba es esencialmente la misma que la del Ejemplo 3.5.16 así que solo daremos un esbozo de la misma.

Sea $(x^{(k)})_k$, $x^{(k)} := (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})$ una sucesión de Cauchy. Entonces, cualquiera sea $\varepsilon > 0$,

$$\sum_{i=1}^{n} |x_i^{(k)} - x_i^{(k+m)}|^p < \varepsilon^p, \qquad \forall m \in \mathbb{N}, \tag{3.7.6}$$

con tal que k sea lo suficientemente grande. Pero entonces como todos los sumandos son positivos, podemos asegurar que, para todo $m \in \mathbb{N}$, $|x_i^{(k)} - x_i^{(k+m)}| < \varepsilon$, y k suficientemente grande. Es decir, cada una de las sucesiones reales $x_i^{(k)}$, $k = 1, 2, \dots, n$, es de Cauchy. Pero $\mathbb{R}$ es un espacio completo por lo que existe $\lim_{k\to\infty} x_i^{(k)} = x_i$ para cada i. Sea $x = (x_1, x_2, \dots, x_n)$. Como $x_i^{(k)}$ converge a x_i para cada i, tomando el límite cuando $m \to \infty$ en (3.7.6) tenemos

$$\sum_{i=1}^{n} |x_i^{(k)} - x_i|^p \leq \varepsilon^p,$$

i.e, $\rho(x^{(k)}, x) \leq \varepsilon$, si n es suficientemente grande. Luego hemos probado que $x^{(k)}$ converge a x y como $x^{(k)}$ era una sucesión fundamental arbitraria, se sigue que $\mathbb{R}^n_p$ es completo.

Para el Ejemplo 3.1.8, i.e., para $\mathbb{R}^n_\infty$ el razonamiento es similar. De hecho, la completitud para todos estos espacios se sigue del Teorema 4.3.2 que probaremos en el próximo capítulo. ■

Problema 3.16 Prueba que el espacio métrico ℓ^∞ del ejemplo 3.1.11 es completo.

Solución: $x^{(k)} := (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}, \dots)$ una sucesión de Cauchy. Usamos (3.7.5) que, en nuestro caso, implica que $\sup_{i\in\mathbb{N}} |x_i^{(k)} - x_i^{(k+m)}| < \varepsilon$, luego para todo $i \in \mathbb{N}$, $|x_i^{(k)} - x_i^{(k+m)}| < \varepsilon$, i.e., cada una de las sucesiones $x_i^{(k)}$, $k = 1, 2, \dots, n$, es de Cauchy. Pero $\mathbb{R}$ es un espacio completo por lo que existe $\lim_{k\to\infty} x_i^{(k)} = x_i$ para cada i. Sea $x = (x_1, x_2, \dots, x_n, \dots)$. Probemos que $x \in \ell^\infty$. Para ello notemos que para todo $i \in \mathbb{N}$

$$|x_i| \leq |x_i - x_i^{(k)}| + |x_i^{(k)}|,$$

pero para cada $i \in \mathbb{N}$, $(x_i^{(k)})$ es de Cauchy, luego es acotada, y si elegimos k suficientemente grande, podemos hacer $|x_i - x_i^{(k)}|$ suficientemente pequeño, luego $(x_i)_i$ es una sucesión acotada, luego $x \in \ell^\infty$. Finalmente, como para todo $i \in \mathbb{N}$, $|x_i - x_i^{(k)}|$ es suficientemente pequeño, entonces $\sup_{i\in\mathbb{N}} |x_i - x_i^{(k)}|$, también lo es, luego $x^{(k)} \to x$. Como $x^{(k)}$ era una sucesión arbitraria de Cauchy, ℓ^∞ es completo. ■

Problema 3.17 Prueba que el espacio métrico definido en el Ejemplo 3.1.13 es completo y separable.

Solución: Sea $\mathbb{X}$ el espacio de las sucesiones reales con la métrica

$$\rho(x, y) = \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{|x_j - y_j|}{1 + |x_j - y_j|}. \tag{3.7.7}$$

En primer lugar, como la función $f(z) = z/(1+z) \leq 1$ para todo $x \geq 0$, tenemos que

$$\frac{1}{2^j} \frac{|x_j - y_j|}{1 + |x_j - y_j|} \leq \frac{1}{2^j} = M_j, \quad \sum_{j=1}^{\infty} M_j = 1 < \infty.$$

Por tanto, por el criterio de comparación de Weierstrass 1.2.8 tenemos que la serie (3.7.7) es convergente cualquiera sean las sucesiones $(x_n)_n$ e $(y_n)_n$ que escojamos.

Está claro que $\rho(x, y) \geq 0$ y solo es cero si $x = y$. Además $\rho(x, y) = \rho(y, x)$. Por otro lado, como $f(x)$ es creciente para $x \geq 0$, entonces, para todos $a, b \in \mathbb{R}$,

$f(|a+b|) \leq f(|a|+|b|)$ y, por tanto, tomando $a = x_j - y_j$ y $b = y_j - z_j$, obtenemos

$$\begin{aligned}\rho(x,z) &= \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{|x_j - z_j|}{1 + |x_j - z_j|} = \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{|x_j - y_j + y_j - z_j|}{1 + |x_j - y_j + y_j - z_j|} \leq \\ &\leq \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{|x_j - y_j| + |y_j - z_j|}{1 + |x_j - y_j| + |y_j - z_j|} = \\ &= \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{|x_j - y_j|}{1 + |x_j - y_j| + |y_j - z_j|} + \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{|y_j - z_j|}{1 + |x_j - y_j| + |y_j - z_j|} \leq \\ &\leq \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{|x_j - y_j|}{1 + |x_j - y_j|} + \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{|y_j - z_j|}{1 + |y_j - z_j|} = \rho(x,y) + \rho(y,z).\end{aligned}$$

Luego la función (3.7.7) es efectivamente una métrica y nuestro espacio es un espacio métrico.

Probemos que $\mathbb{X}$ es completo. Sea $(x^{(n)})_n \in \mathbb{X}$ una sucesión de Cauchy. Entonces (3.7.5) implica que para todo $\varepsilon' > 0$ existe un $N \in \mathbb{N}$ tal que para todos $n, m > N$,

$$\rho(x^{(n)}, x^{(m)}) = \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{|x_j^{(n)} - x_j^{(m)}|}{1 + |x_j^{(n)} - x_j^{(m)}|} < \varepsilon'. \tag{3.7.8}$$

Ahora bien, para cada $k \in \mathbb{N}$ fijo y cualquiera sea $\varepsilon > 0$, tenemos[11], para todos $n, m > N$,

$$\frac{1}{2^k} \frac{|x_k^{(n)} - x_k^{(m)}|}{1 + |x_k^{(n)} - x_k^{(m)}|} \leq \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{|x_j^{(n)} - x_j^{(m)}|}{1 + |x_j^{(n)} - x_j^{(m)}|} < \frac{1}{2^k} \frac{\varepsilon}{1 + \varepsilon},$$

donde hemos elegido $\varepsilon' = \frac{1}{2^k} \frac{\varepsilon}{1+\varepsilon}$ en (3.7.8). Como f es estrictamente creciente para $x > 0$ lo anterior implica que para todo $j \in \mathbb{N}$, $|x_j^{(n)} - x_j^{(m)}| < \varepsilon$. Es decir, para cada $j \in \mathbb{N}$, las sucesiones $(x_j^{(n)})_n \in \mathbb{R}$ son de Cauchy y como $\mathbb{R}$ es completo, son convergentes. Sea $x_j^{(n)} \to x_j$ cuando $n \to \infty$. Está claro que $x \in \mathbb{X}$. Tomando el límite $m \to \infty$ en (3.7.8) y usando que la métrica es una aplicación continua tenemos que $\rho(x^{(n)}, x) \leq \varepsilon$ si n es lo suficientemente grande, luego $x^{(n)} \to x \in \mathbb{X}$, y como $(x^{(n)})_n$ era arbitraria, $\mathbb{X}$ es completo.

Veamos ahora que es separable. Para ello definiremos el conjunto $Q \subset \mathbb{X}$ de todas las sucesiones del tipo $y = (q_1, q_2, \ldots, q_n, 0, 0, \ldots)$, para todo $n \in \mathbb{N}$,

[11] De hecho, de esta misma forma se puede probar que $\rho(x,y)$ es suficientemente pequeño si y solo si $|x_i - y_i|$ es suficientemente pequeño para cada i, donde $x = (x_i)_i$ e $y = (y_i)_i$.

donde para todo k, q_k es racional (ver el Ejemplo 3.4.5)). Como ya vimos Q es un subespacio numerable. Probemos que Q es denso en $\mathbb{X}$. Sea una sucesión cualquiera $(x_n)_n \in \mathbb{X}$. Entonces, para todo $\varepsilon > 0$ existe un $N \in \mathbb{N}$ tal que

$$\sum_{j=N+1}^{\infty} \frac{1}{2^j} \frac{|x_j|}{1+|x_j|} < \frac{\varepsilon}{2}.$$

Como $\mathbb{Q}$ es denso en $\mathbb{R}$, entonces para todo x_k existe un racional q_k tan cerca como se quiera, i.e., para todo $\varepsilon > 0$ y todo k, $|x_k - q_k| < \varepsilon/(2N)$. Luego existe una sucesión $q \in Q$ tal que

$$\sum_{j=1}^{N} \frac{1}{2^j} \frac{|x_j - y_j|}{1+|x_j - y_j|} < \frac{\varepsilon}{2}.$$

Entonces, para todo $x \in \mathbb{X}$ podemos encontrar un $q \in Q$ tal que

$$\rho(x,q) = \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{|x_j - q_j|}{1+|x_j - q_j|} = \sum_{j=1}^{N} \frac{1}{2^j} \frac{|x_j - y_j|}{1+|x_j - y_j|} + \sum_{j=N+1}^{\infty} \frac{1}{2^j} \frac{|x_j|}{1+|x_j|} < \varepsilon,$$

i.e., para todo $x \in \mathbb{X}$, existe un $q \in Q$ tal que $\rho(x,q) < \varepsilon$, luego Q es denso en $\mathbb{X}$. Como $\mathbb{X}$ contiene un subconjunto denso y numerable entonces es separable. ■

Problema 3.18 Estudia bajo qué condiciones (sobre los $a_{i,j}$) la aplicación $A : \mathbb{R}^n \mapsto \mathbb{R}^n$ definida por el sistema de ecuaciones

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

es de contracción en los siguientes casos

1. Si usamos la métrica del Ejemplo 3.1.5.
2. Si usamos la métrica del Ejemplo 3.1.6
3. Si usamos la métrica del Ejemplo 3.1.8

Solución: En el segundo caso tenemos $\mathbb{R}^n_1$:

$$\rho(Ax, Ay) = \sum_{i=1}^{n} \left| \sum_{j=1}^{n} a_{ij} x_j + b_i - \sum_{j=1}^{n} a_{ij} y_j + b_i \right| = \sum_{i=1}^{n} \left| \sum_{j=1}^{n} a_{ij}(x_j - y_j) \right| \leq$$

$$\leq \sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}(x_j - y_j)| \leq \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\max_{j=1,\ldots,n} |a_{ij}| \right) |x_j - y_j| =$$

$$= \sum_{i=1}^{n} \max_{j=1,\ldots,n} |a_{ij}| \sum_{j=1}^{n} |x_j - y_j| = \alpha \sum_{j=1}^{n} |x_j - y_j| = \alpha \rho(x,y).$$

Luego A será contractiva si $\alpha = \sum_{i=1}^{n} \max_{j=1,\ldots,n} |a_{ij}| < 1$.

Para el tercer caso

$$\rho(Ax, Ay) = \max_{i=1,\ldots,n} \left| \sum_{j=1}^{n} a_{ij}x_j + b_i - \sum_{j=1}^{n} a_{ij}y_j + b_i \right| = \max_{i=1,\ldots,n} \left| \sum_{j=1}^{n} a_{ij}(x_j - y_j) \right| \leq$$
$$\leq \max_{i=1,\ldots,n} \sum_{j=1}^{n} |a_{ij}||x_j - y_j| \leq \max_{i=1,\ldots,n} \sum_{j=1}^{n} \max_{j=1,..n} |a_{ij}||x_j - y_j| =$$
$$= \max_{i=1,\ldots,n} \max_{j=1,..,n} |a_{i,j}||x_j - y_j|\, n = \alpha \max_{j=1,\ldots,n} |x_j - y_j| = \alpha\rho(x, y).$$

Luego, A será contractiva si

$$\alpha = n \max_{i,j=1,\ldots,n} |a_{i,j}| < 1 \Leftrightarrow \max_{i,j=1,\ldots,n} |a_{i,j}| < \frac{1}{n}.$$

Finalmente, en el primer caso, usando la desigualdad de Cauchy-Schwarz[12] en $\mathbb{R}^n$, se puede comprobar que

$$\rho(Ax, Ay) = \sqrt{\sum_{i=1}^{n} \left| \sum_{j=1}^{n} a_{ij}(x_j - y_j) \right|^2} \leq \sqrt{\sum_{i,j=1}^{n} |a_{ij}|^2} \rho(x, y),$$

así que si, $\sqrt{\sum_{i,j=1}^{n} |a_{ij}|^2} < 1$, nuestra aplicación es contractiva. No obstante debemos mencionar que si A es una matriz simétrica el resultado anterior se puede mejorar usando el *Teorema espectral* para matrices simétricas. ∎

Problema 3.19 Utiliza el Teorema del punto fijo para probar que el problema de valores iniciales asociado al sistema de ecuaciones diferenciales

$$y_i'(x) = f_i(x, y_1, y_2, \ldots, y_n), i = 1, \ldots, n, \;\; y_1(x_0) = (y_1)_0, \ldots, y_n(x_0) = (y_n)_0,$$

tiene solución única en cierto δ-entorno de x_0. Asumir que las funciones f_i son continuas en cierta región Ω de $\mathbb{R}^{n+1}$ que contiene al punto $(x_0, (y_1)_0, \ldots, (y_n)_0)$, y que se satisfacen una condición de Lipschitz de la forma

$$\exists K > 0;\; \forall x, y_i, \tilde{y}_i \in \Omega,\;\; |f(x, y_1, \ldots, y_n) - f(x, \tilde{y}_1, \ldots, \tilde{y}_n)| \leq K \max_{i=1,\ldots,n} |y_i - \tilde{y}_i|.$$

[12] $|\sum_{k=1}^{n} a_k b_k| \leq \sqrt{\sum_{k=1}^{n} a_k^2} \sqrt{\sum_{k=1}^{n} b_k^2}$

Solución: La prueba consiste en adaptar la prueba del Teorema de Picard 3.5.36 para una ecuación a un sistema de ecuaciones.

Problema 3.20 Sea la ecuación integral de Fredholm de 2º tipo

$$f(x) = \lambda \int_a^b K(x,y) f(y) dy + h(x),$$

donde el *núcleo* $K(x,y)$ de la ecuación integral y h son funciones conocidas y continuas en el cuadrado $D = \{(x,y) \,|\, x \in [a,b],\ y \in [a,b]\}$. Prueba que en el espacio de las funciones continuas $C([a,b])$ con la métrica $\rho(f,g) = \max_{x\in[a,b]} |f(x) - g(x)|$ (ver Ejemplo 3.1.9) la ecuación anterior tiene una única solución si $|\lambda| < (M(b-a))^{-1}$, donde M es el máximo de la función $|K(x,y)|$ en D. **Ayuda:** Utiliza el Teorema del punto fijo.

Solución: Sea el operador $T : C([a,b]) \longrightarrow C([a,b])$, definido por

$$Tf(x) = \lambda \int_a^b K(x,y) f(y) dy + h(x).$$

Tomemos en $C([a,b])$ la norma del máximo (supremo) de forma que tenemos el espacio métrico completo $C^\infty_{[a,b]}$. Calculamos

$$\begin{aligned}
|Tf(x) - Tg(x)| &= \left| \lambda \int_a^b K(x,y)(f(y) - g(y)) dy \right| \\
&\leq |\lambda| \int_a^b |K(x,y)||f(y) - g(y)| dy \\
&\leq |\lambda| \int_a^b M \cdot \max_{y\in[a,b]} |f(y) - g(y)| dy \\
&= |\lambda| M(b-a) \max_{y\in[a,b]} |f(y) - g(y)|,
\end{aligned}$$

donde $M = \max_{x,y\in[a,b]} |K(x,y)|$. Luego

$$\max_{x\in[a,b]} |Tf(x) - Tg(x)| \leq |\lambda| M(b-a) \max_{y\in[a,b]} |f(y) - g(y)|$$

de donde se sigue que

$$\rho(Tf, Tg) \leq |\lambda| M(b-a) \rho(f,g) = \alpha \rho(f,g), \quad \alpha = |\lambda| M(b-a).$$

Si $|\lambda| < (M(b-a))^{-1}$, entonces $\alpha \in (0,1)$ y, por tanto, T es un operador de contracción en un espacio métrico completo, luego podemos aplicar el Teorema de punto fijo de Banach 3.5.31 que nos asegura que existe una única función f^* tal que $Tf^* = f^*$, i.e.,

$$f^*(x) = \lambda \int_a^b K(x,y) f^*(y) dy + h(x),$$

por lo que existe una única solución a la ecuación integral de Fredholm. ∎

Problema 3.21 Prueba que la función $f(x)$ definida sobre $(0,1)$ tal que $f(x) = 1/q$ si $x = p/q$ (fracción reducida) e igual a 0 si x es irracional es continua en cada irracional de $(0,1)$ y discontinua en cada racional de $(0,1)$.

Solución: La función propuesta se conoce como función de Thomae: $f : \mathbb{R} \mapsto [0,1]$, definida como

$$f(x) = \begin{cases} \dfrac{1}{q}, & \text{si } x = \dfrac{p}{q} \in \mathbb{Q}, \text{ con } p \in \mathbb{Z}, q \in \mathbb{N} \text{ y } \mathrm{mcd}(p,q) = 1, \\ 0, & \text{si } x \in \mathbb{R} \backslash \mathbb{Q}. \end{cases}$$

Estudiemos ahora su continuidad. Sea $x_0 = \frac{p}{q} \in \mathbb{Q}$ con $p \in \mathbb{Z}, q \in \mathbb{N}$, coprimos. Luego $f(x_0) = \frac{1}{q}$.

Sea $\alpha \in \mathbb{R}\backslash\mathbb{Q}$, arbitrario. Definimos $x_n = x_0 + \dfrac{\alpha}{n} \in \mathbb{R}\backslash\mathbb{Q}$, $n \in \mathbb{N}$. Entonces

$$|x_0 - x_n| = \frac{\alpha}{n}, \text{ y } \quad |f(x_0) - f(x_n)| = \frac{1}{q}.$$

Luego $x_n \to x_0$ pero $f(x_n) \not\to f(x_0)$, y, por tanto, f no es continua en ningún racional.

Veamos ahora la continuidad en los irracionales. Sea $\varepsilon > 0$ cualquiera, $i \in \mathbb{N}$ y $x_0 \in [0,1]\backslash\mathbb{Q}$, cualquiera. Por la propiedad arquimediana, existe $r \in \mathbb{N}$ tal que $0 < 1/r < \varepsilon$, y existen $k_i \in \mathbb{N}$ tales que

$$\forall\, i = 1, ..., r \ \text{ se tiene } \ 0 < \frac{k_i}{i} < x_0 < \frac{k_i + 1}{i}.$$

Ahora definimos la mínima distancia de x_0 a sus i-ésimas cotas como:

$$d_i := \text{mín}\left\{ \left| x_0 - \frac{k_i}{i} \right|, \left| x_0 - \frac{k_i + 1}{i} \right| \right\}, \quad \text{y sea} \quad \delta = \min_{1 \le i \le r} \{d_i\}.$$

Elijamos $x \in (x_0 - \delta, x_0 + \delta)$. Si x es irracional, entonces $f(x_0) = f(x) = 0$. Si x es racional con fracción irreducible p/q, necesariamente ha de ser $q > r$ y, por tanto:

$$|f(x_0) - f(x)| = \frac{1}{q} < \frac{1}{r} < \varepsilon.$$

Así concluimos que f es continua en los irracionales.

Problema 3.22 El Teorema de Baire 3.6.1 se puede probar a partir del Teorema de las esferas encajadas 3.5.20 para ello hay que encontrar una sucesión apropiada de esferas encajadas[13]. Ello también se puede hacer razonando de la misma forma que en la prueba presentada en el apartado 3.6 y mostrado que existe una esfera S_k tal que $S_k \subset B_k$ pero que no está contenida en B_{k+1}. Usar esta idea para dar una prueba alternativa del Teorema de Baire 3.6.1.

[13]Ver, e.g., G.E. Shilov, *Elementary Real and Complex Analysis*. Dover Publications, New York, 1996.

Capítulo 4

Espacios normados y espacios de Banach

Man muss immer generalisieren (Uno debe siempre generalizar).

Atribuida a Carl G.J. Jacobi

4.1. Espacios vectoriales

Definición 4.1.1 *Sea $\mathbb{V}$ un conjunto de elementos cualesquiera y $\mathbb{K}$ el cuerpo de los números reales $\mathbb{R}$ o complejos $\mathbb{C}$. Definiremos en $\mathbb{V}$ las operaciones* suma "+" *de dos elementos x, y de $\mathbb{V}$ y* multiplicación "·" *de un elemento de $\mathbb{V}$ por un número (real o complejo) $\alpha \in \mathbb{K}$ por un elemento de $\mathbb{V}$. Diremos que $\mathbb{V}$ es un espacio vectorial sobre $\mathbb{K}$ (real o complejo si $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$, respectivamente), si se cumplen las siguientes propiedades (axiomas):*

1. *Para todos x e y,* vectores *de $\mathbb{V}$, el vector suma, $w = x + y$, también es un vector de $\mathbb{V}$ y para todos $x, y, z \in \mathbb{V}$ se cumple que:*

    a) $x + y = y + x$

    b) $(x + y) + z = x + (y + z)$

    c) *Existe un elemento "nulo" de $\mathbb{V}$, tal que $x + 0 = 0 + x = x$*

    d) *Cualquiera sea el vector x de $\mathbb{V}$, existe el elemento $(-x)$ "opuesto" a x, tal que $x + (-x) = (-x) + x = 0$.*

2. *Para todo* x vector *de* $\mathbb{V}$*, el vector que se obtiene al multiplicar por un escalar,* $w = \alpha \cdot x$*, también es un vector de* $\mathbb{V}$ *y para todos* $x, y \in \mathbb{V}$*,* $\alpha, \beta \in \mathbb{K}$ *se cumple que:*

   a) $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$

   b) $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$

   c) $\alpha \cdot (\beta \cdot x) = (\alpha\beta) \cdot x$

   d) $1 \cdot x = x$

Ejemplos.

1. El conjunto de los vectores de $\mathbb{R}^n$ cuando la suma de dos vectores y la multiplicación por un escalar es la estándar.

2. El conjunto de las matrices $m \times n$ cuando la suma de dos matrices y la multiplicación por un escalar es la estándar. Dicho espacio lo denotaremos por $\mathbb{R}^{m\times n}$.

3. El conjunto de los polinomios reales[1] de grado a lo sumo n, que denotaremos por $\mathbb{P}_n$, es decir,

$$\mathbb{P}_n = \{p_n(t) = a_0 + a_1\, t + \cdots + a_n\, t^n, \quad a_0, ..., a_n \text{ números reales.}\},$$

   donde definiremos la suma de dos polinomios y la multiplicación por un escalar de la siguiente forma:

$$p(t) = a_0 + a_1\, t + \cdots + a_n\, t^n, \quad q(x) = b_0 + b_1\, t + \cdots + b_n\, t^n,$$

$$(p + q)(t) := p(t) + q(t) = (a_0 + b_0) + (a_1 + b_1)t + \cdots + (a_n + b_n)t^n,$$

$$(\alpha \cdot p)(t) := \alpha\, p(t) = (\alpha a_0) + (\alpha a_1)t + \cdots + (\alpha a_n)t^n.$$

   Además, $p_n = 0$, si y solo si $a_0 = a_1 = \cdots = a_n = 0$.

4. El conjunto de las funciones continuas en el intervalo $[a, b]$, que denotaremos por $C([a, b])$, cuando la suma de dos funciones f y g y la multiplicación por un escalar α están dadas por

$$(f + g)(t) := f(t) + g(t), \quad (\alpha \cdot f)(t) := \alpha\, \cdot f(t).$$

[1] De forma totalmente análoga se puede definir para el caso complejo.

Definición 4.1.2 *Sea $\mathbb{V}$ un espacio vectorial. Diremos que un subconjunto $H \subset \mathbb{V}$ de elementos de $\mathbb{V}$ es un subespacio vectorial de $\mathbb{V}$ si H es a su vez un espacio vectorial respecto a las mismas operaciones* suma "+" *y* multiplicación "·" *que $\mathbb{V}$.*

Ejemplos.

1. Dado un espacio vectorial $\mathbb{V}$, son subespacios vectoriales "triviales" los subespacios $H = \{0\}$ (conjunto que tiene como único elemento, el nulo) y $H = \mathbb{V}$ (el mismo espacio vectorial).

2. Para $\mathbb{V} = C([a,b])$, $H = \mathbb{P}_n$ es un subespacio vectorial, para cualquier $n = 0, 1, 2, \ldots$ entero no negativo.

3. Para $\mathbb{V} = \mathbb{P}_n$, $H = \mathbb{P}_k$ es un subespacio vectorial para todo $k < n$.

Ejercicio 4.1.3 *Prueba que los espacios ℓ^∞ y ℓ^p $(p \geq 1)$ de los ejemplos 3.1.11 y 3.1.12, respectivamente, son espacios vectoriales.*

Que la suma $x + y \in \ell^\infty$ es inmediato. Para el caso de ℓ^p se sigue de la desigualdad de Minkowski (3.1.3). Para la multiplicación por un escalar es inmediato. Finalmente, se puede comprobar mediante un cálculo directo que en cada caso se cumplen las propiedades de la Definición 4.1.1. ■

Teorema 4.1.4 *Un subconjunto H de elementos de $\mathbb{V}$ es un subespacio vectorial de $\mathbb{V}$ si y solo si se cumple*[a] *que para todos x e y,* vectores de *H y α, $\beta \in \mathbb{K}$ el vector $w = \alpha x + \beta y$ también es un vector de H.*

[a] Usualmente se impone además que el elemento nulo de $\mathbb{V}$ pertenezca a H, pero eso se deduce de la condición $w = \alpha x + \beta y \in H$ tomando $\alpha = \beta = 0$.

La demostración es inmediata de la definición de subespacio vectorial.

Definamos ahora la envoltura lineal $\operatorname{span}(v_1, v_2, \ldots, v_p)$ de los vectores $v_1, v_2, \ldots, v_p$ como el conjunto de todas las combinaciones lineales de dichos vectores:

$$\operatorname{span}(v_1, v_2, \ldots, v_p) = \left\{ \sum_{k=1}^{p} \alpha_k v_k \;\middle|\; \alpha_k \in \mathbb{K},\ k = 1, 2, \ldots, p \right\}.$$

Usando el teorema anterior se deduce el siguiente

Teorema 4.1.5 *Dado un conjunto de vectores* $\{v_1, v_2, ..., v_p\}$ *de un espacio vectorial* $\mathbb{V}$*, el conjunto* $\operatorname{span}(v_1, v_2, ..., v_p)$ *es un subespacio vectorial de* $\mathbb{V}$*. Dicho subespacio vectorial comúnmente se denomina subespacio generado por los vectores* $v_1, v_2, ..., v_p$*.*

Conjuntos linealmente independientes. Bases.

Un conjunto de vectores $v_1, v_2, ..., v_p$ de un espacio vectorial $\mathbb{V}$ se denomina linealmente independiente si la ecuación vectorial

$$x_1 v_1 + x_2 v_2 + \cdots + x_p v_p = 0,$$

tiene como única solución la trivial $x_1 = \cdots = x_p = 0$.

Un conjunto de vectores $v_1, v_2, ..., v_p$ se denomina linealmente dependiente si existen los valores $x_1, x_2, \cdots, x_p$ no todos iguales a cero tales que se verifique la ecuación vectorial

$$x_1 v_1 + x_2 v_2 + \cdots + x_p v_p = 0.$$

Se dice que un conjunto infinito de vectores es linealmente independiente si cualquier subsistema finito del mismo es linealmente independiente. En caso contrario se dice que el sistema es dependiente.

Las siguientes propiedades se pueden verificar fácilmente:

1. *Un conjunto* $S = \{v_1, v_2, ..., v_p\}$ *de dos o más vectores es linealmente dependiente si y solo si al menos uno de los vectores del conjunto es combinación lineal de los demás.*

2. *Un conjunto* $S = \{v_1, v_2, ..., v_p\}$ *de dos o más vectores de* $\mathbb{V}$ *con alguno de los vectores* $v_i = 0$ $(1 \leq i \leq p)$ *es necesariamente un conjunto de vectores linealmente dependientes, es decir si alguno de los vectores de* S *es el vector nulo entonces* S *es un conjunto de vectores linealmente dependientes.*

3. *Dos vectores* v_1 *y* v_2 *de* $\mathbb{V}$ *son linealmente dependientes si y solo si son proporcionales, es decir, si existe un escalar* α *tal que* $v_1 = \alpha v_2$ *o* $v_2 = \alpha v_1$

Los vectores linealmente independientes de un espacio vectorial juegan un papel fundamental en el estudio de los sistemas lineales gracias a la siguiente definición:

Definición 4.1.6 *Dado un subespacio vectorial H del espacio vectorial $\mathbb{V}$ diremos que el conjunto de vectores $B = \{b_1, b_2, ..., b_p\}$ de $\mathbb{V}$ es una base de H si*

i) *B es un conjunto de vectores linealmente independientes*

ii) *$H = \operatorname{span}(b_1, b_2, ..., b_p)$, es decir, B genera a todo H.*

En particular si H coincide con $\mathbb{V}$, entonces B es una base de todo el espacio vectorial $\mathbb{V}$.

Por ejemplo, si tomamos una matriz $n \times n$ invertible, entonces sus columnas $a_1, ..., a_n$ son linealmente independientes y además se tiene que $\mathbb{R}^n = \operatorname{span}(a_1, ..., a_n)$. Por tanto $B = \{a_1, ..., a_n\}$ es una base de $\mathbb{R}^n$. En particular, si $A = I_n$, la matriz identidad $n \times n$, las columnas $e_1, e_2, ..., e_n$ de la misma son una base de $\mathbb{R}^n$ la cual se conoce como base canónica de $\mathbb{R}^n$.

Otro ejemplo lo constituye el conjunto de vectores $S = \{1, t, t^2, ..., t^n\}$ del espacio vectorial $\mathbb{P}_n$. Es fácil comprobar que dichos vectores son linealmente independientes y que $\operatorname{span}(1, t, t^2, ..., t^n) = \mathbb{P}_n$. S se conoce como la base canónica de $\mathbb{P}_n$.

El siguiente teorema es de gran importancia en las aplicaciones.

Teorema 4.1.7 *Si un espacio vectorial $\mathbb{V}$ tiene una base de n vectores $B = \{b_1, b_2, ..., b_n\}$, entonces cualquier conjunto con más de n vectores de $\mathbb{V}$ es linealmente dependiente. Más aún, si un espacio vectorial $\mathbb{V}$ tiene una base de n vectores $B = \{b_1, b_2, ..., b_n\}$, entonces cualquier otra base de $\mathbb{V}$ tendrá que tener n vectores de $\mathbb{V}$.*

Por tanto, el menor número de vectores linealmente independientes que generan un espacio vectorial es una propiedad intrínseca de dicho espacio. Dicho número se denomina *dimensión del espacio vectorial*.

Un espacio vectorial es de dimensión finita n si $\mathbb{V}$ está generado por una base de n elementos, es decir, si $\mathbb{V} = \operatorname{span}(b_1, ..., b_n)$, donde $B = \{b_1, ..., b_n\}$ es una base de $\mathbb{V}$ y lo escribiremos de la forma $\dim V = n$. En el caso que $\mathbb{V} = \{0\}$ sea el espacio vectorial nulo, $\dim\{0\} = 0$. Si $\mathbb{V}$ no puede ser generado por una base finita de vectores, entonces diremos que $\mathbb{V}$ es de dimensión infinita y lo denotaremos por $\dim V = \infty$.

Así, $\dim \mathbb{R}^n = n$, $\dim \mathbb{P}_n = n + 1$, $\dim C([a, b]) = \infty$, y $\dim \ell^p = \infty$.

4.2. Espacios normados y de Banach

Definición 4.2.1 *Un espacio vectorial $\mathbb{X}$ (real o complejo) se denomina espacio normado si para todo $x \in \mathbb{X}$ existe un número real denominado norma, y que denotaremos por $\|x\|$, que cumple con las condiciones*

1. *Para todo $x \in \mathbb{X}$, $\|x\| \geq 0$ y si $\|x\| = 0$ entonces $x = 0$.*
2. *Para todo $x \in \mathbb{X}$ y $\lambda \in \mathbb{R}$ (o $\mathbb{C}$), $\|\lambda x\| = |\lambda| \|x\|$,*
3. *Para todos $x, y \in \mathbb{X}$ se tiene la desigualdad triangular*

$$\|x + y\| \leq \|x\| + \|y\|. \tag{4.2.1}$$

Es evidente que si en un espacio normado $\mathbb{X}$ definimos la función $\rho(x, y) = \|x - y\|$, esta satisface los axiomas de la Definición 3.1.1, i.e., todo espacio normado es un espacio métrico. La función ρ anterior se denomina métrica inducida por la norma.

Definición 4.2.2 *Un espacio normado completo (en la métrica inducida por la norma) se denomina espacio de Banach.*

Ejemplo 4.2.3 *$\mathbb{X} = \mathbb{R}^n$ ($\mathbb{C}^n$) el espacio de las n-tuplas $x = (x_1, x_2, \ldots, x_n)$ con la norma $\|x\| = \sqrt{\sum_{k=1}^n |x_k|^2}$, es un espacio de Banach.*

Ejercicio 4.2.4 *¿Qué ocurre con $\mathbb{X} = \mathbb{R}^n$ ($\mathbb{C}^n$) si usamos las normas $\|x\|_p = \left(\sum_{k=1}^n |x_k|^p\right)^{1/p}$, $p \geq 1$? ¿Y con la norma $\|x\| = \max_{k=1,\ldots,n} |x_k|$?*

En ambos casos los correspondientes espacios son de Banach. ■

Ejemplo 4.2.5 *Sea $\mathbb{X} = C([a, b])$ el espacio de las funciones continuas definidas sobre el segmento $[a, b]$ y definamos la norma $\|f\| = \left(\int_a^b |f(x)|^p dx\right)^{1/p}$, $p \geq 1$. Este espacio es un espacio normado pero no de Banach (¿por qué?).*

Ejemplo 4.2.6 *Sea $\mathbb{X} = C([a, b])$ el espacio de las funciones continuas definidas sobre el segmento $[a, b]$. Definamos la norma $\|f\| = \max_{x \in [a,b]} |f(x)|$.*

Este espacio es un espacio de Banach.

Ejemplo 4.2.7 *Sea $\mathbb{X}$ el espacio de las sucesiones $x = (x_1, x_2, \dots, x_n, \dots)$ tales que $\sum_{k=1}^{\infty} |x_k|^p < +\infty$ con la norma $\|x\| = \left(\sum_{k=1}^{\infty} |x_k|^p\right)^{1/p}$, $p \geq 1$. Dicho espacio lo denotaremos por ℓ^p y es un espacio de Banach.*

Ejemplo 4.2.8 *El espacio $\mathbb{X}$ de las sucesiones $x = (x_1, x_2, \dots, x_n, \dots)$ acotadas con la métrica $\|x\| = \sup_{k\in\mathbb{N}} |x_k|$, es un espacio de Banach (ver el Problema 3.16).*

Esta claro que todo espacio normado es un espacio métrico con la métrica inducida por la norma. Una pregunta inmediata es si el recíproco es cierto, es decir, si todo espacio métrico es normado. Para responder a esta pregunta consideremos el espacio $\mathbb{X}$ de todas las sucesiones reales $x = (x_1, x_2, \dots, x_n, \dots)$ con la métrica (ver Ejemplo 3.1.13)

$$\rho(x, y) = \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{|x_j - y_j|}{1 + |x_j - y_j|}.$$

Esta métrica no puede ser inducida por ninguna norma ya que de ella nunca podremos obtener la propiedad 2 de la Definición 4.2.1. De hecho una consecuencia inmediata de dicha definición es el siguiente lema:

Lema 4.2.9 *Sea $\mathbb{X}$ un espacio normado. Entonces, la métrica ρ inducida por la norma satisface las siguientes condiciones:*

$$\rho(x + z, y + z) = \rho(x, y), \qquad \rho(\lambda x, \lambda y) = |\lambda| \rho(x, y).$$

Como ya vimos, los espacios normados son espacios métricos con la métrica ρ inducida por la norma $\rho(x, y) = \|x - y\|$ y por tanto en ellos podemos definir la convergencia de sucesiones, sucesiones de Cauchy, etc.. Así, por ejemplo, la Definición 3.5.2 se transformaría en la siguiente:

Definición 4.2.10 *Una sucesión $(x_n)_n$ de elementos de un espacio normado $\mathbb{X}$ es convergente, si existe un $x \in \mathbb{X}$ tal que para todo $\varepsilon > 0$ y un $N \in \mathbb{N}$ tal que para todo $n > N$, $\|x_n - \| < \varepsilon$. En ese caso escribiremos que $x_n \overset{n\to\infty}{\longrightarrow} x$ o que $\lim_{n\to\infty} x_n = x$. En caso contrario, si no existe dicho x,* *diremos que $(x_n)_n$ es divergente.*

Además, al tener los espacios normados una estructura algebraica podemos definir, entre otras cosas, las series.

Definición 4.2.11 *Dada una sucesión de elementos $(x_n)_n$ de un espacio normado $\mathbb{X}$ definiremos la sucesión de sumas parciales $(s_n)_n$*

$$s_n = \sum_{k=1}^{n} x_k, \qquad \forall n \in \mathbb{N}.$$

Si s_n converge (en norma) a cierto $s \in \mathbb{X}$, cuando $n \to \infty$, diremos que la serie es convergente en $\mathbb{X}$ y s es su suma. Si converge la serie $\sum_{k=1}^{n} \|x_k\|$, diremos que la serie converge absolutamente.

Teorema 4.2.12 *Sea $\mathbb{X}$ un espacio de Banach (normado y completo). Entonces toda serie absolutamente convergente es convergente.*

Demostración: Sea $(x_n)_n$ absolutamente convergente, i.e., la serie ***numérica*** $\sum_{k=1}^{\infty} \|x_n\|$ converge. Entonces, obviamente $\sum_{k=n+1}^{\infty} \|x_n\| \stackrel{n\to\infty}{\longrightarrow} 0$, o equivalentemente, su sucesión de sumas parciales $(\tilde{s}_n)_n$ es de Cauchy, i.e.,

$$\forall \varepsilon > 0, \; \exists N \in \mathbb{N} \mid \forall n > N, \forall p \in \mathbb{N}, \; |\tilde{s}_{n+p} - \tilde{s}_n| = \|x_{n+1}\| + \cdots + \|x_{n+p}\| < \varepsilon.$$

Entonces, para todo $\varepsilon > 0$, existe un $N \in \mathbb{N}$ tal que, para todo $n > N$ y todo $p \in \mathbb{N}$,

$$\|s_n - s_{n+p}\| = \left\| \sum_{k=n+1}^{n+p} x_n \right\| \leq \sum_{k=n+1}^{n+p} \|x_n\| < \varepsilon,$$

es decir, la sucesión $(s_n)_n$ de las sumas parciales es fundamental. Como $\mathbb{X}$ es completo, entonces existe un $s \in \mathbb{X}$ tal que $\lim_{n\to\infty} s_n = s$. ■

El teorema anterior no es cierto si $\mathbb{X}$ no es completo. Como ejercicio se proponemos al lector que encuentre un contraejemplo.

Ejemplo 4.2.13 *Prueba que si $\mathbb{X}$ es un espacio normado, entonces toda serie absolutamente convergente es convergente si y solo si $\mathbb{X}$ es completo.*

Si $\mathbb{X}$ es completo se sigue del teorema anterior que si la serie es absolutamente convergente es convergente. Probemos el recíproco, i.e., asumamos que toda serie absolutamente convergente es convergente. Sea

$(x_n)_n$ una sucesión de Cauchy cualquiera, i.e.,

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \ \forall n > N, \forall p \in \mathbb{N}, \quad \|x_n - x_{n+p}\| < \varepsilon.$$

Entonces, existe (¿por qué?) una sucesión creciente de números naturales $(n_k)_k$ tales que $n_1 < n_2 < n_3 < \cdots$

$$\|x_{n_{k+1}} - x_{n_k}\| \leq \frac{1}{2^k}, \qquad k = 1, 2, 3, \dots .$$

Además, la serie

$$\sum_{k=1}^{\infty} \|x_{n_{k+1}} - x_{n_k}\| \leq \sum_{k=1}^{\infty} \frac{1}{2^k} < +\infty.$$

Entonces, la serie $\sum_{k=1}^{\infty}(x_{n_{k+1}} - x_{n_k})$, converge a cierto $x \in \mathbb{X}$. Pero esta serie es una serie telescópica, y sus sumas parciales son tales que

$$\sum_{k=1}^{N}(x_{n_{k+1}} - x_{n_k}) = x_{n_N} - x_{n_1} \to x$$

y, por tanto,

$$x = \sum_{k=1}^{\infty}(x_{n_{k+1}} - x_{n_k}) = \lim_{k\to\infty}(x_{n_k} - x_{n_1}) \ \Rightarrow\ \lim_{k\to\infty} x_{n_k} = x + x_{n_1},$$

es decir, la sucesión $(x_{n_k})_k$ es convergente, pero entonces la sucesión original $(x_n)_n$ también converge (ver el punto 2 del Problema 3.13). Como $(x_n)_n$ era una sucesión arbitraria de Cauchy, entonces $\mathbb{X}$ es completo. ■

Definición 4.2.14 *Sea $\mathbb{X}$ un espacio normado y sea M un subespacio vectorial de $\mathbb{X}$. Si M es un espacio normado con la norma de $\mathbb{X}$ restringida a M se dice que M es un subespacio de $\mathbb{X}$. Si M es cerrado en $\mathbb{X}$ entonces se dice que es un subespacio cerrado.*

Definición 4.2.15 *Sea $\mathbb{X}$ un espacio normado. Sea $(e_n)_n$ una sucesión de elementos de $\mathbb{X}$ tal que, para todo $x \in \mathbb{X}$, existe una única sucesión de escalares $(\alpha_n)_n$ tales que $\|x - (\alpha_1 e_1 + \cdots + \alpha_n e_n)\| \overset{n\to\infty}{\longrightarrow} 0$. Dicha sucesión* *se denomina base de Schauder.*

Ejemplo 4.2.16 *Sea $\mathbb{X}$ el espacio ℓ^p de las sucesiones y sea $(e_n)_n$ la sucesión definida por $e_k = (\delta_{i,k})_i$, i.e., la sucesión de vectores de ℓ^p con 1 en la posición $i = k$ y 0 en el resto es una base de Schauder.*

En efecto, para todo $x \in \ell^p$ tenemos $x = (x_1, x_2, x_3, \dots, x_n, \dots)$. Sea $s_n = \sum_{k=1}^n \alpha_k e_k$, con $\alpha_k = x_k$ para todo $k \in \mathbb{N}$. Como $x \in \ell^p$, entonces

$$\lim_{n\to\infty} \|s_n - s\| = \lim_{n\to\infty} \left(\sum_{k=n+1}^{\infty} |x_k|^p \right)^{1/p} \overset{n\to\infty}{\longrightarrow} 0.$$

Está claro que la sucesión de escalares $(\alpha_n)_n$ es única. Para ello basta notar que, en general, la distancia

$$\|s_n - s\|^p = \sum_{k=1}^{n} |\alpha_k - x_k|^p + \sum_{k=n+1}^{\infty} |x_k|^p,$$

solo se puede hacer tan pequeño como se quiera si $\alpha_j = x_j$ para todo $j \in \mathbb{N}$. ■

Proposición 4.2.17 *Si un espacio normado $\mathbb{X}$ tiene una base de Schauder, entonces es separable.*

Demostración: Por simplicidad asumiremos el caso real. i.e., $\mathbb{K} = \mathbb{R}$ (el caso complejo se deja como ejercicio). Sea $(e_n)_n$ una base de Schauder en $\mathbb{X}$ y sea el conjunto

$$Y = \left\{ y = \sum_{j=1}^{N} \beta_j e_j,\ N \in \mathbb{N},\ \beta_i \in \mathbb{Q},\ \ j = 1, 2, \dots, N \right\},$$

i.e., el conjunto de todas las combinaciones lineales finitas con coeficientes racionales de elementos de la base $(e_n)_n$. Obviamente Y es numerable (pues $\mathbb{Q}$ lo es). Sea $x \in \mathbb{X}$ cualquiera y sea $\varepsilon > 0$. Como $(e_n)_n$ es una base de Schauder entonces existen $\alpha_1, \alpha_2, \dots, \alpha_N \in \mathbb{R}$ tales que

$$\left\| x - \sum_{j=1}^{N} \alpha_j e_j \right\| < \frac{\varepsilon}{2}.$$

Como $\mathbb{Q}$ es denso en $\mathbb{R}$, entonces existen $\beta_1, \beta_2, \dots, \beta_N \in \mathbb{Q}$ tales que

$$|\alpha_j - \beta_j| < \frac{\varepsilon}{2N\|e_j\|}.$$

Entonces, para todo $x \in \mathbb{X}$ existe un $y = \sum_{j=1}^{N} \beta_j e_j \in Y$, tal que

$$\begin{aligned}
\|x - y\| = \left\| x - \sum_{j=1}^{N} \beta_j e_j \right\| &= \left\| x - \sum_{j=1}^{N} \alpha_j e_j + \sum_{j=1}^{N} (\alpha_j - \beta_j) e_j \right\| \\
&\leq \left\| x - \sum_{j=1}^{N} \alpha_j e_j \right\| + \left\| \sum_{j=1}^{N} (\alpha_j - \beta_j) e_j \right\| \\
&\leq \left\| x - \sum_{j=1}^{N} \alpha_j e_j \right\| + \sum_{j=1}^{N} |\alpha_j - \beta_j| \|e_j\| < \varepsilon
\end{aligned}$$

■

El recíproco no es cierto en general. P.H. Enflo en 1973 encontró un espacio de Banach, separable que no tiene ninguna base de Schauder.

Finalmente debemos mencionar que todo espacio normado se puede completar y convertirlo en un espacio de Banach. De hecho como consecuencia directa del Teorema 3.5.26 se tiene el siguiente teorema:

Teorema 4.2.18 *Sea $(\mathbb{X}, \|.\|)$ un espacio normado. Entonces existe un espacio de Banach $\widehat{\mathbb{X}}$ y una isometría A de $\mathbb{X}$ en $W \subset \widehat{\mathbb{X}}$, tal que W es denso en $\widehat{\mathbb{X}}$. Además, $\widehat{\mathbb{X}}$ es único excepto isometrías.*

4.3. Espacios normados de dimensión finita

Comenzaremos con un lema técnico.

Lema 4.3.1 *Sean n vectores cualesquiera $x_1, \dots, x_n$ linealmente independientes de un espacio normado $\mathbb{X}$. Entonces, existe un número real $c > 0$ tal que cualesquiera sean los escalares $\alpha_1, \dots, \alpha_n$,*

$$\|\alpha_1 x_1 + \cdots + \alpha_n x_n\| \geq c(|\alpha_1| + \cdots + |\alpha_n|). \tag{4.3.1}$$

Demostración: Sea $s = |\alpha_1| + \cdots + |\alpha_n|$. Si $s = 0$ el lema es trivial así que asumiremos $s > 0$. Dividiendo por s 4.3.1 se sigue que 4.3.1 es equivalente a probar que si $x_1, \ldots, x_n$ son linealmente independientes, entonces existe un número real $c > 0$ tal que cuales quiera sean los los escalares $\beta_1, \ldots,$ β_n, con $\sum_{k=1}^n |\beta_k| = 1$

$$\|\beta_1 x_1 + \cdots + \beta_n x_n\| \geq c.$$

Supongamos que la desigualdad anterior es falsa. Entonces ha de existir (¿por qué?) una sucesión $(y_m)_m \subset \mathbb{X}$ tal que

$$y_m = \beta_1^{(m)} x_1 + \cdots + \beta_n^{(m)} x_n, \qquad \sum_{k=1}^n |\beta_k^{(m)}| = 1, \text{ y } \|y_m\| \overset{m\to\infty}{\longrightarrow} 0.$$

De la condición $\sum_{k=1}^n |\beta_k^{(m)}| = 1$ se sigue que las n sucesiones numéricas $(\beta_k^{(m)})_m$, $k = 1, \ldots, n$, son acotadas. Sea la sucesión $(\beta_1^{(m)})_m$ acotada, entonces por el Teorema de Bolzano-Weierstrass de ella se puede extraer una subsucesión convergente $\beta_1^{(m_j)} \overset{j\to\infty}{\longrightarrow} \beta_1$. Escojamos de cada una de las sucesiones restantes $(\beta_k^{(m)})_m$, $k = 2, \ldots, n$, con las subsucesiones definidas por los índices m_j de antes. Entonces la sucesión $(\beta_2^{(m_j)})_j$ es acotada y por Bolzano-Weierstrass de ella se puede extraer una subsucesión convergente $\beta_2^{(j_l)} \overset{l\to\infty}{\longrightarrow} \beta_2$. Además, si escogemos los índices j_l definidos por esta sucesión, la subsucesión $(\beta_1^{(j_l)})_j \overset{l\to\infty}{\longrightarrow} \beta_1$ (¿por qué?). Continuando este proceso tenemos que existe una subsucesión de índices l_i tales que $\beta_k^{(l_i)} \overset{i\to\infty}{\longrightarrow} \beta_k$ para todos los $k = 1, 2, \ldots, n$. Además, dicha sucesión de índices define una subsucesión $(y_{l_i})_i$ de $(y_m)_m$ tal que

$$y_{l_i} = \sum_{k=1}^n \beta_k^{(l_i)} x_k, \qquad \beta_k^{(l_i)} \overset{i\to\infty}{\longrightarrow} \beta_k.$$

Luego

$$\lim_{i\to\infty} y_{l_i} = \sum_{k=1}^n \beta_k x_k := y \;\text{ y }\; \sum_{k=1}^n |\beta_k| = 1.$$

De lo anterior se sigue que no todos los β_k pueden ser ceros al mismo tiempo. Como los vectores $x_1, \ldots, x_n$ son linealmente independientes y no todos los $\beta_1, \ldots \beta_n$ son cero, entonces $y \neq 0$ (¿por qué?). Ahora bien, como la norma es una aplicación continua (ver Problema 4.2), entonces se tiene

$$\lim_{i\to\infty} y_{l_i} = y \quad \Rightarrow \quad \lim_{i\to\infty} \|y_{l_i}\| = \|y\|,$$

pero como $\|y_m\| \stackrel{m\to\infty}{\longrightarrow} 0$, entonces $\lim_{i\to\infty} \|y_{l_i}\| = 0$, luego $\|y\| = 0$ de donde se sigue que $y = 0$ lo cual es una contradicción. ■

Como corolario tenemos el siguiente teorema de completitud:

Teorema 4.3.2 *Todo subespacio M de dimensión finita de un espacio normado es completo. En particular, todo espacio normado de dimensión finita es completo.*

Demostración: Sea $(y_m)_m$ una sucesión de Cauchy arbitraria de M. Sea $\dim M = n$. Entonces para todo $y_m \in M$, existen los números (únicos) $\alpha_k^{(m)}$, $k = 1, 2, \dots, n$, tales que $y_m = \sum_{k=1}^n \alpha_k^{(m)} e_k$. Como $(y_m)_m$ es de Cauchy, entonces, por un lado

$$\forall \varepsilon > 0, \ \exists N \in \mathbb{N}, \mid \forall m > N, \ \forall p \in \mathbb{N}, \quad \|y_m - y_{m+p}\| < \varepsilon,$$

y por el otro, existe un $c > 0$ (por el lema anterior) tal que

$$\varepsilon > \left\| \sum_{k=1}^n (\alpha_k^{(m)} - \alpha_k^{(m+p)}) e_k \right\| \geq c \sum_{k=1}^n |\alpha_k^{(m)} - \alpha_k^{(m+p)}| \Rightarrow \sum_{k=1}^n |\alpha_k^{(m)} - \alpha_k^{(m+p)}| \leq \frac{\varepsilon}{c},$$

i.e., las sucesiones $(\alpha_k^{(m)})_m$ son de Cauchy y, por tanto, convergen (los $\alpha_k^{(m)}$ son números reales o complejos), a ciertos α_k, $k = 1, 2, \dots, n$. Sea $y = \sum_{k=1}^n \alpha_j e_j \in M$. Entonces

$$\|y_m - y\| = \left\| \sum_{k=1}^n (\alpha_k^{(m)} - \alpha_k) e_k \right\| \leq \sum_{k=1}^n |\alpha_k^{(m)} - \alpha_k| \|e_k\| \stackrel{m\to\infty}{\longrightarrow} 0,$$

lo que implica que $\lim\limits_{m\to\infty} y_m = y \in M$, i.e., M es completo. ■

Corolario 4.3.3 *Todo subespacio M de dimensión finita de un espacio de normado (en particular de Banach) $\mathbb{X}$ es cerrado en $\mathbb{X}$.*

Demostración: M es en sí mismo un espacio normado y por hipótesis es de dimensión finita, luego, por el Teorema 4.3.2, es completo. Pero si $M \subset \mathbb{X}$ es completo, entonces repitiendo la demostración de la primera implicación del Teorema 3.5.14 se sigue que M es cerrado en $\mathbb{X}$ (ver Nota 3.5.15). ■

El siguiente ejemplo muestra que la condición de dimensión finita no se puede eliminar.

Ejemplo 4.3.4 *Sean los vectores linealmente independientes* $x_1, \ldots, x_n \in \mathbb{X}$*,* $\mathbb{X}$ *espacio normado y sea* $M = \operatorname{span}(x_1, x_2, \cdots, x_n)$*. Dado que* $M \subset \mathbb{X}$ *es de dimensión finita, el corolario anterior 4.3.3 implica que* M *es cerrado.*

Veamos que, efectivamente, la condición de dimensión finita es esencial. Para ello consideremos el espacio $C^{\infty}_{[0,1]}$ del Ejemplo 4.2.6 y sea $M = \operatorname{span}(x_1, x_2, x_2, \cdots)$, donde $x_n := x_n(t) = t^{n-1}$, i.e., M es el espacio de todos los polinomios de cualquier grado. M no puede ser cerrado pues hay infinidad de funciones continuas que son puntos de acumulación de M pero no pertenecen a M. Por ejemplo, si elegimos $f(t) = e^t$, entonces, usando el Teorema de Taylor tenemos que $P_n(t) = \sum_{k=0}^{n} \frac{t^n}{n!}$, es tal que

$$\|e^t - P_n(t)\| = \max_{t\in[0,1]} |e^t - P_n(t)| = \max_{t\in[0,1]} \left| \frac{e^{\xi} t^{n+1}}{(n+1)!} \right| = \frac{e^{\xi}}{(n+1)!} \to 0,\ \xi \in (0,1),$$

cuando $n \to \infty$, luego e^t es un punto de acumulación de M pero obviamente no es un polinomio, luego M no puede ser cerrado. ∎

Definición 4.3.5 *Una norma* $\|\cdot\|$ *en un espacio vectorial* $\mathbb{X}$ *es equivalente a otra norma* $\|\cdot\|'$ *si existen dos números reales* a, b *positivos (*$a > 0$*,* $b > 0$*) tales que para todo* $x \in \mathbb{X}$

$$a\|x\|' \leq \|x\| \leq b\|x\|'.$$

Ejercicio 4.3.6 *Sea* $\mathbb{X}$ *un conjunto dado y sean* $\|\cdot\|$ *y* $\|\cdot\|'$ *dos normas equivalentes definidas en* $\mathbb{X}$*. Prueba que toda sucesión de Cauchy en* $(\mathbb{X}, \|\cdot\|)$ *si y solo si lo es en* $(\mathbb{X}, \|\cdot\|')$*. ¿Qué consecuencia tiene esta propiedad?*

Esto es una consecuencia inmediata de la definición anterior dado que si dos normas son equivalentes entonces, si para todo $\epsilon > 0$ tan pequeño como se quiera se tiene $\|x_n - x_m\|' < \epsilon$, entonces la distancia $\|x_n - x_m\| < b\epsilon$, seguirá siendo tan pequeña como se quiera, y si $\|x_n - x_m\| < \epsilon$, entonces $\|x_n - x_m\|' < \epsilon/a$, también seguirá siendo tan pequeña como se quiera. En otras palabras, todas las definiciones que involucran la norma no se ven afectadas si cambiamos una norma por otra equivalente, por ejemplo no solo las sucesiones de Cauchy seguirán siendo de Cauchy, sino que las propiedades topológicas de los espacios y sus subespacios no cambian, etc. ∎

Teorema 4.3.7 *Sea $\mathbb{X}$ un espacio vectorial de dimensión finita. Entonces cualquier norma $\|\cdot\|$ en $\mathbb{X}$ es equivalente a cualquier otra norma $\|\cdot\|'$.*

Es decir, en cualquier espacio normado de dimensión finita la propiedad de convergencia (o divergencia) es independiente de la norma.

Demostración: Sea $\dim \mathbb{X} = n$ y sea $e_1, \dots, e_n$ una base de $\mathbb{X}$. Entonces para todo $x \in \mathbb{X}$ existen (y son únicos) los números $\alpha_1, \dots, \alpha_n$ tales que $x = \sum_{j=1}^n \alpha_j e_j$. Entonces, por el Lema 4.3.1

$$\|x\| \geq c(|\alpha_1| + \cdots + |\alpha_n|) \quad \Rightarrow |\alpha_1| + \cdots + |\alpha_n| \leq \frac{\|x\|}{c}.$$

Pero

$$\|x\|' = \left\| \sum_{j=1}^n \alpha_j e_j \right\|' \leq \sum_{j=1}^n |\alpha_j| \|e_j\|' \leq \max_j \|e_j\|' \sum_{j=1}^n |\alpha_j| \leq \frac{k}{c} \|x\|,$$

donde hemos denotado por k al $\max_j \|e_j\|$. Intercambiando las normas obtenemos $\|x\| \leq \frac{k'}{c'} \|x\|'$, de donde se sigue el resultado. ■

Ya vimos en el apartado 3.5 que en un espacio métrico todo subconjunto compacto (Definición 3.5.5) era cerrado y acotado (Lema 3.5.7) y también vimos que el recíproco no es cierto (ver también el Problema 4.17). Esta afirmación obviamente es válida para los espacios normados al ser estos espacios métricos con la métrica inducida por la norma. Sin embargo, si el espacio normado es de dimensión finita, entonces todo cerrado y acotado es necesariamente compacto. Así tenemos el siguiente teorema:

Teorema 4.3.8 *En un espacio normado $\mathbb{X}$ de dimensión finita, todo subconjunto $M \subset \mathbb{X}$ es compacto si y solo si es cerrado y acotado.*

Demostración: Si M es compacto entonces, por el Lema 3.5.7, es cerrado y acotado. Probemos el recíproco. Sea M cerrado y acotado. Probemos que M es compacto. Sea $n = \dim \mathbb{X}$ y sea $e_1, \dots, e_n$ una base. Sea $(x_m)_m$ una sucesión de elementos de M, entonces $x_m = \sum_{k=1}^n \alpha_k^{(m)} e_k \in \mathbb{X}$. Como M es acotado, x_m también lo es, i.e., existe $K > 0$ tal que, para todo $m \in \mathbb{N}$, $\|x_m\| \leq K$, luego, por el Lema 4.3.1 tenemos que

$$K \geq \|x_m\| = \left\| \sum_{k=1}^n \alpha_k^{(m)} e_k \right\| \geq c \sum_{k=1}^n |\alpha_k^{(m)}|.$$

Luego las sucesiones $(\alpha_k^{(m)})_m$ son acotadas para cada k y, por tanto, existe una subsucesión convergente (Lema de Bolzano-Weierstrass) a α_k, para todo $k = 1, 2, \ldots, n$ –debemos hacer una construcción similar a la de la prueba del lema 4.3.1–. Así, existe una subsucesión $(x_{n_k})_k$ que converge a $x = \sum_{k=1}^n \alpha_k e_k$. Como M es cerrado $x \in M$. Luego toda sucesión $(x_n)_n$ de m tiene una subsucesión convergente en M, i.e., M es compacto. ■

4.4. Aplicaciones lineales

Gracias a la estructura algebraica de los espacios normados podemos definir un caso particular de gran importancias de los operadores discutidas en el apartado **3.2**: las aplicaciones lineales. Como allí $\mathcal{D}(T)$ denotará el dominio (que asumiremos es un subespacio vectorial de $\mathbb{X}$) de la aplicación T e $\mathcal{I}(T)$ la imagen de T. Asumiremos que los espacios $\mathbb{X}$ e $\mathbb{Y}$ son *espacios vectoriales sobre el mismo cuerpo* $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$) y que tenemos el operador $A : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathbb{Y}$.

Definición 4.4.1 *Una aplicación (operador) $A : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathbb{Y}$ es lineal si para todos $\alpha, \beta \in \mathbb{K}$, y todos $x, y \in \mathcal{D}(T)$*

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y).$$

Veamos algunos ejemplos de operadores lineales:

Ejemplo 4.4.2 *El operador identidad $I : \mathbb{X} \mapsto \mathbb{X}$, definido por $y = Ix = x$ para todo $x \in \mathbb{X}$.*

Ejemplo 4.4.3 *El operador nulo $\Theta : \mathbb{X} \mapsto \mathbb{Y}$, definido por $y = \Theta x = 0$ para todo $x \in \mathbb{X}$.*

Ejemplo 4.4.4 *Sea $\mathbb{P}$ el espacio de los polinomios reales $p(t)$ (o complejos) de cualquier grado. Definimos el operador $D : \mathbb{P} \mapsto \mathbb{P}$, $y(t) = Dp(t) = p'(t)$ que no es más que operador de derivación (o derivada).*

Ejemplo 4.4.5 *El operador $T : \mathbb{R}^n \mapsto \mathbb{R}^m$, definido por $y = Tx = A \cdot x$, donde A es una matriz $n \times m$, x e y son los correspondientes vectores de $\mathbb{R}^n$ y $\mathbb{R}^m$ respectivamente, y "$\cdot$" denota la multiplicación usual de matrices.*

Ejemplo 4.4.6 *Sea $C([a, b])$ el espacio de las funciones continuas y sea $g(t) \in C([a, b])$. Definimos el operador $M : C([a, b]) \mapsto C([a, b])$, $y(t) = Mf(t) = g(t)f(t)$ que denominaremos operador multiplicación por $g(t)$.*

Definición 4.4.7 *Llamaremos espacio nulo o núcleo de T al espacio $\mathcal{N}(T)$ de todos los vectores $x \in \mathcal{D}(T)$ tales que $Tx = 0$.*

Teorema 4.4.8 *Sea $T : \mathbb{X} \mapsto \mathbb{Y}$ una aplicación lineal. Entonces*

1. *$\mathcal{I}(T)$ es un subespacio vectorial de $\mathbb{Y}$.*
2. *Si $\dim \mathcal{D}(T) = n < \infty$, entonces $\dim \mathcal{I}(T) \leq n$.*
3. *$\mathcal{N}(T)$ es un subespacio vectorial de $\mathcal{D}(T)$.*

Demostración: 1. Sean $y_1, y_2 \in \mathcal{I}(T)$. Entonces existen $x_1, x_2 \in \mathcal{D}(T)$ tales que $y_1 = Tx_1$ e $y_2 = Tx_2$. Como T es lineal, entonces para todo $\alpha, \beta \in \mathbb{K}$,

$$\mathcal{I}(T) \ni T(\alpha x_1 + \beta x_2) = \alpha T x_1 + \beta T x_2 = \alpha y_1 + \beta y_2,$$

y obviamente $T0 = 0$. Luego por el Teorema 4.1.4, $\mathcal{I}(T)$ es un subespacio vectorial de $\mathbb{Y}$.

2. Sean $y_1, \ldots, y_{n+1} \in \mathcal{I}(T)$, $n + 1$ vectores. Existen $n + 1$ vectores $x_1, \ldots, x_{n+1}$ de $\mathcal{D}(T)$ tales que $Tx_k = y_k$, $k = 1, \ldots, n+1$. Como $\dim \mathcal{D}(T) = n$, existen $n+1$ números $\alpha_1, \ldots, \alpha_{n+1}$, no todos nulos, tales que $\sum_{k=1}^{n+1} \alpha_k x_k = 0$ (¿por qué?). Pero T es lineal luego

$$T\left(\sum_{k=1}^{n+1} \alpha_k x_k\right) = \sum_{k=1}^{n+1} \alpha_k T(x_k) = \sum_{k=1}^{n+1} \alpha_k y_k = 0,$$

i.e., los vectores $y_1, \ldots, y_{n+1} \in \mathcal{I}(T)$ son linealmente dependientes, por tanto $\dim \mathcal{I}(T) \leq n$.

3. Este apartado es similar al primero y se deja como ejercicio. ∎

Una pregunta natural e importante que surge al estudiar los operadores lineales es saber cuándo existe el operador inverso. El siguiente resultado responde a dicha cuestión:

Teorema 4.4.9 *Sea $T : \mathbb{X} \mapsto \mathbb{Y}$ una aplicación lineal con $\mathcal{D}(T) \subset \mathbb{X}$ y $\mathcal{I}(T) \subset \mathbb{Y}$. Entonces*

1. *Existe la aplicación inversa T^{-1} de T, si y solo si $Tx = 0$ implica $x = 0$.*
2. *Si existe T^{-1}, entonces T^{-1} es lineal.*
3. *Si T es invertible y $\dim \mathcal{D}(T) = n < \infty$, entonces $\dim \mathcal{I}(T) = \dim \mathcal{D}(T) = n$.*

Demostración: Ante todo recordemos que para que una aplicación tenga inversa ha de ser inyectiva, i.e., $x_1 \neq x_2 \Rightarrow Tx_1 \neq Tx_2$, o, equivalentemente, $Tx_1 = Tx_2 \Rightarrow x_1 = x_2$.

Comenzaremos probando 2. Supongamos que existe T^{-1}. Probemos que para todos $y_1, y_2 \in \mathcal{I}(T)$ y $\alpha, \beta \in \mathbb{K}$, $T^{-1}(\alpha y_1 + \beta y_2) = \alpha T^{-1} y_1 + \beta T^{-1} y_2$. Como existe T^{-1}, entonces, para todos $y_1, y_2 \in \mathcal{I}(T)$ existen $x_1, x_2 \in \mathcal{D}(T)$ tales que $x_1 = T^{-1} y_1$ y $x_2 = T^{-1} y_2$, luego $y_1 = Tx_1$ y $y_2 = Tx_2$. Como T es lineal

$$T(\alpha x_1 + \beta x_2) = \alpha T x_1 + \beta T x_2 = \alpha y_1 + \beta y_2.$$

Si le aplicamos el operador T^{-1} a la ecuación anterior nos queda

$$T^{-1}(\alpha y_1 + \beta y_2) = \alpha x_1 + \beta x_2 = \alpha T^{-1} y_1 + \beta T^{-1} y_2.$$

Probemos ahora 1. Supongamos que $Tx = 0$ implica $x = 0$. Entonces, cualesquiera sean x_1 y x_2 tales que $Tx_1 = Tx_2$, tenemos

$$0 = Tx_1 - Tx_2 = T(x_1 - x_2) \quad \Rightarrow \quad x_1 - x_2 = 0 \quad \Rightarrow \quad x_1 = x_2,$$

i.e., T es inyectiva, luego existe su inversa T^{-1}. Por el contrario, si T tiene inversa entonces T^{-1} es lineal (por el apartado 2) así que tomando T^{-1} en la ecuación $Tx = 0$ obtenemos $x = T^{-1} 0 = 0$.

Probemos 3. Por el Teorema 4.4.8 $\dim \mathcal{I}(T) \leq \dim \mathcal{D}(T)$. Si aplicamos el mismo teorema a la aplicación lineal T^{-1} obtenemos $\dim \mathcal{I}(T^{-1}) = \dim \mathcal{D}(T) \leq \dim \mathcal{D}(T^{-1}) = \dim \mathcal{I}(T)$, i.e., $\dim \mathcal{I}(T) \geq \dim \mathcal{D}(T)$, de donde se sigue el resultado. ∎

Ejemplo 4.4.10 *Sea* $T : \mathbb{X} \mapsto \mathbb{Y}$ *una aplicación lineal y supongamos que* $\dim \mathbb{X} = n < \infty$. *Prueba que* $\dim \mathbb{X} = n = \dim \mathcal{N}(T) + \dim \mathcal{I}(T)$.

Sea $e_1, \dots, e_k$, $k \le n$ una base del núcleo $\mathcal{N}(T)$ de T. Completemos dichos vectores con los vectores linealmente independientes $e_{k+1}, \dots, e_n$ de $\mathbb{X}$ tales que $e_1, e_2, \dots e_n$ sean una base de $\mathbb{X}$. Probemos que Te_{k+1}, ... Te_n es una base de la imagen $\mathcal{I}(T)$ de T. Ante todo notemos que $\mathrm{span}\,(Te_{k+1}, \dots, Te_n) = \mathcal{I}(T)$. En efecto, para todo $y \in \mathcal{I}(T)$ existe un $x \in \mathcal{D}(T) \subset \mathbb{X}$ tal que $y = Tx$,

$$y = Tx = T(x_1e_1 + \cdots + x_ke_k + x_{k+1}e_{k+1} + x_ne_n) = 0 + x_{k+1}Te_{k+1} + \dots x_nTe_n.$$

Probemos ahora que son linealmente independientes. Supongamos que no lo son, entonces existe al menos un $x_j \ne 0$, $j = k+1, \dots, n$, tal que

$$x_{k+1}Te_{k+1} + \cdots + x_nTe_n = 0 \quad \Rightarrow \quad T(x_{k+1}e_{k+1} + \cdots + x_ne_n) = 0.$$

Pero entonces $z = x_{k+1}e_{k+1} + \cdots + x_ne_n \in \mathcal{N}(T)$, lo cual indica que $z = 0$ (pues z es combinación lineal de vectores que no pertenecen al núcleo). Pero entonces los vectores $e_{k+1}, \dots, e_n$ son linealmente dependientes lo cual es una contradicción, i.e., los vectores $Te_{k+1}, \dots Te_n$ son linealmente independientes y, por tanto, son una base de $\mathcal{I}(T)$. Entonces $\dim \mathcal{N}(T) = k$ y $\dim \mathcal{I}(T) = n - k$, de donde se deduce el resultado. ■

Ejercicio 4.4.11 *Sea* $T : \mathbb{X} \mapsto \mathbb{Y}$ *una aplicación lineal y supongamos que* $\dim \mathbb{X} = \dim \mathbb{Y} = n < \infty$. *Prueba que* $\mathcal{I}(T) = \mathbb{Y}$ *si y solo si* T^{-1} *existe.* ***Ayuda:*** *Usa el resultado del ejemplo anterior.*

Si existe T^{-1}, entonces por el apartado 3 del Teorema 4.4.9 $\dim \mathcal{N}(T) = 0$ y por el 3 $\dim \mathcal{D}(T) = \dim \mathcal{I}(T)$. Del ejemplo anterior tenemos entonces que $\dim \mathbb{X} = \dim \mathcal{I}(T)$, luego $\mathcal{I}(T) = \mathbb{Y}$. Por el contrario, si $\mathcal{I}(T) = \mathbb{Y}$, entonces del ejemplo se deduce que $\dim \mathcal{N}(T) = 0$, por lo que el apartado 1 del Teorema 4.4.9 nos asegura que T es invertible. ■

Este resultado no es cierto para dimensión infinita. En efecto, sea $\mathbb{X}$ el conjunto de las funciones infinitamente derivables en $\mathbb{R}$. Sea $D : \mathbb{X} \mapsto \mathbb{X}$, $y(t) = Dx(t) = x'(t)$ (derivada). Para cada $y(t) \in \mathbb{X}$ podemos definir $x(t) = \int_0^t y(\tau)d\tau$, de forma que $x'(t) = y(t)$ en $\mathbb{R}$, i.e., $\mathcal{I}(D) = \mathbb{X}$. Pero $Dx = 0$ no implica $x = 0$, ya que $x'(t) = 0$ para toda función constante, no necesariamente nula, es decir, no existe el inverso de D (ver el punto 1 del Teorema 4.4.9).

Definición 4.4.12 *Sean $\mathbb{X}$ e $\mathbb{Y}$ dos espacios normados y sea el operador $T : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathbb{Y}$ lineal. T es acotado si existe $c \geq 0$ tal que*

$$\|Tx\| \leq c\|x\|, \qquad \forall x \in \mathcal{D}(T). \tag{4.4.1}$$

Nota 4.4.13 *En la definición anterior se sobrentiende que $\|x\|$ es la norma en $\mathbb{X}$ y $\|Tx\|$ es en $\mathbb{Y}$.*

De la definición anterior se sigue que si T es acotado, entonces para todo $x \neq 0$,

$$\frac{\|Tx\|}{\|x\|} \leq c, \qquad \forall x \in \mathcal{D}(T),\ x \neq 0. \tag{4.4.2}$$

El menor valor de c para el cual (4.4.1) se cumple lo denotaremos por $\|T\|$ y se denomina *norma del operador* lineal T. Tomando supremos en $x \neq 0$ en (4.4.2) e ínfimos en c tenemos

$$\sup_{x\in\mathbb{X}\setminus\{0\}} \frac{\|Tx\|}{\|x\|} \leq \|T\|.$$

Por otro lado, para todo $y \neq 0$

$$\frac{\|Ty\|}{\|y\|} \leq \sup_{x\in\mathbb{X}\setminus\{0\}} \frac{\|Tx\|}{\|x\|} := c',$$

luego $\|Ty\| \leq c'\|y\|$ por lo tanto

$$\|T\| = \inf\{c : \|Ty\| \leq c\|y\|, \quad \forall y \in \mathbb{X}\} \leq c' = \sup_{x\in\mathbb{X}\setminus\{0\}} \frac{\|Tx\|}{\|x\|},$$

de donde se sigue que

$$\|T\| = \sup_{x\in\mathbb{X}\setminus\{0\}} \frac{\|Tx\|}{\|x\|} \quad \Longleftrightarrow \quad \|T\| = \sup_{\|x\|=1} \|Tx\|. \tag{4.4.3}$$

Si $T = 0$ obviamente $\|T\| = 0$. Además de (4.4.1), tomando ínfimos en c se tiene

$$\forall y \in \mathbb{X}, \quad \frac{\|Ty\|}{\|y\|} \leq \|T\| \quad \Longleftrightarrow \quad \|Ty\| \leq \|T\|\|y\|.$$

Ejercicio 4.4.14 *Prueba que la cantidad $\|T\|$ definida en (4.4.3) es una norma, es decir, se cumplen los axiomas de la definición 4.2.1.*

Usando (4.4.3) se sigue que $\|T\| \geq 0$ y solo vale cero si $\sup_{\|x\|=1} \|Tx\| = 0$, i.e., $Tx = 0$ para todo $x \in \mathcal{D}(T)$, luego $T = 0$. Además, como $(\lambda T)x = \lambda(Tx)$ se sigue que $\sup_{\|x\|=1} \|(\lambda T)x\| = |\lambda| \sup_{\|x\|=1} \|Tx\| = |\lambda|\|T\|$. Finalmente, como $\|(T+U)x\| \leq \|Tx\| + \|Ux\|$ para todo $x \in \mathbb{X}$, se tiene, tomando supremos en $\|x\| = 1$ que $\|T+U\| \leq \|T\| + \|U\|$. ■

Ejemplo 4.4.15 *El operador I del Ejemplo 4.4.2 es acotado y $\|I\| = 1$. El operador Θ del Ejemplo 4.4.3 es acotado y $\|\Theta\| = 0$. El operador D del Ejemplo 4.4.4 es no acotado. En efecto, escojamos el espacio $\mathbb{P}$ en $J = [0,1]$ e introduzcamos la norma $\|p\| = \max_{t \in J} |p(t)|$. Como $Dp(t) = p'(t)$, entonces si escogemos la sucesión $p_n(t) = t^n$, $\|p_n\| = 1$, tenemos $\|Dp_n\| = n$, luego $\|Dp_n\|/\|p_n\| = n$, que obviamente no es acotada. Finalmente, para el Ejemplo 4.4.5 de las matrices si usamos, por ejemplo, la norma $\|x\| = \left(\sum_{k=1}^n |x_k|^2\right)^{1/2}$, entonces*

$$\|Tx\| \leq c\|x\|, \qquad c = \sqrt{\sum_{k=1}^{m} \sum_{j=1}^{n} a_{kj}^2} := \|T\|,$$

donde a_{kj} son los elementos de la matriz A.

Ejercicio 4.4.16 *Sea el espacio $C^{\infty}_{[a,b]}$ de las funciones continuas con la norma del máximo del Ejemplo 4.2.6. Sea $x(t) \in C^{\infty}_{[a,b]}$ y sean los los puntos $t_1 < t_2 < \cdots < t_n$ pertenecientes $[a,b]$. Definamos el funcional*

$$f : C^{\infty}_{[a,b]} \mapsto \mathbb{R}, \quad f(x) = \sum_{k=1}^{n} c_k x(t_k), \quad c_1, \ldots, c_n \in \mathbb{R}.$$

Prueba que f es lineal y acotado y que $\|f\| = l_n := \sum_{k=1}^n |c_k|$.

La linealidad se sigue fácilmente de la definición de f. Como $\mathbb{R}$ es un espacio normado con la norma del valor absoluto, entonces

$$|f(x)| \leq \sum_{k=1}^{n} |c_k||x(t_k)| \leq \max_{t \in [a,b]} |x(t)| \sum_{k=1}^{n} |c_k| = l_n \|x\| \quad \Rightarrow \quad \|f\| \leq l_n.$$

Dividamos el intervalo $[a, b]$ en los subintervalos $[a, t_1], \ldots, [t_n, b]$. Construyamos ahora una función $\tilde{x}(t)$ de forma que $\tilde{x}(t_k) = \operatorname{signo}(c_k)$, y sea lineal entre los puntos $(0, 0)$, $(t_1, \operatorname{signo}(c_1)), \ldots, (t_n, \operatorname{signo}(c_n))$ y $(b, 0)$ (ver figura 4.1).

Obviamente $\|\tilde{x}(t)\| = 1$. Además

$$\|f\| = \sup_{\|x\|=1} |f(x)| \geq |f(\tilde{x})| = \left| \sum_{k=1}^{n} c_k \tilde{x}(t_k) \right| = \sum_{k=1}^{n} |c_k| \;\Rightarrow\; \|f\| \geq l_n,$$

de donde se deduce $\|f\| = l_n$. ■

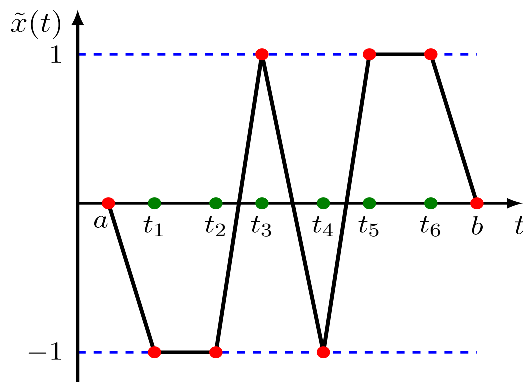


Figura 4.1: La función $\tilde{x}(t)$ con $n = 6$.

Teorema 4.4.17 *Toda aplicación lineal $T : \mathbb{X} \mapsto \mathbb{Y}$ de un espacio normado de dimensión finita $\mathbb{X}$ en otro espacio normado cualquiera $\mathbb{Y}$ es acotada.*

Demostración: Sea $\dim \mathbb{X} = n$ y sea $(e_1, \ldots, e_n)$ una base de $\mathbb{X}$. Para todo $x \in \mathbb{X}$, $x = \sum_{k=1}^{n} x_k e_k$. Entonces

$$\|Tx\| = \left\| T \sum_{k=1}^{n} x_k e_k \right\| \leq \sum_{k=1}^{n} |x_k| \|Te_k\| \leq \max_k \|Te_k\| \sum_{k=1}^{n} |x_k|.$$

Por otro lado, usando el Lema 4.3.1 tenemos que existe un $c > 0$ tal que

$$\|x\| = \left\| \sum_{k=1}^{n} x_k e_k \right\| \geq c \sum_{k=1}^{n} |x_k|.$$

Combinando ambas tenemos

$$\|Tx\| \leq \max_k \|Te_k\| \frac{\|x\|}{c} \quad \Rightarrow \quad \|Tx\| \leq \gamma \|x\|,$$

con $\gamma = \max_k \|Te_k\|/c$. ■

Para terminar este apartado probemos el siguiente teorema sobre aplicaciones lineales continuas.

Teorema 4.4.18 *Sea $T : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathbb{Y}$ una aplicación lineal de un espacio normado $\mathbb{X}$ a otro espacio normado $\mathbb{Y}$. Entonces*

1. *T es continuo si y solo si T es acotado.*
2. *Si T es continuo en algún $x_0 \in \mathcal{D}(T)$, T es continuo en $\mathcal{D}(T)$.*

Demostración: Asumiremos que T no es el operador nulo.

1. Sea T acotado y sea $x_0 \in \mathcal{D}(T)$ cualquiera. Como T es lineal y acotado, entonces

$$\|Tx - Tx_0\| = \|T(x - x_0)\| \leq \|T\|\|x - x_0\|.$$

Entonces, para todo $\varepsilon > 0$, existe un $\delta = \varepsilon/\|T\| > 0$ tal que, para todo x con $\|x - x_0\| < \delta$, $\|Tx - Tx_0\| < \varepsilon$, i.e., T es continuo en $\mathcal{D}(T)$.

Sea T lineal y continuo en $x_0 \in \mathcal{D}(T)$ cualquiera. Entonces para todo $\varepsilon > 0$, existe un $\delta > 0$, tal que, para todo x con $\|x - x_0\| < \delta$, $\|Tx - Tx_0\| < \varepsilon$. Sea $y \neq 0$ en $\mathcal{D}(T)$ cualquiera. Escojamos x tal que

$$x = x_0 + \frac{\delta}{2\|y\|} y \Rightarrow x - x_0 = \frac{\delta}{2\|y\|} y \Rightarrow \|x - x_0\| < \delta \Rightarrow \|Tx - Tx_0\| < \varepsilon.$$

Además, para dichos x tenemos, usando la linealidad de T, que

$$\|Tx - Tx_0\| = \|T(x - x_0)\| = \left\| T \frac{\delta}{2\|y\|} y \right\| = \frac{\delta}{2\|y\|} \|Ty\| \leq \varepsilon \Rightarrow \|Ty\| \leq \frac{2\varepsilon}{\delta} \|y\|,$$

luego T es acotado.

2. Nótese que en la segunda parte de la prueba anterior se probó que si T era continuo en un punto $x_0 \in \mathcal{D}(T)$, entonces era acotado en $\mathcal{D}(T)$. Pero entonces por la primera parte, al ser T acotado en $\mathcal{D}(T)$, es continuo en $\mathcal{D}(T)$. ■

Nota 4.4.19 *El teorema anterior nos indica que la acotación y continuidad para las aplicaciones lineales son conceptos equivalentes.*

Sea $\mathcal{L}(\mathbb{X},\mathbb{Y})$ el espacio de todas las aplicaciones lineales de $\mathbb{X}$ en $\mathbb{Y}$, $\mathbb{X}$ e $\mathbb{Y}$ espacios vectoriales. Definamos en dicho espacio la suma de dos aplicaciones $A+B$ y la multiplicación por un escalar λ de la forma habitual

$$\forall x \in \mathbb{X}, \quad (A+B)x = Ax + By, \quad (\lambda A)x = \lambda(Ax).$$

Con esta definición es fácil ver que $\mathcal{L}(\mathbb{X},\mathbb{Y})$ es un espacio vectorial (el elemento nulo de $\mathcal{L}(\mathbb{X},\mathbb{Y})$ es el operador nulo).

Supongamos ahora que $\mathbb{X}$ e $\mathbb{Y}$ son espacios normados. Definamos el subespacio $\mathcal{B}(\mathbb{X},\mathbb{Y}) \subset \mathcal{L}(\mathbb{X},\mathbb{Y})$ de todas las aplicaciones lineales acotadas (y, por tanto, según el Teorema 4.4.18, continuas). Entonces, como consecuencia del Ejercicio 4.4.14 se sigue que $\mathcal{B}(\mathbb{X},\mathbb{Y})$ es un espacio normado. De hecho, como veremos a continuación, es un espacio de Banach.

Ejercicio 4.4.20 *Prueba que si $\mathbb{X}$ es un espacio normado e $\mathbb{Y}$ es un espacio de Banach, entonces $\mathcal{B}(\mathbb{X},\mathbb{Y})$ es un espacio de Banach. (ver Teorema 6.3.6).*

Sea $(U_n)_n \in \mathcal{B}(\mathbb{X},\mathbb{Y})$ una sucesión de Cauchy. Entonces

$$\forall \varepsilon > 0, \quad \exists N \in \mathbb{N}, \text{ tal que } \forall n, m > N, \quad \|U_n - U_m\| < \varepsilon,$$

y, por tanto, para cada $x \in \mathbb{X}$

$$\|(U_n - U_m)x\| = \|U_n x - U_m x\| < \varepsilon \|x\|, \tag{4.4.4}$$

i.e., para cada $x \in \mathbb{X}$, la sucesión $(U_n x)_n \in \mathbb{Y}$ es de Cauchy, y como $\mathbb{Y}$ es completo entonces $\lim_n U_n x = y$. Definamos el operador U tal que, para cada $x \in \mathbb{X}$, $Ux = \lim_n U_n x$. Tomando límite $n \to \infty$ en

$$U_n(\alpha x + \beta y) = \alpha U_n x + \beta U_n y \quad \Rightarrow \quad U(\alpha x + \beta y) = \alpha U x + \beta U y,$$

luego U es lineal. Tomando el límite $m \to \infty$ en (4.4.4) tenemos

$$\|U_n x - Ux\| \le \varepsilon \|x\| \quad \Rightarrow \quad U - U_n \text{ es acotado.}$$

Como $U_n \in \mathcal{B}(\mathbb{X},\mathbb{Y})$ y $U - U_n \in \mathcal{B}(\mathbb{X},\mathbb{Y})$, su suma $U_n + U - U_n = U$ es acotada. Además, de la desigualdad anterior se sigue que, para todo $x \neq 0$,

$$\frac{\|(U_n - U)x\|}{\|x\|} \le \varepsilon \;\Rightarrow\; \sup_{\|x\| \neq 0} \frac{\|(U_n - U)x\|}{\|x\|} \le \varepsilon \quad \Rightarrow \quad \|U_n - U\| \to 0.$$

Luego, $U_n \to U \in \mathcal{B}(\mathbb{X}, \mathbb{Y})$, de donde se sigue que el espacio de las aplicaciones lineales acotadas $\mathcal{B}(\mathbb{X}, \mathbb{Y})$ cuando $\mathbb{Y}$ es de Banach, es completo (¿por qué?) y, por tanto, de Banach. ∎

4.5. El Teorema de Banach-Steinhaus

Para terminar este capítulo dedicado a los espacios normados vamos a demostrar un resultado muy útil para una sucesión de operadores lineales y acotados: el Teorema de Banach-Steinhaus.

Teorema 4.5.1 (Banach-Steinhaus) *Sea $(T_n)_n$ una sucesión*[a] *de operadores lineales acotados $T_n : \mathbb{X} \mapsto \mathbb{Y}$ de un espacio de Banach $\mathbb{X}$ a otro normado cualquiera $\mathbb{Y}$ tales que la sucesión $(\|T_n x\|)_n$ es acotada para cada $x \in \mathbb{X}$, es decir, para cada $x \in \mathbb{X}$ existe un $c_x \in \mathbb{R}$ tal que, para todo $n = 1, 2, \dots$ se tiene*

$$\|T_n x\| \leq c_x. \tag{4.5.1}$$

Entonces, la sucesión de normas $(\|T_n\|)_n$ es acotada, es decir, existe un $c \geq 0$ tal que, para todo $n = 1, 2, \dots$ se cumple

$$\|T_n\| \leq c. \tag{4.5.2}$$

[a] Se asume nuevamente que todos los T_n tienen el mismo dominio $\mathcal{D}(T_n)$.

Nota 4.5.2 *Este teorema se puede generalizar a una familia de operadores T_i, $i \in I$, no teniendo que ser el conjunto de índices I numerable. Se deja como ejercicio al lector que modifique la prueba para este caso.*

Demostración: Comenzaremos construyendo una familias de subconjuntos $M_k \subset \mathbb{X}$, $k = 1, 2, \dots$, de todos los $x \in \mathbb{X}$ tales que para todo $n \in \mathbb{N}$

$$\|T_n x\| \leq k.$$

Probemos que los M_k son cerrados. Para ello probemos que cualquiera sea $x \in \overline{M_k}$, entonces $x \in M_k$ (ver Proposición 3.5.4). Como $x \in \overline{M_k}$ existe una sucesión $(x_m)_m \in M_k$ tal que $x_m \to x$, pero entonces, para todo m y todo n, como $x_m \in M_k$, usando la continuidad de la norma y que T_n son acotados, luego continuos, obtenemos

$$\|T_n x_m\| \leq k \quad \Rightarrow \quad m \to \infty, \quad \|T_n x\| \leq k \quad \Rightarrow \quad x \in M_k.$$

Por la condición (4.5.1) está claro que para cada $x \in \mathbb{X}$ existe un entero $k_x = [c_k] + 1$, tal que

$$\|T_n x\| \leq c_x \leq [c_x] + 1 = k_x,$$

luego cada $x \in \mathbb{X}$ pertenece a algún M_k, es decir,

$$\mathbb{X} = \bigcup_{k=1}^{\infty} M_k.$$

Pero entonces el Corolario del Teorema de Baire 3.6.2 ($\mathbb{X}$ es completo) nos asegura que al menos uno de los conjuntos M_k contiene un abierto. Supongamos que M_{k_0} es dicho conjunto, i.e., entonces existe una bola $B(x_0, r) \subset M_{k_0}$ tal que para todo $z \in B(x_0, r)$ se tiene $\|T_n z\| \leq k_0$. Elijamos ahora z

$$z = x_0 + \gamma x, \quad \gamma > 0,$$

de forma que $z \in B(x_0, r)$ para todo $x \in \mathbb{X}$, $x \neq 0$. Para ello basta que $\gamma\|x\| < r$, para todo x. Tomemos por ejemplo $\gamma = r/(2\|x\|)$. Entonces, para todo $n = 1, 2, \ldots$, $\|T_n z\| \leq k_0$. Luego, para todo $x \in \mathbb{X}$, y todo $n = 1, 2, \ldots$, tenemos

$$\|T_n x\| = \left\| T_n \left(\frac{z - x_0}{\gamma} \right) \right\| = \frac{1}{\gamma} \|T_n z - T_n x_0\| \leq \frac{1}{\gamma} (\|T_n z\| + \|T_n x_0\|) \leq \frac{2k_0}{\gamma}.$$

Usando que $\gamma = r/(2\|x\|)$ tenemos

$$\frac{\|T_n x\|}{\|x\|} \leq \frac{4k_0}{r} \quad \Rightarrow \quad \sup_{\|x\|=1} \frac{\|T_n x\|}{\|x\|} \leq \frac{4k_0}{r} = c,$$

de donde se sigue que, para todo $n = 1, 2, \ldots$, $\|T_n\| < c$, como se quería probar. ∎

Nótese que la hipótesis (4.5.1) y la consecuencia (4.5.2) del Teorema de Banach-Steinhaus se pueden cambiar por

$$\sup_{n \in \mathbb{N}} \|T_n x\| < \infty, \ \forall x \in \mathbb{X}, \ \text{y} \ \sup_{n \in \mathbb{N}} \|T_n\| < \infty,$$

respectivamente, lo que explica por qué a este teorema se le suela denominar *Principio de acotación uniforme*, pues se obtiene una cota uniforme para la sucesión de normas $\|T_n\|$ a partir de las cotas puntuales $\|T_n x\|$.

Así, una formulación equivalente del Teorema de Banach-Steinhaus es la siguiente:

Principio de acotación uniforme: Sea $(T_n)_n$ una sucesión de operadores lineales acotados $T_n : \mathbb{X} \mapsto \mathbb{Y}$ de un espacio de Banach $\mathbb{X}$ a otro normado cualquiera $\mathbb{Y}$ tales que

$$\sup_{n\in\mathbb{N}} \|T_n x\| < \infty, \ \forall x \in \mathbb{X} \ \Rightarrow \ \sup_{n\in\mathbb{N}} \|T_n\| < \infty. \tag{4.5.3}$$

Nótese que la prueba que U es acotado en el Ejercicio 4.4.20 se puede simplificar notablemente si usamos el Teorema de Banach-Steinhaus 4.5.1. En efecto, como existe el límite $U_n x$ cuando n tiende a infinito para todo $x \in \mathbb{X}$, entonces la sucesión $(U_n x)_n$ es acotada, i.e., existe $c_x \geq 0$ tal que $\|U_n x\| \leq c_x$, pero entonces existe un $c \geq 0$ tal que $\|U_n\| \leq c$. Usando entonces que $\|U_n x\| \leq \|U_n\|\|x\| \leq c\|x\|$ y tomando el límite $n \to \infty$, se tiene que $\|Ux\| \leq c\|x\|$, i.e., U es acotado.

Ejercicio 4.5.3 *Sea $(T_n)_n$ una sucesión de operadores lineales acotados $T_n : \mathbb{X} \mapsto \mathbb{Y}$ con $\mathbb{X}$ espacio de Banach. Prueba que si T_n es puntualmente convergente a un operador T (véase la Definición 3.5.34) entonces la sucesión $(T_n)_n$ está uniformemente acotada, i.e.,*

$$\exists c > 0 \text{ tal que } \forall n \in \mathbb{N}, \quad \|T_n\| < c, \tag{4.5.4}$$

y el operador T es lineal y acotado (continuo).

Efectivamente, al ser $\|T_n x\|$ convergente para cada x, es acotada, luego se tiene (4.5.1), por tanto podemos aplicar el Teorema de Banach-Steinhaus 4.5.1 de donde se sigue (4.5.4).

La linealidad sale de tomar límites $n \to \infty$ en

$$T_n(\alpha x + \beta y) = \alpha T_n x + \beta T_n y.$$

Por otro lado, como para todo n, $\|T_n\| < c$, se sigue que existe un $c > 0$ tal que

$$\forall x \in \mathbb{X}, \forall n \in \mathbb{N}, \quad \|T_n x\| \leq c\|x\|.$$

Tomando el límite cuando $n \to \infty$, y usando la continuidad de la norma, tenemos entonces que existe $c > 0$ tal que

$$\forall x \in \mathbb{X}, \quad \|Tx\| \leq c\|x\|,$$

luego T es acotado (continuo). ■

4.5.1. Aplicación a las series de Fourier

Una aplicación interesante del Teorema de Banach-Steinhaus tiene que que ver con las series de Fourier. Dada una función $f(x)$ periódica de periodo 2π definiremos la serie trigonométrica de Fourier por

$$Sf(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nx) + b_n \sin(nx), \tag{4.5.5}$$

donde los coeficientes vienen dados por las expresiones

$$a_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos(nx)dx, \quad b_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin(nx)dx, \tag{4.5.6}$$

$n = 0, 1, 2, \dots$, asumiendo que las integrales existen. Una pregunta natural es cuándo, para cada[2] $x \in [0, 2\pi)$, las sumas parciales de la serie de Fourier

$$S_n f(x) := \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos(kx) + b_k \sin(kx), \tag{4.5.7}$$

convergen, y, en ese caso, si $S_n f(x) \to f(x)$ en cada x, es decir, si hay convergencia puntual. Nosotros nos ocuparemos de la primera cuestión. Para la segunda el lector puede consultar, por ejemplo, [29, §18.2].

Por ejemplo, encontremos la serie de Fourier la función $f(x) = 1$ si $0 \le x < \pi$ y 0 si $\pi \le x < 2\pi$ (ver figura 4.2). Un cálculo directo nos da $a_0 = 1$,

$$a_n = 0, \quad b_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin nx\, dx = \frac{1}{\pi}\int_0^{\pi} \sin nx\, dx = \frac{1}{\pi}\left(\frac{1-(-1)^n}{n}\right).$$

Luego la serie de Fourier de f es

$$Sf(x) = \frac{1}{2} + \frac{2}{\pi}\sum_{k=0}^{\infty} \frac{\sin(2k+1)x}{2k+1}. \tag{4.5.8}$$

Usando el criterio de Abel-Dirichlet para series numéricas[3] se puede comprobar que la serie anterior converge en todo punto de $[0, 2\pi)$ (de hecho en todo $\mathbb{R}$), incluido el punto de discontinuidad $x = 1/2$ donde toma el valor $1/2$ (ver figura 4.2). Es decir, para la convergencia de la serie de Fou-

[2] Por la periodicidad de f, y $S_n f(x)$, eso equivale a preguntarnos por la convergencia en todo $\mathbb{R}$.

[3] Ver, por ejemplo, [29, Proposición 3 §16.2.3, pág. 376].

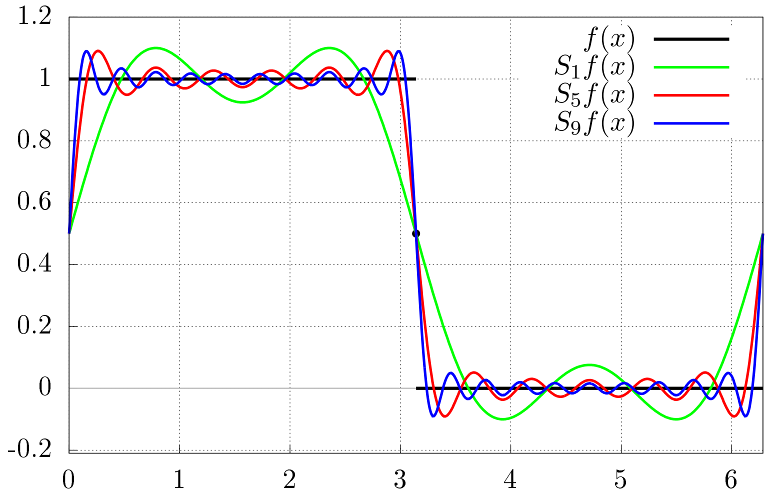


Figura 4.2: La función discontinua f y las sumas parciales de su serie de Fourier (4.5.8) para $n = 1, 5$ y 9.

rier, aunque esta no converja a la función en todo punto, no es necesaria la continuidad de f. Resulta que tampoco es suficiente. Es decir, se tiene el siguiente teorema:

Teorema 4.5.4 *Existen funciones 2π-periódicas continuas cuyas series de Fourier divergen en un punto x_0 dado.*

Para probar el teorema anterior conviene recordar algunas de las propiedades de las series de Fourier.

Sustituyendo las expresiones (4.5.6) de los coeficientes a_n y b_n de la serie de Fourier en (4.5.7) obtenemos

$$S_n f(x) = \frac{1}{\pi} \int_0^{2\pi} f(t) D_n(t-x) dt, \quad D_n(z) = \frac{1}{2} + \sum_{k=1}^{n} \cos(kz), \tag{4.5.9}$$

donde $D_n(z)$ es el núcleo de Dirichlet. Usando la identidad

$$2\cos(kz)\sin z/2 = \sin(k+1/2)z - \sin(k-1/2)z,$$

y sumando de $k = 1$ hasta n obtenemos la siguiente expresión alternativa para el núcleo de Dirichlet:

$$D_n(z) = \frac{\sin(n+1/2)z}{2\sin z/2}, \quad z \neq 0, \quad D_n(0) = n + \frac{1}{2}. \tag{4.5.10}$$

Nótese que $|D_n(z)| \leq n + 1/2$.

Definamos el espacio $\widetilde{C}$ de las funciones continuas en $[0, 2\pi]$ tales que $f(0) = f(2\pi)$ dotado con la norma del supremo: $\|f\| = \max_{x\in[0,2\pi]} |f(x)|$. Dicho espacio es un subespacio cerrado en el espacio $C^{\infty}_{[0,2\pi]}$ (¿por qué?), luego es completo (¿por qué?).

Sin pérdida de generalidad podemos elegir como punto de divergencia al punto $x = 0$.

Definamos la sucesión $(T_n)_n$ de funcionales lineales[4] sobre $\widetilde{C}$

$$T_n f = \frac{1}{\pi}\int_0^{2\pi} f(t)D_n(t)dt.$$

Nótese que $T_n f = S_n f(0)$. Además

$$|T_n f| \leq \frac{1}{\pi}\int_0^{2\pi} |f(t)||D_n(t)|dt \leq \max_{x\in[0,2\pi]} |f(t)| \underbrace{\left(\frac{1}{\pi}\int_0^{2\pi} |D_n(t)|dt\right)}_{=l_n}.$$

Luego, $\|T_n f\| \leq l_n\|f\|$, luego T_n es acotado para cada n y tomando el supremo para todas las f con $\|f\| = 1$, obtenemos $\|T_n\| \leq l_n$.

Demostremos ahora que $\|T_n\| = l_n$. Para ello vamos a encontrar una sucesión $(f_k)_k$ de funciones continuas en $[0, 2\pi]$ con norma 1 tal que para todo n, $T_n f_k \to l_n$ cuando $k \to \infty$, de donde, por la continuidad de la norma (en este caso el valor absoluto) se deduce que $\|T_n\| = l_n$.

Para ello notemos que, para cada $n \in \mathbb{N}$, $D_n(x) = 0$ en los puntos $a_i = \frac{2i\pi}{2n+1}$, $i = 1, 2, \ldots, 2n$. Además, D_n es no negativa en los intervalos $[0, a_1]$, $[a_2, a_3]$, $\ldots$, $[a_{2n-2}, a_{2n-1}]$, $[a_{2n}, 2\pi]$, cuya unión llamaremos Σ_+ y no positiva en $[a_1, a_2]$, $[a_3, a_4]$, $\ldots$, $[a_{2n-1}, a_{2n}]$, cuya unión llamaremos Σ_-. Para cada $n \in \mathbb{N}$ definamos la función g_n igual a 1 en Σ_+ e igual a -1 en Σ_- (ver la gráfica de la izquierda en la figura 4.3). Nótese que $g_n(t)D_n(t) = |D_n(t)|$ y que el número de intervalos donde g_n es negativo es exactamente igual a n.

Vamos a construir para cada $n \in \mathbb{N}$ una sucesión de funciones continuas en $(f_k)_k$ en $[0, 2\pi]$ que valga 1 en $[0, a_1 - \delta_k]$ y $[a_{2n} + \delta_k, 2\pi]$ y que sea lineal entre $a_i - \delta_k$ y $a_i + \delta_k$ para cada uno de los intervalos donde g_n es

[4]Un funcional lineal T no es más que una aplicación lineal $T : \mathbb{X} \mapsto \mathbb{R}$, de un espacio normado $\mathbb{X}$ en $\mathbb{R}$ (o $\mathbb{C}$) donde $\mathbb{R}$ se entiende como un espacio normado con la norma del valor absoluto.

negativo (ver detalle en la gráfica de la derecha de la figura 4.3) donde δ_k es lo suficientemente pequeño y lo escogeremos convenientemente como veremos más adelante. Es *obvio* que $|f_k(x)| \leq |g_n(x)| \leq 1$, luego $\|f_k\| = 1$.

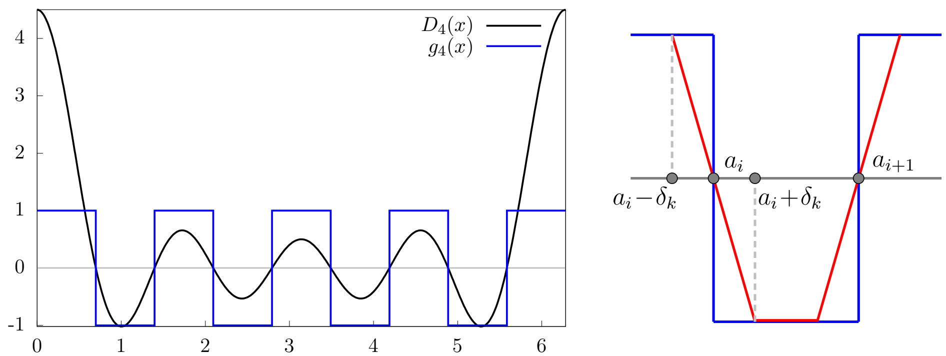


Figura 4.3: Las funciones $D_4(x)$ y $g_4(x)$ (izquierda) y detalle de la función g_n (azul) y de las f_k (rojo) (derecha).

Además, se puede comprobar[5] que para cada $n \in \mathbb{N}$ las funciones f_k y g_n así construidas son tales que

$$\int_0^{2\pi} |f_k(t) - g_n(t)| dt = 2n\delta_k.$$

Ahora bien, para cada $n \in \mathbb{N}$, como $l_n = \frac{1}{\pi}\int_0^{2\pi} |D_n(t)|dt = \frac{1}{\pi}\int_0^{2\pi} g_n(t)D_n(t)dt$, tenemos

$$\begin{aligned}
|T_n f_k - l_n| &\leq \frac{1}{\pi}\int_0^{2\pi} |f_k(t) - g(t)||D_n(t)|dt \\
&\leq \left(\max_{x\in[0,2\pi]} |D_n(x)|\right) \frac{1}{\pi}\int_0^{2\pi} |f_k(t) - g(t)|dt \leq \frac{(2n+1)n}{\pi}\delta_k,
\end{aligned}$$

donde hemos usado que $\max |D_n(x)| = n + 1/2$ para todo $x \in [0, 2\pi]$. Luego si, para cada $n \in \mathbb{N}$ tomamos en la definición de nuestra función f_k, $\delta_k \leq \frac{\pi}{n(2n+1)k}$ tenemos

$$|T_n f_k - l_n| \leq \frac{1}{k} \to 0, \quad k \to \infty \quad \Rightarrow \quad T_n f_k \to l_n,$$

[5]Dicha integral es el área de la región definida por la diferencia de ambas funciones que como se ve en la gráfica derecha de la figura 4.3 es igual a $2\delta_k$ en cada uno de los n intervalos $(a_i - \delta_k, a_{i+1} + \delta_k)$, $i = 1, 3, 5, \ldots, 2n-1$.

y como $\|f_k\| = 1$, entonces $\|T_n\| = l_n$.

Probemos ahora que la sucesión $(l_n)_n$ es no acotada. Comenzamos notando que

$$l_n = \frac{1}{\pi}\int_0^{2\pi} |D_n(t)|dt = \frac{1}{2\pi}\int_0^{2\pi} \frac{|\sin(n+\frac{1}{2})t|}{|\sin\frac{t}{2}|}dt \geq \frac{1}{\pi}\int_0^{2\pi} \frac{|\sin(n+\frac{1}{2})t|}{t}dt.$$

Haciendo el cambio $z = (n+1/2)t$ en la última integral obtenemos que

$$\begin{aligned} l_n &\geq \frac{1}{\pi}\int_0^{(2n+1)\pi} \frac{|\sin z|}{z}dz = \frac{1}{\pi}\sum_{m=0}^{2n}\int_{m\pi}^{(m+1)\pi} \frac{|\sin z|}{z}dz \\ &\geq \frac{1}{\pi}\sum_{m=0}^{2n}\int_{m\pi}^{(m+1)\pi} \frac{|\sin z|}{(m+1)\pi}dz = \frac{1}{\pi^2}\sum_{m=0}^{2n}\frac{1}{m+1}\int_{m\pi}^{(m+1)\pi} |\sin z|dz \\ &= \frac{2}{\pi^2}\sum_{m=0}^{2n}\frac{1}{m+1} \to \infty \text{ si } n \to \infty, \end{aligned}$$

pues la suma obtenida son las sumas parciales de la serie armónica $\sum_{n=1}^{\infty} 1/n$ que es divergente.

Por tanto, la sucesión de normas $\|T_n\|$ no está acotada, es decir, (4.5.2) no se cumple, por tanto, el Teorema de Banach-Steinhaus (véase el Problema 4.19) nos asegura que tiene que existir una función $f \in \widetilde{C}$ (recordemos que $\widetilde{C}$ es completo) tal que $\|T_n f\| \to \infty$ ($T_n f$ es no acotada para alguna f, o si no se cumpliría (4.5.1), véase el Problema 4.19), pero como $T_n f = S_n f(0)$, eso significa que tiene que existir una función f continua en $\widetilde{C}$ tal que su serie de Fourier en cero es divergente. ■

4.6. Problemas

Problema 4.1 Prueba que los espacios $\mathbb{R}^n$, $\mathbb{R}^{n\times m}$, $\mathbb{P}_n$ y $C([a,b])$ son espacios vectoriales.

Solución: Simplemente hay que comprobar que en cada caso se cumplen las propiedades de la Definición 4.1.1.

Problema 4.2 Prueba que para todos x, y de un espacio normado $\mathbb{X}$ se cumple la desigualdad $|\|x\| - \|y\|| \leq \|x - y\|$. Deduce de este resultado que la norma $\|\cdot\| : \mathbb{X} \mapsto [0, \infty)$ es una aplicación continua.

Solución: Tenemos que

$$\|x\| \leq \|x - y\| + \|y\|, \quad \|y\| \leq \|x - y\| + \|x\| \quad \Rightarrow$$

$$\|x\| - \|y\| \leq \|x - y\|, \quad \|y\| - \|x\| \leq \|x - y\| \quad \Rightarrow \quad |\|x\| - \|y\|| \leq \|x - y\|.$$

Esta claro que si $\|x_n - x\| \to 0$, entonces $\|x_n\| \to \|x\|$, de donde se sigue el resultado. ■

Problema 4.3 Prueba que el espacio ℓ^∞ de las sucesiones acotadas $(x_n)_n$ con la *norma* $\|x\| = \sup_{k \in \mathbb{N}} |x_k|$ es un espacio normado. Demuestra que en este espacio el subespacio Y de las sucesiones con un número finito de términos no nulos no es cerrado (y, por tanto, no es completo) en ℓ^∞. **Ayuda:** Usa la sucesión $s_n = (1, 1/2, \ldots, 1/n, 0, 0, \ldots) \in Y$ y la sucesión $x = (1, 1/2, \ldots, 1/n, 1/(n+1), \ldots) \in \overline{Y}$.

Solución: Probar que dicho espacio es normado es similar a la prueba de que ℓ^∞ es métrico (ver Ejemplo 3.1.11). Probemos que $Y \subset \ell^\infty$ no es cerrado. Sea la sucesión $s_n = (1, 1/2, \ldots, 1/n, 0, 0, \ldots) \in Y$. Está claro que para cada n, s_n coincide con la sucesión $x = (1, 1/2, \ldots, 1/n, 1/(n+1), \ldots)$ hasta el término n. Además

$$\|s_n - x\| = \sup_{k \in \mathbb{N}} |s_n^{(k)} - x^{(k)}| = \frac{1}{n+1} \to 0 \quad n \to \infty.$$

Pero $x \notin Y$ y, por tanto, Y no es cerrado pues no contiene a todos sus puntos límites. La incompletitud se sigue del Teorema 3.5.14. ■

Problema 4.4 Sean los espacios normados (no necesariamente de Banach) $\mathbb{X}_1$ y $\mathbb{X}_2$ y sean $\|\cdot\|_1$ y $\|\cdot\|_2$ sus correspondientes normas. Prueba que ambas normas son equivalentes si y solo si $\|x_n\|_1 \to 0$ implica $\|x_n\|_2 \to 0$ y viceversa.

Solución: Si las normas son equivalentes entonces se tiene que existen $a, b > 0$ tales que, para todo $n \in \mathbb{N}$

$$a\|x_n\|_1 \leq \|x_n\|_2 \leq b\|x_n\| \leq \tfrac{b}{a}\|x_n\|_2,$$

de donde se deduce, por el Teorema de las tres sucesiones, que si $\|x_n\|_1 \to 0$, entonces $\|x_n\|_2 \to 0$ y que si $\|x_n\|_2 \to 0$, entonces $\|x_n\|_1 \to 0$.

Veamos la otra implicación. Supongamos que si $\|x_n\|_1 \to 0$ implica $\|x_n\|_2 \to 0$ y viceversa pero que no existe $a > 0$ tal que $a\|x\|_1 \leq \|x\|_2$ par todo $x \in \mathbb{X}$. Entonces, para cada $n \in \mathbb{N}$ existe un $x_n \in \mathbb{X}$ tal que $\|x_n\|_1 = 1$ y $\|x_n\|_2 < 1/n$ ($\frac{1}{n}\|x_n\|_1 > \|x_n\|_2$), lo cual es imposible pues entonces $\|x_n\|_2 \to 0$ pero $\|x_n\|_1 \not\to 0$. Análogamente, si no existe $b > 0$ tal que $\|x\|_2 \leq b\|x\|_1$, entonces para cada n existirá un $x_n \in \mathbb{X}$ tal que $\|x_n\|_2 = 1$ y $\|x_n\|_1 < 1/n$ ($\|x_n\|_2 > n\|x_n\|_1$), luego $\|x_n\|_1 \to 0$ pero $\|x_n\|_2 \not\to 0$, lo cual es una contradicción. Nótese que en la prueba de este resultado no se necesita la completitud del espacio. ■

Problema 4.5 Prueba que los operadores de los ejemplos 4.4.2-4.4.6 son operadores lineales. Decide, usando distintas normas, si el operador del Ejemplo 4.4.6 es acotado.

Solución: Los ejemplos 4.4.2-4.4.5 son inmediatos. Veamos el operador del Ejemplo 4.4.6. En este caso tenemos

$$S(\alpha f(t) + \beta g(t)) = t(\alpha f(t) + \beta g(t)) = \alpha t f(t) + \beta t g(t) = \alpha S f(t) + \beta S g(t),$$

luego es lineal. Veamos si el operador multiplicación por $g(t)$ es acotado. Comencemos con la norma del Ejemplo 4.2.6. Así, sea $m = \text{máx}_{t\in[a,b]} |g(t)|$, entonces

$$\|Mf\| = \max_{t\in[a,b]} |g(t)f(t)| \leq m \max_{t\in[a,b]} |f(t)| = c\|f(t)\|,$$

luego M es acotado.

En el caso de la norma del Ejemplo 4.2.5 tendremos

$$\|Mf\| = \left(\int_a^b |g(x)f(x)|^p dx\right)^{1/p} \leq m\left(\int_a^b |f(x)|^p dx\right)^{1/p} \leq m\|f(t)\|,$$

de donde se sigue que M es acotado. ■

Problema 4.6 Prueba que el operador $T : C^\infty_{[0,1]} \mapsto C^\infty_{[0,1]}$ con $y = Tx$ definido por

$$y(t) = \int_0^1 k(t,\tau)x(\tau)d\tau, \tag{4.6.1}$$

con $k(t,\tau)$ continua en el cuadrado $[0,1] \times [0,1]$ es lineal y acotado y que

$$\|T\| = K = \max_{t\in[0,1]} \int_0^1 |k(t,\tau)|d\tau.$$

Solución: Nótese que por Teorema de la página 56, el operador está bien definido. La linealidad es consecuencia de la linealidad de la integral. De la desigualdad

$$|y(t)| \le \int_0^1 |k(t,\tau)||x(\tau)|d\tau \le \|x\| \int_0^1 |k(t,\tau)|d\tau,$$

se sigue, tomando el supremo en $t \in [0,1]$, que $\|Tx\| \le K\|x\|$, con

$$K = \max_{t\in[0,1]} \int_0^1 |k(t,\tau)|d\tau.$$

Luego, T es acotado y $\|T\| \le K$. Definamos ahora la función

$$u_n(t) = \begin{cases} -1 & t \le -1/n, \\ nt, & -1/n \le t \le 1/n \\ 1, & t \ge 1/n. \end{cases}$$

Sea la sucesión $x_n(t,\tau) = u_n(k(t,\tau))$. Nótese que $\|x_n\| = 1$, y que

$$k(t,\tau)u_n(k(t,\tau)) \ge 0, \text{ e igual a } |k(t,\tau)| \text{ cuando } |k(t,\tau)| \ge 1/n.$$

Denotemos por I_n el conjunto de puntos (subintervalos de $[0,1]$) donde se cumple la desigualdad anterior. Se tiene entonces

$$|Tx_n(t,\tau)| = \int_0^1 k(t,\tau)x_n(t,\tau)d\tau = \left[\int_{I_n} + \int_{[0,1]\setminus I_n}\right] \ge \int_{I_n} k(t,\tau)x_n(t,\tau)d\tau.$$

Pero como

$$\int_{I_n} k(t,\tau)x_n(t,\tau)d\tau = \int_{I_n} |k(t,\tau)|d\tau = \int_0^1 |k(t,\tau)|d\tau - \int_{[0,1]\setminus I_n} |k(t,\tau)|d\tau,$$

y en $[0,1] \setminus I_n$, $|k(t,\tau)| < 1/n$, entonces

$$|Tx_n(t,\tau)| \ge \int_{I_n} |k(t,\tau)|d\tau \ge \int_0^1 |k(t,\tau)|d\tau - \frac{1}{n}.$$

Tomando supremos en n obtenemos[6]

$$\sup_n |Tx_n(t,\tau)| \ge \int_0^1 |k(t,\tau)|d\tau.$$

[6]Ello se debe a la siguiente propiedad del supremo: si para todo $x \in M \subset \mathbb{R}$ y para todo $\varepsilon > 0$, $x \ge a - \varepsilon$, entonces $\sup_{x\in M} x \ge a$. En efecto, si no fuese cierto entonces, si $\sup_{x\in M} x < a$, entonces para todo $x \in M$, $x < a$ y, por tanto, existirá un $\varepsilon > 0$ tal que $x + \varepsilon < a$ (¿por qué?), lo cual es una contradicción.

Tomando en la desigualdad anterior máximo en t tenemos

$$\max_{t\in[0,1]}\left[\sup_n |Tx_n(t,\tau)|\right] \geq \max_{t\in[0,1]}\int_0^1 |k(t,\tau)|d\tau = K.$$

Pero, como $\|x_n(t,\tau)\| = 1$, entonces[7]

$$\|T\| = \sup_{\|x\|=1}\|Tx\| \geq \max_{t\in[0,1]}\left[\sup_n |Tx_n(t,\tau)|\right] \geq \max_{t\in[0,1]}\int_0^1 |k(t,\tau)|d\tau = K,$$

de donde se sigue que $\|T\| = K$.

Nótese que como corolario de lo anterior se tiene que el funcional $f: C^\infty_{[0,1]} \mapsto \mathbb{R}$ con $f(x) = \int_0^1 k(\tau)x(\tau)d\tau$, con $k(\tau)$ continua en $[0,1]$ es acotado y $\|f\| = \int_0^1 |k(\tau)|d\tau$. ■

Problema 4.7 Prueba que el operador $T: L^1_{[0,1]} \mapsto L^1_{[0,1]}$ con $y = Tx$ definido por (4.6.1) es lineal y acotado y que

$$\|T\| = \max_{\tau\in[0,1]}\int_0^1 |k(t,\tau)|dt,$$

donde por $L^1_{[0,1]}$ denotamos el completamiento del conjunto de las funciones integrables en $[0,1]$, i.e., $x \in L^1_{[0,1]}$ si $\int_0^1 |x(t)|dt < +\infty$.

Solución: Que el operador está bien definido es consecuencia del Teorema de la página 56. Para la linealidad véase el Problema 4.6. Para la acotación usamos que

$$\begin{aligned}\|Tx\| &= \int_0^1 |y(t)|dt = \int_0^1 \left|\int_0^1 k(t,\tau)x(\tau)d\tau\right| dt \leq \int_0^1 \left(\int_0^1 |k(t,\tau)||x(\tau)|d\tau\right) dt \\ &= \int_0^1 \left(\int_0^1 |k(t,\tau)|dt\right)|x(\tau)|d\tau \leq \max_{\tau\in[0,1]}\left(\int_0^1 |k(t,\tau)|dt\right)\|x\| = K\|x\| \ \Rightarrow\end{aligned}$$

$\|T\| \leq K$. Sea $\tau_0 \in [0,1]$ el valor donde se alcanza el máximo[8] anterior, $K = \int_0^1 |k(t,\tau_0)|dt$. Como $k(t,\tau)$ es continua en un cerrado y acotado, entonces es uniformemente continua en dicho conjunto (ver el Teorema de Heine 1.3.5). Luego, para todo $\varepsilon > 0$ existe un $\delta > 0$ tal que, si $|t-t'| < \delta$ y $|\tau-\tau'| < \delta$, $|k(t,\tau) - k(t',\tau')| < \varepsilon$. Elijamos ahora un intervalo $I \subset [0,1] = [\tau_1,\tau_2]$, tal que $\tau_1 \leq \tau_0 \leq \tau_2$ con $\tau_2 - \tau_1 < \delta$ y definamos la función

$$\tilde{x}(\tau) = \begin{cases} \dfrac{1}{\tau_2-\tau_1}, & t \in [\tau_1,\tau_2], \\ 0, & \text{en otro caso,}\end{cases} \qquad \text{además } \|\tilde{x}\| = 1.$$

[7] Los x_n son un subconjunto de las x con norma 1.

[8] La función $g(\tau) = \int_0^1 |k(t,\tau)|dt$ es continua en $[0,1]$ (ver Teorema de la página 56).

Entonces, para la norma de T obtenemos

$$\begin{aligned}\|T\| = \sup_{\|x\|=1} \|Tx\| \geq \|T\tilde{x}\| &= \int_0^1 \left| \int_0^1 k(t,\tau)\tilde{x}(\tau)d\tau \right| dt \\ &= \frac{1}{\tau_2 - \tau_1} \int_0^1 \left| \int_{\tau_1}^{\tau_2} k(t,\tau)d\tau \right| dt.\end{aligned}$$

Pero

$$\begin{aligned}\left| \int_{\tau_1}^{\tau_2} k(t,\tau_0)d\tau \right| &\leq \left| \int_{\tau_1}^{\tau_2} k(t,\tau)d\tau \right| + \left| \int_{\tau_1}^{\tau_2} k(t,\tau_0) - k(t,\tau)d\tau \right| \\ &\leq \left| \int_{\tau_1}^{\tau_2} k(t,\tau)d\tau \right| + \int_{\tau_1}^{\tau_2} |k(t,\tau_0) - k(t,\tau)|\, d\tau \leq \left| \int_{\tau_1}^{\tau_2} k(t,\tau)d\tau \right| + \varepsilon(\tau_2 - \tau_1),\end{aligned}$$

luego

$$\left| \int_{\tau_1}^{\tau_2} k(t,\tau)d\tau \right| \geq \left| \int_{\tau_1}^{\tau_2} k(t,\tau_0)d\tau \right| - \varepsilon = (\tau_2 - \tau_1)\big(|k(t,\tau_0)| - \varepsilon\big), \tag{4.6.2}$$

y, por tanto,

$$\frac{1}{\tau_2 - \tau_1} \int_0^1 \left| \int_{\tau_1}^{\tau_2} k(t,\tau)d\tau \right| dt \geq \int_0^1 |k(t,\tau_0)|dt - \varepsilon = K - \varepsilon.$$

Es decir, para todo $\varepsilon > 0$, $\|T\| \geq K - \varepsilon$ y, por tanto, $\|T\| \geq K$, de donde se deduce que $\|T\| = K$ como se quería demostrar. ■

Problema 4.8 Prueba que el operador $T : L^1_{[0,1]} \mapsto C^\infty_{[0,1]}$ con $y = Tx$ definido por (4.6.1) es lineal y acotado y que

$$\|T\| = \max_{\tau,t\in[0,1]} |k(t,\tau)|.$$

Solución: Que el operador está bien definido es consecuencia del Teorema de la página 56. La linealidad ya la hemos visto en el Problema 4.6. De la desigualdad

$$|Tx(t)| = \left| \int_0^1 k(t,\tau)x(\tau)d\tau \right| \leq \max_{\tau\in[0,1]} |k(t,\tau)| \int_0^1 |x(\tau)|d\tau = \max_{\tau\in[0,1]} |k(t,\tau)|\|x\|,$$

tomando el máximo en $t \in [0,1]$ obtenemos $\|Tx\| \leq K\|x\|$, donde $K = \max_{\tau,t\in[0,1]} |k(t,\tau)|$, i.e., T es acotado. Como $k(t,\tau)$ es continua en un cerrado y acotado, existen $t_0, \tau_0 \in [0,1]$ tales que $|k(t_0,\tau_0)| = K$.

Ahora construimos la misma función auxiliar del Problema 4.7 donde τ_0 es el valor encontrado antes. Así,

$$\|T\| = \sup_{\|x\|=1} \|Tx\| \geq \|T\tilde{x}\| = \max_{t\in[0,1]} \left| \int_0^1 k(t,\tau)\tilde{x}(\tau)d\tau \right| = \frac{1}{\tau_2 - \tau_1} \max_{t\in[0,1]} \left| \int_{\tau_1}^{\tau_2} k(t,\tau)d\tau \right|.$$

Usando (4.6.2) se sigue que, para todo $\varepsilon > 0$,

$$\frac{1}{\tau_2 - \tau_1} \left| \int_{\tau_1}^{\tau_2} k(t,\tau) d\tau \right| \geq |k(t,\tau_0)| - \varepsilon$$

que al tomar el máximo en $t \in [0,1]$ no conduce a $\|T\| \geq K - \varepsilon$. Luego, tomando $\varepsilon \to 0$ se tiene que $\|T\| \geq K$ y, por tanto, $\|T\| = K$. ■

Problema 4.9 Prueba que el operador $T : C^2_{[0,1]} \mapsto C^2_{[0,1]}$ con $y = Tx$ definido por (4.6.1) es lineal y acotado y que

$$\|T\| \leq \sqrt{\int_0^1 \int_0^1 |k(t,\tau)|^2 dt d\tau}.$$

En este caso no es sencillo encontrar la norma del operador en el caso general.

Solución: Que el operador está bien definido se sigue del Teorema de la página 56. Para la linealidad nuevamente véase el Problema 4.6. Finalmente, usando que

$$\|Tx\| = \sqrt{\int_0^1 \left| \int_0^1 k(t,\tau)x(\tau)d\tau \right|^2 dt} \leq \sqrt{\int_0^1 \left(\int_0^1 |k(t,\tau)x(\tau)|d\tau \right)^2 dt} \leq$$

$$\overset{\text{Hölder } p=2}{\leq} \sqrt{\int_0^1 \left(\int_0^1 |k(t,\tau)|^2 d\tau \right) \left(\int_0^1 |x(\tau)|^2 d\tau \right) dt} =$$

$$= \sqrt{\int_0^1 \int_0^1 |k(t,\tau)|^2 d\tau dt} \sqrt{\int_0^1 |x(\tau)|^2 d\tau} = \sqrt{\int_0^1 \int_0^1 |k(t,\tau)|^2 d\tau dt}\, \|x\|,$$

se deduce que T es acotado. ■

Problema 4.10 Demuestra que el operador T del Ejemplo 4.4.5 $T : \mathbb{R}^n \mapsto \mathbb{R}^n$, $y = Ax$, $A \in \mathbb{R}^{n\times n}$ es acotado en los siguientes casos:

1. $T : \mathbb{R}^n_\infty \mapsto \mathbb{R}^n_\infty$ (ver Ejemplo 3.1.8). Prueba además que
$$\|T\| = \max_{i=1,\ldots,n} \left(\sum_{k=1}^n |a_{ik}| \right).$$

2. $T : \mathbb{R}^n_1 \mapsto \mathbb{R}^n_1$ (ver Ejemplo 3.1.7). Prueba además que
$$\|T\| = \max_{k=1,\ldots,n} \left(\sum_{i=1}^n |a_{ik}| \right).$$

3. $T : \mathbb{R}_2^n \mapsto \mathbb{R}_2^n$ (ver Ejemplo 3.1.7). Prueba además que

$$\|T\| \leq \sqrt{\sum_{i=1}^{n}\sum_{k=1}^{n}|a_{ik}|^2}.$$

Solución: Tenemos que $y = Tx = Ax$. Comencemos probando 1.

$$\|Tx\| = \max_{i=1,\dots,n}\left|\sum_{k=1}^{n} a_{ik}x_k\right| \leq \left[\max_{i=1,\dots,n}\sum_{k=1}^{n}|a_{ik}|\right]\|x\| = K\|x\| \Rightarrow \|T\| \leq K.$$

Sea i_0 el índice donde se alcanza el valor de K, i.e., $K = \sum_{k=1}^{n}|a_{i_0k}|$, y sea $\tilde{x}$, tal que $\tilde{x}_k = \operatorname{signo}(a_{i_0k})$. Entonces

$$\|T\| = \max_{\|x\|=1}\|Tx\| \geq \|T\tilde{x}\| = \max_{i=1,\dots,n}\left|\sum_{k=1}^{n} a_{ik}\tilde{x}_k\right| \geq \left|\sum_{k=1}^{n} a_{i_0k}\tilde{x}_k\right| = \sum_{k=1}^{n}|a_{i_0k}| = K,$$

de donde se sigue el resultado.

2. En este caso tenemos

$$\begin{aligned}\|Tx\| &= \sum_{i=1}^{n}\left|\sum_{k=1}^{n} a_{ik}x_k\right| \leq \sum_{i=1}^{n}\sum_{k=1}^{n}|a_{ik}||x_k| = \sum_{k=1}^{n}\left[\sum_{i=1}^{n}|a_{ik}|\right]|x_k| \\ &\leq \sum_{k=1}^{n}\left[\max_{k=1,\dots,n}\sum_{i=1}^{n}|a_{ik}|\right]|x_k| = \left[\max_{k=1,\dots,n}\sum_{i=1}^{n}|a_{ik}|\right]\|x\| = K\|x\| \Rightarrow \|T\| \leq K.\end{aligned}$$

Sea k_0 el índice donde se alcanza el valor de K, i.e., $K = \sum_{i=1}^{n}|a_{ik_0}|$, y sea $\tilde{x} = \delta_{k,k_0}$, i.e., tal que $\tilde{x}_k$ vale cero para todo $k \neq k_0$ y uno si $k = k_0$. Entonces

$$\|T\| = \max_{\|x\|=1}\|Tx\| \geq \|T\tilde{x}\| = \sum_{i=1}^{n}\left|\sum_{k=1}^{n} a_{ik}\tilde{x}_k\right| = \sum_{i=1}^{n}|a_{ik_0}| = K.$$

3. En este caso, usando la desigualdad de Hölder (3.7.2), obtenemos

$$\|Tx\|^2 = \sum_{i=1}^{n}\left|\sum_{k=1}^{n} a_{ik}x_k\right|^2 \leq \sum_{i=1}^{n}\left|\sum_{k=1}^{n}|a_{ik}||x_k|\right|^2 \leq \sum_{i=1}^{n}\left[\sum_{k=1}^{n}|a_{ik}|^2\right]\left[\sum_{k=1}^{n}|x_k|^2\right],$$

luego $\|Tx\|^2 \leq \left[\sum_{i=1}^{n}\sum_{k=1}^{n}|a_{ik}|^2\right]\|x\|^2$, de donde se sigue el resultado. En este caso, a diferencia de los anteriores no es sencillo encontrar la norma del operador T. En el caso cuando A es una matriz simétrica, se puede probar que $\|T\| = |\lambda|$, siendo λ el mayor autovalor en valor absoluto de A. ■

Problema 4.11 Decide si el operador del Ejemplo 4.4.5 es acotado si escogemos la norma $\|x\| = \left(\sum_{k=1}^{n} |x_k|^p\right)^{1/p}$, $p \geq 1$. En caso de que sea acotado da una estimación de su norma.

Solución: Este problema es similar al apartado 3 del Problema 4.10.

Problema 4.12 Sean A, B dos operadores lineales y biyectivos $A : \mathbb{X} \mapsto \mathbb{Y}$, y $B : \mathbb{Y} \mapsto \mathbb{Z}$, $\mathbb{X}, \mathbb{Y}, \mathbb{Z}$ espacios vectoriales. Entonces existe el operador T inverso del operador $BA : \mathbb{X} \mapsto \mathbb{Z}$, $T = (BA)^{-1} : \mathbb{Z} \mapsto \mathbb{X}$ y $T = (BA)^{-1} = A^{-1}B^{-1}$.

Solución: En efecto, como A y B son biyectivos, ambos poseen inversos A^{-1} y B^{-1} y los inversos de operadores lineales son lineales (ver Teorema 4.4.9). La composición de operadores lineales también es lineal, por tanto, la aplicación $T = A^{-1}B^{-1} : \mathbb{Z} \to \mathbb{X}$ existe, está bien definida, es biyectiva y lineal. Además, para todo $z \in \mathbb{Z}$, $BA(A^{-1}B^{-1})(z) = z$, por tanto, $T = (BA)^{-1} = A^{-1}B^{-1}$. ■

Problema 4.13 Sean dos operadores A, B de $\mathbb{X} \mapsto \mathbb{X}$, $\mathbb{X}$ espacio normado, lineales y acotados. Prueba que $\|AB\| \leq \|A\|\|B\|$. Nótese que de lo anterior se sigue que, en particular, para todo $n \in \mathbb{N}$, $\|A^n\| \leq \|A\|^n$.

Solución: De las condiciones del problema se tiene que para todo $x \in \mathbb{X}$

$$\|Ax\| \leq \|A\|\|x\|, \quad \text{y} \quad \|Bx\| \leq \|B\|\|x\|.$$

Luego, para $y = Bx \in \mathbb{X}$,

$$\|Ay\| = \|A(Bx)\| \leq \|A\|\|Bx\| \leq \|A\|\|B\|\|x\|.$$

Por tanto, para todo $x \neq 0$,

$$\frac{\|ABx\|}{\|x\|} \leq \|A\|\|B\| \quad \Rightarrow \quad \|AB\| = \sup_{\|x\| \neq 0} \frac{\|ABx\|}{\|x\|} \leq \|A\|\|B\|.$$

■

Problema 4.14 Sea $\mathcal{B}(\mathbb{X})$ el espacio de todas las aplicaciones lineales y acotadas $A : \mathbb{X} \mapsto \mathbb{X}$, $\mathbb{X}$ espacio de Banach. Dadas dos aplicaciones lineales $A, B \in \mathcal{B}(\mathbb{X})$, definamos la aplicación $T : \mathcal{B}(\mathbb{X}) \mapsto \mathcal{B}(\mathbb{X})$ por $TX = AXB$, $X \in \mathcal{B}(\mathbb{X})$. Prueba que el operador T es lineal y acotado (respecto a la norma de operadores) y que $\|T\| \leq \|A\|\|B\|$.

Solución: Ante todo, recordemos que $\mathcal{B}(\mathbb{X})$ es el espacio de las aplicaciones lineales y acotadas $A : \mathbb{X} \mapsto \mathbb{X}$ con la suma definida por $(X + Y)x = (Xx) + (Yx)$ y el producto por un escalar por $(\lambda X)x = \lambda(Xx)$. Probemos que T es una aplicación lineal.

En primer lugar, tenemos que para todo $x \in \mathbb{X}$

$$
\begin{aligned}
(T(X+Y))x =&(A(X+Y)B)x = (A(X+Y))(Bx) = A((X+Y)(Bx)) \\
=&A[(X(Bx)) + (Y(Bx))] \overset{A \text{ lineal}}{=} A[(X(Bx)) + (Y(Bx))] \\
=&AXBx + AYBx = TXx + TYx \ \Rightarrow \ T(X+Y) = TX + TY.
\end{aligned}
$$

Por otro lado, para todo $x \in \mathbb{X}$ se tiene que

$$
\begin{aligned}
(T(\lambda X))x &= (A\lambda XB)x = ((A\lambda X)(Bx) = A((\lambda X)(Bx)) \\
&= A(\lambda(X(Bx))) \overset{A \text{ lineal}}{=} \lambda A(X(Bx)) \\
&= \lambda(AXB)x = \lambda Tx \ \Rightarrow \ T(\lambda X) = \lambda T(X).
\end{aligned}
$$

Que T es acotado es una consecuencia del Problema 4.13 pues

$$
\|TX\| \leq \|A\|\|B\|\|X\| \Rightarrow \frac{\|TX\|}{\|X\|} \leq \|A\|\|B\| \Rightarrow \sup_{\|X\|\neq 0} \frac{\|TX\|}{\|X\|} \leq \|A\|\|B\|,
$$

luego $\|T\| \leq \|A\|\|B\|$. ■

Problema 4.15 Sean $\mathbb{X}, \mathbb{Y}$ espacios normados y $T : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathbb{Y}$ un operador lineal y acotado. Prueba que si $(x_n)_n \in \mathcal{D}(T)$ y $x_n \to x \in \mathcal{D}(T)$, entonces $Tx_n \to Tx$. Prueba además que el espacio nulo $\mathcal{N}(T)$ es cerrado en $\mathcal{D}(T)$. **Ayuda:** Usa el Teorema 4.4.18 y la Proposición 3.5.4.

Solución: La primera parte se deduce del Teorema 4.4.18 pues como T es acotado es continuo. Para la segunda, tomemos $x \in \overline{\mathcal{N}(T)}$ cualquiera. Entonces existe una sucesión $(x_n)_n \in \mathcal{N}(T)$ tal que $x_n \to x$. Pero como $x_n \in \mathcal{N}(T)$, entonces $Tx_n = 0$, de donde se deduce, por la continuidad de T que $Tx = 0$, luego $x \in \mathcal{N}(T)$. Como x era arbitrario $\overline{\mathcal{N}(T)} = \mathcal{N}(T)$, luego $\mathcal{N}(T)$ es cerrado. ■

Problema 4.16 Sea un espacio normado $\mathbb{X}$ y sea $M \subset \mathbb{X}$ un subespacio cerrado de $\mathbb{X}$ distinto del propio $\mathbb{X}$. Prueba que existe un vector $y \in \mathbb{X}$ tal que, para todo $x \in M$, $\|y\| = 1$ y $\|y - x\| > 1/2$.

Solución: El resultado de este problema nos indica que si tenemos un subespacio cerrado M de un espacio normado siempre podemos encontrar elementos del espacio que están *lejos* del subespacio M. Elijamos $y_0 \in \mathbb{X} \setminus M$ y sea

$d = \inf_{x\in M} \|y_0 - x\|$. Probemos que $d > 0$. En efecto, si $d = 0$ entonces existiría una sucesión $(x_n)_n \in M$ tal que $x_n \to y_0$. Pero entonces $y_0 \in M$, pues M es cerrado, lo cual es una contradicción. Elijamos ahora $x_0 \in M$ tal que $\|y_0 - x_0\| < 2d$ (está claro que $\|y_0 - x_0\|$ ha de ser mayor que d) y definamos $y = \frac{y_0 - x_0}{\|y_0 - x_0\|}$. Es *obvio* que $\|y\| = 1$. Sea $x \in M$, entonces, como $x_0 \in M$, $z_0 = x_0 + x\|y_0 - x_0\| \in M$. Entonces

$$\|y - x\| = \frac{\|y_0 - z_0\|}{\|y_0 - x_0\|} \geq \frac{\inf_{z\in M} \|y_0 - z\|}{\|y_0 - x_0\|} = \frac{d}{\|y_0 - x_0\|} > \frac{d}{2d} = \frac{1}{2},$$

de donde se sigue el resultado. ■

Problema 4.17 Sea un espacio normado $\mathbb{X}$ de dimensión infinita. Prueba que la esfera unidad S, i.e. el conjunto de los $x \in \mathbb{X}$ tales que $\|x\| \leq 1$, no es un conjunto compacto[9]. Una consecuencia de este resultado es que si en un espacio normado $\mathbb{X}$ la esfera unidad es compacta, entonces $\mathbb{X}$ es de dimensión finita. **Ayuda:** Usando el resultado del Problema 4.16 construir una sucesión de elementos $(x_n)_n \in S$ tal que, para todos $n \neq m \in \mathbb{N}$, $\|x_n - x_m\| > 1/2$.

Solución: La idea es encontrar una sucesión de elementos $(x_n)_n \in S$ tal que, para todos $n \neq m \in \mathbb{N}$, $\|x_n - x_m\| > 1/2$. Está claro de que de dicha sucesión es imposible construir una subsucesión convergente.

Elegimos $x_1 \in S$ tal que $\|x_1\| = 1$. Sea $M_1 = \operatorname{span}(x_1)$. M_1 es un subespacio cerrado[10] de $\mathbb{X}$, luego el Problema 4.16 nos asegura que existe un $x_2 \in \mathbb{X}$ con $\|x_2\| = 1$, i.e., $x_2 \in S$, $x_2 \notin M_1$, tal que $\|x_2 - x\| > 1/2$, para todo $x \in M_1$, y, en particular $\|x_2 - x_1\| > 1/2$. Sea $M_2 = \operatorname{span}(x_1, x_2)$. Como M_2 es cerrado en $\mathbb{X}$, el Problema 4.16 nos asegura que existe un $x_3 \in S$, $\|x_3\| = 1$, $x_3 \notin M_2$, tal que $\|x_3 - x\| > 1/2$, para todo $x \in M_2$, y, en particular $\|x_3 - x_1\| > 1/2$, $\|x_3 - x_2\| > 1/2$, y así sucesivamente. Nótese que de la construcción anterior tenemos una sucesión infinita de términos distintos tales que para todo $n \in \mathbb{N}$, $\|x_n\| = 1$, i.e., $x_n \in S$. Es decir, tenemos una sucesión infinita de subconjuntos cerrados M_k, $k \in \mathbb{N}$ y una sucesión $(x_n)_n \in S$ tal que, para todos $n \neq m \in \mathbb{N}$, $\|x_n - x_m\| > 1/2$. Luego la esfera S, i.e., el conjunto de los $x \in \mathbb{X}$ tales que $\|x\| \leq 1$, no es un compacto (¿por qué?). El resultado que acabamos de probar fue demostrado por F. Riesz. Nótese que la demostración ha consistido en probar que en la esfera de radio unidad de un espacio de dimensión infinita existen sucesiones que no contienen sucesiones de Cauchy. Los espacios (conjuntos) tales

[9] En dimensión finita como la esfera unidad es cerrada y acotada, entonces es un compacto.

[10] Como ya vimos en el Ejemplo 4.3.4, los subespacios $\operatorname{span}(x_1, \ldots, x_n)$ son cerrados.

que toda sucesión contiene una sucesión de Cauchy se conocen como conjuntos precompactos, por lo que el resultado probado se puede reescribir diciendo: en dimensión infinita la esfera unidad no es precompacta. ■

Problema 4.18 Demuestra que no se puede prescindir de la condición de completitud de $\mathbb{X}$ en el Teorema de Banach-Steinhaus 4.5.1. Para ello considera el subespacio $\mathbb{X} \subset \ell^\infty$ de todas las sucesiones $x = (x_1, x_2, \dots, x_j, \dots)$ que tienen un número finito de términos no nulos, es decir, que $x_j = 0$ para todo $j > J \in \mathbb{N}$ (cada x tiene a lo más las primeras J coordenadas no nulas pudiendo depender la J de cada x). Sea el operador (funcional) $T_n x = f_n(x) = n x_n$. Probar que $\mathbb{X}$ no es completo (ver Problema 4.3), que $|f_n(x)|$ está acotado para cada n y todo $x \in \mathbb{X}$, y que $(\|T_n\|)_n$ es no acotada.

Solución: Por el Problema 4.3 sabemos que $\mathbb{X}$ no es cerrado, luego no es completo (ver Teorema 3.5.14). Ahora bien, por la definición de $\mathbb{X}$ está claro que para cada $n \in \mathbb{N}$, $\|T_n x\| = \sup_{x\in\mathbb{X}} |n x_n|$ es acotado pues, para cada n solo hay un número finito de coordenadas de x que son no nulas, luego existe un c_x tal que $\|T_n x\| \le c_x$. Por otro lado, está claro que para cada n podemos elegir un $\tilde{x} \in \mathbb{X}$ tal que $\|\tilde{x}\| = 1$ (por ejemplo, que solo tenga como coordenadas ceros y unos) y tal que $T_n \tilde{x} = n x_n = n$, entonces $\|T_n\| = \sup_{\|x\|=1} \|T_n x\| \ge \|T_n \tilde{x}\| = n$, luego la sucesión $(\|T_n\|)_n$ no es acotada, luego no se tiene (4.5.2). ■

Problema 4.19 Usando el Teorema de Banach-Steinhaus prueba que si $\mathbb{X}$ es un espacio de Banach, $\mathbb{Y}$ uno normado, $T_n \in \mathcal{B}(\mathbb{X}, \mathbb{Y})$ y $\sup_n \|T_n\| = \infty$, entonces existe un $x_0 \in \mathbb{X}$ tal que $\sup_n \|T_n x_0\| = \infty$. Dicho x_0 se suele denominar como punto o vector de *resonancia*.

Solución: Supongamos que es falso, entonces, para todo $x \in \mathbb{X}$, $\sup_n \|T_n x\| < \infty$, luego la sucesión $\|T_n x\|$ es acotada para cada x. Pero entonces por el Teorema de Banach-Steinhaus existe un $c \ge 0$ tal que, para todo n, $\|T_n\| < c$, lo cual contradice que $\sup_n \|T_n\| = \infty$. ■

Capítulo 5

Espacios de Hilbert

Un viejo matemático francés dijo: "*Una teoría matemática no se considerará completa hasta que la hayas dejado tan clara que puedas explicársela al primer hombre que encuentres en la calle*"

David Hilbert

De su conferencia en el ICM de 1900 en París

5.1. Espacios euclídeos y espacios de Hilbert

En adelante asumiremos que $\mathbb{E}$ es un espacio vectorial complejo, y por $\overline{z}$ denotaremos al complejo conjugado del número complejo z.

Definición 5.1.1 *Se dice que un espacio vectorial $\mathbb{E}$ sobre el cuerpo $\mathbb{C}$ es un espacio euclídeo o prehilbertiano, si dados dos elementos cualesquiera $x, y \in \mathbb{E}$ existe un número denominado* producto escalar *y que denotaremos por $\langle x, y\rangle$ tal que*

1. *Para todos $x, y \in \mathbb{E}$, $\langle x, y\rangle = \overline{\langle y, x\rangle}$.*
2. *Para todos $x, y, z \in \mathbb{E}$, $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$.*
3. *Para todos $x, y \in \mathbb{E}$ y $\lambda \in \mathbb{C}$, $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$*
4. *Para todo $x \in \mathbb{E}$, $x \neq 0$, $\langle x, x\rangle > 0$ y si $\langle x, x\rangle = 0$, entonces $x = 0$.*

La propiedad 1 se conoce como *propiedad de simetría* del producto escalar, mientras que la 2 y 3 implican la linealidad del mismo respecto al primer argumento (de la izquierda).

Ejercicio 5.1.2 *Prueba que, de la definición anterior, se tiene que*

1. *Para todos* $x, y, z \in \mathbb{E}$, $\langle x, y+z\rangle = \langle x, y\rangle + \langle x, z\rangle$.
2. *Para todos* $x, y \in \mathbb{E}$ *y* $\lambda \in \mathbb{C}$, $\langle x, \lambda y\rangle = \overline{\lambda}\langle x, y\rangle$.
3. *Para todo* $x \in \mathbb{E}$, $\langle x, 0\rangle = \langle 0, x\rangle = 0$.
4. *Si* $\langle x, z\rangle = \langle y, z\rangle$ *para todos los* $z \in \mathbb{E}$, *entonces* $x = y$.

Para probar 1 combinamos los puntos 1 y 2 de la definición. Para 2 usamos los puntos 1 y 3 de la definición. Para 3 basta elegir $\lambda = 0$ en el punto 3 de la definición. Finalmente, para probar 4 usamos los puntos 2 y 3 de la definición que nos da $\langle x-y, z\rangle = 0$ para todo $z \in \mathbb{E}$. Eligiendo ahora $z = x-y$ y usando el punto 4 de la definición de espacio euclídeo obtenemos el resultado. ■

Ejemplo 5.1.3 *El ejemplo más sencillo de espacio euclídeo es el espacio* $\mathbb{C}^n$ *con el producto escalar estándar: dados* $x = (x_1, \dots, x_n)$, *e* $y = (y_1, \dots, y_n)$

$$\langle x, y\rangle = \sum_{k=1}^{n} x_k \overline{y_k}.$$

Ejemplo 5.1.4 *Otro ejemplo es* ℓ^2, *el espacio de las sucesiones* $(x_n)_n$ *tales que* $\sum_{k=1}^{\infty} |x_k|^2 < \infty$, *donde dadas dos sucesiones* $x = (x_1, x_2, \dots, x_n, \dots)$, *e* $y = (y_1, y_2, \dots, y_n, \dots)$ *el producto escalar viene dado por*

$$\langle x, y\rangle = \sum_{k=1}^{\infty} x_k \overline{y_k}.$$

Nótese que de la desigualdad de Hölder 3.7.3 se sigue que dicho producto escalar esta bien definido.

Una prueba directa de dicha afirmación es como sigue. Para todos $a, b \geq 0$ se tiene $ab \leq (a^2+b^2)/2$. Tomemos $a = |x_i|/\|x\|$, $b = |y_i|/\|y\|$, donde la norma

$\|\cdot\|$ es la norma de ℓ^2. Entonces

$$\frac{|x_i y_i|}{\|x\|\|y\|} \leq \frac{1}{2}\left(\frac{|x_i|^2}{\|x\|^2} + \frac{|y_i|^2}{\|y\|^2}\right) \Rightarrow \sum_{i=1}^{\infty} \frac{|x_i y_i|}{\|x\|\|y\|} \leq \frac{1}{2}\sum_{i=1}^{\infty}\frac{|x_i|^2}{\|x\|^2} + \frac{1}{2}\sum_{i=1}^{\infty}\frac{|y_i|^2}{\|y\|^2} = 1,$$

de donde se deduce el resultado pues $|\sum_{k=1}^{\infty} x_k \overline{y_k}| \leq \sum_{k=1}^{\infty} |x_k y_k|$.

Ejemplo 5.1.5 *Nuestro tercer ejemplo es el espacio $C([a,b])$ de las funciones continuas en $[a,b]$ cerrado y acotado con el siguiente producto escalar*

$$\langle f, g\rangle = \int_a^b f(x)\overline{g(x)}dx. \tag{5.1.1}$$

Una propiedad importante de los espacios euclídeos es la desigualdad de Cauchy-Schwarz[1]

Teorema 5.1.6 (Desigualdad de Cauchy-Schwarz) *Sea $\mathbb{E}$ un espacio euclídeo. Entonces para todos $f, g \in \mathbb{E}$,*

$$|\langle f, g\rangle|^2 \leq \langle f, f\rangle\langle g, g\rangle. \tag{5.1.2}$$

Demostración: Para demostrarla basta usar que para todo $\lambda \in \mathbb{C}$ y $f, g \in \mathbb{E}$, $\langle \lambda f + g, \lambda f + g\rangle \geq 0$, o equivalentemente,

$$|\lambda|^2\langle f, f\rangle + \lambda\langle f, g\rangle + \overline{\lambda}\langle g, f\rangle + \langle g, g\rangle = |\lambda|^2\langle f, f\rangle + 2\Re(\lambda\langle f, g\rangle) + \langle g, g\rangle \geq 0.$$

Usando que $\Re(\lambda\langle f, g\rangle) \leq |\lambda\langle f, g\rangle)|$

$$0 \leq |\lambda|^2\langle f, f\rangle + 2\Re(\lambda\langle f, g\rangle) + \langle g, g\rangle \leq |\lambda|^2\langle f, f\rangle + 2|\lambda||\langle f, g\rangle| + \langle g, g\rangle.$$

De la desigualdad $|\lambda|^2\langle f, f\rangle + 2|\lambda||\langle f, g\rangle| + \langle g, g\rangle \geq 0$, se sigue que el discriminante de la ecuación cuadrática en $|\lambda|$,

$$|\lambda|^2\langle f, f\rangle + 2|\lambda||\langle f, g\rangle| + \langle g, g\rangle = 0,$$

ha de ser negativo o nulo (¿por qué?). Luego $|\langle f, g\rangle|^2 - \langle f, f\rangle\langle g, g\rangle \leq 0$, de donde se deduce (5.1.2). ∎

[1]Para los casos de $\mathbb{C}^n$, ℓ^2 y $C([a,b])$ discutidos antes esta desigualdad no es más que la desigualdad de Hölder.

Teorema 5.1.7 *Todo espacio euclídeo $\mathbb{E}$ es normado si en él definimos la norma mediante la fórmula $\|f\| = \sqrt{\langle f, f\rangle}$. Además, $|\langle f, g\rangle| \le \|f\| \cdot \|g\|$.*

Demostración: Los dos primeros axiomas de la definición de norma 4.2.1 son inmediatos. Para probar el tercero nótese que

$$\begin{aligned}\|f+g\|^2 &= \langle f+g, f+g\rangle = \langle f, f\rangle + 2\Re(\langle f, g\rangle) + \langle g, g\rangle \\ &\le \langle f, f\rangle + 2|\langle f, g\rangle| + \langle g, g\rangle \le \langle f, f\rangle + 2\sqrt{\langle f, f\rangle\langle g, g\rangle} + \langle g, g\rangle \\ &= (\sqrt{\langle f, f\rangle} + \sqrt{\langle g, g\rangle})^2,\end{aligned}$$

donde hemos usado la desigualdad de Cauchy-Schwarz (5.1.2). Tomando ahora raíces cuadradas se sigue el resultado. ■

De lo anterior se sigue que todo espacio euclídeo $\mathbb{E}$ es un espacio métrico con la métrica inducida por el producto escalar mediante la fórmula

$$\rho(x, y) = \|x - y\| = \sqrt{\langle x - y, x - y\rangle}.$$

Así, por ejemplo, en $\mathbb{C}^n$ tenemos que la norma inducida es

$$\|x\| = \sqrt{\sum_{k=1}^{n} |x_k|^2},$$

en ℓ^2,

$$\|x\| = \sqrt{\sum_{k=1}^{\infty} |x_k|^2},$$

y en $C([a, b])$ es

$$\|f\| = \sqrt{\int_a^b |f(x)|^2 dx}.$$

Además, en cada caso la métrica inducida es la correspondiente a los ejemplos 3.1.5, 3.1.12 ($p = 2$) y 3.1.10 ($p = 2$) del capítulo 3, respectivamente.

Ejercicio 5.1.8 *Prueba que, en la norma inducida por el producto escalar, las operaciones adición de vectores, multiplicación por un escalar y producto escalar de vectores son continuas, i.e., si $x_n \overset{n\to\infty}{\longrightarrow} x$, $y_n \overset{n\to\infty}{\longrightarrow} y$ y $\alpha_n \overset{n\to\infty}{\longrightarrow} \alpha$, entonces $x_n + y_n \overset{n\to\infty}{\longrightarrow} x + y$, $\alpha_n x_n \overset{n\to\infty}{\longrightarrow} \alpha x$, y $\langle x_n, y_n\rangle \overset{n\to\infty}{\longrightarrow} \langle x, y\rangle$.*

La primera parte es consecuencia de la continuidad de la norma. Para la segunda usamos que $\langle x, y\rangle = \langle x - x_n, y\rangle + \langle x_n, y - y_n\rangle + \langle x_n, y_n\rangle$, luego

$$\begin{aligned} |\langle x, y\rangle - \langle x_n, y_n\rangle| &\leq |\langle x, y_n - y\rangle| + |\langle x_n - x, y_n\rangle| \\ &\leq \|x\|\|y_n - y\| + \|x_n - x\|\|y_n\| \to 0, \end{aligned}$$

de donde, usando nuevamente de la continuidad de la norma, y que $(y_n)_n$ es acotada (¿por qué?) se sigue el resultado. ■

Es fácil probar (ver Ejercicio 5.1) que para todos $a, b \in \mathbb{E}$, la norma inducida por el producto escalar cumple con la *ley del paralelogramo*:

$$\|a + b\|^2 + \|a - b\|^2 = 2(\|a\|^2 + \|b\|^2). \tag{5.1.3}$$

De hecho se tiene la siguiente

Proposición 5.1.9 *Un espacio normado $\mathbb{X}$ real es euclídeo si y solo si para todos $x, y \in \mathbb{E}$, se cumple la ley del paralelogramo* (5.1.3).

Demostración: Sea la función $f(x, z) := \langle x, z\rangle = \frac{1}{4}(\|x+z\|^2 - \|x-z\|^2)$, $z \in \mathbb{X}$, y asumamos que se cumple la ley del paralelogramo (5.1.3). Probemos que se cumplen las cuatro propiedades de la Definición 5.1.1.

Está claro de la definición de f que $\langle x, z\rangle = \langle z, x\rangle$, luego se tiene el punto 1 de la Definición 5.1.1.

Calculemos ahora la cantidad

$$f(x + y, z) + f(x - y, z) =$$

$$\frac{1}{4}\left[\underbrace{\|x + y + z\|^2 + \|x - y + z\|^2}_{\text{en (5.1.3) } a = x + z,\ b = y.} - \underbrace{\left(\|x + y - z\|^2 + \|x - y - z\|^2\right)}_{\text{en (5.1.3) } a = x - z,\ b = y.}\right]$$
$$= \frac{1}{2}(\|x + z\|^2 - \|x - z\|^2) = 2f(x, z),$$

i.e.,

$$f(x+y, z) + f(x-y, z) = 2f(x, z) \Rightarrow \langle x+y, z\rangle + \langle x-y, z\rangle = 2\langle x, z\rangle. \tag{5.1.4}$$

Probemos ahora que $\langle \lambda x, z\rangle = \lambda\langle x, z\rangle$. Primero lo probamos para $\lambda = n \in \mathbb{N}$. Está claro que se cumple para $n = 1$. Para $n = 2$ se sigue de (5.1.4) aplicándola con $x = y$ y teniendo en cuenta que $f(0, z) = 0$. Asumamos que la igualdad $f(kx,z) = kf(x,z)$ es cierta para $k = 1, 2, \dots, n$ y demostremos que es cierta para $k = n+1$. En efecto,

$$\begin{aligned} f((n+1)x,z) &= f(nx+x,z) = 2f(nx,z) - f(nx-x,z) \\ &= 2nf(x,z) - (n-1)f(x,z) = (n+1)f(x,z). \end{aligned}$$

Probemos ahora que $\langle \lambda x, z\rangle = \lambda\langle x,z\rangle$ es cierto para $\lambda \in \mathbb{Q}$. Como $f(-x,z) = -f(x,z)$ por definición, basta probarlo para $\lambda = n/m$, $n,m \in \mathbb{N}$. Ahora bien, de la identidad $f(nx,z) = nf(x,z)$, $n \in \mathbb{N}$ implica

$$\|nx+z\|^2 - \|nx-z\|^2 = n(\|x+z\|^2 - \|x-z\|^2), \quad \forall x,z \in \mathbb{X}. \tag{5.1.5}$$

Luego

$$f\left(\frac{n}{m}x, z\right) = \frac{1}{4}\left(\left\|\frac{n}{m}x+z\right\|^2 - \left\|\frac{n}{m}x-z\right\|^2\right) = \frac{\|nx+mz\|^2 - \|nx-mz\|^2}{4m^2},$$

de donde usando (5.1.5) dos veces (la segunda intercambiando en (5.1.5) x y z) tenemos

$$\begin{aligned} f\left(\frac{n}{m}x, z\right) =& \frac{n}{4m^2}(\|x+mz\|^2 - \|x-mz\|^2) = \frac{n}{4m}(\|x+z\|^2 - \|x-z\|^2) \\ =& \frac{n}{m}f(x,z). \end{aligned}$$

Finalmente, usamos que la norma es una función continua y que $\mathbb{Q}$ es denso en $\mathbb{R}$, luego para todo número real λ, existe una sucesión de números racionales λ_n tal que $\lambda = \lim_{n\to\infty}\lambda_n$, entonces

$$f(\lambda x, z) = f\left(\lim_{n\to\infty}\lambda_n x, z\right) = \lim_{n\to\infty} f(\lambda_n x, z) = \lim_{n\to\infty}\lambda_n f(x,z) = \lambda f(x,z),$$

que era lo que se pretendía probar. Así, f cumple con el punto 3 de la Definición 5.1.1.

Probemos que $f(x,z)$ es lineal en su primera variable, i.e., $f(x+y,z) = f(x,z) + f(y,z)$. Para ello reescribimos (5.1.4) como $f(u+v,z) + f(u-v,z) = 2f(u,z)$ y hacemos $u+v = 2x$ y $u-v = 2y$, luego $u = x+y$ y, por tanto, $f(2x,z) + f(2y,z) = 2f(x+y,z)$, de donde, usando la linealidad de la multiplicación por un escalar, se tiene la linealidad de la suma de vectores, i.e., se tiene el punto 2 de la Definición 5.1.1.

Finalmente, por la propia definición de f y de las propiedades de la norma se sigue que $f(x,x) = \langle x,x\rangle = 0$ si y solo si $x = 0$, luego se cumple el punto 4 de la Definición 5.1.1. ∎

La proposición anterior nos dice que la ley del paralelogramo (5.1.3) caracteriza los espacios euclídeos. El caso complejo es algo más complicado de probar. La demostración original se debe a P. Jordan y J. von Neuman (*Annals of Math.* **36**, (1935) págs. 719–723)[2] y se basa en probar que la cantidad

$$\langle x, y\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\right)$$

define un producto escalar en $\mathbb{X}$ complejo, lo cual dejamos como ejercicio al lector.

Definición 5.1.10 *Un espacio euclídeo $\mathbb{E}$ completo[a] se denomina espacio de Hilbert y lo denotaremos por $\mathbb{H}$.*

[a]Es decir, un espacio $\mathbb{E}$ donde cualquier sucesión de Cauchy converge a un vector de $\mathbb{E}$ (en la métrica inducida por el producto escalar).

Definición 5.1.11 *Sea el sistema de vectores $(\phi_n)_n$ (finito o infinito) de un espacio euclídeo $\mathbb{E}$. Diremos que $(\phi_n)_{n=1}^\infty$ es un sistema ortogonal dos a dos si para todos n, m*

$$\langle \phi_n, \phi_m\rangle = \delta_{n,m}\|\phi_n\|^2. \tag{5.1.6}$$

Si además $\|\phi_n\| = 1$ para todo n, se dice que el sistema es ortonormal.

Por ejemplo, el sistema de los vectores canónicos de $\mathbb{C}^n$ $(e_k)_{k=1}^n$, definido por $e_k = \delta_{i,k}$, i.e.,

$$e_1 = (1,0,0,\dots,0,0),\ e_2 = (0,1,0,\dots,0,0),\ \dots\ e_n = (0,0,0,\dots,0,1).$$

es un sistema ortonormal. Análogamente, el sistema $(e_k)_{k=1}^\infty$, definido por $e_k = \delta_{i,k}$, i.e.,

$$e_1 = (1,0,0,0,\dots),\ e_2 = (0,1,0,0,\dots),\ e_3 = (0,0,1,0\dots),\dots$$

es un sistema ortonormal de ℓ^2. Finalmente, el sistema de funciones $\{1\} \cup \{\sin nx, \cos nx\}_{n=1}^\infty$ es un sistema ortogonal dos a dos de $C^2_{[-\pi,\pi])}$, i.e., del espacio euclídeo de las funciones continuas en $[-\pi,\pi]$, con el producto escalar

$$\langle f, g\rangle = \int_{-\pi}^{\pi} f(x)\overline{g(x)}dx.$$

[2]Ver, por ejemplo, el Teorema 11.1 pág. 244 de [22].

Ejercicio 5.1.12 *Prueba que si los vectores (no nulos) $\phi_1, \dots, \phi_n$ de un espacio euclídeo son ortogonales, entonces son linealmente independientes.*

Supongamos que son linealmente dependientes. Entonces existe al menos un escalar $\alpha_k \neq 0$, $k = 1, \dots, n$ tal que $\phi := \alpha_1\phi_1 + \cdots \alpha_n\phi_n = 0$. Usando la linealidad del producto escalar así como que los vectores $\phi_1, \dots, \phi_n$ son ortogonales obtenemos $0 = \langle \phi, \phi_k \rangle = \alpha_k$, $k = 1, \dots, n$, lo cual es una contradicción. ■

Teorema 5.1.13 (Gram-Schmidt) *En un espacio de Hilbert $\mathbb{H}$ de cualquier conjunto (finito o infinito) de vectores linealmente independiente se puede construir un conjunto de vectores ortonormales (ortogonales).*

Demostración: Para probar el teorema tomamos un sistema de vectores linealmente independiente $(\phi_n)_n$ de $\mathbb{H}$ cualquiera y definimos un nuevo sistema de vectores $(\psi_n)_n$ de la siguiente forma:

Como los $(\phi_n)_n$ son linealmente independiente, entonces todos son distintos de 0. Definamos $\psi_1 = \phi_1/\|\phi_1\|$, luego $\|\psi_1\| = 1$. A continuación definimos $\widetilde{\psi}_2 = \phi_2 + \alpha_{2,1}\psi_1$, y elegimos $\alpha_{2,1}$ de forma tal que $\widetilde{\psi}_2$ sea ortogonal al vector ψ_1, i.e. $\langle \widetilde{\psi}_2, \psi_1 \rangle = 0$, de donde se deduce que $\alpha_{2,1} = -\langle \phi_2, \psi_1 \rangle$. Además, como $\operatorname{span}(\psi_1, \widetilde{\psi}_2) = \operatorname{span}(\phi_1, \phi_2)$, entonces $\widetilde{\psi}_2 \neq 0$ (¿por qué?). Luego $\psi_2 = \widetilde{\psi}_2/\|\widetilde{\psi}_2\|$ es ortogonal a ψ_1 y $\|\psi_2\| = 1$. Nótese además que ψ_1 y ψ_2 son linealmente independientes.

Supongamos que de esta forma hemos construido una sucesión de $n-1$ vectores ortonormales dos a dos $\psi_1, \dots, \psi_{n-1}$, $n \geq 2$. Definamos el vector $\widetilde{\psi}_n$, $n \geq 2$,

$$\widetilde{\psi}_n = \phi_n + \sum_{k=1}^{n-1} \alpha_{n,k}\psi_k,$$

y elijamos los coeficientes $\alpha_{n,k}$, $k = 1, \dots, n-1$, de forma tal que $\widetilde{\psi}_n$ sea ortogonal a todos los vectores ψ_k, anteriores, i.e., $\langle \widetilde{\psi}_n, \psi_k \rangle = 0$, $k = 1, 2, \dots, n-1$. Usando la ortogonalidad es fácil comprobar que $\alpha_{n,k} = -\langle \phi_n, \psi_k \rangle$, $k = 1, 2, \dots, n-1$. Además, como $\operatorname{span}(\psi_1, \cdots, \psi_{n-1}, \widetilde{\psi}_n) = \operatorname{span}(\phi_1, \cdots, \phi_{n-1}, \phi_n)$, entonces $\widetilde{\psi}_n \neq 0$, por lo que podemos definir el vector $\psi_n = \widetilde{\psi}_n/\|\widetilde{\psi}_n\|$, que tiene norma uno y es ortogonal a todos los vectores anteriores ψ_k, $k = 1, 2, \dots, n-1$. Y así, sucesivamente.

Es decir, de cualquier conjunto de vectores linealmente independientes podemos construir un conjunto de vectores ortonormales. ■

El proceso anterior se conoce como *proceso de ortogonalización de Gram-Schmidt*. Nótese que de la prueba del teorema anterior se sigue que, para cada $n \geq 2$ ($\widetilde{\psi}_1 = \phi_1$),

$$\widetilde{\psi}_n = \phi_n + \sum_{k=1}^{n-1} \alpha_{n,k}\widetilde{\psi}_k \quad \Rightarrow \quad \phi_n = \widetilde{\psi}_n + \sum_{k=1}^{n-1} \beta_{n,k}\widetilde{\psi}_k,$$

de donde se deduce que, para todo $k = 1, \ldots n-1$,

$$\langle \widetilde{\psi}_n, \widetilde{\psi}_k \rangle = 0 \quad \Longrightarrow \quad \langle \widetilde{\psi}_n, \phi_k \rangle = 0. \tag{5.1.7}$$

La propiedad anterior nos conduce a un interesante resultado pero antes necesitamos introducir las matrices de Gram y sus determinantes.

Sea $G_n := G_n(\phi_1, \ldots, \phi_n)$ la matriz cuyas entradas son los valores $g_{i,j} = \langle \phi_i, \phi_j \rangle$, i.e.,

$$G_n = \begin{pmatrix} \langle \phi_1, \phi_1 \rangle & \langle \phi_1, \phi_2 \rangle & \cdots & \langle \phi_1, \phi_{n-1} \rangle & \langle \phi_1, \phi_n \rangle \\ \langle \phi_2, \phi_1 \rangle & \langle \phi_2, \phi_2 \rangle & \cdots & \langle \phi_2, \phi_{n-1} \rangle & \langle \phi_2, \phi_n \rangle \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \langle \phi_n, \phi_1 \rangle & \langle \phi_n, \phi_2 \rangle & \cdots & \langle \phi_n, \phi_{n-1} \rangle & \langle \phi_n, \phi_n \rangle \end{pmatrix}, \quad n \geq 1. \tag{5.1.8}$$

La matriz G_n anterior se denominan matriz de Gram de los vectores $\phi_1, \ldots, \phi_n$. Los determinantes de G_n, se conocen como determinantes de Gram de los vectores $\phi_1, \ldots, \phi_n$, i.e.,

$$\Delta_n = \det \begin{vmatrix} \langle \phi_1, \phi_1 \rangle & \langle \phi_1, \phi_2 \rangle & \cdots & \langle \phi_1, \phi_{n-1} \rangle & \langle \phi_1, \phi_n \rangle \\ \langle \phi_2, \phi_1 \rangle & \langle \phi_2, \phi_2 \rangle & \cdots & \langle \phi_2, \phi_{n-1} \rangle & \langle \phi_2, \phi_n \rangle \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \langle \phi_n, \phi_1 \rangle & \langle \phi_n, \phi_2 \rangle & \cdots & \langle \phi_n, \phi_{n-1} \rangle & \langle \phi_n, \phi_n \rangle \end{vmatrix}. \tag{5.1.9}$$

Nótese que $\overline{G_n} = G_n^T$, donde G_n^T es la matriz transpuesta de G_n, luego $\overline{\Delta_n} = \overline{\det G_n} = \det \overline{G_n} = \det G_n^T = \det G_n = \Delta_n$, de donde se sigue que Δ_n es real.

Como consecuencia del Teorema 5.1.13 tenemos el siguiente teorema:

Teorema 5.1.14 *Dado un conjunto cualquiera de vectores linealmente independientes ϕ_n, $n \geq 1$, se pueden construir los vectores ortonormales ψ_n,*

$n \geq 1$, *mediante la siguiente expresión explícita:*

$$\psi_n = \frac{\det \begin{vmatrix} \langle \phi_1, \phi_1 \rangle & \langle \phi_1, \phi_2 \rangle & \cdots & \langle \phi_1, \phi_{n-1} \rangle & \phi_1 \\ \langle \phi_2, \phi_1 \rangle & \langle \phi_2, \phi_2 \rangle & \cdots & \langle \phi_2, \phi_{n-1} \rangle & \phi_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \langle \phi_n, \phi_1 \rangle & \langle \phi_n, \phi_2 \rangle & \cdots & \langle \phi_n, \phi_{n-1} \rangle & \phi_n \end{vmatrix}}{\sqrt{\Delta_{n-1}\Delta_n}}, \qquad (5.1.10)$$

donde $\Delta_n := \Delta_n(\phi_1, \dots, \phi_n)$ *para* $n \geq 1$ *es el determinante de Gram* (5.1.9) *de los vectores* $\phi_1, \dots, \phi_n$*, donde se asume que* $\Delta_0 := 1$.

Demostración: En efecto, si denotamos por $\widetilde{\psi}_n$ al numerador de (5.1.10), tenemos que el producto escalar $\langle \widetilde{\psi}_n, \phi_k \rangle = 0$, $k = 1, 2, \dots, n-1$, ya que el determinante resultante de calcular dicho producto escalar tiene dos columnas iguales. Luego, usando (5.1.7), $\langle \widetilde{\psi}_n, \langle \widetilde{\psi}_k \rangle = 0$, $k = 1, 2, \dots, n-1$. Además, $\langle \widetilde{\psi}_n, \phi_n \rangle = \Delta_n$. Por otro lado, de (5.1.10) se sigue que

$$\widetilde{\psi}_n = \Delta_{n-1}\phi_n + \sum_{i=1}^{n-1} \alpha_{n,k}\phi_k.$$

Luego, usando que Δ_n es real tenemos

$$\langle \widetilde{\psi}_n, \widetilde{\psi}_n \rangle = \Delta_{n-1} \langle \phi_n, \widetilde{\psi}_n \rangle = \Delta_{n-1}\Delta_n, \qquad (5.1.11)$$

de donde se sigue (5.1.10). ■

Nótese que de (5.1.11) y teniendo en cuenta que $\Delta_1 = \langle \phi_1, \phi_1 \rangle > 0$ se sigue que $\Delta_n > 0$, para todo $n \geq 1$.

Ejercicio 5.1.15 *Prueba que un conjunto de vectores* $(\phi_i)_{i=1}^p$ *de un espacio euclídeo* $\mathbb{E}$*,* $i = 1, \dots, p$*, son linealmente independientes si y solo si su determinante de Gram* Δ_p *es distinto de cero, i.e.,*

$$\Delta_p = \det \begin{vmatrix} \langle \phi_1, \phi_1 \rangle & \langle \phi_1, \phi_2 \rangle & \cdots & \langle \phi_1, \phi_p \rangle \\ \langle \phi_2, \phi_1 \rangle & \langle \phi_2, \phi_2 \rangle & \cdots & \langle \phi_2, \phi_p \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \phi_p, \phi_1 \rangle & \langle \phi_p, \phi_2 \rangle & \cdots & \langle \phi_p, \phi_p \rangle \end{vmatrix} \neq 0.$$

Probaremos, por contraposición, que los vectores son dependientes si y solo si el determinante de Gram es cero.

Si los vectores son dependientes, existen unos escalares α_k, no todos nulos, tales que $\sum_{k=1}^{p}\alpha_k\phi_k = 0$, pero entonces

$$\sum_{k=1}^{p}\alpha_k\langle\phi_k,\phi_j\rangle = 0, \qquad j=1,\dots,p,$$

luego las filas del determinante de Gram son dependientes y, por tanto, el determinante es cero –por ejemplo, si $\alpha_1 \neq 0$, entonces escribiendo $\langle\phi_1,\phi_j\rangle = \sum_{k=2}^{p}(\alpha_k/\alpha_1)\langle\phi_k,\phi_j\rangle$, obtenemos que la primera fila es combinación lineal de las restantes–. Por el contrario, si el determinante de Gram es cero, entonces las filas tienen que ser linealmente dependientes, luego existirán los escalares α_k, $k=1,\dots,p$, no todos nulos, tales que, para $j=1,\dots,p$

$$\sum_{k=1}^{p}\alpha_k\langle\phi_k,\phi_j\rangle = 0 \quad \Rightarrow \left\langle\sum_{k=1}^{p}\alpha_k\phi_k,\phi_j\right\rangle = 0 \quad \Rightarrow \overline{\alpha_j}\left\langle\sum_{k=1}^{p}\alpha_k\phi_k,\phi_j\right\rangle = 0$$

y, por tanto,

$$0=\sum_{j=1}^{p}\overline{\alpha_j}\left\langle\sum_{k=1}^{p}\alpha_k\phi_k,\phi_j\right\rangle = \left\langle\sum_{k=1}^{p}\alpha_k\phi_k,\sum_{j=1}^{p}\alpha_j\phi_j\right\rangle = \left\|\sum_{k=1}^{p}\alpha_k\phi_k\right\|^2 = 0,$$

es decir, $\sum_{k=1}^{p}\alpha_k\phi_k = 0$, luego los vectores $\phi_1,\dots,\phi_p$ son linealmente dependientes. ■

El siguiente teorema nos será de gran utilidad más adelante.

Teorema 5.1.16 *Si el espacio euclídeo $\mathbb{E}$ es separable, entonces cualquier sistema ortogonal (ortonormal) de $\mathbb{E}$ es numerable.*

Demostración: Asumamos sin pérdida de generalidad que el sistema $(\psi_n)_n$ es ortonormal. Entonces $\|\psi_n-\psi_m\| = \sqrt{2}$ si $n\neq m$. Sea el conjunto de las bolas de radio $1/2$ y centro en cada ψ_n, $B(\psi_n,1/2)$. Estas bolas no se interceptan, luego en cada bola hay un único vector ψ_n de nuestro sistema ortonormal. Sea ahora $(\phi_k)_k$ un conjunto numerable denso en $\mathbb{E}$ (pues éste es separable). Entonces, en cada bola $B(\psi_n,1/2)$ habrá al menos un ϕ_k, luego el número de bolas y, por tanto, el de elementos ψ_n es numerable. ■

En los próximos apartados estudiaremos algunas propiedades de los espacios de Hilbert $\mathbb{H}$ separables, en especial veremos las denominadas series de Fourier en espacios de Hilbert separables. Nótese que, en virtud del teorema anterior, en estos espacios los sistemas ortogonales son numerables.

5.2. Ortogonalidad y Teorema de Riesz

Definición 5.2.1 *Dado un vector $x \in \mathbb{H}$ definiremos la serie de Fourier respecto al sistema ortonormal $(\phi_n)_{n=1}^{\infty}$ a la serie*

$$s := \sum_{n=1}^{\infty} c_n \phi_n, \tag{5.2.1}$$

donde los coeficientes vienen dados por las expresiones

$$c_n = \langle x, \phi_n \rangle, \quad \forall n \geq 1. \tag{5.2.2}$$

Teorema 5.2.2 *Sea H el subespacio vectorial de $\mathbb{H}$ generado por los vectores ortonormales $\phi_1, \phi_2 \ldots, \phi_n$, $n \in \mathbb{N}$, i.e., $H = \operatorname{span}(\phi_1, \phi_2 \ldots, \phi_n)$. Sea $x \in \mathbb{H}$. Entonces,*

$$\min_{q \in H} ||x - q||^2 = ||x||^2 - \sum_{k=1}^{n} |c_k|^2$$

donde c_k son los coeficientes definidos en (5.2.2) y se alcanza cuando q es la suma parcial de la serie de Fourier (5.2.1)

$$q = s_n := \sum_{k=1}^{n} c_k \phi_k.$$

Demostración: Sea $g_n = \sum_{k=1}^{n} a_k \phi_k$. Calculamos

$$\begin{aligned} \|x - g_n\|^2 &= \left\langle x - \sum_{k=1}^{n} a_k \phi_k, x - \sum_{m=1}^{n} a_m \phi_m \right\rangle \\ &= \|x\|^2 - \sum_{k=1}^{n} (\overline{a_k} c_k + a_k \overline{c_k}) + \sum_{k=1}^{n} \sum_{m=1}^{n} a_k \overline{a_m} \langle \phi_k, \phi_m \rangle. \end{aligned}$$

Usando que

$$\langle \phi_k, \phi_m \rangle = \delta_{k,m} \quad \text{y} \quad |a_k - c_k|^2 = |a_k|^2 + |c_k|^2 - \overline{a_k} c_k - a_k \overline{c_k},$$

obtenemos

$$\|x - g_n\|^2 = \|x\|^2 - \sum_{k=1}^{n} |c_k|^2 + \sum_{k=1}^{n} |a_k - c_k|^2.$$

Obviamente de la expresión anterior se deduce que la norma $\|x - g_n\|^2$ es es mínima si y solo si $a_k = c_k$ para todo $k \in \mathbb{N}$, i.e., $g_n = s_n$. ■

Como corolario de lo anterior tenemos que para todo $n \in \mathbb{N}$

$$\|x - s_n\|^2 = \|x\|^2 - \sum_{k=1}^{n} |c_k|^2 \geq 0 \quad \Rightarrow \quad \sum_{k=1}^{\infty} |c_k|^2 \leq \|x\|^2, \tag{5.2.3}$$

por lo que la serie $\sum_{k=1}^{\infty} |c_k|^2$ converge, i.e., $(c_n)_n \in \ell^2$ y, por tanto,

$$\lim_{n\to\infty} |c_n| = 0 \;\Rightarrow\; \lim_{n\to\infty} |\langle x, \phi_n\rangle| = 0 \;\Rightarrow\; \lim_{n\to\infty} \langle x, \phi_n\rangle = 0. \tag{5.2.4}$$

La desigualdad de la derecha en (5.2.3) se conoce como **desigualdad de Bessel**. Nótese además que de la igualdad de la izquierda en (5.2.3) se deduce que una condición necesaria y suficiente para que la serie de Fourier (5.2.1) converja a x (en norma) es que

$$\|x - s_n\| \to 0 \quad \Longleftrightarrow \quad \|x\|^2 = \sum_{k=1}^{\infty} |c_k|^2 = \sum_{k=1}^{\infty} |\langle x, \phi_n\rangle|^2.$$

Esta igualdad se conoce como **igualdad de Parseval** y es, en general, muy complicada de comprobar.

Definición 5.2.3 *Se dice que un sistema de vectores linealmente independientes $(\phi_n)_n$ es completo*[a] *en $\mathbb{X} \subset \mathbb{H}$ si para todo vector $x \in \mathbb{X} \subset \mathbb{H}$ y cualquiera sea $\varepsilon > 0$ existe una combinación lineal*

$$l_n = \sum_{k=1}^{n} \alpha_k \phi_k \quad \textit{tal que} \quad \|x - l_n\| < \varepsilon\,.$$

[a] También se suele denominar sistema total.

En otras palabras cualquier vector $x \in \mathbb{X} \subset \mathbb{H}$ se puede aproximar en norma tanto como se quiera mediante alguna combinación finita de vectores del sistema $(\phi_n)_n$. O, equivalentemente, sea $\operatorname{span}(\phi_1, \phi_2 \dots)$ el conjunto de todas las combinaciones lineales finitas de $(\phi_n)_n$, entonces $(\phi_n)_n$ es completo en $\mathbb{X} \subset \mathbb{H}$ si el conjunto $\operatorname{span}(\phi_1, \phi_2, \dots) \subset \mathbb{X}$ es denso en $\mathbb{X}$. La definición anterior 5.2.3 es equivalente a decir que $\mathbb{X}$ es el menor subespacio vectorial cerrado que contiene al conjunto $\phi_1, \phi_2, \dots$ ($(\phi_n)_n$ genera a todo $\mathbb{X}$).

Definición 5.2.4 *Un sistema ortogonal (ortonormal) completo de $\mathbb{X} \subset \mathbb{H}$ se denomina base ortogonal (ortonormal) de $\mathbb{X} \subset \mathbb{H}$.*

Por ejemplo, los sistemas $(e_k)_k$, $e_k = \delta_{i,k}$, definidos en la página 181 son bases ortogonales completas de $\mathbb{C}^n$ y ℓ^2, respectivamente.

Teorema 5.2.5 (De los sistemas completos) *Sea $\mathbb{H}$ un espacio de Hilbert y sea un sistema ortonormal de vectores $(\phi_n)_{n=1}^{\infty}$ de $\mathbb{H}$. Entonces las siguientes condiciones son equivalentes:*

1. *$(\phi_n)_n$ es completo en $\mathbb{H}$.*
2. *Para todo $x \in \mathbb{H}$, $x = \displaystyle\sum_{k=1}^{\infty} \langle x, \phi_k \rangle \phi_k$.*
3. *Para todo $x \in \mathbb{H}$, se cumple la igualdad de Parseval*
$$\|x\|^2 = \sum_{k=1}^{\infty} |\langle x, \phi_k \rangle|^2. \tag{5.2.5}$$
4. *Si $\langle x, \phi_k \rangle = 0$ para todo $k \in \mathbb{N}$ entonces $x = 0$.*

Demostración: 1) $\Leftrightarrow$ 2) Cualquiera sea $x \in \mathbb{H}$ construimos la serie de Fourier (5.2.1) y sean $s_n = \sum_{k=1}^{n} c_k \phi_k$, $c_k = \langle x, \phi_k \rangle$, sus sumas parciales.

1) $\Rightarrow$ 2) Usando 1) tenemos que para todo $\varepsilon > 0$ existe una combinación lineal l_N de vectores de $(\phi_n)_n$ tales que $\|x - l_N\| < \varepsilon$ (ver Definición 5.2.3). Pero el Teorema 5.2.2 nos dice que el menor de todos los valores de $\|x - l_N\|$ se alcanza cuando $l_N = s_N$, siendo s_N la N-ésima suma parcial de Fourier, i.e., $\|x - s_N\| \leq \|x - l_N\|$. Luego,

$$\forall \varepsilon > 0, \quad \exists N \in \mathbb{N} \text{ tal que } \|x - s_N\| < \varepsilon.$$

Pero para todo $n > N$, $\|x - s_n\| \leq \|x - s_N\|$. En efecto, de la igualdad de la izquierda en (5.2.3) tenemos

$$\begin{aligned} \|x - s_n\|^2 &= \|x\|^2 - \sum_{k=1}^{n} |c_k|^2 = \|x\|^2 - \sum_{k=1}^{N} |c_k|^2 - \sum_{k=N+1}^{n} |c_k|^2 \\ &= \|x - s_N\|^2 - \sum_{k=N+1}^{n} |c_k|^2. \end{aligned}$$

Entonces, para todo $n > N$, $\|x - s_n\| \leq \varepsilon$, es decir, $\lim_{n\to\infty} s_n = x$, de donde se sigue 2). Obviamente 2) implica 1) (¿por qué?).

2)⇒3) Tomando el límite $n \to \infty$ en la igualdad de la izquierda en (5.2.3), i.e., en $\|x - s_n\|^2 = \|x\|^2 - \sum_{k=1}^n |c_k|^2$, y usando 2) ($\lim_{n\to\infty} \|x - s_n\| = 0$), se sigue 3).

3)⇒4) Si para todo $k \in \mathbb{N}$, $\langle x, \phi_k \rangle = 0$, entonces de 3) se deduce que $\|x\| = 0$, luego $x = 0$.

4)⇒2) Recordemos que el sistema $(\phi_n)_n$ es ortonormal. Sea $y_n = \sum_{k=1}^n c_k \phi_k$, donde $c_k = \langle x, \phi_k \rangle$. De la desigualdad de Bessel (5.2.3) se sigue que la serie $\sum_{m=1}^\infty |c_m|^2$ es una serie convergente y, por tanto, la cantidad

$$\|y_n - y_{n+p}\|^2 = \sum_{k=n+1}^{n+p} |c_k|^2,$$

se puede hacer tan pequeña como se quiera, i.e., la sucesión y_n es de Cauchy, luego es convergente pues $\mathbb{H}$ es completo. Sea y su límite. Está claro que $y_n \to y = \sum_{k=1}^\infty \langle x, \phi_k \rangle \phi_k$ ($\|y - y_n\|^2 = \sum_{k=n+1}^\infty |c_k|^2 \to 0$). Probemos que $y = x$.

Usando la desigualdad de Cauchy-Schwarz (5.1.2)

$$|\langle y, \phi_k \rangle - \langle y_n, \phi_k \rangle| = |\langle y - y_n, \phi_k \rangle| \leq \|y_n - y\| \|\phi_k\| = \|y_n - y\|,$$

deducimos que $\lim_{n\to\infty} \langle y_n, \phi_k \rangle = \langle y, \phi_k \rangle$ para todo $k \in \mathbb{N}$. Pero

$$\langle y_n, \phi_k \rangle = \left\langle \sum_{j=1}^n c_j \phi_j, \phi_k \right\rangle = \sum_{j=1}^n \langle x, \phi_j \rangle \langle \phi_j, \phi_k \rangle = \langle x, \phi_k \rangle, \quad \forall k \leq n,$$

luego $\langle x - y_n, \phi_k \rangle = 0$ para todo $k \leq n$ de donde, tomando $n \to \infty$ deducimos que $\langle x - y, \phi_k \rangle = 0$ para todo $k \in \mathbb{N}$, luego por 4) $x - y = 0$. ■

Nota 5.2.6 *La equivalencia entre 1 y 2, así como las implicaciones* $2 \to 3 \to 4$*, son también ciertas para espacios euclídeos cualesquiera (no necesariamente completos). Solo* $4 \to 2$ *requiere la completitud del espacio.*

Del apartado 4 del Teorema 5.2.5 se sigue el siguiente corolario:

Corolario 5.2.7 *Sea el sistema ortonormal completo $(\phi_n)_n$ y sean $x, y \in \mathbb{X} \subset \mathbb{H}$ tales que $\langle x, \phi_k \rangle = \langle y, \phi_k \rangle$ para todo $k \in \mathbb{N}$, entonces $x = y$.*

En otras palabras, dos elementos de $\mathbb{H}$ con iguales coeficientes de Fourier son iguales, por tanto cualquier vector de $\mathbb{H}$ queda **biunívocamente** determinado por sus coeficientes de Fourier.

Definición 5.2.8 *Se dice que un sistema ortonormal $(\phi_n)_n$ es cerrado en un espacio euclídeo $\mathbb{E}$ si para todo vector $x \in \mathbb{E}$ se cumple la igualdad de Parseval* (5.2.5)*, i.e., $\sum_{k=1}^{\infty} |\langle x, \phi_k \rangle|^2 = \|x\|^2$.*

De la definición anterior y el Teorema 5.2.5 se tiene que un un sistema ortonormal $(\phi_n)_n$ es completo en un espacio de Hilbert separable $\mathbb{H}$ si y solo si $(\phi_n)_n$ es cerrado en $\mathbb{H}$. De hecho, de la primera ecuación de la fórmula (5.2.3), se deduce que la afirmación anterior también es válida en espacios euclídeos no necesariamente completos.

Teorema 5.2.9 *Todo espacio de Hilbert $\mathbb{H}$ separable tiene una base ortonormal.*

Demostración: Como $\mathbb{H}$ es separable, existe un conjunto numerable de vectores $(\phi_n)_n$ denso en $\mathbb{H}$. Como $(\phi_n)_n$ es denso en $\mathbb{H}$, entonces obviamente $\operatorname{span}(\phi_1, \phi_2, \dots)$ es denso en $\mathbb{H}$ (recordemos que $\operatorname{span}(\phi_1, \phi_2 \dots)$ es el conjunto de todas las combinaciones lineales finitas de los $(\phi_n)_n$). Si de dicho conjunto $(\phi_n)_n$ eliminamos aquellos vectores ϕ_k que se pueden obtener como combinación lineal de los anteriores ϕ_j, $j < k$, obtenemos un nuevo sistema sistema $(\phi_{n_k})_k$ $(n_1 < n_2 < \cdots)$ de vectores que claramente cumplen que $\operatorname{span}(\phi_1, \phi_2, \dots) = \operatorname{span}(\phi_{n_1}, \phi_{n_2}, \dots)$ y, por tanto, dicho conjunto también es denso en $\mathbb{H}$, luego $(\phi_{n_k})_k$ es un sistema completo de vectores linealmente independientes de $\mathbb{H}$ (¿por qué?). La base ortonormal $(\psi_k)_k$ se obtiene al aplicar a dicho sistema completo de vectores linealmente independientes $(\phi_{n_k})_k$ el proceso de ortogonalización de Gram-Schmidt pues $\operatorname{span}(\psi_1, \psi_2, \dots) = \operatorname{span}(\phi_{n_1}, \phi_{n_2}, \dots)$. ■

El teorema anterior 5.2.9 se puede generalizar a cualquier espacio euclídeo separable (no necesariamente completo).

Como hemos visto, dado un $x \in \mathbb{H}$ y un sistema ortonormal $(\phi_n)_n \in \mathbb{H}$ existen los coeficientes de Fourier de x en dicho sistema ortonormal y estos, como consecuencia de la desigualdad de Bessel (5.2.3), definen una sucesión de ℓ^2. El siguiente resultado es el recíproco de lo anterior.

Teorema 5.2.10 (Riesz-Fischer) *Sea $(\phi_n)_n$ un sistema ortonormal en un espacio de Hilbert $\mathbb{H}$ separable y sean los números c_1, c_2, ..., c_n, ... tales que*

$$\sum_{n=1}^{\infty} |c_n|^2 < +\infty.$$

Entonces, existe un elemento $x \in \mathbb{H}$ cuyos coeficientes de Fourier son precisamente los números c_1, c_2, ..., c_n, ..., y además

$$\sum_{n=1}^{\infty} |c_n|^2 = \|x\|^2, \qquad c_n = \langle x, \phi_n \rangle.$$

Demostración: Sea $x_n = \sum_{k=1}^{n} c_k\phi_k$. Como en la prueba del Teorema 5.2.5 tenemos $\|x_n - x_{n+p}\|^2 = \sum_{k=n+1}^{n+p} |c_k|^2$, luego la sucesión $(x_n)_n$ es de Cauchy y como $\mathbb{H}$ es completo entonces $x_n \stackrel{n\to\infty}{\longrightarrow} x \in \mathbb{H}$, $x = \sum_{k=1}^{\infty} c_k\phi_k$.

Probemos que si $x_n \to x$, entonces para todo $k \in \mathbb{N}$, $c_k = \langle x, \phi_k\rangle$. Por un lado tenemos que para todo $k \le n$, $\langle x_n, \phi_k\rangle = \langle \sum_{j=1}^{n} c_j\phi_j, \phi_k\rangle = c_k$. Luego, usando la continuidad del producto escalar, obtenemos que $c_k = \lim_{n\to\infty}\langle x_n, \phi_k\rangle = \langle x, \phi_k\rangle$, para todo $k \in \mathbb{N}$. Además, como $x_n \stackrel{n\to\infty}{\longrightarrow} x$, de la igualdad de la izquierda en (5.2.3) se obtiene que

$$\|x\|^2 - \sum_{k=1}^{n} |c_k|^2 = \|x - x_n\|^2 \stackrel{n\to\infty}{\longrightarrow} 0,$$

de donde se sigue el teorema. ■

De los teoremas 5.2.5, 5.2.9 y 5.2.10 se deduce que en un espacio de Hilbert separable $\mathbb{H}$ a cada $x \in \mathbb{H}$ le corresponde una serie de Fourier cuyos coeficientes de Fourier están biunívocamente determinados por x y que además se corresponden con un único vector de ℓ^2. Eso nos conduce a un resultado muy interesante pero antes de enunciarlo necesitamos la siguiente definición:

Definición 5.2.11 *Una aplicación U entre dos espacios de Hilbert $\mathbb{H}$ y $\mathbb{H}^*$ se denomina unitaria si es lineal, biyectiva y preserva el producto escalar, i.e.*[a], $\langle x, y\rangle = \langle Ux, Uy\rangle_* = \langle x^*, y^*\rangle_*$.

Dos espacios $\mathbb{H}$ y $\mathbb{H}^*$ son isomorfos si existe una aplicación unitaria $U : \mathbb{H} \mapsto \mathbb{H}^*$ tal que $x^* = Ux$, donde $x \in \mathbb{H}$ y $x^* \in \mathbb{H}^*$.

[a] Se entiende que $\langle \cdot, \cdot\rangle_*$ denota el producto escalar en $\mathbb{H}^*$ que no tiene por que ser el

mismo que en $\mathbb{H}$.

Teorema 5.2.12 (del isomorfismo) *Cualquier espacio de Hilbert separable $\mathbb{H}$ es isomorfo a $\mathbb{C}^n$ o a ℓ^2.*

Demostración: Como $\mathbb{H}$ es separable, en $\mathbb{H}$ existe una base ortonormal numerable (ver Teorema 5.2.9) que denotaremos por $(\phi_n)_{n\in I}$, donde I es un conjunto numerable (finito o infinito). Asumiremos que I es infinito, ($I = \mathbb{N}$) i.e., probaremos el caso cuando $\mathbb{H}$ es isomorfo a ℓ^2 (el caso finito es totalmente análogo).

Sea $x \in \mathbb{H}$ y sea la aplicación $U : \mathbb{H} \mapsto \ell^2$ definida por[3]

$$x^* = Ux = \sum_{k\in I} \langle x, \phi_k\rangle e_k = \sum_{k\in I} x_k e_k \in \ell^2, \quad x_k = \langle x, \phi_k\rangle,$$

donde $(e_n)_n$ denota la base ortonormal canónica de ℓ^2 que ya hemos visto antes. Como el producto escalar $\langle x, y\rangle$ es lineal respecto al elemento de la izquierda (i.e., x), esta claro que U es lineal (probarlo como ejercicio). Para probar que U es unitario usamos, por un lado, que

$$\langle x, y\rangle = \left\langle \sum_{k\in I} x_k\phi_k, \sum_{m\in I} y_m\phi_m \right\rangle = \sum_{k\in I} x_k\overline{y_k},$$

y, por el otro, que el producto escalar en ℓ^2 viene dado por $\langle x^*, y^*\rangle_{\ell^2} = \sum_{k\in I} x_k\overline{y_k}$, de donde se sigue que $\langle x, y\rangle = \langle x^*, y^*\rangle_{\ell^2}$.

Pero si U preserva el producto escalar, y por tanto, la norma (¿por qué?), entonces U es inyectiva. En efecto, si $Tx = Ty$, obtenemos

$$\|x - y\| = \|T(x - y)\|_* = \|Tx - Ty\|_* = 0 \quad \Rightarrow \quad x = y.$$

Probemos ahora que es sobreyectiva (y por tanto, biyectiva). Sea x^* cualquier sucesión $(x_n)_n$ de ℓ^2. Entonces, por el Teorema de Riesz-Fischer 5.2.10 existe un $x \in \mathbb{H}$ cuyos coeficientes de Fourier coinciden con dichos valores x_n y por el Corolario 5.2.7 dicho elemento es único. Es decir, la ecuación $x^* = Ux$ siempre tiene solución. ∎

5.2.1. El Teorema de la proyección ortogonal

[3]De la desigualdad de Bessel se sigue que el vector x^* así definido es de ℓ^2.

Definición 5.2.13 *Sea $M \subset \mathbb{H}$ un subconjunto del espacio de Hilbert $\mathbb{H}$. Denominaremos complemento ortogonal de M, y lo denotaremos por $M^\perp$, al conjunto*

$$M^\perp = \{x \in \mathbb{H};\ \langle x, y\rangle = 0,\ \ \forall y \in M\}.$$

De la definición anterior se sigue que si elegimos $M = \mathbb{H}$, entonces $\mathbb{H}^\perp = \{0\}$, pues el único elemento ortogonal a todos los vectores de $\mathbb{H}$ es el vector nulo. Nótese además que en la Definición 5.2.13 M no tiene que ser necesariamente un subespacio vectorial de $\mathbb{H}$. En el caso especial cuando M sea un subespacio vectorial de $\mathbb{H}$ se tiene que $M \bigcap M^\perp = \{0\}$.

Ejercicio 5.2.14 *Sea $M \subset \mathbb{H}$ un subconjunto (no necesariamente un subespacio vectorial) de un espacio de Hilbert $\mathbb{H}$. Entonces:*

1. *$M^\perp$ es un subespacio vectorial cerrado de $\mathbb{H}$.*
2. *Si $M \subset N$, entonces $N^\perp \subset M^\perp$.*
3. *$\overline{M^\perp} = M^\perp = \overline{M}^\perp$.*

1. Probemos que $M^\perp$ es un espacio lineal. Sean $x_1, x_2 \in M^\perp$ cualesquiera, i.e., para todo $y \in M$, $\langle x_1, y\rangle = \langle x_2, y\rangle = 0$. Entonces, para todo $y \in M$ usando la linealidad del producto escalar tenemos que $\langle \alpha x_1 + \beta x_2, y\rangle = \alpha\langle x_1, y\rangle + \beta\langle x_2, y\rangle = 0$, luego $\alpha x_1 + \beta x_2 \in M^\perp$. Probemos ahora que $M^\perp$ es cerrado. Para ello probaremos que $\overline{M^\perp} = M^\perp$ (ver Proposición 3.2.12), i.e., $M^\perp$ contiene a todos sus puntos límites. Sea x' un punto límite de $M^\perp$. Entonces, existe una sucesión de elementos de $M^\perp$, $(x'_n)_n$, tal que $\lim_{n\to\infty} x'_n = x' \in \mathbb{H}$ (ver Proposición 3.5.4). Por tanto, para todo $y \in M$,

$$\langle x', y\rangle = \left\langle \lim_{n\to\infty} x'_n, y\right\rangle = \lim_{n\to\infty} \langle x'_n, y\rangle = \lim_{n\to\infty} 0 = 0 \quad \Rightarrow \quad x' \in M^\perp.$$

Como x' era arbitrario, se sigue que $\overline{M^\perp} = M^\perp$.

2. Sea $x \in N^\perp$ cualquiera. Entonces, $\langle x, y\rangle = 0$ para todo $y \in N$, pero como $M \subset N$, $\langle x, y\rangle = 0$ para todo $y \in M$, luego $x \in M^\perp$. Como x era arbitrario se sigue que $N^\perp \subset M^\perp$.

3. Que $\overline{M^\perp} = M^\perp$ se sigue de 1 y la Proposición 3.2.12. Probemos que $M^\perp = \overline{M}^\perp$. Como $M \subset \overline{M}$, el punto anterior implica que $\overline{M}^\perp \subset M^\perp$. Probemos ahora que $M^\perp \subset \overline{M}^\perp$. Sea $y \in M^\perp$ cualquiera. Entonces,

$\langle x, y\rangle = 0$ para todo $x \in M$. Sea $x \in \overline{M}$ cualquiera, entonces por la Proposición 3.5.4, existe una sucesión de elementos de M, $(x_n)_n$, tal que $\lim_{n\to\infty} x_n = x \in \overline{M}$. Como $\langle x_n, y\rangle = 0$ para todo n, entonces tomando el límite cuando $n \to \infty$ se obtiene, por la continuidad del producto escalar, que $\langle x, y\rangle = 0$ cualquiera sea $x \in \overline{M}$, es decir, $y \in \overline{M}^{\perp}$, luego $M^{\perp} \subset \overline{M}^{\perp}$ y, por tanto, $M^{\perp} = \overline{M}^{\perp}$. ∎

El siguiente teorema[4] es de especial importancia.

Teorema 5.2.15 (de la proyección ortogonal) *Sea $\mathbb{H}$ un espacio de Hilbert (no necesariamente separable), sea $M \subset \mathbb{H}$ un subespacio cerrado de $\mathbb{H}$, y sea x un vector de $\mathbb{H}$ dado. Entonces existe un único $y \in M$ tal que se cumple[a] $\|y - x\| = \inf_{m\in M} \|x - m\|$. Además, existe dicho $y \in M$ si y solo si $\langle x - y, m\rangle = 0$, para todo $m \in M$.*

[a]Nótese que esta condición es equivalente a decir que $\|x-y\| \le \|x-m\|$, para todo $m \in M$.

Demostración: Si $x \in M$ tomamos $y = x$ y el teorema es trivial. Así que asumiremos que $x \notin M$. Sea $\delta = \inf_{m\in M} \|x - m\|$. Entonces existe una sucesión $(y_n)_n \in M$ tal que $\|x - y_n\| \to \delta$. Probemos que $(y_n)_n$ es de Cauchy. Para ello usamos la regla del paralelogramo (5.1.3)

$$\begin{aligned}\|(y_n - x) + (x - y_{n+p})\|^2 + \|(y_n - x) - (x - y_{n+p})\|^2& \\ = 2\|(y_n - x)\|^2 + 2\|(x - y_{n+p})\|^2 \quad &\Rightarrow \\ \|y_n - y_{n+p}\|^2 = 2\|(y_n - x)\|^2 + 2\|(x - y_{n+p})\|^2 - 4\left\|\frac{y_n}{2} + \frac{y_{n+p}}{2} - x\right\|^2.&\end{aligned}$$

Pero, como M es un subespacio de $\mathbb{H}$, entonces

$$\frac{y_n}{2} + \frac{y_{n+p}}{2} \in M \quad \Rightarrow \quad \left\|\frac{y_n}{2} + \frac{y_{n+p}}{2} - x\right\|^2 \ge \delta^2,$$

luego,

$$\|y_n - y_{n+p}\|^2 \le 2\|(y_n - x)\|^2 + 2\|(x - y_{n+p})\|^2 - 4\delta^2,$$

de donde, usando que $\|(y_n - x)\| \to \delta$, y tomando el límite $n \to \infty$, se tiene que $\|y_n - y_{n+p}\|^2 \to 0$.

[4]Este teorema sigue siendo cierto si M es un subconjunto convexo y completo de $\mathbb{H}$. Recordemos que un conjunto $M \subset \mathbb{H}$ es convexo si para todo $x, y \in M$, y para todo $t \in [0, 1]$ el vector $z = tx + (1 - t)y \in M$.

Como M es cerrado, es completo (ver Teorema 3.5.14), entonces $y_n \to y \in M$. Luego existe un $y \in M$ tal que

$$\delta = \|y - x\| = \inf_{m \in M} \|x - m\|. \tag{5.2.6}$$

Probemos que y es único. Asumamos que existe un $z \neq y$ tal que $\delta = \|z - x\|$. Entonces

$$\begin{aligned}\|y - z\|^2 &= 2\|(y - x)\|^2 + 2\|(x - z)\|^2 - 4\left\|\frac{y}{2} + \frac{z}{2} - x\right\|^2 \\ &= 4\delta^2 - 4\left\|\frac{y}{2} + \frac{z}{2} - x\right\|^2 \leq 4\delta^2 - 4\delta^2 = 0,\end{aligned}$$

pues $\left\|\frac{y}{2} + \frac{z}{2} - x\right\|^2 \geq \delta^2$ ya que $\frac{y}{2} + \frac{z}{2} \in M$. Luego $x = y$ lo que es una contradicción.

Probemos ahora que $\langle x - y, m\rangle = 0$, para todo $m \in M$, i.e., $(x - y)\perp M$. Para ello supongamos que existe un $m \in M$ que no es ortogonal a $x - y$. Asumiremos que $\|m\| = 1$. Sea el vector $m_1 = y + \lambda m$, con $\lambda = \langle x - y, m\rangle \neq 0$, y siendo y es vector de antes que minimiza el error $\|x - m\|$. Entonces

$$\begin{aligned}\|x - m_1\|^2 = \|x - y - \lambda m\|^2 &= \|x - y\|^2 - \langle x - y, \lambda m\rangle - \langle \lambda m, x - y\rangle + |\lambda|^2 \\ &= \|x - y\|^2 - |\lambda|^2 \leq \|x - y\|^2,\end{aligned}$$

i.e., y no es el vector que minimiza el error, lo cual es una contradicción.

Finalmente, notemos que cualquiera sea $m \in M$, si $x - y$ ortogonal a todos los $m \in M$, como $m - y \in M$, entonces usando el Teorema de Pitágoras (ver el Problema 5.4)

$$\|x - m\|^2 = \|(x - y) - (m - y)\|^2 = \|x - y\|^2 + \|y - m\|^2 \geq \|x - y\|^2.$$

Es decir, $\|x - y\| \leq \|x - m\|$ para todo $m \in M$, i.e., y es un vector óptimo, que como ya vimos antes, es único. ■

Un caso particular del teorema anterior es el caso cuando M es un subespacio generando por un número finito de vectores linealmente independientes, $e_1, \ldots, e_p$, de $\mathbb{H}$. En efecto, si $M = \operatorname{span}(e_1, \ldots, e_p)$, por el Corolario 4.3.3, M es cerrado (ver Ejemplo 4.3.4), por lo que podemos aplicar el Teorema 5.2.15, que nos conduce al siguiente resultado:

Corolario *Sea $\mathbb{H}$ un espacio de Hilbert (no necesariamente separable), sea M el espacio generado por los vectores linealmente independientes $e_1, \dots, e_p \in \mathbb{H}$, i.e., $M = \mathrm{span}\,(e_1, \dots, e_p)$, y sea x un vector de $\mathbb{H}$ dado. Entonces existe un único $y \in M$ tal que $\|x - y\| \le \|x - m\|$ para todo $m \in M$. Además existe dicho $y \in M$ si y solo si $\langle x - y, m\rangle = 0$, para todo $m \in M$.*

Otra consecuencia inmediata del Teorema de la proyección ortogonal 5.2.15 es el siguiente teorema usualmente conocido como el Teorema de la suma directa:

Teorema 5.2.16 (de la suma directa) *Sea $M \subset \mathbb{H}$ un subespacio cerrado del espacio de Hilbert (no necesariamente separable) $\mathbb{H}$ y $M^\perp$ su complemento ortogonal. Entonces, todo vector $x \in \mathbb{H}$ admite una única representación de la forma $x = y + y^\perp$ donde $y \in M$ e $y^\perp \in M^\perp$.*

Si todo elemento de $\mathbb{H}$ se puede escribir en la forma $x = y + y^\perp$ donde $y \in M$, M cerrado, e $y^\perp \in M^\perp$, entonces se dice que $\mathbb{H}$ es *suma directa* de M y $M^\perp$, y se escribe como

$$\mathbb{H} = M \oplus M^\perp. \tag{5.2.7}$$

Lo anterior nos permite enunciar el Teorema 5.2.16 de la siguiente forma equivalente:

Teorema de la suma directa *Sea un espacio de Hilbert (no necesariamente separable) $\mathbb{H}$ y sea $M \subset \mathbb{H}$ un subespacio cerrado de $\mathbb{H}$. Entonces $\mathbb{H}$ se puede escribir como suma directa de M y $M^\perp$, i.e., se tiene* (5.2.7).

Demostremos el Teorema 5.2.16.

Demostración: En primer lugar, si $x \in M$ (en particular $x = 0$), entonces basta tomar[5] $y = x$ y $y^\perp = 0$ y se tiene el resultado. Supongamos ahora que $x \notin M$ y que $x \neq 0$ y sea $y \in M$ el vector que cumple que $\|y - x\| = \inf_{m\in M} \|x - m\|$ (que por el Teorema de la proyección ortogonal 5.2.15 anterior sabemos que es único). Entonces, por la segunda parte del Teorema de la proyección ortogonal 5.2.15, el vector $y^\perp = x - y$ es ortogonal a todos los $m \in M$, i.e., $\langle y^\perp, m\rangle = 0$ para todo $m \in M$. Como y

[5] Sea $x \in M$ y supongamos que existe un $y \neq x$ tal que $x = y + y^\perp$. Entonces $y^\perp = x - y \in M$, pero el único elemento que puede estar en M y $M^\perp$ es el vector nulo, luego $x = y$ lo que es una contradicción.

es único, $y^\perp$ también lo es. De lo anterior se sigue que, dado un subespacio cerrado $M \subset \mathbb{H}$ de un espacio de Hilbert $\mathbb{H}$ cualquiera (no necesariamente separable), y para todo $x \in \mathbb{H}$, existen unos únicos $y \in M$ e $y^\perp \in M^\perp$ tales que

$$x = y + y^\perp, \qquad y \in M, \quad y^\perp \in M^\perp. \tag{5.2.8}$$

Es decir, cualquiera sea $x \in \mathbb{H}$ se tiene la descomposición (5.2.7). ■

Por completitud daremos una prueba alternativa del teorema anterior. Sea $x \notin M$, $x \neq 0$ (en caso contrario el resultado es trivial) y definamos el vector $y^\perp = x - y$, donde $y \in M$ es el único vector tal que $\delta = \|x - y\| = \operatorname{ínf}_{m \in M} \|x - m\|$ (que sabemos que existe por el Teorema 5.2.15), luego $y^\perp$ también es único. Probemos que $y^\perp \perp M$, luego $y^\perp \in M^\perp$, de donde se deduce el teorema. Sea $x \notin M$, $x \neq 0$. Entonces, dado un elemento $u \in M$, $u \neq 0$, y un escalar α cualesquiera definamos $m_1 = y + \alpha u \in M$, se tiene

$$\begin{aligned}\|x - y\|^2 = \delta^2 &= \operatorname*{ínf}_{m \in M} \|x - m\|^2 \leq \|x - m_1\|^2 = \|x - y - \alpha u\|^2 = \|y^\perp - \alpha u\|^2 \\ &= \|y^\perp\|^2 + |\alpha|^2 \|u\|^2 - 2\Re \alpha \langle y^\perp, u\rangle.\end{aligned}$$

Si elegimos ahora $\alpha = \overline{\langle y^\perp, u\rangle}/\|u\|^2$, teniendo en cuenta que $\|y^\perp\| = \delta$ (recuérdese que $x - y = y^\perp$), obtenemos que, para todo $u \in M$,

$$\delta^2 \leq \delta^2 - \frac{|\langle y^\perp, u\rangle|^2}{\|u\|^2} \;\Rightarrow\; |\langle y^\perp, u\rangle|^2 \leq 0 \;\Rightarrow\; \langle y^\perp, u\rangle = 0 \quad \Rightarrow \quad y^\perp \perp M.$$

Nota 5.2.17 *Es fácil ver que la noción de suma directa (5.2.7) se puede extender al caso de un número finito o numerable de subespacios M_1, M_2, etc. Esto será de de interés cuando veamos más adelante el Teorema espectral para operadores autoadjuntos y compactos (ver, por ejemplo, el Teorema de Hilbert-Schmidt 6.5.17).*

Nótese que si en (5.2.7) elegimos $M = \mathbb{H}$ obtenemos que $\mathbb{H} = \mathbb{H} \oplus \mathbb{H}^\perp = \mathbb{H} \oplus \{0\}$. Está claro además que si $x \in M \subset \mathbb{H}$, entonces $x \perp M^\perp$, luego $x \in (M^\perp)^\perp$ y, por tanto, $M \subset (M^\perp)^\perp$.

Una consecuencia de la descomposición (5.2.8) para espacios de Hilbert (no necesariamente separables) es la siguiente propiedad:

Ejercicio 5.2.18 *Probar que si M es un subespacio cerrado de $\mathbb{H}$, entonces $(M^\perp)^\perp = M$.*

Como ya vimos cualquiera sea $M \subset \mathbb{H}$, $M \subset (M^\perp)^\perp$. Sea $x \in (M^\perp)^\perp$ cualquiera. Como M es un subespacio cerrado entonces, por (5.2.8) se tiene que $x = y + y^\perp$ con $y \in M \subset (M^\perp)^\perp$ e $y^\perp \in M^\perp$. Por otro lado, $y^\perp = x - y \in (M^\perp)^\perp$ (pues $x \in (M^\perp)^\perp$ por hipótesis e $y \in M \subset (M^\perp)^\perp$) de donde se sigue que $y^\perp \perp M^\perp$, es decir, $\langle y^\perp, y^\perp \rangle = 0$ (recuérdese que $y^\perp \in M^\perp$), i.e., $y^\perp = 0$, luego $x = y \in M$. Como x era arbitrario, se tiene $(M^\perp)^\perp \subset M$, de donde se sigue el resultado. ∎

Como consecuencia del ejercicio anterior y del primer apartado del Ejercicio 5.2.14 se deduce que un subespacio M de un espacio de Hilbert $\mathbb{H}$ es cerrado si y solo si $(M^\perp)^\perp = M$ (ver Problema 5.11).

Una aplicación interesante

El corolario del Teorema 5.2.15 (ver página 196) nos dice que dado un vector $x \in \mathbb{H}$, siempre existe un único vector $y \in M = \operatorname{span}(y_1, \ldots, y_p)$, $y = \sum_{k=1}^p \alpha_k y_k$, que es el más cercano a dicho x, i.e., y minimiza las distancias $\|x - m\|$, para todo $m \in M$. Además, el $\min_{m \in M} \|x - m\|$ se alcanza cuando $x - y$ es ortogonal a cada todos los elementos de M o, equivalentemente, a cada uno de los y_j, $j = 1, \ldots, p$. Luego

$$\left\langle x - \sum_{k=1}^p \alpha_k y_k, y_j \right\rangle = 0, \quad \forall j = 1, \ldots, p.$$

Ello nos conduce al siguiente sistema de n ecuaciones lineales con n incógnitas (los coeficientes α_k, $k = 1, \ldots, p$):

$$\sum_{k=1}^p \langle y_k, y_j \rangle \alpha_k = \langle x, y_j \rangle, \quad j = 1, \ldots, p, \tag{5.2.9}$$

conocido como *sistema normal*. Nótese que la matriz del sistema normal coincide con la transpuesta de la matriz de Gram de los vectores $y_1, \ldots, y_p$ (5.1.8). Dicho sistema (5.2.9) tiene solución (es decir, podremos encontrar los α_i, $i = 1, \ldots, p$) si el determinante de su matriz es distinto de cero. Pero dicho determinante, que coincide con el determinante de Gram (5.1.9), es distinto de cero si y solo si los vectores y_i, $i = 1, \ldots, p$, son linealmente independientes (ver Ejercicio 5.1.15).

Tomemos ahora un $x \in \mathbb{H}$ dado y sea $\delta = \min_{m \in M} \|x - m\| = \|x - y\|$. Entonces, como $y \in M$ y $(x - y) \perp M$,

$$\delta^2 = \|x - y\|^2 = \langle x - y, x - y \rangle = \langle x - y, x \rangle,$$

o, equivalentemente,

$$\sum_{k=1}^{p} \alpha_j \langle y_k, x \rangle + \delta^2 = \langle x, x \rangle.$$

Juntando esta última ecuación con sistema normal (5.2.9) obtenemos el siguiente sistema lineal de $p+1$ ecuaciones con $p+1$ incógnitas:

$$\begin{pmatrix} \langle y_1, y_1 \rangle & \langle y_2, y_1 \rangle & \cdots & \langle y_p, y_1 \rangle & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \langle y_1, y_p \rangle & \langle y_2, y_p \rangle & \cdots & \langle y_p, y_p \rangle & 0 \\ \langle y_1, x \rangle & \langle y_2, x \rangle & \cdots & \langle y_p, x \rangle & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_p \\ \delta^2 \end{pmatrix} = \begin{pmatrix} \langle x, y_1 \rangle \\ \vdots \\ \langle x, y_p \rangle \\ \langle x, x \rangle \end{pmatrix}.$$

Resolviendo el sistema anterior –nótese que su determinante coincide con el del sistema normal (5.2.9)– podemos encontrar no solo los coeficientes α_k y, por tanto, el vector y, donde se alcanza el mínimo $\text{mín}_{m\in M} \|x-m\|$, sino también el valor δ del mismo. De hecho usando la regla de Cramer [24], y teniendo en cuenta que los determinantes de una matriz y su transpuesta coinciden, obtenemos

$$\delta^2 = \frac{\Delta_{p+1}(y_1, \ldots, y_p, x)}{\Delta_p(y_1, \ldots, y_p)},$$

donde $\Delta_n(v_1, \ldots, v_k)$ es el determinante de Gram de los vectores $v_1, \ldots, v_k$ (5.1.9).

Como aplicación de lo anterior veamos el siguiente ejercicio:

Ejercicio 5.2.19 *Sean m números reales $x_1, \ldots, x_m$ y sea una matriz $m \times n$, W con $n < m$ cuyas columnas son linealmente independientes. Encontrar el vector $z \in \mathbb{R}^n$ tal que la norma euclídea $\|x - Wz\|$ sea mínima, i.e., queremos encontrar el mejor estimador lineal de $x \in \mathbb{R}^m$.*

Como $\mathbb{R}^n$ es un espacio de Hilbert podemos usar el Teorema de las proyecciones 5.2.15. Sean $w_k = (w_{1,k}, \ldots, w_{m,k})^T$, $k = 1, \ldots, n$ las columnas de W que son l.i., y sea $M = \text{span}\,(w_1, \ldots, w_n)$. Entonces, existe un $y = Wz$ único tal que $\|x - y\|$ es mínimo. Para encontrar $z = (z_1, \ldots, z_n)$ basta usar el sistema de ecuaciones normal (5.2.9).

Un cálculo directo nos da que la matriz de dicho sistema es $W^T W$ y que el vector correspondiente al término independiente es $W^T x$, donde

W^T es la matriz traspuesta de W. Así que en nuestro caso el sistema de ecuaciones normal se escribe como

$$W^T W z = W^T x \quad \Rightarrow \quad z = (W^T W)^{-1} W^T x,$$

pues al ser las columnas de W linealmente independientes, existe la inversa de $W^T W$. ■

Otras aplicaciones interesantes se pueden encontrar en [12].

5.2.2. El Teorema de representación de Riesz

Para terminar este capítulo demostraremos un teorema sobre funcionales lineales acotados de gran importancia en la teoría de operadores en espacios de Hilbert tal y como veremos en el próximo capítulo.

Definición 5.2.20 *Un funcional lineal es una aplicación lineal $f : \mathbb{H} \mapsto \mathbb{K}$ siendo $\mathbb{K}$ el conjunto $\mathbb{C}$ o $\mathbb{R}$ dependiendo si $\mathbb{H}$ es un espacio de Hilbert real o complejo, respectivamente.*

Teorema 5.2.21 (de representación de Riesz) *Cualquier funcional lineal, $f : \mathbb{H} \mapsto \mathbb{K}$, acotado sobre un espacio de Hilbert $\mathbb{H}$ (no necesariamente separable), se puede representar en términos de un producto escalar, i.e.,*

$$f(x) = \langle x, z \rangle, \tag{5.2.10}$$

donde z esta unívocamente determinado por f, i.e., a cada f le corresponde uno y solo un vector $z \in \mathbb{H}$. Además, se tiene que

$$\|z\| = \|f\|. \tag{5.2.11}$$

Demostración: Si $f = 0$ el teorema es trivial tomando $z = 0$. Supongamos que f no es 0, en cuyo caso obviamente $z \neq 0$. Está claro que si (5.2.10) es cierta, entonces $f(x) = 0$ para todo x tal que $\langle x, z \rangle = 0$, es decir para todo x perteneciente al espacio nulo de f (que denotaremos por $\mathcal{N}$) se tiene que $z \perp x$, i.e., $z \in \mathcal{N}^\perp$. Dado que f es no nula, entonces $\mathcal{N} \neq \mathbb{H}$ (¿por qué?). Pero $\mathcal{N}$ es un subespacio vectorial cerrado[6] de $\mathbb{H}$ (ver Problema

[6] Una prueba directa es como sigue: sea $x \in \overline{\mathcal{N}}$. Existe una $(x_n)_n \in \mathcal{N}$ tal que $x_n \to x$. Pero $f(x_n) \to f(x)$, pues f es acotado. Como $f(x_n) = 0$, $f(x) = 0$, luego $x \in \mathcal{N}$.

4.15) luego es válida la descomposición (5.2.8) y, por tanto, se tiene que $\mathbb{H} = \mathcal{N} \oplus \mathcal{N}^{\perp}$, luego $\mathcal{N}^{\perp} \neq \{0\}$, por tanto existe un $z_0 \in \mathcal{N}^{\perp}$, $z_0 \neq 0$, que, sin pérdida de generalidad, tomaremos normalizado a la unidad $\|z_0\| = 1$. Construyamos un vector v tal que para todo $x \in \mathbb{H}$

$$v = f(x)z_0 - f(z_0)x \quad \Rightarrow \quad f(v) = 0 \quad \Rightarrow \quad v \in \mathcal{N}.$$

Pero entonces, $z_0 \perp v$, pues $z_0 \in \mathcal{N}^{\perp}$ y, por tanto,

$$0 = \langle v, z_0 \rangle = f(x) - f(z_0)\langle x, z_0 \rangle \quad \Rightarrow \quad f(x) = \langle x, \overline{f(z_0)} z_0 \rangle,$$

luego existe un $z \in \mathbb{H}$ tal que se tiene (5.2.10) para todo $x \in \mathbb{H}$.

Demostremos que dicho z es único. Sean $y \neq z$ tal que, para todo $x \in \mathbb{H}$, $f(x) = \langle x, z \rangle = \langle x, y \rangle$. Por el punto 4 del Ejercicio 5.1.2 tenemos que $z = y$ lo que es una contradicción.

Probemos ahora (5.2.11). En primer lugar sabemos que f es acotada y se cumple (5.2.10). Usando la desigualdad de Cauchy-Schwarz (5.1.2) tenemos, efectivamente, $|f(x)| = |\langle x, z \rangle| \leq \|z\|\|x\|$, luego tomando el supremo en los $x \in \mathbb{H}$ con $\|x\| = 1$ tenemos

$$\|f\| := \sup_{\|x\|=1} \frac{|f(x)|}{\|x\|} \leq \|z\|,$$

de donde además se deduce que, para todo $x \in \mathbb{H}$, $|f(x)| \leq \|f\|\|x\|$. Si en esta última desigualdad tomamos $x = z$ (recordemos que $z \neq 0$) obtenemos

$$\|z\|^2 = \langle z, z \rangle = |f(z)| \leq \|f\|\|z\| \quad \Rightarrow \quad \|z\| \leq \|f\|$$

de donde se sigue que $\|f\| = \|z\|$. ■

Conviene hacer notar que en la prueba del Teorema de representación de Riesz 5.2.21 es válida para cualquier espacio de Hilbert $\mathbb{H}$, no necesariamente separable. Esto será importante a la hora de probar el Teorema de Hahn-Banach 7.2.7 como veremos en el capítulo 7.

5.3. Problemas

Problema 5.1 Sea $\mathbb{E}$ un espacio euclídeo y $\|\cdot\|$ la norma inducida por el producto escalar. Prueba que para todos $x, y \in \mathbb{E}$,

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2). \tag{5.3.1}$$

Esta igualdad se suele denominar *ley del paralelogramo*.

Solución: Basta usar que $\|x\|^2 = \langle x, x\rangle$ y usar las propiedades del producto escalar 5.1.1 como en el caso del Problema 5.4. ■

Problema 5.2 Prueba que en el caso complejo un espacio normado es euclídeo si se tiene la ley del paralelogramo (5.3.1). Para ello hay que probar que la cantidad

$$\langle x, y\rangle = \frac{1}{4}\big(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\big)$$

define un producto escalar en $\mathbb{X}$ complejo.

Solución: La prueba es similar al caso real (ver la Proposición 5.1.9). Ver e.g. [22, Teorema 11.1, pág. 244] . ■

Problema 5.3 Usando el Ejercicio 5.1 prueba que los espacios normados $\mathbb{R}^n_p$, ℓ^p y $C^p_{[a,b]}$ correspondientes al ejercicio 4.2.4 y los ejemplos 4.2.7 y 4.2.5 no son espacios euclídeos salvo para $p = 2$.

Solución: Para ver que no son euclídeos usaremos la Proposición 5.1.9, es decir, mostraremos que no se cumple la ley del paralelogramo.

Comenzamos con $\mathbb{R}^n_p$. Sean los vectores $x = (1, 0, ..., 0), y = (0, 1, ..., 0) \in \mathbb{R}^n$. Tenemos por un lado que

$$2(\|x\|_p^2 + \|y\|_p^2) = 2((1^p)^{\frac{2}{p}} + (1^p)^{\frac{2}{p}}) = 4,$$

y por el otro,

$$\|x+y\|_p^2 + \|x-y\|_p^2 = 2^{\frac{2}{p}} + 2^{\frac{2}{p}} = 2\cdot 4^{\frac{1}{p}}.$$

Luego, usando (5.1.3) tenemos que $4 = 2\cdot 4^{\frac{1}{p}}$ lo cual es falso si $p \neq 2$.

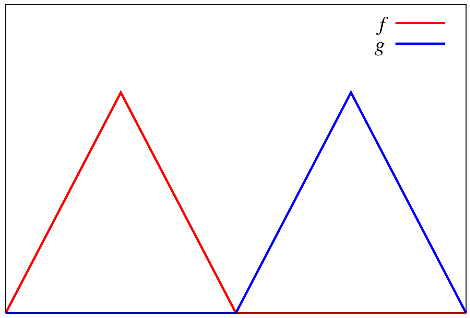


Figura 5.1: Las funciones $f(x)$ y $g(x)$ del Problema 5.3.

Para ℓ^p, podemos las sucesiones $x = (1,0,0,\dots), y = (0,1,0,\dots) \in l^p$ y obtenemos el mismo resultado.

Veamos ahora el espacio $C^p_{[a,b]}$. Lo que haremos es encontrar dos funciones continuas con soporte compacto disjunto y con la misma norma, así su suma y su diferencia también tendrán la misma norma. De esta forma consideramos las siguientes funciones (ver la figura 5.1):

$$f(x) = \begin{cases} x, & x \in [a, \frac{a+b}{4}), \\ \frac{a+b}{2} - x, & x \in [\frac{a+b}{4}, \frac{a+b}{2}), \\ 0, & x \in [\frac{a+b}{2}, b], \end{cases} \quad \text{y} \quad g(x) = \begin{cases} 0, & x \in [a, \frac{a+b}{2}), \\ x - \frac{a+b}{2}, & x \in [\frac{a+b}{2}, \frac{3(a+b)}{4}), \\ b - x, & x \in [\frac{3(a+b)}{4}, b]. \end{cases}$$

Usando las simetrías de f y g para calcular las normas tenemos:

$$\begin{aligned} 2(\|f\|_p^2 + \|g\|_p^2) &= 4\|f\|_p^2 = 4\Big(\int_a^b |f(x)|^p\,dx\Big)^{\frac{2}{p}} = 4\Big(2\int_a^{\frac{a+b}{4}} x^p\,dx\Big)^{\frac{2}{p}} = \\ &= 4\cdot 4^{\frac{1}{p}}\Big(\int_a^{\frac{a+b}{4}} x^p\,dx\Big)^{\frac{2}{p}} \\ \|f+g\|_p^2 + \|f-g\|_p^2 &= 2\|f+g\|_p^2 = 2\Big(\int_a^b |f(x)+g(x)|^p\,dx\Big)^{\frac{2}{p}} = \\ &= 2\Big(4\int_a^{\frac{a+b}{4}} x^p\,dx\Big)^{\frac{2}{p}} = 2\cdot 4^{\frac{2}{p}}\Big(\int_a^{\frac{a+b}{4}} x^p\,dx\Big)^{\frac{2}{p}}. \end{aligned}$$

Luego, como $4\cdot 4^{\frac{1}{p}} \neq 2\cdot 4^{\frac{2}{p}}$ para todo $p \neq 2$, tenemos que

$$2(\|f\|_p^2 + \|g\|_p^2) \neq \|f+g\|_p^2 + \|f-g\|_p^2, \quad \forall p \neq 2.$$

Finalmente, notemos que para $p = 2$ los espacios $\mathbb{R}^n_p$, ℓ^p y $C^p_{[a,b]}$ se corresponden con los ejemplos 5.1.3-5.1.5, respectivamente, donde se cumple la ley del paralelogramo pues en este caso sí que son espacios euclídeos. ■

Problema 5.4 Prueba que si dos elementos x e y de un espacio euclídeo son ortogonales, entonces

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2.$$

La igualdad anterior es una generalización del Teorema de Pitágoras.

Solución: Como x e y son ortogonales $\langle x, y\rangle = \langle y, x\rangle = 0$, luego

$$\begin{aligned}\|x+y\|^2 &= \langle x+y, x+y\rangle = \langle x, x+y\rangle + \langle y, x+y\rangle = \\ &= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle = \langle x, x\rangle + \langle y, y\rangle = \|x\|^2 + \|y\|^2.\end{aligned}$$ ■

Problema 5.5 Sea $\mathbb{H}$ un espacio de Hilbert separable. Prueba que cualquier subespacio cerrado $M \subset \mathbb{H}$ es a su vez un espacio de Hilbert. Además, M tiene una base ortonormal.

Solución: Como $\mathbb{H}$ es completo y $M \subset \mathbb{H}$ es cerrado entonces por el Teorema 3.5.14 M es completo. Pero como $\mathbb{H}$ es separable, entonces por el Ejercicio 3.4.7 M es separable. Entonces, por el Teorema 5.2.9, M tiene una base ortonormal.■

Problema 5.6 Dar una demostración alternativa del Teorema 5.2.16 suponiendo que $\mathbb{H}$ es un espacio de Hilbert separable. Es decir, prueba que si $M \subset \mathbb{H}$ es un subespacio cerrado del espacio de Hilbert separable $\mathbb{H}$ y $M^\perp$ su complemento ortogonal, entonces, todo vector $x \in \mathbb{H}$ admite una única representación de la forma $x = y + y^\perp$ donde $y \in M$ e $y^\perp \in M^\perp$.

Solución: Como $\mathbb{H}$ es separable y $M \subset \mathbb{H}$ es cerrado en $\mathbb{H}$, entonces M es completo (ver Teorema 3.5.14) y separable (ver el Ejercicio 3.4.7), i.e., M es a su vez un espacio de Hilbert separable por lo que existe una base ortonormal completa $(\phi_n)_n$ de M (ver Teorema 5.2.9). Definamos $y = \sum_n \langle x, \phi_n\rangle \phi_n$. Obviamente $y \in M$. Definamos $y^\perp = x - y$. Entonces, para todo n, $\langle y^\perp, \phi_n\rangle = 0$. Ahora bien, cualquiera sea $\tilde{y} \in M$, $\tilde{y} = \sum_n a_n \phi_n$, luego $\langle y^\perp, \tilde{y}\rangle = 0$, es decir, $y^\perp \in M^\perp$. Luego cualquiera sea $x \in \mathbb{H}$, existen $y \in M$ y $y^\perp \in M^\perp$ tales que $x = y + y^\perp$. Probemos que esta descomposición es única. Para ello supongamos que no lo son, i.e., supongamos que existen un $z \in M$, $z \neq y$ tal que $x = z + z^\perp$, $z^\perp \in M^\perp$ (y $z^\perp \neq y^\perp$). Entonces, para todo $n \in \mathbb{N}$, se tiene

$$\langle z, \phi_n\rangle = \langle x - z^\perp, \phi_n\rangle = \langle x, \phi_n\rangle = \langle y + y^\perp, \phi_n\rangle = \langle y, \phi_n\rangle$$

que, por el Corolario 5.2.7, implica $z = y$, lo cual es una contradicción. ■

Problema 5.7 A partir del Teorema 5.2.16 (o del Problema 5.6), prueba que todo sistema ortogonal $(\phi_n)_{n=1}^N$ en un espacio de Hilbert $\mathbb{H}$ separable se puede completar hasta obtener una base ortogonal de $\mathbb{H}$.

Solución: Del Teorema 5.2.9 se sigue que $\mathbb{H}$ tiene una base ortonormal, no obstante el objetivo de este problema no es probar que existe dicha base (eso ya lo afirma el teorema) sino mostrar cómo, dado un conjunto finito de vectores ortonormales completarlos hasta obtener dicha base. Existen varias formas de probar este resultado. Una de ellas, efectivamente, es usar el Teorema 5.2.16.

Sea M el espacio generado por el sistema ortogonal $(\phi_n)_{n=1}^N$. M es cerrado. Por el Teorema 5.2.16, tenemos que todo vector $x \in \mathbb{H}$ se expresa de forma única como una suma de un vector $y \in M$ y un vector $y^\perp \in M^\perp$. Al ser $\mathbb{H}$ separable y $M^\perp$ un subconjunto suyo, tenemos que $M^\perp$ es separable (por el Ejercicio 3.4.7) y cerrado (por el Ejercicio 5.2.14), por tanto $M^\perp$ es un espacio de Hilbert que admite una base ortogonal numerable (Teorema 5.2.9) que denotaremos por $(\phi_k^\perp)_k$. Dado que todo $x \in \mathbb{H}$ se expresa de forma única como $x = y + y^\perp$, y se expresa de forma única en la base $(\phi_n)_{n=1}^N$ y $y^\perp$ en la base $(\phi_k^\perp)_k$, entonces está claro que x se puede escribir de forma única en la base $(\phi_n)_{n=1}^N \bigcup (\phi_k^\perp)_k$. En otras palabras, $(\phi_n)_n \bigcup (\phi_k^\perp)_k$ es una base de $M \oplus M^\perp = \mathbb{H}$, luego hemos completado la base $(\phi_n)_{n=1}^N$ hasta obtener una base del espacio entero $\mathbb{H}$. ∎

Problema 5.8 Sea $\mathbb{H}$ un espacio de Hilbert separable y sea $(\phi_n)_n$ una base ortonormal completa de $\mathbb{H}$. Prueba que para todos $x, y \in \mathbb{H}$, se tiene la igualdad, también denominada igualdad de Parseval

$$\langle x, y\rangle = \sum_{n=1}^{\infty} \langle x, \phi_n\rangle \overline{\langle y, \phi_n\rangle}.$$

Solución: Primero probemos que la serie converge. Para ello notamos que, usando la desigualdad de Hölder (3.7.2) con $p = 2$ (o la desigualdad de Cauchy-Schwarz (5.1.2) en $\mathbb{C}^N$)

$$\left(\sum_{n=1}^{N} |\langle x, \phi_n\rangle \overline{\langle y, \phi_n\rangle}|\right)^2 \leq \sum_{n=1}^{N} |\langle x, \phi_n\rangle|^2 \sum_{n=1}^{N} |\overline{\langle y, \phi_n\rangle}|^2,$$

pero como los coeficientes de Fourier son una sucesión de ℓ^2 (ver la desigualdad de Bessel (5.2.3)), las sumas de la derecha están acotadas para todo N, luego tomando el límite $N \to \infty$ tenemos que la serie es absolutamente convergente, luego es convergente. Por otro lado, sea

$$x_N = \sum_{n=1}^{N} \langle x, \phi_n \rangle \phi_n \quad \text{y} \quad y_N = \sum_{n=1}^{N} \langle y, \phi_n \rangle \phi_n,$$

entonces, como $(\phi_n)_n$ es una base ortonormal completa de $\mathbb{H}$,

$$x_N \to x = \sum_{n=1}^{\infty} \langle x, \phi_n \rangle \phi_n \quad \text{y} \quad y_N \to y = \sum_{n=1}^{\infty} \langle y, \phi_n \rangle \phi_n,$$

cuando $N \to \infty$ (ver Teorema 5.2.5). Un cálculo directo, usando la ortonormalidad de los $(\phi_n)_n$, nos da

$$\langle x_N, y_N \rangle = \sum_{n=1}^{N} \langle x, \phi_n \rangle \overline{\langle y, \phi_n \rangle},$$

de donde, tomando el límite $N \to \infty$, y usando la continuidad del producto escalar obtenemos el resultado. ■

Problema 5.9 Sea $\mathbb{H}$ un espacio de Hilbert separable y sea $(\phi_n)_n$ una base ortonormal completa de $\mathbb{H}$. Prueba, sin usar el Teorema 5.2.5, que para todo $x \in \mathbb{H}$, la serie

$$y = \sum_{n=1}^{\infty} \langle x, \phi_n \rangle \phi_n$$

converge, y que el vector $x - y$ es ortogonal a cada ϕ_n, $n \geq 1$.

Solución: Si usamos el Teorema 5.2.5 el problema es *trivial*, pues del punto 2 de dicho teorema se sigue que la serie es convergente y que su suma es x, luego $x = y$ y, por tanto, $x - y = 0$ que es ortogonal a todos los ϕ_n. Demos una prueba directa sin usar el teorema. Sean $y_n = \sum_{k=1}^{N} \langle x, \phi_k \rangle \phi_k$ las sumas parciales de nuestra serie. Usando la ortonormalidad de $(\phi_k)_k$ tenemos que

$$\|y_{n+p} - y_n\|^2 = \langle y_{n+p} - y_n, y_{n+p} - y_n \rangle = \sum_{k=n+1}^{n+p} |\langle x, \phi_k \rangle|^2.$$

Como los coeficientes de Fourier son una sucesión de ℓ^2 (desigualdad de Bessel (5.2.3)), las sumas de la derecha están acotadas tienden a cero si $n \to \infty$ independientemente de p, luego la sucesión $(y_n)_n$ es de Cauchy y como $\mathbb{H}$ es completo, es convergente. Además, tomando $p \to \infty$ se tiene que $y_n \to y$.

Probemos ahora que para todo n, $(x-y)\perp\phi_n$. Sea $n < N$ cualquiera, entonces

$$\langle x - y_N, \phi_n\rangle = \langle x, \phi_n\rangle - \sum_{k=1}^{N}\langle x, \phi_k\rangle\langle\phi_k, \phi_n\rangle = \langle x, \phi_n\rangle - \langle x, \phi_n\rangle = 0.$$

Tomando en la expresión anterior el límite $N \to \infty$ y usando la continuidad del producto escalar obtenemos que, para todo n, $\langle x - y, \phi_n\rangle$ como se quería probar. De lo anterior y del Corolario 5.2.7 se sigue que $x = y$. ■

Problema 5.10 Sea $\mathbb{H}$ un espacio de Hilbert y sea $M \subset \mathbb{H}$. Definiremos como $\mathrm{span}\,(M)$ al conjunto de todas las combinaciones lineales de vectores de M. Prueba que para todo $M \neq \emptyset$, $V = \mathrm{span}\,(M)$ es denso en $\mathbb{H}$ si y solo si $M^\perp = \{0\}$.

Solución: Probemos que si $\overline{V} = \mathbb{H}$, entonces $M^\perp = \{0\}$. Sea $x \in M^\perp$ cualquiera y supongamos que $\overline{V} = \mathbb{H}$. Entonces $x \in \mathbb{H} = \overline{V}$, i.e., x es un punto de acumulación de V, luego existe $(x_n)_n \in V$ tal que $x_n \to x$. Como $x \in M^\perp$, entonces $x \perp M$ y, por tanto, $x \perp V$, luego para todo $x \in V^\perp$, $\langle x_n, x\rangle = 0$. Tomando el límite cuando $n \to \infty$ en la igualdad anterior y usando la continuidad del producto escalar, tenemos $0 = \langle x_n, x\rangle \to \langle x, x\rangle$, luego $x = 0$. Como x era cualquiera, $M^\perp = \{0\}$.

Supongamos ahora que $M^\perp = \{0\}$. Sea $x \perp V$ cualquiera. Entonces $x \perp M$, i.e., $x \in M^\perp$, pero como $M^\perp = \{0\}$ se tiene que $x = 0$, luego $V^\perp = \{0\}$. Ahora bien, como $\overline{V} \subset \mathbb{H}$ es un cerrado, el Teorema de la suma directa 5.2.16, y que $\overline{V}^\perp = V^\perp$ (ver Ejercicio 5.2.14, punto 3), obtenemos que $\mathbb{H} = \overline{V} \oplus \overline{V}^\perp = \overline{V} \oplus V^\perp = \overline{V} \oplus \{0\}$ y, por tanto, $\overline{V} = \mathbb{H}$. ■

Problema 5.11 Sea $\mathbb{H}$ un espacio de Hilbert y sea M un subespacio de $\mathbb{H}$. Prueba que M es cerrado si y solo si $M = (M^\perp)^\perp$.

Solución: Si $M = (M^\perp)^\perp$, entonces por el ejercicio 5.2.14, M es cerrado. Si M es cerrado por el ejercicio 5.2.18, $M = (M^\perp)^\perp$. Luego M es cerrado si y solo si $M = (M^\perp)^\perp$. ■

Problema 5.12 Sea $\mathbb{H}$ un espacio de Hilbert y sea $M \subset \mathbb{H}$ un subespacio cerrado de $\mathbb{H}$. En este caso sabemos (ver (5.2.7)) que para todo $x \in \mathbb{H}$ existen unos únicos $y \in M$ y $y^\perp \in M^\perp$ tales que $x = y + y^\perp$. Definamos el operador $P : \mathbb{H} \mapsto M$ mediante la fórmula $x \mapsto y = Px$. P se denomina operador de proyección (o proyector ortogonal) de $\mathbb{H}$ sobre M. Prueba que

1. P es lineal, acotado y si $M \neq \{0\}$, $\|P\| = 1$.
2. $P(M^{\perp}) = 0$.
3. $P^2 = P \cdot P = P$ (es idempotente).

Solución: 1. Sea $x_1, x_2 \in M$. Entonces como $x_1 = y_1 + y_1^{\perp}$ y $x_2 = y_2 + y_2^{\perp}$, tenemos $Px_1 = y_1$ y $Px_2 = y_2$. Entonces, como $x_1 + x_2 = (y_1 + y_2) + (y_1^{\perp} + y_2^{\perp})$, y $\lambda x = \lambda y + \lambda y^{\perp}$, tenemos

$$P(x_1 + x_2) = y_1 + y_2 = Px_1 + Px_2, \quad P(\lambda x) = \lambda y = \lambda P x,$$

luego P es lineal. Además, usando el Teorema de Pitágoras (ver Problema 5.4) tenemos que, para todo $x \in \mathbb{H}$,

$$\|x\|^2 = \|y + y^{\perp}\|^2 = \|y\|^2 + \|y^{\perp}\|^2 \geq \|y\|^2 = \|Px\|^2 \quad \Rightarrow \quad \|Px\| \leq \|x\|,$$

luego P es acotado y $\|P\| \leq 1$. Si $M = \{0\}$ es obvio que $Px = 0$ para todo $x \in \mathbb{H}$ (¿por qué?), pero si $M \neq \{0\}$, entonces tomando $x = y \in M$ se tiene $Py = y$, luego $\sup_{x \in \mathbb{H}, \|x\|=1} \|Px\| \geq \sup_{y \in M, \|y\|=1} \|Py\| = 1$, luego $\|P\| \geq 1$, de donde se sigue el resultado.

2. Está claro que de la definición de P, $Px = x$ para todo $x \in M$, además $Px = 0$ para todo $x \in M^{\perp}$, luego $P(M) = M$ y $P(M^{\perp}) = 0$.

3. Finalmente, para todo $x \in \mathbb{H}$, $P^2 x = P(Px) = Py = y = Px$, luego $P^2 = P$. ■

Capítulo 6

Operadores en espacios de Hilbert

> Hay partes importantes de las matemáticas que no se pueden entender en profundidad sin la ayuda de la teoría de operadores.
>
> Stefan Banach
>
> En *"Théorie des Opérations Linéaires"* (1932)

En este capítulo estudiaremos algunas propiedades de los operadores lineales definidos en espacios de Hilbert separables sobre $\mathbb{C}$.

6.1. Definiciones

Definición 6.1.1 *Sea la aplicación (operador) lineal y acotada* $A : \mathbb{E} \mapsto \mathbb{E}'$*,* $\mathbb{E}$*,* $\mathbb{E}'$ *espacios euclídeos. Si existe el operador lineal* $A^* : \mathbb{E}' \mapsto \mathbb{E}$ *tal que para todo* $x \in \mathbb{E}$ *e* $y \in \mathbb{E}'$

$$\langle Ax, y\rangle' = \langle x, A^* y\rangle, \tag{6.1.1}$$

lo denominaremos adjunto de A*.*

En general, por sencillez, asumiremos $\mathbb{E}' = \mathbb{E}$.

Por ejemplo, dado el operador *desplazamiento* (o *shift*) $S_{\curvearrowright} : \ell^2 \mapsto \ell^2$ definido por

$$S_{\curvearrowright}(x_1, x_2, x_3, \dots) = (0, x_1, x_2, \dots), \tag{6.1.2}$$

su adjunto $S^*_{\curvearrowright}$ es el operador $S_{\curvearrowleft}:\ell^2\mapsto\ell^2$ definido por

$$S_{\curvearrowleft}:=S^*_{\curvearrowright}(x_1,x_2,x_3,\dots)=(x_2,x_3,\dots). \tag{6.1.3}$$

Ejercicio 6.1.2 *Sea el operador multiplicación $M:C^2_{[a,b]}\mapsto C^2_{[a,b]}$ definido por*

$$Mx(t)=f(t)x(t),\quad f(t)\in C([a,b]),\quad \forall x(t)\in C^2_{[a,b]}.$$

Prueba que M^* existe y es el operador multiplicación por la función complejo conjugada $\overline{f(t)}$. Prueba que este operador es acotado y que $\|M\|\le \sup_{t\in[a,b]}|f(t)|$. Nótese que si $f(t)$ es real entonces $M^*=M$.

La existencia y forma de M^* se sigue de la igualdad

$$\langle Mx,y\rangle=\int_a^b f(t)x(t)\cdot\overline{y(t)}dt=\int_a^b x(t)\cdot\overline{\overline{f(t)}y(t)}dt.$$

Para probar la segunda afirmación notemos que

$$\|Mx\|^2=\left(\int_a^b |f(t)|^2|x(t)|^2dt\right)\le \sup_{t\in[a,b]}|f(t)|^2\|x\|^2=\left(\sup_{t\in[a,b]}|f(t)|\right)^2\|x\|^2.$$

Luego, $\|M\|=\sup_{\|x\|=1}\|Mx\|\le\sup_{t\in[a,b]}|f(t)|$. ■

Para los espacios euclídeos la existencia del operador adjunto no está garantizada en general, no obstante sí que lo está en el caso de los espacios de Hilbert. De hecho, como consecuencia del Teorema de representación de Riesz 5.2.21 se tiene el siguiente resultado:

Teorema 6.1.3 (Existencia del operador adjunto) *Sea la aplicación (operador) lineal y acotada $A:\mathbb{H}\mapsto\mathbb{H}'$, $\mathbb{H}$, $\mathbb{H}'$ espacios de Hilbert. Entonces existe un único operador $A^*:\mathbb{H}'\mapsto\mathbb{H}$ adjunto a A. Además, A^* es lineal y*

$$\|A^*\|=\|A\|. \tag{6.1.4}$$

Demostración: Para cada $y\in\mathbb{H}'$ definimos el funcional $T_yx:\mathbb{H}\mapsto\mathbb{K}$, $T_yx:=\langle Ax,y\rangle'$. Obviamente, para cada y fijo

$$|T_yx|\le\|Ax\|\|y\|\le\|A\|\|x\|\|y\|\le K\|x\|,\quad \forall y\in\mathbb{H}',$$

i.e., T_y es un funcional acotado, así que el Teorema de Riesz 5.2.21 nos asegura que existe un único vector $y^*\in\mathbb{H}$ tal que $T_yx=\langle x,y^*\rangle$, para todo

$x \in \mathbb{H}$. Así, el operador $A : \mathbb{H} \mapsto \mathbb{H}'$ induce un operador $A^* : \mathbb{H}' \mapsto \mathbb{H}$ que a cada $y \in \mathbb{H}'$ le hace corresponder un único $y^* \in \mathbb{H}$, tal que $A^*y = y^*$, pues $\langle Ax, y\rangle' = \langle x, y^*\rangle = \langle x, A^*y\rangle$. La linealidad de A^* se sigue de la linealidad de A. En efecto, para todos $x \in \mathbb{H}$ e $y, z \in \mathbb{H}'$, se tiene

$$\begin{aligned}\langle x, A^*(\alpha y + \beta z)\rangle &= \langle Ax, \alpha y + \beta z\rangle' = \overline{\alpha}\langle Ax, y\rangle + \overline{\beta}\langle Ax, z\rangle \\ &= \overline{\alpha}\langle x, A^*y\rangle + \overline{\beta}\langle x, A^*z\rangle = \langle x, \alpha A^*y + \beta A^*z\rangle.\end{aligned}$$

Luego $A^*(\alpha y + \beta z) = \alpha A^*y + \beta A^*z$.

Probemos ahora que $\|A^*\| = \|A\|$. De la igualdad $\langle Ax, y\rangle' = \langle x, A^*y\rangle$ y usando que A es acotado tenemos

$$|\langle x, A^*y\rangle| = |\langle Ax, y\rangle'| \leq \|A\|\|x\|\|y\|.$$

Sea $y \neq 0$ tal que[1] $A^*y \neq 0$ y escojamos $x = A^*y$. Entonces tenemos que

$$\|A^*y\|^2 \leq \|A\|\|A^*y\|\|y\| \;\Rightarrow\; \frac{\|A^*y\|}{\|y\|} \leq \|A\| \;\Rightarrow\; \|A^*\| \leq \|A\|.$$

De lo anterior se sigue que A^* es acotado. Pero entonces podemos aplicar el mismo razonamiento intercambiando A y A^* lo que nos conduce a la desigualdad contraria $\|A\| \leq \|A^*\|$. En efecto, asumiendo ahora que $\|Ax\| \neq 0$, y $\|x\| \neq 0$, tenemos

$$|\langle Ax, y\rangle'| = |\langle x, A^*y\rangle| \leq \|A^*\|\|x\|\|y\|, \quad y = Ax \;\Rightarrow$$
$$\|Ax\|^2 \leq \|A^*\|\|Ax\|\|x\| \;\Rightarrow\; \frac{\|Ax\|}{\|x\|} \leq \|A^*\| \;\Rightarrow\; \|A\| \leq \|A^*\|.$$

Por tanto $\|A\| = \|A^*\|$. ■

Veamos algunas propiedades de los operadores adjuntos.

Proposición 6.1.4 *Sean las aplicaciones lineales acotadas $A : \mathbb{H} \mapsto \mathbb{H}'$, $B : \mathbb{H} \mapsto \mathbb{H}'$, $\mathbb{H}$, $\mathbb{H}'$ espacios de Hilbert. Entonces se cumple que*

1. $\langle A^*y, x\rangle = \langle y, Ax\rangle'$, $x \in \mathbb{H}$ *e* $y \in \mathbb{H}'$.
2. $(A + B)^* = A^* + B^*$.
3. $(\alpha A)^* = \overline{\alpha}A^*$.
4. $(A^*)^* = A$.

[1] Si $\|A^*y\| = 0$, entonces usando la desigualdad anterior obtenemos que $\|A\| \geq 0$, lo cual es evidente.

5. $\|A^*A\| = \|AA^*\| = \|A\|^2$.

6. $A^*A = 0$ *si y solo si* $A = 0$.

7. *Si* $\mathbb{H} = \mathbb{H}'$, *entonces* $(AB)^* = B^*A^*$.

Demostración: Como $A, B \in \mathcal{B}(\mathbb{H}, \mathbb{H}')$, existen los adjuntos A^* y B^* de A y B, respectivamente. Para probar 1 usamos que $\langle A^*y, x\rangle = \overline{\langle x, A^*y\rangle} = \overline{\langle Ax, y\rangle'} = \langle y, Ax\rangle'$. Para 2 tenemos

$$\begin{aligned}\langle x, (A+B)^*y\rangle &= \langle (A+B)x, y\rangle' = \langle Ax, y\rangle' + \langle Bx, y\rangle' = \langle x, A^*y\rangle + \langle x, B^*y\rangle \\ &= \langle x, (A^*+B^*)y\rangle' \;\Rightarrow\; (A+B)^* = A^* + B^*.\end{aligned}$$

El punto 3 se sigue de

$$\langle x, (\alpha A)^*y\rangle = \langle (\alpha A)x, y\rangle' = \alpha\langle Ax, y\rangle' = \alpha\langle x, A^*y\rangle = \langle x, \overline{\alpha}A^*y\rangle.$$

Probemos 4. Como $A^* : \mathbb{H}' \mapsto \mathbb{H}$ es acotado, entonces existe su adjunto $(A^*)^* : \mathbb{H} \mapsto \mathbb{H}'$. Entonces,

$$\langle Ax, y\rangle' = \langle x, (A^*)y\rangle = \langle (A^*)^*x, y\rangle' \;\Rightarrow\; (A^*)^* = A.$$

Antes de probar 5 notemos que $A^*A : \mathbb{H} \mapsto \mathbb{H}$ y $AA^* : \mathbb{H}' \mapsto \mathbb{H}'$. Por otro lado, como A y A^* son acotados su producto A^*A también lo es (ver Problema 4.13) y $\|A^*A\| \leq \|A^*\|\|A\| = \|A\|^2$. Probemos ahora que $\|A\|^2 \leq \|A^*A\|$. Para ello usamos que ($\|x\| \neq 0$)

$$\|Ax\|^2 = \langle Ax, Ax\rangle' = \langle A^*Ax, x\rangle \leq \|A^*Ax\|\|x\| \leq \|A^*A\|\|x\|^2 \;\Rightarrow$$

$$\left(\frac{\|Ax\|}{\|x\|}\right)^2 \leq \|A^*A\| \Rightarrow \left(\sup_{\|x\|\neq 0} \frac{\|Ax\|}{\|x\|}\right)^2 \leq \|A^*A\| \Rightarrow \|A\|^2 \leq \|A^*A\|,$$

luego $\|A^*A\| = \|A\|^2$. Finalmente, intercambiando A y A^* obtenemos que $\|(A^*)^*A^*\| = \|A^*\|^2$. Usando 4 y que $\|A^*\| = \|A\|$ se sigue el resultado. La propiedad 6 es consecuencia de 5 y de las propiedades de la norma. Finalmente, para probar 7 usamos que

$$\langle x, (AB)^*y\rangle = \langle ABx, y\rangle = \langle Bx, A^*y\rangle = \langle x, B^*A^*y\rangle,$$

de donde se sigue que $(AB)^* = B^*A^*$. ■

Definición 6.1.5 *Sea $A : \mathbb{H} \mapsto \mathbb{H}$. Si $A = A^*$, se dice que el operador es autoadjunto o hermítico.*

Lo anterior implica que A es autoadjunto si y solo si para todos $x, y \in \mathbb{H}$,

$$\langle Ax, y \rangle = \langle x, A^* y \rangle = \langle x, Ay \rangle.$$

Por ejemplo, el operador multiplicación por una función real del Ejemplo 6.1.2 es autoadjunto, mientras que el operador desplazamiento $S_{\curvearrowright}$ definido por (6.1.2) no lo es.

Definición 6.1.6 *Sea $U : \mathbb{H} \mapsto \mathbb{H}$ biyectivo. Si $U^{-1} = U^*$, se dice que el operador es unitario.*

Nótese que lo anterior es equivalente a decir que $U^*U = UU^* = I$. Además de la definición de adjunto se sigue que

$$\langle Ux, Uy \rangle = \langle x, U^*Uy \rangle = \langle x, y \rangle,$$

i.e., los operadores unitarios preservan el producto escalar (ver la Definición 5.2.11).

Definición 6.1.7 *Sea $A : \mathbb{H} \mapsto \mathbb{H}$. Si $AA^* = A^*A$, se dice que el operador es normal.*

Obviamente tanto los operadores autoadjuntos como los unitarios son normales, pero no al contrario. Tómese como ejemplo el operador $T : \mathbb{H} \mapsto \mathbb{H}$, $T = 2iI$, donde I es el operador identidad (ver Ejemplo 4.4.2). Está claro que $TT^* = T^*T = 4I$, luego es normal, pero no es autoadjunto pues $T^* = -T$ y claramente no es unitario.

Antes de continuar vamos a hacer una breve incursión por el álgebra lineal. En adelante asumiremos que $\mathbb{H} = \mathbb{H}'$. Supongamos que $\mathbb{H}$ es un espacio de Hilbert de dimensión finita. Entonces como ya hemos visto, $\mathbb{H}$ es isomorfo a $\mathbb{C}^n$.

Sea $T : \mathbb{H} \mapsto \mathbb{H}$, un operador lineal y sea $(e_k)_k$ una base de $\mathbb{H}$. Entonces para todo $x \in \mathbb{H}$

$$x = \sum_{k=1}^{n} x_k e_k \ \Rightarrow \ y = Tx = \sum_{k=1}^{n} x_k T e_k.$$

Si

$$Te_k = \sum_{i=1}^{n} a_{i,k} e_i \ \Rightarrow \ y = \sum_{i=1}^{n} \left(\sum_{k=1}^{n} a_{i,k} x_k \right) e_i = \sum_{i=1}^{n} y_i e_i \ \Rightarrow$$

$$y_i = \sum_{k=1}^{n} a_{i,k} x_k, \qquad i = 1, 2, \ldots n.$$

Es decir, si consideramos los vectores $x, y \in \mathbb{C}^n$ con coordenadas x_i, y_i, $i = 1, \ldots, n$, respectivamente, entonces el operador T se puede representar usando una matriz $(a_{i,j})_{i,j=1}^{n}$, i.e., tenemos la aplicación $T : \mathbb{C}^n \mapsto \mathbb{C}^n$, $y = Ax$, donde A es la matriz $n \times n$

$$A = (Te_1 \, Te_2 \, \cdots \, Te_n) = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots & a_{2,n} \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdots & a_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \cdots & a_{n,n} \end{pmatrix},$$

obtenida a partir de la transformaciones de los elementos de la base $(e_k)_k$ de $\mathbb{H}$.

Nota 6.1.8 *Nótese que la equivalencia de un operador lineal $A : \mathbb{H} \mapsto \mathbb{H}$ definido sobre un espacio de dimensión finita, $\dim \mathbb{H} = n < +\infty$, discutida aquí es extensible a cualquier espacio normado (no necesariamente un espacio de Hilbert) de dimensión finita, pues la construcción que se ha hecho solo hace uso de las propiedades de los espacios vectoriales y las aplicaciones lineales y no de las del producto escalar.*

Si tenemos un espacio de Hilbert de dimensión finita hemos visto que es isomorfo a $\mathbb{C}^n$ (ver Teorema 5.2.12). Tomemos en $\mathbb{C}^n$ la base canónica $e_k = \delta_{i,k}$, $k, i = 1, \ldots, n$. Dados $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$, el producto escalar viene dado por

$$\langle x, y \rangle = \sum_{k=1}^{n} x_k \overline{y_k} = x^T \cdot \overline{y},$$

donde por x^T entendemos el vector traspuesto de x. Luego

$$\langle Tx, y \rangle = (Ax)^T \cdot \overline{y} = x^T \cdot (A^T \overline{y}) = x^T \cdot \overline{(\overline{A}^T \, y)} = \langle x, T^* y \rangle.$$

Es decir, si al operador T le corresponde la matriz A a su adjunto T^* le corresponde la matriz $A^* = \overline{A}^T$.

Así, tenemos en dimensión finita que

1. Un operador T es autoadjunto si su correspondiente matriz A en hermítica, i.e., $A = \overline{A}^T$.

2. Un operador T es unitario si su correspondiente matriz A es unitaria, i.e., $A^{-1} = \overline{A}^T \iff A \cdot \overline{A}^T = \overline{A}^T \cdot A = I$.

3. Un operador T es normal si su correspondiente matriz A satisface que $A \cdot \overline{A}^T = \overline{A}^T \cdot A$.

Ejercicio 6.1.9 *Prueba que la matriz del operador identidad es la matriz identidad. ¿Cuál es la matriz del operador $A : \mathbb{R}^n \mapsto \mathbb{R}^n$ tal que $Ae_1 = e_2$, $Ae_2 = e_1$, $Ae_k = e_k$, $k = 3, \dots, n$, donde $(e_k)_{k=1}^n$ es la base canónica de $\mathbb{R}^n$?*

Nótese que lo anterior se puede generalizar al caso de dimensión infinita, solo que en este caso la matriz sería una matriz infinita

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n} & \cdots \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots & a_{2,n} & \cdots \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdots & a_{3,n} & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \cdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \cdots & a_{n,n} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Ejercicio 6.1.10 *Usando la base de Schauder de ℓ^2 (ver Ejemplo 4.2.16) encuentra las matrices del operador S y S^* definidos en (6.1.2) y (6.1.3), respectivamente.*

Ejercicio 6.1.11 *¿Cuál es la matriz del operador $A : \ell^2 \mapsto \ell^2$ tal que $Ae_1 = e_2$, $Ae_2 = e_1$, $Ae_k = e_k$, $k = 3, 4, \dots$, $(e_k)_k$ en la base de Schauder definida en el Ejemplo 4.2.16? ¿Es dicho operador acotado?*

6.2. Operadores autoadjuntos

Veamos algunas propiedades de los operadores autoadjuntos.

Lema 6.2.1 *Sea $A : \mathbb{E} \mapsto \mathbb{E}$, con $\mathbb{E}$ un espacio euclídeo complejo. Entonces, si $\langle Ax, x\rangle = 0$ para todo $x \in \mathbb{E}$, A es el operador nulo $A = \Theta$.*

Demostración: Dados $x, y \in \mathbb{E}$ cualesquiera y $\alpha \in \mathbb{C}$ se tiene

$$\langle A(x+\alpha y), x+\alpha y\rangle = \langle Ax, x\rangle + |\alpha|^2\langle Ay, y\rangle + \overline{\alpha}\langle Ax, y\rangle + \alpha\langle Ay, x\rangle. \quad (6.2.1)$$

Por hipótesis $\langle Av, v\rangle = 0$, $v = x + \alpha y$, luego, de (6.2.1) obtenemos

$$0 = \langle Av, v\rangle = \langle A(x + \alpha y), \alpha x + \alpha y\rangle = \overline{\alpha}\langle Ax, y\rangle + \alpha\langle Ay, x\rangle.$$

Para $\alpha = 1$ tenemos $\langle Ax, y\rangle+\langle Ay, x\rangle = 0$, y para $\alpha = -i$, $\langle Ax, y\rangle-\langle Ay, x\rangle = 0$. Sumando ambas obtenemos $\langle Ax, y\rangle = 0$ para todos $x, y \in \mathbb{E}$, luego $Ax = 0$ para todo $x \in \mathbb{E}$ (punto 4 del Ejercicio 5.1.2) i.e. $A = \Theta$. ■

Teorema 6.2.2 *Sea $\mathbb{H}$ un espacio de Hilbert y sea $A : \mathbb{H} \mapsto \mathbb{H}$, un operador lineal y acotado. Entonces*

1. *Si A es autoadjunto, entonces para todo $x \in \mathbb{H}$, $\langle Ax, x\rangle \in \mathbb{R}$.*
2. *Si $\mathbb{H}$ es complejo y para todo $x \in \mathbb{H}$, $\langle Ax, x\rangle \in \mathbb{R}$, entonces A es autoadjunto.*

Demostración: Para probar el primer punto tenemos

$$\overline{\langle Ax, x\rangle} = \langle x, Ax\rangle = \langle Ax, x\rangle \;\Rightarrow\; \langle Ax, x\rangle \in \mathbb{R},$$

donde la primera igualdad es por la propiedad de simetría del producto escalar y la segunda por ser A autoadjunto. Para la segunda afirmación usamos que para todo $x \in \mathbb{H}$, $\langle Ax, x\rangle \in \mathbb{R}$, entonces, para todo $x \in \mathbb{H}$,

$$\langle Ax, x\rangle = \overline{\langle Ax, x\rangle} = \langle x, Ax\rangle = \langle A^*x, x\rangle \;\Rightarrow\; \langle (A - A^*)x, x\rangle = 0,$$

pero, por el lema anterior, $A - A^* = 0$, luego $A = A^*$. ■

El teorema anterior caracteriza los operadores autoadjuntos en espacios de Hilbert complejos.

Ejercicio 6.2.3 *Probar, usando las definiciones correspondientes, las siguientes afirmaciones:*

1. *Sean A y B operadores autoadjuntos de $\mathbb{H}$ en $\mathbb{H}$, y sea $\alpha \in \mathbb{R}$. Entonces $A+B$ y αA también son autoadjuntos.*
2. *Sea $A : \mathbb{H} \mapsto \mathbb{H}$ un operador y A^* su adjunto. Entonces $A+A^*$ y A^*A son autoadjuntos.*
3. *El espacio de todos los operadores $A : \mathbb{H} \mapsto \mathbb{H}$ normales y acotados es un subespacio vectorial de $\mathcal{B}(\mathbb{H}, \mathbb{H})$.*

Proposición 6.2.4 *El producto AB de dos operadores autoadjuntos A y B es autoadjunto si y solo si A y B conmutan, i.e., $AB = BA$. Si A es autoadjunto A^n, $n = 2, 3, \ldots$ también lo es.*

Demostración: Del punto 7 de la Proposición 6.1.4 se sigue que $(AB)^* = B^*A^* = BA$, luego $(AB)^* = (AB)$ si y solo si $AB = BA$. La segunda parte es inmediata usando inducción y teniendo en cuenta A conmuta consigo mismo. ■

Teorema 6.2.5 *Sea $(T_n)_n$ una sucesión de operadores lineales acotados y autoadjuntos $T_n : \mathbb{H} \mapsto \mathbb{H}$, $\mathbb{H}$ espacio de Hilbert. Si $T_n \to T$ cuando $n \to \infty$, entonces T es un operador lineal acotado y autoadjunto en $\mathbb{H}$.*

Demostración: Como T es el límite de la sucesión T_n, T es lineal y acotado. Ello se sigue de que el espacio $\mathcal{B}(\mathbb{H}, \mathbb{H})$ es un espacio de Banach (ver Ejercicio 4.4.20). No obstante, por completitud, vamos a dar una prueba directa de la acotación. Como la sucesión $(T_n)_n$ es convergente, entonces es acotada, luego existe $M \geq 0$ tal que para todo $n \in \mathbb{N}$, $\|T_n\| \leq M$. Luego, $\|T_n x\| \leq M\|x\|$, para todo $x \in \mathbb{H}$. Entonces $\|Tx\| = \lim_{n\to\infty} \|T_n x\| \leq M\|x\|$, de donde se sigue la acotación de T. La linealidad se sigue de la linealidad de T_n tomando límites cuando $n \to \infty$ (ver Ejercicio 4.4.20).

Como T es lineal y acotado en un espacio de Hilbert, el Teorema 6.1.3 nos asegura que existe su adjunto T^*. Probemos[2] ahora que T es autoad-

[2]Nótese que si consideramos T_n como elementos del espacio de Banach $\mathbb{B}(\mathbb{H}, \mathbb{H})$, entonces $T_n \to T$ implica que T es único (ver Problema 3.10), luego como $T_n^* = T_n$ y $T_n^* \to T^*$, se sigue que $T = T^*$. Otra forma de probarlo es usar que para todo $n \in \mathbb{N}$, $\langle T_n x, y\rangle = \langle x, T_n y\rangle$ y tomar límites cuando $n \to \infty$.

junto, i.e., que $T = T^*$. Comencemos notando que

$$\|T_n^* - T^*\| = \|(T_n - T)^*\| = \|T_n - T\|.$$

La primera igualdad es consecuencia del punto 2 de la Proposición 6.1.4 y la segunda de (6.1.4). En particular lo anterior nos indica que si $T_n \to T$, entonces $T_n^* \to T^*$. Usando la desigualdad triangular

$$\|T - T^*\| \leq \|T - T_n\| + \underbrace{\|T_n - T_n^*\|}_{=0} + \|T_n^* - T^*\| = 2\|T - T_n\| \to 0,$$

luego $\|T - T^*\| = 0$ de donde se sigue que $T = T^*$. ■

Teorema 6.2.6 *Sea $A : \mathbb{H} \mapsto \mathbb{H}$ un operador acotado y autoadjunto en un espacio de Hilbert $\mathbb{H}$. Entonces $\|A\| = \sup_{\|x\|=1} |\langle Ax, x\rangle|$.*

Demostración: Sea $M = \sup_{\|x\|=1} |\langle Ax, x\rangle|$. Tomemos $x \neq 0$, $\|x\| = 1$. Entonces

$$|\langle Ax, x\rangle| \leq \|Ax\|\|x\| \leq \|A\|\|x\|^2 = \|A\| \quad \Rightarrow \quad M \leq \|A\|,$$

donde la implicación se sigue al tomar el supremos sobre todos los $x \in \mathbb{H}$, con $\|x\| = 1$. Nótese que

$$\sup_{\|x\|=1} |\langle Ax, x\rangle| = \sup_{\|x\|\neq 0} \frac{|\langle Ax, x\rangle|}{\|x\|^2} \quad \Rightarrow \quad |\langle Ax, x\rangle| \leq M\|x\|^2, \quad \forall x \neq 0.$$

Sea $x \in \mathbb{H}$ tal que $\|x\| = 1$ y tomemos $y = Ax/\|Ax\|$ ($Ax \neq 0$). Claramente para esta elección de los vectores x e y se tiene $\langle Ax, y\rangle = \|Ax\| = |\langle Ax, y\rangle| = |\Re(\langle Ax, y\rangle)|$. Por otro lado, de (6.2.1), se sigue la identidad

$$\Re(\langle Ax, y\rangle) = \frac{1}{4}\Big(\langle A(x+y), x+y\rangle - \langle A(x-y), x-y\rangle\Big), \quad \forall x, y \in \mathbb{H}. \tag{6.2.2}$$

Luego, para dicha elección, de x e y obtenemos que

$$\|Ax\| = |\langle Ax, y\rangle| \leq \frac{1}{4}\big(|\langle A(x+y), x+y\rangle| + |\langle A(x-y), x-y\rangle|\big).$$

Ahora bien, de la desigualdad anterior, usando que $|\langle Ax, x\rangle| \leq M\|x\|^2$, para todo $x \neq 0$ y la ley del paralelogramo (5.1.3) se tiene ($\|x\| = \|y\| = 1$)

$$\|Ax\| = |\Re(\langle Ax, y\rangle)| \leq \frac{M}{4}\left(\|x+y\|^2 + \|x-y\|^2\right) = \frac{M}{2}\left(\|x\|^2 + \|y\|^2\right) = M.$$

Luego, para todo x con $\|x\| = 1$, $\|Ax\| \leq M$, i.e., $\|A\| \leq M$, por tanto $\|A\| = M$, como se quería probar. ■

Nótese que el Teorema 6.2.6 es válido para cualquier espacio de Hilbert, ya sea real o complejo.

Corolario 6.2.7 *Sean A y B dos operadores autoadjuntos en un espacio de Hilbert $\mathbb{H}$. Si $\langle Ax, x\rangle = \langle Bx, x\rangle$ para todo $x \in \mathbb{H}$, entonces $A = B$. En particular, si $\langle Ax, x\rangle = 0$ para todo $x \in \mathbb{H}$, entonces A es el operador nulo.*

Demostración: Si $\mathbb{H}$ es un espacio de Hilbert complejo podemos usar el Lema 6.2.1. Probémoslo usando en Teorema 6.2.6 válido para cualquier espacio de Hilbert. Así

$$\|A - B\| = \sup_{\|x\|=1} |\langle (A-B)x, x\rangle| = \sup_{\|x\|=1} |\langle Ax, x\rangle - \langle Bx, x\rangle| = \sup_{\|x\|=1} |0| = 0,$$

luego $A = B$. Si $\langle Ax, x\rangle = 0$ para todo $x \in \mathbb{H}$, entonces $\sup_{\|x\|=1} |\langle Ax, x\rangle| = 0$, luego $\|A\| = 0$, i.e., A es el operador nulo del Ejemplo 4.4.3. ■

Ejercicio 6.2.8 *Sea $A : \mathbb{H} \mapsto \mathbb{H}$ un operador acotado y autoadjunto. Prueba que $\|A^2\| = \|A\|^2$.*

Del punto 5 de la Proposición 6.1.4 tenemos que $\|A^*A\| = \|A\|^2$ y como A es autoadjunto $A = A^*$, por tanto $A^*A = A^2$, luego $\|A^2\| = \|A\|^2$. ■

Para terminar este apartado veamos un ejemplo muy útil, como veremos más adelante, a la hora de construir distintos tipos de operadores lineales y acotados en un espacio de Hilbert $\mathbb{H}$.

Ejemplo 6.2.9 *Sea $(\phi_n)_n$ una base ortonormal completa de un espacio de Hilbert $\mathbb{H}$. Sea $(\mu_n)_n \in \mathbb{C}$ una sucesión (finita o infinita) acotada, sea $m = \sup_n |\mu_n|$. Entonces existe un único operador lineal y acotado $T : \mathbb{H} \mapsto \mathbb{H}$, tal que, para todo n,*

$$T\phi_n = \mu_n \phi_n. \tag{6.2.3}$$

Vamos a probarlo para el caso de dimensión infinita puesto que el caso de dimensión finita se puede considerar un caso particular de este cuando $\mu_n = 0$ para todos $n > N \in \mathbb{N}$.

Por el Teorema 5.2.5 sabemos que todo $x \in \mathbb{H}$ admite una única representación en serie de Fourier en la base $(\phi_n)_n$ y además que

$$x = \sum_{k=1}^{\infty} c_k \phi_k, \qquad c_k = \langle x, \phi_k\rangle, \qquad \sum_{n=1}^{\infty} |c_n|^2 < +\infty. \tag{6.2.4}$$

Para cada $x \in \mathbb{H}$ definamos la serie

$$S = \sum_{k=1}^{\infty} \mu_k c_k \phi_k.$$

Probemos que dicha serie converge para cada $x \in \mathbb{H}$. Sean S_n las sumas parciales de S

$$S_n = \sum_{k=1}^{n} \mu_k c_k \phi_k,$$

y sea $j > n$. Entonces,

$$\|S_n - S_j\|^2 = \sum_{k=n+1}^{j} |\mu_n|^2 |c_k|^2 \le m^2 \sum_{k=n+1}^{j} |c_k|^2,$$

que se puede hacer tan pequeño como se quiera si elegimos $j > n > N \in \mathbb{N}$, con N suficientemente grande (¿por qué?). Luego S_n es de Cauchy y como $\mathbb{H}$ es completo, $S_n \to S$. Por tanto podemos definir el operador

$$T : \mathbb{H} \mapsto \mathbb{H}, \quad \text{tal que} \quad \forall x \in \mathbb{H}, \quad Tx = S = \sum_{k=1}^{\infty} \mu_k \langle x, \phi_k \rangle \phi_k. \tag{6.2.5}$$

Además, podemos definir para cada $n \in \mathbb{N}$ la sucesión de operadores $T_n : \mathbb{H} \mapsto \mathbb{H}$, $T_n x = S_n$. Está claro que T_n es lineal (¿por qué?). Probemos que T_n es acotado para cada n y que $\|T_n\| \le m$. En efecto,

$$\|T_n x\|^2 = \sum_{k=1}^{n} |c_k|^2 |\mu_k|^2 \le m^2 \sum_{k=1}^{n} |c_k|^2 \le m^2 \sum_{k=1}^{\infty} |c_k|^2 = m^2 \|x\|^2. \tag{6.2.6}$$

Además, para cada $x \in \mathbb{H}$, $T_n x \to Tx$ pues

$$\|T_n x - Tx\|^2 = \left\| \sum_{k=n+1}^{\infty} \mu_k c_k \phi_k \right\|^2 = \sum_{k=n+1}^{\infty} |\mu_k|^2 |c_k|^2 \le m \sum_{k=n+1}^{\infty} |c_k|^2 \to 0.$$

El operador T así definido es lineal y acotado. La linealidad de T se sigue de la linealidad de los T_n (ver, e.g., la prueba del Ejercicio 4.4.20). La acotación de T se sigue de tomar el límite cuando $n \to \infty$ en (6.2.6) y usar la continuidad de la norma.

Nota: Otra forma de probar la acotación es usando el Teorema de Banach-Steinhaus pues como la sucesión $\|T_n x\|$ es acotada para cada x, entonces existe un c tal que para todo n, $\|T_n\| \leq c$. Tomando $n \to \infty$, se tiene que $\|T\| \leq c$, pues, por el Ejercicio 4.4.20, $\mathcal{B}(\mathbb{H}, \mathbb{H})$ es un espacio de Banach. ■

Nótese que tomando el límite cuando $n \to \infty$ en $\|T_n\| \leq m$ se sigue que $\|T\| \leq m$. Por otro lado,

$$\|T\| = \sup_{\|x\|=1} \|Tx\| \geq \|T\phi_n\| = \|\mu_n \phi_n\| = |\mu_n| \;\Rightarrow\; \forall n \quad \|T\| \geq |\mu_n|.$$

Tomando supremos en n en la desigualdad anterior se deduce que $\|T\| \geq m$, luego $\|T\| = m = \sup_n |\mu_n|$.

Nótese que como caso particular de (6.2.5) se tiene (6.2.3) y que

$$T\left(\sum_{k=1}^{\infty} c_k \phi_k\right) = \sum_{k=1}^{\infty} \mu_k c_k \phi_k. \tag{6.2.7}$$

Además, como T es acotado, existe T^*. Usando que, para todo $x \in \mathbb{H}$, tenemos, por un lado,

$$T^* x = y^* = \sum_{k=1}^{\infty} \nu_k \phi_k \;\Rightarrow\; \langle T^* x, \phi_n \rangle = \nu_n,$$

y, por el otro,

$$\langle T^* x, \phi_n \rangle = \langle x, T\phi_n \rangle = \left\langle \sum_{k=1}^{\infty} c_k \phi_k, \mu_n \phi_n \right\rangle = \overline{\mu_n}\, c_n \;\Rightarrow\; \nu_n = \overline{\mu_n}\, c_n,$$

de donde se sigue que

$$T^* x = T^* \left(\sum_{k=1}^{\infty} c_k \phi_k\right) = \sum_{k=1}^{\infty} \overline{\mu_k}\, c_k \phi_k, \tag{6.2.8}$$

y que, para todo n, $T^* \phi_n = \overline{\mu_n}\, \phi_n$. Además, del Teorema 6.1.3, T^* es acotado y $\|T^*\| = \|T\|$.

Finalmente, como $T^*T\phi_n = |\mu_n|^2\phi_n = TT^*\phi_n$, entonces, por el problema[a] 6.6 se sigue que $T^*T = TT^*$, luego nuestro operador T es normal. Está claro que si la sucesión $(\mu_n)_n$ es real, entonces T es autoadjunto. ∎

[a] Se puede comprobar directamente, usando (6.2.7) y (6.2.8), que para todo $x \in \mathbb{H}$, $T^*Tx = TT^*x$, con x definido mediante la serie (6.2.4).

6.3. Inverso de un operador

Veamos a continuación el problema de invertir un operador lineal A en un espacio normado en general y en un espacio euclídeo en particular. En el caso de dimensión finita este problema es relativamente sencillo. De hecho que exista el inverso A^{-1} de A es equivalente a que A sea inyectivo, sobreyectivo, que exista un B tal que $AB = I$, o que $BA = I$, donde I denota al operador identidad (ver Ejemplo 4.4.2). Además, como ya vimos en el punto 4 del Teorema 4.4.9, en el caso de dimensión finita existirá el inverso de A si $\dim \mathbb{X} = \dim \mathbb{Y}$. Además, si nos restringimos a espacios normados de dimensión finita, A es acotado y si es invertible su inversa también es acotada.

En dimensión infinita la situación es mucho más complicada. Los siguientes ejemplos muestran una variedad de casos.

Ejemplo 1. El operador *operador desplazamiento* $S_{\curvearrowright}$ definido por (6.1.2) es acotado e inyectivo (la ecuación $S_{\curvearrowright}x = y$ tiene una única solución, o bien no tiene), luego tiene inversa según la Definición 3.3.7, sin embargo no es sobreyectivo (la ecuación $S_{\curvearrowright}x = y$ no tiene solución, e.g. elegir $y = (y_1, y_2, \dots)$, $y_1 \neq 0$). Por otro lado, el operador $S_{\curvearrowleft}$ definido en (6.1.3) es sobreyectivo pero no inyectivo (luego no tiene inverso según la Definición 3.3.7). Además, es fácil comprobar que $S_{\curvearrowleft}S_{\curvearrowright} = I$ mientras que $S_{\curvearrowright}S_{\curvearrowleft} \neq I$.

Ejemplo 2. Sea el operador multiplicación por t del Ejemplo 4.4.6, $M : L^2_{[0,1]} \mapsto L^2_{[0,1]}$ (que es un caso particular del operador M del Ejemplo 6.1.2) con $g(t) = t$). Este operador es acotado y además inyectivo. En efecto, $Mx(t) = 0$ si y solo si $x(t) = 0$ en casi todo punto, sin embargo no es sobreyectivo pues, por ejemplo, $Mx(t) = 1$ implica que $x(t) = 1/t$

que no es de $L^2_{[0,1]}$.

Ejemplo 3. Sea el operador de Volterra $V : C^\infty_{[0,1]} \mapsto C([0,1])$,

$$Vx := (Vx)(t) = \int_0^t x(\tau)d\tau, \quad t \in (0,1). \tag{6.3.1}$$

Este operador es acotado y tiene inverso (ver Teorema 4.4.9) pues $Vx = 0$ implica $x(t) = 0$. Además está claro que la ecuación $(Vx)(t) = y(t)$ siempre tiene solución $x(t) = y'(t)$, luego el inverso de V es el operador derivada D que sabemos que no es acotado. Nótese que $\mathcal{I}(V) \neq C([0,1])$ pues si $y \in C([0,1])$ tal que $y(0) \neq 0$ entonces $y \notin \mathcal{I}(V)$.

Ejemplo 4. Sea el operador $A : \ell^2 \mapsto \ell^2$,

$$A(x_1, x_2, \dots, x_n, \dots) = (x_1, x_2/2, \dots, x_n/n \dots). \tag{6.3.2}$$

Está claro que A es acotado pues

$$\|Ax\|^2 \leq \sum_{k=1}^\infty \frac{|x_k|^2}{k^2} \leq \sum_{k=1}^\infty |x_k|^2 = \|x\|^2.$$

Por otro lado, se puede comprobar directamente que el siguiente operador B definido sobre ℓ^2

$$B(x_1, x_2, \dots, x_n, \dots) = (x_1, 2x_2, \dots, nx_n \dots),$$

es tal que $ABx = BAx = x$ para todo $x \in \ell^2$. Luego $B = A^{-1}$, pero está claro que B no es acotado (basta tomar como sucesión $(x_n)_n$, la base de Schauder definida en el Ejemplo 4.2.16). Por otro lado, $I(A) \neq \ell^2$ (no es sobreyectivo), pues si elegimos $y = (1/\sqrt{n^3})_n \in \ell^2$, entonces $Ax = y$ para $x = (1/\sqrt{n})_n$ que no es de ℓ^2. Así A es acotado e invertible pero su inverso no es acotado.

Los ejemplos anteriores nos llevan a la siguiente pregunta: ¿qué definición de inverso de un operador adoptar?

Una opción es usar la propia Definición 3.3.7 que ya vimos en el contexto de los espacios métricos, pero es poco práctica. Es mucho más común una definición inspirada en la propiedad discutida en la Nota 3.3.9. Así, podríamos utilizar la siguiente definición:

Definición: Sea el operador lineal $A : \mathbb{X} \mapsto \mathbb{Y}$, $\mathbb{X}$ e $\mathbb{Y}$ espacios de Banach. Diremos que A es invertible si existe un operador B, $A : \mathbb{Y} \mapsto \mathbb{X}$ tal que $AB = I_{\mathbb{Y}}$, $BA = I_{\mathbb{X}}$.

Nótese que según esta definición el operador desplazamiento $S_{\frown}$ definido en (6.1.2), que es invertible según la Definición 3.3.7 cumple con $S_{\frown}S_{\curvearrowright} = I$ pero $S_{\curvearrowright}S_{\frown} \neq I$, por tanto no tendría inversa. Por otro lado, el operador del Ejemplo 4 cumple con la definición anterior, pero como ya vimos su inverso es un operador no acotado.

Sin embargo hay un resultado muy importante cuya prueba por el momento omitiremos[3] conocido como el *Teorema de Banach de las aplicaciones inversas acotadas* que establece lo siguiente:

Teorema: *Sea $A \in \mathcal{B}(\mathbb{X}, \mathbb{Y})$, con $\mathbb{X}, \mathbb{Y}$ espacios de Banach, tal que el núcleo de A $\mathcal{N}(A) = \{0\}$, y la imagen de A $\mathcal{I}(A) = \mathbb{Y}$, entonces $A^{-1} \in \mathcal{B}(\mathbb{Y}, \mathbb{X})$.*

Dado que estamos interesados en los operadores lineales y acotados, lo anterior nos conduce a redefinir la inversa de un operador $A \in \mathcal{B}(\mathbb{X}, \mathbb{Y})$ de la siguiente forma:

Definición 6.3.1 *Sea el operador $A \in \mathcal{B}(\mathbb{X}, \mathbb{Y})$, $\mathbb{X}$ e $\mathbb{Y}$ espacios de Banach. Diremos que A es invertible si existe un operador $B \in \mathcal{B}(\mathbb{Y}, \mathbb{X})$, tal que $AB = I_{\mathbb{Y}}$, $BA = I_{\mathbb{X}}$.*

Nótese que según esta definición el operador (6.3.2) del Ejemplo 4 no es invertible.

En adelante entenderemos que un operador acotado es invertible si existe su inversa y esta es acotada. Veamos a continuación varias propiedades de los operadores invertibles.

Teorema 6.3.2 (Existencia del operador inverso) *Sea el operador lineal $A : \mathbb{X} \mapsto \mathbb{X}$, $\mathbb{X}$ espacio de Banach. Si $\|A\| < 1$, entonces $I - A$ es invertible*

[3] Ver el Teorema 7.3.4 de la página 292. Como veremos dicho teorema es una consecuencia del Teorema de Banach de la aplicación abierta, que a su vez se sigue del Teorema de Baire 3.6.1 que ya vimos al final del capítulo 3.

y se tiene la igualdad (serie de Neumann)

$$(I-A)^{-1} = \sum_{k=0}^{\infty} A^k, \;\; \textit{donde } A^0 := I \;\; \textit{y} \;\; \|(I-A)^{-1}\| \le \frac{1}{1-\|A\|}.$$

Demostración: Sea la sucesión de operadores $(A_n)_n$, definida por $A_n x := (I+A+A^2+\cdots+A^n)x$, $x \in \mathbb{X}$ (cualquiera). *Obviamente* A_n es un operador acotado (probarlo como ejercicio). Probemos que $(A_n x)_n$ es una sucesión de Cauchy. Para ello calculamos

$$\begin{aligned}
\|A_{n+p}x - A_n x\| &= \|A^{n+1}x+\cdots+A^{n+p}x\| \le \|A^{n+1}x\|+\cdots+\|A^{n+p}x\| \\
&\le (\|A^{n+1}\| + \cdots + \|A^{n+p}\|)\|x\| \le (\|A\|^{n+1} + \cdots + \|A\|^{n+p})\|x\| \\
&\le (\|A\|^{n+1} + \cdots + \|A\|^{n+p} + \cdots)\|x\| \le \frac{\|A\|^{n+1}}{1-\|A\|}\|x\| \overset{n\to\infty}{\longrightarrow} 0. \qquad (6.3.3)
\end{aligned}$$

Como $\mathbb{X}$ es completo (es de Banach) entonces, para cada $x \in \mathbb{X}$ la sucesión $A_n x$ converge a un $y \in \mathbb{X}$ que denotaremos por $y = Tx$. Así, tendremos el operador $T : \mathbb{X} \mapsto \mathbb{X}$ bien definido. Como A es lineal, A_n lo será y, por tanto, T también. Para ello basta tomar el límite cuando $n \to \infty$ en la igualdad $A_n(\alpha x + \beta y) = \alpha A_n x + \beta A_n y$, para todos $\alpha, \beta \in \mathbb{K}$ y $x, y \in \mathbb{X}$.

Probemos ahora que T es acotado. Para ello tomamos el límite $p \to \infty$ en la desigualdad (6.3.3) y usamos que $A_{n+p}x \to Tx$. Así,

$$\|Tx - A_n x\| \le \|x\| \frac{\|A\|^{n+1}}{1-\|A\|}, \qquad (6.3.4)$$

de donde se sigue, tomando el supremo cuando $\|x\| = 1$, que

$$\|T - A_n\| \le \frac{\|A\|^{n+1}}{1-\|A\|} \overset{n\to\infty}{\longrightarrow} 0 \quad \Rightarrow \quad \lim_{n\to\infty} A_n = T.$$

Nótese que de la desigualdad anterior se tiene que $T - A_n$ es un operador lineal acotado, luego T lo será (¿por qué?). Probemos ahora que $T = (I-A)^{-1}$. Ante todo notemos que $I - A$ es un operador lineal y acotado (por tanto continuo, ver Teorema 4.4.18). Calculemos el producto

$$\begin{aligned}
(I-A)Tx &= (I-A)\lim_{n\to\infty} A_n x = \lim_{n\to\infty}(I-A)A_n x = \lim_{n\to\infty}(A_n x - A\cdot A_n x) \\
&= \lim_{n\to\infty}(x - A^{n+1}x).
\end{aligned}$$

Teniendo en cuenta que $\|A\| < 1$ y que $\|A^{n+1}x\| \leq \|A\|^{n+1}\|x\| \overset{n\to\infty}{\longrightarrow} 0$, se sigue que $A^{n+1}x \overset{n\to\infty}{\longrightarrow} 0$ y, por tanto, $(I-A)Tx = x$ para todo x. La prueba de que $T(I-A)x = x$ para todo x es análoga y la omitiremos. Finalmente, para la cota de $\|(I-A)^{-1}\|$ usamos que $A_n \to (I-A)^{-1}$ y que $\|A^k\| \leq \|A\|^k$ (ver Problema 4.13), luego

$$\begin{aligned}\|A_n\| &= \|I + A + \cdots + A^n\| \leq 1 + \|A\| + \|A^2\| + \cdots + \|A^n\| \\ &\leq 1 + \|A\| + \|A\|^2 + \cdots + \|A\|^n \leq \frac{1}{1-\|A\|},\end{aligned}$$

de donde, tomando el límite cuando $n \to \infty$, se tiene el resultado. ∎

Nota: Usando el Ejercicio 4.4.20 se puede simplificar la demostración. En primer lugar, A es acotado, luego todas sus potencias lo son (ver Problema 4.13), y por tanto A_n lo es para todo n. Como $\|A_{n+p}x - A_nx\|/\|x\| \to 0$ cuando $n \to \infty$, se tiene que $\|A_{n+p} - A_n\|$ es de Cauchy, luego como $\mathcal{B}(\mathbb{X},\mathbb{X})$ es un espacio de Banach, entonces $A_n \to T = \sum_{k=0}^{\infty} A^k$, $T \in \mathcal{B}(\mathbb{X},\mathbb{X})$, i.e., T es lineal y acotado. Las demostraciones de que $T = (1-A)^{-1}$, así como la estimación de la norma son idénticas.

El teorema anterior admite una generalización que nos será de gran utilidad más adelante.

Teorema 6.3.3 *Sea el operador lineal y acotado $A : \mathbb{X} \mapsto \mathbb{X}$, $\mathbb{X}$ espacio de Banach. Si $\|A\| < |\lambda|$, $\lambda \in \mathbb{C}$, entonces el operador $(\lambda I - A)$ es invertible y*

$$A_\lambda := (\lambda I - A)^{-1} = \sum_{k=0}^{\infty} \frac{A^k}{\lambda^{k+1}}, \quad \textit{y se cumple que} \quad \|A_\lambda\| \leq \frac{1}{|\lambda| - \|A\|}.$$

Demostración: Asumimos que $\lambda \neq 0$. Entonces, como $\|A\| < |\lambda|$, el operador $\widetilde{A} = A/\lambda$ es tal que $\|\widetilde{A}\| < 1$, y por el Teorema 6.3.2, $1 - \widetilde{A}$ es invertible. Además, claramente el operador λI es invertible y, por tanto, se tiene

$$A_\lambda = \left[\lambda\left(I - \frac{A}{\lambda}\right)\right]^{-1} = \left[(\lambda I)\left(I - \frac{A}{\lambda}\right)\right]^{-1} = \left(I - \frac{A}{\lambda}\right)^{-1}(\lambda I)^{-1}.$$

Luego,

$$A_\lambda = (\lambda I - A)^{-1} = \lambda^{-1}\left(I - \frac{A}{\lambda}\right)^{-1} = \lambda^{-1}\left(I - \widetilde{A}\right)^{-1}.$$

Como $\|A\| < |\lambda|$, entonces $\|\widetilde{A}\| < 1$. Aplicando el Teorema 6.3.2 a $\widetilde{A}$ obtenemos la expresión para A_λ

$$(I - \widetilde{A})^{-1} = \sum_{k=0}^{\infty} \widetilde{A}^k = \sum_{k=0}^{\infty} \frac{A^k}{\lambda^k} \quad \Rightarrow \quad (\lambda I - A)^{-1} = \sum_{k=0}^{\infty} \frac{A^k}{\lambda^{k+1}}.$$

Para probar la desigualdad restante usamos que

$$\|A_\lambda\| = \left\| \sum_{k=0}^{\infty} \frac{A^k}{\lambda^{k+1}} \right\| \leq \frac{1}{|\lambda|} \sum_{k=0}^{\infty} \frac{\|A^k\|}{|\lambda|^k} \leq \frac{1}{|\lambda|} \sum_{k=0}^{\infty} \frac{\|A\|^k}{|\lambda|^k} = \frac{1}{|\lambda|} \frac{1}{1 - \|A\|/|\lambda|},$$

de donde se sigue el resultado. ■

Ejercicio 6.3.4 *Modifica la demostración del Teorema 6.3.2 para probar directamente el teorema anterior.*

Definición 6.3.5 *Sea $A : \mathbb{X} \mapsto \mathbb{X}$ un operador de $\mathbb{X}$ en $\mathbb{X}$, siendo $\mathbb{X}$ un espacio normado. El conjunto de de todos los operadores lineales lo denotaremos por $\mathcal{L}(\mathbb{X})$ y al conjunto de todos los operadores lineales y acotados lo denotaremos por $\mathcal{B}(\mathbb{X})$.*

Un corolario trivial del resultado probado en el Ejercicio 4.4.20 es el siguiente teorema:

Teorema 6.3.6 *El espacio $\mathcal{B}(\mathbb{X})$, $\mathbb{X}$ espacio de Banach, es un espacio de Banach respecto a la norma de los operadores.*

Ejercicio 6.3.7 *Adapta la prueba del Teorema 6.3.2 para probar que toda sucesión de Cauchy de operadores $T_n \in \mathcal{B}(\mathbb{X})$, $\mathbb{X}$ espacio de Banach, es convergente, i.e., se tiene el Teorema 6.3.6.*

En efecto, si T_n es de Cauchy entonces para todo $\varepsilon > 0$, existe un $N \in \mathbb{N}$ tal que para todo $n \in \mathbb{N}$, $n > N$ y cualquiera sea $p \in \mathbb{N}$, $\|T_{n+p} - T_n\| < \varepsilon$. Luego la sucesión $X_n := T_n x$ es de Cauchy

$$\|T_{n+p}x - T_n x\| \leq \|T_{n+p} - T_n\| \|x\| \leq \varepsilon \|x\| \Rightarrow \|T_{n+p}x - T_n x\| \overset{n\to\infty}{\longrightarrow} 0.$$

Como $\mathbb{X}$ es completo $T_n x$ tendrá un límite y para cada x, i.e., podemos definir el operador $T : \mathbb{X} \mapsto \mathbb{X}$ por $Tx = y$. De la misma forma que en la prueba del Teorema 6.3.2 se sigue que T es lineal. Para probar que es acotado usamos la desigualdad anterior:

$$\|T_{n+p}x - T_nx\| \leq \|T_{n+p} - T_n\|\|x\| \Rightarrow \frac{\|T_{n+p}x - T_nx\|}{\|x\|} \leq \varepsilon \Rightarrow \frac{\|Tx - T_nx\|}{\|x\|} \leq \varepsilon,$$

donde hemos tomando el límite $p \to \infty$ y usado la continuidad de la norma. Luego, $T - T_n$ es acotado y, por tanto, T lo es (T_n es acotado). Además, de lo anterior se sigue también, tomando el supremo en x, que $\|T - T_n\| \leq \varepsilon$ para todo $n > N$, i.e., $T_n \overset{n\to\infty}{\longrightarrow} T$. ■

Una aplicación directa del Teorema 6.3.2 es el siguiente resultado:

Teorema 6.3.8 *Sea $\mathbb{X}$ un espacio de Banach. El conjunto $\mathcal{E} \subset \mathcal{B}(\mathbb{X})$ de los operadores invertibles en $\mathbb{X}$ es abierto en $\mathcal{B}(\mathbb{X})$.*

Demostración: Sea $A \in \mathcal{B}(\mathbb{X})$ un operador invertible. Definamos la bola $B(A, 1/\|A^{-1}\|)$. Cualquiera sea $B \in B(A, 1/\|A^{-1}\|)$ tendremos que

$$\|B - A\| < \frac{1}{\|A^{-1}\|} \quad \Rightarrow \quad \|(B - A)A^{-1}\| \leq \|B - A\|\|A^{-1}\| < 1,$$

luego el operador $I - (A - B)A^{-1} = BA^{-1}$ es invertible, por tanto el operador $B = (BA^{-1})A$ lo será[4]. Es decir, todo operador $B \in B(A, 1/\|A^{-1}\|)$ es invertible y, por tanto, dado un operador invertible cualquiera $A \in \mathcal{E}$ existe una bola $B(A, 1/\|A^{-1}\|) \subset \mathcal{E}$ centrada en A, luego $\mathcal{E}$ es abierto. ■

6.4. Operadores compactos

Definición 6.4.1 *Un operador lineal $A : \mathbb{X} \mapsto \mathbb{Y}$, $\mathbb{X}$ e $\mathbb{Y}$ espacios de Banach es compacto si para toda sucesión acotada $(x_n)_n$ de $\mathbb{X}$, la sucesión $(Ax_n)_n$ de $\mathbb{Y}$ tiene una subsucesión convergente.*

Nótese que si A es compacto, A es acotado pues en caso contrario existiría una sucesión acotada $(x_n)_n$ tal que $\|Ax_n\| \overset{n\to\infty}{\longrightarrow} \infty$ y entonces la sucesión $(Ax_n)_n$ no tendría una subsucesión convergente. Así pues, tenemos el siguiente teorema:

[4]En efecto, si A y B son invertibles, entonces AB es invertible y $(AB)^{-1} = B^{-1}A^{-1}$, como se puede comprobar directamente de la Definición 6.3.1.

Teorema 6.4.2 *Todo operador compacto es acotado.*

El recíproco no es cierto.

Ejercicio 6.4.3 *Prueba que el operador identidad $I : \mathbb{H} \mapsto \mathbb{H}$, $\mathbb{H}$ espacio de Hilbert de dimensión infinita no es compacto.*

Escojamos una sucesión ortonormal cualquiera $(\phi_n)_n$ en $\mathbb{H}$. Como $\|\phi_n - \phi_m\|^2 = 2$ para todos $n, m \in \mathbb{N}$, entonces la sucesión $I\phi_n = \phi_n$ no contiene subsucesiones de Cauchy y, por tanto, no tiene subsucesiones convergentes, de donde se sigue el resultado. ■

Teorema 6.4.4 *Sea A un operador lineal $A : \mathbb{X} \mapsto \mathbb{Y}$, $\mathbb{X}$ e $\mathbb{Y}$ espacios normados de dimensión finita. Entonces A es compacto.*

Demostración: En efecto, como $\mathbb{X}$ es de dimensión finita, entonces A es acotado (ver Teorema 4.4.17), luego dada cualquier sucesión acotada $(x_n)_n$, la sucesión $(Ax_n)_n$ es acotada. Sea el conjunto $C_A = \{Ax_n : n \in \mathbb{N}\}$. C_A es acotado, luego su clausura $\overline{C_A}$ es un subconjunto de $\mathbb{Y}$ que es cerrado y acotado, luego por el Teorema 4.3.8, es compacto. Luego, toda sucesión $(y_n)_n \in \overline{C_A}$, tiene una subsucesión convergente $(y_{n_k})_k$, $y_{n_k} \to y \in \overline{C_A}$, es decir, $y \in \mathbb{Y}$. Por tanto, existe una subsucesión convergente $(Ax_{n_k})_k$, luego A es compacto. ■

Nota 6.4.5 *Nótese que si el operador lineal A del Teorema 6.4.4 es acotado, i.e., $A \in \mathcal{B}(\mathbb{X}, \mathbb{Y})$, y su imagen $\mathcal{I}(A)$ es de dimensión finita, entonces, si $(x_n)_n$ es acotada, $(Ax_n)_n$ lo será y razonando como en la prueba de dicho teorema, se deduce que A es compacto (independientemente de la dimensión de $\mathbb{X}$). Es decir, la hipótesis de que $\mathbb{X}$ sea de dimensión finita en el Teorema 6.4.4 no es necesaria si el operador es acotado.*

Ejemplo 6.4.6 *Se puede probar directamente, sin usar el teorema anterior, que cualquier operador lineal $A : \mathbb{H} \mapsto \mathbb{H}$, con $\mathbb{H}$ espacio de Hilbert de dimensión finita es compacto.*

En efecto, del Teorema 4.4.17 sabemos que el operador A es acotado, y del Teorema del isomorfismo 5.2.12 que $\mathbb{H}$ es equivalente (isomorfo) a $\mathbb{C}^N$, con N la dimensión de $\mathbb{H}$. Entonces, dada cualquier sucesión acotada $(x_n)_n$, la sucesión (Ax_n) es acotada y por el Teorema de Bolzano-Weierstrass (en $\mathbb{C}^N$) dicha sucesión contiene al menos una subsucesión convergente, luego A es compacto. ■

Teorema 6.4.7 *El conjunto de todos los operadores compactos $A : \mathbb{H} \mapsto \mathbb{H}$, es un subespacio vectorial de $\mathcal{B}(\mathbb{H})$.*

Demostración: Sea $(x_n)_n \in \mathbb{H}$ una sucesión acotada. Sean A y B, dos operadores compactos. Como A es compacto, existe una subsucesión $(y_k)_k$ de $(x_n)_n$ ($y_k = x_{n_k}$) tal que Ay_k es convergente. Luego, también lo es la sucesión $(\lambda A)y_k$, y por tanto el operador λA es compacto. Por otro lado, como B es compacto, entonces existe una subsucesión $(z_j)_j$ de $(y_k)_k$ (y, por tanto, de $(x_n)_n$) tal que $(Bz_j)_j$ es convergente. Pero entonces $(Az_j)_j$ es convergente (¿por qué?). Luego, dada cualquier sucesión acotada $(x_n)_n$, existe una subsucesión $(z_j)_j$ de $(x_n)_n$ tal que $([A+B]z_j)_j$ es convergente. Luego $A+B$ es compacto. ■

Proposición 6.4.8 *Sea $\mathbb{H}$ un espacio de Hilbert y sean y y z dos elementos dados de $\mathbb{H}$. Sea el operador lineal $T : \mathbb{H} \mapsto \mathbb{H}$ definido por $Tx = \langle x, y\rangle z$. Entonces T es compacto.*

Demostración: Sea $(x_n)_n$ una sucesión acotada de $\mathbb{H}$, i.e., $\|x_n\| \le M$ para todo $n \in \mathbb{N}$ y cierto $M > 0$. Sea la sucesión compleja $a_n = \langle x_n, y\rangle$. Probemos que es acotada

$$|a_n| = |\langle x_n, y\rangle| \le \|x_n\|\|y\| \le M\|y\|.$$

Entonces, por el Teorema de Bolzano-Weierstrass a_n tiene una subsucesión convergente $a_{k_n} = \langle x_{k_n}, y\rangle$. Sea a su límite. Entonces, utilizando la continuidad del producto escalar,

$$Tx_{k_n} = \langle x_{k_n}, y\rangle z \overset{n\to\infty}{\longrightarrow} az,$$

es decir, la sucesión Tx_n contiene una subsucesión convergente. Luego T es compacto. ■

Ejercicio 6.4.9 *Sea $A : \mathbb{H} \mapsto \mathbb{H}$ un operador acotado de* rango finito *en un espacio de Hilbert $\mathbb{H}$, i.e., la imagen de A, $\mathcal{I}(A)$, es de dimensión finita ($\dim \mathcal{I}(A) = n < +\infty$). Prueba, usando la proposición anterior, que A es compacto.*

Vamos a probarlo directamente usando la Proposición 6.4.8. Sea e_1, …, e_n una base ortogonal de $\mathcal{I}(A) \subset \mathbb{X}$. Sea $y \in \mathcal{I}(A)$, cualquiera. Entonces, para cada y existe al menos un x tal que $y = Ax$, y además $y = \sum_{k=1}^{n} \langle Ax, e_k \rangle e_k$. Denotemos por A_k a los operadores $A_k : \mathbb{H} \mapsto \mathbb{H}$, $A_k = \langle Ax, e_k \rangle e_k$, $k = 1, \ldots, n$. Entonces $A = \sum_{k=1}^{n} A_k$. Como A es acotado existe su adjunto A^* (Teorema 6.1.3), luego $A_k x = \langle Ax, e_k \rangle e_k = \langle x, A^* e_k \rangle e_k$. Entonces, por la Proposición 6.4.8 deducimos que los A_k, $k = 1, \ldots, n$ son compactos y como $A = \sum_{k=1}^{n} A_k$, del Teorema 6.4.7 se deduce que A es compacto. ■

Nota: Nótese que al ser A acotado y de rango finito se cumplen las premisas de la Nota 6.4.5 de donde se sigue inmediatamente el resultado del Ejercicio 6.4.9.

Teorema 6.4.10 *Sea $(T_n)_n$ una sucesión de operadores compactos $T_n : \mathbb{H} \mapsto \mathbb{H}$, $\mathbb{H}$ espacio de Hilbert, tal que $T_n \overset{n\to\infty}{\longrightarrow} T$. Entonces T es compacto.*

Demostración: Sea $(x_n)_n \in \mathbb{H}$ una sucesión acotada cualquiera y sea $S = \sup_{n\in\mathbb{N}} \|x_n\|$. Asumiremos que $S > 0$, pues en caso contrario $(x_n)_n$ sería la sucesión nula $(0)_n$. Vamos a probar que, como para todo m, T_m es compacto, entonces existe una subsucesión $(z_j)_j$ de $(x_n)_n$ tal que $\lim_{k\to\infty} T_m z_j \to y_m$ para $m = 1, 2, 3, \ldots$, i.e., la sucesión $(T_m z_j)_j$ es convergente para cada $m \in \mathbb{N}$. La forma de construir dicha sucesión recuerda a la que se usó en la prueba del Lema técnico 4.3.1, y se conoce como *construcción diagonal*[5], y es como sigue:

Sea $n = 1$. Como T_1 es compacto entonces existe una subsucesión $(x_{1,j})_j$ tal que $(T_1 x_{1,j})_j$ es convergente. Como T_2 es compacto entonces existe una subsucesión $(x_{2,j})_j$ de $(x_{1,j})_j$ tal que $(T_2 x_{2,j})_j$ es convergente, y así sucesivamente. De esta forma obtenemos que, para cada $m \in \mathbb{N}$, existe una subsucesión $(x_{m,j})_j$ de $(x_j)_j$ tal que $\lim_{j\to\infty} T_m x_{m,j} \to y_m$, para $m = 1, 2, 3, \ldots$, y además, por construcción, $(x_{n,j})_j$ es una subsucesión de las anteriores $(x_{k,j})_j$, $k = 1, 2, \ldots, n-1$. El siguiente diagrama muestra la

[5]Este tipo de razonamiento fue utilizado por primera vez por Cantor para probar que el conjunto de los reales no es numerable.

construcción anterior:

$$
\begin{array}{lcccccccl}
(x_j)_j & x_1 & x_2 & x_3 & \cdots & x_m & \cdots & x_j & \cdots \\
(x_{1,j})_j \subset (x_j)_j & x_{1,1} & x_{1,2} & x_{1,3} & \cdots & x_{1,m} & \cdots & x_{1,j} & \cdots \Rightarrow T_1 x_{1,j} \xrightarrow{j} y_1 \\
(x_{2,j})_j \subset (x_{1,j})_j & x_{2,1} & x_{2,2} & x_{2,3} & \cdots & x_{2,m} & \cdots & x_{2,j} & \cdots \Rightarrow T_2 x_{2,j} \xrightarrow{j} y_2 \\
(x_{3,j})_j \subset (x_{2,j})_j & x_{3,1} & x_{3,2} & x_{3,3} & \cdots & x_{3,m} & \cdots & x_{3,j} & \cdots \Rightarrow T_3 x_{3,j} \xrightarrow{j} y_3 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & & \quad\vdots \\
(x_{m,j})_j \subset (x_{m-1,j})_j & x_{m,1} & x_{m,2} & x_{m,3} & \cdots & x_{m,m} & \cdots & x_{m,j} & \cdots \Rightarrow T_m x_{m,j} \xrightarrow{j} y_m \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & & \quad\vdots
\end{array}
$$

Sea ahora la sucesión de las “diagonales” $(x_{j,j})_j$. Para cada m fijo la sucesión $(x_{j,j})_{j=m}^{\infty} = (x_{m,m}, x_{m+1,m+1}, \dots)$ es una subsucesión de $(x_{m,j})_j$, luego

$$\lim_{j\to\infty} T_m x_{j,j} = y_m, \quad \text{para} \quad m = 1, 2, 3, \dots .$$

Eso implica que la sucesión $(T_m x_{j,j})_j$ es de Cauchy. Probemos ahora que la sucesión $(T x_{j,j})_j$ es de Cauchy.

Como $T_n \overset{n\to\infty}{\longrightarrow} T$, eso significa que para todo $\varepsilon > 0$ existe un $N_1 \in \mathbb{N}$ tal que, para todo $M > N_1$,

$$\|T - T_M\| < \frac{\varepsilon}{3S}.$$

Pero como ya vimos, la sucesión $(T_m x_{j,j})_j$ es de Cauchy para todo $m \in \mathbb{N}$, luego lo será para $m = M$. Luego, para todo $\varepsilon > 0$ existe un $N_2 > M \in \mathbb{N}$ tal que, para todos $m > n > N_2$

$$\|T_M x_{n,n} - T_M x_{m,m}\| \le \frac{\varepsilon}{3}.$$

Pero entonces, si elegimos $m > n > N_2 > M > N_!$, obtenemos

$$
\begin{aligned}
\|Tx_{n,n} - Tx_{m,m}\| \le{}& \|Tx_{n,n} - T_M x_{n,n}\| + \|T_M x_{n,n} - T_M x_{m,m}\| \\
&+ \|T_M x_{m,m} - T x_{m,m}\| \\
\le{}& \|T - T_M\| \|x_{n,n}\| + \frac{\varepsilon}{3} + \|T_M - T\| \|x_{m,m}\| < \varepsilon,
\end{aligned}
$$

i.e., $(Tx_{j,j})_j$ es de Cauchy y como $\mathbb{H}$ es completo, entonces $(Tx_{j,j})_j$ es convergente, luego T es compacto. ■

El Teorema 6.4.10 se puede parafrasear diciendo que el subespacio de los operadores compactos es cerrado[6] en $\mathcal{B}(\mathbb{H})$ y, por tanto, completo.

[6] Ello se deduce de la Proposición 3.5.4. Además, como $\mathcal{B}(\mathbb{H})$ es completo entonces del Teorema 3.5.14 se sigue que el conjunto de los operadores compactos es completo.

Teorema 6.4.11 *Sea $T : \mathbb{H} \mapsto \mathbb{H}$, un operador compacto, $\mathbb{H}$ espacio de Hilbert. Entonces su adjunto T^* es compacto.*

Demostración: Sea $(x_n)_n$ una sucesión acotada cualquiera no nula, y sea $S = \sup_{n\in\mathbb{N}} \|x_n\|$. Definamos la sucesión $(y_n)_n \in \mathbb{H}$, $y_n = T^* x_n$. Como T es compacto, T es acotado y, por tanto, existe su adjunto T^* y es acotado, así que la sucesión $(y_n)_n$ es acotada. Como T es compacto, entonces existe una subsucesión $(y_{n_k})_k$ de $(y_n)_n$ tal que $(Ty_{n_k})_k \in \mathbb{H}$ es convergente y, por tanto, de Cauchy (pues $\mathbb{H}$ es completo). Pero entonces tomando n_k, n_m suficientemente grandes la cantidad $\|Ty_{n_k} - Ty_{n_m}\|$ la podemos hacer tan pequeña como se quiera (por ejemplo menor que $\varepsilon^2/(2S)$). Entonces,

$$\begin{aligned}
\|y_{n_k} - y_{n_m}\|^2 &= \|T^* x_{n_k} - T^* x_{n_m}\|^2 = \langle T^*(x_{n_k} - x_{n_m}), T^*(x_{n_k} - x_{n_m})\rangle \\
&= \langle TT^*(x_{n_k} - x_{n_m}), (x_{n_k} - x_{n_m})\rangle \\
&\leq \|TT^*(x_{n_k} - x_{n_m})\| \|x_{n_k} - x_{n_m}\| \\
&\leq 2S\|Ty_{n_k} - Ty_{n_m}\| < 2S\frac{\varepsilon^2}{2S} = \varepsilon^2.
\end{aligned}$$

Luego, $(y_{n_k})_k$ es de Cauchy, y como $\mathbb{H}$ es completo, es convergente, i.e., $(T^* x_{n_k})_k$ es convergente y, por tanto, T^* es compacto. ■

6.5. Teoría espectral

En este apartado estudiaremos la teoría espectral de operadores autoadjuntos y compactos en espacios de Hilbert. Dicha teoría es la extensión del conocido problema de autovalores para matrices $n \times n$ al caso infinito.

6.5.1. El espectro de un operador

Comenzaremos, con la definición de autovalores y autovectores[7]

Definición 6.5.1 *Sea $A : \mathbb{X} \mapsto \mathbb{X}$ un operador en un espacio normado $\mathbb{X}$. Los números complejos λ tales que*

$$Ax = \lambda x, \quad x \in \mathbb{X},\ x \neq 0, \tag{6.5.1}$$

[7] También denominados valores propios y vectores propios, respectivamente.

se denominan autovalores de A y los correspondientes x, autovectores de A asociados al autovalor λ.

En un espacio de dimensión finita podemos definir el *espectro* de un operador como el conjunto de los autovalores de su correspondiente matriz en alguna basePuesto que para cualquier matriz $n \times n$ existen n autovalores (pudiendo estar repetidos como ocurre, por ejemplo, con la matriz identidad), en el caso finito es relativamente simple de estudiar (ver Ejercicio 6.5.3). No ocurre lo mismo en el caso infinito.

Por ejemplo, el operador desplazamiento $S_\frown : \ell^2 \mapsto \ell^2$, definido por $S_\frown(x_1, x_2, x_3, \ldots) = (0, x_1, x_2, \ldots)$, no tiene autovalores pues la igualdad $S_\frown x = \lambda x$ implica $x = 0$.

Así pues se precisa de una definición más general

Definición 6.5.2 *Sea $\mathbb{X}$ un espacio de Banach y A una aplicación lineal $A : \mathbb{X} \mapsto \mathbb{X}$. El espectro de A, que denotaremos por $\sigma(A)$, es el conjunto de números complejos tales que el operador $(\lambda I - A)$ es no invertible, i.e., no existe $A_\lambda := (\lambda I - A)^{-1}$.*

El operador $A_\lambda := (\lambda I - A)^{-1}$ se denomina resolvente de A. El conjunto de todos los $\lambda \in \mathbb{C}$ para los cuales el operador A_λ está bien definido se denomina conjunto resolvente de A y se suele denotar por $\rho(A)$. Obviamente el espectro de A, $\sigma(A)$, es el complementario del conjunto resolvente, $\rho(A)$, i.e., $\sigma(A) = \mathbb{C} \setminus \rho(A)$.

La cantidad $r(A) = \sup\{|\lambda| \,;\, \lambda \in \sigma(A)\}$ se denomina radio espectral de A.

Ejercicio 6.5.3 *Sea $A : \mathbb{X} \mapsto \mathbb{X}$, $A \in \mathcal{B}(\mathbb{X})$. Prueba que en dimensión finita, $\dim \mathbb{X} < +\infty$, $\sigma(A)$ es el conjunto de todos los autovalores de A.*

Ante todo notemos que si $\dim \mathbb{X} = n < +\infty$, entonces como vimos en la Nota 6.1.8, cada operador $A : \mathbb{H} \mapsto \mathbb{H}$ se puede representar mediante una matriz $n \times n$. Luego, la ecuación (6.5.1) se trasforma en el sistema lineal

$$Ax = \lambda x \quad \Longleftrightarrow \quad (A - \lambda I)x = 0, \quad x \neq 0,$$

donde I es la matriz identidad $n \times n$. Para que este sistema tenga solución no trivial ha de cumplirse que $\det(A - \lambda I) = 0$, lo cual implica que el operador $A - \lambda I$ no puede ser invertible, luego $\lambda \in \sigma(A)$. Nótese además que $\det(A - \lambda I)$, es un polinomio de grado n en λ, luego la ecuación

$\det(A - \lambda I) = 0$ tiene al menos una solución y como mucho n soluciones distintas. Conviene recordar que si usamos dos bases distintas para representar la matriz del operador A, lo que obtenemos son dos matrices semejantes y por tanto ambas tienen los mismos autovalores[8], i.e., los autovalores de un operador en dimensión finita son independientes de la base usada para representar la matriz de dicho operador. ∎

De la Definición 6.5.2 se sigue que todo autovalor de A pertenece al espectro de A, pero no a la inversa. De hecho existen operadores cuyo espectro no contiene autovalores.

Ejemplo 6.5.4 *Sea el espacio de las funciones reales (o complejas) de $L^2_{[a,b]}$ (i.e., el completamiento del espacio $C^2_{[a,b]}$, véase la Nota 3.5.27). Sea $M : L^2_{[a,b]} \mapsto L^2_{[a,b]}$, el operador multiplicación por una función real $f \in C([a,b])$, $Mx(t) = f(t)x(t)$ (ver Ejemplo 6.1.2). Está claro que M esta bien definido para todo $x \in L^2_{[a,b]}$, es autoadjunto y acotado. Este operador, en general, no tiene autovalores.*

Puesto que $(M - \lambda I)x(t) = (f(t) - \lambda)x(t)$, entonces

$$(M - \lambda I)^{-1}x(t) = \frac{1}{f(t) - \lambda}x(t).$$

El espectro de M es el conjunto de todos los λ tales que $f(t) = \lambda$ para algún $t \in [a,b]$, es decir, $\sigma(M)$ es la imagen de la función $f(t)$, i.e., el rango $\mathcal{I}(M)$. Si la función $f(t)$ es constante, $f(t) = c$, entonces $\lambda = c$ es un autovalor de M (pues $Mx(t) = cx(t)$) que obviamente está en $\sigma(U)$. Pero si, por ejemplo, $f(t)$ es una función estrictamente monótona en $[a,b]$, entonces $\sigma(M) = [f(a), f(b)]$ (por ejemplo, si $f(t) = t$, entonces $\sigma(M) = [a,b]$), pero la ecuación $Mx(t) = f(t)x(t) = \lambda x(t)$ solo se cumple si $x(t) \equiv 0$, i.e., M no tiene autovalores. ∎

Ejercicio 6.5.5 *Sea $A \in \mathcal{B}(\mathbb{X})$, $\mathbb{X}$ espacio de Banach. Prueba que $r(A) \leq \|A\|$. Es decir, $\sigma(A)$ está contenido en el disco cerrado $D = \{z;\, |z| \leq \|A\|\}$.*

Supongamos que existe un $\lambda \in \sigma(A)$ tal que $|\lambda| > \|A\|$. Entonces por el Teorema 6.3.3 el operador $A_\lambda := (\lambda I - A)^{-1}$ está bien definido lo cual

[8] Dos matrices A y B son semejantes, $A \sim B$, si y solo si existe una matriz invertible S tal que $A = SBS^{-1}$. Pero entonces $\det(A - \lambda I) = \det(S(A - \lambda I)S^{-1}) = \det(SAS^{-1} - \lambda I)$ $= \det(B - \lambda I)$.

es una contradicción. Luego si $\lambda \in \sigma(A)$, $|\lambda| \leq \|A\|$ y, por tanto, $\sup\{|\lambda| : \lambda \in \sigma(A)\} \leq \|A\|$, de donde se sigue el resultado. ■

Teorema 6.5.6 *Sea $\mathbb{H}$ un espacio de Hilbert y $A \in \mathcal{B}(\mathbb{H})$ un operador autoadjunto. Entonces todos los autovalores de A (si los tiene) son reales. Además los autovectores correspondientes a autovalores distintos son ortogonales.*

Demostración: Sea λ un autovalor de A y x su correspondiente autovector, que sin pérdida de generalidad asumiremos normalizado $\|x\| = 1$. Entonces, usando que $Ax = \lambda x$ y que A es autoadjunto, tenemos

$$\langle Ax, x\rangle = \langle x, Ax\rangle \Rightarrow \lambda\|x\| = \overline{\lambda}\|x\| \Rightarrow \lambda = \overline{\lambda} \Rightarrow \lambda \in \mathbb{R}.$$

Sea $Ax_1 = \lambda_1 x_1$ y $Ax_2 = \lambda_2 x_2$. Entonces como A es autoadjunto

$$\langle Ax_1, x_2\rangle = \langle x_1, Ax_2\rangle \Rightarrow \lambda_1\langle x_1, x_2\rangle = \lambda_2\langle x_1, x_2\rangle \Rightarrow (\lambda_1 - \lambda_2)\langle x_1, x_2\rangle = 0,$$

de donde se sigue que si $\lambda_1 \neq \lambda_2$ entonces x_1 y x_2 son ortogonales. ■

Teorema 6.5.7 *Sea $\mathbb{H}$ un espacio de Hilbert y $A \in \mathcal{B}(\mathbb{H})$ un operador autoadjunto. Entonces $r(A) = \|A\|$.*

Demostración: Para $A = \Theta$ (operador nulo) el resultado es trivial así que asumiremos que $A \neq \Theta$. Del Ejercicio 6.5.5 sabemos que $r(A) \leq \|A\|$, así que bastará probar que existe un $\lambda \in \sigma(A)$ tal que $|\lambda| = \|A\|$, i.e., $\lambda = \|A\|$ o $\lambda = -\|A\|$.

Por el Teorema 6.2.6 sabemos que $\|A\| = \sup_{\|x\|=1} |\langle Ax, x\rangle|$, luego existe una sucesión de vectores normalizados a la unidad ($\|x_n\| = 1$ para todo n) $(x_n)_n$ de $\mathbb{H}$ y tal que $|\langle Ax_n, x_n\rangle| \overset{n\to\infty}{\longrightarrow} \|A\|$ y, por tanto, sin pérdida de generalidad, podemos asegurar que $\langle Ax_n, x_n\rangle \overset{n\to\infty}{\longrightarrow} \|A\|$, o $\to -\|A\|$.

Sea $\lambda = \|A\|$ (o $-\|A\|$). Probemos que $\lambda \in \sigma(A)$. Ante todo notemos que, como $\lambda^2 = \|A\|^2 = \lambda^2$ y $\|x_n\| = 1$, obtenemos

$$\begin{aligned} 0 \leq \|Ax_n - \lambda x_n\|^2 =& \|Ax_n\|^2 - 2\lambda\langle Ax_n, x_n\rangle + \lambda^2\|x_n\|^2 \\ \leq& \|A\|^2 - 2\lambda\langle Ax_n, x_n\rangle + \lambda^2 = 2\lambda(\lambda - \langle Ax_n, x_n\rangle) \overset{n\to\infty}{\longrightarrow} 0, \end{aligned}$$

es decir, nuestra sucesión $(x_n)_n$ es tal que $\|Ax_n - \lambda x_n\| \overset{n\to\infty}{\longrightarrow} 0$. Asumamos que $\lambda = \|A\|$ (o, en su caso, $-\|A\|$) no está en $\sigma(A)$, i.e., $\lambda \in \rho(A)$. Como

λ está en la resolvente de A entonces para dicho λ el operador $(A-\lambda I)^{-1}$ está bien definido y es acotado. Pero entonces

$$1 = \|x_n\| = \|(A-\lambda I)^{-1}(A-\lambda I)x_n\| \leq \|(A-\lambda I)^{-1}\|\|(A-\lambda I)x_n\| \stackrel{n\to\infty}{\longrightarrow} 0,$$

lo cual es una contradicción. Luego $\lambda = \|A\| \in \sigma(A)$ (o $-\|A\| \in \sigma(A)$). ∎

Nota: El teorema anterior establece que para los operadores autoadjuntos y acotados siempre $\|A\| \in \sigma(A)$ o $-\|A\| \in \sigma(A)$, sin embargo, como hemos visto antes, en dimensión infinita un operador lineal A acotado en general puede no tener autovalores, es decir, los valores $\lambda \in \sigma(A)$ no tienen que ser necesariamente solución de la ecuación (6.5.1). Esto es lo que ocurre, por ejemplo, con el operador desplazamiento $S_\curvearrowright$ (definido por (6.1.2)) que no es autoadjunto, o el operador multiplicación por una función real (ver Ejemplo 6.5.4) que si es autoadjunto. Incluso, un operador compacto puede no tener autovalores. Un ejemplo es el operador $T = S_\curvearrowright A$, donde A es el operador del Ejemplo 4 de la página 223. Dicho operador A es un operador de Hilbert-Schmidt pues $\sum_{n=1}^\infty \|Ax\|^2 = \sum_{n=1}^\infty |x_n|^2/n^2 \leq (\sup_n |x_n|^2) \sum_{n=1}^\infty 1/n^2 < +\infty$ (ver Problema 6.8) y, por tanto, compacto (ver Problema 6.18), así que el operador T es compacto (ver Problema 6.14), sin embargo claramente T no tiene autovalores pues la única solución de la ecuación $Tx = \lambda x$ es $x = 0$. No ocurre así con los operadores autoadjuntos y compactos.

Teorema 6.5.8 *Sea $\mathbb{H}$ un espacio de Hilbert y $A \in \mathcal{B}(\mathbb{H})$ un operador autoadjunto y compacto. Entonces $\lambda = \|A\|$ o $\lambda = -\|A\|$ es un autovalor de A.*

Demostración: Para $A = \Theta$ (operador nulo) el resultado es trivial así que asumiremos que $A \neq \Theta$. Escojamos $\lambda = \|A\|$ (el caso $\lambda = -\|A\|$ es análogo y lo omitiremos). De la prueba del teorema anterior sabemos que si A es un operador autoadjunto y acotado entonces existe una sucesión $(x_n)_n$ de $\mathbb{H}$ tal que $\|x_n\| = 1$ para todo n y que $Ax_n - \lambda x_n \stackrel{n\to\infty}{\longrightarrow} 0$. Como A es compacto, entonces existe una subsucesión $(x_{n_k})_k$ de $(x_n)_n$ tal que $(Ax_{n_k})_k$ es convergente, así $Ax_{n_k} \stackrel{k\to\infty}{\longrightarrow} u$. Además, como $Ax_{n_k} - \lambda x_{n_k} \stackrel{n\to\infty}{\longrightarrow} 0$, se tiene que $\lambda x_{n_k} \stackrel{k\to\infty}{\longrightarrow} u$. Como A es compacto, A es acotado y, por tanto, continuo así que por un lado tendremos

$$\lambda A x_{n_k} = A(\lambda x_{n_k}) \stackrel{k\to\infty}{\longrightarrow} Au \text{ y por otro } \lambda(Ax_{n_k}) \stackrel{k\to\infty}{\longrightarrow} \lambda u \;\Rightarrow\; Au = \lambda u.$$

Finalmente, notemos que

$$\|u\| = \lim_{k\to\infty} \|\lambda x_{n_k}\| = |\lambda| \lim_{k\to\infty} \|x_{n_k}\| = |\lambda| \lim_{k\to\infty} 1 = |\lambda| \neq 0,$$

luego λ es, efectivamente un autovalor. ■

Teorema 6.5.9 *Sea $\mathbb{H}$ un espacio de Hilbert de dimensión infinita, $A \in \mathcal{B}(\mathbb{H})$ un operador compacto y $(\phi_n)_n$ una sucesión ortonormal de $\mathbb{H}$. Entonces* $\lim_{n\to\infty} A\phi_n = 0$.

Demostración: Supongamos que el teorema es falso, entonces para algún $\varepsilon > 0$ ha de existir una subsucesión $(\phi_{n_k})_k$ de$(\phi_n)_n$ tal que $\|A\phi_{n_k}\| > \varepsilon$, para todo $k \in \mathbb{N}$. Como A es compacto y la sucesión $(\phi_{n_k})_k$ es acotada (¿por qué?), entonces existe al menos una subsucesión $(\psi_j)_j$ de $(\phi_{n_k})_k$ convergente, tal que $\lim_{j\to\infty} A\psi_j = \psi \neq 0$. Entonces, por la continuidad del producto escalar,

$$0 \neq \|\psi\|^2 = \langle \psi, \psi \rangle = \lim_{j\to\infty} \langle \psi, A\psi_j \rangle = \lim_{j\to\infty} \langle A^*\psi, \psi_j \rangle = 0,$$

pues $A^*\psi \in \mathbb{H}$ y $\langle A^*\psi, \psi_j \rangle$ es el j-ésimo coeficiente de Fourier c_j de $A^*\psi$ el cual sabemos, por (5.2.4), que tiende a cero si $j \to \infty$, lo cual es una contradicción. ■

Teorema 6.5.10 *Sea $\mathbb{H}$ un espacio de Hilbert, $A \in \mathcal{B}(\mathbb{H})$ un operador compacto y sea $\mathbb{X}_\lambda = \ker(\lambda I - A)$, el autoespacio correspondiente a un autovalor $\lambda \neq 0$ (si lo tiene). Entonces* $\dim \mathbb{X}_\lambda < +\infty$.

Demostración: Sea $A : \mathbb{H} \mapsto \mathbb{H}$ compacto. Supongamos que $\dim \mathbb{X}_k = \infty$ para algún k, con $\lambda_k \neq 0$. Entonces, existe una sucesión de autovectores linealmente independiente $x_1, x_2, \ldots, x_n, \ldots$ que, sin pérdida de generalidad asumiremos ortonormales. Si $n \neq m$, $\|Ax_m - Ax_n\|^2 = |\lambda_k|^2 \|x_m - x_n\|^2 = 2|\lambda_k|^2 > 0$, luego $(Ax_n)_n$ no puede contener ninguna subsucesión convergente (¿por qué?) lo cual contradice la compacidad de[9] A. ■

Teorema 6.5.11 *Sea $\mathbb{H}$ un espacio de Hilbert y $A \in \mathcal{B}(\mathbb{H})$ un operador compacto con un número infinito de autovalores. Entonces, si los ordena-*

[9]También podríamos razonar como sigue: como $Ax_n = \lambda_k x_n$, $n = 1, 2, 3, \ldots$, entonces, por el Teorema 6.5.9, $Ax_n \overset{n\to\infty}{\longrightarrow} 0$, de donde se sigue $\lambda_k x_n \overset{n\to\infty}{\longrightarrow} 0$, lo cual implica que $\lambda_k = 0$, pues $\|x_n\| = 1$ para todo n, que es una contradicción.

mos en modulo de mayor a menor $|\lambda_n| \stackrel{n\to\infty}{\longrightarrow} 0$.

Demostración: En efecto, supongamos que hay infinitos λ_n distintos pero que $|\lambda_n| \stackrel{n\to\infty}{\not\to} 0$. Entonces existe un $\varepsilon > 0$ tal que infinitos λ_{n_k} son tales que $|\lambda_{n_k}| > \varepsilon$. Construyamos con dichos elementos una sucesión que denotaremos por $(\lambda_k)_k$. Como todos los elementos de $(\lambda_k)_k$ son distintos, el Teorema 6.5.6 garantiza que sus correspondientes autovectores son ortogonales x_k, i.e., $\langle x_{k_1}, x_{k_2}\rangle = 0$ si $k_1 \neq k_2$. Si calculamos ahora la norma

$$\|Ax_{k+p} - Ax_k\|^2 = \|\lambda_{k+p}x_{k+p} - \lambda_k x_k\|^2 = |\lambda_{k+p}|^2 + |\lambda_k|^2 > 2\varepsilon^2, \quad \forall k, p \in \mathbb{N},$$

es decir, la sucesión $(Ax_k)_k$ no contiene ninguna subsucesión de Cauchy y, por tanto, no contiene ninguna sucesión convergente (¿por qué?) lo que contradice que A sea un operador compacto. ■

Teorema 6.5.12 *Sea* $\mathbb{H}$ *un espacio de Hilbert y* $A \in \mathcal{B}(\mathbb{H})$ *un operador autoadjunto y compacto. Entonces* A *tiene o bien un número finito de autovalores* λ_n *reales distintos o, si dicho número es infinito, entonces, si los ordenamos en valor absoluto de mayor a menor* $|\lambda_n| \stackrel{n\to\infty}{\longrightarrow} 0$.

Demostración: Asumiremos que $A \neq \Theta$ no es el operador nulo. Del Teorema 6.5.8 se sigue que el conjunto de autovalores no es vacío, luego está compuesto por un número finito o infinito de elementos (que son números reales por el Teorema 6.5.6). Sea λ_1 tal que $|\lambda_1| = \|A\|$ y sea x_1 el correspondiente autovector normalizado a la unidad, i.e. $\|x_1\| = 1$. Sea $\mathbb{H}_1 = \mathbb{H}$ y definamos $\mathbb{H}_2$ el espacio de todos los vectores ortogonales a x_1, i.e.,

$$\mathbb{H}_2 = \{x \in \mathbb{H} \,|\, \langle x, x_1\rangle = 0\},$$

es el complemento ortogonal de x_1. Nótese que por el Ejercicio 5.2.14, $\mathbb{H}_2$ es cerrado, luego es un espacio de Hilbert (¿por qué?). Además, para todo $x \in \mathbb{H}_2$,

$$\langle Ax, x_1\rangle = \langle x, Ax_1\rangle = \lambda_1\langle x, x_1\rangle = 0,$$

es decir, $\mathbb{H}_2$ es invariante respecto a la acción de A.

Sea $A|_{\mathbb{H}_2}$ la restricción de A al espacio $\mathbb{H}_2$. Si $A|_{\mathbb{H}_2} = \Theta$, entonces tendremos un único autovalor λ_1 y el teorema es cierto. En caso contrario, como $A|_{\mathbb{H}_2}$ es autoadjunto y compacto[10] (pues coincide con A en $\mathbb{H}_2$) podemos aplicar el mismo razonamiento de antes, por lo que existirá λ_2 y

[10] Que $A|_{\mathbb{H}_2}$ es autoadjunto es *evidente* pues $A|_{\mathbb{H}_2}$ coincide con A en $\mathbb{H}_2$. Para ver que es compacto basta notar que cualquiera sea la sucesión acotada $(y_n)_n \in \mathbb{H}_2$, $(Ay_n)_n$ contiene una subsucesión convergente, pues A es compacto en $\mathbb{H}$ y $\mathbb{H}_2 \subset \mathbb{H}$ es completo.

$x_2 \neq 0$ tales que $|\lambda_2| = \|A|_{\mathbb{H}_2}\|$, $Ax_2 = \lambda_2 x_2$, donde además, por la definición de norma de operadores, $|\lambda_2| \leq |\lambda_1|$. Ahora construimos el espacio

$$\mathbb{H}_3 = \{x \in \mathbb{H}_2 \,|\, \langle x, x_2 \rangle = 0\} = \{x \in \mathbb{H} \,|\, \langle x, x_1 \rangle = \langle x, x_2 \rangle = 0\},$$

que es invariante respecto a A, y así sucesivamente.

Ese decir, podemos continuar el proceso anterior a no ser que para cierto $n \geq 1$, $A|_{\mathbb{H}_{n+1}} = \Theta$. En ese caso tendremos una sucesión finita de autovalores $(\lambda_k)_{k=1}^n$ con sus correspondientes autovectores normalizados[11] y ortogonales entre sí, $(x_k)_{k=1}^n$, tales que

$$|\lambda_1| = \|A|_{\mathbb{H}_1}\| \geq |\lambda_2| = \|A|_{\mathbb{H}_2}\| \geq \cdots \geq |\lambda_n| = \|A|_{\mathbb{H}_n}\| \neq 0, \quad A|_{\mathbb{H}_{n+1}} = 0.$$

Si $A|_{\mathbb{H}_k} \neq 0$ para todo $k \in \mathbb{N}$, entonces existen infinitos autovalores λ_n (¿por qué?) con sus correspondientes autovectores x_n ortonormalizados (que pueden ser más de uno). Probemos que, en este caso, $|\lambda_n| \stackrel{n\to\infty}{\longrightarrow} 0$. Para ello usamos que $0 = \lim_{n\to\infty} \|Ax_n\|^2 = \lim_{n\to\infty} \langle Ax_n, Ax_n \rangle = \lim_{n\to\infty} |\lambda_n|^2$, donde la primera igualdad es consecuencia del Teorema 6.5.9. Nótese que en el caso que hubiese infinitos autovalores λ_n, también el Teorema 6.5.11 implica que $|\lambda_n| \stackrel{n\to\infty}{\longrightarrow} 0$. ■

Nota 6.5.13 *Del Teorema 6.5.10 sabemos que cada autovalor λ_k solo puede tener asociado un número finito de autovectores, así que en la prueba del Teorema 6.5.12 hemos asumido, por simplicidad, que existe un único autovector asociado a cada λ_k. En caso de que haya p_k autovectores $x_{k,j}$, $j = 1, \ldots, p_k$, asociados a λ_k, definiremos el espacio $\mathbb{H}_{k+1}$ como el espacio de todos los $x \in \mathbb{H}_k$ tales que sean ortogonales dichos $x_{k,j}$, y se razona de la misma manera.*

6.5.2. Teorema espectral

Teorema espectral para operadores autoadjuntos y compactos

Para los operadores autoadjuntos y compactos se tiene el siguiente importante resultado:

[11] Es importante destacar que para cada λ_k, $k = 1, 2, 3, \ldots$ puede haber más de un autovector. Si los correspondientes autovectores asociados cada λ_k no son ortogonales usamos el método de Gram-Schmidt.

Teorema 6.5.14 (Teorema espectral) *Sea $\mathbb{H}$ un espacio de Hilbert y $A \in \mathcal{B}(\mathbb{H})$ un operador autoadjunto y compacto. Existe una sucesión finita o infinita de autovectores ortonormales $(x_n)_n$ de A cuya correspondiente sucesión de autovalores no nulos denotaremos por $(\lambda_n)_n$ tales que, para todo $x \in \mathbb{H}$*

$$Ax = \sum_n \lambda_n \langle x, x_n \rangle x_n, \tag{6.5.2}$$

donde n recorre todos los autovalores no nulos, incluida su multiplicidad[a]*. Además, se tiene que:*

1. *En (6.5.2) aparecen todos los autovalores no nulos de A.*
2. *Si la sucesión de autovalores no nulos $(\lambda_n)_n$ es infinita se puede reordenar de forma que $|\lambda_n| \xrightarrow{n\to\infty} 0$.*
3. *Los correspondientes espacios vectoriales*[b] *$\ker(\lambda_n I - A)$, $\lambda_n \neq 0$, $n = 1, 2, 3, \ldots$ son de dimensión finita, siendo dicha dimensión el número de veces que aparece un mismo λ_n en la fórmula (6.5.2).*

[a]Es decir, si el autovalor $\lambda_k \neq 0$ tiene asociado p_k autovectores, en (6.5.2) habrá un sumando para cada uno de los autovectores. Ver la fórmula equivalente (6.5.4).

[b]El espacio $\ker(\lambda_n I - A)$ no es más que el núcleo del operador $A - \lambda_n I$, i.e., el conjunto de $x \in \mathbb{H}$ tales que $(A - \lambda_n I)x = 0$, i.e., el subespacio vectorial asociado al autovalor λ_n.

Demostración: Probemos que la fórmula (6.5.2) es cierta. Para ello, en primer lugar, notemos que si algún λ_n fuese cero, entonces los correspondientes autovectores no aportarían nada a la suma en (6.5.2). También asumiremos que $A \neq \Theta$, pues en caso contrario el teorema sería trivial.

Utilizando la misma construcción que usamos en la prueba del Teorema 6.5.12 obtenemos una sucesión de autovalores $(\lambda_k)_{k=1}^n$ con sus correspondientes autovectores normalizados $(x_k)_{k=1}^n$ tales que[12]

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|.$$

[12]Para cada λ_k puede haber más de un autovector (aunque siempre un número finito, ver Teorema 6.5.10) en función de la *multiplicidad* p_k del mismo y que escogeremos, sin pérdida de generalidad, ortonormales. Por tanto, en la cadena de desigualdades podemos considerar que cada autovalor se repetirá según su multiplicidad algebraica, es decir, si al autovalor λ_k le corresponden p_k autovectores linealmente independientes, entonces dicho autovalor aparecerá p_k veces.

Además existe una cadena de subespacios cerrados (luego completos) e invariantes respecto a A (véase además la Nota 6.5.13)

$$\mathbb{H} = \mathbb{H}_1 \supseteq \mathbb{H}_2 \supseteq \cdots \supseteq \mathbb{H}_n$$

donde $\mathbb{H}_{k+1} = \{x \in \mathbb{H}_k \,|\, \langle x, x_j \rangle = 0,\ j = 1, 2, \dots, k\}$.

Supongamos que $A|_{\mathbb{H}_{n+1}} = 0$, entonces el proceso acaba (caso finito). Sea, en este caso[13], $y_n = x - \sum_{k=1}^{n} \langle x, x_k \rangle x_k$. Entonces, para todos $j = 1, 2, \dots, n$,

$$\langle y_n, x_j \rangle = \langle x, x_j \rangle - \sum_{k=1}^{n} \langle x, x_k \rangle \underbrace{\langle x_k, x_j \rangle}_{\delta_{j,k}} = \langle x, x_j \rangle - \langle x, x_j \rangle \underbrace{\|x_j\|^2}_{=1} = 0.$$

Luego, $y_n \in \mathbb{H}_{n+1}$ y, por tanto, $Ay_n = 0$, i.e.,

$$0 = Ay_n = Ax - \sum_{k=1}^{n} \langle x, x_k \rangle A x_k \ \Rightarrow\ Ax = \sum_{k=1}^{n} \lambda_k \langle x, x_k \rangle x_k,$$

como se quería probar.

El caso infinito es similar. Ante todo notemos que en este caso tenemos una sucesión λ_n tal que $|\lambda_n| \overset{n\to\infty}{\longrightarrow} 0$ (ver Teorema 6.5.12). Definamos nuevamente el vector $y_n = x - \sum_{k=1}^{n} \langle x, x_k \rangle x_k$. Usando (5.2.3) se sigue que $\|y_n\|^2 \leq \|x\|^2$. Además, como ya vimos $y_n \in \mathbb{H}_{n+1}$, luego, usando que $Ay_n = A|_{\mathbb{H}_{n+1}} y_n$, y $\|A|_{\mathbb{H}_{n+1}}\| = |\lambda_{n+1}|$ obtenemos

$$\|Ay_n\| \leq |\lambda_{n+1}| \|x\| \overset{n\to\infty}{\longrightarrow} 0 \ \Rightarrow\ \lim_{n\to\infty} Ay_n = 0,$$

de donde, como A es acotado, se sigue la igualdad buscada (6.5.2), pues[14]

$$0 = \lim_{n\to\infty} Ay_n = Ax - \lim_{n\to\infty} \sum_{k=1}^{n} \langle x, x_k \rangle A x_k = Ax - \sum_{k=1}^{\infty} \lambda_k \langle x, x_k \rangle x_k.$$

Comprobemos ahora que todos los autovalores de A están presentes en la fórmula (6.5.2). Esta cuestión es de suma importancia pues en la

[13] Si los autovalores tienen multiplicidad mayor que uno hay que incluir en la suma todos los autovectores correspondientes a cada autovalor. En cualquier caso dicha suma es finita pues a cada autovalor no nulo solo le pueden corresponder un número finito de autovectores por el Teorema 6.5.10.

[14] Que la suma infinita está bien definida se sigue de la acotación de Ay_n, aunque también se puede comprobar directamente como se hizo en el Ejemplo 6.2.9.

prueba de (6.5.2) solo aparecen los autovalores obtenidos gracias al Teorema 6.5.8 por lo que procede preguntarse si existen otros autovalores no nulos de A distintos a los anteriores. Comprobemos que eso es imposible. Supongamos que existe un autovalor $\lambda \neq 0$ distinto de los autovalores λ_k que aparecen en la suma (6.5.2), i.e., $\lambda \neq \lambda_k$, para todo k, y sea $x \neq 0$ su correspondiente autovector normalizado a la unidad. Por el Teorema 6.5.6, x es ortogonal a todos los x_k, i.e., $\langle x, x_k \rangle = 0$ para todo k. Pero entonces, aplicando la fórmula (6.5.2) a dicho autovector x tenemos

$$\lambda x = Ax = \sum_{n=1}^{\infty} \lambda_n \langle x, x_n \rangle x_n = \sum_{n=1}^{\infty} \lambda_n \, 0 \, x_n = 0,$$

i.e., o $\lambda = 0$ o $x = 0$, lo que es una contradicción. Que en la fórmula (6.5.2) aparezcan todos los autovalores no nulos de A tiene otra implicación muy importante y es que, formalmente, el proceso de encontrar los autovalores no nulos de A basado en el Teorema 6.5.8 permite encontrarlos ***todos***.

Por otro lado, del Teorema 6.5.10 sabemos que los autoespacios $\mathbb{X}_k = \ker(\lambda_k I - A)$, para cada $\lambda_k \neq 0$, $k = 1, 2, 3, \ldots$, son de dimensión finita. Probemos que, efectivamente, la dimensión de $\mathbb{X}_k = p_k$, donde p_k es el número de veces que aparece λ_k en la fórmula (6.5.2) (es decir, la multiplicidad del autovalor λ_k). Por simplicidad denotemos por $x_k^{(1)}, \ldots, x_k^{(p_k)}$, los p_k autovectores linealmente independientes asociados a λ_k. Supongamos que $\dim \mathbb{X}_k > p_k$, entonces existe al menos un $x \in \mathbb{X}_k$ que es linealmente independiente de los vectores $x_k^{(1)}, \ldots, x_k^{(p_k)}$ y que asumiremos, sin pérdida de generalidad, ortogonal a ellos (¿por qué?). Entonces está claro que dicho x es ortogonal a todos los vectores x_n que aparecen en la fórmula (6.5.2) (ya sea porque los x_n corresponden a autovalores distintos a λ_k o bien sean los $x_k^{(1)}, \ldots, x_k^{(p_k)}$ de antes), i.e., $\langle x, x_n \rangle = 0$ para todo n, luego usando (6.5.2) obtenemos

$$0 \neq \lambda_k x = Ax = \sum_{n} \lambda_n \langle x, x_n \rangle x_n = \sum_{n} \lambda_n \, 0 \, x_n = 0,$$

lo cual es una contradicción. ■

Como hemos visto en el teorema anterior, cuando $\lambda_k \neq 0$, $k = 1, 2, 3, \ldots$, la dimensión de $\mathbb{X}_k = \ker(\lambda_k I - A)$ es igual a p_k, donde p_k es el número de veces que aparece λ_k en la fórmula (6.5.2). Sea $\mathbb{X}_k = \operatorname{span}(x_{1,k}, \cdots, x_{p_k,k})$ el subespacio generado por los autovectores asociados a cada autovalor λ_k (que sabemos que son finito-dimensionales) y que tomaremos ortonormales entre sí (si no lo fuesen, aplicamos el método de Gram-Schmidt).

Sea el operador $P_k : \mathbb{H} \mapsto \mathbb{H}$ definido, para cada $k \in \mathbb{N}$ como

$$P_k x = \sum_{i=1}^{p_k} \langle x, x_{i,k}\rangle x_{i,k} = \sum_{i=1}^{p_k} c_{i,k} x_{i,k}. \tag{6.5.3}$$

Proposición 6.5.15 *El operador $P_k : \mathbb{H} \mapsto \mathbb{H}$ definido en (6.5.3) es el operador de proyección sobre el espacio $\mathbb{X}_k = \ker(\lambda_k I - A)$.*

Demostración: Que la imagen de P_k es $\mathbb{X}_k$ se deduce de la propia definición (6.5.3). Veamos que el operador P_k así definido es autoadjunto e idempotente y, por tanto, es un operador de proyección[15] (ver Problema 6.29).

Comencemos probando que P_k es autoadjunto. Para ello notemos que P_k es suma finita de operadores de rango 1 del tipo $\widehat{P}_i = \langle x, x_{i,k}\rangle x_{i,k}$ que, como hemos visto en la Proposición 6.4.8, son compactos. Entonces, por el Teorema 6.4.7 los P_k son compactos, luego acotados y, por tanto, existen sus correspondientes adjuntos P_k^*. Por otro lado, para todos $x, y \in \mathbb{H}$,

$$\begin{aligned}\langle \widehat{P}_i^* x, y\rangle &= \langle x, \widehat{P}_i y\rangle = \langle x, \langle y, x_i\rangle x_i\rangle = \overline{\langle y, x_i\rangle}\langle x, x_i\rangle = \langle x_i, y\rangle\langle x, x_i\rangle \\ &= \langle\langle x, x_i\rangle x_i, y\rangle = \langle \widehat{P}_i x, y\rangle,\end{aligned}$$

i.e., $\widehat{P}_i$ es autoadjunto, luego P_k lo es.

Que P_k es idempotente es *sencillo* de comprobar pues

$$\begin{aligned}P_k^2 x &= P_k\left(\sum_{i=1}^{p_k} c_{i,k} x_{i,k}\right) = \sum_{i=1}^{p_k} c_{i,k} P_k x_{i,k} = \sum_{i=1}^{p_k} c_{i,k} \sum_{j=1}^{p_k} \underbrace{\langle x_{i,k}, x_{j,k}\rangle}_{=\delta_{j,i}} x_{j,k} \\ &= \sum_{i=1}^{p_k} c_{i,k} x_{i,k} = P_k x.\end{aligned}$$

De hecho, $\mathfrak{I}(P_k) = \ker(\lambda_k I - A)$. ■

Usando los operadores P_k, el Teorema espectral 6.5.14 se puede reescribir como

$$\forall x \in \mathbb{H}, \qquad Ax = \sum_k \lambda_k \left(\sum_{i=1}^{p_k} \langle x, x_{i,k}\rangle x_{i,k}\right), \tag{6.5.4}$$

[15] Para la definición de operador de proyección ver el Problema 5.12.

donde, k recorre todos los autovalores $\lambda_k \neq 0$. La expresión anterior se puede escribir como sigue:

$$Ax = \sum_n \lambda_n P_n x, \tag{6.5.5}$$

donde, n recorre todos los autovalores $\lambda_n \neq 0$ siendo P_n los operadores de proyección sobre los subespacios $\ker(\lambda_n I - A)$ definidos por (6.5.3).

De hecho se tiene el siguiente resultado:

Teorema 6.5.16 *Sea $\mathbb{H}$ un espacio de Hilbert. Todo operador $A \in \mathcal{B}(\mathbb{H})$ autoadjunto y compacto se puede escribir de la forma*

$$A = \sum_n \lambda_n P_n, \tag{6.5.6}$$

donde n recorre todos los autovalores $\lambda_n \neq 0$, incluida su multiplicidad, siendo P_n los operadores de proyección sobre los subespacios $\ker(\lambda_n I - A)$ definidos por (6.5.3).

Demostración: Vamos a probarlo en el caso cuando hay infinitos autovalores pues en caso contrario es *trivial*. Para probar que la serie converge en $\mathcal{B}(\mathbb{H})$ basta probar que la sucesión de sus sumas parciales $A_k = \sum_{n=1}^k \lambda_n P_n$ es de Cauchy, pues el espacio $\mathcal{B}(\mathbb{H})$ es de Banach (ver Teorema 6.3.6). Para ello usamos que[16] $P_n = \langle x, x_n\rangle x_n$, luego

$$\left\| \sum_{n=m+1}^k \lambda_n P_n x \right\|^2 = \left\langle \sum_{n=m+1}^k \lambda_n P_n x, \sum_{i=m+1}^k \lambda_i P_i x \right\rangle = \sum_{n=m+1}^k \lambda_n^2 |\langle x, x_n\rangle|^2$$
$$\leq \lambda_{m+1}^2 \sum_{n=m+1}^k |\langle x, x_n\rangle|^2 \leq \lambda_{m+1}^2 \sum_{n=1}^\infty |\langle x, x_n\rangle|^2 \leq \lambda_{m+1}^2 \|x\|^2.$$

Como $\lambda_m \to 0$, entonces, cualquiera sea $\varepsilon > 0$, si tomamos m suficientemente grande podemos hacer $\lambda_m^2 < \varepsilon^2$. Luego $\left\| \sum_{n=m+1}^k \lambda_n P_n \right\| < \varepsilon$, i.e., A_k es de Cauchy, luego es convergente en la norma de operadores. Para

[16]Recordemos que en la suma hay que tener en cuenta la multiplicidad de los λ_n y que λ_n^2 son una sucesión monótona decreciente que tiende a cero (ver Teorema 6.5.12).

probar la fórmula (6.5.6) calculamos

$$\left\| Ax - \sum_{n=1}^{m} \lambda_n P_n x \right\|^2 = \left\langle \sum_{n=m+1}^{\infty} \lambda_n P_n x, \sum_{n=m+1}^{\infty} \lambda_n P_n x \right\rangle = \sum_{n=m+1}^{\infty} \lambda_n^2 |\langle x, x_n\rangle|^2$$
$$\leq \lambda_{m+1}^2 \|x\|^2 \leq \varepsilon^2 \|x\|^2 \;\Rightarrow\; \left\| \sum_{n=m+1}^{\infty} \lambda_n P_n \right\| < \varepsilon,$$

pues, $\lambda_n \stackrel{n\to\infty}{\longrightarrow} 0$. Luego $A_k \to A$ en $\mathcal{B}(\mathbb{H})$. ■

Conviene hacer notar que el conjunto de autovectores $(x_n)_n$ de un operador autoadjunto y compacto en un espacio de Hilbert $\mathbb{H}$ cualquiera mencionado que aparecen en el Teorema espectral 6.5.14 no tiene por que constituir una base ortonormal completa de $\mathbb{H}$. No obstante, si $\mathbb{H}$ es separable se tiene el siguiente importante resultado:

Teorema 6.5.17 (Hilbert-Schmidt) *Sea $\mathbb{H}$ un espacio de Hilbert separable y $A \in \mathcal{B}(\mathbb{H})$ un operador autoadjunto y compacto. Entonces, existe un sistema ortogonal completo (base ortonormal) de vectores ortonormales $(e_n)_n$ de $\mathbb{H}$ consistente en los correspondientes autovectores de A, incluido los asociados al autovalor $\lambda_0 = 0$, caso que lo tuviera.*

Demostración: Del Teorema espectral 6.5.14 se sigue que existe un conjunto de autovectores ortonormales (finito o infinito) $(\phi_n)_n$ tal que

$$Ax = \sum_{n} \lambda_n \langle x, \phi_n\rangle \phi_n, \qquad \forall x \in \mathbb{H}, \tag{6.5.7}$$

donde n recorre todos los autovalores no nulos $\lambda_n \neq 0$, incluida su multiplicidad.

Como el núcleo de A, $\ker A \subset \mathbb{H}$ (que coincide con el subespacio vectorial generado por los autovectores correspondientes a $\lambda_0 = 0$) es a su vez un espacio de Hilbert separable ($\ker A$ es cerrado –Ejercicio 4.15– y por tanto completo –Teorema 3.5.14– y separable –Ejercicio 3.4.7–) existirá un sistema (numerable) ortonormal completo en $\ker(A)$ que denotaremos por $(\psi_n)_n$ (Teorema 5.2.9). Dicho sistema de vectores estará constituido por los autovectores correspondientes al autovalor $\lambda_0 = 0$. Sea ahora un autovector cualquiera ϕ_m correspondiente al autovalor $\lambda_m \neq 0$. Entonces ϕ_m será ortogonal a todos los ψ_n y el sistema $(\phi_m)_m$ será ortogonal a $(\psi_n)_n$.

Además de (6.5.7) se tiene que para todo $x \in \mathbb{H}$

$$A\left(x - \sum_m \langle x, \phi_m\rangle \phi_m\right) = Ax - \sum_m \langle x, \phi_m\rangle A\phi_m = 0,$$

i.e., $x - \sum_m \langle x, \phi_m\rangle\phi_m \in \ker A$, luego podemos desarrollarlo en serie de Fourier en la base $(\psi_n)_n$ del subespacio $\ker A$ de forma que para todo $x \in \mathbb{H}$ tendremos

$$x - \sum_m \langle x, \phi_m\rangle\phi_m = \sum_n \langle x, \psi_n\rangle \psi_n \Rightarrow x = \sum_m \langle x, \phi_m\rangle\phi_m + \sum_n \langle x, \psi_n\rangle\psi_n,$$

i.e., el sistema ortonormal $(\phi_m)_m \bigcup (\psi_n)_n$ es un sistema completo (Teorema 5.2.5) de donde se sigue el teorema. ■

Nótese que si denotamos por $(\phi_n)_n$ a la sucesión de autovectores ortonormales (incluyendo la multiplicidad) de A, y por $(\psi_m)_m$ la base ortonormal del núcleo de A, que coincide con el autoespacio asociado al autovalor 0, entonces el corolario anterior nos dice que cualquier $x \in \mathbb{H}$ se puede escribir como combinación lineal del sistema ortonormal $(\phi_n)_n \bigcup (\psi_m)_m$ y, por tanto, existe un único $y \in \mathcal{N}(A)$ tal que

$$x = y + \sum_n \langle x, \phi_n\rangle\phi_n, \quad \langle y, \phi_n\rangle = 0, \ \forall n. \tag{6.5.8}$$

Si denotamos por $\mathbb{X}_0 = \ker(A)$ el subespacio vectorial correspondiente al autovalor cero (es decir, el núcleo del operador A) y por $\mathbb{X}_k$ los correspondientes autoespacios asociados a cada autovalor λ_k – es decir los subespacios $\mathbb{X}_k := \ker(\lambda_k I - A)$–, entonces el Teorema de Hilbert-Schmidt se puede reescribir como la suma directa[17]

$$\mathbb{H} = \mathbb{X}_0 \oplus \mathbb{X}_1 \oplus \mathbb{X}_2 \oplus \cdots \mathbb{X}_k \oplus \cdots, \tag{6.5.9}$$

siendo la suma finita o infinita en función de si hay un número finito o infinito de autovalores distintos.

Otra consecuencia del Teorema de Hilbert-Schmidt 6.5.17 es que todo operador lineal $A : \mathbb{H} \mapsto \mathbb{H}$ autoadjunto y compacto en $\mathbb{H}$, espacio de Hilbert separable, se le puede hacer corresponder una matriz (finita o infinita) que además es diagonalizable y en cuya diagonal aparecen los correspondientes autovalores.

[17]Véase la Nota 5.2.17 en la página 197.

Supongamos ahora que dos operadores $A, B \in \mathcal{B}(\mathbb{H})$ comparten una base completa común de autovectores. Es decir, para todo n se tiene que $A\phi_n = \alpha_n\phi_n$ y $B\phi_n = \beta_n\phi_n$. Entonces, $(AB - BA)\phi_n = 0$ para todo n y, por tanto, $(AB - BA) = 0$ (ver Problema 6.6). Osea, si dos operadores comparten una base completa de autovectores, entonces A y B conmutan. Como consecuencia del Teorema de Hilbert-Schmidt 6.5.17 se tiene el recíproco de dicho resultado para operadores autoadjuntos y compactos.

Teorema 6.5.18 *Sea $\mathbb{H}$ un espacio de Hilbert separable y $A, B \in \mathcal{B}(\mathbb{H})$ dos operadores autoadjuntos y compactos. Si A y B conmutan, entonces ambos tienen un sistema completo común de autovectores.*

Demostración: Sea λ un autovalor de A y sea $\mathbb{X}_\lambda$ el subespacio vectorial generado por los autovectores de A correspondientes a λ, i.e., $\mathbb{X}_\lambda := \ker(\lambda I - A)$, $\dim \mathbb{X}_\lambda < +\infty$. Para todo $x \in \mathbb{X}_\lambda$ tenemos $ABx = BAx = \lambda Bx$, i.e., Bx es también es un autovector de A correspondiente a λ, $Bx \in \mathbb{X}_\lambda$, así que el espacio $\mathbb{X}_\lambda \subset \mathbb{H}$ es un subespacio vectorial de $\mathbb{H}$ invariante respecto a B y es a su vez un espacio de Hilbert ($\mathbb{X}_\lambda$, es un subespacio cerrado –ver Problema 4.15– y, por tanto, completo –ver Teorema 3.5.14). Como B es autoadjunto y compacto, el Teorema de Hilbert-Schmidt (Teorema de Hilbert-Schmidt 6.5.17) nos asegura que $\mathbb{X}_\lambda$ tiene una base ortonormal de autovectores de B, que además son autovectores de A pues están en $\mathbb{X}_\lambda$. Repitiendo en proceso para cada uno de los subespacios $\mathbb{X}_\lambda$ de A, incluyendo el caso $\lambda = 0$, obtenemos la correspondiente base de autovectores. La unión de todas ellas es la base común buscada.■

Teorema espectral para operadores normales y compactos

En este apartado formularemos y probaremos el correspondiente teorema espectral para operadores normales y compactos. Comenzaremos demostrando dos propiedades importantes de los mismos.

Teorema 6.5.19 *Sea $\mathbb{H}$ un espacio de Hilbert separable y $T \in \mathcal{B}(\mathbb{H})$ un operador normal y compacto. Entonces T tiene al menos un conjunto numerable de autovalores.*

Demostración: Sea $T : \mathbb{H} \mapsto \mathbb{H}$ normal y compacto. Como consecuencia de los Problema 6.2 y 6.23 sabemos que existen dos operadores $A, B \in$

$\mathcal{B}(\mathbb{H})$, autoadjuntos, compactos[18] y que conmutan entre sí, tales que $T = A + iB$ (ver fórmula (6.6.3)). Por otro lado, el Teorema 6.5.18 establece que existe una base ortonormal completa $(x_n)_n$ de autovectores (incluidos los correspondientes al autovalor nulo) común a los operadores A y B, i.e, existen la sucesión (finita o infinita) de vectores no nulos $(x_n)_n$ y los conjuntos de números reales[19] (finitos o infinitos) $(\alpha_n)_n$ y $(\beta_n)_n$ tales que $Ax_n = \alpha_n x_n$ y $Bx_n = \beta_n x_n$, luego $Tx_n = (\alpha_n + i\beta_n)x_n = \lambda_n x_n$. Luego, T tiene al menos un conjunto numerable de autovalores. ∎

Teorema 6.5.20 *Sea $\mathbb{H}$ un espacio de Hilbert y $T \in \mathcal{B}(\mathbb{H})$, un operador normal. Entonces los autovectores correspondientes a autovalores distintos son ortogonales.*

Demostración: En efecto, sea $Tx_1 = \lambda_1 x_1$ y $Tx_2 = \lambda_2 x_2$ con $\lambda_1 \neq \lambda_2$. Entonces, por un lado tenemos $\langle Tx_1, Tx_2\rangle = \langle x_1, T^*Tx_2\rangle = |\lambda_2|^2\langle x_1, x_2\rangle$, y, por el otro, $\langle Tx_1, Tx_2\rangle = \langle T^*Tx_1, x_2\rangle = |\lambda_1|^2\langle x_1, x_2\rangle$. Restando ambas expresiones obtenemos $(|\lambda_1|^2 - |\lambda_2|^2)\langle x_1, x_2\rangle = 0$, de donde se sigue el resultado. ∎

Una consecuencia del resultado anterior es que, al ser tanto los operadores autoadjuntos como los unitarios casos particulares de los operadores normales, para ellos también se tendrá que los autovectores correspondientes a autovalores distintos son ortogonales, resultado que probamos antes para el caso de los autoadjuntos en el Teorema 6.5.6.

En la prueba del Teorema 6.5.19 vimos que $T = A + iB$, donde los operadores A y B son autoadjuntos, compactos y comparten un sistema ortonormal completo de autovectores cuyos elementos son también autovectores de T. Denotemos, como en el caso del Teorema espectral 6.5.14, por $(x_n)_n$ los autovectores correspondientes a los autovalores $\lambda_n \neq 0$ no nulos[20] (incluida su multiplicidad) de T, y por $(\psi_m)_m$ los correspondientes al autovalor nulo $\lambda_0 = 0$ que son una base del $\ker(A)$. Con esta notación tenemos que $Tx_n = \lambda_n x_n$, $\lambda_n \neq 0$ y $T\psi_m = 0$.

[18]La compacidad se deduce de los Teoremas 6.4.7 y 6.4.11 pues, como vimos en el Problema 6.2, $A = (T + T^*)/2$ y $B = (T - T^*)/2i$.

[19]Dado que los espacios nulos de A y B pueden ser no vacíos, pueden existir vectores x_k tales que $Ax_k = 0$ y/o $Bx_k = 0$.

[20]Nótese que $\lambda_n = 0$ si y solo si $\alpha_n = \beta_n = 0$, donde α_n y β_n son los autovalores de los operadores A y B.

Entonces, usando que para todo $x \in \mathbb{H}$ se tiene el desarrollo

$$x = \sum_n \langle x, x_n \rangle x_n + \sum_m \langle x, \psi_m \rangle \psi_m,$$

y que T es acotado (luego continuo) obtenemos, para todo $x \in \mathbb{H}$, la siguiente fórmula para calcular Tx:

$$Tx = \sum_n \lambda_n \langle x, x_n \rangle x_n, \tag{6.5.10}$$

donde la suma recorre todos los autovalores λ_n no nulos, incluida su multiplicidad[21]. La fórmula (6.5.10) es el análogo de la fórmula espectral (6.5.2) para los operadores normales y compactos.

Nótese además que al ser T compacto, entonces, por un lado, los autoespacios $\mathbb{X}_k = \ker(\lambda_k I - T)$ del operador T son, para cada $\lambda_k \neq 0$, $k = 1, 2, 3, \ldots$, de dimensión finita (Teorema 6.5.10) y, por el otro, que si la sucesión de autovalores $(\lambda_n)_n$ es infinita, entonces se puede reordenar de forma que $|\lambda_n| \overset{n\to\infty}{\longrightarrow} 0$ (Teorema 6.5.11).

Por otro lado, repitiendo el mismo razonamiento que usamos en la prueba del punto 1 del Teorema espectral 6.5.14 se puede comprobar que en la fórmula (6.5.10) están presentes todos los autovalores no nulos de T. Esta cuestión es muy importante pues para obtener la expresión (6.5.10) hemos usado los autovectores de las componentes cartesianas A y B de T y podría ocurrir que existiesen otros autovalores de T no nulos distintos a ellos.

Finalmente, usando el mismo razonamiento que usamos para probar el punto 3 del Teorema espectral 6.5.14 obtenemos que la dimensión del $\ker(\lambda_k I - A)$, $\lambda_k \neq 0$, $k = 1, 2, 3, \ldots$, es el número de veces que aparece λ_k en la fórmula (6.5.10) (es decir, la multiplicidad del autovalor λ_k).

De esta forma tenemos el siguiente resultado para los operadores normales y compactos:

Teorema 6.5.21 (Teorema espectral) *Sea $\mathbb{H}$ un espacio de Hilbert separable y $T \in \mathcal{B}(\mathbb{H})$ un operador normal y compacto. Entonces existe una sucesión numerable (finita o infinita) de autovectores ortonormales $(x_n)_n$ de T cuya correspondiente sucesión de autovalores no nulos denotaremos por $(\lambda_n)_n$ tales que, para todo x de $\mathbb{H}$, Tx se puede calcular por la fór-*

[21] Si un autovalor es cero, ni dicho autovalor ni sus correspondientes autovectores (que pertenecerán al núcleo de T) aparecerán en la suma (6.5.10).

mula (6.5.10) *donde* n *recorre todos los autovalores no nulos, incluida su multiplicidad. Además, se tiene que:*

1. *En* (6.5.10) *aparecen todos los autovalores de* T.

2. *Si la sucesión de autovalores no nulos* $(\lambda_n)_n$ *es infinita se puede reordenar de forma que* $|\lambda_n| \stackrel{n\to\infty}{\longrightarrow} 0$.

3. *Los correspondientes espacios vectoriales* $\ker(\lambda_n I - T)$*,* $\lambda_n \neq 0$*,* $\lambda_n \neq 0$*,* $n = 1, 2, 3, \ldots$ *son de dimensión finita, siendo dicha dimensión el número de veces que aparece un mismo* λ_n *en la fórmula* (6.5.10).

Está claro de la prueba del teorema anterior, que el conjunto de todos los autovectores de un operador normal y compacto $T \in \mathcal{B}(\mathbb{H})$, $\mathbb{H}$ espacio de Hilbert separable, incluido los asociados al autovalor $\lambda = 0$, constituyen una base completa de vectores ortonormales de $\mathbb{H}$. Es decir, se tiene también la descomposición (6.5.9). Otra consecuencia del Teorema espectral 6.5.21 es que todo operador normal y compacto también se puede escribir de la forma (6.5.6).

Antes de terminar este apartado conviene hacer una aclaración. El lector avispado recordará el Ejemplo 6.2.9, donde estudiamos el operador definido por (6.2.5). Si comparamos (6.2.5) con (6.5.2) podemos ver que, efectivamente, existen operadores no necesariamente autoadjuntos (o normales) y compactos que admiten una representación del tipo (6.2.5). En el caso del Ejemplo 6.2.9 basta que $\mu_n \not\to 0$ cuando $n \to \infty$ para que no sea compacto, o que alguno de los μ_n no sea real para que no sea autoadjunto. De hecho, se puede probar una representación *similar* a (6.2.5) para operadores autoadjuntos acotados (no necesariamente compactos) e incluso no acotados. El lector interesado puede consultar, por ejemplo, [11, Capítulos 9 y 10] o [22, Capítulo 13].

Finalmente, queremos hacer notar que el Problema 6.32 muestra una aplicación interesante del Teorema espectral 6.5.14. De hecho, como muestran los Problemas 6.33 y 6.34, es suficiente con saber que existe una descomposición del tipo (6.5.5) para poder resolver un sinnúmero de problemas interesantes.

6.6. Problemas

Problema 6.1 Sea T, $T: L^2_{[0,1]} \mapsto L^2_{[0,1]}$, el operador integral del problema[22] 4.9:

$$y = Tx, \qquad y(t) = \int_0^1 k(t,\tau)x(\tau)d\tau, \tag{6.6.1}$$

con $k(t,\tau)$ continua en el cuadrado $[0,1]\times[0,1]$. Demuestra que el adjunto de dicho operador es el operador $T^*: C([0,1]) \mapsto C([0,1])$, $z = T^*x$

$$z(t) = \int_0^1 \overline{k(\tau,t)}x(\tau)d\tau. \tag{6.6.2}$$

Luego $T^* = T$ si y solo si $\overline{k(\tau,t)} = k(t,\tau)$.

Solución: Usando el Teorema de Fubini tenemos

$$\begin{aligned}\langle Tx, y\rangle &= \int_0^1 \left(\int_0^1 k(t,\tau)x(\tau)d\tau\right)\overline{y(t)}dt = \int_0^1\int_0^1 k(t,\tau)x(\tau)\overline{y(t)}d\tau dt \\ &= \int_0^1 x(\tau)\overline{\left(\int_0^1 \overline{k(t,\tau)}y(t)dt\right)}d\tau = \langle x, T^*y\rangle \;\Rightarrow\; \text{se tiene (6.6.2)}. \quad \blacksquare\end{aligned}$$

Problema 6.2 Sea $\mathbb{H}$ un espacio de Hilbert y $T \in \mathcal{B}(\mathbb{H})$. Entonces existen dos únicos operadores autoadjuntos[23] A, B de $\mathbb{H}$ en $\mathbb{H}$ tales que

$$T = A + iB, \qquad A = \frac{T+T^*}{2}, \quad B = \frac{T-T^*}{2i}. \tag{6.6.3}$$

Solución: Como T es un operador lineal y acotado en un espacio de Hilbert, existe su adjunto T^*. Un cálculo directo usando la Proposición 6.1.4 se tiene que los operadores A y B definidos mediante la fórmula anterior son efectivamente autoadjuntos. Probemos que la descomposición $T = A + iB$ es única. Para ello supongamos que existen otros operadores autoadjuntos $A \neq \tilde{A}$ y $B \neq \tilde{B}$ tales que $T = \tilde{A} + i\tilde{B}$. Entonces $T^* = \tilde{A}^* - i\tilde{B}^* = \tilde{A} - i\tilde{B}$, luego $T + T^* = 2\tilde{A}$ y $T - T^* = 2i\tilde{B}$, de donde se sigue que $A = \tilde{A}$ y $B = \tilde{B}$, lo cual es una contradicción. $\blacksquare$

Problema 6.3 Sea $\mathbb{H}$ un espacio de Hilbert, $T \in \mathcal{B}(\mathbb{H})$ y $A \in \mathcal{B}(\mathbb{H})$ un operador autoadjunto. Prueba que T^*AT es autoadjunto.

[22] Recordemos que por $L^2_{[0,1]}$ denotamos el completamiento del espacio $C^2_{[0,1]}$.
[23] A la descomposición (6.6.3) se denomina descomposición cartesiana de un operador.

Solución: Es inmediato de la definición de operador autoadjunto y la propiedad 7 de la Proposición 6.1.4. ■

Problema 6.4 Sea $\mathbb{H}$ un espacio de Hilbert y $A \in \mathcal{B}(\mathbb{H})$. Se dice que A es antihermítico si $A^* = -A$. Prueba que A es antihermítico si y solo si el operador $B = iA$ es hermítico (autoadjunto).

Solución: Si $A^* = -A$, entonces multiplicando por i y usando la Proposición 6.1.4 tenemos $(iA)^* = iA$, luego B es autoadjunto. Que B autoadjunto implica que A es antihermítico es inmediato. ■

Problema 6.5 Sean $U, V : \mathbb{H} \mapsto \mathbb{H}$, dos operador unitarios en un espacio de Hilbert $\mathbb{H} \neq \{0\}$. Prueba que

1. U es una isometría, i.e., $\|Ux\| = \|x\|$ para todo $x \in \mathbb{H}$.
2. $\|U\| = 1$.
3. Su inverso U^{-1} es unitario.
4. UV es unitario
5. U es normal.

Solución: Recordemos que U es unitario si $U^* = U^{-1}$. Para probar 1 tenemos

$$\|Ux\|^2 = \langle Ux, Ux\rangle = \langle x, U^*Ux\rangle = \langle x, U^{-1}Ux\rangle = \langle x, x\rangle = \|x\|^2.$$

2. Se sigue de lo anterior pues $\|Ux\|/|x\| = 1$, para todo $x \in \mathbb{H} \setminus \{0\}$. Veamos 3. Como U es biyectivo U^{-1} lo es, además, como U es unitario

$$U^{-1} = U^* \Rightarrow (U^{-1})^* = (U^*)^* = U = (U^{-1})^{-1} \;\Rightarrow\; (U^{-1})^{-1} = (U^{-1})^*.$$

4. $(UV)^*(UV) = (V^*U^*)(UV) = V^*(U^*U)V = V^*V = I$. 5. Como $U^{-1} = U^*$, entonces $U^*U = UU^* = I$. ■

Problema 6.6 Sea $(\phi_n)_n$ una base completa del espacio de Hilbert $\mathbb{H}$ y sean dos operadores lineales y acotados A, B definidos sobre $\mathbb{H}$. Prueba que si $A\phi_n = B\phi_n$, para todo n, entonces $A = B$.

Solución: Sea $x \in \mathbb{H}$ cualquiera y sea $(\phi_n)_n$ una base completa de $\mathbb{H}$. Definamos el operador $C = A - B$, que es lineal y acotado (luego continuo). Además C es tal que $C\phi_n = 0$ para todo n. Pero entonces, cualquiera sea $x \in \mathbb{H}$, $x = \sum_n x_n \phi_n$, luego $Cx = C\sum_n x_n\phi_n = \sum_n x_n C\phi_n = \sum_n x_n \cdot 0 = 0$ (donde hemos usado la continuidad de C). Luego $C = 0$ y, por tanto, $A = B$. ■

Problema 6.7 Sean $(e_n)_n$ y $(f_n)_n$ bases ortonormales de los espacios de Hilbert $\mathbb{H}$ y $\mathbb{H}'$, respectivamente y sea $A : \mathbb{H} \mapsto \mathbb{H}'$ lineal y acotado. Prueba que la serie $\sum_{n=1}^{\infty} \|Ae_n\|^2$ converge si y solo si lo hace la serie $\sum_{m=1}^{\infty} \|A^* f_m\|^2$, en cuyo caso se tiene que

$$\sum_{n=1}^{\infty} \|Ae_n\|^2 = \sum_{m=1}^{\infty} \|A^* f_m\|^2. \tag{6.6.4}$$

Deduce, de lo anterior, que la cantidad $\sum_{n=1}^{\infty} \|Ae_n\|^2$ es independiente de la base ortogonal $(e_n)_n$ escogida.

Solución: Como $A \in \mathcal{B}(\mathbb{H}, \mathbb{H}')$, entonces existe el operador adjunto, $A^* \in \mathcal{B}(\mathbb{H}', \mathbb{H})$ y se tiene (6.1.1), para todo $x \in \mathbb{H}$ e $y \in \mathbb{H}'$. Usando la identidad de Parseval (5.2.5) en $\mathbb{H}$ con $x = A^* f_m$ y $\phi_n = e_n$, tenemos, para todo $m = 1, 2, 3, \cdots$,

$$\|A^* f_m\|^2 = \sum_{k=1}^{\infty} |\langle A^* f_m, e_k\rangle|^2 = \sum_{k=1}^{\infty} \left|\langle f_m, Ae_k\rangle'\right|^2 = \sum_{k=1}^{\infty} \left|\langle Ae_k, f_m\rangle'\right|^2, \tag{6.6.5}$$

donde en la última igualdad hemos usado la propiedad de simetría del producto escalar. Análogamente, la identidad de Parseval en $\mathbb{H}'$ nos da

$$\|Ae_n\|^2 = \sum_{j=1}^{\infty} \left|\langle Ae_n, f_j\rangle'\right|^2$$

Sumando en m en (6.6.5) y usando la expresión anterior obtenemos

$$\sum_{m=1}^{N} \|A^* f_m\|^2 = \sum_{m=1}^{N}\sum_{k=1}^{\infty} \left|\langle Ae_k, f_m\rangle'\right|^2 \leq \sum_{m=1}^{\infty}\sum_{k=1}^{\infty} \left|\langle Ae_k, f_m\rangle'\right|^2 = \sum_{n=1}^{\infty} \|Ae_n\|^2.$$

De lo anterior se sigue que si, $\sum_{n=1}^{\infty} \|Ae_n\|^2 < \infty$, entonces $\sum_{m=1}^{\infty} \|A^* f_m\|^2 < \infty$, también lo es. Por otro lado,

$$\sum_{n=1}^{N} \|Ae_n\|^2 = \sum_{n=1}^{N}\sum_{j=1}^{\infty} \left|\langle Ae_n, f_j\rangle'\right|^2 \leq \sum_{n=1}^{\infty}\sum_{j=1}^{\infty} \left|\langle Ae_n, f_j\rangle'\right|^2 = \sum_{j=1}^{\infty} \|A^* f_m\|^2,$$

donde para la última igualdad hemos usado (6.6.5). Luego, si $\sum_{m=1}^{\infty} \|A^* f_m\|^2 < \infty$, entonces $\sum_{n=1}^{\infty} \|Ae_n\|^2 < \infty$. Tomando $N \to \infty$ en las dos desigualdades anteriores obtenemos las desigualdades

$$\sum_{m=1}^{\infty} \|A^* f_m\|^2 \leq \sum_{n=1}^{\infty} \|Ae_n\|^2 \quad \text{y} \quad \sum_{n=1}^{\infty} \|Ae_n\|^2 \leq \sum_{m=1}^{\infty} \|A^* f_m\|^2,$$

respectivamente, de donde se sigue (6.6.4).

Probemos ahora que la suma $\sum_{n=1}^{\infty} \|Ae_n\|^2$ no depende de la base escogida. Sea $(\widetilde{e}_n)_n$ otra base ortonormal de $\mathbb{H}$ distinta de $(e_n)_n$. Entonces, si aplicamos la igualdad anterior primero a las bases $(\widetilde{e}_n)_n$ y $(f_n)_n$, y luego a las bases $(f_n)_n$ y $(e_n)_n$, deducimos que

$$\sum_{n=1}^{\infty} \|A\widetilde{e}_n\|^2 = \sum_{m=1}^{\infty} \|A^* f_m\|^2 = \sum_{n=1}^{\infty} \|Ae_n\|^2 \,,$$

luego, la cantidad $\sum_{n=1}^{\infty} \|Ae_n\|^2$ es independiente de la base ortonormal $(e_n)_n$ escogida. ■

Problema 6.8 Sea $A : \mathbb{H} \mapsto \mathbb{H}'$ un operador lineal y acotado, $\mathbb{H}$ y $\mathbb{H}'$ espacios de Hilbert. Se dice que A es un operador del tipo Hilbert-Schmidt[24] si existe una base ortonormal completa $(\phi_n)_n$ en $\mathbb{H}$ tal que $\sum_{n=1}^{\infty} \|A\phi_n\|^2 < +\infty$. Prueba que:

1. El espacio de todos los operadores del tipo Hilbert-Schmidt son un subespacio vectorial del espacio $\mathcal{B}(\mathbb{H}, \mathbb{H}')$ de los operadores lineales y acotados, i.e, prueba que si A y B son del tipo Hilbert-Schmidt, entonces $A + B$ y λA también lo son.

2. A es del tipo Hilbert-Schmidt si y solo si A^* lo es.

Solución: 1. Para probar que la suma $A + B$ es del tipo Hilbert-Schmidt usamos que

$$\|(A + B)\phi_n\|^2 = \|A\phi_n + B\phi_n\|^2 \leq 2(\|A\phi_n\| + \|B\phi_n\|^2),$$

que se sigue directamente de la ley del paralelogramo (5.1.3) tomando $a = A\phi_n$ y $b = B\phi_n$, y, por tanto, la serie $\sum_{n=1}^{\infty} \|(A+B)\phi_n\|^2$ converge. Que λA es del tipo Hilbert-Schmidt es inmediato ya que $\|(\lambda A)\phi_n\| = |\lambda| \|A\phi_n\|$.

2. Basta usar el resultado del Problema 6.7. ■

[24]Como hemos visto en el problema anterior 6.7 la suma $\sum_{n=1}^{\infty} \|A\phi_n\|^2$ es independiente de la base.

Problema 6.9 Prueba que el operador integral T del Problema 6.1, cuando $k(t,\tau)$ satisface la condición

$$\int_0^1 \int_0^1 |k(t,\tau)|^2 dt d\tau < +\infty, \tag{6.6.6}$$

es un operador del tipo Hilbert-Schmidt (ver Problema 6.8).

Solución: Sea $(\phi_n)_n$ una base ortonormal de $L^2_{[0,1]}$.

$$(T\phi_n)(t) = \int_0^1 k(t,\tau)\phi_n(\tau)d\tau = \langle k_t, \overline{\phi_n}\rangle,$$

donde $k_t(\tau) := k(t,\tau)$, es decir, consideramos t como un parámetro y τ es la variable. Ahora bien, $(\overline{\phi_n})_n$ es a su vez una base completa[25] de $L^2_{[0,1]}$, así que usando la identidad de Parseval (5.2.5) para la función k_t, como función de τ, tenemos

$$\begin{aligned}\sum_{n=1}^{\infty} \|T\phi_n\|^2 &= \sum_{n=1}^{\infty} \int_0^1 |\langle k_t, \overline{\phi_n}\rangle|^2 dt = \int_0^1 \left(\sum_{n=1}^{\infty} |\langle k_t, \overline{\phi_n}\rangle|^2\right) dt = \int_0^1 \|k_t\|^2 dt \\ &= \int_0^1 \left(\int_0^1 |k_t(\tau)|^2 d\tau\right) dt = \int_0^1 \int_0^1 |k_t(\tau)|^2 d\tau dt < +\infty.\end{aligned}$$

Nótese que en la secuencia de igualdades hemos intercambiado la serie con la integral, lo cual no es obvio en general, no obstante si $k(t,\tau)$ es una función continua en un compacto (como es el caso) es cierto, o si $k(t,\tau)$ es integrable podemos usar el Teorema de la convergencia dominada. Finalmente, para la última igualdad podemos usar el Teorema de Fubini que nos asegura que si se tiene (6.6.6), entonces $k_t \in L^2_{[0,1]}$ y además se tiene la última igualdad de la expresión anterior. ∎

Problema 6.10 Prueba que si $A \in \mathcal{B}(\mathbb{H})$ es autoadjunto, entonces cualquier polinomio con coeficientes reales de A, i.e., $P_n(A) = a_0 I + a_1 A + \cdots + a_n A^n$, $a_0, \ldots, a_n \in \mathbb{R}$, también es acotado y autoadjunto.

Solución: Es inmediato del Ejercicio 6.2.3. ∎

Problema 6.11 *Prueba que si $A \in \mathcal{B}(\mathbb{H})$, $\mathbb{H}$ espacio de Hilbert, entonces $\mathcal{N}(A) = (\mathcal{I}(A^*))^{\perp}$. Es decir, el (espacio) núcleo de un operador lineal acotado coincide con el complemento ortogonal de la imagen (rango) de su adjunto A^*.*

[25] Esto es consecuencia, por ejemplo, del apartado 4 del Teorema 5.2.5.

Solución: Como $A \in \mathcal{B}(\mathbb{H})$, entonces existe el adjunto A^* de A. Sea $x \in \mathcal{N}(A)$, i.e., $Ax = 0$. por tanto,

$$\langle x, A^*y\rangle = \langle Ax, y\rangle = 0,$$

i.e., si $x \in \mathcal{N}(A)$, entonces $x \perp A^*y$ para todo $y \in \mathbb{H}$ y, por tanto, $x \in (\mathfrak{I}(A^*))^\perp$, luego, $\mathcal{N}(A) \subseteq (\mathfrak{I}(A^*))^\perp$. Por otro lado, si $x \perp \mathfrak{I}(A^*)$, entonces

$$\forall y \in \mathbb{H}, \quad 0 = \langle x, A^*y\rangle = \langle Ax, y\rangle \;\Rightarrow\; \text{(punto 4 del Ejercicio 5.1.2)} \quad Ax = 0,$$

i.e., $x \in \mathcal{N}(A)$, luego $(\mathfrak{I}(A^*))^\perp \subseteq \mathcal{N}(A)$, de donde se sigue el resultado. ■

Problema 6.12 Sea $\mathbb{H}$ un espacio de Hilbert sobre $\mathbb{C}$, y $A \in \mathcal{B}(\mathbb{H})$. Se dice que A es positivo si $\langle Az, z\rangle \geq 0$ para todo $z \in \mathbb{H}$ y se denota por $A \geq 0$. Prueba las siguientes afirmaciones:

1. Si $A \geq 0$ entonces A es autoadjunto (ver Teorema 6.2.2).
2. Si $A \geq 0$, $B \geq 0$, y $\alpha \geq 0$, entonces $A + B \geq 0$ y $\alpha A \geq 0$.
3. Si $A \geq 0$ y $T \in \mathcal{B}(\mathbb{H})$, entonces $T^*AT \geq 0$.
4. Cualquiera sea $A \in \mathcal{B}(\mathbb{H})$, $A^*A \geq 0$.
5. Sean $A, B \in \mathcal{B}(\mathbb{H})$ tales que $A \geq 0$, $B \geq 0$ y $A + B = \Theta$, entonces $A = B = \Theta$ (Θ es el operador nulo del Ejemplo 4.4.3).
6. Si $A \geq 0$ entonces, para todos $x, y \in \mathbb{H}$ se cumple
$$|\langle Ax, y\rangle|^2 \leq \langle Ax, x\rangle\langle Ay, y\rangle.$$
7. Si $A \geq 0$, entonces $Ax = 0$ si y solo si $\langle Ax, x\rangle = 0$ (usa el punto anterior).

Ayuda para probar 6: Usa que $A \geq 0$, i.e., $\langle Av, v\rangle \geq 0$, para todo $v \in \mathbb{H}$. Elige $v = x + \lambda\langle Ax, y\rangle y$, con $\lambda \in \mathbb{R}$ y razona como en la prueba del Teorema 5.1.6.

Solución: 1. Si A es positivo, entonces $\langle Ax, x\rangle \geq 0$, es decir $\langle Ax, x\rangle \in \mathbb{R}$. Entonces el Teorema 6.2.2 nos asegura que A es autoadjunto. No obstante daremos una prueba general que no necesita del Teorema 6.2.2. Por un lado tenemos la identidad[26]

$$\langle Ax, y\rangle = \frac{1}{4}\Big(\langle A(x+y), x+y\rangle - \langle A(x-y), x-y\rangle \\ + i\big[\langle A(x+iy), x+iy\rangle - \langle A(x-iy), x-iy\rangle\big]\Big),$$

[26]Cuya comprobación, que es directa usando las propiedades del producto escalar (ver, e.g., la ecuación (6.2.1)), dejamos como ejercicio al lector.

y, por el otro, intercambiando x e y, obtenemos

$$\begin{aligned}\langle Ay, x\rangle &= \frac{1}{4}\Big([\langle A(y+x), y+x\rangle - \langle A(y-x), y-x\rangle] \\ &\qquad + i\langle A(y+ix), y+ix\rangle - i\langle A(y-ix), y-ix\rangle\Big) \\ &= \frac{1}{4}\Big([\langle A(x+y), x+y\rangle - \langle A(x-y), x-y\rangle] \\ &\qquad - i\big[\langle A(x+iy), x+iy\rangle - \langle A(x-iy), x-iy\rangle\big]\Big) = \overline{\langle Ax, y\rangle}.\end{aligned}$$

La última igualdad se sigue de comparar el valor obtenido de $\langle Ay, x\rangle$ con en valor de $\langle Ax, y\rangle$ obtenido antes pues, por hipótesis, las expresiones entre corchetes, son reales (recuérdese que $\langle Az, z\rangle \geq 0$ para todo $z \in \mathbb{H}$, en particular para $z = x \pm iy$ y $z = x \pm y$). Luego tenemos,

$$\langle Ax, y\rangle = \overline{\langle Ay, x\rangle} = \langle x, Ay\rangle,$$

de donde se sigue que $A = A^*$.

2. $\langle (A+B)x, x\rangle = \langle Ax, x\rangle + \langle Bx, x\rangle \geq 0$, $\langle (\alpha A)x, x\rangle = \alpha\langle Ax, x\rangle \geq 0$.

3. $\langle T^*ATx, x\rangle = \langle ATx, Tx\rangle = \langle A(Tx), (Tx)\rangle = \langle Az, z\rangle \geq 0$.

4. $\langle A^*Ax, x\rangle = \langle Ax, Ax\rangle \geq 0$.

5. Se sigue de

$$0 = \langle (\Theta)x, x\rangle\langle (A+B)x, x\rangle = \langle Ax, x\rangle + \langle Bx, x\rangle \geq 0,$$

pues, por hipótesis, para todo $x \in \mathbb{H}$, $\langle Ax, x\rangle \geq 0$ y $\langle Bx, x\rangle \geq 0$, luego $\langle Ax, x\rangle = 0$ y $\langle Bx, x\rangle = 0$ para todo $x \in \mathbb{H}$, y del corolario 6.2.7 (pues por el punto 1 A y B son autoadjuntos) se sigue que A y B son el operador nulo.

6. Efectivamente, calculamos $\langle Av, v\rangle \geq 0$, con $v = x + \lambda\langle Ax, y\rangle y$, y $\lambda \in \mathbb{R}$

$$0 \leq \langle Av, v\rangle = \langle Ax, x\rangle + \lambda^2|\langle Ax, y\rangle|^2\langle Ay, y\rangle + \lambda\big(\langle Ax, y\rangle\langle Ay, x\rangle + \overline{\langle Ax, y\rangle}\langle Ax, y\rangle\big).$$

Como A es positivo, entonces es autoadjunto (ver punto 1), luego $\langle Ax, y\rangle\langle Ay, x\rangle = \langle Ax, y\rangle\overline{\langle x, Ay\rangle} = \langle Ax, y\rangle\overline{\langle Ax, y\rangle}$, de donde tenemos

$$\lambda^2|\langle Ax, y\rangle|^2\langle Ay, y\rangle + 2\lambda|\langle Ax, y\rangle|^2 + \langle Ax, x\rangle \geq 0.$$

El miembro derecho de la desigualdad anterior es un polinomio en λ de grado 2, así que su discriminante ha de ser menor o igual a cero, luego

$$4|\langle Ax, y\rangle|^4 \leq 4|\langle Ax, y\rangle|^2\langle Ay, y\rangle\langle Ax, x\rangle,$$

de donde se sigue el resultado[27].

7. Si $Ax = 0$ es obvio que $\langle Ax, x\rangle = 0$. Si $\langle Ax, x\rangle = 0$, entonces el punto 6 anterior nos dice que, para todo $y \in \mathbb{H}$, $\langle Ax, y\rangle = 0$, luego el punto cuatro del Ejercicio 5.1.2 implica que $Ax = 0$. ■

Problema 6.13 Prueba que el operador T definido en el Ejemplo 6.2.9 es positivo si y solo si para todo n, $\mu_n \geq 0$. Otro ejemplo de operador positivo es el operador multiplicación (ver Ejemplo 6.5.4) cuando $f(t) \geq 0$ en $[a, b]$.

Solución: Usando las expresiones (6.2.4) y (6.2.7) obtenemos que, para todo $x \in \mathbb{H}$,

$$\langle Tx, x\rangle = \sum_{k=1}^{\infty} \mu_k |c_k|^2,$$

de donde se sigue que, si $\mu_k \geq 0$ para todo k, T es positivo. Por otro lado, usando (6.2.3), y teniendo en cuenta que $T \geq 0$, tenemos que, para todo $n \in \mathbb{N}$,

$$0 \leq \langle T\phi_n, \phi_n\rangle = \mu_n,$$

de donde se sigue el resultado. Para el caso del operador de multiplicación

$$\langle Tx, x\rangle = \int_a^b f(t)|x(t)|^2 d\tau \geq 0,$$

si $f(t) \geq 0$ en $[a, b]$. ■

Problema 6.14 Sea $\mathbb{H}$ un espacio de Hilbert, $A, B \in \mathcal{B}(\mathbb{H})$ y A (o B) compacto. Entonces los operadores AB y BA son compactos.

Solución: Sea $(x_n)_n$ una sucesión acotada de $\mathbb{H}$. Como B es acotado, entonces la sucesión $(Bx_n)_n$ es acotada y como A es compacto, de la sucesión $(ABx_n)_n$ se puede extraer una subsucesión convergente, luego AB es compacto. Sea ahora $(x_n)_n$ una sucesión acotada de $\mathbb{H}$. Como A es compacto, existe una subsucesión $(Ax_{n_k})_k$ de $(Ax_n)_n$ que converge. Ahora bien, B es acotado, luego es continuo (ver Teorema 4.4.18), por lo que la subsucesión $(BAx_{n_k})_k$ converge, luego BA es compacto. ■

Problema 6.15 Sea $A \in \mathcal{B}(\mathbb{H})$, $\mathbb{H}$ espacio de Hilbert, y $\mathfrak{I}(A) = \mathbb{H}$. Prueba que si A tiene inverso, y dicho inverso es acotado, entonces el adjunto A^* es invertible y $(A^*)^{-1} = (A^{-1})^*$. Si además, A es autoadjunto, entonces A^{-1} también lo es.

[27] Se asume que $A \neq \Theta$.

Solución: Probaremos que $(A^{-1})^*A^*x = A^*(A^{-1})^*x = x$ para todo $x \in \mathbb{H}$ de donde se deduce el resultado. Sea $y \in \mathbb{H}$ cualquiera, entonces, como A^{-1} es acotado, existe $(A^{-1})^*$, y, por tanto

$$\begin{aligned}\langle y, (A^{-1})^*A^*x\rangle &= \langle A^{-1}y, A^*x\rangle = \langle AA^{-1}y, x\rangle = \langle y, x\rangle \\ &= \langle A^{-1}Ay, x\rangle = \langle Ay, (A^{-1})^*x\rangle = \langle y, A^*(A^{-1})^*x\rangle,\end{aligned}$$

de donde, usando la propiedad 4 del Ejercicio 5.1.2 se sigue que $(A^{-1})^*A^*x = x = A^*(A^{-1})^*x$. Si A es autoadjunto, entonces $A = A^*$, luego $(A^{-1})^* = (A^*)^{-1} = A^{-1}$, luego A^{-1} es autoadjunto. ■

Problema 6.16 Prueba que si $A \in \mathcal{B}(\mathbb{H})$ es compacto y $\mathbb{H}$ es de dimensión infinita, entonces, si existe el inverso de A, A^{-1} no puede ser acotado.

Solución: Supongamos que A es invertible y su inversa A^{-1}, es acotada. Entonces, $AA^{-1} = A^{-1}A = I$. Pero por el Ejercicio 6.14 sabemos que si A es compacto entonces AA^{-1} y $A^{-1}A$ tienen que ser compactos, pero I, no es compacto cuando $\mathbb{H}$ es de dimensión infinita (ver Ejercicio 6.4.3). ■

Problema 6.17 Prueba que $A : \mathbb{H} \mapsto \mathbb{H}$, $\mathbb{H}$ espacio de Hilbert, es compacto si y solo si A^*A es compacto.

Solución: Si A es compacto, A^* es compacto (por el Teorema 6.4.11) luego su producto es compacto (basta con que A^* sea acotado, ver Problema 6.14). Supongamos que A^*A es compacto y sea $(x_n)_n$ una sucesión acotada. Entonces existe una subsucesión $(y_k)_k$ de $(x_n)_n$ tal que $(A^*Ay_k)_k$ es convergente y, por tanto, de Cauchy, i.e., si n, m, son lo suficientemente grandes, $\|A^*Ay_n - A^*Ay_m\| < \varepsilon/(2M)$, donde $M = \sup_k \|y_k\|$. Entonces, si n, m son lo suficientemente grandes,

$$\begin{aligned}\|Ay_n - Ay_m\|^2 &= \langle A(y_n - y_m), A(y_n - y_m)\rangle = \langle A^*A(y_n - y_m), y_n - y_m\rangle \\ &\leq \|A^*A(y_n - y_m)\|\|y_n - y_m\| \leq 2M\|A^*A(y_n - y_m)\| < \varepsilon,\end{aligned}$$

donde hemos usado la desigualdad de Cauchy-Schwarz. O sea, $(Ay_k)_k$ es de Cauchy, luego es convergente pues $\mathbb{H}$ es completo. ■

Problema 6.18 Prueba que si A es un operador del tipo Hilbert-Schmidt, entonces es compacto. El recíproco no es cierto. Encuentra un ejemplo de operador compacto que no sea de Hilbert-Schmidt.

Solución: Sea A un operador de tipo Hilbert-Schmidt (ver Problema 6.8) y sea $(e_n)_n$ una base ortonormal de $\mathbb{H}$. Entonces, $x = \sum_{n=1}^{\infty}\langle x, e_n\rangle e_n$. Definamos la sucesión de operadores $A_k : \mathbb{H} \mapsto \mathbb{H}$

$$A_k x = A_k\left(\sum_{n=1}^{\infty}\langle x, e_n\rangle e_n\right) = \sum_{n=1}^{k}\langle x, e_n\rangle Ae_n.$$

Nótese que A_k definidos de esta manera son operadores de rango finito pues $\dim \mathfrak{I}(A_k) \leq k$, para cada $k \in \mathbb{N}$, y por tanto compactos (ver Ejercicio 6.4.9). Probemos que $A_k \to A$ cuando $k \to \infty$ en la norma de operadores, luego por el Teorema 6.4.10, A es compacto. Tenemos que

$$Ax - A_k x = \sum_{n=k+1}^{\infty}\langle x, e_n\rangle e_n \;\Rightarrow\; \|(A - A_k)x\| \leq \sum_{n=k+1}^{\infty}|\langle x, e_n\rangle|\|Ae_n\|.$$

Usando la desigualdad de Hölder (3.7.3) con $p = 2$,

$$\begin{aligned}
\|(A - A_k)x\| &\leq \left(\sum_{n=k+1}^{\infty}|\langle x, e_n\rangle|^2\right)^{\frac{1}{2}}\left(\sum_{n=k+1}^{\infty}\|Ae_n\|^2\right)^{\frac{1}{2}} \\
&\leq \left(\sum_{n=1}^{\infty}|\langle x, e_n\rangle|^2\right)^{\frac{1}{2}}\left(\sum_{n=k+1}^{\infty}\|Ae_n\|^2\right)^{\frac{1}{2}} = \|x\|^2\left(\sum_{n=k+1}^{\infty}\|Ae_n\|^2\right)^{\frac{1}{2}},
\end{aligned}$$

donde, para la última desigualdad hemos usado la identidad de Parseval (5.2.5). Luego, tomado el supremo en $\|x\| = 1$ obtenemos

$$\|A - A_k\| \leq \left(\sum_{n=k+1}^{\infty}\|Ae_n\|^2\right)^{\frac{1}{2}} \to 0, \quad k \to \infty,$$

pues $\sum_{n=k+1}^{\infty}\|Ae_n\|^2$ es la cola de la serie convergente $\sum_{n=1}^{\infty}\|Ae_n\|^2$ ya que A es un operador de tipo Hilbert-Schmidt.

Sea el operador $T : \ell^2 \mapsto \ell^2$ definido en el 6.2.9. Si $\mu_n \to 0$, T es compacto (véase el Problema 6.19). Sea la sucesión $(\mu_n)_n$, $\mu_n = 1/\sqrt{n} \to 0$. Dado que $\sum_{n=1}^{\infty}\|T\phi_n\|^2 = \sum_{n=1}^{\infty}1/n$ diverge, T no puede ser de tipo Hilbert-Schmidt. ∎

Problema 6.19 Prueba que el operador T definido en el Ejemplo 6.2.9 es compacto si y solo si $\mu_n \to 0$.

Solución: Sea $(\phi_n)_n$ una base ortonormal de $\mathbb{H}$ completa de $\mathbb{H}$. Si T es compacto, entonces por el Teorema 6.5.9, $T\phi_n \to 0$, luego, como $T\phi_n = \mu_n\phi_n$, $\mu_n \to 0$.

Supongamos ahora que $\mu_n \to 0$ cuando $n \to \infty$. Entonces, podemos definir una sucesión T_k de operadores de rango finito y, por tanto, compactos (véase Ejercicio 6.4.9), similar a la usada en el Ejemplo 6.18, i.e.,

$$T_k x = T_k\left(\sum_{n=1}^{\infty}\langle x,\phi_n\rangle\phi_n\right) = \sum_{n=1}^{k}\langle x,\phi_n\rangle T\phi_n = \sum_{n=1}^{k}\mu_n\langle x,\phi_n\rangle\phi_n.$$

Entonces

$$\begin{aligned}\|(T-T_k)x\|^2 &\leq \sum_{n=k+1}^{\infty}|\mu_n|^2|\langle x,\phi_n\rangle|^2 \leq \left(\sup_{n>k}|\mu_n|^2\right)\sum_{n=1}^{\infty}|\langle x,\phi_n\rangle|^2\\ &= \left(\sup_{n>k}|\mu_n|^2\right)\|x\|^2 \;\Rightarrow\; \|T-T_k\| \leq \left(\sup_{n>k}|\mu_n|^2\right)^{1/2}.\end{aligned}$$

Pero como $\mu_n \to 0$, entonces $\sup_{n>k}|\mu_n|^2 \to 0$, de donde, por el Teorema 6.4.10, se deduce el resultado. ■

Problema 6.20 Prueba que el operador $T : L^2_{[0,1]} \mapsto L^2_{[0,1]}$, definido en el Problema 6.1, es compacto si se cumple la condición (6.6.6). **Ayuda:** Usa los Problemas 6.9 y 6.18.

Problema 6.21 Sea $A : \mathbb{H} \mapsto \mathbb{H}$, $\mathbb{H}$, espacio de Hilbert real, un operador autoadjunto. Prueba, sin usar el Corolario 6.2.7, que si $\langle Ax, x\rangle = 0$, para todo $x \in \mathbb{H}$, entonces $A = \Theta$, i.e, A es el operador nulo.

Solución: En general, el resultado se sigue del Corolario 6.2.7. Si $\mathbb{H}$ es un espacio de Hilbert complejo, entonces podemos usar el Lema 6.2.1. Hagamos una prueba directa cuando $\mathbb{H}$ es real. Notemos que, para todos $x, y \in \mathbb{H}$, tenemos, definiendo el vector $v = x + y$, que

$$\begin{aligned}0 &= \langle Av, v\rangle = \langle A(x+y), x+y\rangle = \langle Ax, x\rangle + \langle Ay, y\rangle + \langle Ax, y\rangle + \langle Ay, x\rangle\\ &= \langle Ax, y\rangle + \langle Ay, x\rangle = \langle Ax, y\rangle + \langle x, Ay\rangle = \langle Ax, y\rangle + \langle A^*x, y\rangle = 2\langle Ax, y\rangle.\end{aligned}$$

Luego, por la propiedad 4 del Ejercicio 5.1.2 tenemos que $Ax = 0$ para todo $x \in \mathbb{H}$, luego A es el operador nulo $A = \Theta$. ■

Problema 6.22 Prueba que un operador acotado $T : \mathbb{H} \mapsto \mathbb{H}$, $\mathbb{H}$, espacio de Hilbert es normal si y solo si $\|Tx\| = \|T^*x\|$ para todo $x \in \mathbb{H}$.

Solución: En efecto, si T es normal $TT^* = T^*T$, luego, para todo $x \in \mathbb{H}$,

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \langle T^*Tx, x\rangle \overset{TT^*=T^*T}{=} \langle TT^*x, x\rangle = \langle T^*x, T^*x\rangle = \|T^*x\|^2.$$

Nótese que si para todo $x \in \mathbb{H}$, $\|Tx\| = \|T^*x\|$ entonces, de la desigualdad anterior deducimos que

$$\|Tx\|^2 = \langle T^*Tx, x\rangle = \langle TT^*x, x\rangle = \|T^*x\|^2 \;\Rightarrow\; \langle (T^*T - TT^*)x, x\rangle = 0, \quad \forall x \in \mathbb{H}.$$

Pero el operador $A = (T^*T - TT^*)$ es autoadjunto pues $A^* = (T^*T)^* - (TT^*)^* = T^*T - TT^* = A$, luego el Corolario 6.2.7 implica $A = \Theta$, de donde se sigue que $T^*T = TT^*$, i.e., T es normal. ∎

Problema 6.23 Prueba que un operador acotado $T : \mathbb{H} \mapsto \mathbb{H}$, $\mathbb{H}$, espacio de Hilbert, es normal si y solo si los operadores autoadjuntos A y B de su forma cartesiana (6.6.3) conmutan, i.e., $AB = BA$.

Solución: Usando (6.6.3) tenemos

$$TT^* = A^2 + B^2 + i(BA - AB), \quad T^*T = A^2 + B^2 - i(BA - AB) \;\Rightarrow$$

$$TT^* - T^*T = 2i(BA - AB).$$

Luego, si T es normal $TT^* = T^*T$ por tanto $AB = BA$, y si $AB = BA$, entonces $TT^* = T^*T$. ∎

Problema 6.24 Prueba que si $T : \mathbb{H} \mapsto \mathbb{H}$, $T \in \mathcal{B}(\mathbb{H})$, $\mathbb{H}$ espacio de Hilbert, es normal, entonces $\|T^n\| = \|T\|^n$, $n \in \mathbb{N}$.

Solución: Asumiremos que no es el operador nulo $T \neq \Theta$. Está claro que $\|T^n\| \leq \|T\|^n$ (ver Problema 4.13). Probemos que $\|T^n\| \geq \|T\|^n$, de donde se seguiría el resultado. Para ello probaremos que $\|T^n x\| \geq \|Tx\|^n$, para todo $x \in \mathbb{H}$ con $\|x\| = 1$. Como el operador T es acotado, T^* lo es y, por tanto, T^*T también. Nótese que

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \langle T^*Tx, x\rangle \leq \|T^*Tx\|\|x\| = \|T^*(Tx)\| = \|T^2x\|,$$

donde hemos usado que $\|x\| = 1$ y además que como T es normal, entonces $\|Ty\| = \|T^*y\|$ para todo y (ver Problema 6.22). Está claro de lo anterior que $\|T^n x\| \geq \|Tx\|^n$ para $n = 1, 2$. Supongamos que es cierto para $n = m$, i.e., $\|T^m x\| \geq \|Tx\|^m$, y probémoslo para $n = m + 1$. Supongamos que no para todo $x \in \mathbb{H}$ $Tx = 0$ (si no, sería el caso trivial $T = \Theta$). Para todo $x \in \mathbb{H}$ tales que

$Tx \neq 0$ definamos el vector $y = Tx/\|Tx\| \in \mathbb{H}$, $\|y\| = 1$. Tenemos, usando la hipótesis de inducción $\|T^m y\| \geq \|Ty\|^m$, que

$$\|T^{m+1}x\| = \|Tx\|\|T^m y\| \geq \|Tx\|\|Ty\|^m = \|Tx\|^{1-m}\|T^2 x\|^m \geq \|Tx\|^{1-m}\|Tx\|^{2m} \\ = \|Tx\|^{m+1},$$

de donde se sigue que $\|T^n x\| \geq \|Tx\|^n$, luego (recordemos que $\|x\| = 1$) $\|T^n\| \geq \|T\|^n$ y, por tanto $\|T^n x\| = \|Tx\|^n$. ∎

Problema 6.25 Sea $\mathbb{X}$ un espacio de Banach y $A \in \mathcal{B}(\mathbb{X})$. Entonces, la resolvente $\rho(A)$ es un abierto de $\mathbb{C}$ y el espectro $\sigma(A)$ es un compacto de $\mathbb{C}$ contenido en el disco cerrado $D = \{z;\ |z| \leq \|A\|\}$.

Solución: Sea un λ_0 cualquiera perteneciente al conjunto resolvente $\rho(A)$, entonces existe el operador $A_{\lambda_0} := (\lambda_0 I - A)^{-1}$. Sea $\lambda \in \mathbb{C}$ cualquiera. Un cálculo directo nos muestra que

$$I - (\lambda_0 I - A)^{-1}(\lambda I - A) = (\lambda_0 - \lambda)(\lambda_0 I - A)^{-1} \quad \Rightarrow$$

$$S_\lambda := I - (\lambda_0 - \lambda)(\lambda_0 I - A)^{-1} = (\lambda_0 I - A)^{-1}(\lambda I - A).$$

Por otro lado, para todos los λ tales que $|\lambda - \lambda_0| < 1/\|A_{\lambda_0}\|$, se tiene que

$$\|(\lambda_0 - \lambda)(\lambda_0 I - A)^{-1}\| = |\lambda_0 - \lambda|\|A_{\lambda_0}\| < 1.$$

Luego, por el Teorema 6.3.2, para todos dichos λ el operador S_λ es invertible y, por tanto, el operador $\lambda I - A = (\lambda_0 I - A)S_\lambda$ también lo será, i.e., dichos $\lambda \in \rho(A)$. Es decir, para todo $\lambda_0 \in \rho(A)$ existe una bola abierta $B(\lambda_0, 1/\|A_{\lambda_0}\|) \subset \rho(A)$ centrada en λ_0, luego $\rho(A)$ es abierto. Como $\sigma(A) = \mathbb{C} \setminus \rho(A)$, $\sigma(A)$ es cerrado.

Puesto que $\sigma(A)$ es cerrado, para demostrar que es compacto basta probar que es acotado (ver Teorema 4.3.8). Para ello procederemos como sigue: escojamos λ tal que $|\lambda| > \|A\|$, entonces, por el Teorema 6.3.3 $(\lambda I - A)$ es invertible. Luego, dichos $\lambda \notin \sigma(A)$, i.e., $\sigma(A) \subset D = \{z;\ |z| \leq \|A\|\}$. Nótese que la acotación de $\sigma(A)$ también se sigue directamente del Ejercicio 6.5.5. ∎

Problema 6.26 Sea $\mathbb{H}$ un espacio de Hilbert, sea $A : \mathbb{H} \mapsto \mathbb{H}$ un operador lineal en $\mathbb{H}$ y sea $\sigma(A)$ el espectro de A. Prueba que

1. $\sigma(A) = \overline{\sigma(A^*)}$.
2. $\sigma(A^{-1}) = \{\lambda^{-1} :\ 0 \neq \lambda \in \sigma(A)\}$.

3. $[\sigma(A)]^n = \sigma(A^n)$, para $n = 1, 2, 3, \dots$.

Solución: Para el primero usamos que $(A - \lambda I)^* = A^* - \overline{\lambda} I$, de donde se sigue el resultado. Para la segunda implicación tenemos que, si existe A^{-1}, entonces

$$(A^{-1} - \lambda I) = -\lambda A^{-1}(A - \lambda^{-1} I),$$

luego $\lambda \in \sigma(A^{-1})$ si y solo si $\lambda^{-1} \in \sigma(A)$.

Veamos la tercera. Para ello notamos que

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

Probemos que si $(\lambda I - A)$ no es invertible, entonces $\lambda^n I - A^n$ tampoco lo será. Para ello usaremos que un operador es invertible si y solo si su núcleo solo contiene el elemento $x = 0$ (ver Teorema 4.4.9). Dado que $\lambda I - A$ no es invertible, existe $x \neq 0$ tal que $(\lambda I - A)x = 0$, y por tanto, para ese x, $(\lambda^n I - A^n)x = 0$. Luego $\lambda^n \in \sigma(A^n)$, y, por tanto, $[\sigma(A)]^n \subset \sigma(A^n)$. Por otro lado, sea $\mu \in \sigma(A^n)$ y sea λ alguna de las raíces n-ésimas de μ, entonces de la factorización anterior se sigue que al menos alguna de dichas λ está en el espectro de A, $\lambda \in \sigma(A)$, luego $\sigma(A^n) \subset [\sigma(A)]^n$. ■

Problema 6.27 Sea el operador lineal $A : \mathbb{H} \mapsto \mathbb{H}$, $\mathbb{H}$ espacio de Hilbert, y sea λ un autovalor del mismo. Prueba que si A es positivo, entonces $\lambda \geq 0$ y si A es unitario entonces $|\lambda| = 1$.

Solución: Si A es positivo, entonces $0 \leq \langle Ax, x \rangle = \lambda \|x\|$, luego $\lambda \geq 0$. Si U es unitario, $\|x\|^2 = \|Ux\|^2 = |\lambda|^2 \|x\|^2$, de donde se sigue el resultado. ■

Problema 6.28 Sea $T \in \mathcal{B}(\mathbb{H})$ un operador normal, $\mathbb{H}$ espacio de Hilbert. Entonces $Tx = \lambda x$ si y solo si $T^* x = \overline{\lambda} x$.

Solución: Sea $T : \mathbb{H} \mapsto \mathbb{H}$ normal. Como el conjunto de los operadores normales es un subespacio vectorial de $\mathcal{B}(\mathbb{H})$ (ver Ejercicio 6.2.3), entonces $T - \lambda I$ es normal. Usando el resultado del Problema 6.22 se tiene que, para todo x de $\mathbb{H}$,

$$\|Tx - \lambda x\| = \|(T - \lambda I)x\| = \|(T^* - \overline{\lambda} I)x\| = \|T^* x - \overline{\lambda} x\| \quad \Longrightarrow$$

$\|Tx - \lambda x\| = 0$ si y solo si $\|T^* x - \overline{\lambda} x\| = 0$, de donde se sigue el resultado. ■

Problema 6.29 Prueba que un operador lineal acotado es de proyección (ver Problema 5.12) si y solo si es autoadjunto e idempotente.

Solución: Si P es de proyección sobre un cerrado $M \subset \mathbb{H}$, P es lineal, acotado e idempotente (ver Problema 5.12). Probemos que P es autoadjunto. Para ello notemos que, por definición, si $x \in \mathbb{H}$ entonces, $x = y+y^{\perp}$, con $y \in M$, $y^{\perp} \in M^{\perp}$, $Px = y$, y si $z \in \mathbb{H}$, $z = w + w^{\perp}$, $w \in M$, $w^{\perp} \in M^{\perp}$, $Pz = w$. Entonces, para todos $x, z \in \mathbb{H}$,

$$\langle Px, z - Pz\rangle = \langle y, w^{\perp}\rangle = 0, \quad \text{pues } y \in M,\ w^{\perp} \in M^{\perp}.$$

Probemos que $P = P^*$. Para ello notamos que

$$\langle Px, z\rangle = \langle Px, Pz\rangle + \underbrace{\langle Px, z - Pz\rangle}_{=0} = \langle Px, Pz\rangle + \underbrace{\langle x - Px, Pz\rangle}_{=0} = \langle x, Pz\rangle.$$

Supongamos que $P : \mathbb{H} \to \mathbb{H}$ es autoadjunto e idempotente. Definamos el conjunto $M = \{x \in \mathbb{H} : Px = x\}$. Está claro que M es no vacío ($x = 0 \in M$). Como P es acotado, entonces P es continuo, así que si $x_n \to x$, se tiene $Px_n \to Px$. Sea la sucesión $(x_n)_n \in M$. Entonces $x_n = Px_n \to Px = x$, luego $x \in M$, luego M es cerrado. Ahora bien, $Px \in M$, pues $P^2x = P(Px) = Px$ (es idempotente), luego $Py = y$, $y = Px \in M$. Probemos que $x - Px \in M^{\perp}$, luego para todo $x \in \mathbb{H}$ tendremos que $x = y + y^{\perp}$, $y = Px \in M$ y $y^{\perp} \in M^{\perp}$, i.e., P es de proyección sobre el espacio M. Sea $z \in M$ cualquiera. Entonces

$$\langle x - Px, z\rangle = \langle x, z\rangle - \langle Px, z\rangle = \langle x, z\rangle - \langle x, Pz\rangle = \langle x, z\rangle - \langle x, z\rangle = 0,$$

pues P es autoadjunto y $Pz = z$, por construcción. ■

Problema 6.30 Sea $A : \mathbb{H} \mapsto \mathbb{H}$, $\mathbb{H}$ espacio de Hilbert, un operador lineal autoadjunto y compacto y sea P_n el operador de proyección al subespacio de los autovectores asociados al autovalor λ_n. Prueba que para todo n, $AP_n = \lambda_n P_n$. Usando lo anterior prueba que para $k = 1, 2, \dots$, $A^k x = \sum_n \lambda_n^k P_n x$ para todo $x \in \mathbb{H}$. De hecho se tiene que $A^k = \sum_n \lambda_n^k P_n$.

Solución: Que $AP_n x = \lambda_n P_n x$, para todo $x \in \mathbb{H}$, se sigue directamente de (6.5.3). Luego, para todo $x \in \mathbb{H}$, $\|(AP_n - \lambda_n P_n)x\| = 0$, lo que significa que $\|AP_n - \lambda_n P_n\| = \sup_{\|x\|=1} \|(AP_n - \lambda_n P_n)x\| = 0$, luego $AP_n - \lambda_n P_n = 0$, en la norma de operadores. Para la segunda parte, basta notar que,

$$A^2x = A(Ax) = A\left(\sum_n \lambda_n \langle x, x_n\rangle x_n\right) = \sum_n \lambda_n \langle x, x_n\rangle Ax_n = \sum_n \lambda_n^2 \langle x, x_n\rangle x_n,$$

donde hemos usado la fórmula (6.5.2), la continuidad de A y que $Ax_n = \lambda_n x_n$. Para el resto de las potencias A^k se procede de forma análoga mediante inducción. La convergencia en norma de operadores de que $A^k = \sum_n \lambda_n^k P_n$ se demuestra de forma análoga a como se hizo para demostrar (6.5.6) (el caso infinito, pues en el caso de una suma finita es similar al primer apartado del problema). Esta claro además que todos los operadores A^k son autoadjuntos y compactos (¿por qué?). ■

Problema 6.31 Sea $A : \mathbb{H} \mapsto \mathbb{H}$, $\mathbb{H}$ espacio de Hilbert, un operador lineal autoadjunto y compacto. Prueba que si para cierto $k \in \mathbb{N}$, $A^k = \Theta$ (operador nulo), entonces A es el operador nulo. **Ayuda:** Prueba que A^k es autoadjunto y compacto y usa el Teorema espectral 6.5.14 para probar que todos autovalores de A^k son cero, de donde se deduce, de nuevo usando el Teorema espectral que $A = \Theta$.

Solución: Que A^k es autoadjunto se deduce de la Proposición 6.2.4 y que es compacto del Problema 6.14. Como $A^k = \Theta$, entonces los autovalores μ_n de A^k son los mismos de Θ, pero para Θ obviamente la ecuación $\Theta x = 0 = \mu_n x$ tiene como única solución $\mu_n = 0$ para todo $x \neq 0$ de $\mathbb{H}$. Pero entonces, usando (6.5.2), y razonando como en el problema anterior

$$Ax = \sum_n \lambda_n \langle x, x_n \rangle x_n \;\Rightarrow\; A^k x = \sum_n \lambda_n^k \langle x, x_n \rangle x_n,$$

donde λ_n son los autovalores de A, entonces $0 = \mu_n = \lambda_n^k$, luego los autovalores de A son todos nulos y usando nuevamente (6.5.2) obtenemos que $A = 0$. ■

Problema 6.32 Sea $A : \mathbb{H} \mapsto \mathbb{H}$, $\mathbb{H}$ espacio de Hilbert, un operador lineal autoadjunto y compacto y sea $f : \mathbb{R} \mapsto \mathbb{R}$ una función real tal que $f(\lambda) \to 0$ si $\lambda \to 0$, $f(0) = 0$. En estas condiciones definiremos la función $f(A)$ de la siguiente forma

$$f(A) = \sum_{n=1}^{\infty} f(\lambda_n) P_n, \quad A = \sum_{n=1}^{\infty} \lambda_n P_n, \tag{6.6.7}$$

donde P_n son los operadores de proyección ortogonal definidos en (6.5.3). Prueba que el operador $f(A)$ está bien definido y que es lineal, autoadjunto y compacto. Como aplicación encuentra el operador $\sqrt{A}$ de un operador compacto y definido positivo A.

Solución: Probar que $f(A)$ está bien definido es casi igual que la prueba de la fórmula (6.5.6), i.e., hay que probar que la sucesión de sumas parciales $f_i(A) =$

$\sum_{n=1}^{i} f(\lambda_n)P_n$ es de Cauchy. Así, para todos $k > m$,

$$\|f_k(A)(x) - f_m(A)(x)\|^2 = \left\| \sum_{n=m+1}^{k} f(\lambda_n)P_n x \right\|^2$$
$$= \left\langle \sum_{n=m+1}^{k} f(\lambda_n)P_n x, \sum_{i=m+1}^{k} f(\lambda_n)P_i x \right\rangle = \sum_{n=m+1}^{k} f(\lambda_n)^2 |\langle x, x_n\rangle|^2.$$

Como $\lambda_n \to 0$, y $f(\lambda_n) \to 0$ entonces si tomamos n suficientemente grande podemos hacer $|f(\lambda_n)| < \varepsilon$, cualquiera sea $\varepsilon > 0$. Luego, para todo $n > N$, $f(\lambda_n)^2 < \varepsilon^2$, cualquiera sea $\varepsilon > 0$. Así, para todo $n > N$,

$$\|f_k(A)(x) - f_m(A)(x)\|^2 = \sum_{n=m+1}^{k} f(\lambda_n)^2 |\langle x, x_n\rangle|^2 \le \varepsilon^2 \|x\|^2.$$

Luego $\|f_k(A) - f_m(A)\| < \varepsilon$, i.e., $f_i(A) = \sum_{n=1}^{i} f(\lambda_n)P_n$ es de Cauchy, luego es convergente, por tanto

$$\lim_{N\to\infty} \sum_{n=1}^{N} f(\lambda_n)P_n = \sum_{n=1}^{\infty} f(\lambda_n)P_n = f(A),$$

en la norma de operadores. La linealidad se sigue de la linealidad de los operadores $f_k(A)$.

Que $f(A)$ es autoadjunto se deduce directamente del Teorema 6.2.5. Una prueba directa usando la continuidad del producto escalar es como sigue:

$$\langle f(A)x, y\rangle = \sum_{n=1}^{\infty} \langle f(\lambda_n)P_n x, y\rangle = \sum_{n=1}^{\infty} f(\lambda_n)\langle P_n x, y\rangle = \sum_{n=1}^{\infty} f(\lambda_n)\langle x, P_n y\rangle$$
$$= \sum_{n=1}^{\infty} \langle x, f(\lambda_n)P_n y\rangle = \langle x, f(A)y\rangle.$$

Como los P_n son compactos (ver Ejemplo 6.4.8), entonces los operadores $f_k(A) = \sum_{n=1}^{k} f(\lambda_n)P_n$, son compactos para cada k (ver el Teorema 6.4.7), y como vimos $f_k(A) \to f(A)$, luego el Teorema 6.4.10 implica que $f(A)$ es compacto.

Sea ahora A positivo. Entonces $\langle Ax, x\rangle \ge 0$ para todo $x \in \mathbb{H}$. Entonces, por el Teorema espectral (ver fórmula (6.5.2)) tenemos que, para todo $m \in \mathbb{N}$,

$$0 \le \langle Ax_m, x_m\rangle = \sum_{n=1}^{\infty} \lambda_n |\langle x_m, x_n\rangle|^2 = \lambda_m \quad \Longrightarrow \quad \lambda_n \ge 0, \ \forall n \in \mathbb{N}.$$

Nótese que de lo anterior y usando el Teorema de Hilbert-Schmidt (ver expresión (6.5.8)) se tiene que, para todo $x \in \mathbb{H}$,

$$\langle Ax, x\rangle = \sum_{n=1}^{\infty} \lambda_n |\langle x, x_n\rangle|^2 \geq 0 \quad \Longleftrightarrow \quad \lambda_n \geq 0,\ \forall n \in \mathbb{N}.$$

Sea ahora $f : [0, \infty) \mapsto [0, \infty)$, $f(\lambda) = \lambda^\alpha$, $\alpha > 0$. Entonces $f(\lambda) \to 0$ si $\lambda \to 0$, $f(0) = 0$, luego podemos definir la función A^α, para todo $\alpha > 0$,

$$A^\alpha = \sum_{n=1}^{\infty} \lambda_n^\alpha P_n.$$

Si elegimos $\alpha = 1/2$ tenemos,

$$A^{1/2} = \sum_{n=1}^{\infty} \sqrt{\lambda_n} P_n, \quad \sqrt{\lambda_n} \in \mathbb{R}. \tag{6.6.8}$$

Entonces, usando los resultados del Problema 6.30, $A^2 = \sum_n \lambda_n^2 P_n$, luego

$$(\sqrt{A})^2 = \sum_{n=1}^{\infty} (\lambda_n^{1/2})^2 P_n = \sum_{n=1}^{\infty} \lambda_n P_n = A,$$

es decir, la serie (6.6.8) es la raíz cuadrada del operador A definido positivo. ■

Problema 6.33 Sea $A : \mathbb{H} \mapsto \mathbb{H}$, $\mathbb{H}$ espacio de Hilbert, un operador lineal tal que se tiene una descomposición espectral de la forma (6.5.2), i.e.

$$Ax = \sum_n \lambda_n \langle x, x_n\rangle x_n, \quad \forall x \in \mathbb{H}, \tag{*}$$

donde $(x_n)_n$ es una sucesión ortonormal de vectores de $\mathbb{H}$, $(\lambda_n)_n$ es una sucesión numérica (no necesariamente real) acotada. Si y pertenece a la imagen de A, $\mathcal{I}(A)$, entonces una solución de la ecuación $Ax = y$ es $x = \sum_{n=1}^{\infty} \lambda_n^{-1} \langle y, x_n\rangle x_n$.

Solución: Como $y \in \mathcal{I}(A)$, existe $z \in \mathbb{H}$ tal que $y = Az$. Entonces

$$\langle y, x_n\rangle = \langle Az, x_n\rangle = \left\langle \sum_j \lambda_j \langle z, x_j\rangle x_j, x_n \right\rangle = \lambda_n \langle z, x_n\rangle \ \Rightarrow$$

la serie $x = \sum_{n=1}^{\infty} \lambda_n^{-1}\langle y, x_n\rangle x_n$ es convergente. En efecto, las sumas parciales de dicha serie $X_k = \sum_{n=1}^{k} \lambda_n^{-1}\langle y, x_n\rangle x_n$ son de Cauchy, pues si elegimos j suficientemente grande

$$\|X_m - X_j\|^2 = \sum_{n=j+1}^{m} \lambda_n^{-2}|\langle y, x_n\rangle|^2 = \sum_{n=j+1}^{m} |\langle z, x_n\rangle|^2 < \varepsilon^2,$$

pues la sucesión $\langle z, x_n\rangle$ es de ℓ^2 (¿por qué). Como $\mathbb{H}$ es completo entonces X_n es convergente. Como $(\lambda_n)_n$ es acotada, entonces A es acotado (ver el Ejercicio 6.2.9), por tanto continuo. Entonces, como $Ax_n = \lambda_n x_n$ (es consecuencia de (*)),

$$Ax = \sum_{n=1}^{\infty} \lambda_n^{-1}\langle y, x_n\rangle Ax_n = \sum_{n=1}^{\infty}\langle y, x_n\rangle x_n = \sum_{n=1}^{\infty} \lambda_n\langle z, x_n\rangle x_n = Az = y,$$

de donde se deduce el resultado. ■

Problema 6.34 Sea $\mathbb{E}$ un espacio euclídeo y sea $A : \mathbb{E} \mapsto \mathbb{E}$ compacto. Supongamos que A admite la siguiente descomposición espectral:

$$Ax = \sum_{n=1}^{\infty} \lambda_n\langle x, x_n\rangle x_n, \quad x_n \text{ ortonormales dos a dos,} \tag{6.6.9}$$

y donde los λ_n pueden ser números complejos en general. Entonces, para todo $\lambda \notin \{0, \lambda_1, \lambda_2, \dots\}$, el operador $A - \lambda I$ es invertible y

$$(A - \lambda I)^{-1}x = -\frac{x}{\lambda} + \frac{1}{\lambda}\sum_n \lambda_n \frac{\langle x, x_n\rangle}{\lambda_n - \lambda} x_n. \tag{*}$$

Ayuda: Denota por Bx el miembro derecho de la ecuación anterior y prueba que la serie Bx está bien definida y que además, para todo $x \in \mathbb{X}$, se tiene que $(A - \lambda I)Bx = B(A - \lambda I)x = x$.

Solución: Sea B el operador definido por la suma del miembro derecho de (*), es decir,

$$B : \mathbb{H} \mapsto \mathbb{H}, \quad Bx = -\frac{x}{\lambda} + \frac{1}{\lambda}\sum_n \lambda_n \frac{\langle x, x_n\rangle}{\lambda_n - \lambda} x_n.$$

Si Bx está bien definido para cada $x \in \mathbb{E}$, entonces, mediante un cálculo directo, podemos comprobar que, efectivamente, $(A - \lambda I)Bx = B(A - \lambda I)x = x$, luego $A - \lambda I$ es invertible y su inversa es $(A - \lambda I)^{-1} = B$ como queríamos probar.

Probemos que el operador B está bien definido y es acotado. Comenzaremos probando que la serie

$$S = \sum_n \lambda_n \frac{\langle x, x_n \rangle}{\lambda_n - \lambda} x_n,$$

es convergente. Sea la sucesión $(v_n)_n$, $v_n = \sum_{k=1}^n \frac{\langle x, x_n \rangle}{\lambda_n - \lambda} x_n$. Nótese que $(Av_n)_n$ es la sucesión de sumas parciales de S, así que basta probar que dicha sucesión $(Av_n)_n$ es convergente.

Probemos que la sucesión $(Av_n)_n$ es de Cauchy. Así, para todo $m > n$,

$$\begin{aligned} \|Av_n - Av_m\|^2 &= \sum_{k=n+1}^m \left| \frac{\lambda_k \langle x, x_k \rangle}{\lambda_k - \lambda} \right|^2 \leq \sup_j \frac{|\lambda_j|^2}{|\lambda_j - \lambda|^2} \sum_{k=n+1}^m |\langle x, x_k \rangle|^2 \\ &\leq M \sum_{k=n+1}^m |\langle x, x_k \rangle|^2 \stackrel{n\to\infty}{\longrightarrow} 0, \end{aligned}$$

donde hemos usado que la sucesión $(\langle x, x_k \rangle)_n$ es de ℓ^2 (desigualdad de Bessel (5.2.3)) y que, como $\lambda \notin \{0, \lambda_1, \lambda_2, \dots\}$, y $|\lambda_n| \to 0$ (Teorema 6.5.11), la sucesión $(|\lambda_n|^2/|\lambda_n - \lambda|^2)_n$ es acotada. Si $\mathbb{E}$ es un espacio de Hilbert, que $(Av_n)_n$ sea de Cauchy implica que tiene límite y por tanto la serie S es convergente. Falta comprobar que B es acotado, con lo cual quedaría probado el resultado. Para ello nótese que, repitiendo el razonamiento de antes,

$$\|Av_n\|^2 = \sum_{k=1}^n \left| \frac{\lambda_k \langle x, x_k \rangle}{\lambda_k - \lambda} \right|^2 \leq M \sum_{k=1}^n |\langle x, x_k \rangle|^2 \leq M \|x\|^2.$$

Pero

$$\|Bx\| \leq \frac{1}{|\lambda|} (\|x\| + \|Av_n\|) \leq \left(\sqrt{M} + 1\|\right) \|x\|,$$

de donde se sigue que B es acotado.

Veamos el caso cuando $\mathbb{E}$ no sea un espacio de Hilbert. Para ello, nótese que

$$\|v_n\|^2 = \langle v_n, v_n \rangle = \sum_{k=1}^n \left| \frac{\langle x, x_n \rangle}{\lambda_n - \lambda} \right|^2 \leq \sup_j \frac{1}{|\lambda_j - \lambda|^2} \sum_{k=1}^n |\langle x, x_n \rangle|^2 \leq K \|x\|^2,$$

donde hemos usado nuevamente la desigualdad de Bessel (5.2.3), y que, como $\lambda \notin \{0, \lambda_1, \lambda_2, \dots\}$ y $|\lambda_n| \to 0$ la sucesión $(1/|\lambda_n - \lambda|^2)_n$ es acotada. Luego, la sucesión $(v_n)_n$, es acotada.

Como v_n es acotada y A es compacto, entonces existe una subsucesión $(v_{n_k})_k$ de $(v_n)_n$ tal que Av_{n_k} es convergente, pero entonces $(Av_n)_n$ también es convergente (ver el apartado 2 del Problema 3.13), y por tanto, la serie (*) está bien definida, luego el operador B también. Además, como hemos visto antes, B es acotado lo que prueba el resultado. ■

Nota: Nótese que, si A es compacto y admite la descomposición espectral (6.6.9), entonces para todos los $\lambda \notin \{0, \lambda_1, \lambda_2, \dots\}$, existe la inversa de $(A - \lambda I)^{-1}$, luego la resolvente de A es $\rho(A) = \mathbb{C} \setminus \{0, \lambda_1, \lambda_2, \dots\}$.

Capítulo 7

Otros resultados fundamentales del Análisis Funcional

> La ciencia se construye a base de hechos, de la misma manera que una casa se construye con piedras. Pero igual que un montón de piedras no hace una casa, tampoco una acumulación de hechos es una ciencia.
>
> Henri Poincaré
>
> En *"La Science et l'hypothèse"* (1902)

En este capítulo incluiremos varios teoremas sobre espacios normados que, junto al Teorema de Banach-Steinhaus 4.5.1, constituyen resultados clásicos del análisis funcional.

7.1. Funcionales en dimensión finita

Como vimos en el apartado 4.4 (ver la discusión tras la Nota 4.4.19), el espacio de los funcionales lineales $\mathcal{L}(\mathbb{X}, \mathbb{C})$, $\mathbb{X}$ espacio vectorial, es un espacio lineal. Este espacio se suele denotar como $\mathbb{X}^*$ y se denomina *espacio dual algebraico* de $\mathbb{X}$. Si $\mathbb{X}$ es además normado, entonces el Ejercicio 4.4.14 implica que el espacio de los funcionales lineales y acotados $\mathcal{B}(\mathbb{X}, \mathbb{C})$ es un espacio normado (además completo por el Ejercicio 4.4.20). Este espacio se suele denotar como $\mathbb{X}'$ y se denomina espacio dual de $\mathbb{X}$ o dual conjugado de $\mathbb{X}$.

Sea $\mathbb{X}$ un espacio normado de dimensión finita. Sea $E = (e_k)_{k=1}^n$ una base del mismo, i.e., para todo $x \in \mathbb{X}$, $x = \sum_{k=1}^n x_k e_k$. Sea f un funcional lineal $f : \mathbb{X} \mapsto \mathbb{C}$ cualquiera. Entonces

$$f(x) = \sum_{k=1}^{n} x_k f(e_k) = \sum_{k=1}^{n} x_k c_k, \quad c_k = f(e_k),\ k = 1, \dots, n. \tag{7.1.1}$$

Está claro que dicha representación es única y que, dados los números $c_k = f(e_k)$, $k = 1, \dots, n$, estos definen biunívocamente a f. De esta forma podemos definir n funcionales f tomando las n-tuplas siguientes

$$a_1 = \delta_{1,j} = (1, 0, \dots, 0, 0),\ \dots,\ a_n = \delta_{n,j} = (0, 0, \dots, 0, 1), \quad j = 1, 2, \dots, n,$$

que nos conducen, cada una de ellas, a los funcionales f_k, $k = 1, \dots, n$, tales que

$$f_k(e_j) = \delta_{j,k}, \qquad j, k = 1, \dots, n, \tag{7.1.2}$$

donde $\delta_{j,k}$ es el símbolo de Kronecker, i.e., $\delta_{j,k} = 1$ si $j = k$ y cero en otro caso. De hecho se tiene que $f_k(x) = f_k(\sum_{j=1}^n x_j e_j) = x_k$, $k = 1, \dots, n$. Mostremos que dichos funcionales f_k son una base de $\mathbb{X}^*$. Para ello notamos que, para todo $x \in \mathbb{X}$ la ecuación

$$\sum_{k=1}^{n} \alpha_k f_k(x) = 0,$$

tiene como única solución $\alpha_j = 0$, $j = 1, \dots, n$, luego los funcionales f_k son linealmente independientes. Efectivamente, si tomamos $x = e_j$, entonces tenemos

$$0 = \sum_{k=1}^{n} \alpha_k f_k(e_j) = \alpha_j, \quad \forall\, j = 1, \dots, n.$$

Sea ahora $f \in \mathbb{X}^*$ cualquiera. Entonces, por (7.1.1), para todo $x \in \mathbb{X}$, se cumple que

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

donde hemos usado que $f_k(x) = x_k$. O sea, cualquiera sea $f \in \mathbb{X}^*$

$$f = \sum_{k=1}^{n} c_k f_k.$$

Nótese que de lo anterior se sigue que $\dim \mathbb{X}^* = \dim \mathbb{X}$.

La discusión anterior nos dice que, en general, en cualquier espacio de dimensión finita podemos construir un funcional lineal y este siempre será acotado independientemente de norma de $\mathbb{X}$ que elijamos, aunque la propia norma del operador si que puede depender (como de hecho ocurre en muchos casos) de la norma de $\mathbb{X}$. Veamos ahora dos ejemplos reveladores.

Ejemplo 7.1.1 *Sea* $f : \mathbb{R}^n_\infty \mapsto \mathbb{R}$*, i.e.,* $\|x\| = \|x\|_\infty = \max_{k=1,\dots,n} |x_k|$*, definido en* (7.1.1)*. Calculemos la norma de* f*.*

Usando (7.1.1) tenemos que

$$|f(x)| \leq \sum_{k=1}^{n} |x_k||c_k| \leq \max_{k=1,\dots,n} |x_k| \sum_{k=1}^{n} |c_k| = \sum_{k=1}^{n} |c_k| \|x\| \Rightarrow \|f\| \leq \sum_{k=1}^{n} |c_k|.$$

Sea ahora $\tilde{x} = (\mathrm{signo}\,(c_1), \cdots, \mathrm{signo}\,(c_n))$. Claramente $\|\tilde{x}\| = 1$, $f(\tilde{x}) = \sum_{k=1}^{n} \mathrm{signo}\,(c_k) c_k = \sum_{k=1}^{n} |c_k|$ y, por tanto,

$$\|f\| = \sup_{\|x\|=1} |f(x)| \geq |f(\tilde{x})| = \sum_{k=1}^{n} |c_k| \Rightarrow \|f\| = \sum_{k=1}^{n} |c_k|.$$

Nótese que si consideramos $c = (c_1, \dots, c_n) \in \mathbb{R}^n$, entonces $\sum_{k=1}^{n} |c_k| = \|c\|_1$, es la norma 1 definida en el Ejemplo 4.2.4. ■

El ejemplo anterior nos lleva a preguntarnos si se puede considerar que $\mathbb{X}' = (\mathbb{R}^n, \|\cdot\|_1)$ es el espacio dual de $\mathbb{X} = (\mathbb{R}^n, \|\cdot\|_\infty)$. Para ello necesitaremos el concepto de isomorfismo entre espacios normados.

Definición 7.1.2 *Diremos que dos espacios normados* $\mathbb{X}_1$ *y* $\mathbb{X}_2$ *sobre un mismo cuerpo (*$\mathbb{R}$ *o* $\mathbb{C}$*) son isomorfos si existe un operador* $T : \mathbb{X}_1 \mapsto \mathbb{X}_2$*, lineal y biyectivo que preserve la norma i.e.,* $\|Tx\|_2 = \|x\|_1$ *para todos* $x \in \mathbb{X}_1$*.*

Nótese que dado que T preserva la norma, entonces también preserva la distancia distancia entre los elementos, es decir, T es una isometría –ver Definición 3.5.23–. Sin embargo, dado que ambos espacios $\mathbb{X}_1$ y $\mathbb{X}_2$ son normados, no nos basta con que exista una isometría entre ellos (como ocurre con los espacios métricos), sino que necesitamos que dicha isometría preserve la estructura de espacio vectorial –para eso necesitamos la linealidad de la isometría T–. Nótese además que, como consecuencia del Problema 6.5, los operadores unitarios (ver Definición 5.2.11) son isometrías, luego la definición de espacios isomorfos en el contexto de espacios de Hilbert (Definición 5.2.11) es un caso particular de la Definición 7.1.2 para los espacios normados.

Con la definición de isomorfismo anterior 7.1.2 podemos responder a la pregunta que formulamos antes, i.e., ¿cuál es el espacio dual de $\mathbb{X} = (\mathbb{R}^n, \|\cdot\|_\infty)$?

Ejemplo 7.1.3 *El espacio dual de $\mathbb{X} = (\mathbb{R}^n, \|\cdot\|_\infty)$ es $\mathbb{X}' = (\mathbb{R}^n, \|\cdot\|_1)$.*

El Ejemplo 7.1.1 nos indica que podemos identificar el espacio dual de $\mathbb{X} = (\mathbb{R}^n, \|\cdot\|_\infty)$ con $\mathbb{X}' = (\mathbb{R}^n, \|\cdot\|_1)$. Para ello definamos la aplicación $T : \mathbb{X}' \mapsto \mathbb{X}$, $f \in \mathbb{X}'$ definido por $f(x) = \sum_{k=1}^n x_k c_k$, $Tf = c$, $f(x) = \sum_{k=1}^n x_k c_k$, $c = (c_1, \dots, c_k) \in \mathbb{R}^n$. Está claro que dicho T es lineal y biyectivo (probar como ejercicio). Además, como hemos visto en el ejemplo anterior 7.1.1, $\|Tf\| = \|c\|_1 = \|f\|$, es decir, T es una isometría, luego el espacio dual $\mathbb{X} = (\mathbb{R}^n, \|\cdot\|_\infty)$ es isomorfo con $\mathbb{X}' = (\mathbb{R}^n, \|\cdot\|_1)$, i.e., si $\mathbb{X} = (\mathbb{R}^n, \|\cdot\|_\infty)$ entonces el espacio dual $\mathbb{X}'$ se puede identificar con $(\mathbb{R}^n, \|\cdot\|_1)$. ■

De manera análoga al caso considerado en el Ejemplo 7.1.3 anterior se puede comprobar que $(\mathbb{R}_2^n)' = \mathbb{R}_2^n$, es decir, el dual de $\mathbb{R}^n$ con la norma euclídea es el propio $\mathbb{R}^n$. De hecho, esto último es un caso particular del siguiente resultado mucho más general:

Ejemplo 7.1.4 *Si $\mathbb{H}$ un espacio de Hilbert real, entonces $\mathbb{H}' = \mathbb{H}$.*

Sea $\mathbb{H}'$ el espacio de todas las aplicaciones lineales acotadas de $\mathbb{H} \mapsto \mathbb{R}$. Por el Teorema de Riesz sabemos que a cada funcional $f \in \mathbb{H}'$ le podemos asociar un producto escalar, i.e, $f(x) = \langle x, y\rangle$ donde $y \in \mathbb{H}$ es único. Para evidenciar la correspondencia biunívoca entre los $f \in \mathbb{H}'$ y los $y \in \mathbb{H}$ usaremos la notación $f_y(x) = \langle x, y\rangle$.

Sea la aplicación $T : \mathbb{H} \mapsto \mathcal{B}(\mathbb{H}, \mathbb{R}) = \mathbb{H}'$ definida por

$$Ty = f_y, \quad f_y : \mathbb{H} \mapsto \mathbb{R}, \quad f_y(x) = \langle x, y\rangle.$$

Esta aplicación es claramente biyectiva pues para cada $f \in \mathbb{H}'$, existe un único $y \in \mathbb{H}$ tal que $f(x) = \langle x, y\rangle$. Además, T es lineal[a]

$$T(\alpha y + \beta z) = \langle x, \alpha y + \beta z\rangle = \alpha\langle x, y\rangle + \beta\langle x, z\rangle = \alpha T y + \beta T z.$$

Sea $y \neq 0$, $x \in \mathbb{H}$. Se tiene, por un lado, que

$$\|(Ty)(x)\| = |f_y(x)| = |\langle x, y\rangle| \leq \|x\|\|y\| \;\; \Rightarrow \;\; \|Ty\| \leq \|y\|$$

y por el otro, que

$$\|Ty\| = \sup_{\|x\|\neq 0} \frac{\|(Ty)(x)\|}{\|x\|} = \sup_{\|x\|\neq 0} \frac{|\langle x, y\rangle|}{\|x\|} \geq \frac{|\langle y, y\rangle|}{\|y\|} = \|y\|,$$

luego $\|Ty\| \geq \|y\|$ y, por tanto, $\|Ty\| = \|y\|$ para todo $y \in \mathbb{H}$. Es decir, T es una isometría así que el espacio $\mathbb{H}'$ es isomorfo con $\mathbb{H}$, luego $\mathbb{H}' = \mathbb{H}$. ■

[a] Si $\mathbb{H}$ fuese complejo T no sería lineal sino *antilineal*. No obstante T seguiría siendo una aplicación biyectiva y una isometría. En este caso, el espacio dual de $\mathbb{H}$ sería el espacio de Hilbert con el producto escalar $\langle f_y, f_z\rangle = \langle z, y\rangle$. Nótese que el caso de que $\mathbb{H}$ sea real $\langle z, y\rangle = \langle y, z\rangle$, de donde se sigue que $\mathbb{H}' = \mathbb{H}$.

Consideremos ahora el problema de extender[1] el funcional del Ejemplo 7.1.1 a un espacio $\mathbb{R}^m$ de dimensión m mayor que n. Está claro que cualquiera sea m finito el razonamiento anterior se puede repetir y las conclusiones serán las mismas: $\dim \mathbb{X} = \dim \mathbb{X}^* = \dim \mathbb{X}' = m$, además, f será acotado. ¿Que ocurre si queremos extender f a dimensión infinita? En ese caso podría ocurrir perfectamente que la sucesión de sumas parciales, $\sum_{k=1}^{n} |c_k|$, sean no acotadas. En otras palabras, a la hora de extender un funcional lineal hay que tener cuidado si nos interesa que la extensión no solo sea lineal, sino también acotada. En el próximo apartado trataremos con cierto detalle el problema de la extensión de un funcional lineal acotado definido sobre espacios normados.

7.2. El Teorema de Hahn-Banach

Comenzaremos recordando la definición de extensión de un operador (ver la Definición 3.3.10) en el contexto de espacios normados.

[1] Recuérdese la Definición 3.3.10.

Definición 7.2.1 *Sean $\mathbb{X}, \mathbb{Y}$, espacios normados. La extensión de una aplicación $T : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathbb{Y}$, a un subconjunto $M \supset \mathcal{D}(T)$ es una aplicación $\widetilde{T}$ tal que $\widetilde{T}|_{\mathcal{D}(T)} = T$, i.e., $\widetilde{T}x = Tx$ para todo $x \in \mathcal{D}(T)$.*

Imaginemos que tenemos un operador lineal operador y acotado $T : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathbb{Y}$. Esta claro que hay muchas formas de extender T al subconjunto $M \supset \mathcal{D}(T)$, sin embargo la idea no es solo extender T a un conjunto más grande, sino también que se mantengan ciertas propiedades como, por ejemplo, la linealidad y la acotación. En este apartado discutiremos esa posibilidad. Comenzaremos con el siguiente teorema:

Teorema 7.2.2 *Sea $T : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathbb{Y}$ un operador lineal y acotado, $\mathbb{X}$ un espacio normado, $\mathbb{Y}$ de Banach. Entonces T se puede extender a $\overline{\mathcal{D}(T)}$, la clausura de $\mathcal{D}(T)$, i.e., existe un operador lineal y acotado $\widetilde{T} : \overline{\mathcal{D}(T)} \mapsto \mathbb{Y}$ tal que $\widetilde{T}(x) = T(x)$ para $x \in D(T)$ y $\|\widetilde{T}\| = \|T\|$.*

Demostración: Sea $x \in \overline{\mathcal{D}(T)}$ cualquiera. Entonces existe $(x_n)_n$ tal que $x_n \to x$. Como $(x_n)_n$ es convergente, es de Cauchy. Entonces, para todos n, m suficientemente grandes

$$\|Tx_n - Tx_m\| \leq \|T\|\|x_n - x_m\| \to 0 \ \Rightarrow \ (Tx_n)_n \quad \text{es de Cauchy.}$$

Como $\mathbb{Y}$ es completo, $Tx_n \to y \in \mathbb{Y}$. Definamos la aplicación $\widetilde{T}$ que a cada $x \in \overline{\mathcal{D}(T)}$ le hace corresponder el valor $y = \widetilde{T}x$ anterior. Probemos que dicha definición de $\widetilde{T}$ es independiente de la sucesión $(x_n)_n$ elegida. Para ello elijamos una sucesión distinta $(z_n)_n$, $z_n \to x$ y supongamos que $Tx_n \to y$, $Tz_n \to y'$, $y \neq y'$. Entonces,

$$\|y - y'\| = \lim_{n\to\infty} \|Tx_n - Tz_n\| = \lim_{n\to\infty} \|T(x_n - z_n)\| = 0,$$

pues $x_n - z_n \to 0$ y T es acotado (por tanto continuo), de donde se sigue que $y = y'$ lo cual es una contradicción.[2]

Demostremos que $\widetilde{T}$ es lineal. Tomamos $x_n \to x$ y $z_n \to z$, entonces para todos $\alpha, \beta \in \mathbb{C}$, $\alpha x_n + \beta z_n \to \alpha x + \beta z$, y $T(\alpha x_n + \beta z_n) \to y'' \in \mathbb{Y}$,

[2]Se puede probar también de la siguiente forma: a partir las sucesiones $(x_n)_n$ y $(z_n)_n$, construimos la sucesión $w_n = (x_1, z_1, x_2, z_2, \ldots, x_n, w_n, \ldots)$. Está claro que $w_n \to x$ (¿por qué?), luego, como T es acotado, $Tw_n \to Tx = y$, pero entonces las subsucesiones $(Tx_n)_n$ y $(Tz_n)_n$ de $(Tw_n)_n$ tienen que tener el mismo límite, luego $Tz_n \to y$, i.e., el límite de Tx_n es independiente de la sucesión elegida.

$y'' = \widetilde{T}(\alpha x + \beta z)$. Luego

$$\begin{array}{ccc} T(\alpha x_n + \beta z_n) = & \alpha T x_n + \beta T z_n \\ \downarrow & \downarrow \quad \downarrow \\ \widetilde{T}(\alpha x + \beta z) = & \alpha \widetilde{T} x + \beta \widetilde{T} z \end{array}$$

Finalmente, veamos que es acotado. Por un lado, como T es acotado,

$$\|Tx_n\| \leq \|T\|\|x_n\|, \, n \to \infty \Rightarrow \|y\| = \|\widetilde{T}x\| \leq \|T\|\|x\| \;\; \Rightarrow \;\; \|\widetilde{T}\| \leq \|T\|.$$

Por el otro,

$$\|\widetilde{T}\| = \sup_{x \in \overline{\mathcal{D}(T)}\,:\,\|x\|=1} \|\widetilde{T}x\| \geq \sup_{x \in \mathcal{D}(T)\,:\,\|x\|=1} \|\widetilde{T}x\| = \sup_{x \in \mathcal{D}(T)\,:\,\|x\|=1} \|Tx\| = \|T\|,$$

luego $\|\widetilde{T}\| = \|T\|$. ■

Es decir, en general siempre podremos extender un operador lineal definido sobre un subconjunto $M \subset \mathbb{X}$ a su clausura $\overline{M}$ de forma que la extensión siga siendo lineal y acotada (y, por tanto, continua). En particular, todo funcional lineal acotado $f : M \subset \mathbb{X} \mapsto \mathbb{C}$ admite una extensión $F : \overline{M} \subset \mathbb{X} \mapsto \mathbb{C}$ que seguirá siendo lineal y acotada y cumplirá que $\|f\| = \|F\|$.

Imaginemos ahora el siguiente problema. Sea $f : M \subset \mathbb{X} \mapsto \mathbb{C}$ un funcional lineal definido sobre un subespacio vectorial de dimensión finita M de un espacio normado $\mathbb{X}$ (ver la sección anterior). Dicho funcional es acotado pues M es de dimensión finita ¿se puede extender a todo $\mathbb{X}$? Si $\mathbb{X}$ es de dimensión finita está claro que si existe una extensión que sea lineal, entonces será acotada (toda aplicación lineal sobre un espacio de dimensión finita es acotada), pero ¿y si $\mathbb{X}$ es de dimensión infinita?

Para responder a las cuestiones anteriores necesitaremos el Teorema de Hahn-Banach que responde afirmativamente a todas ellas. La forma más general del teorema precisa del concepto de funcional convexo y funcional sublineal:

Definición 7.2.3 *Sea $\mathbb{V}$ un espacio vectorial complejo. Se dice que el funcional $p : \mathbb{V} \mapsto \mathbb{R}$ es un funcional sublineal si se cumple*

$$\forall x, y \in \mathbb{V}, \quad p(x + y) \leq p(x) + p(y), \tag{7.2.1}$$

$$\forall x, y \in \mathbb{V}, \quad p(\alpha x) = \alpha p(x), \quad \forall \alpha \geq 0. \tag{7.2.2}$$

Si $p : \mathbb{V} \mapsto \mathbb{R}$, *para todo* $x \in \mathbb{V}$ *y en vez de* (7.2.2) *se cumple*

$$\forall x, y \in \mathbb{V}, \quad p(\alpha x) = |\alpha| p(x), \quad \forall \alpha \in \mathbb{C}, \tag{7.2.3}$$

se dice que p *es un funcional convexo.*

Por ejemplo, si $\mathbb{X}$ es un espacio normado, $p(x) = a\|x\|$, $a > 0$, es un funcional sublineal y convexo.

Nótese que de (7.2.2) y (7.2.3) se sigue que $p(0) = 0$. Además, en el caso de un funcional convexo de (7.2.1) se tiene que $p(x) \geq 0$, pues $0 = p(x - x) \leq p(x) + p(-x) = 2p(x)$.

El Teorema de Hahn-Banach establece lo siguiente:

Teorema 7.2.4 (Hahn-Banach para un espacio vectorial) *Sea* $\mathbb{V}$ *un espacio vectorial real o complejo y sea* p *un funcional convexo definido sobre* $\mathbb{V}$. *Sea* $M \subset \mathbb{V}$, $M \neq \{0\}$, *un subespacio vectorial y sea* $f : M \mapsto \mathbb{C}$ *un funcional lineal tal que*

$$|f(x)| \leq p(x), \qquad \forall x \in M. \tag{7.2.4}$$

Entonces, existe un funcional lineal $F : \mathbb{V} \mapsto \mathbb{C}$ *tal que*

$$\forall x \in M, \quad F(x) = f(x) \quad y \quad \forall x \in \mathbb{V}, \quad |F(x)| \leq p(x). \tag{7.2.5}$$

Antes de proseguir conviene hacer una aclaración. Como mencionamos en la discusión posterior a la prueba del Teorema 7.2.2, la idea es ver en que condiciones se puede extender un funcional lineal y acotado f a un espacio más grande manteniendo tanto la linealidad como la acotación. Dado que en un espacio vectorial general $\mathbb{V}$ no tiene porque tener definida una norma es necesario introducir alguna función que, de alguna forma nos ayude a *acotar* nuestro funcional. En el Teorema de Hahn-Banach general 7.2.4 la función que juega el papel de la norma es el funcional $p(x)$ de la Definición 7.2.3. En el caso cuando tengamos definida una norma, por ejemplo si estamos considerando un espacio normado o uno de Hilbert, basta usar $p(x) = \|f\|\|x\|$.

Un corolario del Teorema Hahn-Banach 7.2.4 es el siguiente resultado fundamental para espacios normados:

Teorema 7.2.5 (Hahn-Banach) *Sea $\mathbb{X}$ un espacio normado, $M \neq \{0\}$ un subespacio de $\mathbb{X}$ y sea $f : M \mapsto \mathbb{C}$, un funcional lineal acotado. Entonces, existe un funcional lineal $F : \mathbb{X} \mapsto \mathbb{C}$ tal que*

$$\forall x \in M, \quad F(x) = f(x) \quad y \quad \|F\| = \|f\|. \tag{7.2.6}$$

Es decir, f se puede extender a todo $\mathbb{X}$ de forma que la norma de F coincida con la de f.

Demostración: Definamos el funcional $p : \mathbb{X} \mapsto \mathbb{R}$, $p(x) = \|f\|\|x\|$. Está claro que p es un funcional convexo, y que

$$\forall x \in M, \quad |f(x)| \leq \|f\|\|x\| = p(x).$$

Luego se cumplen las condiciones del Teorema de Hahn-Banach 7.2.4, así que existe un funcional lineal F tal que

$$\forall x \in \mathbb{X}, \quad |F(x)| \leq p(x) = \|f\|\|x\| \;\Rightarrow\; \|F\| \leq \|f\|,$$

es decir F es acotado. Por otro lado,

$$\|F\| = \sup_{x\in\mathbb{X}\,:\,\|x\|=1} |F(x)| \geq \sup_{x\in M\,:\,\|x\|=1} |F(x)| = \sup_{x\in M\,:\,\|x\|=1} |f(x)| = \|f\|,$$

de donde se sigue el resultado. ■

Veamos un ejemplo del uso del teorema anterior.

Ejemplo 7.2.6 Sea un espacio normado $\mathbb{X} \neq \{0\}$ cualquiera. Como $\mathbb{X}$ no es el espacio nulo, entonces existe $0 \neq x_0 \in \mathbb{X}$. Definamos $M = \operatorname{span}(x_0)$, i.e., $x = \alpha x_0$, $\alpha \in \mathbb{C}$. Entonces el funcional $f : M \mapsto \mathbb{C}$, $f(x) = f(\alpha x_0) = \alpha$ es un funcional lineal y acotado. La linealidad es inmediata ya que para todos $a, b \in \mathbb{C}$, y $x = \alpha x_0$ e $y = \beta x_0$ tenemos

$$f(ax + by) = f[(a\alpha + b\beta)x_0] = a\alpha + b\beta = af(x) + bf(y).$$

Para la acotación tomamos $0 \neq x \in M$

$$\frac{|f(x)|}{\|x\|} = \frac{|f(\alpha x_0)|}{\|\alpha x_0\|} = \frac{|\alpha|}{|\alpha|\|x_0\|} = \frac{1}{\|x_0\|} \;\Rightarrow\; \|f\| \leq \frac{1}{\|x_0\|}.$$

Luego, el Teorema de Hahn-Banach 7.2.5 establece que nuestro f se puede extender a todo el espacio $\mathbb{X}$ de forma que el funcional extendido F sea lineal y $\|F\| = \|f\|$. ■

El Teorema 7.2.5 resuelve el problema planteado al inicio de este apartado pues nos asegura que siempre es posible extender un funcional lineal definido sobre un subespacio a todo el espacio normado manteniendo invariante la norma del mismo. De hecho el resultado es mucho más relevante de lo que a primera vista puede parecer pues, tal como muestra el Ejemplo 7.2.6, basta con que un espacio normado tenga al menos un elemento no nulo para que exista un subespacio del mismo que admita un funcional lineal acotado no nulo, es decir, en cualquier espacio normado que tenga elementos distintos del vector nulo se pueden definir funcionales lineales no nulos, luego todo espacio normado no nulo $\mathbb{X}$ tiene asociado un espacio dual $\mathbb{X}'$ no nulo. Además, dichos funcionales se pueden extender a espacios normados más grandes que contengan al espacio original.

Para probar el Teorema de Hahn-Banach en un espacio vectorial real general se precisa del Lema de Zorn (equivalente al *Axioma de elección* de la Teoría de conjuntos) que no vamos a discutir en este curso. En vez de ello lo demostraremos para algunos casos de especial interés. La prueba general se puede consultar, por ejemplo, en [1, §11.1, pág. 176–182] o en [11, pág. 219, Teorema 4.3-1].

Teorema 7.2.7 (Hahn-Banach para un espacio vectorial real) *Sea $\mathbb{V}$ un espacio vectorial real y sea p un funcional sublineal definido sobre $\mathbb{V}$. Sea $M \subset \mathbb{V}$, $M \neq \{0\}$, un subespacio vectorial y sea $f : M \mapsto \mathbb{R}$ un funcional lineal tal que*

$$f(x) \leq p(x), \qquad \forall x \in M. \tag{7.2.7}$$

Entonces, existe un funcional lineal $F : \mathbb{V} \mapsto \mathbb{R}$ tal que

$$\forall x \in M, \quad F(x) = f(x) \quad y \quad \forall x \in \mathbb{V}, \quad F(x) \leq p(x). \tag{7.2.8}$$

Demostración: Caso I: Asumiremos que[3] $M \neq \{0\}$ y que $M \neq \mathbb{V}$. Entonces existe un $x_1 \in \mathbb{V}$, $x_1 \neq 0$, que no pertenece a M. Definamos el subespacio

$$M_1 = \{z = x + \alpha x_1 \,:\, \forall x \in M,\ \forall \alpha \in \mathbb{R}\}.$$

[3]En el primer caso $f = 0$ y su extensión trivialmente es $F = 0$, y en el segundo no tiene sentido hablar de la extensión de f.

Nótese que cualquiera sea $z \in M_1$, la representación $z = x + \alpha x_1$ es única.[4]

Vamos a probar que f se puede extender a un funcional F_1 definido en M_1 tal que se tenga (7.2.8). Vamos a suponer que dicha extensión existe. Entonces ha de cumplir que

$$F_1(z) = F_1(x + \alpha x_1) = F_1(x) + \alpha F_1(x_1) = f(x) + \alpha F_1(x_1).$$

Nótese al ser la representación $z = x + \alpha x_1$ única para cada $z \in M_1$, el funcional F_1, caso de que existiese, quedaría determinado por el valor de $F_1(x_1)$. Además, F_1 sería tal que, para todo $z \in M_1$, $F_1(z) \leq p(z)$, luego se tiene que cumplir que

$$F_1(x + \alpha x_1) = f(x) + \alpha F_1(x_1) \leq p(x + \alpha x_1), \quad \forall x \in M, \ \forall \alpha \in \mathbb{R}. \tag{7.2.9}$$

Si $\alpha > 0$, entonces

$$F_1(x_1) \leq \frac{1}{\alpha}\big[p(x + \alpha x_1) - f(x)\big] = p\left(\frac{x}{\alpha} + x_1\right) - f\left(\frac{x}{\alpha}\right) \Rightarrow$$
$$F_1(x_1) \leq p(x_1 + \xi) - f(\xi), \qquad \xi = \frac{x}{\alpha} \in M.$$

Si $\alpha < 0$, $\alpha = -|\alpha|$, entonces

$$F_1(x_1) \geq \frac{1}{|\alpha|}\big[f(x) - p(x - |\alpha| x_1)\big] = f\left(\frac{x}{|\alpha|}\right) - p\left(\frac{x}{|\alpha|} - x_1\right) \Rightarrow$$
$$F_1(x_1) \geq f(\xi') - p(\xi' - x_1), \qquad \xi' = \frac{x}{|\alpha|} \in M.$$

Juntando ambas

$$f(\xi') - p(\xi' - x_1) \leq F_1(x_1) \leq p(\xi + x_1) - f(\xi), \qquad \forall \xi, \xi' \in M. \tag{7.2.10}$$

Está claro que si existe $F_1(x_1)$ tal que (7.2.10) se cumpla, entonces (7.2.9) se cumple y, por tanto, $F_1(z) \leq p(z)$, para todo $z \in M_1$. Luego basta con demostrar que siempre se puede elegir un F_1 tal que $F_1(x_1)$ cumpla (7.2.10). Para ello es necesario que para todos $\xi, \xi' \in M$

$$f(\xi') - p(\xi' - x_1) \leq p(x_1 + \xi) - f(\xi), \tag{7.2.11}$$

[4] Supongamos que no lo es. Entonces existes $x' \in M$ y $\alpha' \in \mathbb{R}$, tales que $z = x' + \alpha' x_1$, luego $M \ni x - x' = (\alpha' - \alpha)x_1 \in \operatorname{span}(x_1)$, lo cual es una contradicción a no ser que $x = x'$ y $\alpha = \alpha'$.

o, equivalentemente,

$$f(\xi)+f(\xi')\leq\ p(x_1+\xi)+p(\xi'-x_1).$$

Usando que f es lineal, que satisface (7.2.7) y que se tiene (7.2.1) obtenemos

$$\begin{aligned}f(\xi)+f(\xi')&=f(\xi+\xi')\leq p(\xi+\xi')=p(\xi+x_1+\xi'-x_1)\\&\leq p(\xi+x_1)+p(\xi'-x_1),\end{aligned}$$

luego se cumple (7.2.11). Si ahora tomamos ínfimos en $\xi\in M$ en (7.2.11) y a continuación supremos en $\xi'\in M$ obtenemos

$$c=\sup_{\xi'\in M}[f(\xi')-p(\xi'-x_1)]\leq\inf_{\xi\in M}[p(x_1+\xi)-f(\xi)]=C.$$

Luego, si elegimos $F_1(x_1)$ tal que $c\leq F_1(x_1)\leq C$, entonces podemos construir el funcional lineal F_1 que extiende f a todo M_1 y que cumple con $F_1(z)\leq p(z)$ en M_1.

Si $\mathbb{V}=M_1$ el teorema queda demostrado. Si no, entonces existe un $x_2\neq 0$ en $\mathbb{V}$ que no está en M_1 y podemos construir el conjunto

$$M_2=\{z=x+\alpha x_2\,:\,\forall x\in M_1,\ \forall\alpha\in\mathbb{R}\},$$

y repitiendo el mismo razonamiento obtenemos un funcional F_2 que extiende F_1 a M_2, luego extiende f a M_2. Si $\mathbb{V}=M_2$, el teorema quedaría probado. Y así sucesivamente.

De esta forma tendremos una sucesión de funcionales F_k definidos sobre los subespacios

$$M_k=\{z=x+\alpha x_k\,:\,\forall x\in M_{k-1},\ \forall\alpha\in\mathbb{R}\},\quad k=1,\dots,n,$$

que extienden a f a cada uno de los subespacios M_k. Si para algún $n\in\mathbb{N}$ se tiene que $\mathbb{V}=M_n$, el teorema queda probado. Veamos que en el caso cuando $\mathbb{V}=\bigcup_{k=1}^{\infty}M_k$ podemos construir el funcional F por inducción. Para ello notamos que cualquiera sea $x\in\mathbb{V}$, $x\in M_n$ para algún $n\in\mathbb{N}$, y dado que $F_n(x)$ extiende a f en M_n, entonces podemos definir, para dicho x, $F(x)=F_n(x)$, de forma que se tiene que $F(x)=F_n(x)\leq p(x)$, como se quería demostrar. Nótese que, en general, si $x\in M_n$, entonces si $m>n$, $x\in M_m$ y podemos definir $F(x)=F_m(x)$, que es una extensión de F_m si $m>n$, y por tanto de f.

Caso II: $\mathbb{V}$ **es un espacio normado separable.** ¿Qué ocurre si $\mathbb{V} \neq \bigcup_{k=1}^{n} M_k$? En este caso asumiremos que $\mathbb{V} = \mathbb{X}$ es un espacio normado separable y f es un funcional acotado en M. Como $\mathbb{X}$ es separable, entonces $\mathbb{X} \setminus M$ es separable, así que existe un subconjunto denso y numerable $V = \{x_1, x_2, \dots\}$ tal que $\overline{V} = \mathbb{X} \setminus M$. Notemos que[5] $\mathbb{X} = M \cup \overline{V} = \overline{M} \cup \overline{V} = \overline{M \cup V}$. Por otro lado, tomando $p(x) = \|f\|\|x\|$ y repitiendo el razonamiento usado en el Caso I, podemos construir la secuencia de subespacios M_1, M_2, ..., M_k, ..., y construir un funcional F que extiende f a $\bigcup_{k=1}^{\infty} M_k = M \cup V$. Pero entonces, como F es acotado en $M \cup V$ podemos utilizar del Teorema 7.2.2 y construir su extensión en $\overline{M \cup V} = \mathbb{X}$. Además, dado que hemos elegido $p(z) = \|f\|\|x\|$ como en la prueba del Teorema de Hahn-Banach 7.2.5, tenemos

$$|F(z_n)| \leq p(z_n) = \|f\|\|z_n\| \quad \xrightarrow{n\to\infty} \quad |F(z)| \leq \|f\|\|z\| = p(z), \quad \forall z \in \mathbb{X}.$$

Caso III: $\mathbb{V}$ **es un espacio de Hilbert cualquiera.** ¿Y si $\mathbb{X}$ no es separable? Entonces podemos restringirnos a un espacio de Hilbert donde, como en el Casi II, elegiremos $p(x) = \|f\|\|x\|$. En este caso podemos usar el Teorema de representación de Riesz para funcionales acotados 5.2.21 válido para cualquier espacio de Hilbert $\mathbb{H}$ (no necesariamente separable). Para ello razonamos como sigue. Primero extendemos $f : M \mapsto \mathbb{R}$ a $\overline{M} \subset \mathbb{H}$ usando el Teorema 7.2.2 y obtenemos el funcional $\tilde{f}$. Pero como $\overline{M}$ es cerrado, entonces es un espacio de Hilbert, luego $\tilde{f}$ admite una única representación como $\tilde{f}(x) = \langle x, y\rangle$, para todo $x \in \overline{M}$, donde $y \in \overline{M}$ es único. La extensión de $\widetilde{f}$ a todo $\mathbb{H}$ la hacemos definiendo $F(x) = \langle x, y\rangle$, para todo $x \in \mathbb{H}$, y eligiendo y el mismo de antes. ∎

Veamos a continuación una serie de consecuencias importantes del Teorema de Hahn-Banach 7.2.5. En adelante por comodidad denotaremos por $\mathbb{K}$ al conjunto $\mathbb{C}$ o $\mathbb{R}$ y por $\mathbb{X}'$ el espacio de todos los funcionales lineales y acotados definidos sobre $\mathbb{X}$, i.e., $\mathbb{X}' = \mathcal{B}(\mathbb{X}, \mathbb{K})$.

Corolario 7.2.8 *Sea $\mathbb{X} \neq \{0\}$ un espacio normado y sea $0 \neq x_0 \in \mathbb{X}$. Entonces, existe un funcional lineal F en $\mathbb{X}$ tal que $F(x_0) = \|x_0\|$ y $\|F\| = 1$.*

Demostración: Sea $M = \{x = \alpha x_0 \ : \ \alpha \in \mathbb{K}\} \subset \mathbb{X}$. Definamos en funcional $f : M \mapsto \mathbb{K}$, $f(x) = f(\alpha x_0) = \alpha\|x_0\|$. Claramente f es lineal y además

$$|f(x)| = |\alpha|\|x_0\| = \|x\| \;\Rightarrow\; \|f\| = 1.$$

[5]Recordemos que, como $M \cup \overline{V}$ es todo el espacio $\mathbb{X}$, la unión de la clausura de M con $\overline{V}$ no adiciona ningún nuevo elemento a $\mathbb{X}$ y, por tanto, $M \cup \overline{V} = \overline{M} \cup \overline{V}$.

Entonces, por el Teorema de Hahn-Banach 7.2.5 existe un funcional $F : \mathbb{X} \mapsto \mathbb{K}$ tal que $F(\alpha x_0) = \alpha \|x_0\|$ en M y $\|F\| = \|f\| = 1$. ■

Una consecuencia de este corolario es que todo espacio normado $\mathbb{X} \neq \{0\}$ tiene asociado un espacio dual no nulo.

Corolario 7.2.9 *Sea $\mathbb{X} \neq \{0\}$ un espacio normado y sean $x_1 \neq x_2 \in \mathbb{X}$. Entonces, existe un funcional lineal F en $\mathbb{X}$ tal que $F(x_1) \neq F(x_2)$.*

Demostración: Se sigue del corolario 7.2.8 tomando $x_0 = x_1 - x_2$. ■

Es decir, en el dual de $\mathbb{X}$ siempre hay funcionales que permiten distinguir (separar) los elementos de $\mathbb{X}$.

Corolario 7.2.10 *Sea $x_0 \in \mathbb{X} \neq \{0\}$, $\mathbb{X}$ espacio normado tal que $f(x_0) = 0$ para todo $f \in \mathbb{X}'$. Entonces $x_0 = 0$.*

Demostración: Supongamos que $x_0 \neq 0$. Por el corolario 7.2.8 existe un $F \in \mathbb{X}'$ de norma $\|F\| = 1$, tal que $F(x_0) = \|x_0\|$, pero como para todo $f \in \mathbb{X}'$ se tiene $f(x_0) = 0$, entonces $F(x_0) = 0$, lo que implica $\|x_0\| = 0$, lo cual es una contradicción, luego $x_0 = 0$. ■

Nótese que de lo anterior se sigue que si $f(x_1) = f(x_2)$ para todo $f \in \mathbb{X}'$, entonces $x_1 = x_2$.

Como ya hemos visto, dado un funcional $f : \mathbb{X} \mapsto \mathbb{K}$,

$$\|f\| = \sup_{x \in \mathbb{X}\,:\, \|x\|=1} |f(x)|.$$

El siguiente corolario es, de alguna forma, el recíproco de lo anterior.

Corolario 7.2.11 *Sea $x_0 \in \mathbb{X} \neq \{0\}$, $\mathbb{X}$ espacio normado. Entonces,*

$$\|x_0\| = \sup_{f \in \mathbb{X}'\,:\, \|f\|=1} |f(x_0)|,$$

y dicho supremo es alcanzable.

Demostración: Si $x_0 = 0$ el resultado es inmediato. Sea $x_0 \neq 0$. Para cualquiera sea $f \in \mathbb{X}'$ tal que $\|f\| = 1$, se tiene que $|f(x_0)| \leq \|f\|\|x_0\| = \|x_0\|$. Luego $\sup_{f \in \mathbb{X}'\,:\, \|f\|=1} |f(x_0)| \leq \|x_0\|$. Pero por el corolario 7.2.8 existe un $F \in \mathbb{X}'$ de norma $\|F\| = 1$, tal que $F(x_0) = \|x_0\|$, i.e., el supremo es alcanzable. ■

Nota: Otra consecuencia inmediata del Teorema de Hahn-Banach es la extensión de la fórmula (7.1.2), válida para espacios de dimensión finita, a cualquier espacio normado $\mathbb{X}$. En efecto, sean $x_1, \dots x_n$ n vectores linealmente independientes y sea el subespacio lineal $M = \operatorname{span}(x_1, \dots, x_n) \subset \mathbb{X}$. Como M es de dimensión finita, existen los funcionales lineales $f_1, \dots, f_n$ tales que $f_k(x_j) = \delta_{j,k}$, $j, k = 1, \dots, n$. Pero entonces, por el Teorema de Hahn-Banach 7.2.5 existen los funcionales $F_k : \mathbb{X} \mapsto \mathbb{K}$ tales que $F_k(x_j) = \delta_{j,k}$, $j, k = 1, \dots, n$.

Veamos una última consecuencia del Teorema de Hahn-Banach.

Teorema 7.2.12 *Sea un subespacio $M \subset \mathbb{X} \neq \{0\}$, $M \neq \mathbb{X}$ y sea $x_0 \in \mathbb{X}$ tal que*

$$d := \rho(x_0, M) = \inf_{x \in M} \|x_0 - x\| > 0. \tag{7.2.12}$$

Entonces, existe un funcional lineal $F : \mathbb{X} \mapsto \mathbb{K}$, tal que $\|F\| = 1$, $F(x_0) = d$, y $F(x) = 0$, para todo $x \in M$.

Demostración: Sea el conjunto $M_1 = \{z = x + \alpha x_0 \ : \ \forall x \in M, \ \alpha \in \mathbb{K}\}$.

Nótese que como $d > 0$, eso implica que $x_0 \notin M$, $x_0 \neq 0$. Además, como ya vimos en la prueba del Teorema de Hahn Banach (ver la nota al pie de página 4 de la página 283) esta descomposición es única, luego podemos definir el funcional $f : M_1 \mapsto \mathbb{K}$, $f(z) = \alpha d$, que es lineal

$$f(az + bz') = f(ax + a\alpha x_0 + bx' + b\alpha' x_0) = (a\alpha + b\alpha')d = af(z) + bf(z'),$$

que se anula en M, $f(x) = 0$ para todo $x \in M$ y tal que $f(x_0) = d$. Además, como para todo $\tilde{x} \in M$, $d \leq \|x_0 - \tilde{x}\|$, tomando $\tilde{x} = -x/\alpha$, con $\alpha \neq 0$, obtenemos

$$|f(z)| = |f(x+\alpha x_0)| = |\alpha| d \leq |\alpha| \|x_0 - \tilde{x}\| = |\alpha| \|x_0 + x/\alpha\| = \|x + \alpha x_0\| = \|z\|,$$

luego $\|f\| \leq 1$. Por la definición de ínfimo tenemos que, cualquiera sea $\varepsilon > 0$, existe un $x' \in M$ tal que $\|x_0 - x'\| < d + \varepsilon$. Así, usando que $f(x_0 - x') = d$, tenemos que, para todo $\varepsilon > 0$,

$$\frac{|f(x_0 - x')|}{\|x_0 - x'\|} > \frac{d}{d + \varepsilon} = 1 - \frac{\varepsilon}{d + \varepsilon} = 1 - \widetilde{\varepsilon} \ \Rightarrow \ \sup_{x' \in M} \frac{|f(x_0 - x')|}{\|x_0 - x'\|} \geq 1,$$

donde la implicación es consecuencia de las propiedades del supremo[6]. Luego, usando que

$$\|f\| = \sup_{z \in M_1 \setminus \{x_0\}} \frac{|f(x_0 - z)|}{\|x_0 - z\|} \geq \sup_{x' \in M} \frac{|f(x_0 - x')|}{\|x_0 - x'\|} \geq 1$$

[6]Ver el pie de página 6 de la página 165.

se sigue que $\|f\| \geq 1$ y, por tanto, $\|f\| = 1$ en M_1. Luego, por el Teorema de Hahn-Banach 7.2.5, existe un funcional $F : \mathbb{X} \mapsto \mathbb{K}$, tal que $F(x) = f(x)$ en M_1, y $\|F\| = \|f\| = 1$, de donde se sigue el resultado. ■

Nota 7.2.13 *El Teorema 7.2.12 se puede considerar como la generalización a espacios normados del Teorema de la proyección ortogonal 5.2.15 para espacios de Hilbert.*

7.3. El Teorema de la aplicación abierta

Como hemos visto en la sección 3.3 una aplicación es continua (Definición 3.3.11) si y solo si la imagen inversa de cualquier subconjunto abierto de $\mathbb{Y}$ es un subconjunto abierto de $\mathbb{X}$ (Proposición 3.3.12). Sin embargo, una aplicación continua puede no transformar abiertos en abiertos. Por ejemplo, $f : \mathbb{R} \mapsto \mathbb{R}$, $f(x) = \sin x$ es continua pero transforma el abierto $(0, 2\pi)$ en el cerrado $[-1, 1]$. Las aplicaciones que transforman abiertos en abiertos se denominan *aplicaciones abiertas*. Así tenemos la siguiente definición:

Definición 7.3.1 *Sean $\mathbb{X}$ e $\mathbb{Y}$ dos espacios normados sobre $\mathbb{K}$. Una aplicación $A : \mathbb{X} \mapsto \mathbb{Y}$ se llama abierta si A transforma abiertos $U \subset \mathbb{X}$ en abiertos $A(U) \subset \mathbb{Y}$.*

Antes de probar el Teorema de la aplicación abierta conviene recordar dos definiciones del álgebra de conjuntos.

Sea $\mathbb{X}$ un espacio vectorial sobre $\mathbb{K} = \mathbb{R}$ o $\mathbb{C}$. Sea $M \subset \mathbb{X}$, $\alpha \in \mathbb{K}$ y $z \in \mathbb{X}$. Se definen los conjuntos $M + z$ y αM como

$$M + z = \{x + z \,:\, \forall x \in M\}, \qquad \alpha M = \{\alpha x \,:\, \forall x \in M\}.$$

Está claro de la definición anterior que si A es una aplicación lineal, entonces $A(M + z) = A(M) + Az$ y que $A(\alpha M) = \alpha A(M)$.

Comenzaremos probando un teorema fundamental conocido como el *Teorema de la bola abierta*.

En adelante $B_{\mathbb{X}} = \{x \in \mathbb{X} : \|x\| < 1\}$ denotará la bola unidad centrada en 0 en $\mathbb{X}$ y $B_{\mathbb{Y}} = \{y \in \mathbb{Y} : \|y\| < 1\}$ la bola unidad centrada en 0 en $\mathbb{Y}$.

Teorema 7.3.2 (de la bola abierta) *Sean $\mathbb{X}$ e $\mathbb{Y}$ dos espacios de Banach sobre $\mathbb{K}$. Sea $A \in \mathcal{B}(\mathbb{X}, \mathbb{Y})$, sobreyectiva, i.e., $A(\mathbb{X}) = \mathbb{Y}$. Entonces, existe un $\delta > 0$ tal que $\delta B_{\mathbb{Y}} \subseteq A(B_{\mathbb{X}})$, es decir, que para todo $y \in \mathbb{Y}$ tal que $\|y\| < \delta$, existe un $x \in \mathbb{X}$ tal que $\|x\| < 1$ y $Ax = y$.*

Demostración: La prueba la haremos en dos partes.

I. En la primera probaremos que existe un número $d > 0$ que cumple que, dado un $\varepsilon > 0$ y un $z \in \mathbb{Y}$, existe un $x \in \mathbb{X}$ tal que $\|Ax - z\| < \varepsilon$, y $\|x\| < \|z\|/d$.

Notemos que cualquiera sea $x \in \mathbb{X}$, $x \in kB_{\mathbb{X}}$, para algún[7] $k \in \mathbb{N}$, y como A es sobreyectiva, entonces para todo $y \in \mathbb{Y}$ existe un $x \in \mathbb{X}$ tal que $y = Ax$. Juntando ambas se tiene que

$$\mathbb{X} = \bigcup_{k=1}^{\infty} kB_{\mathbb{X}} \;\Rightarrow\; \mathbb{Y} = A(\mathbb{X}) = A\left(\bigcup_{k=1}^{\infty} kB_{\mathbb{X}}\right) = \bigcup_{k=1}^{\infty} A(kB_{\mathbb{X}}).$$

Ahora bien, como la unión de todos los $A(kB_{\mathbb{X}})$ es el espacio completo $\mathbb{Y}$, entonces si cambiamos los $A(kB_{\mathbb{X}})$ por sus clausuras no agregamos ningún punto nuevo a la unión y, por tanto,

$$\mathbb{Y} = \bigcup_{k=1}^{\infty} A(kB_{\mathbb{X}}) = \bigcup_{k=1}^{\infty} \overline{A(kB_{\mathbb{X}})}.$$

Pero $\mathbb{Y}$ es de Banach, luego por el corolario del Teorema de Baire 3.6.2 al menos un $\overline{A(kB_{\mathbb{X}})}$ contiene un abierto, i.e., existe un $r > 0$ y un $y_0 \in \mathbb{Y}$ tal que $B(y_0, r) \subset \overline{A(kB_{\mathbb{X}})}$. Notemos que si $y \in \mathbb{Y}$ es tal que $\|y\| < r$, entonces $y + y_0 \in B(y_0, r)$, luego $y + y_0 \in \overline{A(kB_{\mathbb{X}})}$. Como A es acotado (luego continuo), entonces, cualquiera sea $y \in \mathbb{Y}$ tal que $\|y\| < r$, han de existir dos sucesiones $(x'_n)_n$ y $(x''_n)_n$ en $kB_{\mathbb{X}}$, tales que

$$Ax'_n \to y_0, \quad Ax''_n \to y + y_0 \;\Rightarrow\; x_n = x''_n - x'_n, \text{ es tal que } Ax_n \to y.$$

Además, como $(x'_n)_n, (x''_n)_n \in kB_{\mathbb{X}}$, la sucesión $(x_n)_n$ es acotada y satisface la desigualdad

$$\|x_n\| = \|x''_n - x'_n\| \leq \|x''_n\| + \|x'_n\| < 2k.$$

[7]Basta que k sea tal que $\|x\| < k$.

Sea ahora $0 \neq z \in \mathbb{Y}$ cualquiera y definamos y tal que

$$y = \frac{r}{2}\frac{z}{\|z\|}, \quad y \in \mathbb{Y}, \quad \|y\| = \frac{r}{2} < r.$$

Entonces, como acabamos de probar, existirá una sucesión $(x_n)_n$, tal que

$$\|x_n\| < 2k, \quad Ax_n \to \frac{r}{2}\frac{z}{\|z\|} \;\Rightarrow\; A\left(\frac{2\|z\|}{r}x_n\right) \to z.$$

Luego, dado un $\varepsilon > 0$ existe un $n \in \mathbb{N}$ tal que podemos elegir un $x \in \mathbb{X}$

$$x = \frac{2\|z\|}{r}x_n \quad \text{tal que} \quad \|Ax - z\| < \varepsilon$$

y que cumple con

$$\|x\| = \frac{2\|z\|}{r}\|x_n\| < \frac{4k}{r}\|z\| = \frac{1}{d}\|z\|,$$

tal y como se quería probar.

II. Utilizando lo demostrado en **I.** vamos a construir inductivamente una sucesión $s_n \in \mathbb{X}$ de Cauchy que nos permitirá demostrar el teorema.

Sea d el número de la parte I. Sea $dB_{\mathbb{Y}} \subset \mathbb{Y}$ la bola de radio d centrada en 0, donde $d > 0$ es el número mencionado de la parte I. Fijemos $\varepsilon = d/2$ y $z = y \in dB_{\mathbb{Y}}$. Por **I.** existe un $x_1 \in \mathbb{X}$ con norma $\|x_1\| < \|y\|/d < 1$, tal que $\|y - Ax_1\| < d/2$. Sea ahora, $\varepsilon = d/4$ y $z = y - Ax_1$, entonces existe un $x_2 \in \mathbb{X}$ tal que

$$\|(y - Ax_1) - Ax_2\| < \frac{d}{4}, \quad \|x_2\| < \frac{1}{d}\|z\| = \frac{1}{d}\|y - Ax_1\| < \frac{1}{2}.$$

Y así sucesivamente, eligiendo $\varepsilon = d/2^n$, $z = y - Ax_1 - \cdots - Ax_{n-1}$, encontraremos un $x_n \in \mathbb{X}$ tal que

$$\|(y - Ax_1 - \cdots - Ax_{n-1}) - Ax_n\| < \frac{d}{2^n}, \quad \|x_n\| < \frac{1}{2^{n-1}}.$$

En efecto, si asumimos que lo anterior cierto (lo es para $n = 1$ y $n = 2$), entonces si elegimos $\varepsilon = d/2^{n+1}$, $z = y - Ax_1 - \cdots - Ax_n$, por **I.** sabemos que existe $x_{n+1} \in \mathbb{X}$ tal que

$$\|(y - Ax_1 - \cdots - Ax_n) - Ax_{n+1}\| < \frac{d}{2^{n+1}}, \quad \|x_{n+1}\| < \frac{1}{d}\|z\| < \frac{1}{2^n}.$$

Definamos para cada $n \in \mathbb{N}$, $s_n = x_1 + x_2 + \cdots + x_n$. Sea $m > n$

$$\|s_m - s_n\| = \|x_{n+1} + \cdots + x_m\| \leq \sum_{k=n+1}^{m} \|x_k\| \leq \sum_{k=n+1}^{m} \frac{1}{2^{k-1}} \to 0 \quad n, m \to \infty,$$

i.e., s_n es de Cauchy, luego $s_n \to x \in \mathbb{X}$, pues $\mathbb{X}$ es de Banach. Además,

$$\|x\| \leq \sum_{k=1}^{\infty} \|x_k\| < \sum_{k=1}^{\infty} \frac{1}{2^{k-1}} = 2 \;\Rightarrow\; x \in 2B_{\mathbb{X}}.$$

Por otro lado,

$$\|y - As_n\| < \frac{d}{2^n} \overset{n\to\infty}{\longrightarrow} 0 \;\Rightarrow\; As_n \overset{n\to\infty}{\longrightarrow} y \;\Rightarrow\; Ax = y.$$

Es decir, cualquiera sea $y \in dB_{\mathbb{Y}}$ existe un $x \in 2B_{\mathbb{X}}$ tal que $y = Ax$, es decir, cualquier y de $dB_{\mathbb{Y}}$ es imagen de algún $x \in 2B_{\mathbb{X}}$, luego $dB_{\mathbb{Y}} \subseteq A(2B_{\mathbb{X}})$. Usando la linealidad de A se tiene entonces que $\delta B_{\mathbb{Y}} \subseteq A(B_{\mathbb{X}})$, con $\delta = d/2$. ■

Nótese que el teorema anterior indica que la imagen $A(B_{\mathbb{X}})$ de la bola unidad $B_{\mathbb{X}}$ en $\mathbb{X}$ de cualquier aplicación lineal acotada sobreyectiva siempre contiene un abierto en $\mathbb{Y}$. De hecho, una consecuencia del Teorema de la bola abierta 7.3.2 es el Teorema de la aplicación abierta.

Teorema 7.3.3 (de la aplicación abierta) *Sean $\mathbb{X}$ e $\mathbb{Y}$ dos espacios de Banach sobre $\mathbb{K}$. Sea $A \in \mathcal{B}(\mathbb{X}, \mathbb{Y})$, sobreyectiva, i.e., $A(\mathbb{X}) = \mathbb{Y}$. Entonces cualquiera sea el abierto $U \subset \mathbb{X}$, su imagen $A(U) \subset \mathbb{Y}$ es un abierto en $\mathbb{Y}$.*

Demostración: Sea un abierto cualquiera $U \subset \mathbb{X}$ y $x_0 \in U$. Tenemos que probar que cualquiera sea $y_0 \in A(U)$, existe una bola $B(y_0, d) \subset A(U)$, i.e., todos los puntos de $A(U)$ son interiores. Para ello, como A es sobreyectiva, bastará probar que para todo y_0 el conjunto $A(U)$ contiene una bola (abierta) centrada en $y_0 = Ax_0$. Sea $U' = U - x_0$. U' es un abierto que contiene a $x = 0$, luego existe un $r > 0$ tal que $rB_{\mathbb{X}} \subseteq U'$ (U' contiene a la bola abierta $B(0, r)$). Pero entonces, el Teorema de la bola abierta 7.3.2 nos dice que existe un $\delta > 0$ tal que $\delta B_{\mathbb{Y}} \subseteq A(B_{\mathbb{X}})$, luego, usando la linealidad de A

$$\delta B_{\mathbb{Y}} \subseteq A(B_{\mathbb{X}}) \quad \Leftrightarrow \quad r\delta B_{\mathbb{Y}} \subseteq rA(B_{\mathbb{X}}) \quad \Leftrightarrow \quad r\delta B_{\mathbb{Y}} \subseteq A(rB_{\mathbb{X}}) \subseteq A(U'),$$

de donde se deduce que

$$A(U) = A(U' + x_0) = A(U') + Ax_0 \supseteq r\delta B_{\mathbb{Y}} + Ax_0 = r\delta B_{\mathbb{Y}} + y_0,$$

es decir, existe una bola $B(y_0, d) \subset A(U)$ con $d = r\delta$, como y_0 era arbitrario, $A(U)$ es abierto. ∎

Como consecuencia del Teorema de la aplicación abierta 7.3.3 tenemos el Teorema de Banach de las aplicaciones inversas acotadas que ya mencionamos en la página 224.

Teorema 7.3.4 (de la inversa acotada de Banach) *Sea $A \in \mathcal{B}(\mathbb{X}, \mathbb{Y})$, con $\mathbb{X}, \mathbb{Y}$ espacios de Banach, biyectiva, i.e., A es tal que el núcleo de A, $\mathcal{N}(A) = \{0\}$, y la imagen de A, $\mathcal{I}(A) = \mathbb{Y}$, entonces $A^{-1} \in \mathcal{B}(\mathbb{Y}, \mathbb{X})$.*

Demostración: Nótese que como $\mathcal{N}(A) = \{0\}$, A es invertible y A^{-1} es lineal (ver Teorema 4.4.9), y como $\mathcal{I}(A) = \mathbb{Y}$, A es sobreyectivo. Solo hemos de probar que A^{-1} es acotado, o equivalentemente, por el Teorema 4.4.18, que A^{-1} es continua. Pero una aplicación es continua si y solo si su imagen inversa trasforma abiertos en abiertos (ver la Proposición 3.3.12). Sea U un abierto de $\mathbb{X}$, entonces como A es biyectiva $(A^{-1})^{-1}(U) = A(U)$. Pero A, por el Teorema de la aplicación inversa 7.3.3, es una aplicación abierta, luego $A(U)$ es un abierto y, por tanto, A^{-1} es continua, luego acotada. ∎

Nota 7.3.5 *En el capítulo anterior vimos que podíamos tener operadores acotados cuya inversa no era acotada. Por ejemplo, el operador del Ejemplo 4 de la página 223. El lector puede además comprobar que ese mismo operador definido de ℓ^∞ a ℓ^∞ tampoco es sobreyectivo. Otro ejemplo es considerar el mismo operador pero definido sobre el subespacio $\mathbb{X} \subset \ell^\infty$ de sucesiones que tengan solo un número finito de términos no nulos (ver el Problema 4.18). En este último caso lo que falla es que $\mathbb{X}$ no es completo.*

7.4. Operadores cerrados y el Teorema del grafo cerrado

Definición 7.4.1 *Sean $\mathbb{X}, \mathbb{Y}$ espacios normados, y A un operador, $A : \mathcal{D}(A) \subset \mathbb{X} \mapsto \mathbb{Y}$. Se denomina grafo de A al conjunto de pares ordena-*

dos

$$\mathcal{G}(A) := \{(x,y)\,|\,x\in\mathcal{D}(A),\ y = Ax\} = \{(x,Ax)\,|\,x\in\mathcal{D}(A)\}\subset\mathbb{X}\times\mathbb{Y}.$$

Nótese que el espacio $\mathbb{X}\times\mathbb{Y}$ es un espacio vectorial si definimos las operaciones

$$(x_1,y_1)+(x_2,y_2) = (x_1+x_2, y_1+y_2), \quad \alpha(x,y) = (\alpha x, \alpha y), \quad \alpha\in\mathbb{K}.$$

Si en $\mathbb{X}\times\mathbb{Y}$ definimos la aplicación $\|(x,y)\| = \|x\|+\|y\|$, $\mathbb{X}\times\mathbb{Y}$ es un espacio normado. Si además $\mathbb{X}$ e $\mathbb{Y}$ son de Banach, $\mathbb{X}\times\mathbb{Y}$ es de Banach.

Que $\mathbb{X}\times\mathbb{Y}$ con la norma $\|(x,y)\| = \|x\|+\|y\|$ es efectivamente un espacio normado es un ejercicio sencillo que dejamos al lector[8]. Para probar que es de Banach tomemos una sucesión $(z_n)_n\in\mathbb{X}\times\mathbb{Y}$, $z_n = (x_n,y_n)$, de Cauchy. Entonces, para todo $\varepsilon>0$, si n,m son suficientemente grandes, $\|z_n - z_m\|<\varepsilon$. Como

$$\|z_n - z_m\| = \|x_n - x_m\| + \|y_n - y_m\| < \varepsilon \;\Rightarrow\; \|x_n - x_m\|<\varepsilon,\ \|y_n - y_m\|<\varepsilon,$$

luego $(x_n)_n$ y $(y_n)_n$ son de Cauchy, y como $\mathbb{X}$ e $\mathbb{Y}$ son de Banach, ambas tienen límite x e y. El lector puede comprobar entonces que $z = (x,y)\in\mathbb{X}\times\mathbb{Y}$ es el límite de z_n cuando $n\to\infty$.

Definición 7.4.2 *Sean $\mathbb{X},\mathbb{Y}$ espacios normados y sea $A:\mathcal{D}(A)\mapsto\mathbb{Y}$ un operador lineal (no necesariamente acotado) cuyo dominio es $\mathcal{D}(A)\subset\mathbb{X}$. A es un operador cerrado si su grafo $\mathcal{G}(A)$ es cerrado en $\mathbb{X}\times\mathbb{Y}$.*

La siguiente proposición proporciona una caracterización muy útil de los operadores cerrados.

Proposición 7.4.3 (del operador cerrado) *Sean $\mathbb{X},\mathbb{Y}$ espacios normados y sea A un operador lineal (no necesariamente acotado) $A:\mathcal{D}(A)\subset\mathbb{X}\mapsto\mathbb{Y}$. A es cerrado si y solo si cumple que si $x_n\to x$, $x_n\in\mathcal{D}(A)$ y $Ax_n\to y$, entonces $x\in\mathcal{D}(A)$ y $Ax = y$.*

Demostración: Supongamos que $\mathcal{G}(A)$ es cerrado en $\mathbb{X}\times\mathbb{Y}$. Entonces $\mathcal{G}(A) = \overline{\mathcal{G}(A)}$, luego, para todo $z = (x,y)\in\overline{\mathcal{G}(A)}$, $z\in\mathcal{G}(A)$ y, en particular, $y = Ax$. Ahora bien, por la Proposición 3.5.4 a), $z\in\overline{\mathcal{G}(A)}$ si y solo si

[8] Otras normas utilizadas son $\|(x,y)\|_\infty = \max\{\|x\|,\|y\|\}$ y $\|(x,y)\|_2 = \sqrt{\|x\|^2+\|y\|^2}$.

existe una sucesión $z_n = (x_n, Ax_n) \in \mathcal{G}(A)$ tal que $z_n \to z$, en cuyo caso $x_n \to x$, $x \in \mathcal{D}(A)$ y $Ax_n \to Ax = y$. En otras palabras, si $\mathcal{G}(A)$ es cerrado, $x_n \to x$ y $Ax_n \to y$, entonces $x \in \mathcal{D}(A)$ e $y = Ax$.

Supongamos, por el contrario, que $x_n \to x$, $x_n \in \mathcal{D}(A)$ y $Ax_n \to y$ implica que $x \in \mathcal{D}(A)$ y $Ax = y$. Sea $z = (x, y) \in \overline{\mathcal{G}(A)}$ cualquiera. Entonces existen $x_n \in \mathcal{D}(A)$, $x_n \to x$, $y_n = Ax_n \to y$, pero de lo anterior se sigue que $x \in \mathcal{D}(A)$ y que $y = Ax$, luego $z = (x, Ax) \in \mathcal{G}(A)$, i.e., $\overline{\mathcal{G}(A)} = \mathcal{G}(A)$, luego $\mathcal{G}(A)$ es cerrado y, por tanto, A también. ■

Ejemplo 7.4.4 *Un operador cerrado no tiene que ser acotado y uno acotado no tiene que ser cerrado.*

Por ejemplo, el operador derivada $D : \mathcal{D}(D) \subset C^\infty_{[0,1]} \mapsto C^\infty_{[0,1]}$ ($Dx = x'$) del Ejemplo 4.4.4 es no acotado (ver Ejemplo 4.4.15). Tomemos $\mathbb{X} = C^\infty_{[0,1]}$. Escojamos una sucesión $x_n \in \mathcal{D}(D)$ tal que $x_n \to x$ y $Dx_n = x'_n \to y$. Como la convergencia en $C^\infty_{[0,1]}$ es la convergencia uniforme, entonces $x'_n \to y$ implica

$$\int_0^t y(s)ds = \int_0^t \lim_{n\to\infty} x'_n(s)ds = \lim_{n\to\infty} \int_0^t x'_n(s)ds = \lim_{n\to\infty} (x_n(s) - x_n(0))$$
$$= x(t) - x(0) \;\Rightarrow\; x(t) = x(0) + \int_0^t y(s)ds.$$

Así que $x \in \mathcal{D}(D)$ y $Dx = x' = y$, luego la Proposición 7.4.3 implica que D es cerrado.

Para probar la afirmación contraria, tomemos un espacio normado $\mathbb{X}$ cualquiera y definamos el operador $T : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathcal{D}(T)$, $Tx = x$, donde $\mathcal{D}(T)$ es un subespacio denso en $\mathbb{X}$ y distinto de $\mathbb{X}$. Está claro que T es lineal y acotado (de hecho es el operador identidad de $\mathcal{D}(T)$). Elijamos $x \in \mathbb{X}$, $x \notin \mathcal{D}(T)$. Como $\mathcal{D}(T)$ es denso en $\mathbb{X}$ existe una sucesión $(x_n)_n$ tal que $x_n \to x$, sin embargo $x \notin \mathcal{D}(T)$, luego por la Proposición 7.4.3, T no es cerrado. ■

Teorema 7.4.5 (del grafo cerrado) *Sean $\mathbb{X}, \mathbb{Y}$ espacios de Banach y $A : \mathcal{D}(A) \subset \mathbb{X} \mapsto \mathbb{Y}$ un operador lineal cerrado. Entonces, si $\mathcal{D}(A)$ es cerrado en $\mathbb{X}$, A es acotado.*

Demostración: Como $\mathcal{D}(A)$ es cerrado en $\mathbb{X}$ y $\mathcal{G}(A)$ lo es en $\mathbb{X} \times \mathbb{Y}$, entonces ambos son subespacios completos (¿por qué?). Consideremos la aplicación $T : \mathcal{G}(A) \mapsto \mathcal{D}(A)$, $Tz = T(x, Ax) = x$. Está claro que T es lineal pues si $z_1 = (x_1, y_1)$, $z_2 = (x_2, y_2)$, $\alpha, \beta \in \mathbb{K}$,

$$T(\alpha z_1 + \beta z_2) = \alpha x_1 + \beta x_2 = \alpha T z_1 + \beta T z_2.$$

Veamos que T es acotada

$$\|Tz\| = \|T(x, Ax)\| = \|x\| \leq \|x\| + \|Ax\| = \|(x, Ax)\| = \|z\|.$$

Por otro lado, $Tz = T(x, Ax) = x$, luego $Tz = 0$ implica $x = 0$, pero si $x = 0$, $Ax = 0$, luego la única solución de $Tz = 0$ es $z = 0$, por tanto el Teorema 4.4.9 implica que T es invertible (e inyectiva). Pero, de la definición de grafo de A, cualquiera sea $x \in \mathcal{D}(A)$, la ecuación $Tz = T(x, Ax) = x$, siempre tiene solución: $z = (x, Ax)$. Luego T es sobreyectiva (luego es biyectiva), de hecho, $T^{-1} : \mathcal{D}(A) \mapsto \mathcal{G}(A)$, $T^{-1}x = (x, Ax)$. Pero si $T : \mathcal{G}(A) \mapsto \mathcal{D}(A)$ es biyectiva y $\mathcal{D}(A)$ y $\mathcal{G}(A)$ son de Banach, entonces el Teorema de la inversa acotada de Banach 7.3.4 nos dice que T^{-1} es acotado. Luego, existe un $c > 0$ tal que, para todo $x \in \mathbb{X}$,

$$\|T^{-1}x\| = \|(x, Ax)\| \leq c\|x\| \;\Rightarrow\; \|Ax\| \leq \|Ax\| + \|x\| = \|(x, Ax)\| \leq c\|x\|,$$

de donde se deduce que A es acotado. ∎

7.5. Convergencia débil

Definición 7.5.1 *Sea $\mathbb{X}$ un espacio normado, y sea la sucesión $(x_n)_n \in \mathbb{X}$. Se dice que que x_n converge débilmente a $x \in \mathbb{X}$ y lo denotaremos[a] por $x_n \rightharpoonup x$, si para todo $f \in \mathbb{X}' = \mathcal{B}(\mathbb{X}, \mathbb{C})$, $f(x_n) \to f(x)$.*

[a]En muchos textos se usa la notación $x_n \xrightarrow{*} x$, o $x_n \xrightarrow{w} x$, donde w significa *weak* (en inglés, débil).

Nótese que como f es un funcional, $f(x_n)$ y $f(x)$ son, en general, para cada x_n y x, números complejos.

Esta definición difiere de la Definición 4.2.10 pues allí $x_n \to x$ cuando $\|x_n - x\| \to 0$. Si ese es el caso diremos que x_n converge en norma a x (o simplemente que x_n converge a x).

Proposición 7.5.2 *Sea $\mathbb{X}$ un espacio normado, y sea la sucesión $(x_n)_n \in \mathbb{X}$. Entonces*

1. *Si $(x_n)_n$ converge débilmente a $x \in \mathbb{X}$, x es único.*
2. *Si $(x_n)_n$ converge débilmente, entonces x_n es acotada.*
3. *Si $(x_n)_n$ converge en norma a x, entonces también converge débilmente.*

Demostración: 1. Sea $x_n \rightharpoonup x$ y $x_n \rightharpoonup z$, $x \neq z$ y $f(x_n) \to f(x)$ y $f(x_n) \to f(z)$. Pero $f(x_n)$ es una sucesión numérica, $f(x_n)$ tiene un único límite, luego $f(x) = f(z)$, i.e., $f(x-z) = 0$ para todo $f \in \mathbb{X}'$. Entonces el corolario 7.2.10 del Teorema de Hahn-Banach implica que $x = z$.

2. Sea $x_n \rightharpoonup x$. Para cada n definamos los funcionales $\phi_n : \mathbb{X}' \mapsto \mathbb{C}$, $\phi_n(f) = f(x_n)$, donde x_n es el correspondiente elemento de la sucesión $(x_n)_n \in \mathbb{X}$. Está claro que la sucesión $f(x_n)$ converge cuando $n \to \infty$ (¿por qué?), luego es acotada. Así que existe un $M_f > 0$ tal que, cualquiera sea $f \in \mathbb{X}'$ se tiene, para todo n,

$$|\phi_n(f)| = |f(x_n)| \leq M_f, \qquad \forall f \in \mathbb{X}'.$$

Como $\mathbb{X}' = \mathcal{B}(\mathbb{X}, \mathbb{C})$ es completo, podemos aplicar el Teorema de Banach-Steinhaus 4.5.1, luego existe un $M > 0$, tal que para todo $n \in \mathbb{N}$, $\|\phi_n\| \leq M$, pero

$$\|\phi_n\| = \sup_{f \in \mathbb{X}' \,:\, \|f\|=1} |\phi_n(f)| = \sup_{f \in \mathbb{X}' \,:\, \|f\|=1} |f(x_n)| = \|x_n\|,$$

donde la última igualdad se sigue del Corolario 7.2.11. Luego, para todo $n \in \mathbb{N}$, $\|x_n\| \leq M$.

3. Como $x_n \to x$ en norma, tenemos, para todo $f \in \mathbb{X}'$,

$$|f(x_n) - f(x)| = |f(x_n - x)| \leq \|f\| \|x - x_n\| \to 0,$$

de donde se sigue el resultado. ■

Hemos visto que si $x_n \to x$ en norma, entonces lo hace débilmente. El recíproco es falso en general. En efecto, sea $\mathbb{X} = \mathbb{H}$ un espacio de Hilbert separable de dimensión infinita y sea $(e_n)_n$ una base ortonormal de $\mathbb{H}$. Por el Teorema de representación de Riesz 5.2.21 cualquiera sea $f \in \mathbb{H}' =$

$\mathcal{B}(\mathbb{H},\mathbb{C})$, existe un único $z \in \mathbb{H}$, tal que $f(x) = \langle x, z\rangle$. Entonces $f(e_n) = \langle e_n, z\rangle$ que tiende a cero si $n \to \infty$ (ver (5.2.4)), luego $e_n \rightharpoonup 0$, i.e., $(e_n)_n$ converge débilmente a 0. Sin embargo, la propia sucesión $(e_n)_n$ no es de Cauchy pues si $m > n$, $\|e_n - e_m\| = \sqrt{2}$, que no es convergente en la norma de $\mathbb{H}$.

Antes de pasar a discutir el caso cuando $\dim \mathbb{X} < \infty$ conviene mencionar que del Teorema de representación de Riesz 5.2.21 se sigue el siguiente importante resultado:

Teorema 7.5.3 *Sea $\mathbb{H}$ un espacio de Hilbert y sea $(x_n)_n \in \mathbb{H}$. Entonces, $x_n \rightharpoonup x$ si y solo si para todo $z \in \mathbb{H}$, $\langle x_n, z\rangle \stackrel{n\to\infty}{\longrightarrow} \langle x, z\rangle$.*

Demostración: En efecto, por el Teorema de representación de Riesz 5.2.21 sabemos que cualquiera sea $f \in \mathbb{H}' = \mathcal{B}(\mathbb{H},\mathbb{C})$, existe un único $z \in \mathbb{H}$, tal que $f(x) = \langle x, z\rangle$. Luego que $f(x_n) \to f(x)$ es equivalente a decir que $\langle x_n, z\rangle \to \langle x, z\rangle$. ■

Lo anterior nos dice que, en general, en dimensión infinita las convergencias débil y en norma no son necesariamente equivalentes. Sin embargo en dimensión finita sí que lo son. De hecho tenemos el siguiente teorema:

Teorema 7.5.4 *Sea la sucesión $(x_n)_n \in \mathbb{X}$, $\mathbb{X}$ espacio normado. Si $\dim \mathbb{X} = k < \infty$, entonces la convergencia en norma y la débil son equivalentes.*

Demostración: Que la convergencia en norma implica la débil lo hemos probado en la Proposición 7.5.2. Probemos que la débil implica la convergencia en norma. Sea $(e_n)_n$ una base normalizada a la unidad de $\mathbb{X}$, y sea $(x_n)_n$ una sucesión de elementos de $\mathbb{X}$, y $x \in \mathbb{X}$. Entonces, para todo $n \in \mathbb{N}$, podemos escribir

$$x_n = \sum_{i=1}^{k} \alpha_i^{(n)} e_i, \qquad x = \sum_{i=1}^{k} \alpha_i e_i.$$

Sean los funcionales $f_j : \mathbb{X} \mapsto \mathbb{C}$ definidos en (7.1.2) –i.e., $f_j(e_i) = \delta_{i,j}$. Como hemos visto en la sección 7.1, dichos funcionales son una base de $\mathbb{X}'$ y cumplen con que $f_j(x_n) = \alpha_j^{(n)}$, $f(x) = \alpha_j$, $j = 1, 2, \ldots, k$. Además, como $x_n \rightharpoonup x$, entonces cualquiera sea $f \in \mathbb{X}'$, $f(x_n) \to f(x)$ (en particular si $f = f_j$) de donde obtenemos que $\alpha_j^{(n)} \to \alpha_j$, $j = 1, \ldots, k$. Pero entonces

$$\|x_n - x\| = \left\| \sum_{i=1}^{k} (\alpha_i^{(n)} - \alpha_i) e_i \right\| \le \sum_{i=1}^{k} |\alpha_i^{(n)} - \alpha_i| \to 0,$$

i.e., $x_n \to x$ en norma. ■

Para terminar este capítulo demostraremos un teorema que garantiza la convergencia débil a partir de la convergencia débil en un subespacio de $\mathbb{X}'$. Para ello, notemos que la definición de sistema de vectores completo 5.2.3 podemos extenderla a cualquier espacio normado $\mathbb{X}$ en general, y de Banach en particular. Así, diremos que un sistema de vectores linealmente independientes $(\phi_n)_n \in \mathbb{X}$ es completo en $\mathbb{X}$ si para todo vector $x \in \mathbb{X}$ y cualquiera sea $\varepsilon > 0$ existe una combinación lineal finita de los vectores ϕ_n tal que

$$\left\| x - \sum_{k=1}^{n} \alpha_k \phi_k \right\| < \varepsilon\,.$$

Teorema 7.5.5 *Sea la sucesión acotada $(x_n)_n \in \mathbb{X}$, en un espacio normado $\mathbb{X}$. Sea $(\phi_k)_k \subset \mathbb{X}'$ un sistema completo*[a] *de funcionales linealmente independientes en $\mathbb{X}'$. Si $\phi_k(x_n) \to \phi_k(x)$, para todo k, entonces $x_n \rightharpoonup x$.*

[a] Recuérdese que $\mathbb{X}' = \mathcal{B}(\mathbb{X}, \mathbb{C})$ es un espacio de Banach.

Demostración: Queremos probar que $f(x_n) \to f(x)$ para todo $f \in \mathbb{X}'$. Como $(\phi_k)_k$ es un sistema completo, entonces cualquiera sea $f \in \mathbb{X}'$ y $\varepsilon > 0$, existe un $m > 0$, tal que $\Phi_m = \sum_{k=1}^{m} \alpha_k \phi_k$ cumple con

$$\|f - \Phi_m\| < \varepsilon.$$

Como los ϕ_k son acotados, Φ_m es acotado para todo $m \in \mathbb{N}$. Además, para todo m, $\Phi_m(x_n) \overset{n\to\infty}{\longrightarrow} \Phi_m(x)$ (¿por qué?), luego para dicho $\varepsilon > 0$, existe un $N \in \mathbb{N}$, tal que para todo $n > N$,

$$|\Phi_m(x_n) - \Phi_m(x)| < \varepsilon.$$

Luego, para todo $\varepsilon > 0$ existe un $N \in \mathbb{N}$, tal que para todo $n > N$

$$\begin{aligned}|f(x_n) - f(x)| &\leq |f(x_n) - \Phi_m(x_n)| + |\Phi_m(x_n) - \Phi_m(x)| + |\Phi_m(x) - f(x)| \\ &\leq \|f - \Phi_m\|\|x_n\| + \varepsilon + \|\Phi_m - f\|\|x\|.\end{aligned}$$

Como x_n es acotada, su límite también lo es, i.e., existe $M > 0$ tal que para todo n, $\|x_n\| < M$ y $\|x\| \leq M$. Entonces

$$|f(x_n) - f(x)| < (2M + 1)\varepsilon,$$

que podemos hacer tan pequeño como queramos. Como f era arbitrario, se tiene que $x_n \rightharpoonup x$, como se quería probar. ■

7.6. Problemas

Problema 7.1 Prueba que el espacio dual de ℓ^1 es ℓ^∞.

Solución: Sea $e_k = \delta_{i,k}$ la base de Schauder de ℓ^1, entonces, cualquiera sea $x \in \ell^1$, $x = \sum_{k=1}^\infty x_k e_k$. Sea $f \in \ell^{1\prime}$ el espacio dual de $\mathcal{B}(\ell^1, \mathbb{C})$. Usando que f es lineal y acotado (luego continuo) tenemos

$$f(x) = \sum_{k=1}^\infty x_k c_k, \quad c_k = f(e_k).$$

Sea $c = (c_1, c_2, \dots, c_k, \dots)$. Como f es acotado, entonces, para todo $k \in \mathbb{N}$ tenemos,

$$|c_k| = |f(e_k)| \le \|f\|\|e_k\| = \|f\| \;\Rightarrow\; \|c\|_\infty = \sup_{k\in\mathbb{N}} |c_k| \le \|f\| \;\Rightarrow\; c \in \ell^\infty.$$

Por otro lado

$$|f(x)| = \left|\sum_{k=1}^\infty x_k c_k\right| \le \|x\|_1 \|c\|_\infty \;\Rightarrow\; \|f\| = \sup_{\|x\|_1=1} |f(x)| \le \|c\|_\infty \;\Rightarrow\; \|f\| = \|c\|_\infty.$$

Probemos ahora que de cualquier elemento de ℓ^∞ podemos definir un funcional lineal sobre ℓ^1. Sea $b \in \ell^\infty$ cualquiera. Definamos el funcional g sobre ℓ^1 de la forma

$$g(x) = \sum_{k=1}^\infty x_k b_k.$$

Está claro que g es lineal, además, como

$$|g(x)| \le \sum_{k=1}^\infty |x_k||b_k| \le \|b\|_\infty \sum_{k=1}^\infty |x_k| = \|b\|_\infty \|x\|_1,$$

g es acotado. De lo anterior deducimos que $g \in \mathcal{B}(\ell^1, \mathbb{C}) = \ell^{1\prime}$ y que a cada $g \in \ell^\infty$ le corresponde uno y solo un $b \in \ell^\infty$.

Lo anterior nos indica que la aplicación $T : \ell^{1\prime} \mapsto \ell^\infty$, $Tf = c$, es lineal, biyectiva y es tal que $\|Tf\| = \|c\|_\infty = \|f\|$, i.e., T es una isometría por lo que $\ell^{1\prime}$ es isomorfo a ℓ^∞, luego $\ell^{1\prime} = \ell^\infty$. Conviene hacer notar que $\ell^{\infty\prime} \ne \ell^1$. ■

Problema 7.2 Prueba que el espacio dual de ℓ^p, $p > 1$, es ℓ^q, con q tal que $1/q + 1/p = 1$. **Ayuda:** Repite los pasos del Problema 7.1 y usa la desigualdad de Hölder (3.7.3) para acotar $f(x)$. Ver e.g. [11, §2.10-7, pág. 122].

Problema 7.3 Sea $T : \mathbb{X} \mapsto \mathbb{Y}$ un operador lineal acotado, $\mathbb{X}, \mathbb{Y}$ espacios de Banach. Si T es biyectivo entonces existen dos números reales positivos a, b tales que, para todo $x \in \mathbb{X}$, $a\|x\| \leq \|Tx\| \leq b\|x\|$. **Ayuda:** Usa el Teorema de la inversa acotada de Banach 7.3.4.

Solución: Como T es acotado, existe $b > 0$ tal que, para todo $x \in \mathbb{X}$, $\|Tx\| \leq b\|x\|$. Pero por el Teorema de la inversa acotada de Banach 7.3.4, T^{-1} es acotado. Luego

$$\|x\| = \|T^{-1}(Tx)\| \leq \|T^{-1}\|\|Tx\| \;\Rightarrow\; \exists a = \|T^{-1}\|^{-1} > 0 \text{ tal que } a\|x\| \leq \|Tx\|,$$

de donde se deduce el resultado. ■

Problema 7.4 Sean $\mathbb{X}_1 = (\mathbb{X}, \|\cdot\|_1)$ y $\mathbb{X}_2 = (\mathbb{X}, \|\cdot\|_2)$ dos espacios de Banach de dimensión no necesariamente finita. Si existe una constante b tal que $\|x\|_1 \leq b\|x\|_2$ para todo $x \in \mathbb{X}$, entonces existe una constante a tal que $\|x\|_2 \leq a\|x\|_1$, es decir, ambas normas son equivalentes. **Ayuda:** Define $T : \mathbb{X}_2 \mapsto \mathbb{X}_1$, $Tx = x$, prueba que es acotado y usa el Teorema de la inversa acotada de Banach 7.3.4.

Solución: Sea la aplicación $T : \mathbb{X}_2 \mapsto \mathbb{X}_1$, $Tx = x$. Claramente T es biyectiva (es la identidad en $\mathbb{X}$ como espacio vectorial), lineal y acotada pues

$$\|Tx\|_1 = \|x\|_1 \leq b\|x\|_2.$$

Entonces, el Teorema de la inversa acotada de Banach 7.3.4 nos dice que $T^{-1} : \mathbb{X}_1 \mapsto \mathbb{X}_2$, $T^{-1}x = x$, es acotada. Entonces

$$\|x\|_2 = \|T^{-1}x\|_2 \leq a\|x\|_1,$$

de donde se sigue el resultado. ■

Problema 7.5 Sea $A : \mathcal{D}(A) \subset \mathbb{X} \mapsto \mathbb{Y}$ un operador lineal acotado, $\mathbb{X}, \mathbb{Y}$ espacios normados. Prueba que

1. Si $\mathcal{D}(A)$ es cerrado en $\mathbb{X}$, entonces A es cerrado.

2. Si A es cerrado e $\mathbb{Y}$ es de Banach, entonces $\mathcal{D}(A)$ es cerrado en $\mathbb{X}$.

Ayuda: Usa la Proposición 7.4.3.

Solución: 1. Si la sucesión $(x_n)_n \in \mathcal{D}(A)$ converge, i.e., $x_n \to x$, entonces $(Ax_n)_n$ también converge (A es acotado, luego continuo). Pero si $x_n \to x$, entonces, por la Proposición 3.5.4, $x \in \overline{\mathcal{D}(A)} = \mathcal{D}(A)$, pues $\mathcal{D}(A)$ es cerrado. Pero entonces $Ax_n \to Ax = y \in \mathbb{Y}$. Así que aplicando la Proposición 7.4.3 se sigue que A es cerrado.

2. Dado $x \in \overline{\mathcal{D}(A)}$ cualquiera, existe una sucesión $(x_n)_n \in \mathcal{D}(A)$ tal que $x_n \to x$. Como A es acotado, Ax_n es de Cauchy, pues para todos $m > n$,

$$\|Ax_n - Ax_m\| \leq \|A\|\|x_n - x_m\| \to 0,$$

si n es lo suficientemente grande ($(x_n)_n$ es convergente, luego es de Cauchy). Como $\mathbb{Y}$ es completo, $Ax_n \to y \in \mathbb{Y}$. Pero A es cerrado, luego por la Proposición 7.4.3 tenemos que $x \in \mathcal{D}(A)$ y que $Ax = y$. Luego para todo $x \in \overline{\mathcal{D}(A)}$ tenemos que $x \in \mathcal{D}(A)$, i.e., $\mathcal{D}(A) = \overline{\mathcal{D}(A)}$, luego $\mathcal{D}(A)$ es cerrado. ■

Problema 7.6 Prueba que si un operador lineal cerrado A tiene inverso, entonces A^{-1} es cerrado. **Ayuda:** Usa que la aplicación $(x, y) \mapsto (y, x)$ es una isometría.

Solución: Sea $A : \mathcal{D}(A) \mapsto \mathbb{Y}$ cerrado. Si existe A^{-1}, entonces su grafo es

$$\mathcal{G}(A^{-1}) := \{(y, A^{-1}y) \,|\, y \in \mathcal{I}(A)\} = \{(Ax, x) \,|\, x \in \mathcal{D}(A)\} \subset \mathbb{Y} \times \mathbb{X},$$

pues si $y = Ax$, entonces $x = A^{-1}y$. Sea ahora la aplicación $T : \mathbb{X} \times \mathbb{Y} \mapsto \mathbb{Y} \times \mathbb{X}$, $T(x, y) = (y, x)$. Esta aplicación es lineal y biyectiva. Además preserva la norma

$$\|T(x, y)\| = \|(y, x)\| = \|y\| + \|x\| = \|x\| + \|y\| = \|(x, y)\|.$$

Es decir, $\mathcal{G}(A^{-1})$ es isomorfo a $\mathcal{G}(A)$, pero $\mathcal{G}(A)$ es cerrado (pues A es cerrado), luego $\mathcal{G}(A^{-1})$ es cerrado y, por tanto, A^{-1} es cerrado. ■

Problema 7.7 Usando el Teorema del grafo cerrado 7.4.5 prueba el Teorema de la inversa acotada de Banach 7.3.4. Nótese que, dado que el Teorema del grafo cerrado 7.4.5 lo probamos usando del Teorema de la inversa acotada de Banach 7.3.4, ello implica que ambos resultados son equivalentes.

Solución: Sea $A \in \mathcal{B}(\mathbb{X}, \mathbb{Y})$, como $\mathcal{D}(A) = \mathbb{X}$, $\mathcal{D}(A)$ es cerrado. Entonces, por el apartado 1 del Problema 7.5, A es cerrado. Pero como A es biyectivo, entonces existe su inverso A^{-1}, luego por el Problema 7.6, A^{-1} es cerrado. Pero $\mathcal{D}(A^{-1}) = \mathbb{Y}$, luego es cerrado, y podemos aplicar el Teorema del grafo cerrado 7.4.5 que nos asegura que A^{-1} es acotado. ■

Problema 7.8 Sea $\mathbb{H}$ un espacio de Hilbert de dimensión infinita y sea $(e_n)_n \in \mathbb{H}$ una sucesión ortonormal. Entonces $e_n \rightharpoonup 0$

Solución: Tenemos que probar que $f(e_n) \to 0$ para todo $f \in \mathbb{H}'$. Pero por el Teorema 7.5.3 ello es equivalente a probar que $\langle e_n, z\rangle \overset{n\to\infty}{\longrightarrow} 0$ para todo $z \in \mathbb{H}$. Pero $\langle e_n, z\rangle = \overline{\langle z, e_n\rangle} \overset{n\to\infty}{\longrightarrow} 0$ por (5.2.4), de donde se sigue el resultado. ■

Problema 7.9 Sea $A \in \mathcal{B}(\mathbb{X},\mathbb{Y})$, $\mathbb{X},\mathbb{Y}$ espacios de Banach. Prueba que si $x_n \rightharpoonup x$ en $\mathbb{X}$, entonces $Ax_n \rightharpoonup Ax$ en $\mathbb{Y}$.

Solución: Tenemos que probar que $f(Ax_n) \to f(Ax)$ para todo $f \in \mathbb{Y}'$. Definamos, para cada $f \in \mathbb{Y}'$, el funcional $g : \mathbb{X} \mapsto \mathbb{C}$, $g = f \circ A$, $g(x) = f(Ax)$. Como f y A son lineales, g es lineal. Además, $|g(x)| = f(Ax) \le \|f\|\|Ax\| \le \|f\|\|A\|\|x\|$, luego g es acotado, i.e., $g \in \mathbb{X}'$. Entonces, como $x_n \rightharpoonup x$, ello implica que $g(x_n) \to g(x)$, luego, para todo $f \in \mathbb{X}'$, $f(Ax_n) \to f(Ax)$, de donde se sigue el resultado. ■

Problema 7.10 Sea $\mathbb{H}$ un espacio de Hilbert. Prueba que si $x_n \rightharpoonup x$ y $\|x_n\| \to \|x\|$, entonces $x_n \to x$ (en norma).

Solución: Como $x_n \rightharpoonup x$, el Teorema 7.5.3 implica que $\langle x_n, z\rangle \to \langle x, z\rangle$ para todo $z \in \mathbb{H}$, en particular para $z = x$. Entonces

$$\|x_n - x\|^2 = \langle x_n - x, x_n - x\rangle = \|x_n\|^2 + \|x\|^2 - \langle x_n, x\rangle - \langle x, x_n\rangle \to 0.$$

Luego $x_n \to x$. ■

Problema 7.11 Sea $\mathbb{H}$ un espacio de Hilbert y sea $(e_k)_k$ una base ortonormal del mismo. Sea $(x_n)_n \in \mathbb{H}$ una sucesión acotada tal que existe el $\lim_{n\to\infty}\langle x_n, e_k\rangle$ para cada $k \in \mathbb{N}$. Prueba que entonces $(x_n)_n$ es débilmente convergente, i.e., existe un $x \in \mathbb{H}$ tal que $x_n \rightharpoonup x$.

Solución: Por el Teorema 7.5.3 basta probar que existe un $x \in \mathbb{H}$ tal que para todo $z \in \mathbb{H}$ $\langle x_n, z\rangle \overset{n\to\infty}{\longrightarrow} \langle x, z\rangle$. Comenzaremos probando que la sucesión $\langle x_n, z\rangle$ es de Cauchy, i.e., que $|\langle x_n, z\rangle - \langle x_m, z\rangle|$ se puede hacer tan pequeño como se quiera si n y m son lo suficientemente grandes. En primer lugar notemos que

$$|\langle x_n, z\rangle - \langle x_m, z\rangle| \le |\langle x_n, z\rangle - \langle x_n, s\rangle| + |\langle x_n, s\rangle - \langle x_m, s\rangle| + |\langle x_m, s\rangle - \langle x_m, z\rangle|.$$

Veamos que el primer y el tercer sumandos si pueden hacer tan pequeños como se quiera si n y m son lo suficientemente grandes. Para ello usaremos, por un lado, que $(x_n)_n$ es acotada, luego existe un $M > 0$ tal que para todo $n \in \mathbb{N}$ $\|x_n\| < M$ y, por el otro, que como $(e_k)_k$ es una base ortonormal de $\mathbb{H}$, entonces para todo $z \in \mathbb{H}$ existe una combinación lineal finita $s = \sum_{k=1}^{l} c_k e_k$ tal que la diferencia $\|z - s\|$ se puede hacer tan pequeña como se quiera. Usando la desigualdad del Cauchy-Schwarz (5.1.2) se tiene que

$$|\langle x_n, z\rangle - \langle x_n, s\rangle| = |\langle x_n, z - s\rangle| \leq \|x_n\|\|z - s\| \leq M\|z - s\|,$$

por lo que ambos sumandos se pueden hacer tan pequeños como se quiera.

Para acotar el segundo sumando escribimos

$$|\langle x_n, s\rangle - \langle x_m, s\rangle| = |\langle x_n - x_m, s\rangle| \leq \sum_{k=1}^{l} |c_k||\langle x_n - x_m, e_k\rangle|.$$

Vamos a probar que cada sumando de la suma finita anterior se puede puede hacer tan pequeño como se quiera y por tanto la propia suma se puede hacer tan pequeña como se quiera. Para ello notamos que, como por hipótesis existe el $\lim_{n\to\infty}\langle x_n, e_k\rangle$ para cada $k \in \mathbb{N}$, entonces la sucesión $\langle x_n, e_k\rangle$ es de Cauchy y, por tanto, para cada $k \in \mathbb{N}$, la cantidad

$$|\langle x_n - x_m, e_k\rangle| = |\langle x_n, e_k\rangle - \langle x_m, e_k\rangle|,$$

se puede puede hacer tan pequeña como se quiera.

Así pues, para todo $z \in \mathbb{H}$, la sucesión $\langle x_n, z\rangle$ es de Cauchy por tanto, para cada z, existe el límite $\langle x_n, z\rangle$. Definamos $x \in \mathbb{H}$ tal que $\langle x_n, z\rangle \overset{n\to\infty}{\longrightarrow} \langle x, z\rangle$ para cada $z \in \mathbb{H}$. Está claro que dicho x es único[9], por tanto para todo $z \in \mathbb{H}$ existe un $x \in \mathbb{H}$ tal que $\langle x_n, z\rangle \overset{n\to\infty}{\longrightarrow} \langle x, z\rangle$, de donde se sigue que $x_n \rightharpoonup x$.

Nota: Nótese que para resolver el Problema 7.11 ha sido esencial que la sucesión $(x_n)_n$ sea acotada. Esta condición es, en efecto, imprescindible pues si tomamos, por ejemplo, $x_n = ne_n$, entonces claramente para cada $k \in \mathbb{N}$, $\lim_{n\to\infty}\langle x_n, e_k\rangle = 0$, pero x_n no es acotada y por tanto no puede ser convergente (ver el punto 2 de la Proposición 7.5.2). ■

Problema 7.12 Probar el Teorema de Hahn-Banach 7.2.5 para un espacio vectorial complejo suponiendo que se cumple la versión real del mismo (ver el teorema 7.2.7).

[9] Supongamos que $\langle x_n, z\rangle \overset{n\to\infty}{\longrightarrow} \langle x, z\rangle$ y $\langle x_n, z\rangle \overset{n\to\infty}{\longrightarrow} \langle y, z\rangle$, pero $x \neq y$. Entonces, $|\langle x, z\rangle - \langle y, z\rangle| \leq |\langle x, z\rangle - \langle x_n, z\rangle| + |\langle x_n, z\rangle - \langle y, z\rangle| \to 0$, luego $\langle x, z\rangle = \langle y, z\rangle$ para todo z y el punto 4 del Ejercicio 5.1.2 implica $x = y$, lo que es una contradicción.

Solución: Sea $M \subset \mathbb{V}$, $M \neq \{0\}$, un subespacio vectorial y sea $f : M \mapsto \mathbb{C}$ nuestro funcional lineal tal que

$$|f(x)| \leq p(x), \qquad \forall x \in M, \tag{7.6.1}$$

donde p es un funcional sublineal y convexo (ver Definición 7.2.3). La idea es usar Teorema de Hahn-Banach 7.2.7 para un espacio vectorial real, así que nos restringiremos por el momento al caso cuando los vectores de $\mathbb{V}$, y por tanto los de M, los multiplicamos por escalares reales. Definamos sobre M el funcional $u : M \mapsto \mathbb{R}$, $u(x) = \Re(f(x))$, que al ser f lineal y habernos restringido al caso cuando multiplicamos por números reales, es un funcional lineal. Además, $u(x) \leq |f(x)| \leq p(x)$ para todo $x \in M$. Luego, por el Teorema de Hahn-Banach 7.2.7 sabemos que existe un funcional lineal $U : \mathbb{V} \mapsto \mathbb{R}$ tal que $U(x) = u(x)$ en M y $U(x) \leq p(x)$ en $\mathbb{V}$

Ahora bien, como para todo $z \in \mathbb{C}$ se tiene que $z = \Re(z) - i\Re(iz)$, entonces, para todo $x \in M$, podemos escribir nuestro funcional $f : M \mapsto \mathbb{C}$ como

$$f(x) = u(x) - iu(ix).$$

Probemos que el funcional $F : \mathbb{V} \mapsto \mathbb{C}$ definido por $F(x) = U(x) - iU(ix)$ es el funcional buscado. Para ello notemos que, como U es lineal respecto a la multiplicación por escalares reales, F también lo es. Además

$$\begin{aligned} F(ix) &= U(ix) - iU(i(ix)) = U(ix) - iU(-x) = iU(x) + U(ix) \\ &= i[U(x) - iU(ix)] = iF(x), \end{aligned}$$

de donde se sigue que también es lineal respecto a la multiplicación por escalares complejos. Está claro que F restringido a M coincide con f.

Probemos ahora que $F(x) \leq p(x)$ en $\mathbb{V}$. En primer lugar notemos que, como en este caso, $p(x) \geq 0$, entonces si $F(x) = 0$, $F(x) \leq p(x)$. Asumamos que $F(x) \neq 0$. Entonces $F(x) = |F(x)|e^{i\phi}$, luego

$$\begin{aligned} |F(x)| &= F(x)e^{-i\phi} = F(e^{-i\phi}x) = \Re F(e^{-i\phi}x) = U(e^{-i\phi}x) \\ &\leq p(e^{-i\phi}x) = |(e^{-i\phi})|p(x) = p(x). \end{aligned}$$

Así, hemos probado que existe un funcional lineal $F : \mathbb{V} \mapsto \mathbb{C}$ tal que

$$\forall x \in M, \quad F(x) = f(x) \quad \text{y} \quad \forall x \in \mathbb{V}, \quad |F(x)| \leq p(x). \tag{7.6.2}$$

lo que demuestra el Teorema de Hahn-Banach para el caso complejo. ∎

Capítulo 8

El análisis funcional y la mecánica cuántica

> La teoría de los operadores lineales es el aparato matemático de la mecánica cuántica.
>
> Vladimir A. Fock
>
> En *"Fundamentals of Quantum Mechanics"* (1976)

En este capítulo vamos a discutir brevemente las matemáticas de la mecánica cuántica. Comenzaremos con una breve introducción histórica.

8.1. Un poco de historia

El 27 de abril de 1900 Thomson (Lord Kelvin) pronunció en la *Royal Institution* una conferencia titulada[1] *Nineteenth-Century Clouds over the Dynamical Theory of Heat and Light*, donde afirmaba que

> *«La belleza y claridad de la teoría dinámica, que afirma que el calor y la luz son modos de movimiento, en la actualidad están empañadas por dos nubes. I. La primera surgió con la teoría ondulatoria de la luz, [...] involucraba la pregunta, ¿cómo podría la tierra moverse a través de un sólido elástico, tal como es esencialmente el éter lumínico? II. La segunda es el Teorema de Maxwell-Boltzmann sobre la partición de la energía.»*

[1]Un año más tarde apareció publicada con el mismo título en la *Philosophical magazine and Journal of Science*, Six Series, Vol. 2. July 1901, págs. 1–40.

Nadie en aquel momento podía imaginar que esas dos *nubecillas* iban a cambiar radicalmente nuestra concepción de los fenómenos naturales y del mundo que nos rodea y que representarían una revolución comparable a la Revolución copernicana que culminó con la aparición en 1687 de los *Philosophiæ naturalis Principia mathematica* de Newton, ya que la primera conduciría a la *Teoría de la relatividad* de Einstein, y la segunda a la *mecánica cuántica*.

El Teorema de la equipartición de la energía que mencionaba Lord Kelvin establece que la energía calorífica de un conjunto de partículas se divide uniformemente entre todas las posibles componentes de los movimientos de cada una de ellas, o como suelen decir los físicos, en cada uno de los grados de libertad del sistema. Este teorema fue enunciado por Ludwig Boltzmann en 1868 y demostrado en condiciones mucho más generales, por James C. Maxwell en 1879. Desde la publicación del trabajo de Boltzmann, y en especial del de Maxwell, hubo mucha controversia sobre la veracidad del mismo. Una de ellas está relacionada con el modelo del cuerpo negro.

El problema sobre la radiación del cuerpo negro tiene su origen en la pregunta ¿cómo emiten la luz los cuerpos al calentarse? Para resolver esta cuestión, los físicos de finales del siglo XIX idealizaron un cuerpo cualquiera y construyeron un modelo que llamaron *cuerpo negro* –también, "*cavidad*", etc–. La idealización consistía en que el cuerpo negro tenía que absorber –y por tanto emitir– ondas electromagnéticas[2] en todo el espectro de frecuencias.

Uno de los primeros en estudiar el problema de la emisión y absorción de energía fue el físico R. Kirchhoff quien en 1860 introdujo el concepto de cuerpo negro. Describamos brevemente el modelo.

Supongamos que tenemos un cuerpo negro ideal a una temperatura T y sea $U(T)$ la cantidad de energía por unidad de volumen del mismo. Dado que el cuerpo emitirá radiación electromagnética en todas las longitudes (frecuencias) de onda posibles asumiremos que cada longitud de onda –frecuencia– aportará su "granito de arena". Si denotamos la densidad de energía por unidad de volumen y unidad de frecuencia ω por $u(\omega, T)$, entonces

$$U(T) = \int_0^\infty u(\omega, T) d\omega. \tag{8.1.1}$$

[2]Era un hecho bien conocido a finales del XIX que la luz *visible* estaba constituida por ondas electromagnéticas dentro de un cierto rango de frecuencias.

Está claro la fórmula anterior permite encontrar la densidad de energía $U(T)$ conocida $u(\omega, T)$, así que la cuestión es ¿quién es la función $u(\omega, T)$?

En 1896 W. Wein y F. Paschen encontraron (de forma teórica el primero, y experimental el segundo) que cuando las frecuencias ω eran muy altas $u(\omega, T) = \alpha\omega^3 e^{-\beta\omega/T}$. Por otro lado, Lord Rayleigh, en un artículo aparecido en junio de 1900, estableció, usando el principio de equipartición de la energía, que la densidad de energía $u(\omega, T)$ en función de la frecuencia ω era proporcional a ω^2 lo que implicaba, según la fórmula (8.1.1), que la energía $U(T)$ irradiada por un cuerpo negro debía ser infinita, algo que obviamente contradecía los resultados experimentales establecidos. Precisamente a este tipo de resultados era a los que se refería Lord Kelvin con su segunda nubecilla.

Sin embargo, quien encontró la fórmula correcta para $u(\omega, T)$ fue Max Planck, quien en 1889 había reemplazado a Kirchhoff en la Universidad de Berlín tras su jubilación. En 1894 Planck se interesa por el problema de la radiación del cuerpo negro y en octubre de 1900 encuentra la fórmula

$$u(\omega, T) = \frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{\exp\left(\dfrac{\hbar\omega}{kT}\right) - 1}, \tag{8.1.2}$$

donde $\hbar$ era cierta constante desconocida y que describía perfectamente los resultados experimentales conocidos hasta la fecha. Nótese que la fórmula (8.1.2) se convierte en la de Wein-Paschen para frecuencias altas y en la de Rayleigh para frecuencias bajas, respectivamente.

Para dar un significado o explicación física a su fórmula Planck postuló que *los cuerpos absorben y emiten energía mediante "quantas" de energía* $E = \hbar\omega$. Eso le permitió a Planck deducir de "*primeros principios*" la expresión matemática (8.1.2) de la energía irradiada por el cuerpo negro. Estos resultados fueron publicados por Planck el 14 de octubre de 1900, día oficial del nacimiento de la Física cuántica.

El segundo paso en esta historia lo dio Albert Einstein. En 1905, Einstein explica el efecto fotoeléctrico suponiendo que la propia luz se emite en quantos (los hoy llamados fotones) de energía $\hbar\omega$, siendo $\hbar$ la constante de Planck. De esta forma Einstein recuperaba la naturaleza corpuscular de la luz predicha por Newton en su *Optiks* de 1704. Diez años más tarde, en 1915, el físico estadounidense Robert Andrews Millikan confirma experimentalmente que la fórmula de Einstein es correcta y comprueba que, en efecto, $\hbar$ es la misma $\hbar$ de Planck. Años más tarde, en 1923, Arthur H. Compton publicó un artículo donde describió los resultados de su investigación sobre la difusión de rayos X (ondas electromagnéticas como la luz, pero de muy alta frecuencia), lo que hoy día conocemos como efecto Compton. Compton descubrió, experimentalmente, que irradiando con rayos X distintos materiales se observaba un cambio de la longitud de onda (o frecuencia) de los rayos X en función del ángulo de dispersión, algo inexplicable usando la teoría ondulatoria de la luz, pero que se podía explicar si se asumía que la luz estaba constituida por partículas con masa cero pero con energía $E = \hbar\omega$ y momento lineal $p = \hbar\omega/c$, respectivamente, que interaccionaban de forma elástica (es decir que se conservaba la energía y el momento).

El siguiente paso en la historia se debe al danés Niels Bohr. Este, intentando describir el espectro del átomo de hidrógeno, postula que el electrón, para que no caiga al núcleo como predice la física clásica, solo puede estar en ciertas órbitas permitidas (estables). Para ello impone una condición *ad-hoc* que no puede justificarse a partir de la mecánica clásica ni de la electrodinámica. La suposición de Bohr era que, de todas las infinitas órbitas posibles, sólo son posibles aquellas para las cuales el momento angular del electrón $L = m_e v r_n$, siendo m_e la masa del electrón, v, su velocidad y r_n el radio de la n-ésima órbita estable, fuesen múltiplos enteros de $\hbar$, i.e. $L = n\hbar$, donde $\hbar$ era, otra vez, la constante de Planck. Unos cálculos sencillos muestran que, a partir de esta sencilla hipótesis, se puede describir la serie de Balmer para el espectro del átomo de hidrógeno:

$$\nu = \frac{E_m - E_n}{h} = \frac{m_e k_e^2 e^4}{4\pi\hbar^3}\left(\frac{1}{n^2} - \frac{1}{m^2}\right), \qquad n < m = 1, 2, 3, \dots . \tag{8.1.3}$$

Entre las preguntas que se hacían muchos de los físicos de la época se encontraban las siguientes: ¿Cómo es posible que una misma constante aparezca en aparentemente tres fenómenos no relacionados entre sí? ¿Cómo justificar el postulado de Bohr sobre las órbitas del átomo de hidrógeno?

La respuesta a la segunda pregunta la dio inesperadamente el físico francés Luis de Broglie. Este, influenciado por el trabajo de Einstein sobre el efecto fotoeléctrico, postula que la *dualidad* onda-partícula que Einstein había proclamado para la luz también tiene lugar para todas las partículas materiales, como por ejemplo, el electrón. La *genialidad* de de Broglie fue equiparar un electrón a una onda plana. Por ejemplo, si tenemos un electrón de masa m y velocidad v, de Broglie postuló que el momento (impulso) p del electrón era

$$p = mv = \hbar\frac{2\pi}{\lambda} = \frac{\hbar 2\pi}{vT} = \frac{\hbar\omega}{v}.$$

De esta forma De Broglie daba un significado físico a las órbitas de Bohr: éstas eran justo aquellas órbitas tales que el cociente entre su longitud y la longitud de onda del electrón era un número entero, es decir era una analogía completa de las ondas estacionarias sobre un anillo (círculo). Curiosamente los físicos C.J. Davisson y C.H. Kunsman, habían publicado en 1923 los resultados de un experimento (que había pasado desapercibido) que consistía en bombardear una placa de platino con electrones y habían descubierto que estos se difundían a través de la misma de forma que mostraban una figura de difracción. Esto fue confirmado en 1927 por el propio Davisson y su nuevo colaborador L. Germe, e independiente por los físicos G.P. Thompson y A. Reid. Así pues, los electrones también se comportaban como ondas.

La aparición de la primera versión de la mecánica cuántica no tardó mucho en aparecer a manos de Werner Heisenberg, un joven y en ese momento desconocido físico. En la primavera de 1925 Heisenberg, intentando entender las ecuaciones que expresan la posición y la velocidad de un electrón tuvo una idea reveladora que lo cambiaría todo: no era importante cómo se movían los electrones en el átomo (estructura atómica), sino lo que se podía observar de ellos, que en este caso, eran las frecuencias e intensidades de las líneas espectrales. Así, en junio de 1925, dio con la solución al problema solo que las matemáticas necesarias eran algo *raras*. Heisenberg descubrió que las frecuencias e intensidades dejaban de ser simples números, para convertirse en *tablas* de números y que la multiplicación de dichas tablas no era, en general, conmutativa. Los resultados de Heisenberg fueron recibidos en la revista *Zeitschrift für Physik* el 29 de julio de 1925 y publicados poco después. Fue Max Born, que era el supervisor académico de Heisenberg, quien se dio cuenta de que las tablas de Heisenberg no eran más que matrices y junto a Pacual Jordan, escribió el segundo artículo sobre la mecánica cuántica donde explicaba las raras matemáticas de Heisenberg. Born había sido asistente de David Hilbert y, a diferencia de Heisenberg, si que conocía la teoría de matrices. El trabajo de Born y Jordan titulado *Sobre la mecánica cuántica* fue recibido en la redacción de la *Zeitschrift für Physik* el 27 de septiembre de 1925: había nacido lo que se llamaría la *mecánica matricial*. Algo más tarde, el 16 de noviembre de 1925 Heisenberg, junto a Max Born y Pascual Jordan, envían a publicar la teoría completa de la mecánica matricial a la revista *Zeitschrift für Physik* en un artículo titulado *Sobre la mecánica cuántica II* que suele conocerse como el *Dreimännerarbeit* (Trabajo de los tres hombres). En él se desarrolla complemente la *mecánica matricial*, cuyas predicciones se ajustaban muy bien a las observaciones –pocas, por cierto– de la época.

No obstante, la acogida la mecánica matricial fue bastante fría. Se formaron dos grupos: los que la aceptaban encabezados por Born, Bohr, Heisenberg y Pauli, y los que la rechazaban, entre los que se encontraba el propio Einstein. Todo empezó a cambiar cuando en la navidad de 1925 Erwin Schrödinger, que en ese momento era profesor en Zurich, tras estudiar los trabajos de de Broglie encuentra una ecuación de ondas

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\Psi + V(x)\Psi = E\Psi, \tag{8.1.4}$$

que según muchos físicos, Einstein entre ellos, aclararía la situación. En efecto, si $V = 0$ (partícula libre), la solución de la ecuación (8.1.4) era la onda plana de de Broglie, y si V era el potencial del átomo de hidrógeno, se recuperaba la fórmula de Balmer (8.1.3). El artículo fue recibido en *Annalen der Physik* el 27 de enero de 1926. Ese día nació la *mecánica ondulatoria*.

Aunque la ecuación de ondas de Schrödinger (8.1.4) también explicaba muy bien las observaciones experimentales había algo que no quedaba claro y era el significado físico de la *onda* Ψ, solución de (8.1.4). Schrödinger, por ejemplo, opinaba las partículas no eran más que paquetes de ondas y, por tanto, Ψ podría describir a los electrones en sí mismos. Pero resultó que eso era un problema pues Max Born probó, a partir de la propia ecuación de Schrödinger, que si las partículas pudiesen representarse como paquetes de onda, entonces cuando colisionaban entre sí, dichos paquetes se difuminaban lo cual no ocurre en la realidad. Pero Born fue mucho más lejos. Tras estudiar cuidadosamente los datos experimentales de las colisiones de electrones, Born aseguró que la función de onda $\Psi(x)$ daba la probabilidad de que una partícula fuese detectada en la posición x y que dicha probabilidad era proporcional a $|\Psi(x)|^2$, es decir la mecánica ondulatoria, al igual que la matricial como se vería más tarde[3], era

[3]La alegría de los físicos tras el descubrimiento de la ecuación de Schrödinger (8.1.4) duró poco pues el mismo Schrödinger por un lado, y Paul Dirac por el otro, probaron que ambas formulaciones eran matemáticamente equivalentes.

una teoría estadística incluso para describir una única partícula. Esta idea era tan revolucionaria que ni al propio Heisenberg le gustó. El artículo de Born fue recibido en la redacción de la revista *Zeitschrift für Physik* el 25 de junio de 1926. Ese día murió el determinismo en la Física. En palabras del propio Born: *«Yo mismo me inclino a renunciar al determinismo en el mundo de los átomos. Pero esa es una cuestión filosófica para la que los argumentos físicos por sí solos no son decisivos.»*

Pero si ya eso era difícil de asimilar, el propio Heisenberg lo complicó aún más cuando, tras muchas discusiones con Borh en Copenhagen donde estaba de visitante, descubrió en febrero-marzo de 1927 el *Principio de incertidumbre* que lleva su nombre. En palabras del propio Heisenberg: *«El producto de las indeterminaciones para la localización y cantidad de movimiento (bajo el término 'cantidad de movimiento' se entiende el producto de la masa por la velocidad) no podía ser más pequeño que el 'quantum' de acción de Planck $\hbar$.»*

Lo que Heisenberg probó fue que si Δx era el error (o incertidumbre) que se comete al medir la posición del electrón (una partícula en general) y Δp era el error al medir la cantidad de movimiento $p = mv$, entonces

$$\Delta x \Delta p \geq \hbar/2 \qquad \text{donde } \hbar \text{ es la constante de Plank.}$$

Es decir, si nuestra medición de la posición es muy precisa perderemos precisión en la medición de la velocidad y viceversa. El artículo fue recibido en la redacción de *Zeitschrift für Physik* el 23 de marzo de 1927.

Se abría así una de las polémicas más grandes de la historia de la Ciencia de los últimos 100 años: La mecánica cuántica es, por principio, una teoría no determinista. Nuevamente los físicos de la época se dividieron en dos bandos. Por un lado estaba el bando de los que consideraban que la mecánica cuántica era una teoría completa y que la aceptaban incondicionalmente entre los cuales estaban Bohr, Born, Heisenberg, Jordan, Pauli, y los que no, por ejemplo, Einstein y Schrödinger. Esta división quedó plasmada en la famosa 5ª Conferencia Solvay de 1927, que es célebre por las discusiones entre Bohr y Einstein sobre la interpretación de la mecánica cuántica y donde se hizo famosa la frase de Einstein de *«Dios no juega a los dados»*. Cada mañana Einstein le proponía a Bohr un experimento mental que mostraba las contradicciones de la interpretación probabilística de la mecánica cuántica, y este se lo rebatía durante la cena. Lo mismo pasó en la siguiente conferencia en 1931. Fue allí donde a Einstein se le ocurrió un experimento mental que años después derivó en la famosa Paradoja EPR

(Einstein-Podolsky-Rosen) publicada[4] en 1935. La paradoja EPR tuvo en jaque a los físicos durante 30 años hasta que, en 1965, un físico irlandés, John Stewart Bell, ideó una forma de comprobar experimentalmente si Einstein tenía o no razón.

Lo que propuso Bell, desde el punto de vista teórico, es una desigualdad matemática que involucra magnitudes medibles, es decir, que se podría diseñar un experimento para comprobar si una teoría física era local-realista (que era lo que, según Einstein, tenía sentido construir, y que la mecánica cuántica no era), experimento que efectivamente se realizó y demostró que Einstein no tenía razón

Así pues, aunque la mecánica cuántica es una teoría *rara* y contraintuitiva, describe muy bien el mundo de lo infinitamente pequeño y son sus leyes y sus predicciones lo que hace que funcionen casi todos los aparatos electrónicos que usamos hoy día: ordenadores, móviles, televisores inteligentes, internet, etc. Y como toda teoría física tiene unas matemáticas muy precisas detrás.

De hecho, las matemáticas de la mecánica cuántica están estrechamente ligadas al problema de la interpretación. La razón principal, postulada por Bohr, se debe a que una misma teoría física no puede contener dos tipos de postulados, principios o axiomas: los clásicos (mundo macroscópico) y los cuánticos (mundo microscópico). Más aún, los principios de la física clásica deben obtenerse de los axiomas de la mecánica cuántica al pasar al mundo macroscópico donde la Física clásica es aplicable.

A ese respecto Hilbert, en una conferencia titulada *El pensamiento axiomático* pronunciada el año 1917 en Zürich ante la Sociedad Matemática de Suiza, comentó:

> *«...la teoría cuántica moderna, lo mismo que el conocimiento progresivo de la estructura del átomo, ha conducido a leyes que contradicen a la electrodinámica actual, basada esencialmente en las ecuaciones de Mawell, por lo que resulta evidente la necesidad de reformarla radicalmente y de darle nuevos fundamentos y organización.»*
>
> *«Vemos así que en las teorías físicas la supresión de las contradicciones debe lograrse por medio de una modificación de la elección*

[4] A. Einstein, B. Podolsky, N. Rosen, Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?. Physical Review 47, (1935), 777-780.

de los axiomas. La única dificultad que ello trae aparejada es la de elegir los axiomas de tal manera que todas las leyes físicas observadas resulten una consecuencia lógica de los mismos.»

Como veremos en este último capítulo, las matemáticas de la mecánica cuántica no son más que la teoría de operadores autoadjuntos en espacios de Hilbert. No obstante, siendo honestos los operadores que aparecen en la mayoría de los problemas cuánticos son operadores no acotados, y nosotros en estas notas hemos trabajado esencialmente con operadores acotados, o incluso compactos. A ese respecto vale la pena concluir esta introducción histórica con la siguiente frase de Max Born que aparece en un pie de página del tercer capítulo del *Trabajo de los tres hombres* de 1925, capítulo íntegramente escrito por Born:

«Hasta ahora, la teoría de las formas cuadráticas (o hermíticas) [matrices asociadas a los operadores] de infinitas variables se ha desarrollado principalmente para una clase especial (formas "acotadas"). Pero aquí nos interesan las formas no acotadas. Sin embargo, podemos suponer que, en general, las reglas se aplican de la misma manera.»

Para suerte de Born, John von Neumann, se encargó de desarrollar la teoría de los operadores no acotados y demostró que, en efecto, *las reglas se aplican de la misma manera*. Y no solo, también a von Neumann le debemos la primera formulación axiomática de la mecánica cuántica así como uno de los primeros textos sobre esta disciplina: *Mathematische Grundlagen der Quantenmechanik* (Fundamentos Matemáticos de la mecánica cuántica) publicado en 1932.

Para más detalles históricos sobre la mecánica cuántica se pueden consultar las monografías [15] y [20].

8.2. Preliminares

Sea $\mathbb{H}$ un espacio de Hilbert separable. En adelante asumiremos, a no ser que se diga lo contrario, que todos los operadores involucrados pertenecen a $\mathcal{L}(\mathbb{H})$. También, a no ser que se especifique lo contrario, asumiremos que los vectores no nulos Ψ de $\mathbb{H}$ están normalizados a la unidad, i.e., $\|\Psi\| = 1$.

Por comodidad, vamos a usar la notación de Dirac. Así, denotaremos el producto escalar $\langle \Psi, \Phi \rangle$ y los *elementos matriciales* $\langle \Psi, \mathcal{L}\Phi \rangle$ de un operador $\mathcal{L}$ por

$$\langle \Psi | \Phi \rangle := \langle \Psi, \Phi \rangle, \quad \langle \Psi | \mathcal{L} | \Phi \rangle := \langle \Psi | \mathcal{L}\Phi \rangle = \langle \Psi, \mathcal{L}\Phi \rangle, \tag{8.2.1}$$

respectivamente.

En la notación de Dirac se define el adjunto del operador $\mathcal{U}$ al operador $\mathcal{U}^*$ tal que para todos Φ y Ψ

$$\langle \Phi | \mathcal{U} | \Psi \rangle = \langle \Phi | \mathcal{U}\Psi \rangle = \overline{\langle \mathcal{U}\Psi | \Phi \rangle} = \overline{\langle \Psi | \mathcal{U}^*\Phi \rangle} = \overline{\langle \Psi | \mathcal{U}^* | \Phi \rangle}.$$

Recordemos además que un operador $\mathcal{U} : \mathbb{H} \mapsto \mathbb{H}$, $\mathbb{H}$ espacio de Hilbert, se denomina autoadjunto o hermítico si $\mathcal{U} = \mathcal{U}^*$ y unitario si $\mathcal{U}\mathcal{U}^* = \mathcal{U}^*\mathcal{U} = \mathcal{I}$, donde $\mathcal{I}$ el operador identidad $\mathcal{I} : \mathbb{H} \mapsto \mathbb{H}$, $\mathcal{I}x = x$.

Si $\mathcal{U}$ es autoadjunto, entonces, para todos Φ y Ψ

$$\langle \mathcal{U}\Psi | \Phi \rangle = \langle \Psi | \mathcal{U}\Phi \rangle \quad \Leftrightarrow \quad \langle \Phi | \mathcal{U} | \Psi \rangle = \overline{\langle \Psi | \mathcal{U} | \Phi \rangle}.$$

Definición 8.2.1 *Sea $\mathcal{U}$ un operador unitario. La transformación*

$$\Psi \mapsto \psi = \mathcal{U}^*\Psi, \quad \mathcal{L} \mapsto \ell = \mathcal{U}^*\mathcal{L}\,\mathcal{U}, \tag{8.2.2}$$

la denominaremos transformación unitaria de Ψ y $\mathcal{L}$.

Proposición 8.2.2 *Las transformaciones unitarias conservan*

1. *Las relaciones de conmutación de los operadores.*
2. *La propiedad de hermiticidad de un operador.*
3. *Los autovalores.*
4. *Los productos escalares y elementos matriciales de un operador.*

Demostración: 1. Sean $\mathcal{A}$, $\mathcal{B}$ y $\mathcal{L}$ tres operadores tales que $[\mathcal{A}, \mathcal{B}] = \mathcal{L}$ y a, b y ℓ sus correspondientes operadores tras realizar la transformación unitaria (8.2.2). Entonces, como $[\mathcal{A}, \mathcal{B}] = \mathcal{L}$

$$\mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A} = \mathcal{L} \;\Rightarrow\; \mathcal{U}^*\mathcal{A}\mathcal{B}\mathcal{U} - \mathcal{U}^*\mathcal{B}\mathcal{A}\mathcal{U} = \mathcal{U}^*\mathcal{L}\,\mathcal{U} = \ell,$$

pero

$$(\mathcal{U}^*\mathcal{A}\mathcal{U})(\mathcal{U}^*\mathcal{B}\mathcal{U}) - (\mathcal{U}^*\mathcal{B}\mathcal{U})(\mathcal{U}^*\mathcal{A}\mathcal{U}) = ab - ba \;\Rightarrow\; [a, b] = \ell.$$

2. Sea $\mathcal{L} = \mathcal{L}^*$. Sea $\ell = \mathcal{U}^*\mathcal{L}\,\mathcal{U}$, entonces

$$\ell^* = (\mathcal{U}^*\mathcal{L}\,\mathcal{U})^* = \mathcal{U}^*\mathcal{L}^*\,\mathcal{U} = \mathcal{U}^*\mathcal{L}\,\mathcal{U} = \ell.$$

3. Sea $\mathcal{L}\Psi = \lambda\Psi$, entonces

$$(\mathcal{U}^*\mathcal{L}\,\mathcal{U})(\mathcal{U}^*)\Psi = \lambda(\mathcal{U}^*\Psi) \;\Rightarrow\; \ell\psi = \lambda\psi.$$

4. $\langle\Psi_1|\mathcal{L}|\Psi_2\rangle = \langle\Psi_1|\,\mathcal{U}\mathcal{U}^*\mathcal{L}\,\mathcal{U}\mathcal{U}^*|\Psi_2\rangle = \langle\,\mathcal{U}^*\Psi_1|\,\mathcal{U}^*\mathcal{L}\,\mathcal{U}|\,\mathcal{U}^*\Psi_2\rangle = \langle\psi_1|\ell|\psi_2\rangle$. ■

Como hemos visto, si un operador $\mathcal{A} : \mathbb{H} \mapsto \mathbb{H}$ es autoadjunto y compacto entonces tiene asociado un conjunto numerable completo de autovectores (Teorema de Hilbert Schmidt 6.5.17). En adelante, por sencillez, vamos a restringirnos a aquellos operadores que tengan asociados un conjunto numerable de autovectores y que dicho conjunto sea un sistema completo.

Supongamos que $\mathcal{L}$ es uno de tales operadores y sea $(\Psi_n)_n$ su correspondiente conjunto completo de autovectores. Si todos los autovalores son simples entonces, como ya hemos visto, los correspondientes autovectores son ortogonales. En el caso de que tengamos autovalores múltiples sus correspondientes autovectores se pueden ortonormalizar usando el método de Gram-Schmidt que describimos antes (ver Teorema 5.1.13). Así pues asumiremos que $(\Psi_n)_n$ es un sistema ortonormal completo de $\mathbb{H}$.

Sea Φ un vector cualquiera de $\mathbb{H}$, entonces Φ se puede desarrollar en serie de Fourier respecto $(\Psi_n)_n$

$$\Phi = \sum_n c_n\Psi_n, \quad c_n = \langle\Phi|\Psi_n\rangle.$$

En otras palabras, $(\Psi_n)_n$ es una *base ortonormal completa* de $\mathbb{H}$.

Las bases juegan un papel fundamental. En particular, las bases asociadas a operadores autoadjuntos.

Sea $(\Psi_n)_n$ una base ortonormal completa de $\mathbb{H}$ y sea $\mathcal{A}$ un operador lineal, entonces

$$\mathcal{A}\Psi_n \in \mathbb{H} \Rightarrow \; \mathcal{A}\Psi_n = \sum_m A_{m,n}\Psi_m \Rightarrow \; A_{m,n} = \langle\Psi_m|\mathcal{A}|\Psi_n\rangle.$$

A la cantidad $\langle \Psi_m | \mathcal{A} | \Psi_n \rangle$ la denominaremos elemento matricial del operador $\mathcal{A}$ en la base $(\Psi_n)_n$.

Si $(\Psi_n)_n$ es la base asociada a cierto operador autoadjunto $\mathcal{L}$ se dice que la matriz $A = (A_{m,n})$ es la matriz del operador $\mathcal{A}$ en la $\mathcal{L}$-representación. Nótese que la matriz del operador $\mathcal{L}$ en su propia representación (la $\mathcal{L}$-representación) es una matriz diagonal con los autovalores en la diagonal.

Proposición 8.2.3 *El conmutador $[\mathcal{A}, \mathcal{B}]$ de dos operadores autoadjuntos $\mathcal{A}$ y $\mathcal{B}$ es tal que $[\mathcal{A}, \mathcal{B}] = i\mathcal{L}$, con $\mathcal{L}$ autoadjunto.*

Demostración: Supongamos que $[\mathcal{A}, \mathcal{B}] = \mathcal{N}$. Entonces

$$\mathcal{N}^* = ([\mathcal{A}, \mathcal{B}])^* = -[\mathcal{A}, \mathcal{B}] = -\mathcal{N} \;\Rightarrow\; \mathcal{N} = i\mathcal{L},$$

con $\mathcal{L}$ autoadjunto, pues $(i\mathcal{L})^* = -i\mathcal{L}^* = -i\mathcal{L}$. ■

Recordemos dos resultados que serán de gran interés más adelante (ver, para el caso de operadores autoadjuntos y compactos, el Teorema 6.5.18 y la discusión previa) y que asumiremos válidos a lo largo del capítulo.

Proposición 8.2.4 *Si dos operadores $\mathcal{L}$ y $\mathcal{N}$ tienen un sistema completo de autovectores $(\Psi_n)_n$ común, entonces $[\mathcal{L}, \mathcal{N}] = 0$.*

Proposición 8.2.5 *Si dos operadores $\mathcal{L}$ y $\mathcal{N}$ con sistemas completos de autovectores conmutan ($[\mathcal{L}, \mathcal{N}] = 0$), entonces tienen un sistema completo de autovectores $(\Psi_n)_n$ común.*

Definición 8.2.6 *Sea $F(z)$ una función analítica en un entorno de $z = 0$ y sea $F(z) = \sum_{n\geq 0} f_n z^n$ su desarrollo en serie de potencias con radio de convergencia $r > 0$. Definiremos al operador $F(\mathcal{A})$ mediante la serie*

$$F(\mathcal{A}) = \sum_{n\geq 0} f_n \mathcal{A}^n.$$

Mostremos que si[5] $\mathcal{A} \in \mathcal{B}(\mathbb{H})$ esta definición tiene sentido. Primero

[5] Recuérdese que $\mathcal{B}(\mathbb{H})$, el conjunto de todos los operadores lineales y acotados de $\mathbb{H}$ en $\mathbb{H}$, es un espacio de Banach, es decir, completo.

notemos que como $F(z)$ tiene radio de convergencia $r>0$, entonces

$$r=\frac{1}{\overline{\lim_n}\sqrt[n]{|f_n|}}.$$

Sean $F_k(\mathcal{A})=\sum_{n=0}^k f_n\mathcal{A}^n$, las sumas parciales de la serie de F.

$$\|F_k(\mathcal{A})x-F_m(\mathcal{A})x\|\leq\sum_{n=k+1}^m|f_n|\|\mathcal{A}^n x\|\leq\sum_{n=k+1}^m|f_n|\|\mathcal{A}\|^n\|x\| \;\Rightarrow$$

$$\frac{\|F_k(\mathcal{A})x-F_m(\mathcal{A})x\|}{\|x\|}\leq\sum_{n=k+1}^m|f_n|\|\mathcal{A}\|^n \;\Rightarrow$$

$$\|F_k(\mathcal{A})-F_m(\mathcal{A})\|\leq\sum_{n=k+1}^m|f_n|\|\mathcal{A}\|^n.$$

La serie numérica $\sum_{n\geq}|f_n|\|\mathcal{A}\|^n$ converge si $\|\mathcal{A}\|\leq r$, en cuyo caso la sucesión $F_k(\mathcal{A})$ es de Cauchy, por lo que la función $F(\mathcal{A})$ está bien definida. Además, como $F_k(\mathcal{A})$ es de Cauchy, es acotada. De hecho, la serie $F(\mathcal{A})$ es absolutamente convergente y define un operador acotado ($\mathcal{B}(\mathbb{H})$ es de Banach).

Definición 8.2.7 *Sea $\mathbb{H}$ un espacio de Hilbert y $\mathcal{A}\in\mathcal{B}(\mathbb{H})$. La derivada operacional $\partial F(\mathcal{A})/\partial\mathcal{A}$ es el operador que se obtiene mediante la formula*

$$\frac{\partial F(\mathcal{A})}{\partial\mathcal{A}}=\lim_{\varepsilon\to0}\frac{F(\mathcal{A}+\varepsilon\mathfrak{I})-F(\mathcal{A})}{\varepsilon}, \tag{8.2.3}$$

donde $\mathfrak{I}$ es el operador identidad.

Por ejemplo

$$\frac{\partial\mathcal{A}^n}{\partial\mathcal{A}}=n\mathcal{A}^{n-1}.$$

Es más, tomando $\varepsilon>0$ lo suficientemente pequeño de forma que $\|A\|+\varepsilon<r$, podemos probar, de forma muy similar a como se prueba el Teorema 1.2.32, que la serie

$$G(\mathcal{A})=\sum_{n\geq1}nf_n\mathcal{A}^{n-1},$$

es absolutamente convergente y que además,

$$\frac{\partial F(\mathcal{A})}{\partial\mathcal{A}}=G(\mathcal{A}).$$

Lema 8.2.8 *Sean $\mathcal{A}$ y $\mathcal{B}$ tales que $[\mathcal{A},\mathcal{B}] = \mathcal{I}$. Entonces*

$$[\mathcal{A},\mathcal{B}^k] = k\mathcal{B}^{k-1}, \qquad k \geq 1.$$

Demostración: Para $k = 1$ es obvio. Supongamos que es cierto para k y probémosle para $k+1$. Así, calculamos

$$[\mathcal{A},\mathcal{B}^{k+1}] = \mathcal{A}\mathcal{B}^{k+1} - \mathcal{B}^{k+1}\mathcal{A} = (\mathcal{A}\mathcal{B}^k)\mathcal{B} - \mathcal{B}^k(\mathcal{B}\mathcal{A}). \tag{8.2.4}$$

Si ahora usamos (8.2.8) (hipótesis de inducción) tenemos, por un lado, $\mathcal{A}\mathcal{B}^k = k\mathcal{B}^{k-1} + \mathcal{B}^k\mathcal{A}$ y para $k = 1$, por el otro que $\mathcal{B}\mathcal{A} = \mathcal{A}\mathcal{B} - \mathcal{I}$. Sustituyendo ambas expresiones en la fórmula (8.2.4) obtenemos en resultado. ■

Proposición 8.2.9 *Si $F(z)$ es una función analítica en un entorno de $z = 0$ y sean $\mathcal{A}$ y $\mathcal{B}$ tales que $[\mathcal{A},\mathcal{B}] = \mathcal{I}$. Entonces*

$$[\mathcal{A}, F(\mathcal{B})] = \frac{\partial F(\mathcal{B})}{\partial \mathcal{B}}.$$

Demostración: Basta escribir la serie de potencias de F y usar el Lema (8.2.8). ■

8.3. Los axiomas de la mecánica cuántica

En esta sección enunciaremos los principales axiomas o postulados de la mecánica cuántica no relativista sobre un espacio de Hilbert $\mathbb{H}$.

Postulado 8.3.1 *A cada sistema físico se le hace corresponder un espacio de Hilbert separable $\mathbb{H}$ apropiado. Además, para cada $t \in \mathbb{R}$ (parámetro correspondiente al tiempo) el estado queda completamente caracterizado (a excepción de un factor constante) por un vector Ψ normalizado a la unidad de $\mathbb{H}$.*

Es decir, para cada t el estado está determinado por un vector de $\mathbb{H}$ tal que $\|\Psi\| = 1$ ($\Psi = 0$ no representa a ningún estado físico real). El postulado además deja claro que si dos vectores no nulos son tales que $\Psi_1 = c\Psi_2$, $c \neq 0$, entonces ambos describen el mismo estado[6]. De aquí

[6] Para cada $t \in \mathbb{R}$ cualquier vector no nulo Ψ siempre se puede normalizar a la unidad.

también se sigue que, en general, dados los estados $\Psi_1, \ldots, \Psi_k$, la combinación lineal $\Phi = \sum_{k=1}^{n} \alpha_k \Psi_k$ también es un (posible) estado. Conviene aclarar que el tipo de estados a los que alude el postulado 8.3.1 son los denominados estados *puros* y son una idealización de los estados de un sistema cuántico.

Postulado 8.3.2 *A cada magnitud física medible (observable) L se le hace corresponder un operador*[a] *linear hermítico (autoadjunto) $\mathcal{L}$ que actúa en $\mathbb{H}$.*

[a]Los operadores cuánticos de las magnitudes físicas se postulan en la teoría. Tres operadores esenciales son el operador posición o coordenadas $\mathcal{X}$, impulso $\mathcal{P}$, y el hamiltoniano del sistema $\mathcal{H}$.

Postulado 8.3.3 *Sea Ψ el estado del sistema en el momento t justo antes de la medición de la magnitud (observable) L (asociada al operador $\mathcal{L}$). Independientemente de cuál sea el estado original Ψ, el resultado de la medición solo puede ser un autovalor de $\mathcal{L}$.*

Este postulado requiere una aclaración y es que, al hacer una medición de $\mathcal{L}$, el sistema cambia (las mediciones interfieren en el sistema). Así pues, antes de medir L el sistema puede estar en *cualquier* estado Ψ, pero al realizar la medición, ésta cambia al sistema y lo deja en el estado determinado por el vector Ψ_λ que pertenece al autoespacio de $\mathcal{L}$ correspondiente al autovalor λ. Este efecto se conoce como el *colapso de la función de onda*.

Postulado 8.3.4 *El valor esperado $\langle L \rangle$ de una magnitud física L cuando el sistema se encuentra en el estado Ψ viene dado por el elemento matricial*

$$\langle L \rangle = \langle \Psi | \mathcal{L} | \Psi \rangle := \langle \Psi | \mathcal{L} \Psi \rangle.$$

Nótese que, como $\mathcal{L}$ es autoadjunto, entonces[7]

$$\langle \Psi | \mathcal{L} | \Psi \rangle = \overline{\langle \Psi | \mathcal{L}^* | \Psi \rangle} = \overline{\langle \Psi | \mathcal{L} | \Psi \rangle} \;\Rightarrow\; \langle L \rangle \in \mathbb{R}.$$

[7]De hecho esta afirmación se sigue del Teorema 6.2.2.

Postulado 8.3.5 *Los elementos matriciales del los operadores $\mathcal{X}$ de la posición x y $\mathcal{P}$ del momento p, definidos por $\langle\Phi|\mathcal{X}|\Psi\rangle$ y $\langle\Phi|\mathcal{P}|\Psi\rangle$, cualquiera sean Φ y Ψ de $\mathbb{H}$ satisfacen las ecuaciones de evolución*

$$\frac{d}{dt}\langle\Phi|\mathcal{X}|\Psi\rangle = \left\langle \Phi \left| \frac{\partial \mathcal{H}}{\partial \mathcal{P}} \right| \Psi \right\rangle, \quad \frac{d}{dt}\langle\Phi|\mathcal{P}|\Psi\rangle = -\left\langle \Phi \left| \frac{\partial \mathcal{H}}{\partial \mathcal{X}} \right| \Psi \right\rangle, \tag{8.3.1}$$

donde $\mathcal{H}$ es el operador asociado a la función de Hamilton del correspondiente sistema clásico (si es que lo hay).

Este postulado tiene un significado físico *evidente* pues nos indica que el promedio de las magnitudes medibles posición, impulso y energía (hamiltoniano) satisfacen las ecuaciones dinámicas de la formulación hamiltoniana de la mecánica clásica, i.e, en el límite apropiado ($\hbar \to 0$) la mecánica cuántica se transforma en la clásica[8].

Proceden unas aclaraciones. En general el hamiltoniano H de un sistema clásico depende de la coordenada x y el impulso p, por lo que el operador $\mathcal{H}$ se obtiene cambiando las x por los correspondientes operadores $\mathcal{X}$ y p por $\mathcal{P}$. Esto, aunque en apariencia es *trivial*, en general no lo es pues $\mathcal{H}$ debe ser autoadjunto (ya que corresponde a la magnitud física energía). Por ejemplo, si el hamiltoniano clásico contiene el término $W = xp$, entonces, el operador $\mathcal{W} = \mathcal{X}\mathcal{P}$ no puede representar al correspondiente operador cuántico ya que no es hermítico (ver Proposición 6.2.4) pues $\mathcal{X}$ y $\mathcal{P}$ no conmutan. Una posible solución, en este caso, es definir $\mathcal{W} = \frac{1}{2}(\mathcal{X}\mathcal{P} + \mathcal{P}\mathcal{X})$.

Postulado 8.3.6 *Los operadores posición $\mathcal{X}$ e impulso $\mathcal{P}$, satisfacen las relaciones de conmutación*

$$[\mathcal{X}, \mathcal{P}] = i\hbar\mathcal{I}, \tag{8.3.2}$$

donde $\hbar$ es una constante (de Planck) e $i = \sqrt{-1}$.

Del postulado anterior se sigue que los operadores $\mathcal{X}$ y $\mathcal{P}$ no pueden tener un conjunto completo de autovectores (ver las Proposiciones 8.2.4–8.2.5) comunes (no conmutan). Este postulado es el análogo de las relaciones conocidas como *llaves de Poisson* de la mecánica hamiltoniana.

[8]Este hecho se conoce como el *Principio de correspondencia de Bohr*.

8.4. Discusión de los postulados

Supongamos el sistema físico se encuentra en el estado definido por Ψ_n, autovector correspondiente al autovalor λ_n de cierto operador $\mathcal{L}$ asociado a la magnitud física L. Entonces[9]

$$\langle \Psi_n | \mathcal{L} | \Psi_n \rangle = \lambda_n, \quad \langle \Psi_n | \mathcal{L}^k | \Psi_n \rangle = \lambda_n^k.$$

Supongamos ahora que el sistema se encuentra en el estado Φ que es en una superposición de los estados Ψ_k, $k = 1, 2, \ldots, N$, entonces como $\Phi = \sum_k c_k \Psi_k$ tenemos

$$\langle \Phi | \mathcal{L} | \Phi \rangle = \sum_k |c_k|^2 \lambda_k, \quad c_n = \langle \Phi | \Psi_n \rangle. \tag{8.4.1}$$

Lo anterior indica, en virtud de postulado 8.3.4, que la cantidad $|c_k|^2$ es la probabilidad con que se observa el valor λ_k al hacer una medición. Lo anterior implica además que tras la medición el sistema va a parar al estado definido por un vector del espacio generado por los autovectores correspondientes a λ_k. Así pues, en el caso de que el autovalor λ_k sea simple la probabilidad de que el sistema estando en un estado original definido por el vector Φ termine en el estado definido por Ψ_k es

$$\text{Prob}(\Phi \mapsto \Psi_k) = |c_k|^2 = |\langle \Phi | \Psi_n \rangle|^2. \tag{8.4.2}$$

Nótese que esta probabilidad es por tanto invariante ante transformaciones unitarias: $\Psi_k \mapsto \mathcal{U}\Psi_k = \tilde{\Psi}_k$, $\Phi \mapsto \mathcal{U}\Phi = \tilde{\Phi}$ pues $\text{Prob}(\Phi \mapsto \Psi_k) = \text{Prob}(\tilde{\Phi} \mapsto |\tilde{\Psi}_k\rangle)$. Luego el sistema físico es invariante frente a cualquier transformación unitaria.

3. Dada cualquier magnitud física clásica L le podemos adicionar $xp - px$ sin cambiarla. Si transformamos L en su operador $\mathcal{L}$ ya no le podemos adicionar el correspondiente operador $\mathcal{X}\mathcal{P} - \mathcal{P}\mathcal{X}$ pues éste no es nulo (ver postulado 8.3.6).

Tomando las derivadas funcionales

$$\frac{\partial}{\partial \mathcal{X}}(\mathcal{X}\mathcal{P} - \mathcal{P}\mathcal{X}) = \frac{\partial}{\partial \mathcal{P}}(\mathcal{X}\mathcal{P} - \mathcal{P}\mathcal{X}) = 0,$$

i.e., $\mathcal{X}\mathcal{P} - \mathcal{P}\mathcal{X}$ es proporcional al operador identidad $\mathcal{X}\mathcal{P} - \mathcal{P}\mathcal{X} = \alpha\mathcal{I}$. Si además $\mathcal{X}$ y $\mathcal{P}$ son autoadjuntos entonces, necesariamente, $\alpha = i\hbar$ (ver la Proposición 8.2.3) donde $\hbar \in \mathbb{R}$, es decir, *recuperamos* la relación de conmutación del postulado 8.3.6.

[9]Recuérdese que los estados están normalizados a la unidad.

8.4.1. Los proyectores ortogonales y la teoría de mediciones

Como hemos visto en el postulado 8.3.3 cuando en un sistema en cierto estado Ψ hacemos una medición para saber el valor de cierto observable L asociado un operador autoadjunto $\mathcal{L}$ el resultado es uno de los autovalores λ_k de dicho operador y el estado, después de la medición, pasa a ser un vector Ψ_k del autoespacio (recordemos que todos los autovectores están normalizados a la unidad) asociado al autovalor λ_k. Además en el caso de que λ_k sea simple sabemos que la probabilidad de que ello ocurra es $\mathrm{Prob}(\Phi \mapsto \Psi_k) = |c_k|^2 = |\langle \Psi | \Psi_k \rangle|^2$ (ver (8.4.2)). Lo anterior lo podemos escribir usando los *proyectores ortogonales*.

Para ello comenzaremos asumiendo el caso más simple. Imaginemos que tenemos la magnitud L y que el resultado de la medición es el autovalor λ_k que asumiremos simple. Tras la medición el sistema estará en el estado Ψ_k, donde Ψ_k es el autovector asociado a λ_k.

Definamos el operador $\mathcal{P}_k$ de la siguiente forma:

$$\mathcal{P}_k : \mathbb{H} \mapsto \mathbb{H}, \quad \mathcal{P}_k \Psi = \langle \Psi | \Psi_k \rangle \Psi_k.$$

En primer lugar nótese que, efectivamente, $\mathrm{Prob}(\Phi \mapsto \Psi_k) = \langle \Psi | \mathcal{P}_k \Psi \rangle = \|\mathcal{P}_k \Psi\|^2$. Además, de la Proposición 6.5.15 se deduce que $\mathcal{P}_k$ es el operador de proyección (o proyector ortogonal) sobre el subespacio generado por Ψ_k. En particular, $\mathcal{P}_k$ es autoadjunto, i.e., $\mathcal{P}_k^* = \mathcal{P}_k$ e idempotente, i.e., $\mathcal{P}_k^2 := \mathcal{P}_k \circ \mathcal{P}_k = \mathcal{P}_k$.

Probemos que sus autovalores son o bien cero o bien uno. En efecto,

$$\mathcal{P}_k \Psi = \lambda \Psi \;\Rightarrow\; \mathcal{P}_k^2 \Psi = \lambda^2 \Psi \;\Rightarrow\; (\mathcal{P}_k^2 - \mathcal{P}_k)\Psi = (\lambda^2 - \lambda)\Psi = 0,$$

luego, $\lambda^2 - \lambda = \lambda(\lambda - 1) = 0$, de donde se sigue el resultado.

Por otro lado, $\mathcal{P}_k \Psi$ es un autovector asociado al autovalor uno[10] mientras que $(\mathcal{I} - \mathcal{P}_k)\Psi$ es el autovector asociado al autovalor cero y, por tanto, cualquiera sea el vector $\Psi \in \mathbb{H}$, los vectores $\mathcal{P}_k \Psi$ y $(\mathcal{I} - \mathcal{P}_k)\Psi$ son ortogonales (¿por qué?).

¿Qué ocurre si el autovector λ_k es degenerado, i.e., tiene asociado un subespacio de dimensión $K > 1$? En ese caso el proyector $\mathcal{P}_k$, como ya vimos –ver (6.5.3)–, es la suma de los proyectores asociados a cada uno de

[10] $\mathcal{P}_k(\mathcal{P}_k \Psi) = \mathcal{P}_k^2 \Psi = \mathcal{P}_k \Psi = 1(\mathcal{P}_k \Psi)$.

los vectores de la base ortonormal $(\Psi_{k,j})_{j=1}^{K}$ del autoespacio $\mathbb{X}_k$ asociado a λ_k. Es decir,

$$\mathcal{P}_k\Psi = \sum_{j=1}^{K}\langle\Psi|\Psi_{k,j}\rangle\Psi_{k,j}.$$

Es *fácil* comprobar que en este caso se tiene las mismas propiedades que en el caso cuando $K = 1$. También es *fácil* comprobar que, en este caso, la probabilidad de obtener el valor λ_k vuelve a ser Prob$(\Phi \mapsto \Psi_k) = \langle\Psi|\mathcal{P}_k\Psi\rangle$, donde ahora $\mathcal{P}_k$ es el proyector ortogonal sobre todo el subespacio asociado a λ_k. Ambas propiedades se dejan como ejercicio al lector.

Supongamos que tenemos dos autovalores distintos λ_k y $\lambda_{k'}$. Entonces, los autoespacios correspondientes $\mathbb{X}_k$ y $\mathbb{X}_{k'}$ son ortogonales, luego

$$\mathcal{P}_k \circ \mathcal{P}_{k'} = 0, \qquad \forall k \neq k'.$$

Si además el operador $\mathcal{L}$ asociado a la magnitud L tiene un sistema de autovectores que constituyen una base de $\mathbb{H}$ entonces, para todo $\Psi \in \mathbb{H}$ se tiene

$$\Psi = \sum_k \langle\Psi|\Psi_k\rangle\Psi_k = \sum_k \mathcal{P}_k\Psi,$$

de donde deducimos que

$$\mathcal{P}_L := \sum_k \mathcal{P}_k = \mathcal{I}. \tag{8.4.3}$$

La igualdad anterior se suele denominar *descomposición de la identidad*.

Por otro lado, $\mathcal{L} \circ \mathcal{P}_k = \lambda_k\mathcal{P}_k$, de donde se sigue que $(\mathcal{L} - \lambda_k\mathcal{I}) \circ \mathcal{P}_k = 0$. Entonces, usando la descomposición de la identidad (8.4.3), y que $\mathcal{L}$ es acotado (luego continuo) tenemos que

$$\mathcal{L} = \mathcal{L} \circ \mathcal{I} = \mathcal{L} \circ \left(\sum_k \mathcal{P}_k\right) = \sum_k \mathcal{L} \circ \mathcal{P}_k = \sum_k \lambda_k\mathcal{P}_k.$$

Es decir, todo operador autoadjunto cuyo conjunto de autovectores es un sistema completo está completamente determinado por sus autovalores y autovectores. Esto es esencialmente el Teorema espectral 6.5.14 en espacios de Hilbert que ya vimos en la sección 6.5 para los operadores autoadjuntos (o normales) y compactos.

Así pues, el postulado 8.3.3 se puede reescribir en términos de los operadores de proyección de la siguiente forma:

Postulado 8.3.3. *Si tenemos un sistema en cierto estado Ψ y sobre él medimos el valor de cierta magnitud L obtendremos como resultado un autovalor λ_k del operador $\mathcal{L}$ asociado a la magnitud L siendo el estado final del sistema el definido por el vector $\mathcal{P}_k\Psi$, donde $\mathcal{P}_k$ es el proyector ortogonal al subespacio asociado al autovalor λ_k.*

Nótese además que de (8.4.1) se sigue que el resultado de la medición será el autovalor λ_k con probabilidad $\text{Prob}(\Phi \mapsto \Psi_k) = \|\mathcal{P}_k\Psi\|^2$.

8.4.2. Representación de los operadores $\mathcal{X}$ y $\mathcal{P}$

Escojamos como espacio de Hilbert de nuestro sistema el conjunto de las funciones de cuadrado integrable[11] $\mathbb{H} = L^2_\Omega$, $\Psi = \Psi(x)$. Definiremos el operador $\mathcal{X}$ en L^2_Ω como el operador multiplicación por x, i.e., $\mathcal{X} := x\mathcal{I}$. Luego $\mathcal{X}^k\Psi(x) = x^k\Psi(x)$. ¿Quién es $\mathcal{P}$?

Nuestro objetivo es encontrar un operador $\mathcal{P}$ tal que cumpla las relaciones de conmutación (8.3.2).

Del postulado 8.3.6 se sigue que $[\mathcal{P}, \mathcal{X}] = \mathcal{P}\mathcal{X} - \mathcal{X}\mathcal{P} = -i\hbar\mathcal{I}$. Luego, usando una obvia generalización del Lema 8.2.4 (ver Problema 8.2) obtenemos

$$[\mathcal{P}, \mathcal{X}^n] = \mathcal{P}\mathcal{X}^n - \mathcal{X}^n\mathcal{P} = -ni\hbar\mathcal{X}^{n-1} = -i\hbar\frac{\partial \mathcal{X}^n}{\partial \mathcal{X}}.$$

Por tanto, para cualquier función analítica $F(z)$ tenemos

$$[\mathcal{P}, F(\mathcal{X})] = -i\hbar\frac{\partial F(\mathcal{X})}{\partial \mathcal{X}}, \tag{8.4.4}$$

que, en nuestra representación, se puede reescribir como

$$\mathcal{P}F(\mathcal{X})\Psi(x) - F(\mathcal{X})\mathcal{P}\Psi(x) = -i\hbar\frac{\partial F(\mathcal{X})}{\partial \mathcal{X}}\Psi(x),$$

o, equivalentemente, usando que $\mathcal{X}^k\Psi(x) = x^k\Psi(x)$,

$$\mathcal{P}[F(x)\Psi(x)] - F(x)\mathcal{P}\Psi(x) = -i\hbar\frac{\partial F(x)}{\partial x}\Psi(x).$$

[11]El matemático avispado se dará cuenta que el espacio de las funciones de cuadrado contiene funciones bastante *raras*, no obstante, el espacio $\mathcal{D}(\Omega)$ de las funciones infinitamente derivables en Ω con soporte acotado en Ω es denso en L^2_Ω (ver [8, Nota 21.3.5, página 189]) por lo que nos restringiremos siempre a elementos de dicho espacio $\mathcal{D}(\Omega)$.

Escogiendo[12] $\Psi(x) = 1$, obtenemos, sustituyendo $\mathcal{P}1 = \phi(x)$, la expresión

$$\mathcal{P}F(x) = -i\hbar\frac{\partial F(x)}{\partial x} + F(x)\phi(x).$$

Vamos a suponer que $\phi = \frac{\partial \Phi}{\partial x}$, donde $\Phi(x)$ es una función lo suficientemente buena. Entonces el operador $\mathcal{P}$ debe tener la forma

$$\mathcal{P} = -i\hbar\frac{\partial}{\partial x} + \frac{\partial \Phi}{\partial x}\mathcal{I}.$$

Hagamos ahora la transformación unitaria

$$\mathcal{P} \to \mathcal{U}^*\mathcal{P}\mathcal{U}, \;\; \mathcal{U} = \exp(-i\Phi(x)/\hbar)\mathcal{I}, \;\; \mathcal{U}^* = \exp(i\Phi(x)/\hbar)\mathcal{I},$$

que, como sabemos, no cambia ni los elementos matriciales, ni las relaciones de conmutación, ni los autovalores, ni la propiedad de hermiticidad de los operadores (i.e., éstos mantendrán el mismo significado físico de antes). Entonces,

$$\mathcal{P} = e^{i\Phi/\hbar}\left(-i\hbar\frac{\partial}{\partial x} + \frac{\partial \Phi}{\partial x}\mathcal{I}\right)e^{-i\Phi/\hbar}\mathcal{I} = -i\hbar\frac{\partial}{\partial x},$$

que nos da la expresión del operador en L^2_Ω. Como ejercicio al lector dejamos que pruebe la identidad

$$[\mathcal{X}, F(\mathcal{P})] = i\hbar\frac{\partial F(\mathcal{P})}{\partial \mathcal{P}}. \tag{8.4.5}$$

8.4.3. Las ecuaciones de Heisenberg y de Schrödinger

En este apartado vamos a discutir las representaciones de Heisenberg y de Schrödinger para las ecuaciones dinámicas de la mecánica cuántica. Supongamos que tenemos el sistema en cierto estado Φ y sea ℓ el operador de cierta magnitud física que queremos estudiar. Si Φ es independiente del tiempo y ℓ no lo es diremos que estamos trabajando con la *representación de Heisenberg*. Si por el contrario, Φ depende de tiempo y ℓ no, entonces diremos que estamos considerando la *representación Schrödinger* de la mecánica cuántica.

[12]Por el momento solo nos interesa encontrar la expresión del operador independientemente de que luego éste vaya actuar sobre L^2_Ω o $\mathcal{D}(\Omega)$.

Por sencillez, en adelante asumiremos que el hamiltoniano del sistema se expresa mediante la fórmula

$$\mathcal{H} = \widehat{T} + \widehat{V}, \quad \widehat{T} = \frac{\mathcal{P}^2}{2m},$$

y $\widehat{V} = V(\mathcal{X}) = V(x)\mathcal{I}$, sólo depende de la coordenada x.

Como $[\mathcal{P}, \widehat{T}] = 0$, tenemos, usando (8.4.4) que

$$[\mathcal{P}, \mathcal{H}] = [\mathcal{P}, V(\mathcal{X})] = -i\hbar\frac{\partial V(\mathcal{X})}{\partial \mathcal{X}} = -i\hbar\frac{\partial \mathcal{H}}{\partial \mathcal{X}}. \tag{8.4.6}$$

Supongamos ahora que los vectores de estado no dependen del tiempo pero los operadores sí que pueden, en principio, depender del tiempo (es decir, consideremos la representación de Heisenberg). Entonces del postulado 8.3.5 se tiene que

$$\frac{d\mathcal{P}}{dt} = -\frac{\partial \mathcal{H}}{\partial \mathcal{X}},$$

de donde se sigue que

$$\frac{d\mathcal{P}}{dt} = -\frac{i}{\hbar}[\mathcal{P}, \mathcal{H}]. \tag{8.4.7}$$

De forma análoga, pero usando (8.4.5), se deduce la segunda ecuación de Heisenberg

$$\frac{d\mathcal{X}}{dt} = -\frac{i}{\hbar}[\mathcal{X}, \mathcal{H}]. \tag{8.4.8}$$

Las ecuaciones anteriores se conocen como *ecuaciones dinámicas* de la mecánica cuántica en la *representación de Heisenberg*: es decir, cuando las funciones de onda son vectores independientes del tiempo pero los operadores dependen del tiempo.

Obviamente hay otra posibilidad y es que los operadores no dependan del tiempo y las funciones de onda sí. En este caso usando el postulado 8.3.5 y la fórmula (8.4.6) ($\partial\mathcal{H}/\partial\mathcal{X} = i/\hbar[\mathcal{P}, \mathcal{H}]$), obtenemos

$$\frac{d}{dt}\langle\Phi|\mathcal{P}|\Psi\rangle = -\left\langle \Phi \left| \frac{\partial \mathcal{H}}{\partial \mathcal{X}} \right| \Psi \right\rangle = \frac{i}{\hbar}\langle\Phi|[\mathcal{P}, \mathcal{H}]|\Psi\rangle.$$

Luego, por un lado,

$$\left\langle \frac{\partial \Phi}{\partial t} \middle| \mathcal{P} \middle| \Psi \right\rangle + \left\langle \Phi \middle| \mathcal{P} \middle| \frac{\partial \Psi}{\partial t} \right\rangle = \left\langle \frac{\partial \Phi}{\partial t} \middle| \mathcal{P}\Psi \right\rangle + \left\langle \mathcal{P}\Phi \middle| \frac{\partial \Psi}{\partial t} \right\rangle$$

y, por otro,

$$\frac{i}{\hbar}\langle\Phi|[\mathcal{P},\mathcal{H}]|\Psi\rangle = \frac{i}{\hbar}\left(\langle\Phi|\mathcal{P}\mathcal{H}|\Psi\rangle - \langle\Phi|\mathcal{H}\mathcal{P}|\Psi\rangle\right) = \left\langle \mathcal{P}\Phi \middle| \frac{i}{\hbar}\mathcal{H}\Psi \right\rangle + \left\langle \frac{i}{\hbar}\mathcal{H}\Phi \middle| \mathcal{P}\Psi \right\rangle,$$

de donde se sigue que

$$\left\langle \mathcal{P}\Phi \middle| \frac{\partial\Psi}{\partial t} + \frac{i}{\hbar}\mathcal{H}\Psi \right\rangle + \left\langle \frac{\partial\Phi}{\partial t} + \frac{i}{\hbar}\mathcal{H}\Phi \middle| \mathcal{P}\Psi \right\rangle = 0,$$

cualquiera sean los vectores Φ y Ψ, independientemente de las condiciones iniciales $\Phi|_{t=0}$ y $\Psi|_{t=0}$. Esta claro que lo anterior se cumple si los vectores Ψ y Φ satisfacen la ecuación

$$i\hbar\frac{\partial\Psi}{\partial t} = \mathcal{H}\Psi. \tag{8.4.9}$$

La ecuación anterior se denomina ecuación de Schrödinger y es la ecuación de evolución de la mecánica cuántica en la representación de Schrödinger.

Veamos ahora cómo se relacionan las representaciones de Heisenberg y de Schrödinger. Para ello notemos que las ecuaciones dinámicas del postulado 8.3.5 han de cumplirse independientemente de que escojamos la representación de Schrödinger (S) o la de Heisenberg (H) discutidas en el apartado anterior. Además, los observables que medimos deben tener los mismos valores medios en ambas representaciones. Eso implica que ha de existir una transformación unitaria (8.2.2) que pase de S a H y viceversa.

Sean ψ y ℓ la función de estado y el observable, respectivamente, en la representación de Heisenberg y Ψ y $\mathcal{L}$ en la de Schrödinger. Entonces entre ambas existe la relación:

$$\Psi = \mathcal{U}^*\psi, \qquad \mathcal{L} = \mathcal{U}^*\ell\,\mathcal{U}, \quad \mathcal{U}^* = e^{-i\mathcal{H}t/\hbar},$$

donde $\mathcal{H}$ es el operador hamiltoniano del sistema que se asume independiente del tiempo.

En efecto, si ψ no depende del tiempo, entonces

$$\frac{\partial\Psi}{\partial t} = \frac{\partial\,\mathcal{U}^*}{\partial t}\psi = -\frac{i}{\hbar}\mathcal{H}e^{-i\mathcal{H}t/\hbar}\psi = -\frac{i}{\hbar}\mathcal{H}\Psi,$$

que no es más que la ecuación de Schrödinger (8.4.9).

Supongamos que ahora $\mathcal{L}$ no depende de t (estamos en la representación de Schrödinger). Entonces,

$$\frac{\partial \ell}{\partial t} = \frac{\partial \mathcal{U}}{\partial t}\mathcal{L}\,\mathcal{U}^* + \mathcal{U}\underbrace{\frac{\partial \mathcal{L}}{\partial t}}_{=\Theta}\mathcal{U}^* + \mathcal{U}\mathcal{L}\frac{\partial \mathcal{U}^*}{\partial t} = \frac{i}{\hbar}(\mathcal{H}\ell - \ell\mathcal{H}) = \frac{i}{\hbar}[\mathcal{H},\ell]. \quad (8.4.10)$$

Si elegimos ℓ como los operadores $\mathcal{P}$ y $\mathcal{X}$ recuperamos las ecuaciones de Heisenberg (8.4.7) y (8.4.8), respectivamente.

Supongamos que tenemos una base $(\Psi_n)_n$ completa de vectores de $\mathbb{H}$. Entonces, todo vector Ψ de $\mathbb{H}$ lo podemos escribir, como ya hemos visto, de la forma

$$\Psi = \sum_n c_n \Psi_n, \qquad f_n = \langle \Psi | \Psi_n \rangle.$$

Es decir, a cada vector de $\mathbb{H}$ le podemos hacer corresponder su vector $\mathbf{c} = (c_1, c_2, \dots)^T$. Análogamente, a cada operador $\mathcal{L}$ le podemos hacer corresponder una matriz $\mathbf{L}$ con entradas $L_{m,n} = \langle \Psi_m | \mathcal{L} | \Psi_n \rangle$. Luego la ecuación (8.4.10) se puede escribir en la forma

$$\frac{\partial \mathbf{L}}{\partial t} = \frac{i}{\hbar}[\mathbf{H}, \mathbf{L}].$$

donde $\mathbf{H}$ es la matriz correspondiente al hamiltoniano del sistema, i.e., recuperamos la también antes mencionada *mecánica matricial de Heisenberg*.

8.4.4. Integrales de movimiento

De la ecuación (8.4.10) se sigue que en la representación de Heisenberg una magnitud física es independiente del tiempo si el operador asociado a dicha magnitud conmuta con el hamiltoniano. Esta propiedad es además muy significativa desde el punto de vista físico como veremos a continuación.

Definición 8.4.1 *Se dice que un observable $\mathcal{A}$ es una integral de movimiento si*

$$\frac{d}{dt}\langle a \rangle = \frac{d}{dt}\langle \Psi | \mathcal{A} | \Psi \rangle = 0.$$

Es decir, una magnitud es una integral de movimiento si el valor de dicha magnitud se conserva en media.

Calculamos la derivada del elemento matricial

$$\frac{d}{dt}\langle\Psi|\mathcal{A}|\Psi\rangle = \left\langle\frac{\partial\Psi}{\partial t}\middle|\mathcal{A}\middle|\Psi\right\rangle + \left\langle\Psi\middle|\frac{\partial\mathcal{A}}{\partial t}\middle|\Psi\right\rangle + \left\langle\Psi\middle|\mathcal{A}\middle|\frac{\partial\Psi}{\partial t}\right\rangle.$$

Supongamos ahora que estamos en la representación de Schrödinger, i.e., A no depende de t y $|\Psi\rangle$ satisface la ecuación de Schrödinger (8.4.9). Entonces, usando (8.4.9) tenemos

$$\frac{d}{dt}\langle\Psi|\mathcal{A}|\Psi\rangle = \left\langle\Psi\middle|\frac{i}{\hbar}[\mathcal{H},\mathcal{A}]\middle|\Psi\right\rangle. \tag{8.4.11}$$

Entonces, $\frac{d}{dt}\langle\Psi|\mathcal{A}|\Psi\rangle = 0$ si y solo si $[\mathcal{A},\mathcal{H}] = 0$.

Si ahora escogemos la representación de Heisenberg entonces (Ψ no depende del tiempo, pero $\mathcal{A}$ si puede) tenemos

$$\frac{d}{dt}\langle\Psi|\mathcal{A}|\Psi\rangle = \left\langle\Psi\middle|\frac{\partial\mathcal{A}}{\partial t}\middle|\Psi\right\rangle = -\frac{i}{\hbar}\langle\Psi|[\mathcal{H},\mathcal{A}]|\Psi\rangle,$$

donde hemos usado la ecuación de Heisenberg (8.4.10). Es decir, también en la representación de Heisenberg $\frac{d}{dt}\langle\Psi|\mathcal{A}|\Psi\rangle = 0$ si y solo si $[\mathcal{A},\mathcal{H}] = 0$.

8.4.5. La ecuación de Schrödinger y el postulado 8.3.5

Supongamos que la magnitud observable L es independiente del tiempo y Ψ es la solución de la ecuación de Schrödinger (8.4.9), donde $\mathcal{H}$ es el operador hamiltoniano

$$\mathcal{H} = \frac{\mathcal{P}}{2m} + V(\mathcal{X}) = \widehat{T} + \widehat{V}.$$

Entonces, la ecuación (8.4.11) nos da

$$\frac{d}{dt}\langle L\rangle = \frac{i}{\hbar}\langle[\mathcal{H},\mathcal{L}]\rangle. \tag{8.4.12}$$

Sea $\mathcal{L} = \mathcal{X}$. Entonces, usando (8.4.5) tenemos

$$[\mathcal{H},\mathcal{X}] = [\widehat{T},\mathcal{X}] = -i\hbar\frac{\partial\mathcal{H}}{\partial\mathcal{P}}. \tag{8.4.13}$$

Sustituyendo lo anterior en (8.4.12) obtenemos

$$\frac{d}{dt}\langle x\rangle = \left\langle \frac{\partial \mathcal{H}}{\partial \mathcal{P}} \right\rangle.$$

Sea $\mathcal{L} = \mathcal{P}$. Entonces, usando (8.4.4) tenemos

$$[\mathcal{H}, \mathcal{P}] = [\widehat{V}, x] = i\hbar \frac{\partial \mathcal{H}}{\partial \mathcal{X}}, \tag{8.4.14}$$

de donde, usando (8.4.12), se sigue que

$$\frac{d}{dt}\langle p\rangle = -\left\langle \frac{\partial \mathcal{H}}{\partial \mathcal{X}} \right\rangle.$$

Es decir, si la función de estado evoluciona según la ecuación de Schrödinger (8.4.9), entonces las medias de las coordenadas e impulsos se comportan como en la mecánica clásica hamiltoniana.

En las mismas condiciones de antes se puede probar (de forma totalmente análoga) que, partiendo de la expresión

$$\frac{d}{dt}\langle \Phi|\mathcal{L}|\Psi\rangle,$$

se obtienen las fórmulas de evolución del postulado 8.3.5. En efecto, en este caso la ecuación 8.4.11 se trasforma en

$$\frac{d}{dt}\langle \Phi|\mathcal{L}|\Psi\rangle = \left\langle \Phi \left| \frac{i}{\hbar}[\mathcal{H}, \mathcal{L}] \right| \Psi \right\rangle.$$

La primera ecuación del postulado 8.3.5 se obtiene eligiendo $\mathcal{L} = \mathcal{X}$ y usando (8.4.13), y la segunda eligiendo $\mathcal{L} = \mathcal{P}$ y usando (8.4.14).

8.4.6. El test de consistencia

Probemos ahora el *test de consistencia*, es decir, que la norma de cada estado $\|\Psi\|$ no cambia con el tiempo (lo que justifica la interpretación probabilística de Born). Para ello probaremos que $d\|\Psi\|^2/dt = 0$. Usando la ecuación de Schrödinger (8.4.9) tenemos

$$\begin{aligned}\frac{\partial \|\Psi\|^2}{\partial t} &= \frac{\partial}{\partial t}\langle \Psi|\Psi\rangle = \left\langle \frac{\partial \Psi}{\partial t} \middle| \Psi \right\rangle + \left\langle \Psi \middle| \frac{\partial \Psi}{\partial t} \right\rangle \\ &= \left\langle -\frac{i}{\hbar}\mathcal{H}\Psi \middle| \Psi \right\rangle + \left\langle \Psi \middle| -\frac{i}{\hbar}\mathcal{H}\Psi \right\rangle = -\frac{i}{\hbar}\big(\langle \mathcal{H}\Psi|\Psi\rangle - \langle \Psi|\mathcal{H}\Psi\rangle \big) = 0.\end{aligned}$$

8.4.7. El principio de incertidumbre

Sean dos operadores autoadjuntos $\mathcal{A}$ y $\mathcal{B}$. Definamos los operadores

$$\Delta\mathcal{A} = \mathcal{A} - \langle A\rangle \mathcal{I}, \qquad \Delta\mathcal{B} = \mathcal{B} - \langle B\rangle \mathcal{I},$$

donde $\langle A\rangle$ y $\langle B\rangle$ son los valores medios de $\mathcal{A}$ y $\mathcal{B}$ en el estado Ψ. Entonces, como $[\mathcal{A},\mathcal{B}] = i\mathcal{L}$, con $\mathcal{L}$ autoadjunto, se sigue que $[\Delta\mathcal{A},\Delta\mathcal{B}] = i\mathcal{L}$.

Las dispersiones de las magnitudes A y B en el estado Ψ vendrán dadas por

$$\Delta A := \sqrt{\langle\Psi|(\Delta\mathcal{A})^2|\Psi\rangle} = \|\Delta\mathcal{A}\Psi\|, \quad \Delta B := \sqrt{\langle\Psi|(\Delta\mathcal{B})^2|\Psi\rangle} = \|\Delta\mathcal{B}\Psi\|.$$

Si usamos la desigualdad de Cauchy-Schwarz (5.1.6)

$$\|\Delta\mathcal{A}\Psi\|\|\Delta\mathcal{B}\Psi\| \geq |\langle\Delta\mathcal{A}\Psi|\Delta\mathcal{B}\Psi\rangle| \geq |\Im\langle\Delta\mathcal{A}\Psi|\Delta\mathcal{B}\Psi\rangle|$$

Calculemos la parte imaginaria de $\langle\Delta\mathcal{A}\Psi|\Delta\mathcal{B}\Psi\rangle = \langle\Psi|\Delta\mathcal{A}\Delta\mathcal{B}|\Psi\rangle$. Para ello recordemos que A es autoadjunto, luego $\Delta\mathcal{A}$ también lo es pues $\langle A\rangle$ es real, luego

$$\begin{aligned}
\Im\langle\Psi|\Delta\mathcal{A}\Delta\mathcal{B}|\Psi\rangle =& \frac{1}{2i}\left(\langle\Psi|\Delta\mathcal{A}\Delta\mathcal{B}|\Psi\rangle - \overline{\langle\Psi|\Delta\mathcal{A}\Delta\mathcal{B}|\Psi\rangle}\right) \\
=& \frac{1}{2i}\left(\langle\Psi|\Delta\mathcal{A}\Delta\mathcal{B}|\Psi\rangle - \langle\Psi|\Delta\mathcal{B}^*\Delta\mathcal{A}^*|\Psi\rangle\right) \\
=& \frac{1}{2i}\left(\langle\Psi|\Delta\mathcal{A}\Delta\mathcal{B}|\Psi\rangle - \langle\Psi|\Delta\mathcal{B}\Delta\mathcal{A}|\Psi\rangle\right) \\
=& \frac{1}{2i}\langle\Psi|[\Delta\mathcal{A},\Delta\mathcal{B}]|\Psi\rangle = \frac{1}{2}\langle\Psi|\mathcal{L}|\Psi\rangle.
\end{aligned}$$

Como $\mathcal{L}$ es autoadjunto, $l = \langle\Psi|\mathcal{L}|\Psi\rangle$ es un número real; así,

$$\Delta A\Delta B \geq \frac{|l|}{2}.$$

Lo anterior aplicado a los operadores $\mathcal{P}$ y $\mathcal{X}$ (ver postulado 8.3.6) nos conduce al principio de incertidumbre de Heisenberg $\Delta x\Delta p \geq \hbar/2$.

8.4.8. Los estados estacionarios del sistema

Toda la discusión anterior nos conduce a que el estado de un sistema viene dado por un vector de estado Ψ que evoluciona según la ecuación de Schrödinger. Vamos a suponer que el operador hamiltoniano (que es autoadjunto) $\mathscr{H}$ es independiente del tiempo y tiene un conjunto de autovectores Φ_n (independientes de t) completo en $\mathbb{H}$, i.e.,

$$\mathscr{H}\Phi_n = E_n\Phi_n,$$

donde los autovalores E_n del hamiltoniano $\mathscr{H}$ representan los posibles valores de la energía del sistema.

Supongamos ahora que tenemos un estado del sistema correspondiente a la energía E_n, que tiene la forma $\Psi_n = \alpha(t)\Phi_n$. Nótese que $\mathscr{H}\Psi_n = E_n\Psi_n$. Como los Ψ_n son estados del sistema, entonces han de satisfacer la ecuación de Schrödinger, i.e.,

$$i\hbar\frac{\partial\Psi_n}{\partial t} = \mathscr{H}\Psi_n = E_n\Psi_n.$$

Resolviendo con respecto al tiempo la ecuación anterior tenemos

$$\Psi_n = e^{-iE_nt/\hbar}\Phi_n,$$

donde Φ_n no depende explícitamente del tiempo.

Comúnmente a los estados Ψ_n anteriores se les denominan estados estacionarios del sistema. Además, de lo anterior se deduce que la única dependencia del tiempo de los estados estacionarios es el factor $e^{-iE_nt/\hbar}$.

Dos propiedades inmediatas que cumplen los estados estacionarios y que los hacen muy especiales son:

1. $\|\Psi_n\|^2 = \|\Phi_n\|^2$, i.e., la densidad de probabilidad no depende del tiempo.

2. Si un observable A no depende del tiempo, entonces el valor medio de dicho observable, $\langle A\rangle$, para cualquier estado estacionario tampoco depende el tiempo.

8.4.9. Los operadores unitarios y la evolución temporal

La discusión del apartado anterior nos da una pista de una transformación unitaria de especial interés. Concretamente la trasformación que define el operador $\mathcal{U}(t) = e^{-i\mathcal{H}t/\hbar}$, donde $\mathcal{H}$ es el hamiltoniano del sistema.

Como los sistemas son invariantes frente a las transformaciones unitarias ello implica que si definimos el vector $\Psi(t) = \mathcal{U}(t)\Psi(0)$, ambos, $\Psi(t)$ y $\Psi(0)$ han de describir el mismo estado. De lo anterior se deduce que para todo t_0, $\Psi(t+t_0) = \mathcal{U}(t)\Psi(t_0)$, es decir, que los estados físicos son invariantes frente a las traslaciones temporales. Como tomar derivadas respecto a t en $\Psi(t) = \mathcal{U}(t)\Psi(0)$ nos conduce a la ecuación de Schrödinger (8.4.9), podemos deducir que dicha ecuación es una consecuencia de la invarianza respecto a las traslaciones temporales de los sistemas físicos.

Nótese que el razonamiento anterior también es válido para cualquier operador unitario de la forma $\mathcal{U}(t) = e^{-i\mathcal{L}t/\hbar}$, con $\mathcal{L}$ autoadjunto (no necesariamente el operador asociado al hamiltoniano del sistema), en cuyo caso obtendríamos la ecuación $i\hbar\frac{\partial\Psi}{\partial t} = \mathcal{L}\Psi$. Como esta última ecuación tiene la misma forma que la ecuación de Schrödinger (8.4.9), entonces es posible tomar como definición del hamiltoniano al operador $\mathcal{H}$ definido por la transformación unitaria[13] $\mathcal{U}(t) = e^{-i\mathcal{H}t/\hbar}$.

8.5. Problemas

Problema 8.1 Dado tres operadores $\mathcal{A}$, $\mathcal{B}$ y $\mathcal{C}$, prueba la identidad de Jacobi

$$[[\mathcal{A}, \mathcal{B}], \mathcal{C}] + [[\mathcal{B}, \mathcal{C}], \mathcal{A}] + [[\mathcal{C}, \mathcal{B}], \mathcal{A}] = 0.$$

[13] Ver, por ejemplo, Steven Weinberg, *Lectures on Quantum Mechanics*. Cambridge University Press, UK, 2012, página 82.

Problema 8.2 Prueba que si $[\mathcal{A},\mathcal{B}]=\alpha\mathcal{B}$, $\alpha\in\mathbb{C}$, entonces $[\mathcal{A},\mathcal{B}^k]=\alpha k\mathcal{B}^{k-1}$, $k\geq 1$.

Problema 8.3 Prueba la fórmula (8.4.5).

Problema 8.4 Sean $A,B\in\mathcal{B}(\mathbb{H})$. Prueba que si $[\mathcal{A},\mathcal{B}]=0$, entonces $e^{\mathcal{A}}e^{\mathcal{B}}=e^{\mathcal{A}+\mathcal{B}}$. Usa lo anterior para probar que, si $\mathcal{A}$ es un operador autoadjunto, entonces $e^{i\mathcal{A}}$ es unitario.

Problema 8.5 Sea $\mathcal{A}\in\mathcal{B}(\mathbb{H})$. Prueba que

$$e^{\mathcal{A}}a e^{-\mathcal{A}}=a+\frac{1}{1!}[\mathcal{A},a]+\frac{1}{2!}[\mathcal{A},[\mathcal{A},a]]+\cdots$$

Ayuda: Encuentra la EDO que satisface el operador $a(t)=e^{t\mathcal{A}}ae^{-t\mathcal{A}}$, donde a no depende del tiempo y desarrolla la función $a(t)$ en potencias de t.

Problema 8.6 Prueba que el producto de dos operadores autoadjuntos $\mathcal{L}$ y $\mathcal{N}$ siempre se puede escribir como

$$\mathcal{L}\mathcal{N}=\mathcal{A}+\mathcal{B},$$

donde $\mathcal{A}$ es autoadjunto y $\mathcal{B}$ es antihermítico, i.e., $\mathcal{B}^*=-\mathcal{B}$.

¿Qué se nos ha quedado en el tintero? 🤔

> Lo que sabemos es poco. Lo que no sabemos es inmenso.
>
> P. Laplace

Entre los muchos temas interesantes e importantes que se podrían haber tratado en un primer curso de Análisis funcional no podemos dejar de mencionar los siguientes:

1. Estudiar con más detalles los espacios L^p. En particular $L^2_{[0,1]}$ y $L^2_{\mathbb{R}}$ como completamiento de $C^2_{[0,1]}$ y $C^2_{\mathbb{R}}$, respectivamente. Estos espacios son de especial importancia en sus aplicaciones no solo en la propia matemática sino en otras áreas de las ciencias. En especial los espacios L^2. Varias referencias introductorias son [5, Capítulo 2], [14, Capítulo 13] y [23, Capítulo 10].

2. Los espacios L^p y ponerlos en relación con la teoría de la medida. Una primera aproximación es [25, Capítulo 7] y [10]. Para más detalles se puede consultar [18].

3. Aplicaciones al mundo real. Aquí cabe mencionar sin duda el uso de la teoría de operadores y de los espacios L^2 en la mecánica cuántica. Aunque en el capítulo 8 discutimos brevemente el tema hay que aclarar que, desde el punto de vista matemático, allí nos restringimos al caso de operadores acotados. Para el caso general se pueden consultar, aparte de [11, Capítulo 11], las magníficas monografías [9] y [17].

Introducción al Análisis Funcional

Prof. Renato Álvarez-Nodarse

Bibliografía

> De los diversos instrumentos del hombre, el más asombroso es, sin duda, el libro. [...] el libro es una extensión de la memoria y de la imaginación.
>
> Jorge Luis Borges
>
> En *"El libro", Borges Oral* (1979)

Para la confección de estas notas hemos consultado tal cantidad de fuentes bibliográficas que es prácticamente imposible incluirlas todas. Sirva esta breve lista como referencia para el lector interesado en profundizar en el apasionante campo del Análisis Funcional y sus aplicaciones. Se recomienda para el tema de espacios métricos, normados y de Hilbert los libros marcados con "m", "n" y "h", respectivamente. Para saber más de mecánica cuántica se recomiendan los libros marcados con "mc".

[1] G. Bachman G y L. Narici, *Functional Analysis*. Dover Publications, New York, 2000.^{m,n,h}

[2] J-L. Basdevant, J. Dalibard, *Quantum mechanics*. Springer Verlag, Berlin, 2002.^{mc}

[3] S.K. Berberian. *Introduction to Hilbert space.* AMS Chelsea Publishing, 1999.^{h}

[4] R.P. Boas, Jr. *A Primer of Real Functions.* Mathematical Association of America, Washington D.C., 1997.^{m}

[5] L. Debnath y P. Mikusinsk. *Introduction to Hilbert spaces with applications.* Elsevier, 2005.^{h}

[6] Y. Eidelman, V.D. Milman, A. Tsolomitis. *Functional Analysis: An Introduction*. Graduate Studies in Mathematics Vol. 66, AMS, RI, 2004.[m,h]

[7] R. Feynman, R. B. Leighton y M. Sands, *The Feynman Lectures on Physics*. Vol. III. *Quantum mechanics*, New Millennium Edition California Institute of Technology, Basic Books, New York, 2010 (de libre acceso en `https://www.feynmanlectures.caltech.edu`).

[8] C. Gasquet y P. Witomski, *Fourier analysis and applications. Filtering, numerical computation, wavelets.* Translated from the French and with a preface by R. Ryan. Texts Appl. Math., 30, Springer-Verlag, New York, 1999.

[9] B.C. Hall. *Quantum Theory for Mathematicians.* Graduate Texts in Mathematics, Volume 267, Springer, New York, 2013.[mc]

[10] A.N. Kolmogorov y A.V. Fomín. *Elementos de la teoría de funciones y del análisis funcional*, Editorial MIR, Moscú, 1978. (Elements of the Theory of Functions and Functional Analysis. Dover, 1999).[m,h]

[11] E. Kreyszig. *Introductory Functional Analysis with Applications.* Wiley Classics Library Edition, New York, 1989.[m,n]

[12] D. G. Luenberger. *Optimization by Vector Space Methods.* Wiley, New York, 1969.[n,h]

[13] B.D. MacCluer. *Elementary Functional Analysis*. Springer, New York, 2009.[n,h]

[14] R. Meise y D. Vogt. *Introduction to Functional Analysis.* Oxford Graduate Texts in Mathematics, Oxford UK, 1997.

[15] J. Mehra, H. Rechenberg, *The Historical Development of Quantum Theory* (Vols. 1–6). Springer-Verlag, New York, 1982-2001.[mc]

[16] A. Messiah, *Mecánica cuántica.* Vol. I y II. Ed. Tecnos. *Quamtum Mechanics*, Dover, New York, 1999.[mc]

[17] E. Prugovecki, *Quantum Mechanics in Hilbert space*. Academic Press, New York, 1971. 2ª Ed. Dover Publications, 2013.[mc]

[18] H. Royden y P. Fitzpatrick. Real Analysis. (4th Edition) Prentice Hall, New York, 2010.

[19] B.P. Rynne y M.A. Youngson. *Linear Functional Analysis.* (2th Edition) Springer, London, 2008.[n,h]

[20] J.M. Sánchez Ron. *Historia de la física cuántica: Volumen I. El período fundacional (1860-1926)*. Drakontos, Barcelona, 2001. *Historia de la física cuántica. Volumen II. La creación e interpretación de la mecánica cuántica: de Heisenberg al gato de Schrödinger (1925-1935)*. Drakontos, Barcelona, 2025.[mc]

[21] K. Saxe. *Beginning Functional Analysis.* Springer, New York, 2002.[n,h]

[22] M. Schechter. *Principles of functional analysis.* Graduate Studies in Mathematics Vol. 36, AMS, RI, 2002.[b,h]

[23] R. Sen. *A First Course in Functional Analysis. Theory and Applications.* Anthem Press, Londres, 2013.

[24] G. Strang. *Linear Algebra and Its Applications*. (4th Edition) Thomson Brooks-Cole, 2006.

[25] A.E. Taylor. *An Introduction to Functional Analysis.* John Wiley and Sons, New York, 1958.

[26] J. Tinsley Oden y L.F. Demkowicz. *Applied Functional Analysis.* CRC Press, BR, 2018.[m,n]

[27] N. Young. *An introduction to Hilbert Space.* Cambridge University Press, Cambridge, 1988.[n,h]

[28] V.A. Zorich, *Mathematical Analysis I.* (2th Edition) Springer-Verlag, Berlin, 2015.

[29] V.A. Zorich, *Mathematical Analysis II.* (2th Edition) Springer-Verlag, Berlin, 2016.

Introducción al Análisis Funcional

Prof. Renato Álvarez-Nodarse

Índice alfabético