Cálculo Avanzado

Notas de clase

Mariano Suárez-Álvarez

Version: 11 de noviembre de 2025

一丝不苟[1]

[1]«Ni una puntada sin cuidado». El pinyin de la frase es *yī sī bù gǒu*.

Índice general

Primera parte

1 **El principio de inducción** **2**
- 1.1 El principio de inducción . . . 2
- 1.2 El principio de recurrencia . . . 4

2 **El axioma de elección** **15**
- 2.1 El axioma de elección . . . 15
- 2.2 Elementos maximales y el Lema de Zorn . . . 19
- 2.3 Algunas aplicaciones . . . 24

3 **Cardinalidad** **31**
- 3.1 La relación de equipotencia . . . 31
- 3.2 Conjuntos finitos . . . 39
- 3.3 Conjuntos infinitos . . . 51
- 3.4 Conjuntos numerables . . . 58
- 3.5 Operaciones con cardinales . . . 71

4 **Ordinales** **86**
- 4.1 Conjuntos bien ordenados . . . 86
- 4.2 Ordinales . . . 99
- 4.3 El principio de inducción transfinita y sus consecuencias . . . 107
- 4.4 Cardinales . . . 110

Segunda parte

5 Espacios métricos **112**

5.1 Métricas y espacios métricos . . . 112
5.2 Ejemplos . . . 114
5.2.1 Métricas euclídeas . . . 114
5.2.2 Espacios de pre-Hilbert . . . 116
5.2.3 Espacios normados . . . 121
La norma $\|-\|_\infty$ sobre $\mathbb{R}^n$ y $\mathbb{C}^n$, 124. Las normas $\|-\|_p$ sobre $\mathbb{R}^n$ y $\mathbb{C}^n$, 125. La ley del paralelogramo y el teorema de Jordan y von Neumann, 129. El espacio normado $\ell^\infty_{\mathbb{k}}$ y sus subespacios c y c_0, 133. Los espacios normados $\ell^p_{\mathbb{k}}$ con $p \in [1, \infty)$, 135.
5.2.4 Semi-métricas y semi-normas . . . 138
5.2.5 Métricas inducidas por familias de semi-métricas y de semi-normas . . . 145
5.2.6 Productos cartesianos . . . 149
5.2.7 Métricas discretas . . . 153
5.2.8 Métricas de Hamming . . . 157
5.2.9 Ultramétricas . . . 164
Un ejemplo sencillo, 165. El espacio de las sucesiones en un conjunto, 165. La métrica sobre un grupo abeliano asociada a una cadena descendente de subgrupos, 166. La métrica p-ádica sobre $\mathbb{Q}$, 168.

6 La topología de un espacio métrico **176**

6.1 Bolas abiertas . . . 176
6.2 Conjuntos abiertos . . . 182
6.3 Entornos y bases de entornos . . . 184
6.4 La topología de un espacio métrico . . . 190
6.4.1 Bases . . . 192
6.4.2 La topología de un subespacio . . . 198
6.4.3 Métricas topológicamente equivalentes . . . 200
6.4.4 Métricas equivalentes en el sentido de Lipschitz . . . 204
6.5 Conjuntos cerrados . . . 210
6.6 Conjuntos densos y espacios separables . . . 218
6.7 El conjunto derivado . . . 226

7 Convergencia **232**

7.1 Sucesiones convergentes . . . 232
7.2 Subsucesiones . . . 243
7.3 Completitud . . . 248
7.4 Métricas equivalentes en el sentido de Cauchy . . . 265
7.5 Espacios normados de dimensión finita . . . 267

7.6 Equiconvergencia . . . 270

8 Continuidad **274**
8.1 Funciones continuas . . . 274
8.2 Continuidad uniforme . . . 284
8.3 Grupos y espacios vectoriales métricos . . . 288
8.4 Problemas de extensión . . . 291
8.5 Completación . . . 301
8.5.1 Completación de grupos abelianos . . . 308
8.5.2 Completación de espacios normados . . . 311
8.5.3 El cuerpo de los números p-ádicos . . . 312
8.6 Espacios de funciones continuas . . . 314
8.7 El teorema de punto fijo de Banach . . . 317
8.7.1 Sucesiones definidas por recurrencia . . . 319
8.7.2 El carácter de Teichmüller . . . 320
8.7.3 El teorema de Lindelöf–Lipschitz–Picard . . . 323
8.8 Funciones lineales . . . 328
8.9 Conjuntos de continuidad . . . 332

9 Compacidad **338**
9.1 Conjuntos compactos . . . 338
9.2 Caracterizaciones de los conjuntos compactos . . . 344
9.3 El teorema de Tikhonov . . . 350
9.4 Funciones continuas sobre espacios compactos . . . 352
9.5 Números de Lebesgue . . . 358
9.6 Variaciones sobre la idea de compacidad . . . 359
9.7 La convergencia uniforme sobre compactos . . . 365
9.8 El teorema de Ascoli–Arzelà . . . 368
9.8.1 Algunas aplicaciones . . . 372
9.8.2 El teorema de Cauchy–Peano . . . 375
9.9 El teorema de Dini . . . 391
9.10 El teorema de Stone–Weierstrass . . . 392

10 Conexión **404**
10.1 Espacios conexos . . . 404

11 El teorema de Baire **411**
11.1 El teorema de Baire . . . 411
11.2 Algunas aplicaciones . . . 415
11.2.1 Funciones reales . . . 415

11.2.2 Uniones de subespacios y de subconjuntos algebraicos 418
11.2.3 Aproximación diofántica. 421

12 Espacios de Banach **429**
12.1 Espacios de Banach 429
12.2 Funciones lineales continuas 431
12.3 El teorema de Hahn–Banach y el espacio dual de un espacio normado. 436
12.4 Completación de espacios de Banach 452
12.5 El teorema de la función abierta y sus consecuencias 460
12.5.1 El teorema de la función inversa acotada 464
12.5.2 El teorema del gráfico cerrado 465
12.5.3 El teorema de acotación uniforme de Banach–Steinhaus 468

Tercera parte

13 Cálculo diferencial e integral **471**
13.1 Integración. 471
13.1.1 Particiones 471
13.1.2 Funciones escalonadas. 474
13.1.3 La integral de Darboux 479
13.1.4 El criterio de integrabilidad de Darboux 486
13.1.5 Los teoremas fundamentales 496
El lema de separación, 496. El teorema del valor medio, 497. El teorema fundamental del cálculo, 498. La regla de Barrow, 500. La regla de integración por sustitución, 501. La regla de integración por partes, 502.

Bibliografía **507**

Notaciones **513**

Lista de personas **514**

Índice **517**

Primera parte

Capítulo 1

El principio de inducción

1.1. El principio de inducción

Recordemos el llamado *Principio de buena ordenación*:

Axioma 1.1.1. *Un subconjunto no vacío de $\mathbb{N}_0$ posee un menor elemento.*

La consecuencia fundamental de esto es el *Principio de inducción*. Para enunciarlo usamos la siguiente definición.

Definición 1.1.2. Un subconjunto S de $\mathbb{N}_0$ es ***inductivo*** si contiene a 0 y vale que

$$k \in S \implies k+1 \in S$$

cualquiera sea k en $\mathbb{N}_0$

Es claro que $\mathbb{N}_0$ es un subconjunto inductivo de $\mathbb{N}_0$. El *Principio de inducción* nos dice que es, de hecho, el único.

Proposición 1.1.3. *El único subconjunto inductivo de $\mathbb{N}_0$ es $\mathbb{N}_0$ mismo.*

Demostración. Sea S un subconjunto inductivo de $\mathbb{N}_0$ y supongamos que S no es igual a $\mathbb{N}_0$, de manera que la diferencia $\mathbb{N}_0 \setminus S$ es un subconjunto no vacío de $\mathbb{N}_0$. De acuerdo al *Principio de buena ordenación*, entonces, podemos considerar su menor elemento: escribámoslo n. Como $0 \in S$, es $n \neq 0$ y, en particular, el número $n-1$ es un elemento de $\mathbb{N}_0$. Como $n-1$ es menor que n, la forma en que elegimos a n implica que $n-1$ no pertenece a $\mathbb{N}_0 \setminus S$ y, por lo tanto, que pertenece

a S. La hipótesis hecha sobre el conjunto S nos dice entonces que $n = (n-1)+1$ es un elemento de S: esto es absurdo, ya que $n \in \mathbb{N}_0 \setminus S$. Esta contradicción provino de haber supuesto que el conjunto $\mathbb{N}_0 \setminus S$ no es vacío, así que debe serlo. Esto prueba la proposición. □

Casi siempre usamos el *Principio de inducción* vía el siguiente resultado.

Corolario 1.1.4. *Para cada $n \in \mathbb{N}_0$ sea $\mathscr{P}_n$ una afirmación. Si*

- *la afirmación $\mathscr{P}_0$ vale y*
- *para cada elemento n de $\mathbb{N}_0$ vale que $\mathscr{P}_n \implies \mathscr{P}_{n+1}$,*

entonces la afirmación $\mathscr{P}_n$ vale para cada $n \in \mathbb{N}_0$.

Demostración. Cuando las dos condiciones del corolario se cumplen el conjunto $\{n \in \mathbb{N}_0 : \mathscr{P}_n\}$ es inductivo y, por lo tanto, coincide con $\mathbb{N}_0$. □

A veces, sin embargo, es más conveniente la siguiente versión — que es, de todas formas, simplemente un caso particular de la anterior.

Corolario 1.1.5. *Para cada $n \in \mathbb{N}_0$ sea $\mathscr{P}_n$ una afirmación. Si*

- *la afirmación $\mathscr{P}_0$ vale y*
- *para cada elemento n de $\mathbb{N}_0$ vale que $\mathscr{P}_0 \wedge \mathscr{P}_1 \wedge \cdots \wedge \mathscr{P}_n \implies \mathscr{P}_{n+1}$,*

entonces la afirmación $\mathscr{P}_n$ vale para cada $n \in \mathbb{N}_0$.

Demostración. Supongamos que se cumplen las dos hipótesis del corolario, y para cada elemento n de $\mathbb{N}_0$ sea $\mathscr{Q}_n$ la afirmación $\mathscr{P}_0 \wedge \mathscr{P}_1 \wedge \cdots \wedge \mathscr{P}_n$. Tenemos entonces que la afirmación $\mathscr{Q}_0$ vale y que para todo $n \in \mathbb{N}_0$ es $\mathscr{Q}_n \implies \mathscr{P}_{n+1}$, de manera que también

$$\mathscr{Q}_n \implies \mathscr{Q}_n \wedge \mathscr{P}_{n+1} = \mathscr{Q}_{n+1}.$$

De acuerdo al Corolario 1.1.4, entonces, para todo $n \in \mathbb{N}_0$ vale la afirmación $\mathscr{Q}_n$ y, en particular, también lo hace la afirmación $\mathscr{P}_n$. □

1.2. El principio de recurrencia

Más allá de los dos corolarios de la sección anterior, el *Principio de inducción* tiene innumerables aplicaciones. Nos interesa aquí destacar una de ellas, que nos da una forma de construir funciones en situaciones que encontramos muy frecuentemente: el llamado *Principio de recurrencia*

Proposición 1.2.1. *Sea X un conjunto, sea $F : \mathbb{N}_0 \times X \to X$ una función, y sea x_0 un elemento de X. Existe exactamente una función $f : \mathbb{N}_0 \to X$ tal que $f(0) = x_0$ y*

$$f(n+1) = F(n, f(n))$$

para cada $n \in \mathbb{N}_0$.

Demostración. Dividiremos nuestro argumento en varios pasos.

Primer paso. Digamos que una función $g : I \to X$ con codominio en X es ***buena*** si

- hay un entero no negativo $n \in \mathbb{N}$ tal que el dominio I de g es $[\![0, n]\!]$, y
- $g(0) = x_0$ y para cada elemento i de $[\![0, n-1]\!]$ vale que $g(i+1) = F(i, g(i))$,

y sea $\mathscr{B}$ el conjunto de todas las funciones buenas. Notemos que una función buena es un subconjunto de $\mathbb{N}_0 \times X$, así que $\mathscr{B}$ es un subconjunto de $\mathscr{P}(\mathbb{N}_0 \times X)$.

Para cada elemento n de $\mathbb{N}_0$ sea $\mathscr{P}_n$ la afirmación

hay una función buena que tiene a $[\![0, n]\!]$ como dominio. (1.1)

Mostraremos que la afirmación $\mathscr{P}_n$ vale cualquiera sea elemento n de $\mathbb{N}_0$ procediendo por inducción. La función $g_0 : [\![0, 0]\!] = \{0\} \to X$ tal que $g_0(0) = x_0$ es claramente buena y, por lo tanto, la afirmación $\mathscr{P}_0$ es cierta. Supongamos ahora que k es un elemento de $\mathbb{N}_0$ tal que la afirmación $\mathscr{P}_k$ es cierta, de manera que hay una función buena $g : [\![0, k]\!] \to X$. Podemos entonces definir una función $h : [\![0, k+1]\!] \to X$ poniendo, para cada $i \in [\![0, k+1]\!]$,

$$h(i) \coloneqq \begin{cases} g(i) & \text{si } i \leq k; \\ F(k, g(k)) & \text{si } i = k+1. \end{cases}$$

Si i es un elemento de $[\![0, k]\!]$, entonces o bien $i < k$, y en ese caso tenemos que $i + 1 \leq k$ y

$$h(i+1) = g(i+1) = F(i, g(i)) = F(i, h(i))$$

porque la función g es buena, o bien $i = k$, y en ese caso es $i + 1 = k + 1$ y

$$h(i+1) = F(k, g(k)) = F(k, h(k))$$

por la forma en que definimos a h. Esto nos dice que h es una función buena y, por lo tanto, que la afirmación $\mathscr{P}_{k+1}$ es cierta. Esto completa la inducción.

Segundo paso. Consideremos ahora el conjunto

$$f \coloneqq \bigcup_{g \in \mathscr{B}} g,$$

que es un subconjunto de $\mathbb{N}_0 \times X$ y probemos que, de hecho, se trata de una función $\mathbb{N}_0 \to X$.

- Sea n un elemento cualquiera de $\mathbb{N}_0$. De acuerdo a (1.1), sabemos que hay una función buena $g : [\![0, n]\!] \to X$ en $\mathscr{B}$, así que el par ordenado $(n, g(n))$ de $\mathbb{N}_0 \times X$ pertenece a g y, por lo tanto, también al conjunto f.
- Tenemos que mostrar que para todo $n \in \mathbb{N}_0$

  $$\textit{hay exactamente un elemento } u \textit{ de } X \textit{ tal que } (n, u) \textit{ pertenece a } f, \tag{1.2}$$

  y lo hacemos por inducción. Esto es cierto cuando n es 0. En efecto, si u es un elemento de X tal que $(0, u)$ pertenece a f, entonces hay una función buena $g : I \to X$ tal que $(0, u) \in g$ o, equivalentemente, tal que $g(0) = u$: como g es buena, esto implica que $u = x_0$.

  Supongamos ahora que k es un elemento de $\mathbb{N}_0$ tal que la afirmación (1.2) es cierta cuando n es k, y supongamos que hay dos elementos u y v de X tales que los pares $(k+1, u)$ y $(k+1, v)$ pertenecen los dos a f. Mostraremos que $u = v$, ya que esto nos dirá que la afirmación (1.2) es cierta también cuando n es $k+1$.

  De la definición de f podemos deducir que hay dos funciones buenas $g : I \to X$ y $h : J \to X$ tales que $(k+1, u) \in g$ y $(k+1, v) \in h$, esto es, tales que $g(k+1, u)$ y $h(k+1) = v$. Como g y h son buenas, hay enteros no negativos s y t tales que $I = [\![0, s]\!]$ y $J = [\![0, t]\!]$, y como $(k+1, u) \in g$ y $(k+1, v) \in h$ es claro que $s \geq k+1$ y $t \geq k+1$. Ahora bien, como $(k, g(k)) \in g \subseteq f$ y $(k, h(k)) \in h \subseteq f$, la hipótesis inductiva implica que $g(k) = h(k)$. Es entonces

  $$u = g(k+1) = F(k, g(k)) = F(k, h(k)) = h(k+1) = v,$$

  como queremos.

La conclusión de esto es que lo que el conjunto f que definimos arriba es, como dijimos, una función $f : \mathbb{N}_0 \to X$.

Tercer paso. Mostremos ahora que esta función satisface las dos condiciones del enunciado

- La función $g_0 : [\![0, 0]\!] \to X$ tal que $g_0(0) = x_0$ es buena, así que g_0 es un subconjunto de f y $(0, x_0) \in g_0 \subseteq f$, de manera que es $f(0) = x_0$.
- Sea n un elemento de $\mathbb{N}_0$. De acuerdo a (1.2), hay una función buena $g : [\![n+1]\!] \to X$ que tiene por dominio a $[\![n+1]\!]$ y, por lo tanto, los pares $(n, g(n))$ y $(n+1, g(n+1))$ pertenecen a g. Como $g \subseteq f$, también pertenecen a f y esto significa que

  $$f(n+1) = g(n+1) = F(n, g(n)) = F(n, f(n)),$$

  ya que la función g es buena.

Cuarto y último paso. Con todo esto ya establecimos la afirmación de existencia que hace la proposición. Veamos, para terminar, la afirmación de unicidad. Supongamos que $f : \mathbb{N}_0 \to X$ y $f' : \mathbb{N}_0 \to X$ son dos funciones tales que $f(0) = x_0$ y $f(n+1) = F(n, f(n))$ para cada $n \in \mathbb{N}_0$, por un lado, y $f'(0) = x_0$ y $f'(n+1) = F(n, f'(n))$ para cada $n \in \mathbb{N}_0$, por otro. Queremos probar que $f = f'$ y para ello probaremos, procediendo por inducción, que $f(n) = f'(n)$ cualquiera sea n en $\mathbb{N}_0$. Que $f(0) = f'(0)$ es claro, y si k es un elemento de $\mathbb{N}_0$ tal que $f(k) = f'(k)$, entonces tenemos que $f(k+1) = F(k, f(k)) = F(k, f'(k)) = f'(k+1)$. Esto completa la prueba de la proposición. □

Este *Principio de recurrencia* está en la base de muchas construcciones usuales.

Ejemplo 1.2.2. Consideremos la función $F : (n, k) \in \mathbb{N}_0 \times \mathbb{N}_0 \mapsto (n+1)k \in \mathbb{N}_0$. De acuerdo a la Proposición 1.2.1 hay exactamente una función $f : \mathbb{N}_0 \to \mathbb{N}_0$ tal que $f(0) = 1$ y

$$f(n+1) = F(n, f(n)) = (n+1) \cdot f(n)$$

para cada $n \in \mathbb{N}_0$. Es entonces

$$\begin{aligned} f(1) &= 1 \cdot f(0) = 1, \\ f(2) &= 2 \cdot f(1) = 2 \cdot 1, \\ f(3) &= 3 \cdot f(2) = 3 \cdot 2 \cdot 1, \\ f(4) &= 4 \cdot f(3) = 4 \cdot 3 \cdot 2 \cdot 1, \end{aligned}$$

y, más generalmente, es claro que $f(n) = n!$, el factorial de n.

Hay muchas variaciones del *Principio de recurrencia* que se adaptan a las muy diversas situaciones en las que lo usamos. Los siguientes ejercicios presentan algunas de ellas.

Ejercicio 1.2.3. Sea X un conjunto.

(a) Sea $F : \mathbb{N}_0 \times X \times X \to X$ una función, y sean x_0 y x_1 dos elementos de X. Pruebe que hay una y solo una función $f : \mathbb{N}_0 \to X$ tal que $f(0) = x_0$, $f(1) = x_1$ y

$$f(n+2) = F(n, f(n), f(n+1))$$

para cada $n \in \mathbb{N}_0$.

(b) Más generalmente, sea k un entero positivo, sea $F : \mathbb{N}_0 \times X^k \to X$ una función, y sean $x_0, x_1, \ldots, x_{k-1}$ elementos de X, y pruebe que hay una y solo una función $f : \mathbb{N}_0 \to X$ tal que $f(i) = x_i$ para cada $i \in [\![0, k-1]\!]$ y

$$f(n+k) = F(k, f(n), f(n+1), \ldots, f(n+k-1))$$

para cada $n \in \mathbb{N}_0$.

Ejemplo 1.2.4. Consideremos la función

$$F : (n, a, b) \in \mathbb{N}_0 \times \mathbb{N}_0 \times \mathbb{N}_0 \mapsto a + b \in \mathbb{N}_0.$$

De acuerdo a la primera parte del Ejercicio 1.2.3, hay exactamente una función $f : \mathbb{N}_0 \to \mathbb{N}_0$ tal que $f(0) = 0$, $f(1) = 1$ y

$$f(n+2) = F(n, f(n), f(n+1)) = f(n) + f(n+1)$$

para cada $n \in \mathbb{N}_0$. Tenemos entonces que

$$\begin{aligned}
f(2) &= f(0) + f(1) = 0 + 1 = 1,\\
f(3) &= f(1) + f(2) = 1 + 1 = 2,\\
f(4) &= f(2) + f(3) = 1 + 2 = 3,\\
f(5) &= f(3) + f(4) = 2 + 3 = 5,\\
f(6) &= f(4) + f(5) = 3 + 5 = 8,
\end{aligned}$$

y, más generalmente, para cada $n \in \mathbb{N}_0$ el número $f(n)$ no es otro que el n–ésimo número de Fibonacci.

La siguiente forma del *Principio de recurrencia* es particularmente importante.

Proposición 1.2.5. *Sea X un conjunto, sea $\mathscr{S}$ el conjunto de todas las funciones $I \to X$ definidas sobre un subconjunto I de $\mathbb{N}_0$ de la forma $[\![n]\!]$ para algún $n \in \mathbb{N}_0$, y sea $F : \mathscr{S} \to X$ una función. Hay exactamente una función $f : \mathbb{N} \to X$ tal que para todo $n \in \mathbb{N}$ es $f(n) = F(f|_{[\![n-1]\!]})$.*

Veremos más adelante, cuando nos ocupemos de los números ordinales, que esta forma del *Principio de recurrencia* tiene una extensión importante a lo que llamamos el caso *transfinito*.

Demostración. Cuando g es un elemento de $\mathscr{S}$, de manera que hay un entero $n \in \mathbb{N}_0$ tal que $g : [\![n]\!] \to X$, escribamos $G(g) : [\![n+1]\!] \to X$ a la función que en cada elemento i de $[\![n+1]\!]$ toma el valor

$$G(g)(i) = \begin{cases} g(i) & \text{si } i \leq n;\\ F(g) & \text{si } i = n+1. \end{cases}$$

Es claro que esto es un elemento de $\mathscr{S}$ y que $G(g)|_{[\![n]\!]} = g$. Hay, por lo tanto, una función

$$G : \mathscr{S} \to \mathscr{S}.$$

Sea $\phi : [\![0]\!] \to X$ el único elemento de $\mathscr{S}$ que tiene por dominio a $[\![0]\!]$. De acuerdo al *Principio de recurrencia* 1.2.1 hay exactamente una función $h : \mathbb{N}_0 \to \mathscr{S}$ tal que $h(0) = \phi$ y $h(n+1) = G(h(n))$

para cada $n \in \mathbb{N}_0$. Notemos que para todo $n \in \mathbb{N}_0$ el dominio de $h(n)$ es $[\![n]\!]$. Esto es claramente cierto cuando n es 0, ya que $h(0) = \phi$, y si es cierto para algún elemento n de $\mathbb{N}_0$, de manera que $h(n) : [\![n]\!] \to X$, entonces $h(n+1) = G(h(n))$ es una función $[\![n+1]\!] \to X$.

Consideremos ahora la función $f : n \in \mathbb{N} \mapsto h(n)(n) \in X$ y mostremos que

$$\textit{para todo } n \in \mathbb{N} \textit{ es } f|_{[\![n]\!]} = h(n). \tag{1.3}$$

Es $f|_{[\![0]\!]} = h(0)$ porque hay exactamente una función $\varnothing \to X$. Por otro lado, si n es un elemento de $\mathbb{N}$ tal que $f|_{[\![n]\!]} = h(n)$, entonces $\big(f|_{[\![n+1]\!]}\big)|_{[\![n]\!]} = f|_{[\![n]\!]} = h(n) = G(h(n))|_{[\![n]\!]} = h(n+1)|_{[\![n]\!]}$ y $f|_{[\![n+1]\!]}(n+1) = f(n+1) = h(n+1)(n+1)$, así que, como $[\![n+1]\!] = [\![n]\!] \cup \{n+1\}$, tenemos que también $f|_{[\![n+1]\!]} = h(n+1)$. Esto prueba (1.3)

Usando eso, vemos inmediatamente que $f(1) = h(1)(1) = G(h(0))(1) = F(h(0)) = F(f|_{[\![0]\!]})$, y que siempre que n es un elemento de $\mathbb{N}$ tal que $f(n) = F(f|_{[\![n-1]\!]})$ vale que

$$f(n+1) = h(n+1)(n+1) = G(h(n))(n+1) = F(h(n)) = F(f|_{[\![n]\!]}),$$

de manera que, de hecho, es $f(n) = F(f|_{[\![n-1]\!]})$ para todo $n \in \mathbb{N}$. Hemos probado con esto la afirmación de existencia de la proposición.

Ocupémosnos ahora de la afirmación de unicidad. Supongamos que f y f' son dos funciones $\mathbb{N} \to X$ tales que $f(n) = F(f|_{[\![n-1]\!]})$ y $f'(n) = F(f'|_{[\![n-1]\!]})$ para todo $n \in \mathbb{N}$. Queremos probar que $f = f'$, pero empezamos por mostrar que

$$\textit{para todo } n \in \mathbb{N} \textit{ es } f|_{[\![n]\!]} = f'|_{[\![n]\!]}.$$

Esto es evidente cuando n es 0, ya que $f|_{[\![0]\!]}$ y $f'|_{[\![0]\!]}$ son dos funciones $\varnothing \to X$, y hay una sola tal función, y si n es un elemento de $\mathbb{N}$ tal que $f|_{[\![n]\!]} = f'|_{[\![n]\!]}$, entonces es

$$f(n+1) = F(f|_{[\![n]\!]}) = F(f'|_{[\![n]\!]}) = f'(n+1),$$

así que también $f|_{[\![n+1]\!]} = f'|_{[\![n+1]\!]}$. Ahora bien, si n es un elemento cualquiera de $\mathbb{N}$, entonces $f(n) = f|_{[\![n]\!]}(n) = f'|_{[\![n]\!]}(n) = f'(n)$, y esto nos dice que $f = f'$, como queremos. □

Esta forma del *Principio de recurrencia* nos permite construir funciones $f : \mathbb{N}_0 \to X$ tales que la determinación del valor de f en en un entero n «depende» del valor de f en todos los enteros menores que n. Es siguiente es un ejemplo célebre de una función de este tipo.

Ejemplo 1.2.6. Sea $\mathscr{S}$ el conjunto de todas las funciones de la forma $[\![n]\!] \to \mathbb{N}_0$ con $n \in \mathbb{N}_0$, y sea

$F : \mathscr{S} \to \mathbb{N}_0$ la función que en cada elemento $g : [\![n]\!] \to X$ toma el valor

$$F(g) = \begin{cases} 1 & \text{si } n = 0; \\ \displaystyle\sum_{i=1}^{n} g(i)g(n+1-i) & \text{si } n > 0. \end{cases}$$

La proposición que acabamos de probar nos dice que hay exactamente una función $f : \mathbb{N} \to \mathbb{N}_0$ tal que $f(1) = F(f|_{[\![0]\!]}) = 1$ y

$$f(n) = F(f|_{[\![n-1]\!]}) = \sum_{i=1}^{n-1} f(i)f(n+1-i)$$

para cada $n \in \mathbb{N}$. Así, tenemos que

$$\begin{aligned} f(1) &= 1, \\ f(2) &= f(1)f(1) = 1 \cdot 1 = 1, \\ f(3) &= f(1)f(2) + f(2)f(1) = 1 \cdot 1 + 1 \cdot 1 = 2, \\ f(4) &= f(1)f(3) + f(2)f(2) + f(3)f(1) = 1 \cdot 2 + 1 \cdot 1 + 2 \cdot 1 = 5, \\ f(5) &= f(1)f(4) + f(2)f(3) + f(3)f(2) + f(4)f(1) = 1 \cdot 5 + 1 \cdot 2 + 2 \cdot 1 + 5 \cdot 1 = 14, \end{aligned}$$

y, si continuamos calculando, obtenemos la siguiente tabla de valores para f:

n	1	2	3	4	5	6	7	8	9	10	11	12
$f(n)$	1	2	5	14	42	132	429	1 430	4 862	16 796	58 786	208 012

Para cada $n \in \mathbb{N}_0$ el número $f(n)$ es conocido como el n-ésimo ***número de Catalan***, por Eugène Charles Catalan. Referimos al lector al libro [Sta2015] de Richard Stanley, enteramente dedicado a estos números, para obtener más información sobre ellos.

Ejercicio 1.2.7. Muestre que si $f : \mathbb{N} \to \mathbb{N}_0$ es la función que construimos en el Ejemplo 1.2.6, entonces para todo $n \in \mathbb{N}$ vale que

$$f(n) = \frac{1}{n+1}\binom{2n}{n}.$$

Ejercicio 1.2.8. Sea X un conjunto, sean $\alpha, \beta : \mathbb{N}_0 \to X$ dos funciones tales que $\alpha(0) = \beta(0)$, y sea $F : X \times X \to X$ una función. Muestre que hay exactamente una función $f : \mathbb{N}_0 \times \mathbb{N}_0 \to X$ tal que

$$f(n, 0) = \alpha(n) \quad \text{y} \quad f(0, n) = \beta(n)$$

para toda elección de n en $\mathbb{N}_0$, y

$$f(m,n) = F(f(m-1,n), f(m,n-1))$$

para toda elección de m y n en $\mathbb{N}$.

Ejemplo 1.2.9. Consideremos las funciones

$$F : (m,n) \in \mathbb{N}_0 \times \mathbb{N}_0 \to m+n \in \mathbb{N}_0, \quad \alpha : n \in \mathbb{N}_0 \mapsto 1 \in \mathbb{N}_0, \quad \beta : n \in \mathbb{N}_0 \mapsto 1 \in \mathbb{N}_0.$$

Se sigue del ejercicio que hay exactamente una función $f : \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{N}_0$ tal que

$$f(n,0) = f(0,n) = 1 \quad \text{para cada } n \in \mathbb{N}_0$$

y

$$f(m,n) = f(m-1,n) + f(m,n-1) \quad \text{para cada } m, n \in \mathbb{N}.$$

Es, por ejemplo,

$$f(2,1) = f(1,1) + f(2,0) = f(1,0) + f(0,1) + f(2,0) = 3.$$

Calculando, podemos tabular los valores de f:

$f(m,n)$	0	1	2	3	4	5
0	1	1	1	1	1	1
1	1	2	3	4	5	6
2	1	3	6	10	15	21
3	1	4	10	20	35	56
4	1	5	15	35	70	126
5	1	6	21	56	126	252

No es difícil mostrar que, de hecho, es

$$f(m,n) = \binom{m+n}{n}$$

para toda elección de m y n en $\mathbb{N}_0$.

El *Principio de recurrencia* es la herramienta fundamental con la que construimos funciones en la llamada *aritmética de Peano*. El siguiente ejercicio da dos ejemplos sencillos e importantes de eso: la construcción de la suma y de la multiplicación en $\mathbb{N}_0$.

Ejercicio 1.2.10. Sea $\mathscr{S}$ el conjunto de todas las funciones $\mathbb{N}_0 \to \mathbb{N}_0$ y sean $s : n \in \mathbb{N}_0 \mapsto n+1 \in \mathbb{N}_0$ y $z : n \in \mathbb{N}_0 \mapsto 0 \in \mathbb{N}_0$. Pruebe las siguientes afirmaciones.

(a) Hay una y solo una función $\sigma : \mathbb{N}_0 \to \mathscr{S}$ tal que $\sigma(0) = \mathrm{id}_{\mathbb{N}_0}$ y $\sigma(n+1) = s \circ \sigma(n)$ para todo $n \in \mathbb{N}_0$, y que $m + n = \sigma(m)(n)$ cualesquiera sean m y n en $\mathbb{N}_0$.

(b) Hay una y solo una función $\mu : \mathbb{N}_0 \to \mathscr{S}$ tal que $\mu(0) = z$ y $\mu(n+1)(m) = \sigma(m) \circ \mu(n)$ para todo $n \in \mathbb{N}_0$.

Más generalmente, el *Principio de recurrencia* es una herramienta fundamental en toda la matemática y en las ciencias de la computación. Los siguientes ejercicios tiene como objetivo presentar una de las razones de esto.

Ejercicio 1.2.11. Consideremos las funciones

- $s : x_1 \in \mathbb{N}_0 \mapsto x_1 + 1 \in \mathbb{N}_0$,
- $z_k : (x_1, \ldots, x_k) \in \mathbb{N}_0^k \mapsto 0 \in \mathbb{N}_0$ para cada $k \in \mathbb{N}_0$,
- $\pi_m^n : (x_1, \ldots, x_n) \in \mathbb{N}_0^n \mapsto x_m \in \mathbb{N}_0$ para cada elección de $n \in \mathbb{N}$ y de $m \in [\![n]\!]$,

Por otro lado, si $r, k \in \mathbb{N}_0$ y $f : \mathbb{N}_0^r \to \mathbb{N}_0$ y $g_1, \ldots, g_r : \mathbb{N}_0^k \to \mathbb{N}_0$ son funciones, entonces escribamos $f\langle g_1, \ldots, g_r\rangle$ a la función

$$(x_1, \ldots, x_k) \in \mathbb{N}_0^k \mapsto f(g_1(x_1, \ldots, x_k), g_2(x_1, \ldots, x_k), \ldots, g_r(x_1, \ldots, x_k)) \in \mathbb{N}_0.$$

Llamamos a esta función la ***composición generalizada*** de f con $(g_1, \ldots, g_r)$.

(a) Sea k un entero no negativo y sean $g : \mathbb{N}_0^k \to \mathbb{N}_0$ y $h : \mathbb{N}_0^{k+2} \to \mathbb{N}_0$ dos funciones. Muestre que hay exactamente una función $R(g,h) : \mathbb{N}_0^{k+1} \to \mathbb{N}_0$ tal que

$$R(g,h)(x_1, \ldots, x_{k+1}) = \begin{cases} g(x_1, \ldots, x_k) & \text{si } x_{k+1} = 0; \\ h(x_1, \ldots, x_k, x_{k+1} - 1, R(g,h)(x_1, \ldots, x_k, x_{k+1} - 1)) & \text{si } x_{k+1} > 0. \end{cases}$$

(b) Sea $\mathscr{F}$ el conjunto de todas las funciones $\mathbb{N}_0^k \to \mathbb{N}_0$ con $k \in \mathbb{N}_0$. Decimos que un subconjunto D de $\mathscr{F}$ es ***distinguido*** si las siguientes afirmaciones son ciertas:

- Las funciones s, z_k con $k \in \mathbb{N}_0$, y π_m^n con $n \in \mathbb{N}$ y $m \in [\![n]\!]$ pertenecen todas a D.
- Si $r, k \in \mathbb{N}_0$, y $f : \mathbb{N}_0^r \to \mathbb{N}_0$ y $g_1, \ldots, g_r : \mathbb{N}_0^k \to \mathbb{N}_0$ son elementos de D, entonces la composición generalizada $f\langle g_1, \ldots, g_r\rangle : \mathbb{N}_0^k \to \mathbb{N}_0$ es un elemento de D.
- Si $k \in \mathbb{N}_0$ y $g : \mathbb{N}_0^k \to \mathbb{N}_0$ y $h : \mathbb{N}_0^{k+2} \to \mathbb{N}_0$ son dos elementos de D, entonces $R(g,h) : \mathbb{N}_0^{k+1} \to \mathbb{N}_0$ es un elemento de D.

Pruebe que hay subconjuntos de $\mathscr{F}$ que son distinguidos, de manera que tiene sentido considerar la intersección de todos ellos, a la que escribimos PR.

Las funciones que pertenecen al conjunto PR descripto en este ejercicio son las llamadas ***funciones recursivas primitivas***. El objetivo del siguiente ejercicio es mostrar algunos ejemplos no triviales.

Ejercicio 1.2.12.

(a) Muestre que las funciones $\mathsf{sum} \coloneqq R(\pi_1^1, s\langle\pi_3^3\rangle)$ y $\mathsf{mult} \coloneqq R(z_1, \mathsf{sum}\langle\pi_1^3, \pi_3^3\rangle)$ tienen para todo $(x_1, x_2) \in \mathbb{N}_0^2$

$$\mathsf{sum}(x_1, x_2) = x_1 + x_2, \qquad \mathsf{mult}(x_1, x_2) = x_1 \cdot x_2.$$

(b) Muestre que la función $\mathsf{pred} \coloneqq R(z_0, \pi_1^2)$ y $\mathsf{sub} \coloneqq R(\pi_1^1, \mathsf{pred}\langle\pi_3^3\rangle)$ tienen para todo $x_1 \in \mathbb{N}_0$

$$\mathsf{pred}(x_1) = \begin{cases} 0 & \text{si } x_1 = 0; \\ x_1 - 1 & \text{si } x_1 > 0; \end{cases} \qquad \mathsf{sub}(x_1, x_2) = \begin{cases} x_1 - x_2 & \text{si } x_1 \geq x_2; \\ 0 & \text{si } x_1 < x_2. \end{cases}$$

(c) Muestre que las funciones $\mathsf{isZero} \coloneqq R(s\langle z_0\rangle, z_2)$ e $\mathsf{if} \coloneqq R(\pi_2^2, \pi_1^4)\langle\pi_2^3, \pi_3^3, \pi_1^3\rangle$ son tales que para toda elección de x_1, x_2 y x_3 en $\mathbb{N}_0$ vale que

$$\mathsf{isZero}(x_1) = \begin{cases} 1 & \text{si } x_1 = 0; \\ 0 & \text{si } x_1 > 0; \end{cases} \qquad \mathsf{if}(x_1, x_2, x_3) = \begin{cases} x_2 & \text{si } x_1 > 0; \\ x_3 & \text{si } x_1 = 0. \end{cases}$$

(d) Muestre que si $f : \mathbb{N}_0^k \to \mathbb{N}_0$ es una función con k positivo, entonces

$$\mathsf{sigma}_f \coloneqq R(z_{k-1}, \mathsf{sum}\langle f\langle\pi_1^{k+1}, \ldots, \pi_{k-1}^{k+1}, s\langle\pi_k^{k+1}\rangle\rangle, \pi_{k+1}^{k+1}\rangle)$$

tiene, para cada $(x_1, \ldots, x_k) \in \mathbb{N}_0^k$,

$$\mathsf{sigma}_f(x_1, \ldots, x_k) = \sum_{i=1}^{x_k} f(x_1, \ldots, x_{k-1}, i).$$

(e) Consideremos las funciones

$$\begin{aligned} &\mathsf{eq} \coloneqq \mathsf{isZero}\langle\mathsf{sum}\langle\mathsf{sub}, \mathsf{sub}\langle\pi_2^2, \pi_1^2\rangle\rangle\rangle, \\ &\mathsf{check} \coloneqq \mathsf{eq}\langle\pi_1^3, \mathsf{mult}\langle\pi_2^3, \pi_3^3\rangle\rangle, \\ &\mathsf{divide} \coloneqq \mathsf{sigma}_{\mathsf{check}}\langle\pi_1^2, \pi_2^2, \pi_1^2\rangle, \\ &\mathsf{isPrime} \coloneqq \mathsf{sigma}_{\mathsf{divide}}\langle\pi_1^1, \pi_1^1\rangle. \end{aligned}$$

Pruebe que para todo $x \in \mathbb{N}_0$ se tiene que

$$\mathsf{isPrime}(x) = \begin{cases} 1 & \text{si } x \text{ es es primo;} \\ 0 & \text{si no lo es.} \end{cases}$$

(f) El *Postulado de Bertrand* afirma que para todo entero n mayor que 2 hay un número primo p tal que $n < p < 2n$. Esto fue conjeturado por Joseph Bertrand en [Ber1845] y probado por primera vez por Pafnuty Lvovich Chebyshev y luego por Srinivasa Ramanujan en [Ram1919]; más tarde, Paul Erdős dio una demostración elemental en [Erd1932]. Usando este resultado demuestre que para todo entero positivo n hay al menos n números primos menores que 2^n, y luego que la función $f : \mathbb{N}_0 \to \mathbb{N}_0$ tal que para cada $n \in \mathbb{N}_0$ el número $f(n)$ es el $(n+1)$-ésimo número primo es recursiva primitiva.

Ejercicio 1.2.13. Muestre que hay exactamente una función $A : \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{N}_0$ tal que

$$A(0, n) = n + 1,$$
$$A(m + 1, 0) = A(m, 1),$$
$$A(m + 1, n + 1) = A(m, A(m + 1, n))$$

cualesquiera sean m y n en $\mathbb{N}_0$. Esta función es la ***función de Ackermann–Péter***, por Wilhelm Ackermann y Rózsa Péter, aunque fue considerada originalmente por Péter en [Pét1935] y Raphael Robinson, modificando un ejemplo originalmente propuesto por David Hilbert de una función que no es recursiva primitiva. Fue Ackermann, alumno de Hilbert, el que pudo probar esto, en [Ack1928].

Esta función crece extraordinariamente rápido. Por ejemplo,

$$A(4, 3) = 2^{2^{65556}} - 3,$$

un número de 19 733 dígitos. Usando esto, podemos mostrar que se trata de una función que no es recursiva primitiva.

Digamos que A ***mayoriza*** a una función $f : \mathbb{N}_0^k \to \mathbb{N}_0$ si hay un elemento b de $\mathbb{N}_0$ tal que

$$f(a_1, \ldots, a_k) < A(\text{máx}\{a_1, \ldots, a_k\}, b)$$

cualesquiera sean los elementos $a_1, \ldots, a_k$ de $\mathbb{N}_0$.

(a) Muestre que la función A no se mayoriza a sí misma.

(b) Muestre que la función A mayoriza a todas las funciones recursivas primitivas.

Se sigue de esto que A no es primitiva recursiva, por supuesto.

```
R[g_, h_][xs___, 0] := g[xs];
R[g_, h_][xs___, x_?Positive] := h[xs, x - 1, R[g, h][xs, x - 1]];

comp[f_, gs__][xs___] := f @@ ComapApply[{gs}, {xs}]

p[n_, m_][xs___] := {xs}[[m]]  /; Length[{xs}] == n;
z[n_][xs___] := 0  /; Length[{xs}] == n;
s[x_] := x + 1;

sum = R[p[1, 1], comp[s, p[3, 3]]];
mul = R[z[1], comp[sum, p[3, 1], p[3, 3]]];
pred = R[z[0], p[2, 1]];
sub = R[p[1, 1], comp[pred, p[3, 3]]];
iszero = R[comp[s, z[0]], z[2]];
eq = comp[iszero, comp[sum, sub, comp[sub, p[2, 2], p[2, 1]]]];
sigma[f_, k_] := R[
   z[k - 1],
   comp[sum,
    comp[f, Splice@Table[p[k + 1, i], {i, 1, k - 1}], comp[s, p[k + 1, k]]],
    p[k + 1, k + 1]
    ]
   ];
check = comp[eq, p[3, 1], comp[mul, p[3, 2], p[3, 3]]];
divide = comp[sigma[check, 3], p[2, 1], p[2, 2], p[2, 1]];
dos = comp[s, comp[s, z[1]]];
isprime = comp[eq, dos, comp[sigma[divide, 2], p[1, 1], p[1, 1]]];
```

Figura 1.1. Las definiciones del Ejercicio 1.2.12 implementadas en MATHEMATICA.

Capítulo 2

El axioma de elección

2.1. El axioma de elección

Podemos enunciar el *Axioma de elección* en la siguiente forma:

Axioma 2.1.1. *Sea $\mathscr{S}$ una familia de subconjuntos no vacíos de un conjunto X. Hay una función $c : \mathscr{S} \to X$ tal que $c(A) \in A$ para cada elemento A de $\mathscr{S}$.*

Llamamos a esa función c una ***función de elección*** para la familia $\mathscr{S}$. Intuitivamente, lo que dice el axioma es que podemos *elegir* en cada elemento A de $\mathscr{S}$ un elemento $c(A)$ y — esto es crítico — hacer todas esas elecciones de manera *simultánea*. Mostraremos en la Proposición 3.2.3 que cuando la familia $\mathscr{S}$ es finita no necesitamos el axioma para hacer esto, así que este es solo útil cuando esa familia es infinita.

Hay muchas consecuencias del *Axioma de elección* que son útiles en diferentes situaciones. Presentaremos aquí algunas de ellas. De hecho, estos resultados son todos *equivalentes* al *Axioma de elección*, pero no probaremos esto aquí.

Proposición 2.1.2. *Si X e Y son dos conjuntos y $R : X \to \mathscr{P}(Y) \setminus \{\varnothing\}$ es una función, entonces hay una función $f : X \to Y$ tal que $f(x) \in R(x)$ para todo x de X.*

Demostración. En esas condiciones el conjunto $\mathscr{S} := \{R(x) : x \in X\}$ es una familia de subconjuntos no vacíos de Y. El *Axioma de elección* nos dice entonces que hay una función $c : \mathscr{S} \to Y$ tal que $c(A) \in A$ para cada $A \in \mathscr{S}$ y, por lo tanto, una función $f : x \in X \mapsto c(R(x)) \in Y$ que es tal que $f(x) = c(R(x)) \in R(x)$ para todo $x \in X$. Esto prueba la proposición. □

Muchas veces usamos esta proposición en la siguiente forma.

Corolario 2.1.3. *Si $(X_i)_{i\in I}$ es una familia de conjuntos no vacíos, entonces el producto cartesiano $\prod_{i\in I} X_i$ es él mismo un conjunto no vacío.*

Demostración. Sea $(X_i)_{i\in I}$ una familia de conjuntos no vacíos, pongamos $X \coloneqq \bigcup_{i\in I} X_i$, y consideramos la función $R : i \in I \mapsto X_i \in \mathscr{P}(X)$. Estamos entonces en la situación de la Proposición 2.1.2, así que hay una función $f : I \to X$ tal que $f(i) \in X_i$ para todo $i \in I$. La I-upla $(f(i))_{i\in I}$ es entonces un elemento del producto cartesiano $\prod_{i\in I} X_i$ y este, por lo tanto, no es vacío. □

Para nuestros fines la siguiente forma del *Axioma de elección* es particularmente conveniente.

Proposición 2.1.4. *Si $f : X \to Y$ es una función sobreyectiva, entonces hay una función $g : Y \to X$ tal que $f \circ g = \mathsf{id}_Y$ y esta función es inyectiva.*

En otras palabras: toda función sobreyectiva posee una función inversa a derecha.

Demostración. Sea $f : X \to Y$ una función sobreyectiva, de manera que para cada elemento y de Y el conjunto $f^{-1}(y)$ no es vacío. Esto nos dice que $\mathscr{S} \coloneqq \{f^{-1}(y) : y \in Y\}$ es una familia de subconjuntos no vacíos de X y el *Axioma de elección* implica entonces que hay una función $c : \mathscr{S} \to X$ tal que $c(A) \in A$ para cada $A \in \mathscr{S}$. Podemos, por lo tanto, definir la función

$$g : y \in Y \mapsto c(f^{-1}(y)) \in X,$$

y es claro que para todo $y \in Y$ es

$$(f \circ g)(y) = f(g(y)) = f(c(f^{-1}(y))) = y,$$

ya que $c(f^{-1}(y)) \in f^{-1}(y)$. Vemos así que $f \circ g = \mathsf{id}_Y$, como queremos. Finalmente, si y e y' son dos elementos de Y tales que $g(y) = g(y')$, entonces tenemos que

$$y = \mathsf{id}_Y(y) = (f \circ g)(y) = f(g(y)) = f(g(y')) = (f \circ g)(y') = \mathsf{id}_Y(y') = y',$$

y esto nos dice que la función g es inyectiva. □

Ejercicio 2.1.5. Pruebe que toda función inyectiva posee una función inversa a izquierda.

Demos ahora una aplicación especialmente importante del axioma, que es una generalización no trivial del *Principio de recurrencia* que llamamos el *Principio de elección dependiente*.

Proposición 2.1.6. *Sea X un conjunto, sea x_0 un elemento de X, y sea $\rightsquigarrow$ una relación en X. Si*

$$\text{para cada } x \in X \text{ existe } y \in X \text{ tal que } x \rightsquigarrow y,$$

entonces hay una función $f : \mathbb{N}_0 \to X$ tal que $f(0) = x_0$ y $f(n) \rightsquigarrow f(n+1)$ para cada $n \in \mathbb{N}_0$.

Notemos que aquí la relación $\rightsquigarrow$ no tiene por qué ser una relación de orden en X.

Demostración. Supongamos que vale la hipótesis, de manera que para cada $x \in X$ el subconjunto $R(x) \coloneqq \{y \in X : x \rightsquigarrow y\}$ de X no es vacío. El *Axioma de elección* nos permite concluir que hay una función $h : X \to X$ tal que $h(x) \in R(x)$ para cada $x \in X$. De acuerdo al *Principio de recurrencia*, entonces, hay una función $f : \mathbb{N}_0 \to X$ tal que $f(0) = x_0$ y $f(n+1) = h(f(n))$ para todo $n \in \mathbb{N}_0$. Si n es un elemento cualquiera de n, esto nos dice que $f(n+1) = h(f(n)) \in R(f(n))$, de manera que $f(n) \rightsquigarrow f(n+1)$ y, por lo tanto, que esta función f satisface las dos condiciones descriptas en la proposición. □

Veremos muchas aplicaciones de esta proposición. El resultado del siguiente ejercicio es una forma alternativa en la que podemos usarlo.

Ejercicio 2.1.7. Sea X un conjunto, sea x_0 un elemento de X, y sea $R : X \to \mathscr{P}(X) \setminus \{\varnothing\}$ una función. Muestre que hay una función $f : \mathbb{N}_0 \to X$ tal que $f(0) = x_0$ y $f(n+1) \in R(f(x))$ para cada $n \in \mathbb{N}_0$.

El *Principio de recurrencia* admite variaciones útiles. Una es la siguiente.

Proposición 2.1.8. *Sea X un conjunto, sea x_0 un elemento de X y sea $R : \mathbb{N}_0 \times X \to \mathscr{P}(X) \setminus \{\varnothing\}$ una función. Hay una función $f : \mathbb{N}_0 \to X$ tal que $f(0) = x_0$ y $f(n+1) \in R(n, f(x))$ para cada $n \in \mathbb{N}_0$.*

Demostración. Sea $Y \coloneqq \mathbb{N}_0 \times X$. Es claro que hay una función $S : Y \to \mathscr{P}(Y) \setminus \{\varnothing\}$ tal que

$$S((n,x)) = \{(n+1, y) : y \in R(n,x)\}$$

para cada elemento (n,x) de Y, así que el resultado del Ejercicio 2.1.7 nos dice que hay una función $g : \mathbb{N}_0 \to Y$ tal que $g(0) = (0, x_0)$ y $g(n+1) \in S(g(n))$ para cada $n \in \mathbb{N}_0$.

Para cada $k \in \mathbb{N}_0$ podemos escribir (n_k, z_k) al elemento $g(k)$ de Y y considerar la función $f : k \in \mathbb{N}_0 \mapsto z_k \in X$. Afirmamos que

$$n_k = k \text{ para todo } k \in \mathbb{N}_0.$$

Esto es claro si k es 0, ya que $(n_0, z_0) = g(0) = (0, x_0)$, así que $n_0 = 0$. Por otro lado, si k es un

elemento de $\mathbb{N}_0$ tal que $n_k = k$, entonces

$$(n_{k+1}, z_{k+1}) = g(k+1) \in S(g(k)) = S((n_k, z_k)) = S((k, z_k)) = \{(k+1, y) : y \in R(k, z_k)\},$$

de manera que $n_{k+1} = k+1$.

Como $g(0) = (0, x_0)$, es $f(0) = z_0 = x_0$. Por otro lado, para cada $k \in \mathbb{N}_0$ tenemos, gracias a lo que probamos recién, que

$$\begin{aligned}(n_{k+1}, x_{k+1}) = g(k+1) \in S(g(k)) &= S((n_k, z_k)) = S((k, f(k)) \\ &= \{(k+1, y) : y \in R(k, f(k))\},\end{aligned}$$

de manera que $x_{k+1} \in R(k, f(k))$. Vemos así que f satisface las condiciones del enunciado. □

Una segunda variación del *Principio de recurrencia*, más general, es la siguiente.

Proposición 2.1.9. *Sea X un conjunto, y sea $\mathscr{S}$ un conjunto de funciones $[\![0, n]\!] \to X$ con $n \in \mathbb{N}_0$ que satisface las siguientes dos condiciones:*

- *Hay algún elemento de $\mathscr{S}$ con dominio $[\![0, 0]\!]$.*
- *Para cada elemento $\phi : [\![0, n]\!] \to X$ de $\mathscr{S}$ existe otro elemento $\psi : [\![0, n+1]\!] \to X$ de $\mathscr{S}$ tal que $\psi|_{[\![0,n]\!]} = \phi$.*

Hay una función $f : \mathbb{N}_0 \to X$ tal que para todo $n \in \mathbb{N}_0$ es $f|_{[\![0,n]\!]} \in \mathscr{S}$.

Demostración. La segunda condición del enunciado nos dice que para cada elemento $\phi : [\![0, n]\!] \to X$ de $\mathscr{S}$ el conjunto $R(\phi)$ de los elementos de $\psi : [\![0, n+1]\!] \to X$ de $\mathscr{S}$ tal que $\psi|_{[\![0,n]\!]} = \phi$ no es vacío, y la primera que hay un elemento $\phi_0 : [\![0, 0]\!] \to X$ en $\mathscr{S}$. De acuerdo al resultado del Ejercicio 2.1.7, entonces, hay una función $F : \mathbb{N}_0 \to \mathscr{S}$ tal que $F(0) = \phi_0$ y $F(n+1) \in R(F(n))$ para todo $n \in \mathbb{N}$. Mostremos que para todo $d \in \mathbb{N}_0$ vale que

$$F(n+d)(n) = F(n)(n) \text{ para todo } n \in \mathbb{N}_0.$$

Esto es claro cuando $d = 0$ y si, por otro lado, d es un elemento de $\mathbb{N}_0$ tal que $F(n+d)(n) = F(n)(n)$ para todo $n \in \mathbb{N}_0$, entonces para cada $n \in \mathbb{N}_0$ es $F(n+d+1) \in R(F(n+d))$ y, en consecuencia,

$$F(n+d+1)(n) = F(n+d+1)|_{[\![0,n+d]\!]}(n) = F(n+d)(n) = F(n, n).$$

Consideremos la función $f : n \in \mathbb{N}_0 \mapsto F(n)(n) \in X$. Si n es un elemento de $\mathbb{N}_0$, entonces para cada $i \in [\![0, n]\!]$ es $f(i) = F(i)(i) = F(i + (n-i))(i) = F(n)(i)$: esto nos dice que $f|_{[\![0,n]\!]} = F(n) \in \mathscr{S}$. La función f satisface, por lo tanto, la condición de la proposición. □

2.2. Elementos maximales y el Lema de Zorn

Muchas veces usamos el *Axioma de elección* vía un resultado conocido como el *Lema de Zorn*. Nuestro objetivo en esta sección es explicar esto, empezando por la siguiente definición.

Definición 2.2.1. Sea X un conjunto ordenado con respecto a una relación de orden $\preceq$ y sea Y un subconjunto de X.

(*i*) Un elemento y_0 de Y es ***maximal en*** Y si no hay elementos y de Y tales que $y_0 \prec y$.

(*ii*) Un elemento x de X es una ***cota superior*** para Y si para todo elemento y de Y es $y \preceq x$, y es un ***máximo*** de Y si además pertenece a Y.

Los conceptos de elemento maximal y elemento máximo son similares pero distintos.

Ejercicio 2.2.2.

(a) Sea X un conjunto ordenado. Muestre que si un subconjunto Y de X posee un elemento máximo, entonces posee exactamente uno solo y que se trata de su único elemento maximal.

(b) Consideremos el conjunto $X \coloneqq \mathbb{N}_0 \times \mathbb{N}_0$ y sobre X la relación de orden $\preceq$ tal que para cada elección de (a, b) y (c, d) en X vale que

$$(a, b) \preceq (c, d) \iff a \leq c \wedge b \leq d$$

Muestre que el subconjunto $Y \coloneqq \{(1, 0)\} \cup \{(0, n) : n \in \mathbb{N}_0\}$ de X posee un elemento maximal que no es una cota para Y, y no posee un elemento máximo. Encuentre un subconjunto de X que no posea elementos maximales, y otro que posea 10.

(c) Muestre que un subconjunto de un conjunto ordenado puede poseer cotas superiores y no elementos maximales.

En general, los subconjuntos de un conjunto ordenado no poseen elementos maximales, y resulta que muchos problemas pueden reducirse al de encontrar elementos maximales en situaciones específicas. Hay por ello gran interés en encontrar criterios que nos permitan asegurar la existencia de elementos maximales. El *Lema de Zorn* es uno, y ya llegaremos a él, pero antes queremos presentar otros dos que son más sencillos. Todos estos resultados involucran el siguiente concepto.

Definición 2.2.3. Sea X un conjunto ordenado con respecto a una relación de orden $\preceq$. Un subconjunto C de X es una ***cadena*** si cada vez que x e y son dos elementos de Y se tiene que $x \preceq y$ o $y \preceq x$.

Veamos nuestro primer ejemplo de un criterio que garantiza la existencia de un elemento maximal para un conjunto.

Lema 2.2.4. *Sea X un conjunto ordenado con respecto a una relación de orden $\preceq$, sean n un entero positivo, sea $f : [\![n]\!] \to X$ una función, y sea I la imagen de f. Si I es una cadena de X, entonces I posee un elemento máximo.*

En el lenguaje que desarrollaremos en el siguiente capítulo, esto nos dice que *toda cadena finita de un conjunto ordenado posee un elemento máximo*.

Demostración. Probaremos la afirmación del lema procediendo por inducción con respecto al entero positivo n. Si n es 1, entonces la imagen de f es el conjunto $\{f(1)\}$ y este tiene claramente a $f(1)$ como máximo: la afirmación del lema es, por lo tanto, cierta en este caso

Supongamos ahora que m es un entero positivo, que la afirmación del lema es cierta cuando n es m, y sea $f : [\![m+1]\!] \to X$ una función cuya imagen $I := f([\![n+1]\!])$ es una cadena de X. Consideremos la restricción $g := f|_{[\![n]\!]} : [\![n]\!] \to X$. La imagen de g esta contenida en I, así que es una cadena, y entonces la hipótesis inductiva nos permite concluir que posee un elemento máximo: esto significa que hay un elemento i_0 en $[\![n]\!]$ tal que $g(i) \preceq g(i_0)$ para todo $i \in [\![n]\!]$. Ahora bien, como tanto $f(i_0)$ como $f(n+1)$ pertenecen a la imagen de f y este conjunto es una cadena, tenemos que $f(i_0) \preceq f(n+1)$ o $f(n+1) \preceq f(i_0)$: en el primer caso tenemos que $f(i) \preceq f(n+1)$ cualquiera sea i en $[\![n+1]\!]$, de manera que $f(n+1)$ es un máximo de I, y en el segundo caso tenemos que $f(i) \preceq f(i_0)$ cualquiera sea i en $[\![n+1]\!]$, de manera que $f(i_0)$ es un máximo de I. Vemos así que, en cualquier caso, la imagen I de f posee un elemento máximo, y esto completa la inducción. □

El segundo criterio que queremos presentar es una aplicación sencilla del *Principio de elección dependiente* de la sección anterior.

Proposición 2.2.5. *Sea X un conjunto ordenado con respecto a una relación de orden $\preceq$. Si no existe ninguna función $\mathbb{N}_0 \to X$ que sea estrictamente creciente, entonces todo subconjunto no vacío de X posee un elemento maximal.*

Este resultado es conocido como el *Principio de inducción noetheriana*, por Emmy Noether, aunque las ideas subyacentes a este corolario se originan en el trabajo [Hil1890] de David Hilbert sobre la teoría de invariantes, en el que las usó para probar su célebre *Teorema de la base*, que afirma que todo ideal de un anillo de polinomios con coeficientes en un cuerpo es finitamente generado.

Demostración. Sea Y un subconjunto no vacío de X, supongamos que Y no posee ningún elemento maximal, y mostremos que hay entonces una función $\mathbb{N}_0 \to X$ que es estrictamente creciente. Esto probará el corolario. La hipótesis hecha sobre Y significa que para cada elemento x de Y hay otro elemento y tal que $x \prec y$ y, de acuerdo a la Proposición 2.1.6, hay entonces una función $f : \mathbb{N}_0 \to Y$ que es estrictamente creciente. Como la función $n \in \mathbb{N}_0 \mapsto f(n) \in X$ es también estrictamente

creciente, así que esto prueba lo que queremos. □

Veamos por fin el *Lema de Zorn*, que es conocido también como el *Lema de Kuratowski–Zorn*. Kazimierz Kuratowski y Max Zorn lo probaron de manera independiente en [Kur1922] y en [Zor1935], respectivamente.

Proposición 2.2.6. *Sea X un conjunto ordenado con respecto a una relación de orden $\preceq$ que no es vacío. Si toda cadena de X posee una cota superior, entonces X posee un elemento maximal.*

Las pruebas originales de esta proposición se apoyan en el *Principio de inducción transfinita*. Pueden encontrarse distintas demostraciones basadas en esa idea en los libros de texto sobre teoría de conjuntos, como el de Peter Cameron [Cam1999], Horst Herrlich [Her2006], o Jech Tomas [Jec2003] [Jec1973]. Hay, por otro lado, una gran cantidad de demostraciones de la proposición que no se apoyan en la inducción transfinita sino que usan directamente, de diversas formas, el *Axioma de elección* — entre ellas podemos mencionar las que de Hellmuth Kneser [Kne1950], Tibor Szele [Sze1950], Jeffrey Dennis Weston [Wes1957], Paul Halmos [Hal1960], James Dugundji [Dug1966], Jin-iti Nagata [Nag1985], Herman Rubin y Jean Rubin [RR1985], Jonathan Lewin [Lew1991], Serge Lang [Lan2002], George Bergman [Ber1997], Daniel Grayson [Gra2007], Arjun Jain [Jai2012], Koji Nuida [Nui2024], Guillermo Incatasciato y Pedro Sánchez Terraf [IT2024]. Es posible probar la proposición usando en lugar del *Axioma de elección* alguna de sus consecuencias: por ejemplo, la obtienen a partir del *Principio maximal de Hausdorff* Frink Orrin [Fri1952], Walter Rudin [Rud1987], Kenneth Kuttler [Kut1998], Kenneth Brown [Bro2010], Lektor Lektorean [Lek2010], Nico Spronk [Spr2017], y a partir de un teorema sobre la existencia de puntos fijos Nicolas Bourbaki [Bou1949], Claude Berge [Ber1963], Arlen Brown [Bro1970], Serge Lang [Lan1993], Mohssin Zarouali-Darkaoui [ZD2019], y Jerry Shurman [Shu2021].

La siguiente demostración es esencialmente la que dan Hellmuth Kneser en [Kne1950] y Tibor Szele en [Sze1950].

Demostración. Supongamos que X satisface la hipótesis de la proposición y, para llegar a una contradicción, que X no posee elementos maximales.

Primer paso. Escribamos $\mathscr{C}$ al conjunto de todas las cadenas de X y para cada subconjunto A de X y cada elemento a de A escribamos

$$A_a \coloneqq \{x \in A : x \prec a\}.$$

Si C es una cadena en X, entonces el conjunto $U(C)$ de los elementos de $X \setminus C$ que son cotas superiores de C en X no es vacío. En efecto, la hipótesis que hicimos sobre X implica que hay en X una cota superior x para C y que x no es maximal en X, así que hay otro elemento y en X tal que $x \prec y$ y claramente $d \in U(C)$. Tenemos entonces una función

$$U : C \in \mathscr{C} \mapsto U(C) \in \mathscr{P}(X) \setminus \{\varnothing\}$$

y el *Axioma de elección* nos dice que hay una función $\phi : \mathscr{C} \to X$ tal que $\phi(C) \in U(C)$ para cada elemento C de $\mathscr{C}$. Notemos que

para toda cadena C es $\phi(C) \notin C$.

Segundo paso. Digamos que un subconjunto C en X es una ***ϕ-cadena*** si las siguientes dos condiciones se cumplen:

- C es una cadena en X.
- No hay una función $f : \mathbb{N}_0 \to C$ que sea estrictamente decreciente.
- Es $\phi(C_c) = c$ para todo $c \in C$.

Es claro que $\varnothing$ es una ϕ-cadena. Por otro lado, es muy fácil verificar que

si C es una ϕ-cadena, entonces también lo es $C \cup \{\phi(C)\}$.

Probemos que

si C y D son dos ϕ-cadenas distintas, entonces o bien hay un elemento c de C tal que $D = C_c$ o bien hay un elemento d de D tal que $C = D_d$. (2.1)

Sean para ello C y D dos ϕ-cadenas distintas y consideremos el conjunto

$$E \coloneqq \{x \in C \cap D : C_x = D_x\}.$$

Notemos que si x es un elemento de E, entonces para cada $y \in C_x$ es $y \in C_x = C_x \cap D_x \subseteq C \cap D$ y $C_y = (C_x)_y = (D_x)_y = D_x$. Esto nos dice que

si x es un elemento de E, entonces $C_x = D_x \subseteq E$.

Queremos ahora probar que

o bien $E = C$ o bien $E = C_c$ con $c \coloneqq \min E \setminus C$. (2.2)

Para ello, supongamos que $E \neq C$. El conjunto ordenado $C \setminus E$ no es entonces vacío y, como no hay funciones estrictamente decrecientes $\mathbb{N}_0 \to C \setminus E$, la Proposición 2.2.5 implica que podemos considerar su menor elemento $c \coloneqq \min C \setminus E$. Si x es un elemento de C_c, de manera que $x \in C$ y $x \prec c = \min C \setminus E$, entonces claramente $x \in E$: esto nos dice que $C_c \subseteq E$. Supongamos que hay un elemento x en $E \setminus C_c$. Como $x \in E$, es $x \in C$, $C_x = D_x$, y $x \neq c$, ya que $c \notin E$; por otro lado, como $x \in C$, $x \notin C_c$ y C es una cadena, tenemos que $c \prec x$, de manera que $c \in C_x \subseteq E$, y esto es absurdo. Vemos así que el conjunto $E \setminus C_c$ es vacío y, por lo tanto, que $E = C_c$.

De manera completamente similar podemos probar que también

o bien $E = D$ o bien $E = D_d$ con $c \coloneqq \min E \setminus D$. (2.3)

Además de esto, es cierto que

$$\text{o bien } E = C \text{ o bien } E = D. \tag{2.4}$$

En efecto, si no fuera ese el caso, entonces en vista de (4.2) y (4.3) tendríamos que $E = C_c$ y que $E = D_d$, de manera que $c = \phi(C_c) = \phi(D_d) = d$. Pero esto nos dice que $c \in C \cap D$ y, como $C_c = E = D_c$, que $c \in E$, y esto es imposible ya que $c \in C \setminus E$. Finalmente, combinando (4.2), (4.3) y (4.3) vemos que, en definitiva, o bien $C = E = D_d \subseteq D$ o bien $C \supseteq C_c = E = D$, y entonces la afirmación (4.1) es cierta.

Tercer paso. Sea ahora F la unión de todas las ϕ-cadenas de X, y mostremos que

$$\text{si } f \text{ es un elemento de } F \text{ y } C \text{ es una } \phi\text{-cadena tal que } f \in C, \text{ entonces } C_f = F_f.$$

Para ello, sea f un elemento de F y sea C una ϕ-cadena tal que $f \in C$. Como $C \subseteq F$ es claro que $C_f \subseteq F_f$. Sea g un elemento de F_f. Como $g \in F$, hay una ϕ-cadena D tal que $g \in D$ y de (4.1) sabemos que o bien $D \subseteq C$, y en ese caso es claro que $g \in C_f$, o bien hay un elemento h de D tal que $C = D_h$, y en ese caso $g \prec f$ y $f \prec h$, por lo que $g \in D_h = C$ y $g \in C_f$. Vemos así que $F_g = C_f$.

Cuarto y último paso. Probemos ahora que F es una ϕ-cadena.

- Sean f y g dos elementos de F. Hay entonces dos ϕ-cadenas C y D tales que $f \in C$ y $g \in D$ y, de acuerdo a (4.1), es $C \subseteq D$ o $D \subseteq C$. En cualquiera de los dos cases, tenemos que $C \cup D$ es una cadena y, como f y g pertenecen los dos a ella, que $f \preceq g$ o $g \preceq f$. Esto nos dice que F es un subconjunto totalmente ordenado de X, esto es, una cadena.
- Supongamos que $f : \mathbb{N}_0 \to F$ es una función estrictamente decreciente. Como $f(0) \in F$, hay una ϕ-cadena C tal que $f(0) \in C$. Como la función f es estrictamente decreciente, para cada $n \in \mathbb{N}$ es $f(n) \in F_{f(0)} = C_{f(0)} \subseteq C$, y entonces hay una función $n \in \mathbb{N}_0 \mapsto f(n+1) \in C$. Esto es absurdo, ya que esta función es claramente estrictamente decreciente.
- Si f es un elemento de F, entonces hay una ϕ-cadena C tal que $f \in C$ y es $\phi(F_c) = \phi(C_f) = f$.

Ahora bien, como F es una ϕ-cadena, también lo es $F' \coloneqq F \cup \{\phi(F)\}$, como observamos arriba: esto no es posible, ya que F contiene a todas las ϕ-cadenas y ciertamente $\phi(F) \notin F$. Esta es la contradicción que buscábamos. □

2.3. Algunas aplicaciones

El objetivo de esta sección es presentar algunas aplicaciones sencillas del *Lema de Zorn* para mostrar cómo lo usamos generalmente en la práctica.

Ejemplo 2.3.1. Sea $\Bbbk$ un cuerpo. Queremos mostrar que

todo espacio vectorial sobre $\Bbbk$ posee una base.

Sea para ello V un espacio vectorial sobre $\Bbbk$ y escribamos X al conjunto de todos los subconjuntos linealmente independientes de V. Consideramos a X como un conjunto ordenado con respecto a la relación de orden $\subseteq$ dada por la inclusión. Queremos probar que hay en X un elemento maximal y usaremos para ello el *Lema de Zorn*. Es claro que el conjunto X no es vacío, ya que $\varnothing \in X$. Tenemos entonces solamente que mostrar que toda cadena de X posee una cota superior.

Sea entonces $\mathscr{C}$ una cadena de X y escribamos

$$\mathscr{C}^+ \coloneqq \bigcup_{C \in \mathscr{C}} C$$

a su unión. Es evidente que $C \subseteq \mathscr{C}^+$ para cada $C \in \mathscr{C}$, así que para ver que $\mathscr{C}^+$ es una cota superior para C en X es suficiente que mostremos que pertenece a C, esto es, que se trata de un subconjunto linealmente independiente de V. Sea n un entero positivo, sean $v_1, v_2, \ldots, v_n$ elementos dos a dos distintos de $\mathscr{C}^+$, sean $a_1, a_2, \ldots, a_n$ escalares de $\Bbbk$, y supongamos que

$$a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = 0.$$

Como los vectores $v_1, v_2, \ldots, v_n$ están en $\mathscr{C}^+$, la definición de este conjunto implica que hay elementos $C_1, C_2, \ldots, C_n$ de $\mathscr{C}$ tales que $v_1 \in C_1$, $v_2 \in C_2$, $\ldots$, $v_n \in C_n$. Como la imagen de la función $f : i \in [\![n]\!] \mapsto C_i \in X$ es una cadena, ya que está contenida en $\mathscr{C}$, sabemos del Lema 2.2.4 que esa imagen posee un elemento máximo: esto nos dice que hay un elemento j de $[\![n]\!]$ tal que $C_i \subseteq C_j$ para cada $i \in [\![n]\!]$. En particular, los vectores $v_1, v_2, \ldots, v_n$ pertenecen todos a C_j y, como el conjunto C_j es linealmente independiente, esto nos permite concluir que los escalares $a_1, a_2, \ldots, a_n$ son todos nulos. Esto prueba que el conjunto $\mathscr{C}^+$ es un elemento de X, como queremos.

Hemos probado con todo esto que el conjunto ordenado X satisface las condiciones del *Lema de Zorn* 2.2.6, así que posee un elemento maximal. Escribámoslo B. Como sabemos del Álgebra Lineal que un subconjunto linealmente independiente maximal de V es necesariamente una base de V, esto completa la prueba de nuestra afirmación de arriba.

Como en el ejemplo, con frecuencia el siguiente caso particular del *Lema de Zorn* es suficiente para nuestros fines.

Ejercicio 2.3.2. Sea X un conjunto y sea $\mathscr{S}$ un subconjunto no vacío de $\mathscr{P}(X)$, al que consideramos ordenado por la relación de inclusión $\subseteq$. Si $\mathscr{S}$ tal que

$$\textit{cada vez que } \mathscr{C} \textit{ es una cadena no vacía de } \mathscr{S} \textit{ la unión } \bigcup_{C\in\mathscr{C}} C \textit{ pertenece a } \mathscr{S}, \tag{2.5}$$

entonces $\mathscr{S}$ posee elementos maximales.

Demos un segundo ejemplo en el que usamos otra vez el caso particular del *Lema de Zorn* descripto en este ejercicio.

Ejemplo 2.3.3. Sea X un conjunto y mostremos que

toda relación de orden en X está contenida en una relación de orden total en X.

Sea R una relación de orden en X y sea $\mathscr{O}$ el conjunto de todas las relaciones de orden en X que contienen a R. Un orden en X es un subconjunto de $X \times X$, así que podemos ver a $\mathscr{O}$ como un conjunto ordenado con respecto a la relación de orden. Como $R \in \mathscr{O}$, es $\mathscr{O} \neq \varnothing$. Sea, por otro lado, $\mathscr{C}$ una cadena no vacía de $\mathscr{O}$ y sea $T \coloneqq \bigcup_{C\in\mathscr{C}} C$, que es ciertamente un subconjunto de $X \times X$. Mostremos que se trata de un orden en X — como claramente contiene a R, es entonces un elemento de $\mathscr{O}$.

- Sea x un elemento de X. Como $\mathscr{C}$ no es vacía, hay un elemento S en $\mathscr{C}$ y, como S es una relación de orden en X, tenemos que $(x, x) \in S \subseteq T$. Esto nos dice que T es una relación reflexiva.
- Sean x e y dos elementos de X y supongamos que $(x, y) \in T$ y que $(y, x) \in T$, de manera que hay dos elementos C y D de $\mathscr{C}$ tales que $(x, y) \in C$ e $(y, x) \in D$. Como $\mathscr{C}$ es una cadena, la unión $C \cup D$ coincide con uno de C o D y es, en particular, un orden sobre X: como los pares (x, y) e (y, x) están los dos en $C \cup D$, se sigue de eso entonces que $x = y$. La relación T es, por lo tanto, anti-simétrica.
- Sean finalmente x, y y z tres elementos de X tales que los pares (x, y) e (y, z) están en T, de manera que hay dos elementos C y D de $\mathscr{C}$ tales que $(x, y) \in C$ e $(y, z) \in D$. Como $\mathscr{C}$ es una cadena, la unión $C \cup D$ coincide con uno de C o D y es un orden sobre X: como los pares (x, y) e (y, z) están en $C \cup D$, podemos concluir que $(x, z) \in C \cup D \subseteq T$. Vemos así que la relación T es transitiva.

Con todo esto hemos probado que estamos en la situación del Ejercicio 2.3.2 y que, por lo tanto, hay en $\mathscr{O}$ un elemento maximal S. Se trata de una relación de orden en X que contiene a R, y para completar la prueba de nuestra afirmación de arriba tenemos que mostrar que se trata de un orden total en X.

Supongamos que no lo es, de manera que hay dos elementos x e y en X tales que $(x, y) \notin S$

e $(y, x) \notin S$, de manera que, en particular, es $x \neq y$, y consideremos el conjunto

$$S' = S \cup \{(z, w) \in X \times X : (z, x) \in S, (y, w) \in S\}.$$

Como $(x, y) \in S'$, tenemos que $S \subsetneq S'$. Para llegar a una contradicción y terminar la prueba es suficiente entonces que mostremos que S' es una relación de orden sobre el conjunto X.

- Si u es un elemento de X, entonces $(u, u) \in S \subseteq S'$.
- Sean u y v dos elementos de X tales que los pares (u, v) y (v, u) están en S'. Supongamos por un momento que uno de esos dos pares no está en S y, sin pérdida de generalidad, que de hecho $(u, v) \notin S$, de manera que $(u, x) \in S$ e $(y, v) \in S$. Esto implica que también $(v, u) \notin S$, ya que en caso contrario la transitividad de S nos diría que $(y, x) \in S$, contra la hipótesis, así que es $(v, x) \in S$ e $(y, u) \in S$: esto es imposible, ya que entonces (y, u) y (u, x) están en S y, por lo tanto, (y, x) también. Esto nos dice que la relación S' es anti-simétrica.
- Sean, finalmente, u, v y w tres elementos de X tales que los pares (u, v) y (v, w) está en S'.
  - Si los dos pares están en S, entonces ciertamente es $(u, w) \in S \subseteq S'$.
  - Si solo el primer par está en S, entonces la definición de S' implica que $(v, x) \in S$ e $(y, w) \in S$, de manera que la transitividad de S nos permite concluir que $(u, w) \in S$.
  - Si solo el segundo par está en S, entonces tenemos que $(u, x) \in S$, $(y, v) \in S$ y $(v, w) \in S$, así que $(u, w) \in S$.
  - Finalmente, si ninguno de los dos pares está en S, los cuatro pares (u, x), (y, v), (v, x) y (y, w) están en S, así que $(u, w) \in S'$.

  En cualquier caso, entonces, el par (u, w) pertenece a S' y, por lo tanto, la relación S' es transitiva.

Como dijimos, esto completa la prueba de la afirmación (2.5).

Si R es una relación de orden en un conjunto X, llamamos a todo orden total L en X tal que $R \subseteq L$ una ***extensión lineal*** de R. La afirmación que probamos en el Ejemplo 2.3.3 nos dice que todo orden posee extensiones lineales, y el siguiente ejercicio que todo orden está, de hecho, completamente determinado por sus órdenes lineales.

Ejercicio 2.3.4. Sea X un conjunto, sea R una relación de orden sobre X, y sea $\mathscr{L}$ el conjunto de todos los ordenes totales sobre X que contienen a R. Pruebe que $R = \bigcap_{L \in \mathscr{L}} L$.

El propósito del siguiente ejercicio es mostrar que el *Lema de Zorn* implica al *Axioma de elección*.

Ejercicio 2.3.5. Sea X un conjunto y sea $\mathscr{S}$ una familia de subconjuntos no vacíos de X. Sea $\mathscr{P}$ el conjunto de todas las funciones $s : A \to X$ definidas en un subconjunto A de $\mathscr{S}$ tales que $s(a) \in a$

para cada $a \in A$. Si $s : A \to X$ y $t : B \to X$ son dos elementos de $\mathscr{P}$, entonces escribimos $s \preceq t$ si $A \subseteq B$ y $t|_A = s$.

(a) Muestre que $\preceq$ es una relación de orden sobre el conjunto $\mathscr{P}$, que $\mathscr{P}$ no es vacío y que toda cadena de $\mathscr{P}$ está acotada, de manera que el *Lema de Zorn* implica que hay en $\mathscr{P}$ elementos maximales.

(b) Muestre que si $c : A \to X$ es un elemento maximal de $\mathscr{P}$, entonces necesariamente $A = \mathscr{S}$, de manera que c es una función de elección para la familia $\mathscr{A}$.

A veces podemos reemplazar el *Lema de Zorn* por el siguiente resultado, que es conocido como el *Teorema de punto fijo de Bourbaki–Witt*, por Nicolas Bourbaki y Ernst Witt, quienes lo probaron en [Bou1949] y en [Wit1951].

Proposición 2.3.6. *Sea X un conjunto ordenado con respecto a una relación de orden $\preceq$ que tiene un elemento mínimo $\perp$ y en el que toda cadena posee un supremo. Una función $s : X \to X$ tal que $x \preceq s(x)$ para todo $x \in X$ posee un punto fijo, esto es, hay un elemento p de X tal que $s(p) = p$.*

Demostración. Sea $s : X \to X$ una función tal que $x \preceq s(x)$ para todo $x \in X$. Decimos que un subconjunto I de X es ***s-inductivo*** si $\perp \in A$, $s(I) \subseteq I$ e I contiene el supremo de toda cadena de X contenida en I. Claramente hay subconjuntos s-inductivos en X, como X mismo, así que podemos considerar la intersección $\mathscr{I}$ de todos los subconjuntos s-inductivos de X. Es fácil ver que $\mathscr{I}$ es un subconjunto s-inductivo de X, y claramente $\mathscr{I}$ no es vacío, ya que contiene a $\perp$.

Primer paso. Digamos que un elemento c de $\mathscr{I}$ es ***crítico*** si para todo $x \in \mathscr{I}$ vale que

$$x \prec c \implies s(x) \preceq c,$$

y mostremos que cuando ese es el caso se tiene que

$$\text{el conjunto } \mathscr{I}_c \coloneqq \{x \in \mathscr{I} : x \preceq c \vee s(c) \preceq x\} \text{ concide con } \mathscr{I}. \tag{2.6}$$

Sea para eso c un elemento crítico de $\mathscr{I}$. Claramente $\mathscr{I}_c \subseteq \mathscr{I}$, y para probar que $\mathscr{I} \subseteq \mathscr{I}_c$ es suficiente que mostremos que $\mathscr{I}_c$ es un subconjunto s-inductivo de X.

- Como $\perp \preceq c$, es $\perp \in \mathscr{I}_c$.
- Sea x es un elemento de $\mathscr{I}_c$, de manera que $x \prec c$, o $x = c$ o $s(c) \preceq x$. Si $x \prec c$, entonces $s(x) \preceq c$ porque c es crítico y entonces $s(x) \in \mathscr{I}_c$. Si $x = c$, entonces $s(c) \preceq s(c) = s(x)$, así que $s(x) \in \mathscr{I}_c$. Finalmente, si $s(c) \preceq x$, entonces $s(c) \preceq x \preceq s(x)$ y, otra vez, $s(x) \in \mathscr{I}_c$. En cualquiera de los tres casos $s(x)$ está en $\mathscr{I}_c$ y podemos concluir de esto que $s(\mathscr{I}_c) \subseteq \mathscr{I}_c$.
- Sea D una cadena de X contenida en $\mathscr{I}_c$ y sea d el supremo de C en X. Como $\mathscr{I}_c$ está contenido en $\mathscr{I}$, D está contenido en $\mathscr{I}$ y que $\mathscr{I}$ sea s-inductivo implica que $d \in \mathscr{I}$. Ahora bien, si c es una cota superior para D en X, entonces $d \preceq c$ y, por lo tanto, $d \in \mathscr{I}_c$. Si, por el

contrario, c no es una cota superior para D, entonces que D esté contenido en $\mathscr{I}_c$ implica que hay un elemento e de D tal que $s(c) \preceq e \preceq d$, así que también es $d \in \mathscr{I}_c$. Vemos con esto que el conjunto $\mathscr{I}_c$ contiene el supremo en X de toda cadena de X contenida en él.

Esto completa la demostración de (2.6).

Segundo paso. Mostremos ahora que, de hecho, todo elemento de $\mathscr{I}$ es crítico.

Sea C el conjunto de todos los elementos críticos de $\mathscr{I}$, que, por supuesto, está contenido en $\mathscr{I}$. Para probar que $C = \mathscr{I}$, que es lo que queremos, es suficiente entonces que probemos que C es un subconjunto s-inductivo de X, ya que en ese caso tendremos que está contenido en $\mathscr{I}$.

- Como $\bot$ es claramente un elemento crítico de $\mathscr{I}$, tenemos que $\bot \in E$.
- Sea c un elemento de E. Queremos ver que $s(c) \in E$, esto es, que $s(x)$ es crítico, y para ello consideramos un elemento x de $\mathscr{I}$ tal que $x \prec s(c)$ y mostramos que $s(x) \preceq s(c)$. Como $x \in \mathscr{I} = \mathscr{I}_c$, sabemos que $x \prec c$, $x = c$ o $s(c) \preceq x$. La tercera de estas alternativas no ocurre, porque elegimos a x de manera que sea $x \prec s(c)$. Si $x \prec c$, entonces $s(x) \preceq c \preceq s(c)$, ya que c es crítico. Si, en cambio, $x = c$, entonces $s(x) \preceq s(c)$. Vemos así que en cualquier caso es $s(c) \preceq s(c)$, como queremos.
- Sea D una cadena de X contenida en E y sea d el supremo de D en X. Tenemos que probar que d pertenece a E, esto es, que es crítico. Como $D \subseteq E \subseteq \mathscr{I}$ e $\mathscr{I}$ es s-inductivo, sabemos que $d \in \mathscr{I}$. Sea, por otro lado, x un elemento de $\mathscr{I}$ tal que $x \prec d$ y mostremos que $s(x) \preceq d$. No puede ser que x sea una cota superior para $s(D)$: en ese caso tendríamos que para todo $y \in D$ es $y \preceq s(y) \preceq x$, de manera que x es una cota superior para D y, por lo tanto, $d \preceq x$, contra la forma en que elegimos a x. Vemos así que hay un elemento e de D tal que $x \prec s(e)$.

  Como $x \in \mathscr{I} = \mathscr{I}_e$, tenemos que $x \preceq e$ o $s(e) \preceq x$, y la segunda alternativa no ocurre. Si $x \prec e$, entonces $s(x) \preceq e \preceq s(e) \preceq d$, porque e es crítico. Si, en cambio, $x = e$, entonces $s(x) = s(e) \preceq s(e) \preceq d$. Así, es $s(x) \preceq d$ en cualquier caso.

Tercer y último paso. El conjunto $\mathscr{I}$ está totalmente ordenado. En efecto, si x e y son dos elementos de $\mathscr{I}$, entonces x es crítico por lo que hicimos en el Segundo paso e $y \in \mathscr{I}_x$ por lo que hicimos en el Primer paso, así que $y \preceq x$ o $x \prec s(x) \preceq y$.

La hipótesis que hicimos sobre el conjunto ordenado X nos dice que $\mathscr{I}$ posee entonces un supremo en C. Escribámoslo m. Que $\mathscr{I}$ sea a la vez un conjunto s-inductivo y una cadena implica que, de hecho, $m \in I$. Como I es admisible, se sigue de esto que $s(m) \in \mathscr{I}$, así que $m \preceq s(m) \preceq m$, porque m es una cota superior para $\mathscr{I}$. Vemos así que el elemento m es un punto fijo de la función s, y esto prueba la proposición. □

El *Teorema de punto fijo de Bourbaki–Witt* tiene como corolario casi inmediato al siguiente *Teorema de punto fijo de Kuratowski*, por Kazimierz Kuratowski, que lo probó en [Kur1922].

Corolario 2.3.7. *Sea X un conjunto y sea $\mathscr{F}$ un subconjunto no vacío de $\mathscr{P}(X)$ tal que*

$$A \in \mathscr{F} \implies \bigcup_{a \in A} a \in \mathscr{F}.$$

Toda función $f : \mathscr{F} \to \mathscr{F}$ tal que $A \subseteq f(A)$ para todo elemento A de $\mathscr{F}$ posee un punto fijo.

Demostración. Sea $f : \mathscr{F} \to \mathscr{F}$ una función tal que $A \subseteq f(A)$ para todo $A \in \mathscr{F}$, sea A_0 un elemento cualquiera de $\mathscr{F}$, sea $\mathscr{F}' \coloneqq \{A \in \mathscr{F} : A_0 \subseteq A\}$, Consideremos a $\mathscr{F}'$ como un conjunto ordenado con respecto a la relación de inclusión. Es claro que $A_0 \in \mathscr{F}'$ y que, de hecho, A_0 es el mínimo elemento de $\mathscr{F}'$. Por otro lado, para todo $A \in \mathscr{F}'$ tenemos que $A_0 \subseteq A \subseteq f(A)$, así que $f(A) \in \mathscr{F}'$: esto nos dice que podemos considerar la restricción $f' : A \in \mathscr{F}' \mapsto f(A) \in \mathscr{F}'$ de f a $\mathscr{F}'$. Es claro que para todo $A \in \mathscr{F}'$ es $A \subseteq f'(A)$, así que para probar que la función f' y, en consecuencia, la función f posee un punto fijo es suficiente que mostremos que toda cadena en $\mathscr{F}'$ posee un supremo, ya que en ese caso podremos usar el *Teorema de punto fijo de Bourbaki–Witt* para concluir lo que queremos.

Sea C una cadena de $\mathscr{F}'$ y sea $D \coloneqq \bigcup_{c \in C} c$. La hipótesis sobre $\mathscr{F}$ implica que $D \in \mathscr{F}$, y es claro que $A_0 \subseteq D$, así que $D \in \mathscr{F}'$. Es evidente que $c \subseteq D$ para todo $c \in C$, así que D es una cota superior para C en $\mathscr{F}'$. Por otro lado, si E es una cota superior para C en $\mathscr{F}'$, entonces $c \subseteq E$ para todo $c \in C$ y, por lo tanto, $D = \bigcup_{c \in C} c \subseteq E$. Vemos así que D es el supremo de C en $\mathscr{F}'$, y esto completa la demostración del corolario. □

Una aplicación importante del *Teorema de punto fijo de Bourbaki–Witt* es el siguiente resultado, conocido como el *Principio maximal de Hausdorff*, por Felix Hausdorff, que lo probó en 1914 [Hau1914].

Corolario 2.3.8. *Toda cadena de un conjunto ordenado está contenida en una cadena maximal.*

Como $\varnothing$ es una cadena de cualquier conjunto ordenado, esto nos dice, en particular, que todo conjunto ordenado posee cadenas maximales.

Demostración. Sea X un conjunto ordenado, sea C_0 una cadena en X, sea $\mathscr{C}$ el conjunto de todas las cadenas de X que contienen a C_0, y consideremos a $\mathscr{C}$ como un conjunto ordenado con respecto a la relación de inclusión. Notemos que $\mathscr{C} \neq \varnothing$, ya que $C_0 \in \mathscr{C}$, y que, de hecho, C_0 es el mínimo elemento de $\mathscr{C}$.

Sea C una cadena en $\mathscr{C}$ y sea $\overline{C} \coloneqq \bigcup_{c \in C} c$ la unión de los elementos de C. Si x e y son dos elementos de $\overline{C}$, entonces hay dos elementos c y d de C tales que $x \in c$ e $y \in d$ y, como C es una cadena, $c \cup d$ coincide con alguno de c o d y es, en particular, una cadena de X: como x e y están los dos en $c \cup d$, esto implica que $x \preceq y$ o $y \preceq x$. Vemos así que $\overline{C}$ es una cadena de X y, como claramente $C_0 \subseteq c \subseteq \overline{C}$ para todo $c \in C$, que se trata de una cota superior para C en $\mathscr{C}$. Es, de

hecho, el supremo de C en $\mathscr{C}$: en efecto, si D es una cota superior para C en $\mathscr{C}$, entonces $c \subseteq D$ para todo $c \in C$, así que $\overline{C} = \bigcup_{c\in C} c \subseteq D$. Esto prueba que toda cadena en $\mathscr{C}$ posee un supremo.

El *Axioma de elección* nos dice que hay una función de elección $\phi : \mathscr{P}(X) \setminus \{\varnothing\} \to X$, de manera que $\phi(A) \in A$ para cada $A \in \mathscr{P}(X) \setminus \{\varnothing\}$. Si C es un elemento de $\mathscr{C}$ que no es maximal en $\mathscr{C}$, entonces ciertamente el conjunto C^* de todos los elementos y de $X \setminus C$ tales que $C \cup \{y\}$ es una cadena no es vacío y $C \cup \{\phi(C^*)\}$ es una cadena de X que contiene a C_0 y distinta de C. Esto implica que hay una función $s : \mathscr{C} \to \mathscr{C}$ que en cada cadena C de X que contiene a C_0 toma el valor

$$s(C) = \begin{cases} C & \text{si } C \text{ es una cadena maximal;} \\ C \cup \{\phi(C^*)\} & \text{si no lo es.} \end{cases}$$

Es evidente que $C \subseteq s(C)$ para todo elemento C de $\mathscr{C}$, así que el *Teorema de punto fijo de Bourbaki–Witt* 9.2.6 nos dice que hay un elemento C de $\mathscr{C}$ tal que $s(C) = C$. Claramente C es una cadena maximal de X que contiene a C_0. □

Es interesante que este corolario tiene una versión *dual* que es importante.

Definición 2.3.9. Decimos que un subconjunto A de un conjunto ordenado X es una ***anticadena*** si cada vez que x e y son dos elementos distintos de A se tiene que $x \not\leq y$ e $y \not\leq x$.

Así, un subconjunto de un conjunto ordenado es una anticadena cuando no posee dos elementos distintos comparables. Esto es exactamente lo opuesto que una cadena: todos los pares de elementos de una cadena son comparables.

Proposición 2.3.10. *Toda anticadena de un conjunto ordenado está contenida en una anticadena maximal.*

Demostración. Sea X un conjunto ordenado, sea A_0 una anticadena de X, y sea $\mathscr{A}$ el conjunto de todas las anticadenas de X que contienen a A_0 ordenado con respecto a la relación de inclusión. Como $A_0 \in \mathscr{A}$, es $\mathscr{A} \neq \varnothing$. Sea, por otro lado, C una cadena en $\mathscr{A}$ y sea $C^+ \coloneqq \bigcup_{c\in C} c$ la unión de los elementos de C. Si x e y son dos elementos de C^+, entonces hay elementos c y d de C tales que $x \in c$ e $y \in d$ y, como C es una cadena, la unión $c \cup d$ es un elemento de C: esta unión es entonces una anticadena y x e y pertenecen a ella, así que $x \not\leq y$ e $y \not\leq x$. Vemos así que C^+ es un elemento de $\mathscr{A}$, ya que claramente contiene a A_0.

Todo esto nos dice que estamos en la situación del Ejercicio 2.3.2, así que hay en $\mathscr{A}$ elementos maximales. Esto significa que hay en X anticadenas maximales que contienen a A_0, y prueba el corolario. □

Capítulo 3

Cardinalidad

En este capítulo presentaremos la noción de cardinalidad, cuyo objetivo fundamental es el de permitirnos hacer distinciones entre conjuntos de acuerdo a una idea muy precisa de *tamaño*.

3.1. La relación de equipotencia

En la base de todo lo que hacemos está la siguiente relación entre conjuntos.

Definición 3.1.1. Decimos que un conjunto X es ***equipotente*** a un conjunto Y si hay una función biyectiva $X \to Y$, y cuando ese es el caso escribimos $X \sim Y$.

Llamamos a esta relación $\sim$ sobre la clase de todos los conjuntos la relación de ***equipotencia***.

Ejercicio 3.1.2. Sean X_1, X_2, Y_1 e Y_2 conjuntos. Si $X_1 \sim Y_1$, $X_2 \sim Y_2$ y $X_1 \cap X_2 = Y_1 \cap Y_2 = \varnothing$, entonces $(X_1 \cup X_2) \sim (Y_1 \cup Y_2)$.

La propiedad más importante de la relación de equipotencia es la siguiente:

Proposición 3.1.3. *La relación de equipotencia es una relación de equivalencia.*

Demostración. Tenemos que mostrar que se trata de una relación reflexiva, simétrica y transitiva.

- Cualquiera sea el conjunto X, la función identidad $\mathsf{id}_X : X \to X$ es biyectiva y, por lo tanto, tenemos que $X \sim X$. Esto nos dice que la relación de equipotencia es reflexiva.

- Sean X e Y dos conjuntos y supongamos que $X \sim Y$, de manera que hay una función biyectiva $f : X \to Y$. Como esta función es biyectiva, es inversible y admite, por lo tanto, una función inversa $g : Y \to X$. Por supuesto, esta función g tiene a f por función inversa, así que es inversible y, en consecuencia, biyectiva: vemos así que $Y \sim X$ y, en definitiva, que la relación de equipotencia es simétrica.
- Sean, finalmente, X, Y y Z tres conjuntos, y supongamos que $X \sim Y$ e $Y \sim Z$, de manera que hay funciones biyectivas $f : X \to Y$ y $g : Y \to Z$. La composición $g \circ f : X \to Z$ es entonces también biyectiva y, por lo tanto, tenemos que $X \sim Z$. Esto prueba que la relación de equipotencia es transitiva.

Juntando estas tres observaciones, como dijimos, podemos concluir que la relación de equipotencia es una relación de equivalencia, que es lo que afirma la proposición. □

Gracias a esta proposición sabemos que la clase de todos los conjuntos queda partida en clases de equivalencia con respecto a la relación de equipotencia, y esa partición es precisamente la que realiza las distinciones entre conjuntos que queremos estudiar.

Definición 3.1.4. El ***cardinal*** $|X|$ de un conjunto X es su clase de equivalencia con respecto a la relación de equipotencia, de manera que

$$|X| \coloneqq \{Y : X \sim Y\}.$$

Esta definición nos permite muchas veces distinguir entre conjuntos: así, por ejemplo, es claro que una forma de probar que dos conjuntos son distintos es mostrar que tienen cardinales distintos o, equivalentemente, que no son equipotentes — veremos instancias concretas de esto más abajo. Es claro, de todas maneras, que para que esto sea una estrategia viable deberemos desarrollar formas de *calcular* el cardinal de un conjunto.

Antes de ocuparnos de eso, estableceremos algunos resultados importantes que nos permitirán operar con la noción de cardinalidad. Nuestro primer objetivo es construir una relación de orden entre los cardinales.

Definición 3.1.5. Sean X e Y dos conjuntos. Decimos que el cardinal de X es ***menor o igual*** que el cardinal de Y si hay una función inyectiva $X \to Y$, y en ese caso escribimos $X \preceq Y$. Si además los conjuntos X e Y no son equipotentes, podemos escribir $X \prec Y$ para dejarlo de manifiesto.

Cuando tenemos dos conjuntos X e Y tales que $X \preceq Y$, de manera que hay una función inyectiva $f : X \to Y$, la correstricción $f|^I : X \to f(X)$ es biyectiva y, por lo tanto, los conjuntos X y $f(X)$ son equipotentes: así, en esta situación Y contiene un subconjunto del «mismo tamaño» que X, así que podemos pensar que X es «menor o igual» que Y. Esta es la idea que motiva la definición que acabamos de dar.

Podemos dar una descripción alternativa de la relación $\preceq$ usando funciones sobreyectivas.

Proposición 3.1.6. *Sean X e Y dos conjuntos.*
(*i*) *Si hay una función sobreyectiva $Y \to X$, entonces $X \preceq Y$.*
(*ii*) *Si $X \preceq Y$ y el conjunto X no es vacío, entonces hay una función sobreyectiva $Y \to X$.*

Si el conjunto X es vacío y el conjunto Y no lo es, entonces ciertamente vale que $X \preceq Y$ pero no hay ninguna función, sobreyectiva o no, $Y \to X$. Es por esto que es necesaria la hipótesis sobre X en la segunda parte de esta proposición.

Demostración. Si $f : Y \to X$ es una función sobreyectiva, entonces la Proposición 2.1.4 nos dice que hay una función inyectiva $X \to Y$, y esto prueba que la afirmación (*i*) de la proposición.

Ocupémonos ahora de la parte (*ii*). Supongamos que $X \preceq Y$, de manera que hay una función inyectiva $f : X \to Y$, y que el conjunto X no es vacío. Sea x_0 un elemento cualquiera de X y sea $I \coloneqq f(X)$ la imagen de f. La correstricción $f|^I : X \to I$ es una función biyectiva, así que posee una función inversa $g \coloneqq (f|^I)^{-1} : I \to X$. Podemos entonces considerar la función $h : Y \to X$ que en cada elemento y de Y toma el valor

$$h(y) = \begin{cases} g(y) & \text{si } y \in I; \\ x_0 & \text{si } y \notin I. \end{cases}$$

Esta función es sobreyectiva — en efecto, para cada elemento x de X es $f(x) \in I \subseteq I$ y, por lo tanto, $h(f(x)) = g(f(x)) = x$ — así que la afirmación (*ii*) de la proposición es cierta. □

Las dos primeras propiedades que podemos dar de la relación $\preceq$ son muy sencillas.

Proposición 3.1.7. *Sean X, Y y Z tres conjuntos.*
(*i*) *Es $X \preceq X$.*
(*ii*) *Si $X \preceq Y$ e $Y \preceq Z$, entonces $X \preceq Z$.*

Demostración. La primera afirmación es consecuencia inmediata de que la función identidad $\mathrm{id}_X : X \to X$ es inyectiva. Por otro lado, si tenemos que$X \preceq Y$ e $Y \preceq Z$, de manera que hay funciones inyectivas $f : X \to Y$ y $g : Y \to Z$, entonces la función $g \circ f : X \to Z$ es también inyectiva y, por lo tanto, es $X \preceq Z$. Esto prueba la afirmación (*ii*). □

Es útil observar también que la relación $\preceq$ es compatible con la equipotencia, en el sentido de que la afirmación del siguiente ejercicio es cierta.

Ejercicio 3.1.8. Sean X, X', Y, e Y' cuatro conjuntos. Muestre que si $X \sim X'$, $Y \sim Y'$ y $X \prec Y$, entonces $X' \sim Y'$.

La Proposición 3.1.7 nos dice que la relación $\preceq$ entre conjuntos es reflexiva y transitiva: es lo que llamamos un ***pre-orden***. Notemos que ciertamente no es un orden: es evidente que $\{1\} \preceq \{2\}$ y que $\{2\} \preceq \{1\}$, pero los conjuntos $\{1\}$ y $\{2\}$ no son iguales, así que esta relación no es anti-simétrica.

En general, cuando $\triangleleft$ es un pre-orden en un conjunto X es fácil ver que la relación binaria

$$\{(x, y) \in X \times X : x \triangleleft y \text{ e } y \triangleleft x\}$$

en X es una relación de equivalencia. El siguiente teorema identifica cuál es la relación de equivalencia que corresponde al pre-orden $\preceq$ entre conjuntos.

Teorema 3.1.9. *Sean X e Y dos conjuntos. Es $X \preceq Y$ e $Y \preceq X$ exactamente cuando $X \sim Y$.*

Este teorema es conocido generalmente como el *Teorema de Cantor–Bernstein–Schröder*. Fue enunciado originalmente por Georg Cantor en 1887 en [Can1887], aunque sin dar una demostración. En 1895 Cantor publicó el artículo [Can1895], en el que lo obtuvo como corolario de un teorema que, sin embargo, no pudo probar — y que es equivalente al *Axioma de Elección*, como probó más tarde Friedrich Moritz Hartogs. Dos años más tarde, Felix Bernstein y Ernst Schröder dieron sendas demostraciones de manera casi simultánea pero independiente. La demostración de Bernstein, obtenida cuando este tenía 19 años, aparece en su tesis doctoral [Ber1905] y fue comunicada por Cantor en el Congreso Internacional de Matemáticos de 1897 en Zürich. La demostración de Schröder, por su parte, aparece en su trabajo [Sch1898] — de todas formas, Alwin Reinhold Korselt [Kor1911] encontró, justo antes de la muerte de Schröder, un error en esta demostración. En 1908, Ernst Zermelo descubrió que Richard Dedekind había obtenido una demostración del teorema (sin usar el *Axioma de Elección*) ya en 1887, y la publicó junto a otra nueva suya en [Zer1908b]. De hecho, Dedekind obtuvo una segunda demostración del teorema en 1897. Se conocen hoy varias otras demostraciones del teorema, basadas en ideas bien diferentes. Es de notar que Arie Hinkis escribió un libro entero, [Hin2013], sobre diferentes demostraciones de este teorema.

Nosotros probaremos el teorema apoyándonos en el siguiente resultado, que es un caso particular del llamado *Teorema de punto fijo de Knaster–Tarski*, de Bronisław Knaster [Kna1928] y Alfred Tarski [Tar1955]. Esta demostración fue dada originalmente por Stefan Banach.

Proposición 3.1.10. *Sea X un conjunto y sea $\phi : \mathscr{P}(X) \to \mathscr{P}(X)$ una función creciente, esto es, tal que cualesquiera sean los elementos A y B de $\mathscr{P}(X)$ se tenga que*

$$A \subseteq B \implies \phi(A) \subseteq \phi(B)$$

Hay un elemento F de $\mathscr{P}(X)$ tal que $\phi(F) = F$.

Llamamos a F un ***punto fijo*** de la función ϕ.

Demostración. Consideremos la familia $\mathscr{C} \coloneqq \{A \in \mathscr{P}(X) : A \subseteq \phi(A)\}$ y el elemento $F \coloneqq \bigcup_{A\in\mathscr{C}} A$ de $\mathscr{P}(X)$. Es

$$F = \bigcup_{A\in\mathscr{C}} A \subseteq \bigcup_{A\in\mathscr{C}} \phi(A) = \phi\left(\bigcup_{A\in\mathscr{C}} A\right) = \phi(F).$$

Como la función ϕ es creciente esto implica, en particular, que $\phi(F) \subseteq \phi(\phi(F))$, claro, de manera que $\phi(F) \in \mathscr{F}$ y, por lo tanto, que $\phi(F) \subseteq F$. Vemos así que $\phi(F) = F$. □

Con este resultado a nuestra disposición podemos probar el teorema de Cantor–Bernstein–Schröder fácilmente.

Demostración del Teorema 3.1.9. Supongamos primero que $X \sim Y$, de manera que hay una función biyectiva $f : X \to Y$. Como f es inyectiva, es $X \preceq Y$. Por otro lado, como f es biyectiva es inversible, y su función inversa $f^{-1} : Y \to X$ es, en particular, inyectiva: esto nos dice que también $Y \preceq X$. Vemos así que la condición que da el teorema es suficiente.

Probemos ahora que es necesaria. Supongamos que es$X \preceq Y$ e $Y \preceq X$, de manera que hay funciones inyectivas $f : X \to Y$ y $g : Y \to X$, y consideremos la función

$$\phi : A \in \mathscr{P}(X) \mapsto g(Y \setminus f(X \setminus A)) \in \mathscr{P}(X).$$

Se trata de una función creciente: si A y B son dos elementos de $\mathscr{P}(X)$ tales que $A \subseteq B$, entonces $X \setminus A \supseteq X \setminus B$, así que $f(X \setminus A) \supseteq f(X \setminus B)$ y, a su vez, esto implica que $Y \setminus f(X \setminus A) \subseteq Y \setminus f(X \setminus B)$ y, en definitiva, que

$$\phi(A) = g(Y \setminus f(X \setminus A)) \subseteq g(Y \setminus f(X \setminus B)) = \phi(B).$$

De acuerdo a la Proposición 3.1.10, entonces, hay un elemento F de $\mathscr{P}(X)$ tal que $\phi(F) = F$, esto es, un subconjunto F de X tal que $g(Y \setminus f(X \setminus F)) = F$.

Como la función f es inyectiva, la función

$$\bar{f} : a \in X \setminus F \mapsto f(a) \in f(X \setminus F)$$

es biyectiva y, por lo tanto,

$$X \setminus F \sim f(X \setminus F). \tag{3.1}$$

De manera similar, como la función g es inyectiva, la función

$$\bar{g} : b \in Y \setminus f(X \setminus F) \mapsto g(b) \in g(Y \setminus f(X \setminus F))$$

es biyectiva y tenemos que

$$Y \setminus f(X \setminus F) \sim g(Y \setminus f(X \setminus F)) = F. \tag{3.2}$$

Como $X = (X \setminus F) \cup F$ y $Y = f(X \setminus F) \cup \big(Y \setminus f(X \setminus F)\big)$, y ambas uniones son disjuntas, podemos concluir de (3.1) y (3.2), gracias al resultado del Ejercicio 3.1.2, que $X \sim Y$. □

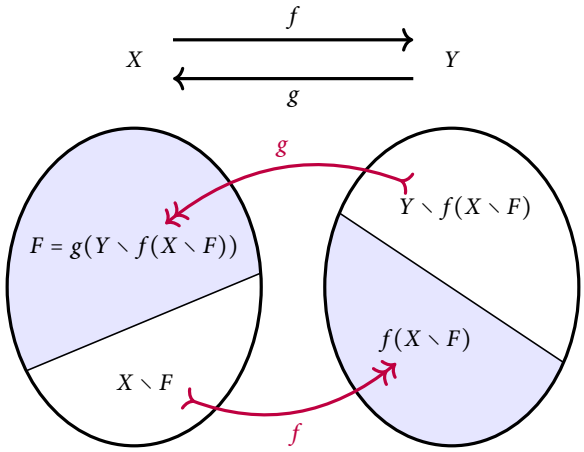


Figura 3.1. La elección del conjunto F en la demostración del Teorema de Cantor–Bernstein–Schröder 3.1.9.

Ejercicio 3.1.11. En la situación de esta demostración describa explícitamente una biyección $X \to Y$.

Una segunda observación fundamental que tenemos que hacer sobre el pre-orden $\preceq$ entre conjuntos es una relación total y que, de hecho, vale el siguiente resultado de tricotomía.

Proposición 3.1.12. *Sean X e Y dos conjuntos. Exactamente una de las siguientes afirmaciones es cierta:*

$$X \prec Y, \qquad X \sim Y, \qquad Y \prec X.$$

Demostración. Se sigue inmediatamente de las definiciones y del *Teorema de Cantor–Bernstein–Schröder* 3.1.9 que como mucho una de esas tres afirmaciones puede ser cierta. Resta probar entonces que alguna de ellas necesariamente es cierta: para ello suponemos que $X \not\preceq Y$, esto es, que no hay ninguna función inyectiva $X \to Y$, y probamos que $Y \preceq X$.

Sea $\mathscr{S}$ el conjunto de todas las funciones inyectivas $s : A \to X$ definidas en un subconjunto A de Y. Este conjunto no es vacío, ya que la función vacía $\varnothing \to X$ es inyectiva. Consideremos en $\mathscr{S}$ la relación binaria $\trianglelefteq$ tal que si $s : A \to X$ y $t : B \to X$ son dos elementos de $\mathscr{S}$ se tiene que

$$s \trianglelefteq t \iff A \subseteq B \text{ y } t|_A = s.$$

Podemos hacer sobre $\trianglelefteq$ las siguientes tres observaciones.

- Si $s : A \to X$ es un elemento de $\mathscr{S}$, entonces $s \trianglelefteq s$: la relación $\trianglelefteq$ es, por lo tanto, reflexiva.
- Sean $s : A \to X$ y $t : B \to X$ dos elementos de $\mathscr{S}$ tales que $s \trianglelefteq t$ y $t \trianglelefteq s$. Esto nos dice que $A \subseteq B$ y $B \subseteq A$, así que, de hecho, $A = B$ y $s = s|_B = t$. Vemos así que la relación $\trianglelefteq$ es anti-simétrica.
- Sean ahora $s : A \to X$, $t : B \to X$ y $u : C \to X$ tres elementos de $\mathscr{S}$ tales que$s \trianglelefteq t$ y $t \trianglelefteq u$, de manera que $A \subseteq B$, $B \subseteq C$, $t|_A = s$ y $u|_B = t$. Las dos inclusiones implican que $A \subseteq C$, y las dos igualdades que $u|_A = (u|_B)|_A = t|_A = s$, así que $s \trianglelefteq u$. La relación $\trianglelefteq$ es, por lo tanto, transitiva.

Todo esto nos permite concluir que la relación $\trianglelefteq$ es una relación de orden sobre el conjunto $\mathscr{S}$.

Queremos mostrar ahora que la relación $\trianglelefteq$ satisface la hipótesis del *Lema de Zorn*. Sea $\mathscr{C}$ un subconjunto de $\mathscr{S}$ que está totalmente ordenado por la relación $\trianglelefteq$, para cada cada elemento s de $\mathscr{C}$ escribamos A_s al dominio de s, y consideremos el conjunto $A \coloneqq \bigcup_{s \in \mathscr{C}} A_s$, que es un subconjunto de Y, y el conjunto

$$f \coloneqq \{(a, s(a)) : s \in \mathscr{C}, a \in A_s\},$$

que es un subconjunto de $A \times X$. Las siguientes dos observaciones prueban que f es, de hecho, una función $A \to X$.

- Sea a un elemento de A. En vista de la definición del conjunto A, hay un elemento s de $\mathscr{C}$ tal que $a \in A_s$ y, por lo tanto, el par $(a, s(a))$ pertenece a f.
- Sea ahora a un elemento de A, sean x y x' dos elementos de X, y supongamos que los pares ordenados (a, x) y (a, x') pertenecen al conjunto f, de manera que hay elementos s y s' de $\mathscr{C}$ tales que $a \in A_s$, $x = s(a)$, $a \in A_{s'}$ y $x' = s'(a)$. Ahora bien, como el conjunto $\mathscr{C}$ está totalmente ordenado por la relación $\trianglelefteq$, es $s \trianglelefteq s'$ o $s' \trianglelefteq s$ y, sin pérdida de generalidad, podemos suponer que la primera de esas afirmaciones es cierta, de manera que $A_s \subseteq A_{s'}$ y $s'|_{A_s} = s$. Tenemos entonces que $x = s(a) = (s'|_{A_s})(a) = s'(a) = x'$.

Probemos ahora que esta función $f : A \to X$ es inyectiva, de manera que es un elemento del conjunto $\mathscr{S}$.

- Sean a y a' dos elementos de A tales que $f(a) = f(a')$. Como los pares ordenados $(a, f(a))$ y $(a', f(a'))$ pertenecen a f, hay elementos s y s' de $\mathscr{C}$ tales que $a \in A_s$, $f(a) = s(a)$, $a \in A_{s'}$ y $f(a') = s'(a')$. Como el conjunto $\mathscr{C}$ está totalmente ordenado con respecto a la relación $\trianglelefteq$, es $s \trianglelefteq s'$ o $s' \trianglelefteq s$ y, sin pérdida de generalidad, podemos suponer que la primera de esas afirmaciones es cierta, de manera que $A_s \subseteq A_{s'}$ y $s'|_{A_s} = s$. Esto implica que a y a' son elementos de $A_{s'}$ y que

  $$s'(a') = f(a') = f(a) = s(a) = (s'|_{A_s})(a) = s'(a),$$

  y, como la función s' es inyectiva, podemos concluir que $a' = a$.

Para terminar de verificar que la relación $\trianglelefteq$ satisface la hipótesis del *Lema de Zorn* tenemos que mostrar que $s \trianglelefteq f$ para todo elemento s de $\mathscr{C}$.

- Sea s un elemento de $\mathscr{C}$. De la definición misma del conjunto A sabemos que $A_s \subseteq A$. Por otro lado, si a es un elemento de A_s, entonces la definición del conjunto f implica que contiene al par ordenado $(a, s(a))$, y esto significa que $f(a) = s(a)$. Vemos así que $f|_{A_s} = s$ y, por lo tanto, que $s \trianglelefteq f$, como queremos.

Ahora bien, como la relación $\trianglelefteq$ en el conjunto $\mathscr{S}$ satisface la hipótesis del *Lema de Zorn*, podemos usar este y concluir que hay un elemento $g : C \to X$ en $\mathscr{S}$ que es maximal para $\trianglelefteq$. Mostraremos que es $C = Y$, de manera que g es una función inyectiva $Y \to X$ y que, por lo tanto, es $Y \preceq X$: esto completará la demostración de la proposición.

Supongamos que, por el contrario, el subconjunto C de Y es propio, de manera que hay un elemento y_0 en $Y \setminus C$. La función g no es sobreyectiva: de serlo sería una biyección, así que sería inversible, la composición de su función inversa $g^{-1} : X \to C$ con la inclusión $C \hookrightarrow Y$ sería una función inyectiva $X \to Y$, y nuestra hipótesis es que $X \not\preceq Y$. Hay, por lo tanto, un elemento x_0 en el conjunto $X \setminus g(C)$. Es claro que hay una función $h : C \cup \{y_0\} \to X$ que en cada elemento y de $C \cup \{y_0\}$ toma el valor

$$h(y) = \begin{cases} g(y) & \text{si } y \in C; \\ x_0 & \text{si } y = y_0. \end{cases}$$

Esta función es inyectiva. Para verlo, supongamos que y e y' son dos elementos de $C \cup \{y_0\}$ tales que $h(y) = h(y')$. Como $x_0 \notin g(C)$, es claro que o bien tanto y como y' están en C o bien ambos son iguales a y_0, y en el primer caso lo que tenemos es que $g(y) = g(y')$, de manera que $y = y'$ porque la función g es inyectiva.

Como la función $h : C \cup \{y_0\} \to X$ es inyectiva, se trata de un elemento de $\mathscr{S}$. Es evidente que $g \trianglelefteq h$ y que $g \neq h$: esto es absurdo, porque g es un elemento maximal con respecto a la relación $\trianglelefteq$. Esta contradicción provino de haber supuesto que el subconjunto C de Y es propio, así que no puede serlo, esto es, es $C = Y$. Como dijimos arriba, esto completa la prueba de la proposición. □

En la demostración de esta proposición usamos el *Lema de Zorn* que, como sabemos, es equivalente al *Axioma de Elección*. De hecho, es posible probar que la proposición que acabamos de probar es ella misma equivalente a ese axioma.

En este punto tenemos una forma precisa de comparar los tamaños de los conjuntos. Una observación importante, debida a Georg Cantor [Can1892], es que no hay un máximo tamaño que un conjunto puede tener.

Proposición 3.1.13. *Sea X un conjunto. Es $X \prec \mathscr{P}(X)$.*

Demostración. La función $x \in X \mapsto \{x\} \in \mathscr{P}(X)$ es claramente inyectiva, así que $X \preceq \mathscr{P}(X)$. Para probar la proposición tenemos que mostrar entonces que $X \not\sim \mathscr{P}(X)$.

Supongamos que, por el contrario, hay una función biyectiva $f : X \to \mathscr{P}(X)$, y consideremos el subconjunto

$$C \coloneqq \{x \in X : x \notin f(x)\}$$

de X. Como la función f es sobreyectiva, hay un elemento x en X tal que $f(x) = C$. Si fuera $x \in C = f(x)$, tendríamos, de acuerdo a la definición de C, que $x \notin C$: esto es absurdo. Vemos así que debe ser $x \notin C = f(x)$: pero debe ser entonces $x \in C$, lo que también es absurdo. En cualquier caso llegamos a una contradicción, y esto provino de haber supuesto que hay una función biyectiva $X \to \mathscr{P}(X)$, así que no la hay. Esto es lo que teníamos que probar. □

Se sigue de esto que si empezamos con un conjunto X cualquiera e iteramos la operación $\mathscr{P}$ obtenemos una sucesión de conjuntos cada vez más grandes:

$$X \prec \mathscr{P}(X) \prec \mathscr{P}(\mathscr{P}(X)) \prec \mathscr{P}(\mathscr{P}(\mathscr{P}(X))) \prec \cdots$$

De hecho, veremos que podemos continuar esta sucesión *transfinitamente* y obtener se esa forma conjuntos inimaginablemente grandes.

3.2. Conjuntos finitos

Para cada $n \in \mathbb{N}_0$ escribimos

$$[\![n]\!] \coloneqq \{i \in \mathbb{N} : i \leq n\},$$

de manera que, en particular, es $[\![0]\!] = \varnothing$. Estos conjuntos nos interesan porque nos permiten hacer la siguiente definición:

Definición 3.2.1. Un conjunto X es ***finito*** si hay un elemento n de $\mathbb{N}_0$ tal que $[\![n]\!] \sim X$.

Es fácil dar ejemplos de conjuntos finitos:

Lema 3.2.2. *Sea n un elemento de $\mathbb{N}_0$. El conjunto $[\![n]\!]$ es finito.*

Demostración. En efecto, la relación de equipotencia es reflexiva, así que $[\![n]\!] \sim [\![n]\!]$. □

No todos los conjuntos son finitos, pero lleva un poco de trabajo exhibir ejemplos de conjuntos que no lo son. Lo haremos en la sección siguiente.

La teoría de los conjuntos finitos no depende del *Axioma de elección*: el «caso finito» de ese axioma pruebe probarse usando solamente el *Principio de inducción*. Aunque no es estrictamente necesario, hagámoslo.

Proposición 3.2.3. *Sea $\mathscr{S}$ una familia* finita *de subconjuntos no vacíos de un conjunto X. Hay una función $c : \mathscr{S} \to X$ tal que $c(A) \in A$ para cada elemento A de $\mathscr{S}$.*

Demostración. Sea n el cardinal de la familia $\mathscr{S}$, de manera que $n \in \mathbb{N}_0$ y hay una función biyectiva $f : [\![n]\!] \to \mathscr{S}$. Mostraremos que la proposición es cierta haciendo inducción con respecto a este entero n. Notemos que cuando n es 0 no hay nada que probar: en ese caso el conjunto $\mathscr{S}$ es vacío y la función vacía $\mathscr{S} \to X$ claramente satisface la condición del enunciado.

Sea entonces n un elemento de $\mathbb{N}_0$, supongamos que la afirmación del enunciado es cierta siempre que el cardinal de la familia $\mathscr{S}$ es $\mathscr{S}$, y consideremos una familia $\mathscr{S}$ de cardinal $n+1$ de subconjuntos de X, de manera que hay una función biyectiva $f : [\![n+1]\!] \to \mathscr{S}$. Es claro que el subconjunto $f([\![n]\!])$ de $\mathscr{S}$ tiene cardinal n, así que la hipótesis inducción nos dice que hay una función $d : f([\![n]\!]) \to X$ tal que $d(A) \in A$ para cada $A \in f([\![n]\!])$. Por otro lado, como el conjunto $f(n+1)$ no es vacío, podemos elegir en él un elemento x y considerar la función $c : \mathscr{S} \to X$ que en cada elemento A de $\mathscr{S}$ toma el valor

$$c(A) = \begin{cases} d(A) & \text{si } A \in f([\![n]\!]); \\ x & \text{en caso contrario.} \end{cases}$$

Esta es una función de elección en $\mathscr{S}$. En efecto, si A es un elemento de $\mathscr{S}$, entonces o bien A pertenece a $f([\![n]\!])$ y $c(A) = d(A) \in A$, o bien A no pertenece a ese conjunto, de manera que $c(A) = x \in f(n+1) = A$. Esto completa la inducción y, por lo tanto, la demostración. □

Una vez que tenemos a mano esta versión finita del *Axioma de elección* podemos, claro, obtener correspondientes versiones finitas de las consecuencias de ese axioma — que podemos probar todas sin usarlo. Así, por ejemplo, el siguiente resultado es la versión finita de la Proposición 2.1.4.

Corolario 3.2.4. *Si $f : X \to Y$ es una función sobreyectiva con codominio finito, entonces hay una función $g : Y \to X$ tal que $f \circ g = \mathsf{id}_Y$ y esta función es inyectiva.*

Demostración. Sea $f : X \to Y$ una función sobreyectiva, de manera que para cada elemento y de Y el conjunto $f^{-1}(y)$ no es vacío, y supongamos que el conjunto Y es finito: $\mathscr{S} \coloneqq \{f^{-1}(y) : y \in Y\}$ es entonces una familia de subconjuntos no vacíos de X y la Proposición 3.2.3 nos dice que hay una función $c : \mathscr{S} \to X$ tal que $c(A) \in A$ para cada $A \in \mathscr{S}$. Podemos, por lo tanto, definir la función $g : y \in Y \mapsto c(f^{-1}(y)) \in X$, y es claro que para todo $y \in Y$ es

$$(f \circ g)(y) = f(g(y)) = f(c(f^{-1}(y))) = y,$$

ya que $c(f^{-1}(y)) \in f^{-1}(y)$. Vemos así que $f \circ g = \mathsf{id}_Y$, como queremos. Finalmente, si y e y' son dos elementos de Y tales que $g(y) = g(y')$, entonces tenemos que

$$y = \mathsf{id}_Y(y) = (f \circ g)(y) = f(g(y)) = f(g(y')) = (f \circ g)(y') = \mathsf{id}_Y(y') = y',$$

y esto nos dice que la función g es inyectiva. □

El siguiente ejercicio da una versión finita de la Proposición 3.1.6.

Ejercicio 3.2.5. Sean X e Y dos conjuntos. Pruebe las siguientes dos afirmaciones sin usar el *Axioma de elección*.

(*i*) Si hay una función sobreyectiva $Y \to X$ y el conjunto X es finito, entonces hay una función inyectiva $X \to Y$.

(*ii*) Si hay una función inyectiva $X \to Y$, el conjunto X no es vacío, y el conjunto Y es finito, entonces hay una función sobreyectiva $Y \to X$.

Queremos ahora clasificar a los conjuntos finitos de acuerdo a su cardinal, y el primer paso para hacerlo es la siguiente observación:

Proposición 3.2.6. *Sean m y n dos elementos de $\mathbb{N}_0$.*

(*i*) *Hay una función sobreyectiva $[\![m]\!] \to [\![n]\!]$ si y solamente si $m \geq n$.*

(*ii*) *Hay una función inyectiva $[\![m]\!] \to [\![n]\!]$ si y solamente si $m \leq n$.*

Demostración. Es fácil ver que cuando $m \geq n$ la función $i \in [\![m]\!] \mapsto \min\{i, n\} \in [\![n]\!]$ es sobreyectiva, así que para ver que la primera afirmación de la proposición es cierta es suficiente probemos inductivamente que para todo $m \in \mathbb{N}_0$ vale que

si n es un elemento de $\mathbb{N}_0$ tal que $m < n$, entonces no hay funciones sobreyectivas $[\![m]\!] \to [\![n]\!]$.

Esto es claro si m es 0: en efecto, si n es un elemento de $\mathbb{N}_0$ tal que $0 < n$, entonces es $[\![0]\!] = \varnothing$ y $[\![n]\!] \neq \varnothing$, así que no hay funciones sobreyectivas $[\![0]\!] \to [\![n]\!]$.

Sea ahora m un elemento de $\mathbb{N}_0$ que hace la afirmación de arriba cierta, sea n un elemento de $\mathbb{N}_0$ tal que $m + 1 < n$, sea $f : [\![m+1]\!] \to [\![n]\!]$ una función, y, para llegar a una contradicción, supongamos que la función f es sobreyectiva. Como $m < n$, la hipótesis inductiva implica que la restricción $f|_{[\![m]\!]} : [\![m]\!] \to [\![n]\!]$ no es sobreyectiva, así que hay un elemento y de $[\![n]\!]$ tal que $f(i) \neq y$ para todo $i \in [\![m]\!]$. Por otro lado, como la función f es sobreyectiva, hay un elemento x de $[\![m+1]\!]$ tal que $f(x) = y$: en vista de la forma en que elegimos a y, entonces, es $x = m + 1$.

Consideremos ahora la función $g : [\![m]\!] \to [\![n-1]\!]$ que en cada elemento i de $[\![m]\!]$ toma el

valor

$$g(i) = \begin{cases} f(i) & \text{si } f(i) < y; \\ f(i) - 1 & \text{si } f(i) \geq y. \end{cases}$$

Notemos que esto tiene sentido:

- si $f(i) < y$, entonces ciertamente $f(i) \in [\![n-1]\!]$, ya que $y \leq n$, y
- si en cambio $f(i) \geq y$, entonces $y < f(i)$, ya que $f(i) \neq y$ porque $i \in [\![m]\!]$, y, por lo tanto, $0 \leq y - 1 < f(i) - 1 \leq n - 1$, así que otra vez $f(i) - 1 \in [\![n-1]\!]$.

Sea j un elemento de $[\![n-1]\!]$.

- Supongamos primero que $j < y$. Como la función f es sobreyectiva, hay un elemento i de $[\![m+1]\!]$ tal que $f(i) = j$, y como $f(m+1) = y \neq j$, es $i \neq m+1$, de manera que $i \in [\![m]\!]$. Es entonces $g(i) = f(i) = j$, ya que $f(i) = j < y$.
- Supongamos ahora que $j \geq y$. Es $j + 1 \in [\![n]\!]$, así que hay un elemento i de $[\![m+1]\!]$ tal que $f(i) = j + 1$, y es $g(i) = f(i) - 1 = j$, ya que $f(i) = j + 1 \geq y$.

En cualquiera de los dos casos vemos que j pertenece a la imagen de g, y esto nos permite concluir que la función $g : [\![m]\!] \to [\![n-1]\!]$ es sobreyectiva. Esto es imposible, ya que, como $m < n - 1$, contradice a la hipótesis inductiva.

Esta contradicción provino de haber supuesto que la función f es sobreyectiva, así que no lo es. Esto completa la inducción y, en consecuencia, la prueba de la afirmación (*i*) de la proposición.

Probemos ahora la afirmación (*ii*). Si $m \leq n$, entonces la inclusión $[\![m]\!] \hookrightarrow [\![n]\!]$ es una función inyectiva, así que la condición dada allí es ciertamente suficiente. Veamos que también es necesaria. Supongamos que $m > n$ y, para llegar a una contradicción, que hay una función inyectiva $[\![m]\!] \to [\![n]\!]$. Como el conjunto $[\![m]\!]$ no es vacío, ya que $m > 0$, el resultado del Ejercicio 3.2.5 nos dice que hay una función sobreyectiva $[\![n]\!] \to [\![m]\!]$: esto es absurdo, ya que contradice la parte (*i*) del lema, que ya probamos. La proposición queda así completamente probada. □

La segunda parte de esta proposición o, más precisamente, la afirmación contrarrecíproca a ella que enunciamos en el siguiente corolario, es conocida como el *Principio del palomar* o *Principio de Dirichlet*, por Peter Gustav Lejeune Dirichlet, que lo usó en su solución de la ecuación de Pell.

Corolario 3.2.7. *Sean m y n elementos de $\mathbb{N}_0$. Una función $[\![m]\!] \to [\![n]\!]$ no es inyectiva si $m > n$.* □

Un segundo corolario de la Proposición 3.2.6 es el siguiente resultado, que nos da la clasificación de los conjuntos de la forma $[\![n]\!]$ con $n \in \mathbb{N}_0$ bajo la relación de equipotencia.

Proposición 3.2.8. *Sean m y n dos elementos de $\mathbb{N}_0$. Los conjuntos $[\![m]\!]$ y $[\![n]\!]$ son equipotentes si y solamente si $n = m$.*

Demostración. La suficiencia de la condición es evidente, ya que la relación de equipotencia es reflexiva, así que bastará que probemos su necesidad. Sean m y n dos elementos de $\mathbb{N}_0$ tales que $[\![m]\!] \sim [\![n]\!]$, de manera que hay una función biyectiva $f : [\![m]\!] \to [\![n]\!]$. Como la función f es sobreyectiva, la primera parte de la Proposición 3.2.6 nos dice que $m \geq n$, y como la función f es inyectiva, la segunda parte de esa proposición nos dice que $m \leq n$: juntando todo vemos que $m = n$, como queremos. □

A su vez, esta proposición nos permite clasificar bajo la relación de equipotencia a todos los conjuntos finitos.

Proposición 3.2.9. *Sea X un conjunto finito. Hay exactamente un elemento n de $\mathbb{N}_0$ tal que $X \sim [\![n]\!]$.*

Demostración. Que hay un elemento de $\mathbb{N}_0$ con esa propiedad es precisamente lo que significa que el conjunto sea finito, así que solamente tenemos probar que hay uno solo. Sean m y n dos elementos de $\mathbb{N}_0$ tales que $X \sim [\![m]\!]$ y $X \sim [\![n]\!]$. El hecho de que la relación de equipotencia es una relación de equivalencia implica que $[\![m]\!] \sim [\![n]\!]$ y entonces la Proposición 3.2.8 nos permite concluir que $m = n$. □

Esta proposición nos dice que todo conjunto finito es equipotente a exactamente un conjunto de la sucesión

$$[\![0]\!], [\![1]\!], [\![2]\!], [\![3]\!], \ldots$$

De ahora en adelante para cada elemento n de $\mathbb{N}_0$ escribiremos n para denotar al cardinal $|[\![n]\!]|$ del conjunto $[\![n]\!]$ y diremos, por ejemplo, que «el cardinal del conjunto $\{1, 2, 3\}$ es 3».

La Proposición 3.2.6, que nos permitió describir completamente la relación de equipotencia entre conjuntos finitos, también nos sirve para describir la relación de pre-orden entre ellos.

Proposición 3.2.10. *Sean X e Y dos conjuntos finitos, y sean m y n sus cardinales. Es $X \prec Y$ si y solamente si $m < n$.*

Demostración. Si $X \preceq Y$, entonces tenemos que $[\![m]\!] \sim X \preceq Y \sim [\![n]\!]$ y, de acuerdo a la segunda parte de la Proposición 3.2.6, es $m \leq n$, y si además $X \not\sim Y$, la Proposición 3.2.8 nos dice que $m \neq n$. Esto prueba la necesidad de la condición que da la proposición.

Por otro lado, si $m < n$, entonces es $[\![m]\!] \preceq [\![n]\!]$ porque la función de inclusión $[\![m]\!] \hookrightarrow [\![n]\!]$ es inyectiva, y $[\![m]\!] \not\sim [\![n]\!]$, así que $X \sim [\![m]\!] \prec [\![n]\!] \sim Y$. La condición de la proposición es, por lo tanto, también suficiente. □

De acuerdo a la Definición 3.2.1, para mostrar que un conjunto es finito es necesario exhibir una biyección entre él y un conjunto de la forma $[\![n]\!]$ con $n \in \mathbb{N}_0$. La siguiente proposición da una condición más débil para ello y, por lo tanto, más fácil de verificar.

Proposición 3.2.11. *Un conjunto X es finito si y solamente si hay una función sobreyectiva $Y \to X$ con dominio finito. Más aún, cuando ese es el caso,*

- *hay enteros no negativos n para los que hay funciones sobreyectivas $[\![n]\!] \to X$,*
- *el menor de ellos es el cardinal del conjunto X, y*
- *si k es ese cardinal, entonces cualquier función sobreyectiva $[\![k]\!] \to X$ es biyectiva.*

Demostración. Si el conjunto X es finito, entonces la función $\mathrm{id}_X : X \to X$ es sobreyectiva y tiene dominio finito: esto muestra que la condición que da la proposición es necesaria.

Veamos ahora que también es suficiente. Supongamos que hay una función sobreyectiva $f : Y \to X$ cuyo dominio es finito, de manera que hay un entero no negativo n y una función biyectiva $g : [\![n]\!] \to Y$. Como la función $f \circ g : [\![n]\!] \to X$ también es sobreyectiva, vemos que el conjunto S de todos los enteros no negativos m tales que hay una función sobreyectiva $[\![m]\!] \to X$ no es vacío y que podemos, por lo tanto, considerar su menor elemento $k \coloneqq \min S$.

Como k pertenece a S, hay una función sobreyectiva $u : [\![k]\!] \to X$. Para probar que el conjunto X es finito — y con ello tanto la suficiencia de la condición de la proposición como las afirmaciones finales que aparecen ahí — es suficiente que probemos que la función u es también inyectiva. Supongamos que no lo es, de manera que hay dos elementos a y b de $[\![k]\!]$ tales que $u(a) = u(b)$ y $a < b$. Como $[\![k]\!]$ tiene en particular dos elementos distintos, es claro que $k \geq 2$, así que podemos considerar el conjunto $[\![k-1]\!]$ y la función $v : [\![k-1]\!] \to X$ que en cada elemento i de $[\![k-1]\!]$ toma el valor

$$v(i) = \begin{cases} u(i) & \text{si } i < b; \\ u(i+1) & \text{si } i \geq b. \end{cases}$$

Sea x un elemento cualquiera de X. Como la función u es sobreyectiva, hay un elemento t de $[\![k]\!]$ tal que $u(t) = x$. Tenemos ahora dos casos a considerar.

- Supongamos primero que $x = u(a)$. Como $a < b \leq k$, el entero a pertenece al dominio de v y es $a < b$, así que tenemos que $x = u(a) = v(a)$.
- Supongamos ahora que $x \neq u(a)$. Como la función u es sobreyectiva, sabemos que hay un elemento c de $[\![k]\!]$ tal que $u(c) = x$. Si $c < b$, entonces c pertenece a $[\![k-1]\!]$ y es $x = u(c) = v(c)$. Si en cambio $c \geq b$, entonces $c - 1 \in [\![k-1]\!]$ porque $b > 1$ y tenemos que $x = u(c) = v(c-1)$, ya que $c - 1 \geq b$

En cualquiera de los dos casos vemos que x pertenece a la imagen de v, y esto nos permite concluir que la función $v : [\![k-1]\!] \to X$ es sobreyectiva y, en consecuencia, que $k-1$ pertenece al conjunto S. Esto es absurdo, ya que k es el menor elemento de S, y esta contradicción provino de haber supuesto que la función u no es inyectiva. Como dijimos arriba, esto prueba la proposición. □

La primera afirmación de la Proposición 3.2.11 puede enunciarse también de la siguiente

manera.

Corolario 3.2.12. *Sean X e Y dos conjuntos. Si $X \preceq Y$ e Y es finito, entonces X es finito.*

Demostración. En efecto, en esa situación sabemos que hay una función sobreyectiva $Y \to X$, así que que X es finito se sigue inmediatamente de la proposición. □

La versión dual de la Proposición 3.2.11 también es cierta.

Ejercicio 3.2.13. Pruebe que un conjunto X es finito si y solamente si hay una función inyectiva $X \to Y$ con codominio finito, y que cuando ese es el caso

- hay enteros no negativos n para los que hay funciones inyectivas $X \to [\![n]\!]$,
- el menor de ellos es el cardinal del conjunto X, y
- si k es ese cardinal, entonces cualquier función inyectiva $X \to [\![k]\!]$ es biyectiva.

La siguiente proposición establece el hecho intuitivamente evidente de que los subconjuntos de un conjunto finito no son más grandes que este y, más aún, de que si son propios son estrictamente más chicos.

Proposición 3.2.14. *Sea X un conjunto finito, sea n el cardinal de X, y sea Y un subconjunto de X. El conjunto Y es finito. Si m es el cardinal de Y, entonces es $m \leq n$ y vale la igualdad si y solamente si $Y = X$.*

Demostración. Sea $f : X \to [\![n]\!]$ una biyección. La función de inclusión $i : Y \hookrightarrow X$ es inyectiva, así que el resultado del Ejercicio 3.2.13 nos dice que el conjunto Y es finito y, como la composición $f \circ i : Y \to [\![n]\!]$ es inyectiva, que el cardinal m de Y es tal que $m \leq n$. Si $m = n$, entonces la tercera afirmación de ese ejercicio nos dice que la función $f \circ i$ es sobreyectiva, y esto implica que $Y = X$: si x es un elemento de X, entonces hay un elemento y de Y tal que $f(x) = (f \circ i)(y) = f(i(y))$ y, como la función f es inyectiva, es $x = i(y) = y \in Y$. Como es claro que, recíprocamente, cuando $X = Y$ es $n = m$, esto prueba la proposición. □

Un caso particular de esta proposición es útil.

Ejercicio 3.2.15. Sea X un conjunto finito y sea x un elemento de X. Muestre que

$$|X \setminus \{x\}| + 1 = |X|.$$

Una aplicación útil de la Proposición 3.2.14 es el siguiente criterio que nos permite decidir si una endofunción de un conjunto finito es biyectiva.

Proposición 3.2.16. *Sea X un conjunto finito. Una función $f : X \to X$ que es inyectiva o sobreyectiva es biyectiva.*

Demostración. Sea m el cardinal de X y sea $f : X \to X$ una función. Para probar la proposición es suficiente que mostremos que f es inyectiva si y solamente si es sobreyectiva.

Supongamos primero que f es inyectiva, y sea I la imagen de f. Como la correstricción $f|^I : X \to I$ es una biyección, sabemos que el conjunto I es finito y que su cardinal es m. Más aún, como I es un subconjunto de X, la Proposición 3.2.14 nos dice que $I = X$, de manera que la función f es sobreyectiva.

Supongamos ahora que la función f es sobreyectiva y, para llegar a una contradicción, que no es inyectiva, de manera que hay dos elementos x y x' en X tales que $f(x) = f(x')$. Consideremos el subconjunto $Y \coloneqq X \setminus \{x'\}$ de X y la restricción $g \coloneqq f|_Y : Y \to X$. Esta función es sobreyectiva: si u es un elemento de X, entonces o bien $u = f(x) \in g(Y)$, o bien $u \neq f(x)$. En este segundo caso hay un elemento v de X tal que $f(v) = u$, ya que la función f es sobreyectiva, y es $v \neq x'$, ya que $f(v) = u \neq f(x) = f(x')$: es entonces $v \in Y$ y $g(v) = f(v) = u$. Vemos así que en cualquier caso v pertenece a la imagen de g.

Escribamos n al cardinal de Y. Como Y es un subconjunto propio de X, sabemos que es finito y que $n < m$. Por otro lado, como la función $g : Y \to X$ es sobreyectiva, sabemos que $n \geq m$. Esto es absurdo, y esta contradicción provino de haber supuesto que la función f no es inyectiva, así que debe serlo. □

La clase de los conjuntos finitos es cerrada con respecto a varias operaciones, y saber esto resulta muy útil en la práctica. Veamos algunos ejemplos.

Proposición 3.2.17. *Sean X e Y dos conjuntos. La unión $X \cup Y$ es un conjunto finito si y solamente si X e Y lo son. Más aún, si ese es el caso y escribimos m y n a los cardinales de X e Y, entonces*

$$|X \cup Y| \leq m + n$$

y la igualdad vale si y solamente los conjuntos X e Y son disjuntos.

Demostración. Si la unión $X \cup Y$ es un conjunto finito, entonces X e Y, que son subconjuntos de esa unión, también lo son. La condición de la proposición es, por lo tanto, necesaria.

Supongamos ahora que los conjuntos X e Y son finitos y escribamos m y n a sus correspondientes cardinales, de manera que hay dos funciones biyectivas $f : [\![m]\!] \to X$ y $g : [\![n]\!] \to Y$. Es claro que hay una función $h : [\![m + n]\!] \to X \cup Y$ que en cada elemento i de $[\![m + n]\!]$ toma el valor

$$h(i) = \begin{cases} f(i) & \text{si } i \leq m; \\ g(i - m) & \text{si } i > m. \end{cases}$$

Esta función es sobreyectiva. Para verlo, sea z un elemento de $X \cup Y$. Si z pertenece a X, entonces como la función f es sobreyectiva hay un elemento i de $[\![m]\!]$ tal que $z = f(i)$ y entonces, como $i \in [\![m+n]\!]$ e $i \leq m$, tenemos que $z = f(i) = h(i)$. Si z pertenece a Y, entonces como la función g es sobreyectiva hay un elemento i de $[\![n]\!]$ tal que $z = g(i)$ y entonces, como $m + i \in [\![m+n]\!]$ y $m + i > m$, tenemos que $z = g(i) = h(m+i)$. En cualquier caso, entonces, z pertenece a la imagen de la función h, así que esta es sobreyectiva, como dijimos. Gracias a la Proposición 3.2.11 podemos concluir que el conjunto $X \cup Y$ es finito y que su cardinal es $|X \cup Y| \leq m + n$. Para terminar, tenemos que probar la afirmación final de la proposición.

Supongamos primero que el cardinal de $X \cup Y$ es $n + m$, de manera que la Proposición 3.2.11 nos dice que la función h es biyectiva, y, contra lo que queremos probar, que hay un elemento z en $X \cup Y$. Como las funciones f y g son sobreyectivas, hay elementos i y j de $[\![m]\!]$ y de $[\![n]\!]$, respectivamente, tales que $f(i) = z$ y $g(j) = n$. Se sigue de esto que $1 \leq i < j + m \leq n + m$ y que $h(i) = h(j+m)$, pero esto es absurdo, ya que la función h es biyectiva. Vemos así que debe ser $X \cap Y = \varnothing$, como queremos.

Supongamos ahora que los conjuntos X e Y son disjuntos y mostremos que la función h es entonces inyectiva: tendremos entonces que es, de hecho, biyectiva y, por lo tanto, que el cardinal de $X \cup Y$ es $m + n$. Sea i y j dos elementos de $[\![m+n]\!]$ tales que $h(i) = h(j)$. Si $h(i) \in X$, entonces $h(j) \in X$ y, de acuerdo a la definición de h, es $i, j \leq m$, así que $f(i) = h(i) = h(j) = f(j)$: como la función f es inyectiva, esto implica que $i = j$. De manera similar, si $h(i) \in Y$, entonces $h(j) \in Y$, $i, j > m$, y por lo tanto, $g(i - m) = h(i) = h(j) = g(j - m)$, de manera que $i - m = j - m$ y, en consecuencia, también $i = j$ en este caso. □

Hay una versión más precisa de esta proposición, conocida como el *Principio de Inclusión–Exclusión*, que dejamos a cargo del lector.

Ejercicio 3.2.18. Si X e Y son dos conjuntos finitos, entonces tanto los conjuntos $X \cup Y$ como $X \cap Y$ son finitos y $|X \cup Y| + |X \cap Y| = |X| + |Y|$.

Por otro lado, la Proposición 3.2.17 tiene una extensión natural al caso de uniones de un número finito arbitrario de conjuntos.

Ejercicio 3.2.19. Sea k un entero no negativo y sean $X_1, X_2, \ldots, X_k$ conjuntos. Pruebe que la unión $X \coloneqq X_1 \cup X_2 \cup \cdots \cup X_k$ es un conjunto finito si y solamente si cada uno de esos k conjuntos es finito, y que si ese es el caso, entonces $|X| \leq |X_1| + |X_2| + \cdots + |X_k|$ y la igualdad vale exactamente cuando los conjuntos $X_1, X_2, \ldots, X_k$ son disjuntos dos a dos.

Hay también una versión del *Principio de Inclusión–Exclusión* para esta situación, pero no la necesitaremos aquí.

La clase de los conjuntos finitos también es cerrada por productos cartesianos.

Proposición 3.2.20. *Sean X e Y dos conjuntos. El producto cartesiano $X \times Y$ es un conjunto finito si y solamente si los conjuntos X e Y lo son, y cuando ese es el caso su cardinal es*

$$|X \times Y| = |X| \cdot |Y|.$$

Demostración. Las funciones

$$(x, y) \in X \times Y \mapsto x \in X, \qquad (x, y) \in X \times Y \mapsto y \in Y$$

son sobreyectivas, así que los conjuntos X e Y son finitos cuando el conjunto $X \times Y$ lo es.

Supongamos ahora que los conjuntos X e Y son finitos y sean m y n sus cardinales, de manera que hay funciones biyectivas $f : X \to [\![m]\!]$ y $g : Y \to [\![n]\!]$. Si $n = 0$, entonces el conjunto Y es vacío, así que $X \times Y$ también lo es, es por lo tanto finito, y su cardinal es $|X \times Y| = 0 = |X| \cdot |Y|$. Podemos suponer, gracias a esto, que en lo que sigue en el entero n es positivo.

Siempre que i y j son elementos de $[\![m]\!]$ y de $[\![n]\!]$, respectivamente, es $(i-1)n + j \in [\![mn]\!]$, así que hay una función

$$h : (x, y) \in X \times Y \mapsto (f(x) - 1)n + g(y) \in [\![mn]\!].$$

Para completar la prueba de la proposición es suficiente que mostremos que esta función es biyectiva.

- Sea k un elemento de $[\![mn]\!]$. Hay dos enteros a y b tales que $k - 1 = an + b$ y $0 \leq b < n$. Como $mn > k = an + b \geq an$ y $n > 0$, tenemos además que $m > a$. Por otro lado, es $a \geq 0$: si fuera $a < 0$ tendríamos que $k - 1 = an + b < -n + b < 0$, lo que es absurdo. Vemos así que $a + 1 \in [\![m]\!]$ y que $b + 1 \in [\![n]\!]$ y que, por lo tanto, hay elementos x e y de X y de Y, respectivamente, tales que $f(x) = a + 1$ y $g(y) = b + 1$. Como entonces

  $$h(x, y) = (f(x) - 1)n + g(y) = an + b + 1 = k,$$

  el entero k pertenece a la imagen de la función h, que es, en consecuencia, sobreyectiva.
- Sean ahora que (x, y) y (x', y') dos elementos de $X \times Y$ tales que $h(x, y) = h(x', y')$, de manera que

  $$(f(x) - 1)n + g(y) = (f(x') - 1)n + g(y'). \tag{3.3}$$

  Esto implica que $g(y) - g(y')$ es divisible por n: como además $-n < g(y) - g(y') < n$, debe ser entonces $g(y) = g(y')$ y, como la función g es inyectiva, $y = y'$. Por otro lado, como ahora sabemos que $g(y) = g(y')$, la igualdad (3.3) y el hecho de que $n > 0$ implican que $f(x) = f(x')$: como la función f también es inyectiva, vemos que $x = x'$ y, en definitiva, que $(x, y) = (x', y')$. Así, la función h es inyectiva. □

Podemos extender este resultado a productos cartesianos de un número arbitrario pero finito de factores.

Ejercicio 3.2.21. Sea k un entero no negativo y sean X_1, X_2, …, X_k conjuntos. Pruebe que el producto cartesiano $X_1 \times X_2 \times \cdots \times X_k$ es finito si y solamente si cada uno de esos k conjuntos lo es, y que cuando ese es el caso su cardinal es $|X_1 \times X_2 \times \cdots \times X_k| = |X_1| \cdot |X_2| \cdot \cdots \cdot |X_k|$.

Notemos que la afirmación de este ejercio permite la posibilidad de que el entero k sea 0: el «producto cartesiano de 0 factores» es un conjunto con exactamente un elemento, la 0-upla vacía.

Otra operación que preserva la clase de los conjuntos finitos es la exponenciación.

Proposición 3.2.22. *Sean X e Y dos conjuntos finitos. El conjunto X^Y de todas las funciones $X \to Y$ es finito y su cardinal es*

$$|Y^X| = |Y|^{|X|}.$$

En esta proposición y en todo lo que sigue convenimos en que $0^0 = 1$.

Demostración. Sean m y n los cardinales de X y de Y, respectivamente. Probaremos la proposición haciendo inducción con respecto a m. Si $m = 0$, entonces el conjunto X es vacío, y Y^X tiene exactamente un elemento, la función vacía $X \to Y$, así que $|Y^X| = 1 = |Y|^{|X|}$.

Supongamos ahora que el entero m es positivo, de manera que el conjunto X no es vacío, y que para cualquier conjunto finito Z de cardinal $m-1$ vale que Y^Z es finito de cardinal $|Y|^{|Z|}$. Sea x_0 un elemento de X y consideremos el conjunto $X' \coloneqq X \setminus \{x_0\}$ y la función

$$\phi : f \in Y^X \mapsto (f|_{X'}, f(x_0)) \in Y^{X'} \times Y.$$

Afirmamos que esta función es biyectiva.

- Si (g, y) un elemento de $Y^{X'} \times Y$, entonces hay una función $f : X \to Y$ que en cada elemento x de X toma el valor

$$f(x) = \begin{cases} y & \text{si } x = x_0; \\ g(x) & \text{si } x \neq x_0. \end{cases}$$

  Es claro que $\phi(g) = (f, y)$, y esto prueba que la función ϕ es sobreyectiva.
- Sean $f : X \to Y$ y $f' : X \to Y$ dos elementos de Y^X tales que $\phi(f) = \phi(f')$, de manera que $(f|_{X'}, f(x_0)) = (f'|_{X'}, f'(x_0))$. Si x es un elemento de X, entonces o bien $x = x_0$ y, por lo tanto, $f(x) = f(x_0) = f'(x_0) = f'(x)$, o bien $x \neq x_0$, $x \in X'$ y

$$f(x) = (f|_{X'})(x) = (f'|_{X'})(x') = f'(x).$$

  Esto nos dice que $f(x) = f'(x)$ cualquiera sea el elemento x de X y, por lo tanto, que $f = f'$.

En vista de esto y de la Proposición 3.2.20 tenemos que

$$|Y^X| = |Y^{X'} \times Y| = |Y^{X'}| \cdot |Y|$$

y, como $|X'| = m - 1$ por el Ejercicio 3.2.15, la hipótesis inductiva nos dice que esto es

$$= n^{m-1} \cdot n = n^m = |Y|^{|X|}.$$

Esto completa la inducción. □

Finalmente, mostremos que el conjunto de partes de un conjunto finito es finito.

Proposición 3.2.23. *El conjunto de partes $\mathscr{P}(X)$ de un conjunto X es finito si y solamente si el conjunto X mismo es finito, y en ese caso es*

$$|\mathscr{P}(X)| = 2^{|X|}.$$

Demostración. La función $x \in X \mapsto \{x\} \in \mathscr{P}(X)$ es inyectiva, así que el conjunto X es finito cuando $\mathscr{P}(X)$ lo es y, por lo tanto, la condición de la proposición es necesaria.

Supongamos ahora que el conjunto X es finito. Sabemos de la Proposición 3.2.22 que el conjunto $[\![2]\!]^X$ es entonces finito y que su cardinal es $2^{|X|}$, así que para completar la prueba de la proposición es suficiente que mostremos que $\mathscr{P}(X) \sim [\![2]\!]^X$.

Hay una función $\phi : [\![2]\!]^X \to \mathscr{P}(X)$ que en cada elemento f de $[\![2]\!]^X$ toma el valor

$$\phi(f) = \{x \in X : f(x) = 2\}.$$

Para hacer lo que queremos mostraremos que esta función es biyectiva.

Si A es un elemento de $\mathscr{P}(X)$, entonces hay una función $f : X \to [\![2]\!]$ tal que

$$f(x) = \begin{cases} 2 & \text{si } x \in A; \\ 1 & \text{si } x \notin A. \end{cases}$$

Es claro que $\phi(f) = A$, y esto nos dice que la función ϕ es sobreyectiva. Por otro lado, supongamos que f y g son dos elementos distintos de $[\![2]\!]^X$, de manera que hay un elemento $x \in X$ tal que $f(x) \neq g(x)$. Si $x \in \phi(f)$, entonces $f(x) = 2$ y, como $g(x) \in [\![2]\!] \setminus \{f(x)\}$, es $g(x) = 1$: esto nos dice que $x \notin \phi(g)$ y, por lo tanto, que $\phi(f) \neq \phi(g)$. Si, por el contrario, es $x \notin \phi(f)$ podemos proceder de exactamente la misma forma y ver que $x \in \phi(g)$ y, otra vez, que $\phi(f) \neq \phi(g)$. La función ϕ es, en definitiva, inyectiva. □

Ejercicio 3.2.24. Muestre que si X un conjunto finito y n es el cardinal de X, entonces el conjunto $S(X)$ de todas las funciones biyectivas $X \to X$ es finito y su cardinal es $n!$.

Ejercicio 3.2.25. Sean X e Y dos conjuntos finito y sean m y n los cardinales de X y de Y, respectivamente. Muestre que el conjunto $\mathsf{Iny}(X, Y)$ de todas las funciones inyectivas $X \to Y$ y el conjunto $\mathsf{Sob}(X, Y)$ de todas las funciones sobreyectivas $X \to Y$ son los dos finitos, y muestre que sus cardinales son

$$|\mathsf{Iny}(X, Y)| = \begin{cases} \dfrac{n!}{(n-m)!} & \text{si } m \leq n; \\ 0 & \text{en caso contrario;} \end{cases}$$

y

$$|\mathsf{Sob}(X, Y)| = \sum_{i=0}^{n-1} (-1)^i \binom{n}{i} (n-i)^m.$$

3.3. Conjuntos infinitos

Nos ocuparemos en esta sección de los conjuntos que no son finitos.

Definición 3.3.1. Un conjunto es ***infinito*** si no es finito.

Demos, antes que nada, un ejemplo de un conjunto infinito.

Proposición 3.3.2. *El conjunto $\mathbb{N}_0$ es infinito.*

Demostración. Empecemos por probar que para cada entero no negativo m vale que

$$\begin{aligned} &\textit{si } f : [\![m]\!] \to \mathbb{N}_0 \textit{ es una función, entonces hay un entero no negativo } k \textit{ tal que} \\ &f(i) \leq k \textit{ para todo } i \in [\![m]\!]. \end{aligned} \tag{3.4}$$

Procedemos por inducción con respecto a m, notando que cuando m es 0 la afirmación es evidente: podemos tomar $k = 0$ en ese caso, ya que el conjunto $[\![m]\!]$ es vacío.

Supongamos ahora que p es un entero no negativo y que la afirmación (3.4) es cierta cuando m es p, y sea $f : [\![p+1]\!] \to \mathbb{N}_0$ una función. Podemos considerar la restricción $g \coloneqq f|_{[\![p]\!]} : [\![p]\!] \to \mathbb{N}_0$ de f a $[\![p]\!]$: la hipótesis inductiva implica que hay un entero no negativo k' tal que $g(i) \leq k'$ para todo $i \in [\![p]\!]$, y podemos poner $k \coloneqq \max\{k', f(p+1)\}$, que también es un entero no negativo. Sea i un elemento de $[\![p+1]\!]$: si $i \leq p$, entonces $f(i) = g(i) \leq k' \leq k$, y si $i = p+1$ entonces $f(i) = f(p+1) \leq k$. Vemos así que $f(i) \leq k$ para todo $i \in [\![p+1]\!]$. Esto prueba que la afirmación (3.4) también es cierta cuando m es $p+1$ y completa la inducción.

Probemos ahora la proposición. En vista de Proposición 3.2.11, para ver que el conjunto $\mathbb{N}_0$ es infinito es suficiente que mostremos que no hay ningún entero no negativo n tal que existe una función sobreyectiva $[\![n]\!] \to \mathbb{N}_0$, y esto es fácil. En efecto, si n es un elemento de $\mathbb{N}_0$ y $f : [\![n]\!] \to \mathbb{N}_0$

es una función, la afirmación (3.4) que probamos arriba nos dice que hay un entero no negativo k tal que $f(i) \leq k$ para todo $i \in [\![n]\!]$ y, por lo tanto, tal que $k+1 \neq f(i)$ cualquiera sea $i \in [\![n]\!]$: la función f no es, por lo tanto, sobreyectiva. □

La siguiente es una propiedad útil del conjunto $\mathbb{N}_0$ que usaremos varias veces.

Lema 3.3.3. *El conjunto $\mathbb{N}_0$ es equipotente a cualquiera de sus subconjuntos infinitos.*

Demostración. Sea X un subconjunto infinito de $\mathbb{N}_0$. Como en particular no es vacío, podemos considerar su menor elemento $x_0 \coloneqq \min X$. Por otro lado, si n es un elemento de X, entonces el conjunto $[\![0, n]\!] \cap X$ es finito, así que es distinto de X y tiene sentido considerar el entero $\min X \setminus [\![0, n]\!]$. Hay, por lo tanto, una función

$$F : n \in X \mapsto \min X \setminus [\![0, n]\!] \in X.$$

El *Principio de recurrencia* 1.2.1 nos dice entonces que hay una función $f : \mathbb{N}_0 \to X$ tal que $f(0) = x_0$ y $f(n+1) = F(f(n))$ para cada $n \in \mathbb{N}_0$. Para probar el lema es suficiente que mostremos que esta función f es biyectiva.

- Si n es un elemento de $\mathbb{N}_0$, entonces $f(n+1) = F(f(n)) = \min X \setminus [\![0, f(n)]\!]$, así que $f(n+1) \notin [\![0, f(n)]\!]$ y, por lo tanto, como $f(n+1)$ ciertamente pertenece a $\mathbb{N}_0$, tenemos que $f(n) < f(n+1)$. Esto implica, como sabemos, que la función f es estrictamente creciente y, en consecuencia, inyectiva.
- Supongamos que la función f no es sobreyectiva, de manera que el conjunto $X \setminus f(\mathbb{N}_0)$ no es vacío y, por lo tanto, que podemos considerar su menor elemento $k \coloneqq \min X \setminus f(\mathbb{N}_0)$. Como $x_0 = f(0) \in f(\mathbb{N}_0)$, es claro que $k \neq x_0 = \min X$ y, como $k \in X$, esto implica que el conjunto $X \cap [\![0, k-1]\!]$ no es vacío. Como además es finito, podemos considerar su mayor elemento $l \coloneqq \max X \cap [\![0, k-1]\!]$. Por supuesto, es $l < k$ y $l \in X$, así que $l \in f(\mathbb{N}_0)$, esto es, hay un elemento r de $\mathbb{N}_0$ tal que $f(r) = l$. Ahora bien, de acuerdo a la propiedad característica de la función f, tenemos que

  $$f(r+1) = F(f(r)) = \min X \setminus [\![0, f(r)]\!] = \min X \setminus [\![0, l]\!] = \min X \setminus [\![0, k-1]\!] = k.$$

  Esto es absurdo, ya que k no pertenece a la imagen de f, y esta contradicción provino de haber supuesto que la función f no es sobreyectiva. Debe serlo, entonces.

Como dijimos, esto completa la demostración del lema. □

La definición que dimos arriba de qué significa que un conjunto sea infinito es *negativa*: un conjunto X es infinito si *no existe* algo — una biyección entre X y un conjunto de la forma $[\![n]\!]$ con $n \in \mathbb{N}_0$. Para muchas cosas es conveniente tener una caracterización *positiva* de los conjuntos infinitos, esto es, que dependa de que algo *exista*. El siguiente resultado nos da una y, además, nos

dice que el cardinal de $\mathbb{N}_0$ es el menor de los cardinales infinitos.

Proposición 3.3.4. *Un conjunto X es infinito si y solamente si $\mathbb{N}_0 \leq X$.*

La condición que da esta proposición es que haya una función inyectiva $\mathbb{N}_0 \to X$ y, de acuerdo a la Proposición 3.1.6, esto es equivalente a que exista una función sobreyectiva $X \to \mathbb{N}_0$.

Demostración. Probaremos primero la necesidad de la condición. Sea X un conjunto infinito. Empezaremos mostrando que

$$\begin{aligned}&\textit{si } n \textit{ es un elemento de } \mathbb{N}_0 \textit{ y } f : [\![n]\!] \to X \textit{ es una función inyectiva, entonces hay}\\ &\textit{una función inyectiva } g : [\![n+1]\!] \to X \textit{ tal que } g|_{[\![n]\!]} = f.\end{aligned} \tag{3.5}$$

Sea para ello n un elemento de $\mathbb{N}_0$ y sea $f : [\![n]\!] \to X$ una función inyectiva. De acuerdo a la Proposición 3.2.11, la función f no es sobreyectiva, así que hay un elemento x de X tal que $f(i) \neq x$ para todo $i \in [\![n]\!]$. Consideremos la función $g : [\![n+1]\!] \to X$ que en cada $i \in [\![n+1]\!]$ toma el valor

$$g(i) = \begin{cases} f(i) & \text{si } i \leq n; \\ x & \text{si } i = n+1. \end{cases}$$

Es claro que la restricción de g a $[\![n]\!]$ es f. Sean, por otro lado, i y j dos elementos de $[\![n]\!]$ tales que $i < j$. Si $j \leq n$, entonces también $i \leq n$ y es $g(i) = f(i) \neq f(j) = g(i)$, ya que la función f es inyectiva. Si, en cambio, $j = n+1$, entonces $i \in [\![n]\!]$ y $g(i) = f(i) \neq x = g(j)$ por la forma en que elegimos a x. Vemos así que la función g es inyectiva. Esto prueba (3.5).

Como consecuencia de (3.5), de que la función vacía $[\![0]\!] \to X$ es inyectiva, y del *Principio de Elección Dependiente*, podemos concluir que hay una sucesión de funciones $(f_n)_{n\geq 0}$ que satisface las siguientes dos condiciones:

- La función f_0 es la función vacía $[\![0]\!] \to X$.
- Para cada $n \in \mathbb{N}_0$ la función $f_n : [\![n]\!] \to X$ es inyectiva.
- Para cada $n \in \mathbb{N}_0$ es $f_{n+1}|_{[\![n]\!]} = f_n$.

Mostremos que, más generalmente, esta sucesión satisface la siguiente tercera condición:

- Si m y n son dos elementos de $\mathbb{N}_0$ tales que $m \leq n$, entonces $f_n|_{[\![m]\!]} = f_m$.

Para hacerlo es suficiente que probemos inductivamente que para todo $k \in \mathbb{N}_0$ vale que

$$\textit{para todo } m \in \mathbb{N}_0 \textit{ es } f_{m+k}|_{[\![m]\!]} = f_m.$$

Esto es evidente si $k = 0$. Si l es un elemento de $\mathbb{N}_0$ tal que para todo $m \in \mathbb{N}_0$ es $f_{m+l}|_{[\![m]\!]} = f_m$, entonces para cada $m \in \mathbb{N}_0$ es $f_{m+(l+1)}|_{[\![m]\!]} = f_{(m+1)+l}|_{[\![m]\!]} = f_m$: esto completa la inducción.

Para completar la prueba de que la condición de la proposición es necesaria, mostraremos que la función

$$f : n \in \mathbb{N}_0 \mapsto f_n(n) \in X$$

es inyectiva. Sean i y j dos elementos de $\mathbb{N}_0$ tales que $f(i) = f(j)$, y consideremos el entero $k \coloneqq \text{máx}\{i, j\}$, que es un elemento de $\mathbb{N}_0$. Como $i \leq k$ y $j \leq k$, tenemos que

$$f(i) = f_i(i) = (f_k)|_{[\![i]\!]}(i) = f_k(i), \qquad f(j) = f_j(j) = (f_k)|_{[\![j]\!]}(j) = f_k(j),$$

de manera que $f_k(i) = f_k(j)$ y, como la función f_k es inyectiva, que $i = j$. Esto prueba lo que queremos.

Tenemos que probar ahora que la condición de la proposición es suficiente. Sea X un conjunto, supongamos que hay una función inyectiva $f : \mathbb{N}_0 \to X$ y, para llegar a una contradicción, que el conjunto X es finito, de manera que hay un elemento n de $\mathbb{N}_0$ y una función biyectiva $g : X \to [\![n]\!]$. Se sigue de esto, claro, que la composición $g \circ f : \mathbb{N}_0 \to [\![n]\!]$ es inyectiva, y esto contradice el resultado del Ejercicio 3.2.13: esta es la contradicción que buscábamos. □

Como la Proposición 3.3.4 es importante, daremos ahora una segunda demostración de que la condición que da es necesaria, esta basada en el *Lema de Zorn*.

Una segunda demostración de la necesidad de la condición de la Proposición 3.3.4. Sea X un conjunto infinito, y consideremos el conjunto $\mathscr{S}$ de todas las funciones inyectivas $I \to X$ definidas en un subconjunto I de $\mathbb{N}_0$. Para cada par de elementos $f : I \to X$ y $g : J \to X$ de $\mathscr{S}$ escribamos $f \preceq g$ si $I \subseteq J$ y $g|_I = f$. Es fácil ver que esta relación $\preceq$ en $\mathscr{S}$ es un orden.

Queremos probar que el conjunto $\mathscr{S}$ dotado de ese orden satisface las hipótesis del *Lema de Zorn* 2.2.6. Antes que nada, es claro que ese conjunto no es vacío, ya que la función vacía $\varnothing \to X$ es inyectiva y está definida en un subconjunto de $\mathbb{N}$.

Sea ahora $\mathscr{C}$ una cadena en $\mathscr{C}$. Como cada elemento de $\mathscr{C}$ es un subconjunto de $\mathbb{N}_0 \times X$, tiene sentido considerar la unión $F \coloneqq \bigcup_{f \in \mathscr{C}} f$, que es un subconjunto de $\mathbb{N}_0 \times X$. Pongamos

$$K \coloneqq \{n \in \mathbb{N}_0 : \text{existe } x \in X \text{ tal que } (n, x) \in F\}$$

y mostraremos que el conjunto F es, de hecho, una función inyectiva $K \to X$.

- Ciertamente, de la definición del conjunto K es claro que $F \subseteq K \times X$, y que para cada $n \in K$ existe $x \in X$ tal que $(n, x) \in F$.
- Supongamos que n es un elemento de K, que x y x' son dos elementos de X, y que los pares ordenados (n, x) y (n, x') pertenecen a F. Hay entonces dos elementos f y g de $\mathscr{C}$ tales que $(n, x) \in f$ y $(n, x') \in g$ y, como el conjunto $\mathscr{C}$ es una cadena de $\mathscr{S}$, tenemos que $f \subseteq g$ o que $g \subseteq f$. En el primer caso los dos pares (n, x) y (n, x') pertenecen a g y, como g es una función, es $x = x'$; en el segundo, esos dos pares pertenecen a f y, como esta también es una función, otra vez es $x = x'$.
- Sean ahora m y n dos elementos de K y supongamos que $F(m) = F(n)$. Como los pares ordenados $(m, F(m))$ y $(n, F(n))$ pertenecen a F, hay dos elementos f y g de $\mathscr{C}$ tales que

$(m, F(m)) \in f$ y $(n, F(n)) \in g$, y como el conjunto $\mathscr{C}$ es una cadena, tenemos que $f \subseteq g$ o $g \subseteq f$. En el primer caso, los dos pares $(m, F(m))$ y $(n, f(n))$ están en g y, como g es una función inyectiva, es $m = n$; en el segundo caso, esos dos pares están en f y, como f también es una función inyectiva, podemos llegar a la misma conclusión.

Vemos así que $F : K \to X$ es un elemento de $\mathscr{S}$. Para terminar de probar que en esta situación podemos usar el *Lema de Zorn* es suficiente que mostremos que F es una cota superior para la cadena $\mathscr{C}$ en $\mathscr{S}$.

- Si $f : I \to X$ es un elemento de $\mathscr{C}$, entonces $I \subseteq F$ y para cada elemento n de I tenemos que $(n, f(n)) \in f \subseteq F$, de manera que $f(n) = F(n) = F|_I(n)$: esto nos dice que $f = F|_I$ y, por lo tanto, que F es una cota superior para la cadena $\mathscr{C}$ en $\mathscr{S}$, como queremos.

En vista de todo esto, el *Lema de Zorn* 2.2.6 nos dice que hay un elemento $h : L \to X$ en $\mathscr{S}$ que es maximal con respecto a la relación de orden $\preceq$. Afirmamos que el dominio L de esta función h es infinito. Para probarlo, supongamos que, por el contrario, el conjunto L es finito. Como $\mathbb{N}_0$ es infinito, se sigue de esto que $L \subsetneq \mathbb{N}_0$ y que, por lo tanto, podemos elegir un elemento n_0 de $\mathbb{N}_0 \setminus L$. Por otro lado, como el conjunto X es infinito, la función h no es sobreyectiva y podemos elegir un elemento x_0 de $X \setminus h(L)$. Hay una función $k : L \cup \{n_0\} \to X$ que en cada elemento n de $L \cup \{n_0\}$ toma el valor

$$k(n) = \begin{cases} h(n) & \text{si } n \in L; \\ x_0 & \text{si } n = n_0. \end{cases}$$

Es fácil ver que esta función es inyectiva, así que se trata de un elemento de $\mathscr{S}$. Más aún, como $L \subseteq L \cup \{n_0\}$ y claramente $k|_L = h$, tenemos que $h \preceq k$ y, como $n_0 \notin L$, que más aún $h \prec k$. Esto es absurdo, ya que h es un elemento maximal del conjunto ordenado $\mathscr{S}$.

La conclusión de esto es que hay una función inyectiva $h : I \to X$ definida en un subconjunto infinito I de $\mathbb{N}_0$. De acuerdo al Lema 3.3.3, hay una función biyectiva $i : \mathbb{N}_0 \to X$ y la composición $h \circ i : \mathbb{N}_0 \to X$ es inyectiva. Vemos así que $\mathbb{N}_0 \preceq X$, como queremos. □

El siguiente resultado generaliza el de la Proposición 3.3.4.

Corolario 3.3.5. *Un conjunto X es infinito si y solamente si hay un conjunto infinito Y tal que $Y \preceq X$.*

Demostración. Como el conjunto $\mathbb{N}_0$ es infinito, que la condición que da el corolario es necesaria es consecuencia de la Proposición 3.3.4. Por otro lado, si X e Y son dos conjuntos tales que $Y \preceq X$ e Y es infinito, entonces esa proposición nos dice que $\mathbb{N}_0 \preceq Y$ y, por lo tanto, que $\mathbb{N}_0 \preceq X$, así que otra vez esa proposición nos permite concluir que el conjunto X es infinito. □

La proposición que acabamos de probar hace fácil probar una segunda caracterización de los conjuntos infinitos que es históricamente importante.

Proposición 3.3.6. *Un conjunto es infinito si y solamente si es equipotente a uno de sus subconjuntos propios.*

Esta caracterización de los conjuntos infinitos es debida a Richard Dedekind y fue propuesta por él en 1888 como *definición*. Tiene el interés de que es la primera caracterización de los conjuntos infinitos que no involucra al conjunto de los números naturales.

Demostración. Para ver que la condición que da la proposición es suficiente alcanza con mostrar que un conjunto finito no es equipotente a ninguno de sus subconjuntos propios, y esto es una de las cosas que afirma la Proposición 3.2.14. Probemos que esa condición también es necesaria.

Supongamos que X es un conjunto infinito, de manera que, de acuerdo a la Proposición 3.3.4, hay una función inyectiva $f : \mathbb{N}_0 \to X$. Sea $I \coloneqq f(\mathbb{N}_0)$ la imagen de la función f. La correstricción $f|^I : \mathbb{N}_0 \to I$ de f a I es una función biyectiva, ya que la función f es inyectiva, y posee, por lo tanto, una función inversa $g \coloneqq (f|^I)^{-1} : I \to \mathbb{N}_0$. Consideremos el conjunto

$$Y \coloneqq X \setminus \{f(0)\},$$

que ciertamente es un subconjunto propio de X. Es fácil ver que hay una función $h : X \to Y$ que en cada elemento x de X toma el valor

$$h(x) = \begin{cases} f(g(x)+1) & \text{si } x \in I; \\ x & \text{si } x \notin I. \end{cases}$$

Notemos que

$$h(I) \subseteq I, \qquad h(X \setminus I) \subseteq X \setminus I. \tag{3.6}$$

Para probar que la condición que da la proposición es necesaria bastará que mostremos que esta función h es biyectiva.

- Sea y un elemento de Y.
  - Si y no pertenece a I, entonces $y = h(y)$ y vemos que y pertenece a la imagen de h.
  - Si, por el contrario, y sí pertenece a I, entonces $g(y) > 0$ porque $y \in Y$, el entero $n \coloneqq g(y) - 1$ es un elemento de $\mathbb{N}_0$ y, como por supuesto $f(n) \in I$, es

    $$y = f(g(y)) = f(n+1) = f(g(f(n)) + 1) = h(f(n)),$$

    así que y también pertenece a la imagen de h en este caso.

  Vemos así que la función h es sobreyectiva.
- Sean ahora x y x' dos elementos de X y supongamos que $h(x) = h(x')$. En vista de (3.6), no puede ser que exactamente uno de x o x' pertenezca a I, así que o bien los dos pertenecen a I o ninguno de los dos lo hace. Si tanto x como x' están en I, entonces

  $$x = f(g(x)) = f(g(f(g(x)+1)) - 1) = f(g(h(x)) - 1)$$

$$= f(g(h(x')) - 1) = f(g(f(g(x') + 1)) - 1) = f(g(x')) = x'.$$

Si, en cambio, ninguno de los dos están en I, entonces $x = h(x) = h(x') = x'$. Podemos concluir con esto que la función h es inyectiva.

Como dijimos, esto completa la prueba de la proposición. □

Es una consecuencia inmediata de la Proposición 3.2.14 que si X es un conjunto finito y no vacío y x es un elemento de X, entonces $X \setminus \{x\} \prec X$. Esta propiedad caracteriza a los conjuntos finitos, de hecho.

Ejercicio 3.3.7. Muestre que si X es un conjunto infinito y F es un subconjunto finito de X, entonces $X \setminus F \sim X$.

La Proposición 3.2.20 nos dice que un producto cartesiano de finitos conjuntos finitos es finito. El siguiente resultado nos dice, más generalmente, cuándo un producto cartesiano es finito.

Proposición 3.3.8. *Sea $(X_i)_{i\in I}$ una familia de conjuntos no vacíos. El producto cartesiano $\prod_{i\in I} X_i$ es finito si y solamente si para cada $i \in I$ el conjunto X_i es finito y, además, el conjunto*

$$D \coloneqq \{i \in I : X_i \text{ tiene al menos dos elementos}\}$$

es finito.

Demostración. Para cada elemento i de I podemos elegir un elemento a_i en X_i, ya que este conjunto no es vacío, y si además i pertenece al conjunto D del enunciado podemos elegir otro elemento b_i en X_i tal que $a_i \neq b_i$, ya que en ese caso X_i tiene al menos dos elementos.

Supongamos que el conjunto $\prod_{i\in I} X_i$ es finito. Si j es un elemento de I, entonces la función

$$x \in \prod_{i\in I} X_i \mapsto x_j \in X_j$$

es sobreyectiva, así que el conjunto X_j es finito. Hay, por otro lado, una función $\psi : D \to \prod_{i\in I} X_i$ tal que para cada $i \in D$ y cada $j \in I$ es

$$\psi(i)_j = \begin{cases} b_j & \text{si } j = i; \\ a_j & \text{si } j \neq i. \end{cases}$$

Esta función es inyectiva: siempre que i e i' son dos elementos distintos de D es $\psi(i) \neq \psi(i')$, ya que $\psi(i)_i = b_i \neq a_i = \psi(i')_i$. Es, por lo tanto, $D \preceq \prod_{i\in I} X_i$ y el conjunto D es finito. Vemos así que la condición de la proposición es necesaria.

Para ver que también es suficiente, supongamos que cualquiera sea el elemento i de I el conjunto X_i es finito y que, además, el conjunto D del enunciado es finito. De acuerdo al resultado

del Ejercicio 3.2.21, el producto $\prod_{i\in D} X_i$ es un conjunto finito. Por otro lado, hay una función

$$\psi : x \in \prod_{i\in I} X_i \mapsto x|_D \in \prod_{i\in D} X_i.$$

Para terminar la demostración de la proposición alcanza con que mostremos que esta función ψ es inyectiva. Sean x y x' dos elementos de $\prod_{i\in I} X_i$ y supongamos que $\psi(x) = \psi(x')$, de manera que es $x_i = x'_i$ para cada $i \in D$. Como además para cada $i \in I \setminus D$ es $x_i = a_i = x'_i$, porque a_i es el único elemento de conjunto X_i, tenemos entonces que, de hecho, es $x = x'$, como queremos. □

3.4. Conjuntos numerables

El estudio sistemático de la teoría de conjuntos tiene uno de sus puntos de partida en la correspondencia entre Georg Cantor y Richard Dedekind de 1873 en la que discutieron y, en definitiva, desarrollaron la noción de numerabilidad. Ese es el tema de esta sección.

Probamos en la sección anterior que un conjunto X es infinito si y solamente si $\mathbb{N}_0 \preceq X$. Esto nos dice que el cardinal de $\mathbb{N}_0$ es el menor de los cardinales infinitos y es esperable que, por ello, tenga un rol especial en la teoría.

Definición 3.4.1. Escribimos $\aleph_0$ al cardinal de $\mathbb{N}_0$, y decimos que un conjunto es ***numerable*** si su cardinal es $\aleph_0$.

De la misma forma en que no todo conjunto es finito, no todo conjunto infinito es numerable. El primero en observar esto fue Cantor mismo, en [Can1874]. Demos un ejemplo, originalmente dado por Cantor en [Can1892].

Proposición 3.4.2. *El conjunto $[\![2]\!]^{\mathbb{N}_0}$ es infinito y no numerable.*

La idea de la demostración que daremos para esto es conocida como el ***argumento diagonal***.

Demostración. Para cada $m \in \mathbb{N}_0$ hay una función $\delta_m : \mathbb{N}_0 \to [\![2]\!]$ que en cada $n \in \mathbb{N}_0$ toma el valor

$$\delta_m(n) = \begin{cases} 1 & \text{si } n = m; \\ 2 & \text{en caso contrario.} \end{cases}$$

Podemos, por lo tanto, considerar la función $f : m \in \mathbb{N}_0 \mapsto \delta_m \in [\![2]\!]^{\mathbb{N}_0}$. Esta función es inyectiva, ya que si m y n son dos elementos distintos de $\mathbb{N}_0$ es $f(m)(m) = \delta_m(m) = 1 \neq 2 = \delta_n(m) = f(n)(m)$, y, por lo tanto, el conjunto $[\![2]\!]^{\mathbb{N}_0}$ es infinito.

Supongamos por un momento que es además numerable, de manera que hay una función sobreyectiva $g : \mathbb{N}_0 \to [\![2]\!]^{\mathbb{N}_0}$. Podemos entonces considerar la función $h : n \in \mathbb{N}_0 \mapsto 3 - g(n)(n) \in [\![2]\!]$. Como h es un elemento de $[\![2]\!]^{\mathbb{N}_0}$ y la función g es sobreyectiva, podemos elegir k en $\mathbb{N}_0$ de manera que sea $h = g(k)$: se tiene entonces que $3 - g(k)(k) = h(k) = g(k)(k)$ o, equivalentemente, que $2 \cdot g(k)(k) = 3$. Esto es claramente imposible, y esta contradicción provino de haber supuesto que el conjunto $[\![2]\!]^{\mathbb{N}_0}$ es numerable, así que no lo es. □

Usando la Proposición 3.4.2 podemos dar un segundo ejemplo mucho más interesante de un conjunto infinito no numerable.

Proposición 3.4.3. *Es $\mathbb{R} \sim [\![2]\!]^{\mathbb{N}_0}$, así que el conjunto $\mathbb{R}$ es infinito y no numerable.*

Escribimos $\mathfrak{c}$ al cardinal de $\mathbb{R}$ y lo llamamos la ***potencia del continuo***. Muchas veces decimos que un conjunto ***tiene la potencia del continuo*** en lugar de decir que su cardinal es $\mathfrak{c}$. Notemos que la proposición nos dice que $\aleph_0 \prec \mathfrak{c}$. Cantor probó que $\mathbb{R}$ no es un conjunto numerable en [Can1874], pero la siguiente demostración es esencialmente la que Cantor dio más tarde en [Can1892].

Primera demostración. Si $\phi : \mathbb{N}_0 \to [\![2]\!]$ es un elemento de $[\![2]\!]^{\mathbb{N}_0}$, entonces la serie

$$\sum_{n=0}^{\infty} \frac{\phi(n) - 1}{3^n}$$

converge absolutamente — esto es consecuencia inmediata del criterio de comparación y de que una serie geométrica de razón 1/3 converge. Podemos entonces definir una función

$$f : \phi \in [\![2]\!]^{\mathbb{N}_0} \mapsto \sum_{n=0}^{\infty} \frac{\phi(n) - 1}{3^n} \in \mathbb{R}.$$

Queremos probar que esta función es inyectiva. Sean $\phi : \mathbb{N}_0 \to [\![2]\!]$ y $\psi : \mathbb{N}_0 \to [\![2]\!]$ dos elementos distintos de $[\![2]\!]^{\mathbb{N}_0}$, de manera que el conjunto $S \coloneqq \{n \in \mathbb{N}_0 : \phi(n) \neq \psi(m)\}$ no es vacío y podemos considerar su menor elemento $m \coloneqq \min S$. Como m pertenece a S, es $|\phi(m) - \psi(m)| = 1$. Por otro lado, para cada $n \in \mathbb{N}_0$ tal que $n \geq m + 1$ es $|\phi(n) - \psi(n)| \leq 1$, así que

$$\left| \sum_{n=m+1}^{\infty} \frac{\phi(n) - \psi(n)}{3^n} \right| \leq \sum_{n=m+1}^{\infty} \frac{|\phi(n) - \psi(n)|}{3^n} \leq \sum_{n=m+1}^{\infty} \frac{1}{3^n} = \frac{1}{2 \cdot 3^m}.$$

Como $\phi(n) = \psi(n)$ para todo entero $n \in \mathbb{N}_0$ tal que $n < m$, es

$$f(\phi) - f(\psi) = \sum_{n=0}^{\infty} \frac{\phi(n) - 1}{3^n} - \sum_{n=0}^{\infty} \frac{\psi(n) - 1}{3^n} = \frac{\phi(m) - \psi(m)}{3^m} + \sum_{n=m+1}^{\infty} \frac{\phi(n) - \psi(n)}{3^n}$$

y entonces

$$|f(\phi) - f(\psi)| = \left| \frac{\phi(m) - \psi(m)}{3^m} + \sum_{n=m+1}^{\infty} \frac{\phi(n) - \psi(n)}{3^n} \right|$$

$$\geq \left|\frac{\phi(m)-\psi(m)}{3^m}\right| - \left|\sum_{n=m+1}^{\infty} \frac{\phi(n)-\psi(n)}{3^n}\right| \geq \frac{1}{3^m} - \frac{1}{2\cdot 3^m} = \frac{1}{3^m} > 0,$$

por lo que $f(\phi) \neq f(\psi)$. Esto prueba lo que queremos. Hemos construido de esta manera una función inyectiva $[\![2]\!]^{\mathbb{N}_0} \to \mathbb{R}$ y, por lo tanto, probado que

$$[\![2]\!]^{\mathbb{N}_0} \preceq \mathbb{R}. \tag{3.7}$$

Si $\phi : \mathbb{N}_0 \to [\![2]\!]$ es un elemento de $[\![2]\!]^{\mathbb{N}_0}$, entonces la serie

$$\sum_{n=0}^{\infty} \frac{\phi(n)-1}{2^{n+1}}$$

converge absolutamente. Más aún, como $0 \leq \phi(n) - 1 \leq 1$ cualquiera sea $n \in \mathbb{N}_0$, tenemos que

$$0 = \sum_{n=0}^{\infty} \frac{0}{2^{n+1}} \leq \sum_{n=0}^{\infty} \frac{\phi(n)-1}{2^{n+1}} \leq \sum_{n=0}^{\infty} \frac{1}{2^{n+1}} = 1.$$

Hay entonces una función

$$g : \phi \in [\![2]\!]^{\mathbb{N}_0} \mapsto \sum_{n=0}^{\infty} \frac{\phi(n)-1}{2^{n+1}} \in [0,1].$$

Mostremos que esta es sobreyectiva. Sea ξ un elemento de $[0,1]$. Pongamos $D \coloneqq \{0,1\}$. Si x es un elemento de $[0,2)$, entonces $2(x - \lfloor x \rfloor) \in [0,2)$ y $\lfloor x \rfloor \in D$, así que hay una función

$$F : (x,y) \in [0,2) \times D \mapsto (2(x - \lfloor x \rfloor), \lfloor x \rfloor) \in [0,2) \times D.$$

De acuerdo al *Principio de recurrencia*, hay una función $d : \mathbb{N}_0 \to [0,2) \times D$ tal que $d(0) = (\xi, 0)$ y $d(n+1) = F(d(n))$ para cada $n \in \mathbb{N}_0$. Hay además dos funciones $\alpha : \mathbb{N}_0 \to [0,2)$ y $\beta : \mathbb{N}_0 \to D$ tales que $d(n) = (\alpha(n), \beta(n))$ para cada $n \in \mathbb{N}_0$. Si $n \in \mathbb{N}_0$, es

$$(\alpha(n+1), \beta(n+1)) = d(n+1) = F(d(n)) = F(\alpha(n), \beta(n)) = (2(\alpha(n) - \lfloor \alpha(n) \rfloor), \lfloor \alpha(n) \rfloor),$$

así que $\alpha(n+1) = 2(\alpha(n) - \lfloor \alpha(n) \rfloor)$ y $\beta(n+1) = \lfloor \alpha(n) \rfloor$, de manera que, de hecho,

$$\frac{\alpha(n+1)}{2} + \beta(n+1) = \alpha(n).$$

Es fácil ver que esto implica que para todo $n \in \mathbb{N}_0$ es

$$\alpha(0) - \sum_{m=1}^{n} \frac{\beta(m)}{2^m} = \frac{\alpha(n)}{2^{n+1}}.$$

Es $(\alpha(0), \beta(0)) = d(0) = (\xi, 0)$, así que $\alpha(0) = \xi$, y para todo $n \in \mathbb{N}_0$ es $0 \leq \alpha(n) < 2$, así que

$$\left|\xi - \sum_{m=1}^{n} \frac{\beta(m)}{2^m}\right| = \left|\alpha(0) - \sum_{m=1}^{n} \frac{\beta(m)}{2^m}\right| = \frac{|\alpha(n)|}{2^{n+1}} < \frac{1}{2^n}.$$

Esto implica inmediatamente que si ϕ es la función $n \in \mathbb{N}_0 \mapsto \beta(m+1)+1 \in [\![2]\!]$, entonces

$$g(\phi) = \sum_{n=0}^{\infty} \frac{\phi(n)-1}{2^{n+1}} = \sum_{n=0}^{\infty} \frac{\beta(n+1)}{2^{n+1}} = \sum_{n=1}^{\infty} \frac{\beta(n)}{2^n} = \xi.$$

Esto prueba, como queremos, que la función $g : [\![2]\!]^{\mathbb{N}_0} \to [0,1]$ es sobreyectiva. Como también es sobreyectiva la función

$$h : t \in (0,1) \mapsto \operatorname{tg} \pi(t - \tfrac{1}{2}) \in \mathbb{R},$$

tenemos que $\mathbb{R} \preceq (0,1) \preceq [0,1] \preceq [\![2]\!]^{\mathbb{N}_0}$. Esto y la desigualdad de (3.7) nos permiten concluir que $[\![2]\!]^{\mathbb{N}_0} \sim \mathbb{R}$, como queremos. □

Queremos dar, antes de seguir, una segunda demostración de la Proposición 3.4.3, esta vez siguiendo de cerca el argumento original que Cantor presenta en [Can1874]. Es útil saber que en el segundo volumen de [Ewa1996] hay una traducción al inglés de ese artículo.

Una segunda demostración de la Proposición 3.4.3. Que el conjunto $\mathbb{R}$ es infinito es claro, porque contiene a $\mathbb{N}_0$, así que tenemos que probar es que no es numerable. Supongamos para esto que hay una función biyectiva $f : \mathbb{N}_0 \to \mathbb{R}$, y sea S el conjunto de todas los 4-uplas (a, b, m, n) de elementos de $\mathbb{R} \times \mathbb{R} \times \mathbb{N}_0 \times \mathbb{N}_0$ tales que $a < f(m) < f(n) < b$ y para las que el conjunto

$$P(a,b,m,n) \coloneqq \{f(i) : i \in \mathbb{N}_0, a < f(i) < b, i \leq \max\{m,n\}\}$$

está contenido en $\{f(m), f(n)\}$. Esta segunda condición significa, intuitivamente, que $f(m)$ y $f(n)$ son los dos «primeros» valores que toma la función f en el intervalo (a,b).

Sean a y b dos elementos de $\mathbb{R}$ tales que $a < b$, y sean

$$u \coloneqq \min\{i \in \mathbb{N} : a < f(i) < b\}, \qquad v \coloneqq \min\{i \in \mathbb{N} : a < f(i) < b, f(i) \neq f(u)\}.$$

Notemos que esto tiene sentido, ya que el intervalo (a,b) tiene por lo menos dos elementos y la función f es sobreyectiva. Como $a < f(v) < b$, es claro que $u \leq v$, y la definición de v implica que, además, $u \neq v$, así que $u < v$; además, claramente es $f(u) \neq f(v)$. Consideremos los enteros

$$m \coloneqq \begin{cases} u & \text{si } f(u) < f(v); \\ v & \text{en caso contrario;} \end{cases} \qquad n \coloneqq \begin{cases} v & \text{si } f(u) < f(v); \\ u & \text{en caso contrario.} \end{cases}$$

Es claro que $a < f(m) < f(n) < b$. Sea, por otro lado, c un elemento de $P(a,b,m,n)$, de manera que $a < c < b$ y hay un elemento i de $\mathbb{N}$ tal que $c = f(i)$ y $i \leq \max\{m,n\}$. Como $f(u) \neq f(v)$ hay dos posibilidades a considerar.

- Puede ser que sea $f(u) < f(v)$, de manera que $m = u$, $n = v$ e $i \leq \max\{m,n\} = v$. Si es $i < v$, entonces la definición de v implica que $c = f(i) = f(u) = f(m)$, ya que $a < f(i) < b$. Si en cambio es $i = v$ tenemos, por supuesto, que $c = f(i) = f(v) = f(n)$.

- Puede ser que sea $f(u) > f(v)$, de manera ahora que $m = v$, $n = u$ e $i \leq \max\{m, n\} = v$. Si $i < v$, entonces $c = f(i) = f(u) = f(n)$, porque $a < f(i) < b$, y si $i = v$, entonces $c = f(i) = f(v) = f(m)$.

Vemos así que $P(a, b, m, n) \subseteq \{f(m), f(n)\}$ y, por lo tanto, que $(a, b, m, n) \in S$.

Esto prueba que el conjunto S no es vacío. Más aún, se sigue de lo que hemos hecho que para cada 4-upla (a, b, m, n) de S el conjunto $R(a, b, m, n) \coloneqq \{(c, d, p, q) \in S : c = f(m), d = f(n)\}$ no es vacío. El *Principio de elección dependiente*, en la forma del resultado del Ejercicio 2.1.7, implica entonces que hay una sucesión

$$\big((a_k, b_k, m_k, n_k)\big)_{k\geq 0}$$

de elementos de S tales que para todo $k \in \mathbb{N}_0$ es $a_{k+1} = f(m_k)$ y $b_{k+1} = f(n_k)$. De la definición de S se sigue ahora que la sucesión $(a_k)_{k\geq 0}$ es estrictamente creciente, que la sucesión $(b_k)_{k\geq 0}$ es estrictamente decreciente, y que $a_k < b_k$ para todo $k \in \mathbb{N}_0$. De esto se deduce, como sabemos, que existe el límite

$$\alpha \coloneqq \lim_{k\to\infty} a_k$$

y que

$$a_k < \alpha < b_k \text{ para todo } k \in \mathbb{N}_0. \tag{3.8}$$

Notemos que, además,

$$\text{para todo } k \in \mathbb{N}_0 \text{ se tiene que } m_k < m_{k+1} \text{ y } n_k < n_{k+1}. \tag{3.9}$$

En efecto, sea $k \in \mathbb{N}_0$. Si fuera $m_k \geq m_{k+1}$, entonces, como

$$a_k < f(m_k) = a_{k+1} < f(m_{k+1}) < b_{k+1} = f(n_k) < b_k, \tag{3.10}$$

tendríamos que $f(m_{k+1}) \in P(a_k, b_n, m_k, n_k) \subseteq \{f(m_k), f(n_k)\}$, pero esto es absurdo, ya que valen las desigualdades de (9.11). De manera similar, si fuera $n_k \geq n_{k+1}$, entonces, como

$$a_k < f(m_k) = a_{k+1} < f(n_{k+1}) < b_{k+1} = f(n_k) < b_k, \tag{3.11}$$

sería $f(n_{k+1}) \in P(a_k, b_n, m_k, n_k) \subseteq \{f(m_k), f(n_k)\}$, contra las desigualdades de (3.11). Esto prueba la afirmación (3.9), y dejamos al lector la tarea fácil de deducir de ella que además

$$\text{para todo } k \in \mathbb{N}_0 \text{ se tiene que } k \leq m_k \text{ y } k \leq n_k.$$

Observemos ahora que

$$\text{para todo } k \in \mathbb{N}_0 \text{ es } f(k) \notin (a_{k+1}, b_{k+1}). \tag{3.12}$$

Si, por el contrario, hubiera un elemento k de $\mathbb{N}_0$ tal que $a_{k+1} < f(k) < b_{k+1}$, tendríamos — ya que la función f es inyectiva, $k \leq m_k < m_{k+1}$ y $k \leq n_k < n_{k+1}$ — que

$$f(k) \notin \{f(m_{k+1}), f(n_{k+1})\} \supseteq P(a_{k+1}, b_{k+1}, m_{k+1}, n_{k+1})$$

y, entonces, que $\max\{m_{k+1}, n_{k+1}\} < k < k+1 \leq m_{k+1}$, lo que es absurdo.

Ahora bien, como la función f es sobreyectiva, hay un elemento j de $\mathbb{N}_0$ tal que $f(j) = \alpha$, y entonces (3.12) nos dice que $\alpha = f(j) \notin (a_{j+1}, b_{j+1})$: esto contradice a (3.8). Esta contradicción provino de haber supuesto que hay una función biyectiva $\mathbb{N}_0 \to \mathbb{R}$, así que no puede haberla. Esto prueba la proposición. □

Ejercicio 3.4.4. En su artículo [Can1874], Cantor prueba la siguiente afirmación más fuerte que la que probamos:

> *si $f : \mathbb{N}_0 \to \mathbb{R}$ es una función inyectiva y (a, b) es un intervalo abierto no vacío de $\mathbb{R}$, entonces hay un elemento c de (a, b) que no pertenece a la imagen de f.*

Demuestre esto.

Que $\mathbb{R}$ sea un conjunto no numerable nos permite probar que muchos otros conjuntos tampoco son numerables. Por ejemplo, como $\mathbb{C}$ contiene a $\mathbb{R}$, es claro que $\mathbb{C}$ no es un conjunto numerable.

Es conveniente hacer la siguiente definición.

Definición 3.4.5. Un conjunto es ***contable*** si es finito o numerable.

La propiedad clave de esta clase de conjuntos es la siguiente.

Proposición 3.4.6. *Sean X e Y dos conjuntos.*

(*i*) *Si $f : X \to Y$ es una función inyectiva y el conjunto Y es contable, entonces también lo es el conjunto X.*

(*ii*) *Si $f : X \to Y$ es una función sobreyectiva y el conjunto X es contable, entonces también lo es el conjunto Y.*

Notemos que si en este enunciado cambiamos la palabra *contable* por *numerable* ninguna de las dos afirmaciones es cierta.

Demostración. (*i*) Sea $f : X \to Y$ una función inyectiva con codominio Y contable, y supongamos que el dominio X es infinito. Esto implica que Y es infinito, como sabemos, y, como es contable, que es numerable. Vemos así que $\mathbb{N}_0 \preceq X$ porque X es infinito y que $X \preceq Y \sim \mathbb{N}_0$. Podemos concluir entonces gracias al Teorema de Cantor–Bernstein–Schröder 3.1.9 que $X \sim \mathbb{N}_0$.

(*ii*) Si $f : X \to Y$ es una función sobreyectiva con dominio contable, entonces sabemos de la

Proposición 2.1.4 que hay una función inyectiva $Y \to X$ y la primera parte de la proposición nos dice que Y es contable. □

Una consecuencia inmediata de esto es el siguiente corolario.

Corolario 3.4.7. *Un subconjunto de un conjunto contable es contable.*

Otra vez, si reemplazamos la palabra *contable* por la palabra *numerable* obtenemos una afirmación falsa. Son estos resultados los que explican la conveniencia de la noción de contabilidad.

Demostración. En efecto, si X es un conjunto contable e Y es un subconjunto de X, entonces la función de inclusión $Y \hookrightarrow X$ es inyectiva, así que la primera parte de la proposición nos dice que el conjunto Y es contable. □

Las dos partes de la Proposición 3.4.6 tienen el siguiente recíproco parcial.

Proposición 3.4.8. *Sea X un conjunto no vacío. Las siguientes afirmaciones son equivalentes.*
(*a*) *El conjunto X es contable.*
(*b*) *Hay una función sobreyectiva $\mathbb{N}_0 \to X$.*
(*c*) *Hay una función inyectiva $X \to \mathbb{N}_0$.*

Demostración. Supongamos que el conjunto X es contable. Si es infinito, entonces es numerable y hay una función biyectiva $\mathbb{N}_0 \to X$ que es, en particular, sobreyectiva. Si, en cambio, es finito, entonces hay un entero positivo m y una función biyectiva $f : [\![m]\!] \to X$, y la función

$$g : n \in \mathbb{N}_0 \mapsto f(\min\{n+1, m\}) \in X$$

es sobreyectiva. Esto muestra que vale la implicacion $(a) \Rightarrow (b)$. Por otro lado, que la implicacion $(b) \Rightarrow (c)$ vale es consecuencia de la Proposición 2.1.4, y que la implicacion $(c) \Rightarrow (a)$ vale sigue de la primera parte de la Proposición 3.4.6. □

Para probar el siguiente grupo de resultados empezamos con la sigueinte observación.

Proposición 3.4.9. *El conjunto $\mathbb{N}_0 \times \mathbb{N}_0$ es numerable.*

Demostración. La función $f : n \in \mathbb{N}_0 \mapsto (n, 0) \in \mathbb{N}_0 \times \mathbb{N}_0$ es inyectiva, así que el conjunto $\mathbb{N}_0 \times \mathbb{N}_0$ es infinito. Por otro lado, el *Teorema fundamental de la aritmética* implica inmediatamente que es inyectiva la función

$$g : (m, n) \in \mathbb{N}_0 \times \mathbb{N}_0 \mapsto 2^m 3^n \in \mathbb{N}_0.$$

Tenemos entonces que $\mathbb{N}_0 \preceq \mathbb{N}_0 \times \mathbb{N}_0$ y que $\mathbb{N}_0 \times \mathbb{N}_0 \preceq \mathbb{N}_0$, de manera que $\mathbb{N}_0 \times \mathbb{N}_0 \sim \mathbb{N}_0$. □

Esta demostración depende del *Teorema de Cantor–Bernstein–Schröder* 3.1.9 y, entonces, del *Axioma de elección*. Una consecuencia de esto es que no nos dice de qué manera obtener una biyección $\mathbb{N}_0 \to \mathbb{N}_0 \times \mathbb{N}_0$. Muchas veces es útil conocer alguna. El resultado del siguiente ejercicio exhibe una biyección de manera explícita.

Ejercicio 3.4.10. Muestre que la función $f_2 : \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{N}_0$ que en cada elemento (m, n) de $\mathbb{N}_0 \times \mathbb{N}_0$ toma el valor

$$f_2(m, n) = \frac{(m + n - 2)(m + n - 1)}{k} + k + 1$$

es biyectiva, y encuentre de manera explícita su función inversa.

La Proposición 3.4.9 tiene la siguiente consecuencia importante.

Proposición 3.4.11. *Si X e Y son conjuntos contables, entonces el producto cartesiano $X \times Y$ también lo es.*

Demostración. Supongamos que X e Y son dos conjuntos contables, de manera que hay funciones sobreyectivas $f : \mathbb{N}_0 \to X$ y $g : \mathbb{N}_0 \to Y$. Como la función

$$(m, n) \in \mathbb{N}_0 \times \mathbb{N}_0 \mapsto (f(m), g(n)) \in X \times Y$$

es sobreyectiva y el conjunto $\mathbb{N}_0 \times \mathbb{N}_0$ es contable, podemos concluir que $X \times Y$ es contable, como queremos. □

Como siempre, podemos extender esto a productos cartesianos finitos.

Ejercicio 3.4.12. Sea k un entero. Muestre que si $X_1, X_2, \ldots, X_k$ son conjuntos contables, entonces el producto cartesiano $X_1 \times X_2 \times \cdots \times X_k$ es contable.

Usando estos resultados sobre productos podemos obtener resultados similares sobre uniones.

Proposición 3.4.13. *Sean X e Y dos conjuntos y sea $F : X \to \mathscr{P}(Y)$ una función. Si*

- *el conjunto X es contable y*
- *para cada $x \in X$ el conjunto $F(x)$ es contable,*

entonces la unión $\bigcup_{x \in X} F(x)$ es contable.

Así, la unión de una familia contable de conjuntos contables es contable.

Demostración. Supongamos que se cumplen las dos condiciones del enunciado, de manera que hay una función sobreyectiva $f : \mathbb{N}_0 \to X$ y para cada $x \in X$ hay una función sobreyectiva $g_x : \mathbb{N}_0 \to F(x)$. Por otro lado, sabemos que el conjunto $\mathbb{N}_0 \times \mathbb{N}_0$ es numerable, así que hay una función biyectiva $h : \mathbb{N}_0 \to \mathbb{N}_0 \times \mathbb{N}_0$ y, en consecuencia, dos funciones $h_1, h_2 : \mathbb{N}_0 \to \mathbb{N}_0$ tales que $h(n) = (h_1(n), h_2(n))$ para todo $n \in \mathbb{N}_0$.

Como $F(x) \subseteq \bigcup_{x \in X} F(x)$ para todo $x \in X$, hay una función

$$k : n \in \mathbb{N}_0 \mapsto g_{f(h_1(n))}(h_2(n)) \in \bigcup_{x \in X} F(x).$$

Para probar la proposición es suficiente que mostremos que es sobreyectiva.

Sea y un elemento del codominio de h, de manera que hay un elemento x de X tal que $y \in F(x)$. Como la función f es sobreyectiva, hay un entero no negativo n_1 tal que $f(n_1) = x$, y como la función g_x es sobreyectiva, hay un entero no negativo n_2 tal que $g_x(n_2) = y$. Finalmente, como la función h es biyectiva, hay un entero no negativo n tal que $h(n) = (n_1, n_2)$ y, por lo tanto, es $h_1(n) = n_1$, $h_2(n) = n_2$, y

$$k(n) = g_{f(h_1(n))}(h_2(n)) = g_{f(n_1)}(n_2) = g_x(n_2) = y.$$

Esto prueba lo que queremos. □

Esto tiene la siguiente consecuencia.

Corolario 3.4.14. *Sean X e Y dos conjuntos y sea $f : X \to Y$ una función. Si*

- *Y es contable y*
- *para cada $y \in Y$ la preimagen $f^{-1}(y)$ es contable,*

entonces el conjunto X es contable.

Demostración. Como $X = \bigcup_{y \in Y} f^{-1}(y)$, esto es consecuencia inmediata de la proposición. □

Una propiedad útil de los conjuntos contables es que no son suficientemente grandes como para «cambiar» el cardinal de un conjunto infinito, en el siguiente sentido preciso.

Proposición 3.4.15. *Sea X un conjunto infinito. Si Y es un conjunto contable, entonces $X \cup Y \sim X$.*

Demostración. Sea Y un conjunto contable, de manera que hay una función sobreyectiva $f : \mathbb{N}_0 \to Y$. Por otro lado, como el conjunto X es infinito, hay una función inyectiva $g : \mathbb{N}_0 \to X$. Sea I la imagen de g y sea $h : I \to \mathbb{N}_0$ la función inversa de la correstricción $g|^I : \mathbb{N}_0 \to I$. Podemos

entonces considerar la función $k : X \to X \cup Y$ que en cada elemento x de X toma el valor

$$k(x) = \begin{cases} x & \text{si } x \notin I; \\ g(h(x)/2) & \text{si } x \in I \text{ y } h(x) \text{ es par;} \\ f((h(x)-1)/2) & \text{si } x \in I \text{ y } h(x) \text{ es impar.} \end{cases}$$

Esta función es sobreyectiva. En efecto, sea z un elemento de $X \cup Y$, tenemos tres casos a considerar.

- Si $z \in X \setminus I$, entonces $z \in X$ y es $k(z) = z$.
- Si $z \in I$, entonces hay un elemento n de $\mathbb{N}_0$ tal que $z = g(n)$ y, por lo tanto, como $g(2n) \in I$ y $h(g(2n)) = 2n$ es par, tenemos que $k(g(2n)) = g(h(g(2n))/2) = g(n) = z$.
- Finalmente, si $z \in Y$, entonces podemos elegir un elemento n de $\mathbb{N}_0$ tal que $f(n) = z$ y es $k(g(2n+1)) = f((h(g(2n+1)) - 1)/2) = f(n) = z$ ya que, por supuesto, $g(2n+1) \in I$ y $h(g(2n+1)) = 2n+1$ es impar,

Como el conjunto X no es vacío, podemos concluir entonces que $X \cup Y \preceq X$. Como es evidente que $X \preceq X \cup Y$, podemos concluir de esto que $X \sim X \cup Y$, como queremos. □

Con las herramientas que ya tenemos a nuestra disposición es fácil probar que muchos conjuntos familiares son contables y, como muchos de ellos son infinitos, numerables. Veamos algunos ejemplos.

Ejemplo 3.4.16.

(*i*) Como la función $(m, n) \in \mathbb{N}_0 \times \mathbb{N}_0 \mapsto m - n \in \mathbb{Z}$ es sobreyectiva, el conjunto $\mathbb{Z}$ es contable.

(*ii*) Como la función $(m, n) \in \mathbb{Z} \times \mathbb{N}_0 \mapsto \dfrac{m}{n+1} \in \mathbb{Q}$ es sobreyectiva, el conjunto $\mathbb{Q}$ es contable.

(*iii*) Si X es un conjunto contable, entonces para cada $k \in \mathbb{N}_0$ el conjunto X^k de todas las secuencias $(x_1, x_2, \ldots, x_k)$ de longitud k de elementos de X es contable, y entonces el conjunto X^* de *todas* las secuencias finitas, de todas las longitudes posibles, es contable.

(*iv*) Sea X un conjunto contable y para cada entero no negativo k escribamos $\mathscr{P}_k(X)$ al conjunto de todos los subconjuntos de X finitos de cardinal k. El conjunto

$$X^{(k)} \coloneqq \{(x_1, x_2, \ldots, x_k) \in X^n : \forall\, i, j \in [\![k]\!],\, i \neq j \implies x_i \neq x_j\}$$

es contable, porque es un subconjunto de X^k, y la función

$$(x_1, x_2, \ldots, x_n) \in X^{(k)} \mapsto \{x_1, x_2, \ldots, x_n\} \in \mathscr{P}_k(X)$$

es sobreyectiva, así que el conjunto $\mathscr{P}_k(X)$ es contable. Se sigue de esto que para cada $l \in \mathbb{N}_0$ el conjunto $\mathscr{P}_{\leq l}(X) \coloneqq \bigcup_{k=0}^{l} \mathscr{P}_k(X)$ es contable, y que también lo es el conjunto $\mathscr{P}_f(X)$ de todas los subconjuntos finitos de X.

Más aún, con lo que tenemos hecho podemos probar el resultado principal del primer artículo

de Cantor.

Ejemplo 3.4.17. Decimos que un número complejo es ***algebraico*** si hay un polinomio no nulo con coeficientes racionales que lo tiene por raíz, y que es ***trascendente*** si no es algebraico. Esta definición fue dada originalmente por Leonard Euler. Johann Heinrich Lambert conjeturó que los números π y e son trascendentes en su célebre trabajo [Lam1768], en el que dio la primera demostración de que π es irracional. El primero que pudo probar que hay números que son transcendantes fue Joseph Liouville, en 1844. Él mismo pudo probar en [Lio1851] que el número

$$\sum_{n=1}^{\infty} 10^{-n!}$$

y muchos otros del mismo tipo son transcendantes. Charles Hermite fue el primero que pudo probar que un número no construido específicamente con ese objetivo es transcendante: probó en [Her1873] que el número e lo es, y con ideas similares Ferdinand von Lindemann pudo probar en [vL1882] que π es transcendante. El libro [Lan1966] de Serge Lang es una buena introducción a la teoría moderna de los números trascendentes.

El árticulo [Can1874] de Cantor, el primero que escribió sobre la teoría de conjuntos, tiene como objetivo dar una demostración de que existen números trascendentes y que, de hecho, «la mayor parte» de los números completos lo son de una naturaleza completamente distinta a lo que se había hecho hasta entonces. Veamos una versión de su argumento.

Para cada entero no negativo d escribamos $\mathbb{Q}[X]_{\leq d}$ al conjunto de todos los polinomios con coeficientes racionales que o son nulos o que no lo son y tienen grado mejor o igual a d. Es claro que la función

$$(a_0, a_1, \ldots, a_d) \in \mathbb{Q}_0^{d+1} \mapsto a_0 + a_1 X + \cdots + a_d X^d \in \mathbb{Q}[X]_{\leq d}$$

es sobreyectiva, así que ese conjunto $\mathbb{Q}[X]_{\leq d}$ es contable. Se sigue de la Proposición 3.4.13, entonces, que también es contable el conjunto

$$\mathbb{Q}[X] = \bigcup_{d \in \mathbb{N}_0} \mathbb{Q}[X]_{\leq d}.$$

Sea ahora $\mathscr{A}$ el conjunto de todos los números algebraicos. No es difícil probar que si α es un elemento de $\mathscr{A}$, entonces hay exactamente un polinomio μ_α en $\mathbb{Q}[X]$, al que llamamos el ***polinomio minimal*** de α, que satisface las siguientes dos condiciones:

- El polinomio μ_α es mónico.
- Cada polinomio de $\mathbb{Q}[X]$ que tiene a α como raíz es divisible por μ_α.

Hay entonces una función

$$\mu : \alpha \in \mathscr{A} \mapsto \mu_\alpha \in \mathbb{Q}[X].$$

Notemos que para cada polinomio p de $\mathbb{Q}[X]$ el conjunto $\mu^{-1}(q)$ es finito — su cardinal es menor o igual que el grado de q — y, por lo tanto, contable. El Corolario 3.4.14 nos dice entonces que el conjunto $\mathscr{A}$ es contable.

Ahora bien, el conjunto $\mathbb{C}$ es infinito y no es numerable, así que no es contable: esto nos permite concluir inmediatamente que el conjunto $\mathscr{A}$ no coincide con $\mathbb{C}$ y, por lo tanto, que hay elementos en $\mathbb{C}$ que no son algebraicos, esto es, que son trascendentes.

Es interesante observar que este argumento prueba que hay números trascendentes pero ¡no nos permite exhibir ninguno! Para ello es necesario seguir los pasos de Liouville, Hermite, von Lindemann, y los continuadores de la teoría que ellos iniciaron.

Es claro que hay subconjuntos infinitos de $\mathbb{R}$ que son numerables y que hay otros que tienen la potencia del continuo. La *Hipótesis del continuo* es la afirmación de que esos son los dos únicos cardinales que puede tener un subconjunto infinito de $\mathbb{R}$.

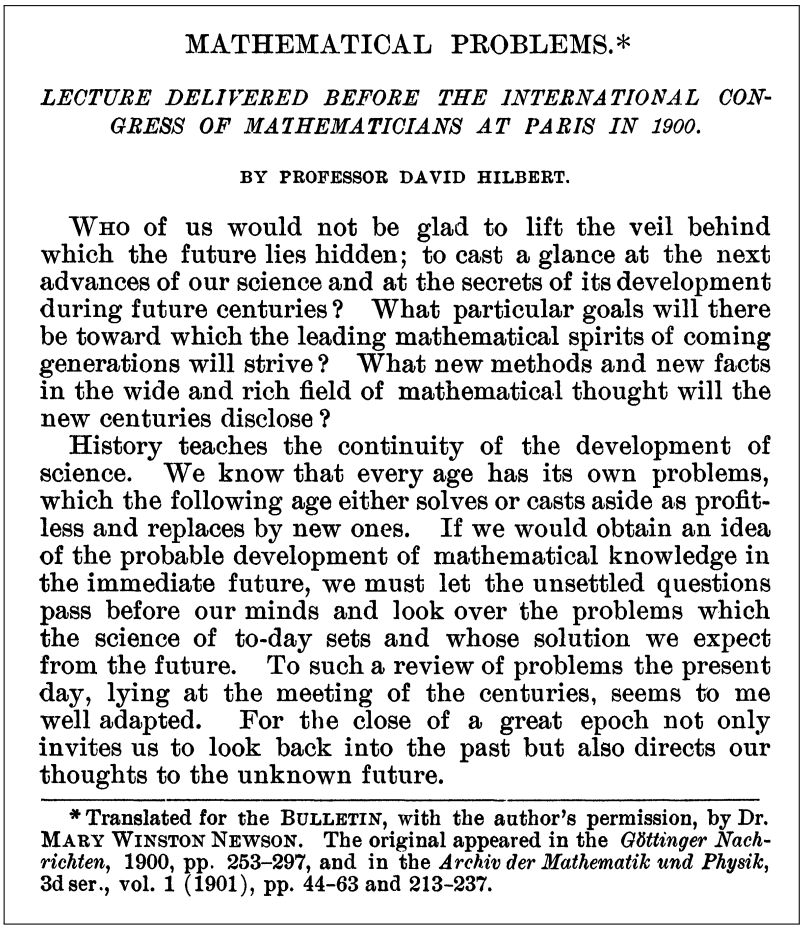
Hipótesis 3.4.18. *Un subconjunto infinito no numerable de $\mathbb{R}$ tiene la potencia del continuo.*

Cantor propuso esta afirmación en su trabajo [Can1877], creía que era cierta, y pasó buena parte de su vida tratando de probarla. En 1900, durante el *Congreso Internacional de Matemáticos*, David Hilbert presentó una célebre lista de 23 problemas cuyas soluciones, en su opinión, podrían orientar el desarrollo de la matemática en el siglo que empezaba, y el primero de ellos es el de probar la *Hipótesis del continuo*. La historia de este problema es larga y complicada, y el final de esa historia es particularmente interesante: Kurt Gödel probó en [Göd1940] que en la teoría usual de conjuntos — la teoría de Zermelo–Fraenkel aumentada del *Axioma de elección* — no es posible probar la negación de la *Hipótesis del continuo*, y Paul Cohen probó en [Coh1963] que tampoco es posible probarla a ella misma en esa teoría — Cohen recibió la *Medalla Fields* en 1966 por ese trabajo. Combinando los dos resultados se obtiene el célebre teorema de independencia de la *Hipótesis del continuo* de la teoría de conjuntos usual.

Se conoce una enorme cantidad de enunciados que son equivalentes a la *Hipótesis del continuo*. El libro clásico [Sie1934] de Wacław Sierpiński es, entre otras cosas, una lista de ejemplos conocidos hacia 1934, pero hay muchos otros. Por ejemplo, de acuerdo a Paul Erdős y Péter Komjáth [EK1990, Kom2001], la *Hipótesis del continuo* es equivalente a cada uno de los siguientes enunciados:

> *Es posible asignar a cada punto del plano un entero de manera que no haya ningún triángulo rectángulo a cuyos vértices corresponda el mismo número.*
>
> *Hay una partición del plano real en tres partes tal que la intersección de cada recta del plano con una esas partes es finita.*

Es de notar que la *Hipótesis del continuo* tiene consecuencias no triviales en la «vida real» matemática — hay afirmaciones muy naturales en topología, teoría de módulos, teoría de la medida, y otras disciplinas que resultan depender de la *Hipótesis del continuo*.



MATHEMATICAL PROBLEMS.*

LECTURE DELIVERED BEFORE THE INTERNATIONAL CONGRESS OF MATHEMATICIANS AT PARIS IN 1900.

BY PROFESSOR DAVID HILBERT.

WHO of us would not be glad to lift the veil behind which the future lies hidden; to cast a glance at the next advances of our science and at the secrets of its development during future centuries? What particular goals will there be toward which the leading mathematical spirits of coming generations will strive? What new methods and new facts in the wide and rich field of mathematical thought will the new centuries disclose?

History teaches the continuity of the development of science. We know that every age has its own problems, which the following age either solves or casts aside as profitless and replaces by new ones. If we would obtain an idea of the probable development of mathematical knowledge in the immediate future, we must let the unsettled questions pass before our minds and look over the problems which the science of to-day sets and whose solution we expect from the future. To such a review of problems the present day, lying at the meeting of the centuries, seems to me well adapted. For the close of a great epoch not only invites us to look back into the past but also directs our thoughts to the unknown future.

* Translated for the BULLETIN, with the author's permission, by Dr. MARY WINSTON NEWSON. The original appeared in the *Göttinger Nachrichten*, 1900, pp. 253–297, and in the *Archiv der Mathematik und Physik*, 3d ser., vol. 1 (1901), pp. 44–63 and 213–237.

Figura 3.2. El principio de la traducción [Hil1902] del artículo [Hil1900] en el que Hilbert presentó su lista de problemas.

3.5. Operaciones con cardinales

En esta sección queremos presentar ciertas «operaciones aritméticas» con cardinales. Empecemos por hacer la siguiente observación.

Ejercicio 3.5.1. Sean X, Y, X' y Y' conjuntos. Muestre que si $X \sim X'$ e $Y \sim Y'$, entonces

$$X \times \{1\} \cup Y \times \{2\} \sim X' \times \{1\} \cup Y' \times \{2\},$$
$$X \times Y \sim X' \times Y',$$
$$X^Y \sim X'^{Y'}.$$

Gracias a esto podemos definir las operaciones que nos interesan.

Definición 3.5.2. Sean $\mathfrak{a}$ y $\mathfrak{b}$ dos cardinales y sean X e Y conjuntos tales que $|X| = \mathfrak{a}$ y $|Y| = \mathfrak{b}$. La ***suma*** $\mathfrak{a} + \mathfrak{b}$, el ***producto*** $\mathfrak{a} \cdot \mathfrak{b}$ y la ***exponenciación*** $\mathfrak{a}^{\mathfrak{b}}$ poniendo

$$\mathfrak{a} + \mathfrak{b} \coloneqq |X \times \{1\} \cup Y \times \{2\}|,$$
$$\mathfrak{a} \cdot \mathfrak{b} \coloneqq |X \times Y|,$$

y

$$\mathfrak{a}^{\mathfrak{b}} \coloneqq |X^Y|.$$

El resultado del ejercicio nos dice que esta definiciones tienen sentido, ya que los cardinales que aparecen a la derecha del signo $\coloneqq$ en cada una de estas definiciones dependen solamente de los cardinales $\mathfrak{a}$ y $\mathfrak{b}$ con los que empezamos y no de la elección de los conjuntos X e Y.

Notemos que definimos la suma en la forma en que lo hicimos porque los conjuntos $X \times \{1\}$ e $Y \times \{2\}$ son disjuntos independientemente de si X e Y lo son o no. De todas formas, es útil observar lo siguiente.

Ejercicio 3.5.3. Pruebe que si X e Y son dos conjuntos disjuntos, entonces $|X| + |Y| = |X \cup Y|$.

Estas operaciones tienen muchas de las propiedades de las operaciones aritméticas usuales.

Ejercicio 3.5.4. Sean $\mathfrak{a}$, $\mathfrak{b}$ y $\mathfrak{c}$ tres cardinales.

(a) Muestre que la suma y el producto de cardinales son operaciones asociativas y conmutativas, y tienen a $0 \coloneqq |[\![0]\!]|$ y a $1 \coloneqq |[\![1]\!]|$, respectivamente, como elementos neutros, y que

$$(\mathfrak{a} + \mathfrak{b}) \cdot \mathfrak{c} = \mathfrak{a} \cdot \mathfrak{c} + \mathfrak{b} \cdot \mathfrak{c}, \qquad \mathfrak{a} \cdot 0 = 0.$$

(b) Pruebe que

$$\mathfrak{a}^1 = \mathfrak{a}, \qquad 1^{\mathfrak{a}} = 1, \qquad \mathfrak{a}^0 = 1, \qquad 0^{\mathfrak{a}} = 0 \text{ si } \mathfrak{a} \neq 0,$$

$$\mathfrak{a}^{\mathfrak{b}+\mathfrak{c}} = \mathfrak{a}^{\mathfrak{b}} \cdot \mathfrak{a}^{\mathfrak{c}}, \qquad \mathfrak{a}^{\mathfrak{b}\cdot\mathfrak{c}} = (\mathfrak{a}^{\mathfrak{b}})^{\mathfrak{c}}, \qquad (\mathfrak{a}\cdot\mathfrak{b})^{\mathfrak{c}} = \mathfrak{a}^{\mathfrak{c}} \cdot \mathfrak{b}^{\mathfrak{c}},$$

(c) Si $\mathfrak{a} \leq \mathfrak{b}$, entonces

$$\mathfrak{a} + \mathfrak{c} \leq \mathfrak{b} + \mathfrak{c}, \qquad \mathfrak{a} \cdot \mathfrak{c} \leq \mathfrak{b} \cdot \mathfrak{c}, \qquad \mathfrak{a}^{\mathfrak{c}} \leq \mathfrak{b}^{\mathfrak{c}},$$

y si además $1 \leq \mathfrak{c}$, entonces también

$$\mathfrak{c}^{\mathfrak{a}} \leq \mathfrak{c}^{\mathfrak{b}},$$

Observación 3.5.5. Es importante notar que que en la parte (*c*) de este ejercicio estamos hablando de la relación $\leq$ y no de su versión estricta $<$. De hecho, la afirmación

si $\mathfrak{a}$ *y* $\mathfrak{b}$ *son cardinales tales que* $\mathfrak{a} < \mathfrak{b}$, *entonces* $2^{\mathfrak{a}} < 2^{\mathfrak{b}}$

es *independiente* de la teoría de conjuntos de Zermelo–Fraenkel — este resultado está íntimamente relacionado con la independencia de la *Hipótesis del continuo* y con una versión más general de esta conocida como la *Hipótesis generalizada del continuo*. Esto significa que no es posible probar en la teoría de conjuntos usual que si un conjunto X tiene «menos» elementos que un conjunto Y entonces también tiene menos subconjuntos que este.

Las operaciones aritméticas entre cardinales «coinciden» con las operaciones aritméticas usuales cuando las restringimos a los cardinales finitos.

Proposición 3.5.6. *Sean m y n dos elementos de $\mathbb{N}_0$. Es*

$$|[\![m]\!]| + |[\![n]\!]| = |[\![m+n]\!]|,$$
$$|[\![m]\!]| \cdot |[\![n]\!]| = |[\![m \cdot n]\!]|,$$
$$|[\![m]\!]|^{|[\![n]\!]|} = |[\![m^n]\!]|.$$

Una de las consecuencias de este resultado es que la convención que hicimos arriba de escribir para cada $m \in \mathbb{N}_0$ simplemente m en lugar de $|[\![m]\!]|$ es compatible con las operaciones aritméticas entre cardinales.

Demostración. Como los conjuntos $[\![m]\!] \times \{1\}$ y $[\![n]\!] \times \{2\}$ son disjuntos, la Proposición 3.2.17 nos dice que

$$|[\![m]\!]| + |[\![n]\!]| = |[\![m]\!] \times \{1\} \cup [\![n]\!] \times \{2\}| = |[\![m]\!] \times \{1\}| + |[\![n]\!] \times \{2\}|$$

y, como las funciones $i \in [\![m]\!] \mapsto (i,1) \in [\![m]\!] \times \{1\}$ e $i \in [\![n]\!] \mapsto (i,2) \in [\![n]\!] \times \{1\}$ son biyectivas, esto es igual a $|[\![m]\!]| + |[\![n]\!]|$. Vemos así que la primera igualdad de la proposición es cierta. Por su

parte, que son ciertas la segunda y la tercera es lo que nos dicen las Proposiciones 3.2.20 y 3.2.22, respectivamente. □

Por otro lado, cuando operamos con cardinales infinitos las cosas son un poco diferentes. Más abajo explicaremos esto en detalle, pero empezamos dando dos ejemplos.

Ejemplo 3.5.7. Mostremos que

$$\aleph_0 + \aleph_0 = \aleph_0, \qquad \mathfrak{c} + \mathfrak{c} = c.$$

El conjunto $\mathbb{N}_0 \times [\![2]\!]$ es claramente infinito y contable, así que es numerable, y esto nos dice precisamente que $\aleph_0 + \aleph_0 = |\mathbb{N}_0| + |\mathbb{N}_0| = \aleph_0$. Por otro lado, las funciones

$$x \in \mathbb{R} \mapsto e^x \in \mathbb{R}, \qquad x \in \mathbb{R} \mapsto -e^b \in \mathbb{R}$$

son inyectivas, así que sus imágenes $(0, +\infty)$ y $(-\infty, 0)$ tienen las dos la potencia del continuo y, como son disjuntas, tenemos que

$$|\mathbb{R} \setminus \{0\}| = |(-\infty, 0) \cup (0, +\infty)| = |(-\infty, 0)| + |(0, +\infty)| = \mathfrak{c} + \mathfrak{c}.$$

Ahora bien, como el conjunto $\{0\}$ es finito, esto nos dice que $\mathfrak{c} = |\mathbb{R}| = |\mathbb{R} \setminus \{0\}| = \mathfrak{c} + \mathfrak{c}$.

Queremos generalizar estos ejemplos a cardinales infinitos cualesquiera. La siguiente proposición es lo que nos permite hacerlo.

Proposición 3.5.8. *Si X es un conjunto infinito, entonces $X \times \mathbb{N}_0 \sim X$.*

Demostración. Consideremos el conjunto $\mathscr{S}$ de todas las funciones inyectivas $A \times \mathbb{N}_0 \to A$ con A un subconjunto infinito de X. Los elementos de $\mathscr{S}$ son subconjuntos de $X \times \mathbb{N}_0 \times X$, así que tiene sentido considerar en $\mathscr{S}$ la relación de orden $\subseteq$ dada por la inclusión. Queremos probar que el conjunto $\mathscr{S}$ así ordenado satisface las hipótesis del *Lema de Zorn*.

- Como el conjunto X es infinito, la Proposición 3.3.4 nos dice que hay un subconjunto numerable A de X. Es entonces $A \sim \mathbb{N}_0 \sim \mathbb{N}_0 \times \mathbb{N}_0 \sim A \times \mathbb{N}_0$ y hay, por lo tanto, una función biyectiva $A \times \mathbb{N}_0 \to A$. Esta función es un elemento de $\mathscr{S}$ y, en particular, este conjunto no es vacío.
- Sea ahora $\mathscr{C}$ una cadena no vacía en $\mathscr{S}$, y pongamos

  $$F \coloneqq \bigcup_{f \in \mathscr{C}} f,$$

  que es un subconjunto de $X \times \mathbb{N}_0 \times X$, y

  $$C \coloneqq \{x \in X : \text{existe } y \in X \text{ tal que } (x, 0, y) \in F\},$$

que es uno de X. Queremos probar que F es una función inyectiva $C \times \mathbb{N}_0 \to C$ y que el conjunto C es infinito, de manera que F es un elemento de $\mathscr{S}$.

- Sean (c, n, y) un elemento de F, de manera que hay un elemento f de $\mathscr{C}$ tal que $(c, n, y) \in f$. Como f está en $\mathscr{S}$, es una función inyectiva $A \times \mathbb{N}_0 \to A$ con A un subconjunto infinito de X y, como $(c, n, y) \in f$, tenemos que $c \in A$ y que $y \in A$. Como entonces las ternas $(c, 0, f(c, 0))$ y $(y, 0, f(y, 0))$ están las dos en f y, por lo tanto, en F, es $c \in C$ e $y \in C$. Esto prueba que $F \subseteq C \times \mathbb{N}_0 \times C$.
- Sea (c, n) un elemento de $C \times \mathbb{N}_0$. Como c está en C, sabemos que hay un elemento y de X tal que $(c, 0, y) \in F$ y, por lo tanto, que hay un elemento f de $\mathscr{C}$ tal que $(c, 0, y) \in f$. Este elemento f es una función inyectiva $A \times \mathbb{N}_0 \to A$ con A algún subconjunto infinito de X. Como $(c, 0, y)$ está en f, c está en A y (c, n) está en $A \times \mathbb{N}_0$, así que la terna $(c, n, f(c, n))$ es un elemento de f y, en consecuencia, de F.
- Sea ahora (c, n) un elemento de $C \times \mathbb{N}_0$, sean y e y' dos elementos de X, y supongamos que (c, n, y) y (c, n, y') pertenecen a F, de manera que hay dos funciones inyectivas $f : A \times \mathbb{N}_0 \to A$ y $g : B \times \mathbb{N}_0 \to B$ con A y B subconjuntos infinitos de X y $(c, n, y) \in f$ y $(c, n, y') \in g$. Como $\mathscr{C}$ es una cadena en $\mathscr{S}$, tenemos que $f \subseteq g$ o $g \subseteq f$. En el primer caso las dos ternas (c, n, y) y (c, n, y') pertenecen a g y, como g es una función $B \times \mathbb{N}_0 \to B$, tenemos que $y = y'$; en el segundo caso esas dos ternas pertenecen a f y, como f es una función $A \times \mathbb{N}_0 \to A$, tenemos otra vez que $y = y'$.
- Sean finalmente (c, n) y (c', n') dos elementos de $C \times \mathbb{N}_0$ e y uno de X, y supongamos que las ternas (c, n, y) y (c', n', y) pertenecen las dos a F. De acuerdo a la definición de F, hay entonces dos elementos f y g de $\mathscr{C}$ tales que $(c, n, y) \in f$ y $(c', n', y) \in g$. Como $\mathscr{C}$ es una cadena, tenemos que $f \subseteq g$ o $g \subseteq f$. En el primer caso, las dos ternas (c, n, y) y (c', n', y) están en g, de manera que $g(c, n) = g(c', n')$ y, como la función g es inyectiva, $(c, n) = (c', n')$; en el segundo, esas dos ternas están en f, de manera que $f(c, n) = f(c', n')$ y, como la función f es inyectiva, otra vez $(c, n) = (c', n')$.

Si $f : A \times \mathbb{N}_0 \to A$ es un elemento de $\mathscr{C}$, entonces claramente $f \subseteq F$: esto nos dice que esta función F que construimos es una cota superior para la cadena $\mathscr{C}$.

En vista de todo esto y del *Lema de Zorn*, podemos concluir que hay una función inyectiva $f : Y \times \mathbb{N}_0 \to Y$ con Y un subconjunto infinito de X que es un elemento maximal de $\mathscr{S}$.

Supongamos por un momento que $Y \not\doteq X$. Como Y es infinito e $Y \subseteq X$, la Proposición 3.4.15 nos dice entonces que $X \setminus Y$ no es contable y, por lo tanto, que es infinito y contiene un subconjunto numerable Z. Como Z es numerable, sabemos ya que hay una función biyectiva $g : Z \times \mathbb{N}_0 \to Z$ y, como $Y \cap Z = \varnothing$, hay una función $h : (Y \cup Z) \times \mathbb{N}_0 \to Y \cup Z$ que en cada elemento (u, n)

de $(Y \cup Z) \times \mathbb{N}_0$ toma el valor

$$h(u,n) = \begin{cases} f(u,n) & \text{si } u \in Y; \\ g(u,n) & \text{si } u \in Z. \end{cases}$$

Es fácil ver que esta función h es inyectiva y, como $Y \cup Z$ es ciertamente un subconjunto infinito de X, que se trata de un elemento de $\mathscr{S}$. Ahora bien, es claro que $f \subsetneq h$, ya que $Z \neq \varnothing$, y esto es absurdo, ya que f es un elemento maximal de $\mathscr{S}$.

Vemos así que necesariamente es $X \sim Y$. Como la función $x \in X \mapsto (x,0) \in X \times \mathbb{N}_0$ es inyectiva, tenemos entonces que

$$X \times \mathbb{N}_0 \sim Y \times \mathbb{N}_0 \preceq Y \sim X \preceq X \times \mathbb{N}_0$$

y el Teorema de Cantor–Bernstein–Schröder 3.1.9 nos permite concluir que $X \times \mathbb{N}_0 \sim X$, que es lo que afirma la proposición. □

Una primera consecuencia de esta proposición es la generalización del Ejemplo 3.5.7 que buscábamos.

Proposición 3.5.9. *Si $\mathfrak{a}$ es un cardinal infinito, entonces $\mathfrak{a} + \mathfrak{a} = \mathfrak{a}$.*

Demostración. Sea $\mathfrak{a}$ un cardinal infinito y sea X un conjunto de cardinal $\mathfrak{a}$. Para probar la proposición es suficiente que mostremos que $X \times [\![2]\!] \sim X$. La función $x \in X \mapsto (x,1) \in X \times [\![2]\!]$ es inyectiva, así que de acuerdo al Teorema de Cantor–Bernstein–Schröder 3.1.9 es suficiente que mostremos que hay una función inyectiva $X \times [\![2]\!] \to X$.

La Proposición 3.5.8 nos dice que $X \times \mathbb{N}_0 \sim X$, así que hay dos funciones $f : X \times \mathbb{N}_0 \to X$ y $g : X \to X \times \mathbb{N}_0$ que son mutuamente inversas. Consideremos las funciones $\xi : X \to X$ y $\nu : X \to \mathbb{N}_0$ tales que $g(x) = (\xi(x), \nu(x))$ para todo $x \in X$, y mostremos que la función

$$h : (x,n) \in X \times [\![2]\!] \mapsto f(\xi(x), 2\nu(x) + n - 1) \in X$$

es inyectiva. Para ello, supongamos que (x,n) y (x',n') son dos elementos de $X \times [\![2]\!]$ tales que $h(x,n) = h(n',n')$, de manera que $f(\xi(x), 2\nu(x) + n - 1) = f(\xi(x'), 2\nu(x') + n' - 1)$. Esto implica, claro que

$$\begin{aligned}(\xi(x), 2\nu(x) + n - 1) &= g(f(\xi(x), 2\nu(x) + n - 1)) = g(f(\xi(x'), 2\nu(x') + n' - 1)) \\ &= (\xi(x'), 2\nu(x') + n' - 1)\end{aligned}$$

y, por lo tanto, que $\xi(x) = \xi(x')$ y que $2\nu(x) + n - 1 = 2\nu(x') + n' - 1$. Como $n - 1$ y $n' - 1$ pertenecen a $\{0,1\}$, podemos deducir de esta última iguadad que $n = n'$ y que $\nu(x) = \nu(x')$, y entonces tenemos además que

$$x = f(g(x)) = f(\xi(x), \nu(x)) = f(\xi(x'), \nu(x')) = f(g(x')) = x'.$$

La función h es, en consecuencia, inyectiva, como queremos. □

Ejercicio 3.5.10. ¿Es sobreyectiva la función h que construimos en esta demostración?

La Proposición 3.5.9 tiene como corolario el siguiente resultado que muestra que la operación de suma de cardinales infinitos es bien distinta que la de cardinales finitos.

Proposición 3.5.11. *Sean* $\mathfrak{a}$ *y* $\mathfrak{b}$ *dos cardinales. Si alguno de los dos es infinito, entonces*

$$\mathfrak{a} + \mathfrak{b} = \max\{\mathfrak{a}, \mathfrak{b}\}.$$

Demostración. Sin pérdida de generalidad podemos suponer que $\mathfrak{a} \leq \mathfrak{b}$ y que $\mathfrak{b}$ es infinito. De acuerdo a la Proposición 3.5.9 y a la monotonía de la suma, tenemos entonces que

$$\mathfrak{b} = 0 + \mathfrak{b} \leq \mathfrak{a} + \mathfrak{b} \leq \mathfrak{b} + \mathfrak{b} = \mathfrak{b}$$

y, en vista del Teorema de Cantor–Bernstein–Schröder 3.1.9, que $\mathfrak{a} + \mathfrak{b} = \mathfrak{b} = \max\{\mathfrak{a}, \mathfrak{b}\}$. □

Queremos ahora obtener resultados similares a estos pero para el producto de cardinales infinitos. Como antes, empecemos con un ejemplo.

Ejemplo 3.5.12. Mostremos que

$$\aleph_0 \cdot \aleph_0 = \aleph_0, \qquad \mathfrak{c} \cdot \mathfrak{c} = \mathfrak{c}.$$

Para ver que vale la primer igualdad, es suficiente recordar que, de acuerdo a la Proposición 3.4.9, es $\mathbb{N}_0 \times \mathbb{N}_0 \sim \mathbb{N}_0$. Para probar la segunda, recordamos que $\mathfrak{c} = 2^{\aleph_0}$ y que, por lo tanto, tenemos que $\mathfrak{c} \cdot \mathfrak{c} = 2^{\aleph_0} \cdot 2^{\aleph_0} = 2^{\aleph_0 \cdot \aleph_0} = 2^{\aleph_0} = \mathfrak{c}$.

Notemos que la segunda afirmación de este ejemplo implica que hay una función biyectiva $\mathbb{R} \to \mathbb{R} \times \mathbb{R}$. Cantor pasó mucho tiempo — de 1871 a 1874 — tratando de probar que una tal función biyectiva no puede existir, hasta que finalmente logró en 1874 probar que, contra lo que esperaba, sí existe. En una célebre carta a Richard Dedekind en la que le comunicaba su descubrimiento, Cantor escribió «Lo veo, pero no lo creo» — «Ich sehe es, aber ich glaube es nicht», en el original[1].

Este descubrimiento iba ciertamente contra las ideas de los matemáticos de la época, y junto con otros de sus resultados de ese tiempo — como la demostración de la existencia de números transcentes que vimos arriba — no tuvieron una buena recepción por parte de muchos matemáticos importantes, como Leopold Kronecker, Henri Poincaré, Hermann Weyl, Luitzen Egbertus Jan Brouwer Ludwig Wittgenstein entre muchos otros. De hecho, la oposición a las ideas de Cantor

[1] En 1877 Dedekind notó que había un error en el argumento de Cantor, causado por el hecho de que algunos números reales poseen *dos* dos desarrollos decimales, pero este es fácilmente salvable.

llegó a ser feroz: Kronecker, por ejemplo, llegó a escribir que Cantor era un charlatán, un renegado, y que sus ideas corrompían a la juventud. Otros, sin embargo, reconocieron rápidamente el valor de su trabajo — David Hilbert famosamente anunció en [Hil1925], hablando de la teoría de conjuntos, que

> *nadie podrá expulsarnos del paraíso que Cantor ha creado*[2].

Como forma de moderar la afirmación contra-intuitiva de que hay funciones sobreyectivas $\mathbb{R} \to \mathbb{R} \times \mathbb{R}$ Cantor argumentaba que se trataba de funciones que no eran continuas y que, por lo tanto, ese hecho no contradice la idea intuitiva de dimensión. De todas formas, más tarde Giuseppe Peano mostró en [Pea1890] que, de hecho, sí hay funciones sobreyectivas $\mathbb{R} \to \mathbb{R} \times \mathbb{R}$ que son continuas. De todas formas, el célebre *Teorema de invariancia del dominio* de Brouwer [Bro1912] afirma que no hay funciones biyectivas $\mathbb{R} \to \mathbb{R} \times \mathbb{R}$ que sean tanto ellas como sus funciones inversas continuas, y esto tiene como consecuencia bien directa que no hay funciones biyectivas continuas $[0,1] \to [0,1] \times [0,1]$.

El resultado del Ejemplo 3.5.12 se generaliza a cardinales infinitos cualesquiera.

Proposición 3.5.13. *Si* $\mathfrak{a}$ *es un cardinal infinito, entonces* $\mathfrak{a} \cdot \mathfrak{a} = \mathfrak{a}$

Demostración. Sea $\mathfrak{a}$ un cardinal infinito y sea X un conjunto de cardinal $\mathfrak{a}$. Para probar la proposición es suficiente que mostremos que $X \times X \sim X$.

Sea $\mathscr{S}$ el conjunto de todas las funciones inyectivas $A \times A \to A$ con A un subconjunto infinito de X, y consideremos a $\mathscr{S}$ ordenado con respecto a la relación de inclusión $\subseteq$ — observemos que esto tiene sentido, ya que los elementos de $\mathscr{S}$ son subconjuntos de $X \times X \times X$. Procediendo como en la demostración que dimos de la Proposición 3.5.8 podemos ver que este conjunto ordenado satisface las hipótesis del *Lema de Zorn*, así que posee elementos maximales. Sea $f : Y \times Y \to Y$ uno de ellos, con Y un subconjunto infinito de X.

Supongamos por un momento que $Y \not\sim X$. Como Y es infinito y $Y \subseteq X$, la Proposición 3.4.15 nos dice entonces que $X \setminus Y$ no es contable y, por lo tanto, que es infinito y contiene un subconjunto numerable Z. Como $\aleph_0 \cdot \aleph_0 = \aleph_0$, hay una función biyectiva $g : Z \times Z \to Z$. Por otro lado, como Z es numerable e Y es infinito, la Proposición 3.5.8 implica que hay una función biyectiva $h : Y \times Z \to Y$.

Pongamos $Y' \coloneqq Y \cup Z$. Hay claramente una función $k : Y' \times Y'$ que en cada elemento (u, v) de $Y' \times Y'$ toma el valor

$$k(u,v) = \begin{cases} f(u,v) & \text{si } (u,v) \in Y \times Y; \\ (h(u,v),v) & \text{si } (u,v) \in Y \times Z; \\ (u,h(v,u)) & \text{si } (u,v) \in Z \times Y; \\ g(u,v) & \text{si } (u,v) \in Z \times Z. \end{cases}$$

[2] «Aus dem Paradies, das Cantor uns geshaffen, soll uns niemand verteiben können» en el original.

Esa fácil ver que esta función es inyectiva, y esto es absurdo, ya que claramente $f \subsetneq k$ y f es un elemento maximal en $\mathscr{S}$.

Esta contradicción provino de haber supuesto que $Y \not\sim X$, así que necesariamente es $Y \sim X$ y, como la función $f : Y \times Y \to Y$ es inyectiva, tenemos que

$$X \times X \sim Y \times Y \leq Y \sim X \leq X \times X$$

y, de acuerdo al Teorema de Cantor–Bernstein–Schröder 3.1.9, que $X \times X \sim X$. Notemos que aquí usamos el hecho de que $X \leq X \times X$: esto es evidente si X es vacío, y si no lo es y x_0 es un elemento de X, entonces también es claro porque la función $x \in X \mapsto (x, x_0) \in X \times X$ es inyectiva. □

Ejercicio 3.5.14. Pruebe en detalle (*i*) que el conjunto ordenado $\mathscr{S}$ que construimos en esta demostración satisface las hipótesis del *Lema de Zorn* y (*ii*) que la función k que construimos ahí es inyectiva.

La Proposición 3.5.13 determina, esencialmente, los cuadrados de todos los cardinales infinitos. Podemos extender esto a todas las potencias finitas.

Ejercicio 3.5.15. Muestre que si α es un cardinal infinito, entonces $\alpha^n = \alpha$ para todo entero positivo n.

Corolario 3.5.16. *El conjunto $\mathscr{P}_f(X)$ de todos los subconjuntos finitos de un conjunto infinito X equipotente con X.*

Demostración. Sea n un elemento de $\mathbb{N}$ y sea $\mathscr{P}_{\leq n}(X)$ el conjunto de todos los subconjuntos de X finitos de cardinal a lo sumo n. Es claro que un subconjunto Y de X pertenece a $\mathscr{P}_{\leq n}(X)$ exactamente cuando hay una función $[\![n]\!] \to X$ que lo tiene por imagen, y esto significa que hay una función $f \in X^{[\![n]\!]} \mapsto f([\![n]\!]) \in \mathscr{P}_{\leq n}(X)$ y que es sobreyectiva. Vemos así que $\mathscr{P}_{\leq n}(X) \leq X^{[\![n]\!]} \sim X$, de acuerdo al Ejercicio 3.5.15, y que hay, por lo tanto, una función sobreyectiva $X \to \mathscr{P}_{\leq n}(X)$.

Gracias a todo esto y al axioma de elección podemos concluir que hay una sucesión $(h_n)_{n\in\mathbb{N}}$ que para cada $n \in \mathbb{N}$ tiene como componente n–ésima una función sobryectiva $h_n : X \to \mathscr{P}_{\leq n}(X)$. Se sigue de eso que hay una función

$$(x, n) \in X \times \mathbb{N}_0 \mapsto h_{n+1}(x) \in \bigcup_{n\in\mathbb{N}} \mathscr{P}_{\leq n}(X) = \mathscr{P}_f(X),$$

y que es sobreyectiva, de manera que tenemos que $X \sim X \times \mathbb{N}_0 \leq \mathscr{P}_f(X)$. Por otro lado, la función $x \in X \mapsto \{x\} \in \mathscr{P}_f(X)$ es inyectiva, así que $X \leq \mathscr{P}_f(X)$ y, en vista del *Teorema de Cantor–Bernstein–Schröder* 3.1.9, tenemos que $X \sim \mathscr{P}_f(X)$, como afirma el corolario. □

La Proposición 3.5.13 nos permite probar que el producto de cardinales infinitos coincide con

su suma.

Proposición 3.5.17. *Sean* $\mathfrak{a}$ *y* $\mathfrak{b}$ *dos cardinales. Si alguno de los dos es infinito y ninguno de los dos es nulo, entonces*

$$\mathfrak{a} \cdot \mathfrak{b} = \max\{\mathfrak{a}, \mathfrak{b}\}.$$

Demostración. Sin pérdida de generalidad podemos suponer que $1 \leq \mathfrak{a} \leq \mathfrak{b}$ y que $\mathfrak{b}$ es infinito. Tenemos entonces que $\mathfrak{b} = 1 \cdot \mathfrak{b} \leq \mathfrak{a} \cdot \mathfrak{b} \leq \mathfrak{b} \cdot \mathfrak{b} \leq \mathfrak{b}$, así que $\mathfrak{a} \cdot \mathfrak{b} = \mathfrak{b} = \max\{\mathfrak{a}, \mathfrak{b}\}$. □

Probemos ahora un resultado similar a los que obtuvimos para la suma y el producto para la exponenciación.

Proposición 3.5.18. *Sean* $\mathfrak{a}$ *y* b *dos cardinales. Si* $\mathfrak{a}$ *es infinito y* $2 \leq \mathfrak{b} \leq 2^{\mathfrak{a}}$*, entonces* $\mathfrak{b}^{\mathfrak{a}} = 2^{\mathfrak{a}}$.

Demostración. Si $\mathfrak{b}$ es un cardinal tal que $2 \leq \mathfrak{b} \leq 2^{\mathfrak{a}}$, entonces

$$2^{\mathfrak{a}} \leq \mathfrak{b}^{\mathfrak{a}} \leq (2^{\mathfrak{a}})^{\mathfrak{a}} = 2^{\mathfrak{a} \cdot \mathfrak{a}} = 2^{\mathfrak{a}},$$

así que, de hecho, es $\mathfrak{b}^{\mathfrak{a}} = 2^{\mathfrak{a}}$. □

Es mucho más difícil obtener resultados como el de esta proposición pero para desigualdades estrictas de cardinales, como remarcamos en la Observación 3.5.5.

Ejercicio 3.5.19. Sean $\mathfrak{a}$, $\mathfrak{b}$ y $\mathfrak{c}$ tres cardinales. Muestre que si $\mathfrak{a} + \mathfrak{b} = \mathfrak{b}$ y $\mathfrak{b} \leq \mathfrak{c}$, entonces $\mathfrak{a} + \mathfrak{c} = \mathfrak{c}$.

Ejercicio 3.5.20. Determine los cardinales de los conjuntos $\mathbb{R}^{\mathbb{R}}$ de todas las funciones $\mathbb{R} \to \mathbb{R}$ y $C(\mathbb{R})$ de todas las funciones $\mathbb{R} \to \mathbb{R}$ que son continuas.

Terminemos esta sección resolviendo dos problemas de álgebra lineal. Primero, mostremos que la dimensión de un espacio vectorial está bien definida en todos los casos — inclusive cuando el espacio no es finitamente generado.

Proposición 3.5.21. *Todas las bases de un espacio vectorial tienen el mismo cardinal.*

Demostración. Sea $\mathbb{k}$ un cuerpo, sea V un espacio vectorial sobre $\mathbb{k}$, y sean B y B' dos bases de V. Si alguna de las dos es finita, entonces el álgebra lineal nos dice que las dos lo son y que B y B' tienen el mismo cardinal. Supongamos entonces que tanto B como B' son infinitas. Probemos que

ningún subconjunto infinito de V *de cardinal menor que el de* B *genera a* V.

Esto implica que $B \leq B'$, claro, y, por simetría, que también $B' \leq B$, así que el *Teorma de Cantor–*

Bernstein–Schröder 3.1.9 nos permite concluir que B y B tienen el mismo cardinal y probar la proposición.

Sea G un subconjunto infinito de V tal que $G \prec B$. Para cada elemento g de G hay un subconjunto finito B_g de B tal que $g \in \langle B_g \rangle$ y una función sobreyectiva $\phi_g : \mathbb{N}_0 \to B$ cuya imagen es exactamente B_g. La imagen I de la función

$$(g, n) \in G \times \mathbb{N}_0 \mapsto \phi_g(n) \in B$$

es precisamente $\bigcup_{g \in G} B_g$ y es claro que $\langle I \rangle \supseteq \langle G \rangle$. Ahora bien, I es un subconjunto de B y

$$I \preceq G \times \mathbb{N}_0 \sim G \prec B$$

ya que el conjunto G es infinito. Esto nos permite concluir que I es un subconjunto propio de B y, como B es una base y, por lo tanto, un subconjunto generador minimal de V, que que $\langle G \rangle \subseteq \langle I \rangle \subsetneq V$. Vemos así que G no genera a V, como queremos. □

Una consecuencia de esto es que la afirmación análoga a la del *Teorema de Cantor–Bernstein–Schröder* para funciones lineales es cierta.

Proposición 3.5.22. *Sea $\Bbbk$ un cuerpo y sean V y W dos espacios vectoriales sobre $\Bbbk$. Si hay funciones lineales inyectivas $V \to W$ y $W \to V$, entonces hay un isomorfismo $V \to W$.*

Notemos que el *Teorema de Cantor–Bernstein–Schröder* 3.1.9 nos permite concluir en esa situación que hay una función biyectiva $V \to W$, pero no que esta función es lineal — y la función que encontramos en la demostración de ese teorema bien puede no serlo!

Demostración. Sean $f : V \to W$ y $g : W \to V$ dos funciones lineales inyectivas, y sean B y B' bases de V y de W, respectivamente. Como la función f es inyectiva, el conjunto $f(B)$ es linealmente independiente y entonces hay una base B_1' de W tal que $f(B) \subseteq B_1'$. Como la función f es inyectiva, su restricción a una función $B \to B_1'$ también lo es, y esto y la Proposición 3.5.21 nos dicen que $B \preceq B_1' \sim B'$. De manera similar pero usando la función g podemos ver que $B' \preceq B$ y, por lo tanto, gracias al *Teorema de Cantor–Bernstein–Schröder* 3.1.9, que $B \sim B'$, de manera que V y W tienen la misma dimensión y son, por ello, isomorfos. □

Observación 3.5.23. Hay clases de objetos para las que vale una versión del *Teorema de Cantor–Bernstein–Schröder* y otras para las que no. Acabamos de mostrar que sí vale para espacios vectoriales. Listemos algunos otros ejemplos para los que ocurre lo mismo.

- Los cuerpos algebraicamente cerrados. Esto es consecuencia de que un cuerpo algebraicamente cerrado queda completamente determinados a menos de isomorfismo por su grado de trascendencia sobre su cuerpo primo y su característica, y la demostración es muy similar

a la que hicimos arriba para los espacios vectoriales.

- Los *grupos abelianos*. Irving Kaplansky da en [Kap1954] una demostración elemental de eso.
- Los *espacios topológicos finitos*. Sean $f : X \to Y$ y $g : Y \to X$ dos funciones continuas inyectivas entre espacios topológicos finitos. Es claro que se trata de biyecciones, y si τ_X y τ_Y son las topologías de X y de Y, respectivamente, entonces las funciones $U \in \tau_Y \to f^{-1}(U) \in \tau_X$ y $U \in \tau_X \to g^{-1}(U) \in \tau_Y$ también son inyectivas, de manera que se trata de biyecciones, así que f y g son homeomorfismos.
- Los *módulos inyectivos sobre un anillo*. Richard Bumby prueba esto en [Bum1965].

Hay muchas otras estructuras en las que el análogo del teorema *no* vale.

- Los *grupos*. Por ejemplo, si F_2 y F_3 son los grupos libres en 2 y en 3 generadores, entonces hay morfismos inyectivos $F_2 \to F_3$ y $F_3 \to F_2$, pero ciertamente esos dos grupos no son isomorfos.
- Los *cuerpos*. Por ejemplo, hay morfismos de anillos $\mathbb{C} \to \mathbb{C}(x)$ y $\mathbb{C}(x) \to \mathbb{C}$ y los cuerpos $\mathbb{C}$ y $\mathbb{C}(x)$ no son isomorfos — uno solo de los dos es algebraicamente cerrado.
- Los *módulos proyectivos sobre un anillo*. Por ejemplo, el álgebra libre $A \coloneqq \mathbb{Q}\langle x, y\rangle$ tiene la propiedad de que hay morfismos de módulos inyectivos $A \to A^\infty$ y $A^\infty \to A$, pero A y A^∞ no son isomorfos — aquí A^∞ denota un módulo libre con base numerable.
- Los *espacios topológicos*. Claramente hay funciones continuas inyectivas $(0,1) \to [0,1]$ y $[0,1] \to (0,1)$, pero los espacios $(0,1)$ y $[0,1]$ no son homeomorfismos — uno solo de los dos es compacto.

Usando lo que sabemos sobre operaciones con cardinales, podemos determinar el cardinal de cualquier espacio vectorial.

Proposición 3.5.24. *Sea $\Bbbk$ un cuerpo, sea V un espacio vectorial sobre $\Bbbk$, y sean κ y β los cardinales de $\Bbbk$ y de V, respectivamente.*

(*i*) *Si β es finito, entonces $|V| = \kappa^\beta$.*
(*ii*) *Si β es infinito, entonces $|V| = \max\{\kappa, \beta\}$.*

Demostración. Sea B una base de V. Para cada función $f : B \to \Bbbk$ sea $S(f) \coloneqq \{b \in B : f(b) \neq 0\}$ y consideremos los conjuntos

$$\Bbbk^{(B)} \coloneqq \{f \in \Bbbk^B : \text{el conjunto } S(f) \text{ es finito}\},$$

que es un subespacio de $\Bbbk^B$ que coincide con $\Bbbk^B$ cuando B es finito. Como B es una base de V, para cada vector v de V hay exactamente un elemento f_v de $\Bbbk^{(B)}$ tal que $v = \sum_{b \in S(f_v)} f(b)b$, y se sigue de esto inmediatamente que tenemos una función $\phi : v \in V \mapsto f_v \in \Bbbk^{(B)}$ que es biyectiva. Esto implica que $V \sim \Bbbk^{(B)}$.

Si β es finito, entonces $V \sim \mathbb{k}^{(B)} = \mathbb{k}^B$, así que $|B| = \kappa^\beta$ y la primera afirmación de la proposición es cierta. Supongamos desde ahora que β es infinito. Si b es un elemento de B y t uno de $\mathbb{k}$, entonces hay una función $\delta_{b,t} : B \to \mathbb{k}$ que en cada elemento c de B toma el valor

$$\delta_{b,t}(c) = \begin{cases} t & \text{si } c = b; \\ 0 & \text{en caso contrario.} \end{cases}$$

La función $b \in B \mapsto \delta_{b,1} \in \mathbb{k}^{(B)}$ es inyectiva, así que $\beta \leq \left|\mathbb{k}^{(B)}\right|$, y si b_0 es un elemento cualquiera de B, entonces la función $t \in \mathbb{k} \mapsto \delta_{b_0,t} \in \mathbb{k}^{(B)}$ es inyectiva, así que también $\kappa \leq \left|\mathbb{k}^{(B)}\right|$. Podemos concluir entonces que

$$\max\{\kappa, \beta\} \leq \left|\mathbb{k}^{(B)}\right|. \tag{3.13}$$

Por otro lado, si F es un subconjunto finito de B, entonces

$$\mathbb{k}^{(B),F} \coloneqq \{f \in \mathbb{k}^B : S(f) \subseteq F\}$$

es un subespacio de $\mathbb{k}^{(B)}$, y si n es el cardinal de F y $s : [\![n]\!] \to B$ es una biyección, entonces la función

$$h \in \mathbb{k}^{[\![n]\!]} \mapsto \sum_{i=1}^{n} h(i) \cdot s(i) \in \mathbb{k}^{(B),F}$$

es un isomorfismo de espacios vectoriales y, en particular, una biyección, así que

$$\left|\mathbb{k}^{(B),F}\right| = |\mathbb{k}^{[\![n]\!]}| = \kappa^n.$$

El Corolario 3.5.16 nos dice que el conjunto $\mathscr{P}_f(B)$ de todos los subconjuntos finitos de B tiene cardinal β, y con lo anterior vemos que

$$|\mathbb{k}^{(B)}| = \left|\bigcup_{F \in \mathscr{P}_f(B)} \mathbb{k}^{(B),F}\right| \leq |\mathscr{P}_f(B)| \cdot \max_{F \in \mathscr{P}_f(B)} \left|\mathbb{k}^{(B),F}\right| = \beta \cdot \max_{n \in \mathbb{N}_0} \kappa^n.$$

Si $\kappa \leq \beta$, entonces tenemos que $\kappa^n \leq \beta^n = \beta$ cualquiera sea el entero positivo n y que $\kappa^0 = 1 \leq \beta$, así que $\beta \cdot \max_{n \in \mathbb{N}_0} \kappa^n \leq \beta \cdot \beta = \beta = \max\{\kappa, \beta\}$. Si, por el contrario, es $\beta < \kappa$, entonces $\kappa^n = \kappa$ para todo $n \in \mathbb{N}$ y $\kappa^0 = 1 \leq \kappa$, de manera que $\beta \cdot \max_{n \in \mathbb{N}_0} \kappa^n = \beta \cdot \kappa = \max\{\kappa, \beta\}$. Junto con la desigualdad (3.13), esto nos permite concluir que $|V| = |\mathbb{k}^{(B)}| = \max\{\kappa, \beta\}$. □

Recordemos que todo espacio vectorial de dimensión finita es isomorfo a su espacio dual. Nuestro siguiente resultado nos dice que esto, de hecho, caracteriza a los espacios vectoriales de dimensión finita.

Proposición 3.5.25. *Sea $\Bbbk$ un cuerpo y sea V un espacio vectorial sobre $\Bbbk$. Si V tiene dimensión infinita, entonces* $\dim V < \dim V^*$*, de manera que, en particular, V y V^* no son isomorfos.*

Este resultado es clásico. Las primeras demostraciones fueron dadas por Grigoriĭ Mikhaĭlovich Fichtenholz y Leonid Vitaľevich Kantorovitch en [FK1934] y por Felix Hausdorff en [Hau1936], y Nathan Jacobson da en [Jac1953] lo que podríamos llamar la demostración usual. La demostración que daremos es debida a Andrea Ferreti [Fer2010].

Demostración. Dividiremos la demostración en dos pasos bien distintos.

Primer paso. Sea $K \subseteq L$ una extensión de cuerpos y supongamos que la afirmación de la proposición es cierta cuando el cuerpo $\Bbbk$ es K: queremos mostrar que también es cierta cuando el cuerpo $\Bbbk$ es L. Sea, con ese objetivo, V un L-espacio vectorial y sea B una base de V. Sean $K^{(B)}$ y $L^{(B)}$ como en la demostración de la Proposición 3.5.24. Ahí vimos que $V \cong L^{(B)}$ en tanto L-espacios vectoriales.

Para cada $b \in B$ sean $\delta_b^K : B \to K$ y $\delta_b^L : B \to L$ los elementos de $K^{(B)}$ y $L^{(B)}$ que en cada elemento c de B toman los valores

$$\delta_b^K(c) = \begin{cases} 1 & \text{si } c = b; \\ 0 & \text{en caso contrario;} \end{cases} \qquad \delta_b^L(c) = \begin{cases} 1 & \text{si } c = b; \\ 0 & \text{en caso contrario.} \end{cases}$$

Es fácil ver que las funciones $b \in B \mapsto \delta_b^K \in K^{(B)}$ y $b \in B \mapsto \delta_b^L \in L^{(B)}$ son inyectivas y que sus imágenes son bases de $K^{(B)}$ y de $L^{(B)}$ sobre K y sobre L, respectivamente. Notemos además que, como $L \subseteq L$, podemos ver al conjunto $K^{(B)}$ como un *subconjunto* de $L^{(B)}$.

Si $\phi : K^{(B)} \to K$ es una función K-lineal, entonces hay exactamente una función L-lineal $\bar{\phi} : L^{(B)} \to L$ tal que $\bar{\phi}|_{K^{(B)}}$: en efecto, esta es la función L-lineal tal que $\bar{\phi}(\delta_b^L) = \phi(\delta_b^K)$ para cada $b \in N$. Obtenemos de esta forma una función

$$\Gamma : \phi \in (K^{(B)})^* \mapsto \bar{\phi} \in (L^{(B)})^*$$

que es evidentemente inyectiva.

Afirmamos que esta función Γ preserva la independencia lineal, esto es, que

> *si C es un subconjunto linealmente independiente de $(K^{(B)})^*$, entonces $\Gamma(C)$ es un subconjunto linealmente independiente de $(L^{(B)})^*$.* (3.14)

Sea C un subconjunto linealmente independiente de $(K^{(B)})^*$ y supongamos que $\Gamma(C)$ es linealmente dependiente. Como la función Γ es inyectiva, esto implica que hay un entero positivo n, elementos $c_1, \ldots, c_n$ dos a dos distintos de C, y escalares $a_1, \ldots, a_n$ en L tales que

$$a_1 \cdot \Gamma(c_1) + a_2 \cdot \Gamma(c_2) + \cdots + a_n \cdot \Gamma(c_n) = 0 \tag{3.15}$$

en $(L^{(B)})^*$. Consideremos la función

$$\varepsilon : w \in K^{(B)} \mapsto (c_1(w), c_2(w), \ldots, c_n(w)) \in K^n,$$

que es K-lineal, ya que las funciones $c_1, c_2, \ldots, c_n$ lo son. Si λ es un elemento del espacio dual $(K^n)^*$ tal que $\lambda \circ \varepsilon = 0$ y escribimos $(e_1, e_2, \ldots, e_n)$ a la base ordenada estándar de K^n, tenemos que para todo $w \in K^{(B)}$ es

$$\begin{aligned} 0 &= \lambda(\varepsilon(w)) \\ &= \lambda(e_1) \cdot c_1(w) + \lambda(e_2) \cdot c_2(w) + \cdots + \lambda(e_2) \cdot c_n(w) \\ &= \big(\lambda(e_1) \cdot c_1 + \lambda(e_2) \cdot c_2 + \cdots + \lambda(e_2) \cdot c_n\big)(w), \end{aligned}$$

de manera que la función lineal $\lambda(e_1)\cdot c_1+\lambda(e_2)\cdot c_2+\cdots+\lambda(e_2)\cdot c_n \in (K^{(B)})^*$ es idénticamente nula y, por lo tanto, como $c_1, c_2, \ldots, c_n$ son elementos linealmente independientes de $(K^{(B)})^*$, tenemos que $\lambda(e_i) = 0$ para cada $i \in [\![n]\!]$, esto es, $\lambda = 0$. Esto nos dice que el anulador de la imagen de la función ε es nulo, así que esa función es sobreyectiva. En particular, hay elementos $w_1, w_2, \ldots, w_n$ en $K^{(B)}$ tales que $\varepsilon(w_i) = e_i$ para cada $i \in [\![n]\!]$. Notemos que esto nos dice que

$$\bar{c}_i(w_j) = c_i(w_j) = \delta_{i,j} \textit{ para toda elección de } i \textit{ y } j \textit{ en } [\![n]\!].$$

De esto y de (3.15) se sigue que cada cada $j \in [\![n]\!]$ es

$$\begin{aligned} 0 &= \big(a_1 \cdot \Gamma(c_1) + a_2 \cdot \Gamma(c_2) + \cdots + a_n \cdot \Gamma(c_n)\big)(w_j) \\ &= a_1 \cdot \Gamma(c_1)(w_j) + a_2 \cdot \Gamma(c_2)(w_j) + \cdots + a_n \cdot \Gamma(c_n)(w_j) \\ &= a_1 \cdot \bar{c}_1(w_j) + a_2 \cdot \bar{c}_2(w_j) + \cdots + a_n \cdot \bar{c}_n(w_j) \\ &= a_j. \end{aligned}$$

Esto completa la demostración de la afirmación (3.14), como consecuencia de ella y de la hipótesis sobre K tenemos que

$$\dim_L V = \dim_L(L^{(B)}) = \dim_K(K^{(B)}) < \dim_K(K^{(B)})^* \leq \dim_L(L^{(B)})^* = \dim_L V^*.$$

Segundo paso. Recordemos que todo cuerpo $\Bbbk$ contiene un menor subcuerpo, el ***cuerpo primo*** de $\Bbbk$, y que este es o finito o isomorfo a $\mathbb{Q}$ y que es, por lo tanto, siempre contable. En vista de lo que hicimos en el Primer paso de esta demostración, para probar la proposición es suficiente entonces que mostremos que es cierta bajo la hipótesis adicional de que el cuerpo $\Bbbk$ que aparece en ella es contable.

Sea entonces $\Bbbk$ un cuerpo contable, sea V un espacio vectorial sobre $\Bbbk$ de dimensión infinita, sea B una base de V, y escribamos κ y β a los cardinales de $\Bbbk$ y de B, respectivamente, de manera que $\kappa \leq \beta$ y $\dim V = \beta$, y sea $\delta \coloneqq \dim V^*$. Si f es un elemento de $\Bbbk^B$, entonces hay exactamente una función lineal $\lambda_f : V \to \Bbbk$ tal que $\lambda_f(b) = f(b)$ para todo $b \in B$, así que hay una función

$\psi : f \in \Bbbk^B \mapsto \lambda_f \in V^*$. Es claro que se trata de una función lineal inyectiva, y la propiedad característica de las bases de un espacio vectorial implica que es además biyectiva. Vemos así que, $|V^*| = |\Bbbk^B| = \kappa^\beta$. La Proposición 3.5.24 nos dice, por otro lado, que $|V| = \max\{\kappa, \beta\} = \beta$. Como β es infinito y $2 \leq \kappa \leq \beta \leq 2^\beta$, la Proposición 3.5.18 nos dice que $|V^*| = \kappa^\beta = 2^\beta > \beta = |V|$ y la Proposición 3.5.24 que

$$\max\{\kappa, \delta\} = |V^*| > |V| = \max\{\kappa, \beta\} = \beta \geq \kappa,$$

de manera que $\dim V^* = \delta > \beta = \dim V$. □

Capítulo 4

Ordinales

4.1. Conjuntos bien ordenados

En esta sección empezaremos con el trabajo necesario para llegar a la definición de los ordinales, estudiando en detalle los conjuntos ordenados que satisfacen la siguiente condición.

Definición 4.1.1. Un conjunto ordenado está ***bien ordenado*** si está totalmente ordenado y cada uno de sus subconjuntos no vacíos posee un elemento mínimo, y cuando ese es el caso decimos que su orden es un ***buen orden***.

El primer ejemplo de conjunto bien ordenado es bien conocido.

Proposición 4.1.2. *El conjunto $\mathbb{N}_0$ dotado de su orden usual es un conjunto bien ordenado.*

Demostración. Esto es precisamente lo que afirma el *Principio de buena ordenación*. □

Es muy fácil verificar la afirmación del siguiente ejercicio.

Ejercicio 4.1.3. Sea X un conjunto bien ordenado con respecto a una relación de orden R, y sea Y un subconjunto de X. La relación $R \cap (Y \times Y)$ es un buen orden sobre el conjunto Y.

Gracias a esto, a partir de cada conjunto bien ordenado podemos obtener muchos otros considerando sus subconjuntos. En particular, tenemos el siguiente corolario.

Corolario 4.1.4. *Sea n un entero no negativo. El conjunto $[\![n]\!]$ ordenado con respecto a la relación de orden usual de los enteros es un conjunto bien ordenado.* □.

Un conjunto bien ordenado está totalmente ordenado, pero los dos conceptos son distintos.

Ejercicio 4.1.5. Muestre que un conjunto ordenado finito está bien ordenado si y solamente si está totalmente ordenado, y dé un ejemplo de un conjunto totalmente ordenado que no está bien ordenado.

Más adelante daremos muchos más ejemplos concretos de conjuntos bien ordenados. Antes de ello mostraremos que conjuntos bien ordenados hay muchos: esa es una de las razones por las que el siguiente resultado es importante.

Teorema 4.1.6. *Todo conjunto posee un buen orden.*

Este resultado es conocido como el *Teorema de Zermelo*, porque Ernst Zermelo fue el primero en probarlo, originalmente en su artículo [Zer1904] y luego, de una forma distinta, en [Zer1908a]. El primero en enunciar el resultado fue Cantor: lo usó en su trabajo [Can1882] sobre los conjuntos infinitos de $\mathbb{R}$, anunciando que daría más tarde una demostración. Cantor no pudo probarlo, y la validez del teorema pasó a ser una hipótesis en sus trabajos y en los de muchos otros. No todos los matemáticos estaban convencidos de la verdad del teorema, de todas formas. Por ejemplo, en 1904 Gyula König anunció una demostración de que el conjunto de los números reales no posee ningún buen orden — Felix Hausdorff encontró un error en su argumento unas semanas más tarde.

Daremos dos demostraciones de este teorema. La primera, que es una variación de la dada originalmente por Zermelo y que usa de manera directa el *Axioma de elección*, es esencialmente la dada por Vladimir Devidé en [Dev1967]. Es interesante observar que es precisamente para poder llevar a cabo un argumento como este que Zermelo introdujo el *Axioma de elección* en la teoría de conjuntos.

Demostración. Sea X un conjunto y sea $\phi : \mathscr{P}(X) \setminus \{\varnothing\} \to X$ una función tal que $\phi(p) \in p$ para cada $p \in \mathscr{P}(X) \setminus \{\varnothing\}$, cuya existencia es consecuencia del *Axioma de elección*. Para cada subconjunto no vacío p de X escribamos $p' \coloneqq p \setminus \{\phi(p)\}$.

Un subconjunto P de $\mathscr{P}(X)$ es *distinguido* si satisface las siguientes tres condiciones:

(C_1) $X \in P$.

(C_2) Si Q es un subconjunto no vacío de P, entonces $\bigcap_{q \in Q} q$ es un elemento de P.

(C_3) Si p es un elemento no vacío de P, entonces $p' \in P$.

Hay subconjuntos distinguidos de $\mathscr{P}(X)$ — de hecho, $\mathscr{P}(X)$ mismo es uno de ellos — así que podemos considerar la intersección de todos ellos. Escribámosla D. Es fácil ver que se trata de un subconjunto distinguido de $\mathscr{P}(X)$.

Supongamos que el conjunto

$$\Pi \coloneqq \{p \in D : \text{existe } q \in D \text{ tal que } p \nsubseteq q \text{ y } q \nsubseteq p\}$$

no es vacío, de manera que podemos considerar la unión

$$G \coloneqq \bigcup_{p \in \Pi} p$$

de sus elementos. Notemos que $G \neq \varnothing$, ya que todos los elementos de Π son conjuntos no vacíos. El subconjunto

$$\mathscr{S} \coloneqq \{p \in D : p \supseteq G\}$$

de D no es vacío, ya que X pertenece a él, así que que D sea distinguido implica que la intersección

$$T \coloneqq \bigcap_{p \in \mathscr{S}} p$$

es un elemento de D. Como $T \supseteq G \neq \varnothing$, podemos considerar su subconjunto T', que también es un elemento de D. Nos proponemos mostrar que el conjunto

$$E \coloneqq \{p \in D : p \subseteq T' \text{ o } T \subseteq p\}$$

es un subconjunto distinguido de $\mathscr{P}(X)$.

- Como D es distinguido, sabemos que $X \in D$ y, como $T \subseteq X$, es $X \in E$.
- Sea ahora Q un subconjunto no vacío de E y consideremos la intersección $r = \bigcap_{q \in Q} q$. Como $E \subseteq D$ y el conjunto D es distinguido, es claro que $r \in D$. Si hay un elemento q_0 en Q tal que $q_0 \subseteq T'$, entonces $r \subseteq q_0 \subseteq T'$ y, por lo tanto, $r \in E$. Si, en cambio, tenemos que $q \nsubseteq T'$ para todo elemento q de Q, entonces que Q esté contenido en E implica que $T \subseteq q$ para todo elemento q de Q y, por lo tanto, que $T \subseteq \bigcap_{q \in Q} q = r$, de manera que, otra vez, es $r \in E$.
- Sea p un elemento no vacío de E, de manera que, por un lado, es $p \in D$ y, por otro, que

  $$\textit{vale que } p \subseteq T' \textit{ o } T \subseteq p. \tag{4.1}$$

  Supongamos que $p' \notin E$. Como ciertamente $p' \in D$, ya que $p \neq \varnothing$ y $p \in D$, que p' no esté en E nos dice que

  $$p' \nsubseteq T' \tag{4.2}$$

  y que

  $$T \nsubseteq p'. \tag{4.3}$$

  De (4.2) es claro que

  $$p \neq T. \tag{4.4}$$

Por otro lado, si fuera $p \subseteq T'$, entonces tendríamos que $p' \subseteq T'$, en contra de (4.2), así que debe ser $p \nsubseteq T'$ y, de acuerdo a (4.1),

$$T \subseteq p. \tag{4.5}$$

Esto y (4.3) implican que $\phi(p) \in T$. Si fuera $p' \subseteq T$ tendríamos que $p = p' \cup \{\alpha(p)\} \subseteq T$ y, junto con (4.5), que $T = P$, contradiciendo (4.4). Vemos así que

$$p' \nsubseteq T. \tag{4.6}$$

Esto y (4.3) implican que $p' \in \Pi$, así que $p' \subseteq G \subseteq T$. Esto es absurdo, ya que contradice a (4.6). Esta contradicción provino de haber supuesto que p' no está en E, así que $p' \in E$.

Vemos así que E es un subconjunto distinguido de $\mathscr{P}(X)$. De la definición de D, entonces, se sigue que $D \subseteq E$ y, como por supuesto $E \subseteq D$, que finalmente es $D = E$. Esto significa que

para todo $p \in D$ es $p \subseteq T'$ o $T \subseteq p$.

En particular, si p es un elemento cualquiera de D, entonces $p \subseteq T'$ o $T' \subseteq T \subseteq p$, y esto nos dice que $T' \notin \Pi$.

Sea ahora p un elemento de Π, de manera que hay un elemento q de D tal que $p \nsubseteq q$ y $q \nsubseteq p$, y supongamos por un momento que $p \nsubseteq T'$. Como $T' \in \Pi$, esto implica que $T' \subseteq p$. Tenemos entonces que $T' \subsetneq p \subseteq G \subseteq T$, así que $p = T$ y, por lo tanto, $q \subseteq G \subseteq T = p \nsubseteq q$: esto es absurdo. La concluir de esto es que $p \subseteq T'$ para todo $p \in \Pi$, de manera que $G \subseteq T'$ y, por lo tanto, $T' \in \mathscr{S}$. Esto es imposible, ya que $T' \subsetneq T = \bigcap_{q \in \mathscr{S}} q$.

Esta contradicción provino de haber supuesto que el conjunto Π que definimos arriba no es vacío, así que debe ser. Esto significa que el conjunto D está totalmente ordenado con respecto a la relación de inclusión.

Para cada elemento x de X podemos considerar el conjunto

$$p_x \coloneqq \bigcap_{\substack{p \in D \\ x \in p}} p,$$

que es un elemento de D. Afirmamos que

$$\phi(p_x) = x \text{ para todo } x \in X.$$

En efecto, si hubiera un elemento x de X tal que $\phi(p_x) \neq x$, entonces tendríamos que $x \in p'_x$: como $p'_x \in D$, esto implica que $p_x \subseteq p'_x$, lo que es absurdo.

Consideremos en X la relación

$$R \coloneqq \{(x, y) \in X \times X : p_x \supseteq p_y\}.$$

Para completar la demostración del teorema mostraremos que R es un buen orden sobre el conjunto X.

- Si x es un elemento de X, entonces claramente $p_x \supseteq p_x$, así que $(x, x) \in R$. La relación R es, por lo tanto, reflexiva.
- Sean x e y dos elementos de X y supongamos que los pares ordenados (x, y) e (y, x) pertenecen los dos a R, de manera que $p_x \supseteq p_y$ y $p_y \supseteq p_x$. Es entonces $p_x = p_y$ y, de acuerdo a la observación que hicimos arriba, $x = \phi(p_x) = \phi(p_y) = y$. Vemos así que la relación R es anti-simétrica.
- Sean x, y y z tres elementos de X y supongamos que los pares ordenados (x, y) e (y, z) están en R, de manera que $p_x \supseteq p_y$ y $p_y \supseteq p_z$. Tenemos entonces que $p_x \supseteq p_z$ y, por lo tanto, que $(x, z) \in R$. Se sigue de esto que la relación R es transitiva.
- Sean x e y dos elementos de X. Como los conjuntos p_x y p_y pertenecen al conjuntos D y D está totalmente ordenado con respecto a la relación de inclusión, tenemos que $p_x \subseteq p_y$ o $p_y \subseteq p_x$. Esto nos dice que R contiene a (x, y) o a (y, x) y nos permite concluir que la relación de orden R es total.
- Sea S un subconjunto no vacío de X, y consideremos los subconjuntos $H \coloneqq \bigcup_{x \in S} p_x$ de X y $Q \coloneqq \{p \in D : H \subseteq p\}$ de D. Este subconjunto Q de D no es vacío, ya que $X \in Q$, así que podemos considerar su intersección $q \coloneqq \bigcap_{p \in Q} p$, que es un elemento de D. Como $S \subseteq H \subseteq q$, es $q \neq \varnothing$ y podemos considerar el elemento $y \coloneqq \phi(q)$ de X. Es $y \in q \in D$, así que $p_y \subseteq q$.

  Supongamos por un momento que $y \notin S$. Para cada $x \in S$ es $p_x \subseteq q'$: si no fuera ese el caso, habría un elemento x de S tal que $p_x \nsubseteq q'$ y, como D está totalmente ordenado y $p_x, q \in D$ tendríamos que $q' \subsetneq p_x \subseteq H \subseteq q$, de manera que $p_x = q$ e $y = \phi(q) = \phi(p_x) = x \in S$. Esto implica que $H \subseteq q'$, así que $q' \in Q$ y $q \subseteq q'$, lo que es absurdo.

  Vemos así que y es un elemento de S. Como $y \in p_y$ e $y = \phi(q) \notin q'$, tenemos que $p_y \nsubseteq q'$: sabemos que el conjunto D está totalmente ordenado por inclusión y que p_y y q' pertenecen a él, así que $q' \subsetneq p_y \subseteq q$ y, en consecuencia, $p_y = q$. Para cada $x \in S$ tenemos que $p_x \subseteq H \subseteq q = p_y$, así que $(y, x) \in R$. Podemos concluir con todo esto que y es el menor elemento del conjunto S. □

La segunda demostración que daremos del *Teorema de Zermelo* 4.1.6 no usa directamente el *Axioma de elección* sino que se apoya en el *Lema de Zorn*.

Demostración. Sea X un conjunto no vacío y consideremos el conjunto $\mathscr{S}$ de todos los subconjuntos no vacíos R de $X \times X$ tales que R es un buen orden sobre algún subconjunto de X. Si x es un elemento cualquiera de X, entonces claramente $\{(x, x)\}$ es un elemento de $\mathscr{S}$, y esto nos dice que $\mathscr{S} \neq \varnothing$.

Si R y S son dos elementos de $\mathscr{S}$, entonces escribimos $R \preceq S$ si se cumplen la siguientes dos condiciones:

- $R \subseteq S$.
- Para toda elección de x e y en X vale que $(x, y) \in S \wedge (y, y) \in R \implies (x, x) \in R$.

Dejamos al lector la tarea de verificar que $\preceq$ es una relación de orden sobre el conjunto $\mathscr{S}$. Nosotros probaremos que $\mathscr{S}$, cuando lo consideramos como un conjunto ordenado con respecto a esa relación, satisface las hipótesis del *Lema de Zorn*.

Sea $\mathscr{C}$ una cadena en $\mathscr{S}$ y consideremos los conjuntos

$$V \coloneqq \bigcup_{R \in \mathscr{C}} R, \qquad U \coloneqq \{x \in X : (x, x) \in V\}.$$

Es $U \subseteq X$ y $V \subseteq X \times X$. Más aún, si (x, y) es un elemento de V, entonces hay un elemento R de $\mathscr{C}$ tal que $(x, y) \in R$: este conjunto R es una relación de orden sobre un subconjunto A de X, así que necesariamente x e y pertenecen a A y, por lo tanto, los pares ordenados (x, x) e (y, y) pertenecen a R y, en consecuencia, a V: vemos así que x e y están los dos en U. Esto prueba que $V \subseteq U \times U$.

Mostremos que V es un orden sobre el conjunto U.

- Si u es un elemento de U, entonces $(u, u) \in R$. Vemos así que la relación V en U es reflexiva.
- Sean ahora u y u' dos elementos de U, y supongamos que los pares (u, u') y (u', u) están los dos en V, de manera que hay dos elementos R y R' de $\mathscr{C}$ tales que $(u, u') \in R$ y $(u', u) \in R'$. Como $\mathscr{C}$ es una cadena en $\mathscr{S}$, su subconjunto finito $\{R, R'\}$ tiene un elemento máximo: escribámoslo S. Este conjunto S es una relación de orden sobre un subconjunto de X y es $(u, u') \in R \subseteq S$ y $(u', u) \in R' \subseteq S$: como S es anti-simétrica, esto nos dice que $u = u'$. La relación V en el conjunto U es, por lo tanto, ella misma también anti-simétrica.
- Sean ahora u, u' y u'' elementos de U y supongamos que los pares (u, u') y (u', u'') están en V, de manera que hay elementos R y R' de $\mathscr{C}$ tales que $(u, u') \in R$ y $(u', u'') \in R'$. Como $\mathscr{C}$ es una cadena en $\mathscr{S}$, su subconjunto finito $\{R, R'\}$ tiene un elemento máximo. Escribámoslo S. Es $(u, u') \in R \subseteq S$ y $(u', u'') \in R' \subseteq S$: como la relación S es un orden sobre un subconjunto de X es transitiva y entonces $(u, u'') \in S \subseteq V$. Esto nos dice que la relación V sobre el conjunto U es transitiva.

Veamos ahora que la relación de orden V sobre el conjunto U es total.

- Sean u y u' dos elementos de U, de manera que hay dos elementos R y R' de $\mathscr{C}$ tales que $(u, u) \in R$ y $(u', u') \in R'$. El conjunto $\{R, R'\}$ es un subconjunto finito de la cadena $\mathscr{C}$, así que posee un elemento máximo. Escribámoslo S. Es $(u, u) \in R \subseteq S$ y $(u', u') \in R' \subseteq S$. Como S es una relación de orden total sobre un subconjunto A de X, es u, $u' \in A$ y alguno de los dos pares (u, u') y (u', u) pertenece a S. Esto implica, claro, que alguno de esos dos pares pertenece a V y, por lo tanto, que V es una relación de orden total en U.

Finalmente, mostremos que V es un buen orden sobre U.

- Sea A un subconjunto no vacío de U. Como A no es vacío, podemos elegir en él un elemento a. Este elemento pertenece a U, así que hay un elemento R de $\mathscr{C}$ tal que $(a, a) \in R$. Ahora bien, R es un buen orden sobre un subconjunto Y de X, y la intersección $Y \cap A$ no es vacía, así que

podemos considerer el menor elemento de $Y \cap A$ con respecto al orden R: escribámoslo a. Por supuesto, a es un elemento de A. Queremos mostrar que es el menor elemento de A con respecto al orden V.

Sea b un elemento de A. Como $b \in U$, hay un elemento R' de $\mathscr{C}$ tal que $(b, b) \in R'$. Sea S el máximo elemento del subconjunto finito $\{R, R'\}$ de $\mathscr{C}$, que existe porque $\mathscr{C}$ es una cadena. Este conjunto S es entonces un buen orden sobre un subconjunto Z de X y, como $(a, a) \in R \subseteq S$ y $(b, b) \in R' \subseteq S$, tenemos que $a, b \in Z$. Como S es un orden total, al menos uno de los pares (a, b) y (b, a) está en S.

- Si $(a, b) \in S$, entonces $(a, b) \in V$.
- Supongamos que, por el contrario, $(a, b) \notin S$, de manera que $(b, a) \in S$. Como $(a, a) \in R$, $(b, a) \in S$, y $R \preceq S$, la segunda condición de la definición de la relación $\preceq$ nos dice que $(b, b) \in R$ y, por lo tanto, que $b \in Y$ y, en definitiva, que $b \in A \cap Y$. Como a es el menor elemento de $Y \cap A$ con respecto a R, tenemos que $(a, b) \in R \subseteq V$.

En cualquier caso, entonces, es $(a, b) \in V$, y esto prueba que a es el menor elemento del conjunto A con respecto al orden V.

Para completar la verificación de que el conjunto ordenado $\mathscr{S}$ satisface la hipótesis del *Lema de Zorn* es suficiente que mostremos que V es una cota superior para la cadena $\mathscr{C}$.

- Sea R un elemento de $\mathscr{C}$. De la definición de V es claro que $R \subseteq V$. Sean, por otro lado, x e y dos elementos de X y supongamos que $(x, y) \in V$ y que $(y, y) \in R$. Como (x, y) está en V, hay un elemento S de $\mathscr{C}$ tal que $(x, y) \in S$. Tanto R como S pertenecen a la cadena $\mathscr{C}$, así que $S \preceq R$ o $R \preceq S$.
  - Si $S \preceq R$, entonces $(x, x) \in S \subseteq R$.
  - Si $S \not\preceq R$, entonces es $R \preceq S$ y, como $(x, y) \in S$ y $(y, y) \in R$, sabemos que $(x, x) \in R$.

  En cualquier caso, entonces, es $(x, x) \in R$ y, por lo tanto, se satisface la segunda condición de la definición de $\preceq$, así que $R \preceq V$.

En vista de todo esto, sabemos que hay en el conjunto ordenado $\mathscr{S}$ elementos maximales. Sea M uno de ellos, y sea Y el subconjunto de X sobre el cual M es un buen orden. Para probar el teorema es suficiente que mostremos que necesariamente es $Y = X$.

Supongamos que esto no es así, de manera que hay un elemento ω en $X \setminus Y$, y consideremos los subconjuntos $Y' \coloneqq Y \cup \{\omega\}$ de X y

$$M' \coloneqq M \cup \{(y, \omega) : y \in Y'\}$$

de $X \times X$. Notemos que

$$M' \cap (\{\omega\} \times Y') = \{(\omega, \omega)\}. \tag{4.7}$$

Es claro que $M \subsetneq M'$, así que para llegar a una contradicción y con eso completar la prueba del

teorema es suficiente que mostremos que M' es un elemento de $\mathscr{S}$ tal que $M \preceq M'$, ya que esto contradice la maximalidad de M.

- Si (x, x') es un elemento de M', entonces o bien $(x, x') \in M$, y en ese caso es $x, x' \in Y \subseteq Y'$, o bien hay un elemento y der Y' tal que $(x, x') = (y, \omega)$, de manera que también $x, x' \in Y$. Esto nos dice que $M' \subseteq Y' \times Y'$.
- Sea y un elemento de Y'. Si $y \in Y$, entonces $(y, y) \in M \subseteq M'$, ya que M es una relación reflexiva sobre Y; si $y \notin Y$, entonces necesariamente $y = \omega$ y de la definición de M' es claro que $(y, y) = (\omega, \omega) \in M'$. Vemos así que M' es una relación reflexiva sobre Y'.
- Sean y e y' dos elementos de Y', y supongamos que los pares ordenados (y, y') e (y', y) están los dos en M'. Si alguno de y o y' es ω, entonces de (4.7) es claro que los dos son iguales a ω, así que $y = y'$. Si en cambio ninguno de esos dos elementos es igual a ω, entonces los dos están en Y, así que (y, y') e (y', y) están los dos en M y, como M es una relación anti-simétrica, tenemos otra vez que $y = y'$. Esto nos dice que M' es una relación anti-simétrica.
- Sean y, y' e y'' tres elementos de Y' y supongamos que los pares (y, y') e (y', y'') están los dos en M'. Si $y = \omega$, entonces de (4.7) vemos que $y' = \omega$ y que $y'' = \omega$, así que ciertamente $(y, y'') \in M'$. Si $y' = \omega$, entonces otra vez de (4.7) vemos que $y'' = \omega$ y que, por lo tanto, $(y, y'') \in M'$. Si $y'' = \omega$, entonces es claro que $(y, y'') \in M'$. Finalmente, si ninguno de los tres elementos y, y' e y'' de Y' es ω, entonces los tres pertenecen a Y y tenemos que $(y, y'), (y', y'') \in M$, así que otra vez $(y, y'') \in M \in M'$, ya que M es una relación transitiva. Vemos así que M' es una relación transitiva sobre el conjunto Y'.

Estas tres observaciones prueban que M' es una relación de orden sobre el subconjunto Y' de X.

- Sean y e y' dos elementos de Y'. Si $y = \omega$ o $y' = \omega$, entonces $(y', y) \in M'$ o $(y, y') \in M'$, respectivamente. Si, en cambio, ni y ni y' es igual a ω, entonces $y, y' \in Y$ y, como M es una relación de orden total sobre Y, alguno de (y, y') o (y', y) está en M y, por lo tanto, en M'. La relación de orden M' sobre Y' es, en consecuencia, total.
- Sea A un subconjunto no vacío de Y'. Si $A = \{\omega\}$, entonces claramente ω es el menor elemento de A con respecto al orden M'. Si $A \neq \{\omega\}$, entonces la intersección $A \cap Y$ no es vacía y posee un menor elemento con respecto al orden M de Y. Escribámoslo a. Si b es un elemento cualquiera de A, entonces o bien $b = \omega$ y es $(a, b) \in M'$, o bien $b \in Y$, y en ese caso $b \in A \cap Y$ y, por lo tanto, $(a, b) \in M \subseteq M'$. Vemos así que a es el menor elemento de A con respecto al orden M'.

Sabemos en este punto que M' es un elemento de $\mathscr{S}$ y claramente $M \subsetneq M'$. Nos resta probar que $M \preceq M'$. Sean x e y dos elementos de X y supongamos que $(x, y) \in M'$ y que $(y, y) \in M$. Si fuera $x = \omega$, de que (x, y) está en M' tendríamos que $y = \omega$ y, por lo tanto, que $(\omega, \omega) \in M$, lo que es absurdo. Debe ser, entonces, $x \in Y$ y, por lo tanto, $(x, x) \in M$.

La prueba del teorema queda así completa. □

Las dos demostraciones que hemos dado del *Teorema de Zermelo* 4.1.6 dependen, directa o indirectamente, del *Axioma de elección*. De hecho, las dos afirmaciones son equivalentes — eso es lo que nos dice el siguiente resultado.

Proposición 4.1.7. *Sea X un conjunto y sea $\preceq$ un buen orden sobre X. Si $\mathscr{S}$ es una familia de subconjuntos no vacíos de X, entonces hay una función $c : \mathscr{S} \to X$ tal que $c(A) \in A$ para todo $A \in \mathscr{S}$.*

Demostración. Sea $\mathscr{S}$ una familia de subconjuntos no vacís de X. Como X está bien ordenado con respecto a $\preceq$, para cada elemento A de $\mathscr{S}$ podemos, ya que A no es vacío, considerar el menor elemento $\min A$ de A con respecto al buen orden $\preceq$. Obtenemos así una función $c : A \in \mathscr{S} \mapsto \min A \in X$ que claramente satisface la condición de la proposición. □

El *Teorema de Zermelo* 4.1.6 nos dice que todo conjunto posee un buen orden. En general, claro, posee muchos, y nos interesa entender qué relaciones hay entre todos ellos. Para hacer eso, empezamos haciendo la siguiente definición bien general.

Definición 4.1.8. Sean X e Y dos conjuntos ordenados. Una función $f : X \to Y$ ***preserva el orden*** si para toda elección de x e y en X se tiene que

$$x \preceq y \iff f(x) \preceq f(y),$$

es un ***isomorfismo*** si además es biyectiva, y un ***automorfismo*** si más aún $Y = X$.

Decimos que los conjuntos ordenados X e Y son ***isomorfos*** si hay un isomorfismo $X \to Y$, y en ese caso escribimos $X \cong Y$.

En esta definición usamos la relación de orden $\preceq$, pero muchas veces es conveniente poder verificar esa condición usando la relación de orden estrícto $\prec$ asociada a $\preceq$.

Ejercicio 4.1.9. Sean X e Y dos conjuntos ordenados. Muestre que una función $f : X \to Y$ preserva el orden si y solamente si para toda elección de x e y en X se tiene que

$$x \prec y \iff f(x) \prec f(y).$$

Por otro lado, en este capítulo estamos por sobre todo interesados en conjuntos totalmente ordenados y para ellos podemos simplificar la condición de la Definición 4.1.8.

Ejercicio 4.1.10. Sean X e Y dos conjuntos *totalmente* ordenados. Muestre que una función biyectiva $f : X \to Y$ preserva el orden si y solamente si para toda elección de x e y en X se tiene

que

$$x \preceq y \implies f(x) \preceq f(y).$$

Un isomorfismo es claramente una función inversible, y el resultado del siguiente ejercicio nos dice que su función inversa también es un isomorfismo.

Ejercicio 4.1.11. Sean X e Y dos conjuntos ordenados. Muestre que si $f : X \to Y$ es un isomorfismo, entonces la función inversa $f^{-1} : Y \to X$ también lo es, y use eso para probar que la relación de isomorfismo entre onjuntos ordenados es una relación de equivalencia.

Después de todos estos preliminares, podemos describir la propiedad más básica de las funciones que preservan un buen orden.

Lema 4.1.12. *Sea X un conjunto bien ordenado. Si $f : X \to X$ es una función que preserva el orden, entonces $x \preceq f(x)$ para todo $x \in X$.*

Demostración. Sea $f : X \to X$ una función que preserva el orden y supogamos, para llegar a una contradicción, que el conjunto $P \coloneqq \{x \in X : f(x) \prec x\}$ no es vacío. Como X está bien ordenado, podemos considerar su menor elemento $m \coloneqq \min P$. Como $m \in P$, tenemos que $m \prec f(m)$ y, como f preserva el orden, esto implica que $f(m) \prec f(f(m))$, de manera que $f(m) \in P$: esto es absurdo, ya que $f(m) \prec m = \min P$. Esto prueba el lema. □

La consecuencia más importante para nosotros que tiene este lema es la siguiente caracterización de los automorfismos de un conjunto bien ordenado.

Proposición 4.1.13. *Sea X un conjunto bien ordenado. El único isomorfismo $X \to X$ es la función identidad $\mathsf{id}_X : X \to X$.*

Demostración. Sea $f : X \to X$ un isomorfismo y supongamos que $f \neq \mathsf{id}_X$, de manera que hay un elemento x de X tal que $f(x) \neq x$. De acuerdo al lema es $x \preceq f(x)$. Por otro lado, como la función $f^{-1} : X \to X$ también preserva el orden el lema nos dice que también $x \preceq f^{-1}(x)$ y, como f preserva el orden, que $f(x) \preceq f(f^{-1}(x)) = x$. Como $\preceq$ es una relación anti-simétrica, tenemos entonces que $x = f(x)$: esto es absurdo. Est prueba la proposición. □

Que los conjuntos bien ordenados tengan tan pocos automorfismos tiene como consecuencia que entre dos de ellos puede haber como mucho un isomorfismo.

Corolario 4.1.14. *Sean X e Y dos conjuntos bien ordenados. Hay a lo sumo un isomorfismo $X \to Y$.*

Demostración. Supongamos que $f : X \to Y$ y $g : X \to Y$ son dos isomorfismos, y sea $g^{-1} : Y \to X$ el isomorfismo inverso a g. La composición $g^{-1} \circ f : X \to X$ es un isomorfismo, así que la proposición nos dice que $g^{-1} \circ f = \mathsf{id}_X$, de manera que $f = g \circ g^{-1} \circ f = g \circ \mathsf{id}_X = g$. □

Definición 4.1.15. Sea X un conjunto bien ordenado. Si u es un elemento de X, entonces el ***segmento inicial*** de X correspondiente a u es el conjunto $\{x \in X : x \prec u\}$.

Siempre vemos a los segmentos iniciales de un conjunto bien ordenado X como conjuntos ordenados con el orden inducido por el de X — y son entonces conjuntos bien ordenados.

Ejercicio 4.1.16. Sean X, Y y Z tres conjuntos bien ordenados. Muestre que si X es isomorfo a un segmento inicial de Y e Y es isomorfo a un segmento inicial de Z, entonces X es isomorfo a un segmento inicial de Z.

Proposición 4.1.17. *Un conjunto bien ordenado no es isomorfismo a ninguno de sus segmentos iniciales.*

Demostración. Sea X un conjunto ordenado, sea Y el segmento inicial de X correspondiente a un elemento u de X, y supongamos que X e Y son isomorfos, de manera que hay un isomorfismo $f : X \to Y$. Es fácil ver que la función $g : x \in X \mapsto f(x) \in X$ preserva el orden: el Lema 4.1.12 nos dice entonces que $u \preceq g(u) = f(u) \in Y$, de manera que $u \preceq f(u) \prec y$, y esto es absurdo. □

Proposición 4.1.18. *Sean X e Y dos conjuntos bien ordenados. Exactamente una de las siguientes afirmaciones es cierta:*

- *X e Y son isomorfos.*
- *X es isomorfo a un segmento inicial del Y.*
- *Y es isomorfo a un segmento inicial del X.*

Demostración. Para cada $u \in X$ escribamos $X(u)$ al segmento inicial de X correspondiente a u, y para cada $v \in Y$ escribamos $U(v)$ al segmento inicial de Y correspondiente a v. Consideremos el conjunto

$$f \coloneqq \{(x, y) \in X \times Y : X(x) \cong Y(y)\}$$

y sea A el conjunto de los elementos de X tales que hay un elemento y en Y tal que $(x, y) \in f$. Mostremos que para cada elección de a y a' en X se tiene que

$$a \prec a' \in A \implies a \in A. \tag{4.8}$$

Sean a y a' dos elementos de X tales que $a \prec a' \in A$, de manera que hay un elemento y de Y y un isomorfismo $h : X(a') \to Y(y)$. Como h es inyectiva y preserva el orden, es claro que podemos restringirla a una función $k : X(a) \to Y(h(a))$ y que esta es inyectiva y preserva el orden. Por otro lado, si z es un elemento de $Y(h(a))$, entonces que h sea sobreyectiva implica que hay un elemento w de $X(a')$ tal que $z = h(w)$: como $h(w) = z \prec h(a)$, es $w \prec a$ y, por lo tanto, $w \in X(a)$ y $k(w) = z$. Esto nos dice que la función k es un isomorfismo, de manera que $(a, h(a)) \in f$ y, por lo tanto, $a \in A$. Esto prueba (12.2).

Claramente el conjunto f está contenido en $A \times Y$. Queremos mostrar que se trata, de hecho, de una función inyectiva $A \to Y$ que preserva el orden.

- La definición del conjunto A implica inmediatameente que para cada elemento a de A hay un elemento y de Y tal que $(a, y) \in f$.
- Sean ahora a un elemento de A e y e y' dos elementos de Y, y supongamos que los pares ordenados (a, y) y (a, y') pertenecen a f, de manera que $X(a) \cong Y(y)$ y $X(a) \cong Y(y')$. Se sigue de esto que $Y(y) \cong Y(y')$. Como Y está totalmente ordenado, tenemos que $y \prec y'$ o $y' \prec y$ o $y = y'$. Si $y \prec y'$ entonces $Y(y)$ es un segmento inicial de $Y(y')$, y si $y' \prec y$ entonces $Y(y')$ es un segmento inicial de $Y(y)$: la Proposición 4.1.17 nos dice que ninguna de estas dos posibilidades puede ocurrir. Debe ser entonces $y = y'$.
- Sean a y b dos elementos de A tales que $a \neq b$ y $f(a) = f(b)$. Como $X(a) \cong Y(f(a))$ y $X(b) \cong Y(f(b))$, esto implica que $X(a) \cong X(b)$. Como a y b son distintos y X está totalmente ordenado, tenemos que $a \prec b$ o $b \prec a$. En el primer caso $X(a)$ es un segmento inicial de $X(b)$ y en el segundo $X(b)$ es un segundo inicial de $X(a)$, y que $X(a)$ sea isomorfo a $X(b)$ contradice en cualquiera de los dos casos a la Proposición 4.1.17.
- Sean a y b dos elementos de A. Supongamos primero que $a \prec b$, de manera que $X(a)$ es un segmento inicial de $X(b)$. Como $X(a) \cong Y(f(a))$ y $X(b) \cong Y(f(b))$, es claro que $Y(f(a))$ es isomorfo a un segmento inicial de $Y(f(b))$. Si fuera $f(b) \prec f(a)$, entonces $Y(f(b))$ sería un segmento inicial de $Y(f(a))$ y, en consecuencia, $Y(f(a))$ sería isomorfo a uno de sus segmentos iniciales, lo que es imposible por la Proposición 4.1.17. Debe ser entonces $f(a) \preceq f(b)$ y, como f es una función inyectiva, $f(a) \prec f(b)$.

  Supongamos ahora que $f(a) \prec f(b)$ y, para llegar a una contradicción, que $b \prec a$. En estas condiciones $Y(f(a))$ es un segmento inicial de $Y(f(b))$ y $X(b)$ es un segmento inicial de $X(a)$. Como $X(a) \cong Y(f(a))$ y $X(b) \cong Y(f(b))$, esto implica que $X(a)$ es isomorfo a uno de sus segmentos iniciales, lo que es absurdo. Debe ser entonces $a \preceq b$ y, como $f(a) \neq f(b)$, de hecho $\prec b$.

  Con esto hemos probado que $a \prec b \iff f(a) \prec f(b)$, de manera que la función f preserva el orden.

Escribamos B a la imagen de la función f y notemos que para cada par de elementos y e y'

de Y vale que

$$y \prec y' \in B \implies y \in B.$$

Si $y \prec y' \in B$, entonces hay un elemento x' de X tal que $y' = f(x')$, de manera que $Y(y') \cong X(x')$. Como $Y(y)$ es un segmento inicial de $Y(y')$, esto implica que hay un elemento x en de $X(x)$ tal que $Y(y) \cong X(x)$ y, por lo tanto, que $y = f(x) \in B$.

Probemos que

$$\textit{si } A \subseteq X \textit{ y } x \coloneqq X \setminus A\textit{, entonces } A = X(a)$$

y que

$$\textit{si } B \subseteq Y \textit{ e } y \coloneqq Y \setminus B\textit{, entonces } B = Y(b).$$

- Supongamos que $A \subsetneq X$, sea $x \coloneqq \min X \setminus A$, y veamos que $A = X(x)$. Si a es un elemento de A y $x \prec a$, entonces (12.2) nos dice que $x \in A$, lo que es absurdo: para todo $a \in A$ debe ser entonces $a \preceq x$ y, como $a \in A$ y $x \notin A$, $a \prec x$. Podemos concluir que $A \subseteq X(x)$. Por otro lado, si $a \in X(x)$, entonces $a \in X$ y $a \prec x = \min X \setminus A$, de manera que $a \in A$: esto nos dice que $X(x) \subseteq A$.
- Supongamos ahora que $B \subsetneq Y$, pongamos $y \coloneqq \min Y \setminus B$, y mostremos que $B = Y(y)$. Si b es un elemento de $Y(y)$, entonces $b \in Y$ y $b \prec y = \min Y \setminus B$, así que $b \in B$. Por otro lado, si b es un elemento de B y $b \notin Y(y)$, entonces $y \preceq b$ y, como $y \notin B$, $y \prec z \in B$: como observamos arriba, esto implica que $y \in B$, lo que es absurdo. Vemos así que $B = Y(y)$.

Después de todos estos preparativos, podemos probar por fin la proposición. Es claro que la correstricción $f|^B : A \to B$ es un isomorfismo, y Esto implica que $A = X$ o $B = Y$. En efecto, si ese no es el caso y ponemos $x \coloneqq \min X \setminus A$ e $y \coloneqq \min B \setminus Y$, entonces $A = X(x)$, $B = Y(y)$ y f es un isomorfismo de $X(x)$ a $Y(y)$, de manera que $(x, y) \in f$ y $x \in A$: esto es absurdo.

Como consecuencia de esto, alguna de las siguientes tres posibilidades debe cumplirse:

- Es $A = X$ y $B = Y$, y la función f es un isomorfismo de X a Y.
- Es $A \subsetneq X$ y $B = Y$, y la función f es un isomorfismo de un segmento inicial de X a Y.
- Es $A = X$ y $B \subsetneq Y$, y la correstricción $f|^B$ es un isomorfismo de X a un segmento inicial de Y.

Esto significa que alguna de las tres afirmaciones de la proposición es cierta. Finalmente, que no puede ser que dos de ellas sean ciertas es consecuencia inmediata de la Proposición 4.1.17. □

4.2. Ordinales

Definición 4.2.1. Decimos que un conjunto X es ***transitivo*** si $X \subseteq \mathscr{P}(X)$.

Ejercicio 4.2.2. Sea X un conjunto. Muestre que las siguientes tres afirmaciones son equivalentes:

(a) El conjunto X es transitivo.
(b) Todos los elementos de X son conjuntos y $\bigcup_{x \in X} x \subseteq X$.
(c) Todos los elementos de X son conjuntos y para todo $x \in X$ es $x \subseteq X$.

Los elementos de un conjunto transitivo son conjuntos, así que tiene sentido considerar la restricción de la relación $\in$ a él.

Definición 4.2.3. Un conjunto X es un ***ordinal*** si es transitivo y la relación $\in$ es un buen orden en X. Escribimos Ord a la clase de todos los ordinales.

Esta definición es debida a John von Neumann, que la propuso en su trabajo [vN1923] — este es el segundo artículo publicado por von Neumann, quien tenía en ese momento 19 años.

Proposición 4.2.4.

(i) *El conjunto vacío $\varnothing$ es un ordinal.*
(ii) *Todo elemento de un ordinal es un ordinal.*
(iii) *Si α y β son ordinales y $\alpha \subsetneq \beta$, entonces $\alpha = \min \beta \setminus \alpha \in \beta$*
(iv) *Si α y β son dos ordinales, entonces $\alpha \subseteq \beta$ o $\beta \subseteq \alpha$.*

Demostración. La afirmación (i) es consecuencia inmediata de las definiciones.

(ii) Sea α un ordinal y sea β un elemento de α. Como α es transitivo, entonces $\beta \subseteq \alpha$ y como α está bien ordenado con respecto a la relación $\in$ también lo está β. Para probar que β es un ordinal, entonces, tenemos que mostrar que es un conjunto transitivo. Sea, para ello, γ un elemento de β y mostremos que $\gamma \subseteq \beta$. Por supuesto, para ello elegimos un elemento δ de γ y mostramos que $\delta \in \beta$.

Como $\gamma \in \beta \subseteq \alpha$, es $\gamma \in \alpha$ y, como α es transitivo, $\gamma \subseteq \alpha$, de manera que $\delta \in \alpha$. Hagamos las siguientes dos observaciones.

- El conjunto $\{\beta, \gamma\}$ tiene un elemento $\in$-minimal: como $\gamma \in \beta$ no puede ser $\beta \in \gamma$, y como $\delta \in \gamma$ tenemos entonces que $\delta \neq \beta$.
- El conjunto $\{\beta, \gamma, \delta\}$ tiene un elemento $\in$-minimal: como $\delta \in \gamma \in \beta$, no puede ser entonces que valga $\beta \in \delta$.

Ahora bien, β y δ son elementos de α y α está totalmente ordenado con respecto a la relación $\in$, así que esto nos permite concluir que $\delta \in \beta$, como queremos.

(*iii*) Sean α y β dos ordinales y supongamos que $\alpha \subsetneq \beta$, de manera que, en particular, podemos considerar el menor elemento γ de $\beta \setminus \alpha$ con respecto al buen orden $\underline{\in}$. Sea $\beta(\gamma)$ el segmento inicial de β correspondiente a γ. Mostremos que $\alpha = \beta(\gamma)$.

- Si $\delta \in \beta(\gamma)$, entonces $\delta \in \beta$ y $\delta \in \gamma = \min \beta \setminus \alpha$, así que $\delta \in \alpha$.
- Sea, por otro lado, δ un elemento de α. Como α es transitivo, es $\delta \subseteq \alpha$. Es $\gamma \notin \alpha$, así que $\gamma \notin \delta$ y $\gamma \neq \delta$. Como γ y δ pertenecen los dos a β y β está totalmente ordenado con respecto a la relación $\underline{\in}$, debe ser entonces $\delta \in \gamma$ y, por lo tanto, $\delta \in \beta(\gamma)$.

Como γ es un elemento de β, es un ordinal y, en particular, es transitivo. Esto implica que $\beta(\gamma) = \{\delta \in \beta : \delta \in \gamma\} \subseteq \gamma$. Por otro lado, si δ es un elemento de γ, entonces $\delta \in \beta$, ya que la transitividad de β nos dice que $\gamma \subseteq \beta$, y, por lo tanto, $\delta \in \beta(\gamma)$: esto nos dice que $\gamma \subseteq \beta(\gamma)$. Juntando todo podemos concluir que $\alpha = \beta(\gamma) = \gamma \in \beta$.

(*iv*) Sean α y β dos cardinales. La intersección $\gamma \coloneqq \alpha \cap \beta$ está contenida en $\mathscr{P}(\alpha) \cap \mathscr{P}(\beta)$, ya que α y β son transitivos, y entonces también en $\mathscr{P}(\alpha \cap \beta) = \mathscr{P}(\gamma)$. Vemos así que el conjunto γ es transitivo. Como es un subconjunto de α, también está bien ordenado con respecto a la relación $\underline{\in}$, así que se trata de un ordinal.

Si fuera $\gamma \neq \alpha$ y $\gamma \neq \beta$, entonces tendríamos que $\gamma \subsetneq \alpha$ y $\gamma \subsetneq \beta$, así que la afirmación (*iii*) nos diría que $\gamma \in \alpha$ y $\gamma \in \beta$: esto implica que $\gamma \in \alpha \cap \beta = \gamma$, y esto es imposible. Debe ser entonces γ igual a α o a β, y esto implica claramente que $\alpha \subseteq \beta$ o $\beta \subseteq \alpha$. □

Esta proposición tiene varios corolarios importantes.

Corolario 4.2.5. *La clase* Ord *no es vacía y está totalmente ordenada con respecto a la relación de inclusión* $\subseteq$.

Demostración. Esto sigue inmediatamente de las afirmaciones (*i*) y (*iv*) de la Proposición 4.2.4. □

Corolario 4.2.6. *Las restricciones de la relación de inclusión* $\subseteq$ *y de la relación de pertenencia* $\underline{\in}$ *a* Ord *coinciden, esto es, si* α *y* β *son ordinales, entonces*

$$\alpha \subseteq \beta \iff \alpha \underline{\in} \beta.$$

Demostración. Sean α y β dos ordinales. Que la implicación $\Rightarrow$ del corolario vale es consecuencia de que β es un conjunto transitivo y que la implicación $\Leftarrow$ vale es parte de la afirmación (*iii*) de la Proposición 4.2.4. □

Corolario 4.2.7. *Para todo ordinal* α *se tiene que* $\alpha = \{\beta \in \mathrm{Ord} : \beta \subsetneq \alpha\}$.

Demostración. Sea α un ordinal. La tercera afirmación de la Proposición 4.2.4 nos dice que todo ordinal β tal que $\beta \subsetneq \alpha$ es un elemento de α. Por otro lado, si β es un elemento de α, entonces la segunda parte de esa proposición nos dice que β es un ordinal y la transitividad de α implica que $\beta \subsetneq \alpha$. Estas dos observaciones prueban el corolario. □

Corolario 4.2.8. *Dos ordinales isomorfos en tanto conjuntos ordenados son iguales.*

Demostración. Supongamos que hay dos ordinales α y β distintos. Como la relación $\subseteq$ es total en Ord, sabemos que $\alpha \subsetneq \beta$ o $\beta \subsetneq \alpha$. Supongamos que, por ejemplo, es $\beta \subsetneq \alpha$. Si γ es un elemento de β, entonces γ pertenece a α y, por lo tanto, al segmento inicial $\alpha(\beta)$ de α correspondiente a β. Por otro lado, es evidente que $\alpha(\beta) \subseteq \beta$, así que β es el segmento inicial $\alpha(\beta)$ de α y, en particular, no es isomorfo a α. Si en cambio de ser $\beta \subsetneq \alpha$ es $\alpha \subsetneq \beta$ podemos proceder de la misma forma. □

Corolario 4.2.9. *La clase* Ord *está bien ordenada con respecto a la relación de inclusión. Más precisamente, si* C *es una subclase no vacía de* Ord*, entonces la intersección* $\bigcap_{C \in \mathsf{C}} C$ *es un ordinal, pertenece a* C *y es, de hecho, el menor elemento de* C*.*

Demostración. Sea C una subclase no vacía de Ord. Como no es vacía, podemos considerar la intersección $M \coloneqq \bigcap_{C \in \mathsf{C}} C$ de sus elementos. Es claro que

$$\mathscr{P}(M) \subseteq \bigcap_{C \in \mathsf{C}} \mathscr{P}(C) = \mathscr{P}\left(\bigcap_{C \in \mathsf{C}} C\right) = \mathscr{P}(M),$$

así que M es un conjunto transitivo; por otro lado, si C es un elemento cualquiera de C, entonces $M \subseteq C$ y, en consecuencia, M está bien ordenado con respecto a la relación $\underline{\in}$. Vemos con esto que M es un ordinal. Es evidente que $M \subseteq C$ para todo $C \in \mathsf{C}$. Para completar la prueba de todas las afirmaciones del corolario, entonces, tenemos que mostrar que $M \in \mathsf{C}$. Pero si no fuera ese el caso, para todo $C \in \mathsf{C}$ tendríamos que $M \subsetneq C$, de manera que $M \in C$ y, por lo tanto, que $M \in \bigcap_{C \in \mathsf{C}} C = M$: esto es absurdo. □

Corolario 4.2.10. *Si* X *es un conjunto no vacío de ordinales, entonces* $\sigma \coloneqq \bigcup_{\alpha \in X} \alpha$ *es un ordinal y es, de hecho, el supremo de* X *en* Ord*:*

- *Es* $\alpha \subseteq \sigma$ *para todo* $\alpha \in X$*.*
- *Si* τ *es un ordinal tal que* $\alpha \subseteq \tau$ *para todo* $\alpha \in X$*, entonces* $\sigma \subseteq \tau$*.*

Demostración. Sea X un conjunto no vacío de ordinales y sea σ el conjunto descripto en el

enunciado del corolario. Es

$$\sigma = \bigcup_{\alpha\in X} \alpha \subseteq \bigcup_{\alpha\in X} \mathscr{P}(\alpha) \subseteq \mathscr{P}\left(\bigcup_{\alpha\in X} \alpha\right) = \mathscr{P}(\sigma),$$

así que σ es un conjunto transitivo. Mostremos que está bien ordenado con respecto a la relación $\subseteq$.

Sea Y un subconjunto no vacío de σ. Como no es vacío, hay un elemento α de X tal que $Y \cap \alpha \neq \varnothing$ y, como α es un ordinal, podemos considerar el elemento $y \coloneqq \min Y \cap \alpha$. Es claro que $y \in Y$. Sea z un elemento de Y distinto de y. Hay un elemento β de X tal que $z \in \beta$, y es $\alpha \subseteq \beta$ o $\beta \subseteq \alpha$, de manera que $\alpha \cup \beta$ es un ordinal que contiene tanto a y como a z. Si fuera $z \in y$, entonces tendríamos, ya que $y \in \alpha$ y α es transitivo, que $y \subseteq \alpha$, de manera que $z \in \alpha \cap Y$, así que $y = \min \alpha \cap Y \in z \in y$, lo que es absurdo. Es entonces $z \neq y$ y $z \notin y$: como la unión $\alpha \cup \beta$ está bien ordenada con respecto a la relación $\subseteq$, vemos entonces que $y \in z$. Vemos así que y es el menor elemento del conjunto Y y, en definitiva, que el conjunto X está bien ordenado con respecto a la relación $\subseteq$, como queríamos.

Todo esto nos dice que σ es un ordinal, y es claro que $\alpha \subseteq \sigma$ cualquiera sea el elemento α de X. Si, finalmente, τ es un ordinal tal que $\alpha \subseteq \tau$ para todo $\alpha \in X$, entonces claramente es $\sigma = \bigcup_{\alpha\in X} \alpha \subseteq \tau$. Esto completa la prueba del corolario. □

Lema 4.2.11. *Todo segmento inicial de un ordinal es un ordinal.*

Demostración. Sea α un ordinal, sea β un elemento de α, y consideremos el segmento inicial $\alpha(\beta)$ de α correspondiente a β. Si γ es un elemento de $\alpha(\beta)$, entonces $\gamma \in \beta$. Sea, por otro lado, si γ es un elemento de β, entonces que α sea transitivo y $\beta \in \alpha$ implican que $\gamma \in \beta \subseteq \alpha$ y, por lo tanto, que $\gamma \in \alpha(\beta)$. Vemos así que el segmento inicial $\alpha(\beta)$ coincide con β, que, como es un elemento de α, es un ordinal. □

Lema 4.2.12. *Si α es un ordinal, entonces $\beta \coloneqq \alpha \cup \{\alpha\}$ también lo es y es, de hecho, el menor elemento de la clase $\{\gamma \in \mathsf{Ord} : \alpha \subsetneq \gamma\}$.*

Demostración. Sea α un ordinal y pongamos $\beta \coloneqq \alpha \cup \{\alpha\}$. Claramente $\beta \subseteq \mathscr{P}(\alpha) \subseteq \mathscr{P}(\beta)$, así que β es transitivo. Sea X un subconjunto de β. Si $X = \{\alpha\}$, entonces ciertamente X tiene un menor elemento. Si $X \neq \{\alpha\}$, entonces $X \cap \alpha \neq \varnothing$, así que podemos considerar el elemento $x \coloneqq \min X \cap \alpha$. Como $x \in \alpha$, es $x \subseteq \alpha$, ya que α es transitivo, y esto implica claramente que $x \subseteq \gamma$ para todo $\gamma \in X$. Finalmente, como $x \in X$, es claro que x es el menor elemento de X.

Vemos así que β es un ordinal. Claramente $\alpha \subsetneq \beta$, porque $\alpha \notin \alpha$. Por otro lado, si γ es un ordinal tal que γ es un ordinal tal que $\alpha \subsetneq \gamma$, entoncs $\alpha \in \gamma$ y, por lo tanto, $\beta = \alpha \cup \{\alpha\} \subsetneq \gamma$. Esto prueba la afirmación final del lema. □

Una consecuencia de este corolario es que la clase Ord no es un conjunto. Si lo fuera, como

es transitiva y está bien ordenada con respecto a la relación $\underline{\in}$, sería un ordinal y tendríamos que $\mathsf{Ord} \in \mathsf{Ord} \cup \{\mathsf{Ord}\} \in \mathsf{Ord}$: esto es absurdo.

El corolario nos permite dar ejemplos de ordinales. Como sabemos que $\varnothing$ es un ordinal, se sigue del corolario que cada uno de los conjuntos

$$\varnothing,$$
$$\{\varnothing\},$$
$$\{\varnothing, \{\varnothing\}\},$$
$$\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\},$$
$$\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}\},$$
$$\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}, \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}\}\},$$
$$\dots$$

es un ordinal. De hecho, como tenemos la función

$$F : \alpha \in \mathsf{Ord} \mapsto \alpha \cup \{\alpha\} \in \mathsf{Ord},$$

el *Principio de recurrencia* 1.2.1 nos dice que hay exactamente una función $f : \mathbb{N}_0 \to \mathsf{Ord}$ tal que $f(0) = \varnothing$ y $f(n+1) = f(n) \cup \{f(n)\}$ para cada $n \in \mathbb{N}_0$. Los valores de esta función son precisamente los ordinales finitos.

Lema 4.2.13. *El conjunto $F \coloneqq \{f(n) : n \in \mathbb{N}_0\}$ es el de todos los ordinales finitos y la función $n \in \mathbb{N}_0 \mapsto f(n) \in F$ es biyectiva y preserva el orden.*

Desde ahora escribiremos simplemente n para denotar al ordinal $f(n)$. Así, 0, 1, y 2 son, por ejemplo, los ordinales $\varnothing$, $\{\varnothing\}$, y $\{\varnothing, \{\varnothing\}\}$.

Demostración. Mostremos que para todo $n \in \mathbb{N}_0$ vale que

$$\textit{si } \alpha \textit{ es un ordinal finito y } n \textit{ es el cardinal de } \alpha\textit{, entonces } \alpha = f(n). \tag{4.9}$$

Si α es un cardinal finito de cardinal 0, entonces necesariamente $\alpha = \varnothing = f(0)$: esto nos dice que esa afirmación es cierta si n es 0.

Sea ahora m un elemento cualquiera de $\mathbb{N}_0$, supongamos que la afirmación (4.9) es cierta cuando n es m, y sea α un ordinal finito de cardinal $m+1$. Como α es finito y está totalmente ordenado con respecto a la relación de pertenencia $\underline{\in}$, tiene un elemento máximo: esto es, hay un elemento β en α tal que $\gamma \in \beta$ para todo $\gamma \in \alpha$ tal que $\gamma \neq \beta$. Vemos así que $\alpha = \beta \cup \{\beta\}$ y, como $\beta \notin \beta$, que el cardinal de β es m. La hipótesis inductiva, entonces, nos dice que $\beta = f(m)$ y entonces que $\alpha = \beta \cup \{\beta\} = f(m) \cup \{f(m)\} = f(m+1)$.

Esto completa la inducción y prueba que el conjunto del lema contiene todos los ordinales finitos. Probemos ahora que todos los ordinales de ese conjunto son finitos. Es suficiente para ello

que mostremos que para todo $n \in \mathbb{N}_0$

$$\textit{el ordinal } f(n) \textit{ es finito de cardinal } n. \tag{4.10}$$

Esto es fácil. En efecto, es evidente que eso es cierto cuando n es 0, y si n es un elemento cualquiera de $\mathbb{N}_0$ tal que el ordinal $f(n)$ es finito e igual a n, entonces por supuesto tenemos que el ordinal $f(n+1) = f(n) \cup \{f(n)\}$ es finito e igual a $n+1$, ya que $f(n) \notin f(n)$.

Con esto hemos probado que la primera afirmación del enunciado es cierta. Que la función $h : n \in \mathbb{N}_0 \mapsto f(n) \in F$ es sobreyectiva es claro, y que es inyectiva es una consecuencia inmediata de (4.10). De la definición de la función f es claro que

$$f(n) \in f(n+1) \textit{ para todo } n \in \mathbb{N}_0. \tag{4.11}$$

Queremos probar que, más generalmente, cualesquiera sean m y n en $\mathbb{N}_0$ vale que

$$m \leq n \implies f(m) \subseteq f(n).$$

Para ello es suficiente que mostremos que para todo elemento d de $\mathbb{N}_0$ vale que

$$\textit{para todo } m \in \mathbb{N}_0 \textit{ es } f(m) \subseteq f(m+d). \tag{4.12}$$

En efecto, si esto vale y m y n son dos elementos de $\mathbb{N}_0$ tales que $m \leq n$, entonces $n - m \geq 0$ y, por lo tanto, $f(m) \subseteq f(m + (n-m)) = f(n)$. Probamos (4.12) por inducción con respecto a d, notando que cuando d es 0 es evidente. Si d es un elemento de $\mathbb{N}_0$ tal que $f(m) \subseteq f(m+d)$ para todo $m \in \mathbb{N}_0$, entonces para cada $m \in \mathbb{N}_0$ tenemos, de acuerdo a (4.11), que

$$f(m) \subseteq f(m+1) \subseteq f((m+1)+d) = f(m+d+1).$$

Como $\mathbb{N}_0$ y F son conjuntos totalmente ordenados, podemos deducir del Ejercicio 4.1.10 que la función f preserva el orden. □

Usando el Corolario 4.2.10 podemos construir un ordinal infinito.

Lema 4.2.14. *Sea $f : \mathbb{N}_0 \to \mathsf{Ord}$ la función tal que $f(0) = \varnothing$ y $f(n+1) = f(n) \cup \{f(n)\}$ para cada $n \in \mathbb{N}_0$. El conjunto $\omega := \{f(n) : n \in \mathbb{N}_0\}$ de todos los ordinales finitos es un ordinal, y es el menor de todos los ordinales infinitos.*

Demostración. Como ω es ciertamente un conjunto de ordinales, el Corolario 4.2.10 nos dice que su unión $\eta := \bigcup_{\alpha \in \omega} \alpha = \bigcup_{n \in \mathbb{N}_0} f(n)$ es un ordinal y que es el supremo de F en Ord.

Si α es un elemento de η, entonces hay un entero no negativo n tal que $\alpha \in f(n)$ y, como $f(n)$ es un ordinal, $\alpha \subseteq f(n)$, de manera que α es finito e igual a $f(m)$ para algún elemento m de $\mathbb{N}_0$.

Esto nos dice que $\eta \subseteq \omega$. Por otro lado, para cada $n \in \mathbb{N}_0$ es $f(n) \in f(n+1) \subseteq \eta$, así que $\eta = \omega$. Esto prueba la primera afirmación del lema.

Sea, por otro lado, β un ordinal infinito. Como los elementos de ω son finitos, tenemos que $\beta \notin \omega$, así que la buena ordenación de Ord con respecto a la relación $\subseteq$ implica que $\omega \subseteq \beta$. Vemos así que ω es el menor ordinal infinito. □

Teorema 4.2.15. *Todo conjunto bien ordenado es isomorfo a exactamente un ordinal.*

Demostración. La afirmación de unicidad es consecuencia inmediata del Corolario 4.2.8, así que nos ocuparemos solamente de la afirmación de existencia. Sea X un conjunto bien ordenado y para cada $x \in X$ escribamos $X(x)$ al segmento inicial de X correspondiente a x. Empecemos por mostrar que

> *si todo segmento inicial de X es isomorfo a un ordinal, entonces X es isomorfo a un ordinal.* (4.13)

Supongamos que para todo $x \in X$ hay un ordinal α tal que $X(x) \cong \alpha$ y consideremos la clase

$$f \coloneqq \{(x, \alpha) \in X \times \mathsf{Ord} : X(x) \cong \alpha\}.$$

Esto es, de hecho, una función. En efecto, si x es un elemento de X y α y β son dos ordinales tales que $X(x) \cong \alpha$ y $X(x) \cong \beta$, entonces tenemos que $\alpha \cong \beta$ y, de acuerdo al Corolario 4.2.8, que $\alpha = \beta$. El *Axioma de reemplazo* nos dice que

$$C \coloneqq \{\alpha \in \mathsf{Ord} : \text{existe } x \in X \text{ tal que } \alpha \cong X(x)\}$$

es un conjunto y, en consecuencia, que tenemos una función $f : X \to C$.

Mostremos que si x e y son dos elementos de X, entonces

$$x < y \iff f(x) \subsetneq f(y). \tag{4.14}$$

- Probemos primero la implicación $\Rightarrow$. Sean para ello x e y dos elementos de X tales que $x < y$. Hay un isomorfismo $g : X(y) \to f(y)$. Como $x < y$, x pertence a $X(y)$ y podemos considerar el segmento inicial $X(y)(x)$ de $X(y)$ correspondiente a x, que coincide con $X(x)$. Por otro lado, $g(x)$ es un elemento del ordinal $f(y)$ y podemos considerar el segmento inicial $f(y)(g(x))$ de $f(y)$ correspondiente a $g(x)$. Como la función g preserva el orden, se restringe a una función $h : X(x) \to f(y)(g(x))$ y esta es un isomorfismo: es entonces $f(x) = f(y)(g(x)) \subsetneq f(y)$, ya que el segmento inicial $f(y)(g(x))$ de $f(y)$ es un ordinal y f es una función.
- Probemos ahora la implicación $\Leftarrow$. Como X está totalmemente ordenado con respecto a la relación de orden $\leq$ y C está totalmente ordenado por la relación $\subseteq$, ya que sus elementos son ordinales, es lo mismo probar la implicación contrarrecíproca, que es

  $$y \leq x \implies f(y) \subseteq f(x).$$

Sean para ello x e y dos elementos de X y supongamos que $y \preceq x$. De acuerdo a (4.14), o es $y = x$ y, por lo tanto, $f(y) = f(x)$, o es $f(y) \subsetneq f(x)$: así, en cualquiera de los dos casos es $f(y) \subseteq f(x)$.

La función $f : X \to C$ es una biyección. Que es sobreyectiva es consecuencia inmediata de la forma en que definimos al conjunto C. Por otro lado, si x e y son dos elementos de X tales que $f(x) = f(y)$, entonces la afirmación (4.14) implica que $x \preceq y$ y que $y \preceq x$, así que $x = y$.

Para completar la prueba de (4.13) es suficiente que mostremos que C es un ordinal. En efecto, en ese caso sabremos que la función f es una biyección que preserva el orden, así que será un isomorfismo entre X y un ordinal.

- Sea α un elemento de C, de manera que hay un elemento x de X y un isomorfismo $k : X(x) \cong \alpha$. Si β es un elemento de α, entonces hay un elemento y de $X(x)$ tal que $k(y) = \beta$: como k es un isomorfismo, se restringe a un isomorfismo $X(y) = X(x)(y) \to \alpha(\beta) = \beta$, de manera que $\beta = f(y) \in C$, ya que β es un ordinal. Vemos así que $\alpha \subseteq C$ y que, por lo tanto, el conjunto C es transitivo.
- Sea S un subconjunto no vacío de C. La preimagen $f^{-1}(S)$ es entonces un subconjunto no vacío de X y podemos considerar su menor elemento $x \coloneqq \min f^{-1}(S)$ y su imagen $s \coloneqq f(x)$ en C. Como $x \in f^{-1}(S)$ y la función f es sobreyectiva, es $s = f(x) \in f(f^{-1}(S)) = S$. Por otro lado, si t es un elemento de S distinto de s, entonces hay exactamente un elemento y de $f^{-1}(S)$ tal que $t = f(y)$: como $f(y) = t \neq s = f(x)$ y la función f es inyectiva, es $y \neq x$ y, por lo tanto, como $y \in f^{-1}(S)$, es $x \prec y$ y, por (4.14), $s = f(x) \subseteq f(y) = t$. Esto nos dice que el conjunto C está bien ordenado con respecto a la relación de inclusión $\subseteq$ y, como sus elementos son ordinales, esta relación coincide con la relación de pertenencia $\in$.

Todo esto prueba la afirmación (4.13). Para probar el teorema, entonces, es suficiente que mostremos que todo segmento inicial del conjunto bien ordenado X es isomorfo a un ordinal.

Supongamos que no es ese el caso, de manera que no es vacío el conjunto P de los elementos x de X tales que el segmento inicial $X(x)$ de X correspondiente a x no es isomorfo a un ordinal. Como X está bien ordenado, podemos considerar entonces el elemento $m \coloneqq \min P$. Como $m \in P$, el segmento inicial $X(m)$ no es isomorfo a un ordinal y, de acuerdo a (4.13), hay un elemento y de $X(m)$ tal que el segmento inicial $X(m)(y)$ no es isomorfo a un ordinal. Como $X(m)(y) = X(y)$, esto implica que $y \in P$, y esto es imposible ya que $y \prec m$. Vemos así que debe ser $P = \varnothing$, y esto completa, como dijimos, la demostración del teorema. □

Como explicamos en su momento, la Proposición 3.2.9 nos da una clasificación de los conjuntos finitos: nos dice que todo conjunto finito es equipotente a exactamente uno de los conjuntos de la lista

$$[\![0]\!], [\![1]\!], [\![2]\!], [\![3]\!], \ldots$$

Por eso es que decimos que el conjunto $\{[\![n]\!] : n \in \mathbb{N}_0\}$ es un sistema completo de representantes de

las clases de equivalencia de los conjuntos finitos bajo la relación de equipotencia. El Teorema 4.2.15 hace lo mismo con los conjuntos bien ordenados: nos dice que todo conjunto bien ordenado es isomorfo a exactamente un elemento de la clase Ord, esto es, que la clase Ord es un sistema completo e irredundante de representantes de las clases de isomorfismo de conjuntos bien ordenados con respecto a la relación de isomorfismo.

4.3. El principio de inducción transfinita y sus consecuencias

Definición 4.3.1. Sea α un ordinal.

- Llamamos al ordinal $\alpha \cup \{\alpha\}$ el ***sucesor*** de α y lo escribimos $\alpha + 1$.
- Decimos que α es un ***ordinal sucesor*** si hay un ordinal β tal que $\alpha = \beta + 1$, y que es un ***ordinal límite*** en caso contrario.

Recordemos que en la sección anterior construimos una función $f : \mathbb{N}_0 \to \mathsf{Ord}$ tal que $f(0) = \varnothing$ y $f(n+1) = f(n) \cup \{f(n)\}$ para todo $n \in \mathbb{N}_0$ que es inyectiva y cuya imagen es precisamente el conjunto de los ordinales finitos y que, cuando es conveniente, escribimos para cada $n \in \mathbb{N}_0$ simplemente n en lugar de $f(n)$. Observemos que usando esta notación y la de la definición que acabamos de dar tenemos que $f(n+1) = n+1$ para todo $n \in \mathbb{N}_0$.

El ordinal $0 = \varnothing$ es claramente un ordinal límite, y si n es un entero positivo, entonces el ordinal n es un ordinal sucesor: es el sucesor del ordinal $n-1$. Por otro lado, el ordinal ω es un ordinal límite. En efecto, si hubiera un ordinal α tal que $\omega = \alpha + 1 = \alpha \cup \{\alpha\}$, tendríamos que $\alpha \in \omega$, así que α sería finito y, por lo tanto, ω también.

Podemos ahora desarrollar una teoría de la «inducción transfinita», en completa analogía a lo que hicimos en la primera sección del capítulo anterior.

Definición 4.3.2. Una clase C de ordinales es ***inductiva*** si satisface las siguientes dos condiciones:

- Si un ordinal α pertenece a C, entonces también lo hace $\alpha + 1$.
- Si α es un ordinal límite y todo ordinal β menor que α pertenece a C, entonces α mismo pertenece a C.

El siguiente resultado es el *Principio de inducción transfinita*.

Proposición 4.3.3. *Toda clase inductiva de ordinales coincide con* Ord.

Demostración. Sea C una clase de ordinales que satisface esas dos condiciones, y supongamos que es distinta de Ord. Como la clase Ord está bien ordenada con respecto a la relación de pertenencia, podemos considerar el menor elemento α de $\mathsf{Ord} \setminus \mathsf{C}$, que es o un ordinal sucesor o un ordinal límite. En el primer caso hay un ordinal β tal que $\alpha = \beta + 1$. Como $\beta \in \alpha = \min \mathsf{Ord} \setminus \mathsf{C}$, entonces $\beta \in \mathsf{C}$ y, en vista de la hipótesis, también $\alpha = \beta + 1 \in \mathsf{C}$: esto es absurdo. Pero entonces α debe ser un ordinal límite y, como es el menor elemento de $\mathsf{Ord} \setminus \mathsf{C}$, todo ordinal menor que α pertenece a C. Esto otra vez es imposible, ya que la hipótesis nos dice entonces que α está en C. Todo esto nos lleva a concluir que, como queremos, la clase C coincide con Ord. □

Como con la inducción usual, casi siempre usamos el *Principio de inducción transfinita* vía el siguiente corolario.

Corolario 4.3.4. *Para cada ordinal α sea $\mathscr{P}_\alpha$ una afirmación y supongamos que se satisfacen las siguiente dos condiciones:*

- *Para todo ordinal α vale $\mathscr{P}_\alpha \implies \mathscr{P}_{\alpha+1}$.*
- *Si α es un ordinal tal que para todo ordinal β menor que α vale $\mathscr{P}_\beta$, entonces $\mathscr{P}_\alpha$.*

La afirmación $\mathscr{P}_\alpha$ vale entonces cualquier sea el ordinal α.

Demostración. En efecto, en esas condiciones la clase C de todos los ordinales α tales que la afirmación $\mathscr{P}_\alpha$ vale es inductiva, así que coincide con Ord. □

El siguiente es el *Principio de recurrencia transfinita*.

Proposición 4.3.5. *Sea α un ordinal, sea X un conjunto, sea $X^{<\alpha}$ el conjunto de todas las funciones $\beta \to X$ definidas en un ordinal β menor que α, y sea $F : X^{<\alpha} \to X$ una función. Hay exactamente una función $f : \alpha \to X$ tal que para cada ordinal β menor que α vale que $f(\beta) = F(f|_\beta)$.*

Demostración. Digamos que una función $g : \beta \to X$ definida en un ordinal menor o igual que α es *buena* si para todo ordinal γ menor que β vale que $g(\gamma) = F(g|_\gamma)$. Notemos que

> *si β es un ordinal menor o igual que α y $g : \beta \to X$ es una función buena, entonces para todo ordinal γ menor que β la restricción $g|_\gamma : \gamma \to X$ es buena.*

En efecto, si $g : \beta \to X$ es una función buena definida sobre un ordinal β menor o igual que α y γ es un ordinal menor que β, entonces para cada ordinal δ menor que γ tenemos que

$$g|_\gamma(\delta) = g(\delta) = F(g|_\delta) = F((g|_\gamma)|_\delta),$$

y esto nos dice que la restricción $g|_\gamma : \gamma \to X$ es buena.

Consideremos ahora el conjunto B de los ordinales β menores que α tales que hay exactamente una función buena $\beta \to X$. Mostremos que

si β es un ordinal y $\beta \subseteq B$, entonces hay exactamente una función buena $\beta \to X$. (4.15)

Sea para ello β un ordinal tal que $\beta \subseteq B$. El ordinal β puede ser un ordinal sucesor o un ordinal límite, y consideramos los dos casos por separado.

- Supongamos primero que β es un ordinal sucesor, de manera que hay otro ordinal γ tal que $\beta = \gamma + 1$. Como $\gamma \in \beta \subseteq B$, hay una exactamente una función buena $g : \gamma \to X$. Podemos entonces considerar la función $h : \beta = \gamma \cup \{\gamma\} \to X$ que en cada elemento δ de β toma el valor

$$h(\delta) = \begin{cases} g(\delta) & \text{si } \delta \in \gamma; \\ F(g) & \text{si } \delta = \gamma. \end{cases}$$

  Esta función es buena. En efecto, si δ es un ordinal menor que β, entonces o bien $\delta \in \gamma$, y en ese caso $\delta \subseteq \gamma$, $h|_\delta = g|_\delta$, y $h(\delta) = g(\delta) = F(g|_\delta) = F(h|_\delta)$ porque g es buena, o bien $\delta = \gamma$, y en ese caso es $g = h|_\delta$ y, por lo tanto, $h(\delta) = F(g) = F(h|_\delta)$. De hecho, esa función $h : \beta \to X$ es la única función buena definida sobre β. Si $h' : \beta \to X$ es otra función buena, entonces la restricción $h'|_\gamma : \gamma \to X$ es buena, así que que γ pertenezca a B implica que $h'|_\gamma = g$ y, por lo tanto, que $h = h'$, ya que también es $h'(\gamma) = F(h'|_\gamma) = F(g) = h(\gamma)$. Todo esto nos dice que si β es un ordinal sucesor entonces hay exactamente una función buena $\beta \to X$.
- Supongamos ahora que β es un ordinal límite. Como $\beta \subseteq B$, para cada ordinal γ menor que β hay exactamente una función buena $g_\gamma : \gamma \to X$, y podemos considerar la unión $g \coloneqq \bigcup_{\gamma \in \beta} g_\gamma$. Para cada $\gamma \in \beta$ es $g_\gamma \subseteq \gamma \times X \subseteq \beta \times X$, así que $g \subseteq \beta \times X$. Veamos que, de hecho, g es una función $\beta \to X$.
  - Sea γ un elemento de β, de manera que γ es un ordinal menor que β. Como $\gamma + 1 \neq \beta$, debe ser $\gamma + 1 \subseteq \beta$, así que $\gamma + 1 \in \beta$ y, por lo tanto, como $\gamma \in \gamma + 1$, el par ordenado $(\gamma, g_{\gamma+1}(\gamma))$ pertenece a $g_{\gamma+1}$ y, en consecuencia, a g.
  - Sean ahora γ un elemento de β y x y x' dos elementos de X tales que los pares ordenados (γ, x) y (γ, x') están los dos en g. Esto significa que hay dos elementos δ y δ' de β tales que $(\gamma, x) \in g_\delta$ y $(\gamma, x') \in g_{\delta'}$, y esto, a su vez, nos dice que $\gamma \in \delta$, $x = g_\delta(\gamma)$, $\gamma \in \delta'$ y $g_{\delta'}(\gamma)$. Como los ordinales están totalmente ordenados por inclusión, es $\delta \subseteq \delta'$ o $\delta' \subseteq \delta$. Si vale la primera de estas dos posibilidades, entonces $g_\delta = g_{\delta'}|_\delta$, ya que la restricción $g_{\delta'}|_\delta : \delta \to X$ es una función buena, δ pertenece al conjunto B, y, por lo tanto, $x = g_\delta(\gamma) = g_{\delta'}|_\delta(\gamma) = g_{\delta'}(\gamma) = x'$. Si vale la segunda podemos concluir de manera similar que $x = x'$.

  Mostremos ahora, para terminar, que esta función $g : \beta \to X$ es buena y que es la única función buena definida sobre β.

– Sea γ un ordinal menor que β. Como β es un ordinal límite, $\gamma + 1$ también es menor que β y $g_{\gamma+1} \subseteq g$: esto implica que para cada ordinal δ menor que $\gamma+1$ es $g(\delta) = g_{\gamma+1}(\delta)$, de manera que $g|_{\gamma+1} = g_{\gamma+1}$ y, por lo tanto, $g|_\gamma = g_{\gamma+1}|_\gamma$ y, como la función $g_{\gamma+1}$ es buena,

$$g(\gamma) = g_{\gamma+1}(\gamma) = F(g_{\gamma+1}|_\gamma) = F(g|_\gamma).$$

Vemos así que la función g es buena.

– Supongamos que hay una función buena $g' : \beta \to X$ distinta de g, de manera que el conjunto $\{\gamma \in \beta : g(\gamma) \neq g'(\gamma)\}$ no es vacío y que podemos, por lo tanto, considerar su menor elemento. Escribámoslo δ. Como δ es menor que β, pertenece al conjunto B y hay, por lo tanto, una única función buena $\delta \to X$: las restricciónes $g|_\delta$ y $g'|_\delta$ son buenas, así que esto nos permite concluir que $g|_\delta = g'|_\delta$ y, como g y g' son funciones buenas, que

$$g(\delta) = F(g|_\delta) = F(g'|_\delta) = g'(\delta).$$

Esto es absurdo, ya que contradice la forma en que elegimos al cardinal δ.

Esto prueba que también cuando el cardinal β es un ordinal límite hay exactamente una función buena $\beta \to X$, y completa la demostración de (4.15)

Queremos mostrar ahora que $B = \alpha$. Supongamos que esto no es así, de manera que el conjunto $\alpha \setminus B$ no es vacío y que podemos, por lo tanto, considerar su menor elemento $\beta := \min \alpha \setminus B$. Todo elemento de β es un elemento de α menor que β, así que es $\beta \subseteq B$. De acuerdo a (4.15), entonces, hay una exactamente una función buena $\beta \to X$: esto nos dice que β pertenece a B, lo que es absurdo. Podemos concluir entonces que $B = \alpha$ y, de acuerdo a (4.15), que hay exactamente una función buena $\alpha \to X$. Esto es precisamente lo que afirma la proposición. □

4.4. Cardinales

Definición 4.4.1. Un ordinal α es un ***cardinal*** si no es equipotente a ninguno de sus elementos.

Lema 4.4.2. *Todo ordinal finito es un cardinal.*

Demostración. Sea α un ordinal finito. Si β es un elemento de α, entonces β es un subconjunto propio de α y, como α es finito, β no es equipotente a α. □

Segunda parte

Capítulo 5

Espacios métricos

5.1. Métricas y espacios métricos

Las siguientes dos definiciones presenta los objetos que nos interesan en este capítulo.

Definición 5.1.1. Sea X un conjunto. Una ***métrica*** sobre X es una función $d : X \times X \to \mathbb{R}$ tal que cualesquiera sean los elementos x, y y z de X valen las siguientes tres condiciones:

($\mathbf{M}_1$) $d(x, y) = 0 \iff x = y$.
($\mathbf{M}_2$) $d(x, y) = d(y, x)$.
($\mathbf{M}_3$) $d(x, y) \leq d(x, z) + d(z, y)$.

También llamamos a las métricas ***distancias***. La desigualdad que aparece en la condición ($\mathbf{M}_3$) es la ***desigualdad triangular***. Normalmente fijamos un conjunto y una métrica sobre él. La siguiente definición hace conveniente hablar de esa situación.

Definición 5.1.2. Un ***espacio métrico*** es un par ordenado (X, d) que tiene como primera componente a un conjunto X y como segunda componente a una métrica $d : X \times X \to \mathbb{R}$ sobre X. Llamamos al conjunto X el ***conjunto subyacente*** al espacio métrico (X, d).

De hecho, casi siempre que esto no introduce ambigüedades cuando un tenemos un espacio métrico (X, d) decimos simplemente que «X es un espacio métrico», dejando la notación para la métrica implícita. La idea de espacio métrico, una abstracción de la idea usual de distancia de la geometría, fue originalmente propuesta por René Maurice Fréchet en [Fré1906], aunque el primero en usar el término *espacio métrico* fue Felix Hausdorff en [Hau1914].

La primera observación que hay que hacer sobre las métricas es que toman valores no negativos.

Frecuentemente esto se incorpora en la definición misma, pero no es estrictamente necesario.

Lema 5.1.3. *Sea X un conjunto y sea $d : X \times X \to \mathbb{R}$ una métrica sobre X. Si x e y son dos elementos de X, entonces $d(x, y) \geq 0$.*

Demostración. Si x e y son dos elementos de X, entonces

$$\begin{aligned} 0 &= d(x,x) && \text{por la condición } (\mathbf{M}_1) \\ &\leq d(x,y) + d(y,x) && \text{por la condición } (\mathbf{M}_3) \\ &\leq d(x,y) + d(x,y) && \text{por la condición } (\mathbf{M}_2) \\ &\leq 2 \cdot d(x,y), \end{aligned}$$

así que $d(x, y) \geq 0$. □

Todo subconjunto de un espacio métrico es, de manera natural, él mismo también un espacio métrico. El siguiente ejercicio explicita cómo.

Ejercicio 5.1.4. Sea X un conjunto, sea $d : X \times X \to \mathbb{R}$ una métrica sobre X, y sea Y un subconjunto de X. Muestre que la restricción $d|_{Y \times Y} : Y \times Y \to \mathbb{R}$ es una métrica sobre Y.

Desde ahora, siempre que tengamos un espacio métrico (X, d) y un subconjunto Y de X, consideraremos a Y como un espacio métrico con respecto a la métrica $d|_{Y \times Y}$ que de obtiene restringiendo la métrica de X a Y.

La noción de espacio métrico es considerablemente abstracta y por una buena razón: su objetivo es cubrir una gran variedad de situaciones muy distintas. Antes de iniciar el estudio de los espacios métricos, entonces, es conveniente presentar una buena cantidad de ejemplos de espacios métricos sobre los que podamos asentar nuestra intuición.

5.2. Ejemplos

Dedicaremos toda esta sección exclusivamente a presentar ejemplos de espacios métricos.

5.2.1. Métricas euclídeas

El primer ejemplo de todos es el que podemos ver como la motivación inicial de toda la teoría.

Proposición 5.2.1.1. *Sea n un entero no negativo. La función*

$$d : (x, y) \in \mathbb{R}^n \times \mathbb{R}^n \mapsto \left(\sum_{i=1}^{n}(x_i - y_i)^2\right)^{1/2} \in \mathbb{R}$$

es una métrica sobre $\mathbb{R}^n$.

Llamamos a esta métrica la ***métrica euclídea*** sobre $\mathbb{R}^n$. Notemos que si x e y son elementos de $\mathbb{R}^n$, entonces para cada $i \in [\![n]\!]$ es $(x_i - y_i)^2 \geq 0$, así que la suma $\sum_{i=1}^{n}(x_i - y_i)^2$ es un número real no negativo que, por lo tanto, tiene raíz cuadrada. Esto nos dice que la expresión que usamos en la definición de la función d de la proposición tiene sentido.

Esta proposición incluye el caso en que el número n es 1: es fácil ver que la métrica euclídea en $\mathbb{R}$ no es otra que la función $d : (x, y) \in \mathbb{R} \times \mathbb{R} \mapsto |x - y| \in \mathbb{R}$.

Demostración. Mostremos que se satisfacen las tres condiciones de la Definición 5.1.1. Sean x, y y z tres elementos de $\mathbb{R}^n$.

- Si $0 = d(x, y) = (\sum_{i=1}^{n}(x_i - y_i)^2)^{1/2}$, entonces $\sum_{i=1}^{n}(x_i - y_i)^2 = 0$ y, como en esta suma todos los términos son no negativos, es $x_i - y_i = 0$ para cada $i \in [\![n]\!]$, de manera que $x = y$. Por otro lado, es claro que si $x = y$ entonces $d(x, y) = (\sum_{i=1}^{n}(x_i - y_i)^2)^{1/2} = 0$.
- Para cada $i \in [\![n]\!]$ es $(x_i - y_i)^2 = (y_i - x_i)^2$, así que

$$d(x, y) = \left(\sum_{i=1}^{n}(x_i - y_i)^2\right)^{1/2} = \left(\sum_{i=1}^{n}(y_i - x_i)^2\right)^{1/2} = d(y, x).$$

- Notemos que si a y b son elementos de $\mathbb{R}^n$, entonces

$$\begin{aligned} 0 &\leq \sum_{i=1}^{n}\sum_{j=1}^{n}(a_i b_j - b_i a_j)^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}(a_i^2 b_j^2 - 2a_i b_i a_j b_j + b_i^2 a_j^2) \\ &= \sum_{i=1}^{n} a_i^2 \cdot \sum_{j=1}^{n} b_j^2 - 2\sum_{i=1}^{n} a_i b_i \cdot \sum_{j=1}^{n} a_j b_j + \sum_{i=1}^{n} b_i^2 \cdot \sum_{j=1}^{n} a_j^2 \\ &= 2\sum_{i=1}^{n} a_i^2 \cdot \sum_{i=1}^{n} b_i^2 - 2\left(\sum_{i=1}^{n} a_i b_i\right)^2, \end{aligned}$$

así que

$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 \leq \sum_{i=1}^{n} a_i^2 \cdot \sum_{i=1}^{n} b_i^2$$

y, por lo tanto,

$$\sum_{i=1}^{n} a_i b_i \leq \left(\sum_{i=1}^{n} a_i^2\right)^{1/2} \cdot \left(\sum_{i=1}^{n} b_i^2\right)^{1/2}. \tag{5.1}$$

Ahora bien, es

$$\begin{aligned} d(x,y)^2 &= \sum_{i=1}^{n}(x_i - y_i)^2 = \sum_{i=1}^{n}(x_i - z_i + z_i - y_i)^2 \\ &= \sum_{i=1}^{n}\big((x_i - z_i)^2 + (z_i - y_i)^2 + 2(x_i - z_i)(z_i - x_i)\big) \\ &= \sum_{i=1}^{n}(x_i - z_i)^2 + \sum_{i=1}^{n}(z_i - y_i)^2 + 2\sum_{i=1}^{n}(x_i - z_i)(z_i - x_i) \\ &= d(x,z)^2 + d(z,y)^2 + 2\sum_{i=1}^{n}(x_i - z_i)(z_i - x_i) \end{aligned}$$

y tomando $a = x - z$ y $b = z - x$ en (5.1) vemos que esto es

$$\begin{aligned} &\leq d(x,z)^2 + d(z,y)^2 + 2\left(\sum_{i=1}^{n}(x_i - z_i)^2\right)^{1/2}\left(\sum_{i=1}^{n}(z_i - y_i)^2\right)^{1/2} \\ &= d(x,z)^2 + d(z,y)^2 + 2 \cdot d(x,z) \cdot d(z,x) \\ &= (d(x,z) + d(z,x))^2. \end{aligned}$$

Como $d(x,y) \geq 0$ y $d(x,z) + d(z,x) \geq 0$, esto implica que $d(x,y) \leq d(x,z) + d(z,x)$, como queremos.

Se sigue de todo esto que la función d es una métrica sobre $\mathbb{R}^n$. □

La definición de la métrica euclídea hace evidente que siempre que x, y y z son tres elementos de $\mathbb{R}^n$ es $d(x+z, y+z) = d(x,y)$. Esta observación resulta ser verdaderamente importante. Le ponemos un nombre a esa propiedad.

Definición 5.2.1.2. Sea G un grupo. Una métrica $d : G \times G \to \mathbb{R}$ sobre G es

- ***invariante por translaciones a izquierda*** si $d(xy, xz) = d(y,z)$ cualesquiera sean x, y, y z en G,
- ***invariante por translaciones a derecha*** si $d(yx, zx) = d(y,z)$ cualesquiera sean x, y, y z

en G,

- y ***invariante por translaciones*** si es al la vez invariante por translaciones a izquierda y a derecha.

Cuando el grupo G es abeliano estas tres definiciones coinciden, por supuesto. Si vemos al espacio vectorial $\mathbb{R}^n$ como un grupo abeliano con respecto a su operación usual de suma, entonces nuestra observación de arriba es que su métrica euclídea es invariante por translaciones.

5.2.2. Espacios de pre-Hilbert

La construcción de la métrica euclídea admite una generalización importante a los espacios vectoriales con producto interno. Recordemos qué es esto.

Definición 5.2.2.1. Sea $\mathbb{k}$ o el cuerpo $\mathbb{R}$ o el cuerpo $\mathbb{C}$, y sea V un espacio vectorial sobre $\mathbb{k}$. Una función $\langle -, - \rangle : V \times V \to \mathbb{k}$ es un ***producto interno*** sobre V si cualesquiera sean los vectores x, y y z de V y los escalares a y b de $\mathbb{k}$ valen las siguientes tres condiciones.

($\mathbf{PI}_1$) $\langle x, y \rangle = \overline{\langle y, x \rangle}$.

($\mathbf{PI}_2$) $\langle x, x \rangle > 0$ si $x \neq 0$.

($\mathbf{PI}_3$) $\langle ax + by, z \rangle = a\langle x, z \rangle + b\langle y, z \rangle$.

La condición ($\mathbf{PI}_1$) implica que para cada $x \in V$ es $\langle x, x \rangle \in \mathbb{R}$ y, por lo tanto, que la desigualdad de la condición ($\mathbf{PI}_2$) tiene sentido. Por otro lado, si x es un vector, entonces o bien $x \neq 0$ y $\langle x, x \rangle > 0$ por ($\mathbf{PI}_2$) o bien $x = 0$ y

$$\langle x, x \rangle = \langle 0, 0 \rangle = \langle 0 \cdot 0 + 0 \cdot 0, 0 \rangle = 0 \cdot \langle 0, 0 \rangle + 0 \cdot \langle 0, 0 \rangle = 0,$$

así que en cualquier caso es $\langle x, x \rangle \geq 0$. Finalmente, la combinación de las condiciones ($\mathbf{PI}_1$) y ($\mathbf{PI}_3$) implica que cualesquiera sean los vectores x, y y z de V y los escalares a y b de $\mathbb{k}$ vale que

$$\langle x, ay + bz \rangle = a\langle x, y \rangle + b\langle x, z \rangle.$$

Mostraremos que todo producto interno en un espacio vectorial real o complejo determina una métrica sobre él, y para ello necesitamos de la desigualdad de la siguiente proposición, conocida como la ***desigualdad de Cauchy–Bunyakovsky–Schwartz***, por Augustin-Louis Cauchy, Viktor Bunyakovsky, y Hermann Schwarz.

Proposición 5.2.2.2. *Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, y sea V un espacio vectorial sobre $\mathbb{k}$ dotado de un producto interno $\langle -, - \rangle : V \times V \to \mathbb{k}$. Si x e y son dos elementos de V, entonces $|\langle x, y \rangle| \leq \langle x, x \rangle^{1/2} \cdot \langle y, y \rangle^{1/2}$.*

Notemos que $\langle x, x \rangle$ e $\langle y, y \rangle$ son números reales no negativos, así que las raíces cuadradas que aparecen aquí tienen sentido.

Demostración. Sean x e y dos elementos de V. Que la desigualdad del lema vale es claro si $y = 0$, ya que en ese caso $\langle x, y\rangle = 0$ e $\langle y, y\rangle = 0$, así que podemos suponer que $y \neq 0$. Para cada elemento t de $\mathbb{R}$ es

$$\begin{aligned} 0 \leq \langle x + ty, x + ty\rangle &= \langle x, x\rangle + t \cdot (\overline{\langle y, x\rangle} + \overline{\langle x, y\rangle}) + t^2\langle y, y\rangle \\ &= \langle x, x\rangle + t \cdot 2\,\mathrm{Re}\langle x, y\rangle + t^2\langle y, y\rangle \end{aligned}$$

y esto nos dice que el polinomio $p(X) \coloneqq \langle x, x\rangle + t \cdot 2\,\mathrm{Re}\langle x, y\rangle + t^2\langle y, y\rangle$, que tiene coeficientes reales y grado 2, toma solamente valores no negativos. El discriminante $4(\mathrm{Re}\langle x, y\rangle)^2 - 4\langle y, y\rangle \cdot \langle y, y\rangle$ de p es entonces no positivo, de manera que $(\mathrm{Re}\langle x, y\rangle)^2 \leq \langle x, x\rangle \cdot \langle y, y\rangle$ y, en consecuencia, $|\mathrm{Re}\langle x, y\rangle| \leq \langle x, x\rangle^{1/2} \cdot \langle y, y\rangle^{1/2}$. Esto completa la prueba del lema. □

Usando el lema es fácil probar lo que queremos.

Proposición 5.2.2.3. *Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, y sea V un espacio vectorial sobre $\Bbbk$ dotado de un producto interno $\langle -, -\rangle : V \times V \to \Bbbk$. La función*

$$d : (v, w) \in V \times V \mapsto \langle v - w, v - w\rangle^{1/2} \in \mathbb{R}$$

es una métrica sobre V que es invariante por translaciones.

Llamamos a la métrica d sobre V descripta en esta proposición la ***métrica asociada al producto interno*** $\langle -, -\rangle$ de V y decimos que el espacio métrico (V, d) es un ***espacio de pre-Hilbert***. Notemos que la expresión $\langle v - w, v - w\rangle^{1/2}$ que aparece en la definición de la función d tiene sentido, ya que la base de esa potencia es no negativa.

Demostración. Ocupémonos de a una de las condiciones de la Definición 5.1.1. Sean x, y, y z tres vectores de V.

- Es $d(x, x) = \langle x - x, x - x\rangle^{1/2} = \langle 0, 0\rangle^{1/2} = 0$. Por otro lado, si x e y son dos vectores de V tales que $d(x, y) = 0$, entonces tenemos que $\langle x - y, x - y\rangle^{1/2} = 0$, de manera que $\langle x - y, x - y\rangle = 0$ y, en vista de **(PI$_2$)**, es $x - y = 0$, esto es, $x = y$.
- Usando **(PI$_1$)** y **(PI$_3$)** vemos que

  $$\langle x - y, x - y\rangle = \langle -(y - x), -(y - x)\rangle = (-1)^2\langle y - x, y - x\rangle = \langle y - x, y - x\rangle$$

  así que $d(x, y) = \langle x - y, x - y\rangle^{1/2} = \langle y - x, y - x\rangle^{1/2} = d(y, x)$.
- Tenemos que

  $$\begin{aligned} d(x, y)^2 = \langle x - y, x - y\rangle &= \langle (x - z) + (z - y), (x - z) + (z - y)\rangle \\ &= \langle x - z, x - z\rangle + \langle z - y, x - z\rangle + \langle x - z, z - y\rangle + \langle z - y, z - y\rangle \\ &= \langle x - z, x - z\rangle + 2\,\mathrm{Re}\langle x - z, z - y\rangle + \langle z - y, z - y\rangle \end{aligned}$$

$$\leq \langle x-z, x-z\rangle + 2|\mathrm{Re}\langle x-z, z-y\rangle| + \langle z-y, z-y\rangle$$

y la desigualdad del Lema 5.2.2.2 nos dice que esto es

$$\begin{aligned}
&\leq \langle x-z, x-z\rangle + 2\langle x-z, x-z\rangle^{1/2} \cdot \langle z-y, z-y\rangle^{1/2} + \langle z-y, z-y\rangle \\
&= \left(\langle x-z, x-z\rangle^{1/2} + \langle z-y, z-y\rangle^{1/2}\right)^2 \\
&= (d(x,z) + d(z,y))^2,
\end{aligned}$$

así que $d(x,y) \leq d(x,z) + d(z,y)$, ya que $d(x,y) \geq 0$ y $d(x,z) + d(z,y) \geq 0$.

Todo esto nos permite concluir que la función d es una métrica sobre el conjunto V, y que es invariante por translaciones es una consecuencia inmediata de su definición. □

Usando este resultado podemos reconstruir la distancia euclídea sobre $\mathbb{R}^n$, y es por eso que la Proposición 5.2.2.3 es una generalización de la Proposición 5.2.1.1.

Ejercicio 5.2.2.4. Sea n un entero no negativo. Pruebe que la función

$$\langle -, -\rangle : (x,y) \in \mathbb{R}^n \times \mathbb{R}^n \mapsto \sum_{i=1}^{n} x_i y_i \in \mathbb{R}$$

es un producto interno sobre el espacio vectorial real $\mathbb{R}^n$, de manera que, de acuerdo a la proposición, la función

$$d : (x,y) \in \mathbb{R}^n \mapsto \mathbb{R}^n \mapsto \left(\sum_{i=1}^{n} (x_i - y_i)^2\right)^{1/2} = \langle x-y, x-y\rangle^{1/2} \in \mathbb{R}$$

es una métrica sobre $\mathbb{R}^n$. Esta es, por supuesto, la métrica euclídea sobre $\mathbb{R}^n$. Llamamos al producto interno sobre $\mathbb{R}^n$ descripto aquí el ***producto interno estándar*** de $\mathbb{R}^n$.

Hay, por supuesto, muchos productos internos sobre $\mathbb{R}^n$ además del estándar, así que podemos construir muchas otras métricas sobre ese espacio vectorial. Nuestros siguientes dos ejercicios describen todos los productos internos que hay sobre $\mathbb{R}^2$ y, más generalmente, sobre $\mathbb{R}^n$.

Ejercicio 5.2.2.5. Muestre que si a, b y c son números reales tales que $a > 0$, $c > 0$ y $ac > b^2$, entonces la función

$$\langle -, -\rangle : (x,y) \in \mathbb{R}^n \times \mathbb{R}^n \mapsto ax_1y_1 + bx_1y_2 + bx_2y_1 + cx_2y_2 \in \mathbb{R}$$

es un producto interno sobre el espacio vectorial real $\mathbb{R}^2$ y que, por lo tanto, hay una función

$$d : (x,y) \in \mathbb{R}^2 \mapsto \mathbb{R}^2 \mapsto (a(x_1-y_1)^2(x_1-y_1) + 2b(x_1-y_1)(x_2-y_2) + c(x_2-y_2)^2)^{1/2} \in \mathbb{R}$$

que es una métrica sobre $\mathbb{R}^2$. Pruebe además que, de hecho, obtenemos de esta forma *todos* los productos internos sobre $\mathbb{R}^2$ y, por lo tanto, todas las métricas sobre $\mathbb{R}^2$ que provienen de productos internos.

Ejercicio 5.2.2.6. Sea n un entero positivo y sea A una matriz de $\mathrm{M}_n(\mathbb{R})$ que es ***definida positiva***, esto es, tal que

para todo vector no nulo x de $\mathbb{R}^n$ es $x^t Ax > 0$.

No es difícil, en principio, decidir si una matriz de $\mathrm{M}_n(\mathbb{R})$ es definida positiva: el *criterio de Sylvester*, por James Joseph Sylvester, nos dice que una matriz de $\mathrm{M}_n(\mathbb{R})$ es definida positiva si y solamente si los n menores principales iniciales de la matriz tienen determinante positivo.

Pruebe las siguientes afirmaciones.

(a) La función $\langle -,- \rangle_A : (x,y) \in \mathbb{R}^n \times \mathbb{R}^n \mapsto x^t Ay \in \mathbb{R}$ es un producto interno sobre el espacio vectorial real $\mathbb{R}^n$, de manera que la función $d_A : (x,y) \in \mathbb{R}^n \times \mathbb{R}^n \mapsto \langle x-y, x-y \rangle_A^{1/2} \in \mathbb{R}$ es una métrica sobre $\mathbb{R}^n$.

(b) Si $\langle -,- \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ es un producto interno sobre el espacio vectorial real, entonces hay una matriz A en $\mathrm{M}_n(\mathbb{R})$ que es definida positiva y tal que $\langle -,- \rangle = \langle -,- \rangle_A$.

(c) Una matriz $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ de $\mathrm{M}_2(\mathbb{R})$ es definida positiva si y solamente si $a > 0$, $d > 0$ y $ac > b^2$.

Ejercicio 5.2.2.7. Sea n un entero no negativo.

(a) Muestre que la función

$$\langle -,- \rangle : (x,y) \in \mathbb{C}^n \times \mathbb{C}^n \mapsto \sum_{i=1}^{n} x_i \overline{y_i} \in \mathbb{C}$$

es un producto interno sobre el espacio vectorial complejo $\mathbb{C}^n$, al que llamamos el ***producto interno estándar*** de $\mathbb{C}^n$, y concluya que la función

$$d : (x,y) \in \mathbb{C}^n \times \mathbb{C}^n \mapsto \left(\sum_{i=1}^{n} |x_i - y_i|^2 \right)^{1/2} \in \mathbb{R}$$

es una métrica sobre $\mathbb{C}^n$.

(b) Determine todos los productos internos sobre el espacio vectorial complejo $\mathbb{C}^2$ y, con ello, todas las métricas sobre $\mathbb{C}^2$ que están determinadas por un producto interno.

Ejercicio 5.2.2.8. Sea n un entero positivo y consideremos ahora el conjunto $D_{\mathbb{C}}(\mathbb{C}^n)$ de todas las métricas sobre $\mathbb{C}^n$ determinadas por productos internos sobre el espacio vectorial complejo $\mathbb{C}^n$. Como $\mathbb{R}$ es un subcuerpo de $\mathbb{C}$, podemos ver al espacio vectorial complejo $\mathbb{C}^n$ como un espacio

vectorial real y, por lo tanto, tiene sentido también considerar el conjunto $D_{\mathbb{R}}(\mathbb{C}^n)$ de todas las métricas sobre $\mathbb{C}^n$ determinadas por productos internos sobre el espacio vectorial *real* $\mathbb{C}^n$. ¿Qué relación hay entre los conjuntos $D_{\mathbb{C}}(\mathbb{C}^n)$ y $D_{\mathbb{R}}(\mathbb{C}^n)$?

Podemos considerar productos internos sobre espacios distintos de $\mathbb{R}^n$ y $\mathbb{C}^n$, claro. Veamos un ejemplo de ello.

Ejemplo 5.2.2.9. Consideremos el espacio vectorial real $C[0,1]$ de todas las funciones continuas $[0,1] \to \mathbb{R}$. Si f y g son dos elementos de V, entonces también lo es el producto fg y, por lo tanto, podemos considerar la integral $\int_0^1 fg$. Esto nos dice que hay una función

$$\langle -,- \rangle : (f,g) \in C[0,1] \times C[0,1] \mapsto \int_0^1 fg \in \mathbb{R}.$$

Es inmediato que esta función satisface las condiciones **(PI$_1$)** y **(PI$_3$)** de la Definición 5.2.2.1. Veamos que también satisface la condición **(PI$_2$)**. Sea f un elemento no nulo de $C[0,1]$. La función $f^2 : [0,1] \to \mathbb{R}$ es no negativa, es no idénticamente nula, y es continua. Hay entonces un elemento t_0 de $[0,1]$ tal que $f(t_0) > 0$, y la continuidad de f implica que hay también un número positivo ε tal que para todo elemento t de $[t_0 - \varepsilon, t_0 + \varepsilon] \cap [0,1]$ vale que $|f(t) - f(t_0)| \leq \frac{1}{2} f(t)$, así que $\frac{1}{2} f(t_0) = f(t_0) - \frac{1}{2} f(t_0) \leq f(t)$. La intersección $[t_0 - \varepsilon, t_0 + \varepsilon] \cap [0,1]$ es un intervalo cerrado al que podemos escribir en la forma $[c,d]$ con $c < d$, y es

$$\begin{aligned}
\langle f, f \rangle &= \int_0^1 f^2 \\
&\geq \int_c^d f^2 && \text{porque } f^2 \geq 0 \text{ y } [c,d] \subseteq [0,1] \\
&\geq \int_c^d \tfrac{1}{4} f(t_0)^2 && \text{porque } f^2(t) \geq \frac{1}{4} f(t_0)^2 \text{ para cada } t \in [c,d] \\
&= \frac{1}{4} f(t_0)^2 (d-c) > 0 && \text{porque } c < d.
\end{aligned}$$

La Proposición 5.2.2.3 nos permite concluir entonces que la función

$$d : (f,g) \in C[0,1] \times C[0,1] \mapsto \left(\int_0^1 fg \right)^{1/2} \in \mathbb{R}$$

es una métrica sobre $C[0,1]$.

5.2.3. Espacios normados

La construcción del apartado anterior de una métrica a partir de un producto interno tiene una generalización importante. Para presentarla necesitamos la siguiente definición.

Definición 5.2.3.1. Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, y sea V un espacio vectorial sobre $\Bbbk$. Una función

$$\|-\| : V \to \mathbb{R}$$

es una ***norma*** sobre V si cualesquiera sean los vectores x e y de V y el escalar t de $\Bbbk$ satisface las siguientes condiciones.

($\mathbf{N}_1$) $\|x\| = 0$ solo si $x = 0$.
($\mathbf{N}_2$) $\|t \cdot x\| = |t| \cdot \|x\|$.
($\mathbf{N}_3$) $\|x + y\| \leq \|x\| + \|y\|$.

Cuando ese es el caso, decimos que el par $(V, \|-\|)$ o, si el contexto determina de qué norma estamos hablando, que X es un ***espacio normado***.

En la situación de esta definición, tenemos que $\|0\| = \|0 \cdot 0\| = |0| \cdot \|0\| = 0$ y que para cada elemento x de X es

$$0 = \|0\| = \|x + (-1)x\| \leq \|x\| + \|(-1)x\| = \|x\| + |-1| \cdot \|x\| = 2 \cdot \|x\|,$$

de manera que $\|x\| \geq 0$. Usaremos estos dos hechos constantemente.

Todo subespacio de un espacio normado es él mismo de manera natural un espacio normado.

Ejercicio 5.2.3.2. Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, sea V un espacio vectorial sobre $\Bbbk$ y sea U un $\Bbbk$-subespacio de V. Muestre que si $\|-\| : V \to \mathbb{R}$ es una norma sobre V, entonces la restricción $\|-\|\big|_U : U \to \mathbb{R}$ es una norma sobre U.

Siempre que estemos en la situación de este ejercicio consideraremos al subespacio como un espacio normado provisto de la norma que se obtiene restringiendo a él la norma de V.

Nos interesa aquí la noción de norma porque vale el siguiente resultado.

Proposición 5.2.3.3. *Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, sea V un espacio vectorial sobre $\Bbbk$, y sea $\|-\| : V \to \mathbb{R}$ una norma sobre V. La función*

$$d : (x, y) \in V \times V \mapsto \|x - y\| \in \mathbb{R}$$

es una métrica sobre V que es invariante por translaciones.

Llamamos a la métrica d sobre V descripta aquí la ***métrica sobre V asociada a la norma $\|-\|$***. Siempre que tenemos un espacio normado $(V, \|-\|)$ vemos a V como un espacio métrico con respecto a la distancia asociada a la norma $\|-\|$.

Demostración. Sean x, y y z tres vectores de V.

- La condiciones $(\mathbf{N}_1)$ y $(\mathbf{N}_2)$ implican que $d(x, y) = \|x - y\| = 0$ si y solamente si $x - y = 0$, esto es, si y solamente si $x = y$, y esto implica que la función d satisface la condición $(\mathbf{M}_1)$.
- De acuerdo a la condición $(\mathbf{N}_2)$ tenemos que

$$d(x, y) = \|x - y\| = \|(-1)(y - x)\| = |-1| \cdot \|y - x\| = d(y, x),$$

  así que la función d satisface la condición $(\mathbf{M}_2)$.
- Finalmente, la función d satisface la condición $(\mathbf{M}_3)$ porque

$$d(x, y) = \|x - y\| = \|(x - z) + (z - y)\| \leq \|x - z\| + \|z - y\| = d(x, z) + d(z, y)$$

  porque la norma satisface la condición $(\mathbf{N}_3)$.

Todo esto implica, de acuerdo a la Definición 5.1.1, que la función d es una métrica sobre V. Que es además invariante por translaciones es una consecuencia inmediata de su definición. □

Notemos que si V es un espacio normado y U es un subespacio de V, entonces la métrica determinada por la norma que se obtiene restringiendo la norma de V a U coincide con la métrica que se obtiene restringiendo a U la métrica sobre V determinada por su norma.

La afirmación del siguiente ejercicio nos dice que la Proposición 5.2.3.3 es una generalización de la Proposición 5.2.2.3.

Ejercicio 5.2.3.4. Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, sea V un espacio vectorial sobre $\Bbbk$, y sea $\langle -, - \rangle : V \times V \to \Bbbk$ un producto interno sobre V. Muestre que la función

$$\|-\| : x \in V \mapsto \langle x, x \rangle^{1/2} \in \mathbb{R}$$

es una norma sobre V, y que la métrica sobre V asociada a esta norma coincide con la métrica sobre V asociada al producto interno $\langle -, - \rangle$. Llamamos a la norma $\|-\|$ la ***norma sobre V asociada al producto interno $\langle -, - \rangle$***.

El interés de la Proposición 5.2.3.3 es que hay normas sobre espacios vectoriales reales o complejos que no son las normas asociadas a ningún producto interno sobre esos espacios. Veremos esto más adelante. Antes de eso, sin embargo, presentaremos algunos ejemplos importantes de espacios normados.

En cada uno de los ejemplos de espacios normados que consideraremos tenemos que verificar

que las condiciones de la Definición 5.2.3.1 se cumplen, y de las tres la que presenta siempre mayores dificultades es la tercera. El siguiente resultado nos da una forma de verificar que esa condición se satisface.

Proposición 5.2.3.5. *Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, y sea V un $\mathbb{k}$-espacio vectorial. Una función $\|-\| : V \to \mathbb{R}$ que que toma valores no negativos y que satisface las condiciones $(\mathbf{N}_1)$ y $(\mathbf{N}_2)$ de la Definición 5.2.3.1 es una norma si y solamente si siempre que x e y son vectores de V tales que $\|x\| = 1$ y $\|y\| = 1$ se tiene que $\|tx + (1-t)y\| \leq 1$ para todo $t \in [0,1]$.*

Demostración. Sea primero $\|-\| : V \to \mathbb{R}$ una norma sobre V. Si x e y son dos vectores de V tales que $\|x\| = 1$ y $\|y\| = 1$, y t es un elemento de $[0,1]$, entonces tenemos que

$$\|tx + (1-t)y\| \leq |t| \cdot \|x\| + |1-t| \cdot \|y\| \leq t + (1-t) = 1,$$

ya que t y $1-t$ pertenecen a $[0,1]$. La condición de la proposición es, por lo tanto, necesaria.

Para probar que es suficiente, supongamos que se cumple y sean x e y dos vectores de V. Tenemos que mostrar que $\|x+y\| \leq \|x\| + \|y\|$. Si alguno de x o y es el vector nulo de V, entonces esta desigualdad es evidente, así que podemos suponer que $x \neq 0$ e $y \neq 0$. Como la condición $(\mathbf{N}_1)$ vale, sabemos entonces que $\|x\| > 0$ y $\|y\| > 0$ y podemos considerar los vectores $x' \coloneqq x/\|x\|$ e $y' \coloneqq y/\|y\|$. Que la condición $(\mathbf{N}_2)$ valga implica que $\|x'\| = \|x/\|x\|\| = \|x\|/\|x\| = 1$ y, de manera similar, que $\|y'\| = 1$, así que la hipótesis implica que $\|tx' + (1-t)y'\| \leq 1$ para todo $t \in [0,1]$. En particular, si elegimos $t = \|x\|/(\|x\| + \|y\|)$, que ciertamente es un elemento de $[0,1]$, entonces

$$\begin{aligned} tx' + (1-t)y' &= \frac{\|x\|}{\|x\| + \|y\|} \cdot \frac{x}{\|x\|} + \left(1 - \frac{\|x\|}{\|x\| + \|y\|}\right) \cdot \frac{x}{\|x\|} \\ &= \frac{1}{\|x\| + \|y\|} \cdot x + \frac{1}{\|x\| + \|y\|} \cdot y = \frac{x+y}{\|x\| + \|y\|}, \end{aligned}$$

de manera que

$$\frac{\|x+y\|}{\|x\| + \|y\|} = \left\| \frac{x+y}{\|x\| + \|y\|} \right\| = \|tx' + (1-t)y'\| \leq 1,$$

y, por supuesto, $\|x+y\| \leq \|x\| + \|y\|$, como queremos. □

Hagamos un par de definiciones generales.

Definición 5.2.3.6. Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, y sea V un $\mathbb{k}$-espacio vectorial. Si x e y son dos vectores de V, entonces llamamos al conjunto

$$[x,y] \coloneqq \{tx + (1-t)y : t \in [0,1]\}$$

el ***segmento*** de recta que une a x con y. Decimos que un subconjunto C de V es ***convexo*** si siempre

que x e y son dos elementos de C se tiene que $[x, y] \subseteq C$.

Usando estas definiciones, podemos dar una versión equivalente de la Proposición 5.2.3.5:

Corolario 5.2.3.7. *Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, sea V un espacio vectorial sobre $\Bbbk$. Una función $\|-\| : V \to \mathbb{R}$ que toma valores no negativos y satisface las condiciones ($\mathbf{N}_1$) y ($\mathbf{N}_2$) de la Definición 5.2.3.1 es una norma si y solamente si el conjunto $B \coloneqq \{x \in V : \|x\| \leq 1\}$ es convexo.* □

Esta interpretación geométrica de la condición ($\mathbf{N}_3$) resulta ser particularmente importante.

Ejercicio 5.2.3.8. Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, sea V un espacio vectorial sobre $\Bbbk$, y sea C un subconjunto convexo de V. Muestre que si n es un entero positivo, $x_1, x_2, \ldots, x_n$ son elementos de C, y $t_1, t_2, \ldots, t_n$ son números reales no negativos tales que $t_1 + t_2 + \cdots + t_n = 1$, entonces $t_1x_1 + t_2x_2 + \cdots + t_nx_n \in C$.

En los apartados siguientes daremos ejemplos importantes de espacios normados.

La norma $\|-\|_\infty$ sobre $\mathbb{R}^n$ y $\mathbb{C}^n$

Empecemos con el que es, probablemente, el ejemplo más sencillo de un espacio normado.

Proposición 5.2.3.9. *Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$ y sea n un entero no negativo. La función*

$$\|-\|_\infty : x \in \Bbbk^n \mapsto \max_{i \in [\![n]\!]} |x_i| \in \mathbb{R}$$

es una norma sobre el $\Bbbk$-espacio vectorial $\Bbbk^n$, de manera que $(\Bbbk^n, \|-\|_\infty)$ es un espacio normado.

Demostración. Veamos que la función $\|-\|_\infty$ satisface las tres condiciones de la Definición 5.2.3.1.

- Si x es un elemento de $\Bbbk^n$ tal que $\|x\|_\infty = 0$, entonces para cada $j \in [\![n]\!]$ tenemos que $|x_j| \leq \max_{i \in [\![n]\!]} |x| = \|x\|_\infty = 0$: esto nos dice que todas las componentes del vector x son nulas, así que $x = 0$.
- Sean x un elemento de $\Bbbk^n$ y t uno del cuerpo $\Bbbk$, y sea $y \coloneqq tx$. Si $j \in [\![n]\!]$, entonces

  $$|y_j| = |tx_j| = |t| \cdot |x_j| \leq |t| \cdot \max_{i \in [\![n]\!]} |x_i| = |t| \cdot \|x\|_\infty$$

  y, por lo tanto, $\|y\|_\infty = \max_{j \in [\![n]\!]} |y_j| \leq |t| \cdot \|x\|_\infty$. Por otro lado, sabemos que hay un elemento k de $[\![n]\!]$ tal que $\|x\|_\infty = \max_{i \in [\![n]\!]} |x_i| = |x_k|$, y entonces tenemos que

  $$|t| \cdot \|x\|_\infty = |t| \cdot |x_k| \leq |tx_k| = |y_k| \leq \max i \in [\![n]\!] |y_k| = \|y\|_\infty.$$

  Vemos así que $\|tx\|_\infty = \|y\|_\infty = |t| \cdot \|x\|_\infty$.

- Sean finalmente x e y dos elementos de $\mathbb{k}^n$, y escribamos $z \coloneqq x + y$. Para cada elemento j de $[\![n]\!]$ tenemos que

$$|z_j| = |x_j + y_j| \leq |x_j| + |y_j| \leq \max_{i \in [\![n]\!]} |x_i| + \max_{i \in [\![n]\!]} |y_i| = \|x\|_\infty + \|y\|_\infty,$$

así que

$$\|x + y\|_\infty = \|z\|_\infty = \max_{j \in [\![n]\!]} |z_j| \leq \|x\|_\infty + \|y\|_\infty. \qquad \square$$

En la construcción de la Proposición 5.2.3.9 la base ordenada estándar del espacio vectorial $\mathbb{k}^n$ juega un rol central, pero podemos hacer una construcción similar empezando con cualquier otra base ordenada de ese espacio.

Ejercicio 5.2.3.10. Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, sea V un espacio vectorial sobre $\mathbb{k}$ de dimensión finita, sea n la dimensión de V, y sea $\mathscr{B} = (e_1, e_2, \ldots, e_n)$ una base ordenada de V. Pruebe que hay exactamente una función $\|-\| : V \to \mathbb{R}$ tal que para cada elección de n escalares $a_1, a_2, \ldots, a_n$ de $\mathbb{k}$ vale

$$\|a_1 e_1 + a_2 e_2 + \cdots + a_n e_n\| = \max\{|a_1|, |a_2|, \ldots, |a_n|\},$$

y que es una norma sobre V.

Las normas $\|-\|_p$ sobre $\mathbb{R}^n$ y $\mathbb{C}^n$

Construiremos en la Proposición 5.2.3.12 de más abajo una familia de normas sobre $\mathbb{R}^n$ y sobre $\mathbb{C}^n$ y para verificar que ellas satisfacen la condición $(\mathbf{N}_3)$ de la Definición 5.2.3.1 usaremos el criterio que nos da la Proposición 5.2.3.5, apoyándonos en el siguiente resultado bien conocido.

Lema 5.2.3.11. *Sea $f : (0, +\infty) \to \mathbb{R}$ una función dos veces diferenciable tal que $f''(t) \geq 0$ para todo $t \in (0, +\infty)$. Si u y v son dos elementos de $(0, +\infty)$ y t uno de $[0, 1]$, entonces*

$$f(tu + (1-t)v) \leq t f(u) + (1-t) f(v).$$

Demostración. Sean u y v dos elementos de $(0, +\infty)$, sea t un elemento de $(0, 1)$, y escribamos $w = tu + (1-t)v$. Si $u = v$, la desigualdad del lema es evidente. Por otro lado, para probar que vale cuando $u \neq v$ es suficiente considerar el caso en que $u < v$, ya que en caso contrario podemos intercambiar los roles de u y v y reemplazar a t por $1 - t$.

Supongamos entonces que $u < v$ y sea $w \coloneqq tu + (1-t)v$. Es $u < w < v$ y

$$t \cdot f(u) + (1-t) \cdot f(v) - f(w) = t \cdot (f(u) - f(w)) + (1-t) \cdot (f(v) - f(w)).$$

El *Teorema del valor medio* nos dice que hay números α y β tales que $u < \alpha < w < \beta < v$,

$$f(u) - f(w) = f'(\alpha) \cdot (u - w) = (1 - t) \cdot f'(\alpha) \cdot (u - v)$$

y

$$f(v) - f(w) = f'(\beta) \cdot (v - w) = -t \cdot f'(\beta) \cdot (u - v),$$

así que

$$\begin{aligned} t \cdot f(u) + (1 - t) \cdot f(v) - f(w) &= t(1 - t) \cdot f'(\alpha) \cdot (u - v) - t(1 - t) \cdot f'(\beta) \cdot (u - v) \\ &= t(1 - t) \cdot \big(f'(\alpha) - f'(\beta)\big) \cdot (u - v). \end{aligned}$$

Como $\alpha < \beta$, el *Teorema del valor medio* implica que hay un número γ tal que $\alpha < \gamma < \beta$ y

$$f'(\alpha) - f'(\beta) = f''(\gamma) \cdot (\alpha - \beta),$$

y entonces tenemos que

$$t \cdot f(u) + (1 - t) \cdot f(v) - f(w) = t(1 - t) \cdot f''(\gamma) \cdot (\alpha - \beta)(u - v).$$

Los productos $t(1 - t)$ y $(\alpha - \beta)(u - v)$ son positivos y es $f''(\gamma) \geq 0$, así que podemos concluir con todo esto que $t \cdot f(u) + (1 - t) \cdot f(v) - f(w) \geq 0$, esto es, que

$$f(tu + (1 - t)v) \leq t \cdot f(u) + (1 - t) \cdot f(v).$$

Esto prueba el lema. □

Con esto a nuestra disposición, podemos dar los ejemplos que queremos.

Proposición 5.2.3.12. *Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, sea n un entero no negativo, y sea p un elemento de $[1, +\infty)$. La función*

$$\|-\|_p : x \in \Bbbk^n \mapsto \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \in [0, +\infty)$$

es una norma sobre el $\Bbbk$-espacio vectorial $\Bbbk^n$, de manera que, en particular, tenemos que

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p. \tag{5.2}$$

siempre que x e y son elementos de $\Bbbk^n$.

La desigualdad (5.2) es conocida como la ***desigualdad de Minkowski***, por Hermann Minkowski.

Demostración. Si x es un elemento de $\mathbb{k}^n$ tal que $\|x\|_p = 0$, entonces para cada $j \in [\![n]\!]$ tenemos que $|x_j|^p \leq \sum_{i=1}^n |x_i|^p$, de manera que

$$|x_j| = (|x_j|^p)^{1/p} \leq \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} = \|x\|_p = 0$$

y, por lo tanto, $x = 0$. La función $\|-\|_p$ satisface entonces la condición $(\mathbf{N}_1)$. Por otro lado, si x es un elemento de $\mathbb{k}^n$ y t uno de $\mathbb{k}$, y ponemos $y \coloneqq tx$, entonces tenemos que

$$\begin{aligned}\|tx\|_p = \|y\|_p &= \left(\sum_{i=1}^p |tx_i|^p\right)^{1/p} = \left(\sum_{i=1}^p |t|^p \cdot |x_i|^p\right)^{1/p} = \left(|t|^p \sum_{i=1}^p \cdot |x_i|^p\right)^{1/p} = |t| \cdot \left(\sum_{i=1}^p \cdot |x_i|^p\right)^{1/p} \\ &= |t| \cdot \|x\|_p.\end{aligned}$$

Vemos con esto que la función $\|-\|_p$ satisface también la condición $(\mathbf{N}_2)$ así que, de acuerdo a la Proposición 5.2.3.5, para ver que se trata de una norma sobre $\mathbb{k}^n$ es suficiente que mostremos que si x e y son dos elementos de $\mathbb{k}^n$ tales que $\|x\|_p = 1$ y $\|y\|_p = 1$, y t un elemento de $[0,1]$, entonces

$$\|tx + (1-t)y\|_p \leq 1.$$

Sean para ello x e y dos elementos de $\mathbb{k}^n$ tales que $\|x\|_p = 1$ y $\|y\|_p = 1$ y sea t un elemento de $[0,1]$. Sea j un elemento de $[\![n]\!]$. Si $x_j \neq 0$ y $y_j \neq 0$, de manera que $|x_j| > 0$ y $|y_j| > 0$, entonces el Lema 5.2.3.11 nos permite concluir que

$$(t \cdot |x_j| + (1-t) \cdot |y_j|)^p \leq t \cdot |x_j|^p + (1-t) \cdot |y_j|^p,$$

ya que la función $t \in (0, +\infty) \mapsto t^p \in (0, +\infty)$ es dos veces diferenciable y tiene derivada segunda no negativa. Como además $|tx_j + (1-t)y_j| \leq t \cdot |x_j| + (1-t) \cdot |y_j|$ y esa función tiene derivada primera no negativa, tenemos que

$$|tx_j + (1-t)y_j|^p \leq t \cdot |x_j|^p + (1-t) \cdot |y_j|^p.$$

Por otro lado, si $x_j = 0$ o $y_j = 0$, entonces esta desigualdad vale por razones triviales, así que vale en todos los casos. Gracias a ella podemos ver que

$$\begin{aligned}\sum_{j=1}^n |tx_j + (1-t)y_j|^p &\leq \sum_{j=1}^n \left(t \cdot |x_j|^p + (1-t) \cdot |y_j|^p\right) = t \sum_{j=1}^n |x_j|^p + (1-t) \sum_{j=1}^n |y_j|^p \\ &= t \cdot \|x\|_p^p + (1-t) \cdot \|y\|_p^p = 1\end{aligned}$$

y, por lo tanto, que

$$\|tx + (1-t)y\|_p = \left(\sum_{j=1}^n |tx_j + (1-t)y_j|^p\right)^{1/p} \leq 1^{1/p} = 1.$$

Como observamos arriba, esto completa la demostración de la proposición. □

Es habitual probar la desigualdad de Minkowski y, con ello, que la función $\|-\|_p$ definida en la Proposición 5.2.3.12 es una norma, de una forma diferente a la de arriba. El siguiente ejercicio da los detalles de esto.

Ejercicio 5.2.3.13. Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, sea p un elemento de $(1, +\infty)$, y sea q el único elemento de $(1, +\infty)$ tal que $1/p + 1/q = 1$.

(a) Muestre que si x e y son elementos de $\mathbb{k}^n$, entonces

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

y que vale aquí la igualdad si y solamente si hay números reales no negativos α y β tales que $\alpha |x_i|^p = \beta |y_i|^q$ para todo $i \in [\![n]\!]$.

Esta es la llamada ***desigualdad de Hölder***, por Otto Hölder.

(b) Deduzca la desigualdad de Minkowski de la desigualdad de Hölder.

Podemos construir normas $\|-\|_p$ sobre un espacio vectorial de dimensión cualesquiera.

Ejercicio 5.2.3.14. Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, sea p un elemento de $[1, +\infty)$, sea V un espacio vectorial sobre $\mathbb{k}$ de dimensión finita, sea n la dimensión de V, y sea $\mathscr{B} = (e_1, e_2, \ldots, e_n)$ una base ordenada de V. Pruebe que hay exactamente una función $\|-\| : V \to \mathbb{R}$ tal que para cada elección de n escalares $a_1, a_2, \ldots, a_n$ de $\mathbb{k}$ vale

$$\|a_1 e_1 + a_2 e_2 + \cdots + a_n e_n\| = \left(\sum_{i=1}^{n} |a_i|^p \right)^{1/p}$$

y que es una norma sobre V.

Hemos construido para cada elemento p de $[1, +\infty]$ una norma $\|-\|_p$ sobre $\mathbb{R}^n$ y sobre $\mathbb{C}^n$. Estas normas dependen continuamente del parámetro p, en el sentido preciso de la afirmación del siguiente ejercicio.

Ejercicio 5.2.3.15. Sea n un entero no negativo, sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, y fijemos un elemento x de $\mathbb{k}^n$. Muestre que la función

$$p \in [1, +\infty) \to \|x\|_p \in \mathbb{R}$$

es continua, creciente, y que además vale que

$$\lim_{p \to +\infty} \|x\|_p = \|p\|_\infty.$$

En la Proposición 5.2.3.12 solamente consideramos números p que pertenecen al intervalo

$[1, +\infty)$. Para los números de $(0,1)$ podemos hacer una construcción similar que no da una norma, sino solamente una métrica.

Ejercicio 5.2.3.16. Sea n un entero no negativo, sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, y sea p un elemento de $(0,1)$.

(a) Muestre que la función

$$d : (x, y) \in \mathbb{k}^n \times \mathbb{k}^n \mapsto \sum_{i=1}^{n} |x_i - y_i|^p \in \mathbb{R}$$

es una métrica sobre $\mathbb{k}^n$ que es invariante por translaciones.

(b) Pruebe que la función $(x, y) \in \mathbb{k}^n \mapsto d(x, y)^{1/p} \in \mathbb{R}$ no es una norma sobre $\mathbb{k}^n$.

La ley del paralelogramo y el teorema de Jordan y von Neumann

El resultado del Ejercicio 5.2.3.4 nos dice que todo producto interno en un espacio vectorial determina una norma sobre él. En general, sin embargo, hay normas que no provienen de esa manera de un producto interno. El siguiente resultado nos permite mostrar esto.

Ejercicio 5.2.3.17. Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, sea V un espacio vectorial sobre $\mathbb{k}$ dotado de un producto interno $\langle -, - \rangle$, y sea $\|-\|$ la correspondiente norma sobre V. Pruebe que cualesquiera sean los elementos x e y de V valen las siguientes afirmaciones:

(*i*) Es $\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$.

(*ii*) Si $\mathbb{k} = \mathbb{R}$, entonces $\langle x, y \rangle = \frac{1}{4}\|x + y\|^2 - \frac{1}{4}\|x - y\|^2$.

(*iii*) Si $\mathbb{k} = \mathbb{C}$, en cambio, entonces $\langle x, y \rangle = \frac{1}{4}\|x + y\|^2 - \frac{1}{4}\|x - y\|^2 + \frac{i}{4}\|x + iy\|^2 - \frac{i}{4}\|x - iy\|^2$.

La primera afirmación es conocida como la *Ley del paralelogramo* — la Figura 5.1 deja en claro el porqué.

La *Ley del paralelogramo* es una condición necesaria para que una norma provenga de un producto interno, y podemos usar esto fácilmente en muchos ejemplos.

Ejemplo 5.2.3.18. Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, sea n un entero mayor que 1, y consideremos los vectores $x \coloneqq (1, 0, \ldots, 0)$ e $y \coloneqq (0, 1, 0, \ldots, 0)$ de $\mathbb{R}^n$. Como

$$\|x + y\|_1^2 + \|x - y\|_1^2 = 2^2 + 2^2 = 8 \neq 4 = 2 \cdot 1 + 2 \cdot 1 = 2\|x\|_1^2 + 2\|y\|_1^2$$

y

$$\|x + y\|_\infty^2 + \|x - y\|_\infty^2 = 1^2 + 1^2 = 2 \neq 4 = 2 \cdot 1 + 2 \cdot 1 = 2\|x\|_\infty^2 + 2\|y\|_\infty^2,$$

ni la norma $\|-\|_1$ ni la norma $\|-\|_\infty$ satisface la *Ley del paralelogramo*, así que ninguna de las dos proviene de un producto interno sobre $\mathbb{k}^n$.

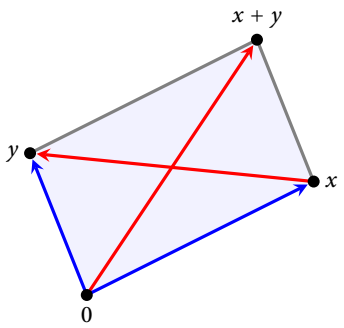


Figura 5.1. La igualdad de la ley del paralelogramo dice que la suma de los cuadrados de las longitudes de las diagonales de un paralelogramo es igual a la suma de los cuadrados de longitudes de los lados.

Ejercicio 5.2.3.19. Sea p un elemento de $(1, +\infty)$ y sea n un entero mayor que 1. Pruebe que la norma $\|-\|_p$ sobre $\mathbb{k}^n$ proviene de un producto interno sobre ese espacio vectorial si y solamente si $p = 2$.

Como dijimos, el Ejercicio 5.2.3.17 nos dice que una condición necesaria para que una norma esté determinada por un producto interno es que satisfaga la *Ley del paralelogramo*. El siguiente teorema, debido a Pascual Jordan y John von Neumann, nos dice que esa condición es, de hecho, también suficiente.

Proposición 5.2.3.20. *Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, y sea V un espacio vectorial normado con norma $\|-\|$. Hay un producto interno sobre V cuya norma asociada es $\|V\|$ si y solamente si esa norma satisface la* Ley del paralelogramo*, esto es, si para todo par de vectores x e y de V se tiene que*

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

Probaremos solamente la afirmación relativa a espacios vectoriales reales.

Demostración. En vista del Ejercicio 5.2.3.17, solamente tenemos que probar que la condición es suficiente. Supongamos que se cumple.

Primer paso. Sean x, x' e y tres vectores de V. La condición implica que

$$\|x + x' + y\|^2 = \|(x + y) + x'\|^2 = 2\|x + y\|^2 + 2\|x'\|^2 - \|x + y - x'\|^2$$

y que

$$\|x + x' - y\|^2 = \|x + (x' - y)\|^2 = 2\|x\|^2 + 2\|x' - y\|^2 - \|x - x' + y\|^2.$$

Restando miembro a miembro estas dos igualdades vemos que

$$\|x + x' + y\|^2 - \|x + x' - y\|^2 = 2\|x + y\|^2 + 2\|x'\|^2 - 2\|x\|^2 - 2\|x' - y\|^2,$$

e intercambiando aquí los roles de x y de x' concluimos que también

$$\|x + x' + y\|^2 - \|x + x' - y\|^2 = 2\|x' + y\|^2 + 2\|x\|^2 - 2\|x' - y\|^2 - 2\|x - y\|^2.$$

Sumando ahora miembro a miembro estas dos últimas igualdades vemos que

$$\|x + x' + y\|^2 - \|x + x' - y\|^2 = \|x + y\|^2 - \|x - y\|^2 + \|x' + y\|^2 - \|x' - y\|^2 \tag{5.3}$$

Segundo paso. Sea $\langle -, - \rangle : V \times V \to \mathbb{R}$ la función que siempre que x e y son elementos de V tiene

$$\langle x, y \rangle = \tfrac{1}{4}\|x + y\|^2 - \tfrac{1}{4}\|x - y\|^2,$$

y mostremos que se trata de un producto interno.

- Si x e y son dos elementos de V, entonces

  $$\langle x, y \rangle = \tfrac{1}{4}\|x + y\|^2 - \tfrac{1}{4}\|x - y\|^2 = \tfrac{1}{4}\|y + x\|^2 - \tfrac{1}{4}\|y - x\|^2 = \langle y, x \rangle$$

  y

  $$\langle x, x \rangle = \tfrac{1}{4}\|x + x\|^2 - \tfrac{1}{4}\|x - x\|^2 = \|x\|^2 \geq 0, \tag{5.4}$$

  así que las condiciones **(PI$_1$)** y **(PI$_2$)** también se cumplen.
- Si x, x', e y son tres elementos de V, entonces vale que

  $$\langle x + x', y \rangle = \langle x, y \rangle + \langle x', y \rangle,$$

  ya que esta igualdad es, en vista de la definición de la función $\langle -, - \rangle$, equivalente a (5.3). Para mostrar que se cumple la condición **(PI$_3$)** resta probar que para toda elección de x e y en V y de $t \in \mathbb{R}$ es

  $$\langle tx, y \rangle = t \cdot \langle x, y \rangle.$$

  Haremos eso en los siguientes dos pasos.

Tercer paso. Para cada $r \in \mathbb{Q}$ sea $P(r)$ la afirmación

$\langle rx, y \rangle = r\langle x, y \rangle$ *cualesquiera sean* x *e* y *en* V.

Queremos mostrar que $P(r)$ vale para todo $r \in \mathbb{Q}$, y lo hacemos en varias etapas:

- Observemos primero que si r es un número racional cualquiera y vale $P(r)$, entonces para cada x, $y \in V$ se tiene que

$$\langle (r+1)x, y\rangle = \langle rx + x, y\rangle = \langle rx, y\rangle + \langle x, y\rangle = r\langle x, y\rangle + \langle x, y\rangle = (r+1)\langle x, y\rangle,$$

  de manera que también vale $P(r+1)$. Como claramente vale $P(1)$, esto implica que, de hecho, vale $P(r)$ para todo $r \in \mathbb{N}$.
- Vale $P(r)$ para todo $r \in \mathbb{Q}$ positivo. En efecto, si $r = \frac{p}{q}$ con p, $q \in \mathbb{N}$, para cada elección de x y de $y \in V$ tenemos, ya que valen $P(p)$ y $P(q)$, que

$$p\langle x, y\rangle = \langle px, y\rangle = \left\langle q \cdot \frac{p}{q}x, y\right\rangle = q\left\langle \frac{p}{q}x, y\right\rangle,$$

  así que $\langle \frac{p}{q}x, y\rangle = \frac{p}{q}\langle x, y\rangle$.
- Vale $P(0)$, ya que

$$\langle 0x, y\rangle = \tfrac{1}{4}\|0x + y\|^2 - \tfrac{1}{4}\|0x - y\|^2 = \tfrac{1}{4}\|y\|^2 - \tfrac{1}{4}\|y\|^2 = 0 = 0\langle x, y\rangle,$$

  y si vale $P(r)$ para algún $r \in \mathbb{Q}$ entonces también vale $P(-r)$: para toda elección de x y de y en V se tiene en ese caso que

$$0 = \langle 0, y\rangle = \langle rx - rx, y\rangle = \langle rx + (-r)x, y\rangle = \langle rx, y\rangle + \langle -rx, y\rangle,$$

  por lo que $\langle -rx, y\rangle = -\langle rx, y\rangle = -r\langle x, y\rangle$. Junto con el paso anterior, esto nos permite concluir que vale $P(r)$ para todo $r \in \mathbb{Q}$, como queríamos.

Cuarto paso. Mostremos que

$$|\langle x, y\rangle| \leq \|x\|\|y\| \text{ siempre que } x \text{ e } y \text{ son elementos de } V. \tag{5.5}$$

Esto es evidente si $y = 0$, así que podemos suponer que no es ese el caso. En vista de lo que probamos en los el paso anterior, sabemos que para cada número racional r es

$$0 \leq \|x - ry\|^2 = \langle x - ry, x - ry\rangle = \langle x, x\rangle - 2r\langle x, y\rangle + r^2\langle y, y\rangle,$$

y esto nos dice que el polinomio $\|y\|^2T^2 - 2\langle x, y\rangle T + \|x\|^2$ de $\mathbb{R}[T]$ no toma valores negativos sobre $\mathbb{Q}$: como es una función continua, no toma entonces valores negativos en ningún punto de $\mathbb{R}$ y, por lo tanto, su discriminante no es positivo, esto es, $\langle x, y\rangle^2 - \|x\|^2\|y\|^2 \leq 0$. La desigualdad (5.5) es consecuencia inmediata de esto.

Quinto paso. Sean x e y dos elementos de V. La función $\zeta : t \in \mathbb{R} \mapsto \langle tx, y\rangle \in \mathbb{R}$ es continua. Para verlo basta observar que, de acuerdo a lo que sabemos ya de la función $\langle -, -\rangle$, se tiene que

$$|\zeta(s) - \zeta(t)| = \big|\langle sx, y\rangle - \langle tx, y\rangle\big| = \big|\langle (s-t)x, y\rangle\big| \leq \|(s-t)x\|\|y\| \leq |s-t|\|x\|\|y\|.$$

La continuidad de ζ y lo que probamos en el Tercer paso implica que la función

$$t \in \mathbb{R} \mapsto \langle tx, y\rangle - t\langle x, y\rangle \in \mathbb{R}$$

es continua y que se anula sobre $\mathbb{Q}$: es, por lo tanto, idénticamente nula esto nos dice que para todo $t \in \mathbb{R}$ vale que $\langle tx, y\rangle = t\langle x, y\rangle$. Vemos así que la función $\langle -, -\rangle$ satisface la condición **(PI$_3$)**.

Sexto y último paso. Concluimos de esta forma que la función $\langle -, -\rangle$ es un producto interno sobre V. De acuerdo a la igualdad (5.4), la norma asociada a este producto interno coincide con la norma de V. Esto prueba la proposición. □

Ejercicio 5.2.3.21. Pruebe la Proposición 5.2.3.20 en el caso de los espacios vectoriales complejos.

El espacio normado $\ell^\infty_{\mathbb{k}}$ y sus subespacios c y c_0

Queremos ahora construir una versión de dimensión infinita del espacio normado $\mathbb{k}^n$ con su norma $\|-\|_\infty$.

Proposición 5.2.3.22. *Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$.*

(*i*) *El conjunto $\ell^\infty_{\mathbb{k}}$ de todos los elementos $x = (x_n)_{n\geq 0}$ de $\mathbb{k}^{\mathbb{N}_0}$ tales que el conjunto $\{|x_n| : n \in \mathbb{N}_0\}$ es acotado es un $\mathbb{k}$-subespacio vectorial de $\mathbb{k}^{\mathbb{N}_0}$.*

(*ii*) *La función*

$$\|-\|_\infty : (x_n)_{n\geq 0} \in \ell^\infty_{\mathbb{k}} \mapsto \sup\{|x_n| : n \in \mathbb{N}_0\} \in \mathbb{R}$$

es una norma sobre $\ell^\infty_{\mathbb{k}}$.

Demostración. Sean $x = (x_n)_{n\geq 0}$ e $y = (y_n)_{n\geq 0}$ dos elementos de $\ell^\infty_{\mathbb{k}}$, sean a y b dos escalares de $\mathbb{k}$, y sea $z := ax + by$. Como x e y están en $\ell^\infty_{\mathbb{k}}$, hay números positivos c y d tales que $|x_n| \leq c$ y $|y_n| \leq d$ para todo $n \in \mathbb{N}_0$, y entonces para todo $n \in \mathbb{N}_0$ tenemos que

$$0 \leq |z_n| = |ax_n + by_n| \leq |a| \cdot |x_n| + |b| \cdot |y_n| \leq |a| \cdot c + |b| \cdot d.$$

Vemos así que el conjunto $\{|z_n| : n \in \mathbb{N}_0\}$ está acotado y, por lo tanto, que z pertenece a $\ell^\infty_{\mathbb{k}}$. Como además es claro que el elemento nulo de $\mathbb{k}^{\mathbb{N}_0}$ está en $\ell^\infty_{\mathbb{k}}$, esto implica que $\ell^\infty_{\mathbb{k}}$ es un subespacio de $\mathbb{k}^{\mathbb{N}_0}$, como afirma la parte (*i*).

Para probar la parte (*ii*) mostremos que la función $\|-\|_\infty$ definida ahí satisface las condiciones de la Definición 5.2.3.1.

- Si $x = (x_n)_{n\geq 0}$ es el elemento nulo del espacio $\ell^\infty_{\mathbb{k}}$, entonces el único elemento del conjunto $\{|x_n| : n \in \mathbb{N}_0\}$ es 0 y, por lo tanto, $\|x\|_\infty = \sup\{|x_n| : n \in \mathbb{N}_0\} = 0$. Si, por otro lado, $x = (x_n)_{n\geq 0}$ es un elemento de $\ell^\infty_{\mathbb{k}}$ tal que $\|x\|_\infty = 0$, entonces para cada $m \in \mathbb{N}_0$ tenemos

que $|x_m| \leq \sup\{|x_n| : n \in \mathbb{N}_0\} = \|x\|_\infty = 0$ y, por lo tanto, que $x_m = 0$: vemos así que x es el elemento nulo de $\ell^\infty_{\Bbbk}$.

- Sea $x = (x_n)_{n\geq 0}$ un elemento de $\ell^\infty_{\Bbbk}$, sea t un escalar de $\Bbbk$. Si t es 0, entonces es claro que $\|tx\|_\infty = 0 = |t| \cdot \|x\|_\infty$. Supongamos entonces que $t \neq 0$ y pongamos $y \coloneqq tx$

  Para cada $n \in \mathbb{N}_0$ es $|y_n| = |tx_n| = |t| \cdot |x_n| \leq |t| \cdot \|x\|_\infty$, así que $\|tx\|_\infty = \|y\|_\infty \leq |t| \cdot \|x\|_\infty$. Esto implica que

  $$\|x\|_\infty = \|t^{-1}y\|_\infty \leq |t^{-1}| \cdot \|y\|_\infty = |t|^{-1} \cdot \|tx\|_\infty$$

  y, por lo tanto que $|t| \cdot \|x\|_\infty \leq \|tx\|_\infty$. Combinando las dos desigualdades que obtuvimos podemos concluir que $\|tx\|_\infty = |t| \cdot \|t\|_\infty$.
- Sean $x = (x_n)_{n\geq 0}$ e $y = (y_n)_{n\geq 0}$ dos elementos de $\ell^\infty_{\Bbbk}$. Para cada $n \in \mathbb{N}_0$ es

  $$|x_n + y_n| \leq |x_n| + |y_n| \leq \|x\|_\infty + \|y\|_\infty,$$

  así que $\|x + y\| = \sup\{|x_n + y_n| : n \in \mathbb{N}_0\} \leq \|x\|_\infty + \|y\|_\infty$.

Vemos así que la función $\|-\|_\infty$ es una norma sobre el espacio vectorial $\ell^\infty_{\Bbbk}$. □

Podemos generalizar esta construcción y pasar de considerar sucesiones con valores en $\Bbbk$ a funciones definidas sobre un conjunto arbitrario.

Ejercicio 5.2.3.23. Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, y sea X un conjunto no vacío. Muestre que el conjunto $\ell^\infty_{\Bbbk}(X)$ de todas las funciones $X \to \Bbbk$ que son acotadas es un $\Bbbk$-subespacio vectorial de $\Bbbk^X$ y muestre que hay una norma $\|-\|_\infty$ sobre $\ell^\infty_{\Bbbk}(X)$ que en cada elemento f de $\ell^\infty_{\Bbbk}(X)$ toma el valor

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

El espacio $\ell^\infty_{\Bbbk}$ tiene tres subespacios importantes, que podemos ver como espacios normados restringiendo la norma de aquél.

Ejercicio 5.2.3.24. Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$. Pruebe las siguientes tres afirmaciones.

(a) El conjunto c de todas los elementos $x = (x_n)_{n\geq 0}$ de $\Bbbk^{\mathbb{N}_0}$ tales que existe $\lim\limits_{n\to\infty} x_n$ es un subespacio de $\ell^\infty_{\Bbbk}$.

(b) El conjunto c_0 de todos los elementos $x = (x_n)_{n\geq 0}$ de $\Bbbk^{\mathbb{N}_0}$ tales que $\lim\limits_{n\to\infty} x_n = 0$ es un subespacio de c.

(c) El conjunto c_{00} de todos los elementos $x = (x_n)_{n\geq 0}$ de $\Bbbk^{\mathbb{N}_0}$ tales que el conjunto $\{n \in \mathbb{N}_0 : x_n \neq 0\}$ es finito es un subespacio de c_{00}.

Ejercicio 5.2.3.25. Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$. Muestre que las normas de los espacios $\ell^\infty_{\Bbbk}$, c, c_0 y c_{00} no provienen de productos internos sobre esos espacios vectoriales.

Los espacios normados $\ell^p_{\mathbb{k}}$ con $p \in [1, \infty)$

Ahora nos proponemos construir versiones de dimensión infinita de los espacios normados $\mathbb{R}^n$ y $\mathbb{C}^n$ dotados de las normas $\|-\|_p$ con $p \in [1, +\infty)$.

Proposición 5.2.3.26. *Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, y sea p un elemento de $[1, +\infty)$.*

(*i*) *El conjunto $\ell^p_{\mathbb{k}}$ de todos los elementos $x = (x_n)_{n\geq 0}$ de $\mathbb{k}^{\mathbb{N}_0}$ tales que la serie $\sum_{n=0}^{\infty}|x_n|^p$ converge es un $\mathbb{k}$-subespacio vectorial de $\mathbb{k}^{\mathbb{N}_0}$.*

(*ii*) *La función*

$$\|-\|_p : (x_n)_{n\geq 0} \in \ell^p_{\mathbb{k}} \mapsto \left(\sum_{n=0}^{\infty}|x_n|^p\right)^{1/p} \in \mathbb{R}$$

es una norma sobre $\ell^p_{\mathbb{k}}$.

Demostración. Veamos primero que $\ell^p_{\mathbb{k}}$ es un $\mathbb{k}$-subespacio de $\mathbb{k}^{\mathbb{N}_0}$.

- Es claro que el elemento nulo de $\mathbb{k}^{\mathbb{N}_0}$ pertenece a $\ell^p_{\mathbb{k}}$.
- Sea $x = (x_n)_{n\geq 0}$ un elemento de $\ell^p_{\mathbb{k}}$, y sea t un escalar de $\mathbb{k}$. Para cada $N \in \mathbb{N}$ es

$$\sum_{n=0}^{N}|tx_n|^p = \sum_{n=0}^{N}|t|^p \cdot |x_n|^p = |t|^p \cdot \sum_{n=0}^{N}\cdot|x_n|^p \leq |t|^p \cdot \sum_{n=0}^{\infty}\cdot|x_n|^p,$$

  ya que esta última serie tiene todos sus términos no negativos, y esto nos dice que la serie $\sum_{n=0}^{\infty}|tx_n|^p$ converge, porque también tiene todos sus términos no negativos y sus sumas parciales acotadas.
- Sean $x = (x_n)_{n\geq 0}$ e $y = (y_n)_{n\geq 0}$ dos elementos de $\ell^p_{\mathbb{k}}$. Si N es un entero positivo, entonces la *Desigualdad de Minkowski* de la Proposición 5.2.3.12 aplicada a los dos vectores $(x_0, x_1, \ldots, x_N)$ e $(y_0, y_1, \ldots, y_N)$ de $\mathbb{k}^{N+1}$ implica que

$$\sum_{n=0}^{N}|x_n + y_n|^p \leq \left(\left(\sum_{n=0}^{N}|x_n|^p\right)^{1/p} + \left(\sum_{n=0}^{N}|y_n|^p\right)^{1/p}\right)^p$$

  y como las funciones $t \in [0, +\infty) \mapsto t^{1/p} \in \mathbb{R}$ y $t \in [0, +\infty) \mapsto t^p \in \mathbb{R}$ son crecientes, esto es

$$\leq \left(\left(\sum_{n=0}^{\infty}|x_n|^p\right)^{1/p} + \left(\sum_{n=0}^{\infty}|y_n|^p\right)^{1/p}\right)^p.$$

  Vemos así que la serie $\sum_{n=0}^{\infty}|x_n + y_n|^p$ converge, ya que tiene sus términos todos no negativos y sus sumas parciales acotadas, y que su suma es tal que

$$\sum_{n=0}^{\infty}|x_n + y_n|^p \leq \left(\left(\sum_{n=0}^{\infty}|x_n|^p\right)^{1/p} + \left(\sum_{n=0}^{\infty}|y_n|^p\right)^{1/p}\right)^p. \tag{5.6}$$

En particular, esto nos permite concluir que $x + y \in \ell^p_{\Bbbk}$.

Todo esto prueba que, como queremos, $\ell^p_{\Bbbk}$ es un $\Bbbk$-subespacio de $\Bbbk^{\mathbb{N}_0}$. Probemos ahora que la función $\|-\|_p$ definida en el enunciado de la proposición es una norma sobre él.

- Es claro que si $x = (x_n)_{n\geq 0}$ es el elemento nulo de $\ell^\infty_{\Bbbk}$, entonces $\|x\|_p = \left(\sum_{n=0}^\infty |x_n|^p\right)^{1/p} = 0$. Supongamos, por otro lado, que $x = (x_n)_{n\geq 0}$ es un elemento de $\ell^\infty_{\Bbbk}$ tal que $\|x\|_p = 0$. Tenemos entonces que para cada $m \in \mathbb{N}_0$ es $|x_m|^p \leq \sum_{n=0}^\infty |x_n|^p = \|x\|_p^p = 0$, así que $x = 0$.
- Si $x = (x_n)_{n\geq 0}$ es un elemento de $\ell^\infty_{\Bbbk}$ y t es un escalar de $\Bbbk$, entonces

$$\begin{aligned}\|tx\|_p &= \left(\sum_{n=0}^\infty |tx_n|^p\right)^{1/p} = \left(\sum_{n=0}^\infty \left(|t|^p \cdot |x_n|^p\right)\right)^{1/p} = \left(|t|^p \cdot \sum_{n=0}^\infty |x_n|^p\right)^{1/p} \\ &= |t|\left(\sum_{n=0}^\infty |x_n|^p\right)^{1/p} = |t| \cdot \|x\|_p.\end{aligned}$$

- Finalmente, vimos arriba que si $x = (x_n)_{n\geq 0}$ e $y = (y_n)_{n\geq 0}$ son dos elementos de $\ell^p_{\Bbbk}$, entonces vale la desigualdad (5.6), y ella implica inmediatamente que $\|x + y\|_p \leq \|x\|_p + \|y\|_p$. □

Obtenemos de esta forma una familia de $\Bbbk$-subespacios $\ell^p_{\Bbbk}$ de $\Bbbk^{\mathbb{N}_0}$, uno para cada elemento p de $[1, +\infty]$. Esta familia es estrictamente creciente.

Ejercicio 5.2.3.27. Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, y sean p y q dos elementos de $[1, +\infty]$ tales que $p < q$. Pruebe las siguientes afirmaciones.

(a) Si $q < \infty$, entonces para cada elemento x de $\ell^p_{\Bbbk}$ se tiene que $\sum_{i=0}^\infty |x_i|^q \leq \|x\|_p^q$, de manera que $x \in \ell^q_{\Bbbk}$ y $\|x\|_q \leq \|x\|_p$, y concluya que $\ell^p_{\Bbbk} \subseteq \ell^q_{\Bbbk}$.
(b) Para cada elemento x de $\ell^p_{\Bbbk}$ se tiene que $|x_i| \leq \|x\|_p$ para todo $i \in \mathbb{N}_0$, de manera que $x \in \ell^\infty_{\Bbbk}$ y $\|x\|_\infty \leq \|x\|_p$ y, por lo tanto, es $\ell^p_{\Bbbk} \subseteq \ell^\infty_{\Bbbk}$.
(c) Es $\ell^p_{\Bbbk} \subsetneq \ell^q_{\Bbbk}$.

Las normas de estos espacios normados de sucesiones no provienen de productos internos salvo en un caso.

Ejercicio 5.2.3.28. Sea p un elemento de $(1, +\infty)$. Pruebe que la norma $\|-\|_p$ sobre $\ell^p_{\Bbbk}$ proviene de un producto interno sobre ese espacio vectorial si y solamente si $p = 2$.

Como hicimos con el espacio $\ell^\infty_{\Bbbk}$, podemos generalizar la construcción del espacio $\ell^p_{\Bbbk}$ a espacios de funciones definidas sobre un conjunto arbitrario.

Ejercicio 5.2.3.29. Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, sea X un conjunto no vacío, y sea p un elemento de $[1, +\infty)$. Decimos que una función $f : X \to \Bbbk$ es ***p-sumable*** si hay un número real c tal que para todo

subconjunto finito F de X vale que

$$\left(\sum_{x\in F}|f(x)|^p\right)^{1/p} \leq c,$$

y cuando ese es el caso escribimos $C(f)$ al conjunto de todos los números reales c que tienen esa propiedad y ponemos $\|f\|_p \coloneqq \text{ínf}\, C(f)$. Muestre que el conjunto $\ell^p_{\Bbbk}(X)$ de todas las funciones p-sumables $X \to \Bbbk$ es un $\Bbbk$-subespacio vectorial de $\Bbbk^X$ y que la función $\|-\|_p : f \in \ell^p_{\Bbbk}(X) \mapsto \|f\|_p \in \mathbb{R}$ es una norma sobre él.

Para cada elemento p de $[1,+\infty]$ hemos construido espacio normado $\ell^p_{\Bbbk}$ de sucesiones. El siguiente ejercicio se ocupa de lo que pasa cuando p pertenece al intervalo $(0,1)$.

Ejercicio 5.2.3.30. Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, y sea p un elemento de $(0,1)$.

(a) Muestre que el conjunto $\ell^p_{\Bbbk}$ de todos los elementos x de $\Bbbk^{\mathbb{N}_0}$ tales que la serie

$$\sum_{i=0}^{\infty}|x_i|^p$$

converge es un $\Bbbk$-subespacio vectorial de $\Bbbk^{\mathbb{N}_0}$. Concluya usando eso que hay una función

$$d : (x,y) \in \ell^p_{\Bbbk} \times \ell^p_{\Bbbk} \mapsto \sum_{i=0}^{\infty}|x_i - y_i|^p \in \mathbb{R},$$

y pruebe que esta es una métrica sobre $\ell^p_{\Bbbk}$.

(b) ¿Qué relaciones de inclusión hay entre los espacios $\ell^p_{\Bbbk}$ con $p \in (0,+\infty]$?

Podemos generalizar todas estas construcciones de espacios normados de sucesiones de escalares para obtener espacios normados de sucesiones de vectores de un espacio normado.

Ejercicio 5.2.3.31. Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, sea V un $\Bbbk$-espacio normado, y escribamos $\|-\|_V$ a la norma de V.

(a) Sea ℓ^∞_V el conjunto de todas las sucesiones $x = (x_n)_{n\geq 0}$ en V tales que el conjunto

$$C(x) \coloneqq \{\|x_n\|_V : n \in \mathbb{N}_0\}$$

está acotado superiormente. Pruebe que ℓ^∞_V es un $\Bbbk$-subespacio del $\Bbbk$-espacio vectorial $V^{\mathbb{N}_0}$ de todas las sucesiones en V, y muestre que la función $\|-\| : x \in \ell^\infty_V \mapsto \sup C(x) \in \mathbb{R}$ es una norma sobre él.

(b) Sea p un elemento de $(0,+\infty)$. Demuestre que el conjunto ℓ^p_V de todas las sucesiones $x = (x_n)_{n\geq 0}$ en V tales que la serie $\sum_{n=0}^{\infty}\|x_n\|_V^p$ converge es un $\Bbbk$-subespacio de $V^{\mathbb{N}_0}$. Pruebe

que si $p \geq 1$, entonces la función

$$\|-\| : x \in V \mapsto \left(\sum_{n=0}^{\infty} \|x_n\|_V^p \right)^{1/p} \in \mathbb{R}$$

es una norma sobre ℓ_V^p, y que si $d \in (0,1)$, entonces la función

$$d : (x,y) \in V \times V \mapsto \sum_{n=0}^{\infty} \|x - y\| \in \mathbb{R}$$

es una métrica sobre ℓ_V^p que es invariante por translaciones.

5.2.4. Semi-métricas y semi-normas

El objetivo de esta subsección es presentar unas formas de construir métricas y normas que son útiles para producir ciertos ejemplos importantes.

Definición 5.2.4.1. Sea X un conjunto. Una ***semi-métrica*** sobre X es una función $d : X \times X \to \mathbb{R}$ tal que para toda elección de x y de y en X valen las siguientes condiciones:

(**SM**$_1$) $d(x,x) = 0$.
(**SM**$_2$) $d(x,y) = d(y,x)$.
(**SM**$_3$) $d(x,y) \leq d(x,z) + d(z,y)$.

Las condiciones (**SM**$_2$) y (**SM**$_3$) son exactamente las mismas que las condiciones (**M**$_2$) y (**SM**$_3$) de la Definición 5.1.1 que dimos de lo que es una métrica — y esto implica, en particular, que podemos probar en la misma forma en que lo hicimos para las métricas que una semi-métrica toma valores no negativos. La diferencia entre las dos definiciones es que una semi-métrica puede tomar el valor nulo en un par (x,y) con $x \neq y$. Un ejemplo extremo que muestra esto es el siguiente: cualquiera sea el conjunto X la función constante nula

$$d : (x,y) \in X \times X \mapsto 0 \in \mathbb{R}$$

es una semi-métrica, pero es una métrica solamente si el conjunto X tiene a lo sumo un elemento.

El siguiente resultado nos da una forma de obtener a partir de un espacio semi-métrico un espacio métrico de manera natural. Usaremos esta construcción para construir la completación de un espacio métrico en la Sección 7.3.

Proposición 5.2.4.2. *Sea X un conjunto y sea $d : X \times X \to \mathbb{R}$ una semi-métrica sobre X.*

(*i*) *El conjunto $R \coloneqq \{(x,y) \in X \times X : d(x,y) = 0\}$ es una relación de equivalencia sobre el conjunto X.*
(*ii*) *Sea $\bar{X} \coloneqq X/R$ el conjunto de las clases de equivalencia de R en X, y para cada elemento x de X*

escribamos $\bar{x}$ a la clase de equivalencia de x con respecto a R. Hay exactamente una función $\bar{d} : \bar{X} \times \bar{X} \to \mathbb{R}$ tal que para cada elección de x y de y en X es $\bar{d}(\bar{x}, \bar{y}) = d(x, y)$, y esa función es una métrica sobre el conjunto $\bar{X}$.

Demostración. Las condiciones **(SM$_1$)**, **(SM$_2$)**, y **(SM$_3$)** de la Definición 5.2.4.1 implican inmediatamente que la relación R en X es reflexiva, simétrica, y transitiva, respectivamente, así que R es una relación de equivalencia sobre X. Consideremos el subconjunto

$$\bar{d} \coloneqq \{((\bar{x}, \bar{y}), d(x, y)) \in (\bar{X} \times \bar{X}) \times \mathbb{R} : x, y \in X\}$$

de $(\bar{X} \times \bar{X}) \times \mathbb{R}$ y mostremos que es, de hecho, una función $\bar{X} \times \bar{X} \to \mathbb{R}$.

- Si u es un elemento de $\bar{X} \times \bar{X}$, entonces hay dos elementos x e y de X tales que $u = (\bar{x}, \bar{y})$ y, por lo tanto, el par $(u, d(x, y))$ es un elemento de $\bar{d}$.
- Sean u un elemento de $\bar{X} \times \bar{X}$ y r y s dos elementos de $\mathbb{R}$ tales que $(u, r) \in \bar{d}$ y $(u, s) \in \bar{d}$. Esto significa que hay elementos x_1, x_2, y_1, e y_2 de X tales que $(u, r) = ((\bar{x}_1, \bar{y}_1), d(x_1, y_1))$ y $(u, s) = ((\bar{x}_2, \bar{y}_2), d(x_2, y_2))$. Como $(\bar{x}_1, \bar{y}_1) = u = (\bar{x}_2, \bar{y}_2)$, tenemos que $\bar{x}_1 = \bar{x}_2$ y que $\bar{y}_1 = \bar{y}_2$ y, en consecuencia, es $d(x_1, x_2) = 0$ y $d(y_1, y_2) = 0$. Se sigue de esto que

  $$d(x_1, y_1) \leq d(x_1, x_2) + d(x_2, y_2) + d(y_2, y_1) = d(x_2, y_2)$$

  y

  $$d(x_2, y_2) \leq d(x_2, x_1) + d(x_1, y_1) + d(y_1, y_2) = d(x_1, y_1),$$

  de manera que es $r = d(x_1, y_1) = d(x_2, y_2) = s$.

Hecho esto, notemos que es evidente que la definición de la función $\bar{d}$ implica inmediatamente que cada elección de x y de y en X es $\bar{d}(\bar{x}, \bar{y}) = d(x, y)$ y, como todo elemento de $\bar{X} \times \bar{X}$ es de la forma $(\bar{x}, \bar{y})$ para alguna elección de x y de y en X, que es la única función con esa propiedad.

Mostremos ahora que la función $\bar{d} : \bar{X} \times \bar{X} \to \mathbb{R}$ es una métrica sobre el conjunto $\bar{X}$.

- Si u es un elemento de $\bar{X}$, de manera que hay un elemento x de X tal que $u = \bar{x}$, entonces $\bar{d}(u, u) = \bar{d}(\bar{x}, \bar{x}) = d(x, x) = 0$. Sean, por otro lado, u y v dos elementos de $\bar{X}$ tales que $\bar{d}(u, v) = 0$. Hay elementos x e y de X tales que $u = \bar{x}$ y $v = \bar{y}$, y es $d(x, y) = \bar{d}(\bar{x}, \bar{y}) = \bar{d}(u, v) = 0$, así que $(x, y) \in R$ y, por lo tanto, $u = \bar{x} = \bar{y} = v$. Esto nos dice que vale la condición **(M$_1$)**.
- Si u y v son dos elementos de $\bar{X}$, entonces hay dos elementos x e y de X tales que $u = \bar{x}$ y $v = \bar{y}$ y, por lo tanto, $\bar{d}(u, v) = \bar{d}(\bar{x}, \bar{y}) = d(x, y) = d(y, x) = \bar{d}(\bar{y}, \bar{x}) = \bar{d}(v, u)$. La función $\bar{d}$ satisface, en consecuencia, la condición **(M$_2$)**.
- Sean, finalmente, u, v, y w tres elementos de $\bar{X}$, de manera que hay tres elementos x, y, y z de X tales que $u = \bar{x}$, $v = \bar{y}$, y $w = \bar{z}$. Es entonces

  $$\bar{d}(u, v) = \bar{d}(\bar{x}, \bar{y}) = d(x, y) \leq d(x, z) + d(z, y) = \bar{d}(\bar{x}, \bar{z}) + \bar{d}(\bar{z}, \bar{y}) = \bar{d}(u, w) + \bar{d}(w, v).$$

La función $\bar{d}$ satisface, por lo tanto, la desigualdad triangular de la condición ($\mathbf{M}_3$).

Esto completa la demostración de la proposición. □

Las métricas tienen la misma relación con las semi-métricas que la que tienen las normas con las semi-normas que presenta la siguiente definición.

Definición 5.2.4.3. Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, y sea V un espacio vectorial sobre $\Bbbk$. Una función $\|-\| : V \to \mathbb{R}$ es una ***semi-norma*** sobre V si cualesquiera sean los vectores x e y de V y el escalar t de $\Bbbk$ satisface las siguientes condiciones.

($\mathbf{SN}_1$) $\|t \cdot x\| = |t| \cdot \|x\|$.

($\mathbf{SN}_2$) $\|x + y\| \leq \|x\| + \|y\|$.

Como pasa con las normas, las semi-normas se anulan en 0 y toman valores no negativos: podemos probar esto de la misma forma en que lo hicimos para las normas en la página 121. Dejamos al lector la tarea de mostrar que una semi-norma da lugar a una semi-métrica:

Ejercicio 5.2.4.4. Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, sea V un espacio vectorial sobre $\Bbbk$, y sea $\|-\| : V \to \mathbb{R}$ una semi-norma sobre V. Pruebe que la función $d : (x, y) \in V \times V \mapsto \|x - y\| \in \mathbb{R}$ es una semi-métrica sobre V.

Todo subespacio de un espacio normado da lugar a una semi-norma sobre este, como describe la siguiente proposición.

Proposición 5.2.4.5. *Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, sea X un $\Bbbk$-espacio normado y sea Y un subespacio de X. La función*

$$\|-\|' : x \in X \mapsto \inf\{\|x + y\| : y \in Y\} \in \mathbb{R}$$

es una semi-norma sobre X.

Notemos que para cada elemento x de X el conjunto $\{\|x + y\| : y \in Y\}$ ciertamente no es vacío y es acotado inferiormente, así que tiene sentido considerar su ínfimo.

Demostración. Verifiquemos las condiciones de la Definición 5.2.4.3.

- Sea x un elemento de X y sea t un escalar no nulo de $\Bbbk$. Para cada $y \in Y$ tenemos que $\|tx + y\| = \|t(x + y/t)\| = |t| \cdot \|x + y/t\| \geq |t| \cdot \|x\|'$, ya que $y/t \in Y$ y $|t| > 0$, y esto implica que $\|tx\|' \geq |t| \cdot \|x\|'$. Como t^{-1} también es un escalar no nulo, esto implica que, en particular $\|x\|' = \|t^{-1}(tx)\| \geq |t^{-1}| \cdot \|tx\|'$ y, por lo tanto, que $|t| \cdot \|x\|' \geq \|tx\|'$. Podemos concluir de esto que $\|tx\|' = |t| \cdot \|x\|'$ para todo escalar no nulo t de $\Bbbk$ y, como esta igualdad también vale cuando t es 0, que vale la condición ($\mathbf{SN}_1$).

- Sean ahora x y x' dos elementos de X. Si y e y' son dos elementos de Y, entonces $y + y' \in Y$, así que

$$\|x + x'\|' \leq \|x + x' + y + y'\| \leq \|x + y\| + \|x' + y'\|$$

o, equivalentemente, $\|x + x'\|' - \|x' + y'\| \leq \|x + y\|$. Esto es cierto cualquiera sea y en Y, así que nos dice que

$$\|x + x'\|' - \|x' + y'\| \leq \inf\{\|x + y\| : y \in Y\} = \|x\|'$$

y, por lo tanto, que $\|x + x'\|' - \|x\|' \leq \|x' + y'\|$. Otra vez, esto es cierto cualquiera sea y' en Y, así que implica que

$$\|x + x'\|' - \|x\|' \leq \inf\{\|x' + y'\| : y' \in Y\} = \|x'\|'.$$

Vemos así que $\|x + x'\|' \leq \|x\|' + \|x'\|'$ y, en definitiva, que vale la condición **(SN$_2$)**. □

El siguiente resultado describe una construcción similar a la de la Proposición 5.2.4.2 pero para semi-normas.

Proposición 5.2.4.6. *Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, sea V un espacio vectorial sobre $\mathbb{k}$, y sea $\|-\| : V \to \mathbb{R}$ una semi-norma sobre V.*

(*i*) *El conjunto $V_0 \coloneqq \{x \in V : \|x\| = 0\}$ es un subespacio vectorial de V, de manera que, en particular, podemos considerar el espacio cociente V/V_0.*

(*ii*) *La función $\|-\|' : \xi \in V/V_0 \mapsto \inf\{\|x\| : x \in \xi\} \in \mathbb{R}$ es una norma sobre V/V_0.*

Demostración. El subconjunto V_0 contiene claramente a 0, así que no es vacío, y es cerrado por multiplicación por escalares y por sumas porque la semi-métrica satisface las condiciones **(SN$_1$)** y **(SN$_2$)**: esto nos dice que es un subespacio vectorial de V.

Si ξ es un elemento del cociente V/V_0, entonces el conjunto $C(\xi) \coloneqq \{\|x\| : x \in \xi\}$ no es vacío y está contenido en $[0, +\infty)$, así que podemos considerar su ínfimo: hay entonces una función

$$\|-\|' : \xi \in V/V_0 \mapsto \inf C(\xi) \in \mathbb{R}.$$

Mostremos que se trata de una norma sobre el espacio vectorial V/V_0.

- Si ξ es el elemento nulo de V/V_0, de manera que $0 \in \xi$, entonces $0 = \|0\|$ pertenece al conjunto $C(\xi)$ y, por supuesto, es $\|\xi\|' = 0$.
- Sea, por otro lado, ξ un elemento de V/V_0 tal que $\|\xi\|' = 0$ y sea x un elemento de V tal que $\xi = x + V$. Si ε es un número positivo, entonces $\inf\{\|x + v\| : v \in V_0\} = \inf C(\xi) = 0 < \varepsilon$, hay un elemento v de V_0 tal que $\|x + v\| < \varepsilon$ y, por lo tanto, es

$$0 \leq \|x\| = \|(x + v) + (-v)\| \leq \|x + v\| + \|-v\| < \varepsilon,$$

ya que $-v \in V_0$. Esto es así cualquiera sea el número positivo ε, así que $\|x\| = 0$ y, por lo tanto, $x \in V_0$ y ξ es V_0, el elemento nulo de V/V_0.

- Sea ξ un elemento de V/V_0 y sea t uno de $\Bbbk$. Si $t = 0$, entonces es claro que $\|t \cdot \xi\|' = |t| \cdot \|\xi\|'$, ya que ambos lados de la igualdad se anulan. Consideremos el caso en que $t \neq 0$. Hay un vector x de V tal que $\xi = x + V_0$, así que $t \cdot \xi = tx + V_0 = tx + tV_0$, ya que t no es nulo, y, por lo tanto,

$$\begin{aligned} C(t \cdot \xi) &= \{\|tx + tv\| : v \in V_0\} = \{|t| \cdot \|x + v\| : v \in V_0\} = |t| \cdot \{\|x + v\| : v \in V_0\} \\ &= |t| \cdot C(\xi). \end{aligned}$$

  Como $|t|$ es un número positivo, esto nos permite concluir que

$$\|t \cdot \xi\|' = \text{ínf}\, C(t \cdot \xi) = |t| \cdot \text{ínf}\, C(\xi) = |t| \cdot \|\xi\|'$$

  también en este caso.

- Sean finalmente ξ y ζ dos elementos de V/V_0. Si c es un elemento del conjunto $C(\xi) + C(\zeta)$, entonces hay elementos x y z de ξ y de ζ tales que $c = \|x\| + \|z\|$ y, como $x + z \in \xi + \zeta$, es

$$\|\xi + \zeta\|' = \text{ínf}\, C(\xi + \zeta) \leq \|x + z\| \leq \|x\| + \|z\| = c.$$

  Esto nos dice que

$$\|\xi + \zeta\|' \leq \text{ínf}(C(\xi) + C(\zeta)) = \text{ínf}\, C(\xi) + \text{ínf}\, C(\zeta) = \|\xi\|' + \|\zeta\|'$$

  así que la función $\|-\|'$ satisface la desigualdad triangular. □

Veamos un ejemplo de cómo podemos usar estas construcciones.

Ejemplo 5.2.4.7. Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, sea X un conjunto, y sea $\mathscr{F}$ un subconjunto no vacío de $\mathscr{P}(X)$ que satisface las siguientes dos condiciones:

$$F, G \in \mathscr{F} \implies F \cup G \in \mathscr{F}, \qquad F \subseteq G \in \mathscr{F} \implies F \in \mathscr{F}.$$

Decimos que un número no negativo c es una ***$\mathscr{F}$-cota*** para una función $f : X \to \Bbbk$ si el conjunto

$$\{x \in X : |f(x)| > c\}$$

pertenece a $\mathscr{F}$, y que una función $f : X \to \Bbbk$ es ***$\mathscr{F}$-acotada*** si el conjunto $C_{\mathscr{F}}(f)$ de los elementos de $[0, +\infty)$ que son $\mathscr{F}$-cotas para f no es vacío. Mostremos que

el conjunto $B_{\mathscr{F}}(X)$ de las funciones $X \to \Bbbk$ que son $\mathscr{F}$-acotadas es un subespacio vectorial del $\Bbbk$-espacio $\Bbbk^X$ de todas las funciones $X \to \Bbbk$. (5.7)

Es claro que 0 es una $\mathscr{F}$-cota para la función nula $x \in X \mapsto 0 \in \Bbbk$, así que está función es $\mathscr{F}$-acotada y, por lo tanto, un elemento de $B_{\mathscr{F}}(X)$. Vemos así que este conjunto no es vacío. Sean, por otro lado, $f, g : X \to \Bbbk$ dos funciones $\mathscr{F}$-acotadas y sean s y t dos elementos de $\Bbbk$. Hay dos números no negativos c y d tales que los conjuntos

$$F \coloneqq \{x \in X : |f(x)| > c\}, \qquad G \coloneqq \{x \in X : |g(x)| > d\}$$

pertenecen a $\mathscr{F}$. Pongamos $e \coloneqq |s| \cdot c + |t| \cdot d$. Si x es un elemento de X que no pertenece a $F \cup G$, entonces

$$|(sf + tg)(x)| \leq |s| \cdot |f(x)| + |t| \cdot |g(x)| \leq |s| \cdot c + |t| \cdot d = e,$$

así que x no pertenece tampoco al conjunto

$$H \coloneqq \{x \in X : |(sf + tg)(x)| > e\}.$$

Esto nos dice que $H \subseteq F \cup G$ y las hipótesis que hicimos sobre el conjunto $\mathscr{F}$ nos permiten concluir que $H \in \mathscr{F}$: el número e es, por lo tanto, una $\mathscr{F}$ cota para la función $sf + tg$ y, por lo tanto, la función $sf + tg$ pertenece a $B_{\mathscr{F}}(X)$. Esto termina la demostración de la afirmación (5.7).

En vista de la definición del espacio $B_{\mathscr{F}}(X)$ hay una función

$$\|-\| : f \in B_{\mathscr{F}}(X) \mapsto \text{ínf}\, C_{\mathscr{F}}(f) \in \mathbb{R}.$$

Probemos que se trata de una semi-norma sobre $B_{\mathscr{F}}(X)$.

- Sea f un elemento de $B_{\mathscr{F}}(X)$ y sea t un escalar de $\Bbbk$. Si $t = 0$, entonces $\|t \cdot f\| = |t| \cdot \|f\|$ simplemente porque ambos lados de la igualdad son nulos. Supongamos que $t \neq 0$. Para todo número positivo c es $\{x \in X : |(t \cdot f)(x)| > |t| \cdot c\} = \{x \in X : |f(x)| > c\} \in \mathscr{F}$, y esto implica inmediatamente que $C(t \cdot f) = |t| \cdot C_{\mathscr{F}}(f)$. A su vez, esta igualdad implica que

  $$\|t \cdot f\| = \text{ínf}\, C(t \cdot f) = \text{ínf}\big(|t| \cdot C(f)\big) = |t| \cdot \text{ínf}\, C(f) = |t| \cdot \|f\|.$$

  Vemos así que la función $\|-\|$ satisface la condición **(SN$_1$)**.
- Sean f y g dos elementos de $B_{\mathscr{F}}(X)$. Si c y d son dos números reales, entonces

  $$\{x \in X : |f(x) + g(x)| > c + d\} \subseteq \{x \in X : |f(x)| > c\} \cup \{x \in X : |g(x)| > d\}.$$

  En vista de las hipótesis que hicimos sobre el conjunto $\mathscr{F}$, esto implica que $c + d \in C_{\mathscr{F}}(f + g)$ siempre que $c \in C_{\mathscr{F}}(f)$ y $d \in C_{\mathscr{F}}(g)$, y, por lo tanto, que $C_{\mathscr{F}}(f) + C_{\mathscr{F}}(g) \subseteq C_{\mathscr{F}}(f + g)$, así que

  $$\begin{aligned}\|f + g\| = \text{ínf}\, C_{\mathscr{F}}(f + g) &\leq \text{ínf}\big(C_{\mathscr{F}}(f) + C_{\mathscr{F}}(g)\big) = \text{ínf}\, C_{\mathscr{F}}(f) + \text{ínf}\, C_{\mathscr{F}}(g) \\ &= \|f\| + \|g\|.\end{aligned}$$

  La función $\|-\|$ satisface entonces la condición **(SN$_2$)**.

Todo esto nos dice que estamos en condiciones de usar el resultado de la Proposición 5.2.4.6. Notemos que una función $f : X \to \Bbbk$ pertenece al subespacio $V_0 \coloneqq \{f \in B_{\mathscr{F}}(X) : \|f\| = 0\}$ de $B_{\mathscr{F}}(X)$ exactamente cuando hay un elemento F de $\mathscr{F}$ tal que $f(x) = 0$ para todo $x \in X \setminus F$ y que, por lo tanto,

> *dos funciones $f, g : X \to \Bbbk$ de $B_{\mathscr{F}}(X)$ son congruentes módulo V_0 exactamente cuando hay un elemento F de $\mathscr{F}$ tal que $f(x) = g(x)$ para todo $x \in X \setminus F$.*

Decimos en ese caso que las funciones f y g son ***congruentes módulo $\mathscr{F}$***, y al espacio normado $B_{\mathscr{F}}(X)/V_0$ lo escribimos $\ell^\infty_\Bbbk(X \text{ mód } \mathscr{F})$.

Variando la elección del conjunto $\mathscr{F}$, esta construcción produce espacios importantes.

- Si $\mathscr{F} = \{\varnothing\}$, entonces el espacio normado $\ell^\infty_\Bbbk(X \text{ mód } \mathscr{F})$ se identifica con el espacio $\ell^\infty_\Bbbk(X)$ del Ejercicio 5.2.3.23. En efecto, en este caso una $\mathscr{F}$-cota para una función $f : X \to \Bbbk$ es lo mismo que una cota, así que $B_{\mathscr{F}}(X) = \ell^\infty_\Bbbk(X)$, la semi-norma $\|-\|$ es una norma, y la relación de equivalencia correspondiente es la de igualdad.
- Supongamos ahora que $\mathscr{F}$ es el conjunto de todos los subconjuntos finitos de X. Una función $X \to \Bbbk$ es $\mathscr{F}$-acotada si y solamente si es acotada, así que $B_{\mathscr{F}}(X) = \ell^\infty_\Bbbk(X)$, pero la semi-norma sobre $B_{\mathscr{F}}(X)$ no es una norma. En efecto, una función $f : X \to \Bbbk$ tiene $\|f\| = 0$ exactamente cuando el conjunto $\{x \in X : f(x) \neq 0\}$ es finito. Esto implica que la relación de equivalencia que usamos para contruir el espacio $\ell^\infty_\Bbbk(X \text{ mód } \mathscr{F})$ no es la de igualdad: dos funciones $X \to \Bbbk$ son equivalentes con respecto a ella si coinciden en un subconjunto de X de complemento finito.
- Supongamos finalmente que $\mathscr{F}$ es el conjunto de todos los subconjuntos contables de X. En este caso no es cierto que una función $X \to \Bbbk$ que es $\mathscr{F}$-acotada sea necesariamente acotada: $B_{\mathscr{F}}(X)$ es el conjunto de las funciones $X \to \Bbbk$ que son acotadas en el complemento de un conjunto contable, y $\ell^\infty_\Bbbk(X \text{ mód } \mathscr{F})$ tiene por elementos las clases de equivalencia de esas funciones: identificamos dos de ellas cuando difieren solamente sobre un conjunto contable de X.

Ejercicio 5.2.4.8. Decimos que un subconjunto N de $\mathbb{R}$ tiene ***medida nula*** si para todo número positivo ε es posible encontrar una sucesión $(I_n)_{n\geq 0}$ de intervalores abiertos $I_n = (a_n, b_n)$ de $\mathbb{R}$ tales que $N \subseteq \bigcup_{n=0}^\infty (a_n, b_n)$ y

$$\sum_{n=0}^{\infty} (b_n - a_n) < \varepsilon.$$

Sea $\mathscr{N}$ el conjunto de todos los subconjuntos de $\mathbb{R}$ que tienen medida nula.

(a) Pruebe que para toda elección de subconjuntos M y N de $\mathbb{R}$ valen

$$M, N \in \mathscr{N} \implies M \cup N \in \mathscr{N}, \qquad M \subseteq N \in \mathscr{N} \implies M \in \mathscr{N},$$

de manera que este conjunto $\mathscr{N}$ satisface las condiciones necesarias para llevar a cabo la construcción del Ejemplo 5.2.4.7.

(b) Pruebe que, más aún, cada vez que $(F_n)_{n\geq 0}$ es una sucesión de elementos de $\mathscr{N}$, la unión $\bigcup_{n=0}^{\infty} F_n$ también es un elemento de $\mathscr{N}$.

5.2.5. Métricas inducidas por familias de semi-métricas y de semi-normas

Hay otra forma de usar semi-métricas y semi-normas para construir métricas sobre conjuntos bien distinta de la que describimos en el apartado anterior. De ella nos ocupamos aquí.

Definición 5.2.5.1. Sea X un conjunto. Decimos que un conjunto $\mathscr{D}$ de semi-métricas sobre X ***separa puntos*** si siempre que x e y son dos elementos distintos de X, entonces hay un elemento d de $\mathscr{D}$ tal que $d(x, y) \neq 0$.

El siguiente resultado es el que estamos buscando.

Proposición 5.2.5.2. *Sea X un conjunto, sea $\mathscr{D}$ una familia numerable de semi-métricas sobre X que separa puntos, sea $\phi : \mathbb{N}_0 \to \mathscr{D}$ una función biyectiva, para cada $n \in \mathbb{N}_0$ escribamos $d_n := \phi(n)$, y sea $(a_n)_{n\geq 0}$ una sucesión de números reales positivos tal que la serie $\sum_{n=0}^{\infty} a_n$ converge.*

(i) Si x e y son dos elementos de X, entonces la serie

$$\sum_{n=0}^{\infty} a_n \frac{d_n(x, y)}{1 + d_n(x, y)} \tag{5.8}$$

converge.

(ii) La función

$$d : (x, y) \in X \times X \mapsto \sum_{n=0}^{n} a_n \frac{d_n(x, y)}{1 + d_n(x, y)} \in \mathbb{R}$$

es una métrica sobre el conjunto X.

Demostración. Si x e y son dos elementos de X, entonces para todo entero no negativo n es

$$0 \leq a_n \frac{d_n(x, y)}{1 + d_n(x, y)} \leq a_n.$$

Esto nos dice que la serie (5.9) del enunciado tiene sus términos no negativos y que está acotada término a término por la serie $\sum_{n=0}^{\infty} a_n$, así que converge.

Mostremos ahora que la función d del enunciado es una métrica sobre el conjunto X.

- Si x es un elemento de X, entonces para todo entero no negativo n tenemos que $d_n(x, x) = 0$,

ya que d_n es una semi-métrica sobre X, y, por lo tanto, es

$$d(x,x) = \sum_{n=0}^{n} a_n \frac{d_n(x,x)}{1+d_n(x,x)} = 0,$$

simplemente porque todos los términos de la serie son nulos. Sean, por otro lado, x e y dos elementos distintos de X. Como la familia $\mathscr{D}$ de semi-métricas separa puntos, hay un elemento δ de $\mathscr{D}$ tal que $\delta(x,y) > 0$ y, como la función ϕ es biyectiva, hay un entero no negativo m tal que $\phi(m) = \delta$. El número a_m es positivo, así que tenemos entonces que

$$0 < a_m \frac{\delta(x,y)}{1+\delta(x,y)} = a_m \frac{\phi(m)(x,y)}{1+\phi(m)(x,y)} \leq \sum_{n=0}^{\infty} a_n \frac{d_n(x,y)}{1+d_n(x,y)} = d(x,y).$$

Esto implica que la función d satisface la condición $(\mathbf{M}_1)$.

- Si x e y son dos elementos de X, entonces es $d_n(x,y) = d_n(y,x)$ para todo entero no negativo n, así que

$$d(x,y) = \sum_{n=0}^{n} a_n \frac{d_n(x,y)}{1+d_n(x,y)} = \sum_{n=0}^{n} a_n \frac{d_n(y,x)}{1+d_n(y,x)} = d(y,x).$$

- Sean x, y y z tres elementos de X. Si α, β y γ son números reales no negativos tales que $\gamma \leq \alpha + \beta$, entonces

$$\frac{\alpha}{1+\alpha} + \frac{\beta}{1+\beta} - \frac{\gamma}{1+\gamma} = \frac{(\alpha+\beta-\gamma)+\alpha\beta(2+\gamma)}{(1+\alpha)(1+\beta)(1+\gamma)} \geq 0,$$

así que

$$\frac{\gamma}{1+\gamma} \leq \frac{\alpha}{1+\alpha} + \frac{\beta}{1+\beta}.$$

Usando esto, vemos que si n es un entero no negativo, entonces

$$\frac{d_n(x,y)}{1+d_n(x,y)} \leq \frac{d_n(x,z)}{1+d_n(x,z)} + \frac{d_n(z,y)}{1+d_n(z,y)},$$

ya que los tres números reales $d_n(x,y)$, $d_n(x,z)$ y $d_n(z,y)$ son no negativos y tales que $d_n(x,y) \leq d_n(x,z) + d_n(z,y)$. Esto implica, a su vez, ya que $a_n > 0$ para todo $n \in \mathbb{N}_0$, que

$$\begin{aligned} d(x,y) &= \sum_{n=0}^{n} a_n \frac{d_n(x,y)}{1+d_n(x,y)} \leq \sum_{n=0}^{n} a_n \left(\frac{d_n(x,y)}{1+d_n(x,y)} + \frac{d_n(x,y)}{1+d_n(x,y)} \right) \\ &= \sum_{n=0}^{n} a_n \frac{d_n(x,y)}{1+d_n(x,y)} + \sum_{n=0}^{n} a_n \frac{d_n(x,y)}{1+d_n(x,y)} \leq d(x,y) + d(y,z). \end{aligned}$$

Vemos así que la función d satisface la desigualdad triangular de la condición $(\mathbf{M}_3)$.

Con todo esto podemos concluir que la función d es una métrica sobre el conjunto X. □

Como podemos obtener semi-métricas de semi-normas, la construcción de esta proposición tiene como consecuencia inmediata que podemos construir métricas a partir de una familia de semi-normas sobre un espacio vectorial. Para ello esa familia tiene que satisfacer la siguiente condición.

Definición 5.2.5.3. Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, y sea V un espacio vectorial sobre $\Bbbk$. Decimos que un conjunto $\mathscr{N}$ de semi-normas sobre V ***separa puntos*** si para todo vector no nulo x de V hay un elemento $\|-\|$ de $\mathscr{N}$ tal que $\|x\| \neq 0$.

El resultado que tenemos en mente es el siguiente:

Corolario 5.2.5.4. *Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, sea V un $\Bbbk$-espacio vectorial, y sea $\mathscr{N}$ una familia numerable de semi-normas sobre V que separa puntos. Sea $\phi : \mathbb{N}_0 \to \mathscr{N}$ una función biyectiva, para cada $n \in \mathbb{N}_0$ escribamos $\|-\|_n \coloneqq \phi(n)$, y sea $(a_n)_{n\geq 0}$ una sucesión de números reales positivos tal que la serie $\sum_{n=0}^{\infty} a_n$ converge.*

(*i*) *Si x e y son dos elementos de X, entonces la serie*

$$\sum_{n=0}^{\infty} a_n \frac{\|x-y\|_n}{1+\|x-y\|_n} \tag{5.9}$$

converge.

(*ii*) *La función*

$$d : (x,y) \in X \times X \mapsto \sum_{n=0}^{\infty} a_n \frac{\|x-y\|_n}{1+\|x-y\|_n} \in \mathbb{R}$$

es una métrica sobre V que es invariante por translaciones.

Esta construcción fue descubierta por René Maurice Fréchet y aparece descripta por primera vez en su tesis doctoral de 1906. Notemos que a partir de una familia de semi-normas obtenemos una métrica, no una norma.

Demostración. Todas las afirmaciones del corolario son consecuencia inmediata de la proposición salvo que la que dice que la distancia d es invariante por translaciones, y esta es evidente. □

Veamos dos ejemplos de esta construcción, dejando los detalles a cargo del lector.

Ejercicio 5.2.5.5. Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, y sea $C_{\Bbbk}(\mathbb{R})$ el espacio vectorial sobre $\Bbbk$ de todas las funciones continuas $\mathbb{R} \to \Bbbk$, y para cada entero positivo n consideremos la función $\|-\|_n : C_{\Bbbk}(\mathbb{R}) \to \mathbb{R}$ que en cada elemento f de $C_{\Bbbk}(\mathbb{R})$ toma el valor

$$\|f\|_n \coloneqq \sup_{x\in[-n,n]} |f(x)|.$$

Notemos que esto tiene sentido, ya que la función f es acotada en el intervalo $[-n, n]$ porque es continua y ese intervalo es compacto.

Pruebe que el conjunto $\mathscr{D} \coloneqq \{\|-\|_n : n \in \mathbb{N}\}$ es un conjunto numerable de semi-normas sobre $C_{\mathbb{k}}(\mathbb{R})$ que separa puntos, y concluya que hay una métrica d sobre $C_{\mathbb{k}}(\mathbb{R})$ tal que

$$d(f,g) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{\sup_{x\in[-n,n]}|f(x)-g(x)|}{1+\sup_{x\in[-n,n]}|f(x)-g(x)|}$$

para cada elección de f y de g en $C_{\mathbb{k}}(\mathbb{R})$ y que se trata de una métrica invariante por translaciones. Pruebe además que está métrica no está inducida por ninguna norma sobre el espacio vectorial $C_{\mathbb{k}}(\mathbb{R})$.

Ejercicio 5.2.5.6. Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, y sea $C^{\infty}(\mathbb{R})$ el espacio vectorial sobre $\mathbb{k}$ de todas las funciones $\mathbb{R} \to \mathbb{k}$ que son infinitamente diferenciables, y para cada par de enteros no negativos m y n consideremos la función $\|-\|_{m,n} : C_{\mathbb{k}}(\mathbb{R}) \to \mathbb{R}$ que en cada elemento f de $C_{\mathbb{k}}(\mathbb{R})$ toma el valor

$$\|f\|_n \coloneqq \sup_{x\in[-n-1,n+1]} |f^{(m)}(x)|.$$

Notemos que, como en el ejercicio anterior, esto tiene sentido, ya que la función $f^{(m)}$ es acotada en el intervalo $[-n, n]$ porque es continua y ese intervalo es compacto.

Pruebe que el conjunto $\mathscr{D} \coloneqq \{\|-\|_{m,n} : m, n \in \mathbb{N}\}$ es un conjunto numerable de semi-normas sobre $C_{\mathbb{k}}(\mathbb{R})$ que separa puntos, y concluya que hay una métrica d sobre $C_{\mathbb{k}}(\mathbb{R})$ que es invariante por translaciones y tal que

$$d(f,g) = \sum_{m,n=1}^{\infty} \frac{1}{2^{m+n}} \frac{\sup_{x\in[-n-1,n+1]}|f^{(m)}(x)-g^{(m)}(x)|}{1+\sup_{x\in[-n-1,n+1]}|f^{(m)}(x)-g^{(m)}(x)|}$$

para cada elección de f y de g en $C_{\mathbb{k}}(\mathbb{R})$.

La construcción del siguiente ejemplo es muy similar a las de los dos ejemplos anteriores y es de uso muy frecuente en teoría de grupos.

Ejemplo 5.2.5.7. Sea G un grupo abeliano, sea $\mathscr{H}$ es un conjunto numerable de subgrupos de G, y para cada elemento H de $\mathscr{H}$ sea d_H una métrica sobre el grupo cociente G/H que es invariante por translaciones. Sea $\phi : \mathbb{N}_0 \to \mathscr{H}$ una biyección y para cada $n \in \mathbb{N}_0$ escribamos $H_n \coloneqq \phi(n)$. Sea finalmente $(a_n)_{n\geq 0}$ una sucesión de números reales positivos tal que la serie $\sum_{n=0}^{\infty} a_n$ converge.

Para cada $n \in \mathbb{N}_0$ la función

$$d_n : (x,y) \in G \times G \mapsto d_{H_n}(x + H_n, y + H_n) \in \mathbb{R},$$

es una semi-métrica sobre G. Podemos considerar, entonces, como en la Proposición 5.2.5.2, la

función

$$d : (x, y) \in G \times G \mapsto \sum_{n=0}^{\infty} a_n \frac{d_n(x, y)}{1 + d_n(x, y)} \in \mathbb{R},$$

ya que la serie converge por las mismas razones que allí, y ver de la misma manera que se trata de una semi-métrica sobre G que es invariante por translaciones. Más aún, cuando

$$\bigcap_{H \in \mathscr{H}} H = 0$$

esta semi-métrica es una métrica. En efecto, si ese es el caso y x e y son dos elementos distintos de G, entonces hay un elemento n de $\mathbb{N}_0$ tal que $x - y \notin H_n$ y, por lo tanto, es $d(x, y) > 0$, ya que

$$d_n(x, y) = d_{H_n}(x + H_n, y + H_n) = d_{H_n}(x - y + H_n, 0 + H_n) > 0.$$

5.2.6. Productos cartesianos

Entre las varias construcciones que hay que nos permiten obtener espacios métricos a partir de otros, el producto cartesiano es una de las más útiles.

Proposición 5.2.6.1. *Sean (X, d_X) e (Y, d_Y) dos espacios métricos. La función*

$$d : ((x, y), (x', y')) \in (X \times Y) \times (X \times Y) \mapsto \max\{d_X(x, x'), d_Y(y, y')\} \in \mathbb{R}$$

es una métrica sobre el conjunto $X \times Y$.

Demostración. Mostremos que la función d satisface las tres condiciones de la Definición 5.1.1.

- Si (x, y) es un elemento de $X \times Y$, entonces

  $$d((x, y), (x, y)) = \max\{d_X(x, x), d_Y(y, y)\} = \max\{0, 0\} = 0.$$

  Por otro lado, si (x, y) y (x', y') son dos elementos de $X \times Y$ tales que $d((x, y), (x', y')) = 0$, entonces

  $$0 \leq d_X(x, x') \leq \max\{d_X(x, x'), d_Y(y, y')\} = d((x, y), (x', y')) = 0$$

  y

  $$0 \leq d_Y(y, y') \leq \max\{d_X(x, x'), d_Y(y, y')\} = d((x, y), (x', y')) = 0,$$

  de manera que $d_X(x, x') = 0$ y $d_Y(y, y') = 0$ y, por lo tanto, $x = x'$, $y = y'$ y, finalmente, $(x, y) = (x'y')$.

- Si (x, y) y (x', y') son dos elementos de $X \times Y$, entonces

$$\begin{aligned} d((x, y), (x', y')) &= \max\{d_X(x, x'), d_Y(y, y')\} = \max\{d_X(x', x), d_Y(y', y)\} \\ &= d((x', y'), (x, y)). \end{aligned}$$

- Sean (x, y), (x', y'), y (x'', y'') tres elementos de $X \times Y$. Como

$$\begin{aligned} d_X(x, x'') &\leq d_X(x, x') + d_X(x', x'') \\ &\leq \max\{d_X(x, x') + d_X(x', x''), d_Y(y, y') + d_Y(y', y'')\} \end{aligned}$$

y

$$\begin{aligned} d_X(y, y'') &\leq d_X(y, y') + d_X(y', y'') \\ &\leq \max\{d_X(x, x') + d_X(x', x''), d_Y(y, y') + d_Y(y', y'')\} \end{aligned}$$

tenemos que

$$\begin{aligned} d((x, y), (x'', y'')) &= \max\{d_X(x, x''), d_Y(y, y'')\} \\ &\leq \max\{d_X(x, x') + d_X(x', x''), d_Y(y, y') + d_Y(y', y'')\}. \end{aligned} \tag{5.10}$$

De manera similar, como

$$\begin{aligned} &d_X(x, x') + d_X(x', x'') \\ &\quad \leq \max\{d_X(x, x'), d_Y(y, y')\} + \max\{d_X(x', x''), d_Y(y', y'')\} \end{aligned}$$

y

$$\begin{aligned} &d_Y(y, y') + d_Y(y', y'') \\ &\quad \leq \max\{d_X(x, x'), d_Y(y, y')\} + \max\{d_X(x', x''), d_Y(y', y'')\}, \end{aligned}$$

tenemos que

$$\begin{aligned} &\max\{d_X(x, x') + d_X(x', x''), d_Y(y, y') + d_Y(y', y'')\} \\ &\quad \leq \max\{d_X(x, x'), d_Y(y, y')\} + \max\{d_X(x', x''), d_Y(y', y'')\} \\ &\quad = d((x, y), (x', y')) + d((x', y'), (x'', y'')). \end{aligned}$$

Esta desigualdad y la de (5.10) nos dicen que

$$d((x, y), (x'', y'')) \leq d((x, y), (x', y')) + d((x', y'), (x'', y'')).$$

Esto completa la demostración. □

Podemos extender la construcción de esta proposición a productos cartesianos de más factores. Si n es un entero no negativo y $(X_1, d_1), (X_2, d_2), \ldots, (X_N, d_n)$ son espacios métricos, entonces el

conjunto $P \coloneqq X_1 \times X_2 \times \cdots \times X_n$ es un espacio métrico con respecto a la función $d : P \times P \to \mathbb{R}$ que tiene

$$d(x, y) = \max\{d_1(x_1, y_1), d_2(x_2, y_2), \cdots, d_n(x_n, y_n)\}$$

para cada par de elementos $x = (x_1, x_2, \ldots, x_n)$ e $y = (y_1, y_2, \ldots, y_n)$ de P. Desde ahora, siempre que tengamos un producto cartesiano finito de espacios métricos lo consideraremos como un espacio métrico con respecto a esta métrica específica, a la que llamamos la ***métrica producto***.

Ejemplo 5.2.6.2. Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, y sea n un entero no negativo. La métrica producto sobre el producto cartesiano $\mathbb{k}^n = \mathbb{k} \times \mathbb{k} \times \cdots \times \mathbb{k}$ que se obtiene de poner en cada uno de los n factores la métrica euclídea no es otra que la métrica que se obtiene de la norma $\|-\|_\infty$ de ese espacio vectorial.

Hay muchas otras métricas que podemos considerar sobre un producto cartesiano. El siguiente ejercicio presenta algunas de ellas.

Ejercicio 5.2.6.3. Sea n es un entero no negativo y sean (X_1, d_1), (X_2, d_2), …, (X_N, d_n) espacios métricos, y consideremos el conjunto $X \coloneqq X_1 \times X_2 \times \cdots \times X_n$. Muestre que para cada elemento p de $[1, +\infty)$ la función $d_p : X \times X \to \mathbb{R}$ que tiene

$$d_p(x, y) = \left(d_1(x_1, y_1)^p + d_2(x_2, y_2)^p + \cdots + d_n(x_n, y_n)^p\right)^{1/p}$$

para cada par de elementos $x = (x_1, x_2, \ldots, x_n)$ e $y = (y_1, y_2, \ldots, y_n)$ de X es una métrica sobre X.

Hay distintas maneras de extender esta construcción de métricas producto a los productos cartesianos infinitos, pero todas solo se aplican en situaciones especiales. Veamos una.

Proposición 5.2.6.4. *Sea $(X_n)_{n\geq 0}$ una sucesión de espacios métricos, para cada $n \in \mathbb{N}_0$ escribamos d_n a la métrica de X_n, y consideremos el conjunto $X \coloneqq \prod_{n=0}^{\infty} X_n$ de todas las sucesiones $(x_n)_{n\geq 0}$ que para cada $n \in \mathbb{N}_0$ tienen $x_n \in X_n$.*

(i) Si $x = (x_n)_{n\geq 0}$ e $y = (y_n)_{n\geq 0}$ son dos elementos de X, entonces la serie

$$\sum_{n=0}^{\infty} \frac{1}{2^n} \frac{d_n(x_n, y_n)}{1 + d_n(x_n, y_n)}$$

converge.

(ii) La función

$$d : ((x_n)_{n\geq 0}, (y_n)_{n\geq 0}) \in X \times X \longmapsto \sum_{n=0}^{\infty} \frac{1}{2^n} \frac{d_n(x_n, y_n)}{1 + d_n(x_n, y_n)} \in \mathbb{R}$$

es una métrica sobre el conjunto X.

Demostración. Sean $x = (x_n)_{n\geq 0}$ e $y = (y_n)_{n\geq 0}$ dos elementos de X. Para cada $n \in \mathbb{N}_0$ es

$$0 \leq \frac{d_n(x_n, y_n)}{1 + d_n(x_n, y_0)} \leq 1,$$

así que la serie

$$\sum_{n=0}^{\infty} \frac{1}{2^n} \frac{d_n(x_n, y_n)}{1 + d_n(x_n, y_n)}$$

tiene todos sus términos no negativos y está acotada término a término por la serie geométrica convergente $\sum_{n=0}^{\infty} 1/2^n$, así que ella misma converge. Esto prueba la primera afirmación de la proposición.

Para probar la segunda, mostramos que la función d descripta en el enunciado satisface las condiciones de la Definición 5.1.1.

- Si $x = (x_n)_{n\geq 0}$ es un elemento de X, entonces $d_x(x_n, x_n) = 0$ para todo $n \in \mathbb{N}_0$, así que

  $$d(x, x) = \sum_{n=0}^{\infty} \frac{1}{2^n} \frac{d_n(x_n, x_n)}{1 + d_n(x_n, x_n)} = 0.$$

  Por otro lado, supongamos que $x = (x_n)_{n\geq 0}$ e $y = (y_n)_{n\geq 0}$ son dos elementos de X tales que $d(x, y) = 0$. Si n es un elemento cualquiera de $\mathbb{N}_0$, entonces el hecho de que la serie que define a $d(x, y)$ tenga todos sus términos no negativos implica que

  $$0 \leq \frac{d_n(x_n, y_n)}{1 + d_n(x_n, y_n)} \leq \sum_{n=0}^{\infty} \frac{1}{2^n} \frac{d_n(x_n, y_n)}{1 + d_n(x_n, y_n)} = d(x, y) = 0,$$

  así que $d_n(x_n, y_n) = 0$ y, por lo tanto, $x_n = y_n$. Vemos así que, de hecho, $x = y$.
- Si $x = (x_n)_{n\geq 0}$ e $y = (y_n)_{n\geq 0}$ son dos elementos de X, entonces

  $$d(x, y) = \sum_{n=0}^{\infty} \frac{1}{2^n} \frac{d_n(x_n, y_n)}{1 + d_n(x_n, y_n)} = \sum_{n=0}^{\infty} \frac{1}{2^n} \frac{d_n(y_n, x_n)}{1 + d_n(y_n, x_n)} = d(y, x).$$

- Sean, finalmente, $x = (x_n)_{n\geq 0}$, $y = (y_n)_{n\geq 0}$, y $z = (z_n)_{n\geq 0}$ tres elementos de X. Si a, b y c son tres números reales, entonces

  $$\frac{a}{a+1} + \frac{b}{b+1} - \frac{c}{c+1} = \frac{ab(c+2) + (a+b-c)}{(a+1)(b+1)(c+1)},$$

  y esto hace evidente que

  $$\frac{c}{c+1} \leq \frac{a}{a+1} + \frac{b}{b+1}$$

cuando $a \geq 0$, $b \geq 0$, $c \geq 0$ y $c \leq a + b$. En particular, para cada entero no negativo n tenemos que

$$\frac{d_n(x_n, y_n)}{1 + d_n(x_n, y_n)} \leq \frac{d_n(x_n, z_n)}{1 + d_n(x_n, z_n)} + \frac{d_n(z_n, y_n)}{1 + d_n(z_n, y_n)},$$

porque las distancias $d_n(x_n, y_n)$, $d_n(x_n, z_n)$, $d_n(z_n, y_n)$ son no negativas y, por supuesto, $d_n(x_n, y_n) \leq d_n(x_n, z_n) + d_n(z_n, y_n)$. Esto implica que

$$\begin{aligned} d(x, y) &= \sum_{n=0}^{\infty} \frac{1}{2^n} \frac{d_n(x_n, y_n)}{1 + d_n(x_n, y_n)} \leq \sum_{n=0}^{\infty} \frac{1}{2^n} \frac{d_n(x_n, z_n)}{1 + d_n(x_n, z_n)} + \sum_{n=0}^{\infty} \frac{1}{2^n} \frac{d_n(z_n, y_n)}{1 + d_n(z_n, y_n)} \\ &= d(x, z) + d(z, y). \end{aligned}$$

□

Como con los productos finitos, hay muchas formas alternativas de poner métricas sobre un producto numerable de espacios métricos.

5.2.7. Métricas discretas

La siguiente proposición nos da un ejemplo un poco extremo pero sin embargo útil de métrica.

Proposición 5.2.7.1. *Sea X un conjunto. La función $d : X \times X \to \mathbb{R}$ que en cada elemento (x, y) de $X \times X$ toma el valor*

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

es una métrica sobre X.

Llamamos a esa métrica la ***métrica discreta*** sobre el conjunto X.

Demostración. Que la función d del enunciado satisface las condiciones $(\mathbf{M}_1)$ y $(\mathbf{M}_2)$ de la Definición 5.1.1 es evidente. Sean, por otro lado, x, y y z tres elementos de X. Si $x = y$, entonces es claro que $d(x, y) = 0 \leq d(x, z) + d(z, y)$. Si, en cambio, $x \neq y$, entonces z es igual a lo sumo a un elemento del conjunto $\{x, y\}$, así que tenemos que $d(x, y) = 1 = d(x, z) + d(z, y)$. Esto nos dice que la función d también satisface la condición $(\mathbf{M}_3)$. □

Ejercicio 5.2.7.2. Muestre que la función $d : \mathbb{N} \times \mathbb{N} \to \mathbb{R}$ que en cada elemento (m, n) de $\mathbb{N} \times \mathbb{N}$ toma el valor

$$d(m, n) = \begin{cases} 0 & \text{si } m = n; \\ 1 + \frac{1}{m} + \frac{1}{m} & \text{si } m \neq n \end{cases}$$

es una distancia sobre $\mathbb{N}$.

Queremos presentar una generalización útil de la métrica discreta sobre un conjunto. El primer paso para ello es dar la siguiente definición.

Definición 5.2.7.3. Sea X un conjunto y sea Γ una relación en X.

- Una función $\gamma : [\![0, n]\!] \to X$ definida en un conjunto de la forma $[\![0, n]\!]$ para algún $n \in \mathbb{N}_0$ es un ***camino*** en Γ si para cada $i \in [\![n]\!]$ es $(f(i-1), f(i)) \in X$, y cuando ese es el caso decimos que el camino γ va de $f(0)$ a $f(n)$ y que tiene ***longitud*** n, y escribimos $\operatorname{len}(\gamma) = n$.
- Decimos que la relación Γ es ***conexa*** si para cada elección de x e y en X hay un camino en Γ que va de x a y.

Así, por ejemplo, la relación total $X \times X$ sobre un conjunto X es claramente conexa.

Proposición 5.2.7.4. *Sea X un conjunto, y sea Γ una relación simétrica y conexa sobre X. La función $d : X \times X \to \mathbb{R}$ que en cada elemento (x, y) de $X \times X$ toma el valor*

$$d(x, y) = \min\{\operatorname{len}(\gamma) : \gamma \text{ es un camino en } \Gamma \text{ de } x \text{ a } y\} \tag{5.11}$$

es una métrica sobre X.

Llamamos a esta métrica d la ***métrica de los caminos*** sobre Γ. Notemos que el conjunto cuyo mínimo estamos tomando en (5.11) no es vacío precisamente porque la relación Γ es conexa. Notemos que esta proposición generaliza a la Proposición 5.2.7.1: si Γ es la relación total $X \times X$, que ciertamente es conexa, entonces la métrica de los caminos sobre Γ es la métrica discreta sobre X.

Demostración. Tenemos que verificar que la función d descripta en el enunciado de la proposición satisface las condiciones de la Definición 5.1.1.

- Si x es un elemento de X, entonces la función $\gamma : [\![0, 0]\!] \to X$ tal que $\gamma(0) = x$ es un camino en Γ de x a x de longitud 0, así que $d(x, x) = 0$. Sean, por otro lado, lado, x e y dos elementos de X tales que $d(x, y) = 0$, de manera que hay un camino $\gamma : [\![0, 0]\!] \to X$ en Γ que va de x a y: esto implica que $x = \gamma(0)$ e $y = \gamma(0)$, así que, por supuesto, es $x = y$.
- Sean x e y dos elementos de X, y escribamos $\Gamma(x, y)$ y $\Gamma(y, x)$ a los conjuntos de todos los caminos en Γ de x a y y de y a x, respectivamente. Como la relación Γ es simétrica, es fácil verificar que si $\gamma : [\![0, n]\!] \to X$ es un elemento de $\Gamma(x, y)$, entonces la función $\gamma^- : i \in [\![0, n]\!] \mapsto \gamma(n - i) \in X$ es un elemento de $\Gamma(y, x)$, y esto implica que tenemos dos funciones

  $$\gamma \in \Gamma(x, y) \mapsto \gamma^- \in \Gamma(y, x), \qquad \gamma \in \Gamma(y, x) \mapsto \gamma^- \in \Gamma(x, y).$$

  Como $(\gamma^-)^- = \gamma$ cualquiera sea el camino γ, es claro que estas dos funciones son mutuamente inversas y, por otro lado, es evidente que $\operatorname{len}(\gamma^-) = \operatorname{len}(\gamma)$ para todo camino γ. Todo

esto implica que

$$\begin{aligned} d(x,y) &= \min\{\mathrm{len}(\gamma) : \gamma \in \Gamma(x,y)\} = \min\{\mathrm{len}(\gamma^-) : \gamma \in \Gamma(y,x)\} \\ &= \min\{\mathrm{len}(\gamma) : \gamma \in \Gamma(y,x)\} = d(y,x). \end{aligned}$$

- Sean, finalmente, x, y y z tres elementos de X, y sean $m \coloneqq d(x,z)$ y $n \coloneqq d(z,y)$. Hay entonces caminos $\gamma : [\![0,m]\!] \to X$ y $\sigma : [\![0,n]\!] \to X$ en Γ que van de x a z y de z a y de longitudes m y n, respectivamente. Hay una función $\tau : [\![0,m+n]\!] \to X$ que en cada elemento i de $[\![0,m+n]\!]$ toma el valor

$$\tau(i) = \begin{cases} \gamma(i) & \text{si } 0 \le i \le m; \\ \sigma(i-m) & \text{si } m < i \le m+n; \end{cases}$$

  Esta función es un camino en Γ. En efecto, si $i \in [\![m+n]\!]$, entonces

  - o bien $i \le m$, y en ese caso $(\tau(i-1), \tau(i)) = (\gamma(i-1), \gamma(i)) \in \Gamma$,
  - o bien $i = m+1$, y en ese caso $(\tau(m), \tau(m+1)) = (\gamma(m), \tau(1)) = (\tau(0), \tau(1)) \in \Gamma$,
  - o bien $i > m+1$, y en ese caso $(\tau(i-1), \tau(i)) = (\sigma(i-1-m), \sigma(i-m)) \in \Gamma$.

  Como τ es un camino en Γ de x a y de longitud $m+n$, es $d(x,y) \le m+n = d(x,z)+d(z,y)$.

Todo esto prueba que la función d es una métrica sobre X. □

Estas métricas de caminos son extremadamente importantes en muchos contextos — particularmente en combinatoria, en teoría de probabilidades, y en teoría de grupos. Veamos una de las maneras en la que aparece en esta última.

Ejemplo 5.2.7.5. Sea G un grupo, sea S un subconjunto de G que genera a G y que es ***simétrico***, esto es, tal que $S = \{s^{-1} : s \in S\}$, y consideremos en G la relación

$$\Gamma_S \coloneqq \{(g, gs) : g \in G, s \in S\}.$$

Se trata de una relación simétrica. En efecto, si g y h son dos elementos de G tales que $(g,h) \in \Gamma_S$, de manera que hay un elemento s de S tal que $h = gs$, entonces sabemos que $s^{-1} \in S$ porque el conjunto S es simétrico y, como $g = hs^{-1}$, esto implica que $(h,g) \in \Gamma_S$.

Sean ahora g y h dos elementos de G, y consideremos el conjunto $S(g,h)$ de todas las secuencias finitas $(s_1, s_2, \ldots, s_n)$ tales que $g^{-1}h = s_1 s_2 \cdots s_n$. Si $c = (s_1, s_2, \ldots, s_n)$ es un elemento de $S(g,h)$, entonces la función

$$\gamma_c : i \in [\![0,n]\!] \mapsto g s_1 s_2 \cdots s_i \in G,$$

es tal que $(\gamma_c(i-1), \gamma_c(i)) = (\gamma_c(i-1), \gamma_c(i-1)s_i) \in \Gamma_S$ para cada $i \in [\![n]\!]$, así que es un camino en Γ_S de $\gamma_c(0) = g$ a $\gamma_c(n) = g s_1 s_2 \cdots s_n = h$. Si escribimos $\Gamma_S(g,h)$ al conjunto de caminos en Γ_S

de g a h, obtenemos de esta forma una función

$$c \in S(g,h) \mapsto \gamma_c \in \Gamma_S(g,h). \tag{5.12}$$

Como el conjunto S genera a G, sabemos que el conjunto $S(g,h)$ no es vacío y esto implica que tampoco lo es el conjunto $\Gamma_S(g,h)$. Vemos con esto que la relación Γ_S es conexa y que, por lo tanto, podemos considerar la métrica $d : G \times G \to \mathbb{R}$ de los caminos en Γ_S, a la que llamamos normalmente la ***métrica de las palabras*** en G con respecto al conjunto generador simétrico S.

Mostremos que función de (5.12) es biyectiva.

- Sean $c = (s_1, s_2, \ldots, s_n)$ y $c' = (s'_1, s'_2, \ldots, s'_{n'})$ dos elementos de $S(g,h)$ tales que las funciones $\gamma_c : [\![0,n]\!] \to G$ y $\gamma_{c'} : [\![0,n']\!] \to G$ son iguales. Esto implica, en primer lugar, que $n = n'$ y, en segundo lugar, que para cada $i \in [\![n]\!]$ es

$$\begin{aligned} s_i &= (gs_1s_2\cdots s_i)^{-1}(gs_1s_2\cdots s_i) = \gamma_c(i-1)^{-1}\gamma_c(i) \\ &= \gamma_{c'}(i-1)^{-1}\gamma_{c'}(i) = (gs'_1s'_2\cdots s'_{i-1})^{-1}(gs'_1s'_2\cdots s'_i) = s'_i, \end{aligned}$$

  de manera que $c = c'$.
- Sea $\gamma : [\![0,n]\!] \to G$ un elemento del conjunto $\Gamma_S(g,h)$, de manera que $\gamma(0) = g$, $\gamma(n) = h$, y para cada $i \in [\![n]\!]$ el par $(\gamma(i-1), \gamma(i))$ pertenece a Γ_S, esto es, hay un elemento s_i de S tal que $\gamma(i) = \gamma(i-1)s_i$, y consideremos la n-upla $c \coloneqq (s_1, s_2, \ldots, s_n)$. Es claro que $\gamma(i) = \gamma(0)s_1s_2\cdots s_i$ para cada $i \in [\![n]\!]$, así $g^{-1}h = \gamma(0)^{-1}\gamma(n) = s_1s_2\cdots s_n$ y que $\gamma = \gamma_c$.

Escribamos ahora e al elemento neutro de G y sea h un elemento cualesquiera de G. Que la función de (5.12) sea biyectiva implica inmediatamente que la distancia $d(e,h)$, que es la menor longitud de un elemento de $\Gamma_S(e,h)$, coincide con la menor longitud de un elemento de $S(e,h)$, que no es otra cosa que la menor longitud de un producto $s_1s_2\cdots s_n$ con factores en el conjunto S e igual a h.

A partir del grupo G y el conjunto generador simétrico S podemos construir el llamado ***diagrama de Cayley*** del par (G,S). Es el grafo que tiene un vértice por cada elemento de G y en el que dos vértices, que corresponden a elementos g y h de G, están conectados por un arco exactamente cuando el par (g,h) pertenece a la relación Γ_S. Así, por ejemplo, si el grupo G es S_4, el grupo simétrico de grado 4, y el conjunto S es el conjunto de transposiciones $\{(12),(23),(34)\}$, que ciertamente genera a S_4 y es simétrico, entonces el diagrama de Cayley de (G,S) es que dibujamos en la Figura 5.2 de la página 157 La distancia entre dos elementos de G es la longitud del camino más corto en este grafo de uno al otro. Por ejemplo, la distancia de la permutación 4312 a la permutación 2134 es 5.

El estudio de estas métricas es una parte fundamental de la llamada *teoría geométrica de grupos*, subdisciplina de la teoría de grupos que surgió a fines de la década de los 1980 a partir del trabajo *Hyperbolic groups* [Gro1987] de Mikhail Gromov.

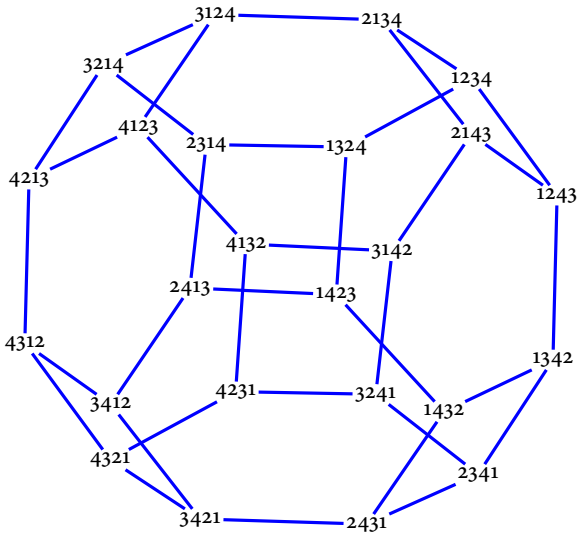


Figura 5.2. El diagrama de Cayley del grupo simétrico de grado 4 con respecto a su conjunto generados $\{(12), (23), (34)\}$. Notemos que el dibujo es lo que queda de un octaedro regular si truncamos cada uno de sus ocho vértices.

Ejercicio 5.2.7.6. Muestre que si G es un grupo, S es un subconjunto generador simétrico de G, y d_S es la métrica de las palabras del Ejemplo 5.2.7.5, entonces para toda elección de g, h, y k en G vale que $d_S(g \cdot h, g \cdot k) = d_S(h, k)$, de manera que esta métrica es ***invariante por translaciones a izquierda***, en el sentido evidente. ¿Es también invariante por translaciones a derecha?

5.2.8. Métricas de Hamming

La siguiente proposición nos da un ejemplo muy sencillo de una métrica que tiene un rol fundamental en la teoría de la información y en la teoría de códigos.

Proposición 5.2.8.1. *Sea A un conjunto y sea n un entero. La función*

$$d : (x, y) \in A^n \times A^n \mapsto |\{i \in [\![n]\!] : x_i \neq y_i\}| \in \mathbb{R}$$

es una métrica sobre el conjunto A^n.

Llamamos a la métrica d la ***métrica de Hamming***, por Richard Hamming, que la introdujo en su artículo [Ham1950], que es el principio de la teoría de los códigos que detectan y corrigen errores.

Demostración. Como siempre, tenemos que mostrar que la función d satisface las condiciones de la Definición 5.1.1.

- Si x es un elemento de A^n, entonces el conjunto $\{i \in [\![n]\!] : x_i \neq x_i\}$ es vacío y su cardinal es nulo, de manera que $d(x,x) = 0$. Por otro lado, si x e y son elementos de A^n tales que $d(x,y) = 0$, entonces el conjunto $\{i \in [\![n]\!] : x_i \neq y_i\}$ es vacío, de manera que $x_i = y_i$ para todo $i \in [\![n]\!]$ y, por lo tanto, $x = y$.
- Si x e y son dos elementos de A^n, entonces es evidente que los conjuntos $\{i \in [\![n]\!] : x_i \neq y_i\}$ y $\{i \in [\![n]\!] : y_i \neq x_i\}$ son iguales, así que sus cardinales $d(x,y)$ y $d(y,x)$ son iguales.
- Sean x, y y z tres elementos de A^n. Si i es un elemento de $[\![n]\!]$ tal que $x_i \neq y_i$, entonces z_i es igual a como mucho un elemento de $\{x_i, y_i\}$, de manera que $x_i \neq z_i$ o $z_i \neq y_i$. Esto nos dice que

  $$\{i \in [\![n]\!] : x_i \neq y_i\} \subseteq \{i \in [\![n]\!] : x_i \neq z_i\}\{i \in [\![n]\!] : z_i \neq y_i\}$$

  y, por lo tanto, que

  $$\begin{aligned} d(x,y) &= |\{i \in [\![n]\!] : x_i \neq y_i\}| \\ &\leq |\{i \in [\![n]\!] : x_i \neq z_i\} \cup \{i \in [\![n]\!] : z_i \neq y_i\}| \\ &\leq |\{i \in [\![n]\!] : x_i \neq z_i\}| + |\{i \in [\![n]\!] : z_i \neq y_i\}| \\ &= d(x,z) + d(z,y). \end{aligned}$$

La función d es, por todo esto, una métrica sobre el conjunto A^n. □

Expliquemos rápidamente cuál es la motivación de Hamming para considerar estas métricas. Alice quiere enviar un mensaje a Bob. Escribamos $\mathscr{M}$ al conjunto de los posibles mensajes que puede enviar Alice, y supongamos que se trata de un conjunto finito y no vacío. La forma en que Alice enviará el mensaje es la siguiente: elegirá un entero n, un conjunto $\mathscr{A}$, el ***alfabeto***, y una función $f : \mathscr{M} \to \mathscr{A}^n$, la ***función de codificación***, y cuando quiera enviar a Bob el mensaje m de $\mathscr{M}$ lo que hará será enviar el elemento $c \coloneqq f(m)$ de $\mathscr{A}^n$. Por supuesto, es importante que la función f sea inyectiva — en caso contrario no es posible que Bob siempre pueda deducir a partir de c cuál es el mensaje m. En esta situación llamamos a la imagen de la función f, que tiene el mismo cardinal que el conjunto de mensajes $\mathscr{M}$, el ***código***.

Ahora bien, hay una dificultad: el canal por el que Alice envía el elemento c de $\mathscr{A}^n$ correspondiente al mensaje m tiene *ruido*, esto es, Bob recibirá un elemento c' de $\mathscr{A}^n$ que no necesariamente coincide con c, sino que puede tener algunas de sus componentes cambiadas. Decimos que

- la función de codificación f ***permite detectar errores*** si Bob puede darse cuenta al recibir c' si hubo o no un error, y que
- la función de codificación f ***permite corregir errores*** si inclusive cuando el elemento c' que recibió Bob no es igual a c puede de todas maneras deducir a partir de c' cuál era el

mensaje m.

En general, una función de codificación f no tiene ninguna de estas dos propiedades, por supuesto. Por ejemplo, si la función f es sobreyectiva, entonces Bob no tiene ninguna forma de saber si hubo un error en la transmisión. De todas formas, si hacemos algunas hipótesis sobre el tipo de ruido que puede ocurrir durante la transmisión, entonces sí podemos hacer cosas útiles.

Ejemplo 5.2.8.2. Supongamos que $\mathscr{M} = \{0, 1\}$, de manera que hay solo dos mensajes, elijamos un entero positivo n, y consideremos el alfabeto $\mathscr{A} = \{0, 1\}$ y la función de codificación $f : \mathscr{M} \to \mathscr{A}^n$ tal que

$$f(0) = 00\cdots00, \qquad f(1) = 11\cdots11.$$

esto es, $f(0)$ y $f(1)$ son los elementos 'constantes' de $\mathscr{A}^n$ de entradas todas iguales a 0 y a 1, respectivamente a 0 y a 1, respectivamente. Esto significa que cuando Alice quiere enviar un 0, manda una secuencia de n ceros, y cuando quiere mandar un 1, manda una secuencia de n unos.

Si Alice usa esa función de codificación, entonces Bob puede hacer lo siguiente: si recibe una secuencia $c' = b_1 b_2 \cdots b_n$ de n bits, puede contar cuántos ceros y cuántos unos hay en ella, y decidir que el mensaje que quiso mandar Alice es el elemento de $\{0, 1\}$ que más veces aparece en c'. Así, por ejemplo, si recibe la secuencia 0010100, entonces interpretará que el mensaje de Alice es 0, y si recibe 101101, que el mensaje de Alice es 1.

Es claro que

> *si hay como mucho $n-1$ errores en c', entonces Bob puede darse cuenta de que hubo un error,*

ya que la única situación en la que Bob no puede darse cuenta del error es aquella en la que todos los n bits del código c fueron cambiados durante la transmisión. Por otro lado,

> *si hay como mucho $(n-1)/2$ errores en c', entonces Bob puede reconstruir c a partir de c'.*

En efecto, si como mucho hay $(n-1)/2$ errores, entonces el bit que aparece más veces en c' es el mismo que aparece más veces en c. Vemos así que bajo hipótesis apropiadas sobre la cantidad de errores esta función de codificación puede tanto detectar errores como corregirlos. Esto tiene, por supuesto, un costo: los mensajes codificados que manda Alice son n veces más largos que el mensaje real.

Ejemplo 5.2.8.3. Consideremos las matrices

$$G \coloneqq \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \in \mathrm{M}_{7,4}(\mathbb{F}_2) \qquad H \coloneqq \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \in \mathrm{M}_{3,7}(\mathbb{F}_2).$$

El conjunto de mensajes que consideramos aquí es $\mathscr{M} \coloneqq \mathbb{F}_2^4$, elegimos $n = 7$ y $\mathscr{A} = \mathbb{F}_2$, y la función de codificación está dada por la multiplicación matricial por la matriz G,

$$f : v \in \mathscr{M} = \mathbb{F}_2^4 \mapsto Gv \in \mathbb{F}_2^7 = \mathscr{A}^7.$$

Así, por ejemplo, si Alice quiere transmitir el mensaje $m = \texttt{1011}$ enviará los 7 bits $c = \texttt{0101011}$. Notemos que el tercero, quinto, sexo y séptimo bits de c son precisamente los de m, y que esto ocurre cualquiera sea m.

Sea ahora c' el elemento de $\mathbb{F}_2^7$ que recibe Bob. Un cálculo directo muestra que HG es la matriz nula, así que si no hubo ningún error en la transmisión y $c' = c$, entonces $Hc' = Hc = HGm = 0$, el vector nulo de $\mathbb{F}_2^3$. Supongamos ahora que en la transmisión hubo *un* error en un bit, de manera que hay un elemento i de $[\![7]\!]$ tal que $c' = c + e_i$ — aquí e_i denota al i–ésimo elemento de la base ordenada estándar de $\mathbb{F}_2^7$. En ese caso tendremos que $Hc' = Hc + He_i = He_i$: en vista de la forma de la matriz H, entonces, es claro que tendremos que $He_i \neq 0$. Más aún, como las 7 columnas de la matriz H son distintas, a partir del valor de Hc' podemos deducir el valor de i.

La conclusión de esto es que

la función de codificación f permite detectar un error de a lo sumo un bit y corregirlo.

En efecto, asumiendo que en c' hay a lo sumo un bit erróneo, Bob puede calcular Hc': si obtiene el vector nulo, entonces sabe que no hubo ningún error, y si obtiene un vector no nulo $b_1b_2b_3$ y lo interpreta como un número i entre 1 y 7 escrito en binario, entonces sabe que el i–ésimo bit de c' es el que fue cambiado durante la transmisión.

La imagen de la función de codificación f es conocida como el ***código de Hamming*** $(7, 4)$, y es el más sencillo de los llamados ***códigos de Hamming***. Nos permite codificar mensajes de 4 bits en 7 bits y detectar y corregir un error de 1 bit.

Sea $f : \mathscr{M} \to \mathscr{A}^n$ una función de codificación, sea $C \coloneqq \mathrm{img}(f)$ el código correspondiente, y sea d la métrica de Hamming sobre $\mathscr{A}^n$. Llamamos al número

$$\delta \coloneqq \mín\{d(x, y) : x, y \in C, x \neq y\}$$

la ***distancia mínima*** del código C. Claramente dos elementos de C que están a distancia de Hamming menor que δ son iguales, y esto tiene las siguientes dos consecuencias:

- Usando la función de codificación f podemos detectar hasta $\delta - 1$ errores. En efecto, si la palabra c' que recibe Bob difiere de un elemento c de C en como mucho $\delta - 1$ componentes, entonces c' no pertenece al conjunto $C \setminus \{c\}$.
- Usando la función de codificación f podemos corregir hasta $(\delta-1)/2$ errores: si un elemento c' de $\mathscr{A}^n$ se obtiene de un elemento c de C cambiando a lo sumo $(\delta - 1)/2$ componentes, entonces c es el único elemento x de C tal que $\delta(c', x) < (\delta - 1)/2$.

Como consecuencia de esto, una parte de la teoría de códigos consiste en encontrar subconjuntos C de $\mathscr{A}^n$ que tengan distancia mínima lo más grande posible y, a la vez, cardinal grande.

Ejemplo 5.2.8.4. El código que corresponde a la función de codificación del Ejemplo 5.2.8.2 es

$$C = \{00\cdots00, 11\cdots11\},$$

y es claro que su distancia mínima es n.

Ejemplo 5.2.8.5. El código que corresponde a la función de codificación del Ejemplo 5.2.8.3 es el conjunto

$$C = \{Gv : v \in \mathbb{F}_2^4\}.$$

Notemos que se trata de un subespacio de $\mathbb{F}_2^7$ de dimensión 4. Es fácil ver que si x e y son elementos de $\mathbb{F}_2^7$, entonces $d(x, y) = d(0, x - y)$, y esto tiene como consecuencia que la distancia mínima de C coincide con el número

$$\text{mín}\{d(0, x) : x \in C, x \neq 0\},$$

al que llamamos el ***peso mínimo*** de C. Claramente este número es la menor cantidad de coordenadas no nulas que tiene un elemento no nulo de C, y es más fácil calcular el peso mínimo que la distancia mínima — para encontrar el primero tenemos que calcular 16 pesos, mientras que para encontrar la segunda tenemos que calcular $\binom{16}{2} = 120$ distancias. Tabulamos los pesos de todos los elementos de C en la Tabla 5.1 de la página 162: es claro que el peso mínimo del código C es 3. Esto se corresponde con la observación que hicimos arriba de que este código permite corregir un error.

Un caso que encontramos frecuentemente es aquel en el que el alfabeto es un cuerpo finito $\mathbb{F}_q$. En ese caso, decimos que un código C de longitud n es ***lineal*** si es un subespacio de $\mathbb{F}_p^n$ y si k es su dimensión y d su distancia mínima, que se trata de un ***código*** $[n, k, d]_q$. Así, el código del Ejemplo 5.2.8.2 es un código $[n, 1, n]_2$, mientras que el código de Hamming del Ejemplo 5.2.8.3 es un código $[7, 4, 3]_2$.

x				Gx							$d(0, Gx)$
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	1	1	1	1	0	0	0	1	4
0	0	1	0	0	1	1	0	0	1	0	3
0	0	1	1	1	0	0	0	0	1	1	3
0	1	0	0	1	0	1	0	1	0	0	3
0	1	0	1	0	1	0	0	1	0	1	3
0	1	1	0	1	1	0	0	1	1	0	4
0	1	1	1	0	0	1	0	1	1	1	4
1	0	0	0	1	1	0	1	0	0	0	3
1	0	0	1	0	0	1	1	0	0	1	3
1	0	1	0	1	0	1	1	0	1	0	4
1	0	1	1	0	1	0	1	0	1	1	4
1	1	0	0	0	1	1	1	1	0	0	4
1	1	0	1	1	0	0	1	1	0	1	4
1	1	1	0	0	0	0	1	1	1	0	3
1	1	1	1	1	1	1	1	1	1	1	7

Cuadro 5.1. Los elementos de $\mathbb{F}_2^4$, los correspondientes elementos de $\mathbb{F}_2^7$ por el código de Hamming $(7, 4)$, y sus distancias al vector nulo.

Los códigos lineales son extremadamente útiles en las aplicaciones. Por ejemplo, la familia de ***códigos lineales de Reed–Solomon***, estudiados originalmente por Irving S. Reed y Gustave Solomon, está en la base del funcionamiento de los CDs, los DVDs, los discos BlueRay, los códigos QR, y gran cantidad de protocolos de comunicación satelital, y fueron usados para mejorar la comunicación con los dos satélites Voyager, que son los únicos objetos hechos por el hombre que han salido del sistema solar.

El siguiente es el ejemplo más sencillo de un código de Reed–Solomon.

Ejercicio 5.2.8.6. Sea q una potencia de un primo, sea $\mathbb{F}_q$ el cuerpo de Galois de cardinal q, sea α una raíz primitiva de $\mathbb{F}_q$, de manera que α genera el grupo multiplicativo $\mathbb{F}_q^\times$ del cuerpo $\mathbb{F}_q$, y sea k un entero menor que q. El conjunto de todos los elementos de $\mathbb{F}_q^q$ de la forma

$$(p(0), p(1), p(\alpha), p(\alpha^2), \ldots, p(\alpha^{q-2}))$$

con p un polinomio de $\mathbb{F}_q[X]$ que es o nulo o tiene grado menor que k es claramente un subespacio de $\mathbb{F}_q^n$. Pruebe que se trata de un código lineal tipo $[q, k, q-k+1]_q$.

Veamos una instancia concreta de esta construcción.

Ejemplo 5.2.8.7. Hay un elemento x en $\mathbb{F}_4$ tal que $x^2 + x + 1 = 0$, y es una raíz primitiva de $\mathbb{F}_4$. En efecto, los cuatro elementos de $\mathbb{F}_4$ son 0, $1 = x^0$, $x = x^1$, y $x^2 = x + 1$. La construcción que describimos nos da entonces un código lineal de tipo $[4, 2, 3]_4$. Los $4^2 = 16$ elementos de ese código — si escribimos 0, 1, 2 y 3 en lugar de 0, 1, x y x^2, respectivamente — son los siguientes vectores de $\mathbb{F}_4^4$:

```
0000  0123  0231  0312  1111  1032  1320  1203
2222  2301  2013  2130  3333  3210  3102  3021
```

Este código nos permite envíar 16 mensajes distintos en palabras de longitud 4 en el alfabeto $\{1, 2, 3, 4\}$ detectando hasta tres errores y corrigiendo hasta un error.

5.2.9. Ultramétricas

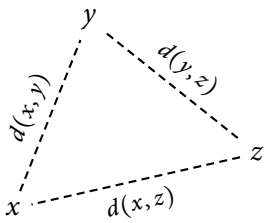
En esta subsección presentaremos algunos ejemplos muy especiales de métricas, que son, de hecho, ultramétricas en el siguiente sentido.

Definición 5.2.9.1. Una métrica $d : X \times X \to \mathbb{R}$ sobre un conjunto X es una ***ultramétrica*** si cualesquiera sean los elementos x, y y z de X vale que

$$d(x, y) \leq \max\{d(x, z), d(z, y)\}.$$

Llamamos a la desigualdad que aparece en esta definición la ***desigualdad ultramétrica*** y un espacio métrico (X, d) cuya métrica es una ultramétrica es, por supuesto, un ***espacio ultramétrico***.

Si X es un espacio métrico y x, y y z son tres puntos de X, entonces podemos imaginar que esos puntos son los vértices de un triángulo, y que las distancias entre ellos son las longitudes de los lados.



Como sabemos, la desigualdad triangular nos dice que la longitud de cada uno de los lados de ese triángulo es menor que la suma de las longitudes de los otros dos. La desigualdad ultramétrica, por su parte, impone la condición de que la longitud de cada uno de los lados sea *igual* a la más grande de las longitudes de los otros dos. Esto tiene como consecuencia que en un espacio ultramétrico «todos los triángulos son isósceles».

Lema 5.2.9.2. *Sea d una ultramétrica sobre un conjunto X y sean x, y y z tres elementos de X. Dos de los números $d(x, y)$, $d(x, z)$ y $d(y, z)$ son iguales.*

Demostración. Si ese no es el caso, entonces una de las tres distancias — digamos $d(x, y)$ — es estrictamente más grande que las otras dos: esto es absurdo, porque la desigualdad ultramétrica nos dice que $d(x, y) \leq \max\{d(x, z), d(y, z)\}$. □

Esto sugiere, claro, que la desigualdad ultramétrica es más fuerte que la desigualdad triangular usual de la condición $(\mathbf{M}_3)$. El siguiente lema afirma precisamente eso.

Lema 5.2.9.3. *Sea X un conjunto. Una función $d : X \times X \to \mathbb{R}$ que satisface las condiciones $(\mathbf{M}_1)$ y $(\mathbf{M}_2)$ de la Definición 5.1.1 y la desigualdad ultramétrica es una métrica sobre X.*

Demostración. Esto es consecuencia inmediata de que para una tal función tenemos que

$$0 = d(x,x) \leq \max\{d(x,y), d(y,x)\} = d(x,y)$$

cualesquiera sean x, y y z en X y, por lo tanto, también que

$$d(x,y) \leq \max\{d(x,z), d(z,y)\} \leq d(x,z) + d(y,z). \qquad \square$$

Un ejemplo sencillo

Daremos cuatro ejemplos de ultramétricas. El primero es el más sencillo.

Proposición 5.2.9.4. *Sea X un subconjunto de $(0, +\infty)$. La función $d : X \times X \to \mathbb{R}$ que en cada elemento (x,y) de $X \times X$ toma el valor*

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

es una ultramétrica sobre X.

Demostración. Sean x, y, y z tres elementos de X. Es $d(x,x) = 0$ y si $d(x,y) = 0$, entonces la definición de d y el hecho de que $0 \notin X$ implican que $x = y$. Que la función d es simétrica es evidente. Finalmente, es $d(x,y) = \max\{x,y\} \leq \max\{\max\{x,z\}, \max\{z,y\}\} = \max\{d(x,z), d(z,y)\}$, ya que $x \leq \max\{x,z\}$ e $y \leq \max\{z,y\}$. Todo esto nos permite concluir que la función d es una ultramétrica sobre X. $\square$

El espacio de las sucesiones en un conjunto

Nuestro segundo ejemplo es menos sencillo, pero más importante.

Proposición 5.2.9.5. *Sea A un conjunto, sea $X \coloneqq A^{\mathbb{N}_0}$, el conjunto de todas las sucesiones en A, y sea p un número real tal que $p > 1$. Si x e y son dos elementos distintos de X pongamos*

$$E(x,y) \coloneqq \{i \in \mathbb{N}_0 : x_j = y_j \text{ para cada } j \in [\![0, i-1]\!]\}$$

y

$$v(x,y) \coloneqq \max E(x,y).$$

La función $d : X \times X \to \mathbb{R}$ que en cada elemento (x,y) de $X \times X$ toma el valor

$$d(x,y) = \begin{cases} 0 & si\ x = y; \\ p^{-v(x,y)-1} & si\ x \neq y \end{cases}$$

es una ultramétrica sobre X.

Notemos que si x e y son elementos distintos de X el conjunto $E(x,y)$ no es vacío, porque contiene a 0, y es finito, así que tiene sentido considerar su elemento máximo.

Demostración. Sean x, y y z tres elementos de X.

- La definición de la función d hace evidente que vale $d(x,y) = 0 \iff x = y$.
- Si $x = y$, entonces $d(x,y) = 0 = d(y,x)$. Si, en cambio, $x \neq y$, entonces los conjuntos $E(x,y)$ y $E(y,x)$ son iguales, así que tienen el mismo elemento máximo, es $v(x,y) = v(y,x)$ y, por lo tanto, $d(x,y) = d(y,x)$.
- Si $x = z$ o $z = y$, entonces un elemento del conjunto $\{d(x,z), d(z,y)\}$ es nulo y el otro es igual a $d(x,y)$, así que $d(x,y) = \max\{d(x,z), d(z,y)\}$. Por otro lado, si $x = y$, entonces $d(x,y) = 0 \leq \max\{d(x,z), d(z,y)\}$, ya que el máximo es un número no negativo. Para probar que la función d satisface la desigualdad ultramétrica resta entonces que consideremos el caso en el que las tres sucesiones x, y y z son distintas dos a dos.

  Pongamos $m \coloneqq v(x,z)$ y $n \coloneqq v(z,y)$ y, sin pérdida de generalidad, supongamos que $m \leq n$, de manera que también $d(x,z) = p^{-m-1} \geq p^{-n-1} = d(z,y)$ y, en consecuencia, $\max\{d(x,z), d(z,y)\} = d(x,z)$. Como $m \in E(x,z)$, para cada $j \in [\![0, m-1]\!]$ es $x_j = z_j$, y como $n \in E(z,y)$, para cada $j \in [\![0, m-1]\!]$ tenemos que $j \in [\![0, n-1]\!]$ y, por lo tanto que $z_j = y_j$. Vemos así que para cada $j \in [\![0, m-1]\!]$ es $x_j = y_j$, de manera que $m \in E(x,y)$, $v(x,y) \geq m$ y

$$d(x,y) = p^{-v(x,y)-1} \leq p^{-m-1} = p^{-v(x,z)-1} = d(x,z) = \max\{d(x,z), d(z,y)\}.$$

Podemos concluir con todo esto que la función d es una ultramétrica sobre el conjunto X. □

La métrica sobre un grupo abeliano asociada a una cadena descendente de subgrupos

El tercer ejemplo que damos es una forma general e importante de construir ultramétricas sobre grupos abelianos.

Proposición 5.2.9.6. *Sea G un grupo abeliano, sea $(H_n)_{n\geq 0}$ una sucesión de subgrupos de G tal que $H_0 = G$, $H_{n+1} \subseteq H_n$ para todo $n \in \mathbb{N}_0$ y $\bigcap_{n=0}^{\infty} H_n = 0$, y para cada elemento x de G consideremos el conjunto $I(x) \coloneqq \{n \in \mathbb{N}_0 : x \in H_n\}$.*

(*i*) *Si x es un elemento no nulo de G, entonces el conjunto $I(x)$ es finito y no vacío y podemos, por lo tanto, considerar su mayor elemento $v(x) \coloneqq \max I(x)$.*

(*ii*) *La función $d : G \times G \to \mathbb{R}$ que en cada elemento (x,y) de $G \times G$ toma el valor*

$$d(x,y) = \begin{cases} 0 & \text{si } x = y; \\ 2^{-v(x-y)} & \text{si } x \neq y \end{cases}$$

es una ultramétrica sobre G que es invariante por translaciones.

Demostración. Sea x un elemento no nulo de G y consideremos el conjunto $I(x)$ definido en el enunciado de la proposición. Como $H_0 = G$, claramente $0 \in I(x)$ y, por lo tanto, $I(x) \neq \varnothing$. Por otro lado, como $x \notin \bigcap_{n=0}^{\infty} H_n$, hay un elemento n_0 de $\mathbb{N}_0$ tal que $x \notin H_{n_0}$: para todo entero n tal que $n \geq n_0$ tenemos entonces que $H_n \subseteq H_{n_0}$, así que $x \notin H_n$ y $n \notin I(x)$. Esto nos dice que $I(x) \subseteq [\![0, n_0 - 1]\!]$ y, por lo tanto, que el conjunto $I(x)$ es finito. Esto prueba la afirmación (i) de la proposición. Podemos entonces considerar el entero $v(x) \coloneqq \max I(x)$ y es, de hecho,

$$I(x) = [\![0, v(x)]\!]. \tag{5.13}$$

En efecto, si n es un elemento de $[\![0, v(x)]\!]$, entonces $H_n \supseteq H_{v(x)} \ni x$, así que $n \in I(x)$. Por otro lado, si n es un elemento de $I(x)$, entonces ciertamente $n \leq v(x)$, así que $n \in [\![0, v(x)]\!]$.

Mostremos ahora que la función d es una ultramétrica invariante por translaciones.

- Es claro que $d(x, x) = 0$ para todo elemento x de G. Por otro lado, si x e y son dos elementos distintos de G, entonces hay un elemento n de $\mathbb{N}_0$ tal que $x - y \notin H_n$ y, por lo tanto, $I(x, y) \subseteq [\![0, n-1]\!]$ y $d(x, y) = 2^{-v(x-y)} \geq 2^{-n+1} > 0$.
- Si x e y son dos elementos distintos de G, entonces para todo $n \in \mathbb{N}_0$ vale

  $$x - y \in H_n \iff y - x \in H_n,$$

  ya que H_n es un subgrupos de G, y esto implica que $I(x-y) = I(y-x)$, $v(x-y) = v(y-x)$, y, por lo tanto, que $d(x, y) = 2^{-v(x-y)} = 2^{-v(y-x)} = d(y, x)$. Por supuesto, la igualdad $d(x, y) = d(y, x)$ también vale si $x = y$.
- Sean x, y, z tres elementos de G. Si hay dos entre esos tres elementos que son iguales, entonces uno de los tres números $d(x, y)$, $d(x, z)$, y $d(z, y)$ es nulo, y entonces vale la desigualdad

  $$d(x, y) \leq \max\{d(x, z), d(z, y)\}.$$

  Supongamos que, por el contrario, los tres elementos x, y, y z son distintos dos a dos. Afirmamos que

  $$I(x - y) \supseteq I(x - z) \cap I(z - y).$$

  En efecto, si n es un elemento de $\mathbb{N}_0$ que está en $I(x - z)$ y en $I(z - y)$, de manera que $x - z \in H_n$ y $z - y \in H_n$, entonces $x - y = (x - z) + (z - y) \in H_n$ porque H_n es un subgrupo de G, así que $n \in I(x - y)$. En vista de (5.13), se sigue de esto que

  $$[\![0, v(x - y)]\!] \supseteq [\![0, v(x - z)]\!] \cap [\![0, v(z - y)]\!]$$

y esto implica claramente que $v(x-y) \geq \min\{v(x-z), v(z-y)\}$, de manera que

$$d(x,y) = 2^{-v(x-y)} \leq \max\{2^{-v(x-z)}, 2^{-v(z-y)}\} = \max\{d(x,z), d(z,y)\}.$$

Vemos así que la función d satisface la desigualdad ultramétrica.

- Sean, finalmente, x, y y z tres elementos de G. Si $x = y$, entonces también $x+z = y+z$ y, por lo tanto, $d(x,y) = 0 = d(x+z, y+z)$. Si, en cambio, $x \neq y$, entonces también $x+z \neq y+z$ y es claro que $I(x-y) = I((x+z)-(y+z))$, así que $v(x-y) = v((x+z)-(y+z))$ y, por lo tanto,

$$d(x,y) = 2^{-v(x-y)} = 2^{-v((x+z)-(y+z))} = d(x+z, y+z).$$

Esto nos dice que la función d es invariante por translaciones.

Esto completa la demostración de la proposición. □

La métrica p-ádica sobre $\mathbb{Q}$

Nuestro cuarto ejemplo es más complicado de presentar, pero es extremadamente importante. Fijemos un número primo p. Si n es un entero no nulo, entonces el conjunto $\{k \in \mathbb{N}_0 : p^n \mid n\}$ es finito y no vacío, así que podemos considerar la función

$$v_p : n \in \mathbb{Z} \setminus \{0\} \mapsto \max\{k \in \mathbb{N}_0 : p^k \mid n\} \in \mathbb{N}_0.$$

Llamamos a esta función la ***valuación p-ádica*** sobre $\mathbb{Z}$. Notemos que para todo entero n vale que

$$v_p(n) \textit{ es el único elemento } k \textit{ de } \mathbb{N}_0 \textit{ tal que hay un entero } n' \textit{ coprimo con } p \textit{ tal que } n = n'p^k.$$

El siguiente lema describe las propiedades fundamentales de la función v_p.

Lema 5.2.9.7. *Sean m y n dos enteros no nulos. Es*

$$v_p(m \cdot n) = v_p(m) + v_p(n)$$

y, si $m + n \neq 0$,

$$v_p(m+n) \geq \min\{v_p(m), v_p(n)\}.$$

Demostración. Hay enteros m' y n' coprimos con p tales que $m = m'p^{v_p(m)}$ y $n = n'p^{v_p(n)}$. Por un lado, esto implica que $mn = m'n'p^{v_p(m)+v_p(n)}$ y, como $m'n'$ es coprimo con p, que $v_p(mn) = v_p(m) + v_p(n)$. Por otro, si ponemos $r := \min\{v_p(m), v_p(n)\}$, entonces tenemos que

$$p^r \mid (m'p^{v_p(m)-r} + n'p^{v_p(n)-r})p^r = m+n$$

de manera que $v_p(m+n) \geq r = \min\{v_p(m), v_p(n)\}$. □

La valuación p-ádica tiene una única extensión natural al cuerpo de los números racionales.

Lema 5.2.9.8. *Hay exactamente una función $\bar{v}_p : \mathbb{Q} \setminus \{0\} \to \mathbb{Z}$ cuya restricción a $\mathbb{Z} \setminus \{0\}$ es v_p y tal que para cada par de números racionales r y s vale que*

$$v_p(r \cdot s) = v_p(r) \cdot v_p(s)$$

y, si $r + s \neq 0$,

$$v_p(r+s) \geq \min\{v_p(r), v_p(s)\}.$$

Llamamos a esta función $\bar{v}_p$ la ***valuación p-ádica*** sobre $\mathbb{Q}$ y, desde ahora, la escribimos simplemente v_p — esto no introduce ninguna ambigüedad, ya que se trata de una extensión de la función v_p, claro.

Demostración. Si r es un elemento no nulo de $\mathbb{Q}$, entonces podemos elegir dos enteros a y b de manera que sea $b \neq 0$ $r = a/b$: afirmamos que la diferencia $v_p(a) - v_p(b)$ depende solamente de r y no de la forma en que elegimos a a y a b. En efecto, si a' y b' son dos enteros tales que $b' \neq 0$ y $r = a'/b'$, entonces $ba' = b'a$ y, por lo tanto, $v_p(b)+v_p(a') = v_p(ba') = v_p(b'a) = v_p(b')+v_p(a)$, de manera que $v_p(a) - v_p(b) = v_p(a') - v_p(b')$.

Esto implica que hay una función $\bar{v}_p : \mathbb{Q} \setminus \{0\} \to \mathbb{Z}$ tal que $\bar{v}_p(a/b) = v_p(a) - v_p(b)$ para toda elección de a y b en $\mathbb{Z}$ tal que $b \neq 0$ y $r = a/b$, y es claro que hay exactamente una. Mostremos que esta función $\bar{v}_p$ tiene las propiedades descriptas en el lema.

- Si n es un entero, entonces tenemos $1 \neq 0$ y $n = n/1$, así que $\bar{v}_p(n) = v_p(n) - v_p(1) = v_p(n)$, ya que $v_p(1) = 0$. Vemos así que la restricción de $\bar{v}_p$ a $\mathbb{Z}$ es v_p.
- Sean r y s dos elementos de $\mathbb{Q} \setminus \{0\}$, y sean a, b, c, y d enteros tales que $b \neq 0$, $d \neq 0$, $r = a/b$ y $s = c/d$, de manera que $\bar{v}_p(r) = v_p(a) - v_p(b)$ y $\bar{v}_p(s) = v_p(c) - v_p(d)$. Como $bd \neq 0$ y $rs = ac/bd$, tenemos que

  $$\bar{v}_p(rs) = v_p(ac) - v_p(bd) = v_p(a) + v_p(c) - v_p(b) - v_p(d) = \bar{v}_p(r) + \bar{v}_p(s).$$

  Supongamos que $r + s \neq 0$. Es $bd \neq 0$ y $r + s = (ad + bc)/bd$, así que

  $$\begin{aligned}
  \bar{v}_p(r+s) &= v_p(ad+bc) - v_p(bd) \\
  &\geq \min\{v_p(ad), v_p(bc)\} - v_p(bd) && \text{ya que } ad+bc \neq 0 \\
  &= \min\{v_p(ad) - v_p(bd), v_p(bc) - v_p(bd)\} \\
  &= \min\{\bar{v}_p(ad/bd), \bar{v}_p(bc/bd)\} \\
  &= \min\{\bar{v}_p(r), \bar{v}_p(s)\}
  \end{aligned}$$

Esto completa la prueba del lema. □

Ejercicio 5.2.9.9. Sea $\mathbb{P}$ el conjunto de los números primos. Pruebe que si x es un elemento no nulo de $\mathbb{Q}$, entonces el conjunto $I(x) \coloneqq \{p \in \mathbb{P} : v_p(x) \neq 0\}$ es finito y si ponemos $\varepsilon = x/|x|$, que es un elemento de $\{\pm 1\}$, entonces

$$x = \varepsilon \prod_{p \in I(x)} p^{v_p(x)}.$$

A partir de la valuación p-ádica podemos construir lo que llamamos el ***valor absoluto p-ádico*** sobre $\mathbb{Q}$, que tiene propiedades similares a las del valor absoluto usual.

Lema 5.2.9.10. *Sea $|-|_p : \mathbb{Q} \to \mathbb{R}$ la función que en cada elemento r de $\mathbb{Q}$ toma el valor*

$$|r|_p = \begin{cases} 0 & \text{si } r = 0; \\ p^{-v_p(r)} & \text{si } r \neq 0. \end{cases}$$

Para cada elección de r y s en $\mathbb{Q}$ valen las siguientes afirmaciones:

(*i*) $|r|_p = 0$ *si y solamente si* $r = 0$.
(*ii*) $|r \cdot s|_p = |r|_p \cdot |s|_p$.
(*iii*) $|-r| = |r|_p$.
(*iv*) $|r + s|_p \leq \max\{|r|_p, |s|_p\} \leq |r|_p + |s|_p$.

Demostración. Sean r y s dos elementos de $\mathbb{Q}$. La afirmación (*i*) es evidente en vista de la definición de la función $|-|_p$. La afirmación (*ii*), por su parte, es también evidente si alguno de r o s es nulo, ya que ambos lados de la igualdad que aparece en ella son nulos en ese caso, y si, en cambio, es $r \neq 0$ y $s \neq 0$, entonces

$$|r \cdot s|_p = p^{-v_p(r \cdot s)} = p^{-v_p(r)} \cdot p^{-v_p(s)} = |r|_p \cdot |s|_p.$$

Usado esto, vemos que $|-r|_p = |(-1) \cdot r|_p = |-1|_p \cdot |r|_p = |r|_p$, ya que $|-1|_p = p^{-v_p(-1)} = 1$: esto prueba la afirmación (*iii*). Finalmente, si alguno de los números r, s o $r+s$ es nulo, entonces la desigualdad de (*iv*) es evidente, y si los tres son no nulos, entonces esa desigualdad también vale ya que

$$\begin{aligned} |r + s|_p = p^{-v_p(r+s)} &\leq p^{-\min\{v_p(r), v_p(s)\}} = p^{\max\{-v_p(r), -v_p(s)\}} = \max\{p^{-v_p(r)}, p^{-v_p(s)}\} \\ &= \max\{|r|_p, |s|_p\}. \end{aligned}$$

□

Usando ahora el valor absoluto p-ádico podemos construir un espacio ultramétrico importante, que es el tercer ejemplo que queremos.

Proposición 5.2.9.11. *Sea p un número primo. La función*

$$d_p : (r,s) \in \mathbb{Q}\times\mathbb{Q} \mapsto |r-s|_p \in \mathbb{R}$$

es una ultramétrica sobre $\mathbb{Q}$ que es invariante por translaciones.

Llamamos a esta ultramétrica d la ***ultramétrica p-ádica*** sobre $\mathbb{Q}$. Por supuesto, como $\mathbb{Z}$ es un subconjunto de $\mathbb{Q}$, podemos también considerar la restricción de esta ultramétrica a $\mathbb{Z}$.

Demostración. Sean r, t, y t tres números racionales. De la definición de la función d es claro que

$$d_p(r,s) = 0 \iff |r-s|_p = 0 \iff r-s = 0 \iff r = s.$$

Por otro lado, es

$$d_p(r,s) = |r-s|_p = |(-1)\cdot(s-r)|_p = |-1|_p \cdot |s-r|_p = 1\cdot d_p(s,r) = d_p(s,r)$$

y

$$\begin{aligned} d_p(r,s) &= |r-s|_p = |(r-t)+(t-s)|_p \leq \max\{|r-t|_p, |t-s|_p\} \\ &= \max\{d_p(r,t), d_p(t,s)\}. \end{aligned}$$

Finalmente, que la función d_p es invariante por translaciones es evidente. □

Las ideas que llevan a este ejemplo pueden generalizarse de varias maneras. Veamos una.

Ejercicio 5.2.9.12. Sea $\Bbbk$ un cuerpo y sea p un elemento irreducible de $\Bbbk[X]$. Para cada polinomio f de $\Bbbk[X]$ el conjunto $\{k \in \mathbb{N}_0 : p^k \mid f\}$ es finito, así que podemos considerar su mayor elemento y escribirlo $v_p(f)$. Muestre que hay una ultramétrica $d : \Bbbk[X]\times\Bbbk[X] \to \mathbb{R}$ sobre $\mathbb{R}[X]$ tal que siempre que f y g son dos elementos distintos de $\Bbbk[X]$ se tiene que $d(f,g) = e^{-v_p(f-g)}$, y que esta es invariante por translaciones.

La métrica p-ádica sobre $\mathbb{Q}$ induce por restricción una métrica sobre $\mathbb{Z}$, por supuesto, y llamamos a esa restricción la métrica p-ádica de $\mathbb{Z}$.

Ejercicio 5.2.9.13. Sea p un número primo, y para cada $n \in \mathbb{N}_0$ sea $H_n \coloneqq p^n\mathbb{Z}$, el subgrupo de $\mathbb{Z}$ de todos los múltiplos de p^n. Claramente es $H_0 = \mathbb{Z}$, $H_{n+1} \subseteq H_n$ para todo $n \in \mathbb{N}_0$, y es fácil verificar que $\bigcap_{n=0}^{\infty} H_n = 0$. La sucesión $(H_n)_{n\geq 0}$ de subgrupos de $\mathbb{Z}$ determina entonces, de acuerdo a la Proposición 5.2.9.6, una métrica sobre $\mathbb{Z}$. Pruebe que está métrica coincide con la métrica p-ádica.

Terminemos este apartado poniendo en contexto las construcciones que hicimos arriba.

Definición 5.2.9.14. Sea K un cuerpo. Una función $|-| : K \to [0, +\infty)$ es un ***valor absoluto*** si cada vez que x e y son elementos de K valen las siguientes condiciones:

- $|x| = 0 \iff x = 0$.
- $|x + y| \leq |x| + |y|$.
- $|x \cdot y| = |x| \cdot |y|$.

En la situación de esta definición es $|1|^2 = |1^2| = |1|$ y $|1| \neq 0$ porque $1 \neq 0$, así que $|1| = 1$. De manera similar, $|-1|^1 = |(-1)^2| = |1| = 1$ y $|-1| \neq 0$ porque $-1 \neq 0$, así que $|-1| = 1$. Por otro lado, si q es un elemento no nulo de K, entonces $1 = |1| = |q \cdot q^{-1}| = |q| \cdot |q^{-1}|$, así que $|q^{-1}| = |q|^{-1}$. Esto y la multiplicatividad de $|-|$ nos permiten concluir que $|q^n| = |q|^n$ para toda $q \in K$ y todo $n \in \mathbb{Z}$.

Todo cuerpo admite al menos un valor absoluto, el descripto en el siguiente ejemplo.

Ejemplo 5.2.9.15. Sea K un cuerpo. La función $|-|_0 : K \to [0, +\infty)$ que en cada elemento x de K toma el valor

$$|x|_0 = \begin{cases} 0 & \text{si } x = 0; \\ 1 & \text{si } x \neq 0 \end{cases}$$

es un valor absoluto sobre K, al que llamamos el ***valor absoluto trivial***.

Sobre el cuerpo $\mathbb{Q}$ tenemos, además del valor absoluto trivial, el valor absoluto usual, que escribimos aquí $|-|_\infty$, y para cada número primo p el valor absoluto p-adico $|-|_p$. El siguiente teorema de Alexander Ostrowski [Ost1916] nos dice que estos son, esencialmente, los únicos valores absolutos que hay sobre $\mathbb{Q}$.

Proposición 5.2.9.16. *Sea $|-| : \mathbb{Q} \to [0, +\infty)$ un valor absoluto sobre el cuerpo $\mathbb{Q}$. Exactamente una de las siguientes tres afirmaciones es cierta:*

(*a*) *Es $|x| = |x|_0$ para todo $x \in \mathbb{Q}$.*
(*b*) *Hay un número positivo λ tal que $|x| = |x|_\infty^\lambda$ para todo $x \in \mathbb{Q}$.*
(*c*) *Hay exactamente un número primo p y un número positivo λ tal que $|x| = |x|_p^\lambda$ para todo $x \in \mathbb{Q}$.*

Demostración. Recordemos el resultado del Ejercicio 5.2.9.9: para cada elemento no nulo r de $\mathbb{Q}$ el subconjunto $I(r) \coloneqq \{p \in \mathbb{P} : v_p(r) \neq 0\}$ del conjunto $\mathbb{P}$ de los números primos es finito y si ponemos $\varepsilon(r) \coloneqq r/|r|_\infty$, entonces es $r = \varepsilon(r) \prod_{p \in I(r)} p^{v_p(r)}$ y, por lo tanto,

$$|r|_* = \prod_{p \in I(r)} |p|_*^{v_p(r)}.$$

Hay tres posibilidades:

- o bien $|-|_*$ coincide con $|-|_0$ y, en particular, $|n|_* = 1$ para todo entero n mayor que 1,
- o bien hay un entero n mayor que 1 tal que $|n|_* > 1$,
- o bien para todo entero n mayor que 1 es $|n|_* \leq 1$.

Supongamos que no estamos en la primera de las tres, de manera que no vale la afirmación (a) de la proposición, y analicemos por separado las otras dos. Tenemos entonces dos casos a estudiar.

Primer caso. Supongamos que hay un elemento n de $\mathbb{N}$ tal que $|n|_* > 1$. Tiene que ser $n > 1$, ya que $|1|_* = 1$. Sean b y k dos enteros tales que $k \geq 1$ y $b \geq 2$. Hay entero positivo m y una m-upla $(c_0, c_1, \ldots, c_m)$ de elementos de $[\![0, m-1]\!]$ tales que $c_m \neq 0$ y $n^k = \sum_{i=0}^{m} c_i b^i$. Como $c_m \geq 1$, es $n^k = \sum_{i=0}^{m} c_i b^i \geq b^m$, así que

$$m \leq k \log_b n.$$

Por otro lado, para cada $i \in [\![0, m]\!]$ es

$$|c_i|_* = |\underbrace{1 + 1 + \cdots + 1}_{c_i \text{ sumandos}}|_* \leq |1|_* + |1|_* + \cdots + |1|_* = c_i \leq b - 1,$$

así que $|c_i b^i|_* \leq (b-1) \cdot |b|_*^i$. Esto implica que

$$|n|_*^k = |n^k|_* = \left|\sum_{i=0}^{m} c_i b^i\right|_* \leq \sum_{i=0}^{m} |c_i b^i|_* \leq (m+1) \cdot (b-1) \cdot \max\{|b|_*^i : i \in [\![0, m]\!]\}$$

Si $|b|_* \leq 1$, entonces el máximo que aparece aquí es 1, y si $|b|_* > 1$, ese máximo es $|b|_*^m$: vemos así que ese máximo es menor o igual que $\max\{1, |b|_*^m\}$. Es entonces

$$|n|_*^k \leq (m+1) \cdot (b-1) \cdot \max\{1, |b|_*^m\} \leq (k \log_b n + 1) \cdot (b-1) \cdot \max\{1, |b|_*^{k \log_b n}\}$$

y, por lo tanto,

$$|n|_* \leq \left[(k \log_b n + 1) \cdot (b-1)\right]^{1/k} \cdot \max\{1, |b|_*^{\log_b n}\}.$$

Como

$$\lim_{k \to \infty} \left[(k \log_b n + 1) \cdot (b-1)\right]^{1/k} = 1,$$

esto nos permite deducir que $1 < |n|_* \leq \max\{1, |b|_*^{\log_b n}\}$, así que, de hecho, $1 < |n|_* \leq |b|_*^{\log_b n}$. Vemos con esto que $|b|_* > 1$ y que

$$\log_n |n|_* \leq \log_b n \cdot \log_n |b|_* = \log_b |b|_*.$$

Podemos intercambiar los roles de n y de b, y concluir que, de hecho, $\log_n|n|_* = \log_b|b|_*$. Vemos así que hay un número positivo λ tal que para todo entero n tal que $n > 1$ es $\log_n|n|_* = \lambda$ y $|n|_* = n^\lambda$ Si ahora q es un elemento no nulo de $\mathbb{Q}$ cualesquiera, entonces

$$|q|_* = |\varepsilon|_* \prod_{p\in I(q)} |p|_*^{v_p(q)} = \prod_{p\in I(q)} p^{\lambda v_p(q)} = |q|_\infty^\lambda$$

ya que $\varepsilon \in \{\pm 1\}$ y, por lo tanto, $|\varepsilon|_* = 1$. Vemos así que vale la afirmación (b) de la proposición.

Segundo caso. Supongamos ahora que $|n|_* \leq 1$ para todo entero n mayor que 1. Como estamos suponiendo además que $|-|_* \neq |-|_0$, hay un entero positivo n tal que $|n|_* < 1$. Tenemos entonces que

$$1 > |n|_* = \prod_{p\in I(n)} |p|_*^{v_p(n)}$$

y, como $|p|_* \leq 1$ para todo $p \in I(n)$, vemos que hay un número primo p tal que $|p|_* < 1$. Si hubiera dos números primos p y q tales que $|p|_* < 1$ y $|q|_* < 1$, entonces habría, por un lado, un entero positivo k tal que $|p|_*^k < 1/2$ y $|q|_*^k < 1/2$ y, por otro, como los números p^k y q^k son coprimos, dos enteros u y v tales que $up^k + vq^k = 1$, de manera que

$$1 = |1|_* = |up^k + vq^k|_* \leq |u| \cdot |p|_*^k + |v| \cdot |q|_*^k < 1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = 1,$$

lo que es absurdo.

Vemos así que hay exactamente un número primo p tal que $|p|_* < 1$ y que, por lo tanto, es $|q|_* = 1$ para todo número primo q distinto de p. Sea $\lambda := -\log_p|p|_*$, que es un número positivo. Si r es un número racional, entonces lo que hemos hecho implica que

$$|r|_* = \prod_{q\in I(r)} |q|_*^{v_q(r)} = |p|_*^{v_p(r)} = p^{\lambda\cdot v_p(r)} = |r|_p^\lambda.$$

Vemos así que vale en este caso vale la afirmación (c) de la proposición. □

Usando la siguoente definición podemos enunciar el *Teorema de Ostrowski* en una forma muy conveniente.

Definición 5.2.9.17. Sea K un cuerpo. Un valor absoluro $|-| : K \to [0, +\infty)$ sobre K es ***equivalente*** a otro $|-|' : K \to [0, +\infty)$ si hay un número positivo λ tal que $|x|' = |x|^\lambda$ para todo $x \in K$, y en ese caso escribimos $|-| \sim |-|'$.

Dejamos al lector la sencilla tarea de verificar que la relación $\sim$ sobre el conjunto de todos los valores absolutos de un cuerpo es una relación de equivalencia. Usando este lenguaje, podemos enunciar el *Teorema de Ostrowski* de la siguiente manera:

Corolario 5.2.9.18. *Todo valor absoluto sobre $\mathbb{Q}$ es equivalente al valor absoluto trivial $|-|_0$, al valor absoluto usual $|-|_\infty$, o al valor absoluto p-ádico $|-|_p$ para algún número primo p.* □

Las clases de equivalencia de valores absolutos sobre un cuerpo son los ***lugares*** de este. El *Teorema de Ostrowski*, entonces, describe el conjunto de los lugares de $\mathbb{Q}$.

Ejercicio 5.2.9.19. Es inmediato verificar que si $|-| : K \to [0, +\infty)$ es un valor absoluto sobre un cuerpo K y L es un subcuerpo de K, entonces la restricción $|-|\big|_L : L \to [0, +\infty)$ de $|-|$ a L es un valor absoluto sobre L. Describa el conjunto de las clases de equivalencia de valores absolutos sobre el cuerpo $\mathbb{C}(t)$ de las funciones racionales con coeficientes en $\mathbb{C}$ que se restringen al valor absoluto trivial de $\mathbb{C}$.

Capítulo 6

La topología de un espacio métrico

6.1. Bolas abiertas

Fijemos un espacio métrico (X, d).

Definición 6.1.1. Si x es un punto de X y r es un número real positivo, entonces el conjunto

$$B_r(x) \coloneqq \{y \in X : d(x, y) < r\}$$

es la ***bola abierta centrada en x de radio r***.

Cuando queremos dejar de manifiesto en qué espacio métrico estamos trabajando, escribiremos $B_r^X(x)$ en lugar de simplemente $B_r(x)$. Esto es útil, por ejemplo, en el siguiente resultado que describe las bolas abiertas de los subespacios del espacio X.

Lema 6.1.2. *Sea Y un subespacio de X. Si x es un elemento de Y y r es un número positivo, entonces* $B_r^Y(x) = Y \cap B_r^X(x)$.

Notemos que, en general, en esta situación las bolas abiertas de Y no son bolas abiertas de X. Por ejemplo, el conjunto $[0, 1)$ es una bola abierta del subespacio $[0, +\infty)$ de $\mathbb{R}$, de centro 0 y radio 1, pero no es una bola abierta de $\mathbb{R}$.

Demostración. Esto es consecuencia inmediata de las definiciones. □

El siguiente ejercicio explicita una propiedad fundamental de las bolas abiertas de un espacio métrico que usaremos todo el tiempo.

Ejercicio 6.1.3. Sea x un punto de X.

(a) Muestre que si n es un entero positivo, $r_1, r_2, \ldots, r_n$ son números positivos, y ponemos $r \coloneqq \min\{r_1, r_2, \ldots, r_n\}$, entonces r es un número positivo y

$$B_r(x) = B_{r_1}(x) \cap B_{r_2}(x) \cap \cdots \cap B_{r_n}(x).$$

Esto nos dice que la intersección de una familia finita de bolas abiertas centradas en x es ella misma una bola abierta centrada en x. En general, eso no es cierto si la familia es infinita.

(b) Muestre que $\bigcap_{n\in\mathbb{N}} B_{1/n}(x) = \{x\}$, y dé ejemplos de subconjuntos no vacíos R de $(0, +\infty)$ tales que $\inf R > 0$ y la intersección $\bigcap_{r\in R} B_r(x)$ no sea una bola abierta centrada en x.

En general, una bola abierta no determina de manera unívoca ni su centro ni su radio.

Ejercicio 6.1.4. Dé ejemplos de espacios métricos en los que haya bolas abiertas con más de un centro y bolas abiertas con más de un radio.

Veamos algunos ejemplos.

Ejemplo 6.1.5. Consideremos el espacio métrico $\mathbb{R}$ con su métrica euclídea. Si x es un elemento de $\mathbb{R}$ y r es un número positivo, entonces

$$B_r(x) = (x - r, x + r).$$

Notemos que las bolas abiertas de $\mathbb{R}$ son exactamente los intervalos abiertos. En efecto, acabamos de observar que toda bola abierta es un intervalo abierto y, por otro lado, si a y b son dos números reales tales que $a < b$, entonces es fácil verificar que

$$(a, b) = B_{\frac{b-a}{2}}\left(\frac{a+b}{2}\right).$$

Ejemplo 6.1.6. Sobre el espacio vectorial real $\mathbb{R}^n$ construimos muchas normas y, por lo tanto, muchas métricas. En la Figura 6.1 de la página 178 dibujamos las bolas abiertas $B_1(0)$ para varias de ellas, y en la Figura 6.2 de la página 179 las correspondientes bolas en $\mathbb{R}^3$.

Ejemplo 6.1.7. Sean (X, d_X) y (Y, d_Y) dos espacios métricos, y sea d la métrica sobre el conjunto $X \times Y$ que describe la Proposición 5.2.6.1, de manera que

$$d((x, y), (x', y')) = \max\{d_X(x, x'), d_Y(y, y')\}$$

cualesquiera sean los puntos (x, y) y (x', y') de $X \times Y$. Si (x, y) es un elemento de $X \times Y$ y r es

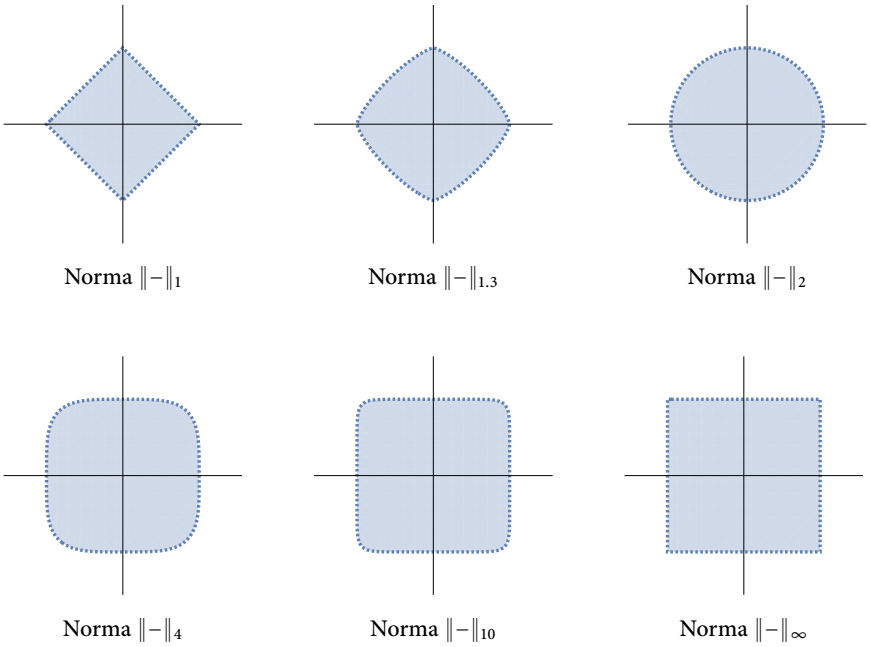


Figura 6.1. La bola abierta $B_1(0)$ en $\mathbb{R}^2$ para algunas de las normas $\|-\|_p$.

un número positivo, la bola $B_r((x,y))$ es entonces el conjunto

$$B_r((x,y)) = \{(x',y') \in X \times Y : \max\{d_X(x,x'), d_Y(y,y')\} < r\}.$$

y es fácil ver que, de hecho, tenemos

$$B_r((x,y)) = B_r^X(x) \times B_r^Y(y).$$

Si consideramos en cambio un elemento p de $[1,+\infty)$ y la métrica d_p del Ejercicio 5.2.6.3, entonces esto no es cierto: la bola $B_r((x,z))$ con respecto a la métrica d_p no es, en general, el producto cartesiano de una bola de X por una bola de Y.

Ejemplo 6.1.8. Si la métrica d de nuestro espacio es discreta, entonces las bolas abiertas son muy sencillas: si x es un punto de X y r es un número positivo, entonces

$$B_r(x) = \begin{cases} \{x\} & \text{si } r < 1; \\ X & \text{si } r \geq 1. \end{cases}$$

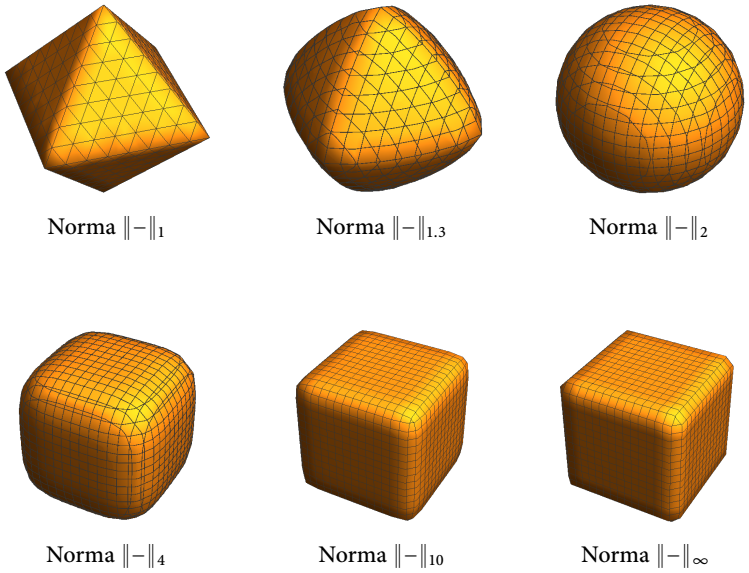


Figura 6.2. La bola abierta $B_1(0)$ en $\mathbb{R}^3$ para algunas de las normas $\|-\|_p$.

Ejemplo 6.1.9. Sea G un grupo, sea S un subconjunto simétrico de G que genera a G como grupo, y sea d la métrica de las palabras sobre G relativa al conjunto S, como en el Ejemplo 5.2.7.5. Si e es el elemento neutro de G y r es un número positivo, entonces la bola abierta $B_r(e)$ es el conjunto de los elementos de G que pueden escribirse como un producto de menos que r elementos de S. Así, por ejemplo, $B_1(e) = \{e\}$, $B_2(e) = S$ y $B_3(e) = \{s_1 s_2 : s_1, s_2 \in S\}$.

Si el conjunto S es finito — de manera que el grupo G es finitamente generado — entonces claramente el conjunto $B_r(e)$ es finito cualquiera sea el número positivo r y su cardinal es, como mucho, $|S|^r$. Tiene sentido entonces considerar la función

$$v_S : n \in \mathbb{N} \mapsto |B_n(e)| \in \mathbb{N}.$$

Esta función es la ***función de crecimiento*** del grupo G con respecto al conjunto generador S, y una parte importante de la teoría geométrica de grupos, iniciada por Mikhail Gromov en [Gro1987], consiste en estudiar al grupo a través de su función de crecimiento. Así, por ejemplo, un teorema de Gromov de [Gro1981] dice que si hay un polinomio $p \in \mathbb{R}[X]$ tal que $v_S(n) \leq p(n)$ para todo $n \in \mathbb{N}$ — decimos en ese caso que el grupo tiene ***crecimiento polinomial*** con respecto a S — entonces el grupo G posee un subgrupo nilpotente de índice finito.

Ejercicio 6.1.10.

(a) Muestre que si G es un grupo finito, S es un subconjunto generador simétrico, y $v_S : \mathbb{N} \to \mathbb{N}$

es la función de crecimiento de G correspondiente a S, entonces hay números reales a y b tales que $v(n) \leq a$ para todo entero n mayor que b.

(b) Sea n un entero positivo y sea $(e_1, e_2, \ldots, e_n)$ la base ordenada estándar del grupo $\mathbb{Z}^n$. Muestre que el conjunto $S \coloneqq \{\pm e_i : i \in [\![n]\!]\}$ es simétrico y genera a $\mathbb{Z}^n$, y determine la correspondiente función de crecimiento.

Ejemplo 6.1.11. Sea n un enter positivo, sea $\mathscr{A}$ un conjunto, y consideremos el conjunto $\mathscr{A}^n$ dotado de la métrica de Hamming. Si x es un elemento de $\mathscr{A}^n$ y r es un número positivo, entonces la bola $B_r(x)$ es el conjunto de los elementos de $\mathscr{A}^n$ que difieren de x en menos que r componentes.

Ejercicio 6.1.12. Sea X un subconjunto de $(0, +\infty)$, y sea $d : X \times X \to \mathbb{R}$ la ultramétrica sobre X que describe la Proposición 5.2.9.4, de manera que $d(x, y) = \max\{x, y\}$ siempre que x e y son dos elementos distintos de X. Pruebe que si x es un elemento de X y r es un número positivo, entonces

$$B_r(x) = \begin{cases} \{x\} & \text{si } x \geq r; \\ X \cap (0, r) & \text{si } x < r. \end{cases}$$

Ejercicio 6.1.13. Sea p un número primo y consideremos sobre $\mathbb{Q}$ y sobre $\mathbb{Z}$ las ultramétricas p-ádica de la Proposición 5.2.9.11.

(a) Pruebe que toda bola abierta de $\mathbb{Z}$ centrada en un entero n es de la forma

$$\{m \in \mathbb{Z} : n \equiv m \mod p^r\}$$

para algún entero no negativo r.

(b) Describa las bolas abiertas de $\mathbb{Q}$ con respecto a la ultramétrica p-ádica. En particular, observe que la bola $B_1(0)$ es el conjunto $\{a/b : a, \in \mathbb{Z}, p \nmid b\}$ que normalmente escribimos en la forma $\mathbb{Z}_{(p)}$ y llamamos la ***localización*** de $\mathbb{Z}$ en el ideal primo (p).

Los espacios involucrados en estos últimos dos ejercicios son ultramétricos. Si nos detenemos a observar la forma de sus bolas abiertas podemos darnos cuenta de que las bolas abiertas de un espacio ultramétrico cualquiera tienen la siguiente propiedad muy especial.

Lema 6.1.14. *Supongamos que (X, d) es un espacio ultramétrico. Dos bolas abiertas de X son o disjuntas o comparables por inclusión.*

Demostración. Sean B y B' dos bolas abiertas de X, de manera que hay dos puntos x e y en X y dos números positivos r y s tales que $B = B_r(x)$ y $B' = B_s(y)$, y supongamos que $B \cap B' \neq \varnothing$, de manera que podemos elegir un elemento z en esa intersección. Sin pérdida

de generalidad podemos suponer que $r < s$. Si u es un elemento de la bola abierta $B_r(x)$, entonces $d(z,u) \leq \max\{d(z,x), d(x,u)\} < r$, ya que $d(z,x) < r$, porque $z \in B_r(x)$, y $d(x,u) < r$. Esto implica que $d(y,u) \leq \max\{d(y,z), d(z,u)\} < s$, ya que $d(y,z) < s$, porque $z \in B_s(y)$, y $d(z,u) < r < s$. Vemos así que $B = B_r(x) \subseteq B_s(y) = B'$. □

Muchas de las métricas que consideramos están definidas sobre grupos abelianos y son invariantes por translaciones. En ese caso, podemos hacer la siguiente observación.

Proposición 6.1.15. *Sea G un grupo abeliano, sea 0 el elemento neutro de G, y sea d una métrica sobre G que es invariante por translaciones. Si x es un elemento de G y r es un número positivo, entonces $B_r(x) = x + B_r(0)$.*

Aquí, como es usual, como es usual, para cada elemento x de G y cada subconjunto A de G escribimos

$$x + A \coloneqq \{x + a : a \in A\}.$$

La proposición nos dice que cuando tenemos una métrica invariante por translaciones sobre un grupo abeliano en esta situación es suficiente con conocer las bolas abiertas centradas en 0, ya que a todas las demás las podemos obtener trasladando a aquellas.

Demostración. Claramente, si x es un elemento de G y r es un número positivo, entonces para todo elemento y de G vale que

$$y \in B_r(x) \iff d(x,y) < r \iff d(0, y-x) < r \iff y - x \in B_r(0) \iff y \in x + B_r(0),$$

y la afirmación de la proposición sigue de esto. □

Ejercicio 6.1.16. Muestre que toda bola abierta en un espacio vectorial normado es un subconjunto convexo de este.

6.2. Conjuntos abiertos

Usando las bolas abiertas de nuestro espacio métrico X podemos definir la clase más importante de subconjuntos de X.

Definición 6.2.1. Sea S un subconjunto de X.

- Un punto x de X es un ***punto interior*** de S si hay un número positivo r tal que $B_r(x) \subseteq S$.
- El ***interior*** de S es el conjunto, que escribimos S° o $\operatorname{int}(S)$, de todos los puntos de X que son interiores a S.
- El conjunto S es ***abierto*** si $S = S^\circ$.

A veces escribimos $\operatorname{int}_X(S)$ en lugar de simplemente $\operatorname{int}(S)$ para dejar de manifiesto el espacio métrico en el que estamos trabajando. Notemos que, como $x \in B_r(x)$ cualquiera sean $x \in X$ y $r \in (0, +\infty)$, es claro que $S^\circ \subseteq S$ para todo subconjunto S de X. En vista de esto, para ver que un conjunto es abierto es suficiente con mostrar que está contenido en su interior.

Antes que nada, demos ejemplos de conjuntos abiertos.

Proposición 6.2.2. *Sea x un elemento de X y sea r un número positivo. La bola abierta $B_r(x)$ es un conjunto abierto.*

Demostración. Sea y un elemento de $B_r(x)$, de manera que $d(x,y) < r$, y sea $s \coloneqq r - d(x,y)$, que es un número positivo. Si z es un elemento de $B_s(y)$, de manera que $d(y,z) < r$, entonces

$$d(x,z) \leq d(x,y) + d(y,z) < d(x,y) + r = d(x,y) = r,$$

así que $z \in B_r(x)$. Esto nos dice que $B_s(z) \subseteq B_r(x)$ y, por lo tanto, que $y \in B_r(x)^\circ$. Podemos concluir entonces que $B_r(x) \subseteq B_r(x)^\circ$ y, en definitiva, que la bola abierta $B_r(x)$ es un conjunto abierto de X. □

El siguiente resultado da las propiedades básicas de la función $S \in \mathscr{P}(X) \mapsto S^\circ \in \mathscr{P}(X)$.

Proposición 6.2.3. *Si S y T son dos subconjuntos de X, entonces*

(*i*) $S^\circ \subseteq S$,
(*ii*) $S \subseteq T \implies S^\circ \subseteq T^\circ$,
(*iii*) $(S^\circ)^\circ = S^\circ$,
(*iv*) $S^\circ \cap T^\circ = (S \cap T)^\circ$,
(*v*) $S^\circ \cup T^\circ \subseteq (S \cup T)^\circ$.

La inclusión de (*v*) no es, en general, una igualdad. Por ejemplo, si $X = \mathbb{R}$ con su métrica euclídea, $S = [0,1]$ y $T = [1,2]$, entonces $S^\circ \cup T^\circ = (0,1) \cup (1,2)$ y $(S \cup T)^\circ = (1,2)$, y estos dos

conjuntos son distintos. De manera similar, la implicación recíproca a la de (*ii*) es en general falsa: en $\mathbb{R}$ podemos elegir $S = [0,1]$ y $T = (0,1)$ para verlo.

Demostración. Sean S y T dos subconjuntos de X.

(*i*) Si x es un elemento de S°, entonces hay un número positivo r tal que $B_r(x) \subseteq S$, y entonces tenemos que $x \in B_r(x) \subseteq S$. Vemos así que $S^\circ \subseteq S$.

(*ii*) Supongamos que $S \subseteq T$. Si x es un elemento de S°, de manera que hay un número positivo r tal que $B_r(x) \subseteq S$, entonces $B_r(x) \subseteq S \subseteq T$ y, por lo tanto, $x \in T^\circ$. Esto prueba que $S^\circ \subseteq T^\circ$.

(*iii*) De (*i*) sabemos que $S^\circ \subseteq S$, así que (*ii*) implica que $(S^\circ)^\circ \subseteq S^\circ$. Sea, por otro lado, x un elemento de S°, de manera que hay un número positivo r tal que $B_r(x) \subseteq S$ y, por lo tanto, $B_r(x) = B_r(x)^\circ \subseteq S^\circ$, de manera que $x \in (S^\circ)^\circ$. Esto prueba que $S^\circ \subseteq (S^\circ)^\circ$ y, juntando todo, que $S^\circ = (S^\circ)^\circ$.

(*iv*) Como $S \cap T \subseteq S$ y $S \cap T \subseteq T$, sabemos de (*ii*) que $(S \cap T)^\circ \subseteq S^\circ$ y que $(S \cap T)^\circ \subseteq T^\circ$, así que, $(S \cap T)^\circ \subseteq S^\circ \cap T^\circ$. Por otro lado, si x es un elemento de $S^\circ \cap T^\circ$, entonces x es un elemento de S° y existe un número positivo r tal que $B_r(x) \subseteq S$, y x es un elemento de T° y existe un número positivo s tal que $B_s(x) \subseteq T$: el número $t \coloneqq \min\{s,t\}$ es un número positivo y $B_t(x) \subseteq B_s(x) \subseteq S$ y $B_t(x) \subseteq B_r(x) \subseteq T$, de manera que $B_t(x) \subseteq S \cap T$ y, por lo tanto, $x \in (S \cap T)^\circ$. Esto prueba que $S^\circ \cap T^\circ \subseteq (S \cap T)^\circ$ y, en definitiva, que $S^\circ \cap T^\circ = (S \cap T)^\circ$, como afirma la proposición.

(*v*) Sea x un elemento de $S^\circ \cup T^\circ$, de manera que $x \in S^\circ$ o $x \in T^\circ$: hay entonces un número positivo r tal que $B_r(x) \subseteq S$ o $B_r(x) \subseteq T$ y, en cualquier caso, $B_r(x) \subseteq S \cup T$, por lo que $x \in (S \cup T)^\circ$. Se sigue de esto que $S^\circ \cup T^\circ \subseteq (S \cup T)^\circ$. □

La siguiente es una consecuencia importante de la Proposición 6.2.3.

Proposición 6.2.4. *El interior de un subconjunto de X es abierto y es, de hecho, el mayor subconjunto abierto de X contenido en él.*

Demostración. Sea S un subconjunto de X. Sabemos que S° es un abierto de X y que está contenido en S. Por otro lado, si U es un abierto contenido en S, de manera que $U \subseteq S$, entonces tenemos que $U = U^\circ \subseteq S^\circ$. Esto prueba la proposición. □

En un grupo abeliano dotado de una métrica invariante por translaciones, la clase de los conjuntos abiertas es cerrada por translaciones en el siguiente sentido.

Proposición 6.2.5. *Sea G un grupo abeliano y sea d una métrica sobre G que es invariante por translaciones. Un subconjunto U de G es abierto si y solamente si el conjunto $x + U$ es abierto cualquiera sea el elemento x de G.*

Demostración. Sea U un subconjunto de G. Si para todo elemento x de U el conjunto $x + U$ es abierto, entonces en particular el conjunto $U = 0 + U$ es abierto, y esto nos dice que la condición de la proposición es suficiente.

Para ver que es necesaria, supongamos que el conjunto U es abierto y sea x un elemento de G. Si y es un elemento de $x + U$, entonces hay un elemento u de U tal que $y = x + u$ y, como U es abierto, hay un número positivo r tal que $B_r(u) \subseteq U$: es fácil ver que $B_r(y) = x + B_r(u)$ y, por lo tanto, que $B_r(y) = x + B_r(u) \subseteq x + U$, de manera que y es un punto interior de U. □

Ejercicio 6.2.6. Sea G un grupo abeliano y sea d una métrica sobre G que es invariante por translaciones. Muestre que si S es un subconjunto de G y x es un elemento de G, entonces $(x + S)^\circ = x + S^\circ$.

6.3. Entornos y bases de entornos

Definimos qué es un conjunto abierto de un espacio métrico usando las bolas abiertas de este. Para muchas cosas es conveniente darnos un poco más de flexibilidad, y ese es el objetivo de la siguiente definición.

Definición 6.3.1. Un ***entorno*** de un punto x de X es un subconjunto N de X tal que $x \in N^\circ$.

Cuando queremos explicitar el espacio métrico en el que estamos trabajando decimos «un entorno de x en X» en lugar de simplemente «un entorno de x». Es importante observar que un entorno de un punto no es necesariamente un conjunto abierto.

Proposición 6.3.2. *Un subconjunto de un espacio métrico es abierto si y solamente si es un entorno de cada uno de sus puntos.*

Demostración. En efecto, esa condición dice que todo punto del conjunto pertenece al interior de este. □

Las propiedades de los conjuntos abiertos y de la operación de tomar interior por supuesto dan lugar a propiedades similares de los entornos de un punto.

Ejercicio 6.3.3. Sea x un punto de X. Muestre que (*i*) la intersección de un conjunto finito de entornos de x es un entorno de x y que (*ii*) la unión de una familia arbitraria de entornos de x es un entorno de x.

Un punto tiene, en general, muchos entornos. Para la mayor parte de las cosas que hacemos con entornos, sin embargo, nos alcanza con considerar muchos menos.

Definición 6.3.4. Una ***base de entornos*** de un punto x de X es un conjunto $\mathscr{N}$ de entornos de x que tiene la siguiente propiedad:

para cada entorno U de x hay un elemento N de $\mathscr{N}$ tal que $N \subseteq U$.

Una de las formas en las que usamos esta noción es vía la siguiente observación.

Proposición 6.3.5. *Sea x un punto de X y sea $\mathscr{N}$ una base de entornos de x. Un conjunto U es un entorno de x si y solamente si hay un elemento N de $\mathscr{N}$ tal que $N \subseteq U$.*

Demostración. Sea U un subconjunto de X. Si U es un entorno de x, entonces que $\mathscr{N}$ sea una base de entornos de x implica que hay un elemento N de $\mathscr{N}$ tal que $N \subseteq U$. Recíprocamente, si hay un elemento N de $\mathscr{N}$ tal que $N \subseteq U$, entonces tenemos que $x \in N^\circ \subseteq U^\circ$ y, por lo tanto, U es un entorno de x. □

Veamos algunos ejemplos de bases de entornos.

Proposición 6.3.6. *Sea x un punto de X. Los siguientes conjuntos son bases de entornos de x:*

(*i*) *el conjunto $\mathscr{N}_1$ de todos los entornos de x;*
(*ii*) *el conjunto $\mathscr{N}_2$ de todos los abiertos de X que contienen a x;*
(*iii*) *el conjunto $\mathscr{N}_3$ de todas las bolas abiertas centradas en x;*
(*iv*) *el conjunto $\mathscr{N}_4 \coloneqq \{B_r(x) : r \in R\}$, con R un subconjunto de $(0, +\infty)$ tal que* ínf $R = 0$.

Demostración. Es claro que los elementos de cada uno de los cuatro conjuntos descriptos en la proposición son entornos de x. Sea U un entorno de x. Para ver que esos cuatro conjuntos son bases de entornos de x, es suficiente que hagamos las siguientes observaciones.

- Es claro que $U \in \mathscr{N}_1$ y que $U \subseteq U$.
- Es $U^\circ \in \mathscr{N}_2$ y $U^\circ \subseteq U$.
- Como $x \in U^\circ$, hay un número positivo r tal que $B_r(x) \subseteq U$ y es $B_r(x) \in \mathscr{N}_2$.
- Como $x \in U^\circ$, hay un número positivo r tal que $B_r(x) \subseteq U$, y como ínf $R < r$, hay un elemento s de R tal que $s < r$, así que $B_s(x) \subseteq B_r(x) \subseteq U$ y $B_s(x) \in \mathscr{N}_4$. □

Un corolario remarcable de esta proposición es el siguiente.

Corolario 6.3.7. *Todo punto de un espacio métrico posee una base de entornos numerable.*

Decimos por esto que todo espacio métrico satisface el ***primer axioma de la numerabilidad***. Esta observación sencilla es extremadamente importante.

Demostración. Sea x un punto de X. Como el conjunto $\{1/n : n \in \mathbb{N}\}$ está contenido en $(0, +\infty)$ y su ínfimo es 0, la proposición nos dice que $\{B_{1/n}(x) : n \in \mathbb{N}\}$ es una base de entornos de un punto x de X. Por supuesto, este conjunto es numerable. □

Ejercicio 6.3.8. Sean (X, d_X) y (Y, d_Y) dos espacios métricos y dotemos al producto cartesiano $X \times Y$ o bien de la métrica sobre el conjunto $X \times Y$ que describe la Proposición 5.2.6.1, o bien de la métrica d_p correspondiente a un elemento p de $[1, +\infty)$ que describe el Ejercicio 5.2.6.3. Pruebe que para todo punto $z = (x, y)$ de $X \times Y$ los conjuntos

$$\{B_r^X(x) \times B_s^Y(y) : r, s \in (0, +\infty)\}$$

y

$$\{B_r^X(x) \times B_r^Y(y) : r \in (0, +\infty)\}$$

son bases de entornos de x en $X \times Y$.

Ejercicio 6.3.9. Sean X y Y dos espacios métricos y dotemos al producto cartesiano $X \times Y$ de la métrica que describe la Proposición 5.2.6.1. Pruebe que si U y V son subconjuntos abiertos de X y de Y, entonces $U \times V$ es un subconjunto abierto de $X \times Y$, y que, en general, no todo subconjunto abierto de $X \times Y$ se obtiene de esta forma.

El conjunto de los entornos de un punto es cerrado por intersecciones finitas pero, en general, esto no es cierto de una base de entornos cualquiera. Nuestro siguiente resultado es lo máximo que podemos decir sobre eso.

Proposición 6.3.10. *Sea $\mathscr{N}$ una base de entornos de un punto x de X. Si M y N son dos elementos de $\mathscr{N}$, entonces hay un elemento P de $\mathscr{N}$ tal que $P \subseteq M \cap N$.*

Demostración. Sean M y N dos elementos de $\mathscr{N}$. Tenemos entonces que $x \in M^\circ$ y $x \in N^\circ$ y, por lo tanto, que $x \in M^\circ \cap N^\circ = (M \cap N)^\circ$, así que $M \cap N$ es un entorno de x. Como $\mathscr{N}$ es una base de entornos de x, hay entonces un elemento P de M tal que $P \subseteq M \cap N$. □

Cuando tenemos un grupo abeliano dotado de una métrica invariante por translaciones, las bases de entornos del elemento neutro determinan las bases de entornos de todos los demás puntos. Esto suele ser muy práctico.

Proposición 6.3.11. *Sea G un grupo abeliano, sea d una métrica sobre G que es invariante por translaciones, y sea $\mathscr{N}_0$ una base de entornos del elemento neutro 0 de G. Para cada elemento x de G el conjunto $\{x + N : N \in \mathscr{N}_0\}$ es una base de entornos de x.*

Demostración. Sea x un elemento de G y sea $\mathscr{N} \coloneqq \{x + N : N \in \mathscr{N}_0\}$. Si N es un elemento de $\mathscr{N}_0$, entonces hay un número positivo r tal que $B_r(0) \subseteq N$ y, por lo tanto, $B_r(x) = x + B_r(0) \subseteq x + N$. Esto nos dice que todo elemento de $\mathscr{N}$ es un entorno de x. Por otro lado, si U es un entorno de x, entonces hay un número positivo r tal que $x + B_r(0) = B_r(x) \subseteq U$ y, como $\mathscr{N}_0$ es una base de entornos de 0, hay un elemento N de $\mathscr{N}_0$ tal que $N \subseteq B_r(0)$: tenemos entonces que $x + N \subseteq x + B_r(0) \subseteq U$. Esto nos dice que $\mathscr{N}$ es una base de entornos de x. □

Veamos un ejemplo importante de la situación de esta proposición.

Ejemplo 6.3.12. Sea G un grupo abeliano, sea $(H_n)_{n\geq 0}$ una sucesión de subgrupos de G tal que $H_0 = G$, $H_{n+1} \subseteq H_n$ para todo $n \in \mathbb{N}_0$, y $\bigcap_{n=0}^{\infty} H_n = 0$, y consideremos sobre G la métrica d invariante por translaciones que describe la Proposición 5.2.9.6. Afirmamos que

el conjunto $\{H_n : n \in \mathbb{N}_0\}$ *es una base de entornos de* 0 *en* G,

y para verlo es suficiente hacer las siguientes dos observaciones.

- Sea n un elemento de $\mathbb{N}_0$. Si $x \in B_{1/2^n}(0)$, entonces o bien $x = 0$ y, por supuesto, $x \in H_n$, o bien $x \neq 0$, y, como $2^{-v(x)} = d(0,x) < 2^{-n}$, es $v(x) > n$ y, por lo tanto, $x \in H_{v(x)} \subseteq H_n$. Esto nos dice que H_n es un entorno de 0 en G.
- Sea, por otro lado, U un entorno de 0 en G, de manera que hay un número positivo r tal que $B_r(0) \subseteq U$. Sea n un entero no negativo tal que $1/2^n < r$. Si x es un elemento de H_n, entonces o bien $x = 0$ y $d(0,x) = 0 < r$, o bien $x \neq 0$, $v(x) \geq n$, y, por lo tanto, $d(0,x) = 2^{-v(x)} \leq 2^{-n} < r$: esto nos dice que $H_n \subseteq U$.

Como consecuencia de esto y de la Proposición 6.3.11, podemos concluir que para todo elemento x de G vale que el conjunto $\{x + H_n : n \in \mathbb{N}_0\}$ es una base de entornos de x en G.

En general, «entender» un espacio métrico incluye tener una buena idea de cómo son las bases de entornos de sus puntos. Demos un par de ejemplos de esto en forma de dos ejercicios.

Ejercicio 6.3.13. Sea X un conjunto, sea $\mathscr{D}$ una familia numerable de semi-métricas sobre X que separa puntos, sea $\phi : \mathbb{N}_0 \to \mathscr{D}$ una función biyectiva, para cada $n \in \mathbb{N}_0$ escribamos $d_n \coloneqq \phi(n)$, sea $(a_n)_{n\geq 0}$ una sucesión de números positivos tal que la serie $\sum_{n=0}^{\infty} a_n$ converge, y consideremos sobre X la métrica d descripta en la Proposición 5.2.5.2.

Sea x un punto de X, para cada elección de n en $\mathbb{N}_0$ y de r en $(0, +\infty)$ consideremos el conjunto

$$B_r^n(x) \coloneqq \{y \in X : d_n(x,y) < r\},$$

que es la bola abierta centrada en x de radio r con respecto a la semi-métrica d_m, y sea

$$\mathscr{S} \coloneqq \{B^n_{1/(m+1)}(x) : m, n \in \mathbb{N}_0\}.$$

Pruebe que el conjunto $\mathscr{B}$ de las intersecciones finitas de elementos de $\mathscr{S}$ es una base de entornos de x en X, pero que el conjunto $\mathscr{S}$ mismo no es, en general, una base de entornos de x. Para esto último, puede ponerse en la situación del Ejemplo 5.2.5.5.

Es importante observar que la base de entornos $\mathscr{B}$ que describe este ejercicio no depende ni de la biyección $\phi : \mathbb{N}_0 \to \mathscr{D}$ ni de la sucesión $(a_n)_{n\geq 0}$ usadas para construir la métrica d, sino solamente de la familia $\mathscr{D}$ de semi-métricas sobre X con la que empezamos.

Ejercicio 6.3.14. Sea A un conjunto y consideremos sobre el conjunto $X \coloneqq A^{\mathbb{N}_0}$ de todas las sucesiones en A la métrica descripta en la Proposición 5.2.9.5. Sea $\lambda : \mathbb{N}_0 \to \mathbb{N}_0$ una biyección y cada vez que x e y son elementos distintos de X pongamos

$$E^\lambda(x,y) \coloneqq \{i \in \mathbb{N}_0 : x_{\lambda(j)} = y_{\lambda(j)} \text{ para cada } j \in [\![0, i-1]\!]\}.$$

(a) Muestre que si x e y son elementos distintos de X, entonces el conjunto $E^\lambda(x,y)$ es finito y no vacío, de manera que, en particular, podemos considerar su máximo elemento

$$v^\lambda(x,y) \coloneqq \max E^\lambda(x,y).$$

(b) Sea x un elemento de X y para cada $n \in \mathbb{N}_0$ sea

$$B^\lambda(x,n) \coloneqq \{x\} \cup \{y \in X \setminus \{x\} : v^\lambda(x,y) \geq n\}.$$

Pruebe que el conjunto $\mathscr{B} \coloneqq \{B^\lambda(x,n) : n \in \mathbb{N}_0\}$ es una base de entornos de x en X.

Notemos que si x es un punto de X y n un elemento de $\mathbb{N}_0$, entonces

- la bola $B_{1/(n+1)}(x)$ centrada en x de radio $1/(n+1)$ es el conjunto de todas las sucesiones en A que coinciden con x en al menos las primeras n componentes, mientras que
- el conjunto $B^\lambda(x,n)$ tiene por elementos las sucesiones en A que coinciden con x en al menos las primeras componentes, pero tomando a estas *de acuerdo al orden descripto por la función* λ.

Muchas propiedades importantes de los espacios métricos consisten en que cierta condición se satisfaga localmente, y por ello entendemos que todo punto tenga un entorno que la satisfaga o, a veces, que todo punto tenga toda una base de entornos cada uno de los cuales la satisface. Veremos más adelante varios ejemplos de esto — aquí presentaremos solamente el siguiente.

Definición 6.3.15. Decimos que un espacio vectorial real o complejo dotado de una métrica invariante por translaciones es ***localmente convexo*** si cada uno de sus puntos posee una base de entornos cuyos elementos son conjuntos convexos.

La invariancia por translaciones de la métrica implica que es suficiente verificar esta condición en un punto, de hecho.

Ejercicio 6.3.16. Muestre que un espacio vectorial real o complejo dotado de una métrica invariante por translaciones es localmente convexo si y solamente si su elemento nulo posee una base de entornos cuyos elementos son conjuntos convexos.

La siguiente proposición da muchos ejemplos de espacios con esta propiedad.

Proposición 6.3.17. *Un espacio normado es localmente convexo.*

Demostración. En efecto, todo punto de un espacio normado posee una base de entornos cuyos elementos son bolas abiertas y el resultado del Ejercicio 6.1.16 nos dice que estas son subconjuntos convexos. □

Otra fuente importante de ejemplos de espacios vectoriales métricos localmente convexos es la construcción del Corolario 5.2.5.4.

Ejercicio 6.3.18. Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, sea V un $\mathbb{k}$-espacio vectorial, sea $\mathscr{N}$ una familia numerable de semi-normas sobre V que separa puntos, sea $\phi : \mathbb{N}_0 \to \mathscr{N}$ una función biyectiva, para cada $n \in \mathbb{N}_0$ escribamos $\|-\|_n \coloneqq \phi(n)$, sea $(a_n)_{n\geq 0}$ una sucesión de números reales positivos tal que la serie $\sum_{n=0}^{\infty} a_n$ converge, y consideremos sobre V la métrica invariante por translaciones d que en cada elemento (x, y) de $V \times V$ toma el valor

$$d : (x, y) \in X \times X \mapsto \sum_{n=0}^{\infty} a_n \frac{\|x - y\|_n}{1 + \|x - y\|_n} \in \mathbb{R},$$

descripta en el Corolario 5.2.5.4. Pruebe que con respecto a esta métrica el espacio V es localmente convexo.

Proposición 6.3.19. *Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, y sea p un elemento de $(0,1)$. El espacio vectorial métrico $\ell^p_{\mathbb{k}}$ que presentamos en el Ejercicio 5.2.3.30 no es localmente convexo.*

Demostración. Supongamos que, por el contrario, el espacio $\ell^p_{\mathbb{k}}$ sí es localmente convexo y que, en particular, su bola unidad $B_1(0)$ contiene un entorno convexo U de 0. Como U es un entorno de 0, hay por lo tanto un número positivo r tal que $B_{2r}(0) \subseteq U$.

Para cada $n \in \mathbb{N}_0$ sea e_n el elemento de $\ell^p_{\mathbb{k}}$ cuya única componente no nula es la n-ésima, que es igual a 1. Si m es un elemento de $\mathbb{N}_0$, entonces $r^{1/p} e_n \in B_{2r}(0) \subseteq U$ para cada $n \in [\![0, m]\!]$, así que

la convexidad de U y el resultado del Ejercicio 5.2.3.8 implican que el vector

$$x_m := \frac{1}{m+1}\sum_{n=0}^{m} r^{1/p} e_n$$

es un elemento de U y, en consecuencia, de $B_1(0)$. Esto nos dice que para cada $m \in \mathbb{N}_0$ es

$$1 > d(0, x_m) = r\sum_{n=0}^{m} \frac{1}{(m+1)^p} = r \cdot (m+1)^{1-p},$$

y esto es absurdo, ya que $1 - p > 0$. Esta contradicción prueba la proposición. □

6.4. La topología de un espacio métrico

La colección de todos los conjuntos abiertos del espacio métrico X juega un rol importante, así que le damos un nombre.

Definición 6.4.1. La ***topología*** del espacio métrico (X, d) es el conjunto τ_X de todos los subconjuntos abiertos de X.

Si necesitemos explicitar cuál es la métrica del espacio métrico escribimos $\tau_{X,d}$ en lugar de τ_X.

Las propiedades más elementales de la topología de un espacio métrico son las que lista el siguiente resultado.

Proposición 6.4.2.
(*i*) *$\varnothing$ y X son elementos de τ_X.*
(*ii*) *Si U y V son elementos de τ_X, entonces $U \cap V$ también lo es.*
(*iii*) *Si $\mathscr{U}$ es un subconjunto de τ_X, entonces $\bigcup_{U\in\mathscr{U}} U \in \tau_X$.*

Demostración. Como $\varnothing \subseteq \varnothing^\circ$, es claro que $\varnothing$ es un subconjunto abierto de X. Por otro lado, si x es un elemento cualquiera de X, entonces ciertamente $B_1(x) \subseteq X$, así que $x \in X^\circ$: esto nos dice que $X \subseteq X^\circ$ y, por lo tanto, que X es un abierto de X. Esto prueba la afirmación (*i*).

Si U y V son elementos de τ_X, de manera que $U = U^\circ$ y $V = V^\circ$, entonces la Proposición 6.2.3 nos dice que $(U \cap V)^\circ = U^\circ \cap V^\circ = U \cap V$, de manera que $U \cap V \in \tau_X$. Vemos así que vale la afirmación (*ii*).

Finalmente, para probar la afirmación (*iii*), consideremos un subconjunto $\mathscr{U}$ de τ_X. Si x es un elemento de $\bigcup_{U\in\mathscr{U}} U$, entonces hay un elemento V de $\mathscr{U}$ tal que $x \in V$ y, como V es abierto en X, hay un número positivo r tal que $B_r(x) \subseteq V \subseteq \bigcup_{U\in\mathscr{U}} U$. Esto nos dice que $\bigcup_{U\in\mathscr{U}} U \subseteq (\bigcup_{U\in\mathscr{U}} U)^\circ$

y, por lo tanto, que $\bigcup_{U\in\mathscr{U}} U \in \tau_X$. □

La tercera parte de esta proposición nos dice que la unión de una familia arbitraria de abiertos de X es abierta. La segunda, por su parte, implica — vía un argumento inductivo bien sencillo — que la intersección de una familia *finita* de abiertos de X es abierta. Esto es lo máximo que podemos esperar en general, como muestra el siguiente ejemplo.

Ejemplo 6.4.3. Consideremos el espacio métrico $\mathbb{R}$. Para cada entero positivo n el conjunto $(-1/n, 1/n)$ es abierto, ya que no es otra cosa que la bola abierta $B_{1/n}(0)$, y la intersección

$$\bigcap_{n=1}^{\infty}(-1/n, 1/n) = \{0\}$$

no es un abierto de $\mathbb{R}$.

Una aplicación de estas propiedades de la topología de un espacio métrico es la siguiente caracterización del interior de un conjunto es muchas veces útil para identificarlo.

Proposición 6.4.4. *El interior de un subconjunto S de X es la unión de todos los subconjunto abiertos de X contenidos en S.*

Demostración. Sea S un subconjunto de X y sea V la unión de todos los subconjuntos abiertos de X contenidos en S. Es consecuencia de la tercera parte de la Proposición 6.4.2 que V es un subconjunto abierto de X, y es claro que está contenido en S, así que tenemos que $V = V^\circ \subseteq S^\circ$. Por otro lado, S° es un abierto de X contenido en X, tenemos que $S^\circ \subseteq U$. Juntando todo vemos que $V = S^\circ$, que es lo que afirma la proposición. □

Todo subconjunto de un espacio métrico X con métrica discreta es abierto, así que la topología de X es el conjunto de partes $\mathscr{P}(X)$. De acuerdo a la siguiente definición, decimos en ese caso que la topología es discreta.

Definición 6.4.5. Sea X un espacio métrico. Decimos que la topología τ_X es ***discreta*** si $\tau_X = \mathscr{P}(X)$.

Hay espacios métricos con topología discreta que no tienen métricas discretas.

Ejercicio 6.4.6. Muestre que un espacio métrico X tiene topología discreta si y solamente si para todo elemento x de X es $\inf\{d(x, y) : y \in X \setminus \{x\}\} > 0$.

En vista de este resultado, la topología del subespacio $\{1/n : n \in \mathbb{N}\}$ de $\mathbb{R}$ es discreta.

6.4.1. Bases

Una consecuencia de la forma en que definimos el interior de un conjunto es que podemos construir cualquier conjunto abierto a partir de bolas.

Proposición 6.4.1.1. *Todo conjunto abierto es unión de bolas y, de hecho, es la unión de todas las bolas abiertas que contiene.*

Demostración. Sea U un subconjunto abierto de X, sea $\mathscr{U}$ el conjunto de todas las bolas abiertas contenidas en U, y sea $V \coloneqq \bigcup_{B\in\mathscr{U}} B$. Tenemos que probar que $U = V$. Es claro que $V \subseteq U$, ya que cada elemento de $\mathscr{U}$ está contenido en U. Por otro lado, si x es un elemento de U, entonces que U sea abierto implica que hay un número positivo r tal que $B_r(x) \subseteq U$, así que $B_r(x) \in \mathscr{U}$ y, por lo tanto, $x \in B_r(x) \subseteq \bigcup_{B\in\mathscr{U}} B = V$. Esto nos dice que $U \subseteq V$. □

Esta proposición nos dice que el conjunto de todas las bolas es una base de la topología de X, en el sentido de la siguiente definición.

Definición 6.4.1.2. Un subconjunto $\mathscr{B}$ de la topología τ_X de X es una ***base*** de τ_X si todo elemento de τ_X es unión de elementos de $\mathscr{B}$.

Veamos otros ejemplos.

Ejemplo 6.4.1.3. El conjunto de los intervalos abiertos de $\mathbb{R}$ es una base de su topología. En efecto, un subconjunto de $\mathbb{R}$ es un intervalo abierto exactamente cuando es una bola abierta.

Ejercicio 6.4.1.4. Sean X y Y dos espacios métricos y dotemos al producto cartesiano $X \times Y$ de la métrica que describe la Proposición 5.2.6.1. Muestre que los conjuntos

$$\{B_r^X(x) \times B_s^Y(y) : x \in X, y \in Y, r, s \in (0, +\infty\}$$

y

$$\{B_r^X(x) \times B_r^Y(y) : x \in X, y \in Y, r \in (0, +\infty\}$$

es una base de la topología de $X \times Y$.

En general, hay muchas bases para la topología de un espacio métrico. La siguiente proposición nos da una forma de reconocerlas.

Proposición 6.4.1.5. *Un subconjunto $\mathscr{B}$ de la topología τ_X de X es una base de τ_X si y solamente si para todo abierto U de X y para todo elemento x de U hay un elemento B de $\mathscr{B}$ tal que $x \in B \subseteq U$.*

Demostración. Sea $\mathscr{B}$ un subconjunto de τ_X, y supongamos primero que la condición de la proposición se satisface. Sea U un subconjunto abierto de X. Para cada elemento x de U la hipótesis nos dice que hay un elemento B_x de $\mathscr{B}$ tal que $x \in B_x \subseteq U$: esto implica inmediatamente que $U \subseteq \bigcup_{x\in U} B_x \subseteq U$, de manera que, de hecho, $U = \bigcup_{x\in U} B_x$. Vemos así que $\mathscr{B}$ es una base de la topología τ_X.

Supongamos ahora que el conjunto $\mathscr{B}$ es una base de la topología τ_X y sean U un subconjunto abierto de X y x un elemento de U. Como $\mathscr{B}$ es una base de τ_X y U es abierto, hay un subconjunto $\mathscr{U}$ de $\mathscr{B}$ tal que $U = \bigcup_{B\in\mathscr{U}} B$ y, por lo tanto, hay un elemento B de $\mathscr{U}$ y, en consecuencia, de $\mathscr{B}$, tal que $x \in B \subseteq U$. Esto nos dice que la condición de la proposición es necesaria. □

Si ya tenemos una base para la topología de un espacio, el siguiente ejercicio nos da un criterio más eficiente que el de la Proposición 6.4.1.5 para verificar que otro conjunto de abiertos es una base para la topología de un espacio métrico.

Ejercicio 6.4.1.6. Sean $\mathscr{B}$ y $\mathscr{B}'$ dos subconjuntos de la topología τ_X del espacio métrico X. Pruebe que si

- $\mathscr{B}$ es una base de τ_X y
- para cada elemento U de $\mathscr{B}$ y cada elemento x de U hay un elemento B de $\mathscr{B}'$ tal que $x \in B \subseteq U$,

entonces $\mathscr{B}'$ es una base de τ_X.

Tener una base para la topología de un espacio métrico es conveniente porque nos permite describir todos los abiertos del espacio, y es razonable pensar que el cardinal mínimo de una base de la topología del espacio es una medida de la complejidad de este. Un caso extremo es aquel en el que hay alguna base finita: el siguiente lema muestra que ese caso no es particularmente muy interesante.

Lema 6.4.1.7. *Si la topología de un espacio métrico X tiene una base finita, entonces X es finito y todos sus subconjuntos son abiertos.*

Demostración. Supongamos que hay un base $\mathscr{B}$ para la topología de X que es finita.

Sea x un elemento de X. La intersección I_x de la familia $\mathscr{B}_x \coloneqq \{B \in \mathscr{B} : x \in B\}$ es un abierto de X, ya que los elementos de esa familia son abiertos y la familia es finita. Es claro que $x \in I_x$. Si y es un elemento de X distinto de x, entonces el número $r \coloneqq d(x, y)$ es positivo, así que la bola $B_r(x)$ es un abierto de X que contiene a x y, como B es una base de la topología de X, hay un elemento B de $\mathscr{B}$ tal que $x \in B \subseteq B_r(x)$: como $B \in \mathscr{B}_x$, tenemos que $I_x \subseteq B$ y, como $B \subseteq B_r(x) \not\ni y$, tenemos que $y \notin I_x$. Podemos concluir de esto que $I_x = \{x\}$ es un abierto de X.

Ahora bien, si V es un subconjunto cualquiera de X, entonces claramente $V = \bigcup_{x\in V}\{x\}$, y esto implica que V es abierto. Esto prueba que todo subconjunto de X es abierto. Por otro lado, si

x es un elemento de X, entonces ahora sabemos que $\{x\}$ es un abierto de X y, como $\mathscr{B}$ es una base, que hay un elemento B de $\mathscr{B}$ tal que $x \in B \subseteq \{x\}$: debe ser, $B = \{x\}$ y, por lo tanto, $\{x\} \in \mathscr{B}$. Vemos así que hay una función $x \in X \mapsto \{x\} \in \mathscr{B}$. Como se trata claramente de una función inyectiva, el conjunto X es finito. □

En vista de esto, el primer caso interesante es aquel en el que la topología del espacio métrico posee una base numerable.

Definición 6.4.1.8. Decimos que el espacio métrico X satisface el ***segundo axioma de la numerabilidad*** si la topología de X posee una base contable.

Veremos muchos ejemplos de espacios que satisfacen esta condición más adelante. Contentémonos con uno por ahora.

Ejemplo 6.4.1.9. Mostremos que el espacio métrico $\mathbb{R}$ satisface el segundo axioma de la numerabilidad. Sea $\mathscr{B}$ el conjunto de todos los intervalos abiertos (a, b) con extremos racionales. Es fácil ver que la función

$$(r, s) \in \mathbb{Q} \times \mathbb{Q}_+ \mapsto (r, r + s) \in \mathscr{B}$$

es biyectiva, así que este conjunto $\mathscr{B}$ es numerable. Probemos que se trata de una base para la topología de $\mathbb{R}$. Como sabemos que el conjunto de todas las bolas abiertas de $\mathbb{R}$ es una base, el resultado del Ejercicio 6.4.1.6 implica que es suficiente que mostremos que

> *si (a, b) es un intervalo abierto de $\mathbb{R}$ y x es un elemento de (a, b), entonces hay dos números racionales r y s tales que $x \in (r, s) \subseteq (a, b)$.*

Esto es inmediato: en efecto, como $a < x < b$, sabemos que hay números racionales r y s tales que $a < r < x$ y $x < s < b$, y entonces tenemos que $x \in (r, s) \in (a, b)$.

Si un espacio métrico satisface el segundo axioma de la numerabilidad, entonces podemos estimar su cardinal y el de su topología.

Proposición 6.4.1.10. *Si el espacio métrico X satisface el segundo axioma de la numerabilidad, entonces X y su topología τ_X tienen cardinal menor o igual que la potencia del continuo.*

Demostración. Supongamos que la topología τ_X de X admite una base numerable $\mathscr{B}$, de manera que la función

$$\phi : \mathscr{U} \in \mathscr{P}(\mathscr{B}) \mapsto \bigcup_{U \in \mathscr{U}} U \in \tau_X$$

es sobreyectiva. Esto implica, como sabemos, que el cardinal de τ_X es menor o igual que el

de $\mathscr{P}(\mathscr{B})$ y, como $|\mathscr{B}| \leq \aleph_0$, tenemos entonces que $|\tau_X| \leq |\mathscr{P}(\mathscr{B})| = 2^{|\mathscr{B}|} \leq 2^{\aleph_0} = \mathfrak{c}$.

Hay, por otro lado, una función

$$\psi : x \in X \mapsto \{U \in \mathscr{B} : x \in U\} \in \mathscr{P}(\mathscr{B})$$

y es inyectiva. En efecto, si x e y son dos elementos distintos de X, entonces hay un elemento B de $\mathscr{B}$ tal que $B_{d(x,y)}(x) \subseteq B$, es $B \in \psi(x)$ y, por lo tanto,

$$\bigcap_{U\in\psi(x)} U \subseteq B_{d(x,y)}(x) \not\ni y, \qquad y \in \bigcap_{U\in\psi(y)} U$$

de manera que $\bigcap_{U\in\psi(x)} U \neq \bigcap_{U\in\psi(y)} U$ y, por lo tanto, $\psi(x) \neq \psi(y)$. Como la función ψ es inyectiva, podemos concluir que $|X| \leq |\mathscr{P}(\mathscr{B})| = 2^{|\mathscr{B}|} \leq 2^{\aleph_0} = \mathfrak{c}$. □

Dar una base de una topología es una forma conveniente de describirla. A veces, podemos ser inclusive más económicos usando la siguiente noción.

Definición 6.4.1.11. Un subconjunto $\mathscr{S}$ de la topología τ_X de un espacio métrico X es una ***subbase*** de τ_X si el conjunto de las intersecciones finitas de elementos de $\mathscr{S}$ es una base de τ_X.

Cuando decimos «intersección finita» en este contexto nos referimos a la intersección de un subconjunto finito y *no vacío* de $\mathscr{S}$.

Demos un ejemplo sencillo de esto.

Ejemplo 6.4.1.12. El conjunto

$$\mathscr{S} \coloneqq \{(-\infty, a) : a \in \mathbb{R}\} \cup \{(a, +\infty) : a \in \mathbb{R}\}$$

es una subbase para la topología euclídea de $\mathbb{R}$. En efecto, para cada elección de a y b en $\mathbb{R}$ tal que $a < b$ tenemos que $(a, b) = (-\infty, b) \cap (a, +\infty)$, y sabemos que el conjunto de los intervalos abiertos es una base de la topología euclídea.

La razón por la que nos interesan las bases es que nos permiten comparar métricas de manera conveniente.

Proposición 6.4.1.13. *Sea X un conjunto, sea $\mathscr{D}$ una familia numerable de semi-métricas sobre X que separa puntos, sea $\phi : \mathbb{N}_0 \to \mathscr{D}$ una biyección, sea $(a_n)_{n\geq 0}$ una sucesión de números reales positivos tal que la serie $\sum_{n=0}^{\infty} a_m$ converge, y consideremos la función*

$$d : (x, y) \in X \times X \mapsto \sum_{n=0}^{\infty} a_n \frac{\phi(n)(x, y)}{1 + \phi(n)(x, y)} \in \mathbb{R},$$

que, de acuerdo a la Proposición 5.2.5.2, es una métrica sobre X. Para cada elección de x en X, de δ

en $\mathscr{D}$ y de ε en $(0,+\infty)$ pongamos $B(x,\delta,r) \coloneqq \{y \in X : \delta(x,y) < r\}$. El conjunto

$$\mathscr{S} \coloneqq \{B(x,\delta,r) : x \in X, \delta \in \mathscr{D}, r \in (0,+\infty)\}$$

es una subbase de la topología del espacio métrico (X,d).

El punto más importante de esta proposición es que la subbase $\mathscr{S}$ que describe no depende no de la elección de la biyección ϕ ni de la elección de la sucesión $(a_n)_{n\geq 0}$ y que, por lo tanto, la topología del espacio métrico (X,d) está completamente determinada por el conjunto numerable $\mathscr{D}$ de semi-métricas.

Demostración. Sea $\mathscr{B}$ el conjunto de las intersecciones finitas de elementos de $\mathscr{S}$. Tenemos que probar que $\mathscr{B}$ es una base para a topología τ_X del espacio métrico (X,d), y usamos para ello el criterio que nos da el Ejercicio 6.4.1.6.

Primero mostramos que los elementos de $\mathscr{B}$ son abiertos con respecto a d. Para ello es suficiente que probemos que los elementos de $\mathscr{S}$ son abiertos, claro, porque sabemos que la topología de X es cerrada por intersecciones finitas. Sea entonces x un elemento de X, δ un elemento de $\mathscr{D}$ y r un número positivo, y probemos que el conjunto $B(x,\delta,r)$ es abierto con respecto a la métrica d. Sea y un elemento de $B(x,\delta,r)$, de manera que $\delta(x,y) < r$. Sea $s \coloneqq a_m \cdot (r - \delta(y,x))$, que es un número positivo. Como la función ϕ es biyectiva, hay exactamente un entero m tal que $\delta = \phi(m)$. Si z es un elemento de X tal que $d(y,z) < s$, entonces

$$\begin{aligned}
\delta(x,z) &= \delta(x,y) + \delta(y,z) = \delta(x,y) + \frac{1}{a_m} \cdot a_m \phi(m)(y,z) \\
&\leq \delta(x,y) + \frac{1}{a_m} \cdot a_m \frac{\phi(m)(y,z)}{1+\phi(m)(y,z)} \leq \delta(x,y) + \frac{1}{a_m} \sum_{n=0}^{\infty} a_n \frac{\phi(n)(y,z)}{1+\phi(n)(y,z)} \\
&= \delta(x,y) + \frac{1}{a_m} \cdot d(y,z) < \delta(x,y) + (r - \delta(y,x)) = r.
\end{aligned}$$

Esto nos dice que $B_s(y) \subseteq B(x,\delta,r)$ y, en definitiva, prueba que el conjunto $B(x,\delta,r)$ es abierto.

En segundo lugar, consideramos un elemento x de X, un número positivo r, la bola abierta $B_r(x)$ centrada en x de radio r con respecto a la métrica d, y mostramos que hay un elemento B de $\mathscr{B}$ tal que $x \in B \subseteq B_r(x)$. Como la serie $\sum_{n=0}^{\infty} a_n$ converge y tiene términos positivos, hay un entero positivo N tal que $0 < \sum_{n=N+1}^{\infty} a_n < r/2$. Pongamos

$$B \coloneqq \bigcap_{n=0}^{N} B(x, \phi(n), r/2(N+1)a_n).$$

Si y es un elemento de B, entonces $\phi(n)(x,y) < r/2(N+1)a_n$ para cada $n \in [\![0,n]\!]$ y, por lo tanto,

$$d(x,y) = \sum_{n=0}^{N} a_n \frac{\phi(n)(x,y)}{1+\phi(n)(x,y)} + \sum_{n=N+1}^{\infty} a_n \frac{\phi(n)(x,y)}{1+\phi(n)(x,y)} \leq \sum_{n=0}^{N} a_n \frac{r}{2(N+1)a_n} + \sum_{n=N+1}^{\infty} a_n$$

$$< \frac{r}{2} + \frac{r}{2} = r.$$

Esto nos dice que B, que es un elemento de $\mathscr{B}$ que claramente contiene a x, está contenido en $B_r(x)$. Esto completa la demostración de la proposición. □

En general no es posible describir todos los conjuntos abiertos de un espacio métrico salvo en situaciones triviales. Una excepción a esto es la siguiente.

Proposición 6.4.1.14. *Todo abierto de $\mathbb{R}$ es la unión de una famila contable de intervalos abiertos disjuntos dos a dos.*

Demostración. Sea U un subconjunto abierto de $\mathbb{R}$. Si x e y son elementos de U, escribamos $x \vee y \coloneqq \text{máx}\{x, y\}$ y $x \wedge y \coloneqq \text{mín}\{x, y\}$, y consideremos el conjunto

$$R \coloneqq \{(x, y) \in U : [x \wedge y, x \vee y] \subseteq U\}$$

Mostremos que se trata de una relación de equivalencia en U.

- Sea x un elemento de U. Como $[x \wedge x, x \vee x] = \{x\} \subseteq U$, tenemos que $(x, x) \in R$.
- Si (x, y) es un elemento de R, entonces $[y \wedge x, y \vee x] = [x \wedge y, x \vee y] \subseteq U$, así que $(y, x) \in R$.
- Supongamos que x, y y z son tres elementos de U tales que $(x, y) \in R$ e $(y, z) \in R$, de manera que los intervalos $[x \wedge y, x \vee y]$ y $[y \wedge z, y \vee z]$ están contenidos en U. Los tres números x, y y z pueden ordenarse en orden creciente, y alguna de las desigualdades de la primera columna de la siguiente tabla vale, y en todos los casos podemos calcular la unión $[x \wedge y, x \vee y] \cup [y \wedge z, y \vee z]$ y verificar que contiene al intervalo $[x \wedge z, x \vee z]$.

	$[x \wedge y, x \vee y] \cup [y \wedge z, y \vee z]$
$x \leq y \leq z$	$[x, z]$
$x \leq z \leq y$	$[x, y]$
$z \leq x \leq y$	$[z, y]$
$z \leq y \leq x$	$[z, x]$
$y \leq z \leq x$	$[y, x]$
$y \leq x \leq z$	$[y, z]$

Esto implica claramente que en cualquier caso es $[x \wedge z, x \vee z] \subseteq U$ y $(x, z) \in R$.

Como R es una relación de equivalencia, el conjunto U es la unión disjunta de las clases de equivalencia de R. Para probar la proposición mostraremos que cada una de esas clases de equivalencia es un intervalo abierto y que el conjunto U/R de esas clases es contable.

Sea C una clase de equivalencia de R, y pongamos

$$\alpha \coloneqq \text{ínf } C, \qquad \beta \coloneqq \sup C,$$

haciendo la convención de que $\alpha = -\infty$ si el conjunto C no está acotado inferiormente, y que $\beta = +\infty$ si C no está acotado superiormente. Queremos probar que $c = (\alpha, \beta)$.

- Sea t un elemento de (α, β). Como $t > \alpha = \text{ínf } C$, hay un elemento a de C tal que $a < t$, y como $t < \beta = \sup C$, hay un elemento b de C tal que $t < b$. Como a y b están en C, $(a, b) \in R$ y, por lo tanto, $t \in [a, b] = [a \wedge b, a \vee b] \subseteq U$. Además $[a \wedge t, a \vee t] = [a, t] \subseteq [a, b] \subseteq U$ y entonces $(a, t) \in R$ y, en definitiva, $t \in C$. Vemos así que $(\alpha, \beta) \subseteq C$.
- Sea t un elemento de C. Como $C \subseteq U$ y U es abierto, hay un número positivo ε tal que $(t - \varepsilon, t + \varepsilon) \subseteq U$. Es $[(t - \varepsilon/2) \wedge t, (t - \varepsilon/2) \vee t] = [t - \varepsilon/2, t] \subseteq (t - \varepsilon, t + \varepsilon) \subseteq U$, tenemos que $(t - \varepsilon/2, t) \in R$ y, por lo tanto, $t - \varepsilon/2 \in C$. Como $t - \varepsilon/2 < t$, esto implica que $\alpha < t$. De manera similar podemos mostrar que $t < \beta$, así que $t \in (\alpha, \beta)$.

Para terminar, probemos que el conjunto de clases U/R es contable.

- Todo intervalo abierto de $\mathbb{R}$ contiene números racionales, así que para cada clase C de U/R la intersección $C \cap \mathbb{Q}$ no es vacío, y el *Axioma de elección* nos dice que hay una función $\gamma : U/R \to \mathbb{Q}$ tal que $\gamma(C) \in C$ para cada $C \in U/R$. Esta función es inyectia: si C y D son dos elemtos de U/R tales que $\gamma(C) = \gamma(D)$, entonces $\gamma(C) \in C \cap D$ y, como C y D son clases de equivalencia, es $C = D$. Como la función γ es inyectiva, el conjunto U/R es contable. □

6.4.2. La topología de un subespacio

Es extremadamente importante tener en cuenta que la condición de que un conjunto sea abierto es relativa al espacio métrico dentro del cual lo estamos considerando. Si Y es un subconjunto de nuestro espacio métrico X, entonces podemos considerar a Y, como siempre, como un espacio métrico con respecto a la métrica que se obtiene de la de X restringiéndola a Y: si S es un subconjunto de Y, entonces las afirmaciones «S es un subconjunto abierto de X» y «S es un subconjunto abierto de Y» no significan lo mismo. Así, por ejemplo, el conjunto $[0, 1)$ no es un subconjunto abierto de $\mathbb{R}$ pero sí es un subconjunto del subespacio $[0, 1]$ de $\mathbb{R}$.

La siguiente proposición caracteriza los subconjuntos abiertos de un subespacio.

Proposición 6.4.2.1. *Sea Y un subespacio de X. Un subconjunto U de Y es abierto en Y si y solamente si hay un subconjunto abierto V de X tal que $U = Y \cap V$.*

En términos de las topologías de X y de Y podemos enunciar este resultado como la igualdad

$$\tau_Y = \{Y \cap U : U \in \tau_X\}.$$

Demostración. Sea U un subconjunto de Y.

Supongamos primero que hay un subconjunto abierto V de X tal que $U = Y \cap V$ y mostremos que U es un subconjunto abierto de Y: esto probará la suficiencia de la condición dada por la proposición. Sea x un elemento de U. Como $x \in U = Y \cap V$, tenemos que $x \in V$ y, como V es un subconjunto abierto de X, hay un número positivo r tal que $B_r^X(x) \subseteq V$. Como entonces es $B_r^Y(x) = Y \cap B_r^X(x) \subseteq Y \cap V = U$, vemos que x pertenece al interior de U en Y. Esto es así cualquiera sea el elemento x de U, así que el conjunto U es abierto en Y.

Para probar ahora que la condición de la proposición es necesaria, consideremos un subconjunto abierto U de Y. Si x es un punto de U, entonces hay un número positivo r_x tal que $B_{r_x}^Y(x) \subseteq U$, y podemos considerar el subconjunto $V \coloneqq \bigcup_{x\in U} B_{r_x}(x)$ de X. Es claro que V es abierto en X, ya que es una unión de bolas abiertas. Por otro lado, es

$$Y \cap V = Y \cap \bigcup_{x\in U} B_{r_x}(x) = \bigcup_{x\in U} (Y \cap B_{r_x}(x)) = \bigcup_{x\in U} B_{r_x}^Y(x) \subseteq U$$

y para cada elemento y de U tenemos que

$$y \in B_{r_y}^Y(y) = Y \cap B_{r_y}(y) \subseteq Y \cap \bigcup_{x\in U} B_{r_x}(x) = Y \cap V,$$

de manera que $U = Y \cap V$, como queremos. □

Como consecuencia de esto, tenemos lo siguiente:

Proposición 6.4.2.2. *Sea Y un subespacio de X. Si S es un subconjunto de Y, entonces*

$$\operatorname{int}_Y(S) \subseteq \operatorname{int}_X(S) \cap Y$$

y

$$\operatorname{int}_Y(S) = \operatorname{int}_X(S \cup (X \setminus Y)) \cap Y.$$

La primera inclusión de esta proposición es, en general, estricta. Por ejemplo, si Y es un subconjunto no vacío de interior vacío de X, entonces $\operatorname{int}_Y(Y) = Y \neq \varnothing$ mientras que $\operatorname{int}_X(X) \cap Y = \varnothing$.

Demostración. Sea S un subconjunto de Y. Como $\operatorname{int}_X(S)$ es abierto en X, $\operatorname{int}_X(S) \cap Y$ es abierto en Y y como este conjunto está claramente contenido en S, tenemos que $\operatorname{int}_Y(S) \subseteq \operatorname{int}_X(S) \cap Y$.

Como $S \subseteq S \cup (X \setminus Y)$, tenemos que $\operatorname{int}_X(S) \cap Y \subseteq \operatorname{int}_X(S \cup (X \setminus Y)) \cap Y$, y entonces

$$\operatorname{int}_Y(S) \subseteq \operatorname{int}_X(S \cup (X \setminus Y)) \cap Y.$$

Por otro lado, como $\operatorname{int}_X(S \cup (X \setminus Y))$ es un abierto de X, la intersección $\operatorname{int}_X(S \cup (X \setminus Y)) \cap Y$ es un abierto de Y y, como

$$\operatorname{int}_X(S \cup (X \setminus Y)) \cap Y \subseteq (S \cup (X \setminus Y)) \cap Y = S \cap Y \cup (X \setminus Y) \cap Y = S,$$

así que $\text{int}_X(S \cup (X \setminus Y)) \cap Y \subseteq \text{int}_Y(S)$. Esto prueba la igualdad de la proposición. □

Otra consecuencia de la Proposición 6.4.2.1 es que es fácil construir bases para los subespacios de un subespacio.

Proposición 6.4.2.3. *Sea X un espacio métrico y sea Y un subespacio de X. Si $\mathscr{B}$ es una base de la topología de X, entonces el conjunto $\mathscr{B}_Y \coloneqq \{B \cap Y : B \in \mathscr{B}\}$ es una base de la topología de Y.*

Demostración. Sea $\mathscr{B}$ una base de la topología de X, sea $\mathscr{B}_Y$ el conjunto descripto en el enunciado de la proposición, sea U un subconjunto abierto de Y, y sea x un elemento de U. La Proposición 6.4.2.1 nos dice que hay un abierto V de X tal que $U = V \cap Y$. Como claramente x pertenece a V, que $\mathscr{B}$ sea una base de la topología de X implica que hay un elemento B de $\mathscr{B}$ tal que $x \in B \subseteq V$ y tenemos entonces que $x \in B \cap Y \subseteq V \cap Y = U$. Esto prueba que el conjunto $\mathscr{B}_Y$ es una base de la topología de Y, como queremos. □

Una consecuencia inmediata de esto es que la propiedad de satisfacer el segundo axioma de la numerabilidad es heredada por todos los subespacios.

Corolario 6.4.2.4. *Los subespacios de un espacio métrico que satisface el segundo axioma de la numerabilidad satisfacen el segundo axioma de la numerabilidad.*

Demostración. Sea X un espacio cuya topología posee una base contable $\mathscr{B}$ y sea Y un subespacio de X. Sabemos que el conjunto $\mathscr{B}_Y \coloneqq \{B \cap Y : B \in \mathscr{B}\}$ es una base de la topología de Y y, como la función $B \in \mathscr{B} \mapsto B \cap Y \in \mathscr{B}_Y$ es sobreyectiva, ese conjunto es contable. □

6.4.3. Métricas topológicamente equivalentes

Sobre un conjunto hay generalmente muchas métricas distintas, pero es bien posible que dos métricas den lugar a la misma topología.

Definición 6.4.3.1. Sea X un conjunto. Decimos que dos métricas d y d' sobre X son ***topológicamente equivalentes*** si las topologías correspondientes $\tau_{X,d}$ y $\tau_{X,d'}$ son iguales, y en ese caso escribimos $d \sim d'$.

Es claro que la relación de equivalencia topológica es una relación de equivalencia sobre el conjunto de todas las métricas sobre un conjunto.

Comparar directamente las topologías correspondientes a dos métricas es generalmente inviable. En la práctica, usamos el criterio de la siguiente proposición para establecer que dos métricas son topológicamente equivalentes.

Proposición 6.4.3.2. *Sea X un conjunto, sean d y d' dos métricas sobre X, y sean $\tau_{X,d}$ y $\tau_{X,d'}$ las correspondientes topologías. Si para todo elemento x de X y todo número positivo r hay un número positivo s tal que $B'_s(x) \subseteq B_r(x)$, entonces entonces $\tau_{X,d} \subseteq \tau_{X,d'}$.*

Aquí, cuando x es un elemento de X y r es un número positivo, estamos escribiendo $B_r(x)$ y $B'_r(x)$ a las bolas abiertas de centradas en x de radio r con respecto a d y a d', respectivamente.

Demostración. Supongamos que se satisface la hipótesis de la proposición, y sea U un subconjunto de X que es abierto para la métrica d. Si x es un elemento de U, hay entonces un número positivo r tal que $B_r(x) \subseteq U$ y, en vista de la hipótesis, hay un número positivo s tal que $B'_s(x) \subseteq B_r(x) \subseteq U$: esto nos dice que x es interior a U con respecto a la métrica d' y, como esto es así cualquiera sea el elemento x de U, que U es abierto con respecto a d'. Esto prueba que $\tau_{X,d} \subseteq \tau_{X,d'}$. □

Veamos un ejemplo de equivalencia topológica.

Proposición 6.4.3.3. *Sea X un conjunto y sea $d : X \times X \to \mathbb{R}$ una métrica sobre d. La función*

$$d' : (x, y) \in X \times X \mapsto \min\{d(x, y), 1\} \in \mathbb{R}$$

es una métrica sobre X que es topológicamente equivalente a d.

Demostración. Mostremos primero que la función d' es una métrica sobre X.

- Si x es un elemento de X, entonces $d'(x, x) = \min\{d(0, 0), 1\} = \min\{0, 1\} = 0$. Por otro lado, si x e y son dos elementos de X tales que $d'(x, y) = 0$, entonces tenemos que $\min\{d(x, y), 1\} = d'(x, y) = 0$ y, como $1 \not\leq 0$, que $d(x, y) = 0$, de manera que $x = y$.
- Si x e y son dos elementos de X, entonces

$$d'(x, y) = \min\{d(x, y), 1\} = \min\{d(y, x), 1\} = d'(y, x).$$

- Sean x, y, y z tres elementos de X. Si $d(x, z) \geq 1$ o $d(z, y) \geq 1$, entonces uno de los dos números $\min\{d(x, z), 1\}$ y $\min\{d(z, y), 1\}$ es igual a 1 y, como el otro es no negativo, tenemos que

$$\begin{aligned} d'(x, z) + d'(z, y) &= \min\{d(x, z), 1\} + \min\{d(z, y), 1\} \geq 1 \geq \min\{d(x, y), 1\} \\ &= d'(x, y). \end{aligned}$$

  Si, en cambio, es $d(x, z) < 1$ y $d(z, y) < 1$, entonces

$$\begin{aligned} d'(x, z) + d'(z, y) &= \min\{d(x, z), 1\} + \min\{d(z, y), 1\} = d(x, z) + d(z, y) \\ &\geq d(x, y) \geq \min\{d(x, y), 1\} = d'(x, y). \end{aligned}$$

  Vemos así que en cualquier caso es $d'(x, y) \leq d'(x, z) + d'(y, z)$.

Probemos ahora que las métricas d y d' son topológicamente equivalentes. Cuando x es un elemento de X y r es un número positivo, escribamos $B_r(x)$ y $B'_r(x)$ a las bolas abiertas de centradas en x de radio r con respecto a d y a d', respectivamente.

- Si x es un elemento de X y r es un número positivo, entonces $B_r(x) \subseteq B'_r(x)$. En efecto, si y es un elemento de $B_r(x)$, entonces $d(x,y) < r$ y, por lo tanto, $d'(x,y) = \min\{d(x,y),1\} < r$.
- Si x es un elemento de X, r es un número positivo, y ponemos $s \coloneqq \min\{r,1\}$, que también es un número positivo, entonces $B'_s(x) \subseteq B_r(x)$. En efecto, si y es un elemento de $B'_s(x)$, entonces $\min\{d(x,y),1\} = d'(x,y) < s = \min\{r,1\} \leq 1$: esto implica que $d(x,y) < 1$ y, por ello, que $d(x,y) = \min\{d(x,y),1\} < s = \min\{r,1\} \leq r$.

Estas dos observaciones y la Proposición 6.4.3.2 nos permiten concluir que $\tau_{X,d} = \tau_{X,d'}$, de manera que las métricas d y d' son topológicamente equivalentes. □

Notemos que esta proposición nos dice, en particular, que una métrica cualquiera es topológicamente equivalente a una métrica que es *acotada*. Hay muchas otras formas de verlo. El siguiente ejercicio da una más.

Ejercicio 6.4.3.4. Sea X un conjunto y sea $d : X \times X \to \mathbb{R}$ una métrica sobre d. La función

$$d' : (x,y) \in X \times X \mapsto \frac{d(x,y)}{1+d(x,y)} \in \mathbb{R}$$

es una métrica sobre X que es topológicamente equivalente a d.

Ejercicio 6.4.3.5. Sea X un conjunto y sea $d : X \times X \to \mathbb{R}$ una métrica sobre X.

(a) Encuentre condiciones sobre una función $f : \mathbb{R} \to \mathbb{R}$ suficientes para que la composición $f \circ d$ sea una métrica sobre X.

(b) Encuentre condiciones sobre una función $f : \mathbb{R} \to \mathbb{R}$ suficientes para que la composición $f \circ d$ sea una métrica sobre X topológicamente equivalente a d.

Ejercicio 6.4.3.6. Sean X y Y dos espacios métricos, sea d_∞ la métrica sobre $X \times Y$ que describe la Proposición 5.2.6.1, y para cada elemento p de $[1,+\infty)$ sea d_p la métrica sobre ese conjunto correspondiente a p descripta en el Ejercicio 5.2.6.3. Pruebe que todas estas métricas sobre $X \times Y$ son topológicamente equivalentes.

Demos algunas equivalencias topológicas entre métricas que son importantes.

Proposición 6.4.3.7. *Sea X un conjunto, sea $\mathscr{D}$ una familia numerable de semi-métricas sobre X que separa puntos, sean ϕ, $\psi : \mathbb{N}_0 \to \mathscr{D}$ dos biyecciones, sean $(a_n)_{n\geq 0}$ y $(b_n)_{n\geq 0}$ dos sucesiones de números reales positivos tales que las series $\sum_{n=0}^{\infty} a_m$ y $\sum_{n=0}^{\infty} b_m$ convergen, y consideremos las*

funciones

$$d : (x, y) \in X \times X \mapsto \sum_{n=0}^{\infty} a_n \frac{\phi(n)(x, y)}{1 + \phi(n)(x, y)} \in \mathbb{R}$$

y

$$d' : (x, y) \in X \times X \mapsto \sum_{n=0}^{\infty} b_n \frac{\psi(n)(x, y)}{1 + \psi(n)(x, y)} \in \mathbb{R}$$

que, de acuerdo a la Proposición 5.2.5.2, son métricas sobre X. Las métricas d y d' son topológicamente equivalentes.

Como consecuencia de esto, la topología que determinan esas métricas sobre X coinciden, y esto nos dice que esa topología depende solamente de la familia de semi-metricas $\mathscr{D}$ y no de la elección de sucesión $(a_n)_{n\geq 0}$ y de la biyección $\phi : \mathbb{N}_0 \to \mathscr{D}$ usadas para construir la métrica a partir de $\mathscr{D}$. Es por esto que nos referimos normalmente a esa topología como la ***topología determinada por la familia de semi-métricas $\mathscr{D}$***.

Demostración. Esto es consecuencia inmediata que la subbases para la topologías sobre X determinadas por d y d' descriptas en la Proposición 6.4.1.13 coinciden. □

El siguiente caso particular de la proposición es especialmente importante.

Corolario 6.4.3.8. *Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, sea V un $\mathbb{k}$-espacio vectorial, y sea $\mathscr{N}$ una familia numerable de semi-normas sobre V que separa puntos. Sean $\phi, \psi : \mathbb{N}_0 \to \mathscr{N}$ dos funciones biyectivas, sean $(a_n)_{n\geq 0}$ y $(b_n)_{n\geq 0}$ dos sucesiones de números reales positivos tales que las series $\sum_{n=0}^{\infty} a_n$ y $\sum_{n=0}^{\infty} b_n$ convergen, y consideremos las funciones*

$$d : (x, y) \in X \times X \mapsto \sum_{n=0}^{\infty} a_n \frac{\phi(n)(x - y)}{1 + \phi(n)(x - y)} \in \mathbb{R}$$

y

$$d' : (x, y) \in X \times X \mapsto \sum_{n=0}^{\infty} b_n \frac{\psi(n)(x - y)}{1 + \psi(n)(x - y)} \in \mathbb{R},$$

que, de acuerdo al Corolario 5.2.5.4, son métricas sobre X invariantes por translaciones. Las métricas d y d' son topológicamente equivalentes. □

6.4.4. Métricas equivalentes en el sentido de Lipschitz

Muchas veces para ver que dos métricas son topológicamente equivalentes mostramos que, de hecho, son equivalentes en el sentido de la siguiente definición, ya que, como veremos, esta es una condición más fuerte.

Definición 6.4.4.1. Sea X un conjunto y sean d y d' dos métricas sobre X. Decimos que d es ***equivalente en el sentido de Lipschitz*** a d' si hay un número positivo c tal que

$$\frac{1}{c} \cdot d(x,y) \leq d'(x,y) \leq c \cdot d(x,)$$

para toda elección de x e y en X, y en ese caso escribimos $d \sim_{\text{Lip}} d'$.

Antes que nada, veamos que esta relación entre métricas es una relación de equivalencia, justificando así su nombre.

Proposición 6.4.4.2. *Sea X un conjunto. La relación $\sim_{\text{Lip}}$ sobre el conjunto de todas las métricas sobre X es una relación de equivalencia.*

Demostración. Mostremos que la relación $\sim_{\text{Lip}}$ es reflexiva, simétrica y transitiva.

- Si d es una métrica sobre X, entonces claramente $1/1 \cdot d(x,y) \leq d(x,y) \leq 1 \cdot d(x,y)$ cualesquiera sean los elementos x e y de X, así que $d \sim_{\text{Lip}} d$. Esto nos dice que la relación $\sim_{\text{Lip}}$ es reflexiva.
- Si d y d' son dos métricas sobre X y $d \sim_{\text{Lip}} d'$, de manera que hay un número positivo c tal que $1/c \cdot d(x,y) \leq d'(x,y) \leq c \cdot d(x,y)$ para toda elección de x e y en X, entonces tenemos que $1/c \cdot d'(x,y) \leq d(x,y) \leq c \cdot d'(x,y)$ para toda elección de x e y en X y, por lo tanto, $d' \sim_{\text{Lip}} d$. Vemos con esto que la relación $\sim_{\text{Lip}}$ es simétrica.
- Sean d, d' y d'' tres métricas sobre X y supongamos que $d \sim_{\text{Lip}} d'$ y $d' \sim_{\text{Lip}} d''$, de manera que hay dos números positivos c y c' tales que $1/c \cdot d(x,y) \leq d'(x,y) \leq c \cdot d(x,y)$ y $1/c' \cdot d'(x,y) \leq d''(x,y) \leq c' \cdot d(x,y)$ para toda elección de x e y en X. Se sigue de esto que $1/cc' \cdot d'(x,y) \leq d''(x,y) \leq cc' \cdot d(x,y)$ para toda elección de x e y en X y, por lo tanto, que $d \sim_{\text{Lip}} d''$. La relación $\sim_{\text{Lip}}$ es, en consecuencia, transitiva. □

Muchas veces es conveniente la caracterización de la equivalencia de métricas en el sentido de Lipschitz que nos da el siguiente ejercicio.

Ejercicio 6.4.4.3. Sea X un conjunto. Pruebe que dos métricas d y d' sobre X son equivalentes en el sentido de Lipschitz si y solamente si hay números positivos a y b tales que $d(x,y) \leq a \cdot d'(x,y)$ y $d'(x,y) \leq b \cdot d(x,y)$ cualesquiera sean los elementos x e y de X.

Mostremos que, como dijimos, la equivalencia en el sentido de Lipschitz es una condición más fuerte que la de equivalencia topológica.

Proposición 6.4.4.4. *Sea X un conjunto y sean d y d' dos métricas sobre X. Si d y d' son equivalentes en el sentido de Lipschitz, entonces también son topológicamente equivalentes.*

Demostración. Supongamos que d y d' son equivalentes en el sentido de Lipschitz, de manera que hay un número positivo c tal que

$$\frac{1}{c} \cdot d(x, y) \leq d'(x, y) \leq c \cdot d(x,)$$

para toda elección de x e y en X. Si x es un elemento de X y r es un número positivo, entonces los números rc y r/c son positivos, y tenemos que

- si $y \in B_{r/c}(x)$, entonces $d'(x, y) \leq c \cdot d(x, y) < c \cdot r/c = r$, así que $y \in B'_r(x)$, y
- si $y \in B_{r/c}(x)$, entonces $d(x, y) \leq c \cdot d'(x, y) < c \cdot r/c = r$, así que $y \in B_r(x)$,

de manera que $B_{r/c}(x) \subseteq B'_r(x)$ y $B'_{r/c}(x) \subseteq B_r(x)$. Esto y la Proposición 6.4.3.2 nos permite concluir que las métricas d y d' son topológicamente equivalentes. □

Cuando tenemos dos métricas que provienen de normas sobre un espacio vectorial, podemos expresar la condición de que aquellas sean equivalentes en el sentido de Lipschitz en términos de estas.

Definición 6.4.4.5. Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, sea V un espacio vectorial sobre $\Bbbk$, y sean $\|-\|$ y $\|-\|'$ dos normas sobre V. Decimos que $\|-\|$ es ***equivalente*** a $\|-\|'$ si hay un número positivo c tal que

$$\frac{1}{c} \cdot \|x\| \leq \|x\|' \leq c \cdot \|x\|$$

para todo elemento x de V, y en ese caso escribimos $\|-\| \sim \|-\|'$.

Antes que nada, justifiquemos el nombre de esta relación entre normas.

Ejercicio 6.4.4.6. Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, y sea V un espacio vectorial sobre $\Bbbk$.

(a) Pruebe que la relación $\sim$ sobre el conjunto de todas las normas sobre V es una relación de equivalencia.

(b) Muestre que dos normas $\|-\|$ y $\|-\|'$ sobre V para las que hay números positivos a y b tales que $\|x\|' \leq a \cdot \|x\|$ y $\|x\| \leq b \cdot \|x\|'$ para todo elemento x de V son equivalentes.

La Definición 6.4.4.5 nos interesa porque vale lo siguiente.

Proposición 6.4.4.7. *Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, sea V un espacio vectorial sobre $\mathbb{k}$, sean $\|-\|$ y $\|-\|'$ dos normas sobre V, y sean d y d' las métricas sobre V que corresponden a ellas. Si las normas $\|-\|$ y $\|-\|'$ son equivalentes, entonces las métricas d y d' son equivalentes en el sentido de Lipschitz y, en particular, topológicamente equivalentes.*

Demostración. Supongamos que las normas $\|-\|$ y $\|-\|'$ son equivalentes, de manera que hay un número positivo c tal que

$$\frac{1}{c} \cdot \|x\| \leq \|x\|' \leq c \cdot \|x\|$$

para todo elemento x de V. Si x e y son dos elementos de V, entonces tenemos que

$$\frac{1}{c} \cdot d(x,y) = \frac{1}{c} \cdot \|x-y\| \leq \|x-y\|' = d'(x,y)$$

y que

$$d'(x,y) = \|x-y\|' \leq c \cdot \|x-y\| = c \cdot d(x,y),$$

y esto nos dice que las métricas d y d' son equivalentes en el sentido de Lipschitz. □

Para cada elemento p de $[1,+\infty]$ construimos una norma $\|-\|_p$ sobre $\mathbb{R}^n$ y sobre $\mathbb{C}^n$. El siguiente resultado nos dice que toda las normas que obtenemos de esta forma son equivalentes y que, por lo tanto, determinan métricas sobre esos espacios que son equivalentes en el sentido de Lipschitz.

Proposición 6.4.4.8. *Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, sea n un entero no negativo, y sean p y q dos elementos de $[1,+\infty]$. Las normas $\|-\|_p$ y $\|-\|_q$ sobre $\mathbb{k}^n$ son equivalentes.*

Veremos más adelante, en la Proposición 7.5.1 que, de hecho, *todas* las normas sobre $\mathbb{k}^n$ son equivalentes.

Demostración. Como la relación de equivalencia en el conjunto de todas las normas sobre $\mathbb{k}^n$ es una relación de equivalencia, para probar la proposición es suficiente que mostremos que $\|-\|_p \sim \|-\|_\infty$ para todo $p \in [1,+\infty)$.

Sea p un elemento de $[1,+\infty)$. Si x es un vector de $\mathbb{k}^n$, entonces $|x_j| \leq \max_{i \in [\![n]\!]} |x_i| = \|x\|_\infty$ para todo $j \in [\![n]\!]$, así que

$$\|x\|_p^p = \sum_{i=1}^n |x_i|^p \leq \sum_{i=1}^n \|x\|_\infty^p = n \cdot \|x\|_\infty^p,$$

de manera que $\|x\|_p \leq n^{1/p} \cdot \|x\|_\infty$. De manera similar, para cada $i \in [\![n]\!]$ es

$$|x_i|^p = |x_i|^p \leq \sum_{i=1}^{n} |x_i|^p = \|x\|_p^p,$$

así que

$$\|x\|_\infty^p = \left(\max_{i \in [\![n]\!]} |x_i|\right)^p = \max_{i \in [\![n]\!]} |x_i|^p \leq \|x\|_p^p$$

y, por lo tanto, $\|x\|_\infty \leq \|x\|_p$. Las dos desigualdad que obtuvimos y la segunda parte del Ejercicio 6.4.4.6 nos permiten concluir que las normas $\|-\|_\infty$ y $\|-\|_p$ son equivalentes. □

Ejercicio 6.4.4.9. Sean (X, d_X) y (Y, d_Y) dos espacios métricos, sea d_∞ la métrica sobre $X \times Y$ que describe la Proposición 5.2.6.1, y para cada elemento p de $[1, +\infty)$ sea d_p la métrica sobre ese conjunto correspondiente a p descripta en el Ejercicio 5.2.6.3. Pruebe que todas estas métricas son equivalentes en el sentido de Lipschitz.

Para cada $p \in [1, +\infty]$ construimos en la Subsección 5.2.3 un espacio normado $\ell_{\Bbbk}^p$. No tiene sentido preguntarse si sus normas son equivalentes o no, simplemente porque esos espacios son distintos dos a dos. Lo que sí podemos hacer es considerar el subconjunto

$$\Bbbk^\infty \coloneqq \{x \in \Bbbk^{\mathbb{N}_0} : \text{el conjunto } \{i \in \mathbb{N}_0 : x_i \neq 0\} \text{ es finito}\},$$

que es un subespacio de $\Bbbk^{\mathbb{N}_0}$ que está contenido en $\ell_{\Bbbk}^p$ para todo $p \in [1, +\infty]$, y comparar las normas sobre él que se obtienen restingiendo las de esos espacios.

Proposición 6.4.4.10. *Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$. Si p y q son dos elementos distintos de $[1, +\infty]$, entonces las métricas sobre $\Bbbk^\infty$ correspondientes a las normas $\|-\|_p$ y $\|-\|_q$ no son equivalentes.*

Esto muestra que un espacio vectorial de dimensión infinita — no es difícil ver que la dimensión de $\Bbbk^\infty$ es $\aleph_0$ — posee en general una gran cantidad de normas no equivalentes.

Demostración. Sea primero p un elemento de $[1, +\infty)$ y mostremos que las métricas d_p y d_∞ sobre $\Bbbk^\infty$ correspondientes a las normas $\|-\|_p$ y $\|-\|_\infty$ no son equivalentes. Para ello, es suficiente que mostremos que la bola abierta $B_1^p(0)$ de centro 0 y radio 1 con respecto a la distancia d_p no es abierta con respecto a la distancia d_∞. Supongamos que, por el contrario, sí lo es y, como ciertamente $0 \in B_1^p(0)$, que hay un número positivo r tal que $B_r^\infty(0) \subseteq B_1^p(0)$. Es fácil ver que

$$\lim_{n \to \infty} \left(\frac{r \cdot n}{2}\right)^{1/p} = +\infty,$$

así que hay un entero positivo N tal que $((r \cdot N)/2)^{1/p} > 1$. Consideremos el elemento y de $\Bbbk^{\mathbb{N}_0}$

que para cada $i \in \mathbb{N}_0$ tiene

$$y_i = \begin{cases} r/2 & \text{si } i < N; \\ 0 & \text{si } i \geq N. \end{cases}$$

Es claro que $y \in \Bbbk^\infty$ y que $\|y\|_\infty = r/2$, de manera que $y \in B_r^\infty(0)$. Por otro lado, es

$$\|y\|_p = \left(\sum_{i=0}^{\infty} |y_i|^p\right)^{1/p} = (r \cdot n/2)^{1/p} > 1,$$

por lo que $y \notin B_1^p(0)$. Esto es absurdo ya que $B_r^\infty(0) \subseteq B_1^p(0)$, y la contradicción provino de haber supuesto que el conjunto $B_1^p(0)$ es abierto con respecto a la distancia d_∞, así que no lo es.

Sean ahora p y q dos elementos de $[1, +\infty)$ tales que $p < q$. Vimos en el Ejercicio 5.2.3.27 que $\ell_\Bbbk^p \subsetneq \ell_\Bbbk^q$, así que podemos elegir un elemento a de $\ell_\Bbbk^q$ que no pertenece a $\ell_\Bbbk^p$. Esto significa que a es una sucesión $(a_i)_{i\geq 0}$ de elementos de $\Bbbk$ tal que la serie $\sum_{i=0}^\infty |a_i|^q$ converge y la serie $\sum_{i=0}^\infty |a_i|^p$ no. Claramente es $a \neq 0$, así que podemos suponer que $\|a\|_q = 1/2$ — si ese no es el caso, simplemente reemplazamos a a por $a/\|2a\|_q$.

Supongamos que las métricas d_p y d_q sobre el espacio $\Bbbk^\infty$ son equivalentes, de manera que la bola $B_1^p(0)$ correspondiente a la métrica d_p es abierta con respecto a la métrica d_q. En particular, como 0 ciertamente pertenece a esa bola, hay un número positivo r tal que $B_1^q(0) \subseteq B_1^p(0)$.

Ahora bien, como la serie de términos positivos $\sum_{i=0}^\infty |a_i|^p$ no converge, la sucesión de sus sumas parciales crece a $+\infty$ y existe, por lo tanto, un entero positivo N tal que $\sum_{i=0}^N |a_i|^p > 1/r$. Sea y el elemento de $\Bbbk^\infty$ que para cada $i \in \mathbb{N}_0$ tiene

$$(y_n)_i = \begin{cases} ra_i & \text{si } i \leq N; \\ 0 & \text{si } i > N. \end{cases}$$

Tenemos que

$$\|y\|_q = \left(\sum_{i=0}^{\infty} |y_i|^q\right)^{1/q} = \left(\sum_{i=0}^{N} |y_i|^q\right)^{1/q} = \left(\sum_{i=0}^{N} |ra_i|^q\right)^{1/q} \leq r\left(\sum_{i=0}^{\infty} |a_i|^q\right)^{1/q} = r \cdot \|a\|_q = r \cdot \frac{1}{2} < r,$$

así que $y \in B_r^q(0)$, y que

$$\|y\|_p = \left(\sum_{i=0}^{\infty} |y_i|^p\right)^{1/p} = \left(\sum_{i=0}^{N} |y_i|^p\right)^{1/p} = \left(\sum_{i=0}^{N} |ra_i|^p\right)^{1/p} = r\left(\sum_{i=0}^{N} |a_i|^p\right)^{1/p} > r \cdot \frac{1}{r} = 1,$$

de manera que $y \notin B_1^p(0)$. Esto es absurdo, ya que la forma en que elegimos a r implica que $B_r^q(0) \subseteq B_1^p(0)$, y esta contradicción provino de haber supuesto que las métricas d_p y d_q sobre $\Bbbk^\infty$ son equivalentes, así que no lo son. □

Hay muchas situaciones en las que la noción de equivalencia de métricas en el sentido de Lipschitz es importante inclusive si correspondiente noción de equivalencia topológica no da ninguna información. Por ejemplo, en la situación del siguiente ejercicio las métricas consideradas dan todas lugar a topologías discretas.

Ejercicio 6.4.4.11. Sea G un grupo.

(a) Sean S y S' dos subconjuntos finitos y simétricos que generan a G. Pruebe que

- las métricas de las palabras d_S y $d_{S'}$ correspondientes a S y a S' son equivalentes en el sentido de Lipschitz y que
- si $v_S, v_{S'} : \mathbb{N} \to \mathbb{N}$ son las funciones de crecimiento de G con respecto a S y a S', como en el Ejemplo 6.1.9, entonces hay un número c tal que $v_S(n/c) \leq v_{S'}(n) \leq n_S(cn)$ para todo $n \in \mathbb{N}_0$.

Deduzca de esto que el grupo G tiene crecimiento polinomial con respecto a S si y solamente si tiene crecimiento polinomial con respecto a S'.

Esto implica que cuando G posee un subconjuntos generador finito y simétrico con respecto al que tiene crecimiento polinomial tiene, de hecho, crecimiento polinomial con respecto a todos sus subconjuntos generadores finitos y simétricos: decimos en ese caso que el grupo ***tiene crecimiento polinomial***.

(b) Supongamos que el grupo G tiene crecimiento polinomial. Muestre que si S es un subconjunto generador finito y simétrico de G, entonces el conjunto

$$A_S \coloneqq \{\alpha \in (0, +\infty) : \text{hay un número positivo } c \text{ tal que } v_S(n) \leq cn^\alpha \text{ para todo } n \in \mathbb{N}_0\}$$

no es vacío, de manera que podemos considerar el número ínf A_S. Pruebe además que ese número ínf A_S depende solamente de G y no de la elección del subconjunto generador finito y simétrico S, de manera que podemos escribirlo simplemente deg G. Llamamos a ese número el ***grado de crecimiento*** de G. Mikhail Gromov probó en [Gro1981] que este número es un entero.

(c) Muestre que el grupo $\mathbb{Z}^n$ tiene crecimiento polinomial y que su grado de crecimiento es n.

6.5. Conjuntos cerrados

Fijemos, como siempre, un espacio métrico (X, d).

Definición 6.5.1. Sea S un subconjunto de X.

- Un punto x de X es un ***punto de adherencia*** de S si para todo número positivo r se tiene que $S \cap B_r(x) \neq \varnothing$.
- La ***clausura*** de S, que escribimos $\overline{S}$ o $\mathrm{cl}(S)$, es el conjunto de todos los puntos de X que son de adherencia de X.
- El conjunto S es ***cerrado*** si $S = \overline{S}$.

Cuando queramos dejar de manifiesto cuál es el espacio métrico en el que estamos trabajando escribiremos $\mathrm{cl}_X(S)$ en lugar de simplemente $\mathrm{cl}(S)$ o $\overline{S}$. Se sigue inmediato de las definiciones que todo conjunto está contenido en su clausura, así que un conjunto S es cerrado si y solo si $S \supseteq \overline{S}$.

Definimos los puntos de adherencia de un conjunto usando las bolas de la métrica, pero tenemos también la siguiente caracterización en términos de entornos.

Proposición 6.5.2. *Sea x un punto de X, sea $\mathscr{N}$ una base de entornos de x en X, y sea S un subconjunto de X. El punto x pertenece a la clausura de S si y solamente si todo elemento de $\mathscr{N}$ tiene intersección no trivial con S.*

Demostración. Si x pertenece a $\overline{S}$ y N es un elemento de $\mathscr{N}$, entonces $x \in N^\circ$, de manera que hay un número positivo r tal que $B_r(x) \subseteq N$, y, por lo tanto, $S \cap N \supseteq S \cap B_r(x) \neq \varnothing$. Esto prueba que la condición de la proposición es necesaria.

Veamos que es suficiente. Supongamos que la condición se cumple y sea r un número positivo. Como $B_r(x)$ es un entorno de x en X y $\mathscr{N}$ es una base de entornos de x en X, hay un elemento N de $\mathscr{N}$ tal que $N \subseteq B_r(x)$, y tenemos que $S \cap B_r(x) \supseteq S \cap N \neq \varnothing$. Esto prueba lo que queremos. □

La siguiente es otra caracterización útil de los puntos de adherencia de un conjunto, ahora en términos de la métrica.

Proposición 6.5.3. *Sea x un punto de X y sea S un subconjunto de X. El punto x pertenece a la clausura de S si y solamente si $d(x, S) = 0$.*

Demostración. Si x pertenece a la clausura de S, entonces para todo número positivo r la intersección $B_r(x) \cap S$ no es vacía, de manera que hay un elemento s de S tal que $d(x, S) \leq d(x, s) < r$: esto implica que $d(x, S) = 0$. Recíprocamente, si $d(x, S) = 0$, entonces para todo número positivo r es $\inf\{d(x, s) : s \in S\} = d(x, S) = 0 < r$, así que hay un elemento s de S tal que $d(x, s) < r$ y, por lo

tanto, $B_r(x) \cap S \neq \varnothing$: esto nos dice que x es un punto de adherencia de S. □

Antes de seguir, demos ejemplos no triviales de conjuntos cerrados.

Proposición 6.5.4. *Sea x un elemento de X y sea r un número no negativo. El conjunto*

$$\overline{B}_r(x) \coloneqq \{y \in X : d(x,y) \leq r\}$$

es cerrado.

Es importante observar que la barra en la notación $\overline{B}_r(x)$ *no* denota una clausura. Llamamos a $\overline{B}_r(x)$ la ***bola cerrada centrada en x de radio r***. Notemos que el radio de esa bola puede ser nulo y que $\overline{B}_0(x) = \{x\}$.

Demostración. Si y es un elemento de $X \setminus \overline{B}_r(x)$, entonces $d(x,y) > r$, el número $s \coloneqq d(x,y) - r$ es positivo, y $B_s(y) \cap \overline{B}_r(x) = \varnothing$: en efecto, si hubiera un punto z en esa intersección, tendríamos que $d(y,z) < s$, $d(x,z) \leq r$ y, por lo tanto, que $d(x,y) \leq d(x,z) + d(z,y) < s + r = d(x,y)$, lo que es imposible. Vemos así que ningún elemento de $X \setminus \overline{B}_r(x)$ es un punto de clausura de $\overline{B}_r(x)$, así que $\overline{\overline{B}_r(x)} \subseteq \overline{B}_r(x)$ y, en consecuencia, que el conjunto $\overline{B}_r(x)$ es cerrado. □

La siguiente es una propiedad importante de los espacios métricos.

Proposición 6.5.5. *Sea X un espacio métrico y sea x un elemento de X. El conjunto $\mathscr{N}$ de los entornos cerrados de x es una base de entornos de x en X.*

Decimos por esto que un espacio métrico es ***regular***.

Demostración. En efecto, si N es un entorno de x en X, entonces hay un número positivo r tal que $\overline{B}_{r/2}(x) \subseteq B_r(x) \subseteq N$ y $\overline{B}_{r/2} \in \mathscr{N}$. □

La siguiente observación muy simple también es importante.

Ejercicio 6.5.6. Sea X un espacio métrico, y consideremos al producto cartesiano $X \times X$ como un espacio métrico con respecta a la métrica de la Proposición 5.2.6.1. Pruebe que

$$\Delta_X \coloneqq \{(x,x) \in X \times X : x \in X\}$$

es un subconjunto cerrado de $X \times X$. Llamamos a Δ_X la ***diagonal*** de $X \times X$.

Veamos las propiedades más elementales de la función $S \in \mathscr{P}(X) \mapsto \overline{S} \in \mathscr{P}(X)$.

Proposición 6.5.7. *Si S y T son dos subconjunto de X, entonces*

(*i*) $S \subseteq \overline{S}$,
(*ii*) $S \subseteq T \implies \overline{S} \subseteq \overline{T}$,
(*iii*) $\overline{\overline{S}} = \overline{S}$,
(*iv*) $\overline{S \cup T} = \overline{S} \cup \overline{T}$,
(*v*) $\overline{S \cap T} \subseteq \overline{S} \cap \overline{T}$.

En general, la inclusión de (*v*) no es una igualdad: por ejemplo, si $X = \mathbb{R}$, $S = [0,1)$ y $T = (1,2]$, entonces $\overline{S \cap T} = \overline{\varnothing} = \varnothing$ y $\overline{S} \cap \overline{T} = [0,1] \cap [1,2] = \{1\}$.

Demostración. Sean S y T dos subconjuntos de X.

(*i*) Si x es un elemento de S, entonces para todo número positivo tenemos que $x \in S \cap B_r(x)$, de manera que $S \cap B_r(x) \neq \varnothing$: esto nos dice que $x \in \overline{S}$ y, por lo tanto, que $S \subseteq \overline{S}$.

(*ii*) Supongamos que $S \subseteq T$. Si x es un elemento de $\overline{S}$, entonces para cada número positivo r tenemos que $B_r(x) \cap T \supseteq B_r(x) \cap S \neq \varnothing$, de manera que $x \in \overline{T}$. Vemos así que $\overline{S} \subseteq \overline{T}$.

(*iii*) Como $S \subseteq \overline{S}$ por (*i*), sabemos de (*ii*) que $\overline{\overline{S}} \subseteq \overline{S}$. Si por otro lado, x es un elemento de $\overline{S}$, entonces para cada número positivo r tenemos que $\overline{S} \cap B_r(x) \supseteq S \cap B_r(x) \neq \varnothing$, de manera que $x \in \overline{\overline{S}}$. Podemos concluir de esto que $\overline{S} \subseteq \overline{\overline{S}}$ y, en definitiva, que $\overline{S} = \overline{\overline{S}}$.

(*iv*) Como $S \subseteq S \cup T$ y $T \subseteq S \cup T$, sabemos de (*ii*) que $\overline{S} \subseteq \overline{S \cup T}$ y $\overline{T} \subseteq \overline{S \cup T}$, así que $\overline{S} \cup \overline{T} \subseteq \overline{S \cup T}$. Sea, por otro lado, x un elemento de X que no pertenece a $\overline{S} \cup \overline{T}$. Como $x \notin \overline{S}$ hay un número positivo r tal que $S \cap B_r(x) = \varnothing$, y como $x \notin \overline{T}$ hay un número positivo s tal que $T \cap B_r(x) = \varnothing$. Si ponemos $t \coloneqq \min\{r, s\}$, que es un número positivo, tenemos entonces que

$$(S \cup T) \cap B_t(x) = S \cap B_t(x) \cup T \cap B_t(x) \subseteq S \cap B_t(r) \cup T \cap B_t(s) = \varnothing,$$

de manera que $x \notin \overline{S \cup T}$. Esto prueba que $\overline{S \cup T} \subseteq \overline{S} \cup \overline{T}$.

(*v*) Si x es un elemento de $\overline{S \cap T}$, entonces para todo número positivo r tenemos que $S \cap B_r(x) \subseteq (S \cup T) \cap B_r(x) \neq \varnothing$ y $T \cap B_r(x) \subseteq (S \cup T) \cap B_r(x) \neq \varnothing$: esto nos dice que $x \in \overline{S}$ y $x \in \overline{T}$, así que $x \in \overline{S} \cap \overline{T}$. □

Una consecuencia de esto es la siguiente caracterización de la clausura de un conjunto.

Proposición 6.5.8. *La clausura de un conjunto es un conjunto cerrado y es, de hecho, el menor conjunto cerrado que lo contiene.*

Demostración. Sea S un subconjunto de X. La parte (*iii*) de la Proposición 6.5.7 nos dice que $\overline{\overline{S}} = \overline{S}$, así que la clausura $\overline{S}$ es un conjunto cerrado. Por otro lado, si F es un conjunto cerrado de X tal que $S \subseteq F$, entonces la parte (*ii*) de la proposición nos dice que $\overline{S} \subseteq \overline{F} = F$. □

Ejercicio 6.5.9. Dé ejemplos de que una bola abierta puede ser cerrada y de que una bola cerrada puede ser abierta, y muestre que la clausura de una bola abierta está contenida en la bola cerrada del mismo centro y mismo radio, pero que no necesariamente coincide con ella.

El siguiente resultado es la versión de la Proposición 6.4.2 para los subconjuntos cerrados.

Proposición 6.5.10. *Sea $\mathscr{F}_X$ el conjunto de todos los subconjuntos cerrados de X.*
(*i*) *$\varnothing$ y X son elementos de $\mathscr{F}_X$.*
(*ii*) *Si F y G son elementos de $\mathscr{F}_X$, entonces $F \cup G$ también lo es.*
(*iii*) *Si $\mathscr{G}$ es un subconjunto de $\mathscr{F}_X$, entonces $\bigcap_{F\in\mathscr{G}} F \in \mathscr{F}_X$.*

Un argumento inductivo evidente a partir de la afirmación (*ii*) prueba que la unión de una familia finita de subconjuntos cerrados de X es cerrada.

Demostración. Es claro que $\overline{\varnothing} = \varnothing$ y que $\overline{X} = X$, así que $\varnothing, X \in \mathscr{F}_X$, como afirma (*i*). Por otro lado, si F y G son dos elementos de $\mathscr{F}_X$, entonces $\overline{F} = F$ y $\overline{G} = G$, así que de acuerdo a la cuarta parte de la Proposición 6.5.7 tenemos que $\overline{F \cup G} = \overline{F} \cup \overline{G} = F \cup G$, así que $F \cup G \in \mathscr{F}_X$. Esto nos dice que vale la afirmación (*ii*).

Sea, finalmente, $\mathscr{G}$ un subconjunto de $\mathscr{F}_X$ y pongamos $G := \bigcap_{F\in\mathscr{G}} F$. Para cada $F \in \mathscr{G}$ tenemos que $G \subseteq F$ y, como F es cerrado, que $\overline{G} \subseteq \overline{F} = F$: esto implica que $\overline{G} \subseteq \bigcap_{F\in G} F = G$, así que G es cerrado. Esto prueba la afirmación (*iii*). □

Ejercicio 6.5.11. Sean X y Y dos espacios métricos, y consideremos al producto cartesiano $X \times Y$ como un espacio métrico con respecto a la métrica que describe la Proposición 5.2.6.1. Pruebe que si F y G son subconjuntos cerrados de X y de Y, respectivamente, entonces $F \times G$ es un subconjunto cerrado de $X \times Y$ pero que, en general, no todos los subconjuntos cerrados de $X \times Y$ se obtienen de esa forma.

Usando la Proposición 6.5.10 podemos dar una caracterización de la clausura dual de la que nos da la Proposición 6.4.4 para el interior.

Proposición 6.5.12. *La clausura de un subconjunto de X es la intersección de todos los subconjuntos cerrados de X que lo contienen.*

Demostración. Sea S un subconjunto de X y sea F la intersección de todos los subconjuntos cerrados de X que contienen a S. La Proposición 6.5.10 nos dice que este conjunto F es cerrado en X y es claro que está contenido en todos los cerrados que contiene a S, así que es el menor cerrado de X que contiene a F, es decir, $F = \overline{S}$. □

No es cierto que la unión de una familia *arbitraria* de subconjuntos cerrados de X sea cerrada.

Por ejemplo, si S es un subconjunto de X que no es cerrado, entonces claramente $S = \bigcup_{x\in S}\{x\}$, y sabemos que el conjunto $\{x\}$ es cerrado cualesquiera sea el elemento x de X. Hay, de todas formas, situaciones en las que podemos garantizar que una unión de cerrados es cerrada. Veamos una.

Definición 6.5.13. Un conjunto $\mathscr{F}$ de subconjuntos de X es ***localmente finito*** si todo punto x de X posee un entorno N tal que el conjunto $\{F \in \mathscr{F} : F \cap N \neq \varnothing\}$ es finito.

Por ejemplo, la familia de intervalos $\{[n, n+1] : n \in \mathbb{Z}\}$ es localmente finita en $\mathbb{R}$.

Proposición 6.5.14. *La unión de una familia localmente finita de subconjuntos cerrados es cerrada.*

Demostración. Sea $\mathscr{F}$ una familia localmente finita de subconjuntos cerrados de X y sea $H \coloneqq \bigcup_{F\in\mathscr{F}} F$ su unión. Sea x un elemento de X que no pertenece a H. La hipótesis implica que hay un entorno N de X tal que el conjunto $I \coloneqq \{F \in \mathscr{F} : F \cap N \neq \varnothing\}$ es finito. Sea F un elemento de $\mathscr{F}$.

- Si $F \notin I$, entonces $F \cap N = \varnothing$.
- Si $F \in I$, entonces como $x \notin H \supseteq F$, hay un entorno N_F de x tal que $F \cap N_F = \varnothing$.

Consideremos el conjunto

$$M \coloneqq N \cap \bigcap_{F\in I} N_F.$$

Como se trata de una intersección finita de entornos de x, este conjunto M es un entorno de x. Por otro lado, si F es un elemento de $\mathscr{F}$, entonces

$$F \cap M = F \cap N \cap \bigcap_{F\in I} N_F \subseteq \left\{ \begin{matrix} F \cap N & \text{si } F \notin I; \\ F \cap N_F & \text{si } F \in I \end{matrix} \right\} = \varnothing,$$

y esto implica que

$$M \cap H = M \cap \bigcup_{F\in\mathscr{F}} F = \bigcup_{F\in\mathscr{F}} M \cap F = \varnothing$$

y, por lo tanto, que $x \notin \overline{H}$. Esto prueba que $\overline{H} \subseteq H$ y, en consecuencia, que el conjunto H es cerrado. □

Las operaciones de interior y clausura están estrechamente relacionadas: eso es lo que nos dice la siguiente proposición.

Proposición 6.5.15. *Si S es un subconjunto de X, entonces $X \setminus \overline{S} = (X \setminus S)^\circ$.*

Demostración. Sea S un subconjunto de X. Si x es es un elemento de $X \setminus \overline{S}$, entonces $x \notin \overline{S}$ y, por lo tanto, hay un entorno N de x tal que $N \cap S = \varnothing$: esto nos dice que $N \subseteq (X \setminus S)$, claro, y, por lo tanto, que el punto x pertenece a $(X \setminus S)^\circ$. Vemos así que $X \setminus \overline{S} \subseteq (X \setminus S)^\circ$.

Recíprocamente, si x es un elemento de $(X \setminus S)^\circ$, de manera que hay un entorno N de x tal que $N \subseteq X \setminus S$, entonces claramente $X \cap S = \varnothing$ y, por lo tanto, $x \notin \overline{S}$, esto es, $x \in X \setminus \overline{S}$. Esto prueba que $(X \setminus S)^\circ \subseteq X \setminus \overline{S}$ y, con lo anterior, la proposición. □

La consecuencia fundamental de esto es que los conjuntos abiertos y los conjuntos cerrados se determinan unos a los otros.

Proposición 6.5.16. *Un subconjunto S de X es cerrado si y solamente si su complemento $X \setminus S$ en X es abierto.*

Demostración. Sea S un subconjunto de X. Si S es cerrado, entonces $X \setminus S = X \setminus \overline{S} = (X \setminus S)^\circ$ es un abierto, y si $X \setminus S$ es abierto, entonces

$$S = X \setminus (X \setminus S) = X \setminus (X \setminus S)^\circ = X \setminus (X \setminus \overline{S}) = \overline{S},$$

de manera que S es cerrado. □

Ejercicio 6.5.17. Podríamos haber usado la Proposición 6.5.16 como *definición* de qué es un conjunto cerrado: decidir que un subconjunto S de X es cerrado si su complemento $X \setminus S$ es abierto. Usando esta definición como punto de partida redemuestre todos los resultados de esta sección.

La Proposición 6.4.2.1 describe los subconjuntos abiertos de un subespacio. Podemos dar una descripción idéntica para los cerrados:

Proposición 6.5.18. *Sea Y un subconjunto de X. Un subconjunto F de Y es cerrado en Y si y solamente si hay un subconjunto cerrado G de X tal que $F = Y \cap G$.*

Demostración. Sea F un subconjunto de Y. Si F es cerrado en Y, entonces $Y \setminus F$ es abierto en Y y, de acuerdo a la Proposición 6.4.2.1, hay un abierto U de X tal que $Y \setminus F = Y \cap U$ o, equivalentemente, $F = Y \setminus (Y \cap U) = Y \cap (X \setminus U)$. Como en ese caso $X \setminus U$ es cerrado en X, esto prueba que la condición de la proposición es necesaria.

Por otro lado, si hay un cerrado G de X tal que $F = Y \cap G$, entonces $X \setminus G$ es abierto en Y y $Y \setminus F = Y \setminus Y \cap G = Y \cap (X \setminus G)$ es abierto en Y, de manera que F es cerrado en Y. La condición de la proposición es, por lo tanto, también suficiente. □

Usando la proposición que acabamos de probar, podemos describir la operación de clausura

en los subespacios de X.

Proposición 6.5.19. *Sea Y un subespacio de X. Si S es un subconjunto de Y, entonces*

$$\mathrm{cl}_Y(S) = \mathrm{cl}_X(S) \cap Y.$$

Observemos que tenemos aquí una igualdad, a diferencia de lo que pasa con el interior relativo a un subespacio — como vimos en la Proposición 6.4.2.2.

Demostración. Sea S un subconjunto de Y. Como $\mathrm{cl}_X(S)$ es cerrado en X, $\mathrm{cl}_X(S) \cap Y$ es cerrado en Y y, como claramente esta intersección contiene a S, tenemos que $\mathrm{cl}_Y(S) \subseteq \mathrm{cl}_X(S) \cap Y$.

Sea x un elemento de $\mathrm{cl}_X(S) \cap Y$ y sea U un entorno de x en Y, de manera que $x \in \mathrm{int}_Y(U)$. Como $\mathrm{int}_Y(U)$ es abierto en Y, hay un abierto V de X tal que $\mathrm{int}_Y(U) = V \cap Y$. Como $x \in V$ y $x \in \mathrm{cl}_X(S)$, es $V \cap S \neq \varnothing$ y, por lo tanto,

$$U \cap S \supseteq \mathrm{int}_Y(U) \cap S = V \cap Y \cap S = V \cap S \neq \varnothing.$$

Vemos así que $x \in \mathrm{cl}_Y(S)$ y, por lo tanto, que $\mathrm{cl}_X(S) \cap Y \subseteq \mathrm{cl}_Y(S)$. □

En muchos contextos, la siguiente definición es importante.

Definición 6.5.20. Un subconjunto S de X es ***localmente cerrado*** si hay un subconjunto abierto U de X y un subconjunto cerrado F de X tales que $S = U \cap G$.

Usamos esa calificación porque esa condición es equivalente a la afirmación (b) de la siguiente proposición.

Proposición 6.5.21. *Sea S un subconjunto de X. Las siguientes tres afirmaciones son equivalentes:*

(a) *S es localmente cerrado en X.*

(b) *Para cada punto x de X hay un entorno N de x tal que $S \cap N$ es cerrado en N.*

(c) *S es abierto en $\overline{S}$.*

Demostración. $(a) \Rightarrow (b)$ Supongamos que S es localmente cerrado en X, de manera que hay un subconjunto abierto U y un subconjunto cerrado F en X tales que $S = U \cap F$. Sea x un punto de X.

- Si $x \notin S$, entonces $N \coloneqq X \setminus F$ es un entorno de x y su intersección con S es

  $$S \cap N = S \cap (X \setminus F) = U \cap F \cap (X \setminus F) = \varnothing,$$

  que es un cerrado de N.
- Si, en cambio, $x \in S$, entonces $x \in U$, así que $N \coloneqq U$ es un entorno de x y la intersección $S \cap N = U \cap F \cap U = F \cap U = F \cap N$ es un cerrado de N, ya que F es cerrado en X.

En cualquiera de los dos casos encontramos un entorno N de x tal que $S \cap N$ es un cerrado de N.

$(b) \Rightarrow (c)$ Supongamos que vale la condición (b) de la proposición. Para cada elemento x de S hay entonces un entorno N_x de x tal que la intersección $S \cap N_x$ es un cerrado de N_x y, por lo tanto, hay un cerrado F_x de X tal que $S \cap N_x = F_x \cap N_x$. Como $N_x \supseteq N_x^\circ$, tenemos entonces que

$$S \cap N_x^\circ = S \cap N_x \cap N_x^\circ = F_x \cap N_x \cap N_x^\circ = F_x \cap N_x^\circ,$$

de manera que $S \cap N_x^\circ$ es cerrado en N_x°. De acuerdo a la Proposición 6.5.19, esto implica que

$$S \cap N_x^\circ = \mathrm{cl}_{N_x^\circ}(S \cap N_x^\circ) \ = \mathrm{cl}_X(S \cap N_x^\circ) \cap N_x^\circ \subseteq \mathrm{cl}_X(S) \cap N_x^\circ \subseteq S \cap N_x^\circ$$

y, por lo tanto, que $S \cap N_x^\circ = \mathrm{cl}_X(S) \cap N_x^\circ$. Ahora bien, como $S \subseteq \bigcup_{x\in S} N_x^\circ$, tenemos que

$$S = S \cap \bigcup_{x\in S} N_x^\circ = \bigcup_{x\in S}(S \cap N_x^\circ) = \bigcup_{x\in S}(\mathrm{cl}_X(S) \cap N_x^\circ) = \mathrm{cl}_X(S) \cap \bigcup_{x\in S} N_x^\circ$$

y, como la unión $\bigcup_{x\in S} N_x^\circ$ es un abierto de X, esto nos permite concluir que S es un abierto de $\overline{S}$.

$(c) \Rightarrow (a)$ Si S es abierto en $\overline{S}$, entonces hay un abierto U de X tal que $S = U \cap \overline{S}$ y, como $\overline{S}$ es cerrado en X, esto nos dice que S es localmente cerrado. □

El siguiente ejercicio da otras caracterizaciones de los conjuntos localmente cerrados.

Ejercicio 6.5.22. Sea S un subconjunto de X. Pruebe que las siguientes afirmaciones son equivalentes:

(*a*) S es localmente cerrado en X.
(*b*) La diferencia $\overline{S} \setminus S$ es cerrada en X.
(*c*) Hay dos subconjuntos abiertos U_1 y U_2 de X tales que $S = U_1 \setminus U_2$.
(*d*) Hay dos subconjuntos cerrados F_1 y F_2 de X tales que $S = F_1 \setminus F_2$.

Como pasa con los subconjuntos abiertos, los subconjuntos cerrados de un grupo abeliano dotado de una métrica invariante por translaciones son preservados por las translaciones.

Ejercicio 6.5.23. Sea G un grupo abeliano y sea d una métrica sobre G que es invariante por translaciones.

(a) Muestre que un subconjunto S de G es cerrado si y solamente si el conjunto $x + S$ es cerrado cualquiera sea el elemento x de G.
(b) Pruebe que para todo subconjunto S de G y todo elemento x de G es $\overline{x + S} = s + \overline{S}$.

6.6. Conjuntos densos y espacios separables

Usando el operador de clausura podemos definir una propiedad importante.

Definición 6.6.1. Un subconjunto D del espacio métrico X es ***denso*** si $\overline{D} = X$.

Hay muchas formas de caracterizar los subconjuntos densos.

Proposición 6.6.2. *Sea D un subconjunto de X y sea $\mathscr{B}$ una base de la topología de X. Las siguientes afirmaciones son equivalentes:*

(*a*) *El conjunto D es denso.*
(*b*) *Todo subconjunto abierto no vacío de X contiene un elemento de D.*
(*c*) *Todo elemento no vacío de $\mathscr{B}$ contiene un elemento de D.*
(*d*) *El interior del complemento de D es vacío.*

Demostración. $(a) \Rightarrow (b)$ Supongamos que D es denso, de manera que $\overline{D} = X$. Si U es un abierto no vacío de X y x es un punto de U, entonces U es un entorno de x y x pertenece a $\overline{D}$, así que $U \cap D \neq \varnothing$.

$(b) \Rightarrow (c)$ Esta implicación es evidente, ya que los elementos de $\mathscr{B}$ son abiertos de X.

$(c) \Rightarrow (d)$ Supongamos que $(X \setminus D)^\circ$ no es vacío. Como $\mathscr{B}$ es una base de la topología, $(X \setminus D)^\circ$ es la unión de los elementos de $\mathscr{B}$ que contiene y hay, por lo tanto, un elemento no vacío U de $\mathscr{B}$ contenido en $(X \setminus D)^\circ$ y, por lo tanto, disjunto de D. Vemos así que hay elementos no vacíos de $\mathscr{B}$ que no contienen elementos de D y, en consecuencia, que vale la implicación contrarrecíproca a la implicación $(c) \Rightarrow (d)$.

$(d) \Rightarrow (a)$ Supongamos que el interior de $X \setminus D$ es vacío. Si x es un punto de X y U es un entorno de x, entonces U°, que es un abierto no vacío, no está contenido en $X \setminus D$ y contiene, por lo tanto, un elemento de D: esto implica que $x \in \overline{D}$. Vemos así que $X \subseteq \overline{D}$. □

Es útil saber que la propiedad de densidad es «transitiva» en el siguiente sentido.

Proposición 6.6.3. *Sea X un espacio métrico y sea Y un subespacio de X. Si D es un subconjunto denso de Y e Y es un subconjunto denso de X, entonces D también es un subconjunto denso de X.*

Demostración. Si D es un subconjunto denso de Y e Y uno de X, entonces la Proposición 6.5.19 nos dice que $Y = \mathrm{cl}_Y(D) = \mathrm{cl}_X(D) \cap Y$, así que $Y \subseteq \mathrm{cl}_X(D)$ y, por lo tanto, $X = \mathrm{cl}_X(Y) \subseteq \mathrm{cl}_X(D)$. □

Es claro que X es un subconjunto denso de X. Lo interesante es encontrar subconjuntos densos pequeños. El siguiente ejercicio muestra que solo los hay *muy* pequeños en situaciones triviales.

Ejercicio 6.6.4. Pruebe que si un espacio métrico X posee un subconjunto denso finito D, entonces, de hecho, X is finito y $D = X$.

Esto nos dice que el primer caso posiblemente interesante es aquel en el que hay subconjuntos densos numerables.

Definición 6.6.5. Decimos que un espacio métrico es ***separable*** si posee un subconjunto denso contable.

La separabilidad es una propiedad topológica: depende solamente de la topología determinada por la métrica del espacio y no realmente de la métrica.

La gracia de esta definición es que hay muchos ejemplos. Todo espacio contable es evidentemente separable — esto incluye a $\mathbb{Q}$ con cualquiera de sus métricas y a los espacios métricos que se obtienen de grupos finitamente generados con sus métricas de las palabras. Veamos algunos no contables.

Ejemplo 6.6.6. El conjunto $\mathbb{Q}$ es denso en $\mathbb{R}$. En efecto, sabemos que el conjunto de todos los intervalos abiertos de $\mathbb{R}$ es una base de su topología y que todo intervalo abierto contiene un número racional. Como $\mathbb{Q}$ es un conjunto numerable, esto nos dice, en particular, que $\mathbb{R}$ es un espacio separable.

Para dar más ejemplos es útil la siguiente observación.

Ejercicio 6.6.7. Sean X e Y dos espacios métricos.

(a) Pruebe que si D y E son subconjuntos densos de X e Y, respectivamente, entonces $D \times E$ es un subconjunto denso de $X \times Y$. Extienda este resultado al caso de productos cartesianos de finitos factores.

(b) Pruebe que el producto cartesiano de finitos espacios no vacíos es separable si y solamente si cada uno de los factores es separable

Usado esto obtenemos inmediatamente el siguiente ejemplo.

Ejemplo 6.6.8. Cualquiera sea el entero positivo n el subconjunto $\mathbb{Q}^n$ es denso en $\mathbb{R}^n$ y, por lo tanto, el espacio $\mathbb{R}^n$ es separable. Gracias al resultado del Ejercicio 6.6.7, esto es consecuencia inmediata del Ejemplo 6.6.6 y de que la métrica sobre $\mathbb{R}^n$ es la métrica producto sobre$\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}$ dotando a cada factor de la distancia euclídea.

Veamos un ejemplo más interesante.

Ejemplo 6.6.9. Sea α un número real que no es racional y mostremos que el conjunto

$$D \coloneqq \{a + \alpha b : a, b \in \mathbb{Z}\}$$

es denso en $\mathbb{R}$. Si reemplazamos a α por $-\alpha$ este conjunto no cambia, así que podemos suponer en lo que sigue que α es positivo.

Supongamos por un momento que $r \coloneqq \text{ínf}\, D \cap (0, +\infty)$ es un número positivo — notemos que la intersección $D \cap (0, +\infty)$ contiene a $\mathbb{N}$ y está acotada inferiormente, así que tiene sentido considerar su ínfimo. Si r no perteneciera a $D \cap (0, +\infty)$, podríamos elegir dos elementos s y t de D tales que $r \leq s < t < (1 + \frac{1}{2})r$: como D es un subgrupo de $\mathbb{R}$, tendríamos entonces que $t - s$ es un elemento de D tal que $0 < t - s < (1 + \frac{1}{2})r - r = \frac{1}{2}r$, y esto es absurdo. Vemos así que r es necesariamente un elemento de $D \cap (0, +\infty)$.

Sea ahora d un elemento cualquiera de $D \cap (0, +\infty)$. Pongamos $n \coloneqq \lfloor d/r \rfloor$, que es un entero positivo porque $d \geq r$. Tenemos que $n \leq d/r < n+1$ y, ya que r es positivo, que $0 \leq d - nr < r$: como $d - nr$ pertenece a D y $D \cap (0, r) = \varnothing$, esto implica que $d - nr = 0$, esto es, que $d = nr$. Vemos así que $D \cap (0, +\infty) = \{nr : n \in \mathbb{N}\}$. En particular, como 1 y α pertenecen a $D \cap (0, +\infty)$, esto implica que hay enteros positivos m y n tales que $1 = mr$ y $\alpha = nr$: esto es absurdo, porque nos dice que $\alpha = n/m$ y el número α no es racional. Podemos concluir gracias a esto que ínf $D \cap (0, +\infty) = 0$.

Sea ahora (u, v) un intervalo abierto cualquiera de $\mathbb{R}$. La diferencia $v - u$ es un número positivo: como ínf $D \cap (0, +\infty) = 0$, hay un elemento d de D tal que $0 < d < v - u$. Sea $n \coloneqq \lfloor v/d \rfloor$, de manera que $v/d - 1 < n \leq v/d$. Es entonces $u < v - d < nd < v$, así que nd, que es un elemento de D, está en el intervalo (u, v). Vemos así que D interseca no trivialmente a todo intervalo abierto de $\mathbb{R}$, así que es denso en $\mathbb{R}$.

Podemos generalizar este resultado. Consideremos ahora α y β dos elementos de $\mathbb{R}$ que son linealmente independientes sobre $\mathbb{Z}$, esto es, que siempre que m y n son enteros tales que $m\alpha + n\beta = 0$ se tiene que $m = n = 0$, y mostremos que

el conjunto $D \coloneqq \{m\alpha + n\beta : m, n \in \mathbb{Z}\}$ *es denso en* $\mathbb{R}$.

La independencia lineal sobre $\mathbb{Z}$ de α y β implica que esos dos números son distintos de 0 y podemos suponer que son positivos, ya que reemplazarlos por su valor absoluto no cambia ni la independencia lineal ni el conjunto D. Por otro lado, esa independencia lineal implica inmediatamente que el cociente β/α es un número irracional, así que lo que ya hicimos nos dice que el conjunto $\{m + n\beta/\alpha : m, n \in \mathbb{Z}\}$ es denso en $\mathbb{R}$. Si (u, v) es un intervalo abierto, entonces también lo es $(u/\alpha, v/\alpha)$ y hay, por lo tanto, enteros m y n tales que $u/\alpha < m + n\beta/\alpha < v/\alpha$: es entonces $u < m\alpha + n\beta < v$, así que $D \cap (u, v) \neq \varnothing$. Esto prueba lo que queremos.

Consideremos ahora algunos espacios normados de dimensión infinita.

Ejemplo 6.6.10. Sea p un elemento de $[1,+\infty)$ y consideremos el espacio *real* normado $\ell^p_{\mathbb{R}}$ de la Proposición 5.2.3.26 y el conjunto D de todas sucesiones $(x_n)_{n\geq 0}$ tales que $x_n \in \mathbb{Q}$ para todo $n \in \mathbb{N}_0$ y el conjunto $\{n \in \mathbb{N}_0 : x_n \neq 0\}$ es finito. Es fácil ver que D es un conjunto numerable Queremos probar que, además,

> *el conjunto D es un subconjunto denso de $\ell^p_{\mathbb{R}}$ y que, en particular, este espacio normado es separable.*

Sea $x = (x_n)_{n\geq 0}$ un elemento de $\ell^p_{\mathbb{R}}$ y sea ε un número positivo. Como la serie $\sum_{n=0}^{\infty}|x_n|^p$ converge, hay un entero positivo N tal que $\sum_{n=N+1}^{\infty}|x_n|^p < \varepsilon^p/2$. Por otro lado, como el conjunto $\mathbb{Q}$ es denso en $\mathbb{R}$, para cada $i \in [\![0, N]\!]$ hay un número racional q_i tal que $|x_i - q_i| < \varepsilon^p/2(N+1)$. Sea $y = (y_n)_{n\geq 0}$ el elemento de $\mathbb{R}^{\mathbb{N}_0}$ que para cada $n \in \mathbb{N}_0$ tiene

$$y_n = \begin{cases} q_n & \text{si } i \leq N; \\ 0 & \text{si } i > N. \end{cases}$$

Es claro que y pertenece a D. Por otro lado, es

$$\begin{aligned}\sum_{n=0}^{\infty}|x_n - y_n|^p &= \sum_{n=0}^{N}|x_n - y_n|^p + \sum_{n=N+1}^{\infty}|x_n - y_n|^p = \sum_{n=0}^{N}|x_n - q_n|^p + \sum_{n=N+1}^{\infty}|x_n|^p \\ &< \sum_{n=0}^{N}\frac{\varepsilon^p}{2(N+1)} + \frac{\varepsilon^p}{2} = \varepsilon^p,\end{aligned}$$

de manera que $d_p(x,y) = \|x - y\|_p < \varepsilon$. Esto prueba que el conjunto D es denso, como queremos.

Ejercicio 6.6.11. Pruebe que para todo elemento p de $[1,+\infty)$ el espacio *complejo* $\ell^p_{\mathbb{C}}$ es separable.

Veamos un ejemplo de un espacio que *no* es separable.

Ejemplo 6.6.12. Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, y consideremos el $\mathbb{k}$-espacio $\ell^\infty_{\mathbb{k}}$ de la Proposición 5.2.3.22 y sea D un subconjunto denso de $\ell^\infty_{\mathbb{k}}$. Para cada subconjunto P de $\mathbb{N}_0$ sea $x^P = (x^P_n)_{n\geq 0}$ el elemento de $\mathbb{k}^{\mathbb{N}_0}$ que para cada $n \in \mathbb{N}_0$ tiene

$$x^P_n = \begin{cases} 1 & \text{si } n \in P; \\ 0 & \text{si } n \notin P. \end{cases}$$

Es claro que $x^P \in \ell^\infty_{\mathbb{k}}$ cualquiera sea $P \in \mathscr{P}(\mathbb{N}_0)$ y que, como el conjunto D es denso, que la intersección $B_{1/2}(x^P) \cap D$ no es vacía. Hay, por lo tanto, una función $\phi : \mathscr{P}(\mathbb{N}_0) \to D$ tal que $\phi(P) \in B_{1/2}(x^P) \cap D$ cualquiera sea $P \in \mathscr{P}(\mathbb{N}_0)$. Esta función es, además, inyectiva: si P y Q son dos elementos distintos de $\mathscr{P}(\mathbb{N}_0)$, entonces hay un elemento n en la diferencia simétrica $(P \smallsetminus Q) \cup (Q \smallsetminus P)$ y es $|x^P_n - x^Q_n| = 1$, de manera que $d_\infty(x^P, x^Q) \geq 1$ y, en particular,

$B_{1/2}(x^P) \cap B_{1/2}(x^Q) = \varnothing$: como $\phi(P) \in B_{1/2}(x^P)$ y $\phi(Q) \in B_{1/2}(x^Q)$, vemos con esto que $\phi(P) \neq \phi(Q)$. Ahora bien, como la función ϕ es inyectiva, tenemos que $\mathfrak{c} = |\mathscr{P}(\mathbb{N}_0)| \leq |D|$ y, por lo tanto, que el conjunto D no es numerable.

El argumento que usamos en este ejemplo funciona en general.

Ejercicio 6.6.13. Sea X un espacio métrico y sea $\mathscr{U}$ un conjunto de abiertos no vacíos y disjuntos dos a dos. Muestre que todo subconjunto denso de X tiene cardinal mayor o igual que el de $\mathscr{U}$.

Ejercicio 6.6.14. Sea $(X_n)_{n\geq 0}$ una sucesión de espacios métricos no vacíos y consideremos al conjunto $X \coloneqq \prod_{n=0}^{\infty} X_n$ como un espacio métrico con respecto a la métrica descripta en la Proposición 5.2.6.4. Pruebe que el espacio X es separable si y solamente si para todo $n \in \mathbb{N}_0$ el espacio métrico X_n es separable.

Ejercicio 6.6.15. Sea A un conjunto, sea $X \coloneqq A^{\mathbb{N}_0}$ el conjunto de todas las sucesiones en A, y consideremos sobre X la ultramétrica d de la Proposición 5.2.9.5. Pruebe el espacio métrico (X, d) es separables si y solamente si el conjunto A es finito.

La siguiente es una de las razones por las que nos interesan los espacios separables.

Proposición 6.6.16. *Un espacio métrico es separable si y solamente si su topología posee una base contable.*

Demostración. Sea X un espacio métrico y supongamos primero que la topología de X posee una base contable $\mathscr{B}$. El *Axioma de elección* nos dice que hay una función $\phi : \mathscr{B} \setminus \{\varnothing\} \to X$ tal que $\phi(U) \in U$ para todo $U \in \mathscr{B} \setminus \{\varnothing\}$. El conjunto $D \coloneqq \operatorname{img}(\phi)$ es claramente contable e interseca cada elemento no vacío de la base $\mathscr{B}$, así que es denso en X: vemos así que el espacio X es separable y que, por lo tanto, la condición de la proposición es suficiente.

Supongamos ahora que X es un espacio métrico y que D es un subconjunto de X que es denso y contable. Mostremos que el conjunto

$$\mathscr{B} \coloneqq \{B_{1/n}(x) : x \in D, n \in \mathbb{N}\}$$

es una base de la topología de X. Es claro que los elementos de $\mathscr{B}$ son abiertos.

Sea, por otro lado, U un subconjunto abierto de X y sea x un elemento de U, de manera que hay un número positivo r tal que $B_r(x) \subseteq U$. Como D es denso, hay un elemento q de D en el abierto $B_{r/2}(x)$, así que $x \in B_{r/2}(q)$. Si y es un elemento de $B_{r/2}(q)$, entonces $d(y, x) \leq d(y, q) + d(q, x) < r/2 + r/2 = r$, así que $y \in B_r(x) \subseteq U$. Vemos así que $x \in B_{1/4}(q) \subseteq U$ y, por lo tanto, que el conjunto $\mathscr{B}$ es una base de la topología de X. □

La consecuencia más importante de esta proposición es nuestro siguiente resultado.

Proposición 6.6.17. *Un subespacio de un espacio separable es separable.*

Demostración. Sea X un espacio separable y sea Y un subespacio de X. Como X es separable, la Proposición 6.6.16 nos dice que hay una base contable $\mathscr{B}$ de la topología de X. Como el conjunto $\mathscr{B}_Y \coloneqq \{B \cap Y : B \in \mathscr{B}\}$ es una base para la topología del subespacio Y y claramente $|\mathscr{B}_Y| \leq |\mathscr{B}|$, vemos que Y tiene una base contable, así que aquella proposición nos dice que Y es separable. □

El lector debería intentar probar este corolario de manera directa, sin involucrar el concepto de bases.

Ejercicio 6.6.18. Sea X un espacio métrico. Pruebe que la siguientes afirmaciones son equivalentes:
(*a*) X es separable.
(*b*) X no posee un subconjunto discreto no numerable.
(*c*) X no posee un conjunto no numerable de abiertos disjuntos dos a dos.

Otro resultado poco evidente es el siguiente:

Proposición 6.6.19. *Sea X un espacio métrico separable. Si $\mathscr{U}$ es un conjunto de abiertos de X tal que $X = \bigcup \mathscr{U}$, entonces hay un subconjunto numerable $\mathscr{V}$ de $\mathscr{U}$ tal que $X = \bigcup \mathscr{V}$.*

Demostración. Sea $\mathscr{U}$ un conjunto de abiertos de X tal que $X = \bigcup \mathscr{U}$. Como el espacio X es separable, la Proposición 6.6.16 nos dice que hay una base contable $\mathscr{B}$ para la topología de X. Como $\mathscr{B}$ es una base, para todo elemento U de $\mathscr{U}$ es $U = \bigcup\{B \in \mathscr{B} : B \subseteq U\}$, así que si ponemos

$$\mathscr{B}' \coloneqq \bigcup_{U \in \mathscr{U}} \{B \in \mathscr{B} : B \subseteq U\},$$

entonces

$$\bigcup \mathscr{B}' = \bigcup_{U \in \mathscr{U}} \bigcup_{U \in \mathscr{U}} \{B \in \mathscr{B} : B \subseteq U\} = \bigcup_{U \in \mathscr{U}} U = X.$$

Para cada elemento B de $\mathscr{B}'$ hay un elemento U de $\mathscr{U}$ tal que $B \subseteq U$, así que el *Axioma de elección* nos dice que hay una función $c : \mathscr{B}' \to \mathscr{U}$ tal que $B \subseteq c(B)$ para todo $B \in \mathscr{B}'$. El subconjunto $\mathscr{V} \coloneqq \operatorname{img}(c)$ de $\mathscr{U}$ es contable, porque $\mathscr{B}'$ está contenido en $\mathscr{B}$ y $\mathscr{B}$ es contable, y

$$\bigcup \mathscr{V} = \bigcup_{B \in \mathscr{B}'} c(B) \supseteq \bigcup_{B \in \mathscr{B}'} B = X.$$

Esto prueba la proposición. □

La situación de esta proposición es importante, así que amerita hacer algunas definiciones.

Definición 6.6.20. Sea X un conjunto.

- Un subconjunto $\mathscr{U}$ de $\mathscr{P}(X)$ es un ***cubrimiento*** de X si $\bigcup \mathscr{U} = X$, y cuando ese es el caso un ***subcubrimiento*** de $\mathscr{U}$ es un cubrimiento de X que está contenido en $\mathscr{U}$.
- Si el conjunto X es un espacio métrico, entonces un cubrimiento de X es un ***cubrimiento abierto*** si todos sus elementos son subconjuntos abierto de X.

Definición 6.6.21. Un espacio métrico es ***de Lindelöf*** si cada uno de sus cubrimientos abiertos posee un subcubrimiento contable.

El nombre de esta propiedad recuerda a Ernst Leonard Lindelöf.

La Proposición 6.6.19 es, en este lenguaje, la mitad de la siguiente:

Proposición 6.6.22. *Un espacio métrico es separable si y solamente si es un espacio de Lindelöf.*

Demostración. Que la condición es necesaria es lo que afirma la Proposición 6.6.19, así que nos queda solamente probar que es suficiente.

Sea X un espacio métrico de Lindelöf. Si n es un entero positivo, entonces es claro que el conjunto $\mathscr{U}_n \coloneqq \{B_{1/n}(x) : x \in X\}$ es un cubrimiento abierto, así que la hipótesis nos dice que hay un cubrimiento abierto y contable $\mathscr{V}_n$ contenido en $\mathscr{U}_n$. Para completar la demostración de la proposición mostraremos que el conjunto $\mathscr{B} \coloneqq \bigcup_{n\in\mathbb{N}} \mathscr{V}_n$, que ciertamente es contable, es una base de la topología de X.

Sea U un subconjunto abierto de X y sea x un elemento de U. Hay un entero positivo n tal que $B_{1/n}(x) \subseteq U$ y, como el conjunto $\mathscr{V}_{2n}$ es un cubrimiento de X, hay un elemento V de $\mathscr{V}_{2n}$ tal que $x \in V$. Como V está en $\mathscr{U}_{2n}$, hay un punto y de X tal que $V = B_{1/2n}(y)$, y como $x \in V = B_{1/2n}(y)$, para todo elemento z de V tenemos que

$$d(x,z) \leq d(x,y) + d(y,z) < 1/2n + 1/2n = 1/n.$$

Esto nos dice que $x \in V \subseteq B_{1/2}(x) \subseteq U$ y, en definitiva, nos permite concluir que el conjunto $\mathscr{B}$ es una base de la topología de X. □

Una de las formas en las que la condición de Lindelöf es útil es a través del siguiente resultado.

Proposición 6.6.23. *Un espacio métrico X es de Lindelöf si y solamente si todo conjunto $\mathscr{F}$ no vacío de subconjuntos cerrados de X tal que*

$$\text{todo subconjunto contable y no vacío de } \mathscr{F} \text{ tiene intersección no vacía} \tag{6.1}$$

tiene intersección no vacía.

Cuando un subconjunto $\mathscr{F}$ de $\mathscr{P}(X)$ satisface la condición (6.1) decimos que tiene la ***propiedad de intersección numerable***.

Demostración. Sea primero X un espacio métrico de Lindelöf, sea $\mathscr{F}$ un conjunto no vacío de subconjuntos cerrados de X que satisface la condición del ejercicio y supongamos, para llegar a una contradicción, que $\bigcap \mathscr{F} = \varnothing$. El conjunto $\mathscr{U} \coloneqq \{X \setminus F : F \in \mathscr{F}\}$ tiene entonces todos sus elementos abiertos y es un cubrimiento de X ya que

$$\bigcup_{U\in\mathscr{U}} U = \bigcup_{F\in\mathscr{F}} (X \setminus F) = X \setminus \bigcap_{F\in\mathscr{F}} F = X \setminus \varnothing = X.$$

Como el espacio X es de Lindelöf, hay entonces un subconjunto contable $\mathscr{V}$ de $\mathscr{U}$ que es un cubrimiento de X. Ahora bien, esto implica que $A \coloneqq \{X \setminus V : V \in \mathscr{V}\}$ es un subconjunto contable de $\mathscr{F}$ y

$$\bigcap A = \bigcap_{V\in\mathscr{V}} (X \setminus V) = X \setminus \bigcup_{V\in\mathscr{V}} V = X \setminus X = \varnothing.$$

Esto es absurdo. Esta contradicción prueba que la condición de la proposición es necesaria.

Veamos ahora que es suficiente. Supongamos que todo conjunto no vacío de cerrados de X que tiene la propiedad de intersección numerable tiene intersección no vacía, sea $\mathscr{U}$ un cubrimiento abierto de X, y, para llegar a una contradicción, supongamos que $\mathscr{U}$ no posee subcubrimientos contables. Claramente el espacio X no es vacío, así que $\mathscr{U}$ no es vacío. El conjunto $\mathscr{F} \coloneqq \{X \setminus U : U \in \mathscr{U}\}$ es entonces no vacío y sus elementos son cerrados. Más aún, si H es un subconjunto contable de $\mathscr{F}$, entonces

$$X \setminus \bigcap_{F\in H} F = \bigcup_{F\in H} (X \setminus F) \neq X,$$

ya que $\{X \setminus F : F \in H\}$ es un subconjunto contable de $\mathscr{U}$, y esto implica que $\bigcap_{F\in H} F \neq \varnothing$. Vemos así que $\mathscr{F}$ es un conjunto no vacío de cerrados de X que tiene la propiedad de intersección numerable y que, por lo tanto, tiene intersección no vacía. Esto es absurdo, ya que entonces

$$X = \bigcup_{U\in\mathscr{U}} U = \bigcup_{F\in\mathscr{F}} (X \setminus F) = X \setminus \bigcap_{F\in\mathscr{F}} F \subsetneq X.$$

Vemos así que el cubrimiento $\mathscr{U}$ debe contener subcubrimientos numerables y que, como queríamos probar, la condición de la proposición es suficiente. □

6.7. El conjunto derivado

Un punto x pertenece a la clausura de un subconjunto S de X si para todo entorno N de x la intersección $N \cap S$ no es vacía. Cuando ese es el caso, tiene sentido preguntarse por el cardinal de esa intersección.

Lema 6.7.1. *Sea S un subconjunto de X y sea x un elemento de $\overline{S}$. Si hay un entorno N de x tal que la intersección $S \cap N$ es finita, entonces hay un entorno M de x tal que $S \cap M = \{x\}$.*

Notemos que, en particular, x pertenece al conjunto S.

Demostración. Sea N un entorno de x, de manera que hay un número positivo r tal que $B_r(x) \subseteq N$ y supongamos que la intersección $S \cap N$ es finita. El conjunto $D \coloneqq \{r\} \cup \{d(x,y) : y \in S \cap N \setminus \{x\}\}$ es entonces un subconjunto finito de $(0, +\infty)$, así que $s \coloneqq \frac{1}{2} \inf D$ es un número positivo y $M \coloneqq B_s(x)$ es un entorno de x. Si z es un elemento de $M \cap S$, entonces $d(x,z) < s < r$, así que $z \in B_r(x) \subseteq N$, y $d(x,z) < s < d(x,y)$ para todo $y \in S \cap N \setminus \{x\}$, así que $y \notin S \cap N \setminus \{x\}$: debe ser, por lo tanto, $z = x$. Vemos así que $M \cap S \subseteq \{x\}$. Como x pertenece a $\overline{S}$, la intersección $M \cap S$ no es vacía, así que podemos concluir que, de hecho, $M \cap S = \{x\}$. □

Usando la definición siguiente, en la situación del lema el punto x es un punto aislado de S.

Definición 6.7.2. Decimos que un elemento x de un subconjunto S de X es ***aislado*** en S si hay un entorno N de x tal que $N \cap S = \{x\}$.

Se sigue de todo esto que si x es un elemento de $\overline{S}$ que no es un punto aislado de S, entonces para todo entorno N de x la intersección $N \cap S$ tiene que ser infinita. Le ponemos nombre a esta situación.

Definición 6.7.3. Sea S un subconjunto de X.

- Un punto x de X es un ***punto de acumulación*** de S si todo entorno de x contiene infinitos elementos de S.
- El ***conjunto derivado*** de S, al que escribimos S', es el conjunto de todos los puntos de acumulación de S.

Para que un punto x sea un punto de acumulación de un conjunto S en todos sus entornos tiene que haber, en principio, infinitos elementos de S pero, de hecho, es suficiente que haya uno distinto de x.

Proposición 6.7.4. *Sea S un subconjunto de X. Un punto x de X pertenece a S' si y solamente si todo entorno de x contiene un elemento de S distinto de x.*

Demostración. Sea x un elemento de X. Si x pertenece a S', entonces todo entorno de x contiene infinitos elementos de S, así que, en particular, contiene uno que es distinto de x. Esto nos dice que la condición de la proposición es necesaria.

Probemos ahora que es suficiente. Supongamos que esa condición se satisface y sea N un entorno de x. Para cada elemento y de $S \setminus \{x\}$ consideremos el subconjunto

$$R(y) \coloneqq \{z \in S \setminus \{x\} : d(x,z) < d(x,y)/2\}$$

de $S \setminus \{x\}$. Este conjunto no es vacío: si $y \in X \setminus \{x\}$, entonces $d(x,y)$ es un número positivo, así que la hipótesis implica que en $B_{d(x,y)/2}(x)$, que es un entorno de x, hay elementos de S distintos de x, y ellos son precisamente los elementos de $R(y)$.

Obtenemos de esta forma una función $R : S \setminus \{x\} \to \mathscr{P}(S \setminus \{x\}) \setminus \{\varnothing\}$. Por otro lado, como N es un entorno de x, hay un número positivo r tal que $B_r(x) \subseteq N$ y, como $B_r(x)$ es también un entorno de x, hay elemento y_0 de S en $B_r(x)$ distinto de x. De acuerdo al *Principio de elección dependiente* en la forma que le dimos en el Ejercicio 2.1.7, hay entonces una función $f : \mathbb{N}_0 \to S \setminus \{x\}$ tal que $f(0) = y_0$ y $f(n+1) \in R(f(n))$ para todo $n \in \mathbb{N}_0$. Esto último implica que

$$0 < d(x, f(n+1)) < \frac{1}{2} d(x, f(n))$$

para todo $n \in \mathbb{N}_0$, y un argumento inductivo que omitimos permite deducir a partir de eso que, más generalmente, siempre que m y n son dos elementos de $\mathbb{N}_0$ tales que $m \leq n$ vale que

$$0 < d(x, f(n)) < \frac{1}{2^{n-m}} d(x, f(m)).$$

Por un lado, esto nos dice que para todo entero no negativo n es que

$$d(x, f(n)) < \frac{1}{2^n} d(x, f(0)) = \frac{1}{2^n} d(x, y_0) < r$$

y, por lo tanto, $f(n) \in B_r(x) \subseteq N$. Vemos así que la imagen de la función f está contenida en $S \cap N$. Por otro lado, esa función es inyectiva. En efecto, si m y n son dos elementos de $\mathbb{N}_0$ tales que $m < n$, entonces

$$d(x, f(n)) < \frac{1}{2^{n-m}} d(x, f(m)) < d(x, f(m),$$

así que $f(n) \neq f(m)$. Todo esto nos permite concluir que el entorno N de x contiene infinitos elementos de S y, como esto es así cualquiera sea ese entorno de x, que $x \in S'$. □

Es esperable que haya una relación cercana entre la clausura de un conjunto y su conjunto derivado. De hecho, tenemos lo siguiente:

Proposición 6.7.5. *Sea S un subconjunto de X. Es $\overline{S} = S \cup S'$ y S es cerrado si y solamente si $S' \subseteq S$.*

Demostración. Sabemos que $S \subseteq \overline{S}$ y es claro que todo punto de acumulación de S es un punto de adherencia de S, de manera que también $S' \subseteq \overline{S}$. Esto implica que $S \cup S' \subseteq \overline{S}$. Por otro lado, si x es un elemento de $\overline{S}$, entonces o bien hay un entorno N de x tal que la intersección $N \cap S$ es finita y, de acuerdo al Lema 6.7.1, x pertenece a S, o bien todo entorno de x contiene infinitos elementos de S y x pertenece a S'. Esto prueba la primea afirmación de la proposición.

Si S es cerrado, entonces $S \cup S' = \overline{S} \subseteq S$, así que $S' \subseteq S$, y si $S' \subseteq S$, entonces $\overline{S} = S \cup S' = S$, así que S es cerrado. Esto prueba la segunda afirmación. □

Ejercicio 6.7.6. Sean S y T dos subconjuntos de X. Pruebe las siguientes afirmaciones:

(a) $\varnothing' = \varnothing$.
(b) $x \in S' \implies x \in (S \setminus \{x\})'$.
(c) $(S \cup T)' = S' \cup T'$.
(d) $S \subseteq T \implies S' \subseteq T'$.

La siguiente es una observación bastante menos evidente que todas las que llevamos hecho.

Proposición 6.7.7. *El conjunto derivado de cualquier subconjunto de X es cerrado.*

Demostración. Sea S un subconjunto de X. De acuerdo a la Proposición 6.7.5, para probar que S' es cerrado es suficiente con que mostremos que $S'' \subseteq S'$. Sea x un elemento de S'' y sea N un entorno de x. Como N° es un entorno de x y $x \in S''$, hay en N° un elemento y de S' distinto de x. Ahora bien, N° también es un entorno de y y, como $y \in S'$, hay en N° infinitos elementos de S. Vemos así que todo entorno de x contiene infinitos elementos de S, así que $x \in S'$. □

Las siguientes dos propiedades que puede tener un conjunto suelen ser importantes.

Definición 6.7.8. Un subconjunto S de X es ***denso en sí mismo*** si $S \subseteq S'$ y es ***perfecto*** si además es cerrado.

Un conjunto es denso en sí mismo exactamente cuando no tiene puntos aislados. Por otro lado, un subconjunto S es perfecto si y solamente si $S = S'$. Es importantes observar de que la perfección de un conjunto depende del espacio en el que lo consideremos: por ejemplo, $\mathbb{Q} \cap [0,1]$ es un subconjunto perfecto de $\mathbb{Q}$ pero no de $\mathbb{Q}$.

Para exhibir conjuntos perfectos usaremos la siguiente noción.

Definición 6.7.9. Decimos que un punto x de X es un ***punto de condensación*** de un subconjunto S de X si para todo entorno N de x la intersección $S \cap N$ no es contable.

Proposición 6.7.10. *Sea X es un espacio métrico separable. Si S es un subconjunto no contable de X, entonces el conjunto C de los puntos de condensación de S es no vacío y perfecto, y la diferencia $S \setminus C$ es contable.*

Demostración. Sea X un espacio métrico separable, sea S un subconjunto no contable de X, y sea $\mathscr{B}$ una base contable de la topología de X.

Supongamos que S no posee puntos de condensación. Si x es un elemento de X, entonces x no es un punto de condensación y, como $\mathscr{B}$ es una base de la topología de X, hay un elemento B_x de $\mathscr{B}$ tal que $x \in B_x$ y $B_x \cap S$ es contable. Sea $\mathscr{U} \coloneqq \{B_x : x \in X\}$. Se trata de un conjunto contable, porque es un subconjunto de $\mathscr{B}$, y como $x \in B_x$ para todo $x \in X$, es $X = \bigcup \mathscr{U}$. Ahora bien, para cada elemento U de $\mathscr{U}$ tenemos que $S \cap U$ es contable y es

$$S = S \cap X = S \cap \bigcup \mathscr{U} = \bigcup_{U \in \mathscr{U}} (S \cap U),$$

así que S es contable. Esto es absurdo.

Vemos así que el conjunto C de los puntos de condensación de S no es vacío. Mostremos ahora que es perfecto.

- Sea x un elemento de C y sea n un entero positivo. Como x es un punto de condensación de S, la intersección $S \cap B_{1/n}(x)$ no es contable. Como

  $$B_{1/n}(x) = \{x\} \cup \bigcup_{m=n}^{\infty} (B_{1/m}(x) \setminus B_{1/(m+1)}(x)),$$

  tenemos que

  $$S \cap B_{1/n}(x) = (S \cap \{x\}) \cup \bigcup_{m=n}^{\infty} \big(S \cap (B_{1/m}(x) \setminus B_{1/(m+1)}(x))\big).$$

  El conjunto $S \cap B_{1/n}(x)$ no es contable y $S \cap \{x\}$ es finito, así que hay un entero m mayor o igual que n tal que la intersección $T \coloneqq S \cap (B_{1/m}(x) \setminus B_{1/(m+1)}(x))$ no es contable. Por lo que probamos ese conjunto T tiene un punto de condensación y en X. Como $T \subseteq S$, es claro que $y \in C$. Por otro lado, como $T \subseteq X \setminus B_{1/(m+1)}(x)$ y este último conjunto es cerrado, tenemos que $y \in \overline{T} \subseteq X \setminus B_{1/(m+1)}(x)$, de manera que $y \neq x$.

  Vemos así que para todo entero positivo n el conjunto $C \cap B_{1/n}(x)$ tiene un elemento distinto de x, así que $x \in C'$. Así, tenemos que $C \subseteq C'$ y, por lo tanto, que C es denso en sí mismo.

- Probemos ahora que C es cerrado. Sea x un elemento de $\overline{C}$ y sea U un entorno abierto de x. Como x es un punto de adherencia de C, hay un elemento y de C en U, y como y es un punto de condensación de S, la intersección $S \cap U$ no es contable. Vemos así que todo entorno abierto de x tiene no contables elementos de S, así que x es un punto de condensación de S, esto es, $x \in C$.

Para terminar, tenemos que probar que la diferencia $S \setminus C$ es contable.

Sea n un entero positivo, sea $S_n \coloneqq \{x \in S : B_{1/n}(x) \cap C = \varnothing\}$, y supongamos que S_n no es contable. De acuerdo a lo que ya hicimos, hay entonces un punto y de condensación de S_n. Como $S_n \subseteq S$, ese punto también es de condensación de S y, por lo tanto, $y \in C$. Ahora bien, como $B_{1/n}(y)$ es un entorno de y, hay un elemento $x \in S_n \cap B_{1/n}(y0)$ y, en particular, $y \in B_{1/n}(x) \cap C$: esto es absurdo, ya que x está en S_n.

Vemos así que para todo entero positivo n el conjunto S_n es contable y, por lo tanto, que también lo es la unión $\bigcup_{n\in\mathbb{N}} S_n$. Esa unión está claramente contenida en $S \setminus C$; recíprocamente, si x es un punto de $S \setminus C$, entonces que C sea cerrado implica que hay un entero positivo n tal que $B_{1/n}(x)) \cap C = \varnothing$ y, en consecuencia, tenemos que $x \in S_n$. Vemos así que $S \setminus C = \bigcup_{n\in\mathbb{N}} S_n$, así que $S \setminus C$ es contable, como queremos. □

El siguiente ejemplo muestra cómo podemos usar este resultado.

Ejemplo 6.7.11. Un subconjunto S no contable de $(0, +\infty)$ tiene un punto de acumulación positivo. Si C es el conjunto de los puntos de condensación de S, entonces $S \setminus C$ es contable y, por lo tanto, distinto de S: se sigue de esto que hay un elemento de S en C y, por lo tanto, que hay un elemento de C que es positivo.

Terminemos esta sección con dos problemas históricamente importantes: el estudio de estos resultados es una de las motivaciones que llevaron a Georg Cantor a desarrollar la teoría de ordinales y cardinales.

Ejercicio 6.7.12. Sea X un espacio métrico y consideremos la función $D : S \in \mathscr{P}(X) \mapsto S' \in \mathscr{P}(X)$. De acuerdo al *Principio de recurrencia* 1.2.1, para cada subconjunto S de X hay entonces una sucesión $(S^{(n)})_{n\geq 0}$ de subconjuntos de X que tiene $S^{(0)} = S$ y tal que $S^{(n+1)} = (S^{(n)})'$ para todo $n \in \mathbb{N}_0$.

(a) Para cada entero no negativo n construya un subconjunto S de $\mathbb{R}$ tal que $S^{(n)} \neq \varnothing$ y $S^{(n+1)} = \varnothing$.

(b) Construya un subconjunto S de $\mathbb{R}$ tal que $S^{(n)} \neq \varnothing$ para todo $n \in \mathbb{N}_0$ y $\bigcap_{n=0}^{\infty} S^{(n)} = \varnothing$.

(c) Construya finalmente un subconjunto S de $\mathbb{R}$ tal que $\bigcap_{n=0}^{\infty} S^{(n)} \neq \varnothing$.

Ejercicio 6.7.13. Como en el ejercicio anterior, sea X un espacio métrico y consideremos la función $D : S \in \mathscr{P}(X) \mapsto S' \in \mathscr{P}(X)$. Elijamos además un ordinal límite α de cardinal mayor que el de X. Para cada subconjunto S de X hay, de acuerdo al *Principio de recurrencia transfinita* 4.3.5, exactmente una función $f_S : \alpha \to \mathscr{P}(X)$ que tiene $f_S(0) = S$ y que en cada elemento β de α toma el valor

$$f_S(\beta) = \begin{cases} D(f_S(\gamma)) & \text{si } \beta \text{ es el sucesor de un ordinal } \gamma; \\ \bigcap\limits_{\gamma\in\beta} f_S(\gamma) & \text{si } \beta \text{ es un ordinal límite.} \end{cases}$$

(a) Muestre que cada vez que β y γ son dos elementos de α tales que $\beta \leq \gamma$ se tiene que $f_S(\beta) \supseteq f_S(\gamma)$, de manera que la función f_S es decreciente.

(b) Pruebe que si hay un elemento β de α tal que $f_S(\beta+1) = f_S(\beta)$, entonces, de hecho, es $f_S(\gamma) = f_S(\beta)$ para todo elemento γ de α tal que $\beta \leq \gamma$.

(c) Muestre que el conjunto $\{\beta \in \alpha : f_S(\beta+1) = f_S(\beta)\}$ no es vacío, de manera que podemos considerar su menor elemento. Llamamos a ese ordinal el ***rango de Cantor–Bendixson***, por Georg Cantor y Ivar Otto Bendixson, y lo escribimos $\mathrm{cbr}(S)$.

(d) Muestre que $\mathrm{cbr}(S)$ depende solamente de S y no de la elección del ordinal límite α que hicimos al principio, y que $f(\mathrm{cbr}(S))$ es el subconjunto perfecto más grande contenido en S.

Capítulo 7

Convergencia

7.1. Sucesiones convergentes

Este capítulo está enteramente dedicado al estudio de la noción presentada por la siguiente definición, que es tiene un rol central en la teoría de los espacios métricos.

Definición 7.1.1. Decimos que una sucesión $x = (x_n)_{n\in\mathbb{N}_0}$ en X ***converge*** a un punto ξ de X si para todo número positivo ε hay un entero n_0 tal que para todo $n \in \mathbb{N}_0$ vale

$$n \geq n_0 \implies d(x_n, \xi) < \varepsilon.$$

Cuando ese es el caso decimos que la sucesión es ***convergente***, y que el punto ξ es un ***límite*** de la sucesión.

La primera observación que tenemos que hacer es la siguiente.

Proposición 7.1.2. *Una sucesión convergente posee exactamente un límite.*

En vista de esto, siempre que tengamos una sucesión $(x_n)_{n\geq 0}$ que es convergente escribiremos sin ambigüedad

$$\lim_{n\to\infty} x_n$$

para denotar a su único límite.

Demostración. Sea $(x_n)_{n\geq 0}$ una sucesión con valores en X y sean ξ y ζ dos de sus límites. Si ε es un número positivo, entonces que la sucesión converja a ξ implica que hay un entero no negativo n_1

tal que para todo $n \in \mathbb{N}_0$ vale $n \geq n_1 \implies d(x_n, \xi) < \varepsilon/2$, y que la sucesión converja a ζ implica que hay un entero no negativo n_2 tal que para todo $n \in \mathbb{N}_0$ vale $n \geq n_2 \implies d(x_n, \zeta) < \varepsilon/2$.

Pongamos $m \coloneqq \max\{n_1, n_2\}$, que es un elemento de $\mathbb{N}_0$. Como $m \geq n_1$ es $d(x_m, \xi) < \varepsilon/2$ y como $m \geq n_2$ es $d(x_m, \zeta) < \varepsilon/2$, así que

$$d(\xi, \zeta) \leq d(\xi, x_m) + d(x_m, \zeta) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Vemos así que el número no negativo $d(\xi, \zeta)$ es menor que todo número positivo, así que es nulo: esto implica que $\xi = \zeta$. □

Es fácil dar ejemplos de sucesiones convergentes.

Ejemplo 7.1.3. La sucesión $(1/(n+1))_{n\geq 0}$ en $\mathbb{R}$ converge a 0. En efecto, si ε es un número real positivo, entonces hay un elemento n_0 de $\mathbb{N}$ tal que $1/n_0 < \varepsilon$, y para todo elemento de $\mathbb{N}_0$ tal que $n \geq n_0$ tenemos que

$$d\left(\frac{1}{n+1}, 0\right) = \left|\frac{1}{n+1} - 0\right| = \frac{1}{n+1} < \frac{1}{n_0} < \varepsilon.$$

Ejemplo 7.1.4. Una sucesión constante $(x_n)_{n\geq 0}$ en X converge. En efecto, sea ξ es el elemento de X tal que $x_n = \xi$ para todo $n \in \mathbb{N}_0$. Si ε es un número positivo, entonces para todo entero n vale $n \geq 0 \implies d(x_n, \xi) < \varepsilon$.

Ejemplo 7.1.5. Sea A un conjunto, sea $X \coloneqq A^{\mathbb{N}_0}$ el conjunto de las sucesiones en A, y consideremos la ultramétrica $d : X \times X \to \mathbb{R}$ de la Proposición 5.2.9.5. Sea $x = (x_n)_{n\geq 0}$ una sucesión en X, de manera que para cada $n \in \mathbb{N}_0$ la componente x_n es una sucesión $(x_{n,i})_{i\geq 0}$ en A, y sea $\xi = (\xi_i)_{i\geq 0}$ un elemento de X. Queremos probar que

$$\begin{aligned} &\textit{la sucesión } x \textit{ converge a } \xi \textit{ si y solamente si para todo } k \in \mathbb{N}_0 \textit{ el conjunto} \\ &I_k \coloneqq \{n \in \mathbb{N}_0 : x_{n,k} \neq \xi_k\} \textit{ es finito.} \end{aligned} \tag{7.1}$$

Supongamos primero que x converge a ξ y sea k un elemento de $\mathbb{N}_0$. La convergencia implica hay un entero n_0 tal que $d(x_n, \xi) < e^{-k}$ para todo $n \in \mathbb{N}_0$ tal que $n \geq n_0$. Si n es un elemento de $\mathbb{N}_0$ tal que $n \geq n_0$, esto nos dice que o bien $x_n = \xi$ o bien $x_n \neq \xi$ y

$$\max\{i \in \mathbb{N}_0 : x_{n,j} = \xi_j \text{ para cada } j \in [\![0, i]\!]\} > k,$$

y en cualquiera de los dos casos es $x_{n,k} = \xi_k$. Vemos así que el conjunto I_k está contenido en $[\![0, n_0 - 1]\!]$ y que es, por lo tanto, finito. Esto prueba que la condición de (7.1) es necesaria.

Supongamos ahora que para todo elemento k de $\mathbb{N}_0$ el conjunto I_k es finito y sea ε un número

positivo. Como $\lim_{n\to\infty} e^{-n} = 0$, podemos elegir r en $\mathbb{N}_0$ de manera que $e^{-r} < \varepsilon$. Pongamos

$$n_0 \coloneqq 1 + \max \bigcup_{k=0}^{r} I_k.$$

Esto tiene sentido porque el conjunto cuyo máximo estamos tomando es una unión finita de conjuntos finitos de enteros. Sea n un elemento de $\mathbb{N}_0$ tal que $n \geq n_0$.

- Si $x_n = \xi$, entonces $d(x_n, \xi) = 0$, por supuesto.
- Si, en cambio, $x_n \neq \xi$, entonces para cada $l \in [\![0, r]\!]$ es $n > \max I_l$, así que $n \notin I_l$ y, por lo tanto, $x_{n,l} = \xi_l$. Esto nos dice que en este caso

  $$r \in E(x_n, \xi) = \{i \in \mathbb{N}_0 : x_{n,j} = \xi_j \text{ para cada } j \in [\![0, i]\!]\},$$

  así que $r \leq \max E(x_n, \xi)$ y $d(x_n, \xi) = e^{-\max E(x_n,\xi)} < e^r < \varepsilon$.

Hemos mostrado que en cualquier caso es $d(x_n, \xi) < \varepsilon$, así que la sucesión x converge a ξ, como queremos, y que la condición de (7.1) es suficiente .

El siguiente ejercicio da una forma alternativa de describir el resultado de este ejemplo.

Ejercicio 7.1.6. En la situación del ejercicio, muestre que la sucesión $(x_n)_{n\geq 0}$ converge a ξ si y solamente si para cada $k \in \mathbb{N}_0$ la sucesión $(x_{n,k})_{n\geq 0}$ es a larga constante de valor ξ_k.

Para dar un ejemplo de una sucesión que no converge — y mostrar con ello que la definición no es trivial — usaremos la siguiente definición.

Definición 7.1.7.

- Un subconjunto de X es ***acotado*** si está contenido en una bola abierta de X.
- Una sucesión $(x_n)_{n\geq 0}$ es ***acotada*** si su imagen $\{x_n : n \in \mathbb{N}_0\}$ es un subconjunto acotado.

Ejercicio 7.1.8. Pruebe que una unión finita de conjuntos acotados es ella misma un conjunto acotado y que, en particular, todo conjunto finito es acotado.

Ejercicio 7.1.9. Pruebe que si un subconjunto S de un espacio métrico X es acotado, entonces para todo punto x de X hay un número positivo tal que $S \subseteq B_r(x)$.

La acotación de una sucesión es una condición necesaria para su convergencia.

Proposición 7.1.10. *Una sucesión convergente es acotada.*

Demostración. Sea $(x_n)_{n\geq 0}$ una sucesión en X que converge y sea ξ su límite. Que converja implica que hay un entero no negativo n_0 tal que para todo $n \in \mathbb{N}_0$ vale $n \geq n_0 \implies d(x_n, \xi) < 1$. El conjunto $\{d(x_n, \xi) : n \in [\![0, n_0]\!], x_n \neq \xi\} \cup \{1\}$ es finito y sus elementos son positivos así que posee un elemento máximo y este es positivo. Escribámoslo r. Si n es un elemento cualquiera de $\mathbb{N}_0$, entonces o bien $x_n = \xi$, y en ese caso $d(x_n, \xi) = 0 < 2r$, o bien $x_n \neq \xi$ y $n < n_0$, y en ese caso $d(x_n, \xi) \leq r < 2r$, o bien $x_n \neq \xi$ y $n \geq n_0$, y en ese caso $d(x_n, \xi) < 1 < 2r$. Vemos así que la sucesión toma valores en la bola $B_{2r}(\xi)$ y que es, por lo tanto, acotada. □

Usando esto podemos exhibir fácilmente ejemplos de sucesiones que no convergen.

Ejemplo 7.1.11. La sucesión $(n)_{n\geq 0}$ en $\mathbb{R}$ no converge. En efecto, si ξ es un elemento cualquiera de ξ, r es un número positivo, y ponemos $m \coloneqq \lceil r + |\xi| \rceil$, tenemos que

$$d(x_m, \xi) = |m - \xi| = \left| \lceil r + |\xi| \rceil - \xi \right| \geq \lceil r + |\xi| \rceil - \xi \geq r + |\xi| - \xi \geq r,$$

y esto nos dice que la sucesión no toma todos sus valores en la bola $B_r(\xi)$.

La acotación es una condición necesaria para la convergencia pero no es suficiente.

Ejemplo 7.1.12. La sucesión $((-1)^n)_{n\geq 0}$ en $\mathbb{R}$ es acotada pero no converge En efecto, sea ξ un número real y sea n_0 un entero no negativo, y sea m el resto de dividir a n_0 por 2 — es $m \in \{0, 1\}$, $n_0 + m$ es par y $n_0 + m + 1$ es impar.

- Si $\xi \geq 0$, entonces $d(x_{n_0+m+1}, \xi) = |(-1)^{n_0+m+1} - \xi| = |-1 - \xi| = 1 + |\xi|$.
- Si $\xi \leq 0$, entonces $d(x_{n_0+m}, \xi) = |(-1)^{n_0+m} - \xi| = |1 - \xi| = 1 + |\xi|$.

Vemos así que en cualquier caso hay un entero n tal que $n \geq n_0$ y $x_n \notin B_{|\xi|}(\xi)$, así que la sucesión no converge a ξ.

Es importante tener en cuenta que el hecho de que una sucesión converja o no depende de la métrica que estamos considerando.

Ejemplo 7.1.13. Consideremos la métrica 3-ádica sobre $\mathbb{Q}$ que describimos en la Proposición 5.2.9.11 y la sucesión $x = (x_n)_{n\geq 0}$ en $\mathbb{Q}$ que para cada $n \in \mathbb{N}_0$ tiene

$$x_n = 1 + 3 + 3^2 + \cdots + 3^n = \frac{3^{n+1} - 1}{3 - 1}.$$

Sea ε un número positivo y sea

$$n_0 \coloneqq \max\left\{0, \left\lceil -\frac{\ln \varepsilon}{\ln 3} - 1 \right\rceil\right\},$$

que es 0 si $\varepsilon \geq 3$. Para todo $n \in \mathbb{N}_0$ tenemos que

$$v_3(x_p + 1/2) = v_3\left(\frac{3^{n+1}-1}{3-1} + \frac{1}{2}\right) = v_3\left(\frac{3^{n+1}}{2}\right) = n+2,$$

así que $d(x_n, -1/2) = 3^{-n-1}$. Si es $n \geq n_0$, entonces

- o bien $\varepsilon \geq 3$, de manera que $n_0 = 0$ y $d(x_n, -1/2) = 3^{-n-1} \leq 3 \leq \varepsilon$,
- o bien $\varepsilon < 3$, de manera que $n \geq n_0 \geq -\ln \varepsilon / \ln 3 - 1$ y $d(x_n, -1/2) = 3^{-n-1} \leq 3^{-n_0-1} \leq \varepsilon$.

Vemos así que la sucesión x converge a $-1/2$, esto es, que

$$\lim_{n\to\infty}\left(1 + 3 + 3^2 + \cdots + 3^n\right) = -\frac{1}{2}$$

o, usando la notación usual de las series, que

$$\sum_{n=0}^{\infty} 3^n = -\frac{1}{2}.$$

Notemos que si consideramos la sucesión x en el conjunto $\mathbb{Q}$ pero ahora dotado de su métrica usual, entonces no converge, ya que no es acotada.

Ejercicio 7.1.14. Sean X e Y dos espacios métricos, y consideremos al producto cartesiano $X \times y$ como un espacio métrico de la manera usual, con su métrica suma. Sean $x = (x_n)_{n\geq 0}$ e $y = (y_n)_{n\geq 0}$ una sucesión en X y una en Y, respectivamente, y sea $z = (z_n)_{n\geq 0}$ la sucesión en $X \times Y$ que para cada $n \in \mathbb{N}_0$ tiene $z_n = (x_n, y_n)$. Pruebe que la sucesión z converge a un punto (ξ, ζ) en $X \times Y$ si y solamente si las sucesiones x e y convergen a ξ y a ζ en X y en Y, respectivamente.

Extienda este resultado a productos cartesianos finitos de espacios métricos.

Usaremos todo el tiempo las siguientes dos observaciones.

Proposición 7.1.15. *Sea X un espacio métrico, sea $x = (x_n)_{n\geq 0}$ una sucesión en X, y sea ξ un elemento de X.*

(*i*) *Si la sucesión x converge a ξ en X, entonces la sucesión $(d(x_n, \xi))_{n\geq 0}$ converge a 0 en $\mathbb{R}$.*

(*ii*) *Si hay una sucesión $(a_n)_{n\geq 0}$ en $\mathbb{R}$ que converge a 0 y tal que $d(x_n, \xi) \leq a_n$ para todo $n \in \mathbb{N}_0$, entonces la sucesión x converge a ξ.*

Notemos que la parte (*ii*) de esta proposición nos dice, en particular, que vale la implicación recíproca a la de la parte (*i*).

Demostración. (*i*) Supongamos que la sucesión x converge a ξ, y sea ε un número positivo. Hay un entero no negativo n_0 tal que para todo entero n vale $n \geq n_0 \implies d(x_n, \xi) < \varepsilon$, y esto implica

que para todo entero n tal que $n \geq 0$ es $d(d(x_n, \xi), 0) = |d(x_n, \xi) - 0| = d(x_n, \xi) < \varepsilon$. Esto nos dice que la sucesión $(d(x_n, \xi))_{n\geq 0}$ converge a 0 en $\mathbb{R}$.

(*ii*) Supongamos ahora que hay una sucesión $(a_n)_{n\geq}$ en $\mathbb{R}$ que converge a 0 y tal que $d(x_n, \xi) \leq a_n$ para todo $n \in \mathbb{N}_0$. Si ε es un número positivo, hay entonces un entero no negativo n_0 tal que para todo entero n vale $n \geq n_0 \implies d(a_n, 0) < \varepsilon$, y esto implica que para cada entero n tal que $n \geq n_0$ es $d(x_n, \xi) \leq a_n = |a_n - 0| = d(a_n, 0) < \varepsilon$. Vemos así que la sucesión $(x_n)_{n\geq 0}$ converge a ξ. □

El siguiente resulta es ocasionalmente muy útil.

Proposición 7.1.16. *Sea X un espacio métrico y sea $\pi : \mathbb{N}_0 \to \mathbb{N}_0$ una función inyectiva. Si $(x_n)_{n\geq 0}$ es una sucesión en X que converge, entonces la sucesión $(x_{\pi(n)})_{n\geq 0}$ también converge y al mismo límite.*

El caso particular de esto en el que la función π es una biyección nos dice que podemos permutar arbitrariamente las componentes de una sucesión convergente sin perturbar su convergencia o su límite.

Demostración. Sea $(x_n)_{n\geq 0}$ una sucesión en X que converge, sea x su límite, y sea ε un número positivo. Como $(x_n)_{n\geq 0}$ converge a x, hay un entero no negativo n_0 tal que para todo $n \in \mathbb{N}_0$ vale $n \geq n_0 \implies d(x_n, x) < \varepsilon$, y esto nos dice que el subconjunto $I \coloneqq \{n \in \mathbb{N}_0 : d(x_n, x) < \varepsilon\}$ de $\mathbb{N}_0$ es tal que la diferencia $\mathbb{N}_0 \setminus I$ está contenida en $[\![0, n_0]\!]$ y, en particular, es finita. Como la función π es inyectiva, esto implica que el conjunto $\mathbb{N}_0 \setminus \pi^{-1}(I) = \pi^{-1}(\mathbb{N}_0 \setminus I)$ es finito, así que hay un entero no negativo n_1 tal que $\mathbb{N}_0 \setminus \pi^{-1}(I) \subseteq [\![0, n_1]\!]$.

Sea ahora n un entero tal que $n > n_1$. La forma en que elegimos n_1 implica que $n \notin \mathbb{N}_0 \setminus \pi^{-1}(I)$, de manera que $n \in \pi^{-1}(I)$ y, por lo tanto $\pi(n) \in I$ y $d(x_{\pi(n)}, x) < \varepsilon$. Esto prueba que la sucesión $(x_{\pi(n)})_{n\geq 0}$ converge a x y, en definitiva, la proposición. □

Es útil tener formas alternativas de ver la convergencia de una sucesión. La siguiente definición introduce lenguaje conveniente para ello.

Definición 7.1.17. Supongamos que tenemos un predicado P que se aplica a las sucesiones en X. Decimos que una sucesión $(x_n)_{n\geq 1}$ ***satisface a la larga*** al predicado P si hay un entero no negativo n_0 tal que la sucesión $(x_{n+n_0})_{n\geq 0}$ satisface al predicado P.

Así, por ejemplo, si S es un subconjunto de X, entonces podemos considerar el predicado $P(x) \coloneqq$ «x tiene sus valores en S». Una sucesión $(x_n)_{n\geq 0}$ satisface a la larga al predicado P si hay un entero no negativo n_0 tal que para todo $n \in \mathbb{N}_0$ vale $n \geq n_0 \implies x_n \in S$. En este caso decimos normalmente que la sucesión ***está a la larga en*** S.

Ejercicio 7.1.18. Sea p un predicado que se aplica a los elementos de X, y sea

$$P(x) \coloneqq \text{para todo } n \in \mathbb{N}_0 \text{ vale } p(x_n)\text{»}.$$

Muestre que una sucesión $(x_n)_{n\geq 0}$ satisface a la larga al predicado P si y solamente si el conjunto $\{n \in \mathbb{N}_0 : \neg P(x_n)\}$ es finito.

Ejercicio 7.1.19. Muestre que una sucesión está acotada si y solamente si está a la larga acotada, y que converge si y solamente si a la larga converge.

Ejercicio 7.1.20. Muestre que una sucesión con finitos valores que converge es a la larga constante.

Usando este lenguaje, podemos dar una forma muy conveniente a la definición de convergencia.

Proposición 7.1.21. *Sea ξ un punto de X y sea $\mathscr{N}$ una base de entornos de ξ. Una sucesión en X converge a ξ si y solamente si está a la larga en cada elemento de $\mathscr{N}$.*

Notemos que esto es cierto cualquiera sea la base de entornos $\mathscr{N}$ de ξ. Muchas veces aprovechamos esto de la siguiente manera: tenemos dos bases de entornos $\mathscr{N}$ y $\mathscr{N}'$ para ξ, mostramos que una sucesión está a la larga en todo elemento de $\mathscr{N}$, de manera que, por la proposición, converge, y concluimos, otra vez por la proposición, que la sucesión está a la larga en todo elemento de $\mathscr{N}'$.

Demostración. Supongamos primero que $(x_n)_{n\geq 0}$ es una sucesión en X que converge a ξ y sea N un elemento de $\mathscr{N}$, de manera que hay un número positivo ε tal que $B_\varepsilon(\xi) \subseteq N$. Como la sucesión converge a ξ, hay un entero no negativo n_0 tal que para todo $n \in \mathbb{N}_0$ vale $n \geq n_0 \implies d(x_n, \xi) < \varepsilon$. Esto nos dice que si n es un elemento de $\mathbb{N}_0$ tal que $n \geq n_0$, entonces $d(x_n, \xi) < \varepsilon$ y, por lo tanto, $x_n \in B_\varepsilon(\xi) \subseteq N$. Vemos así que la sucesión está a la larga en N y, en consecuencia, que la condición que da la proposición es necesaria.

Para ver que también es suficiente, supongamos que $x = (x_n)_{n\geq 0}$ es una sucesión que la satisface y sea ε un número positivo. Como $B_\varepsilon(\xi)$ es un entorno de ξ y $\mathscr{N}$ es una base de entornos de ξ, hay un elemento N de $\mathscr{N}$ tal que $N \subseteq B_\varepsilon(\xi)$. La hipótesis nos dice entonces que la sucesión x está a la larga en N, esto es, que hay un entero un negativo n_0 tal que para todo $n \in \mathbb{N}_0$ vale $n \geq n_0 \implies x_n \in N$. En particular, si n es un elemento de $\mathbb{N}_0$ tal que $n \geq n_0$, entonces tenemos que $x_n \in N \subseteq B_\varepsilon(\xi)$, así que $d(x_n, \xi) < \varepsilon$. Vemos así que la sucesión x converge a ξ. □

Una consecuencia de la Proposición 7.1.21 es que la convergencia de sucesiones de un espacio está completamente determinada por la topología de este. El siguiente corolario de la proposición explicita esto.

Corolario 7.1.22. *Una sucesión x en X converge a un punto ξ si y solamente si está a la larga en todo abierto de X que contiene a ξ.*

Esto implica, claro, que para saber si una sucesión en X converge a un punto es suficiente conocer todos los abiertos de X.

Demostración. Esto es consecuencia de la Proposición 7.1.21, ya que para todo punto ξ de X el conjunto $\mathscr{N}$ de todos los abiertos de X que contienen a ξ es una base de entornos de ξ. □

Todo esto nos dice que la topología determina las sucesiones convergentes y sus límites. Mostremos ahora que, recíprocamente, las sucesiones convergentes y sus límites determinan la topología.

Proposición 7.1.23.

(*i*) *Un subconjunto U de X es abierto si y solamente si toda sucesión en X que converge a un punto de U está a la larga en U.*

(*ii*) *Un subconjunto F de X es cerrado si y solamente si toda sucesión convergente que a la larga esta en F tiene su límite en F.*

Demostración. (*i*) Supongamos que U es un subconjunto abierto de X y sea x una sucesión en X que converge a un punto ξ de U. Como U es un entorno de ξ, la Proposición 7.1.21 nos dice que x está a la larga en U. Vemos así que la condición que da la parte (*i*) de la proposición es necesaria.

Supongamos ahora que U es un subconjunto de X tal que toda sucesión convergente que tiene su límite en U está a la larga en U, y supongamos que U no es abierto, de manera que hay un punto ξ de U que no es interior a U. Para cada entero no negativo n la bola $B_{1/(n+1)}(x)$ es un entorno de ξ, así que $B_{1/(n+1)}(\xi) \not\subseteq U$ y, por lo tanto, hay un elemento $x_n \in B_{1/(n+1)}(\xi) \setminus U$. Obtenemos de esta forma una sucesión $x \coloneqq (x_n)_{n\geq 0}$.

Esa sucesión converge a ξ. En efecto, si ε es un número positivo, entonces hay un entero no negativo n_0 tal que $1/(n_0+1) < \varepsilon$, y entonces cada vez que n es un elemento de $\mathbb{N}_0$ tal que $n \geq n_0$ tenemos que $x_n \in B_{1/(n+1)}(\xi)$, por lo que $d(x_n, \xi) < 1/(n+1) \leq 1/(n_0+1) < \varepsilon$.

Ahora bien, como la sucesión x converge a ξ, que es un elemento de U, la hipótesis que hicimos sobre U implica que x está a la larga en U. Hay, por lo tanto, un entero no negativo n_1 tal que para todo $n \in \mathbb{N}_0$ vale $n \geq n_1 \implies x_n \in U$. En particular, tenemos que $x_{n_1} \in U$: esto es absurdo, porque x_{n_1} es un elemento de $B_{1/(n_1+1)}(\xi) \setminus U$. Esta contradicción provino de haber supuesto que el conjunto U no es abierto, así que debe serlo. Esto prueba que la condición que da la parte (*i*) de la proposición es suficiente.

(*ii*) Sea F un subconjunto cerrado de X, sea $x = (x_n)_{n\geq 0}$ una sucesión convergente que a la larga está en F, y supongamos que el límite ξ de x no pertenece a F. Como ξ pertenece a $X \setminus F$ y este conjunto es abierto, lo que ya probamos nos dice que la supongamos x está a la larga en $X \setminus F$,

y esto es absurdo ya que a la larga está en F. Vemos así que debe ser $\xi \in F$.

Supongamos ahora que F es un subconjunto de X que contiene el límite de toda sucesión convergente de X que está a la larga en F. Sea ξ un elemento de $\overline{F}$. Para cada entero no negativo n el conjunto $B_{1/(n+1)}(\xi)$ es un entorno de ξ, así que podemos elegir un elemento x_n en el conjunto $B_{1/(n+1)}(\xi) \cap F$. Obtenemos de esta forma una sucesión $x \coloneqq (x_n)_{n\geq 0}$ con valores en F, y es claro que x converge a ξ: la hipótesis que hicimos sobre F implica entonces que $\xi \in F$. Vemos de esta forma que $\overline{F} \subseteq F$, por lo que podemos concluir que el conjunto F es cerrado. □

Todo esto nos da un criterio de equivalencia entre métricas:

Corolario 7.1.24. *Sea X un conjunto y sean d y d' dos métricas sobre X. Las siguientes afirmaciones son equivalentes:*

(*a*) *Las métricas d y d' son topológicamente equivalentes.*

(*b*) *Una sucesión converge con respecto a d si y solamente si converge con respecto a d' y, cuando ese es el caso, tiene el mismo límite con respecto a d que con respecto a d'.*

En particular, métricas topológicamente equivalentes tienen las mismas sucesiones convergentes.

Demostración. La implicación $(a) \Rightarrow (b)$ es consecuencia del Corolario 7.1.22. Supongamos ahora que la condición (b) se cumple, sea U un subconjunto abierto con respecto a la métrica d, y mostremos que U es también abierto con respecto a la métrica d' — por la simetría de la situación, esto es suficiente para probar la implicación $(b) \Rightarrow (a)$. Sea x una sucesión que converge a un punto ξ de U con respecto a la métrica d'. La hipótesis implica que x converge a ξ también con respecto a la métrica d y, como U es abierto con respecto a esa métrica, la Proposición 7.1.23 nos dice que x está a la larga en U. Vemos así que toda sucesión que converge a un punto de U con respecto a d' está a la larga en U, así que U es un abierto con respecto a esa métrica d', como queremos. □

Una aplicación del hecho de que podemos describir la clausura de un conjunto usando sucesiones es que podemos acotar el cardinal de los espacios separables.

Proposición 7.1.25. *El cardinal de un espacio métrico separable es a lo sumo el del continuo.*

Demostración. Sea X un espacio métrico separable, de manera que hay en X un subconjunto denso contable D. El conjunto C de todas las sucesiones convergentes con valores en D es un subconjunto de $D^{\mathbb{N}_0}$, así que su cardinal es menor o igual que $\mathfrak{c}$. Hay, por otro lado, una función $x \in C \mapsto \lim x \in X$ y la imagen de esta función es $\overline{D} = X$, así que es sobreyectiva. Esto implica, claro, que $|X| \leq |C| \leq \mathfrak{c}$. □

Usando sucesiones convergentes podemos describir también el interior, la clausura, y el conjunto derivado de un conjunto. Dejamos los detalles de esto a cargo del lector.

Ejercicio 7.1.26. Sea S un subconjunto de X y sea ξ un elemento de X. Pruebe las siguientes afirmaciones:

(a) ξ es un punto interior de S si y solamente si toda sucesión convergente en X que converge a ξ está a la larga en S.

(b) ξ es un punto de adherencia de S si y solamente si hay una sucesión convergente con valores en S que converge a ξ.

(c) ξ es un punto de acumulación de S si y solamente si hay una sucesión $(x_n)_{n\geq 1}$ con valores en S que converge a ξ y tal que $x_n \neq \xi$ para todo $n \in \mathbb{N}_0$.

Una consecuencia inmediata de la caracterización de los conjuntos cerrados usando sucesiones es la siguiente forma de acotar el límite de una sucesión convergente.

Ejercicio 7.1.27. Muestre que si una sucesión x toma valores en una bola abierta o cerrada B y converge, entonces su límite pertenece a la correspondiente bola cerrada. Deduzca de ello que si K es un número real y $x = (x_n)_{n\geq 0}$ es una sucesión en un espacio normado V que converge a un punto ξ y tal que $\|x_n\| \leq K$ para todo $n \in \mathbb{N}_0$, entonces $\|\xi\| \leq K$.

Hay otras nociones de convegencia además de la de sucesiones que estudiamos en esta sección. Los siguientes dos ejercicios presentan las dos más importantes.

Ejercicio 7.1.28. Hagamos las siguientes definiciones:

- Un ***conjunto dirigido*** es un conjunto Λ dotado de una relación de pre-orden $\preceq$ que satisface la siguiente condición:

  si α y β son dos elementos de Λ, entonces hay un elemento γ de Λ tal que $\alpha \preceq \gamma$ y $\beta \preceq \gamma$.

- Una ***red*** en un conjunto X es una función $\phi : \Lambda \to X$ cuyo dominio es un conjunto dirigido.
- Decimos que una red $\phi : \Lambda \to X$ en un conjunto X está ***a la larga*** en un subconjunto U de X si hay un elemento α de Λ tal que $\phi(\beta) \in U$ para todo $\beta \in \Lambda$ con $\alpha \preceq \beta$.
- Finalmente, decimos que una red $\phi : \Lambda \to X$ en un espacio metrico converge a un punto ξ de X si está a la larga en todo entorno de ξ en X.

Estas definiciones fueron dadas originalmente por Eliakim Hastings Moore y Herman Smith en el artículo [MS1922], que tiene por título «*A General Theory of Limits*», pero el nombre de red fue propuesto por John L. Kelley en [Kel1950].

(a) Consideremos en $\mathbb{N}_0$ la relación $\preceq$ de orden usual. Muestre que una red $\phi : \mathbb{N}_0 \to X$ en un espacio métrico X converge a un punto ξ de X si y solamente si la sucesión $(\phi(n))_{n\geq 0}$ converge a ξ en X. Esto muestra que la noción de convergencia de redes generaliza a la de

convergencia de sucesiones.

(b) Muestre que una red en un espacio métrico converge a lo sumo a un punto.

Demos un ejemplo de cómo podemos usar la noción de convergencia de redes. Sea I un conjunto cualquiera, sea $f : I \to \mathbb{R}$ una función, veamos al conjunto $\mathscr{P}_f(I)$ de las partes finitas de I como un conjunto ordenado con respecto a la relación de inclusión — es fácil ver que es unconjunto dirigido, de hecho — y consideremos la función

$$s_f : F \in \mathscr{P}_f(I) \mapsto \sum_{i \in F} f(i) \in \mathbb{R},$$

que es una red en $\mathbb{R}$. Decimos que la función f es ***sumable*** si esta red converge, y en ese case llamamos al límite de esa red la ***suma*** de d.

(c) Muestre que una función $f : \mathbb{N}_0 \to \mathbb{R}$ es sumable si y solamente si la serie $\sum_{n=0}^{\infty} f(n)$ converge absolutamente.

(d) Muestre que si una función $f : I \to \mathbb{R}$ que toma valores no negativos es sumable, entonces el conjunto $\{i \in I : f(i) > 0\}$ es contable.

Ejercicio 7.1.29. Otra vez hacemos algunas definiciones.

- Un ***pre-filtro*** en un conjunto X es una famila no vacía $\mathscr{F}$ de subconjuntos no vacíos de X que satisface la siguiente condición:

  > *si S y T son elementos de $\mathscr{F}$, entonces hay un elemento U de $\mathscr{F}$ tal que $U \subseteq S \cap T$.*

- Decimos que un pre-filtro $\mathscr{F}$ en un espacio métrico X ***converge*** a un punto ξ de X si todo entorno de ξ en X contiene un elemento de $\mathscr{F}$.

Estas definiciones fueron originalmente propuestas por Henri Cartan en [Car1937], y son la base del tratamiento de la noción de convergencia dada por Nicolas Bourbaki en su *Topologie Générale* [Bou2007a, Bou2007b].

(a) Muestre que si X es un conjunto, entonces $\mathscr{F}(X) \coloneqq \{F \in \mathscr{P}(X) : X \setminus F \text{ es finito}\}$ es un pre-filtro en X, al que llamamos el ***pre-filtro de Fréchet*** en X.

(b) Sea $x = (x_n)_{n \geq 0}$ una sucesión en un espacio métrico X. Muestre que el conjunto $\mathscr{F} \coloneqq \{\{x_i : i \in F\} : F \in \mathscr{F}(\mathbb{N}_0)\}$ es un pre-filtro en X y que la sucesión x converge a un punto ξ de X si y solamente si el pre-filtro $\mathscr{F}$ converge a ξ.

7.2. Subsucesiones

Definición 7.2.1. Sea $x = (x_n)_{n\geq 0}$ una sucesión en un conjunto X. Si $f : \mathbb{N}_0 \to \mathbb{N}_0$ es una función estrictamente creciente de $\mathbb{N}_0$ a $\mathbb{N}_0$, entonces la ***subsucesión*** de x correspondiente a f es la sucesión $(y_n)_{n\geq 0}$ que para cada $n \in \mathbb{N}_0$ tiene componente $y_n = x_{f(n)}$.

Si vemos a la sucesión x como la función $x : n \in \mathbb{N}_0 \mapsto x_n \in X$, entonces la subsucesión de x correspondiente a f es simplemente la composición $a \circ f$.

Notemos que si x es una sucesión en un conjunto X, y es la subsucesión de x correspondiente a una función estrictamente creciente $f : \mathbb{N}_0 \to \mathbb{N}_0$, y z es la subsucesión de y correspondiente a una segunda función estrictamente creciente $g : \mathbb{N}_0 \to \mathbb{N}_0$, entonces z es la subsucesión de x correspondiente a la función $f \circ g : \mathbb{N}_0 \to \mathbb{N}_0$, que es, claro, estrictamente creciente — es importante notar el orden en que componemos estas funciones. Así, una subsucesión de una subsucesión de x es una subsucesión de x.

Ejercicio 7.2.2. Sea P un predicado que se aplica a los elementos de un conjunto X. Muestre que si x es una sucesión en X que a la larga satisface al predicado P, entonces toda subsucesión de X a la larga satisface al predicado P.

El resultado de este ejercicio tiene la siguiente consecuencia inmediata.

Proposición 7.2.3. *Toda subsucesión de una subsucesión convergente x es ella misma convergente y su límite coincide con el de x.*

Demostración. Sea x una sucesión en X que converge a un elemento ξ de X, de manera que x está a la larga en todo entorno de ξ. Si y es una subsucesión de x, entonces y también está a la larga en todo entorno de ξ, así que y converge a ξ. □

Ejercicio 7.2.4. Sea X un espacio métrico y sea $x = (x_n)_{n\geq 0}$ una sucesión en X. Pruebe que x converge a un punto ξ de X si y solamente si las subsucesiones $(x_{2n})_{n\geq 0}$ y $(x_{2n+1})_{n\geq 0}$ de x convergen a ξ.

Cuando trabajamos con subsucesiones, el siguiente lema es extremadamente útil.

Lema 7.2.5.

(*i*) *Si $f : \mathbb{N}_0 \to \mathbb{N}_0$ es una función estrictamente creciente, entonces para todo $n \in \mathbb{N}_0$ se tiene que $f(n) \geq n$.*

(*ii*) *Si I es un subconjunto infinito de $\mathbb{N}_0$, entonces hay una función estrictamente creciente $f : \mathbb{N}_0 \to \mathbb{N}_0$ cuya imagen es I.*

Demostración. (*i*) Sea $f : \mathbb{N}_0 \to \mathbb{N}_0$ una función estrictamente creciente. Claramente vale que $f(0) \geq 0$, simplemente porque $f(1)$ es un elemento de $\mathbb{N}_0$. Por otro lado, si n es un elemento cualquiera de $\mathbb{N}_0$ tal que $f(n) \geq n$, entonces que f sea estrictamente creciente implica que $n \leq f(n) < f(n+1)$ y, como $f(n+1)$ es un entero, que, de hecho, $n+1 \leq f(n+1)$. Esto implica que la afirmación (*i*) es cierta.

(*ii*) Sea I un subconjunto infinito de $\mathbb{N}_0$. Si n es un elemento de $\mathbb{N}_0$, entonces el conjunto $R(n) \coloneqq I \setminus [\![0, n]\!]$ no es vacío y está contenido en $\mathbb{N}_0$, así que tiene sentido considerar su menor elemento. Hay, por lo tanto, una función $F : n \in \mathbb{N}_0 \mapsto \min I \setminus [\![0, n]\!] \in \mathbb{N}_0$. De acuerdo al *Principio de recurrencia* 1.2.1, hay una función $f : \mathbb{N}_0 \to \mathbb{N}_0$ tal que $f(0) = \min I$ y $f(n+1) = F(f(n))$ para todo $n \in \mathbb{N}_0$. Mostremos que esta función satisface las condiciones que queremos.

- La función f es estrictamente creciente. En efecto, si n es un elemento de $\mathbb{N}_0$, entonces

  $$f(n+1) = F(f(n)) = \min I \setminus [\![0, f(n)]\!],$$

  así que $f(n+1) \notin [\![0, f(n)]\!]$, de manera que, de hecho, $f(n+1) > f(n)$.
- La imagen de f está contenida en I, ya que $f(0) = \min I \in I$ y para cada entero positivo n tenemos que $f(n) = F(f(n-1)) = \min I \setminus [\![0, f(n-1)]\!] \in I$.
- La imagen de f es exactamente I. Sea y un elemento de I. El conjunto $I \cap [\![0, y-1]\!]$ es finito. Sea n su cardinal: mostremos que $f(n) = y$, haciendo inducción con respecto a n.
  - Si $n = 0$, entonces el conjunto $I \cap [\![0, y-1]\!]$ es vacío, así que todo elemento de I es mayor o igual a y. Como además $y \in I$, esto nos dice que $y = \min I = f(0)$.
  - Supongamos ahora que para todo elemento y de I tal que $|I \cap [\![0, y-1]\!]| = n$ es $y = f(n)$, y sea z un elemento de I tal que $|I \cap [\![0, z-1]\!]| = n+1$. En particular, el conjunto $I \cap [\![0, z-1]\!]$ es finito y no vacío, así que podemos considerar su elemento máximo w. Es

    $$I \cap [\![0, w]\!] = (I \cap [\![0, w-1]\!]) \cup \{w\} = I \cap [\![0, z-1]\!] \tag{7.2}$$

    y $w \notin I \cap [\![0, w-1]\!]$, así que $|I \cap [\![0, w-1]\!]| = n$, y la hipótesis inductiva nos dice que $f(n) = w$. Como $w \in I \cap [\![0, z-1]\!]$, es $w < z$, así que $z \in I \setminus [\![0, w]\!]$. Por otro lado, como vale (7.2) sabemos que

    $$I \setminus [\![0, w]\!] = I \setminus (I \cap [\![0, w]\!]) = I \setminus (I \cap [\![0, z-1]\!]) = I \setminus [\![0, z-1]\!],$$

    así que todos los elementos de $I \setminus [\![0, w]\!]$ son mayores o iguales que z. Se sigue de esto que

    $$f(n+1) = F(f(n)) = \min I \setminus [\![0, f(n)]\!] = \min I \setminus [\![0, w]\!] = z.$$

Esto completa la inducción. □

Veamos un ejemplo de cómo podemos usar este lema.

Proposición 7.2.6. *Sea ξ un punto de X. Si x es una sucesión que no converge a ξ, entonces hay un entorno N de ξ y una subsucesión de x con valores en $X \setminus N$.*

Demostración. Sea $x = (x_n)_{n\geq 0}$ una sucesión en X que no converge a ξ. No es cierto que x está a la larga en todo entorno de ξ, así que hay un entorno N de ξ tal que el conjunto $\{n \in \mathbb{N}_0 : x_n \notin N\}$ es infinito. El lema nos dice que hay una función estrictamente creciente $f : \mathbb{N}_0 \to \mathbb{N}_0$ cuya imagen es I y la subsucesión $(x_{f(n)})_{n\geq 0}$ de x correspondiente a f claramente toma valores en $X \setminus N$. Esto prueba la proposición. □

Queremos estudiar el conjunto de los límites de las subsucesiones convergentes de una sucesión dada. Antes que nada, como siempre, le ponemos nombre.

Definición 7.2.7. Sea x una sucesión en X. Un punto ξ de X es un ***punto límite*** de x si hay una subsucesión de x que converge a ξ. Escribimos $L(x)$ al conjunto de todos los puntos límite de x.

Veamos algunos ejemplos.

Ejemplo 7.2.8. Todas las subsucesión de la sucesión $x \coloneqq (n)_{n\geq 0}$ en $\mathbb{R}$ son no acotadas, así que ninguna de ellas converge: esto nos dice que $L(x) = \varnothing$.

Ejemplo 7.2.9. Toda subsucesión convergente de $x \coloneqq ((-1)^n)_{n\geq 0}$ es eventualmente constante, ya que toma finitos valores, y es claro entonces que su límite es un elemento de $\{\pm 1\}$. Como las subsucesiones que corresponden a las funciones $n \in \mathbb{N}_0 \mapsto 2n \in \mathbb{N}_0$ y $n \in \mathbb{N}_0 \mapsto 2n+1 \in \mathbb{N}_0$ convergen a 1 y a -1, podemos concluir entonces que $L(x) = \{\pm 1\}$.

Ejemplo 7.2.10. Sabemos que hay una función biyectiva $\phi : \mathbb{N}_0 \to \mathbb{Q}$. Consideremos la sucesión $x \coloneqq (\phi(n))_{n\geq 0}$ en $\mathbb{R}$ y mostremos que $L(x) = \mathbb{R}$.

Sea ξ un elemento de $\mathbb{R}$. Si n es un elemento de $\mathbb{N}_0$, entonces el conjunto

$$R(n) \coloneqq \{m \in \mathbb{N}_0 : m > n, |\phi(m) - \xi| < 1/(n+1)\}$$

no es vacío. En efecto, la intersección $A(n) \coloneqq (\xi - 1/(n+1), \xi + 1/(n+1)) \cap \mathbb{Q}$ es infinita, así que no está contenida en el conjunto finito $B(n) \coloneqq \phi(\llbracket 0, n \rrbracket)$, y es $R(n) = \phi^{-1}(A(n) \setminus B(n))$. Hay entonces una función $f : \mathbb{N}_0 \to \mathbb{N}_0$ tal que $f(0) = 0$ y $f(n+1) \in R(f(n))$ para todo $n \in \mathbb{N}_0$. Consideremos la sucesión $y \coloneqq (\phi(f(n)))_{n\geq 0}$. Si n es un elemento de $\mathbb{N}_0$, entonces $f(n+1) \in R(f(n))$ y, por lo tanto, $f(n+1) > f(n)$ y $|\phi(f(n+1)) - \xi| < 1/(n+2)$. Esto nos dice, por un lado, que la función f es estrictamente creciente y, por lo tanto, que la sucesión y es una subsucesión de x, y, por otro, que converge a ξ.

Ejercicio 7.2.11. Muestre que un espacio métrico X es separable si y solamente si hay una sucesión x en X tal que $L(x) = X$.

Podemos caracterizar los puntos límite de una sucesión fácilmente:

Proposición 7.2.12. *Sea $x = (x_n)_{n\geq 0}$ una sucesión en X, sea ξ un punto de X, y sea $\mathscr{N}$ una base de entornos de ξ. El punto ξ es un punto límite de x si y solamente si para todo elemento N de $\mathscr{N}$ el conjunto $\{n \in \mathbb{N}_0 : x_n \in N\}$ es infinito.*

Así, un punto es un punto límite de una sucesión si esta pasa infinitas veces por cada entorno de aquel.

Demostración. Supongamos primero que ξ es un punto límite de x, de manera que hay una función estrictamente creciente $f : \mathbb{N}_0 \to \mathbb{N}_0$ tal que la subsucesión $y \coloneqq (x_{f(n)})_{n\geq 0}$ converge a ξ. Sea N un elemento de $\mathscr{N}$. Que y converja a ξ implica que y está a la larga en N, así que el conjunto $I \coloneqq \{n \in \mathbb{N}_0 : x_{f(n)} \notin N\}$ es finito. El conjunto $J \coloneqq \{n \in \mathbb{N}_0 : x_{f(n)} \in N\}$ coincide con $\mathbb{N}_0 \setminus I$, así que es infinito y, por lo tanto, también lo es $K \coloneqq \{n \in \mathbb{N}_0 : x_n \in N\}$, ya que $f(J) \subseteq N$ y la función f es inyectiva. Vemos así que la condición de la proposición es necesaria.

Supongamos ahora que la condición de la proposición se cumple. Si m es un elemento de $\mathbb{N}_0$, entonces hay un elemento N de $\mathscr{N}$ tal que $N \in B_{1/(m+1)}(\xi)$, y entonces el conjunto $\{n \in \mathbb{N}_0 : x_n \in B_{1/(m+1)}(\xi)\}$ es infinito, ya que contiene a $\{n \in \mathbb{N}_0 : x_n \in N\}$. Esto implica, claro, que el conjunto

$$R(m) \coloneqq \{n \in \mathbb{N}_0 : n > m, x_n \in B_{1/(m+1)}(\xi)\}$$

no es vacío. De acuerdo al *Principio de elección dependiente* 2.1.6, hay entonces una función $f : \mathbb{N}_0 \to \mathbb{N}_0$ tal que $f(0) = 0$ y $f(n+1) \in R(f(n))$ para todo $n \in \mathbb{N}_0$. Esto nos dice, en primer lugar, que $f(n+1) > f(n)$ para todo $n \in \mathbb{N}_0$, así que la función f es estrictamente creciente, $(x_{f(n)})_{n\geq 0}$ es una subsucesión de x, y $f(n) \geq n$ para todo $n \in \mathbb{N}$. Además de eso, tenemos que para cada $n \in \mathbb{N}_0$ es $x_{f(n+1)} \in B_{1/(f(n)+1)}(\xi)$, esto es,

$$d(x_{f(n+1)}, \xi) < \frac{1}{f(n)+1} \leq \frac{1}{n+1},$$

así que la subsucesión $(x_{f(n)})_{n\geq 0}$ converge a ξ. □

El conjunto de puntos límite de una sucesión es de una naturaleza similar al de los puntos de acumulación de un conjunto. Nuestra siguiente proposición explica esto.

Proposición 7.2.13. *Sea $x = (x_n)_{n\geq 0}$ una sucesión y sea $S \coloneqq \{x_n : n \in \mathbb{N}_0\}$ su imagen en X. Es $S' \subseteq L(x)$ y la inclusión es estricta si y solamente si hay un número real ζ y un número real positivo ε*

tales que $\{n \in \mathbb{N}_0 : x_n = \zeta\}$ *es infinito y* $\{n \in \mathbb{N}_0 : 0 < d(x_n, \zeta) < \varepsilon\}$ *es vacío.*

Se sigue de esto, por ejemplo, que si la sucesión x toma cada valor un número finito de veces, entonces vale la igualdad $S' = L(x)$.

Demostración. Sea ξ un elemento de S'. Si n es un entero no negativo, entonces que ξ esté en S' implica que el conjunto $S \cap B_{1/(n+1)}(\xi)$ es infinito, así que la preimagen de ese conjunto por la función sobreyectiva $m \in \mathbb{N}_0 \mapsto x_m \in S$ es infinita, y, en definitiva, el conjunto

$$R(n) \coloneqq \{m \in \mathbb{N}_0 : m > n, x_m \in B_{1/n}(\xi)\}$$

no es vacío. De acuerdo al *Principio de elección dependiente* 2.1.6, hay entonces una función $f : \mathbb{N}_0 \to \mathbb{N}_0$ tal que $f(0) = 0$ y $f(n+1) \in R(f(n))$ para todo $n \in \mathbb{N}_0$. Esto nos dice que $f(n+1) > f(n)$ para todo $n \in \mathbb{N}_0$, de manera que la función f es estrictamente creciente y $y = (x_{f(n)})_{n\geq 0}$ es una subsucesión de x. Por otro lado, nos dice que

$$x_{f(n)} \in B_{1/(f(n)+1)}(\xi) \subseteq B_{1/(n+1)}(\xi)$$

para todo $n \in \mathbb{N}_0$, de manera que esa subsucesión y converge a ξ y, por lo tanto, ξ pertenece a $L(x)$. Esto prueba que $S' \subseteq L(x)$.

Supongamos que $S' \subsetneq L(x)$, de manera que hay un elemento ζ de $L(x)$ que no pertenece a S'. Como ζ pertenece a $L(x)$, hay una función estrictamente creciente $f : \mathbb{N}_0 \to \mathbb{N}_0$ tal que la subsucesión $y \coloneqq (x_{f(n)})_{n\geq 0}$ converge a ζ, y como ζ no pertenece a S', sabemos que hay un número positivo ε tal que $B_\varepsilon(\zeta) \cap S \subseteq \{\zeta\}$. Ahora bien, la subsucesión y converge a ζ y toma valores en S, así que está a la larga en $B_\varepsilon(\zeta) \cap S$. Esto nos dice que y es a la larga constante de valor ζ, de manera que hay un entero n_0 tal que $x_{f(n)} = \zeta$ para todo $n \geq n_0$ y, por lo tanto, que el conjunto $I \coloneqq \{n \in \mathbb{N}_0 : x_{f(n)} = \zeta\}$ es infinito. Como la función f es inyectiva y la imagen de I por f es precisamente $\{n \in \mathbb{N}_0 : x_n = \zeta\}$, este último conjunto es infinito. Por otro lado, el conjunto $\{n \in \mathbb{N}_0 : 0 < d(x_n, \zeta) < \varepsilon\}$ es vacío: si hubiera un elemento n en él, entonces x_n pertenecería a a $B_\varepsilon(\zeta) \cap S \subseteq \{\zeta\}$ y sería distinto de ζ, lo que es absurdo.

Esto muestra que la condición dada por la proposición para que la inclusión $S' \subseteq L(x)$ sea estricta es necesaria. Mostremos, para terminar, que también es suficiente. Supongamos que se cumple, de manera que hay un número ζ y un número positivo ε tales que el conjunto $I \coloneqq \{n \in \mathbb{N}_0 : x_n = \zeta\}$ es infinito y el conjunto $J \coloneqq \{n \in \mathbb{N}_0 : 0 < d(x_n) < \varepsilon\}$ es vacío. Como el conjunto I es infinito, sabemos que hay una función estrictamente creciente $f : \mathbb{N}_0 \to \mathbb{N}_0$ que lo tiene por imagen y claramente la subsucesión $(x_{f(n)})_{n\geq 0}$ de x que corresponde a f es constante de valor ζ, así que converge a ζ: vemos así que $\zeta \in L(x)$. Si ζ perteneciera a S', la intersección $B_\varepsilon(\zeta) \cap S$ sería infinita: tendría entonces un elemento distinto de ζ, de manera que habría un entero no negativo m tal que $0 < d(x_m, \zeta) < \varepsilon$: esto es absurdo, ya que en ese caso tendríamos que m pertenece a $J = \varnothing$. □

7.3. Completitud

Vimos que la acotación es una condición necesaria para la convergencia de una sucesión. La siguiente es otra condición necesaria, mucho más fuerte.

Definición 7.3.1. Decimos que una sucesión $(x_n)_{n\geq 0}$ en un espacio métrico satisface la ***condición de Cauchy***, y también que ***es de Cauchy***, si para todo número positivo ε hay un entero n_0 tal que siempre que m y n son elementos de $\mathbb{N}_0$ vale

$$m, n \geq n_0 \implies d(x_m, x_n) < \varepsilon.$$

Podemos reescribir esta condición en términos de una base de entornos de 0 en $\mathbb{R}$.

Ejercicio 7.3.2. Sea $\mathscr{N}$ una base de entornos de 0 en $\mathbb{R}$. Pruebe que una sucesión $(x_n)_{n\geq 0}$ en un espacio métrico satisface la condición de Cauchy si y solamente si para todo elemento N de $\mathscr{N}$ hay un entero n_0 tal que siempre que m y n son elementos de $\mathbb{N}_0$ vale

$$m, n \geq n_0 \implies d(x_m, x_n) \in N.$$

Ejercicio 7.3.3. Sea X un espacio métrico, sea $x = (x_n)_{n\geq 0}$ una sucesión en X, y, para cada $n \in \mathbb{N}_0$, sea $A_n \coloneqq \{x_m : m \in \mathbb{N}_0, m \geq n\}$. Pruebe que la sucesión x satisface la condición de Cauchy si y solamente la sucesión de números reales $(\operatorname{diam} A_n)_{n\geq 0}$ converge a 0.

Mostremos que la condición de Cauchy es, como dijimos, más fuerte que la de acotación.

Proposición 7.3.4. *Una sucesión de Cauchy es acotada.*

Demostración. Sea $x = (x_n)_{n\geq 1}$ una sucesión de Cauchy en un espacio métrico X. Hay un entero no negativo n_0 tal que siempre que m y n son elementos de $\mathbb{N}_0$ vale $m, n \geq n_0 \implies d(x_m, x_n) < 1$, y esto implica claramente que el conjunto $A \coloneqq \{x_n : n \in \mathbb{N}_0, n \geq n_0\}$ está contenido en la bola $B_1(x_{n_0})$ y que es, por lo tanto, acotado. Como la imagen de la sucesión x es la unión del conjunto finito $\{x_i : i \in [\![0, n_0]\!]$, que es acotado, y A, es acotada. La sucesión x es, en consecuencia, acotada. □

Es importante observar que que una sucesión satisfaga la condición de Cauchy es algo que depende de la métrica del espacio y no solamente de su topología.

Ejemplo 7.3.5. Sobre el conjunto $\mathbb{N}$ consideremos la métrica discreta d y la función

$$d' : (m, n) \in \mathbb{N} \times \mathbb{N} \mapsto \left| \frac{1}{m} - \frac{1}{n} \right| \in \mathbb{R},$$

que también es una métrica sobre $\mathbb{N}$, como es fácil ver. Si m y n son elementos de $\mathbb{N}$ y $m \in B_{1/n^2}(n)$, entonces $-1/n^2 < 1/m - 1/n < 1/n^2$, así que $n - 1 < n^2/m < n^2 + 1$ y, por lo tanto, $m = n$. Vemos con esto que la topología determinada por la métrica d' es discreta, de manera que las dos métricas d y d' son topológicamente equivalentes.

Consideremos la sucesión $x = (x_n)_{n\geq 0}$ en $\mathbb{N}$ que tiene $x_n = n + 1$ para todo $n \in \mathbb{N}_0$. Es evidente que x no satisface la condición de Cauchy con respecto a la métrica d, ya que no hay dos elementos m y n de $\mathbb{N}_0$ tales que $d(x_m, x_n) < 1/2$. La sucesión es, por otro lado, una sucesión de Cauchy con respecto a la métrica d'. En efecto, si ε es un número positivo y ponemos $n_0 \coloneqq \lceil 2/\varepsilon \rceil$, entonces siempre que m y n son elementos de $\mathbb{N}_0$ tales que $m, n \geq n_0$ tenemos que

$$d(x_m, x_n) = \left| \frac{1}{m+1} - \frac{1}{n+1} \right| \leq \frac{1}{m+1} + \frac{1}{n+1} \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

El siguiente criterio es frecuentemente útil.

Proposición 7.3.6. *Sea X un espacio métrico y sea $\sum_{n=0}^{\infty} c_n$ una serie de términos no negativos que converge. Una sucesión $(x_n)_{n\geq 0}$ en X tal que $d(x_{n+1}, x_n) \leq c_n$ para todo $n \in \mathbb{N}_0$ satisface la condición de Cauchy.*

Demostración. Sea $(x_n)_{n\geq 0}$ en X tal que $d(x_{n+1}, x_n) \leq c_n$ para todo $n \in \mathbb{N}_0$ y sea ε un número positivo. Como la serie $\sum_{n=0}^{\infty} c_n$ converge, la sucesión de sus restos $(\sum_{k=m}^{\infty} c_k)_{m\geq 0}$ converge a 0 y, en particular, satisface la condición de Cauchy: hay, por lo tanto, un entero no negativo n_0 tal que para toda elección de m y n vale

$$m, n \geq n_0 \implies \left| \sum_{k=m}^{\infty} c_k - \sum_{n=k}^{\infty} c_k \right| < \varepsilon.$$

Sean m y n dos elementos de $\mathbb{N}_0$ tales que $m, n \geq n_0$. Es

$$d(x_m, x_n) \leq \sum_{i=\min\{m,n\}}^{\max\{m,n\}-1} d(x_{i+1}, x_i) \leq \sum_{i=\min\{m,n\}}^{\max\{m,n\}-1} c_i = \left| \sum_{i=n}^{\infty} c_i - \sum_{i=m}^{\infty} c_i \right| < \varepsilon,$$

y esto nos dice que la sucesión $(x_n)_{n\geq 0}$ satisface la condición de Cauchy. □

El siguiente es un criterio de convergencia para las sucesiones de Cauchy.

Proposición 7.3.7. *Una sucesión en un espacio métrico que satisface la condición de Cauchy y posee una subsucesión convergente converge al mismo límite que esta.*

Demostración. Sea X un espacio métrico, sea $(x_n)_{n\geq 0}$ una sucesión en X que satisface la condición de Cauchy, y supongamos que $f : \mathbb{N}_0 \to \mathbb{N}_0$ es una función estrictamente creciente tal que la subsucesión $(x_{f(n)})_{n\geq 0}$ de X converge a un punto ξ de X. Sea ε un número positivo. La hipótesis sobre f implica que hay un entero no negativo n_1 tal que para todo $n \in \mathbb{N}_0$ vale $n \geq n_1 \implies d(x_{f(n)}, \xi) < \varepsilon/2$. Por otro lado, como x satisface la condición de Cauchy, hay un entero no negativo n_2 tal que para toda elección de m y n en $\mathbb{N}_0$ vale $m, n \geq n_2 \implies d(x_m, x_m) < \varepsilon/2$.

Sea $n_0 \coloneqq \max\{n_1, n_2\}$. Si n es un elemento de $\mathbb{N}_0$ tal que $n \geq n_0$, tenemos entonces que $n \geq n_1$ y $f(n) \geq n \geq n_0 \geq n_2$, así que

$$d(x_n, \xi) \leq d(x_n, x_{f(n)}) + d(x_{f(n)}, \xi) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Esto nos dice que la sucesión $(x_n)_{n\geq 0}$ converge a ξ. □

Otro criterio de convergencia útil es el que da el siguiente ejercicio.

Ejercicio 7.3.8. Sea X un espacio métrico. Si $x = (x_n)_{n\geq 0}$ e $y = (y_n)_{n\geq 0}$ son dos sucesiones en X que satisfacen la condición de Cauchy, la sucesión x converge a un punto ξ de X, y

$$\lim_{n\to\infty} d(x_n, y_n) = 0,$$

entonces la sucesión y también converge y al mismo punto.

El siguiente ejercicio es otro criterio de convergencia: nos dice que una sucesión de Cauchy converge si está «cerca» de una sucesión que converge.

Ejercicio 7.3.9. Sea X un espacio métrico y sean $x = (x_n)_{n\geq 0}$ e $y = (y_n)_{n\geq 0}$ dos sucesiones en X tales que la sucesión de números reales $(d(x_n, y_n))_{n\geq 0}$ converge a 0. Pruebe que la sucesión x satisface la condición de Cauchy si la sucesión y lo hace, y que si y además converge, entonces x también lo hace y al mismo límite.

La razón por la que estamos interesados en la condición de Cauchy es que se trata de una condición necesaria para la convergencia de una sucesión.

Proposición 7.3.10. *Una sucesión convergente satisface la condición de Cauchy.*

Demostración. Sea $x = (x_n)_{n\geq 0}$ una sucesión convergente en un espacio métrico X, sea ξ su límite, y sea ε un número positivo. Como x converge a ξ, hay un entero n_0 tal que para todo $n \in \mathbb{N}_0$ vale $n \geq n_0 \implies d(x_n, \xi) < \varepsilon/2$. Se sigue de esto que si m y n son dos elementos de $\mathbb{N}_0$ tales que $m, n \geq n_0$, entonces $d(x_m, x_n) \leq d(x_m, \xi) + d(\xi, x_n) < \varepsilon/2 + \varepsilon/2 = \varepsilon$. Vemos así que la sucesión x satisface la condición de Cauchy. □

Así, por ejemplo, usando la condición de Cauchy podemos volver a obtener el resultado del Ejemplo 7.1.12 pero con menos esfuerzo.

Ejemplo 7.3.11. Consideremos la sucesión $x = (x_n)_{n\geq 0}$ que tiene $x_n = (-1)^n$ para todo $n \in \mathbb{N}_0$. Cualquiera sea el elemento n_0 de $\mathbb{N}_0$ los enteros $m := n_0$ y $n := n_0 + 1$ son tales que $m, n \geq n_0$ y $d(x_m, x_n) = 2$, y esto implica que la sucesión x no satisface la condición de Cauchy y, por lo tanto, que no converge.

El siguiente resultado, que nos dice que en $\mathbb{R}$ la condición de Cauchy es suficiente para garantizar la convergencia de una sucesión, es uno de los teoremas fundamentales del *Análisis*.

Proposición 7.3.12. *Toda sucesión de Cauchy en $\mathbb{R}$ converge.*

Demostración. Sea $x = (x_n)_{n\geq 0}$ una sucesión de números reales que satisface la condición de Cauchy. Mostraremos que x converge en varios pasos.

Primer paso. Como la sucesión x es acotada, es claro que para cada $n \in \mathbb{N}_0$ el conjunto $C_n := \{x_m : m \in \mathbb{N}_0, m \geq n\}$ es acotado y que podemos considerar las sucesiones $\alpha = (\alpha_n)_{n\geq 0}$ y $\beta = (\beta_n)_{n\geq 0}$ que tienen $\alpha_n = \text{ínf } C_n$ y $\beta_n = \sup C_n$ para todo $n \in \mathbb{N}_0$. Es claro que siempre que m y n son elementos de $\mathbb{N}_0$ vale $m \leq n \implies C_n \subseteq C_m$, y esto implica que

$$m \leq n \implies \alpha_m \leq \alpha_n \leq \beta_n \leq \beta_m$$

cualesquiera sean m y n en $\mathbb{N}_0$. Mostremos que

$$\begin{aligned} &\textit{para todo número positivo } \varepsilon \textit{ hay un entero } n_0 \textit{ tal que para tod } n \in \mathbb{N}_0 \textit{ vale} \\ &n \geq n_0 \implies \beta_n - \alpha_n \leq \varepsilon. \end{aligned} \tag{7.3}$$

Sea ε un número positivo. Como la sucesión x es de Cauchy, hay un entero n_0 tal que para todo $m, n \in \mathbb{N}_0$ vale $m, n \geq n_0 \implies |x_m - x_n| < \varepsilon/2$. Si n es un elemento de $\mathbb{N}_0$ tal que $n \geq n_0$, tenemos entonces que para todo $m \in \mathbb{N}_0$ tal $m \geq n$ tenemos que $x_n - \varepsilon/2 < x_m < x_n + \varepsilon/2$, así que $x_n - \varepsilon/2$ y $x_n + \varepsilon/2$ son una cota inferior y una cota superior para el conjunto C_n y, por lo tanto $x_n - \varepsilon/2 \leq \alpha_n$ y $\beta_n \leq x_n + \varepsilon/2$ y, por lo tanto, $0 \leq \beta_n - \alpha_n \leq \varepsilon$.

Segundo paso. El conjunto $A := \{\alpha_n : n \in \mathbb{N}_0\}$ no es vacío y está acotado superiormente por β_0, mientras que el conjunto $B := \{\beta_n : n \in \mathbb{N}_0\}$ no es vacío y está acotado inferiormente por α_0: podemos entonces considerar los números $\xi := \sup A$ y $\zeta := \text{ínf } B$. Para todo $m \in \mathbb{N}_0$ es $\alpha_m \leq \beta_n$ cualquiera sea $n \in \mathbb{N}_0$, así que $\alpha_m \leq \text{ínf } B = \zeta$. Como esto es cierto cualquiera sea m en $\mathbb{N}_0$, implica que $\xi = \sup A \leq \zeta$. Queremos probar que, de hecho, es $\xi = \zeta$.

Sea ε un número positivo. Hay elementos n_1 y n_2 tales que $\xi - \varepsilon/3 < \alpha_{n_1}$ y $\beta_{n_2} < \zeta + \varepsilon/3$. Por otro lado, sabemos de (7.3) que hay un elemento n_0 de $\mathbb{N}_0$ tal que $\beta_n - \alpha_n \leq \varepsilon/3$ para todo entero n

tal que $n \geq n_0$. Si ponemos $m \coloneqq \max\{n_0, n_1, n_2\}$, tenemos entonces que $\xi - \varepsilon/3 < \alpha_{n_1} \leq \alpha_m$, $\beta_m \leq \beta_{n_2} < \zeta + \varepsilon/3$, y $\beta_m - \alpha_m \leq \varepsilon/3$: esto implica que

$$0 \leq \xi - \zeta \leq (\alpha_m + \varepsilon/3) - (\beta_m - \varepsilon/3) = (\alpha_m - \beta_m) + 2\varepsilon/3 \leq \varepsilon.$$

Esto es cierto cualquiera sea el número positivo ε, así que $\xi = \zeta$, como queremos.

Tercer y último paso. Probemos finalmente que la sucesión x converge a ξ. Sea ε un número positivo. Hay elementos n_1 y n_2 de $\mathbb{N}_0$ tales que que $\xi - \varepsilon < \alpha_{n_1}$ y $\beta_{n_2} < \zeta + \varepsilon$. Pongamos $n_0 \coloneqq \max\{n_1, n_2\}$. Como $C_{n_0} = \{x_m : m \in \mathbb{N}_0, m \geq n_0\}$ y

$$\xi - \varepsilon < \alpha_{n_1} \leq \alpha_{n_0} = \inf C_{n_0} \leq \sup C_{n_0} = \beta_{n_0} \leq \beta_{n_2} < \zeta + \varepsilon/2 = \xi + \varepsilon,$$

para todo elemento n de $\mathbb{N}_0$ tal que $n \geq n_0$ es $\xi - \varepsilon < x_n < \xi + \varepsilon$, esto es, $d(x_n, \xi) < \varepsilon$. Esto completa la demostración. □

En general, sin embargo, la condición de Cauchy no es suficiente para la convergencia.

Ejemplo 7.3.13. Sea X un espacio métrico y sea Y un subespacio de X que no es cerrado en X. Hay, por lo tanto, una sucesión x en X que toma valores en Y, que converge, y cuyo límite ξ no pertenece a Y. En particular, esa sucesión x satisface la condición de Cauchy. Ahora bien, podemos ver a x como una sucesión en el subespacio Y y es claro que en tanto sucesión en Y también satisface la condición de Cauchy. En Y, sin embargo, no converge: si tuviera un límite en tanto sucesión en Y, ese también sería su límite en tanto sucesión en X, así que debería ser ζ, y esto es imposible porque ζ no pertenece a Y.

Esto motiva la siguiente definición.

Definición 7.3.14. Decimos que un espacio métrico X es ***completo*** si toda sucesión en X que satisface la condición de Cauchy converge.

Hay dos casos particulares que reciben un nombres especiales.

Definición 7.3.15. Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$.

- Un ***espacio de Banach*** es un espacio vectorial normado $(V, \|-\|)$ que es completo con respecto a la métrica sobre V inducida por su norma.
- Un ***espacio de Hilbert*** es un espacio vectorial con producto interno $(V, \langle -, - \rangle)$ que es completo con respecto a la métrica sobre V inducida por su producto interno.

Los nombres recuerdan a David Hilbert y a Stefan Banach.

Veamos ejemplos de espacios completos, empezando por algunos triviales.

Ejemplo 7.3.16. Un espacio métrico con métrica discreta es completo. De hecho, un espacio métrico X tal que para todo elemento x de X es $r \coloneqq \inf\{d(x,y) : x, y \in X, x \neq y\} > 0$ es completo. En efecto, supongamos que X es un espacio que satisface esa condición y sea $x = (x_n)_{n\geq 0}$ una sucesión en X que satisface la condición de Cauchy. Hay, por lo tanto, un entero n_0 tal que para cada elección de m y n en $\mathbb{N}_0$ vale $m, n \geq n_0 \implies d(x_m, x_n) < r/2$. En particular, para todo $n \in \mathbb{N}_0$ tal que $n \geq n_0$ tenemos que $d(x_{n_0}, x_n) < r/2$, de manera que $x_{n_0} = x_n$. Vemos así que la sucesión x es a la larga constante y, por lo tanto, convergente.

Notemos que un espacio métrico en el que todos los puntos son aislados no es necesariamente completo. Así, por ejemplo, el subespacio $\{1/n : n \in \mathbb{N}\}$ de $\mathbb{R}$ tiene todos sus puntos aislados pero no es completo.

Nuestro segundo ejemplo, como vimos arriba, es mucho menos trivial.

Ejemplo 7.3.17. El espacio métrico $\mathbb{R}$ es completo. En efecto, esto es lo que nos dice la Proposición 7.3.12

Para extender este resultado a los espacios $\mathbb{R}^n$ con $n \in \mathbb{N}$ cualquiera usamos la siguiente observación.

Proposición 7.3.18. *Si X e Y son espacios métricos completos, entonces el producto cartesiano $X \times Y$ también es completo.*

Como es habitual, un argumento inductivo evidente extiende este resultado al caso de productos finitos arbitrarios de espacios métricos completos.

Demostración. Sean (X, d) e (Y, d') dos espacios métricos completos y sea $z = (z_n)_{n\geq 0}$ una sucesión de Cauchy en $X \times Y$. Para cada $n \in \mathbb{N}_0$ sean x_n e y_n los elementos de X y de Y tales que $z_n = (x_n, y_n)$. Sea ε un número positivo. Que z es una sucesión de Cauchy implica que hay un entero n_0 tal que para toda elección de m y n en $\mathbb{N}_0$ vale $m, n \geq n_0 \implies d(z_m, z_n) < \varepsilon$. Si n es un elemento de $\mathbb{N}_0$, entonces tenemos que

$$d(x_m, x_n) \leq d(x_m, x_n) + d'(y_m, y_n) = d(z_m, z_n) < \varepsilon$$

y

$$d'(y_m, y_n) \leq d(x_m, x_n) + d'(y_m, y_n) = d(z_m, z_n) < \varepsilon.$$

Esto prueba que las sucesiones $(x_n)_{n\geq 0}$ y $(y_n)_{n\geq 0}$ en X y en Y son de Cauchy así que, de acuerdo a la hipótesis, ambas convergen. Sean ξ y ζ sus límites.

Para terminar, mostremos que la sucesión z converge al punto (ξ, ζ) de $X \times Y$. Sea ε un número positivo. Como las sucesiones $(x_n)_{n\geq 0}$ y $(y_n)_{n\geq 0}$ convergen en X y en Y a ξ y a ζ, hay enteros n_1 y n_2 tales que para todo $n \in \mathbb{N}_0$ vale $n \geq n_1 \implies d(x_n, \xi) < \varepsilon/2$ y $n \geq n_2 \implies d'(y_n, \zeta) < \varepsilon/2$.

Pongamos $n_0 \coloneqq \max\{n_1, n_2\}$. Si n es un elemento de $\mathbb{N}_0$ tal que $n \geq n_0$, entonces $n \geq n_1$ y $n \geq n_2$, así que $d(x_n, \xi) < \varepsilon/2$, $d'(y_n, \zeta) < \varepsilon/2$ y, por lo tanto,

$$d(z_n, (\xi, \zeta)) = d(x_n, \xi) + d'(y_n, \zeta) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Esto completa la demostración. □

Una aplicación inmediata de esta proposición es la siguiente.

Ejercicio 7.3.19. Use la Proposición 7.3.18 para probar que el conjunto $\mathbb{C}$ de los números complejos dotado de su métrica usual es completo.

Otro corolario significativo de la proposición es el siguiente:

Corolario 7.3.20. *Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$. Para todo $n \in \mathbb{N}_0$ el espacio $\Bbbk^n$ dotado de su norma $\|-\|_1$ es completo, así que $(\Bbbk^n, \|-\|_1)$ es un espacio de Banach., así que $(\Bbbk^n, \|-\|_1)$ es un espacio de Banach.*

Demostración. En efecto, $\Bbbk^n$ dotado de su norma $\|-\|_1$ es el producto cartesiano de n copias del espacio $\Bbbk$ dotado de su norma usual, y sabemos de la Proposición 7.3.12 o del Ejercicio 7.3.19 que este último es completo. □

Para ver que los espacios $\mathbb{R}^n$ y $\mathbb{C}^n$ son completos con respecto a las métricas inducidas por otras normas usamos la siguiente observación.

Proposición 7.3.21. *Sea X un conjunto y sean d y d' dos métricas sobre X. Si d y d' son uniformemente equivalentes y el espacio métrico (X, d) es completo, entonces el espacio métrico (X, d) es completo.*

Demostración. Supongamos que las métricas d y d' son uniformemente equivalentes, de manera que hay un número positivo c tal que para todo elemento (x, y) de $X \times X$ es

$$\frac{1}{c} \cdot d(x, y) \leq d'(x, y) \leq c \cdot d(x, y),$$

supongamos que el espacio métrico (X, d) es completo, y sea $x = (x_n)_{n \geq 0}$ una sucesión de Cauchy en el espacio métrico (X, d').

Sea ε un número positivo. Hay un entero n_0 tal que siempre que m y n son elementos de $\mathbb{N}_0$ vale $m, n \geq n_0 \implies d'(x_m, x_n) < \varepsilon/c$, y se sigue de esto que si m y n son elementos de $\mathbb{N}_0$ tales que $m, n \geq n_0$, entonces $d(x_m, x_n) \leq c \cdot d'(x_m, x_n) < \varepsilon$. Vemos así que la sucesión x es de Cauchy con respecto a la métrica d y, como el espacio métrico (X, d) es completo, que converge con respecto a d. Como las métricas d y d' son topológicamente equivalentes, podemos concluir que la sucesión x converge también con respecto a la métrica d', y esto prueba que el espacio métrico (X, d') es completo. □

El siguiente es un ejemplo que muestra que en la Proposición 7.3.21 no podemos reemplazar la hipótesis de que las métricas sean uniformemente equivalentes por la de que sean simplemente topológicamente equivalentes.

Ejemplo 7.3.22. Consideremos la función $f : x \in \mathbb{R} \mapsto \arctan x \in \mathbb{R}$. Se trata de una función diferenciable y es $f'(x) = (1+x^2)^{-1} > 0$ para todo $x \in \mathbb{R}$, así que f es estrictamente creciente y, en particular, inyectiva. Sea

$$d' : (x, y) \in \mathbb{R} \times \mathbb{R} \mapsto |f(x) - f(y)| \in \mathbb{R}.$$

Mostremos d' es una métrica sobre $\mathbb{R}$.

- Es claro que $d(x, x) = 0$ cualquiera sea el elemento x de $\mathbb{R}$. Si, por otro lado, x e y son dos elementos de $\mathbb{R}$ tales que $d(x, y) = 0$, esto es, tales que $|f(x) - f(y)| = 0$, entonces $f(x) = f(y)$ y, como la función f es inyectiva, es $x = y$.
- Si x e y son dos elementos de $\mathbb{R}$, entonces $d(x, y) = |f(x)-f(y)| = |f(y)-f(x)| = d(y, x)$.
- Finalmente, si x, y, y z son tres elementos de $\mathbb{R}$, entonces

$$d(x, y) = |f(x) - f(y)| \leq |f(x) - f(z)| + |f(z) - f(y)| = d(x, z) + d(z, y).$$

En segundo lugar, probemos que la métrica d' es topológicamente equivalente a la métrica euclídea d de $\mathbb{R}$ usando el criterio que nos da la Proposición 6.4.3.2. Sea x un elemento de $\mathbb{R}$ y sea ε un número positivo.

- Como la función f es continua, hay un número positivo δ tal que para todo número real y vale $|y - x| < \delta \implies |f(y) - f(x)| < \varepsilon$, y esto nos dice precisamente que $B_\delta(x) \subseteq B'_\varepsilon(x)$.
- La función $t \in (-\pi/2, \pi/2) \mapsto \tan t \in \mathbb{R}$ es continua y $f(x) \in (-\pi/2, \pi/2)$, así que hay un número positivo η tal que para todo u de $(-\pi/2, \pi/2)$ vale $|u-f(x)| < \eta \implies |\tan u-x| < \varepsilon$. En particular, si y es un número tal que $|f(y) - f(x)| < \eta$ tenemos que $|y - x| < \varepsilon$: esto nos dice que $B'_\eta(x) \subseteq B_\varepsilon(x)$.

Tenemos en este momento dos métricas d y d' sobre $\mathbb{R}$ que son topológicamente equivalente y sabemos que el espacio métrico $(\mathbb{R}, d)$ es completo. Mostremos que $(\mathbb{R}, d')$ no lo es exhibiendo una sucesión en $\mathbb{R}$ que satisface la condición de Cauchy con respecto a d' pero que no converge.

Sea x la sucesión $(n)_{n\geq 0}$. Es claro que x no converge con respecto a la métrica d, ya que no es acotada, y entonces, como d y d' son topológicamente equivalente, vemos que x no converge con respecto a la métrica d'. Probemos finalmente que x satisface la condición de Cauchy con respecto a d'. Sea ε un número positivo. Como

$$\lim_{n\to\infty} \arctan n = \frac{\pi}{2},$$

hay un entero n_0 tal que $|\arctan n - \pi/2| < \varepsilon/2$ siempre que n es un elemento de $\mathbb{N}_0$ tal que $n \geq n_0$, y esto implica que si m y n son dos enteros tales que $m, n \geq n_0$, entonces

$$d'(m, n) = |\arctan m - \arctan n| = |(\arctan m - \pi/2) - (\arctan n - \pi/2)|$$

$$\leq |\arctan m - \pi/2| + |\arctan n - \pi/2| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

El siguiente es el corolario de la proposición que queríamos.

Corolario 7.3.23. *Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, sea $n \in \mathbb{N}_0$, y sea $p \in [1, +\infty]$. El espacio $\mathbb{k}^n$ dotado de su norma $\|-\|_p$ es completo, de manera que $(\mathbb{k}^n, \|-\|_p)$ es un espacio de Banach.*

Demostración. En efecto, sabemos de la Proposición 6.4.4.8 que las normas $\|-\|_1$ y $\|-\|_p$ son equivalentes, así que inducen sobre $\mathbb{k}^n$ distancias uniformemente equivalentes, y del Corolario 7.3.20 que $\mathbb{k}^n$ es completo con respecto a la que induce la primera, así que la Proposición 7.3.21 nos permite concluir que también lo es con respecto a la que induce la segunda. □

Mostraremos más adelante, en la Proposición 7.5.1, que *todas* las normas sobre $\mathbb{R}^n$ son equivalentes, y entonces sabremos que $\mathbb{R}^n$ es completo con respecto a cualquiera de sus normas.

Ejercicio 7.3.24. Pruebe directamente que $\mathbb{R}^n$ es completo con respecto a la métrica inducida por la norma $\|-\|_p$, cualquiera sea $p \in [1, +\infty]$.

Ejercicio 7.3.25. Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, sea p un elemento de $[1, +\infty)$, sea V un espacio vectorial sobre $\mathbb{k}$ de dimensión finita, sea n la dimensión de V, y sea $\mathscr{B} = (e_1, e_2, \ldots, e_n)$ una base ordenada de V. Sabemos del Ejercicio 5.2.3.14 que hay una función $\|-\| : V \to \mathbb{R}$ tal que para cada elección de n escalares $a_1, a_2, \ldots, a_n$ de $\mathbb{k}$ vale

$$\|a_1e_1 + a_2e_2 + \cdots + a_ne_n\| = \left(\sum_{i=1}^{n} |a_i|^p\right)^{1/p}$$

y que es una norma sobre V. Pruebe que V dotado de esa norma es un espacio de Banach.

La siguiente observación es una fuente de ejemplos de espacios completos y una explicación general de las observaciones que hicimos en el Ejemplo 7.3.11.

Proposición 7.3.26.
(*i*) *Un subespacio cerrado de un espacio métrico completo es completo.*
(*ii*) *Un subespacio completo de un espacio métrico es cerrado.*

En particular, en un espacio métrico completo los subconjuntos completos son precisamente los subconjuntos cerrados.

Demostración. Sea primero Y un subespacio cerrado de un espacio métrico completo X, y sea x una sucesión de Cauchy en Y. Como x es también una sucesión de Cauchy en X y X es completo, sabemos que x converge en X. Sea ξ su límite allí. Como Y es un subconjunto cerrado de X y la sucesión x toma valores en Y, sabemos que el límite ξ de esa sucesión está en Y: es claro, entonces, que la sucesión x converge a ξ en tanto sucesión en Y. Esto prueba que el espacio métrico Y es completo y, por lo tanto, la primera afirmación de la proposición.

Sea ahora Y un subespacio completo de un espacio métrico cualquiera X, y sea x una sucesión en Y que converge en X a un punto ξ de X. Esto implica que x satisface la condición de Cauchy en tanto sucesión en Y y, como Y es completo, que x converge en Y a un punto ξ' de Y. Por supuesto, x también converge a ξ' en X y, como tiene un único límite, podemos concluir que $\xi = \xi' \in Y$. Vemos así que el conjunto Y es cerrado en X. Esto prueba la segunda afirmación de la proposición. □

En *Análisis*, muchas veces usamos la Proposición 7.3.12 indirectamente, vía el siguiente resultado conocido como el *Principio de los intervalos encajados*.

Proposición 7.3.27. *Sea $(I_n)_{n\geq}$ una sucesión de intervalos cerrados de $\mathbb{R}$, par cada $n \in \mathbb{N}_0$ sean a_n y b_n los elementos de $\mathbb{R}$ tales que $I_n = [a_n, b_n]$, y supongamos que las siguientes condiciones se cumplen:*

- *Para todo $n \in \mathbb{N}_0$ es $I_{n+1} \subseteq I_n$.*
- *La sucesión $(b_n - a_n)_{n\geq}$ converge a 0.*

La intersección $\bigcap_{n\in\mathbb{N}_0} I_n$ no es vacía y tiene exactamente un elemento.

Demostración. La primera condición implica claramente que $a_n \leq a_{n+1} \leq b_{n+1} \leq b_n$ para todo $n \in \mathscr{N}_0$. El conjunto $A \coloneqq \{a_n : n \in \mathbb{N}_0\}$ esta acotado superiormente por b_0 y el conjunto $B \coloneqq \{b_n : n \in \mathbb{N}_0\}$ está acotado inferiormente por a_0, y ambos son no vacíos, así que podemos considerar los números $\alpha \coloneqq \sup A$ y $\beta \coloneqq \text{ínf } B$. Para todo $n \in \mathbb{N}_0$ es claro que b_n es una cota superior para A, de hecho, y entonces tenemos que $\alpha \leq b_n$. Esto nos dice, por su parte, que α es una cota inferior para X, así que $\alpha \leq \beta$. El número $\beta - \alpha$ es entonces no negativo.

Sea ε un número positivo. Como la sucesión $(b_n - a_n)_{n\geq 0}$ converge a 0, hay un entero no negativo m tal que $b_m - a_m < \varepsilon$. Como $a_m \in A$ y $b_m \in B$, es $a_m \leq \alpha$ y $\beta \leq b_m$ y, por lo tanto, $0 \leq \beta - \alpha \leq b_m - a_m < \varepsilon$. Vemos así que el número no negativo $\beta - \alpha$ es menor que todo número positivo y, en consecuencia, que es nulo.

Como $a_n \leq \alpha \leq b_n$ para todo $n \in \mathbb{N}_0$, ya que $a_n \in A$, $\alpha = \sup A$, $b_n \in B$ y $\beta = \text{ínf } B$, tenemos que $\alpha \in [a_n, b_n] = I_n$. Esto nos dice que α pertenece a la intersección $\bigcap_{n\in\mathbb{N}_0} I_n$, así que esta no es vacía. Por otro lado, supongamos que x es un elemento de esa intersección. Tenemos entonces que para todo $n \in \mathbb{N}_0$ es $x \in I_n[a_n, b_n]$, es decir, que $a_n \leq x \leq b_b$: se sigue de esto que x es una cota superior para A y una cota interior para B y, por lo tanto, que $\alpha \leq x \leq b = \alpha$, de la manera que

$x = \alpha$. Esto prueba que α es el único elemento de la intersección $\bigcap_{n\in\mathbb{N}_0} I_n$, y completa la prueba de la proposición. □

Hemos probado esta proposición desde cero — empezando simplemente las propiedades básicas de $\mathbb{R}$ — pero a esta altura podemos hacerlo más fácilmente usando el hecho de que $\mathbb{R}$ es un espacio métrico completo.

Ejercicio 7.3.28. Pruebe la Proposición 7.3.27 usando la Proposición 7.3.12.

Que esto es posible sugiere que podemos generalizar la Proposición 7.3.27 a espacios métricos completos arbitrarios. De hecho, con un poco más de trabajo podemos obtener una caracterización de esos espacios:

Proposición 7.3.29. *Un espacio métrico X es completo si y solamente si todo conjunto no vacío $\mathscr{F}$ de cerrados de X tal que*

- *la intersección de todo subconjunto finito y no vacío de $\mathscr{F}$ no es vacía, y*
- *para todo número positivo r hay un elemento F en $\mathscr{F}$ tal que* $\operatorname{diam} F < r$

tiene intersección no vacía de cardinal 1.

Demostración. Supongamos primero que el espacio métrico es completo y sea $\mathscr{F}$ un conjunto no vacío de cerrados de X que satisface las dos condiciones del enunciado. Probemos primero que

$$\begin{aligned}&\text{si } (F_n)_{n\geq 0} \text{ es una sucesión de cerrados de } X \text{ tal que } \operatorname{diam} F_n < 1/(n+1) \text{ y}\\ &\textstyle\bigcap_{i=0}^{n} F_n \neq \varnothing \text{ para todo } n \in \mathbb{N}_0, \text{ entonces hay un elemento } \xi \text{ de } X \text{ tal que}\\ &\textstyle\bigcap_{n=0}^{\infty} F_n = \{\xi\}.\end{aligned} \tag{7.4}$$

Sea para ello $(F_n)_{n\geq 0}$ una sucesión de cerrados de X tal que $\operatorname{diam} F_n < 1/(n+1)$ y $G_n := \bigcap_{i=0}^{n} F_n \neq \varnothing$ para todo $n \in \mathbb{N}_0$. Para cada $n \in \mathbb{N}_0$ elijamos un elemento x_n de G_n. Notemos que si m y n son dos elementos de $\mathbb{N}_0$ tales que $m \geq n$, entonces $x_n \in G_n = \bigcap_{i=0}^{n} F_n \subseteq \bigcap_{i=0}^{m} F_i = G_m$. Esto nos dice que

para todo $m \in \mathbb{N}_0$ la sucesión $(x_n)_{n\geq 0}$ está a la larga en G_m.

Sea ε un número positivo. Hay un entero no negativo n_0 tal que $1/(n_0+1) < \varepsilon/2$. Si m y n son dos elementos de $\mathbb{N}_0$ tales que $m, n \geq n_0$, entonces vimos que $x_m, x_n \in G_{n_0} \subseteq F_{n_0}$ y, como $\operatorname{diam} F_{n_0} < 1/(n_0+1) < \varepsilon$, tenemos que $d(x_m, x_n) < \varepsilon$. Esto nos dice que la sucesión $(x_n)_{n\geq 0}$ satisface la condición de Cauchy y, como X es completo, que converge a un punto ξ de X. Para todo $m \in \mathbb{N}_0$ la sucesión $(x_n)_{n\geq 0}$ está a la larga en G_m y G_m es cerrado, así que el límite ξ de esa sucesión está en G_m, se sigue de esto que

$$\xi \in \bigcap_{m=0}^{\infty} G_m = \bigcap_{m=0}^{\infty} \bigcap_{i=0}^{m} F_m = \bigcap_{m=0}^{\infty} F_m.$$

Si ζ es un elemento de $\bigcap_{m=0}^{\infty} F_m$, entonces para todo $m \in \mathbb{N}_0$ tenemos que $\xi, \zeta \in F_m$, así que $d(\xi, \zeta) \leq \operatorname{diam} F_m < 1/(m+1)$, y esto implica que $d(\xi, \zeta) = 0$, esto es, que $\zeta = \xi$. Así, ξ es el único elemento de la intersección $\bigcap_{m\in\mathbb{N}_0} F_m$ y tenemos que

$$\bigcap_{m=0}^{\infty} F_m = \{\xi\}.$$

Esto prueba la afirmación 7.4.

Ahora bien, para cada $n \in \mathbb{N}_0$ hay un elemento F_n de $\mathscr{F}$ tal que $\operatorname{diam} F_n < 1/(n+1)$ y es cerrado, ya que todos los elementos de $\mathscr{F}$ lo son. Como para todo $n \in \mathbb{N}_0$ la intersección $\bigcap_{i=0}^{n} F_n$ no es vacía por la segunda hipótesis que hicimos sobre $\mathscr{F}$, la afirmación que acabamos de probar nos dice que hay un punto ξ de X tal que

$$\bigcap_{F\in\mathscr{F}} F \subseteq \bigcap_{m=0}^{\infty} F_m = \{\xi\}. \tag{7.5}$$

Supongamos por un momento que esta inclusión es estricta, de manera que hay un elemento H de $\mathscr{F}$ tal que $\xi \notin H$. La sucesión $(H \cap F_n)_{n\geq 0}$ es una sucesión de cerrados de X que tiene $\bigcap_{i=0}^{n}(H \cap F_n) = H \cap \bigcap_{i=0}^{n} F_n \neq \varnothing$, ya qu e$\{H\} \cup \{F_i : i \in [\![0, n]\!]\}$ es un subconjunto no vacío de $\mathscr{F}$. La afirmación (7.4) implica entonces que hay un elemento ξ' de X tal que

$$\varnothing = H \cap \{\xi\} = H \cap \bigcap_{m=0}^{\infty} F_m = \bigcap_{m=0}^{\infty} (H \cap F_m) = \{\xi'\}.$$

Esto es, por supuesto, absurdo. La contradicción provino de haber supuesto que la inclusión de (7.5) es estricta, así que debe ser $\bigcap_{F\in\mathscr{F}} F = \{\xi\}$. Todo esto prueba que la condición que da la proposición para que el espacio X sea completo es necesaria.

Probemos ahora que esa condición es suficiente. Supongamos que se cumple, y sea $(x_n)_{n\geq 0}$ una sucesión de Cauchy en X. Para cada $n \in \mathbb{N}_0$ sea F_n la clausura del conjunto $\{x_m : m \in \mathbb{N}_0, m \geq n\}$ y pongamos $\mathscr{F} \coloneqq \{F_n : n \in \mathbb{N}_0\}$.

Si I es un subconjunto finito y no vacío de $\mathbb{N}_0$ y μ es el máximo elemento de I, entonces

$$\bigcap_{i\in I} F_i = \bigcap_{i\in I} \overline{\{x_m : m \in \mathbb{N}_0, m \geq i\}} = \overline{\bigcap_{i\in I} \{x_m : m \in \mathbb{N}_0, m \geq i\}} = \overline{\{x_m : m \in \mathbb{N}_0, m \geq \mu\}} \neq \varnothing.$$

Por otro lado, si r es un número positivo, entonces que la sucesión $(x_n)_{n\geq 0}$ satisfaga la condición de Cauchy nos dice que hay un entero no negativo n_0 tal que para todo $m, n \in \mathbb{N}_0$ vale $m, n \geq n_0 \implies d(x_n, x_m) < r/2$, y esto implica inmediatamente que $\dim F_{n_0} \leq r/2 < r$. Vemos asi que el conjunto $\mathscr{F}$ satisface las dos condiciones que aparecen en la proposición, así que la hipótesis implica que hay un punto ξ de X tal que $\bigcap_{n=0}^{\infty} F_n = \{\xi\}$.

Sea ε un número positivo. Como vimos arriba, hay un entero no negativo tal que $\operatorname{diam} F_n < \varepsilon$ y, como ξ y x_n pertenecen a F_n, tenemos que que $d(\xi, x_n) \leq \operatorname{diam} F_n < \varepsilon$, de manera que $x_n \in B_\varepsilon(\xi)$. Vemos así que la sucesión $(x_n)_{n\geq 0}$ pasa por todos los entornos de ξ, así que posee una subsucesión que converge a ξ: como es de Cauchy, la Proposición 7.3.7 nos permite concluir que, de hecho, la sucesión $(x_n)_{n\geq 0}$ converge a ξ. Esto prueba que el espacio X es completo. □

Terminemos esta sección con algunos ejemplos importantes de espacios métricos completos.

Proposición 7.3.30. *Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, y sea p un elemento de $[1, +\infty)$. El espacio normado $\ell^p_{\mathbb{k}}$ es un espacio de Banach.*

Demostración. Sea $x = (x_n)_{n\geq 0}$ una sucesión de Cauchy en $\ell^p_{\mathbb{k}}$, de manera que para cada $n \in \mathbb{N}_0$ la n-ésima componente es un elemento $x_n = (x_{n,i})_{i\geq 0}$ de $\ell^p_{\mathbb{k}}$. Como la sucesión x satisface la condición de Cauchy, es acotada y hay, por lo tanto, un número no negativo K tal que $\|x_n\|_p \leq K$ para todo $n \in \mathbb{N}_0$.

Sea k un elemento de $\mathbb{N}_0$. Como la sucesión x es de Cauchy, hay un entero no negativo n_0 tal que siempre que m y n son elementos de $\mathbb{N}_0$ vale $m, n \geq n_0 \implies d(x_m, x_n) < \varepsilon$. Esto implica que siempre m y n son elementos de $\mathbb{N}_0$ tales que $m, n \geq 0$ se tiene que

$$|x_{m,k} - x_{n,k}| \leq \left(\sum_{i\geq 0} |x_{m,i} - x_{n,i}|^p\right)^{1/p} = \|x_m - x_n\| = d(x_m, x_n) < \varepsilon.$$

Vemos así que la sucesión $(x_{n,k})_{n\geq 0}$ en $\mathbb{k}$ es de Cauchy y, como $\mathbb{k}$ es un espacio métrico completo, que converge. Escribamos ξ_k a su límite, y consideremos la sucesión $\xi := (\xi_i)_{i\geq 0}$, que es un elemento de $\mathbb{k}^{\mathbb{N}_0}$. Vamos a mostrar que ξ pertenece, de hecho, a $\ell^p_{\mathbb{k}}$ y que la sucesión x converge a ξ en $\ell^p_{\mathbb{k}}$. Esto probará la proposición.

- Sea N un entero no negativo, y para cada elemento $z = (x_n)_{n\geq 1}$ de $\mathbb{k}^{\mathbb{N}_0}$ escribamos z^N al elemento $(z_0, z_2, \ldots, z_N)$ de $\mathbb{k}^{N+1}$. Notemos que para cada $n \in \mathbb{N}_0$ es

$$\|x_n^N\|_p = \left(\sum_{i=0}^{N} |x_{n,i}|^p\right)^{1/p} \leq \left(\sum_{i=0}^{\infty} |x_{n,i}|^p\right)^{1/p} = \|x_n\|_p \leq K. \tag{7.6}$$

  Como para cada $i \in [\![0, N]\!]$ la sucesión $(x_{n,i})_{n\geq 0}$ converge a ξ_i en $\mathbb{k}$, el resultado del Ejercicio 7.1.14 nos dice que la sucesión $(x_n^N)_{n\geq 0}$ converge en $\mathbb{k}^{N+1}$ al punto ξ^N con respecto a la métrica de que se obtiene viendo a este espacio como el producto cartesiano $\mathbb{k} \times \mathbb{k} \times \cdots \times \mathbb{k}$. Esta métrica producto es la métrica d_1 de $\mathbb{k}^{N+1}$, y esta es equivalente a la métrica d_p, así que esto nos dice que la sucesión $(x_n^N)_{n\geq 0}$ converge a ξ^N en el espacio normado $(\mathbb{k}^{N+1}, \|-\|_p)$. De acuerdo a (7.6), tenemos que $\|x_n^N\|_p \leq K$ para todo $n \in \mathbb{N}_0$, así que el Ejercicio 7.1.27 nos permite concluir que

$$\sum_{i=0}^{N} |\xi_i|^p = \|\xi^N\|_p^p \leq K^p.$$

  Esto es cierto cualquiera sea el entero no negativo N, así que la serie de términos no negativos $\sum_{i=0}^{\infty} |\xi_i|^p$ tiene sus sumas parciales acotadas y, por lo tanto, converge. Esto prueba que ξ es un elemento de $\ell^p_{\mathbb{k}}$.

- Sea ε un número positivo.
  - Como la sucesión $(x_n)_{n\geq 0}$ es de Cauchy en $\ell^p_{\Bbbk}$, hay un entero no negativo n_1 tal que

    $$\|x_m - x_n\|_p < \varepsilon \text{ si } m, n \geq n_1.$$

  - Como x_{n_1} y ξ están en $\ell^p_{\Bbbk}$, las series $\sum_{i=0}^{\infty}|x_{n_1,i}|^p$ y $\sum_{i=0}^{\infty}|\xi_i|^p$ convergen y, por lo tanto, hay un entero no negativo N tal que $\sum_{i=N+1}^{\infty}|x_{n_1,i}|^p < \varepsilon^p$ y $\sum_{i=N+1}^{\infty}|\xi_i|^p < \varepsilon^p$, y esto implica que

    $$\|x_{n_1} - x^N_{n_1}\|_p = \left(\sum_{i=N+1}^{\infty}|x_{n_1,i}|^p\right)^{1/p} < \varepsilon, \qquad \|\xi - \xi^N\|_p = \left(\sum_{i=N+1}^{\infty}|\xi_i|^p\right)^{1/p} < \varepsilon.$$

  - Vimos arriba que la sucesión $(x^N_n)_{n\geq 0}$ converge a ξ^N en $(\Bbbk^{N+1}, \|-\|_p)$, así que hay un entero no negativo n_2 tal que

    $$\|x^N_n - \xi^N\|_p < \varepsilon \text{ si } n \geq n_2.$$

  Sea ahora $n_0 \coloneqq \max\{n_1, n_2\}$. Si n es un elemento de $\mathbb{N}_0$ tal que $n \geq n_0$, entonces $n \geq n_1$, así que $\|x^N_n - x^N_{n_1}\|_p \leq \|x_n - x_{n_1}\|_p < \varepsilon$, y $n \geq n_2$, así que $\|x^N_n - \xi^N\|_p < \varepsilon$, y esto implica que

  $$\begin{aligned}\|x_n - \xi\|_p &\\ &\leq \|x_n - x_{n_1}\|_p + \|x_{n_1} - x^N_{n_1}\|_p + \|x^N_{n_1} - x^N_n\|_p + \|x^N_n - \xi^N\|_p + \|\xi^N - \xi\|_p \qquad (7.7)\\ &< \varepsilon + \varepsilon + \varepsilon + \varepsilon + \varepsilon = 5\varepsilon.\end{aligned}$$

  Esto prueba que la sucesión $(x_n)_{n\geq 0}$ converge a ξ en ℓ^p.

Aquí estamos usando la notación $\|-\|_p$ tanto para la norma de $\ell^p_{\Bbbk}$ como para la norma de $\Bbbk^{N+1}$ y, en particular, en la desigualdad (7.7) estamos usando la desigualdad triangular para esas dos normas y, además, el hecho de que la norma de un elemento $(\alpha_0, \alpha_1, \ldots, \alpha_N)$ de $\Bbbk^{N+1}$ coincide con la norma del elemento $(\alpha_0, \alpha_1, \ldots, \alpha_N, 0, 0, \ldots)$ de $\ell^p_{\Bbbk}$. Dejamos al lector la tarea de verificar con cuidado que todo lo que escribimos tiene sentido y, más aún, es cierto. □

Ejercicio 7.3.31. Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, y sea X un conjunto. Pruebe que el espacio normado $\ell^\infty_{\Bbbk}(X)$ del Ejercicio 5.2.3.23 es un espacio de Banach. En particular, tomando $X = \mathbb{N}_0$, claro, esto nos dice que el espacio $\ell^\infty_{\Bbbk}$ es de Banach.

Es en general no muy sencillo probar que un espacio normado es completo. Muchas veces hacemos eso usando el siguiente criterio.

Proposición 7.3.32. *Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$. Un $\Bbbk$-espacio normado X es completo si y solamente si para toda sucesión $(x_n)_{n\geq 0}$ en X tal que la serie $\sum_{n=0}^{\infty}\|x_n\|$ converge la serie $\sum_{n=0}^{\infty} x_n$ converge.*

En otras palabras, un espacio normado es completo exactamente cuando toda serie absolutamente convergente converge.

Demostración. Sea X un espacio normado. Supongamos primero que X es completo, sea $(x_n)_{n\geq 0}$ una sucesión en X tal que la serie $\sum_{n=0}^{\infty}\|x_n\|$ converge en $\mathbb{R}$, y para cada elemento n de $\mathbb{N}_0$ sea $y_n \coloneqq \sum_{i=0}^{n} x_n$. Como para cada $n \in \mathbb{N}_0$ tenemos que $\|y_{n+1} - y_n\| = \|x_{n+1}\|$ y la serie $\sum_{n=0}^{\infty}\|x_{n+1}\|$ converge, la Proposición 7.3.6 nos dice que la sucesión $(y_n)_{n\geq 0}$ satisface la condición de Cauchy y, como el espacio X es completo, converge. Esto significa precisamente que la serie $\sum_{n=0}^{\infty} x_n$ converge, y vemos así que la condición de la proposición es necesaria.

Veamos ahora que esa condición también es suficiente. Sea $(x_n)_{n\geq 0}$ una sucesión en X que satisface la condición de Cauchy, de manera que, en particular, para cada $k \in \mathbb{N}_0$ el conjunto

$$R(k) \coloneqq \{l \in \mathbb{N}_0 : l > k \wedge (u, v \geq l \implies \|x_u - x_v\| < 2^{-k})\}$$

no es vacío. Hay, por lo tanto, una función $f : \mathbb{N}_0 \to \mathbb{N}_0$ tal que $f(0) = 0$ y $f(n+1) \in R(n)$ para cada $n \in \mathbb{N}_0$. Esto implica, en primer lugar, que la función f es estrictamente creciente y, en segundo lugar, que siempre que n es un elemento de $\mathbb{N}_0$, entonces $f(n) < f(n+1) < f(n+2)$ y, como $f(n+1) \in R(f(n))$,

$$\|x_{f(n+2)} - x_{f(n+1)}\| \leq \frac{1}{2^{f(n)}} \leq \frac{1}{2^n}.$$

Esto, el hecho de que la serie $\sum_{n=0}^{\infty} 1/2^n$ converge, y la hipótesis nos permiten concluir que la serie

$$\sum_{n=0}^{\infty}(x_{f(n+2)} - x_{f(n+1)})$$

converge en X. Para cada $N \in \mathbb{N}_0$ es $\sum_{n=0}^{N}(x_{f(n+2)} - x_{f(n+1)}) = x_{f(N+1)} - x_{f(1)}$, así que la convergencia de esa serie signifiva que la sucesión $(x_{f(n+1)} - x_{f(1)})_{n\geq 0}$ converge en X y, por lo tanto, que también lo hace la subsucesión $(x_{f(n+1)})_{n\geq 0}$ de $(x_n)_{n\geq 0}$. Como esta última sucesión satisface la condición de Cauchy y posee una subsucesión convergente, es ella misma convergente. Esto prueba que el espacio normado X es completo. □

Veamos una aplicación sencilla de este criterio.

Proposición 7.3.33. *Sea X un espacio de Banach y sea Y un subespacio cerrado de X. Hay una función $\|-\| : X/Y \to \mathbb{R}$ tal que*

$$\|x + Y\|_{X/Y} = \inf\{\|x + y\| : y \in Y\}$$

para toda elección de x en X, es una norma sobre X/Y, y hace de X/Y un espacio de Banach.

HACER: Qué pasa si Y no es cerrado?

Demostración. La Proposición 5.2.4.5 nos dice que la función

$$\|-\|' : x \in X \mapsto \inf\{\|x+y\| : y \in Y\} \in \mathbb{R}$$

es una semi-norma sobre X. Si x es un elemento de X tal que $\|x\|' = 0$, entonces hay una sucesión $(y_n)_{n\geq 0}$ de elementos de Y tal que $\|x + y_n\| \leq 1/(n+1)$ para todo $n \in \mathbb{N}_0$, de manera que, en particular, la sucesión $(y_n)_{n\geq 0}$ converge a x y, como toma valores en el subespacio Y y este es cerrado en X, es $x \in Y$. Es claro, por otro lado, que $\|y\|' = 0$ para todo elemento y de Y, así que lo que tenemos es que

$$Y = \{x \in X : \|x\|' = 0\}.$$

La Proposición 5.2.4.6, entonces, nos dice que la función

$$\|-\|_{X/Y} : \xi \in X/Y \mapsto \inf\{\|x\| : x \in \xi\} \in \mathbb{R}$$

es una norma sobre el espacio cociente X/Y. Esto implica, claro, que para cada elemento x de X es $\|x + X\|_{X/Y} = \inf\{\|x+y\| : y \in Y\}$. Para probar la proposición tenemos que mostrar que cuando dotamos al espacio X/Y de esta norma $\|-\|_{X/Y}$ es completo, y lo haremos usando el criterio de la Proposición 7.3.32.

Sea $(\xi_n)_{n\geq 0}$ una sucesión en X/Y tal que la serie $\sum_{n=0}^{\infty}\|\xi_n\|_{X/Y}$ converge en $\mathbb{R}$. Para cada $n \in \mathbb{N}_0$ es $\|\xi_n\|_{X/Y} < 2^{-n} + 2\|\xi_n\|_{X/Y}$, así que la definición de la norma $\|-\|_{X/Y}$ implica que hay un elemento x_n en ξ_n tal que $\|x_n\|_X \leq 2^{-n} + 2\|\xi_n\|_{X/Y}$. Como la serie $\sum_{n=0}^{\infty}(2^{-n} + 2\|\xi_n\|_{X/Y})$ converge en $\mathbb{R}$ y el espacio X es completo, la Proposición 7.3.32 nos dice que la serie $\sum_{n=0}^{\infty} x_n$ converge en X. Sea x su suma y sea $\xi := x + Y$ la clase de x en X/Y. Para cada $N \in \mathbb{N}_0$ es

$$\sum_{n=0}^{N}\xi_n - \xi = \sum_{n=0}^{N} x_n - x + Y,$$

así que

$$\left\|\sum_{n=0}^{N}\xi_n - \xi\right\|_{X/Y} \leq \left\|\sum_{n=0}^{N} x_n - x\right\|$$

Como la serie $\sum_{n=0}^{\infty} x_n$ converge a x en X, esto implica que la serie $\sum_{n=0}^{\infty}\xi_n$ converge a ξ en X/Y. La Proposición 7.3.32 nos permite concluir que el espacio cociente X/Y es completo con respecto a la norma $\|-\|_{X/Y}$. □

El siguiente ejemplo muestra una situación concreta en la que usamos la construcción de la Proposición 7.3.33.

Ejemplo 7.3.34. Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, sea X un conjunto, y sea $\mathscr{F}$ un subconjunto no vacío de $\mathscr{P}(X)$ que satisface las siguientes dos condiciones:

- Si $\mathscr{G}$ es un subconjunto contable de $\mathscr{F}$, entonces $\bigcup \mathscr{G} \in \mathscr{F}$.
- Si F es un elemento de $\mathscr{F}$ y G es un subconjunto de F, entonces $G \in \mathscr{F}$.

Por ejemplo, $\mathscr{F}$ podría ser el conjunto de todos los subconjuntos numerables de X o, si X es $\mathbb{R}$, el conjunto de todos los subconjuntos de $\mathbb{R}$ de medida nula que estudiamos en el Ejercicio 5.2.4.8.

Estamos en la situación del Ejemplo 5.2.4.7 y podemos considerar el $\Bbbk$-espacio semi-normado $B_{\mathscr{F}}(X)$ de las funciones $X \to \Bbbk$ que son $\mathscr{F}$-acotadas; escribamos $\|-\|_{\mathscr{F}}$ a su semi-norma. Si como allí ponemos

$$V_0 \coloneqq \{f \in B_{\mathscr{F}}(X) : \|f\|_{\mathscr{F}} = 0\}$$

que es un subespacio de $B_{\mathscr{F}}(X)$, entonces podemos considerar el espacio cociente

$$\ell^\infty_\Bbbk(X \text{ mód } \mathscr{F}) \coloneqq B_{\mathscr{F}}(X)/V_0$$

dotado de la norma que induce la seminorma de $B_{\mathscr{F}}(X)$. Queremos probar que

$$\ell^\infty_\Bbbk(X \text{ mód } \mathscr{F}) \textit{ es un espacio de Banach,}$$

y lo haremos usando el criterio que nos da la Proposición 7.3.32.

Sea $(\phi_n)_{n\geq 0}$ una sucesión en $\ell^\infty_\Bbbk(X \text{ mód } \mathscr{F})$ tal que la serie $\sum_{n=0}^\infty \|\phi_n\|$ converge en $\mathbb{R}$. Para cada elemento n de $\mathbb{N}_0$ es

$$\|\phi_n\| = \text{ínf}\{\|f\|_{\mathscr{F}} : f \in \phi_n\} < 2^{-n} + 2\|\phi_n\|,$$

así que hay un elemento f_n en ϕ_n tal que $\|f_n\|_{\mathscr{F}} \leq 2^{-n} + 2\|\phi_n\|$: esto significa que el ínfimo del conjunto de las $\mathscr{F}$-cotas de f_n es menor o igual que $2^{-n} + 2\|\phi_n\|$ y hay, por lo tanto, un elemento F_n de $\mathscr{F}$ tal que

$$|f_n(x)| \leq 2 \cdot 2^{-n} + 2\|\phi_n\| \text{ para todo elemento } x \text{ de } X \smallsetminus F_n. \tag{7.8}$$

Sea $F \coloneqq \bigcup_{n=0}^\infty F_n$, que es un elemento de $\mathscr{F}$. De (7.8) sabemos que $(f_n|_{X\smallsetminus F})_{n\geq 0}$ es una sucesión de funciones acotadas $X \smallsetminus F \to \Bbbk$, esto es, de elementos de $\ell^\infty_\Bbbk(X \smallsetminus F)$, y que para cada $n \in \mathbb{N}_0$ es

$$\|f_n|_{X\smallsetminus F}\|_\infty \leq 2 \cdot 2^{-n} + 2\|\phi_n\|.$$

Como la serie $\sum_{n=0}^\infty (2 \cdot 2^{-n} + 2\|\phi_n\|)$ converge y, de acuerdo al Ejercicio 7.3.31, el espacio normado $\ell^\infty_\Bbbk(X \smallsetminus F)$ es completo, sabemos que la serie $\sum_{n=0}^\infty f_n|_{X\smallsetminus F}$ converge allí a una funcion acotada $f : X \smallsetminus F \to \Bbbk$.

Consideremos ahora la funcion $\tilde{f} : X \to \Bbbk$ que en cada elemento x de X toma el valor

$$\tilde{f}(x) = \begin{cases} f(x) & \text{si } x \in X \smallsetminus F; \\ 0 & \text{si } x \in F. \end{cases}$$

Es claro que esta funcion $\tilde{f}$ es acotada, así que es $\mathscr{F}$-acotada y, por lo tanto, es un elemento de $B_{\mathscr{F}}(X)$. Sea $\phi \coloneqq \tilde{f} + V_0$ la clase de $\tilde{f}$ en $\ell^\infty_{\Bbbk}(X \text{ mód } \mathscr{F})$. Si N es un elemento de $\mathbb{N}_0$, entonces

$$\tilde{f} - \sum_{n=0}^{N} f_n \in \phi - \sum_{n=0}^{N} \phi_n$$

y la restricción de $\tilde{f} - \sum_{n=0}^{N} f_n$ a $X \setminus F$ es la funcion $f - \sum_{n=0}^{N} f_n|_{X\setminus F}$, que es acotada, así que

$$\left\| \phi - \sum_{n=0}^{N} \phi_n \right\| \leq \left\| \tilde{f} - \sum_{n=0}^{N} f_n \right\|_{\mathscr{F}} \leq \left\| f - \sum_{n=0}^{N} f_n|_{X\setminus F} \right\|_\infty .$$

Como la serie $\sum_{n=0}^{\infty} f_n|_{X\setminus F}$ converge en $\ell^\infty_{\Bbbk}(X \setminus F)$ a f, esto nos permite concluir que la serie $\sum_{n=0}^{\infty} \phi_n$ converge en $\ell^\infty_{\Bbbk}(X \text{ mód } \mathscr{F})$ a ϕ. Esto completa la demostración de la afirmación 7.3.34.

7.4. Métricas equivalentes en el sentido de Cauchy

Usando la condición de Cauchy podemos construir una nueva relación de equivalencia entre métricas. Para hacerlo, empezamos con una nueva notación.

Definición 7.4.1. Si (X, d) es un espacio métrico, escribimos $\mathscr{C}(X)$ — o, cuando es necesario poner de manifiesto la métrica, $\mathscr{C}(x, d)$ — al conjunto de todas las sucesiones en X que satisfacen la condición de Cauchy con respecto a la métrica d.

Notemos que si Y es un subespacio de un espacio métrico Y, entonces claramente $\mathscr{C}(Y)$ es precisamente el subconjunto de $\mathscr{C}(X)$ de todos los elementos que toman valores en Y.

Definición 7.4.2. Sea X un conjunto. Dos métricas d y d' sobre X son ***equivalentes en el sentido de Cauchy*** si $\mathscr{C}(X, d) = \mathscr{C}(X, d')$, y en ese caso escribimos $d \sim_{\text{Cau}} d'$.

Es evidente, en vista de la forma de esta definición, que la relación $\sim_{\text{Cau}}$ es una relación de equivalencia en el conjunto de todas las métricas sobre un conjunto.

Proposición 7.4.3.

(*i*) *Dos métricas equivalentes en el sentido de Lipschitz, son equivalentes en el sentido de Cauchy.*

(*ii*) *Dos métricas equivalentes en el sentido de Cauchy, son topológicamente equivalentes.*

(*iii*) *Dos métricas completas son equivalentes en el sentido de Cauchy si y solamente si son topológicamente equivalentes.*

Demostración. Sea X un conjunto y sean d y d' dos métricas sobre X.

(*i*) Supongamos primero que d y d' son equivalentes en el sentido de Lipschitz, de manera que hay un número positivo c tal que $c^{-1} \cdot d(x,y) \leq d'(x,y) \leq c \cdot d(x,y)$ para toda elección de x y de y en X. Sea $(x_n)_{n\geq 0}$ una sucesión en X que satisface la condición de Cauchy con respecto a la métrica d y sea ε un número positivo. Como la sucesión satisface la condición de Cauchy con respecto a d, hay un entero no negativo n_0 tal que para cada m, $n \in \mathbb{N}_0$ vale $m, n \geq n_0 \implies d(x_m, x_n) < \varepsilon/c$, y esto implica que si m y n son dos elementos de $\mathbb{N}_0$ tales que $m, n \geq n_0$ es $d'(x_m, x_n) \leq c \cdot d(x_m, x_n) < \varepsilon$. Vemos así que la sucesión $(x_n)_{n\geq 0}$ también satisface la condición de Cauchy con respecto a la métrica d'. Esto prueba que $\mathscr{C}(X,d) \subseteq \mathscr{C}(X,d')$ y, por simetría, que de hecho $\mathscr{C}(X,d) = \mathscr{C}(X,d')$, de manera que las métricas d y d' son equivalentes en el sentido de Cauchy.

(*ii*) Supongamos ahora que las métricas d y d' son equivalentes en el sentido de Cauchy. Para probar que d y d', de acuerdo al Corolario 7.1.24 y gracias a la simetría de la relación de equivalencia en el sentido de Cauchy, es suficiente que mostremos que toda sucesión convergente con respecto a d converge con respecto a d' y al mismo límite.

Sea entonces $(x_n)_{n\geq 0}$ una sucesión en X que converge con respecto a la métrica d a un punto x. La sucesión $(y_n)_{n\geq 0}$ que para cada $n \in \mathbb{N}_0$ tiene

$$y_n = \begin{cases} x_{n/2} & \text{si } n \text{ es par;} \\ x & \text{si } n \text{ es impar} \end{cases}$$

también converge a x con respecto a d y, en particular, satisface la condición de Cauchy con respecto a d. La hipótesis hipótesis implica entonces que la sucesión $(y_n)_{n\geq 0}$ también satisface la condición de Cauchy con respecto a la métrica d'. Como su subsucesión $(y_{2n+1})_{n\geq 0}$ claramente converge a x, la Proposición 7.3.7 nos dice que $(y_n)_{n\geq 0}$ también converge a x con respecto a d'. Por supuesto, esto implica que la subsucesión $(y_{2n})_{n\geq 0}$ converge a x con respecto a d', y esa subsucesión no es otra cosa que la sucesión $(x_n)_{n\geq 0}$ con la que empezamos. Esto, por lo tanto, prueba lo que queremos.

(*iii*) Supongamos finalmente que las métricas d y d' son completas. Si d y d' son equivalentes en el sentido de Cauchy, entonces son topológicamente equivalentes, como acabamos de ver. Supongamos que d y d' son topológicamente equivalentes. Si $(x_n)_{n\geq 0}$ es una sucesión en X que satisface la condición de Cauchy con respecto a d, entonces esa sucesión converge con respecto a d, ya que la métrica d es completa, y, por lo tanto, también converge con respecto a la métrica d', porque las métricas son topológicamente equivalentes, así que, en particular, satisface la condición de Cauchy con respecto a d'. Esto muestra que $\mathscr{C}(X,d) \subseteq \mathscr{C}(X,d')$ y, por simetría, que, de hecho, $\mathscr{C}(X,d) = \mathscr{C}(X,d')$. Podemos concluir entonces que las métricas d y d' son equivalentes en el sentido de Cauchy, como queremos. □

En general, las implicaciones recíprocas de las de las dos primeras partes de la Proposición 7.4.3 no son ciertas.

Ejercicio 7.4.4. Pruebe que las funciones

$$d_1 : (x, y) \in \mathbb{R} \times \mathbb{R} \mapsto |x - y| \in \mathbb{R}$$

y

$$d_2 : (x, y) \in \mathbb{R} \times \mathbb{R} \mapsto |x^3 - y^3| \in \mathbb{R}$$

son métricas sobre $\mathbb{R}$ que son equivalentes en el sentido de Cauchy pero no en el sentido de Lipschitz.

Ejercicio 7.4.5. Pruebe que las funciones

$$d_1 : (x, y) \in \mathbb{R} \times \mathbb{R} \mapsto |x - y| \in \mathbb{R}$$

y

$$d_2 : (x, y) \in \mathbb{R} \times \mathbb{R} \mapsto |\arctan x - \arctan y| \in \mathbb{R}$$

son métricas sobre $\mathbb{R}$ que son topológicamente equivalentes pero no en el sentido de Cauchy.

7.5. Espacios normados de dimensión finita

Como una primera aplicación de la noción de completitud queremos probar el siguiente resultado, que es básico en toda la teoría.

Proposición 7.5.1. *Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$. Si V es un espacio vectorial de dimensión finita sobre $\mathbb{k}$, entonces todas las normas sobre V son equivalentes y V es completo con respecto a cualquiera de ellas.*

Demostración. Digamos que un espacio vectorial V sobre $\mathbb{k}$ es ***bueno*** si todas sus normas son equivalentes y es completo con respecto a cualquiera de ellas. Probemos que todo $\mathbb{k}$-espacio vectorial de dimensión finita es bueno haciendo inducción con respecto a dimensión. Es claro que los espacios de dimensión 0 son buenos, claro.

Primer paso. Sea V un $\mathbb{k}$-espacio vectorial de dimensión 1 y sea e un elemento no nulo de V. Sean primero $\|-\|$ y $\|-\|'$ dos normas sobre V. Si x es un elemento de V, entonces hay un escalar t en $\mathbb{k}$ tal que $x = te$ y, por lo tanto, $\|x\| = \|te\| = |t| \cdot \|e\|$ y $\|x\|' = \|te\|' = |t| \cdot \|e\|'$, así que

$$\|x\| = \frac{\|e\|}{\|e\|'} \cdot \|x\|'.$$

Esto implica que las normas $\|-\|$ y $\|-\|'$ son equivalentes.

Fijemos ahora una norma $\|-\|$ sobre V, y sea $(x_n)_{n\geq 0}$ una sucesión de Cauchy en V. Para cada $n \in \mathbb{N}_0$ hay exactamente un elemento t_n de $\Bbbk$ tal que $x_n = t_n e$. Si ε es un número positivo, entonces hay un entero no negativo n_0 tal que para toda elección de m y n en $\mathbb{N}_0$ vale $m, n \geq n_0 \implies \|x_m - x_n\| < \varepsilon\|e\|$, y esto implica que si $m, n \in \mathbb{N}_0$ son tales que $m, n \geq n_0$ es

$$|t_m - t_n| = \frac{|t_m - t_n|\|e\|}{\|e\|} = \frac{\|t_m e - t_m e\|}{\|e\|} = \frac{\|x_m - x_n\|}{\|e\|} < \varepsilon.$$

Vemos así que la sucesión $(t_n)_{n\geq 0}$ en $\Bbbk$ satisface la condición de Cauchy en $\Bbbk$ y, como $\Bbbk$ es un espacio métrico completo, que converge a un elemento τ de $\Bbbk$.

Sea ε un número positivo. Como la sucesión $(t_n)_{n\geq 0}$ converge a τ, hay un entero no negativo n_0 tal que para todo $n \in \mathbb{N}_0$ vale $n \geq n_0 \implies |t_n - \tau| < \varepsilon/\|e\|$, y esto implica que para todo $n \in \mathbb{N}$ tal que $n \geq n_0$ es

$$\|x_n - \tau e\| = \|t_n e - \tau e\| = |t_n - \tau| \cdot \|e\| < \varepsilon.$$

Vemos así que la sucesión $(x_n)_{n\geq 0}$ converge en V a τe y, por lo tanto, que el espacio V es completo con respecto a la norma $\|-\|$. Podemos concluir con todo esto que el espacio V es bueno.

Segundo paso. Sean ahora k un entero positivo mayor que 1, supongamos que todo $\Bbbk$-espacio vectorial de dimensión menor que k es bueno, sea V un $\Bbbk$-espacio vectorial de dimensión k, y elijamos una base ordenada $\mathscr{B} = (e_1, e_2, \ldots, e_k)$ de V. Sabemos del Ejercicio 5.2.3.14 que hay exactamente una función $\|-\|_1 : V \to \mathbb{R}$ tal que

$$\|a_1 e_1 + a_2 e_2 + \cdots + a_k e_k\| = |a_1| + |a_2| + \cdots + |a_k|,$$

y que esa función es una norma sobre V. Más aún, en el Ejercicio 7.3.25 mostramos que V es completo con respecto a esa norma.

Sea $\|-\|$ una norma cualquiera sobre V, y consideremos el número

$$K \coloneqq \text{máx}\{\|e_1\|, \|e_2\|, \ldots, \|e_k\|\}.$$

Si x es un elemento de V, y $a_1, a_2, \ldots, a_k$ son los escalares de $\Bbbk$ tales que $x = a_1 e_1 + a_2 e_2 + \cdots + a_k e_k$, entonces tenemos que

$$\begin{aligned}\|x\| &= \|a_1 e_1 + a_2 e_2 + \cdots + a_k e_k\| \leq |a_1| \cdot \|e_1\| + |a_2| \cdot \|e_2\| + \cdots + |a_k| \cdot \|e_k\| \\ &\leq |a_1| \cdot K + |a_2| \cdot K + \cdots + |a_k| \cdot K = K \cdot (|a_1| + |a_2| + \cdots + |a_k|) = K \cdot \|x\|_1.\end{aligned}$$

Sea, por otro lado, l un elemento de $[\![k]\!]$. El subespacio $W \coloneqq \langle e_1, \ldots, \widehat{e_l}, \ldots, e_k\rangle$ de V generado por los elementos de la base ordenada $\mathscr{B}$ distintos del l-ésimo tiene dimensión menor que k: si lo vemos como un subespacio normado de V, entonces la hipótesis inductiva nos dice que es completo y, de acuerdo a la Proposición 7.3.26, cerrado en V. En particular, como e_l no pertenece

a V, hay un número positivo r_l tal que $B_{r_l}(e_l) \cap W = \varnothing$. Se sigue de esto que para toda elección de n escalares $a_1, a_2, \ldots, a_k$ en $\Bbbk$ tales que $a_l \neq 0$ tenemos que

$$r_l \leq \left\| e_l + \sum_{i \in [\![k]\!] \smallsetminus \{l\}} \frac{a_i}{a_l} e_i \right\|$$

y, por lo tanto,

$$r_l \cdot |a_l| \leq \left\| \sum_{i=1}^{k} a_i e_i \right\|.$$

Consideremos el número $r \coloneqq \min\{r_1, r_2, \ldots, r_k\}$, que es positivo. Si x un elemento cualquiera de V, y $a_1, a_2, \ldots, a_k$ son los escalares de en $\Bbbk$ tales que $x = a_1 e_1 + a_2 e_2 + \cdots + a_k e_k$, entonces para cada elemento l de $[\![k]\!]$ tenemos que o bien $a_l \neq 0$ y, por que hicimos recién,

$$r \cdot |a_l| \leq r_l \cdot |a_l| \leq \|x\|,$$

o bien $a_l = 0$, y en este caso también $\|x\| \geq r \cdot |a_l|$. Vemos así que, de hecho,

$$\frac{r}{k} \cdot \|x\|_1 = \frac{r}{k} \cdot (|a_1| + |a_2| + \cdots + |a_k|) \leq \|x\| \quad \text{cualquiera sea } x \in V.$$

Con todo esto hemos probado que para todo elemento x de V vale

$$\frac{r}{k} \cdot \|x\|_1 \leq \|x\| \leq K \cdot \|x_1\|$$

y, por lo tanto, que la norma $\|-\|$ de V es equivalente a la norma $\|-\|_1$. Esto nos permite concluir, en primer lugar, que todas las normas de V son equivalentes a la norma $\|-\|_1$ y, por lo tanto, que son todas equivalentes, y, en segundo lugar, que V es completo con respecto a todas ellas, porque es completo con respecto a $\|-\|_1$. Así, el espacio V es bueno, y esto completa la inducción. □

Una consecuencia directa de esto es la siguiente observación:

Corolario 7.5.2. *Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$. Todo subespacio vectorial de dimensión finita de un $\Bbbk$-espacio normado es cerrado en él.*

Demostración. Sea V un $\Bbbk$-espacio normado y sea W un subespacio vectorial de V de dimensión finita. Podemos ver a W como un espacio normado restringiendo la norma de V a él y la Proposición 7.5.1 nos dice que es entonces un espacio métrico completo. De acuerdo a la Proposición 7.3.26, entonces, W es un subconjunto cerrado de V. □

7.6. Equiconvergencia

Queremos mostrar que el espacio $c_{\mathbb{k}}$ de las sucesiones convergentes en $\mathbb{k}$, que es un subespacio de $\ell^\infty_{\mathbb{k}}$ y, por lo tanto, un espacio normado, es completo. Para ello haremos algunas consideraciones de carácter general antes.

Sea $\mathfrak{X}$ un conjunto de sucesiones convergentes en un espacio métrico X, para cada elemento x de $\mathfrak{X}$ sea α_x el límite de x, y sea ε un número positivo. Si $x = (x_n)_{n\geq 0}$ es un elemento de $\mathfrak{X}$, entonces x converge a α_n y, por lo tanto, hay un entero no negativo n_0^x tal que para todo $n \in \mathbb{N}_0$ vale $n \geq n_0^x \implies d(x_n, \alpha_x) < \varepsilon$. Hay entonces una *función* $n_0 : x \in \mathfrak{X} \mapsto n_0^x \in \mathbb{N}_0$ tal que para todo $x \in \mathfrak{X}$ y todo $n \in \mathbb{N}_0$ vale $n \geq n_0(x) \implies d(x_n, \alpha_n) < \varepsilon$. El siguiente ejemplo muestra que, en general, no es posible encontrar una función con esta propiedad que sea constante.

Ejemplo 7.6.1. Para cada $m \in \mathbb{N}_0$ consideremos la sucesión $x_m \coloneqq ((1 - 1/(m+2))^n)_{n\geq 0}$, y sea $\mathfrak{X} \coloneqq \{x_m : m \in \mathbb{N}_0\}$. Todos los elementos de $\mathfrak{X}$ convergen a 0 en $\mathbb{R}$: si m es un entero no negativo y ε es un número real positivo, entonces para todo $n \in \mathbb{N}_0$ vale

$$|x_{m,n} - 0| = (1 - 1/(m+2))^n < \varepsilon \iff n > \frac{\ln \varepsilon}{\ln(1 - 1/(m+2))}, \tag{7.9}$$

de manera que

$$n_0^m \coloneqq \max\left\{\left\lceil \frac{\ln \varepsilon}{\ln(1 - 1/(m+2))} \right\rceil, 0\right\},$$

es un elemento de $\mathbb{N}_0$ tal que para todo $n \in \mathbb{N}_0$ vale $n \geq n_0^m \implies d(x_{m,n}, 0) < \varepsilon$.

Ahora bien, no existe ningún entero no negativo n_0 tal que para todo $m \in \mathbb{N}_0$ y todo $n \in \mathbb{N}_0$ valga $n \geq n_0 \implies d(x_{m,n}, 0) < \frac{1}{2}$: si n_0 fuera un tal elemento de $\mathbb{N}_0$, entonces tendríamos que $d(x_{m,n_0}) < \frac{1}{2}$ para todo $m \in \mathbb{N}_0$ y, de acuerdo a (7.9), que

$$n_0 > \frac{\ln \frac{1}{2}}{\ln(1 - 1/(m+2))}$$

para todo $m \in \mathbb{N}_0$, lo que — como el lector puede verificar fácilmente — es absurdo.

Esto motiva la siguiente definición.

Definición 7.6.2. Sea $\mathfrak{X}$ una conjunto de sucesiones convergentes en un espacio métrico X, y para cada elemento x de $\mathfrak{X}$ sea α_x el límite de x. Decimos que la familia $\mathfrak{X}$ es ***equiconvergente*** si para todo número positivo ε hay un entero no negativo n_0 tal que para todo $n \in \mathbb{N}_0$ y todo elemento x de $\mathfrak{X}$ vale $n \geq n_0 \implies d(x_n, \alpha_n) < \varepsilon$.

La condición de equiconvergencia es importante, entre muchas otras razones, porque vale el

siguiente resultado.

Proposición 7.6.3. *Sea X un espacio métrico completo y para cada elección de m y n en $\mathbb{N}_0$ sea $x_{m,n}$ un elemento de X. Si*

- *para cada $m \in \mathbb{N}_0$ la sucesión $x_m = (x_{m,n})_{n\geq 0}$ converge en X a un punto α_m,*
- *el conjunto $\{x_m : m \in \mathbb{N}_0\}$ es equiconvergente,*
- *para cada $n \in \mathbb{N}_0$ la sucesión $(x_{m,n})_{m\geq 0}$ converge en X a un punto β_n,*

entonces las sucesiones $(\alpha_m)_{m\geq 0}$ y $(\beta_n)_{n\geq 0}$ convergen en X al mismo límite.

Este resultado es conocido como el *Teorema de Moore–Osgood*, por Eliakim Hastings Moore y William Fogg Osgood. La conclusión de la proposición es que vale la igualdad

$$\lim_{n\to\infty} \lim_{m\to\infty} x_{m,n} = \lim_{m\to\infty} \lim_{n\to\infty} x_{m,n}$$

y, en particular, que todos estos límites existen.

Demostración. Supongamos que se cumplen las tres hipótesis.

Primer paso. Primero mostraremos que la sucesión $(\alpha_m)_{m\geq 0}$ converge. Sea ε un número positivo. La hipótesis de equiconvergencia nos dice que hay un entero no negativo n_0 tal que

$$\textit{siempre que } m \textit{ y } n \textit{ están en } \mathbb{N}_0 \textit{ vale } n \geq n_0 \implies d(x_{m,n}, \alpha_m) < \frac{\varepsilon}{3}.$$

Por otro lado, la sucesión $(x_{m,n_0})_{m\geq 0}$ converge a β_{n_0}, así que satisface la condición de Cauchy y hay, en consecuencia, un entero no negativo m_0 tal que siempre que m y m' son elementos de $\mathbb{N}_0$ vale $m, m' \geq m_0 \implies d(x_{m,n_0}, x_{m',n_0}) < \varepsilon/3$. Esto implica que si m y m' son dos elementos cualesquiera de $\mathbb{N}_0$, entonces

$$d(\alpha_m, \alpha_{m'}) \leq d(\alpha_m, x_{m,n_0}) + d(x_{m,n_0}, x_{m',n_0}) + d(x_{m',n_0}, \alpha_{m'}) < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Vemos así que la sucesión $(\alpha_m)_{m\geq 0}$ satisface la condición de Cauchy, así que, como el espacio X es completo, converge a algún punto α de X.

Segundo y último paso. Mostraremos ahora que la sucesión $(\beta_n)_{n\geq 0}$ también converge a α. Sea ε un número positivo. La equiconvergencia implica que hay un entero no negativo n_0 tal que

$$\textit{siempre que } m \textit{ y } n \textit{ están en } \mathbb{N}_0 \textit{ vale } n \geq n_0 \implies d(x_{m,n}, \alpha_m) < \frac{\varepsilon}{3}.$$

Sea n un entero tal que $n \geq n_0$.

- Como la sucesión $(\alpha_m)_{m\geq 0}$ converge a ξ, hay un entero no negativo m_1 tal que para todo $m \in \mathbb{N}_0$ vale $m \geq m_1 \implies d(\alpha_m, \xi) < \varepsilon/3$.

- Como la sucesión $(x_{m,n})_{m\geq 0}$ converge a β_n, hay un entero no negativo m_2 tal que para todo $m \in \mathbb{N}_0$ vale $m \geq m_2 \implies d(x_{m,n}, \beta_n) < \varepsilon/3$.

Si $m_0 \coloneqq \max\{m_1, m_2\}$, entonces tenemos que

$$d(\beta_n, \xi) \leq d(\beta_n, x_{m_0,n}) + d(x_{m_0,n}, \alpha_{m_0}) + d(\alpha_{m_0}, \xi) < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Esto prueba que la sucesión $(\beta_n)_{n\geq 0}$ converge a ξ, como afirma la proposición. □

Notemos que la hipótesis que aparece en esta proposición de que el espacio X sea completo es necesaria.

Ejemplo 7.6.4. Consideremos el subespacio $X \coloneqq (0, +\infty)$ de $\mathbb{R}$ y sea $x_{m,n} \coloneqq 1/(m+1) + 1/(n+1)$ siempre que m y n están en $\mathbb{N}_0$. Es claro que para todo $m \in \mathbb{N}_0$ la sucesión $x_m = (x_{m,n})_{n\geq 0}$ converge a $\alpha_m \coloneqq 1/(m+1)$ y, más aún, que el conjunto $\{x_m : m \in \mathbb{N}_0\}$ es equiconvergente: si ε es un número positivo y n_0 es un entero positivo tal que $1/n_0 < \varepsilon$, entonces siempre que m y n son elementos de $\mathbb{N}_0$ tales que $n \geq n_0$ tenemos que $|x_{m,n} - \alpha_n| = 1/(n+1) < 1/n_0 < \varepsilon$. Finalmente, es claro que para cada $n \in \mathbb{N}_0$ la sucesión $(x_{m,n})_{m\geq 0}$ converge en X a $\beta_n = 1/(n+1)$. Sin embargo, las sucesiones $(\alpha_m)_{m\geq 0} = (1/(m+1))_{m\geq 0}$ y $(\beta_n)_{n\geq 0} = (1/(n+1))_{n\geq 0}$ no convergen en X.

Probemos ahora lo que queremos.

Proposición 7.6.5. *Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$. Los espacios $c_{\mathbb{k}}$ y $c_{0,\mathbb{k}}$ de las sucesiones de $\mathbb{k}^{\mathbb{N}_0}$ que convergen y que convergen a 0, dotados de la norma que tienen por ser subespacios de $\ell^\infty_{\mathbb{k}}$, son espacios de Banach.*

Demostración. Tenemos que probar que $c_{\mathbb{k}}$ y $c_{0,\mathbb{k}}$ son completos con respecto a las normas que obtienen de la de $\ell^\infty_{\mathbb{k}}$ por restricción. El Ejercicio 7.3.31 nos dice que $\ell^\infty_{\mathbb{k}}$ es completo, así que para hacer eso es suficiente, de acuerdo a la Proposición 7.3.26, que mostremos que $c_{\mathbb{k}}$ y $c_{0,\mathbb{k}}$ son subespacios cerrados de $\ell^\infty_{\mathbb{k}}$.

Para cada $n \in \mathbb{N}_0$ sea $x_n = (x_{m,n})_{m\geq 0}$ un elemento de $c_{\mathbb{k}}$, y supongamos que la sucesión $x = (x_n)_{n\geq 0}$ converge en $\ell^\infty_{\mathbb{k}}$ a un punto $\xi = (\xi_m)_{m\geq 0}$ de $\ell^\infty_{\mathbb{k}}$. Para mostrar que $c_{\mathbb{k}}$ es cerrado en $\ell^\infty_{\mathbb{k}}$, tenemos que probar que ξ es un elemento de $c_{\mathbb{k}}$.

- Sea ε un número positivo. Como la sucesión x converge a ξ, hay un entero no negativo n_0 tal que para todo $n \in \mathbb{N}_0$ vale $n \geq n_0 \implies \|x_n - \xi\|_\infty < \varepsilon$, y esto implica que siempre que m y n son elementos de $\mathbb{N}_0$ vale $n \geq n_0 \implies |x_{m,n} - \xi_m| < \varepsilon$. Esto nos dice que el conjunto $\{(x_{m,n})_{n\geq 0} : m \in \mathbb{N}_0\}$ es equiconvergente y que para cada $m \in \mathbb{N}_0$ la sucesión $(x_{m,n})_{n\geq 0}$ converge a ξ_m.
- Para cada $n \in \mathbb{N}_0$ la sucesión $x_n = (x_{m,n})_{m\geq 0}$ esta en $c_{\mathbb{k}}$, así que converge a un elemento β_n de $\mathbb{k}$.

Como $\mathbb{k}$ es un espacio métrico completo, la Proposición 7.6.3 nos permite concluir que las sucesión $\xi = (\xi_m)_{m\geq 0}$ y $(\beta_n)_{n\geq 0}$ convergen en $\mathbb{k}$ y al mismo límite. En particular, esto nos dice que ξ es un elemento de $c_{\mathbb{k}}$, como queríamos.

Supongamos ahora $n \in \mathbb{N}_0$ sea $x_n = (x_{m,n})_{m\geq 0}$ un elemento de $c_{0,\mathbb{k}}$, y supongamos que la sucesión $x = (x_n)_{n\geq 0}$ converge en $\ell^\infty_{\mathbb{k}}$ a un punto $\xi = (\xi_m)_{m\geq 0}$ de $\ell^\infty_{\mathbb{k}}$. Lo que acabamos de hacer nos dice que la sucesión ξ pertenece a $c_{\mathbb{k}}$ y que su límite es $\lim_{n\to\infty} \lim_{m\to\infty} x_{m,n} = \lim_{n\to\infty} 0 = 0$, así que ξ pertenece a $c_{0,\mathbb{k}}$. Esto prueba que el subespacio $c_{0,\mathbb{k}}$ de $\ell^\infty_{\mathbb{k}}$ es cerrado. □

Capítulo 8

Continuidad

8.1. Funciones continuas

Hasta ahora estudiamos a los espacios métricos de a uno por vez. En este capítulo estudiaremos funciones entre ellos. Salvo que nos veamos forzados por el contexto o por nuestras intenciones, escribiremos a todas las métricas de todos los espacios usando la notación d.

Definición 8.1.1. Sean (X, d_X) e (Y, d_Y) dos espacios métricos. Decimos que una función $f : X \to Y$ es ***continua*** en un punto ξ de X si

> *para todo número positivo ε hay un número positivo δ tal que para todo elemento x de X vale $d_X(x, \xi) < \delta \implies d_Y(f(x), f(\xi)) < \varepsilon$.*

La condición de esta definición está enunciada en términos de las métricas de los espacios métricos involucrados, pero podemos dar condiciones equivalentes en términos de entornos y de sucesiones.

Proposición 8.1.2. *Sean X e Y dos espacios métricos, sea $f : X \to Y$ una función, sea ξ un punto de X, y sea $\mathcal{N}$ una base de entornos de $f(\xi)$ en Y. Las siguientes afirmaciones son equivalentes:*

(a) *La función f es continua en ξ.*
(b) *Para todo elemento N de $\mathcal{N}$ el conjunto $f^{-1}(N)$ es un entorno de ξ.*
(c) *Cada vez que $(x_n)_{n\geq 0}$ es una sucesión en X que converge a ξ la sucesión $(f(x_n))_{n\geq 0}$ converge a $f(\xi)$.*

Demostración. Supongamos primero que la función f es continua en ξ y sea N un elemento de $\mathscr{N}$. Como N es un entorno de $f(\xi)$, hay un número positivo ε tal que $B_\varepsilon(f(\xi)) \subseteq N$ y, como la función f es continua en ξ, hay un número positivo δ tal que para todo elemento x de X vale $d(x,\xi) < \delta \implies d(f(x), f(\xi)) < \varepsilon$. Esto implica que para cada elemento x de $B_\delta(\xi)$ es $f(x) \in B_\varepsilon(f(\xi)) \subseteq N$ y, por lo tanto, que $B_\delta(\xi) \subseteq f^{-1}(N)$, de manera que $f^{-1}(N)$ es un entorno de ξ. Esto prueba la implicación $(a) \Rightarrow (b)$.

Supongamos ahora que la función f satisface la condición (b), sea $(x_n)_{n\geq 0}$ una sucesión en X que converge a ξ, y sea N un entorno de $f(\xi)$. La hipótesis nos dice que $f^{-1}(N)$ es un entorno de x, así que la sucesión $(x_n)_{n\geq 0}$ está a la larga en $f^{-1}(N)$, y esto implica inmediatamente que la sucesión $(f(x_n))_{n\geq 0}$ está a la larga en N. Vemos así que la sucesión $(f(x_n))_{n\geq 0}$ converge a $f(\xi)$ y, en definitiva, que vale la implicación $(b) \Rightarrow (c)$.

Finalmente, supongamos que se satisface la condición (c), sea ε un número positivo y, para llegar a una contradicción, supongamos que para todo número positivo δ hay un elemento x de X tal que $d(x,\xi) < \delta$ y $d(f(x), f(\xi)) \geq \varepsilon$. Esto implica, en particular, que para cada entero no negativo n hay un elemento x_n de X tal que $d(x,\xi) < 1/(n+1)$ y $d(f(x_n), f(\xi)) \geq \varepsilon$, claro. Obtenemos de esta forma una sucesión $(x_n)_{n\geq 0}$ que converge a ξ. La hipótesis nos dice que la sucesión $(f(x_n))_{n\geq 0}$ converge a $f(\xi)$ y, por lo tanto, que hay un entero no negativo m tal que $d(f(x_m), f(\xi)) < \varepsilon$: esto es absurdo, en vista de la forma en que elegimos al punto x_m. Esta contradicción provino de haber supuesto que la función f no es continua en ξ, así que debe serlo. Esto prueba que vale la implicación $(c) \Rightarrow (a)$. □

Casi siempre nos interesan las funciones entre espacios métricos que son continuas en todos los puntos de su dominio.

Definición 8.1.3. Sean X e Y dos espacios métricos. Una función $f : X \to Y$ es ***continua*** si es continua en cada punto de X.

A veces escribimos $f : (X, d_X) \to (Y, d_Y)$ a una función continua $f : X \to Y$ si queremos enfatizar las métricas de los espacios.

Veamos algunos ejemplos.

Proposición 8.1.4. *Sea X un espacio métrico.*

(i) Si x es un punto de X, entonces la función $f : y \in X \mapsto d(y,x) \in \mathbb{R}$ es continua.

(ii) Si S es un subconjunto no vacío de X, entonces la función $g : y \in X \mapsto d(y,S) \in \mathbb{R}$ es continua.

Demostración. (i) Sea x un elemento de X y consideremos la función $f : y \in X \mapsto d(y,x) \in \mathbb{R}$. Sea y un elemento de X y sea ε un número positivo. Si y' es un elemento de X tal que $d(y,y') < \varepsilon$, entonces $d(y',x) \leq d(y',y) + d(y,x) < \varepsilon + d(y,x)$ y, por lo tanto, $d(y',x) - d(y,x) < \varepsilon$, y $d(y,x) \leq d(y,y') + d(y',x) < \varepsilon + d(y',x)$, así que $d(y,x) - d(y',x) < \varepsilon$: vemos de esta forma

que $|f(y') - f(y)| = |d(y', x) - d(y, x)| < \varepsilon$. Esto prueba que la función f es continua en y.

(*ii*) Sea ahora S un subconjunto no vacío de X y sea $g : y \in X \mapsto d(y, S) \in \mathbb{R}$. Sea y un elemento de X, sea ε un número positivo, y supongamos que y' es un elemento de X tal que $d(y', y) < \varepsilon/2$. Si s es un elemento de S, entonces $d(y', s) \leq d(y', y) + d(y, s) \leq \varepsilon/2 + d(y, S)$, así que $d(y', S) = \sup\{d(y', s) : s \in S\} \leq \varepsilon/2 + d(y, S)$ y, por lo tanto, que $d(y', S) - d(y, S) \leq \varepsilon/2 < \varepsilon$. Por simetría, también tenemos que $d(y, S) - d(y', S) < \varepsilon/2$ y, juntando las dos desigualdad que obtuvimos, que $|g(y') - g(y)| = |d(y', S) - d(y, S)| < \varepsilon$. Esto demuestra que la función g es continua en el punto y. □

Podemos caracterizar la continuidad de muchas formas.

Proposición 8.1.5. *Sean X e Y dos espacios métricos y sea $f : X \to Y$ una función. Las siguientes afirmaciones son equivalentes:*

(*a*) *La función f es continua.*
(*b*) *Si U es un subconjunto abierto de Y, entonces el conjunto $f^{-1}(U)$ es abierto en X.*
(*c*) *Si F es un subconjunto cerrado de Y, entonces el conjunto $f^{-1}(U)$ es cerrado en X.*
(*d*) *Para todo subconjunto S de Y es $f^{-1}(S^\circ) \subseteq f^{-1}(S)^\circ$.*
(*e*) *Para todo subconjunto S de X es $f(\overline{S}) \subseteq \overline{f(S)}$.*

Demostración. (*a*) $\Rightarrow$ (*b*) Supongamos que la función f es continua y sea U un subconjunto abierto de Y. Si x es un punto de $f^{-1}(U)$, entonces $f(x) \in U$ y, como U es un entorno de $f(x)$, la continuidad de f en x implica, de acuerdo a la Proposición 8.1.2, que $f^{-1}(U)$ es un entorno de x. Vemos así que $f^{-1}(U)$ es un entorno de cada uno de sus puntos, así que es un conjunto abierto.

(*b*) $\Rightarrow$ (*c*) Supongamos que vale la afirmación (*b*) y sea F un subconjunto cerrado de Y. El conjunto $Y \setminus F$ es entonces abierto y la hipótesis implica que $X \setminus f^{-1}(F) = f^{-1}(Y \setminus F)$ es un subconjunto abierto de X, de manera que su complemento en X, $f^{-1}(F)$ es cerrado en X.

(*c*) $\Rightarrow$ (*e*) Supongamos que vale la afirmación (*c*) y sea S un subconjunto de X. Como $\overline{f(S)}$ es un subconjunto cerrado de Y, la hipótesis implica que $f^{-1}(\overline{f(S)})$ es un subconjunto cerrado de X. Ahora bien, es $f(S) \subseteq \overline{f(S)}$, así que $S \subseteq f^{-1}(f(S)) \subseteq f^{-1}(\overline{f(S)})$ y, por lo tanto, $\overline{S} \subseteq f^{-1}(\overline{f(S)})$ y, finalmente, $f(\overline{S}) \subseteq \overline{f(S)}$.

(*e*) $\Rightarrow$ (*d*) Supongamos que vale la afirmación (*d*) y sea S un subconjunto de Y. Como $f(f^{-1}(Y \setminus S)) \subseteq Y \setminus S$, la hipótesis implica que $f(\overline{f^{-1}(Y \setminus S)}) \subseteq \overline{f(f^{-1}(Y \setminus S))} \subseteq \overline{Y \setminus S}$ y, por lo tanto, que $\overline{f^{-1}(Y \setminus S)} \subseteq f^{-1}(\overline{Y \setminus S})$, así que

$$X \setminus f^{-1}(\overline{Y \setminus S}) \subseteq X \setminus \overline{f^{-1}(Y \setminus S)}. \tag{8.1}$$

Ahora bien, gracias a la Proposición 6.5.15 sabemos que

$$f^{-1}(S^\circ) = f^{-1}(Y \setminus \overline{Y \setminus S}) = X \setminus f^{-1}(\overline{Y \setminus S}),$$

y que

$$f^{-1}(S)^\circ = X \setminus \overline{X \setminus f^{-1}(S)} = X \setminus \overline{f^{-1}(Y \setminus S)},$$

así que (8.1) nos dice que $f^{-1}(S^\circ) \subseteq f^{-1}(S)^\circ$.

$(d) \Rightarrow (a)$ Supongamos finalmente que vale la afirmación (d) y sea ξ un elemento de X. Si N es un entorno de $f(\xi)$, de manera que $f(\xi) \in N^\circ$, entonces la hipótesis nos dice que $\xi \in f^{-1}(N^\circ) \subseteq f^{-1}(N)^\circ$, así que $f^{-1}(N)$ es un entorno de ξ. De acuerdo a la Proposición 8.1.2, la función f es continua en ξ y, como esto es así cualquiera sea el punto ξ de X, continua. □

Una consecuencia de la caracterización (b) de la continuidad que nos da la Proposición 8.1.5 es que que una función $f : X \to Y$ entre espacios métricos sea continua o no depende solamente de las topologías de X y de Y y no de las métricas de esos espacios.

También podemos caracterizar la continuidad de una función en términos de la preservación de la convergencia de sucesiones.

Proposición 8.1.6. *Sean X e Y dos espacios métricos. Una función $f : X \to Y$ es continua si y solamente si para toda sucesión convergente $(x_n)_{n\geq 0}$ de X la sucesión $(f(x_n))_{n\geq 0}$ de Y converge.*

Notemos que no hacemos aquí ninguna hipótesis sobre cuáles son los límites de las sucesiones que estamos considerando.

Demostración. Sea $f : X \to Y$ una función. Supongamos primero que f es continua y sea $(x_n)_{n\geq 0}$ una sucesión convergente en X. Si ξ es el límite de esa sucesión, entonces f es continua en ξ y, de acuerdo a la Proposición 8.1.2, la sucesión $(f(x_n))_{n\geq 0}$ converge. Esto nos dice que la condición de la proposición es necesaria.

Veamos ahora que la condición es suficiente. Supongamos que la función f no es continua, de manera que hay un punto ξ de X en el que la función f no es continua. De acuerdo a la Proposición 8.1.2, hay entonces una sucesión $(x_n)_{n\geq 0}$ en X que converge a ξ tal que la sucesión $(f(x_n))_{n\geq 0}$ en Y no converge a $f(\xi)$ y, de acuerdo a la Proposición 7.2.6, hay un entorno N de $f(\xi)$ y una función estrictamente creciente $h : \mathbb{N}_0 \to \mathbb{N}_0$ tal que $f(x_{h(n)}) \notin N$ para todo $n \in \mathbb{N}_0$. Consideremos la sucesión $(y_n)_{n\geq 0}$ en X que tiene para cada $n \in \mathbb{N}_0$

$$y_n = \begin{cases} x_{h(n)} & \text{si } n \text{ es par;} \\ \xi & \text{si } n \text{ es impar.} \end{cases}$$

Esta sucesión converge a ξ. En efecto, sea ε un número positivo. Que $(x_n)_{n\geq 0}$ converja a ξ nos dice que hay un entero no negativo n_0 tal que para todo $n \in \mathbb{N}_0$ vale $n \geq n_0 \implies d(x_n, \xi) < \varepsilon$. Ahora bien, si n es un elemento de $\mathbb{N}_0$ tal que $n \geq n_0$, entonces o bien n es impar, y $d(y_n, \xi) = d(\xi, \xi) < \varepsilon$, o bien n es par, $h(n) \geq n \geq 0$ ya que h es estrictamente creciente, y $d(y_n, \xi) = d(x_{h(n)}, \xi) < \varepsilon$.

Por otro lado, la sucesión $(f(y_n))_{n\geq 0}$ en Y no converge. Para verlo es suficiente que probemos que no satisface la condición de Cauchy. Como N es un entorno de ξ, hay un número positivo r tal que $B_r(\xi) \subseteq N$. Si n_0 es un elemento cualquiera de $\mathbb{N}_0$ y k es el resto de dividir a n_0 por dos, entonces $n_0 + k$ y $n_0 + k + 1$ son dos elementos de $\mathbb{N}_0$ mayores o iguales que n_0, el primero es par, el segundo impar, y, por lo tanto, $d(y_{n_0+k}, y_{n_0+k+1}) = d(x_{h(n_0+k)}, \xi) \geq r$, ya que $x_{h(n_0+k)} \notin N \supseteq B_r(\xi)$.

Todo esto prueba que si la función f no es continua, entonces hay una sucesión convergente $(y_n)_{n\geq 0}$ en X tal que la sucesión $(f(y_n))_{n\geq 0}$ no converge en X, y prueba que la condición de la proposición es suficiente. □

Usando esta caracterización de la continuidad podemos probar fácilmente el siguiente resultado útil.

Proposición 8.1.7. *Sean X e Y dos espacios métricos. Si $f : X \to Y$ es una función continua, entonces el **gráfico** de f, el subconjunto*

$$\Gamma(f) \coloneqq \{(x, f(x)) \in X \times Y : x \in X\}$$

de $X \times Y$ es cerrado.

Aquí estamos viendo al producto cartesiano $X \times Y$ como un espacio métrico son su métrica producto usual.

Demostración. Sea $f : X \to Y$ una función continua, y sea $(p_n)_{n\geq 0}$ una sucesión en $\Gamma(f)$ que converge en $X \times Y$ a un punto p. Para cada $n \in \mathbb{N}_0$ hay un elemento x_n de X tal que $p_n = (x_n, f(x_n))$, y hay elementos x e y de X y de Y, respectivamente, tales que $p = (x, y)$. Como la sucesión $(p_n)_{n\geq 0}$ converge a p, tenemos que las sucesiones $(x_n)_{n\geq 0}$ y $(f(x_n))_{n\geq 0}$ convergen en X y en Y a x y a y, y como la función f es continua, la segunda converge también a $f(x)$: esto nos dice que $y = f(x)$ y, por lo tanto, que el punto p pertenece a $\Gamma(f)$. Este conjunto es, por esto, un subconjunto cerrado de $X \times Y$. □

El siguiente ejemplo muestra que la condición que da la Proposición 8.1.7 es necesaria pero no suficiente para la continuidad de la función. Veremos más adelante en la Proposición **??** una condición sobre Y que repara esto.

Ejemplo 8.1.8. Consideremos la función $f : \mathbb{R} \to \mathbb{R}$ que en cada elemento x de $\mathbb{R}$ toma el valor

$$f(x) = \begin{cases} \dfrac{1}{x} & \text{si } x \neq 0; \\ 0 & \text{si } x = 0. \end{cases}$$

El gráfico de f es el conjunto

$$\{(0,0)\} \cup \{(x, y) \in \mathbb{R} \times \mathbb{R} : xy - 1 = 0\},$$

y esto es cerrado en $\mathbb{R}\times\mathbb{R}$ porque los dos conjuntos cuya unión estamos considerando son cerrados allí: el primero porque es finito, y el segundo porque es la preimagen de $\{0\}$ por la función continua $(x, y) \in \mathbb{R} \times \mathbb{R} \mapsto xy \in \mathbb{R}$. Claramente la función f no es continua.

La clase de las funciones continuas es cerrada por varias operaciones. Probablemente la más importante es la de composición.

Proposición 8.1.9. *Sean X, Y, y Z tres espacios métricos. Si $f : X \to Y$ y $g : Y \to Z$ son funciones continuas, entonces la composición $g \circ f : X \to Y$ también es continua.*

Demostración. Sean $f : X \to Y$ y $g : Y \to Z$ dos funciones continuas. Si U es un abierto de Z, entonces $f^{-1}(U)$ es un abierto de Y y, por lo tanto, $(g \circ f)^{-1}(U) = g^{-1}(f^{-1})(U))$ es un abierto de X. Esto nos dice que la función $g \circ f$ es continua. □

Esto también es cierto punto a punto, en el sentido del siguiente ejercicio.

Ejercicio 8.1.10. Sean X, Y, y Z tres espacios métricos y sean $f : X \to Y$ y $g : Y \to Z$ dos funciones. Si f es continua en un punto ξ de X y g es continua en el punto $f(\xi)$, entonces la composición $g \circ f$ es continua en el punto ξ.

Otra cosa que podemos hacer con las funciones continuas es restringirlas.

Ejercicio 8.1.11. Sean X e Y dos espacio métrico.

(a) Si $f : X \to Y$ es una función continua y Z es un subespacio de X, entonces la restricción $f|_Z : Z \to Y$ es una función continua.
(b) Si $f : X \to Y$ es una función continua y Z es un subespacio de Y tal que $f(X) \subseteq Z$, entonces la correstricción $f|^Z : X \to Z$ es una función continua.

El siguiente ejercicio se ocupa de algunas construcción que podemos hacer con funciones continuas y productos cartesianos.

Ejercicio 8.1.12. Sean X, Y, Z y W cuatro espacios métricos.

(a) Muestre que si $f : X \to Y$ y $g : Z \to W$ son dos funciones continuas, entonces la función

$$f \times g : (x, y) \in X \times Z \mapsto (f(x), g(y)) \in Y \times W$$

también es continua.
(b) Muestre que la función $x \in X \mapsto (x, x) \in X \times X$ es continua.
(c) Muestre que si $f : X \to Y$ y $g : X \to Z$ son funciones continua, entonces también lo es

$$x \in X \mapsto (f(x), g(x)) \in Y \times Z.$$

La siguiente definición introduce una de las posibles clases de funciones que podemos considerar «isomorfismos» entre espacios métricos.

Definición 8.1.13. Sean X e Y dos espacios métricos. Una función $f : X \to Y$ es un ***homeomorfismo*** si es continua y hay una función continua $g : Y \to X$ tal que $g \circ f = \mathrm{id}_X$ y $f \circ g = \mathrm{id}_Y$.

Así, un homeomorfismo es una función continua y biyectiva cuya función inversa es continua.

Ejercicio 8.1.14. Sea X un espacio métrico. Muestre que la composición de dos homeomorfismos $X \to X$ es un homeomorfismos, y que el conjunto $\mathrm{Homeo}(X)$ de todos los homeomorfismos $X \to X$ dotada de la operación de composición $\circ$ es un grupo. Llamamos a este grupo, por supuesto, el ***grupo de homeomorfismos*** de X — se trata de un invariante extremadamente importante del espacio X.

Una observación útil es que podemos expresar la equivalencia topológica de dos métricas en términos de homeomorfismos.

Proposición 8.1.15. *Sea X conjunto. Dos métricas d y d' sobre X sobre topológicamente equivalentes si y solamente si la función $\mathrm{id}_X : (X, d) \to (X, d')$ es un homeomorfismo.*

Demostración. Por la simetría de la situación, es suficiente que probemos que

$$\tau_{X,d'} \subseteq \tau_{X,d} \iff \text{la función } \mathrm{id}_X : (X, d') \to (X, d) \text{ es continua,}$$

y esto es evidente, ya que esa función es continua exactamente cuando para todo subconjunto de X abierto con respecto a d la preimagen $U = \mathrm{id}_X^{-1}(U)$ es abierta con respeto a d'. □

Usando esta observación, podemos mostrar fácilmente que una función continua y biyectiva no es necesariamente un homeomorfismo. Veamos un ejemplo de esto.

Ejemplo 8.1.16. Sea d la métrica euclídea sobre $\mathbb{R}$ y sea d' la métrica discreta sobre $\mathbb{R}$. La función $\mathrm{id}_X : (\mathbb{R}, d') \to (\mathbb{R}, d)$ es biyectiva y continua pero no es un homeomorfismo.

El siguiente es un ejemplo bastante más interesante.

Ejemplo 8.1.17. Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, sean p y q dos elementos de $[1, +\infty]$ tales que $p < q$, sea $\Bbbk^\infty$ el subespacio vectorial de $\Bbbk^{\mathbb{N}_0}$ que consideramos en la Proposición 6.4.4.10, y sean d_p y d_p las métricas sobre $\Bbbk^\infty$ que se obtienen restringiendo a él las normas de ℓ^p y de ℓ^q, respectivamente. La Proposición 6.4.4.10 nos dice que las métricas d_p y d_q no son equivalentes. Esto significa que la función $\mathrm{id}_{\Bbbk^\infty} : (\Bbbk^\infty, d_p) \to (\Bbbk^\infty, d_q)$, que claramente es biyectiva, no es un homeomorfismo. Es, sin embargo, continua: es una restricción de la inclusión $\ell^p_\Bbbk \to \ell^q_\Bbbk$ y veremos en la Proposición 8.8.8

que esta es continua.

La siguiente proposición es extremadamente útil: nos permite acotar una función en un entorno de cualquier punto en el que es continua.

Proposición 8.1.18. *Sea X un espacio métrico, sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, sea V un $\mathbb{k}$-espacio normado, sea x un punto de X, y sea $f : X \to V$ una función que es continua en x.*

(*i*) *Hay un número positivo r tal que el subconjunto $\{\|f(x')\| : x' \in B_r(x)\}$ de $\mathbb{R}$ es acotado.*

(*ii*) *Si $f(x) \neq 0$, entonces hay un número positivo s tal que para todo elemento x' de $B_s(x)$ es $\frac{1}{2}\|f(x)\| < \|f(x')\|$.*

Demostración. (*i*) Como f es continua en x, hay un número positivo r tal que para todo $x \in X$ vale $d(x', x) < r \implies d(f(x'), f(x)) < 1$. Sea $C \coloneqq \|f(x)\| + 1$. Si x' es un elemento de $B_r(x)$, entonces $d(x', x) < r$ y, por lo tanto, $\|f(x') - f(x)\| = d(f(x'), f(x)) < 1$, de manera que es $\|f(x')\| \leq \|f(x') - f(x)\| + \|f(x)\| < C$. Vemos así que el número C es una cota superior para el conjunto $\{\|f(x')\| : x' \in B_r(x)\}$.

(*ii*) Supongamos ahora que $f(x) \neq 0$, de manera que $\frac{1}{2}\|f(x)\|$ es un número positivo. La continuidad de f en x implica entonces que hay un número positivo s tal que para todo elemento x' de $B_s(x)$ es $\|f(x') - f(x)\| < \frac{1}{2}\|f(x)\|$. Esto implica que si x' es un elemento de $B_s(x)$, entonces

$$\|f(x)\| \leq \|f(x) - f(x')\| + \|f(x')\| < \tfrac{1}{2}\|f(x)\| + \|f(x')\|$$

y, por lo tanto, que $\frac{1}{2}\|f(x)\| < \|f(x')\|$. □

El siguiente resultado conocido generalmente como el «álgebra de las funciones continuas»: nos dice que la clase de funciones continuas con valores en un espacio normado es cerrada por las operaciones algebraicas naturales.

Proposición 8.1.19. *Sea X un espacio métrico, sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, y sea V un $\mathbb{k}$-espacio normado.*

(*i*) *Si $f : X \to V$ es una función continua y x es un elemento de X, entonces hay un número positivo r tal que el subconjunto $\{\|f(x')\| : x' \in B_r(x)\}$ de $\mathbb{R}$ es acotado.*

(*ii*) *Si $f, g : X \to V$ son funciones continuas, entonces la función*

$$f + g : x \in X \mapsto f(x) + g(x) \in V$$

es continua.

(*iii*) *Si $f : X \to V$ y $\lambda : X \to \mathbb{k}$ son funciones continuas, entonces la función*

$$\lambda f : x \in X \mapsto \lambda(x) f(x) \in V$$

es continua.

(*iv*) *Si* $f : X \to \Bbbk$ *es una función continua tal que* $f(x) \neq 0$ *para todo* $x \in X$*, entonces la función*

$$1/f : x \in X \mapsto 1/f(x) \in \Bbbk$$

es continua.

(*v*) *Si* $f, g : X \to \mathbb{R}$ *son dos funciones continuas, entonces las*

$$\text{máx}\{f, g\} : x \in X \mapsto \text{máx}\{f(x), g(x)\} \in \mathbb{R}$$

y

$$\text{mín}\{f, g\} : x \in X \mapsto \text{mín}\{f(x), g(x)\} \in \mathbb{R}$$

son continuas.

(*vi*) *Si* $f : X \to V$ *es una función continua, entonces la función*

$$\|f\| : x \in X \mapsto \|f(x)\| \in \mathbb{R}$$

es continua.

Demostración. (*i*) Sea $f : X \to V$ una función continua y sea x un elemento de X. Como f es continua en x, hay un número positivo r tal que para todo $x \in X$ vale

$$d(x', x) < r \implies d(f(x'), f(x)) < 1.$$

Sea $C \coloneqq \|f(x)\| + 1$. Si x' es un elemento de $B_r(x)$, entonces $d(x', x) < r$ y, por lo tanto, $\|f(x') - f(x)\| = d(f(x'), f(x)) < 1$, de manera que es

$$\|f(x')\| \leq \|f(x') - f(x)\| + \|f(x)\| < C.$$

Vemos así que el número C es una cota superior para el conjunto $\{\|f(x')\| : x' \in B_r(x)\}$.

(*ii*) Sean $f, g : X \to V$ dos funciones continuas, sea x un elemento de X, y sea ε un número positivo. Como f y g son continuas en x, hay un número positivo δ tal que para todo elemento x' de X vale

$$d(x', x) < \delta \implies \|f(x') - f(x)\| < \frac{\varepsilon}{2} \wedge \|g(x') - g(x)\| < \frac{\varepsilon}{2},$$

y para cada elemento x' de $B_\delta(x)$ tenemos entonces que

$$\|(f+g)(x') - (f+g)(x)\| \leq \|f(x') - f(x)\| + \|g(x') - g(x)\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Esto nos dice que la función $f + g$ es continua en x.

(*iii*) Sean $f : X \to V$ y $\lambda : X \to \Bbbk$ dos funciones continuas, sea x un elemento de X, y sea ε un número positivo. Como la función f es continua en x, la primera parte de la proposición

nos dice que hay dos números positivos r y K tales que $\|f(x')\| < K$ siempre que $x' \in B_r(x)$. Por otro lado, como las funciones f, g y λ son continuas, hay un número positivo η tal que para todo elemento x' de X valen

$$d(x', x) < \eta \implies \|f(x') - f(x)\| < \frac{\varepsilon}{2(1 + |\lambda(x)|)}$$

y

$$d(x', x) < \eta \implies \|\lambda(x') - \lambda(x)\| < \frac{\varepsilon}{2K}.$$

Sea ponemos $\delta \coloneqq \text{mín}\{r, \eta\}$, entonces para cada elemento x' de $B_\delta(x)$ tenemos que

$$\begin{aligned}
\|(\lambda f)(x') - (\lambda f)(x)\| &= \|\lambda(x')f(x') - \lambda(x)f(x)\| \\
&= \|(\lambda(x') - \lambda(x))f(x') + \lambda(x)(f(x') - f(x))\| \\
&\leq |\lambda(x') - \lambda(x)| \cdot \|f(x')\| + |\lambda(x)| \cdot \|f(x') - f(x)\| \\
&\leq \frac{\varepsilon}{2K} \cdot K + |\lambda(x)| \cdot \frac{\varepsilon}{2(1 + |\lambda(x)|)} < \varepsilon.
\end{aligned}$$

Esto nos dice que la función λf es continua en x.

(*iv*) Sea $f : X \to \mathbb{k}$ una función continua tal que $f(x) \neq 0$ para todo $x \in X$, y sea x un elemento de X. Como la función f es continua en x, la Proposición 8.1.18 nos dice que hay un número positivo s tal que $|f(x')| > \frac{1}{2}|f(x)|$ para todo $x' \in B_s(x)$. Por otro lado, como la función f es es continua en x, hay un número positivo r tal que para todo $x' \in X$ vale

$$d(x', x) < r \implies |f(x') - f(x)| < \frac{\varepsilon \cdot |f(x)|^2}{2}.$$

Si ponemos $\delta \coloneqq \text{mín}\{r, s\}$, entonces para todo elemento x' de $B_\delta(x)$ tenemos que

$$\left|\frac{1}{f(x')} - \frac{1}{f(x)}\right| = \frac{|f(x) - f(x')|}{|f(x)||f(x')|} < \varepsilon.$$

Vemos así que la función $1/f$ es continua en x.

(*v*) Sean $f, g : X \to \mathbb{R}$ dos funciones continuas, sea $h : x \in X \mapsto \text{máx}\{f(x), g(x)\} \in \mathbb{R}$ la función que en el enunciado de la proposición escribimos $\text{máx}\{f, g\}$, sea x un elemento de X, y sea ε un número positivo. Como las funciones f y h son continuas, hay un número positivo δ tal que para todo elemento y de X vale

$$d(y, x) < \delta \implies |f(y) - f(x)| < \varepsilon \wedge |g(y) - g(x)| < \varepsilon.$$

Sea y un elemento de X tal que $d(y, x) < \delta$. Es $|f(y) - f(x)| < \varepsilon$ y $|g(y) - g(x)| < \varepsilon$, así que, en particular, $f(y) < f(x) + \varepsilon$ y $g(y) < g(x) + \varepsilon$ y, por lo tanto,

$$h(y) = \text{máx}\{f(y), g(y)\} < \text{máx}\{f(x) + \varepsilon, g(x) + \varepsilon\} < \text{máx}\{f(x), g(x)\} + \varepsilon = h(x) + \varepsilon.$$

De manera similar, es $f(x) - \varepsilon < f(y)$ y $g(x) - \varepsilon < g(y)$, as así que

$$h(x) - \varepsilon = \max\{f(x), g(x)\} - \varepsilon = \max\{f(x) - \varepsilon, g(x) - \varepsilon\} < \max\{f(y), g(y)\} = h(y).$$

Estas dos desigualdades nos dicen que $|h(y) - h(x)| < \varepsilon$, y todo esto nos permite concluir que la función h es continua en x.

La parte (*iii*) de la proposición nos dice que las funciones $-f$ y $-g$ son continuas en X, así que lo que ya probamos implica que $\max\{-f, -g\}$ es continua y, otra vez por la parte (*iii*), que $-\max\{-f, -g\}$ es continua. Esta función coincide con la función $\min\{f, g\}$, así que esta también es continua.

(*vi*) Sea $f : X \to V$ una función continua, sea $\|f\| : x \in X \mapsto \|f(x)\| \in \mathbb{R}$, sea x un elemento de X, y sea ε un número positivo. Como f es continua, hay un número positivo δ tal que para todo elemento y de X vale $d(y, x) < \delta \implies \|f(y) - f(x)\| < \varepsilon$. Sea y un elemento de X tal que $d(y, x) < \delta$. Tenemos entonces que $\|f(y)\| - \|f(x)\| \leq \|f(y) - f(x)\| < \varepsilon$ y que $\|f(x)\| - \|f(y)\| \leq \|f(x) - f(y)\| < \varepsilon$, de manera que $\big|\|y\| - \|x\|\big| < \varepsilon$. Esto nos permite concluir que la función $\|f\|$ es continua. □

8.2. Continuidad uniforme

Que una función $f : X \to Y$ entre dos espacios métricos sea continua significa que para cada número positivo ε y cada punto x de X podemos encontrar un número positivo $\delta(x, \varepsilon)$ tal que $f(B_{\delta(x,\varepsilon)}(x)) \subseteq B_\varepsilon(f(x))$. Es claro que en general la elección de $\delta(x, \varepsilon)$ no puede sino depender de ε. En general también esa elección va a depender del punto x.

Ejemplo 8.2.1. Consideremos la función $f : t \in \mathbb{R} \mapsto t^3 \in \mathbb{R}$ y sea ε un número positivo. Para cada $x \in \mathbb{R}$ y cada número positivo δ que la función f sea estrictamente creciente implica que $f(B_\delta(x)) \subseteq B_\varepsilon(f(x))$ si y solamente si $x^3 - \varepsilon \leq (x - \delta)^3$ y $(x + \delta)^3 \leq x^3 + \varepsilon$, y esto ocurre si y solamente si $\delta \leq \min\{x - (x^3 - \varepsilon)^{1/3}, (x^3 + \varepsilon)^{1/3} - x\}$. Esto nos dice que, en la situación de arriba, podemos elegir como $\delta(x, \varepsilon)$ a cualquier número positivo menor que

$$\delta_{\max}(x, \varepsilon) := \min\{x - (x^3 - \varepsilon)^{1/3}, (x^3 + \varepsilon)^{1/3} - x\}.$$

Es fácil ver que

$$\lim_{x \to \infty} \delta_{\max}(x, \varepsilon) = 0,$$

y esto implica que no podemos elegir $\delta(x, \varepsilon)$ independiente de x.

Que sí sea posible elegir $\delta(x, \varepsilon)$ independiente del punto x suele ser un dato importante sobre la función f.

Definición 8.2.2. Sean X e Y dos espacios métricos. Una función $f : X \to Y$ es ***uniformemente continua*** si para todo número positivo ε hay un número positivo δ tal que para toda elección de x e y en X vale $d(x, y) < \delta \implies d(f(x), f(y)) < \varepsilon$.

Esta condición es más fuerte que la de continuidad.

Proposición 8.2.3. *Una función uniformemente continua es continua.*

Demostración. Sean X e Y dos espacios métricos y sea $f : X \to Y$ una función uniformemente continua. Sea x_0 un elemento de X y sea ε un número positivo. Como la función f es uniformemente continua, hay un número positivo δ tal que para toda elección de x y de x' en X vale $d(x, x') < \delta \implies d(f(x), f(x')) < \varepsilon$. En particular, si x es un elemento de X tal que $d(x, x_0) < \delta$, entonces es $d(f(x), f(x_0) < \varepsilon$, y esto nos dice que la función f es continua en x_0. Esto es así cualquiera sea el elemento x_0 de X, así que la función f es continua. □

Hay varias condiciones útiles que implican la continuidad uniforme. Veamos un par de ellas.

Definición 8.2.4. Sean X e Y dos espacios métricos.

(*i*) Decimos que una función $f : X \to Y$ satisface la ***condición de Lipschitz*** si hay un número real no negativo C tal que para toda elección de x y x' en X vale

$$d(f(x), f(x')) \leq C \cdot d(x, x').$$

El número C es entonces una ***constante de Lipschitz*** para f.

(*ii*) Decimos que una función $f : X \to Y$ satisface la ***condición de Hölder*** si hay números reales no negativos C y α tales que para toda elección de x y x' en X vale

$$d(f(x), f(x, x')) \leq C \cdot d(x, x')^{\alpha}.$$

El número X es entonces un ***exponente de Hölder*** para f.

Los nombres de estas condiciones recuerdan a Rudolf Lipschitz y a Otto Hölder.

Proposición 8.2.5. *Sean X e Y dos espacios métricos.*

(*i*) *Una función $f : X \to Y$ que satisface la condición de Lipschitz es uniformemente continua.*

(*ii*) *Una función $f : X \to Y$ que satisface la condición de Hölder es uniformemente continua.*

Demostración. (*i*) Sea $f : X \to Y$ una función que satisface la condición de Lipschitz, y sea C una constante de Lipschitz para f. Si ε es un número positivo, entonces siempre que x y x' son elementos de X tales que $d(x, x') < \varepsilon/C$ tenemos que $d(f(x), f(x')) \leq C \cdot d(x, x') < \varepsilon$. Esto nos dice que la función f es uniformemente continua.

(*ii*) Sea ahora $f : X \to Y$ una función que satisface la condición de Hölder, y sean C y α dos números no negativos tales que $d(f(x), f(x')) \leq C \cdot d(x, x')^{\alpha}$. Sea ε un número positivo y pongamos $\delta \coloneqq \varepsilon^{1/\alpha}/C$. Si x y x' son dos elementos de X, entonces

$$d(f(x), f(x')) \leq C \cdot d(x, x')^{\alpha} < C \cdot \delta^{\alpha} = \varepsilon.$$

Vemos así que la función f es uniformemente. □

Notemos que la condición de Lipschitz es más fuerte que la de Hölder.

Ejemplo 8.2.6. Mostremos que la función $f : t \in [0,1] \mapsto \sqrt{t} \in \mathbb{R}$ satisface la condición de Hölder pero no la de Lipschitz. Para ver que satisface la condición de Hölder, basta observar que si x y x' son dos elementos distintos de $[0,1]$, entonces $0 < \sqrt{|x - x'|} \leq \left|\sqrt{x} + \sqrt{x'}\right|$, ya que

$$|x - x'| \leq |x| + |x'| = x + x' \leq x + x' + 2\sqrt{xx'} = \left|\sqrt{x} + \sqrt{x'}\right|^2,$$

y, por lo tanto,

$$|f(x) - f(x')| = \left|\sqrt{x} - \sqrt{x'}\right| = \frac{|x - x'|}{\left|\sqrt{x} + \sqrt{x'}\right|} = \frac{\sqrt{|x - x'|}}{\left|\sqrt{x} + \sqrt{x'}\right|} \cdot |x - x'|^{1/2} \leq |x - x'|^{1/2}.$$

Para ver que la función no satisface la condición de Lipschitz, supongamos, para llegar a una contradicción, que hay un número no negativo C tal que $|f(x) - f(x')| \leq C \cdot |x - x'|$ siempre que x y x' son elementos de $[0,1]$. El número $1/(1 + C)^2$ pertenece a $[0,1]$, y entonces tenemos que

$$\frac{1}{1 + C} = \left|\sqrt{\frac{1}{(1 + C)^2}} - \sqrt{0}\right| \leq C \cdot \left|\frac{1}{(1 + C)^2} - 0\right| = \frac{C}{(1 + C)^2},$$

de manera que $1 + C \leq C$, lo que es absurdo.

Decir que una función satisface la condición de Hölder con exponente 0 es lo mismo que decir que es acotada, y que satisface la condición de Hölder con exponente 1 es lo mismo que decir que satisface la condición de Lipschitz. El siguiente resultado describe que las funciones que satisfacen la condición de Hölder con exponente mayor que 1 no son, en general, interesantes.

Proposición 8.2.7. *Sean X e Y dos espacios métricos, y supongamos que*

> *siempre que x y x' son dos puntos de X y n es un entero positivo es posible elegir $n+1$ puntos $x_0, x_1, \ldots, x_n$ en X de manera que sea $x_0 = x$, $x_n = x'$ y $d(x_{i-1}, x_i) < 1/n$ para cada $i \in [\![n]\!]$.*

Una función $X \to Y$ que satisface la condición de Hölder con exponente mayor que 1 es constante.

Notemos que la hipótesis sobre X se satisface si X es $\mathbb{R}^n$, y que no se satisface si X tiene algún punto aislado.

Demostración. Sea $f : X \to Y$ una función, sean C y α dos números no negativos, y supongamos que $d(f(x), f(x')) \leq C \cdot d(x, x')^\alpha$ para cada elección de x y de x' en X y que $\alpha > 1$.

Sean x y x' dos puntos de X y sea n un entero positivo. La hipótesis implica que hay $n+1$ puntos $x_0, x_1, \ldots, x_n$ en X tales que $x_0 = x$, $x_n = x'$ y $d(x_{i-1}, x_i) < 1/n$ para cada $i \in [\![n]\!]$, y, por lo tanto, que

$$d(f(x), f(x')) \leq \sum_{i=1}^{n} d(f(x_{i-1}), f(x_i)) \leq \sum_{i=1}^{n} C \cdot d(x_{i-1}, x_i)^\alpha \leq \sum_{i=1}^{n} C \cdot \frac{1}{n^\alpha} = C \cdot n^{1-\alpha}.$$

Esto es cierto cualquiera sea el entero positivo n, y sabemos que $\lim_{n\to\infty} n^{1-\alpha} = 0$ porque $\alpha > 1$, así que podemos concluir que $d(f(x), f(x')) = 0$, esto es, que $f(x) = f(x')$. Vemos así que la función f es constante. □

Veamos algunos ejemplos importantes de funciones uniformemente continuas.

Ejemplo 8.2.8. Sea X un espacio métrico y sea x_0 un punto de X. Veamos que

> *la función $f : x \in X \mapsto d(x, x_0) \in \mathbb{R}$ es uniformemente continua.*

En efecto, si x y x' son dos puntos de X, entonces la desigualdad triangular nos dice que $d(x, x_0) - d(x', x_0) \leq d(x, x')$ y que $d(x', x_0) - d(x, x_0) \leq d(x, x')$, así que

$$\left|d(x, x_0) - d(x', x_0)\right| \leq d(x, x').$$

Vemos así que la función f satisface la condición de Lipschitz con constante de Lipschitz 1 y que, en particular, es uniformemente continua.

Ejemplo 8.2.9. Sea X un espacio métrico. Afirmamos que

> *la métrica $d : X \times X \to \mathbb{R}$ es uniformemente continua*

cuando dotamos a $X \times X$ de su métrica producto. Sean (x, x') e (y, y') dos puntos de $X \times X$.

Es $d(x,x') \leq d(x,y) + d(y,y') + d(y',x')$, así que $d(x,x') - d(y,y') \leq d(x,y) + d(y',x')$ y, por simetría, también $d(y,y') - d(x,x') \leq d(y,x) + d(x',y')$. Como la métrica d es simétrica, podemos concluir de esto que

$$\big|d(x,x') - d(y,y')\big| \leq d(x,y) + d(y',x') \leq 2 \cdot d\big((x,x'),(y,y')\big).$$

La función $d : X \times X \to \mathbb{R}$ satisface entonces la condición de Lipschitz con constante 2 y es, en consecuencia, uniformemente continua.

Ejercicio 8.2.10. Sea X un espacio métrico y sea S un subconjunto no vació de X. Pruebe que la función $f : x \in X \mapsto d(x,S) \in \mathbb{R}$ es uniformemente continua.

Ejemplo 8.2.11. Si $\Bbbk$ es o $\mathbb{R}$ o $\mathbb{C}$ y es V un $\Bbbk$-espacio normado, entonces

la norma $\|-\| : V \to \mathbb{R}$ es una función uniformemente continua.

En efecto, para todo elemento x de V es $\|x\| = d(x,0)$, así que esto es consecuencia del resultado del Ejemplo 8.2.8.

Ejemplo 8.2.12. Sea p un número primo, consideremos sobre $\mathbb{Q}$ la métrica p-ádica d_p, y sea $|-|_p : \mathbb{Q} \to \mathbb{Z}$ el valor absoluto p-ádico del Lema 5.2.9.10. Si r y s son elementos de $\mathbb{Q}$, entonces $|r|_p \leq |s|_p + |r-s|_p$ y $|s|_p \leq |r|_p + |s-r|_p$: como además $|s-r|_p = |r-s|_p$, esto nos permite concluir que

$$\big||r|_p - |s|_p\big| \leq |r-s|_p = d_p(r,s).$$

Esto nos dice que la función $|-|_p$ satisface la condición de Lipschitz y, en particular, que es uniformemente continua.

8.3. Grupos y espacios vectoriales métricos

Definición 8.3.1. Un ***grupo métrico*** es un par (G,d) en el que G es un grupo y $d : G \times G \to \mathbb{R}$ es una métrica sobre G con respecto a la que las funciones

$$m : (x,y) \in G \times G \mapsto xy \in G, \qquad i : x \in G \mapsto x^{-1} \in G$$

son continuas.

Proposición 8.3.2. *Sea G un grupo. Si $d : G \times G \to \mathbb{R}$ es una métrica sobre G que es invariante por translaciones a izquierda y a derecha, entonces (G, d) es un grupo métrico y, de hecho, las funciones*

$$m : (x, y) \in G \times G \mapsto xy \in G, \qquad i : x \in G \mapsto x^{-1} \in G$$

son uniformemente continuas con respecto a la métrica d.

Demostración. Sea $d : G \times G \to \mathbb{R}$ una métrica sobre G que es invariante por translaciones. Si x, x', y e y' son elementos de G, entonces la invariancia de d implica que

$$\begin{aligned} d(m(x, y), m(x', y')) = d(xy, x'y') &= d(x'^{-1}(xy)y^{-1}, x'^{-1}(x'y')y^{-1}) = d(x'^{-1}x, y'y^{-1}) \\ &\leq d(x'^{-1}x, e) + d(e, y'y^{-1}) = d(x, x') + d(y, y') \\ &\leq 2 \cdot \max\{d(x, x'), d(y, y')\} = 2 \cdot d((x, y), (x', y')). \end{aligned}$$

Esto nos dice que la función m satisface la condición de Lipschitz, así que es uniformemente continua. Por otro lado, si x e y son dos elementos de G, entonces

$$d(i(x), i(y)) = d(x^{-1}, y^{-1}) = d(xx^{-1}y, xy^{-1}y) = d(y, x) = d(x, y).$$

Esto nos dice que la función i es una isometría y, en particular, que es uniformemente continua. □

La métrica de un grupo métrico no tiene por qué ser invariante por translaciones, pero siempre podemos encontrar otra topológicamente equivalente a ella que sea invariante por translaciones a izquierda.

Proposición 8.3.3. *Sea (G, d) un grupo métrico. Hay una métrica $d' : G \times G \to \mathbb{R}$ que es invariante por translaciones y topológicamente equivalente a d.*

Demostración. Digamos que un entorno U de e es ***simétrico*** si $i(U) = U$. Notemos que para ello es suficiente que se tenga $i(U) \subseteq U$, ya que en ese caso tenemos también que $U = i(i(U)) \subseteq i(U)$. Afirmamos que

todo entorno de e contiene un entorno simétrico de e.

En efecto, sea U un entorno de e. Como la función i es continua y $i(e) = e$, el conjunto $i^{-1}(U)$ es un entorno de e y, por lo tanto, también lo es la intersección $V \coloneqq U \cap i^{-1}(U)$. Si x es un elemento de V, entonces $i(x) \in U$ e $i(i(x)) = x \in U$, así que $i(x) \in i^{-1}(U)$: vemos así que $i(x) \in V$ y, por lo tanto, que el entorno V es simétrico.

Primer paso. Sea $\mathscr{S}$ el conjunto de todos los pares ordenados (n, U) con $n \in \mathbb{N}_0$ y U un entorno simétrico de e en G tal que $U \subseteq B_{1/(n+1)}(e)$.

- Como $B_1(e)$ es un entorno de e en G, hay un entorno simétrico U_0 de e en G tal que $U_0 \subseteq B_{1/n}(e)$ y el par $(0, U_0)$ es un elemento de $\mathscr{S}$.
- Sea (n, U) un elemento de $\mathscr{S}$. Afirmamos que el conjunto

$$R(n, U) \coloneqq \{(m, V) \in \mathscr{S} : m > n, V \cdot V \subseteq U\}$$

no es vacío. En efecto, como la función m es continua, U es un entorno de e, y $m(e, e) = e$, hay un entorno W de (e, e) en $G \times G$ tal que $m(W) \subseteq U$. Hay entonces un número entero m mayor que n tal que $B_{1/m}(e) \times B_{1/m}(e) \subseteq W$ y, por lo que hicimos arriba, hay un entorno simétrico V de e tal que $V \subseteq B_{1/(m+1)}(e)$. El par (m, V) es entonces un elemento de $\mathscr{S}$ y, más aún, pertenece a $R(n, U)$. Esto prueba nuestra afirmación.

De acuerdo *Principio de elección dependiente* en la forma que le da el Ejercicio 2.1.7 vemos con esto que hay una función $f : \mathbb{N}_0 \to \mathscr{S}$ tal que $f(0) = (0, U_0)$ y $f(i+1) \in R(f(i))$ para todo $i \in \mathbb{N}_0$. Si para cada $i \in \mathbb{N}_0$ escribimos (n_i, U_i) al par ordenado $f(i)$, entonces la sucesión $(n_i)_{i \geq 0}$ es estrictamente creciente y para cada $i \in \mathbb{N}_0$ tenemos que

$$U_{i+1} \subseteq U_{i+1} \cdot U_{i+1} \subseteq U_i \subseteq B_{1/(n_i+1)}(e).$$

Notemos que, en particular, esto nos dice que

$$\bigcap_{i=0}^{\infty} U_i \subseteq \bigcap_{i=0}^{\infty} B_{1/(n_i+1)}(e) = \{e\}.$$

Segundo paso. Consideremos el conjunto $D \coloneqq \{a/2^n : n \in \mathbb{N}, a \in [\![2^n - 1]\!]\}$.

Sea q un elemento de D. Hay exactamente una forma de elegir un entero positivo n y un elemento impar a de $[\![2^n - 1]\!]$ de manera que sea $q = a/2^n$, y la teoría del desarrollo en base 2 implica que hay un único subconjunto no vacío $P(q)$ de $[\![0, n-1]\!]$ tal que $a/2^n = \sum_{i \in P(q)} 2^{-i}$. Si $a_1, a_2, \ldots, a_k$ son los elementos de $P(q)$ listados en orden estrictamente decreciente, ponemos

$$V_q \coloneqq U_{a_1} \cdot U_{a_2} \cdot \cdots \cdot U_{a_k}.$$

De esta manera obtenemos para cada elemento q del conjunto D un entorno abierto V_q de e en G. Queremos probar ahora que

si q y r son dos elementos de D y $q < r$, entonces $V_q \subseteq V_r$.

□

8.4. Problemas de extensión

Sean X e Y dos espacios métricos y sea Z un subespacio de X. Sabemos que si $f : X \to Y$ es una función continua, entonces la restricción $f|_Z : Z \to Y$ también es continua, y esto nos dice que hay una función

$$\mathrm{res}_Z : f \in C(X,Y) \mapsto f|_Z \in C(Z,Y).$$

En general esta función no es ni sobreyectiva ni inyectiva, y nuestro problema en esta sección es tratar de encontrar hipótesis bajo las cuales podamos garantizar que sí lo son.

Ejemplo 8.4.1. Sean ahora $X = \mathbb{R}$, $Z = [0, +\infty)$, e $Y = \mathbb{R}$, y consideremos las funciones

$$f : t \in X \mapsto t \in \mathbb{R}, \qquad g : t \in X \mapsto |t| \in \mathbb{R}.$$

Claramente $f|_Z = g_Z$ pero $f \neq g$. Esto nos dice que la función $\mathrm{res}_Z : C(X,\mathbb{R}) \to C(Z,\mathbb{R})$ no es inyectiva.

Ejemplo 8.4.2. Sean $X \coloneqq [-1,1]$, $Z \coloneqq [-1,0) \cup (0,1]$ e $Y \coloneqq \mathbb{R}$. La función $f : t \in Z \mapsto t/|t| \in \mathbb{R}$ es continua, pero no hay ninguna función $g : X \to \mathbb{R}$ tal que $g|_Z = f$. En efecto, supongamos que la hay. Como $(1/n)_{n\geq 1}$ y $(-1/n)_{n\geq 1}$ son sucesiones en X que convergen a 0, la continuidad de g implica que las sucesiones $(g(1/n))_{n\geq 1}$ y $(g(-1/n))_{n\geq 1}$ convergen a $g(0)$ y, en particular, tienen límites iguales. Ahora bien, para todo $n \in \mathbb{N}$ los números $1/n$ y $-1/n$ están en Z, así que $g(1/n) = g|_Z(1/n) = f(1/n) = 1$ y $g(-1/n) = g|_Z(-1/n) = f(-1/n) = -1$: las sucesiones $(1/n)_{n\geq 1}$ y $(-1/n)_{n\geq 1}$ son entonces constantes de valor 1 y -1, respectivamente, y, por lo tanto, sus límites no son iguales. Esto es absurdo. Vemos así que la función $\mathrm{res}_Z : C(X,\mathbb{R}) \to C(Z,\mathbb{R})$ no es sobreyectiva.

Ocupémonos primero de cómo garantizar que la función de restricción sea inyectiva.

Proposición 8.4.3. *Sean X e Y dos espacios métricos y sea Z un subconjunto de X. Si $f, g : X \to Y$ son dos funciones continuas y $f|_Z = g|_Z$, entonces $f|_{\overline{Z}} = f|_{\overline{Z}}$.*

Daremos dos demostraciones distintas de esta proposición. En la primera usaremos la caracterización topológica de la continuidad que nos da la Proposición 8.1.5.

Demostración. Sean $f, g : X \to Y$ dos funciones continuas y supongamos que $f|_Z = g|_Z$. Sabemos del Ejercicio 6.5.6 que el subconjunto $\Delta_Y \coloneqq \{(y,y) \in Y \times Y : y \in Y\}$ de $Y \times Y$ es cerrado, y del Ejercicio 8.1.12 que la función $h : x \in X \mapsto (f(x), g(x)) \in Y \times Y$ es continua, así que el conjunto $h^{-1}(\Delta_Y)$ es cerrado en X. La hipótesis nos dice que $Z \subseteq h^{-1}(\Delta_Y)$, así que que este último conjunto sea cerrado implica que también $\overline{Z} \subseteq h^{-1}(\Delta_Y)$. Esto nos dice que $(f(x), g(x)) = h(x) \in \Delta_Y$, esto es,

que $f(x) = g(x)$ para todo $x \in \overline{Z}$. □

En la segunda demostración usamos la caracterización de la continuidad en términos de sucesiones.

Demostración. Sean $f, g : X \to Y$ dos funciones continuas, supongamos que $f|_Z = g|_Z$, y sea x un elemento de $\overline{Z}$, de manera que hay una sucesión $(x_n)_{n\geq 0}$ con valores en Z que converge a x. Como las función f y g son continuas, sabemos que las sucesiones $(f(x_n))_{n\geq 0}$ y $(g(x_n))_{n\geq 0}$ convergen a $f(x)$ y $g(x)$, respectivamente. Por otro lado,la hipótesis implica que $f(x_n) = g(x_n)$ para todo $n \in \mathbb{N}_0$, así que esas dos sucesión coinciden. Vemos así que es $f(x) = g(x)$. Como esto es así cualquiera sea el elemento x de $\overline{Z}$, podemos concluir que $f|_{\overline{Z}} = g_{\overline{Z}}$. □

El resultado de la Proposición 8.4.3 es el mejor posible, como muestra el siguiente ejemplo.

Ejemplo 8.4.4. Sea X un espacio métrico, sea Z un subconjunto cerrado de X tal que $Z \subsetneq X$, y consideremos las funciones $f : x \in X \mapsto 0 \in \mathbb{R}$ y $g : x \in X \mapsto d(x, Z) \in \mathbb{R}$, que son continuas. Es claro que $f|_Z = g|_Z$ y, como Z es cerrado, que $f(x) \neq g(x)$ para todo elemento x de $X \setminus Z$.

Un caso particular de la Proposición 8.4.3 que usamos frecuentemente es el siguiente.

Corolario 8.4.5. *Sean X e Y dos espacios métricos y sean $f, g : X \to Y$ dos funciones continuas. Si hay un subconjunto denso D de X tal que $f|_D = g|_D$, entonces $f = g$.*

Demostración. Esto es consecuencia inmediata de la proposición, ya que $\overline{D} = X$. □

Juntando todo, obtenemos una posible solución a uno de nuestros problemas.

Corolario 8.4.6. *Sea X un espacio métrico y sea D un subconjunto de X. La función*

$$f \in C(X, Y) \mapsto f|_D \in C(D, Y)$$

es inyectiva para todo espacio métrico Y si y solamente si el subconjunto D es denso en X.

Demostración. Que la condición es suficiente es precisamente lo que afirma el Corolario 8.4.5. Supongamos que para todo espacio métrico Y la función $f \in C(X, Y) \mapsto f|_D \in C(D, Y)$ es inyectiva, de manera que, en particular, la función $f \in C(X, \mathbb{R}) \mapsto f|_D \in C(D, \mathbb{R})$ es inyectiva. Las funciones $x \in X \mapsto 0 \in \mathbb{R}$ y $x \in X \mapsto d(x, D) \in \mathbb{R}$ son continuas y tiene la misma restricción a D — la función nula — así que son entonces iguales. Esto significa que para todo elemento x de X es $d(x, D) = 0$, así que $x \in \overline{D}$: vemos así que $\overline{D} = X$ y, por lo tanto, que D es un subconjunto denso de X. □

El siguiente resultado es una consecuencia sencilla pero importante del Corolario 8.4.5.

Proposición 8.4.7. *Sean X e Y dos espacios métricos. Una función continua $f : X \to Y$ tiene imagen densa en Y si y solamente si satisface la siguiente condición:*

> *si Z es un espacio métrico Z y $u, v : Y \to Z$ son dos funciones continuas tales que $u \circ f = v \circ f$, entonces $u = v$.*

En el lenguaje de la teoría de categorías, esta proposición nos dice que una función continua $f : X \to Y$ es un ***epimorfismo*** si y solamente si su imagen es densa en Y.

Demostración. Supongamos primero que $f : X \to Y$ es una función continua con imagen densa en Y. Si Z es un espacio métrico y $u, v : Y \to Z$ son dos funciones continuas tales que $u \circ f = v \circ f$, entonces $u|_{f(X)} = v|_{f(X)}$ y, como $f(X)$ es un subconjunto denso de Y, el Corolario 8.4.5 nos permite concluir que $u = v$.

Supongamos ahora que $f : X \to Y$ es una función continua que satisface la condición del enunciado, y consideremos las funciones $u : y \in Y \mapsto 0 \in \mathbb{R}$ y $v : y \in Y \mapsto d(y, f(X)) \in \mathbb{R}$. Las dos son continuas, y es claro que $u \circ f = v \circ f$, así que la hipótesis nos permite concluir que $u = v$, esto es, que $d(y, f(X)) = 0$ para todo $y \in Y$. Como sabemos, esto implica que $\overline{f(X)} = Y$, esto es, que la imagen $f(X)$ es densa en Y. □

Consideremos ahora el problema de la sobreyectividad de la función

$$\mathrm{res}_Z : f \in C(X, Y) \mapsto f|_Z \in C(Z, Y),$$

al que podemos plantear de como el siguiente ***problema de extensión***: dados dos espacios métricos X e Y y un subconjunto Z de X, y

> *dada una función continua $g : Z \to Y$, encontrar, si esto es posible, una función continua $f : X \to Y$ tal que $f|_Z = g$.*

Es fácil dar una condición necesaria y suficiente para que esto sea posible.

Proposición 8.4.8. *Sean X e Y dos espacios métricos, sea Z un subconjunto de X y sea $g : Z \to Y$ una función continua. Hay una función continua $f : \overline{Z} \to Y$ tal que $f|_Z = g$ si y solamente si para toda sucesión $(x_n)_{n\geq 0}$ en Z que converge en X la sucesión $(g(x_n))_{n\geq 0}$ en Y converge.*

Cuando una tal función existe, está bien determinada: esto es consecuencia del Corolario 8.4.5.

Demostración. Supongamos que hay una función continua $f : \overline{Z} \to Y$ tal que $f|_Z = g$ y sea $(x_n)_{n\geq 0}$ una sucesión en Z que converge en X a un punto ξ. Este punto ξ pertenece entonces a $\overline{Z}$ y, como la función f es continua, la sucesión $(f(x_n))_{n\geq 0}$ converge a $f(\xi)$ y, en particular, converge. Esto muestra que la condición de la proposición es necesaria.

Supongamos ahora que la condición de la proposición se satisface, y mostremos que

$$\begin{aligned}&\textit{si } (x_n)_{n\geq 0} \textit{ y } (x'_n)_{n\geq} \textit{ son dos sucesiones en } Z \textit{ que convergen en } X \textit{ al mismo punto,}\\ &\textit{entonces las sucesiones } (f(x_n))_{n\geq 0} \textit{ y } (f(x'_n))_{n\geq 0} \textit{ convergen en } Y \textit{ a mismo punto.}\end{aligned} \tag{8.2}$$

Sean para ello $x = (x_n)_{n\geq 0}$ y $x' = (x'_n)_{n\geq 0}$ dos sucesiones en Z que convergen en X al mismo punto ξ de X, y consideremos la sucesión $y = (y_n)_{n\geq 0}$ en Z que tiene para cada $n \in \mathbb{N}_0$

$$y_n = \begin{cases} x_n & \text{si } n \text{ es par;} \\ x'_n & \text{si } n \text{ es impar.} \end{cases}$$

Como las subsucesiones $(x_{2n})_{n\geq 0}$ y $(x'_{2n+1})_{n\geq 0}$ de x y de x' convergen a ξ, el resultado del Ejercicio 7.2.4 nos permite concluir que la sucesión y converge y la hipótesis entonces que también lo hace la sucesión $(f(y_n))_{n\geq 0}$. Sea ζ el límite de esta última. Todo esto implica, claro, que las subsucesiones $(f(x_{2n}))_{n\geq 0} = (f(y_{2n}))_{n\geq 0}$ y $(f(x_{2n+1}))_{n\geq 0} = (f(y_{2n+1}))_{n\geq 0}$ de $(f(y_n))_{n\geq 0}$ también convergen a ζ. Esto implica que las sucesiones $(f(x_n))_{n\geq 0}$ y $(f(x'_n))_{n\geq 0}$, que convergen, tienen el mismo límite ζ. Vemos así que la afirmación (8.2) es cierta.

Sea ahora $\mathscr{C}$ el conjunto de todas las sucesiones en Z que convergen en X, y cada vez que $c = (x_n)_{n\geq 0}$ es un elemento de $\mathscr{C}$ escribamos $g(c)$ a la sucesión $(g(x_n))_{n\geq 0}$. De acuerdo a la hipótesis podemos considerar el conjunto

$$f \coloneqq \{(\lim c, \lim g(c)) \in \overline{Z} \times Y : c \in \mathscr{C}\}.$$

Queremos probar que se trata de una función $\overline{Z} \to Y$.

- Si ξ es un elemento de $\overline{Z}$, entonces hay una sucesión $c = (x_n)_{n\geq 0}$ en Z que converge a ξ: es $c \in \mathscr{C}$, así que $(\xi, \lim g(x)) \in f$.
- Sean ξ un elemento de $\overline{Z}$ y sean y e y' dos elementos de Y tales que (ξ, y), $(\xi, y') \in f$, de manera que hay dos sucesiones c y c' en $\mathscr{C}$ tales que $\lim c = \xi$, $\lim g(c) = y$, $\lim c' = \xi$ y $\lim g(c') = y'$. La afirmación (8.2) nos dice que entonces $y = y'$.

Veamos que $f|_Z = g$.

- Si z es un elemento de Z, entonces la sucesión $c = (x_n)_{n\geq 0}$ que tiene $x_n = z$ para todo $n \in \mathbb{N}_0$ pertenece a $\mathscr{C}$ y es $(z, g(z)) = (\lim c, \lim g(c)) \in f$, de manera que $f(z) = g(z)$.

Para terminar, tenemos que probar que la función $f : \overline{Z} \to Y$ es continua.

Sea x un punto de $\overline{Z}$ y sea U un subconjunto abierto de Y que contiene a $f(x)$. Para probar que la función f es continua en el punto x de $\overline{Z}$ es suficiente que mostremos que $f^{-1}(U)$ es un entorno de x. De acuerdo a la Proposición 6.5.5 hay un entorno cerrado V de $f(x)$ en Y tal que $V \subseteq U$. Mostremos que

$$\textit{hay un entorno abierto } M \textit{ de } x \textit{ en } X \textit{ tal que } g(M \cap Z) \subseteq V. \tag{8.3}$$

Supongamos, para llegar a una contradicción, que esto no es así, de manera que para todo $n \in \mathbb{N}_0$ hay un punto x'_n en $B^X_{1/(n+1)}(x) \cap Z$ tal que $g(x'_n) \notin V$. La sucesión $(x'_n)_{n\geq 0}$ converge a x. Como $(x_n)_{n\geq 0}$ también converge a x, la sucesión $(x''_n)_{n\geq 0}$ que para cada $n \in \mathbb{N}_0$ tiene

$$x''_n = \begin{cases} x_n & \text{si } n \text{ es par;} \\ x'_n & \text{si } n \text{ es impar} \end{cases}$$

también converge a x. La hipótesis nos dice entonces que la sucesión $(g(x''_n))_{n\geq 0}$ converge en Y a un punto y. Esto implica que su subsucesión $(g(x_{2n}))_{n\geq 0}$ converge a y y, como esta también es una subsucesión de $(g(x_n))_{n\geq 0}$, que $y = f(x)$. Vemos así que la sucesión $(g(x'_{2n+1}))_{n\geq 0}$ converge a $f(x)$, así que la sucesión $(g(x'_n))_{n\geq 0}$, que converge, converge a $f(x)$. Esto es absurdo, porque tenemos que $g(x'_n) \notin V$ para todo $n \in \mathbb{N}_0$ y V es un entorno de $f(x)$. Esto prueba (8.3).

Sea ahora z un elemento de $M \cap \overline{Z}$. Hay una sucesión $(z_n)_{n\geq 0}$ en Z que converge a z en X y, como M es un entorno de z en X, esta sucesión está a la larga en M. La sucesión $(g(z_n))_{n\geq 0}$ converge a $f(z)$ y,de acuerdo a (8.3), está en V, que es cerrado, así que $f(z) \in V \subseteq U$ y $z \in f^{-1}(U)$. Esto prueba que el entorno $M \cap \overline{Z}$ de x en $\overline{Z}$ está contenido en $f^{-1}(U)$ y, en definitiva, que $f^{-1}(U)$ es un entorno de x, como queremos. □

La Proposición 8.4.8 nos da una condición necesaria y suficiente para que una función continua definida en un subconjunto de un espacio métrico pueda extenderse a una función continua sobre la clausura de ese subconjunto, pero esa condición no es fácil de verificar directamente. El siguiente resultado da una condición más fuerte que la implica.

Proposición 8.4.9. *Sean X e Y dos espacios métricos, sea Z un subconjunto de X, y sea $g : Z \to Y$ una función continua. Si g es uniformemente continua y el espacio Y es completo, entonces hay una y solo una función continua $f : \overline{Z} \to Y$ tal que $f|_Z = g$, y esta función f es uniformemente continua.*

Demostración. Supongamos que la función g es uniformemente continua y que el espacio Y es completo y mostremos que entonces la condición de la Proposición 8.4.8 se cumple: esto probara la primera de las dos afirmaciones de la proposición. Sea $(x_n)_{n\geq 0}$ una sucesión en Z que converge en X y mostremos que la sucesión $(g(x_n))_{n\geq 0}$ entonces converge. Para ello, como el espacio Y es completo, es suficiente con mostrar que la sucesión $(g(x_n))_{n\geq 0}$ satisface la condición de Cauchy.

Sea ε un número positivo. Como la función $g : Z \to Y$ es uniformemente, hay un número positivo δ tal que para cada elección de x e y en Z vale $d(x, y) < \delta \implies d(f(x), f(y)) < \varepsilon$. Por otro lado, como la sucesión $(x_n)_{n\geq 0}$ converge en X, satisaface la condición de Cauchy y, por lo tanto, hay un número entero positivo n_0 tal que siempre que m y n son elementos de $\mathbb{N}_0$ vale $m, n \geq n_0 \implies d(f(x_m), f(x_m)) < \delta$. Si ahora m y n son dos elementos de $\mathbb{N}_0$ tales que $m, n \geq 0$, entonces la forma en que elegimos n_0 implica que $d(f(x_m), f(x_n)) < \delta$ y, a su vez, la forma en que elegimos δ implica que $d(f(x_m), f(x_n)) < \varepsilon$. Vemos así que la sucesión $(f(x_n))_{n\geq 0}$ satisaface

la condición de Cauchy, como queremos.

Para completar la demostración tenemos que mostrar que la función continua $f : \overline{Z} \to Y$ tal que $f|_Z = g$ que ahora sabemos existe es uniformente continua. Sea ε un número positivo. Como la función g es uniformente continua, hay un número positivo δ tal que para toda elección de z y z' en Z vale $d(z, z') < \delta \implies d(g(z), g(z')) < \varepsilon/3$. Sean u y v dos elementos de $\overline{Z}$ tales que $d(u, v) < \delta/3$. Hay sucesiones $(u_n)_{n\geq 0}$ y $(v_n)_{n\geq 0}$ en Z que convergen a u y a v en X, y las sucesiones $(g(u_n))_{n\geq 0}$ y $(g(v_n))_{n\geq 0}$ convergen en Y a $f(u)$ y a $f(v)$, respectivamente. En particular, hay un entero positivo m tal que $d(u, u_m) < \delta/3$, $d(v, v_m) < \delta/3$, $d(f(u), g(u_m)) < \varepsilon/3$ y $d(f(v), g(v_m)) < \varepsilon/3$. Como

$$d(u_m, v_m) \leq d(u_m, u) + d(u, v) + d(v, v_m) < \frac{\delta}{3} + \frac{\delta}{3} + \frac{\delta}{3} = \delta,$$

así que la forma en que elegimos a δ implica que $d(g(u_m), g(v_m)) < \varepsilon/3$ y, por lo tanto, que

$$d(f(u), f(v)) \leq d(f(u), g(u_m)) + d(g(u_m), g(v_m)) + d(g(v_m), f(v)) < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

Esto nos dice que la función f es uniformente continua. □

Notemos que la continuidad uniforme que aparece como hipótesis en la Proposición 8.4.9 no es una condición necesaria para poder usar la Proposición 8.4.8. Así, si X e Y son espacios métricos, $f : X \to Y$ es una función que *no* es uniformemente continua, y Z es un subconjunto denso de X, entonces la restricción $g \coloneqq f|_Z : Z \to Y$ no es uniformemente continua — esto es consecuencia de la Proposición 8.4.9 — y sin embargo cada vez que $(x_n)_{n\geq 0}$ es una sucesión en Z que converge en X la sucesión $(g(x_n))_{n\geq 0}$ converge en Y.

Lo que hemos hecho nos da condiciones para extender una función continua definida sobre un subconjunto Z de un espacio métrico X a la clausura $\overline{Z}$ de Z en X. El siguiente paso natural es extender esa extensión ahora a todo el espacio X. Este es un problema extremadamente delicado, sobre todo porque depende mucho de la naturaleza del codominio de las funciones que estamos estudiando. Aquí nos contentaremos con presentar dos resultado básico en esa dirección.

Empezamos con la siguiente observación sencilla, que es un caso particular del *Lemma de Urysohn* de Pavel Samuilovich Urysohn [Ury1925], uno de los resultados furdamentales de la topología general.

Proposición 8.4.10. *Sea X un espacio métrico y sean A y B dos subconjuntos no vacíos, cerrados, y disjuntos. Hay una función continua $f : X \to [0,1]$ tal que $A \subseteq f^{-1}(0)$ y $B \subseteq f^{-1}(1)$.*

Porque los espacios métricos tienen esta propiedad decimos que son ***normales***: podemos *separar* en ellos dos subconjuntos cerrados y disjuntos usando una función continua. En la demostración probaremos que hay, de hecho, una función continua $f : X \to [0,1]$ que tiene $A = f^{-1}(0)$

y $B = f^{-1}(1)$, lo que es una propiedad más fuerte que la normalidad — decimos por esto que los espacios métricos son ***perfectamente normales***.

La demostración de esta proposición es muy sencilla porque estamos trabajando con espacios métricos. El resultado original de Urysohn caracteriza los espacios topológicos generales que son normales, y esto es considerablemente más difícil.

Demostración. Como $A \cap B = \varnothing$ y A y B son cerrados, para todo elemento x de X tenemos que $d(x, A) + d(x, B)$. Esto implica que hay una función $f : X \to [0,1]$ que en cada elemento x de X toma el valor

$$f(x) = \frac{d(x, A)}{d(x, A) + d(x, B)}.$$

Se sigue inmediatamente de las Proposiciones 8.1.18 y 8.1.19 que se trata de una función continua.

Si x es un elemento de X, entonces claramente $f(x) = 0$ si y solamente si $d(x, A) = 0$ y, como A es cerrado, esto ocurre exactamente cuando x es un elemento de A. De manera similar, $f(x) = 1$ si y solamente si $d(x, B) = 0$ y, como B es cerrado, esto ocurre exactamente cuando x es un elemento de B. Para probar la proposición, entonces, es suficiente que mostremos que la función f es continua. □

El segundo resultado de extensión que consideraremos es el siguiente *Teorema de extensión de Tietze*, debido a Heinrich Tietze [Tie1914].

Proposición 8.4.11. *Sea X un espacio métrico, sea Z un subespacio cerrado de X, y sea $g : Z \to [a, b]$ una función continua. Hay una función continua $f : X \to [a, b]$ tal que $f|_Z = g$.*

Demostración. Es fácil ver que es suficiente probar el caso particular de la proposición en el que el intervalo $[a, b]$ es $[-1, 1]$, ya que el caso general se deduce de él — dejamos a cargo del lector verificar esto.

Sea $\mathscr{S}$ el conjunto de todas las funciones $\phi : [\![0, n]\!] \to C$ definidas en un conjunto de la forma $[\![0, n]\!]$ para algún $n \in \mathbb{N}_0$ y tales que

$$|\phi(i)(x)| \leq \frac{2^{i-1}}{3^i} \qquad \text{para todo } i \in [\![0, n]\!] \text{ y todo } x \in X$$

y

$$\left| g(x) - \sum_{i=0}^{n} \phi(i)(x) \right| \leq \frac{2^n}{3^n} \quad \text{para todo } x \in Z.$$

Hagamos las siguientes dos observaciones.

- La función $\phi_0 : [\![0, 0]\!] \to C$ tal que $\phi_0(0)(x) = 0$ para todo $x \in X$ es un elemento de $\mathscr{S}$. En efecto, por un lado es $|\phi_0(0)(x)| = |0| \leq 1/2 = 2^{0-1}/3^0$ para todo $x \in X$ y, por otro, es $|g(x) - \phi_0(0)(x)| = |f(x)| \leq 1 = 2^0/3^0$ para todo $x \in Z$.

- Supongamos que $\phi : [\![0, n]\!] \to C$ es un elemento de $\mathscr{S}$ y consideremos los conjuntos

$$A \coloneqq \left\{ z \in Z : g(x) - \sum_{i=0}^{n} \phi(i)(x) \leq -\frac{2^n}{3^{n+1}} \right\}$$

y

$$B \coloneqq \left\{ z \in Z : g(x) - \sum_{i=0}^{n} \phi(i)(x) \geq \frac{2^n}{3^{n+1}} \right\}.$$

Como la función $g - \sum_{i=0}^{n} \phi(i)|_Z : Z \to \mathbb{R}$ es continua, estos dos conjuntos son cerrados en Z y, por lo tanto, en X. Consideremos la función

$$h : x \in X \mapsto -\frac{2^n}{3^{n+1}} + \frac{2^{n+1}}{3^{n+1}} \cdot \frac{d(x, A)}{d(x, A) + d(x, B)} \in \mathbb{R}.$$

Esta función es continua, vale $-2^n/3^{n+1}$ sobre A y $2^n/3^{n+1}$ sobre B, y es $|h(x)| \leq 2^n/3^{n+1}$ para cada $x \in X$. Si x es un elemento de A, entonces $h(x) = -2^n/3^{n+1}$ y

$$-\frac{2^n}{3^n} \leq g(x) - \sum_{i=0}^{n} \phi(i)(x) \leq -\frac{2^n}{3^{n+1}},$$

así que

$$-\frac{2^{n+1}}{3^{n+1}} = -\frac{2^n}{3^n} + \frac{2^n}{3^{n+1}} \leq g(x) - \sum_{i=0}^{n} \phi(i)(x) - h(x) \leq -\frac{2^n}{3^{n+1}} + \frac{2^n}{3^{n+1}} = 0,$$

y si x es un elemento de B, entonces $h(x) = 2^n/3^{n+1}$,

$$\frac{2^n}{3^{n+1}} \leq g(x) - \sum_{i=0}^{n} \phi(i)(x) \leq \frac{2^n}{3^n},$$

y

$$-\frac{2^{n+1}}{3^{n+1}} = \frac{2^n}{3^{n+1}} - \frac{2^n}{3^{n+1}} \leq g(x) - \sum_{i=0}^{n} \phi(i)(x) - h(x) \leq \frac{2^n}{3^n} - \frac{2^n}{3^{n+1}} = \frac{2^{n+1}}{3^{n+1}}.$$

Por otro lado, si x es un elemento de $Z \setminus A \cup B$, entonces

$$\leq |g(x) - \sum_{i=0}^{n} \phi(i)(x) - h(x)| \leq \left| g(x) - \sum_{i=0}^{n} \phi(i)(x) \right| + |h(x)| \leq \frac{2^n}{3^{n+1}} + \frac{2^n}{3^{n+1}} = \frac{2^{n+2}}{3^{n+1}}.$$

Para todo $x \in Z$ vale entonces

$$\left| g(x) - \sum_{i=0}^{n} \phi(i)(x) - h(x) \right| \leq \frac{2^{n+1}}{3^{n+1}},$$

y juntando todo vemos la función $\psi : [\![0, n+1]\!] \to C$ que en cada elemento i de $[\![0, n+1]\!]$ toma el valor

$$\psi(i) = \begin{cases} \phi(i) & \text{si } i \leq n; \\ h & \text{si } i = n+1 \end{cases}$$

es un elemento de $\mathscr{S}$.

La Proposición 2.1.9 nos dice, en vista de todo esto, que hay una función $\phi : \mathbb{N}_0 \to C$ tal que para todo $n \in \mathbb{N}_0$ es $\phi|_{[\![0,n]\!]} \in \mathscr{S}$. Para cada $n \in \mathbb{N}_0$ escribamos $f_n \coloneqq \phi(n)$. Tenemos entonces que

$$|f_n(x)| \leq \frac{2^{n-1}}{3^n} \qquad \text{para todo } n \in \mathbb{N}_0 \text{ y todo } x \in X \tag{8.4}$$

y

$$\left| g(x) - \sum_{i=0}^{n} f_i(x) \right| \leq \frac{2^n}{3^n} \quad \text{para todo } n \in \mathbb{N}_0 \text{ y todo } x \in Z. \tag{8.5}$$

De (8.4) se deduce inmediatamente que para todo elemento x de X converge en $\mathbb{R}$ la serie $\sum_{n=0}^{\infty} f_n(x)$ converge *absolutamente*, así que, en particular, hay una función

$$f : x \in X \mapsto \sum_{n=0}^{\infty} f_n(x) \in \mathbb{R},$$

y por (8.5) sabemos que $f(x) = g(x)$ para todo $x \in Z$, de manera que $f|_Z = g$. Para probar la proposición, entonces, es suficiente que mostremos que la función f es continua.

Sea x_0 un elemento de X y sea ε un número positivo. Hay un entero positivo n_0 tal que $2^{n_0}/3^{n_0} < \varepsilon/3$, ya que $0 < 2/3 < 1$, y la función $s \coloneqq \sum_{i=0}^{n_0} f_i : X \to \mathbb{R}$ es continua, así que hay un número positivo δ tal que para todo elemento x de X vale

$$d(x, x_0) < \delta \implies |s(x) - s(x_0)| < \frac{\varepsilon}{3}.$$

Para cada elemento x de X tenemos que

$$|f(x) - s(x)| = \left| \sum_{n=0}^{\infty} f_n(x) - \sum_{n=0}^{n_0} f_n(x) \right| = \left| \sum_{n=n_0+1}^{\infty} f_n(x) \right|$$

y como la serie converge absolutamente, esto es

$$= \sum_{n=n_0+1}^{\infty} |f_n(x)| \leq \sum_{n=n_0+1}^{\infty} \frac{2^n}{3^n} = \frac{2^{n_0}}{3^{n_0}} < \frac{\varepsilon}{3}.$$

Esto implica que para cada elemento x de X tal que $d(x, x_0) < \delta$ es

$$|f(x) - f(x_0)| \leq |f(x) - s(x)| + |s(x) - s(x_0)| + |s(x_0) - f(x_0)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Vemos con esto que la función f es continua, como queremos. □

Esta proposición implica que una función continua $g : Z \to \mathbb{R}$ es definida sobre un subconjunto cerrado Z de un espacio métrico X que es acotada tiene una extensión continua a todo el espacio X y que esta extensión es acotada. En efecto, podemos en ese caso considerar los núme-

ros $a \coloneqq \inf_{x\in X} f(x)$ y $b \coloneqq \sup_{x\in X} f(x)$ y correstringir g al intervalo $[a,b]$. De hecho, también es cierto para funciones no necesariamente acotadas.

Ejercicio 8.4.12. Sea X un espacio métrico, sea Z un subconjunto cerrado de X, y sea $g : X \to \mathbb{R}$ una función continua. Pruebe que hay una función continua $f : X \to \mathbb{R}$ tal que $f|_Z = g$.

En el *Teorema de Tietze* 8.4.11 es importante que las funciones que consideramos toman valores en $\mathbb{R}$ o en intervalos de $\mathbb{R}$. En efecto, un enunciado análogo para funciones con valores en un espacio métrico arbitrario es falso. Veamos un ejemplo sencillo de esto.

Ejemplo 8.4.13. Consideremos el espacio métrico $X \coloneqq [0,1]$, su subespacio cerrado $Z \coloneqq \{0,1\}$, y la función identidad $g \coloneqq \mathrm{id}_Z : Z \to Z$ de Z, que ciertamente es continua. Afirmamos que

no hay ninguna función continua $f : X \to Z$ tal que $f|_Z = g$.

En efecto, supongamos que hay una función continua $f : X \to Z$ tal que $f|_Z = g$. Su composición inclusión $Z \hookrightarrow \mathbb{R}$ es entonces una función continua $h : [0,1] \to \mathbb{R}$ tal que $h(0) = 0$ y $h(1) = 1$ y, de acuerdo al *Teorema de los valores intermedios*, hay un elemento ξ de $(0,1)$ tal que $h(\xi) = 1/2$. Esto es claramente absurdo, ya que la imagen de la función h está contenida en Z.

Hay muchos ejemplos del mismo tipo, y eso motiva la siguiente definición.

Definición 8.4.14. Sea X un espacio métrico. Un subconjunto Z de X es un ***retracto*** de X si la función identidad $\mathrm{id}_Z : Z \to Z$ tiene una extensión continua a todo el espacio X, y toda tal extensión es una ***retracción*** de X a Z.

Notemos que un retracto es necesariamente cerrado un subconjunto cerrado: en efecto, si $f : X \to Z$ es una extensión continua de id_Z, entonces el conjunto $\{x \in X : f(x) = x\}$ es cerrado en X y coincide con Z. En este lenguaje, el Ejemplo 8.4.13 nos dice que $\{0,1\}$ no es un retracto de $[0,1]$. Muchos problemas importantes de la topología pueden reducirse al de decidir si algún subespacio es o no un retracto.

Ejemplo 8.4.15. Sea n un entero positivo y consideremos la bola unidad cerrada $B \coloneqq \overline{B}_1(0)$ del espacio euclídeo $\mathbb{R}^n$. El subespacio cerrado

$$S \coloneqq \{x \in \mathbb{R}^n : \|x\| = 1\}$$

de B, al que llamamos la ***esfera unidad***, no es un retracto de B. Este es un resultado altamente no trivial, estrechamente relacionado con el *Teorema de punto fijo de Brouwer* de Luitzen Egbertus Jan Brouwer. Normalmente las demostraciones de esto requieren de las herramientas de la topología algebraica — la demostración que da Allen Hatcher en [Hat2002, Corollary 2.15] es un ejemplo de esto — aunque hay otras demotraciones más elementales, como la que presentan Martin Aigner

y Günther Ziegler en [AZ1999], que se basa en el *Lema de Sperner*, un resultado puramente combinartorio, o la que presenta John Milnor en [Mil1978], que depende solamente del cálculo del cálculo de funciones de variables.

8.5. Completación

En general, un espacio métrico no es completo: esto significa que hay sucesiones de Cauchy en él que no tienen límite — en otras palabras, «hay puntos que le faltan». Hay una forma canónica de *agregar* al espacio puntos de manera de que el resultado sea un espacio completo. En esta sección estudiaremos esto.

Fijemos un espacio métrico X.

Lema 8.5.1. *Sea $\mathscr{C}(X)$ el conjunto de todas las sucesiones en X que satisfacen la condición de Cauchy.*

(*i*) *Si $x = (x_n)_{n\geq 0}$ e $y = (y_n)_{n\geq 0}$ son elementos de $\mathscr{C}(X)$, entonces la sucesión $(d(x_n, y_n))_{n\geq 0}$ en $\mathbb{R}$ converge. Hay, por lo tanto, una función*

$$D : (x, y) \in \mathscr{C}(X) \times \mathscr{C}(X) \mapsto \lim_{n\to\infty} d(x_n, y_n) \in \mathbb{R}.$$

(*ii*) *Esta función es una semi-métrica sobre el conjunto $\mathscr{C}(X)$, en el sentido de la Definición 5.2.4.1.*

Demostración. (*i*) Sean $x = (x_n)_{n\geq 0}$ e $y = (y_n)_{n\geq 0}$ dos elementos de $\mathscr{C}(X)$. Sea ε un número positivo. Como x e y satisfacen la condición de Cauchy hay un entero positivo n_0 tal que para toda elección de m y n en $\mathbb{N}_0$ vale $m, n \geq n_0 \implies d(x_m, x_n) < \varepsilon \wedge d(y_m, y_n) < \varepsilon$. Sean m y n dos elementos de $\mathbb{N}_0$ tales que $m, n \geq n_0$. Es

$$d(x_m, y_m) \leq d(x_m, x_n) + d(x_n, y_n) + d(y_m, y_n) < \frac{\varepsilon}{2} + d(x_n, y_n) + \frac{\varepsilon}{2},$$

de manera que $d(x_m, y_m) - d(x_n, y_n) < \varepsilon$. Por simetría también es $d(x_n, y_n) - d(x_m, y_m) < \varepsilon$, así que $|d(x_m, y_m) - d(x_n, y_n a)| < \varepsilon$. Esto prueba que la sucesión $(d(x_n, y_n))_{n\geq 0}$ en $\mathbb{R}$ satisface la condición de Cauchy y, como $\mathbb{R}$ es un espacio métrico completo, que converge. Vemos así que la primera afirmación del lema es cierta.

Mostremos ahora que la función D definida ahí es una semi-métrica sobre el conjunto $\mathscr{C}(X)$.

- Si $x = (x_n)_{n\geq 0}$ es un elemento de $\mathscr{C}(X)$, entonces $d(x_n, x_n) = 0$ para todo $n \in \mathbb{N}$, así que, por supuesto, $D(x, x) = \lim_{n\to\infty} d(x_n, y_n) = 0$.
- Si $x = (x_n)_{n\geq 0}$ e $y = (y_n)_{n\geq 0}$ son dos elementos de $\mathscr{C}(X)$, entonces para todo elemento n de $\mathbb{N}_0$ tenemos que $d(x_n, y_n) = d(y_n, x_n)$, así que

$$D(x, y) = \lim_{n\to\infty} d(x_n, y_n) = \lim_{n\to\infty} d(y_n, x_n) = D(y, x).$$

- Sean $x = (x_n)_{n\geq 0}$, $y = (y_n)_{n\geq 0}$, y $z = (z_n)_{n\geq 0}$ tres elementos de $\mathscr{C}(X)$. Para cada elemento n de $\mathbb{N}_0$ tenemos que $d(x_n, y_n) \leq d(x_n, z_n) + d(z_n, y_n)$, así que

$$\begin{aligned} D(x,y) &= \lim_{n\to\infty} d(x_n, y_n) \leq \lim_{n\to\infty} \big(d(x_n, z_n) + d(z_n, y_n)\big) \\ &= \lim_{n\to\infty} d(x_n, z_n) + \lim_{n\to\infty} d(z_n, y_n) = D(x,z) + D(z,y). \end{aligned}$$ □

Consideremos la relación

$$R \coloneqq \{(x,y) \in \mathscr{C}(X) \times \mathscr{C}(X) : D(x,y) = 0\}$$

sobre el conjunto $\mathscr{C}(X)$. De acuerdo a la primera parte de la Proposición 5.2.4.2, se trata de una relación de equivalencia sobre $\mathscr{C}(X)$. Escribamos $\bar{\mathscr{C}}(X)$ al conjunto cociente $\mathscr{C}(X)/R$, y para cada elemento x de $\mathscr{C}(X)$, escribamos $\bar{x}$ a la clase de equivalencia de x en $\mathscr{C}(X)$ con respecto a R. La segunda parte de la Proposición 5.2.4.2 nos dice que hay exactamente una función $\bar{D} : \bar{\mathscr{C}}(X) \times \bar{\mathscr{C}}(X) \to \mathbb{R}$ tal que

$$\bar{D}(\bar{x}, \bar{y}) = D(x,y)$$

cualesquiera sean los elementos x e y de $\mathscr{C}(X)$, y que esta función $\bar{D}$ es una métrica sobre el conjunto $\bar{\mathscr{C}}(X)$.

Si x es un elemento de X, entonces podemos considerar la sucesión constante $(x)_{n\geq 0}$, que evidentemente es un elemento de $\mathscr{C}(X)$, y, en consecuencia, su clase de equivalencia con respecto a la relación R, que es un elemento de $\bar{\mathscr{C}}(X)$: la escribiremos $\iota(x)$. Obtenemos de esta forma una función $\iota : X \to \bar{\mathscr{C}}(X)$.

Lema 8.5.2. *La función $\iota : X \to \bar{\mathscr{C}}(X)$ es una inyección isométrica, de manera que*

$$\bar{D}(\iota(x), \iota(y)) = d(x,y)$$

cualesquiera sean los elementos x e y de X, y su imagen es un subconjunto denso de $\bar{\mathscr{C}}(X)$.

Demostración. De acuerdo a las definiciones, si x e y son dos elementos de X, entonces

$$\bar{D}(\iota(x), \iota(y)) = D((x)_{n\geq 0}, (y)_{n\geq 0}) = \lim_{n\to\infty} d(x,y) = d(x,y).$$

Esto prueba la primera afirmación del lema. Para ver la segunda, sea ξ un elemento de $\bar{\mathscr{C}}(x)$, de manera que hay un elemento $x = (x_n)_{n\geq 0}$ tal que $\xi = \bar{x}$, y sea ε un número positivo. Como la sucesión x pertenece a $\mathscr{C}(X)$ satisface la condición de Cauchy y hay, por lo tanto, un entero positivo n_0 tal que para toda elección de m y n en $\mathbb{N}_0$ vale $m, n \geq n_0 \implies d(x_m, x_n) < \varepsilon/2$. Esto

implica que

$$\bar{D}(\xi, \iota(x_{n_0})) = D(x, (x_{n_0})_{n\geq 0}) = \lim_{n\to\infty} d(x_n, x_{n_0}) \leq \frac{\varepsilon}{2} < \varepsilon$$

y, por lo tanto, que $\operatorname{img}(\iota) \cap B_\varepsilon(\xi) \neq \varnothing$. Vemos así que el conjunto $\operatorname{img}(\iota)$ es denso en $\bar{\mathscr{C}}(X)$. □

Este lema y el siguiente resultado general nos permitirán probar que el espacio métrico $\bar{\mathscr{C}}(X)$ es completo.

Lema 8.5.3. *Sea Y un espacio métrico y sea D un subconjunto denso de Y. Si toda sucesión con valores en D que satisface la condición de Cauchy converge en Y, entonces el espacio Y es completo.*

Demostración. Supongamos que toda sucesión con valores en D que satisface la condición de Cauhcy converge en Y, y sea $y = (y_n)_{n\geq 0}$ una sucesión en Y que satisface la condición de Cauchy. La densidad de D en Y implica que para cada $n \in \mathbb{N}_0$ hay un elemento z_n de D tal que $d(y_n, z_n) < 1/(n+1)$. Sea ε un número positivo. Que la sucesión y satisfaga la condición de Cauchy nos dice que hay un entero positivo n_1 tal que para toda elección de m y n en $\mathbb{N}_0$ vale $m, n \geq n_1 \implies d(y_m, y_n) < \varepsilon/3$. Hay, por otro lado, un entero positivo n_2 tal que $1/(n_2+1) < \varepsilon/3$. Si ponemos $n_0 \coloneqq \max\{n_1, n_2\}$, entonces siempre que m y n son dos elementos de $\mathbb{N}_0$ tales que $m, n \geq n_0$ tenemos que

$$d(z_m, z_n) \leq d(z_m, y_m) + d(y_m, y_n) + d(y_n, z_n) < \frac{1}{m+1} + \frac{\varepsilon}{3} + \frac{1}{n+1} < \varepsilon.$$

Vemos con todo esto que la sucesión $(z_n)_{n\geq 0}$, que tiene valores en D, satisface la condición de Cauchy: la hipótesis nos permite concluir entonces que converge a un punto ζ de Y. Como claramente $\lim_{n\to\infty} d(y_n, z_n) = 0$, el resultado del Ejercicio 7.3.8 nos permite concluir que la sucesión y converge en Y y que lo hace al punto ζ. □

Como dijimos, esto nos permite dar el siguiente paso en nuestra construcción.

Lema 8.5.4. *El espacio métrico $\bar{\mathscr{C}}(X)$ es completo.*

Demostración. Esto es consecuencia inmediata de los Lemas 8.5.2 y 8.5.3. □

En este punto hemos construido, empezando a partir del espacio métrico X, una inyección isométrica $\iota : X \to \bar{\mathscr{C}}(X)$ que tiene imagen densa. En los términos de la siguiente definición, esa función es una completación.

Definición 8.5.5. Sean X e Y dos espacios métricos. Decimos que una función $\iota : X \to Y$ es una ***completación*** del espacio X si es una inyección isométrica con imagen densa y el espacio Y es completo.

Usando este lenguaje, podemos resumir la construcción que hicimos de la siguiente manera:

Proposición 8.5.6. *Para todo espacio métrico X la función $\iota : X \to \bar{\mathscr{C}}(X)$ que construimos arriba es una completación de X.* □

El siguiente ejercicio da un ejemplo mucho más trivial de completación.

Ejercicio 8.5.7. Muestre que si X es un espacio métrico completo, entonces la función $\mathrm{id}_X : X \to X$ es una completación de X.

La propiedad fundamental de las completaciones de un espacio es la siguiente.

Proposición 8.5.8. *Sean X e Y dos espacios métricos y sea $\iota : X \to Y$ una completación de X. Cada vez que Z es un espacio métrico completo y $f : X \to Z$ es una función uniformemente continua hay exactamente una función continua $\bar{f} : Y \to Z$ tal que $\bar{f} \circ \iota = f$, y esta función es uniformemente continua.*

$$\begin{array}{ccc} X & \xrightarrow{\ f\ } & Z \\ {\scriptstyle \iota}\downarrow & \nearrow_{\bar{f}} & \\ Y & & \end{array}$$

Esta es la llamada ***propiedad universal*** de la completación.

Demostración. Sea Z un espacio métrico completo y sea $f : X \to Z$ una función uniformemente continua. El conjunto $D \coloneqq \iota(X)$ es denso en Y, la correstricción $\iota|^D : X \to D$ es biyectiva, y podemos por, lo tanto, considerar la función $g \coloneqq f \circ (\iota|^D)^{-1} : D \to Z$.

Esta función es uniformemente continua. Para verlo, sea ε un número positivo. Como la función f es uniformemente continua, hay un número positivo δ tal que siempre que x y x' son elementos de X vale $d(x, x') < \delta \implies d(f(x), f(x')) < \varepsilon$. Esto implica que si u y u' son dos elementos de D tales que $d(u, u') < \delta$, entonces $d((\iota|^D)^{-1}(u), (\iota|^D)^{-1}(u')) = d(u, u') < \delta$ porque la función ι es una inyección isométrica y, por lo tanto,

$$d(g(u), g(u')) = d(f((\iota|^D)^{-1}(u)), f((\iota|^D)^{-1}(u'))) < \varepsilon.$$

Como la función g es uniformemente continua, el espacio Z es completo, y $\bar{D} = Y$, la Proposición 8.4.9 nos permite concluir que hay una función continua $\bar{f} : Y \to Z$ tal que $\bar{f}|_D = g$, y que

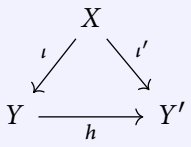
esta función es uniformemente continua. Para cada elemento x de X es $\iota(x) \in D$, así que

$$\bar{f}(\iota(x)) = \bar{f}|_D(\iota(x)) = g(\iota(x)) = f((\iota|^D)^{-1}(\iota(x)) = f(x),$$

y esto nos dice que $\bar{f} \circ \iota = f$. Todo esto prueba la afirmación de existencia de la proposición y la de uniformidad. La de unicidad es consecuencia inmediata del Corolario 8.4.5. □

Todo espacio métrico admite muchas completaciones, pero todas son «esencialmente la misma»: la siguiente proposición explicita qué queremos decir con esto.

Proposición 8.5.9. *Sea X un espacio métrico. Si $\iota : X \to Y$ e $\iota' : X \to Y'$ son dos completaciones de X, entonces hay exactamente un función continua $h : Y \to Y'$ tal que $\iota' = h \circ \iota$ y se trata de una isometría.*



Demostración. Sean $\iota : X \to Y$ e $\iota' : X \to Y'$ dos completaciones de X. Como el espacio Y' es completo, la función ι es una completación de X, y la función ι' es uniformemente continua — ya que es una inyección isométrica — la Proposición 8.5.8 nos dice que hay una función continua $h : Y \to Y'$ tal que $h \circ \iota = \iota'$. De manera similar, como el espacio Y es completo, la función ι' es una completación de X, y la función ι es uniformemente continua la Proposición 8.5.8 nos dice que hay una función continua $k : Y' \to Y$ tal que $k \circ \iota' = \iota$.

Notemos que

$$(k \circ h) \circ \iota = k \circ (h \circ \iota) = k \circ \iota' = \iota = \mathsf{id}_Y \circ \iota.$$

Esto implica que las funciones continuas $k \circ h : Y \to Y$ e $\mathsf{id}_Y : Y \to Y$ coinciden sobre la imagen de la función ι y, como esa imagen es densa en Y, que son, de hecho, iguales. Por supuesto, como también

$$(h \circ k) \circ \iota' = h \circ (k \circ \iota') = h \circ \iota = \iota' = \mathsf{id}_{Y'} \circ \iota',$$

podemos deducir de la misma forma que $h \circ k = \mathsf{id}_{Y'}$. Vemos así que las funciones h y k son mutualmente inversas y, en particular, que la función h es biyectiva.

Para terminar, tenemos que mostrar que la función h es una isometría. Sean y y z dos elementos de Y. Como la imagen de la función ι es densa en Y, hay sucesiones $(y_n)_{n\geq 0}$ y $(z_n)_{n\geq 0}$ en X tales que las sucesiones $(\iota(y_n))_{n\geq 0}$ e $(\iota(z_n))_{n\geq 0}$ convergen a y y a z, respectivamente. Como la función h es continua, las sucesiones $(\iota'(y_n))_{n\geq 0} = (h(\iota(y_n)))_{n\geq 0}$ y $(\iota'(z_n))_{n\geq 0} = (h(\iota(z_n)))_{n\geq 0}$ convergen

a $h(y)$ y a $h(z)$. Esto implica que la sucesión $(d(y_n, z_n))_{n\geq 0} = (d(\iota(y_n), \iota(z_0)))_{n\geq 0}$ converge a $d(y,z)$ y que la sucesión $(d(y_n, z_n))_{n\geq 0} = (d(\iota'(y_n), \iota'(z_0)))_{n\geq 0}$ converge a $d(h(y), h(z))$: vemos así que $d(h(y), h(z)) = d(y,z)$. □

Para todo espacio métrico X hemos construido una completación $\iota : X \to \bar{\mathscr{C}}(X)$, pero hay situaciones en las que podemos encontrar otras completaciones de manera mucho más directa: la gracia de la Proposición 8.5.9 es que lo que encontramos siempre es esencialmente lo mismo. Demos un ejemplo de esto.

Proposición 8.5.10. *Sea Y un espacio métrico completo. Si X es un subespacio de Y y $\overline{X}$ es la clausura de X en Y, entonces la función de inclusión $\iota : X \hookrightarrow \overline{X}$ es una completación de X.*

Demostración. El espacio métrico $\overline{X}$ es completo, porque es un subespacio cerrado del espacio completo Y, y es evidente que la función de inclusión $\iota : X \hookrightarrow \overline{X}$ es una inyección isométrica de imagen densa. □

Usando este resultado podemos construir dar una segunda construcción de una completación de cada espacio métrico.

HACER: Esto depende de C_v, que está en la sección siguiente…

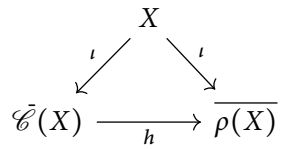
Proposición 8.5.11. *Sea X un espacio métrico no vacío, sea x_0 un punto de X, y sea $C_b(X, \mathbb{R})$ el espacio métrico de todas las funciones continuas y acotadas $X \to \mathbb{R}$.*

(i) Para cada elemento x de X la función $\rho_x : y \in X \mapsto d(y,x) - d(y, x_0) \in \mathbb{R}$ es continua y acotada.

(ii) La función $\rho : x \in X \mapsto \rho_x \in C_b(X, \mathbb{R})$ es una inyección isométrica.

(iii) Si $\overline{\rho(X)}$ es la clausura de la imagen de la función ρ en $C_b(X, \mathbb{R})$, entonces la correstricción $\iota : x \in X \mapsto \rho_x \in \overline{\rho(X)}$ de la función ρ a $\overline{\rho(X)}$ es una completación de X.

Por supuesto, la Proposición 8.5.9 implica que hay una isometría $h : \bar{\mathscr{C}}(X) \to \overline{\rho(X)}$ unívocamente determinada que hace conmutar el diagrama



y, por lo tanto, que este espacio $\overline{\rho(X)}$ es, esencialmente, el mismo que $\bar{\mathscr{C}}(X)$.

Demostración. (*i*) Sea x un elemento de X. Es

$$|\rho_x(y) - \rho_x(y')| = \big|(d(y,x) - d(y,x_0)) - (d(y',x) - d(y',x_0))\big|$$

$$\begin{aligned} &\leq \left|d(y,x) - d(y',x_0)\right| + \left|d(y',x) - d(y',x_0)\right| \\ &\leq d(y,y') + d(y,y') = 2d(y,y') \end{aligned}$$

cualesquiera sean y e y' en X, y esto nos dice que la función ρ_x satisface la condición de Lipschitz y, en particular, que es continua. Por otro lado, si y es un elemento de X, entonces

$$|\rho_x(y)| = |d(y,x) - d(y,x_0)| \leq d(x,x_0),$$

así que la función ρ_x es acotada.

(*ii*) Sean x y x' dos elementos de X. Para todo elemento y de X tenemos que

$$\begin{aligned} |\rho_x(y) - \rho_{x'}(y)| &= \left|(d(y,x) - d(y,x_0)) - (d(y,x') - d(y,x_0))\right| \\ &= |d(y,x) - d(y,x')| \\ &\leq d(x,x'), \end{aligned}$$

Como, por otro lado, es

$$|\rho_x(x) - \rho_{x'}(x)| = \left|(d(x,x) - d(x,x_0)) - (d(x,x') - d(,x_0))\right| = d(x,x'),$$

vemos que $d(\rho_x, \rho_{x'}) = \sup\{d(\rho_x(y), \rho_{x'}(y)) : y \in X\} = d(x,x')$.

(*iii*) La función ι descripta en la proposición es una inyección isométrica con imagen densa, así que es una completación de su dominio. □

Notemos que esta construcción de una completación de un espacio métrico es mucho más eficiente que la primera que hicimos! En general, sin embargo, el proceso de construir una completación de un espacio es *productivo*, en el sentido de que el resultado final, el codominio de la completación, es un espacio «nuevo». Para muchas cosas se hace necesario tener descripciones concretas de ese espacio, y para ello la descripción en término de sucesiones de Cauchy es mucho más práctica.

Proposición 8.5.12. *Sean X e Y dos espacios métricos, y sean $\iota : X \to Z$ e $\iota' : Y \to W$ una completación de X y una de Y, respectivamente. La función*

$$\iota'' : (x,y) \in X \times Y \mapsto (\iota(x), \iota'(y)) \in Z \times W$$

es una completación de $X \times Y$.

Demostración. Es claro que la función ι'' es una inyección isométrica, y el resultado de la primera parte del Ejercicio 6.6.7 nos permite concluir que su imagen es densa en su condominio, así que se trata de una completación. □

8.5.1. Completación de grupos abelianos

Cuando completamos un espacio que tiene alguna estructura adicional, la completación tiene una estructura similar. Nuestro siguiente objetivo es explicitar esto en el caso de los grupos abelianos.

Proposición 8.5.1.1. *Sea G un grupo abeliano dotado de una métrica d que es invariante por translaciones y sea $\iota : G \to \overline{G}$ una completación de G. Hay exactamente una función continua $\overline{m} : \overline{G} \times \overline{G} \to \overline{G}$ que hace del conjunto $\overline{G}$ un grupo abeliano de manera tal que la función ι sea un morfismo de grupos. La métrica de $\overline{G}$ es invariante por translaciones con respecto a esa estructura de grupo.*

Desde ahora siempre que consideremos la completación de un grupo abeliano con respecto a una métrica invariante por translaciones la consideraremos como un grupo con respecto a la estructura de grupo que describe esta proposición.

Demostración. Sea $\iota : G \to \overline{G}$ una completación de G. De acuerdo a la Proposición 8.5.12, la función

$$\iota \times \iota : (x, y) \in G \times G \mapsto (\iota(x), \iota(y)) \in \overline{G} \times \overline{G}$$

es una completación de $G \times G$, vimos en Ejemplo **??** que la función

$$m : (x, y) \in G \times G \mapsto x + y \in G$$

es uniformemente continua, y como la función $\iota : G \to \overline{G}$ es una inyección isométrica, también lo es la composición $\iota \circ m$. La propiedad universal de la completación, entonces, nos dice que hay exactamente una función $\overline{m} : \overline{G} \times \overline{G} \to \overline{G}$ tal que

$$\overline{m} \circ \iota \times \iota = \iota \circ m. \qquad \begin{array}{ccc} G \times G & \xrightarrow{m} & G \\ {\scriptstyle \iota_2}\downarrow & & \downarrow{\scriptstyle \iota} \\ \overline{G} \times \overline{G} & \xrightarrow{\overline{m}} & \overline{G} \end{array}$$

Podemos evaluar la función $\overline{m}$ de la siguiente manera: si x e y son dos elementos de $\overline{G}$, entonces hay sucesiones $(x_n)_n \geq 0$ e $(y_n)_{n\geq 0}$ con valores en G tales que $(\iota(x_n))_{n\geq 0}$ e $(\iota(y_n))_{n\geq 0}$ convergen en $\overline{G}$ a x y a y, respectivamente, y la sucesión $(\iota(x_n + y_n))_{n\geq 0}$ converge en $\overline{G}$ y su límite es

$$\overline{m}(x, y) = \lim_{n\to\infty} \iota(x_n + y_n).$$

Queremos probar que la función $\overline{m}$ es una operación sobre el conjunto $\overline{G}$ que hace de él un grupo abeliano.

- Sean x, y, y z tres elementos de $\overline{G}$. Hay entonces sucesiones $(x_n)_{n\geq 0}$, $(y_n)_{n\geq 0}$, y $(z_n)_{n\geq 0}$ en G tales que las sucesiones $(\iota(x_n))_{n\geq 0}$, $(\iota(y_n))_{n\geq 0}$, y $(\iota(z_n))_{n\geq 0}$ convergen en $\overline{G}$ a x, a y, y a z,

respectivamente. Por su parte, las sucesiones $(\iota(x_n + y_n))_{n\geq 0}$ y $(\iota(y_n + z_n))_{n\geq 0}$ convergen en $\overline{G}$ a $\overline{m}(x, y)$ y a $\overline{m}(y, z)$, y entonces las sucesiones

$$(\iota((x_n + y_n) + z_n))_{n\geq 0}, \qquad (\iota(z_n + (y_n + z_n)))_{n\geq 0} \tag{8.6}$$

convergen en $\overline{G}$ a $(x+y)+z$ y a $x+(y+z)$. Ahora bien, como la operación de G es asociativa, tenemos que para todo $n \in \mathbb{N}_0$ es $(x_n + y_n) + z_n = x_n + (y_n + z_n)$ y, por lo tanto, que las dos sucesiones de (8.6) son iguales y, por supuesto, sus límites también, de manera que $\overline{m}(\overline{m}(x, y), z) = \overline{m}(x, \overline{m}(y, z))$. Esto nos dice que la operación $\overline{m}$ en $\overline{G}$ es asociativa.

- Sean x e y dos elementos de $\overline{G}$, y sean $(x_n)_{n\geq 0}$ e $(y_n)_{n\geq 0}$, sucesiones en G tales que las sucesiones $(\iota(x_n))_{n\geq 0}$ e $(\iota(y_n))_{n\geq 0}$ convergen en $\overline{G}$ a x y a y. Las sucesiones $(\iota(x_n+y_n))_{n\geq 0}$ e $(\iota(y_n+x_n))_{n\geq 0}$ convergen entonces en $\overline{G}$ a $\overline{m}(x, y)$ y a $\overline{m}(x, y)$ y, como son iguales porque $x_n + y_n = y_n + x_n$ para todo $n \in \mathbb{N}_0$, esto nos permite concluir que $\overline{m}(x, y) = \overline{m}(y, x)$. Vemos así que la operación $\overline{m}$ es conmutativa.
- Sea x un elemento de $\overline{G}$ y sea $(x_n)_{n\geq 0}$ una sucesión en G tal que la sucesión $(\iota(x_n))_{n\geq 0}$ converge a x en $\overline{G}$. La sucesión constante $(\iota(0))_{n\geq 0}$ en $\overline{G}$ a $\iota(0)$, porque la función ι es continua, y la sucesión $(\iota(x_n+0))_{n\geq 0}$ converge a $\overline{m}(x, 0)$. Como $x_n+0 = x_n$ para todo $n \in \mathbb{N}_0$, esta última sucesión coincide con $(\iota(x_n))_{n\geq 0}$ y, por lo tanto, tenemos que $\overline{m}(x, \iota(0)) = x$. Podemos concluir con esto que el elemento $\iota(0)$ de $\overline{G}$ es un elemento neutro para la operación $\overline{m}$.
- Sea finalmente x un elemento de $\overline{G}$ y sea $(x_n)_{n\geq 0}$ una sucesión en G tal que la sucesión $(\iota(x_n))_{n\geq 0}$ converge en $\overline{G}$ a x. Sea ε un número positivo. Como la sucesión $(\iota(x_n))_{n\geq 0}$ converge en $\overline{G}$, satisface la condominio de Cauchy y, como la función ι es una inyección isométrica, también la satisface la sucesión $(x_n)_{n\geq 0}$. Hay entonces un entero positivo n_0 tal que para toda elección de m y n en $\mathbb{N}_0$ vale $m, n \geq n_0 \implies d(x_m, x_n) < \varepsilon$. Ahora bien, si m y n son dos elementos de $\mathbb{N}_0$ tales que $m, n \geq n_0$, entonces la invariancia por translaciones de la métrica de G implica que

  $$d(-x_m, -x_n) = d(-x_m + (x_m + x_n), -x_n + (x_m + x_n)) = d(x_n, x_m) = d(x_m, x_n) < \varepsilon.$$

  Vemos así que la sucesión $(-x_n)_{n\geq 0}$ en G satisface la condición de Cauchy, así que la sucesión $(\iota(-x_n))_{n\geq 0}$ también y, por lo tanto, esta converge a un elemento y de $\overline{G}$. La sucesión $(\iota(x_n + (-x_n)))_{n\geq 0}$ converge entonces a $\overline{m}(x, y)$: como $x_n + (-x_n) = 0$ para todo $n \in \mathbb{N}_0$, esta sucesión no es otra que la sucesión constante $(\iota(0))_{n\geq 0}$, que converge a $\iota(0)$, que es el elemento neutro de $\overline{G}$. Vemos así que hay un elemento y de $\overline{G}$ tal que $\overline{m}(x, y) = \iota(0)$. La conclusión de esto es que todo elemento de $\overline{G}$ posee un elemento opuesto con respecto a la operación $\overline{m}$.

Lo que hemos hecho prueba que, como dijimos, el conjunto $\overline{G}$ es un grupo abeliano con respecto a la operación $\overline{m}$ y elemento neutro $\iota(0)$. Desde ahora escribiremos a su operación + y a su elemento neutro simplemente 0, como hacemos con todos los grupos abelianos.

Mostremos ahora que la función $\iota : G \to \overline{G}$ es un morfismo de grupos.

- Sean x e y dos elementos de G. Las sucesiones constantes $(x)_{n\geq 0}$ e $(y)_{n\geq 0}$ son tales que las sucesiones $(\iota(x))_{n\geq 0}$, $(\iota(y))_{n\geq 0}$, y $(\iota(x+y))_{n\geq 0}$ convergen en $\overline{G}$ a $\iota(x)$, a $\iota(y)$, y a $\iota(x+y)$, respectivamente, así que la descripción de la operación $\overline{m}$ que dimos arriba implica que la sucesión $(\iota(x+y))_{n\geq 0}$ converge en $\overline{G}$ a $\iota(x)+\iota(y)$. Esto nos dice que $\iota(x+y) = \iota(x)+\iota(y)$.

Probemos que, más aún, la operación $\overline{m}$ es la única operación continua sobre $\overline{G}$ que hace que la función ι sea un morfismo de grupos.

- Supongamos que $\overline{m}' : \overline{G}\times\overline{G} \to \overline{G}$ es otra función continua que hace de $\overline{G}$ un grupo abeliano tal que la función ι es un morfismo de grupos. Sean x e y dos elementos de $\overline{G}$. Podemos elegir sucesiones $(x_n)_{n\geq 0}$ e $(y_n)_{n\geq 0}$ en G de manera que las sucesiones $(\iota(x_n))_{n\geq 0}$ e $(\iota(y_n))_{n\geq 0}$ converjan en $\overline{G}$ a x y a y. Sabemos que $\overline{m}(x,y)$ es el límite de la sucesión $(\iota(x_n+y_n))_{n\geq 0}$ en $\overline{G}$. Por otro lado, la sucesión $((\iota(x_n),\iota(y_n)))_{n\geq 0}$ converge en $\overline{G}\times\overline{G}$ a (x,y), así que la continuidad de la función $\overline{m}'$ implica que la sucesión $(\overline{m}'(\iota(x_n),\iota(y_n)))_{n\geq 0}$ converge a $\overline{m}'(x,y)$. Ahora bien, como la función ι es un morfismo de grupos $G\to\overline{G}$ cuando hacemos de $\overline{G}$ un grupo con la operación $\overline{m}'$, para cada $n\in\mathbb{N}_0$ tenemos que $\overline{m}'(\iota(x_n),\iota(y_n)) = \iota(x_n+y_n)$: esto implica que la sucesión $(\overline{m}'(\iota(x_n),\iota(y_n)))_{n\geq 0}$ tiene el mismo límite que $(\iota(x_n+y_n))_{n\geq 0}$ y, por lo tanto, que $\overline{m}'(x,y) = \overline{m}(x,y)$. Vemos así que las funciones $\overline{m}$ y $\overline{m}'$ coinciden.

Con todo esto hemos probado la primera afirmación de la proposición. Para probar la segunda y, con ello, completar la demostración mostremos que la métrica de $\overline{G}$ es invariante por translaciones con respecto a la estructura de grupo que le da $\overline{m}$ a $\overline{G}$.

Sean x, y, y z tres elementos de $\overline{G}$. Hay entonces sucesiones $(x_n)_{n\geq 0}$, $(y_n)_{n\geq 0}$, y $(z_n)_{n\geq 0}$ en G tales que las sucesiones $(\iota(x_n))_{n\geq 0}$, $(\iota(y_n))_{n\geq 0}$, y $(\iota(z_n))_{n\geq 0}$ convergen en $\overline{G}$ a x, a y, y a z, respectivamente. La sucesión $(d(\iota(x_n),\iota(y_n)))_{n\geq 0}$ converge en $\mathbb{R}$ a $d(x,y)$, porque la métrica de $\overline{G}$ es una función continua. De manera similar, como las sucesiones $(\iota(x_n+z_n))_{n\geq 0}$, y $(\iota(y_n+z_n))_{n\geq 0}$ convergen en $\overline{G}$ a $x+z$ y a y_z, la sucesión $(d(\iota(x_n+z_n),\iota(y_n+z_n)))_{n\geq 0}$ converge a $d(x+z,y+z)$. Finamente, como la función ι es una inyección isométrica y la métrica de G es invariante por translaciones, para cada $n\in\mathbb{N}_0$ tenemos que

$$d(\iota(x_n),\iota(y_n)) = d(x_n,y_n) = d(x_n+z_n,y_n+z_n) = d(\iota(x_n+z_n),\iota(y_n+z_n)).$$

Juntado todo, entonces, vemos que $d(x,y) = d(x+z,y+z)$. □

La propiedad universal de la completación que describe la Proposición 8.5.8 tiene el siguiente refinamiento en el caso de la completación de grupos abelianos dotados de métricas invariantes por translaciones.

Proposición 8.5.1.2. *Sea G un grupo abeliano dotado de una métrica d que es invariante por translaciones y sea $\iota : G\to\overline{G}$ una completación de G. Cada vez que H es un grupo abeliano dotado de una métrica invariante por translaciones con respecto a la cual es un espacio métrico completo*

y $f : G \to H$ es un morfismo de grupos uniformemente continuo hay exactamente un morfismo de grupos continuo $\bar{f} : \bar{G} \to H$ tal que $\bar{f} \circ \iota = f$, y este morfismo de grupos $\bar{f}$ es uniformemente continuo.

Demostración. Sea H un grupo abeliano dotado de una métrica invariante por translaciones con respecto a la cual es un espacio métrico completo y sea $f : G \to H$ un morfismo de grupos uniformemente continuo. La propiedad universal de la completación que nos da la Proposición 8.5.8 nos dice que hay exactamente una función continua $\bar{f} : \overline{G} \to H$ tal que $\bar{f} \circ \iota = f$, y que esta función es uniformemente continua. Para probar la proposición, entonces, es suficiente que mostremos que esta función $\bar{f}$ es un morfismo de grupos.

Sean x e y dos elementos de $\overline{G}$, y sean $(x_n)_{n\geq 0}$ e $(y_n)_{n\geq 0}$ sucesiones en G tales que las sucesiones $(\iota(x_n))_{n\geq 0}$ e $(\iota(y_n))_{n\geq 0}$ convergen a x y a y en $\overline{G}$. Usando la continuidad de la función $\bar{f}$ podemos hacer las siguientes dos observaciones.

- Las sucesiones $(\bar{f}(\iota(x_n)))_{n\geq 0}$ y $(\bar{f}(\iota(y_n)))_{n\geq 0}$ convergen en H a $\bar{f}(x)$ y a $\bar{f}(y)$ y, como la suma de H es una función continua, la sucesión $(\bar{f}(\iota(x_n)) + \bar{f}(\iota(y_n)))_{n\geq 0}$ converge entonces en H a $\bar{f}(x) + \bar{f}(y)$,
- La sucesión $(\iota(x_n + y_n))_{n\geq 0}$ converge en $\overline{G}$ a $x + y$, así que la sucesión $(\bar{f}(\iota(x_n + y_n)))_{n\geq 0}$ converge a $\bar{f}(x + y)$.

Ahora bien, para cada $n \in \mathbb{N}_0$ tenemos que

$$\bar{f}(\iota(x_n)) + \bar{f}(\iota(y_n)) = f(x_n) + f(y_n) = f(x_n + y_n) = \bar{f}(\iota(x_n + y_n)),$$

que $\bar{f} \circ \iota = f$ y f es un morfismo de grupos, así que podemos concluir que $\bar{f}(x) + \bar{f}(y) = \bar{f}(x + y)$, de manera que la función $\bar{f}$ es un morfismo de grupos, como queremos. □

8.5.2. Completación de espacios normados

Proposición 8.5.2.1. *Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$ y sea V un $\mathbb{k}$-espacio normado. Si $\iota : V \to W$ es una completación de V, entonces hay exactamente una estructura de $\mathbb{k}$-espacio normado sobre W que induce sobre él su estructura de espacio métrico y tal que la función ι es lineal.*

Demostración. □

8.5.3. El cuerpo de los números p-ádicos

Ejemplo 8.5.3.1. Sea p un número primo, sean v_p, N_p, y d_p la valuación p-ádica, el valor absoluto p-ádico, y la correspondiente ultramétrica d_p, los tres definidos sobre $\mathbb{Q}$, y sea $\iota : \mathbb{Q} \to \overline{\mathbb{Q}}_p$ una completación del espacio métrico $(\mathbb{Z}, d_p)$.

Mostremos que

hay una función $\overline{N}_p : \overline{\mathbb{Q}}_p \to \mathbb{R}$ tal que $\overline{N}_p \circ \iota = N_p$.

Para esto es suficiente, de acuerdo a la Proposición 8.5.8, con que mostremos que la función $N_p : \mathbb{Q} \to \mathbb{R}$ es uniformemente continua. Sea ε un número positivo. Si r y s son dos elementos de $\mathbb{Q}$ tales que $d_p(r,s) < \varepsilon$, entonces o bien $r = s$ y en ese caso $N_p(r) - N_p(s)$

Sabemos que el conjunto $\iota(\mathbb{Q})$ es denso en $\overline{\mathbb{Q}}_p$. Mostremos que, de hecho, vale que

para todo elemento α de $\overline{\mathbb{Q}}_p$ y todo entero no negativo n hay un entero a_n tal que $0 \leq a_n < p^n - 1$ y $d(\alpha, a_n) \leq p^{-n}$.

Probemos esto. Sean α un elemento de $\overline{\mathbb{Q}}_p$ y sea n un entero no negativo. Como $\iota(\mathbb{Q})$ es denso en $\overline{\mathbb{Q}}_p$, hay enteros u y v tales que $v > 0$ y $d(\alpha, u/v) < p^{-n}$.

Probemos esto. Sea ξ un elemento de $\overline{\mathbb{Z}}_p$ y sea ε un número positivo. Como $\iota(\mathbb{Z})$ es denso en $\overline{\mathbb{Z}}_p$, hay un entero n tal que $d(\xi, \iota(n)) < \varepsilon$. Si n es negativo, entonces $v_p(n+n) = v_p(2n) \geq v_p(n)$, así que $d_p(n,-n) \leq p^{-v_p(n)}$ y, por lo tanto,

$$\begin{aligned} d(\xi, \iota(-n)) &\leq \text{máx}\{d(\xi, \iota(n)), d(\iota(n), \iota(-n))\} \\ &= \text{máx}\{d(\xi, \iota(n)), d(n,-n)\} \\ &\leq \text{máx}\{\varepsilon, p^{-v_p(n)}\} \end{aligned}$$

Sea $\alpha = (a_i)_{n\geq 0}$ una sucesión arbitraria de elementos de $[\![0, p-1]\!]$, y para cada $n \in \mathbb{N}_0$ escribamos $\sigma_n(\alpha) \coloneqq \sum_{i=0}^{n} a_i p^i$. Obtenemos de esta forma una sucesión $(\sigma_n(\alpha))_{n\geq 0}$ de elementos de $\mathbb{N}_0$. Esta sucesión satisface la condición de Cauchy en el espacio métrico $(\mathbb{Z}, d_p)$: si ε es un número positivo, entonces hay un entero positivo n_0 tal que $1/p^{n_0+1} < \varepsilon$, y siempre que m y n son elementos de $\mathbb{N}_0$ tales que $m \geq n \geq n_0$ tenemos que o bien $\sigma_m(\alpha) = \sigma_n(\alpha)$ y, por lo tanto, $d_p(\sigma_m(\alpha), \sigma_n(\alpha)) = 0 < \varepsilon$, o bien $\sigma_m(\alpha) \neq \sigma_n(\alpha)$,

$$v_p(\sigma_m(\alpha) - \sigma_n(\alpha)) = v_p\left(\sum_{i=0}^{m} a_i p^i - \sum_{i=0}^{n} a_i p^i\right) = v_p\left(\sum_{i=n+1}^{m} a_i p^i\right) > n+1,$$

y, por lo tanto, $d_p(\sigma_m(\alpha), \sigma_n(\alpha)) < p^{-n-1} < \varepsilon$. Como la función ι es una inyección isométrica, la sucesión $(\iota(\alpha_n))_{n\geq 0}$ en $\overline{\mathbb{Z}}_p$ también satisface la condición de Cauchy y, como el espacio $\overline{\mathbb{Z}}_p$ es completo, converge a un elemento de $\overline{\mathbb{Z}}_p$, al que escribimos

$$\sum_{i=0}^{\infty} \alpha_i p^i,$$

ya que la sucesión $(\sigma_n(\alpha))_{n\geq 0}$ no es otra que la de las sumas parciales de esta serie. Obtenemos de esta forma una función

$$\sigma : \alpha \in [\![0, p-1]\!]^{\mathbb{N}_0} \mapsto \sum_{i=0}^{\infty} \alpha_i p^i \in \overline{\mathbb{Z}}_p.$$

Queremos probar que esta función σ es sobreyectiva. Sea ξ un elemento de $\overline{\mathbb{Z}}_p$. Como la función $\iota : \mathbb{Z} \to \overline{\mathbb{Z}}_p$ tiene imagen densa, para cada entero no negativo n hay un elemento α_n de $\mathbb{Z}$ tal que $d_p(\iota(\alpha_n), \xi) < p^{-n-1}$, y la teoría del desarrollo en base p nos dice hay exactamente una secuencia $(a_{n,i})_{i\geq 0}$ de elementos de $[\![0, p-1]\!]$ tal que

> *hay exactamente un entero no negativo N_n tal que $a_{n,N_n} \neq 0$ y $a_{n,i} = 0$ siempre que i es un elemento de $\mathbb{N}_0$ tal que $i > N_n$, y se tiene que $\alpha_n = \sum_{i=0}^{N_n} a_i p^i$.*

8.6. Espacios de funciones continuas

Proposición 8.6.1. *Sea X un conjunto, sea Y un espacios métricos, y sea $C_n(X,Y)$ el conjunto de todas las funciones acotadas $X \to Y$.*

(*i*) *Si $f, g : X \to Y$ son dos funciones acotadas, entonces el subconjunto $\{d(f(x), g(x)) : x \in X\}$ de $\mathbb{R}$ es acotado y, por lo tanto, podemos considerar su supremo.*

(*ii*) *La función*

$$d : (f,g) \in C_b(X,Y) \times C_b(X,Y) \mapsto \sup\{d(f(x), g(x)) : x \in X\}$$

es una métrica sobre el conjunto $C_b(X,Y)$.

(*iii*) *Si el espacio Y es completo, entonces también lo es $C_b(X,Y)$.*

Demostración. Sean $f, g : X \to Y$ dos funciones acotadas, de manera que hay un elemento y_0 de Y y un número positivo R tal que $f(x) \in B_R(y_0)$ y $g(x) \in B_R(y_0)$ cualquiera sea el elemento x de X. Para cada elemento x de X tenemos entonces que

$$d(f(x), g(x)) \leq d(f(x), y_0) + d(y_0, g(x)) < 2R,$$

así que el conjunto $\{d(f(x), g(x)) : x \in X\}$ está contenido en $[0, 2R]$ y es, por lo tanto, acotado. Esto prueba la primera de las afirmaciones de la proposición.

Mostremos ahora que la función d definida en la segunda afirmación es una métrica.

- Si $f : X \to Y$ es un elemento cualquiera de $C_b(X,Y)$, entonces $d(f(x), f(x)) = 0$ para todo $x \in X$, así que $d(f,f) = \sup\{d(f(x), f(x)) : x \in X\} = 0$. Por otro lado, si $f, g : X \to Y$ son dos elementos distintos de $C_b(X,Y)$, entonces hay un elemento x_0 de X tal que $f(x_0) \neq g(x_0)$ y, por lo tanto,

  $$0 < d(f(x_0), g(x_0)) \leq \sup\{d(f(x), g(x)) : x \in X\} = d(f,g).$$

  Todo esto prueba que la función d satisface la condición $(\mathbf{M}_1)$.
- Si $f, g : X \to Y$ son dos elementos de $C_b(X,Y)$, entonces para todo $x \in X$ es $d(f(x), g(x)) = d(g(x), f(x))$, así que

  $$d(f,g) = \sup\{d(f(x), g(x)) : x \in X\} = \sup\{d(g(x), f(x)) : x \in X\} = d(g,f).$$

  Vemos así que la función d satisface la condición $(\mathbf{M}_2)$.
- Sean, finalmente, $f, g, h : X \to Y$ tres elementos de $C_b(X,Y)$. Para cada elemento z de X es

  $$\begin{aligned} d(f(z), g(z)) &\leq d(f(z), h(z)) + d(h(z), g(z)) \\ &\leq \sup\{d(f(x), h(x)) : x \in X\} + \sup\{d(h(x), g(x)) : x \in X\} \end{aligned}$$

$$= d(f,h) + d(h,g),$$

así que

$$d(f,g) = \sup\{d(f(x),g(x)) : x \in X\} \leq d(f,h) + d(h,g).$$

Esto nos dice que la función d satisface la condición **(M$_3$)**.

Para terminar la demostración de la proposición, supongamos que el espacio Y es completo y probemos que que el espacio métrico $C_b(X,Y)$ es completo. Sea $(f_n)_{n\geq 0}$ una sucesión en $C_b(X,Y)$ que satisface la condición de Cauchy.

Sea x un elemento de X y sea ε un número positivo. Que la sucesión $(f_n)_{n\geq 0}$ satisfaga la condición de Cauchy nos dice que hay un entero positivo n_0 tal que para toda elección de m y n en $\mathbb{N}_0$ vale $m,n \geq n_0 \implies d(f_m,f_n) < \varepsilon$. Si m y n son dos elementos de $\mathbb{N}_0$ tales que $m,n \geq n_0$, entonces tenemos que

$$d(f_m(x),f_n(x)) \leq d(f_m,f_n) < \varepsilon,$$

y esto nos dice que la sucesión $(f_n(x))_{n\geq 0}$ en Y satisface la condición de Cauchy. Como el espacio Y es completo, esa sucesión tiene un límite, que podemos escribir $f(x)$. Obtenemos de esta forma una función $f : x \in X \mapsto f(x) \in Y$. Mostraremos que f pertenece a $C_b(X,Y)$ y que la sucesión $(f_n)_{n\geq 0}$ converge a f en $C_b(X,Y)$.

- Como la sucesión $(f_n)_{n\geq 0}$ satisface la condición de Cauchy, hay un entero positivo n_1 tal que para toda elección de m y n en $\mathbb{N}_0$ vale $m,n \geq n_1 \implies d(f_n,f_m) < 1$. Por otro lado, la función $f_{n_0} : X \to Y$ es acotada, así que hay un elemento y_0 de Y y un número positivo R tal que $d(f_{n_0}(x),y_0) < R$ para todo $x \in X$.

  Sea x un elemento de X. Como la sucesión $(f_n(x))_{n\geq 0}$ converge a $f(x)$, hay un entero positivo n_2 tal que $d(f_n(x),f(x)) < 1$ siempre que n es un entero tal que $n \geq n_2$. Si ponemos $k \coloneqq \text{máx}\{n_1,n_2\}$, entonces tenemos que

  $$d(f(x),y_0) \leq d(f(x),f_k(x)) + d(f_k(x),f_{n_0}(x)) + d(f_{n_0}(x),y_0) < 1 + 1 + R.$$

  Esto es así cualquiera sea x en X, así que la función f es acotada, esto es, un elemento de $C_b(X,Y)$.
- Sea ε un número positivo. Como la $(f_n)_{n\geq 0}$ satisface la condición de Cauchy, hay un entero positivo n_0 tal que siempre que m y n están en $\mathbb{N}_0$ vale $m,n \geq n_0 \implies d(f_n,f_m) < \varepsilon/3$. Sea n un entero tal que $n \geq n_0$. Si x es un elemento de X, entonces, como la sucesión $(f_n(x))_{n\geq 0}$ converge a $f(x)$, hay un entero m mayor que n_0 tal que $d(f_m(x),f(x)) < \varepsilon/3$ y tenemos entonces que

  $$d(f_n(x),f(x)) \leq d(f_n(x),f_m(x)) + d(f_m(x),f(x)) < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \frac{2\varepsilon}{3}.$$

Esto es cierto cualquiera sea x en X, así que

$$d(f_n, f) = \sup\{d(f_n(x), f(x)) : x \in X\} \leq \frac{2\varepsilon}{3} < \varepsilon.$$

La proposición queda con esto probada. □

En el transcurso de esta demostración probamos el siguiente resultado.

Lema 8.6.2. *Sea X un conjunto, sea Y un espacio métrico. Si $(f_n)_{n\geq 0}$ es una sucesión en $C_b(X, Y)$ que converge a un elemento $f : X \to Y$ de $C_b(X, Y)$, entonces para cada elemento x de X es*

$$f(x) = \lim_{n\to\infty} f_n(x).$$

Demostración. En efecto, si la sucesión $(f_n)_{n\geq 0}$ converge en $C_b(X, Y)$, entonces satisface la condición de Cauchy allí y en la demostración de la proposición vimos que su límite es precisamente la función descripta en el enunciado del lema. □

El espacio $C_b(X, Y)$ es, en general, demasiado grande para las aplicaciones, pero contiene subespacios importantes. El ejemplo más importante es el que da la siguiente proposición.

Proposición 8.6.3. *Sean X e Y dos espacios métricos. El subconjunto $C_b(X, Y)$ de $C_b(X, Y)$ de todas las funciones acotadas $X \to Y$ que son continuas es un subespacio cerrado de $C_b(X, Y)$ que es completo.*

Demostración. Como el espacio $C_b(X, Y)$ es completo, para probar que su subespacio $C_b(X, Y)$ es completo es suficiente, de acuerdo a la Proposición 7.3.26, que mostremos que es cerrado en $C_b(X, Y)$. Sea para ello $(f_n)_{n\geq 0}$ una sucesión en $C_b(X, Y)$ que converge en $C_b(X, Y)$ a una función $f : X \to Y$: tenemos que probar que esta función f es continua. Para ello consideramos una sucesión $(x_n)_{n\geq 0}$ en X que converge a un punto ξ y mostramos que la sucesión $(f(x_n))_{n\geq 0}$ converge a $f(\xi)$.

Para cada $m \in \mathbb{N}_0$ la sucesión $\xi_m \coloneqq (f_n(x_m))_{n\geq 0}$ converge a $f(x_m)$. De hecho, el conjunto de sucesiones $\{\xi_m : m \in \mathbb{N}_0\}$ es equiconvergente en el sentido de la Definición 7.6.2. Para verlo, sea ε un número positivo: como la sucesión $(f_n)_{n\geq 0}$ converge en $C_b(X, Y)$ a f, hay un entero positivo n_0 tal que para todo $n \in \mathbb{N}_0$ vale $n \geq n_0 \implies d(f_n, f) < \varepsilon$ y, por lo tanto, siempre que m y n son elementos de $\mathbb{N}_0$ tales que $n \geq n_0$ es $d(\xi_{m,n}, f(x_m)) = d(f_n(x_m), f(x_m)) \leq d(f_n, f) < \varepsilon$.

Por otro lado, es claro que para cada $n \in \mathbb{N}_0$ la sucesión $(f_n(x_m))_{m\geq 0}$ converge a $f_n(\xi)$, ya que la función f_n es continua. La Proposición 7.6.3 nos permite entonces concluir a partir de todo esto que las sucesiones $(f_n(\xi))_{n\geq 0}$ y $(f(x_m))_{m\geq 0}$ convergen y al mismo límite. Como la primera converge a $f(\xi)$, la segunda también lo hace y esto prueba lo que queremos. □

8.7. El teorema de punto fijo de Banach

Antes de continuar, queremos presentar una aplicación concreta e importante de la teoría que llevamos construida. Empezamos por el siguiente resultado célebre, conocido como el *Teorema de punto fijo de Banach*, por Stefan Banach, que lo enunció originalmente en [Ban1922a].

Proposición 8.7.1. *Sea X un espacio métrico completo no vacío y sea $f : X \to X$ una función. Si hay un número L en $[0,1)$ tal que*

$$d(f(x), f(y)) \leq L \cdot d(x, y)$$

para toda elección de x e y en X, entonces hay exactamente un elemento ξ en X tal que $f(\xi) = \xi$.

Demostración. Supongamos que hay un elemento $L \in [0,1)$ que satisface la hipótesis de la proposición. La función f es continua, ya que satisface la condición de Lipschitz. Como X no es vacío, el *Principio de recurrencia* 1.2.1 nos dice que hay una sucesión $(x_n)_{n\geq 0}$ en X tal que $x_{n+1} = f(x_n)$ para cada $n \in \mathbb{N}_0$. Esto implica que para cada $n \in \mathbb{N}_0$ es

$$d(x_{n+2}, x_{n+1}) = d(f(x_{n+1}), f(x_n)) \leq L \cdot d(x_{n+1}, x_n)$$

y, a su vez, esto nos permite concluir, vía un argumento inductivo evidente, que

$$d(x_{n+1}, x_n) \leq L^n \cdot d(x_1, x_0)$$

cualquiera sea $n \in \mathbb{N}_0$. Como la serie $\sum_{n=0}^{\infty}(L^n \cdot d(x_1, x_0))$ de términos no negativos converge, la Proposición 7.3.6 nos dice que la sucesión $(x_n)_{n\geq 0}$ satisface la condición de Cauchy y, como el espacio X es completo, que converge a un punto ξ de X. Como la función f es continua, la sucesión $(f(x_n))_{n\geq 0}$ converge a $f(\xi)$: como $f(x_n) = x_{n+1}$ para todo $n \in \mathbb{N}_0$, esta sucesión $(f(x_n))_{n\geq 0}$ coincide con $(x_{n+1})_{n\geq 0}$, que es una subsucesión de $(x_n)_{n\geq 0}$ y, por lo tanto, converge también a ξ. Esto nos permite concluir que $f(\xi) = \xi$ y prueba la afirmación de existencia de la proposición.

Para probar la de unicidad, supongamos que ζ es otro elemento de X tal que $f(\zeta) = \zeta$. Tenemos entonces que $d(\xi, \zeta) = d(f(\xi), f(\zeta)) \leq L \cdot d(\xi, \zeta)$, así que $0 \leq (1-L) \cdot d(\xi, \zeta) \leq 0$. Como $1-L > 0$, esto implica que $d(\xi, \zeta) = 0$, esto es, que $\xi = \zeta$. □

Es usual hacer la siguiente definición para referirnos a las funciones que satisfacen la hipótesis del *Teorema de Banach*.

Definición 8.7.2. Sea X un espacio métrico. Un función $f : X \to X$ es una ***contracción*** si hay un número L en $(0,1)$ tal que

$$d(f(x), f(y)) \leq L \cdot d(x, y)$$

cualesquiera sean x e y en X, y en ese caso decimos que f es una contracción de ***razón*** L.

El *Teorema de Banach* nos dice, entonces, que una contracción de un espacio métrico completo posee un y solo un punto fijo.

Ejercicio 8.7.3. Sea X un espacio métrico completo y sea $f : X \to X$ una función sobreyectiva tal que hay un número k mayor que 1 tal que

$$d(f(x), f(y)) \geq k \cdot d(x, y)$$

cualesquiera sean los elementos x e y de X. Pruebe que hay exactamente un punto fijo de f en X.

Corolario 8.7.4. *Sea X un espacio métrico completo no vacío y sea $f : X \to X$ una función. Si hay un entero positivo n tal que la composición f^n es una contracción, entonces f posee exactamente un punto fijo en X.*

Demostración. Supongamos que hay un entero positivo n tal que la composición f^n es una contracción y, por lo tanto, de acuerdo a la Proposición 8.7.1, que hay exactamente un elemento ξ de X tal que $f^n(\xi) = \xi$. Como $f^n(f(\xi)) = f^{n+1}(\xi) = f(f^n(\xi)) = f(\xi)$, el punto $f(\xi)$ también es un punto fijo de f^n y, por lo tanto, $f(\xi) = \xi$. Esto prueba la afirmación de existencia del corolario. Por otro lado, si ζ es otro punto fijo de f, entonces ciertamente es $f^n(\zeta) = \zeta$, así que $\zeta = \xi$. □

Una observación importante es que el punto fijo de una contracción depende de manera continua de ella: la siguiente es una forma de formalizar esta afirmación.

Proposición 8.7.5. *Sean X e Y dos espacios métricos y supongamos que X es completo. Sea $f : X \times Y \to X$ una función continua y sea L un elemento de $(0,1)$ tal que para toda elección de x y x' en X y de y en X vale*

$$d(f(x, y), f(x', y)) \leq L \cdot d(x, x').$$

(i) Para cada $y \in Y$ hay exactamente un elemento x_y de X tal que $f(x_y, y) = x_y$.
(ii) La función $\phi : y \in Y \mapsto x_y \in X$ es continua.

Demostración. Para cada elemento y de Y la función $f_y : x \in X \mapsto f(x, y) \in X$ es una contracción, así que la Proposición 8.7.1 nos dice que hay exactamente un elemento x_y de X tal que $f(x_y, y) = x_y$. Esto prueba la primera parte de la proposición.

Sea $(y_n)_{n\geq 0}$ una sucesión en Y que converge a un punto y de Y. Es

$$d(x_{y_n}, x_y) = d(f(x_{y_n}, y_n), f(x_y, y))$$

$$\leq d(f(x_{y_n}, y_n), f(x_y, y_n)) + d(f(x_y, y_n), f(x_y, y))$$
$$\leq L \cdot d(x_{y_n}, x_y) + d(f(x_y, y_n), f(x_y, y)),$$

así que

$$d(x_y, x_{y_n}) \leq \frac{1}{1-L} d(f(x_y, y_n), f(x_y, y)).$$

Ahora bien, como la función f es continua y sucesión $((x_y, y_n))_{n\geq 0}$ converge a (x_y, y), la sucesión $(d(f(x_y, y_n), f(x_y, y)))_{n\geq 0}$ converge a 0 y, por lo tanto, la sucesión $(x_{y_n})_{n\geq 0}$ converge a x_y. Esto nos dice que la función ϕ del enunciado es continua. □

En lo que resta de esta sección daremos tres aplicaciones bien diferentes del teorema.

8.7.1. Sucesiones definidas por recurrencia

Usando el *Principio de recurrencia* 1.2.1 podemos construir sucesiones de números $(\xi_n)_{n\geq 0}$ en las que cada componente ξ_n está determinada de alguna manera por las componentes anteriores $\xi_0, \xi_1, \ldots, \xi_{n-1}$. El siguiente ejemplo, que es una aplicación sencilla del *Teorema de punto fijo de Banach*, muestra como podemos construir a veces sucesiones en las que cada componente está determinada por todas las que la siguen.

Ejemplo 8.7.1.1. Para cada elemento $x = (x_n)_{n\geq 0}$ de $\ell^\infty_{\mathbb{R}}$ escribamos $T(x)$ a la sucesión que para cada $n \in \mathbb{N}_0$ tiene componente n-ésima

$$T(x)_n = \frac{n}{n+1} + \sum_{m=1}^{\infty} \frac{\sqrt{x_{m+n}^2 + 1}}{2m^2}.$$

Para cada $n \in \mathbb{N}_0$ es

$$|T(x)_n| = \frac{n}{n+1} + \sum_{m=1}^{\infty} \frac{\sqrt{x_{m+n}^2 + 1}}{2m^2} \leq 1 + \sqrt{\|x\|_\infty^2 + 1} \cdot \sum_{m=1}^{\infty} \frac{1}{2m^2} \leq 1 + \frac{\pi^2}{12} \cdot \sqrt{\|x\|_\infty^2 + 1}$$

y esto nos dice que la sucesión $T(x)$ es acotada, esto es, un elemento de $\ell^\infty_{\mathbb{R}}$. Tenemos entonces una función $T : \ell^\infty_{\mathbb{R}} \to \ell^\infty_{\mathbb{R}}$.

Sean ahora $x = (x_n)_{n\geq 0}$ e $y = (y_n)_{n\geq 0}$ dos elementos de $\ell^\infty_{\mathbb{R}}$. El *Teorema del valor medio* nos dice que para cada $n \in \mathbb{N}_0$ hay un número z_n entre x_n e y_n tal que

$$\left|\sqrt{x_n^2 + 1} - \sqrt{y_n^2 + 1}\right| = \left|\frac{z_n}{\sqrt{z_n^2 + 1}} \cdot (x_n - y_n)\right| \leq |x_n - y_n|.$$

Esto implica que para cada $n \in \mathbb{N}_0$ es

$$|T(x)_n - T(y)_n| = \left|\sum_{m=1}^{\infty} \frac{\sqrt{x_{m+n}^2 + 1}}{2m^2} - \sum_{m=1}^{\infty} \frac{\sqrt{y_{m+n}^2 + 1}}{2m^2}\right|$$

$$\leq \sum_{m=1}^{\infty} \frac{1}{2m^2} \left| \sqrt{x_{m+n}^2 + 1} - \sqrt{y_{m+n}^2 + 1} \right|$$
$$= \sum_{m=1}^{\infty} \frac{1}{2m^2} |x_{m+n} - y_{m+n}| \leq \frac{\pi^2}{12} \cdot \|x - y\|_\infty .$$

así que $\|T(x) - T(y)\|_\infty \leq \pi^2/12 \cdot \|x - y\|_\infty$. Vemos así que la función T es una contracción de razón $\pi^2/12 \approx 0{,}82\ldots$ y, como el espacio $\ell_{\mathbb{K}}^\infty$ es completo, el *Teorema de punto fijo de Banach* 8.7.1 nos dice que T posee exactamente un punto fijo en $\ell_{\mathbb{R}}^\infty$. En vista de la definición de la función T, esto significa que hay una y solo una sucesión acotada $(\xi_n)_{n\geq 0}$ de números reales tales que para cada entero no negativo n tiene

$$\xi_n = \frac{n}{n+1} + \sum_{m=1}^{\infty} \frac{\sqrt{\xi_{n+m}^2 + 1}}{2m^2}.$$

El lector probablemente podría entretenerse un buen rato intentando calcular el número ξ_0.

8.7.2. El carácter de Teichmüller

Sea p un número primo y consideremos al función $f : x \in \mathbb{Z} \mapsto x^p \in \mathbb{Z}$. Cada entero x_0 es la 0-ésima componente de una sucesión $(x_n)_{n\geq 0}$ que tiene $x_{n+1} = f(x_n)$ para todo $n \in \mathbb{N}_0$. Por ejemplo, si p es 5 y x_0 es 2, las primeras diez componentes de la sucesión son

$$\begin{array}{r}
2 \\
32 \\
33\,554\,432 \\
42\,535\,295\,865\,117\,307\,932\,921\,825\,928\,971\,026\,432 \\
139\,234\,637\,988 \cdots 313\,788\,274\,787\,963\,704\,661\,873\,320\,009\,853\,338\,386\,432 \\
523\,282\,787\,918 \cdots 610\,455\,326\,513\,450\,434\,370\,695\,277\,350\,467\,175\,186\,432 \\
392\,357\,762\,942 \cdots 138\,750\,356\,976\,170\,535\,402\,511\,232\,578\,336\,359\,186\,432 \\
929\,846\,473\,947 \cdots 375\,670\,614\,566\,811\,017\,687\,621\,221\,837\,682\,279\,186\,432 \\
695\,114\,331\,497 \cdots 690\,750\,860\,071\,945\,173\,935\,926\,496\,134\,411\,879\,186\,432 \\
162\,286\,058\,782 \cdots 911\,858\,830\,949\,449\,159\,746\,336\,067\,618\,059\,879\,186\,432 \\
112\,565\,697\,399 \cdots 020\,356\,158\,325\,247\,203\,020\,463\,925\,036\,299\,879\,186\,432
\end{array} \tag{8.7}$$

Notemos que aquí estamos escribiendo solamente los primeros y últimos dígitos de estos enteros porque son enormes: el décimo número de esta lista es $2^{9\,765\,625}$, que tiene 2 939 747 dígitos. Mirando con atención estas dos tablas podemos observar que los dígitos decimales de los números de la sucesión $(x_n)_{n\geq 0}$ parecen estabilizarse. Queremos probar que esto efectivamente ocurre.

Consideremos la sucesión de enteros $(x_n)_{n\geq 0}$ que tiene $x_0 = 2$ y $x_{n+1} = x_n^5$ para todo $n \in \mathbb{N}_0$, y sea N un entero positivo. Dos enteros positivos a y b tienen los mismos últimos N dígitos

decimales si y solamente si $a \equiv b \mod 10^N$, y el *Teorema chino del resto* nos dice que esto ocurre exactamente cuando $a \equiv b \mod 2^N$ y $a \equiv b \mod 5^N$, y entonces para mostrar que los últimos N dígitos de los números de la sucesión $(x_n)_{n\geq 0}$ se estabilizan tenemos que probar que

> *hay enteros positivos l_N y n_0 tales que $x_n \equiv l_N \mod 2^N$ y $x_n \equiv l_N \mod 5^N$ para todo entero n mayor que n_0.*

De hecho, como 2^N y 5^N son enteros coprimos, el *Teorema chino del resto* implica que es suficiente probar que

> *hay enteros positivos a_N y n_1 tales que $x_n \equiv a_N \mod 2^N$ siempre que n es un entero mayor que n_1.*

y

> *hay enteros positivos b_N y n_2 tales que $x_n \equiv b_N \mod 5^N$ siempre que n es un entero mayor que n_2.* (8.8)

La primera de estas afirmaciones es fácil de probar. En efecto, para todo $n \in \mathbb{N}_0$ tal que $5^n \geq N$ es $x_n = 2^{5^n} \equiv 0 \mod 2^N$, así que podemos poner $b_N \coloneqq 0$ y $n_1 \coloneqq \lceil \log_5 N \rceil$. La segunda afirmación es mucho más interesante. El primer paso para probarla es el siguiente resultado sobre el anillo de los enteros p-ádicos.

Proposición 8.7.2.1. *Sea p un número primo y sea x un elemento de $\mathbb{Z}_p$, el anillo de los enteros p-ádicos. La sucesión $(x_n)_{n\geq 0}$ que tiene $x_0 = x$ y $x_{n+1} = x_n^p$ para todo $n \in \mathbb{N}_0$ converge en $\mathbb{Z}_p$ a un elemento $\omega(x)$, que es único elemento ξ de $\mathbb{Z}_p$ tal que $\xi^p = \xi$ y $\xi \equiv x \mod p\mathbb{Z}_p$.*

Claramente el elemento ξ no es nulo si x no pertenece a $p\mathbb{Z}_p$ — esto es, si x es una unidad de $\mathbb{Z}_p$ — y ξ es en ese caso, por lo tanto, una raíz $(p-1)$-ésima de la unidad del anillo $\mathbb{Z}_p$.

Demostración. El *Pequeño teorema de Fermat* nos dice que $u^p - u \in p\mathbb{Z} \subseteq p\mathbb{Z}_p$ para todo $u \in \mathbb{Z}$. Como $\mathbb{Z}$ es denso en $\mathbb{Z}_p$ y la función $h : u \in \mathbb{Z}_p \mapsto u^p - u \in \mathbb{Z}_p$ es continua, esto implica que $h(\mathbb{Z}_p) = h(\overline{\mathbb{Z}}) \subseteq \overline{h(\mathbb{Z})} \subseteq p\mathbb{Z}_p$.

Consideremos en $\mathbb{Z}_p$ el conjunto $B_{1/p}(x)$, que coincide con el conjunto $x + p\mathbb{Z}_p$. Si y es un elemento de $B_{1/p}(x)$, entonces $y^p - y \in p\mathbb{Z}_p$, así que $y^p \in y + p\mathbb{Z}_p = x + p\mathbb{Z}_p = X$. Vemos así que hay una función

$$F : y \in B_{1/p}(0) \mapsto y^p \in B_{1/p}(0).$$

Mostremos que esta función es una contracción de razón $1/p$. Si u y v son dos elementos distintos de $B_{1/p}(x)$, entonces en $\mathbb{Z}_p$ tenemos que

$$u^p - v^p = (u - v) \cdot \sum_{i=0}^{p-1} u^i v^{p-1-i}.$$

Como $u \equiv v \mod p\mathbb{Z}_p$, para cada $i \in [\![0, p-1]\!]$ tenemos que $u^i v^{p-1-i} \equiv u^{p-1} \mod p\mathbb{Z}_p$ y, por lo

tanto, que

$$\sum_{i=0}^{p-1} u^i v^{p-1-i} \equiv p \cdot u^{p-1} \equiv 0 \mod p\mathbb{Z}_p.$$

Esto implica que si $|-|_p$ es el valor absoluto p-adico de $\mathbb{Z}_p$, entonces

$$\left|\sum_{i=0}^{p-1} u^i v^{p-1-i}\right|_p \leq \frac{1}{p}$$

y, por lo tanto, ya que ese valor absoluto es multiplicativo,

$$|F(u) - F(v)|_p = |u^p - v^p|_p \leq \frac{1}{p} \cdot |u - v|_p.$$

Como la bola $B_{1/p}(x)$ es cerrada en $\mathbb{Z}_p$ y este último espacio es completo, la bola es completa. El *Teorema de punto fijo de Banach* 8.7.1, entonces, nos permite concluir que la función F tiene exactamente un punto fijo ξ en $B_{1/p}(x)$, y que este es el límite de la sucesión $(x_n)_{n\geq 0}$ descripta en el enunciado de la proposición. Esto significa que ξ es el único elemento de $B_{1/p}(x)$ — y por lo tanto el único elemento de $\mathbb{Z}_p$ congruente módulo $p\mathbb{Z}_p$ con x — tal que $\xi^p = \xi$. □

Como consecuencia de esta proposición vemos que hay exactamente una función

$$\omega : \mathbb{Z}_p^\times \to \Omega_{p-1}$$

del grupo $\mathbb{Z}_p^\times$ de las unidades de $\mathbb{Z}_p$ al grupo Ω_{p-1} de las raíces $(p-1)$-ésimas de la unidad de $\mathbb{Z}_p$ tal que $\omega(x) \equiv x \mod p\mathbb{Z}_p$ para todo $x \in \mathbb{Z}_p^\times$. Esta función es el ***carácter de Teichmüller***, por Oswald Teichmüller, y tiene un rol importante en la teoría de los números p-ádicos. La demostración usual de que esta función existe se hace usando el llamado *Lema de Hensel*.

Ejercicio 8.7.2.2. Muestre que el carácter de Teichmüller es un homomorfismo de grupos.

Podemos ahora volver al problema original.

Proposición 8.7.2.3. *Sean p y q dos números primos distintos, consideremos la sucesión de enteros $(x_n)_{n\geq 0}$ que tiene $x_0 = q$ y $x_{n+1} = x_n^p$ para todo $n \in \mathbb{N}_0$, y sea N un entero positivo. Hay enteros positivos a_N y n_0 tales que $x_n \equiv a_N \mod p^N$ para todo $n \in \mathbb{N}_0$ tal que $n \geq n_0$.*

Demostración. La Proposición 8.7.2.1 nos dice que hay exactamente un elemento ξ de $\mathbb{Z}_p$ tal que $\xi^p = \xi$ y $\xi \equiv q \mod p\mathbb{Z}_5$, y que la sucesión $(x_n)_{n\geq 0}$ converge en $\mathbb{Z}_p$ a ξ. En particular, como $\xi + p^N\mathbb{Z}_p$ es un entorno de ξ, hay un entero positivo n_0 tal que $x_{n_0} \in \xi + p^N\mathbb{Z}_p$ y, por lo tanto, $x_{n_0} \equiv \xi \mod p^N\mathbb{Z}_p$. Por otro lado, si n es un elemento de $\mathbb{N}_0$ tal que $x_n \equiv \xi \mod p^N\mathbb{Z}_p$, entonces

$x_{n+1} = x_n^2 \equiv \xi^2 = \xi \mod p^N\mathbb{Z}_p$. Esto nos dice que, de hecho, $x_n \equiv \xi \mod p^N\mathbb{Z}_p$ para todo entero n tal que $n \geq n_0$, y prueba la proposición. □

En particular, esta proposición nos dice que la afirmación (8.8) es cierta y, en consecuencia, que es cierto que los dígitos decimales de la sucesión (8.7) se estabilizan.

8.7.3. El teorema de Lindelöf–Lipschitz–Picard

Fijemos un entero positivo n, sea Ω un subconjunto abierto de $\mathbb{R} \times \mathbb{R}^n$, sea $f : \Omega \to \mathbb{R}^n$ una función, y sea (t_0, x_0) un elemento de Ω. Queremos encontrar un número positivo ρ y una función $x : (t_0 - \rho, t_0 + \rho) \to \mathbb{R}^n$ que satisfaga las siguientes condiciones:

- La función x es derivable en todo su dominio y para todo $t \in (t_0 - \rho, t_0 + \rho)$ el punto $(t, x(t))$ pertenece al conjunto Ω.
- Para todo $t \in (t_0 - \rho, t_0 + \rho)$ es $x'(t) = f(t, x(t))$.
- Finalmente, $x(t_0) = x_0$.

Decimos en ese caso que la función x es una solución del ***problema de valor inicial***

$$\begin{cases} x'(t) = f(t, x(t)), \\ x(t_0) = x_0. \end{cases}$$

En general, este problema no tiene solución. Hay distintos resultados que imponen condiciones de naturaleza diversa sobre la función f que garantizan la existencia de soluciones. A continuación presentaremos uno de ellos, el *Teorema de Lindelöf–Lipschitz–Picard*, por Ernst Lindelöf, Rudolf Lipschitz y Émile Picard.

Proposición 8.7.3.1. *Sea n un entero positivo, sea (t_0, x_0) un elemento de $\mathbb{R} \times \mathbb{R}^n$, sean r y s dos números positivos, y consideremos el conjunto abierto*

$$\Omega := \{(t, x) \in \mathbb{R} \times \mathbb{R}^n : |t - t_0| < r, \|x - x_0\| < s\}.$$

Sea $f : \Omega \to \mathbb{R}^n$ una función continua y acotada tal que hay una número positivo L para el que se tiene que

$$\|f(t, x) - f(t, x')\| \leq L \cdot \|x - x'\|$$

para toda elección de $t \in [t_0 - r, t_0 + r]$ y de x, $x' \in \overline{B}_s(x_0)$, y pongamos

$$M := 1 + \sup\{\|f(t, x)\| : (t, x) \in \overline{\Omega}\}, \qquad \rho := \min\left\{r, \frac{1}{2L}, \frac{s}{M}\right\}.$$

Hay una solución del problema de valor inicial

$$\begin{cases} x'(t) = f(t, x(t)), \\ x(t_0) = x_0 \end{cases}$$

definida sobre el intervalo $(t_0 - \rho, t_0 + \rho)$.

Demostración. Pongamos $b \coloneqq M \cdot \rho$, y consideremos el conjunto

$$X \coloneqq C(B_\rho(t_0), B_b(x_0))$$

de todas las funciones continuas $B_\rho(t_0) \to B_b(x_0)$ dotado de su métrica usual. Notemos que este espacio X es completo, ya que la bola $B_b(x_0)$ es acotada.

Primer paso. Sea $\phi : B_\rho(t_0) \to B_b(x_0)$ un elemento de X. Para cada elemento t de $B_\rho(t_0)$ es

$$(t, \phi(t)) \in B_\rho(t_0) \times B_b(x_0) \subseteq B_r(t_0) \times B_s(x_0) = \Omega,$$

así que podemos considerar la función $\tau \in B_\rho(t_0) \mapsto f(\tau, \phi(\tau)) \in \mathbb{R}^n$. Esta función es continua, así que es integrable en cada subintervalo cerrado de $B_\rho(t_0)$, y, por lo tanto, para cada $t \in B_\rho(t_0)$ podemos considerar el vector

$$\mathscr{F}(\phi)(t) = x_0 + \int_{t_0}^{x_0} f(\tau, \phi(\tau))\,\mathrm{d}\tau.$$

Obtenemos de esta forma una función $F(\phi) : B_\rho(t_0) \to \mathbb{R}^n$, y el *Teorema fundamental del cálculo* nos dice que esta función es derivable en su dominio, así que, en particular, es allí continua. Para cada elemento t de $B_\rho(t_0)$ tal que $t \geq t_0$ tenemos que

$$\begin{aligned}\|F(\phi)(t) - x_0\| = \left\| \int_{t_0}^{x_0} f(\tau, \phi(\tau))\,\mathrm{d}\tau \right\| &\leq \int_{t_0}^{t} \|f(\tau, \phi(\tau))\|\,\mathrm{d}\tau \leq \int_{t_0}^{t} \|f(\tau, \phi(\tau))\|\,\mathrm{d}\tau \\ &\leq \int_{t_0}^{t} M\,\mathrm{d}\tau = M \cdot |t - t_0| < M \cdot a = b.\end{aligned}$$

De manera similar, si t es un elemento de $B_\rho(t_0)$ tal que $t < t_0$, entonces

$$\begin{aligned}\|F(\phi)(t) - x_0\| = \left\| \int_{t_0}^{x_0} f(\tau, \phi(\tau))\,\mathrm{d}\tau \right\| &\leq \int_{t}^{t_0} \|f(\tau, \phi(\tau))\|\,\mathrm{d}\tau \leq \int_{t}^{t_0} \|f(\tau, \phi(\tau))\|\,\mathrm{d}\tau \\ &\leq \int_{t}^{t_0} M\,\mathrm{d}\tau = M \cdot |t - t_0| < M \cdot \rho = b.\end{aligned}$$

Esto nos dice que $F(\phi)(t) \in B_b(x_0)$ para todo $t \in B_\rho(t_0)$ y, por lo tanto, que podemos correstringir el codominio de $F(\phi)$ y verla como una función $F(\phi) : B_\rho(t_0) \to B_b(x_0)$ y, en consecuencia, como un elemento del espacio X. Tenemos así una función $F : X \to X$.

Segundo paso. Sean $\phi, \psi : B_\rho(t_0) \to B_b(x_0)$ dos elementos de X, y sea t un elemento de $B_\rho(t_0)$. Si $t \geq t_0$, entonces

$$\|F(\phi)(t) - F(\psi)(t)\| = \left\| \int_{t_0}^{t} f(\tau, \phi(\tau))\,\mathrm{d}\tau - \int_{t_0}^{t} f(\tau, \psi(\tau))\,\mathrm{d}\tau \right\|$$

$$
\begin{aligned}
&= \left\| \int_{t_0}^{t} \big(f(\tau, \phi(\tau)) - f(\tau, \psi(\tau)) \big) \, \mathrm{d}\tau \right\| \\
&\leq \int_{t_0}^{t} \big\| f(\tau, \phi(\tau)) - f(\tau, \psi(\tau)) \big\| \, \mathrm{d}\tau \\
&\leq \int_{t_0}^{t} L \cdot \big\| \phi(\tau) - \psi(\tau) \big\| \, \mathrm{d}\tau \\
&\leq \int_{t_0}^{t} L \cdot \big\| \phi - \psi \big\| \, \mathrm{d}\tau \\
&= L \cdot |t - t_0| \cdot \| \phi - \psi \| < L \cdot a \cdot \| \phi - \psi \|.
\end{aligned}
$$

Un cálculo similar muestra que la desigualdad $\|F(\phi)(t) - F(\psi)(t)\| < L \cdot a \cdot \|\phi - \psi\|$ también vale si t es un elemento de $B_\rho(t_0)$ tal que $t < t_0$, así que lo que tenemos es que

$$
\|F(\phi) - F(\psi)\| \leq L \cdot \rho \cdot \|\phi - \psi\| \leq \frac{1}{2} \cdot \|\phi - \psi\|
$$

Vemos así que la función $F : X \to X$ es una contracción de razón $1/2$.

Segundo paso. El *Teorema de punto fijo de Banach* 8.7.1 nos permite concluir entonces que hay en X exactamente un punto fijo de la función F. Sea $\phi : B_\rho(t_0) \to B_b(x_0)$ ese punto fijo, de manera que $\phi = F(\phi)$, esto es, para cada $t \in B_\rho(t_0)$ es

$$
\phi(t) = x_0 + \int_{t_0}^{t} f(\tau, \phi(\tau)) \, \mathrm{d}\tau. \tag{8.9}
$$

Como ϕ pertenece al conjunto X, vimos arriba que para todo $t \in B_\rho(t_0)$ tenemos que $(t, \phi(t)) \in \Omega$. Por otro lado, observamos arriba que la función F toma por valores funciones que son derivables en todo el intervalo $B_\rho(t_0)$: derivando a ambos lados de la igualdad (8.9) y teniendo en cuenta al *Teorema fundamental del cálculo*, vemos que para todo $t \in B_\rho(t_0)$ es

$$
\phi'(t) = f(t, \phi(t)).
$$

Finalmente, cuando $t = t_0$ la igualdad (8.9) nos dice que $\phi(t_0) = x_0$. Vemos así que la función ϕ es una solución del problema de valor inicial descripto en el enunciado de la proposición sobre el intervalo $(t_0 - \rho, t_0 + \rho)$. Esto prueba la proposición. □

Ejemplo 8.7.3.2. Pongamos $n := 1$, sea r un número mayor que 1, consideremos el subconjunto abierto $\Omega := (-r, r) \times (-r, r)$ de $\mathbb{R} \times \mathbb{R}$ y consideremos la función

$$
f : (t, x) \in \Omega \mapsto x \in \mathbb{R}
$$

y el punto $(t_0, x_0) := (0, 1)$ de $\mathbb{R} \times \mathbb{R}$. Es claro que cualesquiera sean $t \in (-r, r)$ y $x, x' \in (-r, r)$

vale

$$|f(t,x) - f(t,x')| \le |x - x'|,$$

así que en la situación de la Proposición 8.7.3.1 podemos tomar $L = 1$. Sean, por otro lado,

$$M := 1 + \sup\{|f(t,x)| : (t,x) \in \Omega\} = 1 + r, \qquad \rho := \min\left\{r, \frac{1}{2L}, \frac{r}{M}\right\} = \frac{1}{2}.$$

De acuerdo a la Proposición 8.7.3.1, hay una solución del problema de valor inicial

$$\begin{cases} x'(t) = x(t), \\ x(0) = 1 \end{cases} \tag{8.10}$$

definida sobre un intervalo $(-1/2, 1/2)$. De hecho, sabemos de la demostración del *Teorema de punto fijo de Banach* y de la de la Proposición 8.7.3.1 que podemos obtener esa solución de la siguiente manera. Si $\phi_0 : (-1/2, 1/2) \to B_{M/2}(1)$ es una función continua cualquiera, podemos construir la sucesión $(\phi_n)_{n\ge 0}$ tal que $\phi_{n+1} = F(\phi)$ para todo $n \in \mathbb{N}_0$: la función ϕ es el límite de esa sucesión en el espacio $C(B_{1/2}(0) < B_{M/2}(1))$.

Elijamos $\phi_0 : t \in B_{1/2}(0) \mapsto 1 \in B_{M/2}(1)$. Para cada $t \in B_{1/2}(0)$ es, entonces,

$$\begin{aligned}
\phi_1(t) &= F(\phi_0)(t) = 1 + \int_0^t \phi_0(\tau)\,\mathrm{d}\tau = 1 + \int_0^t \mathrm{d}\tau = 1 + t, \\
\phi_2(t) &= F(\phi_1)(t) = 1 + \int_0^t \phi_1(\tau)\,\mathrm{d}\tau = 1 + \int_0^t (1+\tau)\mathrm{d}\tau = 1 + t + \frac{t^2}{2!}, \\
\phi_3(t) &= F(\phi_2)(t) = 1 + \int_0^t \phi_2(\tau)\,\mathrm{d}\tau = 1 + \int_0^t \left(1 + \tau + \frac{\tau^2}{2!}\right)\mathrm{d}\tau = 1 + t + \frac{t^2}{2!} + \frac{t^3}{3!},
\end{aligned}$$

y es claro que, más generalmente, para cada $n \in \mathbb{N}_0$ es

$$\phi_n(t) = \sum_{i=0}^{n} \frac{t^i}{i!}.$$

En efecto, es fácil probar esta igualdad haciendo un razonamiento inductivo que dejamos a cargo del lector. Se sigue de esto que hay una solución $\phi : (-1/2, 1/2) \to \mathbb{R}$ del problema de valor inicial (8.10) que tiene, para cada $t \in (-1/2, 1/2)$,

$$\phi(t) = \lim_{n\to\infty} \phi_n(t) = \lim_{n\to\infty} \sum_{i=0}^{n} \frac{t^i}{i!} = \sum_{i=0}^{\infty} \frac{t^i}{i!} = e^t.$$

El siguiente ejercicio implica que es, de hecho, la única solución en ese intervalo.

Ejercicio 8.7.3.3. Pruebe que la situación de la Proposición 8.7.3.1 hay exactamente una solución del problema de valor inicial allí considerado definida en el intervalo $(t_0 - \rho, t_0 + \rho)$.

8.8. Funciones lineales

Una función entre espacios métricos es continua si es continua en *todos* los puntos de su dominio. Hay, sin embargo, una situación extremadamente importante en la que esa suficiente con que sea continua en *un* punto.

Proposición 8.8.1. *Sean G y H dos grupos abelianos dotados de métricas invariantes por translaciones. Una función $f : G \to H$ que es un homomorfismo de grupos es continua si y solamente si es continua en un punto de G.*

Recordemos que en esta situación, siempre que N es un entorno de un elemento x de G e y es un elemento de G, el conjunto $(y - x) + N$ es un entorno de y.

Demostración. Sea $f : G \to H$ un homomorfismo de grupos. Es claro que la condición de la proposición es necesaria para que la función f sea continua. Para ver que también es suficiente, supongamos que la función f es continua en un elemento y de G y mostremos que también es continua en cualquier otro elemento de G.

Sea x un elemento de G y sea N un entorno de $f(x)$. El conjunto $(f(y) - f(x)) + N$ es un entorno de $f(y)$, así que la hipótesis sobre f implica que el conjunto $f^{-1}((f(y) - f(x)) + N)$ es un entorno de y y

$$P \coloneqq (x - y) + f^{-1}((f(y) - f(x)) + N)$$

un entorno de x. Si z es un elemento de P, entonces hay un elemento u de $(f(y) - f(x) + N)$ tal que $z = (x - y) + u$ y es

$$f(z) = (f(x) - f(y)) + f(u) \in (f(x) - f(y)) + (f(y) - f(x)) + N = N,$$

de manera que $z \in f^{-1}(N)$. Vemos así que $P \subseteq f^{-1}(N)$ y, por lo tanto, que $f^{-1}(N)$ es un entorno de x. Esto prueba que la función f es continua en x. □

Podemos especializar esta proposición al caso de las funciones lineales entre espacios vectoriales con métricas invariantes por translaciones.

Corolario 8.8.2. *Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, y sean V y W dos espacios vectoriales sobre $\Bbbk$ dotados de métricas invariantes por translaciones. Una función lineal $f : V \to W$ es continua si y solamente si es continua en el vector nulo de V.* □

Demostración. En efecto, una función lineal es, en particular, un homomorfismo de los grupos abelianos subyacentes a los espacios vectoriales, así que esto es consecuencia inmediata de la proposición. □

Si restringimos nuestra atención a los espacios normados, entonces podemos dar un criterio más preciso.

Corolario 8.8.3. *Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$ y sean V y W dos $\mathbb{k}$-espacios normados. Una función lineal $f : V \to W$ es continua si y solamente si hay un número real no negativo C tal que $\|f(x)\| \leq C \cdot \|x\|$ para todo $x \in V$.*

Demostración. Como las métricas de V y W que corresponden a sus normas son invariantes por translaciones, en vista del Corolario 8.8.2 tenemos que mostrar que

$$\text{una función lineal } f : V \to W \text{ es continua en } 0 \text{ si y solamente si hay un número real no negativo } C \text{ tal que } \|f(x)\| \leq C\|x\| \text{ para todo } x \in V. \tag{8.11}$$

Supongamos primero que $f : V \to W$ es una función lineal continua. Como $B_1^W(0)$ es un entorno de 0 en W, el conjunto $f^{-1}(B_1^W(0))$ es un entorno de 0 en V y hay, por lo tanto, un número positivo r tal que $B_r^V(0) \subseteq f^{-1}(B_1^W(0))$. Pongamos $C \coloneqq 2/r$. Si x es un elemento no nulo de V, entonces $rx/2\|x\|$ es un elemento de $B_r^V(0)$, así que esto nos dice que $f(rx/2\|x\|) \in B_1^W(0)$, esto es, que $\|rf(x)/2\|x\|\| \leq 1$, y esto implica que $\|f(x)\| \leq 2\|x\|/r = C \cdot \|x\|$. Como esta última desigualdad también vale si x es 0, vemos que la condición de (8.11) es necesaria.

Para probar que también es necesaria, supongamos que $f : V \to W$ es una función lineal y que hay un número no negativo C tal que $\|f(x)\| \leq C \cdot \|x\|$ para todo $x \in V$. Si N es un entorno de 0 en W, entonces hay un número positivo r tal que $B_r^W(0) \subseteq N$, y para todo vector x de $B_{r/C}(0)$ tenemos que $\|f(x)\| \leq C \cdot \|x\| < C \cdot r/C = r$: esto nos dice que $B_{r/C}(0) \subseteq f^{-1}(N)$ y, por lo tanto, que $f^{-1}(N)$ es un entorno de 0 en V. Vemos así que la función f es continua en 0, como queremos. □

Gracias al Corolario 8.8.3 podemos hacer la siguiente definición.

Definición 8.8.4. Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, sean V y W dos $\mathbb{k}$-espacios normados, y sea $f : V \to W$ una función lineal continua. La ***norma*** de f es el número

$$\|f\| \coloneqq \inf\{C \in \mathbb{R} : \|f(x)\| \leq C \cdot \|x\| \text{ para todo } x \in V\}.$$

En efecto, ese corolario nos dice que el conjunto cuyo ínfimo estamos tomando aquí no es vacío — notemos que que está acotado inferiormente por 0 es evidente.

Podemos escribir la condición de este corolario en una forma un poco distinta:

Corolario 8.8.5. *Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$ y sean V y W dos $\mathbb{k}$-espacios normados. Una función lineal $f : V \to W$ es continua si y solamente si el conjunto $f(B_1^V(0))$ de W es acotado.*

Demostración. Sea $f : V \to W$ una función lineal. Si f es continua, entonces el Corolario 8.8.3 nos dice que hay un número real no negativo C tal que $\|f(x)\| \leq C \cdot \|x\|$ para todo $x \in V$. En particular, si $x \in B_1^V(0)$, entonces $\|f(x)\| \leq C \cdot \|x\| < C$: esto nos dice que el conjunto $f(B_1^V(0))$ está contenido en $B_C^W(0)$ y que, por lo tanto, es acotado.

Recíprocamente, si ese conjunto es acotado, de manera que hay un número positivo C tal que $f(B_1^V(0)) \subseteq B_C^W(0)$, entonces para todo vector no nulo x de V tenemos que $\|f(x/2\|x\|)\| \leq C$, de manera que $\|f(x)\| \leq 2C \cdot \|x\|$. Como esta desigualdad también vale si el vector x es nulo, el Corolario 8.8.3 nos dice que la función f es continua. □

Veamos algunos ejemplos de funciones lineales continuas.

Proposición 8.8.6. *Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, sea V un $\mathbb{k}$-espacio normado de dimensión finita, y sea W un $\mathbb{k}$-espacio normado arbitrario. Toda función lineal $f : V \to W$ es continua.*

Demostración. Sean $\|-\|_V$ y $\|-\|_W$ las normas de V y de W, respectivamente, y sea $f : V \to W$ una función lineal. Mostremos que la función

$$\|-\|'_V : x \in V \mapsto \|x\|_V + \|f(x)\|_W \in \mathbb{R}$$

es una norma en V.

- Si x es un elemento de V tal que $\|x\|'_V = 0$, entonces $\|x\|_V \leq \|x\|_V + \|f(x)\|_W = \|x\|'_V = 0$ y, por lo tanto, $x = 0$.
- Si x es un elemento de V y t uno de $\mathbb{k}$, entonces

$$\begin{aligned}\|tx\|'_V &= \|tx\|_V + \|f(tx)\|_W = |t| \cdot \|x\|_V + \|tf(x)\|_W = |t| \cdot \|x\|_V + |t| \cdot \|f(x)\|_W \\ &= |t| \cdot (\|x\|_V + \|f(x)\|_W) = |t| \cdot \|x\|'_V.\end{aligned}$$

- Si x e y son dos elementos de V, entonces

$$\begin{aligned}\|x + y\|'_V &= \|x + y\|_V + \|f(x + y)\|_W \leq \|x\|_V + \|y\|_V + \|f(x) + f(y)\|_W \\ &\leq \|x\|_V + \|y\|_V + \|f(x)\|_W + \|f(y)\|_W = \|x\|'_V + \|y\|'_V.\end{aligned}$$

Ahora bien, como el espacio V tiene dimensión finita, todas sus normas son equivalentes y, en particular, hay una constante positiva c tal que $\|x\|'_V \leq c \cdot \|x\|_V$ para todo $x \in V$. Esto implica, claro que para todo elemento x de V es

$$\|f(x)\|_W \leq \|x\|_V + \|f(x)\|_W = \|x\|'_V \leq c \cdot \|x\|_V$$

y, por lo tanto, que la función f es continua. □

El resultado de esta proposición es, de hecho, una caracterización de los espacios normados

de dimensión finita entre todos los espacios normados.

Proposición 8.8.7. *Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$. Si V es un $\Bbbk$-espacio normado de dimensión infinita finita, entonces hay una función lineal $f : V \to \Bbbk$ que no es continua.*

Demostración. Sea V un $\Bbbk$-espacio normado de dimensión infinita, y sea $\mathscr{B}$ una base de V. Como V tiene dimensión infinita, hay una función inyectiva $\phi : \mathbb{N}_0 \to \mathscr{B}$. Sea I la imagen de ϕ y sea $\psi : I \to \mathbb{N}_0$ la función inversa de la correstricción $\phi|^I : \mathbb{N}_0 \to I$. Como $\mathscr{B}$ es una base del espacio vectorial V, hay exactamente una función lineal $f : V \to \Bbbk$ que en cada elemento x de $\mathscr{B}$ toma el valor

$$f(x) = \begin{cases} 2\|x\|\psi(x) & \text{si } x \in I; \\ 0 & \text{si } x \in \mathscr{B} \setminus I. \end{cases}$$

Mostremos que esta función no es continua: para ello es suficiente, de acuerdo al Corolario 8.8.5, que mostremos que el conjunto $\{f(x) : x \in B_1^V(0)\}$ de $\Bbbk$ no es acotado. Esto es fácil: para cada $n \in \mathbb{N}_0$ el vector $x \coloneqq \phi(n)/2\|\phi(n)\|$ está en $B_1(V)$ y es

$$\|f(x)\| = \left\| \frac{f(\phi(n))}{2\|\phi(n)\|} \right\| = \left\| \frac{2\|\phi(n)\|n}{2\|\phi(n)\|} \right\| = n. \qquad \square$$

Veamos ahora un ejemplo de función continua entre espacios normados de dimensión infinita.

Proposición 8.8.8. *Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, y sean p y q dos elementos de $[1, +\infty]$ tales que $p < q$. La inclusión $\iota : \ell^p_\Bbbk \hookrightarrow \ell^q_\Bbbk$ es continua y su norma es* 1.

Recordemos que en el Ejercicio 5.2.3.27 vimos que en esta situación $\ell^p_\Bbbk$ es un subespacio de $\ell^q_\Bbbk$, así que hay efectivamente una función de inclusión.

Demostración. Sea $x = (x_n)_{n\geq 0}$ un elemento no nulo de $\ell^p_\Bbbk$ y pongamos $t \coloneqq \|x\|_p$. Para cada $n \in \mathbb{N}_0$ es $0 \leq |x_n/t| \leq 1$ y, por lo tanto, $|x_n/t|^q \leq |x_n/t|^p$, así que

$$\frac{\|x\|_q^q}{t^q} = \left\| \frac{x}{t} \right\|_q^q = \sum_{n=0}^{\infty} \left| \frac{x_n}{t} \right|^q \leq \sum_{n=0}^{\infty} \left| \frac{x_n}{t} \right|^p = \left\| \frac{x}{t} \right\|_p^p = 1.$$

Esto implica que $\|x\|_q \leq t = \|x\|_p$. Hemos probado esta desigualdad cuando x es un elemento no nulo de $\ell^p_\Bbbk$, y claramente vale también cuando $x = 0$, así que la inclusión $\iota : \ell^p_\Bbbk \to \ell^q_\Bbbk$ es continua y su norma es $\|\iota\| \leq 1$.

Queremos ver que, de hecho, $\|\iota\| = 1$. Sea ε un real positivo. Sabemos que $\ell^p_\Bbbk \subsetneq \ell^q_\Bbbk$, así que hay un elemento $x = (x_n)_{n\geq 0}$ de $\ell^q_\Bbbk$ tal que $\|x\|_q = 1$ y la serie $\sum_{n=0}^{\infty} |x_n|^p$ diverge. Esto implica,

claramente, que

$$\lim_{N\to\infty} \frac{\sum_{n=0}^{N}|x_i|^q}{\sum_{n=0}^{N}|x_i|^p} = 0.$$

Esto implica que hay un entero positivo N tal que

$$(1-\varepsilon) < \sum_{n=0}^{N}|x_i|^q < \varepsilon \cdot \sum_{n=0}^{N}|x_i|^p. \tag{8.12}$$

Sea $y = (y_n)_{n\geq 0}$ el elemento de $\mathbb{k}^{\mathbb{N}_0}$ que para cada $n \in \mathbb{N}_0$ tiene

$$y_n = \begin{cases} x_n & \text{si } n \leq N; \\ 0 & \text{si } n > N. \end{cases}$$

Es claro que $y \in \ell^p_{\mathbb{k}}$, y la desigualdad (8.12) nos dice que

$$(1-\varepsilon)^{1/q} \leq \|y\|_q \leq \varepsilon \cdot \|y\|_p^{p/q}.$$

Es $\|y\|_p \leq \|x\|_p$ HACER. □

8.9. Conjuntos de continuidad

Por supuesto, una función entre espacios métricos puede ser continua en algunos puntos de su dominio y no en otros. A veces es importante entender el conjunto de los puntos donde sí lo es.

Definición 8.9.1. Un subconjunto de un espacio métrico es un ***conjunto*** G_δ si es la intersección de un conjunto contable de abiertos de X, y es un ***conjunto*** F_σ si es la unión de un conjunto contable de cerrados de X.

Las *Leyes de De Morgan* y la Proposición 6.5.16 implican, claro, que un subconjunto de X es un conjunto G_δ y si solamente si su complemento en X es un conjunto F_σ.

Veamos algunos ejemplos.

Ejemplo 8.9.2. Mostremos que

todo conjunto cerrado de un espacio métrico es un conjunto G_δ. (8.13)

Sea X un espacio métrico, sea F un subconjunto cerrado de X, y para cada entero positivo n sea $U_n := \{x \in X : f(x,F) < 1/n\}$. Como la función $x \in X \mapsto d(x,F) \in \mathbb{R}$ es continua, es claro que U_n es un subconjunto abierto de X para todo $n \in \mathbb{N}_0$. Por otro lado, la intersección $\bigcap_{n=1}^{\infty} U_n$ coincide

con F. En efecto, un punto x de X pertenece a esa intersección exactamente cuando $d(x, F) = 0$ y, como F es cerrado en X, sabemos que esto ocurre si y solamente si x pertenece a F.

Por dualidad, a partir de (8.13) vemos que también es cierto que

todo conjunto abierto de un espacio métrico es un conjunto F_σ.

Ejercicio 8.9.3. Muestre que el conjunto de los subconjuntos G_δ de un espacio métrico es cerrado por uniones finitas e intersecciones contables.

Observación 8.9.4. Las clases G_δ y F_σ son el principio de la llamada ***jerarquía de Borel***. Estas son las dos sucesiones $(\Sigma_n^0)_{n\geq 1}$ y $(\Pi_n^0)_{n\geq 1}$ de subconjuntos de un espacio métrico que se construyen recursivamente de la siguiente manera:

- Σ_1^0 es el conjunto de todos los abiertos de X y Π_1^0 el de todos los cerrados de X.
- Para cada $n \in \mathbb{N}$ el conjunto Σ_{n+1}^0 es el conjunto de todas las uniones contables de elementos de $\bigcup_{i=1}^n \Pi_i^0$, y Π_{n+1}^0 es el conjunto de todas las intersecciones contables de elementos de $\bigcup_{i=1}^n \Sigma_i^0$.

Así, Σ_2^0 es el conjunto de todos los conjuntos F_σ de X y Π_2^0 el de todos los conjuntos G_δ de X. Las dos sucesiones pueden, de hecho, continuarse a secuencias transfinitas $(\Sigma_\alpha^0)_{\alpha\in \mathsf{Ord}}$ y $(\Pi_\alpha^0)_{\alpha\in \mathsf{Ord}}$. Las primeras componentes de estas secuencias son importantes en análisis — sobre todo en la teoría de la medida — y las jerarquías completas son uno de los objetos de estudio de la llamada teoría descriptiva de conjuntos

El lector puede encontrar información sobre esto en los libros [Kec1995] y [Sri1998] de Alexander Kechris y Shashi Mohan Srivastava, respectivamente.

Estas definiciones nos interesan aquí aquí por el siguiente resultado.

Proposición 8.9.5. *Sean X e Y dos espacios métricos. El conjunto de puntos de continuidad de una función $f : X \to Y$ es un subconjunto G_δ de X, y el de sus puntos de discontinuidad es un subconjunto F_σ de X.*

Esto nos da una condición fuerte sobre un subconjunto de un espacio métrico para que sea el conjunto de puntos de continuidad de una función continua.

Demostración. Es suficiente probar la primera afirmación de la proposición, ya que la segunda es consecuencia inmediata de ella.

Sea $f : X \to Y$ una función, y para cada entero positivo n sea A_n el conjunto de los puntos x de X tales que hay un número positivo δ tal que para toda elección de y y z en $B_\delta(x)$ se tiene que $d(f(y), f(z)) < 1/n$. Para probar la proposición mostraremos que para todo entero positivo n el

conjunto A_n es abierto y que la intersección $\bigcap_{n=1}^{\infty} A_n$ es precisamente el conjunto de puntos de continuidad de f.

- Sea n un entero positivo y sea x un elemento de A_n, de manera que hay un número positivo δ tal que siempre que y y z son elementos de $B_\delta(x)$ es $d(f(y), f(z)) < 1/n$. Sea x' un elemento de $B_\delta(x)$ y sea $\delta' \coloneqq \delta - d(x, x')$, que es un número positivo. Si y y z son dos elementos de $B_{\delta'}(z')$, entonces $d(y, x) \leq d(y, x) + d(x', x) < (\delta - d(x, x')) + d(x, x') = \delta$, así que $y \in B_\delta(x)$, y, por una razón similar, $z \in B_\delta(x)$: la forma en que elegimos δ implica entonces que $d(f(y), f(z)) < 1/n$. Vemos así que $B_\delta(x) \subseteq A_n$ y, por lo tanto, que el conjunto A_n es abierto.
- Sea x un elemento de la intersección $\bigcap_{n=1}^{\infty} A_n$ y sea ε un número positivo. Hay un entero positivo n tal que $1/n < \varepsilon$ y, como x pertenece al conjunto A_n, un número positivo δ tal que siempre que y y z son elementos de $B_\delta(x)$ se tiene que $d(f(y), f(z)) < 1/n$. En particular, para todo elemento y de $B_\delta(x)$ tenemos que $d(f(y), f(x)) < 1/n < \varepsilon$. Esto nos permite concluir que la función f es continua en x.
- Finalmente, sea x un punto de X en el que la función f es continua, y sea n un entero positivo. La continuidad en x de f nos dice que hay un número positivo δ tal que siempre que y es un elemento de $B_\delta(x)$ es $d(f(y), f(x)) < 1/2n$. Si ahora y y z son dos elementos de $B_\delta(x)$, entonces $d(f(y), f(z)) \leq d(f(y), f(x)) + d(f(x), f(z)) < 1/2n + 1/2n = 1/n$. Vemos así que el punto x pertenece a A_n. Esto es así cualquiera sea el entero positivo n, así que vemos que todo punto de continuidad de f pertenece a la intersección $\bigcap_{n=1}^{\infty} A_n$. □

Usando la idea de esta demostración podemos cuantificar cuán lejos de ser continua es una función en un punto. Eso es lo que hace la función ω_f del siguiente ejercicio.

Ejercicio 8.9.6. Sean X e Y dos espacios métricos y sea $f : X \to Y$ una función. Para cada $x \in X$ y cada número positivo δ podemos considerar

$$\Omega_f(x, \delta) \coloneqq \sup\{d(f(y), f(z)) : y, z \in B_\delta(x)\},$$

que puede ser un elemento de $[0, +\infty)$ o, si el conjunto que estamos considerando no está acotado superiormente, $+\infty$. Podemos interpretar a $\Omega_f(x, \delta)$ como una medida de la *oscilación* de la función f en la bola $B_\delta(x)$. Definimos una función $\omega_f : X \to [0, +\infty]$ poniendo, para cada elemento x de X,

$$\omega_f(x) \coloneqq \text{ínf}\{\omega_f(x, \delta) : \delta \in (0, +\infty)\}.$$

Pruebe que la función f es continua en un punto x de X si y solamente si $\omega_f(x) = 0$.

No es cierto que todo subconjunto G_δ de un espacio métrico X sea el conjunto de puntos de continuidad de una función $X \to Y$. Sí es cierto, sin embargo, el siguiente caso particular de esa

afirmación.

Proposición 8.9.7. *Todo subconjunto G_δ de $\mathbb{R}$ es el conjunto de puntos de continuidad de una función $\mathbb{R} \to \mathbb{R}$.*

Demostración. Sea A un subconjunto G_δ de $\mathbb{R}$, de manera que hay una sucesión $(U_n)_{n\geq 0}$ de subconjuntos abiertos de $\mathbb{R}$ tal que $A = \bigcap_{n=0}^{\infty} U_n$; sin pérdida de generalidad podemos suponer que $U_0 = \mathbb{R}$. Para cada $n \in \mathbb{N}_0$ pongamos $V_n \coloneqq \bigcap_{i=0}^{n} U_n$. La sucesión $(V_n)_{n\geq 0}$ es otra sucesión de subconjuntos abiertos de $\mathbb{R}$ y claramente $A = \bigcap_{n=0}^{\infty} V_n$. Notemos que $V_0 = \mathbb{R}$ y que $V_m \subseteq V_n$ siempre que m y n son elementos de $\mathbb{N}_0$ tales que $m \geq n$.

Para cada $n \in \mathbb{N}$ pongamos

$$D_n \coloneqq V_n \setminus V_{n+1}, \qquad E_n \coloneqq D_n \setminus (D_n^\circ \cap \mathbb{Q}).$$

Si m y n son elementos de $\mathbb{N}_0$ tales que $m < n$, entonces $D_m \cap V_n \subseteq D_m \cap V_{m+1} = \varnothing$ y $D_n \subseteq V_n$, así que $D_m \cap D_n = \varnothing$. Esto nos dice que los conjuntos de la sucesión $(D_n)_{n\geq 0}$ son disjuntos dos a dos y, como $E_n \subseteq D_n$ para todo $n \in \mathbb{N}_0$, que también lo son los de la sucesión $(E_n)_{n\geq 0}$. Notemos que

$$\bigcup_{n=0}^{\infty} D_n = \bigcup_{n=0}^{\infty} (V_n \setminus V_{n+1}) = V_0 \setminus \bigcap_{n=0}^{\infty} V_n = \mathbb{R} \setminus A,$$

de manera que, como $\bigcup_{n=0}^{\infty} E_n \subseteq \bigcup_{n=0}^{\infty} D_n$,

$$A \cap \bigcup_{n=0}^{\infty} E_n = \varnothing. \tag{8.14}$$

Por otro lado, para todo $n \in \mathbb{N}_0$ es $E_n^\circ = \varnothing$. En efecto, si hubiera un elemento y de $\mathbb{R}$ y un número positivo δ tal que $B_\delta(y) \subseteq E_n = D_n \setminus (D_n^\circ \cap \mathbb{Q})$, tendríamos que $B_\delta(y) \subseteq D_n$, de manera que $B_\delta(y) \subseteq D_n^\circ$, y que

$$\varnothing = B_\delta(y) \cap D_n^\circ \cap \mathbb{Q} = B_\delta(y) \cap \mathbb{Q},$$

lo que es absurdo.

Hay una función $f : \mathbb{R} \to \mathbb{R}$ tal que

$$f(x) = \begin{cases} \dfrac{1}{n+2} & \text{si } x \in E_n \text{ para algún } n \in \mathbb{N}_0; \\ 0 & \text{si } x \notin \bigcup_{n=0}^{\infty} E_n. \end{cases}$$

Vamos a mostrar que el conjunto A es precisamente el conjunto de puntos de continuidad de f.

- Sea x un elemento de A. En vista de (8.14), es $f(x) = 0$. Sea ε un número positivo y sea n_0 un elemento de $\mathbb{N}_0$ tal que $1/(n_0 + 2) < \varepsilon$. El conjunto V_{n_0} es un abierto que contiene a x, así que es un entorno de ese punto, y

$$V_{n_0} = A \cup \bigcup_{i=n_0}^{\infty} (V_i \setminus V_{i+1}).$$

Sea y un elemento de U_{n_0}.

- Si y pertenece a A, entonces la igualdad (8.14) implica que $f(y) = 0$ y, por lo tanto, que es $d(f(y), f(x))0 < \varepsilon$.
- Si, por el contrario, hay un entero i tal que $i \geq n_0$ e $y \in V_i \setminus V_{i+1}$, entonces o bien $y \in E_i$ y $f(y) = 1/(i+2)$, de manera que $d(f(y), f(x)) = 1/(i+2) \leq 1/(n_0+2) < \varepsilon$, o bien $y \notin E_i$, de manera que $f(y) = 0$ y $d(f(y), f(x)) = 0 < \varepsilon$.

En cualquier caso, entonces, $d(f(y), f(x)) < \varepsilon$, y así que la función f es continua en x.

- Sea ahora x un elemento de $\mathbb{R} \setminus A = \bigcup_{n=0}^{\infty} D_n$, de manera que hay exactamente un entero no negativo n tal que $x \in D_n$. O bien x pertenece a $D_n^\circ \cap \mathbb{Q}$ o no, y analizamos las dos posibilidades por separado.
  - Supongamos que $x \in D_n^\circ \cap \mathbb{Q}$. En este caso $x \notin E_n$, así que $f(x) = 0$. Sea ε un número positivo. Como $x \in D_n^\circ$, hay un número positivo r tal que $(x - r, x + r) \subseteq D_n$. Si ponemos $s \coloneqq \min\{\varepsilon, r\}$, que es un número positivo, entonces $(x - s, x + s) \subseteq D_n$ y hay en este último intervalo un número irracional y. Por supuesto, y está en $B_\varepsilon(x)$ y, como $y \in D_n \setminus \mathbb{Q} \subseteq D_n \setminus (D_n^\circ \cap \mathbb{Q}) = E_n$, es $f(y) = 1/(n+2)$ y $d(f(y), f(x)) \geq 1/(n+2)$. Esto prueba que la función f no es continua en x.
  - Supongamos ahora que $x \notin D_n^\circ \cap \mathbb{Q}$, de manera que, de hecho, $x \in D_n \setminus (D_n^\circ \cap \mathbb{Q}) = E_n$ y $f(x) = 1/(n+2)$. Supongamos por un momento que la función f es continua en x. Existe entonces un número positivo δ tal que para todo elemento y de $B_\delta(x)$ es $d(f(y), f(x)) < 1/(n+2)(n+3)$. Ahora bien, como el interior de E_n es vacío, hay un elemento y de $B_\delta(x)$ que no pertenece a E_n y, por lo tanto, tal que $f(y) \neq 1/(n+2)$. En vista de cómo definimos a la función f, esto implica que $f(y) \geq 1/(n+3)$ o $f(y) \leq 1/(n+1)$. En el primer caso es

$$d(f(y), f(x)) \geq \frac{1}{n+2} - \frac{1}{n+3} = \frac{1}{(n+2)(n+3)}$$

    y en el segundo

$$d(f(y), f(x)) \geq \frac{1}{n+1} - \frac{1}{n+2} = \frac{1}{(n+1)(n+2)} > \frac{1}{(n+2)(n+3)}.$$

    Esto es absurdo.

Vemos así que en cualquier caso la función f no es continua en x.

La proposición queda así probada. □

La siguiente es una caracterización interesante de los conjuntos G_δ.

Proposición 8.9.8. *Sea X un espacio métrico. Un subconjunto A de X es un conjunto G_δ si y solamente si hay una métrica $d_A : A \times A \to \mathbb{R}$ sobre A que es topológicamente equivalente a la restricción a A de la métrica de X y con respecto a la cual A es completo.* □

Decimos por esto que los subconjuntos G_δ de X son precisamente los subconjuntos de X que son ***completamente metrizables***.

Capítulo 9

Compacidad

9.1. Conjuntos compactos

En la Sección 6.6 explicamos que un espacio métrico es de Lindelöf si todos sus cubrimientos abiertos poseen subcubrimientos contables. En esta sección estudiaremos una condición mucho más fuerte, que presentamos en la siguiente definición.

Definición 9.1.1. Un espacio métrico X es ***compacto*** si todo cubrimiento abierto de X posee un subcubrimiento finito. Por otro lado, un subconjunto de un espacio métrico es ***compacto*** si, visto como un espacio métrico, es compacto.

Antes que nada, veamos dos ejemplos.

Ejemplo 9.1.2. Sean a y b dos números reales tales que $a < b$. Mostremos que

el intervalo $[a, b]$ es compacto.

Sea $\mathscr{U}$ un cubrimiento abierto de $[a, b]$, y sea T el conjunto de todos los elementos t de $[a, b]$ tales que $\mathscr{U}$ contiene un subconjunto finito $\mathscr{V}$ tal que $[a, t] \subseteq \bigcup \mathscr{V}$.

Este conjunto T no es vacío. En efecto, como $\mathscr{U}$ es un cubrimiento de $[a, b]$, hay un elemento U de $\mathscr{U}$ tal que $a \in U$ y como U es abierto, hay un número positivo r tal que $[a, a + r) \subseteq U$. El conjunto $\mathscr{V} \coloneqq \{U\}$, que es finito, es tal que $[a, t] \subseteq \bigcup \mathscr{V}$ para cada $t \in [a, a + r)$, y esto nos dice que $[a, a + r) \subseteq T$. Como además T está claramente acotado superiormente, podemos considerar su supremo $\tau \coloneqq \sup T$. Notemos que $a < \tau \leq b$.

Supongamos que $\tau < b$. En ese caso $\tau \in (a, b)$ y, como $\mathscr{U}$ es un cubrimiento de $[a, b]$, hay un elemento V de $\mathscr{U}$ tal que $\tau \in V$ y un número positivo s tal que $(\tau - s, \tau + s) \subseteq V$. Como $\tau = \sup T$,

hay un elemento t de T tal que $\tau - s/2 < t$, y esto implica que hay un subconjunto finito $\mathscr{V}$ de $\mathscr{U}$ tal que $[a, \tau - s/2] \subseteq [a, t] \subseteq \bigcup \mathscr{V}$. Entonces $\mathscr{V}' \coloneqq \mathscr{V} \cup \{V\}$ es un subconjunto finito de $\mathscr{U}$ tal que $[a, \tau + s/2] \subseteq \bigcup \mathscr{V}'$ y, por lo tanto, $\tau + s/2 \in T$: esto es absurdo, ya que $\tau = \sup T$.

Vemos así que $b = \sup T$. Como $\mathscr{U}$ cubre a $[a, b]$ hay un elemento W de $\mathscr{U}$ tal que $b \in W$, y como W es abierto hay un número positivo u tal que $(b - u, b] \subseteq W$. Como $b = \sup T$, hay un elemento v de T tal que $b - u < v$ y, por lo tanto, un subconjunto finito $\mathscr{W}$ de $\mathscr{U}$ tal que $[a, v] \subseteq \bigcup \mathscr{W}$. El conjunto $\mathscr{W}' \coloneqq \mathscr{W} \cup \{W\}$ es un subconjunto finito de $\mathscr{U}$ y claramente $[a, b] \subseteq \mathscr{U}$. Hemos probado que todo cubrimiento abierto de $[a, b]$ posee un subcubrimiento finito, así que $[a, b]$ es un espacio compacto.

Ejercicio 9.1.3. Pruebe que un espacio métrico cuya topología es discreta es compacto si y solamente si es finito.

Ejercicio 9.1.4. Pruebe que un espacio métrico X es compacto si y solamente si todo conjunto $\mathscr{F}$ de cerrados de X tal que

$$\textit{todo subconjunto finito y no vacío de } \mathscr{F} \textit{ tiene intersección no vacía} \tag{9.1}$$

tiene intersección no vacía. Esto es la versión de la Proposición 6.6.23 que probamos sobre espacios de Lindelöf para espacios compactos. Decimos que un subconjunto $\mathscr{F}$ de $\mathscr{P}(X)$ tiene la ***propiedad de intersección finita*** si satisface la condición (9.1).

Es conveniente poder decidir si un subconjunto de un espacio métrico X es compacto trabajando exclusivamente con los abiertos de X.

Proposición 9.1.5. *Sea X un espacio métrico. Un subconjunto Y de X es compacto si y solamente si todo conjunto $\mathscr{U}$ de abiertos de X tal que $Y \subseteq \bigcup \mathscr{U}$ tiene un subconjunto finito $\mathscr{V}$ tal que $Y \subseteq \bigcup \mathscr{V}$.*

Demostración. Sea Y un subconjunto compacto de X, y sea $\mathscr{U}$ un conjunto de abiertos de X tal que $Y \subseteq \bigcup \mathscr{U}$. Si ponemos $\mathscr{U}_Y \coloneqq \{U \cap Y : U \in \mathscr{U}\}$, entonces $\mathscr{U}_Y$ es un conjunto de abiertos de Y tal que $Y = \bigcup \mathscr{U}_Y$, esto es, es un cubrimiento abierto del espacio Y y, como este es compacto, hay un subconjunto finito $\mathscr{W}$ de $\mathscr{U}_Y$ tal que $Y = \bigcup \mathscr{W}$. La función $\rho : U \in \mathscr{U} \mapsto U \cap Y \in \mathscr{U}_Y$ es sobreyectiva, así que hay un subconjunto finito $\mathscr{V}$ de $\mathscr{U}$ tal que $\rho(\mathscr{V}) = \mathscr{W}$, y esto claramente implica que $Y \cap \bigcup \mathscr{V} = \bigcup_{V \in \mathscr{V}} (Y \cap V) = \bigcup \mathscr{W} = Y$, de manera que $Y \subseteq \bigcup \mathscr{V}'$. Esto prueba que la condición de la proposición es necesaria.

Supongamos ahora que Y es un subconjunto de X que hace cierta la condición de la proposición, y sea $\mathscr{U}$ un cubrimiento abierto del espacio Y. Para cada elemento U de $\mathscr{U}$ podemos elegir un subconjunto abierto U' de X tal que $U = Y \cap U'$, y considerar el conjunto $\mathscr{U}' \coloneqq \{U' : U \in \mathscr{U}\}$. Es claro que $\mathscr{U}'$ es un conjunto de abiertos de X tal que $Y \cap \bigcup \mathscr{U}' = \bigcup_{U \in \mathscr{U}'} (Y \cap U) = \bigcup \mathscr{U} = Y$,

de manera que $Y \subseteq \bigcup \mathscr{U}'$. La hipótesis nos dice entonces que hay un subconjunto finito $\mathscr{V}'$ de $\mathscr{U}'$ tal que $Y \subseteq \bigcup \mathscr{V}'$. El conjunto $\mathscr{V} \coloneqq \{Y \cap V : V \in \mathscr{V}'\}$ es entonces un subconjunto finito de $\mathscr{U}$ y es $Y = \bigcup \mathscr{V}'$. Esto prueba que Y es un subconjunto compacto de X. □

Esta caracterización de los subconjuntos compactos tiene la siguiente consecuencia.

Proposición 9.1.6. *Un subconjunto cerrado de un espacio métrico compacto es compacto.*

Demostración. Sea X un espacio métrico compacto, sea Y un subconjunto cerrado de X, y sea $\mathscr{U}$ un conjunto de abiertos de X tal que $Y \subseteq \bigcup \mathscr{U}$. Como Y es cerrado en X, todos los elementos de $\mathscr{U}' \coloneqq \mathscr{U} \cup \{X \setminus Y\}$ son subconjuntos abiertos de X. Como además $\bigcup \mathscr{U}' = \bigcup \mathscr{U} \cup (X \setminus Y) \supseteq Y \cup (X \setminus Y) = X$, vemos que $\mathscr{U}$ es un cubrimiento abierto de X: como X es compacto, hay entonces un subconjunto finito $\mathscr{V}'$ de $\mathscr{U}'$ tal que $X = \bigcup \mathscr{V}$. Es fácil ver que el conjunto $\mathscr{V} \coloneqq \mathscr{U} \cap \mathscr{V}'$ es un subconjunto finito de $\mathscr{U}$ tal que $Y \subseteq \bigcup \mathscr{V}$. Esto prueba que Y es un subconjunto compacto de Y. □

Ejercicio 9.1.7. Muestre que la unión de una familia finita de subconjuntos compactos de un espacio métrico es compacta.

La compacidad es una condición que, para muchas cosas, aproxima los espacios con los que trabajamos a conjuntos finitos. La manifestación más inmediata de esto es la siguiente observación.

Proposición 9.1.8. *Un espacio compacto es acotado.*

Demostración. Sea X un espacio compacto. Si X es vacío, es claro que es acotado. Supongamos que, por el contrario, hay un elemento x en X. El conjunto $\mathscr{U} \coloneqq \{B_n(x) : n \in \mathbb{N}_0\}$ es un cubrimiento abierto de X, así que la hipótesis nos dice que $\mathscr{U}$ posee un subcubrimiento finito, esto es, que hay un subconjunto finito F de $\mathbb{N}_0$ tal que el subconjunto $\{B_n(x) : x \in F\}$ de $\mathscr{U}$ también es un cubrimiento de X. Como el conjunto F es finito, podemos considerar su máximo elemento $m \coloneqq \max F$. Para todo elemento n de F es $n \leq m$, así que $B_n(x) \subseteq B_m(x)$ y, por lo tanto, $X = \bigcup_{n \in F} B_n(x) \subseteq B_m(x)$. Esto nos dice que el espacio X es acotado. □

Que un espacio sea acotado significa que lo podemos cubrir con una bola, pero en general el radio de esa bola tiene que ser grande. Una condición más fuerte que podemos imponer sobre un espacio es que podamos cubrirlo con finitas bolas tan pequeñas como queramos. Así, por ejemplo, si X es un espacio métrico con infinitos puntos y métrica discreta, entonces podemos cubrir a X con una bola de radio 2, pero es imposible cubrirlo con finitas bolas de radio 1: usando la siguiente definición, decimos que X es acotado pero no totalmente acotado.

Definición 9.1.9. Un espacio métrico es ***totalmente acotado*** si para todo número positivo r hay un conjunto finito F de X tal que $X = \bigcup_{x\in F} B_r(X)$. Un subconjunto de un espacio métrico es ***totalmente acotado*** si, visto como un espacio métrico, es totalmente acotado.

Podemos decidir que un subconjunto de un espacio métrico es totalmente acotado sin considerarlo como un espacio: esto es lo que nos dice es siguiente ejercicio.

Ejercicio 9.1.10. Sea X un espacio métrico y sea Y un subconjunto de X. Las siguientes afirmaciones son equivalentes.

(*a*) El subconjunto Y es totalmente acotado.
(*b*) Para todo número positivo r hay un subconjunto finito F de Y tal que $Y \subseteq \bigcup_{x\in F} B_r(x)$.
(*c*) Para todo número positivo r hay un subconjunto finito F de X tal que $Y \subseteq \bigcup_{x\in F} B_r(x)$.

La clausura de un conjunto acotado es ella misma acotada. Para conjuntos totalmente acotados vale un resultado similar.

Proposición 9.1.11. *La clausura de un conjunto totalmente acotado de un espacio métrico es totalmente acotada.*

Demostración. Sea X un espacio métrico, sea Y un subconjunto totalmente acotado de X, y sea r un número positivo. Como Y es totalmente acotado en X, hay un subconjunto finito F de X tal que $Y \subseteq \bigcup_{x\in F} B_{r/2}(x)$. Tenemos entonces que

$$\overline{Y} \subseteq \overline{\bigcup_{x\in F} B_{r/2}(x)} = \bigcup_{x\in F} \overline{B_{r/2}(x)} \subseteq \bigcup_{x\in F} B_r(x)$$

y, por lo tanto, el conjunto Y es totalmente acotado. □

Un espacio métrico totalmente acotado es acotado, claro: la condición implica que lo podemos cubrir con un número finito de bolas de radio 1 y cada una de ellas es acotada, así que la conclusión sigue del resultado del Ejercicio 7.1.8. Por otro lado, de la misma forma que todo subespacio de un espacio acotado es acotado, tenemos la siguiente proposición.

Proposición 9.1.12. *Todo subespacio de un espacio métrico totalmente acotado es totalmente acotado.*

Demostración. Sea X un espacio métrico totalmente acotado, sea Y un subespacio de X, y sea r un número positivo. Como X es totalmente acotado, hay un subconjunto finito H de X tal que $X = \bigcup_{x\in H} B^X_{r/2}(x)$. Sea $G \coloneqq \{x \in H : B^X_{r/2}(x) \cap Y \neq \varnothing\}$, de manera que hay una función $c : G \to Y$ tal que $c(x) \in B^X_{r/2}(x) \cap Y$. Pongamos $F \coloneqq c(G)$. Por otro lado, para cada elemento x de G es

$c(x) \in B^X_{r/2}(x)$, así que $B^X_{r/2}(x) \subseteq B^X_r(c(x))$: se sigue de esto que

$$\begin{aligned} Y = Y \cap X = Y \cap \bigcup_{x\in H} B^X_{r/2}(x) = \bigcup_{x\in H}(Y \cap B^X_{r/2}(x)) = \bigcup_{x\in G}(Y \cap B^X_{r/2}(x)) \\ \subseteq \bigcup_{x\in G}(Y \cap B^X_r(c(x))) = \bigcup_{y\in F}(Y \cap B^X_r(y)) = \bigcup_{y\in F} B^Y_r(y). \end{aligned}$$

Esto nos dice que F es un subconjunto finito de Y tal que $Y = \bigcup_{y\in F} B^X_r(y)$ y prueba que Y es totalmente acotado. □

El siguiente ejercicio se ocupa de qué pasa con los conjuntos totalmente acotados cuando cambiamos una métrica por otra que le es equivalente.

Ejercicio 9.1.13.

(a) Sea X un conjunto y sean d y d' dos métricas sobre X que son uniformemente equivalentes. Pruebe que el espacio métrico (X, d) es totalmente acotado si y solamente si el espacio métrico (X, d') es totalmente acotado.

(b) Muestre que si en la parte (a) reemplazamos la hipótesis de que las métricas d y d' son uniformemente equivalentes por la de que son topológicamente equivalentes la conclusión no sigue valiendo.

Proposición 9.1.14. *Un espacio métrico compacto es totalmente acotado.*

Demostración. Sea X un espacio métrico compacto y sea r un número positivo. El conjunto $\mathscr{U} \coloneqq \{B_r(x) : x \in X\}$ es claramente un cubrimiento abierto de X, así que la hipótesis implica que $\mathscr{U}$ posee un subcubrimiento finito, esto es, que hay un subconjunto finito F de X tal que $\mathscr{V} = \{B_r(x) : x \in F\}$ es un cubrimiento de X y, por lo tanto, $X = \bigcup_{x\in F} B_r(x)$. Vemos así que X es totalmente acotado. □

Se sigue de esta proposición y de la siguiente que los espacios compactos son separables.

Proposición 9.1.15. *Un espacio métrico totalmente acotado es separable*

Demostración. Sea X un espacio métrico totalmente acotado. Para cada entero positivo n hay un subconjunto finito F_n de X tal que $X = \bigcup_{x\in F_n} B_{1/n}(x)$. Es claro que el conjunto $F \coloneqq \bigcup_{n\in\mathbb{N}} F_n$ es contable. Sea, por otro lado, U un subconjunto abierto no vacío de X, de manera que hay un punto y en U. Como U es abierto, hay un número entero positivo n tal que $B_{1/n}(y)$ está contenido en U. Por otro lado, como $y \in \bigcup_{x\in F_n} B_{1/n}(x)$, hay un elemento x de F_n tal que $y \in B_{1/n}(x)$.

Tenemos entonces que $x \in B_{1/n}(y) \subseteq U$ y, por lo tanto, que $U \cap F \neq \varnothing$. Vemos así que F interseca a todo abierto no vacío de X y, por lo tanto, que es denso. □

Otra propiedad importante de los espacios compactos es que toda sucesión en ellos tiene puntos límites.

Proposición 9.1.16. *Toda sucesión en un espacio métrico compacto tiene una subsucesión convergente.*

Probaremos más abajo que, de hecho, esta propiedad caracteriza a los espacios compactos.

Demostración. Sea X un espacio métrico compacto, y sea $x = (x_n)_{n\geq 0}$ una sucesión en X. Supongamos que x no posee ninguna subsucesión convergente. De acuerdo a la Proposición 7.2.12, esto implica cada elemento y de X posee un entorno abierto U_y tal que el conjunto $I_y \coloneqq \{n \in \mathbb{N}_0 : x_n \in U_y\}$ es finito. El conjunto $\mathscr{U} \coloneqq \{U_y : y \in X\}$ es claramente un cubrimiento abierto de X, así que la hipótesis de que X es compacto implica que $\mathscr{U}$ posee un subcubrimiento finito. En otras palabras, hay un subconjunto finito F de X tal que $X = \bigcup_{y\in F} U_y$. Para cada $n \in \mathbb{N}_0$ tenemos que $x_n \in X = \bigcup_{y\in F} U_y$, así que existe $y \in F$ tal que $x_n \in U_y$ y, por lo tanto, $n \in U_y$. Vemos de esta forma que $\mathbb{N}_0 \subseteq \bigcup_{y\in F} I_y$ y esto es absurdo, ya que esta es una unión finita de conjuntos finitos. Esta contradicción nos permite concluir que la sucesión x necesariamente tiene una subsucesión convergente. □

Obtenemos de la proposición una nueva condición necesaria para la compacidad de un espacio.

Corolario 9.1.17. *Un espacio métrico compacto es completo.*

Demostración. Sea X un espacio métrico compacto y sea x una sucesión en X que satisface la condición de Cauchy. De acuerdo a la Proposición 9.1.16, x posee una subsucesión convergente y entonces ella misma converge por la Proposición 7.3.7. □

Corolario 9.1.18. *Todo subconjunto infinito de un espacio métrico compacto posee un punto de acumulación.*

Demostración. Sea X un espacio métrico y sea S un subconjunto infinito de X, de manera que hay una función inyectiva $h : \mathbb{N}_0 \to S$. La Proposición 9.1.16 nos dice que la sucesión $(h(n))_{n\geq 0}$ posee una sucesión convergente, de manera que hay una función estrictamente creciente $f : \mathbb{N}_0 \to \mathbb{N}_0$ tal que la subsucesión $(h(f(n)))_{n\geq 0}$ converge a un punto ξ de X. Si N es un entorno de ξ, entonces esa subsucesión está a la larga en N y, por lo tanto, hay un entero positivo n_0 tal que $h(f(n)) \in N \cap Y$ para todo entero n tal que $n \geq n_0$. Como la función $h \circ f$ es inyectiva, el conjunto $\{h(f(n)) : n \in \mathbb{N}_0, n \geq 0\}$ es infinito y, como está contenido en $N \cap Y$, este último

también. El punto ξ es, en consecuencia, un punto de acumulación de S. □

Ejercicio 9.1.19. Pruebe que una sucesión en un espacio métrico compacto que tiene un único punto límite converge.

9.2. Caracterizaciones de los conjuntos compactos

En este punto tenemos varias condiciones necesarias para la compacidad de un espacio. Es importante que algunas combinaciones de ellas resultan ser condiciones suficientes: la más importante de todas es, sin duda, la dada en nuestro siguiente resultado.

Proposición 9.2.1. *Un espacio métrico es compacto si y solamente si es completo y totalmente acotado.*

Demostración. Que la condición es necesaria es consecuencia de la Proposición 9.1.14 y el Corolario 9.1.17. Mostremos que también es suficiente. Sea X un espacio métrico completo y totalmente acotado, sea $\mathscr{U}$ un cubrimiento abierto de X y, para llegar a una contradicción, supongamos que $\mathscr{U}$ no posee ningún subcubrimiento finito.

Sea Q el conjunto de todos los elementos (r,x) de $(0,+\infty)\times X$ tales que no hay un subconjunto finito $\mathscr{V}$ de $\mathscr{U}$ tal que $B_r(x)\subseteq\bigcup\mathscr{V}$, y para cada elemento (r,x) de Q escribamos

$$R(r,x) \coloneqq \{(s,y)\in Q : s = r/2, y\in B_r(x)\}.$$

El conjunto Q no es vacío: como X es totalmente acotado, es acotado y hay, por lo tanto, un punto x y un número positivo r tal que $B_r(x) = X$ y, como $\mathscr{U}$ no posee ningún subcubrimiento finito de X, esto nos dice que $(r,x)\in Q$.

Supongamos que (r,x) es un elemento de Q. Como X es totalmente acotado, la bola $B_r(x)$ es totalmente acotada y hay un subconjunto finito F de $B_r(x)$ tal que $B_r(x)\subseteq\bigcup_{y\in F} B_{r/2}(y)$. Si para cada elemento y de F hubiera un subconjunto finito $\mathscr{V}_y$ de $\mathscr{U}$ tal que $B_{r/2}(y)\subseteq\bigcup\mathscr{V}_y$, entonces el conjunto $\mathscr{V}\coloneqq\bigcup_{y\in F}\mathscr{V}_y$ sería un subconjunto finito de $\mathscr{U}$ tal que

$$B_r(x)\subseteq\bigcup_{y\in F} B_{r/2}(y)\subseteq\bigcup_{y\in F}\bigcup_{V\in\mathscr{V}_y} V = \bigcup_{V\in\mathscr{V}} V = \bigcup\mathscr{V},$$

lo que es imposible porque (r,x) pertenece a Q. Vemos así que hay un elemento y de F tal que no hay un subconjunto finito $\mathscr{V}$ de $\mathscr{U}$ tal que $B_{r/2}(y)\subseteq\bigcup\mathscr{V}$ y, por lo tanto, $(r/2,y)\in Q$. Más aún, como $y\in B_r(x)$, tenemos que $(r/2,y)\in R(r,x)$.

Hemos probado que $Q\neq\varnothing$ y que $R(r,x)\neq\varnothing$ para todo elemento (r,x) de Q. De acuerdo al *Principio de elección dependiente*, entonces, hay una función $f:\mathbb{N}_0\to Q$ tal que $f(n+1)\in R(f(n))$

para todo $n \in \mathbb{N}_0$. Si para todo $n \in \mathbb{N}_0$ escribimos $f(n) = (r_n, x_n)$, con $r_n \in (0, \infty)$ y $x_n \in X$, entonces esto significa que $r_{n+1} = r_n/2$ y $x_{n+1} \in B_{r_n}(x)$ cualquiera sea n en $\mathbb{N}_0$. Es fácil ver esto implica que $r_n = r_0/2^n$ y $d(x_{n+1}, x_n) < r_0/2^n$ para todo $n \in \mathbb{N}_0$: como la serie $\sum_{n=0}^{\infty} r_0/2^n$ tiene sus términos positivos y converge, la Proposición 7.3.6 nos permite concluir que la sucesión $(x_n)_{n\geq 0}$ es de Cauchy. Como el espacio X es completo, esa sucesión converge a un punto ξ de X.

Como $\mathscr{U}$ es un cubrimiento de X, hay un elemento U de $\mathscr{U}$ tal que $\xi \in U$, y como U es abierto, hay un número positivo t tal que $B_t(\xi) \subseteq U$. Como la sucesión $(x_n)_{n\geq 0}$ converge a ξ, está a la larga en $B_{t/2}(\xi)$ y hay, por lo tanto, un entero positivo k tal que $r_k = r_0/2^k < t/2$ y $x_k \in B_{t/2}(\xi)$. Como (r_k, x_k) es un elemento de Q, no hay un subconjunto finito de $\mathscr{U}$ que cubra a $B_{r_k}(x_k)$. Esto es absurdo, ya que $B_{r_k}(x_k) \subseteq B_t(\xi) \subseteq U$: en efecto, para cada elemento z de $B_{r_k}(x_k)$ es $d(z, \xi) \leq d(z, x_k) + d(x_k, \xi) < r_k + t/2 < t$

Esta contradicción provino de haber supuesto que el cubrimiento $\mathscr{U}$ de X no posee subcubrimientos finitos. Vemos así que el espacio X es compacto, como afirma la proposición. □

Ejercicio 9.2.2. Pruebe que toda sucesión en un espacio totalmente acotado posee una subsucesión que satisface la condición de Cauchy y use eso para dar una nueva demostración de que la condición que da la Proposición 9.2.1 para que un espacio sea compacto es suficiente.

Nos proponemos describir los subconjuntos compactos de $\mathbb{R}^n$ usando la Proposición 9.2.1: para ello necesitamos entender cuáles son los subconjuntos totalmente acotados de ese espacio.

Proposición 9.2.3. *Sea V un espacio normado real de dimensión finita. Un subconjunto de V es totalmente acotado si y solamente si es acotado.*

Este resultado es una generalización del *Principio de Arquímedes* que nos dice que para todo número real positivo x hay un entero positivo n tal que $x \leq n$.

Demostración. Ya sabemos que la condición es necesaria, así que tenemos que ver que también es suficiente. Sea $(e_1, e_2, \ldots, e_n)$ una base ordenada de V, sea

$$K \coloneqq \|e_1\| + \|e_2\| + \cdots + \|e_n\|,$$

y notemos que para toda elección de $a_1, a_2, \ldots, a_n$ en $\mathbb{R}$ y todo $i \in [\![n]\!]$ es

$$|a_i| \cdot \|e_i\| \leq \|a_1 e_1 + a_2 e_2 + \cdots + a_n e_n\|$$

de manera que

$$|a_i| \leq \frac{\|a_1 e_1 + a_2 e_2 + \cdots + a_n e_n\|}{\|e_i\|} \leq \frac{\|a_1 e_1 + a_2 e_2 + \cdots + a_n e_n\|}{K}.$$

Sea X un subconjunto acotado de V, de manera que hay un número positivo R tal que $X \subseteq B_R(0)$, sea r un número real positivo, sea $s \coloneqq r/2K$, y consideremos el conjunto

$$F \coloneqq \{sa_1e_1 + sa_2e_2 + \cdots + sa_ne_n : a_1, a_2, \ldots, a_n \in \mathbb{Z}, |a_1|, |a_2|, \ldots, |a_n| \leq R/Ks + 1\}.$$

Sea x un elemento de X y sean $x_1, x_2, \ldots, x_n$ los números reales tales que $x = x_1e_1 + x_2e_2 + \cdots + x_ne_n$. Para cada $i \in [\![n]\!]$ pongamos $a_i \coloneqq \lfloor x_i/s \rfloor$, de manera que $0 \leq x_i - sa_i < s$. Si $i \in [\![n]\!]$, entonces $|x_i| \leq \|x\|/K \leq R/K$ y, por lo tanto,

$$-R/Ks - 1 \leq -|x_i|/s - 1 \leq x_i/s - 1 < a_i \leq x_i/s \leq |x_i|/s \leq R/Ks + 1.$$

Esto implica que el punto $y \coloneqq sa_1e_1 + sa_2e_2 + \cdots + sa_ne_n$ pertenece a F. Por otro lado,

$$\begin{aligned}\|x - y\| &= \big\|(x_1e_1 + x_2e_2 + \cdots + x_ne_n) - (sa_1e_1 + sa_2e_2 + \cdots + sa_ne_n)\big\| \\ &\leq |x_1 - sa_1| \cdot \|e_1\| + |x_2 - sa_2| \cdot \|e_2\| + \cdots + |x_n - sa_n| \cdot \|e_n\| \\ &\leq s \cdot (\|e_1\| + \|e_2\| + \cdots + \|e_n\|) = s \cdot K = \frac{r}{2} < r,\end{aligned}$$

de manera que $x \in B_r(y)$. Hemos probado que

$$X \subseteq \bigcup_{y \in F} B_r(y),$$

y, en definitiva, que el subconjunto X de V es totalmente acotado. □

Esta proposición se ocupa solamente de espacios normados reales, pero el mismo resultado vale para los espacios normados complejos.

Ejercicio 9.2.4. Pruebe que un subconjunto acotado de un espacio normado complejo de dimensión finita es totalmente acotado.

Juntando todo, obtenemos el *Teorema de Heine–Borel*, por Eduard Heine y Émile Borel.

Proposición 9.2.5. *Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, y sea V un espacio normado sobre $\Bbbk$ de dimensión finita. Un subconjunto de V es compacto si y solamente si es cerrado y acotado.*

Demostración. Si X es un subconjunto compacto de V, entonces el Corolario 9.1.17 nos dice que X es un subespacio completo de V, así que la Proposición 7.3.26 nos dice que X es cerrado en V y la Proposición 9.1.8 nos dice que X es acotado. Vemos así que la condición de la proposición es necesaria. Por otro lado, si X es un subconjunto cerrado y acotado de $\mathbb{R}^n$, entonces la Proposición 7.3.26 nos dice que X es cerrado, ya que V es completo, y la Proposición 9.2.3 nos dice que X es totalmente acotado: podemos concluir gracias a la Proposición 9.2.1 que X es compacto. □

Una consecuencia inmediata de este resultado es el siguiente *Teorema de Bolzano–Weierstrass*, por Bernard Bolzano y Karl Weierstrass.

Proposición 9.2.6. *Sea n un entero no negativo. Toda sucesión acotada en $\mathbb{R}^n$ posee una subsucesión convergente.*

Demostración. Sea x una sucesión acotada en $\mathbb{R}^n$, de manera que hay un número positivo tal que x toma todos sus valores en la bola cerrada $\overline{B}_R(0)$. El *Teorema de Heine–Borel* nos dice que esa bola es cerrada, así que la Proposición 9.1.16 nos permite concluir que la sucesión x posee una subsucesión convergente. □

Una forma equivalente de enunciar este resultado es la siguiente:

Proposición 9.2.7. *Un subconjunto infinito y acotado de $\mathbb{R}^n$ posee un punto de acumulación.*

Demostración. Sea S un subconjunto infinito y acotado de $\mathbb{R}^n$. Como es infinito, hay una función inyectiva $f : \mathbb{N}_0 \to S$, y como es acotado, la sucesión $(f(n))_{n\geq 0}$ es acotada. La Proposición 9.2.6 nos dice entonces que esta sucesión posee una subsucesión convergente y es claro que su límite es un punto de acumulación del conjunto S. □

Veamos cómo podemos usar esta proposición para obtener un resultado útil.

Ejemplo 9.2.8. Sea α un número real irracional. Queremos volver a probar la afirmación

$$\textit{el conjunto } D \coloneqq \{m + n\alpha : m, n \in \mathbb{Z}\} \textit{ es denso en } \mathbb{R}$$

del Ejemplo 6.6.9. El conjunto $D \cap [0,1]$ es infinito. Si no lo fuera, la función

$$n \in \mathbb{Z} \mapsto -\lfloor n\alpha \rfloor + n\alpha \in D \cap [0,1]$$

no sería inyectiva y habría dos enteros distintos m y n tales que $m\alpha - \lfloor m\alpha \rfloor = n\alpha - \lfloor n\alpha \rfloor$, de manera que

$$\alpha = \frac{\lfloor m\alpha \rfloor - \lfloor n\alpha \rfloor}{m - n},$$

lo que es, por supuesto, imposible. La Proposición 9.2.7 nos dice entonces que el conjunto $D \cap [0,1]$ posee un punto de acumulación ξ. Sea ahora x un número real y sea ε un número positivo. Como ξ es un punto de acumulación de $D \cap [0,1]$, hay dos elementos u y v de D tales que $u < v$ y $|u - \xi| < \varepsilon/2$ y $|v - \xi| < \varepsilon/2$, de manera que $0 < v - u < \varepsilon$. El número

$$y \coloneqq \left\lfloor \frac{x}{|v-u|} \right\rfloor \cdot (v - u)$$

pertenece a D y es

$$|y-x| = \left| \left\lfloor \frac{x}{|v-u|} \right\rfloor \cdot (v-u) - x \right| = \left| \left\lfloor \frac{x}{|v-u|} \right\rfloor - \frac{x}{v-u} \right| \cdot |v-u| < \varepsilon.$$

Esto nos implica que el conjunto D es denso en $\mathbb{R}$, como queremos.

Ejercicio 9.2.9. Sea $S \coloneqq \{z \in \mathbb{C} : |z| = 1\}$ la circunferencia unidad en $\mathbb{C}$ y sea α un número real irracional. Muestre que el subconjunto $D \coloneqq \left\{(e^{it}, e^{\alpha i t}) : t \in \mathbb{R}\right\}$ de $S \times S$ es denso en $S \times S$.

La propiedad de los espacios $\mathbb{R}^n$ que describe la Proposición 9.2.5 es muy especial. Le ponemos un nombre:

Definición 9.2.10. Decimos que un espacio métrico X tiene la ***propiedad de Heine–Borel*** si los subconjuntos cerrados y acotados de X son compactos.

Vemos así que todo espacio normado de dimensión finita tiene la propiedad de Heine–Borel y resulta, de hecho, que son los únicos espacios normados que la tienen.

Proposición 9.2.11. *Sea V un espacio normado. Si V tiene dimensión infinita, entonces la bola cerrada $\overline{B}_1(0)$ no es compacta y, por lo tanto, V no tiene la propiedad de Heine–Borel.*

Dejamos a cargo del lector verificar que, de hecho, en esa situación cualquiera sea el número positivo r la bola cerrada $\overline{B}_r(0)$ no es compacta.

Demostración. Supongamos que la bola $\overline{B}_1(0)$ es compacta, de manera que es totalmente acotada. Hay entonces un subconjunto finito F de B tal $B_1(0) \subseteq \overline{B}_1(0) \subseteq \bigcup_{x\in F} B_{1/2}(x)$ Para cada $x \in F$ es $B_{1/2}(x) = x + B_{1/2}(0)$, así que si escribimos W al subespacio de V generado por F, entonces

$$B_1(0) \subseteq \bigcup_{x\in F} (x + B_{1/2}(0)) \subseteq W + B_{1/2}(0)$$

y, por lo tanto, $B_2(0) \subseteq W + B_1(0)$. Un argumento inductivo evidente deduce de esto que, de hecho, para todo entero positivo n es $B_{2^n}(0) \subseteq W + B_1(0)$, y, por lo tanto, que

$$V = \bigcup_{n\in\mathbb{N}} B_{2^n}(0) \subseteq W + B_1(0).$$

Sea v un elemento de V. Para todo $0 \in \mathbb{R}$ hay elementos $w_n \in W$ y $b_n \in B_1(0)$ tales que $nv = w_n + b_n$ y, por lo tanto, $v - w_n/n = b_n/n$. Como $\|b_n/n\| < 1/n$ para todo $n \in \mathbb{N}_0$, es claro que la sucesión $(b_n/n)_{n\geq 0}$ converge a 0, así que la sucesión $(w_n)_{n\geq 0}$ converge a v. Ahora bien, como W es un espacio normado de dimensión finita, es completo y, por lo tanto, es cerrado en V, así que el límite v de la sucesión $(w_n)_{n\geq 0}$ pertenece a W. Vemos así que $V = W$ y, por lo tanto, que V tiene dimensión finita. Esto prueba la proposición. □

Daremos muchos ejemplos de espacios compactos más abajo, pero antes de eso necesitamos algunos criterios para poder reconocerlos.

Proposición 9.2.12. *Un espacio métrico X es compacto si y solamente si toda sucesión en X posee una sucesión convergente.*

Demostración. Sea X un espacio métrico, supongamos que X es compacto, y sea $x = (x_n)_{n\geq 0}$ una sucesión en X. Supongamos que x no posee ninguna subsucesión convergente. De acuerdo a la Proposición 7.2.12, esto implica cada elemento y de X posee un entorno abierto U_y tal que el conjunto $I_y \coloneqq \{n \in \mathbb{N}_0 : x_n \in U_y\}$ es finito. El conjunto $\mathscr{U} \coloneqq \{U_y : y \in X\}$ es claramente un cubrimiento abierto de X, así que la hipótesis de que X es compacto implica que $\mathscr{U}$ posee un subcubrimiento finito. En otras palabras, hay un subconjunto finito F de X tal que $X = \bigcup_{y\in F} U_y$. Para cada $n \in \mathbb{N}_0$ tenemos que $x_n \in X = \bigcup_{y\in F} U_y$, así que existe $y \in F$ tal que $x_n \in U_y$ y, por lo tanto, $n \in U_y$. Vemos de esta forma que $\mathbb{N}_0 \subseteq \bigcup_{y\in F} I_y$ y esto es absurdo, ya que esta es una unión finita de conjuntos finitos. Esta contradicción nos permite concluir que la sucesión x necesariamente tiene una subsucesión convergente y, por lo tanto, que la condición de la proposición es necesaria.

Sea ahora X es un espacio métrico que satisface la condición de la proposición. Una sucesión en X que satisface la condición de Cauchy posee entonces una subsucesión convergente y la Proposición 7.3.7 implica que, de hecho, converge: esto nos dice que X es completo.

Supongamos, por otro lado, que X no es totalmente acotado, de manera que hay un número positivo r tal que no hay ningún subconjunto finito F de X tal que $X = \bigcup_{x\in F} B_r(x)$. Sea $\mathscr{S}$ el conjunto de todas las funciones $\phi : [\![0, n]\!] \to X$ con $n \in \mathbb{N}_0$ tales que cada vez que i y j son elementos distintos de $[\![0, n]\!]$ es $d(\phi(i), \phi(j)) \geq r$.

Sea $\phi : [\![0, n]\!] \to X$ un elemento de $\mathscr{S}$. De acuerdo a la hipótesis hay un elemento ξ en $X \setminus \bigcup_{i=0}^{n} B_r(\phi(i))$, y podemos entonces considerar la función $\psi : [\![0, n+1]\!] \to X$ que en cada elemento i de $[\![0, n+1]\!]$ toma el valor

$$\psi(i) = \begin{cases} \phi(i) & \text{si } i < n+1; \\ \xi & \text{si } i = n+1. \end{cases}$$

Esta función pertenece al conjunto $\mathscr{S}$. En efecto, si i y j son dos elementos de $[\![0, n+1]\!]$ tales que $i < j$, entonces o bien $j < n+1$ y $d(\psi(i), \psi(j)) = d(\phi(i), \phi(j)) \geq r$, o bien $i < j = n_1$ y $d(\psi(i), \psi(j)) = d(\phi(i), \xi) \geq r$ porque $\xi \notin \bigcup_{i=0}^{n} B_r(\phi(i))$. Por otro lado, es claro que $\psi|_{[\![0,n]\!]} = \phi$. Finalmente, como X no es totalmente acotado no es vacío y hay, por lo tanto, una función $\phi : [\![0, 0]\!] \to X$ y está claramente en $\mathscr{S}$.

Todo esto nos dice que estamos en la situación de la Proposición 2.1.9 y esta nos permite concluir que hay una función $h : \mathbb{N}_0 \to X$ tal que $h|_{[\![0,n]\!]} \in \mathscr{S}$ para todo $n \in \mathbb{N}_0$. La sucesión $(h(n))_{n\geq 0}$ tiene, por hipótesis, una subsucesión convergente, esto es, hay una función estrictamente creciente $f : \mathbb{N}_0 \to \mathbb{N}_0$ tal que la sucesión $(h(f(n))_{n\geq 0}$ converge y, en particular, satisface la condición de Cauchy. Esto implica que hay un entero n_0 tal que siem-

pre que $m, n \in \mathbb{N}_0$ vale $m, n \geq 0 \implies d(h(f(m)), h(f(n))) < r$. En particular, tenemos que $d(h(f(n_0)), h(f(n_0+1))) < r$. Por otro lado, como $h|_{[\![0, f(n_0+1)]\!]} \in \mathscr{S}$, tenemos que que $d(h(f(n_0)), h(f(n_0+1))) \geq r$. Esto es absurdo, y la contradicción provino de haber supuesto que el espacio X no es totalmente acotado.

Vemos así que X es completo y totalmente acotado, así que es compacto. □

9.3. El teorema de Tikhonov

Queremos probar un caso particular de un célebre teorema de Andrey Nikolayevich Tikhonov [Tyc1935] — el teorema afirma que un producto cartesiano arbitrario de espacios topológicos compactos es compactos, y nosotros solamente consideramos aquí productos cartesianos numerables de espacios métricos.

Proposición 9.3.1. *Sea $((X_n, d_n))_{n\geq 0}$ una sucesión de espacios métricos, y consideremos al conjunto $X \coloneqq \prod_{n=0}^{\infty} X_n$ dotado de la métrica $d : X \times X \to \mathbb{R}$ descripta en la Proposición 5.2.6.4, de manera que para cada elección de $x = (x_n)_{n\geq 0}$ y de $y = (y_n)_{n\geq 0}$ en X es*

$$d(x, y) = \sum_{n=0}^{\infty} \frac{1}{2^n} \frac{d_n(x_n, y_n)}{1 + d_n(x_n, y_n)}.$$

Si para cada $n \in \mathbb{N}_0$ el espacio X_n es compacto, entonces X es compacto.

Demostración. Sea $x = (x_m)_{m\geq 0}$ una sucesión de puntos en X, de manera que para cada $m \in \mathbb{N}_0$ es $x_m = (x_{m,n})_{n\geq 0}$. Para probar la proposición es suficiente que mostremos que x posee una subsucesión convergente.

Primer paso. Escribamos $\mathscr{F}$ al conjunto de todas las funciones $\mathbb{N}_0 \to \mathbb{N}_0$ que son estrictamente crecientes. Cuando f y g son dos elementos de $\mathscr{F}$, escribamos $f \preceq g$ si hay un elemento h de $\mathscr{F}$ tal que $f = g \circ h$ — notemos que en ese caso, como la función g es inyectiva, hay exactamente una función h que satisface esa condición. Digamos, por otro lado, que una función $\phi : [\![0, k]\!] \to \mathscr{F}$ definida en un conjunto de la forma $[\![0, k]\!]$ con $k \in \mathbb{N}_0$ es ***buena*** si

$$\text{para cada } i \in [\![0, k-1]\!] \text{ es } \phi(i+1) \preceq \phi(i) \tag{9.2}$$

y

$$\text{para cada } i \in [\![0, k]\!] \text{ y cada } j \in [\![0, i]\!] \text{ la sucesión } (x_{\phi(i)(m),j})_{m\geq 0} \text{ converge,} \tag{9.3}$$

y hagamos las siguientes dos observaciones.

- La sucesión $(x_{m,0})_{m\geq 0}$ toma valores en el espacio X_0 y este es compacto, así que hay una función h en $\mathscr{F}$ tal que la sucesión $(x_{h(m),0})_{m\geq 0}$ converge. Se sigue de esto que la función $\phi : [\![0,0]\!] \to \mathscr{F}$ que tiene $\phi(0) = \mathrm{id}_{\mathbb{N}_0}$ es buena. En efecto, aquí k es 1, así que la condición (9.2) se cumple trivialmente, y la condición (9.3) se cumple por la forma en que elegimos la función h.
- Supongamos que k es un elemento de $\mathbb{N}_0$ y que $\phi : [\![0,k]\!] \to \mathscr{F}$ es una función buena. La sucesión $(x_{\phi(k)(m),k+1})_{m\geq 0}$ toma valores en X_{k+1}, que es compacto, así que hay un elemento h de $\mathscr{F}$ tal que la sucesión $(x_{\phi(k)(h(m)),k+1})_{m\geq 0}$ converge. Es claro que hay exactamente una función $\psi : [\![0,k+1]\!] \to \mathscr{F}$ la función tal que $\psi|_{[\![0,k]\!]} = \phi$ y $\psi(k+1) = \phi(k)\circ h$. Veamos que es buena.
  - Si i es un elemento de $[\![0,k]\!]$, entonces o bien $i < k$ y es $\psi(i+1) = \phi(i+1) \preceq \phi(i) = \psi(i)$ porque ϕ es buena, o bien $i = k$ y $\psi(k+1) = \phi(k)\circ h \preceq \phi(k) = \psi(k)$ por la forma en que definimos la relación $\preceq$. Esto nos dice que ψ satisface la condición (9.2).
  - Sea i un elemento de $[\![0,k+1]\!]$.
    * Si $i \leq k$, entonces $\psi(i) = \phi(i)$ y, por lo tanto, para cada $j \in [\![0,i]\!]$ la sucesión $(x_{\psi(i)(m),j})_{m\geq 0}$ converge porque coincide con $(x_{\phi(i)(m),j})_{m\geq 0}$ y la función ϕ es buena.
    * Si, en cambio, $i = k+1$, entonces para cada $j \in [\![0,i]\!]$ o bien $j < i$, de manera que $j \in [\![0,k]\!]$ y $(x_{\psi(k+1)(m),j})_{m\geq 0}$ es una subsucesión de $(x_{\phi(k)(m),j})_{m\geq 0}$, que converge, así que ella misma converge, o bien $j = i$, de manera que la sucesión $(x_{\psi(k+1)(m),k+1})_{m\geq 0}$ converge por la forma en que elegimos a la función h.

    Esto nos dice que la función ψ satisface la condición (9.3).

En vista de esto y de acuerdo a la Proposición 2.1.9 podemos concluir que hay una función $\phi : \mathbb{N}_0 \to \mathscr{F}$ tal que para todo $k \in \mathbb{N}_0$ la restricción $\phi|_{[\![0,k]\!]}$ es buena. Esto implica que

- para cada $k \in \mathbb{N}_0$ es $\phi(k+1) \preceq \phi(k)$, de manera que hay exactamente una función h_k en $\mathscr{F}$ tal que $\phi(k+1) = \phi(k)\circ h_k$, y que
- para cada $k \in \mathbb{N}_0$ y cada $i \in [\![0,k]\!]$ la sucesión $(x_{\phi(k)(m),i})_{m\geq 0}$ converge.

Segundo paso. Consideremos ahora la función $\delta : k \in \mathbb{N}_0 \mapsto \phi(k)(k) \in \mathbb{N}_0$. Se trata de un elemento de $\mathscr{F}$: en efecto, si $k \in \mathbb{N}_0$, entonces

$$\delta(k+1) = \phi(k+1)(k+1) = \phi(k)(h_k(k+1)) > \phi(k)(h_k(k)) \geq \phi(k)(k) = \delta(k),$$

ya que las funciones h_k y $\phi(k)$ son estrictamente crecientes. Afirmamos que, además,

> *para todo $k \in \mathbb{N}_0$ es $\delta \preceq \phi(k)$, de manera que la sucesión $(x_{\delta(m),k})_{m\geq 0}$ converge porque la sucesión $(x_{\phi(k)(m),k})_{m\geq 0}$ lo hace.*

□

9.4. Funciones continuas sobre espacios compactos

Las funciones continuas sobre espacios compactos tienen algunas propiedades especiales que hacen que estos sean importantes en las aplicaciones. En esta sección presentaremos algunas.

La primera, extremadamente útil, es que una función continua sobre un espacio compacto es automáticamente uniformemente continua.

Proposición 9.4.1. *Sean X e Y dos espacios métricos y supongamos que X es compacto. Una función continua $f : X \to Y$ es uniformemente continua.*

Demostración. Sea $f : X \to Y$ una función continua que no es uniformemente continua, de manera que hay un número positivo ε tal que para todo entero no negativo n hay un par de puntos x_n e y_n en X tales que $d(x_n, y_n) < 1/(n+1)$ y $d(f(x_n), f(y_n)) \geq \varepsilon$. Como el espacio X es compacto, la sucesión $(x_n)_{n\geq 0}$ posee una subsucesión convergente, así que hay una función estrictamente creciente $h : \mathbb{N}_0 \to \mathbb{N}_0$ tal que la sucesión $(x_{h(n)})_{n\geq 0}$ converge a un punto ξ de X. Como

$$d(x_{h(n)}, y_{h(n)}) < \frac{1}{h(n)+1} < \frac{1}{n+1}$$

para todo $n \in \mathbb{N}_0$, es claro que la sucesión $(y_{h(n)})_{n\geq 0}$ también converge a ξ. Ahora bien, como la función f es continua en ξ, hay un número positivo δ tal que para toda elección de x en X vale $d(x, \xi) < \delta) \implies d(f(x), f(\xi)) < \varepsilon/2$. Por otro lado, como las sucesiones $(x_{h(f)})_{n\geq 0}$ e $(y_{h(f)})_{n\geq 0}$ convergen a ξ, hay un entero positivo n_0 tal que para todo entero n mayor que n_0 es $d(x_{h(n)}, \xi) < \delta$ y $d(y_{h(n)}, \xi) < \delta$. Sea m un entero positivo mayor que n_0 y tal que $1/m < \delta$. Tenemos entonces que $h(m) \geq m \geq n_0$ y, en consecuencia, que $d(x_{h(m)}, \xi) < \delta$, $d(y_{h(m)}, \xi) < \delta$, y

$$\varepsilon \leq d(f(x_{h(m)}), f(y_{h(m)})) \leq d(f(x_{h(m)}), f(\xi)) + d(f(\xi), f(y_{h(m)})) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Esto es absurdo, y esta contradicción prueba la proposición. □

La segunda propiedad importante de los espacios compactos en relación a las funciones continuas es que estas los "preservan", en el sentido preciso de la siguiente proposición.

Proposición 9.4.2. *Sean X e Y dos espacios métricos y sea $f : X \to Y$ una función continua. Si K es un subconjunto compacto de X, entonces $f(K)$ es un subconjunto compacto de Y.*

Daremos dos demostraciones de esto, basadas cada una en una caracterización distinta de la compacidad.

Primera demostración. Sea K un subconjunto compacto de X y sea $\mathscr{U}$ un cubrimiento de $f(K)$ por abiertos de Y. El conjunto $\{f^{-1}(U) : U \in \mathscr{U}\}$ es entonces un cubrimiento de K y sus elementos son abiertos porque la función f es continua: como K es compacto, hay entonces un subconjunto finito $\mathscr{V}$ de $\mathscr{U}$ tal que $\{f^{-1}(V) : V \in \mathscr{V}\}$ es también un cubrimiento abierto de K. Esto implica inmediatamente que $\mathscr{V}$ es un cubrimiento de $f(K)$ y prueba que el conjunto $f(K)$ es compacto, como queremos. □

Segunda demostración. Sea K un subconjunto compacto de X y sea $(y_n)_{n\geq 0}$ una sucesión en Y con valores en el conjunto $f(K)$. Para cada $n \in \mathbb{N}_0$ hay entonces un elemento x_n de K tal que $y_n = f(x_n)$, y podemos considerar la sucesión $(x_n)_{n\geq 0}$ en K. Como el conjunto K es compacto, hay una función estrictamente creciente $h : \mathbb{N}_0 \to \mathbb{N}_0$ tal que la subsucesión $(x_{h(n)})_{n\geq 0}$ de $(x_n)_{n\geq 0}$ converge y, como la función f es continua, esto implica que la subsucesión $(y_{h(n)})_{n\geq 0} = (f(x_{h(n)}))_{n\geq 0}$ converge en Y. Esto prueba que el conjunto $f(K)$ es compacto. □

Ejercicio 9.4.3. De una tercera demostración de la proposición, usando esta vez la caracterización de los espacios compactos como aquellos que son completos y totalmente acotados.

Este resultado tiene muchos corolarios importantes.

Corolario 9.4.4. *Sean X e Y dos espacios métricos y supongamos que X es compacto. Toda función continua $X \to Y$ es acotada.*

Demostración. En efecto, la imagen de toda tal función es un subconjunto de Y que es compacto y, por lo tanto, acotado. □

Podemos usar la proposición para dar un criterio sencillo pero útil para decidir si un espacio es compacto.

Corolario 9.4.5. *Sean X e Y dos espacios métricos y sea $f : X \to Y$ una función continua. Si f es sobreyectiva y X es compacto, entonces Y es compacto.*

Demostración. Esto es consecuencia inmediata de la proposición. □

El siguiente corolario da un criterio extremadamente útil para probar que una función es un homeomorfismo.

Corolario 9.4.6. *Sean X e Y dos espacios métricos y sea $f : X \to Y$ una función continua y biyectiva. Si X es compacto, entonces f es un homeomorfismo.*

Demostración. Supongamos que el espacio X es compacto. Tenemos que probar que la función inversa $g \coloneqq f^{-1} : Y \to X$ es continua, y para ello es suficiente con mostrar que para cada subconjunto cerrado F de X el conjunto $g^{-1}(F)$ es cerrado en Y. Ahora bien, si F es un subconjunto cerrado de X, entonces, como X es compacto, F es compacto y la proposición nos dice que el conjunto $g^{-1}(F) = f(F)$ es compacto en Y, así que, en particular, es cerrado. □

Otra consecuencia fundamental de la proposición es la siguiente generalización del *Teorema de Bolzano–Weierstrass*, por Bernard Bolzano y Karl Weierstrass, también conocido como el *Teorema de los valores extremos*.

Corolario 9.4.7. *Sea X un espacio métrico no vacío y sea $f : X \to \mathbb{R}$ una función continua. Si el espacio X es compacto, entonces f es acotada y, de hecho, hay un punto x de X tal que $f(x) \geq f(y)$ para todo elemento y de X.*

Demostración. Supongamos que el espacio X es compacto. La proposición nos dice entonces que $f(X)$ es un subconjunto compacto de $\mathbb{R}$ y, en particular, que es cerrado y acotado. Podemos entonces considerar el número $\mu \coloneqq \sup f(X)$ y sabemos que, de hecho, $\mu \in f(X)$. Esto último nos dice que hay un elemento x de X tal que $f(x) = \mu$ y, por lo tanto, tal que $f(x) \geq f(y)$ para todo elemento y de X. □

Ejercicio 9.4.8. Sea X un espacio métrico compacto y sea $f : X \to \mathbb{R}$ una función continua. Pruebe las siguientes afirmaciones. Si y es un número real tal que $f(x) \neq t$ para todo $x \in X$, entonces hay un número positivo ε tal que $|f(x) - t| > \varepsilon$ para todo $x \in X$.

Ejercicio 9.4.9. Sea X un espacio métrico. Pruebe que si S y T son dos subconjuntos compactos y disjuntos de X, entonces hay elementos s y t de S y de T, respectivamente, tales que $d(s,t) = d(S,T)$.

El *Teorema de Bolzano–Weierstrass* tiene la siguiente recíproca parcial que nos da una nueva caracterización de los espacios métricos compactos, esta vez en términos de funciones.

Proposición 9.4.10. *Un espacio métrico X es compacto si y solamente si toda función continua $X \to \mathbb{R}$ es acotada.*

Notemos que esto tiene como consecuencia el hecho un poco sorprendente de que si toda función real continua sobre un espacio métrico es acotada entonces toda función real continua sobre ese espacio tiene máximo.

Demostración. Ya sabemos que la condición de la proposición es necesaria, así que solamente tenemos que probar que es suficiente. Supongamos que X es un espacio métrico que no es compacto, de manera que hay una sucesión $(x_n)_{n\geq 0}$ en X que no posee ninguna subsucesión convergente, y consideremos el conjunto $S \coloneqq \{x_n : n \in \mathbb{N}_0\}$. Queremos probar que este conjunto S es cerrado en X.

Sea x un punto de adherencia de S en X, de manera que hay una función $h : \mathbb{N}_0 \to \mathbb{N}_0$ tal que la sucesión $(x_{h(n)})_{n\geq 0}$ converge a x, y supongamos por un momento que la imagen $h(\mathbb{N}_0)$ de la función h es infinita, de manera que, de acuerdo al Lema 7.2.5, hay una función estrictamente creciente $k : \mathbb{N}_0 \to \mathbb{N}_0$ que tiene la misma imagen que h. Sea ε un número positivo. Como $(x_{h(n)})_{n\geq 0}$ converge a x, hay un entero positivo n_0 tal que para todo $n \in \mathbb{N}_0$ vale $n \geq n_0 \implies d(x_{h(n)}, x) < \varepsilon$. Si ponemos $J \coloneqq \{n \in \mathbb{N}_0 : d(x_{h(n)}, x) < \varepsilon\}$, entonces esto nos dice que $\mathbb{N}_0 \setminus J$ es finito, ya que está contenido en $[\![0, n_0 - 1]\!]$. Como $k(\mathbb{N}_0) = h(\mathbb{N}_0) = h(J) \cup h(\mathbb{N}_0 \setminus J)$ y el conjunto $h(\mathbb{N}_0 \setminus J)$ es finito, hay un entero no negativo n_1 tal que $k(n) \in h(J)$ para todo entero n mayor que n_1. Así, si n es un entero mayor que n_1, entonces hay otro entero no negativo m en J tal que $k(n) = h(m)$ y, por lo tanto, $d(x_{k(n)}, x) = d(x_{h(m)}, x) < \varepsilon$). Vemos así que la sucesión $(x_{h(n)})_{n\geq 0}$ converge a x: esto es imposible, porque se trata de una subsucesión de $(x_n)_{n\geq 0}$, que no tiene ninguna subsucesión convergente. Concluimos así que que la imagen de la función h es finita, de manera que la sucesión $(x_{h(n)})_{n\geq 0}$ toma solamente finitos valores y, de acuerdo al resultado del Ejercicio 7.1.20, su límite, esto es, x, es uno de esos valores. Esto prueba que x pertenece a S y, en definitiva, que el conjunto S es cerrado.

Más aún, este argumento prueba que el conjunto derivado S' es vacío. En efecto, si x fuera un punto de S', entonces podríamos elegir una función $h : \mathbb{N}_0 \to \mathbb{N}_0$ tal que la sucesión $(x_{h(n)})_{n\geq 0}$ converge a x y tiene sus componentes distintas dos a dos, pero lo que hicimos muestra que esto es imposible, porque la sucesión $(x_{h(n)})_{n\geq 0}$ solamente toma finitos valores.

El conjunto S es contable e infinito — si no lo fuera, la sucesión $(x_n)_{n\geq 0}$ tendría una subsucesión constante — y entonces hay una función biyectiva $l : S \to \mathbb{N}_0$. Como S no tiene puntos de acumulación, esta función l es continua y, como S es cerrado en X, el *Teorema de extensión de Tietze* en la forma que da el Ejercicio 8.4.12, nos dice que hay una función continua $f : X \to \mathbb{R}$ tal que $f|_S = l$. Como l no es acotada, tampoco lo es f. Esto prueba, como queremos, que la condición de la proposición es suficiente para que X sea compacto. □

Hay otras formas en la que podemos hacer la construcción de una función no acotada en la demostración de esta proposición. El siguiente ejercicio describe una.

Ejercicio 9.4.11. Sea X un espacio métrico que no es compacto, de manera que hay una sucesión $(x_n)_{n\geq 0}$ en X que no tiene ninguna subsucesión convergente.

(a) Pruebe que hay una función estrictamente creciente $h : \mathbb{N}_0 \to \mathbb{N}_0$ tal que para toda elección de i y j en $\mathbb{N}_0$ vale $x_{h(i)} = x_{h(j)} \iff i = j$.

(b) Pruebe que para cada $n \in \mathbb{N}_0$ el número

$$\delta_n \coloneqq \text{ínf}\{d(x_{h(n)}, x_{h(m)}) : m \in \mathbb{N}_0 \setminus \{n\}\}$$

es estrictamente positivo.

(c) Para cada $n \in \mathbb{N}_0$ consideremos la función

$$f_n : x \in X \mapsto \text{máx}\left\{1 - \frac{3}{\delta_n} d(x_{h(n)}, x), 0\right\} \in \mathbb{R}.$$

Pruebe que para cada $n \in \mathbb{N}_0$ la función f_n es continua, y que cada elemento x de X posee un entorno N tal que el conjunto

$$\{n \in \mathbb{N}_0 : \text{hay un punto } y \text{ en } N \text{ tal que } f_n(y) \neq 0\}$$

tiene como mucho un elemento.

(d) Muestre que para cada punto x de X la serie $\sum_{n=0}^{\infty} n \cdot f_n(x)$ converge, y que la función

$$f : x \in X \mapsto \sum_{n=0}^{\infty} n \cdot f_n(x) \in \mathbb{R}$$

que obtenemos de esa forma es continua y no acotada.

Veamos una aplicación del *Teorema de los valores extremos* 9.4.7.

Ejemplo 9.4.12. Probemos el *Teorema fundamental del álgebra*:

todo polinomio no constante de $\mathbb{C}[X]$ *posee una raíz en* $\mathbb{C}$.

Sea p un elemento no constante de $\mathbb{C}[X]$, de manera que hay un entero positivo d y números complejos $a_0, \ldots, a_d$ tales que $a_d \neq 0$ y $p = a_d X^d + a_{d-1} X^{d-1} + \cdots + a_1 X + a_0$. Sea $\alpha \coloneqq \text{máx}\{1, |a_0|, \ldots, |a_d|\}$ y sea

$$R \coloneqq \text{máx}\left\{1, \frac{2\alpha d}{|a_d|}, \left(\frac{2|p(0)|}{|a_d|}\right)^{1/d}\right\}.$$

Si z es un número complejo tal que $|z| \geq R$, entonces $|a_k|/|z^{d-k}| \leq \alpha/R$ para todo $k \in [\![0, d-1]\!]$ y

$$\begin{aligned} |p(z)| &= |a_d z^d + a_{d-1} z^{d-1} + \cdots + a_1 z + a_0| = |z^d| \cdot \left|a_d + a_{d-1}\frac{1}{z} + \cdots + a_1 \frac{1}{z^{d-1}} + a_0 \frac{1}{z^d}\right| \\ &\geq |z^d| \cdot \left(|a_d| - \frac{|a_{d-1}|}{|z|} - \cdots - \frac{|a_1|}{|z^{d-1}|} - \frac{|a_0|}{|z^d|}\right) \geq R^d \cdot \left(|a_d| - \frac{\alpha d}{R}\right) \geq \frac{|a_d|}{2} R^d \geq |p(0)|. \end{aligned}$$

Esto implica que $\text{ínf}\{|p(z)| : z \in \mathbb{C}, |z| \geq R\} \geq |p(0)|$ y, por lo tanto, que

$$\text{ínf}\{|p(z)| : z \in \mathbb{C}\} = \text{mín}\big\{\text{ínf}\{|p(z)| : z \in \mathbb{C}, |z| \geq R\}, \text{ínf}\{|p(z)| : z \in \mathbb{C}, |z| \leq R\}\big\}$$

$$\geq \min\bigl\{|p(0)|, \inf\{|p(z)| : z \in \mathbb{C}, |z| \leq R\}\bigr\}$$
$$= \inf\{|p(z)| : z \in \mathbb{C}, |z| \leq R\}.$$

Como el conjunto $B \coloneqq \{z \in \mathbb{C} : |z| \leq R\}$ es compacto y la función $z \in \mathbb{C} \mapsto |p(z)| \in \mathbb{R}$ es continua, el *Teorema de Bolzano–Weierstrass* nos dice que hay un elemento z_0 en B tal que

$$|p(z_0)| = \inf\{|p(z)| : z \in \mathbb{C}, |z| \leq R\} = \inf\{|p(z)| : z \in \mathbb{C}\}$$

y, por lo tanto, que $|p(z_0)| \leq |p(z)|$ para todo número complejo z.

Supongamos por un momento que $p(z_0) \neq 0$. El polinomio $p(X)$ posee una raíz exactamente cuando el polinomio $q(X) \coloneqq p(X + z_0)/p(z_0)$ posee una raíz en $\mathbb{C}$. El grado de q es d, es $q(0) = 1$, y

$$1 = |q(0)| \leq |q(z)| \text{ para todo } z \in \mathbb{C}. \tag{9.4}$$

En particular, hay escalares $b_1, \ldots, b_d \in \mathbb{C}$ tales que $b_d \neq 0$ y $q = b_d X^d + b_{d-1}X^{d-1} + \cdots + b_1 X + 1$.

El conjunto $\{i \in [\![d]\!] : b_j = 0 \text{ para cada } j \in [\![i-1]\!]\}$ no es vacío, porque contiene a 1. Sea n su mayor elemento, de manera que $b_1 = \cdots = b_{n-1} = 0$ y $b_n \neq 0$. Sea ζ un elemento de $\mathbb{C}$ tal que $\zeta^n = -1/b_n$, necesariamente no nulo, y consideremos el polinomio

$$r \coloneqq b_n^{-1}(b_d X^{d-n} + b_{d-1}X^{d-n-1} + \cdots + b_{n+1}X^1),$$

que es nulo cuando $n = d$. En cualquier caso, la función $z \in \mathbb{C} \mapsto r(z) \in \mathbb{C}$ es continua y vale 0 en 0, así que hay un número positivo δ tal que para todo $z \in \mathbb{C}$ vale

$$|z| < \delta \implies |r(z)| < \frac{1}{2}.$$

Ahora bien, en vista de la forma en que elegimos a n es claro que

$$\begin{aligned} q &= b_d X^d + b_{d-1}X^{d-1} + \cdots + b_n X^n + 1 \\ &= \frac{b_d X^{d-n} + b_{d-1}X^{d-n-1} + \cdots + b_{n+1}X^1}{b_n} \cdot b_n X^n + b_n X^n + 1 \\ &= r(X) \cdot b_n X^n + b_n X^n + 1. \end{aligned}$$

Sea t un número real tal que $0 < t < \min\{\delta/|\zeta|, 1/2\}$. Como

$$q(t\zeta) = -r(t\zeta) \cdot t^n - t^n + 1,$$

tenemos que

$$|q(t\zeta)| \leq |r(t\zeta)| \cdot |t^n| + |1 - t^n| \leq \frac{1}{2} \cdot t^n + 1 - t^n = 1 - \frac{1}{2}t^n < 1,$$

y esto es absurdo, ya que contradice (9.4). Esta contradicción provino de haber supuesto arriba que $p(z_0) \neq 0$, así que el número z_0 debe ser una raíz del polinomio p: esto prueba lo que queremos.

9.5. Números de Lebesgue

El siguiente concepto es útil con frecuencia.

Definición 9.5.1. Sea X un espacio métrico. Decimos que un número positivo r es un ***número de Lebesgue*** de un cubrimiento $\mathscr{U}$ de X si para todo subconjunto no vacío S de X tal que $\operatorname{diam} S < r$ hay un elemento U de $\mathscr{U}$ tal que $S \subseteq U$.

El nombre recuerda a Henri Léon Lebesgue.

Proposición 9.5.2. *Todo cubrimiento abierto de un espacio métrico compacto posee un número de Lebesgue.*

Demostración. Sea X un espacio métrico compacto y sea $\mathscr{U}$ un cubrimiento abierto de X. Para cada elemento x de X hay un elemento U_x de $\mathscr{U}$ tal que $x \in U_x$ y, como U_x es abierto, un número real positivo r_x tal que $B_{r_x}(x) \subseteq U_x$. El conjunto $\{B_{r_x/2}(x) : x \in X\}$ es un cubrimiento abierto de X, que este espacio sea compacto implica que tiene un subconjunto finito F tal que $X \subseteq \bigcup_{x\in F} B_{r_x/2}(x)$. Consideremos el número $r \coloneqq \frac{1}{2}\min\{r_x : x \in F\}$, que es positivo y mostremos que r es un número de Lebesgue para $\mathscr{U}$.

Sea S un subconjunto no vacío de X tal que $\operatorname{diam} S < r$ y sea s un elemento de S. Como $S \subseteq X = \bigcup_{x\in F} B_{r_x/2}(x)$, hay un elemento x de F tal que $s \in B_{r_x/2}(x)$. Si t es un elemento cualquiera de S, entonces

$$d(t,x) \leq d(t,s) + d(s,x) < r + \frac{r_x}{2} \leq r_x,$$

de manera que $t \in B_{r_x}(x) \subseteq U_x$. Vemos así que $S \subseteq U_x \in \mathscr{U}$. Esto prueba lo que queremos. □

Esta proposición tiene el siguiente recíproco parcial.

Proposición 9.5.3. *Muestre que un espacio métrico X es compacto si y solamente si es totalmente acotado y todos sus cubrimientos abierto poseen un número de Lebesgue.*

Demostración. Como todo espacio métrico compacto es totalmente acotado y vale la Proposición 9.5.2, sabemos que la condición de la proposición es necesaria. Para ver que es suficiente alcanza con mostrar que un espacio que la cumple es necesariamente completo.

Sea X un espacio métrico que no es completo, de manera que hay una sucesión $(x_n)_{n\geq 0}$ en X que satisface la condición de Cauchy pero no converge. Notemos que si x es un punto de X, entonces la sucesión $(d(x,x_n))_{n\geq 0}$ converge en $\mathbb{R}$, ya que satisface la condición de Cauchy: en efecto, si ε es un número positivo hay un entero no negativo n_0 tal que para toda elección de m y n en $\mathbb{N}_0$ vale $m, n \geq 0 \implies d(x_m, x_n) < \varepsilon/2$, y entonces cada vez que m y n son enteros mayores

que n_0 tenemos que $|d(x,x_n) - d(x,x_m)| \leq d(x_n,x_m) < \varepsilon$. Esto nos dice que hay una función

$$\delta : x \in X \mapsto \lim_{n\to\infty} d(x,x_n) \in \mathbb{R}.$$

Notemos que δ no se anula: si x fuera un punto de X tal que $\delta(x) = 0$, entonces claramente x sería un límite de la sucesión $(x_n)_{n\geq 0}$, y esta no tiene ninguno.

La función δ es uniformemente continua: si ε es un número positivo y x e y son dos elementos de X tales que $d(x,y) < \varepsilon/2$, entonces $|d(x,x_n) - d(y,x_n)| \leq d(x,y) < \varepsilon/2$ para todo $n \in \mathbb{N}_0$ y, por lo tanto,

$$|\delta(x) - \delta(y)| = \left|\lim_{n\to\infty} d(x,x_n) - \lim_{n\to\infty} d(y,x_n)\right| = \lim_{n\to\infty} \left|d(x,x_n) - d(y,x_n)\right| \leq \frac{\varepsilon}{2} < \varepsilon.$$

En particular, por supuesto, la función es continua, y para cada $n \in \mathbb{N}_0$ el conjunto

$$U_n \coloneqq \{x \in X : \delta(x) > 1/(n+1)\}$$

es abierto en X. Como la función δ no se anula, el conjunto $\mathscr{U} \coloneqq \{U_n : n \in \mathbb{N}_0\}$ es un cubrimiento abierto de X. Queremos mostrar que $\mathscr{U}$ no posee un número de Lebesgue.

Sea r un número positivo. Como la sucesión $(x_n)_{n\geq 0}$ satisface la condición de Cauchy hay un entero positivo N tal que el conjunto $S \coloneqq \{x_n : n \in \mathbb{N}_0, n \geq N\}$ tiene diámetro menor que r. Supongamos por un momento que hay un elemento de $\mathscr{U}$ que contiene a S, esto es, que hay un elemento k de $\mathbb{N}_0$ tal que $S \subseteq U_k$. Ahora bien, como la sucesión satisface la condición de Cauchy, hay un entero positivo M tal que $d(x_m,x_n) < 1/(k+2)$ siempre que m y n son enteros mayores o iguales que M. En particular, esto implica que $d(x_M,x_n) < 1/(k+2)$ siempre que $n \geq M$ y, por lo tanto, que $\delta(x_M) \leq 1/(k+1) < 1/k + 1$, de manera que $x_M \notin U_k$: esto es absurdo, ya que $x_M \in S$. Vemos así que el número rf no es un número de Lebesgue para el cubrimiento $\mathscr{U}$ y, por lo tanto, que este cubrimiento no posee ningún número de Lebesgue. □

9.6. Variaciones sobre la idea de compacidad

No todos los espacios métricos son compactos, pero hay algunas variaciones sobre la misma idea que son útiles en muchas situaciones. Presentaremos dos en esta sección.

Definición 9.6.1. Un espacio métrico es ***localmente compacto*** si cada uno de sus puntos posee un entorno compacto.

Podemos dar ejemplos fácilmente. Así, todo espacio métrico con métrica discreta es localmente

compacto, ya que todas sus bolas abiertas de radio 1/2 son compactas. Por otro lado, un espacio compacto es localmente compacto simplemente porque es él mismo un entorno compacto de cada uno de sus puntos. La siguiente proposición nos da ejemplos más sustanciales.

Proposición 9.6.2. *Un espacio normado es localmente compacto si y solamente si tiene dimensión finita.*

Demostración. Si V es un espacio normado de dimensión finita y x es un elemento de V, entonces la bola cerrada $\overline{B}_1(x)$ es un entorno de x y es cerrada y acotada en V, así que, de acuerdo al *Teorema de Heine–Borel* 9.2.5, es compacta. La condición de la proposición es, por lo tanto, suficiente.

Sea, por otro lado, V un espacio normado localmente compacto y sea N un entorno compacto de 0 en V. Hay entonces un número positivo r tal que $B_r(0) \subseteq N$ y, como N es cerrado, $\overline{B}_r(0) \subseteq N$: como N es compacto, vemos así que $\overline{B}_r(0)$ también lo es y podemos concluir gracias a la Proposición 9.2.11 que V tiene dimensión finita. □

Ejemplo 9.6.3. El espacio $\mathbb{Q}$ con su topología usual no es localmente compacto. De hecho, ninguno de sus puntos posee ningún entorno compacto.

En efecto, sea q un número racional y sea N un entorno de q en $\mathbb{Q}$, de manera que hay un número positivo ε tal que $(q-\varepsilon, q+\varepsilon) \cap \mathbb{Q} \subseteq N$. Como el conjunto $(q-\varepsilon, q+\varepsilon) \cap \mathbb{Q}$ es contable y $(q-\varepsilon, q+\varepsilon)$ no, hay un elemento ξ de $(q-\varepsilon, q+\varepsilon)$ que no es racional. Por otro lado, como $\mathbb{Q}$ es denso en $\mathbb{R}$, hay una sucesión $(x_n)_{n\geq 0}$ con valores en $(q-\varepsilon, q+\varepsilon) \cap \mathbb{Q}$ y, por lo tanto, en N, que converge a ξ. Notemos que entonces esa sucesión no tiene ninguna subsucesión convergente en $\mathbb{Q}$ y que, por lo tanto, el conjunto N no es compacto.

La definición de compacidad local pide la existencia de *un* entorno compacto para cada punto, pero implica que hay toda una base de entornos compactos.

Proposición 9.6.4. *Cada punto de un espacio métrico localmente compacto posee una base de entornos compactos.*

Demostración. Sea X un espacio métrico localmente compacto y sea x un punto de X, de manera que hay un entorno compacto N de x, de manera que, en particular, hay un número positivo r tal que $B_r(x) \subseteq N$. Mostremos que el conjunto $\mathscr{N}$ de todos los entornos cerrados de x contenidos en N es una base de entornos de X.

Sea M un entorno de x, de manera que hay un número positivo s tal que $B_s(x) \subseteq M$, y consideremos el número $t \coloneqq \frac{1}{2}\min\{r, s\}$, que es positivo. Como $t < r$, tenemos que $\overline{B}_t(x) \subseteq B_r(x) \subseteq N$ y esto implica, ya que N es compacto, que $\overline{B}_t(x)$ es un entorno compacto de x y, por lo tanto, un elemento de $\mathscr{N}$. Como además $\overline{B}_t(x) \subseteq B_s(x) \subseteq M$, ya que $t < s$, esto prueba que $\mathscr{N}$ es una base de entornos de x, como queremos. □

Una consecuencia inmediata de esto es:

Corolario 9.6.5. *Todo subespacio abierto o cerrado de un espacio métrico localmente compacto es localmente compacto.*

Demostración. Sea X un espacio métrico localmente compacto, y sean primero F un subespacio cerrado de X y x un punto de F. Como X es localmente compacto, hay un entorno N de x en X que es compacto y entonces $N \cap F$ es un entorno de x en F que es compacto, ya que es cerrado en el compacto N.

Sea ahora U un subespacio abierto de X y sea x un punto de U. Como U es un entorno de x en X y X es localmente compacto, la Proposición 9.6.4 nos dice que hay un entorno compacto N de x en X tal que $N \subseteq U$. Este conjunto N es entonces un entorno compacto de x en U, y esto prueba que U es un espacio localmente compacto. □

Otro corolario útil de la proposición es el siguiente.

Corolario 9.6.6. *La topología de un espacio métrico localmente compacto y separable posee una base contable cuyos elementos tienen clausura compacta.*

Demostración. Sea X un espacio métrico localmente compacto y separable. De acuerdo a la Proposición 6.6.16, hay una base $\mathscr{B}$ de la topología de X que es contable. Sea $\mathscr{B}'$ el subconjunto de $\mathscr{B}$ de los elementos que tienen clausura compacta. Claramente $\mathscr{B}'$ es contable. Para probar el corolario es suficiente que mostremos que $\mathscr{B}'$ es una base de la topología de X.

Sea U un subconjunto abierto de X y sea x un punto de U. Como X es localmente compacto, la Proposición 9.6.4 nos dice que x posee una base de entornos compactos y hay, por lo tanto, un entorno compacto N de x tal que $N \subseteq U$. Por otro lado, como $\mathscr{B}$ es una base de la topología de X, hay un elemento B de $\mathscr{B}$ tal que $B \subseteq N$. Como N es compacto en X, es cerrado y, como consecuencia de ello, $\overline{B} \subseteq N$: esto implica que la clausura $\overline{B}$ es compacta y que $B \in \mathscr{B}'$. Como $x \in B \subseteq N$, esto prueba que $\mathscr{B}'$ es una base de la topología de X, como queremos. □

La siguiente proposición da una propiedad útil de los espacios localmente compactos.

Proposición 9.6.7. *Sea X un espacio localmente compacto. Si K es un subconjunto compacto de X y U es un abierto de X tal que $K \subseteq U$, entonces hay un conjunto compacto L de X tal que $K \subseteq L^\circ$ y $L \subseteq U$.*

Demostración. Sea K un subconjunto compacto de X y sea U un subconjunto abierto de X tal que $K \subseteq U$. Como cada punto de X tiene una base de entornos compactos, para cada punto x de K podemos elegir un entorno compacto N_x de x tal que $N_x \subseteq U$. El conjunto $\{N_x^\circ : x \in K\}$ es

entonces un cubrimiento abierto de K, así que que K sea compacto nos dice que hay un subconjunto finito F de K tal que $K \subseteq \bigcup_{x\in F} N_x^\circ$. Pongamos $L := \bigcup_{x\in F} \overline{N_x^\circ}$. Como el conjunto F es finito, es claro que L es cerrado. Por otro lado, para cada $x \in F$ es $N_x^\circ \subseteq N_x$, así que $\overline{N_x^\circ} \subseteq N_x$ y, por lo tanto, tenemos que $L = \bigcup_{x\in F} \overline{N_x^\circ} \subseteq \bigcup_{x\in F} N_x \subseteq U$. Como L es cerrado y la unión $\bigcup_{x\in F} N_x$ es compacta, vemos que L es compacto. Finalmente, es

$$K \subseteq \bigcup_{x\in F} N_x^\circ \subseteq \bigcup_{x\in F} \overline{N_x^\circ} = L$$

y la unión $\bigcup_{x\in F} N_x^\circ$ es abierta, así que $K \subseteq L^\circ$. □

La condición de Lindelöf y la condición de compacidad imponen que todo cubrimiento abierto del espacio tenga un subcubrimiento de cardinal acotado. Presentaremos ahora una generalización extremadamente importante de la compacidad que no se basa en la existencia de subcubrimientos sino de refinamientos, en el siguiente sentido.

Definición 9.6.8. Sea X un conjunto. Un cubrimiento abierto $\mathscr{U}$ de X es un ***refinamiento*** de otro cubrimiento abierto $\mathscr{V}$ de X si todo elemento de $\mathscr{U}$ está contenido en un elemento de $\mathscr{V}$, y en ese caso escribimos $\mathscr{U} \preceq \mathscr{V}$.

Cuando un cubrimiento abierto refina a otro sus abiertos son «más chicos» que los de este. Es claro que su $\mathscr{U}$ es un subcubrimiento de un cubrimiento abierto $\mathscr{V}$, entonces $\mathscr{U}$ es un refinamiento de $\mathscr{V}$, así que esta relación generaliza a la de subcubrimientos. No coinciden, por supuesto, $\{(-\infty, 1), (-1, +\infty)\}$ es un refinamiento del cubrimiento $\{\mathbb{R}\}$ de $\mathbb{R}$ que no es un subcubrimiento.

Ejercicio 9.6.9. Sea X un conjunto. Muestre que la relación $\preceq$ de refinamiento es una relación de orden sobre el conjunto de todos los cubrimientos abiertos de X.

Recordemos de la Definición 6.5.13 que cuando X es un espacio métrico decimos que subconjunto $\mathscr{F}$ de $\mathscr{P}(X)$ es ***localmente finito*** si todo punto x de X posee un entorno N tal que el conjunto $\{F \in \mathscr{F} : F \cap N \neq \varnothing\}$ es finito.

Definición 9.6.10. Un espacio métrico X es ***paracompacto*** si todo cubrimiento abierto de X admite un refinamiento localmente finito.

Esta definición fue propuesta originalmente por Jean Dieudonné en [Die1944].

Todo espacio compacto es paracompacto: en efecto, cada uno de sus cubrimientos posee un subcubrimiento finito y este es un refinamiento de aquel que es evidentemente localmente finito.

Ejercicio 9.6.11. Muestre que un subespacio cerrado de un espacio métrico paracompacto es paracompacto.

Proposición 9.6.12. *Un espacio métrico localmente compacto y separable es paracompacto.*

La demostración de esto es un gran ejemplo del tipo de cosas que podemos hacer gracias a la numerabilidad en topología.

Demostración. Sea X un espacio métrico localmente compacto y separable.

Primer paso. Sea $\mathscr{U}$ un cubrimiento abierto de X. Como X es separable, el Corolario 9.6.6 nos dice que hay una base contable $\mathscr{B}$ de la topología de X cuyos elementos tienen clausura compacta. Sea $\phi : \mathbb{N}_0 \to \mathscr{B}$ una función sobreyectiva.

Si x es un elemento de X, entonces hay un elemento U_x de $\mathscr{U}$ tal que $x \in U_x$ y, como U_x es abierto y $\mathscr{B}$ es una base de la topología de X, hay un elemento $v(x)$ de $\mathbb{N}_0$ tal que $x \in \phi(v(x)) \subseteq U_x$. El conjunto $\mathscr{V} \coloneqq \{\phi(v(x)) : x \in X\}$ es un subconjunto de $\mathscr{B}$, así que es contable y sus elementos son abiertos de X y tienen clausura compacta. Como $x \in \phi(v(x))$ para todo $x \in X$, es $X \subseteq \bigcup_{x\in X} \phi(v(x)) = \bigcup \mathscr{V}$, así que $\mathscr{V}$ es un cubrimiento de X. Finalmente, si V es un elemento de $\mathscr{V}$, entonces hay un punto x de X tal que $V = \phi(v(x))$ y, por lo tanto, $V \subseteq U_x \in \mathscr{U}$: esto nos dice que $\mathscr{V}$ es un refinamiento de $\mathscr{U}$.

Como la relación de refinamiento es transitiva, podemos concluir de esto que para probar que X es paracompacto basta mostrar que

> *todo cubrimiento abierto contable de X cuyos elementos tienen clausura compacta posee un refinamiento localmente finito.*

Segundo paso. Sea ahora $\mathscr{U}$ un cubrimiento abierto contable de X cuyos elementos tienen clausura compacta. Como $\mathscr{U}$ es contable, hay una función sobreyectiva $\phi : \mathbb{N}_0 \to \mathscr{U}$. Para cada $n \in \mathbb{N}_0$ pongamos $U_n \coloneqq \bigcup_{i=0}^{n} \phi(i)$, que es un abierto de X. Notemos que el conjunto $\mathscr{U}' \coloneqq \{U_n : n \in \mathbb{N}_0\}$ es un cubrimiento abierto de X, ya que

$$\bigcup \mathscr{U}' = \bigcup_{n=0}^{\infty} U_n = \bigcup_{n=0}^{\infty} \bigcup_{i=0}^{n} \phi(i) = \bigcup_{n=0}^{\infty} \phi(n) = \bigcup \mathscr{U} = X,$$

y que

$$U_m \subseteq U_n \text{ siempre que } m \text{ y } n \text{ son elementos de } \mathbb{N}_0 \text{ tales que } m \leq n. \tag{9.5}$$

Sea $n \in \mathbb{N}_0$. Es

$$\overline{U_n} = \overline{\bigcup_{i=0}^{n} \phi(i)} = \bigcup_{i=0}^{n} \overline{\phi(i)},$$

y esta última unión es compacta, ya que es finita y para cada $i \in [\![0, n]\!]$ el conjunto $\phi(i)$ tiene clausura compacta, así que el conjunto $\overline{U_n}$ es compacto. Como $\mathscr{U}'$ es un cubrimiento abierto de X, hay un subconjunto finito F de $\mathscr{U}''$ que es un cubrimiento de $\overline{U_n}$ y, por lo tanto, hay un subconjunto

finito G de $\mathbb{N}_0$ tal que $\{U_m : m \in G\}$ es un cubrimiento de $\overline{U_n}$. Si ponemos $\mu \coloneqq \max G \cup \{n+1\}$, entonces (9.5) implica que, de hecho, $\overline{U_n} \subseteq U_\mu$ y $\mu > n$.

Esto prueba que para cada $n \in \mathbb{N}_0$ el conjunto $R(n) \coloneqq \{m \in \mathbb{N}_0 : m > n, \overline{U_n} \subseteq U_m\}$ no es vacío. El *Principio de elección dependiente*, entonces, nos dice que hay una función $f : \mathbb{N}_0 \to \mathbb{N}_0$ tal que $f(0) = 0$ y $f(n+1) \in R(n)$ para todo $n \in \mathbb{N}_0$. Esta función es claramente estrictamente creciente. Para cada $n \in \mathbb{N}_0$ sea $V_n \coloneqq U_{f(n)}$, y consideremos el conjunto $\mathscr{V} \coloneqq \{V_n : n \in \mathbb{N}_0\}$.

Es claro que $\mathscr{V}$ es contable. Los elementos de $\mathscr{V}$ son abiertos y tienen clausura compacta, ya que son elementos de $\mathscr{U}'$. Para cada elemento x de X existe un elemento n de $\mathbb{N}_0$ tal que $x \in U_n \subseteq U_{f(n)} = V_n$, ya que $\mathscr{U}'$ es un cubrimiento de X y la función f es estrictamente creciente, y esto nos dice que $\mathscr{V}$ es un cubrimiento de X. Finalmente, si $\mathscr{W}$ es un refinamiento de $\mathscr{V}$ y W es un elemento de $\mathscr{W}$, entonces hay un elemento V de $\mathscr{V}$ tal que $W \subseteq V$ y un elemento n de $\mathbb{N}_0$ tal que $V = V_n = U_{f(n)}$, y es $W \subseteq V = U_{f(n)} \in \mathscr{U}$: esto nos dice que $\mathscr{W}$ es un refinamiento de $\mathscr{U}$.

La conclusión de todo esto es que para probar que X es paracompacto es suficiente que mostremos que

> *todo cubrimiento abierto de X de la forma $\{V_n : n \in \mathbb{N}_0\}$ con $\overline{V_n}$ compacto y $\overline{V_n} \subseteq V_{n+1}$ para todo $n \in \mathbb{N}_0$ admite un refinamiento localemente finito.*

Tercer paso. Sea $\mathscr{U} = \{U_n : n \in \mathbb{N}_0\}$ un cubrimiento abierto de X tal que para todo $n \in \mathbb{N}_0$ la clausura $\overline{U_n}$ es compacta y está contenida en U_{n+1}. Pongamos $U_{-1} \coloneqq \varnothing$ y $U_{-2} \coloneqq \varnothing$.

- Como $\overline{U_0}$ está contenido en el abierto U_1, para cada $x \in \overline{U_0}$ hay un número positivo $r_{x,0}$ tal que $B_{r_{x,0}}(x) \subseteq U_1$. El conjunto $\{B_{r_{x,0}}(x) : x \in \overline{U_0}\}$ es un cubrimiento abierto de $\overline{U_0}$: como este es compacto, hay un subconjunto finito F_0 de $\overline{U_0}$ tal que $\mathscr{V}_0 \coloneqq \{B_{r_{r,0}}(x) : x \in F_0\}$ es un cubrimiento abierto finito de $\overline{U_0}$. Pongamos $K_0 \coloneqq \overline{U_0}$.
- Sea ahora n un elemento de $\mathbb{N}$. El conjunto $K_n \coloneqq \overline{U_n} \setminus U_{n-1}$ es un cerrado contenido en el compacto $\overline{U_n}$, así que es él mismo compacto, y está contenido en $W_n \coloneqq U_{n+1} \setminus \overline{U_{n-2}}$. Para cada $x \in K_n$ hay un número positivo $r_{x,n}$ tal que $B_{r_{x,n}}(x) \subseteq W_n$, ya que W_n es abierto, y el conjunto $\{B_{r_{x,n}}(x) : x \in K_n\}$ es un cubrimiento abierto de K_n. Hay, por lo tanto, un subconjunto finito F_n de K_n tal que $\mathscr{V}_n \coloneqq \{B_{r_{x,n}}(x) : x \in F_x\}$ es un cubrimiento abierto de K_n.

Notemos que para todo $n \in \mathbb{N}_0$ el conjunto $\mathscr{V}_n$ es un cubrimiento abierto finito de K_n y que sus elementos están todos contenidos en $U_{n+1} \setminus \overline{U_{n-2}}$. Por otro lado, para todo $n \in \mathbb{N}_0$ es $U_n \subseteq \overline{U_n} \subseteq U_{n+1}$, así que $U_n \subseteq (\overline{U_n} \setminus U_{n-1}) \cup U_{n-1}$: un argumento inductivo evidente prueba a partir de eso que

$$X = \bigcup_{n=0}^{\infty} U_n = \overline{U_0} \cup \bigcup_{n=1}^{\infty} (\overline{U_n} \setminus U_{n-1}) = \bigcup_{n=0}^{\infty} K_n. \tag{9.6}$$

Sea $\mathscr{V} \coloneqq \bigcup_{n=0}^{\infty} \mathscr{V}_n$. Para terminar la prueba de la proposición es suficiente que mostremos que $\mathscr{V}$ es un cubrimiento abierto localmente finito de X que refina a $\mathscr{U}$.

- Es claro que los elementos de $\mathscr{V}$ son subconjuntos abiertos de X.

- Si x es un punto de X, entonces en vista de (9.6) sabemos que hay un elemento n de $\mathbb{N}_0$ tal que $x \in K_n$ y, por lo tanto, que x pertenece a un elemento de $\mathscr{V}_n$. Esto nos dice que $\mathscr{V}$ es un cubrimiento abierto de X.
- Si V es un elemento de $\mathscr{V}$, entonces hay un elemento n de $\mathbb{N}_0$ tal que $V \in \mathscr{V}_n$ y, por lo tanto, $V \subseteq U_{n+1} \in \mathscr{U}$. Esto nos dice que $\mathscr{V}$ es un refinamiento del cubrimiento $\mathscr{U}$.
- Sea, finalmente, x un elemento de X. Sabemos que hay un elemento n de $\mathbb{N}_0$ tal que $x \in K_n$. Si m es un entero tal que $m \geq n+3$ y V es un elemento de $\mathscr{V}_m$, entonces $V \subseteq W_m = U_{m+1} \setminus \overline{U_{m-2}}$, así que $V \cap U_{n+1} \subseteq V \cap \overline{U_{m-2}} = \varnothing$. Esto nos dice que el conjunto $\{V \in \mathscr{V} : V \cap U_{n+1} \neq \varnothing\}$ está contenido en $\bigcup_{i=0}^{n+2} \mathscr{V}_n$ y que es, en particular, finito. Como U_{n+1} es un entorno de x, esto prueba que $\mathscr{V}$ es un cubrimiento localmente finito de X. □

De hecho, vale el siguiente teorema:

Teorema 9.6.13. *Todo espacio métrico es paracompacto.* □

Esto fue probado originalmente por Arthur Harold Stone en [Sto1948]. La demostración más sencilla que se conoce es la dada por Mary Ellen Rudin en [Rud1969] — ocupa exactamente una carilla! Todas las demostraciones conocidas dependen explícitamente del *Axioma de elección* — vía el *Principio de buena ordenación* de Zermelo. Se sabe, de hecho, que no es posible probar el teorema en la teoría de conjuntos de Zermelo–Fraenkel, ni siquiera aumentándola con el principio de elección dependiente. El lector podrá encontrar la demostración «usual» de este teorema en la Sección 41 del libro [Mun2000] de James Munkres. Nosotros la omitiremos.

9.7. La convergencia uniforme sobre compactos

El hecho de que todas las funciones continuas sobre un espacio compacto sean acotadas tiene la siguiente consecuencia.

Proposición 9.7.1. *Sean X e Y dos espacios métricos, sea $C(X,Y)$ el conjunto de todas las funciones continuas $X \to Y$, y supongamos que X es compacto.*

(*i*) *Si $f, g : X \to Y$ son dos funciones continuas, entonces la función $x \in X \mapsto d(f(x), g(x)) \in \mathbb{R}$ es continua y acotada, de manera que podemos considerar el número*

$$d(f,g) \coloneqq \sup\{d(f(x), g(x)) : x \in X\}.$$

(*ii*) *La función $d : C(X,Y) \times C(X,Y) \to \mathbb{R}$ que obtenemos de esta forma es una métrica sobre el conjunto $C(X,Y)$.*

(*iii*) *Si el espacio métrico Y es completo, entonces también lo es $C(X,Y)$ con respecto a esta métrica.*

Demostración. Si f y g son dos elementos de $C(X,Y)$, entonces $x \in X \mapsto d(f(x),g(x)) \in \mathbb{R}$ es continua, y que es acotada es entonces consecuencia del Corolario 9.4.7. Esto prueba la parte (*i*) de la proposición. Por otro lado, se acuerdo al Corolario 9.4.5, toda función continua $X \to Y$ tiene imagen compacta, así que es acotada: esto nos dice que el conjunto $C(X,Y)$ coincide con el conjunto $C_b(X,Y)$ de las funciones continuas $X \to Y$ que son acotadas. Que la función d de la parte (*ii*) de la proposición es una métrica y que con respecto a ella $C(X,Y)$ es un espacio métrico completo, entonces, es consecuencia de la Proposición 8.6.1. □

Es más difícil hacer del conjunto $C(X,Y)$ de todas las funciones continuas $X \to Y$ un espacio métrico cuando el espacio X no es compacto. Discutamos una forma de hacerlo.

Definición 9.7.2. Sea X un espacio métrico. Una ***exhaución por compactos*** de X es una sucesión $(K_n)_{n\geq 0}$ de subconjuntos compactos de X tal que $K_n \subseteq K_{n+1}^\circ$ para todo $n \in \mathbb{N}_0$ y $X = \bigcup_{n=0}^\infty K_n$.

Por ejemplo, la sucesión $(B_{n+1}(0))_{n\geq 0}$ es una exhaución de $\mathbb{R}^d$.

Proposición 9.7.3. *Sea X un espacio métrico. Si X posee una exhaución por compactos, entonces X es localmente compacto. Por otro lado, las siguientes afirmaciones son equivalentes:*

(*a*) *X posee una exhaución por compactos.*
(*b*) *X es localmente compacto e igual a la unión de una familia contable de subconjuntos compactos.*
(*c*) *X es localmente compacto y satisface la condición de Lindelöf.*

Recordemos que un espacio satisface la condición de Lindelöf si y solamente si es separable y si y solamente posee una base numerable.

Demostración. Supongamos que $(K_n)_{n\geq 0}$ es una exhaución por compactos de X. Si x es un elemento de X, hay entonces un elemento n de $\mathbb{N}_0$ tal que $x \in K_n \subseteq K_{n+1}^\circ$ y, por lo tanto, el conjunto K_{n+1} es un entorno compacto de x. Esto prueba la primera parte de la proposición, y la implicación (*a*) $\Rightarrow$ (*b*).

Supongamos ahora que X es localmente compacto y que $(K_n)_{n\geq 0}$ es una sucesión de subconjuntos compactos de X tal que $X = \bigcup_{n=0}^\infty K_n$, y consideremos un cubrimiento abierto contable $\mathscr{U}$ de X. Para cada $n \in \mathbb{N}_0$ hay entonces un subconjunto finito $\mathscr{V}_n$ de $\mathscr{U}$ tal que $K_n \subseteq \bigcup \mathscr{V}_n$, y es claro que si ponemos $\mathscr{V} := \bigcup_{n=0}^\infty \mathscr{V}_n$ entonces $\mathscr{V}$ es un subcubrimiento de $\mathscr{U}$ que es contable. Vemos así que X satisface la condición de Lindelöf y, por lo tanto, que vale la implicación (*b*) $\Rightarrow$ (*c*).

Finalmente, supongamos que X es localmente compacto y que satisface la condición de Lindelöf. Para cada punto x de X podemos elegir un entorno compacto N_x de X. El conjunto $\{N_x^\circ : x \in X\}$ es entonces un cubrimiento abierto de X y la condición de Lindelöf implica que

podemos elegir un conjunto contable F de X tal que $\{N_x^\circ : x \in F\}$ es un cubrimiento de X. Sea $\phi : \mathbb{N}_0 \to F$ una función sobreyectiva y para cada $n \in \mathbb{N}_0$ escribamos $U_n \coloneqq N_{\phi(n)}^\circ$. Obtenemos así una sucesión $(U_n)_{n\geq 0}$ de abierto de X con clausura compacta que cubren a X.

Sea $\mathscr{K}$ el conjunto de los compactos de X y digamos que una función $\phi : [\![0, n]\!] \to \mathscr{K}$ es buena si para cada $i \in [\![0, n]\!]$ es $U_i \subseteq \phi(i)^\circ$ y para cada $i \in [\![n]\!]$ es $\phi(i-1) \subseteq \phi(i)^\circ$.

- Como U_0 tiene clausura compacta, la Proposición 9.6.7 nos dice que hay un conjunto compacto K_0 tal que $U_0 \subseteq K_0^\circ$. Podemos considerar entonces la función $\phi : [\![0, 0]\!] \to \mathscr{K}$ tal que $\phi(0) = K_0$, y es claro que es buena.
- Supongamos ahora que n es un elemento de $\mathbb{N}_0$ y que $\phi : [\![0, n]\!] \to \mathscr{K}$ es una función buena. El conjunto $L \coloneqq \overline{U}_{n+1} \cup \phi(n)$ es compacto, así que la Proposición 9.6.7 nos dice que hay un subconjunto compacto K de X tal que $L \subseteq K^\circ$. La función $\psi : [\![0, n+1]\!] \to \mathscr{K}$ que en cada elemento i de $[\![0, n+1]\!]$ toma el valor

$$\psi(i) = \begin{cases} \phi(i) & \text{si } i < n+1; \\ K & \text{si } i = n+1 \end{cases}$$

  es claramente buena.

Estas dos observaciones nos permiten concluir que hay una función $\phi : \mathbb{N}_0 \to \mathscr{K}$ tal que para cada $n \in \mathbb{N}_0$ la restricción $\phi|_{[\![0,n]\!]}$ es buena. Si ponemos $K_n \coloneqq \phi(n)$ para cada $n \in \mathbb{N}_0$, entonces obtenemos una sucesión $(K_n)_{n\geq 0}$ de subconjuntos compactos de X tal que $K_n \subseteq K_{n+1}^\circ$ para todo $n \in \mathbb{N}_0$ y $X \subseteq \bigcup_{n=0}^\infty U_n \subseteq \bigcup_{n=0}^\infty K_n$. Vemos así que X admite una exhaución por compactos. □

No todo espacio admite exhauciones por compactos. Para verlo es suficiente con considerar espacios métricos no localmente compactos, como $\mathbb{Q}$ con su métric usual o espacios normados de dimensión infinita.

La razón por la que todo esto nos interesa es que hay una forma natural de metrizar los conjuntos de funciones continuas sobre un espacio métrico que admite una exhaución por compactos.

Proposición 9.7.4. *Sea X un espacio métrico que admite una exhaución por compactos $(K_n)_{n\geq 0}$, sea Y un espacio métrico, y sea $(a_n)_{n\geq 0}$ una sucesión de números reales positivos tal que la serie $\sum_{n=0}^\infty a_n$ converge.*

(*i*) *Si $f, g : X \to Y$ son dos funciones continuas y n es un elemento de $\mathbb{N}_0$, entonces el conjunto $\{d(f(x), g(x)) : x \in K_n\}$ es acotado y podemos considerar el número*

$$d_n(f, g) \coloneqq \sup\{d(f(x), g(x)) : x \in K_n\}.$$

(*ii*) *Para cada $n \in \mathbb{N}_0$ la función $d_n : C(X, Y) \times C(X, Y) \to \mathbb{R}$ que obtenemos de esta forma es una semi-métrica.*

(iii) *La familia de semi-métricas* $\{d_n : n \in \mathbb{N}_0\}$ *separa puntos en* $C(X,Y)$ *y la función*

$$d : (f,g) \in C(X,Y) \times C(X,Y) \mapsto \sum_{n=0}^{\infty} a_n \frac{d_n(f,g)}{1+d_n(f,g)} \in \mathbb{R}$$

es una métrica sobre el conjunto $C(X,Y)$.

(iv) *Una sucesión* $(f_n)_{n\geq 0}$ *en* $C(X,Y)$ *converge con respecto a la métrica* d *si y solamente si*

> *para cada subconjunto compacto* K *de* X *la sucesión* $(f_n|_K)_{n\geq 0}$ *converge en* $C(K,Y)$ *con respecto a la métrica de la convergencia uniforme.*

(v) *Si el espacio* Y *es completo, entonces el espacio métrico* $C(X,Y)$ *que obtenemos de esta forma es completo.*

(vi) *La topología determinada por la métrica* d *y el conjunto* $\mathscr{C}$ *de las sucesiones en* $C(X,Y)$ *que satisfacen la condición de Cauchy con respecto a* d *dependen solamente de* X *y no de la elección de la exhaución por compactos* $(K_n)_{n\geq 0}$ *o de la sucesión* $(a_n)_{n\geq 0}$ *usadas para construir la métrica* d.

Demostración. HACER. □

9.8. El teorema de Ascoli–Arzelà

Muchas veces es conveniente trabajar no con conjuntos compactos sino con conjuntos relativamente compactos, en el siguiente sentido.

Definición 9.8.1. Un subconjunto de un espacio métrico es ***relativamente compacto*** si tiene clausura compacta.

Un conjunto Y relativamente compacto de un espacio métrico X es totalmente acotado. En efecto, Y es un subconjunto de $\overline{Y}$, que es totalmente acotado porque es compacto, así que es él mismo totalmente acotado. Si X es completo, entonces vale la afirmación recíproca.

Ejercicio 9.8.2. Pruebe que los subconjuntos relativamente compactos de un espacio métrico son exactamente los subconjuntos totalmente acotados.

Notemos que un subconjunto de $\mathbb{R}^n$ es relativamente compacto si y solamente si es acotado.

Lema 9.8.3. *Un subconjunto* Y *de un espacio métrico es relativamente compacto si y solamente si toda sucesión en* Y *posee una subsucesión que converge en* X.

Demostración. Sea X un espacio métrico y sea Y un subconjunto de Y. Si Y es relativamente compacto y $(y_n)_{n\geq 0}$ es una sucesión con valores en Y, entonces la sucesión toma valores en el conjunto compacto $\overline{Y}$ de X y, por lo tanto, posee una subsucesión convergente en X. Esto nos dice que la condición del lema es necesaria.

Para probar que también es suficiente, supongamos que se cumple y sea $(x_n)_{n\geq 0}$ una sucesión en $\overline{Y}$. Para cada $n \in \mathbb{N}_0$ hay un elemento y_n de Y tal que $d(x_n, y_n) < 1/(n+1)$, y la hipótesis implica que hay una función estrictamente creciente $h : \mathbb{N}_0 \to \mathbb{N}_0$ tal que la sucesión $(y_{h(n)})_{n\geq 0}$ converge en X a un punto ξ. Como para todo $n \in \mathbb{N}_0$ es

$$d(x_{h(n)}, \xi) \leq d(x_{h(n)}, y_{h(n)}) + d(y_{h(n)}, \xi) < \frac{1}{h(n)+1} + d(y_{h(n)}, \xi) < \frac{1}{n+1} + d(y_{h(n)}, \xi)$$

así que claramente la sucesión $(x_{h(n)})_{n\geq 0}$ converge a ξ. Esto prueba que el conjunto $\overline{Y}$ es compacto y, por lo tanto, que el conjunto Y es relativamente compacto. □

Queremos caracterizar, bajo condiciones apropiadas, los subconjuntos relativamente compactos de un espacio de funciones $C(X, Y)$. Para ello necesitamos los conceptos introducidos en la siguiente definición.

Definición 9.8.4. Sean X e Y dos espacios métricos, sea $C(X, Y)$ el espacio métrico de todas las funciones continuas acotadas $X \to Y$ dotado de la métrica de la convergencia uniforme, y sea $\mathscr{F}$ un subconjunto de $C(X, Y)$. Para cada elemento x de X escribimos

$$\mathscr{F}(x) \coloneqq \{f(x) : f \in \mathscr{F}\},$$

y decimos que $\mathscr{F}$ es ***equicontinuo*** si para punto x de X y cada número positivo ε hay un número positivo δ tal que para todo elemento f de $\mathscr{F}$ y todo punto y de X vale

$$d(x, y) < \delta \implies d(f(x), f(y)) < \varepsilon,$$

y que es ***uniformemente equicontinuo*** si cada número positivo ε hay un número positivo δ tal que para todo elemento f de $\mathscr{F}$ y todo par de puntos x y y de X vale

$$d(x, y) < \delta \implies d(f(x), f(y)) < \varepsilon,$$

El resultado que buscamos es el siguiente *Teorema de Arzelà–Ascoli*, por Cesare Arzelà [Arz1885] and Giulio Ascoli [Asc1884].

Proposición 9.8.5. *Sean X e Y dos espacios métricos, sea $\mathscr{F}$ un subconjunto de $C(X, Y)$, y supongamos que X es compacto y que Y es completo. Las siguientes afirmaciones son equivalentes:*

(*a*) *$\mathscr{F}$ es relativamente compacto en $C(X, Y)$.*

(*b*) *$\mathscr{F}$ es equicontinuo y para todo punto x de X el conjunto $\mathscr{F}(x)$ es relativamente compacto*

en Y.

(*c*) *$\mathscr{F}$ es uniformemente equicontinuo y para todo punto x de X el conjunto $\mathscr{F}(x)$ es relativamente compacto en Y.*

Demostración. Supongamos que el conjunto $\mathscr{F}$ es relativamente compacto en $C(X, Y)$, y sean x un elemento de X y ε un número positivo. Para cada elemento f de $\mathscr{F}$ hay un número positivo δ_f tal que $f(B_{\delta_f}(x)) \subseteq B_{\varepsilon/3}(f(x))$. El conjunto $\{B_{\varepsilon/3}(f) : f \in \mathscr{F}\}$ es un cubrimiento abierto de $\mathscr{F}$, así que hay un subconjunto finito F de $\mathscr{F}$ tal que $\mathscr{F} \subseteq \bigcup_{f\in F} B_{\varepsilon/3}(f)$. Sea $\delta \coloneqq \min\{\delta_f : f \in F\}$, que es un número positivo. Sea y un elemento de $B_\delta(x)$ y g uno de $\mathscr{F}$. Por la forma en que elegimos F hay un elemento f de F tal que $d(f, g) < \varepsilon/3$ y, como $\delta \leq \delta_f$, tenemos entonces que para todo elemento x' de $B_\delta(x)$ es

$$d(g(x), g(x')) \leq d(g(x), f(x)) + d(f(x), f(x')) + d(f(x'), g(x')) < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Esto nos dice que el conjunto $\mathscr{F}$ es equicontinuo. Por otro lado, si x es un elemento de X, entonces la función $e_x : f \in C(X, Y) \mapsto f(x) \in Y$ es continua, así que $e_x(\overline{F})$ es compacto en Y y, como $\mathscr{F}(x) = e_x(\mathscr{F}) \subseteq e_x(\overline{F})$, la clausura de $\mathscr{F}(x)$ es compacta. Esto prueba la implicación $(a) \Rightarrow (b)$.

Supongamos ahora que el conjunto $\mathscr{F}$ es equicontinuo y que para todo elemento x de X el conjunto $\mathscr{F}(x)$ es relativamente compacto, y mostremos que $\mathscr{F}$ es uniformemente equicontinuo. Esto probará la implicación $(b) \Rightarrow (c)$. Sea ε un número positivo. Como el conjunto $\mathscr{F}$ es equicontinuo, para cada punto x de X hay un número positivo δ_x tal que $f(B_{\delta_x}(x)) \subseteq B_{\varepsilon/2}(f(x))$ para todo elemento f de $\mathscr{F}$. El conjunto $\{B_{\delta_x/2}(x) : x \in X\}$ es un cubrimiento abierto del espacio X, que es compacto, así que hay un subconjunto finito F de X tal que $X = \bigcup_{x\in F} B_{\delta_x/2}(x)$. Pongamos $\delta \coloneqq \frac{1}{2}\min\{\delta_x : x \in F\}$, que es un número positivo. Sean x' y x'' dos elementos de X tales que $d(x', x'') < \delta$ y sea f un elemento de $\mathscr{F}$. La forma en que elegimos F implica que hay un elemento x de F tal que $d(x, x') < \delta_x/2$, así que $d(x, x'') \leq d(x, x') + d(x', x'') < \delta_x/2 + \delta < \delta_x$ y, por lo tanto,

$$d(f(x'), f(x'')) \leq d(f(x'), f(x)) + d(f(x), f(x'') < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

ya que tanto x' como x'' están en $B_{\delta_x}(x)$. Esto prueba que el conjunto $\mathscr{F}$ es uniformemente equicontinuo, como queremos.

Para terminar, probemos la implicación $(c) \Rightarrow (a)$. Supongamos que el conjunto $\mathscr{F}$ es uniformemente equicontinuo y que para cada elemento x de X el conjunto $\mathscr{F}(x)$ es relativamente compacto en Y, y mostremos que $\mathscr{F}$ es relativamente compacto en $C(X, Y)$. Como Y es completo, el espacio $C(X, Y)$ es completo, así que la clausura $\overline{\mathscr{F}}$ completa. Para ver que $\overline{\mathscr{F}}$ es un subconjunto compacto de $C(X, Y)$, entonces, es suficiente que probemos que es totalmente acotada y, de acuerdo a la Proposición 9.1.11, para ello, a su vez, alcanza con que probemos que $\mathscr{F}$ es totalmente acotado. Eso haremos.

Sea r un número positivo. Como $\mathscr{F}$ es uniformemente equicontinuo hay un número positivo δ tal que $f(B_\delta(x)) \subseteq B_{r/5}(f(x))$ para toda elección de x en X y de f en $\mathscr{F}$. Como X es compacto y $X = \bigcup_{x\in X} B_\delta(x)$, hay un subconjunto finito F de X tal que

$$X = \bigcup_{x\in F} B_\delta(x). \tag{9.7}$$

Para cada $x \in F$ el conjunto $\mathscr{F}(x)$ es relativamente compacto en Y, así que la unión $\bigcup_{x\in F} \mathscr{F}(x)$ también lo es y, por lo tanto, esta unión es totalmente acotada: hay entonces un subconjunto finito G de Y tal que

$$\bigcup_{x\in F} \mathscr{F}(x) \subseteq \bigcup_{y\in G} B_{r/5}(y).$$

Sea Π el conjunto de todas las funciones $F \to G$, que es ciertamente finito, y para cada elemento $\pi : F \to G$ de Π consideremos el conjunto

$$\mathscr{F}_\pi \coloneqq \{f \in \mathscr{F} : \text{para cada } x \in F \text{ es } d(f(x), \pi(x)) < r/5\}.$$

Queremos probar que

$$\mathscr{F} = \bigcup_{\pi\in\Pi} \mathscr{F}_\pi. \tag{9.8}$$

Sea f un elemento de $\mathscr{F}$. Si z es uno de los elementos de F, entonces

$$f(z) \in \mathscr{F}(z) \subseteq \bigcup_{x\in F} \mathscr{F}(x) \subseteq \bigcup_{y\in G} B_{r/5}(y)$$

y, por lo tanto, el subconjunto $\tilde{\pi}(x) \coloneqq \{y \in G : d(f(x), y) < r/5\}$ de G no es vacío. Se sigue de esto que hay una función $\pi : F \to G$ tal que $\pi(x) \in \tilde{\pi}(x)$ para todo $x \in F$ y, por lo tanto, tal que $f \in \mathscr{F}_\pi$. Esto prueba la igualdad (9.8).

Sea ahora π un elemento de Π y sean f y g dos elementos de $\mathscr{F}_\pi$. Si x es un elemento cualquiera de X, entonces sabemos de (9.7) que hay un elemento x_0 de F tal que $d(x, x_0) < \delta$, y entonces

$$\begin{aligned} d(f(x), g(x)) &\le d(f(x), f(x_0)) + d(f(x_0), \pi(x_0)) + d(\pi(x_0), g(x_0)) + d(g(x_0), g(x) \\ &< \frac{r}{5} + \frac{r}{5} + \frac{r}{5} + \frac{r}{5} = \frac{4}{5}r. \end{aligned}$$

Vemos así que $\operatorname{diam} \mathscr{F}_\pi \le 4r/5$ y, por lo tanto, que hay un elemento f_π de $C(X, Y)$ tal que $\mathscr{F}_\pi \subseteq B_r(f_\pi)$. Podemos concluir de todo esto que $\mathscr{F} \subseteq \bigcup_{\pi\in\Pi} B_r(f_\pi)$ y, en definitiva, que el conjunto $\mathscr{F}$ es totalmente acotado en $C(X, Y)$, como queremos. □

9.8.1. Algunas aplicaciones

Ejemplo 9.8.1.1. Sea $\mathscr{F}$ un subconjunto de $C([0,1],\mathbb{R})$ tal que

- hay un elemento t_0 de $[0,1]$ y un número positivo L tal que $|f(t_0)| \leq L$ para todo elemento f de $\mathscr{F}$, y
- todos los elementos de $\mathscr{F}$ son derivables en $(0,1)$, y hay un número positivo K tal que $|f'(t)| \leq K$ para todo $t \in (0,1)$ y todo elemento f de $\mathscr{F}$.

Veamos que el conjunto $\mathscr{F}$ es relativamente compacto en $C([0,1],\mathbb{R})$ usando el *Teorema de Arzelà–Ascoli*.

- Sea t un elemento de $[0,1]$. Para cada elemento f de $\mathscr{F}$ hay un número ξ_f entre t_0 y t tal que $f(t) - f(t_0) = f'(\xi_f)\cdot(t-t_0)$ y entonces

  $$|f(t)| \leq |f(t_0)| + |f'(\xi_f)|\cdot|t-t_0| \leq K + L.$$

  Esto nos dice que el conjunto $\mathscr{F}(t)$ es acotado en $\mathbb{R}$ y, por lo tanto, relativamente compacto.
- Sea ε un número positivo y pongamos $\delta \coloneqq \varepsilon/K$. Si s y t son dos elementos de $[0,1]$ tales que $|s-t| < \delta$ y f es un elemento de $\mathscr{F}$, entonces hay un número ξ entre s y t tal que

  $$|f(s) - f(t)| = |f'(\xi)|\cdot|s-t| \leq K\cdot\frac{\varepsilon}{K} = \varepsilon.$$

  Vemos así que el conjunto $\mathscr{F}$ es uniformemente equicontinuo.

La Proposición 9.8.5 nos permite concluir de esto que el conjunto $\mathscr{F}$ es relativamente compacto en $C([0,1],\mathbb{R})$ y, en particular, que toda sucesión con valores en $\mathscr{F}$ posee una subsucesión convergente en $C([0,1],\mathbb{R})$.

En este ejemplo suponemos que los elementos del conjunto $\mathscr{F}$ son derivables, pero podemos llegar a una conclusión similar suponiendo solamente que satisfacen la condición de Lipschitz.

Ejercicio 9.8.1.2. Sea $[a,b]$ un intervalo de $\mathbb{R}$ y sea $\mathscr{F}$ un conjunto de funciones $[a,b] \to \mathbb{R}^n$. Muestre que si hay una constante positiva K tal que

$$\|f(t) - f(t')\| \leq K\cdot|t-t'|$$

siempre que $f \in \mathscr{F}$ y $t, t' \in [a,b]$, entonces el conjunto $\mathscr{F}$ está contenido en $C([a,b],\mathbb{R}^n)$ y es allí relativamente compacto.

En los siguientes ejemplos nos ocuparemos de funciones que satisfacen la siguiente condición:

Definición 9.8.1.3. Sean X e Y dos espacios métricos. Una función $f : X \to Y$ es ***compacta*** si para cada subconjunto acotado B de X el conjunto $f(B)$ es relativamente compacto en Y.

Normalmente solo nos ocupamos de funciones compactas cuando hablamos de funciones lineales entre espacios normados.

Lema 9.8.1.4. *Una función linear entre espacios normados que es compacta es continua.*

Demostración. Sean X e Y dos espacios normados y sea $f : X \to Y$ una función lineal compacta. Si B es la bola unidad de X, entonces por supuesto B es un subconjunto acotado de X, así que la hipótesis implica que $f(B)$ es un subconjunto relativamente compacto de‘Y y, por lo tanto, acotado. Como sabemos, esto nos permite concluir que la función f es continua. □

En general, las funciones lineales continuas entre espacios normados no son compactas. Por ejemplo, si X es un espacio normado de dimensión infinita, entonces la función identidad $\mathrm{id}_X : X \to X$ es continua pero no compacta — recordemos que la bola unidad cerrada de X no es compacta.

Ejercicio 9.8.1.5.

(a) Sean X, Y, y Z tres espacios normados. Pruebe que la composición $g \circ f$ de dos funciones lineales continuas $f : X \to Y$ y $g : Y \to Z$ es compacta si alguna de las dos lo es.

(b) Muestre que el subconjunto $K(X)$ de $L(X)$ de todas las funciones lineales compactas $X \to X$ es un ideal propio de $L(X, X)$.

El *Teorema de Arzelà–Ascoli* es particularmente útil para mostrar que una función lineal es compacta. Veamos un ejemplo importante de esto, el ***operador de Volterra***.

Ejemplo 9.8.1.6. Sea $f : [0,1] \to \mathbb{R}$ una función continua. Para cada $t \in \mathbb{R}$ tiene sentido la integral

$$\int_0^t f(s)\,\mathrm{d}s,$$

y, por lo tanto, hay una función

$$T(f) : t \in [0,1] \mapsto \int_0^t f(s)\,\mathrm{d}s \in \mathbb{R}.$$

Esta función es continua. Esto es claro si la función f es idénticamente nula, así que es suficiente que consideremos el caso en el que $\|f\|_\infty > 0$. Sea ε un número positivo y sea $\delta \coloneqq \varepsilon/\|f\|_\infty$. Si t y t' son dos elementos de $[0,1]$ tales que $|t - t'| < \delta$ y es $t' < t$, entonces

$$\left|T(f)(t) - T(f)(t')\right| = \left|\int_0^t f(s)\,\mathrm{d}s - \int_0^{t'} f(s)\,\mathrm{d}s\right|$$

$$\leq \int_{t'}^{t} |f(s)|\,\mathrm{d}s \leq \|f\|_\infty \cdot |t - t'| < \|f\|_\infty \cdot \frac{\varepsilon}{\|f\|_\infty} = \varepsilon,$$

y la misma desigualdad vale si $t' > t$. Vemos así tenemos una función

$$T : f \in C([0,1],\mathbb{R}) \to T(f) \in C([0,1],\mathbb{R}).$$

Es fácil verificar que se trata de una función lineal y es, de hecho, continua. En efecto, si f es un elemento de $C([0,1],\mathbb{R})$, entonces $\|T(f)\|_\infty \leq \|f\|_\infty$, ya que para todo elemento t de $[0,1]$ es

$$|T(f)(t)| = \left|\int_0^t f(s)\,\mathrm{d}s\right| \leq \int_0^t |f(s)|\,\mathrm{d}s \leq t \cdot \|f\|_\infty \leq \|f\|_\infty.$$

Mostremos que la función T es compacta. Sea B un subconjunto no vacío y acotado de $C([0,1],\mathbb{R})$, de manera que hay un número positivo R tal que $B \subseteq B_R(0)$.

- Para cada elemento f de B es $T(f)(0) = 0$.
- Sea f un elemento de B. El *Teorema fundamental del cálculo* nos dice que $T(f)$ es derivable en $(0,1)$ y que para cada $t \in (0,1)$ es $|T(f)'(t)| = |f(t)| \leq \|f\|_\infty \leq R$.

Vemos con esto que estamos en las condiciones del Ejemplo 9.8.1.1, sí que el conjunto B es relativamente compacto en $C([0,1],\mathbb{R})$.

Ejemplo 9.8.1.7. Sean X e Y dos espacios de Banach, sea $T : X \to Y$ una función lineal y continua, y sea $T^* : Y^* \to X^*$ la función adjunta a T. Queremos probar que

si la función T es compacta, entonces también lo es T^*.

Supongamos que T es compacta. Sea B la bola unidad de X, de manera que $T(B)$ es un subconjunto relativamente compacto de Y y $K \coloneqq \overline{T(B)}$ es compacto. Sea B' la bola unidad de Y'.

Para cada elemento $f : Y \to \Bbbk$ de B' podemos considerar la restricción $f|_K : K \to \Bbbk$, que es una función continua. Tenemos entonces un subconjunto $\mathscr{F} \coloneqq \{f|_K : f \in B'\}$ de $C(K,\Bbbk)$. Veamos que $\mathscr{F}$ es uniformemente acotado y equicontinuo.

- HACER.

El *Teorema de Arzelà–Ascoli* 9.8.5 nos dice entonces que $\mathscr{F}$ es un subconjunto relativamente compacto de $C(K,\Bbbk)$. Sea ahora $(\phi_n)_{n\geq 0}$ una sucesión convalores en $T^*(B')$, de manera que hay una sucesión $(\psi_n)_{n\geq 0}$ en B' tal que $\phi_n = T^*(\psi_n)$ para cada $n \in \mathbb{N}_0$. Como $\mathscr{F}$ es relativamente compacto, la sucesión $(\psi_n|_K)_{n\geq 0}$ posee una subsucesión que converge en $C(K,\Bbbk)$, esto es, hay una función estrictamente creciente $h : \mathbb{N}_0 \to \mathbb{N}_0$ tal que $(\psi_{h(n)}|_K)_{n\geq 0}$ converge en $C(K,\Bbbk)$. Como la función $T : X \to Y$ es continua, esto implica que la sucesión $((\psi_{h(n)} \circ T)|_B)_{n\geq 0}$ converge uniformemente en $C(B,\Bbbk)$ y esta sucesión no es otra que $(\phi_{h(n)}|_B)_{n\geq 0} = (T^*(\phi_{h(n)})|_B)_{n\geq 0}$. Finalmente, sabemos que que la sucesión $(\phi_{h(n)}|_B)_{n\geq 0}$ converja uniformemente implica que la

sucesión $(\phi_{h(n)}|_B)_{n\geq 0}$ converge en X'. Hemos probado que toda sucesión en $T^*(B')$ posee una subsucesión que converge en X' y esto implica que $T^*(B')$ es relativamente compacto.

9.8.2. El teorema de Cauchy–Peano

Queremos presentar una aplicación importante y considerablemente más sustancial que las anteriores del *Teorema de Arzelà–Ascoli* a la teoría de las ecuaciones diferenciales. Antes de llegar a ella, sin embargo, tenemos que hacer algunas preparaciones.

Primero tenemos que generalizar la noción de diferenciabilidad. Decimos que una función $x : (a, b) \to \mathbb{R}^n$ definida en un intervalo abierto (a, b) es ***continua a trozos*** o de ***clase*** C^0_p si hay una sucesión finita y estrictamente creciente $(t_1, \ldots, t_k)$ de elementos de (a, b) tal que

- la función x es continua en cada uno de los intervalos $(a, t_1), (t_1, t_2), \ldots, (t_{k-1}, t_k)$, y (t_k, b),
- para cada $i \in [\![k]\!]$ los límites $\lim_{t\nearrow t_i} x(t)$ y $\lim_{t\searrow t_i} x(t)$ existe.

De manera similar, una función $x : [a, b] \to \mathbb{R}^n$ definida ahora en un intervalo cerrado es ***continua a trozos*** o de clase C^0_p si su restricción al intervalo abierto (a, b) es continua a trozos y además los límites $\lim_{t\searrow a} x(t)$ y $\lim_{t\nearrow b} x(t)$ existen.

Sobre este tipo de funciones necesitamos saber lo siguiente:

Lema 9.8.2.1. *Una función $x : [a, b] \to \mathbb{R}^n$ que es continua a trozos es integrable.* □

Todas las integrales de las que hablamos en esta sección son en el sentido de Riemann.

Decimos que una función $x : (a, b) \to \mathbb{R}^n$ ***es derivable con continuidad a trozos*** o que tiene ***clase*** C^1_p si es continua en (a, b) y hay una sucesión finita y estrictamente creciente $(t_1, t_2, \ldots, t_k)$ de elementos de (a, b) tal que las siguientes condiciones se satisfacen:

- La función x tiene derivada en cada uno de los intervalos (a, t_1), (t_1, t_2), …, (t_{k-1}, t_k), y (t_k, b), y esa derivada es allí continua.
- Para cada $i \in [\![k]\!]$ los límites laterales $\lim_{t\nearrow t_i} x'(t)$ y $\lim_{t\searrow t_i} x'(t)$ existen.

Cuando ese es el caso escribimos $x' : [a, b] \to \mathbb{R}^n$ a la función que en cada uno de los intervalos abiertos $(a, t_1), (t_1, t_2), \ldots, (t_{k-1}, t_k)$ y (t_k, b) coincide con la derivada de x, y que tiene $x'(t_i) = 0$ para cada $i \in [\![k]\!]$. Esta función es continua a trozos.

Por otro lado, una función $x : [a, b] \to \mathbb{R}^n$ definida en un intervalo cerrado ***es derivable con continuidad a trozos*** o es de ***clase*** C^1_p si su restricción al intervalo abierto (a, b) es derivable con continuidad a trozos y los límites $\lim_{t\searrow a} x'(t)$ y $\lim_{t\searrow b} x'(t)$ existen. Como antes, en esta situación escribimos $x' : [a, b] \to \mathbb{R}^n$ a la función que coincide con la derivada de x en cada uno de los intervalos$(a, t_1), (t_1, t_2), \ldots, (t_{k-1}, t_k)$ y (t_k, b), y que tiene $x'(a) = 0$, $x'(t_i) = 0$ para cada $i \in [\![k]\!]$, y $x'(b) = 0$. Esta función x' es continua a trozos sobre $[a, b]$.

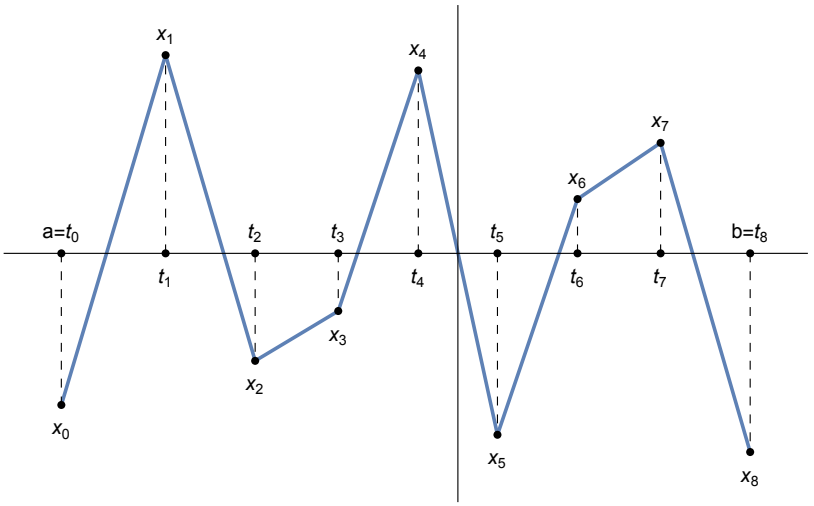


Figura 9.1. Una función lineal a trozos.

Ejemplo 9.8.2.2. Sea $[a, b]$ un intervalo cerrado, sea $(t_0, t_1, \ldots, t_k)$ una secuencia finita y estrictamente creciente de elementos de $[a, b]$ tal que $t_0 = a$ y $t_k = b$, y sea $(x_0, x_1, \ldots, x_k)$ una secuencia arbitraria de números reales de la misma longitud. La función $x : [a, b] \to \mathbb{R}$ que tiene $x(a) = x_0$ y

$$x(t) = \frac{(x_i - x_{i-1})t + (t_i x_{i-1} - t_{i-1} x_i)}{t_i - t_{i-1}}$$

siempre que $i \in [\![k]\!]$ y $t \in (t_{i-1}, t_i]$ es derivable con continuidad a trozos en el intervalo $[a, b]$. Como esta función es lineal en cada uno de los intervalos $[t_{i-1}, t_i]$ con $i \in [\![k]\!]$ decimos que es ***linear a trozos***. Dibujamos un ejemplo de este tipo de función en la Figura 9.1 de la página 376.

Será importante para nosotros saber que el *Teorema fundamental del cálculo* es cierto para funciones de clase C^1_p.

Lema 9.8.2.3. *Si $x : [a, b] \to \mathbb{R}^n$ es una función de clase C^1_p, entonces*

$$\int_a^b x'(\tau)\, \mathrm{d}\tau = x(b) - x(a). \qquad \square$$

Our next step is to generalize the notion of solution of a differential equation. Let Ω be a non-empty subset of $\mathbb{R} \times \mathbb{R}^n$ and let $f : \Omega \to \mathbb{R}$ be a function. If ε is a positive number, then we say that a function $x : (a, b) \to \mathbb{R}$ is an ***ε-approximate solution*** to the equation

$$x' = f(t, x)$$

if it is of class C^1_p, for all $t \in (a,b)$ we have that $(t, x(t)) \in \Omega$, and for each $t \in (a,b)$ such that x is differentiable at t we have that

$$|x'(t) - f(t, x(t))| < \varepsilon.$$

Lema 9.8.2.4. *Let Ω be a non-empty subset of $\mathbb{R} \times \mathbb{R}^n$, let $f : \Omega \to \mathbb{R}$ be a function, and let ε be a positive number. if $x : (a,b) \to \mathbb{R}^n$ is an ε-approximate solution to the equation*

$$x' = f(t,x)$$

and t_0 is an element of $a,b)$, then for all $t \in (a,b)$ we have

$$\left| x(t) - x(t_0) - \int_{t_0} f(\tau, x(\tau))\, \mathrm{d}\tau \right| \leq |t - t_0| \cdot \varepsilon$$

Demostración. Let $x : (a,b) \to \mathbb{R}^n$ be an ε-approximate solution to the equation in the lemma, so that for all $t \in (a,b)$ we have $|x'(t) - f(t,x(t))| < \varepsilon$. As x if of class C^1_p, using the *Fundamental Theorem of Calculus* we see immediately that for all $t \in (a,b)$ we have

$$\begin{aligned} \left| x(t) - x(t_0) - \int_{t_0}^{t} f(\tau, x(\tau))\, \mathrm{d}\tau \right| &= \left| \int_{t_0}^{t} x'(\tau), \mathrm{d}\tau - \int_{t_0} f(\tau, x(\tau))\, \mathrm{d}\tau \right| \\ &\leq \int_{t_0}^{t} \left| x'(\tau) - f(\tau, x(\tau)) \right| \mathrm{d}\tau \leq \varepsilon \cdot |t - t_0|, \end{aligned}$$

and this is what the lemma claims. □

The most important fact about approximate solutions to differential equations is that they exist under very mild hypothesis:

Proposición 9.8.2.5. *Let (t_0, x_0) be a point in $\mathbb{R} \times \mathbb{R}^n$, let r and s be two positive numbers, and let*

$$\Omega \coloneqq \{(t,x) \in \mathbb{R} \times \mathbb{R}^n : |t - t_0| < r, |x - x_0| < s\}.$$

Let $f : \overline{\Omega} \to \mathbb{R}^n$ be a continuous function, and let

$$M \coloneqq \sup\{|f(t,x)| : (t,x) \in \overline{\Omega}\}, \qquad \rho \coloneqq \min\{r, s/2M\}.$$

If ε is a positive number, then there exists a function $x : [t_0 - \rho, t_0 + \rho] \to \mathbb{R}$ such that

- *x is of class C^1_p,*
- *$x(t_0) = x_0$,*
- *$|x(s) - x(t)| \leq M \cdot |s - t|$ whenever s and t are in $[t_0 - \rho, t_0 + \rho]$, and*
- *for all $t \in [t_0 - \rho, t_0 + \rho]$ such that x is differentiable at t the point $(t, x(t))$ is in the domain of f and $|x'(t) - f(t, x(t))| < \varepsilon$.*

In particular, x is an ε-approximate solution to the differential equation

$$x' = f(t, x)$$

such that $x(t_0) = x_0$.

Demostración. Let ε be a positive number. Since the function $f : \overline{\Omega} \to \mathbb{R}^n$ is continuous and $\overline{\Omega}$ is compact, the function f is in fact uniformly continuous on $\overline{\Omega}$. There exists then a positive number δ such that whenever (t_1, x_1) and (t_2, x_2) are points in $\overline{\Omega}$ we have that

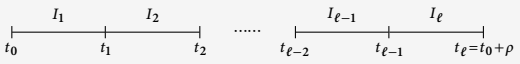
$$|t_1 - t_2| < \delta, |x_1 - x_2| < \delta \implies |f(t_1, x_1) - f(t_2, x_2)| < \varepsilon.$$

Let ℓ be a positive integer such that

$$\frac{\rho}{\ell} < \min\left\{\delta, \frac{\delta}{M}\right\}.$$

We will organize our construction in several steps.

First step. For each $k \in [\![0, \ell]\!]$ we put $t_k := t_0 + \frac{\rho k}{\ell}$, and for each $k \in [\![\ell]\!]$ we consider the interval $I_k := [t_{k-1}, t_k]$. The intervals $I_1, \ldots, I_\ell$ that we obtain in this way cover the interval $[t_0, t_0 + \rho]$:



We will construct a sequence $\xi_0, \ldots, \xi_\ell$ of $\ell + 1$ points of $\mathbb{R}^n$ recursively.

- We start by putting $\xi_0 := x_0$.
- Next, we suppose that k is an element of $[\![0, \ell - 1]\!]$ and that we have already constructed the point ξ_k in such a way that $|\xi_k - x_0| \leq M\rho k/\ell$. As $M\rho k/\ell \leq s$, we can put

  $$\xi_{k+1} := \xi_k + \frac{\rho}{\ell} \cdot f(t_k, \xi_k).$$

  We then have that

  $$|\xi_{k+1} - x_0| \leq |\xi_{k+1} - \xi_k| + |\xi_k - x_0| \leq \frac{\rho}{\ell} \cdot \big|f(t_k, \xi_k)\big| + \frac{M\rho k}{\ell} \leq \frac{M\rho(k+1)}{\ell},$$

  and this allows us to continue the construction.

The points $\xi_0, \ldots, \xi_\ell$ that we obtain in this way are such that $\xi_0 = x_0$ and

$$\xi_k = \xi_{k-1} + \frac{\rho}{\ell} \cdot f(t_{k-1}, \xi_{k-1}) \quad \text{for each } k \in [\![\ell]\!]. \tag{9.9}$$

Let us notice that

$$|\xi_k - \xi_{k'}| \leq \sum_{i=k}^{k'-1} |\xi_i - \xi_{i+1}| = \sum_{i=k}^{k'-1} \frac{\rho}{\ell} \cdot |f(t_i, \xi_i)| \leq (k' - k) \cdot \frac{\rho}{\ell} \cdot M = M \cdot (t_k - t_{k'})$$

whenever k and k' are two elements of $[\![0, \ell]\!]$ such that $k \le k'$.

Second step. Using (9.9) we can see that there is a function $u : [t_0, t_0 + \rho] \to \mathbb{R}^n$ such that

$$u(t) = \xi_{k-1} + \frac{\ell}{\rho}(\xi_k - \xi_{k-1})(t - t_{k-1}) \quad \text{whenever } k \in [\![\ell]\!] \text{ and } t \in I_k.$$

This function is continuous and, since it is linear on each of the intervals $I_1, \dots, I_\ell$, of class C^1_p. It has $u(t_k) = \xi_k$ for each $k \in [\![0, \ell]\!]$ and, in particular, $u(t_0) = x_0$. From its definition we see immediately that

$$u'(t) = f(t_{k-1}, \xi_{k-1}) \qquad \text{whenever } k \in [\![\ell]\!] \text{ and } t \in (t_{k-1}, t_k)$$

and thus, according to the *Mean Value Inequality* **??** and the defintion of M,

$$|u(s) - u(t)| \le M \cdot |s - t| \quad \text{whenever } k \in [\![\ell]\!] \text{ and } s, t \in (t_{k-1}, t_k).$$

Let s and t be two elements of $[t_0, t_0 + \rho]$ such that $s \le t$. There are then two elements k and k' of $[\![\ell]\!]$ such that $k \le k'$, $s \in I_k$ and $t \in I_{k'}$. If $k = k'$, then

$$|u(s) - u(t)| \le M \cdot |s - t|.$$

If instead $k < k'$, then

$$\begin{aligned} |u(s) - u(t)| &\le |u(s) - u(t_k)| + |u(t_k) - u(t_{k'-1})| + |u(t_{k'-1} - u(t)| \\ &\le M \cdot |s - t_k| + |\xi_k - \xi_{k'-1}| + M \cdot |t - t_{k'-1}| \\ &\le M \cdot |s - t_k| + M \cdot |t_{k'-1} - t_k| + M \cdot |t - t_{k'-1}| \\ &= M \cdot |s - t|, \end{aligned}$$

because $s \le t_k \le t_{k'-1} \le t$.

Finally, let t be a point in $[t_0, t_0 + \rho]$ that belongs to the interior of the interval I_k for some $k \in [\![\ell]\!]$. What we have just proved tells us that

$$|u(t) - x_0| = |u(t) - u(t_0)| \le M \cdot |t - t_0| \le M \cdot \rho < \frac{s}{2},$$

and therefore the point $(t, u(t))$ is in the domain of the function f. Moreover, we have that $|t - t_{k-1}| \le \rho/\ell < \delta$ and

$$|u(t) - \xi_{k-1}| = \frac{\ell}{\rho} \cdot |\xi_k - \xi_{k-1}| \cdot |t - t_{k-1}| = |f(t_{k-1}, \xi_{k-1})| \cdot |t - t_{k-1}| \le M \cdot \frac{\rho}{\ell} < \delta,$$

so the way we chose δ implies that

$$|u'(t) - f(t, u(t))| = |f(t_{k-1}, \xi_{k-1}) - f(t, u(t))| < \varepsilon.$$

Third step. We have shown that there is a function $u : [t_0, t_0 + \rho] \to \mathbb{R}^n$ such that

- u is continuous and of class C^1_p,
- $u(t_0) = x_0$,
- $|u(s) - u(t)| \leq M \cdot |s - t|$ whenever s and t are in $[t_0, t_0 + \rho]$, and
- for all $t \in [t_0, t_0 + \rho]$ such that u is differentiable at t the point $(t, u(t))$ is in the domain of f and $|u'(t) - f(t, u(t))| < \varepsilon$.

Exactly in the same way we can construct a function $v : [t_0 - \rho, t_0] \to \mathbb{R}^n$ such that

- v is continuous and of class C^1_p,
- $v(t_0) = x_0$,
- $|v(s) - v(t)| \leq M \cdot |s - t|$ whenever s and t are in $[t_0 - \rho, t_0]$, and
- for all $t \in [t_0 - \rho, t_0]$ such that v is differentiable at t the point $(t, v(t))$ is in the domain of f and $|v'(t) - f(t, v(t))| < \varepsilon$.

As $u(t_0) = v(t_0)$, we can define a function $x : [t_0 - \rho, t_0 + \rho] \to \mathbb{R}$ such that putting, for each t in its domain,

$$x(t) = \begin{cases} v(t) & \text{if } t_0 - \rho \leq t \leq t_0; \\ u(t) & \text{if } t_0 < t \leq t_0 + \rho. \end{cases}$$

It is clear that x is continuous and of class C^1_p, that it has $x(t_0) = x_0$, and that for all $t \in [t_0 - \rho, t_0 + \rho]$ such that x is differentiable at t the point $(t, x(t))$ is in the domain of f and $|x'(t) - f(t, x(t))| < \varepsilon$. Finally, if s and t are two elements of $[t_0 - \rho, t_0 + \rho]$ such that $s < t$, then we have $s < t \leq t_0$, and then

$$|x(s) - x(t)| = |v(s) - v(t)| \leq M \cdot |s - t|,$$

or $t_0 \leq s < t$, and then

$$|x(s) - x(t)| = |u(s) - u(t)| \leq M \cdot |s - t|,$$

or $s \leq t_0 \leq t$, and then

$$|x(s) - x(t)| \leq |v(s) - v(t_0)| + |u(t_0) - u(t)| \leq M \cdot |s - t_0| + M \cdot |t_0 - t| = M \cdot |s - t|.$$

We thus see that the function x satisfies all the conditions described in the statement of the proposition. □

The approximate solution that we constructed in this proof is a piecewise linear function. It is often called an ***Euler polygon***, and the method we used to construct it is the ***Euler method***. In the first step of the proof we constructed it on the interval $[t_0, t_0 + \rho]$, using ℓ points $a = t_0, \ldots, t_\ell = b$ each distant from the next one exactly by ρ/ℓ. We call ℓ the ***number of steps*** of the construction,

and the number ρ/ℓ the ***step size***.

Ejemplo 9.8.2.6. Let us consider the one-dimensional initial value problem

$$\begin{cases} x' = x, \\ x(0) = 1, \end{cases} \tag{9.10}$$

so that we can take $\Omega \coloneqq \mathbb{R} \times \mathbb{R}$ and $f : (t, x) \in \Omega \mapsto x \in \mathbb{R}$. We fix a positive integer N and look for the Euler polygonal for this equation on the interval $[0,1]$ with N steps and, therefore, step size equal to $1/N$. To do that we need to find the numbers $\xi_0, \ldots, \xi_N$, which are determined by the conditions that $\xi_0 = 1$ and

$$\xi_k = \xi_{k-1} + \frac{1}{N} \cdot f\left(\frac{k-1}{N}, \xi_{k-1}\right) = \left(1 + \frac{1}{N}\right) \cdot \xi_{k-1} \qquad \text{for each } k \in [\![N]\!].$$

One can immediately solve this recurrence equation and find that, in fact, we have

$$\xi_k = \left(1 + \frac{1}{N}\right)^k \qquad \text{for each } k \in [\![N]\!].$$

and thus the Euler polygonal we are after is the function $x_N : [0,1] \to \mathbb{R}$ that for all $k \in [\![N]\!]$ and each $t \in [(k-1)/N, k/N]$ has

$$x_N(t) = \left(1 + \frac{1}{N}\right)^{k-1} \left(1 + t - \frac{k-1}{N}\right).$$

This implies that in fact for all $t \in [0,1]$ we have

$$x_N(t) = \left(1 + \frac{1}{N}\right)^{\lfloor Nt+1 \rfloor - 1} \left(1 + t - \frac{\lfloor Nt+1 \rfloor - 1}{N}\right).$$

Let us fix an element t in $[0,1]$. For all $N \in \mathbb{N}$ we have

$$x_N(t) = \left(1 + \frac{1}{N}\right)^{Nt} \frac{1 + \frac{Nt - \lfloor Nt+1 \rfloor + 1}{N}}{\left(1 + \frac{1}{N}\right)^{Nt - \lfloor Nt-1 \rfloor + 1}}.$$

As $1 \leq Nt - \lfloor Nt+1 \rfloor + 1 \leq 2$, it is easy to see that the denominator and the numerator of the fraction appearing in the right hand side here both have limit 1 as N increases, and it follows then that

$$\lim_{N \to +\infty} x_N(t) = e^t.$$

We thus see that as the number of steps N increases the Euler polygonal x_N converges pointwise to the real solution of the initial value problem (9.10). In fact, one can compute explicitly those polygonals and it turns out that in this example they are very close to the real solution even for small values of N. In Figure 9.2 we have drawn the solution and the polygonal x_{10}.

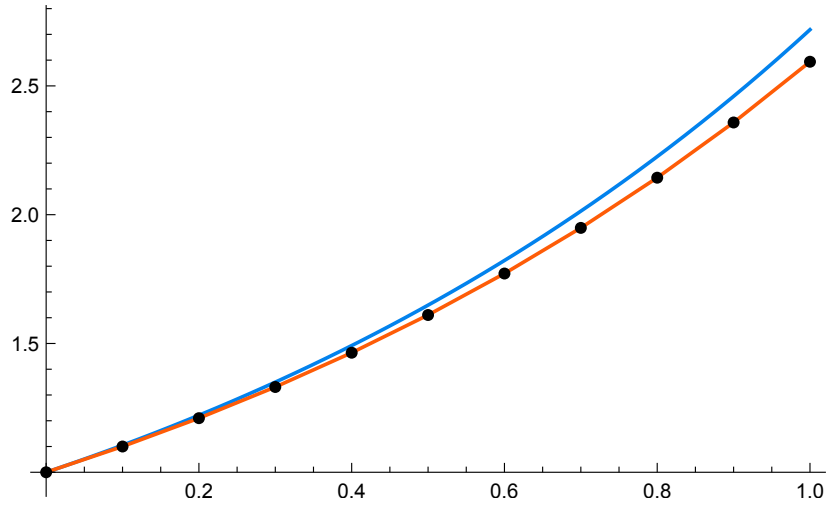


Figura 9.2. The Euler polygon for the equation $x' = x$ with initial condition $x(0) = 1$ on the interval $[0, 1]$ with 10 steps, and the exact solution. The maximal difference between the two curves occurs where $t = 1$ and is less that 0,124

It is certainly not the case that Euler polygonals are good approximations to the real solutions of the initial value problems that give them rise. In Figure 9.3 we have drawn the real solution to the initial value problem

$$\begin{cases} x' = x - 2\sin t, \\ x(0) = 1, \end{cases}$$

and the Euler polygons on $[0, 4]$ computed with a few different numbers of steps. It is clear that in this example the approximation to the real solution given by these polygons becomes better and better as the number of steps used in their construction increases, but that surely when that number is small the polygons differ greatly from the real solution.

Ejemplo 9.8.2.7. Let us consider the two-dimensional initial value problem

$$\begin{cases} x' = -y, \\ y' = x, \\ x(0) = 1, \\ y(0) = 0. \end{cases}$$

We want to construct an approximate solution using the Euler method on the interval $[0, L]$ with ℓ steps. We have to find the points $\xi_k = (x_k, y_k)$ with $i \in [\![0, \ell]\!]$ such that $x_0 = 1$, $y_0 = 0$ and

$$x_k = x_{k-1} - \frac{L}{\ell} y_{k-1}, \qquad y_k = y_{k-1} + \frac{L}{\ell} x_{k-1}$$

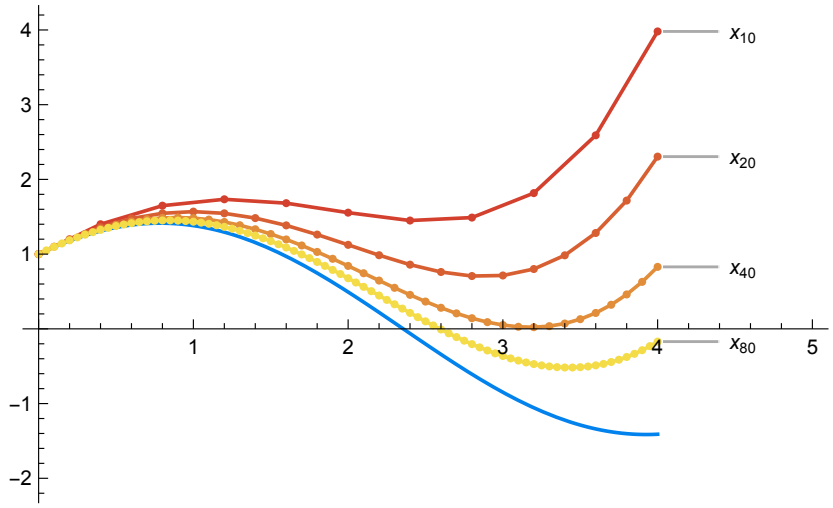


Figura 9.3. The Euler polygons for the equation $x' = x - 2\sin t$ with initial condition $x(0) = 1$ on the interval $[0, 4]$ with 10, 20, 40 and 80 steps, and the exact solution.

for each $k \in [\![\ell]\!]$. If for each $i \in [\![0, \ell]\!]$ we consider the *complex* number $z_k := x_k + iu_k$, then what we have is that $z_0 = 1$ and

$$z_k = z_{k-1} \cdot \left(1 + i\frac{L}{\ell}\right)$$

for each $k \in [\![\ell]\!]$. It is therefore clear that for all $k \in [\![0, k]\!]$ we have

$$x_k + iy_k = \left(1 + i\frac{L}{\ell}\right)^k.$$

If we put $\theta := \arctan L/\ell$, then for all $k \in \mathbb{Z}$ we have

$$\left(1 + i\frac{L}{\ell}\right)^k = \left(1 + \frac{L^2}{\ell^2}\right)^{k/2} (\cos k\theta + i \sin k\theta).$$

Putting everything together, we see that

$$x_k = \left(1 + \frac{L^2}{\ell^2}\right)^{k/2} \cos k\theta \qquad y_k = \left(1 + \frac{L^2}{\ell^2}\right)^{k/2} \sin k\theta$$

The approximate solution that we get in this way is the function $x : [0, L] \to \mathbb{R}^2$ such that

$$x(t) = (x_{k-1}, y_{k-1}) + \frac{\ell}{\rho}(t - t_{k-1})(x_k - x_{k-1}, y_k - y_{k-1})$$

whenever $k \in [\![\ell]\!]$ and $\frac{L}{\ell}(k-1) \le t \le \frac{L}{\ell}k$.

In Figure 9.4 we have plotted this polygonal and the exact solution.

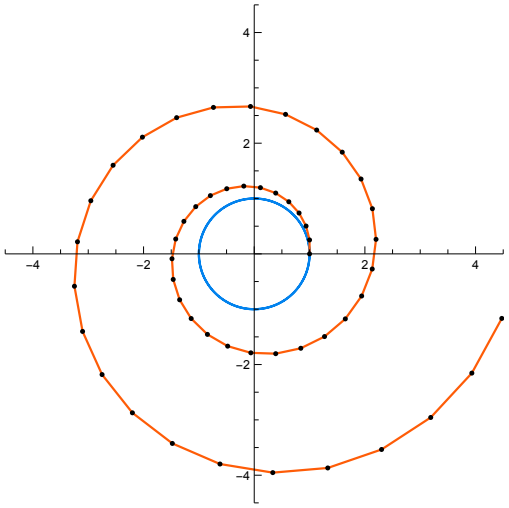


Figura 9.4. The Euler polygon and the exact solution for the intial value problem of Ejemplo 9.8.2.7.

In general, in the situation of Proposición 9.8.2.5 it is *not* true that as the number of steps increases the Euler polygons converge in any sense to a solution of the initial value problem problem — they may not converge at all, in fact! The reader can find an example of this phenomenon in Exercise 12 of the first chapter of the book [CL1955].

We want to prove next, though, that when we are in the hypotheses of the *Picard–Lindelöf Theorem* **??** the Euler polygons do converge to exact solutions and thus provide a way to approximate arbitrarily well the latter. The key fact needed in order to prove this is the following observation that bounds the difference between any two approximate solutions to a differential equation.

Proposición 9.8.2.8. *Let Ω be an open subset of $\mathbb{R} \times \mathbb{R}^n$, let $f : \Omega \to \mathbb{R}^n$ be a continuous function that satisfies the Lipschitz condition with constant K, and let ε_1 and ε_2 be positive numbers. If x_1, $x_2 : (a, b) \to \mathbb{R}^n$ are an ε_1- and an ε_2-approximate solution to the equation*

$$x' = f(t, x), \tag{9.11}$$

respectively, t_0 is an element of (a, b), and δ is a positive number such that $|x_1(t_0) - x_2(t_0)| < \delta$, then for all $t \in (a, b)$ we have that

$$|x_1(t) - x_2(t)| \leq \delta e^{K|t-t_0|} + \frac{\varepsilon_1 + \varepsilon_2}{K}(e^{K|t-t_0|} - 1).$$

Demostración. Let $x_1, x_2 : (a,b) \to \mathbb{R}^n$ be an ε_1- and an ε_2-approximate solution to the equation (9.11), let t_0 is an element of (a,b), and let δ is a positive number such that $|x_1(t_0) - x_2(t_0)| < \delta$. As the function $|x_1 - x_2|$ is continuous on (a,b), we may consider the function $R : (a,b) \to \mathbb{R}$ that on each $t \in (a,b)$ takes the value

$$R(t) = \int_{t_0}^{t} |x_1(\tau) - x_2(\tau)| \, \mathrm{d}\tau.$$

Let t be an element of $[t_0, b)$. As x_1 and x_2 are ε_1- and ε_2-approximate solutions to (9.11), there is a finite subset F of (a,b) such that the functions x_1 and x_2 are differentiable on $(a,b) \setminus F$ and, moreover, for all $\tau \in [t_0, b) \setminus F$ and all $i \in \{1,2\}$ we have

$$\left|x_i'(\tau) - f(\tau, x_i(\tau))\right| < \varepsilon_i.$$

As the functions x_1 and x_2 are of class C_p^1, we can integrate on $[t_0, t]$ and find that for each $i \in \{1,2\}$

$$\begin{aligned}\left|x_i(t) - x_i(t_0) - \int_{t_0}^{t} f(\tau, x_i(\tau)) \, \mathrm{d}\tau\right| &\leq \left|\int_{t_0}^{t} x_i'(\tau) \, \mathrm{d}\tau - \int_{t_0}^{t} f(\tau, x_i(\tau)) \, \mathrm{d}\tau\right| \\ &\leq \int_{t_0}^{t} \left|x_i'(\tau) - f(\tau, x_i(\tau))\right| \mathrm{d}\tau \\ &\leq \varepsilon_i (t - t_0).\end{aligned}$$

It follows from this that

$$\begin{aligned}&\left|\Big(x_1(t) - x_2(t)\Big) - \Big(x_1(t_0) - x_2(t_0)\Big) - \int_{t_0}^{t} \Big(f(\tau, x_1(\tau)) - f(\tau, x_2(\tau))\Big) \, \mathrm{d}\tau\right| \\ &\quad\leq \left|x_1(t) - x_1(t_0) - \int_{t_0}^{t} f(\tau, x_1(\tau)) \, \mathrm{d}\tau\right| + \left|x_2(t) - x_2(t_0) - \int_{t_0}^{t} f(\tau, x_2(\tau)) \, \mathrm{d}\tau\right| \\ &\quad\leq (\varepsilon_1 + \varepsilon_2)(t - t_0),\end{aligned}$$

and thus that

$$\begin{aligned}|x_1(t) - x_2(t)| &\leq |x_1(t_0) - x_2(t_0)| + \int_{t_0}^{t} \left|f(\tau, x_1(\tau)) - f(\tau, x_2(\tau))\right| \mathrm{d}\tau + (\varepsilon_1 + \varepsilon_2)(t - t_0) \\ &\leq \delta + K \int_{t_0}^{t} \left|x_1(\tau) - x_2(\tau)\right| \mathrm{d}\tau + (\varepsilon_1 + \varepsilon_2)(t - t_0) \qquad (9.12)\end{aligned}$$

In terms of the function R that we defined above, this means that

$$-K \cdot R(t) + R'(t) \leq \delta + (\varepsilon_1 + \varepsilon_2)(t - t_0)$$

or, equivalently, that

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(e^{-K(t-t_0)} R(t)\Big) \leq \delta e^{-K(t-t_0)} + (\varepsilon_1 + \varepsilon_2)(t - t_0) e^{-K(t-t_0)}.$$

Integrating now on $[t_0, t]$, and recalling that $R(t_0) = 0$, we see that

$$\begin{aligned} e^{-K(t-t_0)}R(t) &\leq \int_{t_0}^{t} \left(\delta e^{-K(t-t_0)} + (\varepsilon_1 + \varepsilon_2)(t-t_0)e^{-K(t-t_0)}\right) \mathrm{d}\tau \\ &= \frac{\delta}{K}(1 - e^{-K(t-t_0)}) - \frac{\varepsilon_1 + \varepsilon_2}{K^2}(1 + K(t-t_0))e^{-K(t-t_0)} + \frac{\varepsilon_1 + \varepsilon_2}{K^2}. \end{aligned}$$

Multiplying by the exponential this becomes

$$R(t) \leq \frac{\delta}{K}(e^{K(t-t_0)} - 1) - \frac{\varepsilon_1 + \varepsilon_2}{K^2}(1 + K(t-t_0)) + \frac{\varepsilon_1 + \varepsilon_2}{K^2}e^{K(t-t_0)}.$$

Going back to the inequality (9.12) and using this we conclude that

$$|x_1(t) - x_2(t)| \leq \delta + K \cdot R(t) + (\varepsilon_1 + \varepsilon_2)(t-t_0) \leq \delta e^{K(t-t_0)} + \frac{\varepsilon_1 + \varepsilon_2}{K}(e^{K(t-t_0)} - 1),$$

which is the inequality that appears in the statement of the proposition because we have $t \in [t_0, b)$. If instead we have $t \in (a, t_0]$ we can proceed in the same way. □

With this result we can now easily prove the convergence of approximate solutions to the exact solution of an initial value problem when the hypothesis of the *Picard–Lindelöf Theorem* holds:

Proposición 9.8.2.9. *Let Ω be an open subset of $\mathbb{R} \times \mathbb{R}^n$, and let $f : \Omega \to \mathbb{R}^n$ be a continuous function satisfies the Lipschitz condition with constant K. Let (t_0, x_0) be a point in Ω and let r and s be positive numbers such that the set*

$$Q \coloneqq \{(t, x) \in \mathbb{R} \times \mathbb{R}^n : |t - t_0| \leq r, |x - x_0| \leq s\}$$

is contained in Ω. Let $(\varepsilon_k)_{k\geq 1}$ be a sequence of positive numbers that converges to 0, and for each $k \in \mathbb{N}$ let $x_k : [t_0 - r, t_0 + r] \to \mathbb{R}$ be an ε_k-approximate solution to the initial value problem

$$\begin{cases} x' = f(t, x), \\ x(t_0) = x_0. \end{cases}$$

The sequence $(x_k)_{k\geq 1}$ converges uniformly to a function $x : [t_0 - r, t_0 + r] \to \mathbb{R}^n$ whose restriction to $(t_0 - r, t_0 + r)$ is a solution to that initial value problem.

Demostración. Let ε be a positive number. As the sequence $(\varepsilon_k)_{k\geq 1}$ converges to 0, there is a positive integer k_0 such that for all integers k such that $k \geq k_0$ we have

$$\varepsilon_k < \frac{K}{e^{2Kr} - 1} \cdot \frac{\varepsilon}{2}.$$

If now k and k' are two integers such that $k, k' \geq k_0$, then we have $x_k(t_0) = x_{k'}(t_0)$ and it follows

from Proposición 9.8.2.8 that for all $t \in [t_0 - r, t_0 + r]$ we have

$$|x_k(t) - x_{k'}(t)| \leq \frac{\varepsilon_k + \varepsilon_{k'}}{K}(e^{K|t-t_0|} - 1) \leq \frac{e^{2Kr} - 1}{K}(\varepsilon_k + \varepsilon_{k'}) < \varepsilon.$$

We see with this that the sequence $(x_k)_{k\geq 1}$ satisfies the Cauchy condition uniformly on $[t_0-r, t_0+r]$, so that it converges uniformly to a function $x : [t_0-r, t_0+r] \to \mathbb{R}^n$. This limit function is continuous because all the functions in the sequence are continuous. To finish the proof we will show that this function x is a solution to the initial value problem that appears in the statement of the proposition.

Let ε be a positive number and let t be an element of $[t_0 - r, r_0 + r]$. Since the sequence $(x_k)_{k\geq 1}$ converges uniformly to x, there is a positive integer k_0 such that for all integers k such that $k \geq k_0$ and all $\tau \in [t_0 - r, t_0 + r]$ we have $|x_k(\tau) - x(\tau)| < \varepsilon/(3 + 3rK)$, and this implies that for all $\tau \in [t_0 - r, t_0 + r]$

$$|f(\tau, x_k(\tau)) - f(\tau, x(\tau))| < K \cdot |x_k(\tau) - x(\tau)| < K \cdot \frac{\varepsilon}{3 + 3rK} = \frac{\varepsilon}{3r}.$$

Also, as the sequence $(\varepsilon_k)_{k\geq 1}$ converges to 0, there is positive integer k_1 such that $\varepsilon_k < \varepsilon/3r$ for all integers k such that $k \geq k_1$. All this together with Lema 9.8.2.4 inplies that if we put $k \coloneqq \max\{k_1, k_2\}$ then

$$\begin{aligned}
&\left|x(t) - x_0 - \int_{t_0}^{t} f(\tau, x(\tau))\,\mathrm{d}\tau\right| \\
&\quad \leq |x(t) - x_k(t)| + \left|x_k(t) - x_0 - \int_{t_0}^{t} f(\tau, x_k(\tau))\,\mathrm{d}\tau\right| + \int_{t_0}^{t} \left|f(\tau, x_k(\tau)) - f(\tau, x(\tau))\right| \mathrm{d}\tau \\
&\quad \leq \frac{\varepsilon}{3 + 3rK} + |t - t_0| \cdot \varepsilon_k + |t - t_0| \cdot \frac{\varepsilon}{3r} < \varepsilon.
\end{aligned}$$

As ε is here an arbitrary positive number, we can conclude that, in fact for all $t \in [t_0 - r, t_0 + r]$ we have that

$$x(t) - x_0 - \int_{t_0}^{t} f(\tau, x(\tau))\,\mathrm{d}\tau$$

and this implies, as we know, that the function x is of class C^1 on $(t_0 - r, t_0 + r)$ and a solution to the initial value problem in the proposition. □

We can use these results to approximate solutions to initial value problems to a given accuracy. Let us give an example of this.

Ejemplo 9.8.2.10. Let us consider the initial value problem

$$\begin{cases} x' = \cos t + \sin x, \\ x(0) = 1, \end{cases}$$

and let us suppose that $x : [0, 2\pi] \to \mathbb{R}$ is the solution defined on the interval $[0, 2\pi]$. We want to find an approximate solution $y : [0, 2\pi] \to \mathbb{R}$ such that $|y(t) - x(t)| < 10^{-3}$ for all $t \in [0, 2\pi]$.

The function $f : (t, x) \in \mathbb{R} \times \mathbb{R} \to \cos t + \sin x \in \mathbb{R}$ that appears in the right hand side of the differential equation satisfies the Lipschitz condition with $K = 1$, and has $|f(t, x)| \leq 2$ for all $(t, x) \in \mathbb{R} \times \mathbb{R}$. If ε is a positive number and y is an arbitrary ε-approximate solution to the initial value problem, then since x is also an ε-approximate solution and $x(0) = y(0)$, we know from Proposición 9.8.2.8 that

$$|y(t) - x(t)| \leq 2\varepsilon(e^{2\pi} - 1)$$

for all $t \in [0, 2\pi]$. We therefore have to choose ε so that $2\varepsilon(e^{2\pi} - 2) < 10^{-3}$. On the other hand, for all $(t_1, x_1), (t_2, x_2) \in [0, 2\pi] \times \mathbb{R}$ we have

$$|t_1 - t - 2| < \varepsilon, |x_1 - x_2| < \varepsilon \implies |f(t_1, x_1) - f(t_2, x_2)| < \varepsilon,$$

so going back to the beginning of our proof of Proposición 9.8.2.5 we see that if we choose $\ell \in \mathbb{N}$ such that

$$\frac{2\pi}{\ell} < \frac{\varepsilon}{2}$$

the Euler polygon that we constructed there will be satisfy our condion. A little calculation show that we need to take ℓ to be at least 13 433 241 for this to hold.

This is a rather big number, and the reason for this is that the bounds that we get from Proposición 9.8.2.8 are usually way too big. Indeed, in this case one can compute tha taking $\ell \sim 5000$ is enough to get the error between the Euler polygon and the exact solution smaller than 10^{-3}: the value of ℓ that we found above is thus almost 2 500 times larger than what is really necessary. In practice one uses other methods to decide the step size that is going to be used to construct approximations.

Proposición 9.8.2.11. *Sea (t_0, x_0) un punto $\mathbb{R} \times \mathbb{R}^n$, sean r y s dos números positivos, y sea*

$$\Omega \coloneqq \{(t, x) \in \mathbb{R} \times \mathbb{R}^n : t_0 \leq t \leq t_0 + r, \|x - x_0\| < s\}.$$

Sea $f : \overline{\Omega} \to \mathbb{R}^n$ una función continua y sean

$$M \coloneqq \sup\{|f(t, x)| : (t, x) \in \overline{\Omega}\}, \qquad \rho \coloneqq \min\{r, s/2M\}.$$

Hay una función continua $x : [t_0, t_0 + \rho] \to \mathbb{R}$ que es derivable en $(t_0, t_0 + r)$ y tal que $(t, x(t)) \in \Omega$ para todo $t \in (t_0, t_0 + r)$ y

$$\begin{cases} x'(t) = f(t, x(t)) & \text{para cada } t \in (t_0, t_0 + r), \\ x(t_0) = x_0. \end{cases}$$

Demostración. La Proposición 9.8.2.5 nos dice que para cada $k \in \mathbb{N}$ hay una función continua $x_k : [t_0 - \rho, t_0 + \rho] \to \mathbb{R}^n$ de clase C^1_p tal que

(P_1) $x_k(t_0) = x_0$,

(P_2) $|x_k(s) - x_k(t)| \leq M \cdot |s - t|$ siempre que s y t están en $[t_0, t_0 + \rho]$, y

(P_3) para cada $t \in [t_0 - \rho, t_0 + \rho]$ tal que x_k es differentiable en t el punto $(t, x_k(t))$ está en el dominio de f y $|x_k'(t) - f(t, x_k(t))| < 1/k$.

Las primeras dos de estas condiciones implican que $|x_k(t)| \leq |x_0| + 2\rho M$ para todo $k \in \mathbb{N}$ y todo $t \in [t_0 - \rho, t_0 + \rho]$, así que la sucesión $(x_k)_{k\geq 1}$ está uniformemente acotada. Por otro lado, la segunda de estas condiciones y el Ejercicio 9.8.1.2 nos dicen que esa sucesión es uniformemente equicontinua. El *Teorema de Arzelà–Ascoli* 9.8.5 nos permite entonces concluir que hay una función estrictamente creciente $h : \mathbb{N} \to \mathbb{N}$ tal que la sucesión $(x_{h(k)})_{k\geq 1}$ converge uniformemente a una función $x : [t_0, t_0 + \rho] \to \mathbb{R}^n$.

Para todo $k \in \mathbb{N}$ es $x_{h(k)}(t_0) = x_0$, así que evidentemente es $x(t_0) = x_0$. De manera similar, si t es un elemento de $[t_0, t_0 + \rho]$, entonces para todo $k \in \mathbb{N}$ tenemos que

$$|x_{h(k)}(t) - x_0| = |x_{h(k)}(t) - x_{h(k)}(t_0)| \leq M \cdot |t - t_0| < \frac{s}{2},$$

así que tomando límites vemos que $|x(t) - x_0| < s/2$: tenemos entonces que $(t, x(t)) \in \Omega$ para todo $t \in (t_0 - \rho, t_0 + \rho)$.

Sea ε un número positivo. Como la función f es continua en $\overline{\Omega}$, que es un conjunto compacto, es uniformemente continua y hay un número positivo δ tal que para cada (τ, ξ), $(\tau', \xi') \in \overline{\Omega}$ es

$$|\tau - \tau'| < \delta, |\xi - \xi'| < \delta \implies |f(\tau, \xi) - f(\tau', \xi')| < \frac{\varepsilon}{3\rho}, \tag{9.13}$$

y podemos elegir δ de manera que además sea $\delta < \varepsilon/3$. Por otro lado, como la sucesión $(x_{h(k)})_{k\geq 1}$ converge uniformemente a x hay un número positivo K tal que siempre que k es un entero vale

$$k \geq K \implies |x(t) - x_{h(k)}(t)| < \delta \quad \text{for all } t \in [t_0, t_0 + \rho], \tag{9.14}$$

y podemos elegir K de manera que además sea $1/K < \varepsilon/3\rho$.

Sea ahora k un entero tal que $k \geq K$ y sea s un elemento de $[t_0 - \rho, t_0 + \rho]$. De (9.14) sabemos que

$$|x(s) - x_{h(k)}(s)| < \delta < \frac{\varepsilon}{3}.$$

Como la función $x_{h(k)}$ es de clase C^1_p, sabemos además que

$$x_{h(k)}(s) = x_0 + \int_{t_0}^{s} x'_{h(k)}(\tau)\, \mathrm{d}\tau,$$

y como satisface la condición (P_3) tenemos que

$$\left|\int_{t_0}^{s}\Big(x'_{h(k)}(\tau)-f(\tau,x_{h(k)}(\tau))\Big)\,\mathrm{d}\tau\right| \le \int_{t_0}^{s}\Big|(x'_{h(k)}(\tau)-f(\tau,x_{h(k)}(\tau))\Big|\,\mathrm{d}\tau \tag{9.15}$$
$$\le \frac{1}{h(k)}\cdot|s-t_0| \le \frac{\rho}{k} \le \frac{\rho}{K} \le \frac{\varepsilon}{3}.$$

Notemos que como la función $x_{h(k)}$ es de clase C^1_p la función

$$t\in[t_0-\rho,t_0+\rho]\mapsto x'_{h(k)}(t)-f(t,x_{h(k)}(t))\in\mathbb{R}^n$$

es continua a trozos y, por lo tanto, integrable, así que (9.15) tiene sentido.

Finalmente, usando (9.14) y (9.13) vemos que

$$\left|\int_{t_0}^{s}\Big(f(\tau,x_{h(k)}(\tau))-f(\tau,x(\tau))\Big)\,\mathrm{d}\tau\right| \le \int_{t_0}^{s}\Big|f(\tau,x_{h(k)}(\tau))-f(\tau,x(\tau))\Big|\,\mathrm{d}\tau$$
$$\le |s-t_0|\cdot\frac{\varepsilon}{3\rho}\le\frac{\varepsilon}{3}.$$

Juntando todo podemos concluir que

$$\begin{aligned}&\left|x(s)-x_0-\int_{t_0}^{s}f(\tau,x(\tau))\,\mathrm{d}\tau\right|\\ &\quad\le|x(s)-x_{h(k)}(s)|+\left|x_{h(k)}(s)-x_0-\int_{t_0}^{s}x'_{h(k)}(\tau)\,\mathrm{d}\tau\right|\\ &\qquad+\left|\int_{t_0}^{s}\Big(x'_{h(k)}(\tau)-f(\tau,x_{h(k)}(\tau))\Big)\,\mathrm{d}\tau\right|+\left|\int_{t_0}^{s}\Big(f(\tau,x_{h(k)}(\tau))-f(\tau,x(\tau))\Big)\,\mathrm{d}\tau\right|\\ &\quad\le\frac{\varepsilon}{3}+0+\frac{\varepsilon}{3}+\frac{\varepsilon}{3}=\varepsilon.\end{aligned}$$

Como el número positivo ε y el elemento s de $[t_0-\rho,t_0+\rho]$ que aparecen aquí son arbitrarios, se sigue de todo esto que

$$x(t)=x_0+\int_{t_0}^{t}f(\tau,x(\tau))\,\mathrm{d}\tau\qquad\text{for all } t\in[t_0-\rho,t_0+\rho].$$

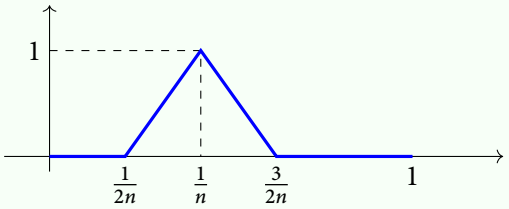
El *Teorema fundamental del cálculo* nos permite deducir de esto que la función x es de clase C^1 en $(t_0,t_0+\rho)$ y que para todo elemento t de ese intervalo es $x'(t)=f(t,x(t))$. □

9.9. El teorema de Dini

En general, una sucesión de funciones entre espacios métricos que converge puntualmente no converge uniformemente, ni siquiera si su dominio es un espacio compacto.

Ejemplo 9.9.1. Para cada $n \in \mathbb{N}_0$ consideremos la función $f_n : [0,1] \to \mathbb{R}$ que en cada elemento x de $[0,1]$ toma el valor

$$f(x) = \begin{cases} n\left(x - \frac{1}{2n}\right) & \text{si } x \in [\frac{1}{2n}, \frac{1}{n}); \\ 2n\left(\frac{3}{2n} - x\right) & \text{si } x \in [\frac{1}{n}, \frac{3}{2n}]; \\ 0 & \text{en cualquier otro caso.} \end{cases}$$



Es fácil ver que la sucesión $(f_n)_{n\geq 0}$ converge puntualmente a la función nula en $C([0,1],\mathbb{R})$, pero que no lo hace uniformemente.

El siguiente resultado, el *Teorema de Dini*, debido a Ulisse Dini, da una condición que, junto con la convergencia puntual, implica la convergencia uniforme.

Proposición 9.9.2. *Sea X un espacio métrico compacto. Una sucesión de funciones continuas $X \to \mathbb{R}$ que converge puntualmente a una función continua $X \to \mathbb{R}$ de manera monótona lo hace uniformemente.*

Demostración. Sea $(f_n)_{n\geq 0}$ una sucesión en $C(X,\mathbb{R})$, sea $f : X \to \mathbb{R}$ una función continua, supongamos que para cada $x \in X$ la sucesión $(f_n(x))_{n\geq 0}$ crece — si decreciera sería todo similar — y converge a $f(x)$, y sea ε un número positivo. Consideremos para cada entero no negativo n el conjunto

$$E_n \coloneqq \{x \in X : f(x) < f_n(x) + \varepsilon\}$$

que, como f y f_n son funciones continuas, es abierto. Si x es un elemento de X, entonces la sucesión $(f_n(x))_{n\geq 0}$ converge a $f(x)$, así que hay un entero no negativo m tal que $|f_n(x) - f(x)| < \varepsilon$ y, en particular, $f(x) < f_m(x) + \varepsilon$, de manera que $x \in E_m$: esto nos dice que $\{E_n : n \in \mathbb{N}_0\}$ es un cubrimiento abierto de X. Como X es compacto, hay entonces un subconjunto finito y no vacío F de $\mathbb{N}_0$ tal que $X = \bigcup_{n\in F} E_n$. Sea $N \coloneqq \max F$. Si n es un elemento cualquiera de F y x es un elemento de E_n, entonces $f(x) < f_n(x) + \varepsilon \leq f_N(x) + \varepsilon$, ya que la sucesión $(f_n(x))_{n\geq 0}$ es creciente y $n \leq N$, así que $x \in E_N$: esto nos dice que $E_n \subseteq E_N$ y, por lo tanto, que de hecho $X = \bigcup_{n\in F} E_n = E_N$.

Sea ahora x un elemento cualquiera de X y sea m un entero tal que $m \geq N$. Como la sucesión $(f_n(x))_{n\geq 0}$ crece a $f(x)$ y $x \in E_N$, tenemos que $f_m(x) \leq f(x) < f_N(x) + \varepsilon \leq f_m(x) + \varepsilon$, así

que $-\varepsilon < 0 \leq f(x) - f_m(x) < \varepsilon$, esto es, $|f(x) - f_m(x)| < \varepsilon$. Esto prueba que la sucesión $(f_n)_{n\geq 0}$ converge uniformemente a f, como queremos. □

Sin la hipótesis de que el espacio X sea compacto la afirmación de la Proposición 9.9.2 no es cierta.

Ejercicio 9.9.3. Para cada $n \in \mathbb{N}_0$ consideremos la función $f_n : t \in (0,1) \mapsto (1+nx)^{-1} \in \mathbb{R}$. Muestre que la sucesión $(f_n)_{n\geq 0}$ tiene todas sus componentes continuas, que converge puntualmente a una función continua de manera monótona, y que esa convergencia no es uniforme.

Notemos que la convergencia puntual monótona en la Proposición 9.9.2 no es suficiente para garantizar la continuidad del límite.

Ejercicio 9.9.4. Para cada $n \in \mathbb{N}_0$ consideremos la función $f_n : t \in [0,1] \mapsto t^n \in \mathbb{R}$. Muestre que la sucesión $(f_n)_{n\geq 0}$ tiene todas sus componentes continuas, que converge puntualmente de manera monótona, y que el límite no es una función continua — de manera que, en particular, la convergencia no es uniforme, claro.

Finalmente, observemos que en la Proposición 9.9.2 no podemos reemplazar la hipótesis de que para cada punto x de X la sucesión $(f_n(x))_{n\geq 0}$ sea monótona por la de que sea a la larga monótona: la sucesión del Ejemplo 9.9.1 es, de hecho, a la larga constante, pero no converge uniformemente.

Ejercicio 9.9.5. Sea X un espacio métrico compacto, sea Y un espacio métrico cualquiera, sea $(f_n)_{n\geq 0}$ una sucesión de funciones continuas y acotadas $X \to Y$, y sea $f : X \to Y$ una función continua. Muestre que si la sucesión $(f_n)_{n\geq 0}$ converge puntualmente a f y para todo elemento x de X la sucesión $(d(f_n(x), f(x)))_{n\geq 0}$ es decreciente, entonces la sucesión $(f_n)_{n\geq 0}$ converge uniformemente a f.

9.10. El teorema de Stone–Weierstrass

Sea X un espacio métrico y sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$. El conjunto $C(X,\mathbb{k})$ de las funciones continuas y acotadas $X \to \mathbb{k}$ es un $\mathbb{k}$-espacio normado real con respecto a la norma $\|-\|_\infty$ de la convergencia uniforme. El hecho de que sus elementos sean funciones con valores en $\mathbb{k}$ y que en $\mathbb{k}$ podamos multiplicar tiene la siguiente consecuencia.

Proposición 9.10.1. *Sea X un espacio métrico y sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$.*

(*i*) *Si f y g son dos elementos de $C(X,\mathbb{k})$, entonces la función $fg : x \in X \mapsto f(x)g(x) \in \mathbb{k}$ es*

continua y acotada, de manera que se trata de un elemento de $C(X, \Bbbk)$.

(*ii*) *La función*

$$\mu : (f, g) \in C(X, \Bbbk) \times C(X, \Bbbk) \mapsto fg \in C(X, \Bbbk)$$

que obtenemos de esta forma es $\Bbbk$*-bilineal y asociativa, y hace del* $\Bbbk$*-espacio vectorial* $C(X, \Bbbk)$ *una* $\Bbbk$*-álgebra unitaria cuyo elemento unidad es la función constante* $\mathbf{1} : x \in X \mapsto 1 \in \Bbbk$*, que tiene norma* 1.

(*iii*) *Cualesquiera sean* f *y* g *en* $C(X, \Bbbk)$ *vale que*

$$\|fg\|_\infty \leq \|f\|_\infty \|g\|_\infty$$

y la función μ *es continua.*

Demostración. Sean f y g dos elementos de $C(X, \Bbbk)$, sea x_0 un elemento de X, y sea ε un número positivo. Como f y g son continuas en X, hay un número positivo δ tal que para todo elemento x de $B_\delta(x_0)$ es $|f(x) - f(x_0)| < \varepsilon/(1 + 2\|g\|_\infty)$ y $|g(x) - g(x_0)| < \varepsilon/(1 + 2\|f\|_\infty)$. Esto implica que para cada elemento x de $B_\delta(x_0)$ es

$$\begin{aligned} |(fg)(x) - (fg)(x_0)| &= |f(x)g(x) - f(x_0)g(x_0)| \\ &\leq |f(x)g(x) - f(x)g(x_0)| + |f(x)g(x_0) - f(x_0)g(x_0)| \\ &= |f(x)| \cdot |g(x) - g(x_0)| + |f(x) - f(x_0)| \cdot |g(x_0)| \\ &< \|f\|_\infty \cdot \frac{\varepsilon}{1 + 2\|f\|_\infty} + \frac{\varepsilon}{1 + 2\|f\|_\infty} \cdot \|g\|_\infty < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Vemos así que la función fg es continua, y esto prueba la afirmación (*i*) de la proposición.

Que la afirmación (*ii*) es cierta es consecuencia de varios cálculos directos muy sencillos que dejamos a cargo del lector. Veamos que también es cierta la afirmación (*iii*). Para cada elemento x de X es

$$|(fg)(x)| = |f(x)g(x)| = |f(x)| \cdot |g(x)| \leq \|f\|_\infty \cdot \|g\|_\infty,$$

así que $\|fg\|_\infty = \sup_{x \in X} |(fg)(x)| \leq \|f\|_\infty \cdot \|g\|_\infty$. □

En los términos de la siguiente definición, la proposición que acabamos de probar nos dice que $C(X, \Bbbk)$ es una $\Bbbk$-álgebra normada y, de hecho, una $\Bbbk$-álgebra de Banach, ya que sabemos que es completa porque $\Bbbk$ lo es.

Definición 9.10.2. Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$. Una ***$\Bbbk$-álgebra normada*** es un $\Bbbk$-espacio normado A dotado de

una multiplicación $\cdot : A \times A \to A$ que es $\Bbbk$-bilineal, asociativa, y tal que

$$\|xy\| \leq \|x\| \cdot \|y\|$$

cualesquiera sean x e y en A. Decimos que esa $\Bbbk$-álgebra normada es ***unitaria*** si hay un elemento $\mathbf{1}$ en A que es una identidad para el producto de A y tal que $\|\mathbf{1}\| = 1$, y que es una ***$\Bbbk$-álgebra de Banach*** si es completa con respecto a su norma.

La primera observación importante que hay que hacer sobre las álgebras normadas es que sus multiplicaciones son funciones continuas.

Proposición 9.10.3. *Sea A una $\Bbbk$-álgebra de Banach unitaria. La función $\cdot : A \times A \to A$ es continua.*

Demostración. Sea (x, y) un elemento de $A \times A$, sea ε un número positivo, y pongamos

$$\delta \coloneqq \frac{\varepsilon}{\|x\|_\infty + \|y\|_\infty}.$$

Si (x', y') es un elemento de $A \times A$ tal que $d((x, y), (x', y')) < \delta$, entonces $\|x' - x\|_\infty < \delta$ y $\|y' - y\|_\infty < \delta$ y, por lo tanto,

$$\begin{aligned}
\|x'y' - xy\|_\infty &\leq \|x'y' - x'y\|_\infty + \|x'y - xy\|_\infty \\
&\leq \|x'\|_\infty \cdot \|y' - y\|_\infty + \|x' - x\|_\infty \cdot \|y\|_\infty \\
&< \|x'\|_\infty \cdot \frac{\varepsilon}{\|x\|_\infty + \|y\|_\infty} + \frac{\varepsilon}{\|x\|_\infty + \|y\|_\infty} \cdot \|y\|_\infty \\
&= \varepsilon.
\end{aligned}$$

Esto nos dice que la función del enunciado es continua. □

La teoría de las álgebras normadas y de Banach es interesante porque muestra constantemente una interacción entre las propiedades algebraicas y las propiedades métricas de las álgebras que estudia. Veamos un ejemplo de esto.

Proposición 9.10.4. *Sea A una $\Bbbk$-álgebra de Banach unitaria. Si x es un elemento de X tal que $\|x - \mathbf{1}\| < 1$, entonces x es inversible.*

Demostración. Sea x un elemento de A tal que el número $\rho \coloneqq \|\mathbf{1} - x\|$ es menor que 1. Consideremos la serie $\sum_{n\geq 0}(\mathbf{1} - x)^n$ en A. Para cada entero no negativo tenemos que $\|(\mathbf{1} - x)^n\| \leq \|\mathbf{1} - x\|^n = \rho^n$ y, como la serie $\sum_{n\geq 0} \rho^n$ converge en $\mathbb{R}$, sabemos que la serie $\sum_{n\geq 0}(\mathbf{1} - x)^n$ converge en A. Sea y su suma.

Ahora bien, para cada entero no negativo N tenemos que

$$\mathbf{1} - x \cdot \sum_{n=0}^{N} (\mathbf{1} - x)^n = \mathbf{1} - (\mathbf{1} - (\mathbf{1} - x)) \cdot \sum_{n=0}^{N} (\mathbf{1} - x)^n = (1 - x)^{N+1},$$

así que

$$\left\| \mathbf{1} - x \cdot \sum_{n=0}^{N} (\mathbf{1} - x)^n \right\| = \|(1 - x)^{N+1}\| \leq \|1 - x\|^{N+1} = \rho^{N+1}$$

y, como $(\rho^n)_{n\geq 0}$ converge a 0 y la multiplicación de A es continua, tenemos que

$$x \cdot y = x \cdot \lim_{n\to\infty} \sum_{n=0}^{N} (\mathbf{1} - x)^n = \lim_{n\to\infty} \left(x \cdot \sum_{n=0}^{N} (\mathbf{1} - x)^n \right) = \mathbf{1}.$$

De manera similar podemos mostrar que $y \cdot x = \mathbf{1}$, así que el elemento y es un inverso para x en A, por lo que x es inversible en A, como queremos. □

Como consecuencia de esto, el conjunto de los elementos inversibles de un álgebra de Banach es abierto.

Corolario 9.10.5. *Sea A una $\mathbb{k}$-álgebra de Banach unitaria. El conjunto $A^\times$ de los elementos inversibles de A es abierto.*

Demostración. Sea x un elemento inversible de A, de manera que hay un elemento y de A tal que $xy = yx = \mathbf{1}$. Ciertamente y no es nulo, así que podemos considerar el número positivo $\delta \coloneqq 1/\|y\|$. Si z es un elemento de $B_\delta(x)$, entonces

$$\|yz - \mathbf{1}\| = \|yz - yx\| \leq \|y\| \cdot \|z - x\| < \|y\| \cdot \delta = 1,$$

así que la proposición nos dice que yz es inversible en A y, por lo tanto, que z lo es. Vemos así que $B_\delta(x) \subseteq A^\times$ y, en definitiva, que el conjunto $A^\times$ es abierto. □

La razón por la que pudimos probar la Proposición 9.10.4 es que cuando t es un elemento de $(-1, 1)$ los números

$$1 + t + t^2 + \cdots + t^n$$

convergen cuando n crece a $(1 - t)^{-1}$. Esta idea es bien flexible.

Ejercicio 9.10.6. Sea A una $\mathbb{C}$-álgebra de Banach.

(a) Muestre que para todo elemento x de A convergen en A las series

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

Escribamos $\exp(x)$, $\cos(x)$ y $\sin(x)$ a sus sumas.

(b) Pruebe que si x e y son dos elementos de A tales que $xy = yx$, entonces valen las siguientes igualdades:

$$\exp(x+y) = \exp(x)\cdot\exp(y), \qquad \cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y),$$
$$\exp(ix) = \cos(x) + i\sin(x), \qquad \sin(x+y) = \sin(x)\cos(y) + \cos(x)\sin y.$$

El siguiente lema nos servirá más adelante en la demostración de la Proposición 9.10.9.

Lema 9.10.7. *Hay exactamente una sucesión $(p_n)_{n\geq 0}$ en $\mathbb{R}[X]$ tal que $p_0 = 0$ y*

$$p_{n+1} = p_n + \frac{1}{2}(X - p_n^2)$$

para todo $n \in \mathbb{N}_0$ y la sucesión $(p_n|_{[0,1]})_{n\geq 0}$ converge en $C([0,1],\mathbb{R})$ a la función $t \in [0,1] \mapsto \sqrt{t} \in \mathbb{R}$ de manera monótona creciente.

Demostración. Que hay efectivamente una tal sucesión de elementos de $\mathbb{R}[X]$ es una consecuencia inmediata del *Principio de recurrencia* 1.2.1, y es $p_1 = 0 + \frac{1}{2}(X - 0^2) = \frac{1}{2}X$.

Sea t un elemento de $[0,1]$. Es $p_0(t) = 0 \leq \frac{1}{2}t = p_1(t) \leq \sqrt{t}$. Sea, por otro lado, n un elemento de $\mathbb{N}_0$ tal que $p_n(t) \leq p_{n+1}(t) \leq \sqrt{t}$. Como $p_{n+1}(t)^2 \leq t$, tenemos que

$$p_{n+1}(t) \leq p_{n+1}(t) + \frac{1}{2}(t - p_{n+1}(t)^2) = p_{n+2}(t).$$

Como $p_{n+1}(t) \leq \sqrt{t} \leq 2 - \sqrt{t}$, es $\sqrt{t} + p_{n+1}(t) \leq 2$ y, como $p_n(t) \leq \sqrt{t}$, es

$$t - p_{n+1}^2 = (\sqrt{t} + p_{n+1})(\sqrt{t} - p_{n+1}) \leq 2(\sqrt{t} - p_{n+1})$$

y, por lo tanto, $p_{n+1} = p_{n+1} + \frac{1}{2}(t - p_{n+1}(t)^2) \leq \sqrt{t}$. Gracias al principio de inducción, entonces, podemos concluir que para todo $n \in \mathbb{N}_0$ es

$$p_n(t) \leq p_{n+1}(t) \leq \sqrt{t}.$$

Vemos así que la sucesión $(p_n(t))_{n\geq 0}$ converge a un número positivo $p(t)$ menor o igual que $\sqrt{t}$. Como

$$p_{n+1}(t) = p_n(t) + \frac{1}{2}(t - p_n^2(t))$$

para todo $n \in \mathbb{N}_0$, tomando límites a ambos lados de la igualdad vemos que

$$p(t) = p(t) + \frac{1}{2}(t - p(t)^2),$$

así que, por supuesto, es $p(t) = \sqrt{t}$. Como la sucesión $(p_n)_{n\geq 0}$ converge entonces puntualmente sobre $[0,1]$ a $t \in [0,1] \mapsto \sqrt{t} \in \mathbb{R}$ de manera monótona, el *Teorema de Dini* 9.9.2 nos dice que la convergencia es, de hecho, uniforme. □

A continuación enunciamos y probamos el *Teorema de Stone–Weierstrass*. Karl Weierstrass probó originalmente en 1889 un caso particular, que enunciamos abajo como el Corolario 9.10.11, y mucho más tarde Marshall Stone dio en [Sto1937] la versión considerablemente más general que presentamos aquí.

Definición 9.10.8. Sea X un espacio métrico, sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$. Decimos que un subconjunto A de $C(X,\mathbb{k})$

- ***separa puntos de*** X si cada vez que x e y son dos elementos distintos de X hay un elemento f de A tal que $f(x) \neq f(y)$, y
- es ***cerrado por conjugación*** si siempre que f es un elemento de A la función $x \in X \mapsto \overline{f(x)} \in \mathbb{k}$ también está en A.

Notemos que cuando $\mathbb{k}$ es $\mathbb{R}$ todo subconjunto de $C(X,\mathbb{k})$ es cerrado por conjugación.

Proposición 9.10.9. *Sea X un espacio métrico compacto y sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$. Una subálgebra cerrada A de $C(X,\mathbb{k})$ coincide con $C(X,\mathbb{k})$ si satisface las siguientes tres condiciones:*

- **1** *pertenece a A.*
- *A separa los puntos de X.*
- *A es cerrada por conjugación.*

Como ciertamente $C(X,\mathbb{k})$ es una subálgebra cerrada de sí misma y satisface esas tres condiciones, lo que nos dice esta proposición es que es, de hecho, la única.

Demostración. Supongamos que A es una subálgebra cerrada de $C(X,\mathbb{k})$ que satisface las tres condiciones del enunciado.

Primer paso. Mostremos que

si $f \in A$, entonces $|f| \in A$.

Sea f un elemento de A. La función $g \coloneqq f\bar{f}/\|f\|_\infty^2$ también está en A y claramente $g(x) \in [0,1]$ para todo $x \in X$. Sea $(p_n)_{n\geq 0}$ la sucesión de polinomios del Lema 9.10.7 y consideremos la

sucesión $(p_n(g))_{n\geq 0}$ de elementos de A. Como para todo $t \in [0,1]$ la sucesión $(p_n(t))_{n\geq 0}$ crece monótonamente a $\sqrt{t}$, es claro que para todo $x \in X$ la sucesión $(p_n(g)(x))_{n\geq 0}$, que coincide con $(p_n(g(x)))_{n\geq 0}$ crece monótonamente a $\sqrt{g(x)}$. Como esta última es continua, el *Teorema de Dini* 9.9.2 nos dice que la sucesión $(p_n(g))_{n\geq 0}$, que toma valores en A, converge uniformemente a $\sqrt{g}$: vemos así que $|f|/\|f\|_\infty = \sqrt{g}$ está en A y, por lo tanto, que $|f|$ también.

Observemos ahora que si f y g son dos elementos de A que toman valores reales, entonces

$$\text{máx}\{f,g\} = \frac{f+g}{2} + \frac{|f-g|}{2} \in A, \qquad \text{mín}\{f,g\} = \frac{f+g}{2} - \frac{|f-g|}{2} \in A.$$

Segundo paso. Mostremos ahora que

> *si $f \in C(X,\mathbb{k})$ es una función con valores reales, ε es un número positivo, y x es un elemento de X, entonces hay un elemento g de A también con valores reales tal que $g(x) = f(x)$ y $g \leq f + \varepsilon$.*

Sea f un elemento de $C(X,\mathbb{k})$ con valores reales. Para cada punto y de X distinto de x hay un elemento h_y en A tal que $h_y(x) \neq h_y(y)$ y podemos considerar el elemento $k_y := |h_y - h_y(x)|$ de A, que toma valores reales y es tal que $k_y(x) \neq k_y(x)$, y poner

$$g_y = f(x) + \frac{f(y)-f(x)}{k_y(y)-k_y(x)}(k_y - k_y(x)),$$

que es un elemento de A con valores reales tal que $g_y(x) = f(x)$. Como $g_y(y) < f(y) + \varepsilon$, hay un número positivo δ_y tal que $g_y(z) < f(z) + \varepsilon$ para todo elemento z de $B_{\delta_y}(y)$. Pongamos además $g_x := f(x)\mathbf{1}$: como $g_x(x) < f(x) + \varepsilon$, hay un número positivo δ_x tal que $g_x(z) < f(z) + \varepsilon$ para todo $z \in B_{\delta_x}(x)$.

El conjunto $\{B_{\delta_y}(\delta_y) : y \in X\}$ es un cubrimiento abierto de X, así que hay un subconjunto finito F de X tal que $X = \bigcup_{y\in F} B_{\delta_y}(y)$. Sean n el cardinal de F y sean $y_1, \ldots, y_n$ los elementos de F. Por lo que hicimos arriba, la función

$$g := \text{mín}\{g_{y_1}, \ldots, g_{y_n}\}$$

es un elemento de A. Es

$$g(x) = \text{mín}\{g_{y_1}(x), \ldots, g_{y_n}(x)\} = \text{mín}\{f(x), \ldots, f(x)\} = f(x).$$

Por otro lado, si z es un elemento cualquiera de X, entonces hay un elemento i de $[\![n]\!]$ tal que $z \in B_{\delta_{y_i}}(y_i)$ y entonces $g(z) \leq g_{y_i}(z) < f(z) + \varepsilon$. Vemos así que $g < f + \varepsilon$. Vemos así que la afirmación que queremos probar es cierta.

Tercer paso. Queremos probar que

> *si $f \in C(X,\mathbb{k})$ es una función con valores reales y ε es un número positivo, entonces hay un elemento g de A también con valores reales tal $\|f - g\| < \varepsilon$.*

Sea f un elemento de $C(X, \Bbbk)$ con valores reales y sea ε un número positivo. De acuerdo a lo que hicimos en el paso anterior, para cada elemento x de X hay un elemento g_x de A a valores reales tal que $g_x(x) = f(x)$ y $g_x < f + \varepsilon/2$ y, por lo tanto, hay un número positivo η_x tal que $g_x(z) > f(z) - \varepsilon/2$ para todo $z \in B_{\eta_x}(x)$. Como antes, la compacidad de X implica que hay un subconjunto finito G de X tal que $X = \bigcup_{x \in G} B_{\eta_x}(x)$. Sea m el cardinal de G, sean $x_1, \ldots, x_m$ sus elementos, y consideremos la función

$$g \coloneqq \text{máx}\{g_{x_1}, \ldots, g_{x_m}\},$$

que pertenece a A y es tal que $g < f + \varepsilon/2$. Si z es un elemento de X, entonces hay un elemento j de $[\![m]\!]$ tal que $z \in B_{\eta_{x_j}}(x_j)$ y, por lo tanto,

$$f(z)/2 - \varepsilon < g_{x_i}(z) \leq g(z) < f(z) + \varepsilon/2.$$

Vemos así que $\|f - g\| < \varepsilon$, como queremos.

Cuarto y último paso. Si el cuerpo $\Bbbk$ es $\mathbb{R}$, entonces lo que hicimos en el tercer paso nos dice que la subálgebra A es densa en $C(X, \Bbbk)$.

Supongamos ahora que $\Bbbk$ es $\mathbb{C}$, sea f un elemento cualquiera de $C(X, \mathbb{C})$, y sea ε un número positivo. Las funciones $\text{Re}(f)$ e $\text{Im}(f)$ están entonces en $C(X, \mathbb{C})$ y toman valores reales, así que lo que hicimos en el tercer paso nos dice que hay elementos g y h en A, los dos a valores reales, tales que $\|\text{Re}(f) - g\| < \varepsilon/2$ y $\|\text{Im}(f) - h\| < \varepsilon/2$, y esto implica que

$$\|f - (g + ih)\| \leq \|\text{Re}(f) - g\| + \|i(\text{Im}(f) - h)\| < \varepsilon.$$

Vemos así que A es densa en $C(X, \Bbbk)$ también en este caso. Como A es cerrada, vemos que $A = C(X, \Bbbk)$, como queremos. □

En las aplicaciones construimos álgebras que satisfacen las hipótesis de la Proposición 9.10.9 usando el siguiente resultado.

Lema 9.10.10. *Sea X un espacio métrico compacto, sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$. La clausura $\overline{A}$ de una subálgebra A de $C(X, \Bbbk)$ es una subálgebra de $C(X, \Bbbk)$, y es unitaria si A lo es y es es cerrada por conjugación si A lo es.*

Demostración. Sea A una subálgebra de $C(X, \Bbbk)$. Sean x e y dos elementos de $\overline{A}$, sean a y b dos elementos de $\Bbbk$, y sea ε un número positivo. Hay elementos f' e g' en A tales que

$$\|f - f'\| < \frac{\varepsilon}{2(1 + |a| + \|g\|)}, \qquad \|g - g'\| < \text{mín}\left\{\frac{\varepsilon}{2(1 + |b| + \|f\|)}, 1\right\},$$

y es

$$\|(af+bg)-(af'+bg')\| \le |a|\cdot\|f-f'\| + |b|\cdot\|g-g'\| < \frac{\varepsilon}{2}+\frac{\varepsilon}{2} = \varepsilon$$

y, como $\|g'\| \le \|g\| + 1$,

$$\begin{aligned}\|fg - f'g'\| &\le \|f(g-g')\| + \|(f-f')g'\| \le \|f\|\cdot\|g-g'\| + \|f-f'\|\cdot\|g'\| \\ &< \frac{\varepsilon}{2}+\frac{\varepsilon}{2} = \varepsilon.\end{aligned}$$

Esto nos dice que $af + bg$ y fg son elementos de $\overline{A}$ y, por lo tanto, que $\overline{A}$ es una subálgebra de $C(X,\Bbbk)$. Esto prueba la primera afirmación del lema.

Si A es unitaria, de manera que $\mathbf{1}$ está en A, entonces es claro que $\overline{A}$ también es unitaria, ya que $A \subseteq \overline{A}$. Supongamos, por otro lado, que A es cerrada por conjugación, que f es un elemento de $\overline{A}$, y sea ε un número positivo. Hay entonces un elemento f' de A tal que $\|f - f'\| < \varepsilon$ y es claro que $\|\overline{f} - \overline{f}'\| < \varepsilon$: como $\overline{f}' \in A$, esto implica que $\overline{f} \in \overline{A}$ y, por lo tanto, que $\overline{A}$ es cerrada por conjugación. □

Usando este lema y la Proposición 9.10.9 podemos probar el siguiente resultado de Weierstrass.

Corolario 9.10.11. *Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$. El conjunto de las funciones polinomiales es denso en $C([0,1],\Bbbk)$*

Demostración. En electo, el conjunto de las funciones polinomiales en $C([0,1],\Bbbk)$ es una subálgebra que contiene las funciones constantes, que separa puntos porque contiene la función inclusión $[0,1] \to \Bbbk$, y que es cerrada por conjugación, así que su clausura es una subálgebra cerrada de $C([0,1],\Bbbk)$ con las mismas tres propiedades y, de acuerdo a la Proposición 9.10.9, es igual a $C([0,1],\Bbbk)$. □

Usando la Proposición 9.10.9 podemos obtener varios otros resultados del mismo tipo. Veamos algunos ejemplos de esto.

Corolario 9.10.12. *Sea $S \coloneqq \{z \in \mathbb{C} : |z| = 1\}$, la circunferencia unidad de $\mathbb{C}$, y para cada $n \in \mathbb{Z}$ consideremos la función $e_n : z \in S \mapsto z^n \in \mathbb{C}$. El subespacio vectorial de $C(S,\mathbb{C})$ generado por el conjunto $\{e_n : n \in \mathbb{Z}\}$ es denso.*

Notemos que los elementos de ese subespacio son los llamados ***polinomios de Laurent***.

Demostración. Si m y n son dos elementos de $\mathbb{Z}$, entonces $e_m \cdot e_n = e_{m+n}$: esto implica inmediatamente que el subespacio A de $C(S,\mathbb{C})$ generado por el conjunto $\{e_n : n \in \mathbb{Z}\}$ es una $\mathbb{C}$-subálgebra. Es $\mathbf{1} = e_0$, la función e_1 es inyectiva puntos, y $\overline{e_n} = e_{-n}$ para todo $n \in \mathbb{N}_0$: vemos así que la subálgebra A es unitaria, separa puntos, y es cerrada por conjugación. Se sigue de esto que se trata de una

subálgebra densa. □

Frecuentemente usamos el Corolario 9.10.12 que acabamos de probar indirectamente, vía el Corolario 9.10.15 que daremos abajo. Para enunciarlo necesitamos el Lema 9.10.14, en cuya demostración usaremos el resultado del siguiente ejercicio.

Ejercicio 9.10.13. Sea δ un número positivo menor que 1. Muestre que si α es un numero real tal que $\cos\alpha > 1 - \delta^2/2$, entonces hay un entero k tal que $|\alpha + 2k\pi| < \delta$.

Lema 9.10.14. *Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, y sea $C_{\mathrm{per}}(\mathbb{R},\Bbbk)$ es espacio vectorial de todas las funciones continuas $f:\mathbb{R}\to\Bbbk$ tales que $f(t+2\pi)=f(t)$ para todo $t\in\mathbb{R}$.*

(*i*) *$C_{\mathrm{per}}(\mathbb{R},\Bbbk)$ es una subálgebra cerrada y unitaria del álgebra de Banach $C(\mathbb{R},\Bbbk)$ de todas las funciones continuas y acotadas $\mathbb{R}\to\Bbbk$, así que, en particular, es un álgebra de Banach con respecto a la norma $\|-\|:C_{\mathrm{per}}(\mathbb{R},\Bbbk)\to\mathbb{R}$ que en cada elemento f de $C_{\mathrm{per}}(\mathbb{R},\Bbbk)$ vale*

$$\|f\| = \sup\{|f(t)| : t\in\mathbb{R}\}.$$

(*ii*) *Sea $S := \{z\in\mathbb{C} : |z| = 1\}$, y sea $p : t\in\mathbb{R}\mapsto e^{it}\in S$. Para cada elemento f de $C(S,\mathbb{C})$ la función $f\circ p$ está en $C_{\mathrm{per}}(\mathbb{R},\mathbb{C})$. La función*

$$F : f\in C(S,\mathbb{C})\mapsto f\circ p\in C_{\mathrm{per}}(\mathbb{R},\mathbb{C})$$

que obtenemos de esa forma es un isomorfismo isométrico de álgebras de Banach.

Demostración. Si $f:\mathbb{R}\to\Bbbk$ es una función continua y periódica de periodo 2π, entonces la restricción $f|_{[0,2\pi]}$ es continua sobre $[0,2\pi]$, así que es acotada, y para todo $t\in\mathbb{R}$ es $t-\lfloor t/2\pi\rfloor 2\pi\in[0,2\pi]$, de manera que

$$|f(t)| = |f(t-\lfloor t/2\pi\rfloor 2\pi)| \le \sup\{|f(s)| : s\in[0,2\pi]\}.$$

Esto nos dice que $C_{\mathrm{per}}(\mathbb{R},\Bbbk)$ es un subconjunto de $C(\mathbb{R},\Bbbk)$, y es claro que es un subespacio vectorial cerrado por la multiplicación de $C(\mathbb{R},\Bbbk)$ y que contiene a la función $\mathbf{1}$. Como $C(\mathbb{R},\Bbbk)$ es un espacio de Banach, para probar la afirmación (*i*) del corolario es suficiente que mostremos que $C_{\mathrm{per}}(\mathbb{R},\Bbbk)$ es un subespacio cerrado de $C(\mathbb{R},\Bbbk)$.

Sea $(f_n)_{n\ge0}$ una sucesión en $C_{\mathrm{per}}(\mathbb{R},\Bbbk)$ que converge en $C(\mathbb{R},\Bbbk)$ a una función $f:\mathbb{R}\to\Bbbk$. Si t es un elemento cualquiera de $\mathbb{R}$, entonces $f_n(t+2\pi)=f_n(t)$ para todo $n\in\mathbb{N}_0$ y, por lo tanto,

$$f(t+2\pi) = \lim_{n\to\infty} f_n(t+2\pi) = \lim_{n\to\infty} f_n(t) = f(t).$$

Esto nos dice que f pertenece a $C_{\mathrm{per}}(\mathbb{R},\Bbbk)$ y prueba lo que queremos.

Como la función p de (ii) es continua y periódica de periodo 2π, es claro que $f \circ p$ es continua y periódica de periodo 2π cualquiera sea el elemento f de $C(S, \mathbb{C})$. Que la función F del enunciado es lineal y un homomorfismo de álgebras es claro, y es inyectiva porque la función p es sobreyectiva.

Sea ahora $g : \mathbb{R} \to \mathbb{C}$ un elemento de $C_{\text{per}}(\mathbb{R}, \mathbb{C})$, y consideremos el conjunto

$$f \coloneqq \{(e^{it}, g(t)) \in S \times \mathbb{C} : t \in \mathbb{R}\}.$$

Si z es un elemento de S, entonces hay un número real t tal que $z = e^{it}$, y entonces tenemos que $(z, g(t)) = (e^{it}, g(t)) \in f$. Por otro lado, si z es un elemento de S y w, w' son dos elementos de $\mathbb{C}$ tales que los pares (z, w) y (z, w') están en f, entonces hay dos números reales t y t' tales que $(z, w) = (e^{it}, g(t))$ y $(z, w') = (e^{it'}, g(t'))$: como $e^{it} = e^{it'}$, hay un entero n tal que $t = t' + 2k\pi$ y, por lo tanto, $w = g(t) = g(t' + 2k\pi) = g(t') = w'$. Todo esto nos dice que el conjunto f es, de hecho, una función $f : S \to \mathbb{C}$, y es claro que $f \circ p = g$. Para ver que la función F es sobreyectiva es suficiente, entonces, que mostremos que la función f es continua.

Sea z_0 un elemento de S y sea ε un número positivo. Sea t_0 un número real tal que $z_0 = e^{it_0}$. Como la función g es continua en t_0, hay un número positivo δ menor que 1 y tal que $|g(t) - g(t_0)| < \varepsilon$ siempre que $t \in \mathbb{R}$ es tal que $|t - t_0| < \delta$. Sea ahora z un elemento de S tal que $|z - z_0| < \delta$, y sea $t \in \mathbb{R}$ tal que $z = e^{it}$. Es

$$|z - z_0|^2 = |e^{it} - e^{it_0}|^2 = |\cos t + i \operatorname{sen} t - \cos t_0 - i \operatorname{sen} t|^2 = 2(1 - \cos(t - t_0)),$$

así que

$$\cos(t - t_0) = 1 - \frac{|z - z_0|^2}{2} > 1 - \frac{\delta^2}{2}.$$

Notemos que

si $\alpha \in \mathbb{R}$ es tal que $\cos \alpha > 1 - \delta^2/2$, entonces hay un entero k tal que $|\alpha + 2k\pi| < \delta$.

□

Corolario 9.10.15. *Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, y sea $C_{\text{per}}(\mathbb{R}, \Bbbk)$ es espacio vectorial de todas las funciones continuas $f : \mathbb{R} \to \Bbbk$ tales que $f(t + 2\pi) = f(t)$ para todo $t \in \mathbb{R}$.*

(*i*) *Para cada $n \in \mathbb{Z}$ consideremos la función*

$$e_n : t \in \mathbb{R} \mapsto e^{int} \in \mathbb{C}.$$

El subespacio de $C_{\text{per}}(\mathbb{R}, \mathbb{C})$ generado por $\{e_n : n \in \mathbb{Z}\}$ es una subálgebra densa.

(*ii*) *Para cada $n \in \mathbb{N}_0$ consideremos las funciones*

$$c_n : t \in \mathbb{R} \mapsto \cos nt \in \mathbb{R}, \qquad s_n : t \in \mathbb{R} \mapsto \operatorname{sin} nt \in \mathbb{R}.$$

El subespacio de $C_{\text{per}}(\mathbb{R}, \mathbb{R})$ generado por $\{c_n, s_n : n \in \mathbb{N}_0\}$ es una subálgebra densa.

Ejemplo 9.10.16. El subespacio de $C([0,2\pi],\mathbb{C})$ generado por las funciones

$$\phi_n(t) = e^{int}$$

es denso.

Ejercicio 9.10.17. Muestre que una función continua $f : [0,2\pi] \to \mathbb{C}$ tal que

$$\int_0^{2\pi} f(t)e^{int}\,\mathrm{d}t = 0$$

para todo $n \in \mathbb{Z}$ es idénticamente nula.

Ejemplo 9.10.18. Sean X e Y dos espacios métricos compactos. El subespacio de $C(X \times Y, \Bbbk)$ generado por las funciones de la forma

$$(x,y) \in X \times Y \mapsto f(x)g(y) \in \Bbbk$$

con $f \in C(X,\Bbbk)$ y $g \in C(Y,\Bbbk)$ es denso.

Ejemplo 9.10.19. Sea X el espacio real normado de las funciones continuas $f : [0,+\infty) \to \mathbb{R}$ tales que $\lim_{t\to\infty} f(x) = 0$ dotado de la métrica de la convergencia uniforme. Queremos mostrar que

el conjunto de las funciones de la forma $e^{-x}p(x)$ con $p \in \mathbb{R}[X]$ es denso en X.

La subálgebra generada por las funciones de la forma $e^{-nx}x^m$ con $m, n \in \mathbb{N}$, es densa.

Capítulo 10

Conexión

10.1. Espacios conexos

Definición 10.1.1. Sea X un espacio métrico.

- Una ***separación*** de X es un par ordenado (A, B) de subconjuntos abiertos de X tal que $A \cap B = \varnothing$ y $A \cup B = X$, y decimos que esa separación es ***trivial*** si alguno de A o B es vacío.
- Decimos que el espacio X es ***conexo*** si no es vacío y toda separación de X es trivial.

Notemos que si (A, B) es una separación de un espacio X e Y es un subespacio de X, entonces $(A \cap Y, B \cap Y)$ es una separación de Y. Usaremos esto varias veces en lo que sigue.

Hay varias formas útiles de ver la noción de conexión.

Proposición 10.1.2. *Sea X un espacio métrico. Las siguientes afirmaciones son equivalentes:*

(*a*) *El espacio X es conexo.*
(*b*) *Hay exactamente dos subconjuntos de X abiertos y cerrados.*
(*c*) *Toda función continua $X \to \{0,1\}$ es constante y X no es vacío.*

Aquí vemos al conjunto $\{0,1\}$ como un espacio métrico dotado de su métrica discreta.

Demostración. (*a*) $\Rightarrow$ (*b*) Supongamos que X es conexo. Es claro que $\varnothing$ y X son dos subconjuntos de X abiertos y cerrados, y son distintos. Por otro lado, si A es un subconjunto de X que es abierto y cerrado, entonces $B \coloneqq X \setminus A$ también es un abierto de X, y es $A \cup B = X$ y $A \cap B = \varnothing$: como X es conexo alguno de A o B es vacío, así que o $A = \varnothing$ o $A = X$. Vemos así que los únicos subconjuntos

abiertos y cerrados de X son $\varnothing$ y X, y que son distintos.

$(b) \Rightarrow (c)$ Supongamos que hay exactamente dos subconjuntos de X abiertos y cerrados. Es claro que X no es vacío, ya que si lo fuera tendría exactamente un subconjunto, contra la hipótesis, y esto implica que los subconjuntos abiertos y cerrados de X son X y $\varnothing$. Sea, por otro lado, $f : X \to \{0,1\}$ una función continua. Como $\{0\}$ es un subconjunto abierto y cerrado de $\{0,1\}$, el conjunto $f^{-1}(0)$ es abierto y cerrado en X: es entonces igual a X o a $\varnothing$, y esto nos dice que la función f es constante de valor 0 o 1, respectivamente.

$(c) \Rightarrow (a)$ Supongamos que X no es vacío y que toda función continua $X \to \{0,1\}$ es constante, y sean A y B dos subconjuntos abiertos de X tales que $A \cup B = X$ y $A \cap B = \varnothing$. Ahora bien, La función característica $\chi_A : X \to \{0,1\}$ del subconjunto A es continua: en efecto, para verlo es suficiente con mostrar que la preimagen por χ_A de toda bola abierta de $\{0,1\}$ es abierta en X, y hay en $\{0,1\}$ tres bolas, $\{0\}$, $\{1\}$ y $\{0,1\}$, preimágenes son A, B y X, respectivamente. La hipótesis nos dice entonces que esta función χ_A es constante, y esto significa que o bien $A = \varnothing$ o bien $A = X$, de manera que $B = \varnothing$. Vemos así que el espacio X es conexo. □

Podemos enunciar la tercera afirmación de esta proposición de una forma un poco diferente:

Ejercicio 10.1.3. Muestre que un espacio métrico X es conexo si y solamente si hay exactamente dos funciones continuas $X \to \{0,1\}$.

Es fácil dar ejemplos de espacios que *no* son conexos, ya que para probar que un espacio no es conexo es suficiente con exhibir una separación no trivial.

Ejemplo 10.1.4. El subespacio $X \coloneqq (0,1) \cup (1,2)$ de $\mathbb{R}$ no es conexo, ya que $\big((0,1),(1,2)\big)$ es una separación no trivial de X.

Ejemplo 10.1.5. Sea X un conjunto. Si vemos a X como un espacio métrico con respecto a su métrica discreta, entonces

X es conexo si y solamente si tiene exactamente un elemento.

En efecto, toda función $X \to \{0,1\}$ es continua y el cardinal del conjunto de las funciones continuas $X \to \{0,1\}$ es $2^{|X|}$, que es 2 si y solamente si $|X| = 1$.

Es más difícil exhibir ejemplos de espacios conexos, ya que para ello tenemos que mostrar que no admiten ninguna separación no trivial. Hagámoslo en un ejemplo importante.

Proposición 10.1.6. *$\mathbb{R}$ es un espacio métrico conexo.*

Demostración. Sea (A, B) una separación de $\mathbb{R}$ y, sin pérdida de generalidad, supongamos que $0 \in A$. Para probar que el espacio $\mathbb{R}$ es conexo es suficiente que mostremos que $A = \mathbb{R}$, ya que en ese caso es $B = \varnothing$.

Como $0 \in A$ y A es abierto en $\mathbb{R}$, sabemos que el conjunto

$$R \coloneqq \{r \in (0, +\infty) : [0, r) \subseteq A\}$$

no es vacío. Supongamos por un momento que R está acotado superiormente, de manera que podemos considerar el número $\rho \coloneqq \sup R$. Notemos que $[0, \rho) \subseteq A$. En efecto, si t es un elemento de $[0, \rho)$, entonces $t < \rho$ y la definición de ρ nos dice que hay un elemento r de R tal que $t < r$, de manera que $t \in [0, r) \subseteq A$. Por otro lado, ρ no pertenece a A: en caso contrario, como A es abierto, habría un número positivo ε tal que $(\rho - \varepsilon, \rho + \varepsilon) \subseteq A$ y, por lo tanto, tal que $[0, \rho + \varepsilon) \subseteq A$ y, en consecuencia, $\sup R < \rho + \varepsilon \in R$, lo que es absurdo. Se sigue de esto que ρ pertenece a B y, como B es abierto, que hay un número positivo δ menor que ρ tal que $(\rho - \delta, \rho + \delta) \subseteq B$. Esto es absurdo, ya que en ese caso es

$$\rho - \delta/2 \in [0, \rho) \cap (\rho - \delta, \rho + \delta) \subseteq A \cap B = \varnothing.$$

Esta contradicción provino de haber supuesto que el conjunto R está acotado superiormente, así que no lo esta. Esto nos permite concluir que $[0, +\infty) \subseteq A$. En efecto, si t es un elemento de $[0, +\infty)$, entonces t no es una cota superior para R y hay un elemento r de R tal que $t < r$ y, por lo tanto, $t \in [0, r) \subseteq A$.

De la misma forma, considerando el conjunto $L \coloneqq \{l \in (0, +\infty) : (-l, 0] \subseteq A$, podemos mostrar que $(-\infty, 0] \subseteq A$ y, en definitiva, que $A = \mathbb{R}$, como queremos. □

Veamos ahora algunas formas de construir espacios conexos. Primero, observemos que la conexión es preservada por las funciones continuas.

Proposición 10.1.7. *Sean X e Y dos espacios métricos y sea $f : X \to Y$ una función continua. Si X es conexo, entonces $f(X)$ también lo es.*

Demostración. En efecto, si $f(X)$ no es conexo, de manera que hay una separación no trivial (A, B) de $f(X)$, entonces $(f^{-1}(A), f^{-1}(B))$ es una separación no trivial de X y, por lo tanto, X tampoco es conexo. □

Usando esto podemos producir más ejemplos.

Ejemplo 10.1.8. Sean a y b dos números reales tales que $a < b$. Cada una de las funciones que aparecen listadas en la primera columna de la siguiente tabla es continua, así que las correspondientes

imágenes, que listamos en la segunda columna, son espacios conexos.

f	$f(\mathbb{R})$
$t \in \mathbb{R} \mapsto a + \frac{b-a}{\pi} \arctan t \in \mathbb{R}$	(a, b)
$t \in \mathbb{R} \mapsto a + 2\frac{b-a}{\pi} \lvert\arctan t\rvert \in \mathbb{R}$	$[a, b)$
$t \in \mathbb{R} \mapsto b - 2\frac{b-a}{\pi} \lvert\arctan t\rvert \in \mathbb{R}$	$(a, b]$
$t \in \mathbb{R} \mapsto \text{máx}\{a, \text{mín}\{b, t\}\} \in \mathbb{R}$	$[a, b]$
$t \in \mathbb{R} \mapsto a + \lvert t\rvert \in \mathbb{R}$	$[a, +\infty)$
$t \in \mathbb{R} \mapsto a - \lvert t\rvert \in \mathbb{R}$	$(-\infty, a]$
$t \in \mathbb{R} \mapsto a + e^x \in \mathbb{R}$	$(a, +\infty)$
$t \in \mathbb{R} \mapsto a - e^x \in \mathbb{R}$	$(-\infty, a)$

Esto prueba que todo intervalo no vacío de $\mathbb{R}$ es conexo.

Segundo, es claro que la unión de conjuntos conexos no es en general conexa, pero sí lo es si estos tienen dos a dos intersección no trivial.

Proposición 10.1.9. *Sea X un espacio métrico no vacío y sea $\mathscr{C}$ una familia de subconjuntos conexos de X tal que $X = \bigcup \mathscr{C}$. Si $C \cap D \neq \varnothing$ siempre que C y D son elementos de $\mathscr{C}$, entonces X es conexo.*

Demostración. Supongamos que se cumple la hipótesis de la proposición, sea (A, B) una separación de X y, sin pérdida de generalidad, supongamos que $A \neq \varnothing$. Pongamos $\mathscr{C}_A \coloneqq \{C \in \mathscr{C} : C \subseteq A\}$ y $\mathscr{C}_B \coloneqq \{C \in \mathscr{C} : C \subseteq B\}$. Para cada elemento C de $\mathscr{C}$ el par $(A \cap C, B \cap C)$ es una separación de C, así que, como C es conexo, alguno de $A \cap C$ o $B \cap C$ es vacío: esto nos dice que C está contenido o en A o en B, es decir, que C pertenece a $\mathscr{C}_A$ o a $\mathscr{C}_B$. Vemos así que $\mathscr{C} = \mathscr{C}_A \cup \mathscr{C}_B$.

Como A no es vacío, podemos elegir un elemento a de A y, en consecuencia, un elemento C_0 de $\mathscr{C}$ tal que $a \in C_0$. Como $A \cap C_0 \neq \varnothing$, tenemos que $C_0 \subseteq A$. Para cada elemento C de $\mathscr{C}_B$ tenemos que $\varnothing \neq C_0 \cap C \subseteq A \cap B = \varnothing$, y esto es absurdo: el conjunto $\mathscr{C}_B$ es, por lo tanto, vacío, y es $\mathscr{C} = \mathscr{C}_A$ y $X = \bigcup \mathscr{C} = \bigcup \mathscr{C}_A \subseteq A$, de manera que $A = X$ y $B = \varnothing$. Esto prueba que toda separación de X es trivial, así que X es conexo. □

Usando este resultado podemos construir una relación importante.

Lema 10.1.10. *Sea X un espacio métrico. El conjunto*

$$R_X \coloneqq \{(x, y) \in X \times X : \textit{hay un subconjunto conexo } C \textit{ de } X \textit{ tal que } x, y \in C\}$$

es una relación de equivalencia en X.

Demostración. Verifiquemos que la relación R_X satisface las condiciones para ser una relación de equivalencia.

- Si x es un elemento x de X, entonces el subconjunto $\{x\}$ de X es conexo y $x \in \{x\}$, así que $(x, x) \in R_X$.
- Si (x, y) es un elemento de R_X, entonces hay un subconjunto conexo C de X tal que x e y están en C y, por lo tanto, es claro que $(y, x) \in R_X$.
- Sean (x, y) e (y, z) dos elementos de R_X, de manera que hay subconjuntos conexos C y D de X tales que $x, y \in C$ e $y, z \in D$. Como $y \in C \cap D$, la intersección $C \cap D$ no es vacía y la Proposición 10.1.9 nos dice que $C \cup D$ es un subconjunto conexo de X: como $x, z \in C \cup D$, vemos que $(x, z) \in R_X$. □

Como toda relación de equivalencia, esta relación determina una partición del conjunto en el que está definida.

Definición 10.1.11. Las ***componentes conexas*** de un espacio métrico X son las clases de equivalencia de la relación de equivalencia R_X del Lema 10.1.10.

Podemos describir las componentes conexas de un espacio de la siguiente manera alternativa:

Proposición 10.1.12. *Sea X un espacio métrico. Todo elemento de X está contenido en un subconjunto conexo maximal, que es su componente conexa.*

Notemos que esto implica, en particular, que dos subconjuntos conexos maximales de un espacio métrico que son distintos son, de hecho, disjuntos.

Demostración. Sea x un elemento de X y sea C la componente conexa de x en X. Para probar la proposición es tenemos que mostrar que C es un subconjunto conexo de X que contiene a todo subconjunto conexo de X que contiene a x. Sea R_X la relación de equivalencia del Lema 10.1.10.

- Para cada elemento y de C es $(x, y) \in R_X$, así que hay un subconjunto conexo C_y de X tal que $x, y \in C_y$ y, como C_y es conexo, es claro que $(x, z) \in R_X$ para todo elemento z de C_y, de manera que $C_y \subseteq C$. Esto implica que $C \subseteq \bigcup_{y \in C} C_y \subseteq C$, de manera que $C = \bigcup_{y \in C} C_y$. Por otro lado, la intersección de dos elementos cualesquiera de $\{C_y : y \in C\}$ no es vacía, ya que contiene a x, y entonces la Proposición 10.1.9 nos dice que la unión $C = \bigcup_{y \in C} C_y$ es conexa.
- Si D es un subconjunto conexo de X tal que $x \in D$, entonces para todo $y \in D$ es $(x, y) \in R_X$, así que $D \subseteq C$. □

Una caso extremo posible e importante es aquel en el que todas las componentes conexas de un espacio son tan chicas como pueden serlo.

Definición 10.1.13. Un espacio métrico es ***totalmente desconexo*** si todas sus componentes conexas tienen exactamente un elemento.

Como las componentes conexas de un espacio son conexas, para ver que un espacio es totalmente desconexo es suficiente con mostrar que sus subconjuntos conexos tienen todos un solo punto.

Ejemplo 10.1.14. En el Ejemplo 10.1.5 vimos que un espacio con métrica discreta es conexo si y solamente si tiene exactamente un elemento, y esto implica inmediatamente que un espacio con métrica discreta es totalmente desconexo. Más generalmente,

> *un espacio métrico discreto — esto es, que no tiene puntos de acumulación — es totalmente discreto.*

En efecto, sea X un espacio métrico discreto, sea C un subconjunto de X con más de un elemento, y sea x un elemento de C. Como x no es un punto de acumulación de X, hay un número positivo r tal que $B_r(x) = \{x\}$ y entonces $X \setminus \{x\} = B \setminus \overline{B}_{r/2}(x)$ es un abierto de X y $(\{x\}, C \setminus \{x\})$ es una separación no trivial de C, de manera que C no es conexo.

No es necesario que un espacio sea discreto para que sea totalmente desconexo. El ejemplo más sencillo de esto es el siguiente.

Ejemplo 10.1.15. El subespacio $\mathbb{Q}$ de $\mathbb{R}$ es totalmente desconexo. En efecto, si C es un subconjunto de $\mathbb{Q}$ con más de un punto y x e y son dos elementos de C tales que $x < y$, entonces hay un número irracional z tal que $x < z < y$ y $(C \cap (-\infty, z), C \cap (z, +\infty))$ es una separación no trivial de C.

De hecho, un subconjunto totalmente desconexo puede ser inclusive perfecto. El ejemplo más famoso de esto es el conjunto de Cantor.

Ejemplo 10.1.16. Sea $X \coloneqq \{0,1\}^{\mathbb{N}_0}$. Si $x = (x_n)_{n\geq 0}$ es un elemento de X, entonces la serie de términos positivos

$$\sum_{n=0}^{\infty} \frac{2x_n}{3^{n+1}}$$

está acotada término a término por la serie $\sum_{n=0}^{2} /3^{n+1}$, y esta última converge y suma 1: esto nos dice que la primera serie también converge y que su suma pertenece a $[0,1]$. Vemos así que hay una función

$$\phi : (x_n)_{n\geq 0} \in X \mapsto \sum_{n=0}^{\infty} \frac{2x_n}{3^{n+1}} \in [0,1].$$

El ***conjunto de Cantor*** es la imagen de ϕ. Escribámoslo C. Queremos probar que

> *el conjunto de Cantor es perfecto y totalmente desconexo.*

Es fácil ver que la función $D : X \times X \to \mathbb{R}$ tal que

$$D(x, y) = \sum_{n=0}^{\infty} \frac{|x_n - y_n|}{3^{n+1}}$$

cualesquiera sean los elementos $x = (x_n)_{n\geq 0}$ e $y = (y_n)_{n\geq 0}$ de X es una métrica,

Capítulo 11

El teorema de Baire

11.1. El teorema de Baire

El objetivo de este capítulo es estudiar el siguiente teorema y algunas de sus consecuencias. Es conocido como el *Teorema de Baire*, porque René-Louis Baire lo probó para el espacio métrico $\mathbb{R}^n$ en [Bai1899], aunque William Fogg Osgood habría probado antes, en 1987, en el caso particular en el que $n = 1$. La versión general que damos aquí, para espacios métricos completos arbitrarios, fue dada por Felix Hausdorff en 1914.

Proposición 11.1.1. *La intersección de una familia contable de abiertos densos de un espacio métrico completo es densa*

Demostración. Sea X un espacio métrico completo, sea $\mathscr{U}$ una familia contable de abiertos densos de X, sea $v : \mathbb{N}_0 \to \mathscr{U}$ una función sobreyectiva, y para cada $n \in \mathbb{N}_0$ escribamos $U_n \coloneqq v(n)$. Queremos probar que la intersección $\bigcap \mathscr{U}$ es un subconjunto denso de X y para ello elegimos un subconjunto abierto y no vacío V de X y mostramos que $V \cap \bigcap \mathscr{U} \neq \varnothing$.

Sea $\mathscr{S}$ el conjunto de todos los elementos (k, x, r) de $\mathbb{N}_0 \times X \times (0, +\infty)$ tales que $r < 1/(k+1)^2$. Si (k, x, r) es un elemento de $\mathscr{S}$, entonces el conjunto

$$R(k, x, r) = \{(k+1, y, s) \in \mathscr{S} : s < r/2, \overline{B}_s(y) \subseteq U_k \cap B_{r/2}(x)\}$$

no es vacío. En efecto, como U_k es un abierto denso de X, la intersección $U_k \cap B_{r/2}(x)$ es abierta y no vacía, así que hay un número positivo s y un elemento y de X tales que $\overline{B}_s(y) \subseteq U_k \cap B_{r/2}(x)$ y $s < r/2$. Esto nos dice que $(k+1, y, s)$ es un elemento de $R(k, x, r)$.

Ahora bien, como el abierto U_0 es denso, la intersección $V \cap U_0$ es abierta y no vacía, así

que hay un punto x_0 de X y un número positivo r_0 tales que $\overline{B}_{r_0}(x) \subseteq V \cap U_0$ y $r_0 < 1$. La terna $(0, x_0, r_0)$ es un elemento del conjunto $\mathscr{S}$ y, de acuerdo al *Principio de elección dependiente*, hay una función $f : \mathbb{N}_0 \to \mathscr{S}$ tal que $f(0) = (0, x_0, r_0)$ y $f(n+1) \in R(f(n))$ para todo $n \in \mathbb{N}_0$. Si para cada $n \in \mathbb{N}_0$ escribimos (k_n, x_n, r_n) a la terna $f(n)$, entonces tenemos que para todo $n \in \mathbb{N}_0$ vale $k_n = n$, $\overline{B}_{r_{n+1}}(x_{n+1}) \subseteq U_n \cap B_{r_n/2}(x_n)$, y $r_{n+1} < r_n/2$. En particular, es $d(x_{n+1}, x_n) < r_n < r_0/2^n$ para todo $n \in \mathbb{N}_0$ y, como la serie $\sum_{n=0}^{\infty} r_0/2^n$ converge, la sucesión $(x_n)_{n\geq 0}$ satisface la condición de Cauchy. El espacio X es completo, así que esa sucesión converge a un punto ξ de X.

Sea m un elemento de $\mathbb{N}_0$. Para cada $d \in \mathbb{N}_0$ tenemos que

$$d(x_{m+d}, x_m) \leq \sum_{i=0}^{d-1} d(x_{m+i+1}, x_{m+i}) < \sum_{i=0}^{d-1} \frac{r_{m+i}}{2} < \sum_{i=0}^{d-1} \frac{r_m}{2^{i+1}} = \frac{2^d - 1}{2^d} \cdot r_m < r_m.$$

Esto nos dice que la sucesión $(x_{m+d})_{d\geq 0}$ toma valores en el conjunto cerrado $\overline{B}_{r_m}(x_m)$, así que su límite, que es ξ está en él. Si $m = 0$, esto nos dice que $\xi \in \overline{B}_{r_0}(x_0) \subseteq V \cap U_0$, así que ξ está en V y en U_0. Por otro lado, si m es un entero positivo esto nos dice que $\xi \in \overline{B}_{r_m}(x_m) \subseteq U_m$. Vemos así que ξ pertenece a $V \cap \bigcap \mathscr{U}$, así que este conjunto no es vacío. □

Hay una forma alternativa de esta proposición que muchas veces es más cómoda en las aplicaciones. Para presentarla necesitamos la siguiente definición.

Definición 11.1.2. Un subconjunto de un espacio métrico es ***nunca denso*** si el interior de su clausura es vacío.

La siguiente proposición explica la elección de este término.

Proposición 11.1.3. *Sea X un espacio métrico y sea M un subconjunto de X. Las siguientes afirmaciones son equivalente:*

(*a*) *El conjunto M es nunca denso.*
(*b*) *El conjunto $X \setminus \overline{M}$ es denso.*
(*c*) *Para todo abierto no vacío U de X hay un abierto no vacío V de U tal que $V \cap M = \varnothing$.*

Demostración. (*a*) ⇒ (*b*) Si M es nunca denso, entonces $\overline{M}^{\circ}$ es vacío, aquí que todo abierto no vacío de X interseca a $X \setminus \overline{M}$ y, por lo tanto, este último conjunto es denso en X.

(*b*) ⇒ (*c*) Si el conjunto $X \setminus \overline{M}$ es denso en X y U es un abierto no vacío de X, entonces $V \coloneqq U \cap (X \setminus \overline{M})$ es un subconjunto abierto no vacío de U y claramente $V \cap M = \varnothing$.

(*c*) ⇒ (*a*) Supongamos que M no es nunca denso, de manera que hay un abierto no vacío U de X tal que $U \subseteq \overline{M}$. Para todo abierto no vacío V de U tenemos entonces que $V \subseteq \overline{M}$ y, por lo tanto, que $V \cap M \neq \varnothing$. □

Notemos que que un subconjunto de un espacio métrico sea nunca denso depende del espacio

ambiente en el que lo consideramos. Así, por ejemplo, ningún subconjunto no vacío nunca denso de un espacio métrico es nunca denso en si mismo.

Ejercicio 11.1.4. Sea X un espacio métrico.

(a) Pruebe que todo subconjunto de un subconjunto nunca denso de X es nunca denso.

(b) Pruebe que si A es un subconjunto de X y $\mathscr{F}$ es una familia de subconjuntos nunca densos de X tal que $A \subset \bigcup \mathscr{F}$, entonces hay una familia $\mathscr{F}'$ de subconjuntos nunca densos de A que tiene cardinal menor o igual que el de $\mathscr{F}$ y tal que $A = \bigcup \mathscr{F}'$.

Ejercicio 11.1.5. Sea X un espacio métrico y sea x un elemento de X. El conjunto $\{x\}$ es nunca denso si y solamente si x es un punto de acumulación de X.

Usando la noción de conjuntos nunca densos podemos dar una forma alternativa al *Teorema de Baire* 11.1.1.

Proposición 11.1.6. *La unión de una familia contable de subconjuntos nunca densos de un espacio métrico completo tiene interior vacío.*

Muchas veces solo necesitamos saber que la unión de una tal familia es un subconjunto propio del espacio, de hecho.

Demostración. Sea X un espacio métrico completo y sea $\mathscr{M}$ una familia contable de subconjuntos nunca densos de X. Para cada elemento M de $\mathscr{M}$ es $\overline{M}^\circ = \varnothing$, así que $X \setminus \overline{M}$ es un abierto denso de X. El conjunto $\mathscr{U} \coloneqq \{X \setminus \overline{M} : M \in \mathscr{M}\}$ es entonces una familia contable de abiertos densos de X y la Proposición 11.1.1 nos dice que el conjunto

$$\bigcap_{M \in \mathscr{M}} (X \setminus \overline{M}) = X \setminus \bigcup_{M \in \mathscr{M}} \overline{M}$$

es denso en X, de manera que el interior de $\bigcup_{M \in \mathscr{M}} \overline{M}$ y, por lo tanto, el de $\bigcup_{M \in \mathscr{M}} M$ es vacío. □

Las Proposiciones 11.1.1 y 11.1.6 tienen muchas consecuencias directas importantes. Demos algunas.

Corolario 11.1.7. *Un espacio métrico no vacío completo y sin puntos aislados no es contable.*

Demostración. Sea X un espacio métrico sin puntos aislados que es numerable. De acuerdo al resultado del Ejercicio 11.1.5, para todo elemento x de X el conjunto $\{x\}$ es nunca denso, y la Proposición 11.1.6 nos dice entonces que el conjunto $\bigcup_{x \in X} \{x\}$ tiene interior vacío. Como ese conjunto es X, vemos que debe ser $X = \varnothing$. Esto prueba el corolario. □

Notemos que este corolario da una demostración de que $\mathbb{R}$ no es contable. Por otro lado, el

corolario implica que no hay ninguna métrica sobre $\mathbb{Q}$ que sea topológicamente equivalente a la métrica usual y con respecto a la cual $\mathbb{Q}$ es completo. Decimos por ello que $\mathbb{Q}$ no es ***completamente metrizable***.

Corolario 11.1.8. *Un espacio métrico completo numerable posee infinitos puntos aislados.*

Demostración. Sea X un espacio métrico completo, sea A el conjunto de los puntos aislados de X, y supongamos que este conjunto A es finito. Si a es un elemento de A, entonces el conjunto $\{a\}$ es abierto: esto implica que $A = \bigcup_{a\in A}\{a\}$ es un abierto de X y, por lo tanto, que $Y \coloneqq X \setminus A$ es cerrado en X y, por lo tanto, completo.

Sea y un elemento de Y y sea U un entorno de y. Como y no es un punto aislado de X, U contiene infinitos puntos de X y, en particular, como A es finito, infinitos puntos de Y. Vemos así que y es un punto de acumulación de Y y, por lo tanto, que Y no posee puntos aislados. Como Y es completo, el Corolario 11.1.7 nos dice que Y no es contable y, en consecuencia, X tampoco. Esto prueba el corolario. □

Corolario 11.1.9. *Una familia contable $\mathscr{F}$ de subconjuntos cerrados de un espacio métrico completo X tal que $X = \bigcup \mathscr{F}$ tiene un elemento con interior no vacío.*

Demostración. Sea X un espacio métrico completo. Si $\mathscr{F}$ es una familia contable de subconjuntos cerrados de X con interior vacío, entonces la Proposición 11.1.6 nos dice que el interior de $\bigcup \mathscr{F}$ es vacío, así que esta unión no puede ser igual a X. □

El *Teorema de Baire* muchas veces es llamado el *Teorema de la categoría de Baire*. La palabra *categoría* en ese nombre hace referencia a la siguiente clasificación de los subconjuntos de un espacio métrico.

Definición 11.1.10. Un subconjunto de un espacio métrico es ***magro*** o ***de primera categoría*** si es la unión de una familia contable de subconjuntos nunca densos, y es ***de segunda categoría*** en caso contrario.

Con este vocabulario, podemos enunciar el *Teorema de Baire* en la siguiente forma.

Corolario 11.1.11. *Un espacio métrico completo es de segunda categoría.*

Demostración. Esto es consecuencia inmediata de la Proposición 11.1.6. □

Ejercicio 11.1.12. Sea X un espacio métrico. Pruebe que un subconjunto de un subconjunto magro en X es magro en X, y que la unión de una familia contable de subconjuntos magros de X es un

subconjunto magro de X.

Ejercicio 11.1.13. Muestre que si $\mathscr{U}$ es una familia contable de abiertos de un espacio métrico completo, entonces

$$\left(\overline{\bigcap_{U\in\mathscr{U}} U}\right)^\circ = \left(\bigcap_{U\in\mathscr{U}} \overline{U}\right)^\circ.$$

11.2. Algunas aplicaciones

11.2.1. Funciones reales

El teorema de Baire es útil para mostrar, por un lado, que ciertas funciones no pueden ser «muy malas» en distintos sentidos y, al mismo tiempo, que hay muchas funciones «muy malas». El siguiente es un ejemplo de lo primero.

Ejemplo 11.2.1.1. Es fácil ver que la función

$$t \in \mathbb{R} \mapsto \left\{ \begin{cases} t^2 \operatorname{sin} \frac{1}{t^2} & \text{si } t \neq 0; \\ 0 & \text{si } t = 0 \end{cases} \right\} \in \mathbb{R}$$

es derivable en cada punto de su dominio y que

$$f'(t) = \begin{cases} 2t \operatorname{sin} \dfrac{1}{t^2} - \dfrac{2}{t} \cos \dfrac{1}{t^2} & \text{si } t \neq 0; \\ 0 & \text{si } t = 0. \end{cases}$$

Es claro que f' no es acotada en ningún entorno de 0 o, equivalentemente, que

$$\limsup_{t\to 0} |f'(t)| = +\infty.$$

Queremos probar que esto no puede ocurrir en demasiados puntos de $\mathbb{R}$. Sea $f : \mathbb{R} \to \mathbb{R}$ una función que es derivable en cada punto de su dominio, de manera que podemos considerar su derivada $f' : \mathbb{R} \to \mathbb{R}$, y consideremos el conjunto

$$U \coloneqq \left\{ t \in \mathbb{R} : \limsup_{s\to t} |f'(s)| < +\infty \right\}.$$

Este conjunto es abierto. En efecto, supongamos que t_0 es un elemento de U, de manera que

$$\inf_{r>0} \sup\{|f(t)| : t \in (t_0 - r, t_0 + r)\} = \limsup_{s\to t_0} |f'(s)| < +\infty$$

y, por lo tanto, hay un numero positivo r tal que el conjunto $\{|f'(t)| : t \in (t_0 - r, t_0 + r)\}$ está acotado superiormente. Si t_1 es un elemento cualquiera de $(t_0 - r, t_0 + r)$, entonces hay un número positivo s tal que $(t_1 - s, t_1 + s) \subseteq (t_0 - r, t_0 + r)$ y es claro que el conjunto $\{|f'(t)| : t \in (t_1 - s, t_1 + s)\}$ está acotado superiormente, de manera que

$$\limsup_{s \to t_1} |f'(s)| = \inf_{r>0} \sup\{|f'(t)| : t \in (t_1 - r, t_1 + r)\} < \infty$$

y, por lo tanto, t_1 pertenece a U. Vemos así que $(t_0 - r, t_0 + r) \subseteq U$ y esto prueba que, como dijimos, el conjunto U es abierto.

Sea, por otro lado, I un intervalo cerrado de $\mathbb{R}$ y para cada un entero positivo k consideremos el conjunto

$$F_k \coloneqq \{t \in I : |f(t) - f(s)| \leq k|t - s| \text{ para todo } s \in I\}.$$

Sea $(t_n)_{n \geq 0}$ una sucesión con valores en F_k que converge en $\mathbb{R}$ a un punto τ. Como $F_k \subseteq I$ e I es cerrado, sabemos que τ pertenece a I. Por otro lado, si s es un elemento de I, entonces tenemos que $|f(t_n) - f(s)| \leq k|t_n - s|$ para todo $n \in \mathbb{N}_0$, así que tomando límites a ambos lados de la desigualdad y usando el hecho de que la función f es continua vemos que $|f(\tau) - f(s)| \leq k|\tau - s|$. Esto es así cualquiera sea el elemento s de I, así que $\tau \in F_k$. Esto prueba que el conjunto F_k es cerrado en I. Afirmamos que además es

$$\bigcup_{k=1}^{\infty} F_k = I. \tag{11.1}$$

Para verlo, sea t un elemento de I. Como f es derivable en t, hay un número positivo δ tal que

$$s \in (t - \delta, t + \delta) \cap I, s \neq t \implies \left| \frac{f(s) - f(t)}{s - t} - f'(t) \right| < 1.$$

Esto implica que para cada elemento s de $(t - \delta, t + \delta) \cap I$ distinto de t es

$$f'(t) - 1 < \frac{f(s) - f(t)}{s - t} < f'(t) + 1$$

y, en consecuencia,

$$\left| \frac{f(s) - f(t)}{s - t} \right| < \text{máx}\{|f'(t) + 1|, |f'(t) - 1|\}.$$

Por otro lado, la función

$$s \in I \setminus (t - \delta, t + \delta) \mapsto \left| \frac{f(s) - f(t)}{s - t} \right| \in \mathbb{R}$$

es continua y su dominio es un subconjunto compacto de $\mathbb{R}$, así que hay un número positivo K tal que

$$\left|\frac{f(s)-f(t)}{s-t}\right| \leq K \quad \text{para todo } s \in I \setminus (t-\delta, t+\delta).$$

Finalmente, hay un entero k tal que $k \geq \max\{|f'(t)+1|, |f'(t)-1|, K\}$ y lo que hemos hecho nos dice que

$$|f(s)-f(t)| \leq k|s-t| \quad \text{para todo } s \in I.$$

Vemos así que t pertenece al conjunto F_k. Esto prueba la igualdad (11.1).

Como I es un espacio métrico completo, la igualdad (11.1) que acabamos de probar y el *Teorema de Baire* 11.1.6 nos permiten concluir que hay un entero positivo m tal que el cerrado F_m tiene interior no vacío y, por lo tanto, hay un intervalo abierto no vacío (a,b) contenido en F_m y, en particular, en I.

Si t es un elemento de (a,b), entonces tenemos que $|f(t)-f(s)| \leq m|t-s|$ para todo $s \in (a,b)$, y esto implica inmediatamente que $|f'(t)| \leq m$. Vemos así que todos los puntos de (a,b) tienen un entorno sobre el cual la función f' está acotada y, por lo tanto, que $(a,b) \subseteq U$: esto nos dice que $U \cap I \neq \varnothing$. Como I es un intervalo cerrado arbitrario de $\mathbb{R}$, podemos concluir que el abierto U es denso en $\mathbb{R}$.

Hemos probado que

> *si una función $f : \mathbb{R} \to \mathbb{R}$ es derivable, entonces el conjunto de puntos de $\mathbb{R}$ que poseen un entorno en el que f' es acotada es abierto y denso en $\mathbb{R}$.*

o, equivalentemente, que el conjunto de puntos que no tienen ningún entorno en el que la derivada de f está acotada es un cerrado nunca denso de $\mathbb{R}$.

Usemos ahora el *Teorema de Baire* para probar que hay muchas funciones que no tienen derivada en ningún punto de su dominio.

Ejemplo 11.2.1.2. Escribamos C al espacio métrico de todas las funciones continuas $[0,1] \to \mathbb{R}$ dotado de la métrica de la convergencia uniforme, y para cada entero positivo n consideremos el subconjunto A_n de C de todas las funciones continuas $f : [0,1] \to \mathbb{R}$ tales que

> *para todo $t \in [0,1]$ hay un elemento s de $[0,1]$ tal que $|f(s)-f(t)| > n|s-t|$.*

Fijemos un entero positivo n y probemos que el conjunto A_n es abierto y denso en C.

-

11.2.2. Uniones de subespacios y de subconjuntos algebraicos

Recordemos el siguiente hecho del *Álgebra lineal*

Lema 11.2.2.1. *Sea $\Bbbk$ un cuerpo infinito. Un espacio vectorial sobre $\Bbbk$ no es la unión de un conjunto finito de subespacios propios.*

Demostración. Sea V in espacio vectorial sobre $\Bbbk$ y supongamos, para llegar a una contradicción, que hay conjunto finitos de subespacios propios de V cuya unión es V. Sea $\mathscr{U}$ uno entre todos ellos de cardinal mínimo, sea n ese cardinal, y sean $U_1, U_2, \ldots, U_n$ los elementos de $\mathscr{U}$.

Si fuera $U_1 \subseteq U_2 \cup \cdots \cup U_n$, entonces tendríamos que $V = U_1 \cup U_2 \cup \cdots \cup U_n = U_2 \cup \cdots \cup U_n$ y esto es imposible porque el conjunto $\{U_2, \ldots, U_n\}$ de subespacios propios de V tiene cardinal menor que n. Podemos elegir entonces un vector u en $U_1 \setminus (U_2 \cup \cdots \cup U_n)$. Por otro lado, como U_1 es un subespacio propio de V, podemos elegir un vector v en $V \setminus U_1$.

Consideremos el conjunto $L \coloneqq \{tu + v : t \in \Bbbk\}$.

- Si w es un elemento de $L \cap U_1$, entonces hay un escalar t en $\Bbbk$ tal que $tu + v = w \in U_1$: esto implica que $v = w - tu \in U_1$, lo que es absurdo. Vemos así que $L \cap U_1 = \varnothing$.
- Sea, por otro lado, i un elemento de $[\![2, n]\!]$ y supongamos que hay en la intersección $L \cap U_i$ dos elementos distintos w y w'. Hay entonces dos escalares t y t en $\Bbbk$ tales que $tu+v = w \in U_i$ y $t'u + v = w' \in U_i$, así que $(t - t')u = w - w' \in U_i$. Como $u \notin U_i$, esto implica que $t = t'$, pero esto es absurdo, ya que $tu + v = w \neq w' = t'u + v$. Vemos así que $|L \cap U_i| \leq 1$.

Estas dos observaciones nos permiten concluir que

$$\begin{aligned} |L| &= |L \cap (U_1 \cup U_2 \cup \cdots \cup U_n)| = |(L \cap U_1) \cup (L \cap U_2) \cup \cdots \cup (L \cap U_n)| \\ &\leq |L \cap U_1| + |L \cap U_2| + \cdots + |L \cap U_n| \leq 0 + 1 + \cdots + 1 = n - 1. \end{aligned}$$

Esto es absurdo, ya que la función $t \in \Bbbk \mapsto tu + v \in L$ es inyectiva y el cuerpo $\Bbbk$ es infinito. □

Ejercicio 11.2.2.2. Pruebe que, más generalmente, si V es un espacio vectorial sobre un cuerpo $\Bbbk$ que es la unión de una familia $\mathscr{U}$ posiblemente infinita de subespacios propios, entonces $|\mathscr{U}| \geq |\Bbbk| + 1$.

Usando el *Teorema de Baire* podemos obtener la siguiente extensión de este resultado.

Proposición 11.2.2.3. *Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$. Un espacio de Banach sobre $\Bbbk$ no es la unión de una familia contable de subespacios vectoriales cerrados y propios.*

En particular, claro, esto se aplica cuando el espacio de Banach es $\mathbb{R}^n$ o $\mathbb{C}^n$.

Demostración. Sea V un espacio de Banach sobre $\Bbbk$ y sea $\mathscr{U}$ una familia contable de subespacios vectoriales de V cerrados tal que $V = \bigcup \mathscr{U}$. La Proposición 11.1.6 nos dice que hay un elemento U de $\mathscr{U}$ que tiene interior no vacío. Hay entonces un vector x de V y un número positivo r tal que $B_r(x) \subseteq U$. Como claramente $x \in U$, esto implica que $B_r(0) = B_r(x) - x \subseteq U$ y, como U es un subespacio de V, que

$$V = \bigcup_{t\in(0,+\infty)} \big(t \cdot B_r(0)\big) \subseteq U,$$

de manera que, de hecho, $U = V$. Esto prueba la proposición. □

Ejercicio 11.2.2.4. Un ***subespacio afín*** de un espacio vectorial V es un subconjunto de la forma $x + U$ con $x \in V$ y U un subespacio de V — en otras palabras, un subespacio afín es una coclase de un subespacio vectorial. Pruebe que un espacio de Banach sobre $\mathbb{R}$ o sobre $\mathbb{C}$ no es la unión de una familia contable de subespacios afines cerrados y propios.

Veamos una «versión no lineal» de este resultado.

Definición 11.2.2.5. Sea $\Bbbk$ un cuerpo. Para cada subconjunto F de $\Bbbk[X_1, X_2, \ldots, X_n]$, el anillo de polinomios en variables $X_1, X_2, \ldots, X_n$ con coeficientes en $\Bbbk$, ponemos

$$V(F) = \{(x_1, x_2, \ldots, x_n) \in \Bbbk^n : f(x_1, x_2, \ldots, x_n) = 0 \text{ para todo elemento } f \text{ de } F\}.$$

Decimos que un subconjunto Z de $\Bbbk^n$ es ***algebraico*** si hay un subconjunto F de $\Bbbk[X_1, X_2, \ldots, X_n]$ tal que $Z = V(F)$.

Si el conjunto F es finito y $f_1, f_2, \ldots, f_k$ son sus elementos, entonces escribimos generalmente $V(f_1, f_2, \ldots, f_k)$ en lugar de $V(F)$. Un subconjunto de $\Bbbk^n$ es algebraico si es el conjunto de ceros comunes de los elementos de algún conjunto de polinomios.

Lema 11.2.2.6. *Sea n un número entero positivo, sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, y sea f un elemento del anillo de polinomios $\Bbbk[X_1, X_2, \ldots, X_n]$. El conjunto $V(f)$ es cerrado en $\Bbbk^n$, y su interior es vacío si f no es nulo.*

Demostración. Sea f un elemento de $\Bbbk[X_1, X_2, \ldots, X_n]$. La función

$$\phi : (x_1, x_2, \ldots, x_n) \in \Bbbk^n \mapsto f(x_1, x_2, \ldots, x_n) \in \Bbbk$$

es continua y $\{0\}$ es un subconjunto cerrado de $\Bbbk$, así que el $V(f)$, que coincide con $\phi^{-1}(\{0\})$, es un subconjunto cerrado de $\Bbbk^n$.

Vamos a probar que

si $V(f)$ tiene interior no vacío entonces f es nulo

procediendo por inducción con respecto al grado total de f. Esto es evidente si el grado total de f es 0, ya que en ese caso f es constante. Supongamos entonces que f tiene grado total positivo y que hay un punto x de $\mathbb{k}^n$ y un número positivo r tal que $f(y) = 0$ para todo elemento y de $B_r(x)$.

Sea $(e_1, e_2, \ldots, e_n)$ la base ordenada estándar de $\mathbb{k}^n$. Si i es un elemento cualquiera de $[\![n]\!]$ e y es un elemento de $B_r(x)$, entonces

$$\frac{\partial f}{\partial x_i}(y) = \lim_{h \to 0} \frac{f(y + he_i) - f(y)}{h} = 0,$$

ya que $f(y) = 0$ y $f(y + he_i) = 0$ siempre que h es un número tal que $|h| < r - d(x, y)$, porque en ese caso $y + te_i \in B_r(x)$: esto nos dice que la derivada $\partial f/\partial x_i$, que es un polinomio de grado menor que el de f, se anula sobre $B_r(x)$, de manera que el interior de $V(\partial f/\partial f_i)$ no es vacío y, en vista de la hipótesis inductiva evidente, $\partial f/\partial f_i = 0$. Vemos así que el polinomio f tiene todas sus derivadas parciales primeras nulas, así que f es constante. Como el conjunto $V(f)$ no es vacío, hay puntos de $\mathbb{k}^n$ donde f se anula y, por lo tanto, f es el polinomio nulo. Esto completa la inducción y, en definitiva, la demostración del lema. □

Apoyándonos en este lema podemos probar el siguiente resultado, al que podemos ver como una extensión polinomial de la Proposición 11.2.2.3.

Proposición 11.2.2.7. *Sea n un número entero positivo y sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$. El conjunto $\mathbb{k}^n$ no es la unión de una familia contable de subconjuntos algebraicos propios de $\mathbb{k}^n$.*

Demostración. En vista de la Proposición 11.1.6, para probar esto es suficiente que mostremos que todo subconjunto algebraico propio y no vacío de $\mathbb{k}^n$ es cerrado y tiene interior vacío. Sea para ello Z un subconjunto algebraico propio de $\mathbb{k}^n$ y sea F un subconjunto de $\mathbb{k}[X_1, X_2, \ldots, X_n]$ tal que $Z = V(F)$.

- Como $Z = V(F) = \bigcap_{f \in F} V(f)$, el Lema 11.2.2.6 nos permite concluir inmediatamente que Z es un subconjunto cerrado de $\mathbb{k}^n$.
- Si $Z = \varnothing$, entonces claramente $Z^\circ = \varnothing$. Supongamos que $Z \neq \varnothing$. Como además $Z \subsetneq \mathbb{k}^n$, hay en F un elemento g que no es constante. Como $Z \subseteq V(g)$, el Lema 11.2.2.6 nos dice que $Z^\circ \subseteq V(g)^\circ = \varnothing$. □

11.2.3. Aproximación diofántica

Como el conjunto $\mathbb{Q}$ es denso en $\mathbb{R}$, sabemos que todo número real puede aproximarse tanto como queramos por números racionales. Un problema clásico de la teoría de números es el de estudiar *cuán* bien podemos hacer esa aproximación. Una forma de hacer esto es tratar de estimar cuál es el error que cometemos al aproximar un número por un número racional en términos del denominador de este último.

Definición 11.2.3.1. La ***medida de irracionalidad de Liouville–Roth*** de un número real x es el supremo $\mu(x)$ del conjunto de los números reales μ tales que hay un número positivo c e infinitos pares (p, q) de enteros coprimos tales que $q > 0$ y

$$\left| x - \frac{p}{q} \right| \leq \frac{c}{q^{\mu}}.$$

El nombre recuerda a Joseph Liouville y a Klaus Friedrich Roth. Por supuesto, para ver que esta definición tiene sentido tenemos que probar que el conjunto de los números cuyo supremo estamos tomando no es vacío, y haremos eso más abajo. Notemos que si ese conjunto no es acotado — y bien puede no serlo — es $\mu(x) = +\infty$. Mientras más grande sea el número $\mu(x)$ mejor podemos aproximar al número x con números racionales de denominadores relativamente chicos.

El siguiente resultado determina la medida de irracionalidad de los números racionales.

Proposición 11.2.3.2. *Sea x un número racional. Es $\mu(x) = 1$. De hecho,*

(*i*) *hay infinitos pares (p, q) de enteros coprimos tales que $q > 0$ y*

$$\left| x - \frac{p}{q} \right| \leq \frac{1}{q},$$

(*ii*) *y para cada par de números positivos ε y c hay solamente un número finito de pares (p, q) de enteros coprimos tales que $q > 0$ y*

$$\left| x - \frac{p}{q} \right| \leq \frac{c}{q^{1+\varepsilon}}. \tag{11.2}$$

Demostración. Sean r y s números enteros coprimos tales que $s > 0$ y $x = r/s$.

Como r y s son coprimos, hay pares de enteros (p, q) tales que $qr - ps = 1$ y, de hecho, hay infinitos. Sea (p, q) uno de ellos y supongamos que $q > 0$. Esa igualdad implica que p y q son coprimos y que $0 < r/s - p/q = 1/qs \leq 1/q$, así que $|x - p/q| \leq 1/q$. Si, en cambio, es $q < 0$, entonces $-r/s + p/q = -1/qs$, así que $|x - (-p)/(-q)| = 1/(-q)s \leq 1/(-q)$. Vemos así que la afirmación (*i*) de la proposición es cierta.

Sean ahora ε y c dos números positivos. Como ε es positivo, hay un entero positivo N tal que $cs/N^{\varepsilon} < 1/2$. Si un par de enteros (p,q) tal que $q \geq N$ satisface la desigualdad (11.2), entonces $|qr-ps| < cs/q^{\varepsilon} < 1/2$, así que debe ser $qr = ps$, ya que $qr-ps$ es un entero: vemos así que $p/q = r/s$ y, en consecuencia, $(p,q) = (r,s)$ porque $\operatorname{mcd}\{p,q\} = 1$, $\operatorname{mcd}\{r,s\} = 1$, $s > 0$ y $q > 0$. Por otro lado, si (p,q) es un par de enteros coprimos tal que $0 < q < N$ que satisface la desigualdad (11.2), entonces $|qx-p| \leq c/q^{\varepsilon} \leq c$, así que p pertenece al conjunto $\mathbb{Z} \cap (qx-c, qx+c)$, que está contenido en $\mathbb{Z} \cap (-c, Nc+x)$, que es finito. Juntado todo, vemos que hay un número finito de pares (p,q) de enteros coprimos que satisfacen la desigualdad 11.2. □

Esta proposición implica, en particular, que cuando el número x es racional el conjunto cuyo supremo estamos tomando en la Definición 11.2.3.1 no es vacío. Veamos que tampoco lo es cuando x es irracional.

Proposición 11.2.3.3. *Sea x un número real irracional. Hay infinitos pares (p,q) de enteros coprimos tales que $q > 0$ y*

$$\left| x - \frac{p}{q} \right| < \frac{1}{q^2},$$

de manera que, en particular, $\mu(x) \geq 2$.

Este resultado es conocido como el *Teorema de aproximación de Dirichlet*, por Peter Gustav Lejeune Dirichlet. La demostración que damos es la de él — tiene el interés histórico de ser uno de los primeros usos explícitos del *Principio del palomar* o *Principio de Dirichlet*.

Demostración. Para cada número real t escribamos $\{t\} \coloneqq t - \lfloor t \rfloor$ a la ***parte fraccionaria*** de t. Notemos que es suficiente probar la proposición en el caso particular en que x es un número positivo.

Primer paso. Sea N un entero positivo. Los $N+1$ números $\{0\}$ $\{x\}, \{2x\}, \ldots, \{Nx\}$ están todos en $[0,1)$, así que están en la unión de los N intervalos disjuntos dos a dos

$$\left[0, \frac{1}{N}\right), \quad \left[\frac{1}{N}, \frac{2}{N}\right), \quad \left[\frac{2}{N}, \frac{3}{N}\right), \quad \ldots \left[\frac{N-1}{N}, 1\right).$$

Hay, por lo tanto, dos elementos distintos m y n de $[\![0, N]\!]$ tales que $|\{mx\} - \{nx\}| < 1/N$.

Supongamos que $\{mx\} \geq \{nx\}$ y pongamos $r \coloneqq \lfloor mx \rfloor - \lfloor nx \rfloor$. Es

$$|(mx - nx) - r| = |(mx - \lfloor mx \rfloor) - (nx - \lfloor nx \rfloor)| = |\{mx\} - \{nx\}| < \frac{1}{N},$$

así que el entero r está a distancia menor que $1/N$ de $(m-n)x$. Esto implica que hay un entero a distancia menor que $1/N$ de $|m-n|x$: es r o $-r$, dependiendo de si $m-n$ es positivo o negativo. Por simetría podemos obtener la misma conclusión si $\{mx\} \leq \{nx\}$.

Si ponemos $q \coloneqq |m - n|$, que es un elemento de $[\![N]\!]$, entonces todo esto nos permite concluir que hay un entero p tal que $|qx - p| < 1/N$ y es $|x - p/q| < 1/qN \leq 1/q^2$.

Segundo paso. Sea $\mathscr{S}$ el conjunto de los pares (p, q) de enteros coprimos tales que $q > 0$ y $|x - p/q| < 1/q^2$. Para probar la proposición tenemos que mostrar que este conjunto es infinito.

- Como $|x - \lfloor x \rfloor/1| = |x - \lfloor x \rfloor| < 1/1^2$, el par $(\lfloor x \rfloor, 1)$ pertenece a $\mathscr{S}$, así que este conjunto no es vacío.
- Sea (p, q) es un elemento de $\mathscr{S}$. Como el número x es irracional es $qx - p \neq 0$ y, por lo tanto, hay un entero positivo N tal que $1/N < |qx - p|$. Vimos en el primer paso que hay un par de enteros (p', q') con $q' > 0$ tal que $|q'x - p'| < 1/N$ y $|x - p'/q'| < 1/q'^2$. Si $d \coloneqq \operatorname{mcd}\{p', q'\}$ y p'' y q'' son los enteros tales que $p' = dp''$ y $q' = dq''$, entonces tenemos que $|x - p''/q''| = |x - p'/q'| < 1/q'^2 \leq 1/q''^2$, así que (p'', q'') es un elemento de $\mathscr{S}$, y $|q''x - p''| = |q'x - p'|/d < |qx - p|$. Vemos así que no es vacío el conjunto

$$R(p, q) \coloneqq \{(r, s) \in \mathscr{S} : |rx - s| < |qx - p|\}.$$

El *Principio de elección dependiente* nos permite concluir a partir de estas dos observaciones que hay una función $f : \mathbb{N}_0 \to \mathscr{S}$ tal que $f(0) = (\lfloor x \rfloor, 1)$ y $f(n + 1) \in R(f(n))$ para todo $n \in \mathbb{N}_0$. Escribamos (p_n, q_n) al par ordenado $f(n)$ para cada $n \in \mathbb{N}_0$. La sucesión $(|q_n x - p_n|)_{n \geq 0}$ es entonces estrictamente decreciente, y esto implica que la función f es inyectiva. Vemos así que el conjunto $\mathscr{S}$ es infinito, como queremos. □

Adolf Hurwitz mejoró en [Hur1891] el resultado de la Proposición 11.2.3.3: probó que para todo número irracional x hay infinitos pares (p, q) de enteros coprimos con $q > 0$ y

$$\left| x - \frac{p}{q} \right| < \frac{1}{\sqrt{5} \cdot q^2}.$$

Esto significa que para cada número irracional el conjunto $\mathscr{M}(x)$ de todos los números M tales que hay infinitos pares (p, q) de enteros coprimos con $q > 0$ y

$$\left| x - \frac{p}{q} \right| < \frac{1}{M \cdot q^2}$$

no es vacío, ya que contiene a $\sqrt{5}$, y que podemos considerar el número $M(x) \coloneqq \sup \mathscr{M}(x)$, al que llamamos la ***constante de Markov*** de x, por Andrei Markov. Este supremo puede ser infinito: por ejemplo, $M(e) = \infty$.

El teorema de Hurwitz nos dice que $M(x) \geq \sqrt{5}$ cualquiera sea el número irracional x. De hecho, Hurwitz probó que el número $\sqrt{5}$ es el número más grande que podemos poner aquí: si M es cualquier número mayor que $\sqrt{5}$ y $\phi \coloneqq (1 + \sqrt{5})/2$ es la razón áurea, entonces hay solamente

finitos pares de números coprimos (p, q) tales que $q > 0$ y

$$\left|\phi - \frac{p}{q}\right| < \frac{1}{M \cdot q^2},$$

así que $M(\phi) = \sqrt{5}$. Sorprendentemente, esto nos dice que la razón áurea ϕ es el número irracional que peor podemos aproximar por números racionales. De hecho, no es el único: se puede mostrar que un número irracional ψ tiene $M(\psi) = \sqrt{5}$ si y solamente si es *equivalente* a ϕ, esto es, si hay cuatro enteros a, b, c y d tales que $|ad - bc| = 1$ y

$$\psi \coloneqq \frac{a\phi + b}{c\phi + d}.$$

Hurwitz también probó que si x es un número irracional que no es equivalente a ϕ, entonces $M(x) \geq \sqrt{8}$, y que $\sqrt{8}$ es el número más grande que podemos poner aquí: los números equivalentes a $\sqrt{2}$, es decir, de la forma

$$\frac{a\sqrt{2} + b}{c\sqrt{2} + d}$$

para alguna elección de cuatro enteros a, b, c y d tales que $|ad - bc| = 1$, y solo ellos, tienen constante de Markov igual a $\sqrt{8}$. Este procedimiento puede repetirse, y obtenemos así una sucesión de números que empieza

$$\sqrt{5}, \quad \sqrt{8}, \quad \frac{\sqrt{221}}{5}, \quad \frac{\sqrt{1517}}{13}, \quad \dots$$

Estos son los ***números de Lagrange***, por Joseph-Louis Lagrange. Increíblemente, tenemos una fórmula para estos números: el n-ésimo número de Lagrange es

$$L_n = \sqrt{9 - \frac{4}{m_n^2}},$$

con m_n el n-ésimo menor entero positivo m tal que la ***ecuación de Markov***

$$m^2 + x^2 + y^2 = 3mxy$$

tiene una solución (x, y) con componentes enteras positivas. El lector puede encontrar mucha información sobre esta bellísima teoría en el libro *The book of numbers* [CG1996] de John H. Conway y Richard K. Guy, y en el libro [Cas1957] de John William Scott Cassels sobre aproximación diofántica.

Una consecuencia de las Proposiciones 11.2.3.2 y 11.2.3.3 es que la medida de irreducibilidad $\mu(x)$ de todo número real x pertenece a $\{1\} \cup [2, +\infty]$. El siguiente resultado implica que los números algebraicos tienen medida de irracionalidad finita.

Proposición 11.2.3.4. *Si x es un número real que es raíz de un polinomio irreducible con coeficientes enteros de grado n, entonces $\mu(x) \leq n$.*

Demostración. Sea x un número real, sea n un entero positivo, sean a_0, a_1, …, a_n enteros, y supongamos que x es raíz del polinomio $f(X) = a_0 + a_1X + \cdots + a_nX^n$ de $\mathbb{Z}[X]$ y que este polinomio f es irreducible. Sean ε y c dos números positivos y sea S el conjunto de todos los pares (p, q) de enteros coprimos tales que $q > 0$ y

$$\left| x - \frac{p}{q} \right| \leq \frac{c}{q^{n+\varepsilon}}.$$

Para probar la proposición es suficiente que mostremos que el conjunto S es finito.

Notemos que la irreducibilidad de f implica que la derivada $f' : \mathbb{R} \to \mathbb{R}$ no se anula en x. Como además esa derivada es una función continua, hay más aún números positivos λ y K tales que $0 < |f'(y)| < 1/K$ para todo elemento y de $(x - \lambda, x + \lambda)$.

Como ε es positivo, hay un entero positivo N tal que $c/N^\varepsilon < \min\{K, \lambda\}$. Supongamos que (p, q) es un elemento de S tal que $q \geq N$ y $f(p/q) \neq 0$. Es entonces $|x - p/q| \leq c/q^{n+\varepsilon} \leq c/N^\varepsilon < \lambda$, así que $p/q \in (x - \lambda, x + \lambda)$. Tenemos además que

$$\left| f\left(\frac{p}{q}\right) \right| = \frac{\left| \sum_{i=0}^{n} a_i p^i q^{n-1} \right|}{q^n} \geq \frac{1}{q^n},$$

ya que el denominador de esta fracción es un entero no nulo, y, en vista del *Teorema del valor medio*, hay un número entre x y p/q tal que

$$f\left(\frac{p}{q}\right) = f\left(\frac{p}{q}\right) - f(x) = \left(\frac{p}{q} - x\right) \cdot f'(\xi).$$

Como ξ también pertenece a $(x - \lambda, x + \lambda)$, es $f'(\xi) \neq 0$ y, por lo tanto,

$$\left| x - \frac{p}{q} \right| = \frac{|f(p/q)|}{|f'(\xi)|} \geq \frac{K}{q^n} > \frac{c}{q^n N^\varepsilon} \geq \frac{c}{q^{n+\varepsilon}}.$$

Esto es absurdo porque (p, q) pertenece a S. Esta contradicción nos permite concluir que el conjunto S es la unión de los dos conjuntos

$$S_1 := \{(p, q) \in S : q < N\}, \qquad S_2 := \{(p, q) \in S : f(p/q) = 0\}.$$

Si (p, q) es un elemento de S_1, entonces $|qx - p| \leq cq/q^{n+\varepsilon} < cN$, así que p es uno de los enteros del intervalo $(qx - cN, qx + cN)$: vemos así que S_1 tiene como mucho $N \cdot (\lceil qx + cN \rceil - \lfloor qx - cN \rfloor)$ elementos. Como, por otro lado, es claro que S_2 tiene como mucho n elementos, vemos que el conjunto S es finito, como queremos. □

Usando esta proposición podemos probar que hay números reales que no son algebraicos, esto es, que son transcendentes.

Proposición 11.2.3.5. *Sea $(a_n)_{n\geq 0}$ una sucesión de elementos de $[\![9]\!]$. La medida de irreducibilidad de*

$$x \coloneqq \sum_{n=0}^{\infty} a_n \cdot 10^{-n!}$$

es $+\infty$, así que este es un número transcendente.

Los números que obtenemos de esta manera son los llamados ***números de Liouville***, por Joseph Liouville, que en 1844 pudo mostrar por primera vez que hay números transcendentes verificando de esencialmente la misma forma en que lo haremos nosotros que sus números lo son. Los primeros 720 dígitos decimales de uno de ellos son, por ejemplo,

$$\begin{aligned}
&0{,}240\,003\,000\,000\,000\,000\,000\,009\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\\
&\quad 000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,006\\
&\quad 000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\\
&\quad 000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\\
&\quad 000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\\
&\quad 000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\\
&\quad 000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\\
&\quad 000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\\
&\quad 000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\\
&\quad 000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\\
&\quad 000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\\
&\quad 000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\\
&\quad 000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,000\,008\ \ldots
\end{aligned}$$

Números como este, en que los dígitos decimales no nulos están más y más espaciados se llaman ***números lacunares***, porque tienen *lagunas*, y generalmente tienen propiedades muy especiales.

Demostración. Notemos que es claro que la serie que define a x converge: sus términos son positivos y menores uno a uno que los de la serie geométrica $\sum_{n=0}^{\infty} 9/10^{-n-1}$. Esto implica, además, que $x \in (0,1)$. Para cada entero no negativo N sea

$$r_N \coloneqq \sum_{n=0}^{N} a_n \cdot 10^{-n!},$$

la N-ésima suma parcial de la serie, que es un número racional, y sean p_N y q_N los enteros coprimos

tales que $q_N > 0$ y $r_N = p_N/q_N$. Como $10^{N!} \cdot x$ es un entero, es $q_N \leq 10^{N!}$, y tenemos entonces que

$$\begin{aligned} |x - r_N| &= \sum_{n=N+1}^{\infty} \frac{a_n}{10^{n!}} \leq \sum_{n=N+1}^{\infty} \frac{9}{10^{n!}} = \frac{1}{10^{(N+1)!}} \cdot \sum_{n=N+1}^{\infty} \frac{9}{10^{n!-(N+1)!}} \\ &\leq \frac{1}{10^{(N+1)!}} \cdot \sum_{n=0}^{\infty} \frac{9}{10^{(N+n+1)!-(N+1)!}} \leq \frac{1}{10^{(N+1)!}} \cdot \sum_{n=0}^{\infty} \frac{9}{10^n} = \frac{10}{10^{(N+1)!}} \leq \frac{10}{q_N^{N+1}}, \end{aligned}$$

ya que para todo $n \in \mathbb{N}_0$ es

$$(N+n+1)! - (N+1)! = \sum_{i=1}^{n} \big[(N+i+1)! - (N+i)! \big] \geq n.$$

Notemos que la función $N \in \mathbb{N}_0 \mapsto (p_N, q_N) \in \mathbb{N} \times \mathbb{N}$ es inyectiva, ya que la sucesión $(r_N)_{N \geq 0}$ es estrictamente creciente porque los términos de la serie que define a x son estrictamente positivos.

Ahora bien, para cada número positivo μ, tenemos que

$$\left| x - \frac{p_N}{q_N} \right| \leq \frac{10}{q_N^{N+1}} \leq \frac{10}{q_N^{\mu}}$$

siempre que N es un entero tal que $N \geq \mu$, y esto nos dice que $\mu(x) \geq \mu$. Podemos concluir de esto que $\mu(x) = +\infty$ y, por la Proposición 11.2.3.4, que el número x es transcendente. □

La Proposición 11.2.3.4 de Liouville fue mejorada de manera sucesiva por Axel Thue [Thu1909], Carl Ludwig Siegel [Sie1921] Freeman Dyson [Dys1947] y Klaus Friedrich Roth [Rot1955], y el resultado final, el llamado *Teorema de Thue–Siegel–Roth*, es el siguiente:

Teorema 11.2.3.6. *Un número irracional algebraico tiene medida de irracionalidad* 2. □

Este teorema nos dice que los números algebraicos — racionales e irracionales — están entre los que peor podemos aproximar por números racionales.

Definición 11.2.3.7. Un número real x es ***diofantino*** si para todo número positivo ε hay otro c tal que

$$\left| x - \frac{p}{q} \right| \geq \frac{c}{q^{2+\varepsilon}}$$

para toda toda elección de p en $\mathbb{Z}$ y de q en $\mathbb{N}$.

Queremos probar que hay relativamente pocos números diofantinos, en el siguiente sentido preciso:

Proposición 11.2.3.8. *El conjunto de los números diofantinos es magro en* $\mathbb{R}$.

Demostración. Sea k un entero positivo y consideremos el conjunto

$$M_k \coloneqq \left\{ x \in \mathbb{R} : \left| x - \frac{p}{q} \right| \geq \frac{1}{k \cdot q^3} \text{ siempre que } p \in \mathbb{Z} \text{ y } q \in \mathbb{N} \right\}.$$

Se trata de un conjunto cerrado. En efecto, supongamos que $(x_n)_{n \geq 0}$ es una sucesión en M_k que converge a un número ξ en $\mathbb{R}$, y sean p y q dos enteros tales que $q > 0$. Como la sucesión toma valores en M_k, tenemos que $|x_n - p/q| \geq 1/kq^3$ y entonces, tomando límite a ambos lados de esa desigualdad, vemos que también $|\xi - p/q| \geq 1/kq^2$. Esto nos permite concluir que ξ pertenece a M_k y, por lo tanto, que M_k es un subconjunto cerrado de $\mathbb{R}$, como dijimos.

Si x es un número racional, entonces hay un par (p, q) de enteros coprimos tal que $q > 0$ y $x = p/q$, y es $|x - p/q| = 0 < 1/kq^3$, de manera que $x \notin M_k$. Vemos así que $M_k \cap \mathbb{Q} = \varnothing$, de manera que M_k, que es cerrado, tiene interior vacío: vemos así que M_k es nunca denso en $\mathbb{R}$ y, por lo tanto, que el conjunto

$$M \coloneqq \bigcup_{k=1}^{\infty} M_k$$

es magro en $\mathbb{R}$. Para probar la proposición es suficiente que mostremos que este conjunto M contiene al conjunto de todos los números diofantinos de $\mathbb{R}$.

Sea x un número diofántico, de manera que, en particular, hay un número positivo c tal que para toda elección de p en $\mathbb{Z}$ y q en $\mathbb{N}$ es $|x - p/q| \geq c/q^3$: como hay un entero positivo k tal que $c \geq 1/k$, esto nos dice entonces que $x \in M_k \subseteq M$. □

Capítulo 12

Espacios de Banach

12.1. Espacios de Banach

Empecemos haciendo dos observaciones sobre espacios normados en general.

Proposición 12.1.1. *Sea X un espacio normado. Si Y es un subespacio de X, entonces la clausura $\overline{Y}$ de Y también es un subespacio de X.*

Demostración. Sea Y un subespacio de X, sean x e y dos elementos de $\overline{Y}$, y sean a y b dos escalares de $\mathbb{k}$. Para ver que $\overline{Y}$ es un subespacio de X es suficiente que mostremos que $ax + by$ pertenece a $\overline{Y}$, porque ciertamente 0, que pertenece a Y, pertenece a $\overline{Y}$. Esto es fácil: en efecto, si ε es un número positivo menor que 2, entonces hay elementos x' e Y' de Y tales que $\|x - x'\| < \varepsilon/(2+|a|)$ y $\|y - y'\| < \varepsilon/(2+|b|)$ y, por lo tanto, $ax' + by' \in Y$ y

$$\|(ax+by)-(ax'+by')\| \leq |a|\cdot\|x-x'\| + |b|\cdot\|y-y'\| < |a|\cdot\frac{\varepsilon}{2+|a|} + |b|\cdot\frac{\varepsilon}{2+|b|}$$
$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

Lema 12.1.2. *Un subespacio propio de un espacio normado tiene interior vacío.*

Demostración. Sea X un espacio normado y sea Y un subespacio de X que tiene interior no vacío, de manera que hay un elemento x de X y un número positivo r tales que $B_r(x) \subseteq Y$. Como x está en Y e Y es un subespacio, tenemos entonces que también $B_r(0) = B_r(x) - x \subseteq Y$. A su vez, esto

implica, también porque Y es un subespacio, que para todo $t \in \mathbb{R}$ es $B_{tr}(0) = t \cdot B_r(0) \subseteq Y$ y, en consecuencia, que

$$X = \bigcup_{t \geq 0} B_{rt}(0) \subseteq Y.$$

Vemos así que Y no es un subespacio propio. □

Proposición 12.1.3. *Un espacio de Banach de dimensión infinita tiene dimensión no numerable.*

Veremos más adelante que, de hecho, un espacio de Banach de dimensión infinita tiene dimensión al menos igual a la potencia del continuo. https://math.stackexchange.com/a/1538842/274

Demostración. Sea X un espacio de Banach de dimensión numerable, de manera que hay una base $\mathscr{B}$ de X que es numerable, sea $\phi : \mathbb{N}_0 \to \mathscr{B}$ una función biyectiva, y para cada $n \in \mathbb{N}_0$ pongamos $x_n \coloneqq \phi(n)$. Para cada $n \in \mathbb{N}_0$ consideremos el subespacio $F_n \coloneqq \langle x_0, x_1, \ldots, x_n \rangle$. Como $\mathscr{B}$ genera a X como espacio vectorial, es claro que

$$X = \bigcup_{n \geq 0} F_n$$

y el *Teorema de Baire* nos dice que hay un elemento m de $\mathbb{N}_0$ tal que el interior de $\overline{F_m}$ no es vacío. Esto es absurdo porque, por un lado, como F_m tiene dimensión finita, la Proposición 7.5.1 nos dice que F_m es completo con respecto a la norma que le induce la de X y, por lo tanto, que es un subconjunto cerrado de X y, por otro, como F_n tiene dimensión finita y X no, es claro que F_m es un subespacio propio de X y, de acuerdo al Lema 12.1.2, tiene interior vacío.

Esta contradicción provino de haber supuesto que hay espacios de Banach de dimensión numerable, así que no los hay: esto implica claramente que un espacio de Banach de dimensión infinita tiene dimensión no numerable. □

Esta proposición nos dice, por ejemplo, que no hay ninguna norma sobre el espacio vectorial real $\mathbb{R}[X]$ de los polinomios reales que haga de él un espacio de Banach.

12.2. Funciones lineales continuas

Entre espacios vectoriales estudiamos las funciones lineales, y si los espacios son además normales, entonces podemos considerar las que son además continuas. Eso es lo que haremos en esta sección.

Definición 12.2.1. Si X e Y son dos espacios normados, entonces escribimos $L(X, Y)$ al conjunto de todas las funciones $X \to Y$ que son a la vez lineales y continuas. Cuando Y es $\mathbb{k}$, con su norma usual, entonces escribimos X' en lugar de $L(X, \mathbb{k})$: llamamos a X' el ***espacio dual continuo*** de X.

Empecemos por la siguiente caracterización extremadamente útil de las funciones lineales que son continuas.

Proposición 12.2.2. *Sean X e Y dos espacios normados y sea $T : X \to Y$ una función lineal. Las siguientes afirmaciones son equivalentes:*

(*a*) *T es continua.*
(*b*) *T es continua en 0.*
(*c*) *T es acotada sobre la bola $\overline{B}_1(0)$ de X.*
(*d*) *Hay un número positivo c tal que $\|T(x)\| \leq c\|x\|$ para todo $x \in X$.*
(*e*) *T satisface la condición de Lipschitz.*
(*f*) *T es uniformemente continua.*

Demostración. Observemos que las implicaciones $(a) \Rightarrow (b)$, $(e) \Rightarrow (f)$ y $(f) \Rightarrow (a)$ son evidentes.

$(b) \Rightarrow (c)$ Supongamos que la función T es continua en 0, de manera que, en particular, como $T(0) = 0$, hay un número positivo δ tal que $f(B^X_\delta(0)) \subseteq B^Y_1(0)$. Si x es un elemento de $\overline{B}^X_1(0)$, entonces $\delta x/2 \in B^X_\delta(0)$, así que $\delta T(x)/2 = T(\delta x/2) \in B^Y_1(0)$, así que $\|T(x)\| < 2/\delta$. Esto nos dice que la función T es acotada sobre $\overline{B}^X_1(0)$.

$(c) \Rightarrow (d)$ Supongamos que hay un número positivo c tal que $\|T(x)\| \leq c$ para todo elemento x de $\overline{B}^X_1(0)$. Si x es un elemento no nulo cualquiera de X, entonces $x/\|x\|$ pertenece a $\overline{B}^X_1(0)$ y, por lo tanto, $\|T(x/\|x\|)\| \leq c$, así que $\|T(x)\| \leq c\|x\|$. Como esta desigualdad vale también cuando $x = 0$, vemos que vale la afirmación (d).

$(d) \Rightarrow (e)$ Si hay un número positivo c tal que $\|T(x)\| \leq c\|x\|$ para todo elemento x de X, entonces para toda elección de x e y en X tenemos que

$$\|T(x) - T(y)\| = \|T(x - y)\| \leq c\|x - y\|,$$

y esto nos dice que T satisface la condición de Lipschitz. □

En la afirmación (c) de esta proposición consideramos la bola cerrada $\overline{B}^X_1(0)$, pero podríamos considerar la bola abierta.

Ejercicio 12.2.3. Sean X e Y dos espacios normados. Muestre que una función lineal $T : X \to Y$ es continua si y solamente si es acotada sobre la bola $B_1^X(0)$.

Hay una situación especial en la que la continuidad de una función lineal es automática:

Proposición 12.2.4. *Sean X e Y dos espacios normados y sea $T : X \to Y$ una función lineal. Si X tiene dimensión finita, entonces T es continua.*

Demostración. Supongamos que el espacio X tiene dimensión finita, y sea $(x_1, \ldots, x_n)$ una base ordenada de X. Hay exactamente una función $\|-\| : X \to \mathbb{R}$ que para cada elección de $a_1, \ldots, a_n$ en $\Bbbk$ tiene

$$\|a_1x_1 + \cdots + a_nx_n\| = \text{máx}\{|a_i| : i \in [\![n]\!]\},$$

y se trata de una norma sobre X. Como X tiene dimensión finita, todas las normas sobre X son equivalentes, así que a los fines de probar que la función T es continua podemos suponer que la norma de X es precisamente $\|-\|$. Eso haremos.

Consideremos el número $c \coloneqq \text{máx}\{\|T(x_i)\| : i \in [\![n]\!]\}$. Si x es un elemento cualquiera de X, y $a_1, \ldots, a_n$ son los escalares de $\Bbbk$ tales que $x = a_1x_1 + \cdots + a_nx_n$, entonces

$$\begin{aligned}\|T(x)\| &= \|T(a_1x_1 + \cdots + a_nx_n)\| = \|a_1T(x_1) + \cdots + a_nT(x_n)\| \\ &\leq |a_1| \cdot \|T(x_1)\| + \cdots + |a_n| \cdot \|T(x_n)\| \leq \|x\| \cdot c + \cdots + \|x\| \cdot c = nc \cdot \|x\|.\end{aligned}$$

De acuerdo a la Proposición 12.2.2, esto implica que la función T es continua, como queremos. □

Esta proposición nos dice que toda función lineal entre espacios normados que tiene dominio de dimensión finita es continua. Esta propiedad, de hecho, caracteriza a los espacios normados de dimensión finita, como muestra el siguiente ejemplo.

Ejemplo 12.2.5. Sea X un espacio normado de dimensión infinita, y sea $\mathscr{B}$ una base de X. Como $\mathscr{B}$ es infinita, hay una función inyectiva $\phi : \mathbb{N}_0 \to \mathscr{B}$: para cada $n \in \mathbb{N}_0$ escribamos x_n a $\phi(x)$. Como el conjunto $\mathscr{B}$ es una base de X y la función ϕ es inyectiva, sabemos que hay exactamente una función lineal $T : X \to \Bbbk$ que en cada elemento x de $\mathscr{B}$ toma el valor

$$T(x) = \begin{cases} n\|x_n\| & \text{si } x = x_n \text{ para algún } n \in \mathbb{N}_0; \\ 0 & \text{si } x \notin \phi(\mathbb{N}_0). \end{cases}$$

Es claro que la sucesión $(x_n/n\|x_n\|)_{n\geq 0}$ converge a 0. Como

$$\left|T\left(\frac{x_n}{n\|x_n\|}\right)\right| = \frac{|T(x_n)|}{n\|x_n\|} = 1,$$

la sucesión $(T(x_n/n\|x_n\|))_{n\geq 0}$ claramente no converge a $T(0)$ y, por lo tanto, la función T no es continua.

En general, si X e Y son espacios normados, hay muchas más funciones lineales $X \to Y$ que funciones lineales y continuas $X \to Y$. Veamos un ejemplo de esto.

Ejemplo 12.2.6. Consideremos el espacio $\ell^1_{\mathbb{k}}$ y para cada $n \in \mathbb{N}_0$ sea e_n el elemento de $\ell^1_{\mathbb{k}}$ cuya única componente no nula es la n-esima, que es igual a 1. La función $n \in \mathbb{N}_0 \mapsto e_n \in \ell^1_{\mathbb{k}}$ es claramente inyectiva y su imagen es linealmente independiente.

Supongamos que $f, g : \ell^1_{\mathbb{k}} \to \mathbb{k}$ son dos funciones lineales y continuas tales que $f(e_n) = g(e_n)$ para todo $n \in \mathbb{N}_0$, y sea $x = (x_n)_{n\geq 0}$ un elemento cualquiera de $\ell^1_{\mathbb{k}}$. Para cada $n \in \mathbb{N}_0$ es

$$\left\| x - \sum_{i=0}^{n} x_i e_i \right\| = \sum_{i=n+1}^{\infty} |x_i|$$

y, como la serie $\sum_{i=0}^{\infty} |x_i|$ converge, es claro que

$$\lim_{n\to\infty} \sum_{i=0}^{n} x_i e_i = x$$

y, como las funciones f y g son lineales y continuas, que

$$\begin{aligned} f(x) &= f\left(\lim_{n\to\infty} \sum_{i=0}^{n} x_i e_i\right) = \lim_{n\to\infty} f\left(\sum_{i=0}^{n} x_i e_i\right) = \lim_{n\to\infty} \sum_{i=0}^{n} x_i f(e_i) = \lim_{n\to\infty} \sum_{i=0}^{n} x_i g(e_i) \\ &= \lim_{n\to\infty} g\left(\sum_{i=0}^{n} x_i e_i\right) = g\left(\lim_{n\to\infty} \sum_{i=0}^{n} x_i e_i\right) = g(x). \end{aligned}$$

Vemos de esta forma que la función

$$\Phi : f \in L(\ell^1_{\mathbb{k}}, \mathbb{k}) \mapsto (f(e_n))_{n\geq 0} \in \mathbb{k}^{\mathbb{N}_0}$$

es inyectiva, así que

$$|L(\ell^1_{\mathbb{k}}, \mathbb{k})| \leq \mathfrak{c}^{\aleph_0} = \left(2^{\aleph_0}\right)^{\aleph_0} = 2^{\aleph_0 \cdot \aleph_0} = 2^{\aleph_0} = \mathfrak{c}.$$

Veremos más adelanete, en la Proposición **??**, que $\dim \ell^1_{\mathbb{k}} \geq \mathfrak{c}$: usando esto y las Proposiciones 3.5.24 y 3.5.25 podemos concluir entonces que

$$|(\ell^1_{\mathbb{k}})^*| \geq \dim(\ell^1_{\mathbb{k}})^* > \dim \ell^1_{\mathbb{k}} \geq \mathfrak{c} = |L(\ell^1_{\mathbb{k}}, \mathbb{k})|$$

y, por lo tanto, que hay más funciones lineales $\ell^1_{\mathbb{k}} \to \mathbb{k}$ que funciones lineales y continuas $\ell^1_{\mathbb{k}} \to \mathbb{k}$.

Si X e Y son espacios normados, entonces el conjunto $L(X, Y)$ de todas las funciones lineaes

y continuas $X \to Y$ es, de manera natural, un espacio vectorial.

Ejercicio 12.2.7. Sean X e Y dos espacios normados. Pruebe que el subconjunto $L(X, Y)$ del espacio vectorial $\mathrm{Hom}(X, Y)$ de todas las funciones lineales $X \to Y$ es, de hecho, un subespacio.

El espacio vectorial $L(X, Y)$ de funciones lineales y continuas entre dos espacios normados es él mismo un espacio normado de manera natural.

Proposición 12.2.8. *Sean X e Y dos espacios normados.*

(*i*) *Si T es un elemento de $L(X, Y)$, entonces el subconjunto $T(\overline{B}_1^X(0))$ de Y es acotado.*

(*ii*) *La función*

$$\|-\|_{L(X,Y)} : T \in L(X, Y) \mapsto \sup\{\|T(x)\| : x \in \overline{B}_1^X(0)\} \in \mathbb{R}$$

es una norma sobre el espacio vectorial $L(X, Y)$.

(*iii*) *Siempre que T es un elemento de $L(X, Y)$ y x uno de X es*

$$\|T(x)\|_Y \leq \|T\|_{L(X,Y)} \cdot \|x\|_X.$$

Siempre que hablamos del espacio $L(X, Y)$ lo consideramos como un espacio normado con respecto a la norma que describe esta proposición.

Demostración. La afirmación (*i*) es consecuencia inmediata de la Proposición 12.2.2 con que empezamos esta sección. Mostremos que la función $\|-\|_{L(X,Y)}$ de (*ii*) es una norma.

- Si $T : X \to Y$ es la función nula, el elemento 0 de $L(X, Y)$, entonces $T(x) = 0$ para todo $x \in \overline{B}_1^X(0)$, así que, por supuesto, $\|T\|_{L(X,Y)} = \sup\{\|T(x)\| : x \in \overline{B}_1^X(0)\} = 0$.
- Sea $T : X \to Y$ un elemento de $L(X, Y)$ tal que $\|T\|_{L(X,Y)} = 0$. Si x es un elemento no nulo cualquiera de X, entonces $x/\|x\|$ es un elemento de $\overline{B}_1^X(0)$ y, por lo tanto,

$$\frac{\|T(x)\|}{\|x\|} = \left\| T\left(\frac{x}{\|x\|}\right)\right\| \leq \sup\{\|T(y)\| : y \in \overline{B}_1^X(0)\} = \|T\|_{L(X,Y)} = 0,$$

  de manera que $T(x) = 0$. Vemos así que $T = 0$.
- Sea $T : X \to Y$ un elemento de $L(X, Y)$ y sea t un escalar de $\mathbb{k}$. Si $t = 0$, entonces es claro que $\|tT\|_{L(X,Y)} = |t| \cdot \|T\|_{L(X,Y)}$, porque ambos lados de la igualdad se anulan. Supongamos entonces que $t \neq 0$. Para cada elemento x de $\overline{B}_1^X(0)$ es

$$\|(tT)(x)\| = \|t \cdot T(x)\| = |t| \cdot \|T(x)\| \leq |t| \cdot \|T\|_{L(X,Y)},$$

  así que

$$\|tT\|_{L(X,Y)} = \sup\{\|(tT)(x)\| : x \in \overline{B}_1^X(0)\} \leq |t| \cdot \|T\|_{L(X,Y)}.$$

Esto vale cualquiera sea T en $L(X,Y)$ y t en $\Bbbk \setminus \{0\}$, así que también tenemos que

$$\|T\|_{L(X,Y)} = \|t^{-1}(tT)\|_{L(X,Y)} \leq |t^{-1}| \cdot \|tT\|_{L(X,Y)} = |t|^{-1} \cdot \|tT\|_{L(X,Y)},$$

de manera que

$$|t| \cdot \|T\|_{L(X,Y)} \leq \|tT\|_{L(X,Y)}.$$

Combinando las dos desigualdades que obtuvimos, vemos que $\|tT\|_{L(X,Y)} = |t| \cdot \|T\|_{L(X,Y)}$.

- Sean ahora $S, T : X \to Y$ dos elementos de $L(X,Y)$. Si x es un elemento de $\overline{B}_1^X(0)$, entonces

$$\|(S+T)(x)\| = \|S(x) + T(x)\| \leq \|S(x)\| + \|T(x)\| \leq \|S\|_{L(X,Y)} + \|T\|_{L(X,Y)},$$

así que

$$\|S+T\|_{L(X,Y)} = \sup\{\|(S+T)(x)\| : x \in \overline{B}_1^X(0)\} \leq \|S\|_{L(X,Y)} + \|T\|_{L(X,Y)}.$$

Probemos, para terminar, la afirmación (*iii*). Sea $T : X \to Y$ un elemento de $L(X,Y)$ y sea x un elemento de X. Si x no es nulo, entonces $x/\|x\|$ es un elemento de $\overline{B}_1^X(0)$, así que $\|T(x)\|/\|x\| = \|T(x/\|x\|\| \leq \|T\|_{L(X,Y}$ y, por lo tanto, $\|T(x)\| \leq \|T\|_{L(X,Y)} \cdot \|x\|$. Es claro que esta última desigualdad también vale cuando x sí es nulo. □

Es importante que cuando X e Y son espacios de Banach el espacio normado $L(X,Y)$ también lo es. De hecho, vale más:

Proposición 12.2.9. *Sean X e Y dos espacios normados. Si Y es un espacio de Banach, entonces $L(X,Y)$ también lo es.*

Demostración. Supongamos que Y es un espacio de Banach, y sea $(T_n)_{n\geq 0}$ una sucesión en $L(X,Y)$ que satisface la condición de Cauchy. Para cada $x \in X$ es

$$\|T_n(x) - T_m(x)\| \leq \|T_n - T_m\| \cdot \|x\|,$$

así que la sucesión $(T_n(x))_{n\geq 0}$ en Y satisface la condición de Cauchy y, por lo tanto, converge. Vemos así que hay una función $T : X \to Y$ tal que

$$T(x) = \lim_{n\to\infty} T_n(x)$$

para todo $x \in X$. Esta función es lineal: en efecto, si x e y son dos elementos de X y a y b son dos escalares de $\Bbbk$, entonces para cada $n \in \mathbb{N}_0$ es $T_n(ax+by) = aT_n(x) + bT_n(y)$ y, como las operaciones de un espacio normado son continuas, es

$$T(ax+by) = \lim_{n\to\infty} T_n(ax+by) = \lim_{n\to\infty} \big(aT_n(x) + bT_n(y)\big)$$

$$= a \lim_{n\to\infty} T_n(x) + b \lim_{n\to\infty} T_n(y) = aT(x) + bT(y).$$

Como la sucesión (T_n) satisface la condición de Cauchy en $L(X, Y)$, es acotada y, por lo tanto, hay un número positivo K tal que $\|T_n\| \leq K$ para todo n. Si x es un elemento cualquiera de $\overline{B}_1^X(0)$, entonces $\|T_n(x)\| \leq K$ para todo $n \in \mathbb{N}_0$, así que, como la norma de Y es una función continua, tenemos que

$$\|T(x)\| = \lim_{n\to\infty} \|T_n(x)\| \leq K.$$

Vemos así que la función lineal T es acotada sobre $\overline{B}_1^X(0)$ y, por lo tanto, que es continua.

Esto nos dice que T es un elemento de $L(X, Y)$. Para terminar, mostremos que la sucesión $(T_n)_{n\geq 0}$ converge a T en $L(X, Y)$. Sea ε un número positivo. Como la sucesión $(T_n)_{n\geq 0}$ satisface la condición de Cauchy en $L(X, Y)$, hay un entero positivo n_0 tal que $\|T_m - T_n\| < \varepsilon/2$ siempre que m y n son enteros mayores que n_0. Sea n un entero mayor que n_0. Si x es un elemento de $\overline{B}_1^X(0)$, entonces para todo entero m mayor que n_0 es

$$\|T_m(x) - T_n(x)\| \leq \|T_n - T_m\| \cdot \|x\| < \frac{\varepsilon}{2}$$

y la sucesión $(T_m(x))_{m\geq 0}$ converge a $T(x)$ y la función $y \in X \mapsto \|y - T_n(x)\| \in \mathbb{R}$ es continua, así que

$$\|(T - T_n)(x)\| = \|T(x) - T_n(x)\| \leq \frac{\varepsilon}{2}.$$

Esto es así cualquiera sea x en $\overline{B}_1^X(0)$ así que $\|T - T_n\| \leq \varepsilon/2 < \varepsilon$. Esto prueba lo que queremos. □

12.3. El teorema de Hahn–Banach y el espacio dual de un espacio normado

Una observación fundamental del álgebra lineal es que toda función lineal $U \to Y$ definida sobre un subespacio U de un espacio vectorial X posee una extensión lineal a todo X. En esta sección estudiaremos este problema en la situacion en la que los espacios son espacios de Banach y las funciones lineales que nos interesan son continuas.

Empezamos con un resultado que es, esencialmente, álgebra lineal.

Proposición 12.3.1. *Sea X un espacio vectorial real, sea U un subespacio de X, sea $f : U \to \mathbb{R}$ una*

función lineal, y sea $p : X \to \mathbb{R}$ *una función tal que*

$$p(x+y) \leq p(x) + p(y), \qquad p(tx) \leq t \cdot p(x)$$

siempre que x *e* y *son elementos de* X *y* t *es un número real no negativo. Si para todo elemento* x *de* U *es* $f(x) \leq p(x)$*, entonces hay una función lineal* $\bar{f} : V \to \mathbb{R}$ *que extiende a* f *y tal que* $\bar{f}(x) \leq p(x)$ *para todo* $x \in X$*.*

Es importante observar que esta proposición habla exclusivamente de espacios vectoriales reales.

Demostración. Sea $\mathscr{S}$ el conjunto de todos los pares (V, g) que tienen como primera componente un subespacio V de X que contiene a U y como segunda componente una función lineal $g : V \to \mathbb{R}$ tal que $g|_U = f$ y $g(x) \leq p(x)$ para todo $x \in V$. Consideremos la relación $\preceq$ sobre el conjunto $\mathscr{S}$ tal que siempre que (V, g) y (V', g') son elementos de $\mathscr{S}$ vale

$$(V, g) \preceq (V', g') \iff V \subseteq V' \wedge g'|_V = g.$$

Es fácil ver que $\preceq$ es una relación de orden sobre $\mathscr{S}$. El conjunto $\mathscr{S}$ no es vacío, ya que (U, f) es un elemento de $\mathscr{S}$. Queremos mostrar que $\mathscr{S}$ satisface la hipótesis del *Lema de Zorn* 2.2.6.

Sea $\mathscr{C}$ un subconjunto no vacío totalmente ordenado de $\mathscr{S}$ y pongamos

$$W \coloneqq \bigcup_{(V,g)\in\mathscr{C}} V, \qquad h \coloneqq \bigcup_{(V,g)\in\mathscr{C}} g.$$

Mostremos que W es un subespacio de X, que h es un subconjunto de $W \times \mathbb{R}$ que es, de hecho, una función lineal $W \to \mathbb{R}$, que el par (W, h) pertenece a $\mathscr{S}$, y que es una cota superior para $\mathscr{C}$ en $\mathscr{S}$.

- Como el conjunto $\mathscr{C}$ no es vacío, hay un elemento (V_0, g_0) en $\mathscr{C}$ y, por lo tanto, $0 \in V_0 \subseteq W$. Sean, por otro lado, x e y dos elementos de W y sean a y b dos escalares de $\mathbb{R}$. Como x e y están en W, hay dos elementos (V, g) y (V', g') en $\mathscr{C}$ tales que $x \in V$ e $y \in V'$, y como $\mathscr{C}$ está totalmente ordenado con respecto a $\preceq$ es $(V, g) \preceq (V', g')$ o $(V', g') \preceq (V, g)$. Supongamos que vale la primera de estas dos posibilidades — si no fuera ese el caso y valiera la segunda podríamos proceder de exactamente la misma forma. Tenemos entonces que $V \subseteq V'$, así que x e y están los dos en V' y, como V' es un subespacio de X, tenemos que $ax + by \in V' \subseteq W$. Esto prueba que W es un subespacio de X.
- Si (x, y) es un elemento de h, de manera que hay un elemento (V, g) de $\mathscr{C}$ tal que $(x, y) \in g$ y, como g es una función $V \to \mathbb{R}$, esto nos dice que $(x, y) \in W \times \mathbb{R}$. Vemos así que $h \subseteq W \times \mathbb{R}$.
- Si x es un elemento de W, entonces hay un elemento (V, g) de $\mathscr{C}$ tal que $x \in V$ y, por lo tanto, $(x, g(x)) \in g \subseteq h$. Sean, por otro lado, x un elemento de W e y e y' dos elementos de $\mathbb{R}$ tales que los pares (x, y) y (x, y') están los dos en h. Hay entonces dos elementos (V, g)

y (V', g') en $\mathscr{C}$ tales que $(x, y) \in g$ y $(x, y') \in g'$. Como $\mathscr{C}$ está totalmente ordenado con respecto a $\preceq$ es $(V, g) \preceq (V', g')$ o $(V', g') \preceq (V, g)$, y podemos suponer, como antes, que la primera posibilidad ocurre, de manera que, en particular, $g \subseteq g'$ y, en consecuencia, $(x, y), (x, y') \in g$: como g es una función $V' \to \mathbb{R}$, esto nos permite concluir que $y = y'$. Todo esto prueba que h es una función $W \to \mathbb{R}$.

- Sean x e y dos elementos de W y sean a y b dos escalares de $\mathbb{R}$. Hay dos elementos (V, g) y (V', g') en $\mathscr{C}$ tales que $x \in V$ e $y \in V'$, y como $\mathscr{C}$ está totalmente ordenado con respecto a $\preceq$ es $(V, g) \preceq (V', g')$ o $(V', g') \preceq (V, g)$ y podemos, de hecho, que vale la primera de estas dos alternativas. En ese caso tenemos que x e y están en V', así que $ax + by$ también está en V', y, por lo tanto, los pares ordenados $(x, g(x))$, $(y, g(y)$ y $(ax+by, ag(x)+bg(y)$ están en g y, en consecuencia, en h. Se sigue de esto que

  $$h(ax + by) = ag(x) + bg(y) = ah(x) + bh(y).$$

  La función h es, por todo esto, lineal.
- Como $\mathscr{C}$ no es vacío, podemos elegir un elemento (V_0, g_0) en $\mathscr{C}$. Es $U \subseteq V_0 \subseteq W$. Por otro lado, si x es un elemento de U, entonces $g(x) = g|_U(x) = f(x)$, así que $(x, f(x)) \in g \subseteq h$ y, por lo tanto, $h(x) = f(x)$: esto nos dice que $h|_U = f$. Finalmente, si x es un elemento de W, entonces hay un elemento (V, g) de $\mathscr{C}$ tal que $x \in V$ y, por lo tanto, $h(x) = g(x) \leq p(x)$. Esto prueba que el par (W, h) pertenece al conjunto $\mathscr{S}$.
- Sea, finalmente, (V, g) un elemento de $\mathscr{C}$. Por la forma en que definimos a W y a h, es claro que $V \subseteq W$ y que $g \subseteq h$, y esto último implica que $g = g \cap (V \times \mathbb{R}) = h \cap (V \times \mathbb{R}) = h|_V$. Esto nos dice que $(V, g) \preceq (W, h)$ y, en definitiva, que (W, h) es una cota superior para $\mathscr{C}$ en $\mathscr{S}$.

Todo esto prueba que el conjunto ordenado $\mathscr{S}$ satisface las condiciones del *Lema de Zorn* 2.2.6 y que, por lo tanto, ese conjunto ordenado posee elementos maximales. Sea (V, g) uno de ellos.

Supongamos por un momento que V es un subespacio propio de X, de manera que hay un elemento x_0 de X que no está en V. Consideremos el subespacio

$$V' \coloneqq V + \langle x_0 \rangle, \tag{12.1}$$

y notemos que esa suma es directa. Si x e y son dos elementos de V, entonces

$$g(x) + g(y) = g(x + y) \leq p(x + y) = p(x - x_0 + y + x_0) \leq p(x - x_0) + p(y + x_0),$$

así que

$$g(x) - p(x - x_0) \leq p(y + x_0) - g(y). \tag{12.2}$$

Esto nos dice, entre otras cosas, que el conjunto $\{g(x) - p(x - x_0) : x \in V\}$ está acotado superiormente. Como no vacío, podemos considerar su supremo: escribámoslo α. La desigualdad (12.2)

nos dice que α es una cota inferior para el conjunto $\{p(y+y_0)-g(y) : y \in V\}$, y entonces tenemos que

$$g(x)-\alpha \leq p(x-x_0), \qquad g(x)+\alpha \leq p(x+x_0) \tag{12.3}$$

cualquiera sea x en V. Como la suma de (12.1) es directa, hay exactamente una función lineal $g' : V' \to \mathbb{R}$ tal que $g'(x) = g(x)$ para todo $x \in V$ y $g'(x_0) = \alpha$. Es claro que $V \subseteq V'$ y que $g'|_V = g$. Sea y un elemento de V', de manera que hay un elemento x de V y un número real t tal que $y = x + tx_0$. Si $t > 0$, entonces la segunda desigualdad de (12.3) nos dice que

$$g'(y) = g'(x+tx_0) = g(x)+t\alpha = tg(t^{-1}x)+t\alpha \leq tp(t^{-1}x+x_0) = p(x+tx_0) = p(y),$$

si $t < 0$, entonces la primera desigualdad de (12.3) nos dice que

$$g'(y) = g'(x+tx_0) = g(x)+t\alpha = -tg(-t^{-1}x)+t\alpha \leq -tp(-t^{-1}x-x_0) = p(x+tx_0) = p(y),$$

y si $t = 0$, entonces $g'(y) = g'(x) = g(x) \leq p(x) = p(y)$. Vemos así que el par (V', g') es un elemento de $\mathscr{S}$ tal que $(V, g) \prec (V', g')$, y esto es imposible porque (V, g) es un elemento maximal de $\mathscr{S}$. Vemos así que debe ser $V = X$.

Hemos probado que hay un elemento en $\mathscr{S}$ de la forma (X, g), de manera que hay una función lineal $g : X \to \mathbb{R}$ tal que $g|_U = f$ y $g(x) \leq p(x)$ para todo $x \in X$. Esto es precisamente lo que afirma la proposición. □

Usando esta proposición podemos probar el siguiente resultado, uno de los teoremas fundamentales del análisis funcional, el *Teorema de Hahn–Banach*, por Hans Hahn y Stefan Banach, que lo probaron de manera independiente en [Hah1927] y en [Ban1929a, Ban1929b].

Proposición 12.3.2. *Sea X un espacio vectorial, sea $p : X \to \mathbb{R}$ una seminorma sobre X, y sea U un subespacio de X. Si $f : U \to \Bbbk$ es una función lineal tal que $|f(x)| \leq p(x)$ para todo $x \in U$, entonces hay una función lineal $\bar{f} : X \to \Bbbk$ tal que $\bar{f}|_U = f$ y $|\bar{f}(x)| \leq p(x)$ para todo $x \in X$.*

Notemos que aquí el espacio X puede ser real o complejo.

Demostración. Supongamos primerp que X es un espacio vectorial *real*. Sea $f : U \to \Bbbk$ una función lineal tal que $|f(x)| \leq p(x)$ para todo $x \in U$. Como p es una seminorma sobre X, satisface las condiciones de la Proposición 12.3.1. La hipótesis sobre f implica que para todo $x \in U$ es $f(x) \leq |f(x)| \leq p(x)$ y entonces esa proposición nos dice que hay una función lineal $\bar{f} : X \to \mathbb{R}$ que extiende a f y tal que $\bar{f}(x) \leq p(x)$ para todo $x \in U$. Notemos que también es $-\bar{f}(x) = \bar{f}(-x) \leq p(-x) = p(x)$ para todo $x \in X$, así que, de hecho, $|\bar{f}(x)| \leq p(x)$ para todo $x \in X$. Esto prueba la proposición en el caso en que el espacio vectorial X es real.

Supongamos ahora que X es un espacio vectorial *complejo* y escribamos $X_{\mathbb{R}}$ al espacio vectorial real que se obtiene de X restringiendo escalares de $\mathbb{C}$ a $\mathbb{R}$. Podemos ver a la función $p : X \to \mathbb{R}$ como una función $p : X_{\mathbb{R}} \to \mathbb{R}$, por supuesto, y es claro que esta es una seminorma sobre el espacio real $X_{\mathbb{R}}$. El $\mathbb{C}$-subespacio U de X es también un $\mathbb{R}$-subespacio de $X_{\mathbb{R}}$, y como la función $f : U \to \mathbb{C}$ es $\mathbb{C}$-lineal, la función

$$\mathrm{Re}(f) : x \in U \mapsto \mathrm{Re}(f(x)) \in \mathbb{R}$$

es $\mathbb{R}$-lineal. Más aún, para cada elemento x de U tenemos que

$$|\mathrm{Re}(f)(x)| = |\mathrm{Re}(f(x))| \leq |f(x)| \leq p(x).$$

La Proposición 12.3.1 nospermite concluir que hay una función $\mathbb{R}$-lineal $g : X_{\mathbb{R}} \to \mathbb{R}$ tal que $g|_U = \mathrm{Re}(f)$ y $g(x) \leq p(x)$ para todo $x \in X$. Consideremos ahora la función

$$\bar{f} : x \in X \mapsto g(x) - ig(ix) \in \mathbb{C}.$$

Esta función es $\mathbb{C}$-lineal: si x e y son dos elementos de X y a y b son dos números complejos, con partes reales a_1 y b_1 y partes imaginarias a_2 y b_2, respectivamente, entonces

$$\begin{aligned}
\bar{f}(ax + by) &= g(ax + by) - ig(iax + iby) \\
&= g(a_1x + ia_2x + b_1y + ib_2y) - ig(ia_1x - a_2x + ib_1y - b_2y) \\
&= a_1g(x) + a_2g(ix) + b_1g(y) + b_2g(iy) - ia_1g(ix) + ia_2g(x) - ib_1g(iy) + ib_2g(y) \\
&= a_1(g(x) - ig(ix)) + a_2(g(ix) + ig(x)) + b_1(g(y) - ig(iy)) + b_2(g(iy) + ig(y)) \\
&= a_1\bar{f}(x) + a_2i\bar{f}(x) + b_1\bar{f}(y) + b_2if(y) \\
&= a\bar{f}(x) + b\bar{f}(y).
\end{aligned}$$

Por otro lado, si x es un elemento de X, entonces hay un número complejo α tal que $|\alpha| = 1$ y $|\bar{f}(x)| = \alpha\bar{f}(x) = \bar{f}(\alpha x) = g(\alpha x) - ig(i\alpha x)$, así que el número $g(i\alpha x)$ es nulo y, por lo tanto,

$$|\bar{f}(x)| = \alpha\bar{f}(x) = \bar{f}(\alpha x) = g(\alpha x) - ig(i\alpha x) = g(\alpha x) \leq p(\alpha x) = p(x).$$

Esto prueba la proposición en el caso en el que el espacio vectorial X es complejo. □

El siguiente es el caso especial más importante de la Proposición 12.3.2.

Proposición 12.3.3. *Sea X un espacio normado, sea U un subespacio de X, y sea $f : U \to \mathbb{k}$ una función lineal y continua. Hay una extensión $\bar{f} : X \to \mathbb{k}$ lineal y continua tal que $\|\bar{f}\|_{X'} = \|f\|_{U'}$.*

Demostración. Como la función $f : U \to \mathbb{k}$ es continua, tenemos que $|f(x)| \leq \|f\|_{U'} \cdot \|x\|$ para todo $x \in U$. La función $p : x \in X \mapsto \|f\|_{U'} \cdot \|x\| \in \mathbb{R}$ es claramente una seminorma sobre X y es $|f(x)| \leq p(x)$ para todo elemento x de U, así que la Proposición 12.3.2 nos dice que hay una función lineal $\tilde{f} : X \to \mathbb{k}$ que extiende a f y tal que $|f(x)| \leq p(x) = \|f\|_{U'} \cdot \|x\|$ para todo elemento x de X. Esta desigualdad nos dice que la función $\tilde{f}$ es continua y que su norma en X' es tal que

$$\|\tilde{f}\|_{X'} \leq \|f\|_{U'}.$$

De hecho, vale aquí la igualdad: si t es un número real menor que $\|f\|_{U'}$, entonces hay un elemento x de U tal que $\|x\| \leq 1$ y $|f(x)| > t$, y como $f(x) = \tilde{f}(x)$, vemos que $\|\tilde{f}\|_{X'} > t$. □

Una consecuencia de esto es que el espacio dual de un espacio normado separa puntos:

Corolario 12.3.4. *Sea X un espacio normado. Si x e y son dos elementos distintos de X, entonces hay un elemento ϕ de X' tal que $\phi(x) \neq \phi(y)$.*

Demostración. Sean x e y dos elementos distintos de X, de manera que $x - y \neq 0$ y el subespacio $U \coloneqq \langle x - y \rangle$ de X tiene dimensión 1. Hay exactamente una función lineal $f : U \to \mathbb{k}$ tal que $f(x - y) = 1$, y es continua, ya que U tiene dimensión finita. La Proposición 12.3.3 nos dice entonces hay que una función lineal y continua $\phi : X \to \mathbb{k}$ que extiende a f y, por lo tanto, tal que $\phi(x) - \phi(y) = \phi(x - y) = f(x - y) \neq 0$. Esto prueba el corolario. □

Observemos que en álgebra lineal probamos un resultado completamente similar: si V es un espacio vectorial cualquiera, entonces el espacio dual V^* de V separa puntos en V. Lo que dice el corolario es que cuando V es un espacio normado entonces es suficiente el espacio dual continuo para separar puntos de V.

El siguiente resultado es una forma distinta del Corolario 12.3.4.

Proposición 12.3.5. *Sea X un espacio normado y sea x un elemento no nulo de X. Hay una función lineal y continua $\phi : X \to \mathbb{k}$ tal que $\phi(x) = \|x\|$ y $\|\phi\|_{X'} = 1$.*

Demostración. Consideremos el espacio $U \coloneqq \langle x \rangle$, que tiene dimensión 1. Hay exactamente una función lineal $f : U \to \mathbb{k}$ tal que $f(x) = \|x\|$ y es continua porque U tiene dimensión finita. De hecho, como $|f(tx)| = \|tx\|$ para todo $t \in \mathbb{k}$, tenemos que $\|f\|_{U'} = 1$. La Proposición 12.3.3 nos dice entonces que hay una función lineal y continua $\phi : X \to \mathbb{k}$ que extiende a f y tal que $\|\phi\|_{X'} = \|f\|_{U'} = 1$. □

Una consecuencia interesante de esta proposición es la siguiente observación que nos dice que el espacio $\ell^\infty_{\mathbb{k}}$ contiene copias de todos los espacios normados separables.

Proposición 12.3.6. *Todo espacio normado separable es linealmente isométrico a un subespacio de $\ell^\infty_{\mathbb{k}}$.*

En particular, esto nos dice que el espacio $\ell^\infty_{\mathbb{k}}$ es extremadamente complicado!

Demostración. Sea X un espacio normado separable. Su esfera unidad

$$S \coloneqq \{x \in X : \|x\|_X = 1\}$$

es entonces también un espacio métrico separable y hay, por lo tanto, una sucesión $(x_n)_{n\geq 0}$ con valores en S tal que la clausura del conjunto $\{x_n : n \in \mathbb{N}_0\}$ es S; notemos que S es cerrado. De acuerdo a la Proposición 12.3.5, hay también una sucesión $(\phi_n)_{n\geq 0}$ en X' tal que $\phi_n(x_n) = \|x_n\| = 1$ y $\|\phi_n\|_{X'} = 1$ para todo $n \in \mathbb{N}_0$.

Si x es un elemento cualquiera de X, entonces $|\phi_n(x)| \leq \|\phi_n\|_{X'} \cdot \|x\|_X = \|x\|_X$, así que la sucesión $(\phi_n(x))_{n\geq 0}$ es un elemento de $\ell^\infty_{\mathbb{k}}$ de norma menor o igual que $\|x\|_X$. Esto nos dice que hay una función

$$\alpha : x \in X \mapsto (\phi_n(x))_{n\geq 0} \in \ell^\infty_{\mathbb{k}},$$

que claramente es lineal, y que $\|\alpha(x)\|_{\ell^\infty_{\mathbb{k}}} \leq \|x\|_X$ para todo $x \in X$, así esta función α es continua y tiene norma

$$\|\alpha\|_{L(X,\ell^\infty_{\mathbb{k}})} \leq 1.$$

Para probar la proposición es suficiente que mostremos que α es una inyección isométrica.

Sea x un elemento de X. Sabemos que $\|\alpha(x)\|_{\ell^\infty_{\mathbb{k}}} \leq \|x\|_X$ y tenemos que probar que vale la igualdad. Si x es nulo no hay nada que hacer, así que supongamos que no lo es. Para cada número positivo ε menor que 1 hay un entero positivo n tal que $\|x_n - x/\|x\|_X\|_X < \varepsilon$, y es

$$\left|\phi_n\left(\frac{x}{\|x\|_X}\right) - 1\right| = \left|\phi_n\left(\frac{x}{\|x\|_X}\right) - \phi_n(x_n)\right| \leq \|\phi_n\|_{X'} \cdot \left\|\frac{x}{\|x\|_X} - x_n\right\| < \varepsilon,$$

de manera que

$$\left\|\alpha\left(\frac{x}{\|x\|_X}\right)\right\|_{\ell^\infty_{\mathbb{k}}} = \sup_{m\in\mathbb{N}_0}\left|\phi_n\left(\frac{x}{\|x\|_X}\right)\right| \geq \left|\phi_n\left(\frac{x}{\|x\|_X}\right)\right| \geq 1 - \varepsilon.$$

Esto implica que $\|\alpha(x/\|x\|_X)\|_{\ell^\infty_{\mathbb{k}}} \geq 1$ y, por lo tanto, que $\|\alpha(x)\|_{\ell^\infty_{\mathbb{k}}} \geq \|x\|_X$, como queremos. □

Podemos generalizar esto a espacios normados no separables, si reemplazamos a $\ell^\infty_{\mathbb{k}}$ por espacios adecuados.

Ejercicio 12.3.7. Sea X un espacio normado y sea D un subconjunto denso de X. Pruebe que hay una inyección isométrica $X \to \ell^\infty_{\Bbbk}(D)$ con codominio en el espacio normado de todas las funciones acotadas $D \to \Bbbk$, como en el Ejercicio 5.2.3.23.

El Corolario 12.3.4 nos dice que podemos separar puntos de un espacio normado usando elementos de su dual. Nuestro siguiente resultado es muy similar: nos dice que podemos separar los puntos del complemento de un subespacio cerrado de un espacio normado de este.

Proposición 12.3.8. *Sea X un espacio normado y sea Y un subespacio cerrado y propio de X. Hay un a función lineal y continua $\phi : X \to \Bbbk$ tal que $\|\phi\| = 1$ y $\phi(y) = 0$ para todo elemento y de Y.*

Demostración. Como Y es un subespacio propio, hay un elemento x_0 de X tal que $\|x_0\|_X = 1$ y que no está en Y, y como Y es cerrado, el número $r \coloneqq d(Y, x_0)$ es positivo. Consideremos el subespacio $Z \coloneqq Y + \langle x_0 \rangle$. Como x_0 no está en Y, esta suma es directa y hay, por lo tanto, una función lineal $\psi : Z \to \Bbbk$ tal que $\psi(y + tx_0) = t$ para toda elección de y en Y y de t en $\Bbbk$.

Sea z un elemento de Z, de manera que hay un elemento y_0 de Y y un escalar t de $\Bbbk$ tal que $z = y_0 + yx_0$. Como Y es un subspacio de X, tenemos que

$$\begin{aligned} |t| \cdot r = |t| \cdot \text{ínf}\{\|y - x_0\|_X y \in Y\} = \text{ínf}\{\|ty - tx_0\|_X y \in Y\} = \text{ínf}\{\|y - tx_0\|_X y \in Y\} \\ \leq \|y_0 + tx_0\|_X = \|z\|_X \end{aligned}$$

y, por lo tanto, $|\psi(z)| = |t| \leq r^{-1} \cdot \|z\|_X$. Esto nos dice que la función $\psi : Z \to \Bbbk$ es lineal y continua, y la Proposición 12.3.3 nos permite concluir que hay una función lineal y continua $\bar{\psi} : X \to \Bbbk$ que extiende a ψ. Como ψ no es nula, $\bar{\psi}$ tampoco lo es, y es claro que la función $\phi \coloneqq \bar{\psi}/\|\bar{\psi}\|_{X'}$ se anula sobre Y y tiene norma 1. □

Una de las consecuencias más directas del hecho de que el espacio dual de un espacio vectorial V separa puntos en V es que un cierto homorfismo canónico $V \to V^{**}$ de V a su doble dual es inyectivo. Aquí podemos probar lo mismo pero para los duales continuos.

Proposición 12.3.9. *Sea X un espacio normado. Hay exactamente una función lineal $\Phi : X \to X''$ tal que*

$$\Phi(x)(\phi) = \phi(x)$$

cualesquiera sean el vector x de W y el funcional ϕ de V', y se trata de una inyección isométrica.

Observemos que como X es un espacio normado, entonces X' y, por lo tanto, también X'' son espacios normados: es por esto que la afirmación de esta proposición tiene sentido.

Demostración. Sea x un elemento de X. La función $\Phi_x : \phi \in X' \mapsto \phi(x) \in \mathbb{k}$ es claramente lineal y para todo elemento ϕ de X' tiene $|\Phi_x(\phi)| = |\phi(x)| \leq \|\phi\|_{X'} \cdot \|x\|_X$, así que es continua y su norma en X'' es tal que $\|\Phi_x\|_{X''} \leq \|x\|_X$. De hecho, vale aquí la igualdad: la Proposición 12.3.5 nos dice que hay un elemento ϕ de X' tal que $\phi(x) = \|x\|_X$ y $\|\phi\|_{X'} = 1$, y es $|\Phi_x(\phi)| = |\phi(x)| = \|x\|_X = \|x\|_X \cdot \|\phi\|_{X'}$, de manera que $\|\Phi_x\| \geq \|x\|_X$.

Es fácil ver que la función $\Phi : x \in X \mapsto \Phi_x \in X''$ es lineal, y lo que acabamos de hacer nos dice que $\|\Phi(x)\|_{X''} = \|x\|_X$ para todo $x \in X$, así que Φ es continua y, más aún, es una inyección isométrica. Como claramente $\Phi(x)(\phi) = \phi(x)$ para todo $x \in X$ y todo ϕ en Φ, esto prueba la proposición. □

Esta proposición nos dice que para todo espacio normado X tenemos una inyección isométrica $X \to X''$ de X a su doble dual continua. Esta función es, por supuesto, inyectiva, pero no es necesariamente sobreyectiva.

Definición 12.3.10. Decimos que un espacio normado X es ***reflexivo*** si la inyección isométrica canónica $X \to X''$ es sobreyectiva.

Es importante que la condición aquí es que la inyección isométrica canónica sea un isomorfismo y no simplemente que exista una isometría $X \to X''$. En efecto, Robert James mostró en [Jam1951] un ejemplo de un espacio de Banach no reflexivo que es isométrico a su doble dual.

Es fácil dar ejemplos no particularmente interesantes:

Proposición 12.3.11. *Todo espacio normado de dimensión finita es reflexivo.*

Demostración. Sea X un espacio normado de dimensión finita. El álgebra lineal nos dice que hay exactamente una función lineal $\Phi : X \to X^{**}$ tal que $\Phi(x)(\phi) = \phi(x)$ siempre que $x \in X$ y $\phi \in X^*$, y que es un isomorfismo. Como X tiene dimensión finita, entonces $X' = X^*$ y $X'' = X^{**}$ y, por lo tanto, esta función Φ es la misma que la de la Proposición 12.3.9, así que esta última es sobreyectiva y el espacio X es reflexivo. □

Queremos ahora presentar una familia importante de ejemplos de espacios normados reflexivos de dimensión infinitai: los espacios $\ell^p_{\mathbb{k}}$ con $p \in (1, +\infty)$. Lo primero que tenemos que hacer para ello es identificar sus espacios dualed continuos y lo hacemos en dos pasos. La siguiente proposición exhibe muchos el elementos del espacio $(\ell^p_{\mathbb{k}})'$.

Proposición 12.3.12. *Sea p un elemento de $(1, +\infty)$ y sea q el único elemento de $(1, +\infty)$ tal que*

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

(*i*) *Si* $x = (x_n)_{n\geq 0}$ *e* $y = (y_n)_{n\geq 0}$ *son un elemento de* $\ell^p_{\Bbbk}$ *y de* $\ell^q_{\Bbbk}$*, respectivamente, entonces la serie*

$$\sum_{n=0}^{\infty} x_n y_n$$

converge absolutamente en $\Bbbk$*.*

(*ii*) *Sea* $y = (y_n)_{n\geq 0}$ *un elemento de* $\ell^q_{\Bbbk}$*. Hay una función*

$$\phi_y : (x_n)_{n\geq 0} \in \ell^p_{\Bbbk} \mapsto \sum_{n=0}^{\infty} x_n y_n \in \Bbbk,$$

es lineal y continua, y su norma es $\|\phi_y\|_{(\ell^p_{\Bbbk})'} = \|y\|_{\ell^q_{\Bbbk}}$*.*

Demostración. (*i*) Sean $x = (x_n)_{n\geq 0}$ e $y = (y_n)_{n\geq 0}$ un elemento de $\ell^p_{\Bbbk}$ y uno de $\ell^q_{\Bbbk}$, respectivamente. De acuerdo a la desigualdad de Hölder que vimos en el Ejercicio 5.2.3.13, para todo $N \in \mathbb{N}_0$ es

$$\sum_{n=0}^{N} |x_n y_n| \leq \left(\sum_{i=0}^{N} |x_n|^p\right)^{1/p} \left(\sum_{i=0}^{N} |y_n|^q\right)^{1/q} \leq \|x\|_{\ell^p_{\Bbbk}} \cdot \|y\|_{\ell^q_{\Bbbk}}, \tag{12.4}$$

y esto implica que la serie $\sum_{n=0}^{\infty} |x_n y_n|$ tiene sus sumas parciales acotadas, de manera que converge y, por lo tanto, que la serie $\sum_{n=0}^{\infty} x_n y_n$ converge absolutamente.

(*ii*) Sea ahora $y = (y_n)_{n\geq 0}$ un elemento de $\ell^q_{\Bbbk}$. Que hay efectivamente una función ϕ_y como en el enunciado es consecuencia directa de la afirmación (*i*) que acabamos de probar. Esa función es lineal. En efecto, si $x = (x_n)_{n\geq 0}$ y $x' = (x'_n)_{n\geq 0}$ son dos elementos de $\ell^p_{\Bbbk}$ y a y b son dos escalares de $\Bbbk$, entonces sabemos que $ax + bx' = (ax_n + bx'_n)_{n\geq 0}$ es un elemento de $\ell^p_{\Bbbk}$ y, por lo tanto, que las series

$$\sum_{n=0}^{\infty} x_n y_n, \qquad \sum_{n=0}^{\infty} x'_n y_n \qquad \sum_{n=0}^{\infty} (ax_n + bx'_n) y_n$$

convergen, de manera que

$$\phi_y(ax + bx') = \sum_{n=0}^{\infty} (ax_n + bx'_n) y_n = a \sum_{n=0}^{\infty} x_n y_n + b \sum_{n=0}^{\infty} x'_n y_n = a\phi_y(x) + b\phi_y(x').$$

Por otro lado, para cada elemento $x = (x_n)_{n\geq 0}$ la desigualdad (12.4) nos permite ver que

$$|\phi_y(x)| = \left|\sum_{n=0}^{\infty} x_n y_n\right| \leq \sum_{n=0}^{\infty} |x_n y_n| = \lim_{N\to\infty} \sum_{n=0}^{\infty} |x_n y_n| \leq \|x\|_{\ell^p_{\Bbbk}} \cdot \|y\|_{\ell^q_{\Bbbk}},$$

de manera que la función ϕ_y es continua y es $\|\phi_y\|_{(\ell^p_{\Bbbk})'} \leq \|y\|_{\ell^q_{\Bbbk}}$. Mostremos que, de hecho, vale aquí la igualdad.

Para cada $n \in \mathbb{N}_0$ elijamos un número complejo α_n tal que $|\alpha_n| = 1$ y $\alpha_n y_n = |y_n|$, y para cada entero no negativo m consideremos la sucesión $x_m = (x_{m,n})_{n\geq 0}$ en $\Bbbk$ que para cada $n \in \mathbb{N}_0$ tiene

$$x_{m,n} = \begin{cases} \alpha_n |y_n|^{q/p} & \text{si } n \leq m; \\ 0 & \text{si } n > m. \end{cases}$$

Es evidente que si m es un entero no negativo, entonces x_m es un elemento de $\ell^p_\Bbbk$, ya que tiene finitas componentes no nulas, y que

$$\|x_m\|_{\ell^p_\Bbbk} = \left(\sum_{n=0}^{\infty} |x_{m,n}|^p\right)^{1/p} = \left(\sum_{n=0}^{m} |\alpha_n|^p |y_n|^q\right)^{1/p} = \left(\sum_{n=0}^{m} |y_n|^q\right)^{1/p},$$

de manera que

$$\lim_{m\to\infty} \|x_m\|_{\ell^p_\Bbbk} = \|y\|^{q/p}_{\ell^q_\Bbbk}.$$

Por otro lado, si m es un elemento cualquiera de $\mathbb{N}_0$, entonces

$$\phi_y(x_m) = \sum_{n=0}^{\infty} x_{m,n} y_n = \sum_{n=0}^{m} \alpha_n |y_n|^{q/p} y_n = \sum_{n=0}^{m} |y_n|^{1+q/p} = \sum_{n=0}^{m} |y_n|^q,$$

así que

$$\|\phi_y\|_{(\ell^p_\Bbbk)'} \cdot \|y\|^{q/p}_{\ell^q_\Bbbk} = \lim_{m\to\infty} \|\phi_y\|_{(\ell^p_\Bbbk)'} \cdot \|x_m\|_{\ell^p_\Bbbk} \geq \lim_{m\to\infty} |\phi_y(x_m)| = \|y\|^q_{\ell^q_\Bbbk}$$

y, por lo tanto, $\|\phi_y\|_{(\ell^p_\Bbbk)'} \geq \|y\|^{q-q/p}_{\ell^q_\Bbbk} \geq \|y\|_{\ell^q_\Bbbk}$. Podemos concluir entonces que $\|\phi_y\|_{(\ell^p_\Bbbk)'} = \|y\|_{\ell^q_\Bbbk}$, como dijimos. □

Nuestro siguiente resultado nos dice que los elementos del dual de $\ell^p_\Bbbk$ que describe la Proposición 12.3.12 son, de hecho, todos.

Proposición 12.3.13. *Sea p un elemento de $(1, +\infty)$, sea q el único elemento de $(1, +\infty)$ tal que $1/p + 1/q = 1$, y para cada elemento y de $\ell^q_\Bbbk$ sea $\phi_y : \ell^p_\Bbbk \to \Bbbk$ el funcional descripto en la Proposición 12.3.12. La función $\Phi : y \in \ell^q_\Bbbk \mapsto \phi_y \in (\ell^p_\Bbbk)'$ es una isometría lineal.*

Demostración. Es fácil verificar que esa función Φ es lineal y, como la Proposición 12.3.12 nos dice que $\|\phi_y\|_{(\ell^p_\Bbbk)'} = \|y\|_{\ell^q_\Bbbk}$ cualquiera sea el elemento y de $\ell^q_\Bbbk$, esa función es una inyección isométrica. Para probar la proposición, entonces, tenemos que probar que es además sobreyectiva.

Sea $\phi : \ell^p_\Bbbk \to \Bbbk$ una función lineal y continua. Para cada $n \in \mathbb{N}_0$ sea e_n el elemento de $\ell^p_\Bbbk$ cuya única componente no nula es la n-ésima, que es igual a 1, y consideremos el escalar $y_n \coloneqq \phi(e_n)$.

Si $x = (x_n)_{n\geq 0}$ es un elemento arbitrario de $\ell^p_{\mathbb{k}}$, entonces para cada $m \in \mathbb{N}_0$ es

$$\left\| x - \sum_{n=0}^{m} x_n e_n \right\|_{\ell^p_{\mathbb{k}}} = \left(\sum_{n=m+1}^{\infty} |x_n|^p \right)^{1/p}$$

así que, como la serie $\sum_{n=0}^{\infty} |x_n|^p$ converge y la función $t \in [0, +\infty) \to t^{1/p} \in \mathbb{R}$ es continua, tenemos que

$$x = \lim_{m\to\infty} \sum_{n=0}^{m} x_n e_n$$

y, como la función ϕ es lineal y continua, que también

$$\phi(x) = \phi\left(\lim_{m\to\infty} \sum_{n=0}^{m} x_n e_n \right) = \lim_{m\to\infty} \sum_{n=0}^{m} x_n \phi(e_n) = \lim_{m\to\infty} \sum_{n=0}^{m} x_n y_n.$$

Vemos así que para todo $x = (x_n)_{n\geq 0}$ en $\ell^p_{\mathbb{k}}$ la serie $\sum_{n=0}^{\infty} x_n y_n$ converge y, más aún, es

$$\phi(x) = \sum_{n=0}^{\infty} x_n y_n.$$

Para probar que la función Φ del enunciado es sobreyectiva, entonces, es suficiente que mostremos que la sucesión $y = (y_n)_{n\geq 0}$ pertenece a $\ell^q_{\mathbb{k}}$.

Para cada $n \in \mathbb{N}_0$ hay un número complejo α_n tal que $|\alpha_n| = 1$ y $\alpha_n y_n = |y_n|$, y podemos considerar el número $z_n \coloneqq \alpha_n |y_n|^{q-1}$. Si m es un elemento cualquiera de $\mathbb{N}_0$, entonces

$$\begin{aligned}
\sum_{n=0}^{m} |y_n|^q &= \left| \sum_{n=0}^{m} |y_n|^q \right| = \left| \sum_{n=0}^{m} \alpha_n |y_n|^{q-1} y_n \right| = \left| \sum_{n=0}^{m} z_n \phi(e_n) \right| = \left| \phi\left(\sum_{n=0}^{m} z_n e_n \right) \right| \\
&\leq \|\phi\| \cdot \left\| \sum_{n=0}^{m} z_n e_n \right\|_{\ell^p_{\mathbb{k}}} = \|\phi\|_{(\ell^p_{\mathbb{k}})'} \cdot \left(\sum_{n=0}^{m} |z_n|^p \right)^{1/p} = \|\phi\|_{(\ell^p_{\mathbb{k}})'} \cdot \left(\sum_{n=0}^{m} |y_n|^{(q-1)p} \right)^{1/p} \\
&= \|\phi\|_{(\ell^p_{\mathbb{k}})'} \cdot \left(\sum_{n=0}^{m} |y_n|^q \right)^{1/p},
\end{aligned}$$

así que

$$\left(\sum_{n=0}^{m} |y_n|^q \right)^{1/q} = \left(\sum_{n=0}^{m} |y_n|^q \right)^{1-1/p} \leq \|\phi\|_{(\ell^p_{\mathbb{k}})'}.$$

Esto implica claramente que la serie $\sum_{n=0}^{\infty} |y_n|^q$ converge, de manera que y pertenece a $\ell^q_{\mathbb{k}}$, como queremos. □

La segunda observación que tenemos que hacer es puramente formal.

Proposición 12.3.14. *Sean X e Y dos espacios normados y sea $f : X \to Y$ una función lineal y continua.*

(*i*) *Si $\phi : Y \to \mathbb{k}$ es un elemento de Y', entonces $\phi \circ f : X \to \mathbb{k}$ es un elemento de X' y es*

$$\|\phi \circ f\|_{X'} \leq \|\phi\|_{Y'} \cdot \|f\|_{L(X,Y)}.$$

(*ii*) *La función $f^t : \phi \in Y' \mapsto \phi \circ f \in X'$ es lineal y continua, y es*

$$\|f^t\|_{L(Y',X')} = \|f\|_{L(X,Y)}.$$

Llamamos a la función f^t la ***función transpuesta*** de f.

Demostración. (*i*) Que la función $\phi \circ f$ es un elemento de X' es claro. Si ϕ es un elemento de Y' y x uno de X, entonces

$$|(\phi \circ f)(x)| = |\phi(f(x))| \leq \|\phi\|_{Y'} \cdot \|f(x)\|_Y \leq \|\phi\|_{Y'} \cdot \|f\|_{L(X,Y)} \cdot \|x\|_X,$$

y esto nos dice que $\|\phi \circ f\|_{X'} \leq \|\phi\|_{Y'} \cdot \|f\|_{L(X,Y)}$.

(*ii*) Que la función f^t es lineal es claro, y es continua porque para todo elemento ϕ de Y' es

$$\|f^t(\phi)\|_{X'} = \|\phi \circ f\|_{X'} \leq \|f\|_{L(X,Y)} \cdot \|\phi\|_{Y'},$$

como acabamos de ver. Esto nos dice, más aún, que $\|f^t\|_{L(Y',X')} \leq \|f\|_{L(X,Y)}$. Para terminar, probemos que aquí vale, de hecho, la igualdad. Por supuesto, esto es evidente si $\|f\|_{L(X,Y)}$ es nulo, así que basta que consideremos el caso en el que no lo es.

Sea c un número positivo menor que $\|f\|_{L(X,Y)}$, de manera que hay un elemento x de X tal que $\|x\|_X \leq 1$ y $\|f(x)\|_Y > c$. De acuerdo a la Proposición 12.3.5, hay un elemento ϕ en Y' tal que $\phi(f(x)) = \|f(x)\|_Y$ y $\|\phi\|_{Y'} = 1$, y entonces

$$\|f^t\|_{L(Y',X')} \geq \|f^t(\phi)\|_{X'} \geq |f^t(\phi)(x)| = |\phi(f(x))| = \|f(x)\|_Y > c.$$

Esto prueba que $\|f^t\|_{L(Y',X')} \geq \|f\|_{L(X,Y)}$, como queremos. □

Esta operación de transposición tiene las propiedades usuales del álgebra lineal.

Ejercicio 12.3.15. Sean X, Y, y Z espacios normados.

(a) La función transpuesta de la función identidad $\mathrm{id}_X : X \to X$ es la función identidad $\mathrm{id}_{X'} : X' \to X'$.

(b) Si $f : X \to Y$ y $g : Y \to Z$ son funciones lineales y continuas, entonces $g \circ f : X \to Z$ también lo es y su función transpuesta es $(g \circ f)^f = f^t \circ g^t$.

(c) Si $f : X \to Y$ es un homeomorfismo lineal, entonces la función transpuesta $f^t : Y' \to X'$ también lo es.

(d) Si $f : X \to Y$ es una isometría lineal, entonces la función transpuesta $f^t : Y' \to X'$ también lo es.

La siguiente observación es propia de nuestro contexto.

Proposición 12.3.16. *Si $f : X \to Y$ es una isometría lineal entre espacios normados, entonces la función transpuesta $f^t : Y' \to X'$ también lo es.*

Demostración. Sea $f : X \to Y$ una isometría lineal entre espacios normados y sea ϕ un elemento de Y'. Como $\|f\|_{L(X,Y)} = 1$, la Proposición 12.3.14 nos dice que $\|f^t(\phi)\|_{X'} \leq \|\phi\|_{Y'}$. Tenemos que probar que vale aquí la igualdad. Esto es claro si ϕ es 0, así que podemos suponer que no lo es.

Sea c un número positivo menor que $\|\phi\|$. Hay entonces un elemento y de $\overline{B}_1^Y(0)$ tal que $|\phi(y)| > c$ y, como f es una isometría, hay un elemento x de $\overline{B}_1^X(0)$ tal que $f(x) = y$. Tenemos entonces que $|f^t(\phi)(x)| = |\phi(f(x))| = |\phi(y)| > c$, así que $\|f^t(\phi)\|_{X'} > c$. Vemos así que $\|f^t(\phi)\|_{X'} \geq \|\phi\|_{Y'}$, como queremos. □

Con todos estos resultados a nuestra disposición podemos dar el ejemplo que buscamos.

Proposición 12.3.17. *Sea p un elemento de $(1, +\infty)$. El espacio $\ell_{\mathbb{k}}^p$ es reflexivo.*

Demostración. Sea q el elemento de $(1, +\infty)$ tal que $1/p + 1/q = 1$. La Proposición 12.3.13 describe funciones

$$\Phi_p : \ell_{\mathbb{k}}^q \to (\ell_{\mathbb{k}}^p)', \qquad \Phi_q : \ell_{\mathbb{k}}^p \to (\ell_{\mathbb{k}}^q)',$$

que isometrías lineales, y la Proposición 12.3.16 nos dice que sus funciones transpuestas

$$\Phi_p^t : (\ell_{\mathbb{k}}^p)'' \to (\ell_{\mathbb{k}}^q)' \qquad \Phi_q^t : (\ell_{\mathbb{k}}^q)'' \to (\ell_{\mathbb{k}}^p)',$$

también son isometrías lineales, y se sigue de esto que la composición

$$\Psi \coloneqq (\Phi_p^t)^{-1} \circ \Phi_q : \ell_{\mathbb{k}}^p \to (\ell_{\mathbb{k}}^p)''$$

es una isometría lineal.

Sean $x = (x_n)_{n\geq 0}$ un elemento de $\ell_{\mathbb{k}}^p$ y ϕ uno de $(\ell_{\mathbb{k}}^p)'$. Como Φ_p es una isometría, hay un elemento $y = (y_n)_{n\geq 0}$ de $\ell_{\mathbb{k}}^q$ tal que $\phi = \Phi_p(y)$, y esto significa que para todo $z = (z_n)_{n\geq 0}$ en $\ell_{\mathbb{k}}^p$ es

$$\phi(z) = \sum_{n=0}^{\infty} y_n z_n.$$

Se sigue de esto que

$$\Psi(x)(\phi) = (\Phi_p^t)^{-1}(\Phi_q(x))(\phi) = (\Phi_p^{-1})^t(\Phi_q(x))(\phi) = \Phi_q(x)(\Phi_p^{-1}(\phi)) = \Phi_q(x)(y)$$

$$= \sum_{n\geq 0} x_n y_n = \phi(x)$$

y, por lo tanto, que la función Ψ coincide con la inyección isométrica canónica de la Proposición 12.3.9. Esta última es, en consecuencia, sobreyectiva y esto nos permite concluir que el espacio $\ell^p_{\mathbb{k}}$ es reflexivo. □

Para terminar con nuestra discusión de los espacios reflexivos, mostraremos un espacio normado que *no* es reflexivo. Empezamos con la siguiente observación.

Proposición 12.3.18. *Un espacio normado es separable si su dual continuo es separable.*

Demostración. Sea X un espacio normado y supongamos que su espacio dual continuo X' es separable. Sea

$$S \coloneqq \{\phi \in X' : \|\phi\|_{X'} = 1\}$$

la esfera unidad de X'. Como X es separable, S también lo es, y hay una sucesión $(\phi_n)_{n\geq 1}$ de elementos de S tal que la clausura de $\{\phi_n : n \in \mathbb{N}_0\}$ en X' contiene a S.

Para cada $n \in \mathbb{N}_0$ el funcional ϕ_n tiene norma 1, así que hay un elemento x_n de X tal que $\|x_n\|_X \leq 1$ y $|\phi_n(x_n)| > 1/2$. Si ponemos $y_n \coloneqq x_n/\|x_n\|$, lo que es posible porque ciertamente $x_n \neq 0$, entonces $\|y_n\| = 1$ y $|\phi(y_n)| = |\phi(x_n)|/\|x_n\| \geq |\phi(x_n)| > 1/2$. Sea Y la clausura del subespacio $\langle x_n : n \in \mathbb{N}_0\rangle$. Este subespacio Y es separable: el conjunto de las combinaciones lineales de elementos de $\{x_n : n \in \mathbb{N}_0\}$ con coeficientes racionales es contable y es denso en Y. Para probar la proposición, entonces, es suficiente que mostremos que $X = Y$.

Supongamos que ese no es el caso, de manera que Y es un subespacio cerrado y propio de X. De acuerdo a la Proposición 12.3.8, hay un elemento ϕ de X' tal que $\|\phi\|_{X'} = 1$ y $f(y) = 0$ para todo elemento y de T. Ahora bien, como ϕ tiene norma 1, pertenece a la esfera S y hay, por lo tanto, un entero positivo n tal que $\|\phi - \phi_n\| < 1/4$. Se sigue de esto que

$$\begin{aligned}
|\phi(y_n)| &= |\phi_n(y_n) - \phi_n(y_n) + \phi(y_n)| \\
&\geq |\phi_n(y_n)| - |\phi_n(y_n) - \phi(y_n)| \\
&\geq |\phi_n(y_n)| - \|\phi - \phi_n\| \cdot \|y_n\|_X \\
&> \frac{1}{2} - \frac{1}{4} > 0,
\end{aligned}$$

y esto es absurdo porque y_n está en Y y ϕ se anula sobre Y. Esta contradicción provino de haber supuesto que Y es un subespacio propio de X, así que debe ser $X = Y$, como queremos. □

Esto nos da una forma de mostrar ejemplos de espacios no reflexivos: basta encontrar un espacio normado separable cuyo espacio dual no es separable. La siguiente proposición nos da

uno.

Proposición 12.3.19. *Hay una isometría $\ell^\infty_{\mathbb{k}} \to (\ell^1_{\mathbb{k}})'$*

Demostración. Sea $x = (x_n)_{n\geq 0}$ un elemento de $\ell^\infty_{\mathbb{k}}$. Si $y = (y_n)_{n\geq 0}$ es un elemento de $\ell^1_{\mathbb{k}}$, entonces para todo m en $\mathbb{N}_0$ es

$$\sum_{n=0}^{n} |x_n y_n| = \sum_{n=0}^{n} |x_n| \cdot |y_n| \leq \|x\|_{\ell^\infty_{\mathbb{k}}} \cdot \sum_{n=0}^{n} |y_n| \leq \|x\|_{\ell^\infty_{\mathbb{k}}} \cdot \|y\|_{\ell^1_{\mathbb{k}}},$$

así que la serie $\sum_{n=0}^{\infty} x_n y_n$ converge absolutamente y su suma es tal que

$$\left| \sum_{n=0}^{n} x_n y_n \right| \leq \sum_{n=0}^{\infty} |x_n y_n| \leq \|x\|_{\ell^\infty_{\mathbb{k}}} \cdot \|y\|_{\ell^1_{\mathbb{k}}}. \tag{12.5}$$

Esto implica que hay una función

$$\phi_x : (y_n)_{n\geq 0} \in \ell^1_{\mathbb{k}} \mapsto \sum_{n=0}^{\infty} x_n y_n \in \mathbb{k}.$$

Esta función es claramente continua y, de acuerdo a (12.5), es $\|\phi_x\|_{(\ell^1_{\mathbb{k}})'} \leq \|x\|_{\ell^\infty_{\mathbb{k}}}$. Vale, de hecho, la igualdad. Esto es evidente si x es nulo, así que podemos suponer que no lo es. En ese caso, si c es un número positivo menor que $\|x\|_{\ell^\infty_{\mathbb{k}}}$, de manera que hay un entero positivo k tal que $|x_k| > c$, y e_k es el elemento de $\ell^1_{\mathbb{k}}$ cuya única componente no nula es la k-ésima, que es igual a 1, entonces $\|e_k\|_{\ell^1_{\mathbb{k}}} = 1$ y $|\phi_x(e_k)| = |x_k| > c$.

Esto nos dice que la función

$$\Phi : x \in \ell^\infty_{\mathbb{k}} \mapsto \phi_x \in (\ell^1_{\mathbb{k}})',$$

que claramente es lineal, es una inyección isométrica. Para probar la proposición, entonces, es suficiente que mostremos que esta función Φ es sobreyectiva.

Sea ϕ un elemento de $(\ell^1_{\mathbb{k}})'$. Para cada n en $\mathbb{N}_0$ sea e_n es el elemento de $\ell^1_{\mathbb{k}}$ cuya única componente no nula es la n-ésima, que es igual a 1, y pongamos $x_n \coloneqq \phi(e_n)$. Para cada $n \in \mathbb{N}_0$ es

$$|x_n| = |\phi(e_n)| \leq \|\phi\|_{(\ell^1_{\mathbb{k}})'} \cdot \|e_n\|_{\ell^1_{\mathbb{k}}} = \|\phi\|_{(\ell^1_{\mathbb{k}})'}$$

y, por lo tanto, la sucesión $x = (x_n)_{n\geq 0}$ es un elemento de $\ell^\infty_{\mathbb{k}}$. Sea, por otro lado, $y = (y_n)_{n\geq 0}$ un elemento cualquiera de $\ell^1_{\mathbb{k}}$. Para cada $m \in \mathbb{N}_0$ es

$$\left\| y - \sum_{n=0}^{m} y_n e_n \right\| = \sum_{n=m+1}^{\infty} |y_n|$$

y, como la serie $\sum_{n=0}^{\infty}|y_n|$ converge, tenemos que

$$y = \lim_{m\to\infty}\sum_{n=0}^{m} y_n e_n.$$

Como la función ϕ es lineal y continua, entonces, es

$$\begin{aligned}\phi(y) &= \phi\left(\lim_{m\to\infty}\sum_{n=0}^{m} y_n e_n\right) = \lim_{m\to\infty}\phi\left(\sum_{n=0}^{m} y_n e_n\right) = \lim_{m\to\infty}\sum_{n=0}^{m} y_n\phi(e_n) = \lim_{m\to\infty}\sum_{n=0}^{m} x_n y_n\\ &= \sum_{n=0}^{\infty} x_n y_n = \phi_x(y).\end{aligned}$$

Vemos así que $\phi = \phi_x$, y esto prueba que la función Φ de arriba es sobreyectiva, que es lo que queríamos. □

Como consecuencia de esta proposición obtenemos el ejemplo que buscamos:

Corolario 12.3.20. *El espacio $\ell^1_{\mathbb{k}}$ no es reflexivo.*

Demostración. Como $\ell^\infty_{\mathbb{k}}$ no es separable, la Proposición 12.3.18 nos dice que el espacio $(\ell^\infty_{\mathbb{k}})'$ no es separable. Ahora bien, como hay una isometría $\ell^\infty_{\mathbb{k}} \to (\ell^1_{\mathbb{k}})'$, hay una isometría $(\ell^1_{\mathbb{k}})'' \to (\ell^\infty_{\mathbb{k}})'$ y podemos concluir que el espacio $(\ell^1_{\mathbb{k}})''$ no es separable. Esto implica que este último no es isométrico a $\ell^1_{\mathbb{k}}$, ya que este sí es separable, y, por lo tanto, que $\ell^1_{\mathbb{k}}$ no es reflexivo. □

12.4. Completación de espacios de Banach

Cuando tenemos un espacio normado que *no* es de Banach, podemos completarlo. Expliquemos precisamente qué queremos decir con esto.

Definición 12.4.1. Una ***completación*** de un espacio normado X es una función $\iota : X \to Y$ con codominio en un espacio de Banach que es lineal e isométrica, y que tiene imagen densa.

Notemos que una completación de un espacio normado X en tanto espacio normado es, en particular, una completación de X en tanto espacio métrico.

Si X es un espacio normado, entonces sabemos que X posee completaciones en tanto espacio métrico, y la siguiente proposición nos dice que cualquiera de ellas puede enriquecerse hasta ser una completación de X en tanto espacio normado. En particular, por supuesto, esto implica que todo espacio normado posee completaciones.

Proposición 12.4.2. *Sea X un espacio normado, y sea $\iota : X \to Y$ una completación de X en tanto espacio métrico. Hay exactamente una forma de hacer de Y un espacio normado de manera que la función ι sea lineal y cuando lo hacemos la métrica sobre Y asociada a su norma coincide con la métrica original de Y.*

Notemos que la afirmación final de la proposición implica que Y es, con respecto a esa estructura de espacio normado, un espacio de Banach. Daremos una segunda demostración de esta proposición, completamente distinta, en la página 460.

Demostración. Dividamos el argumento, que es largo pero sencillo, en varios pasos.

Primer paso. Sean x e y dos elementos de Y. Como la imagen de ι es densa en Y, hay sucesiones $(x_n)_{n\geq 0}$ e $(y_n)_{n\geq 0}$ en X tales que las sucesiones $(\iota(x_n))_{n\geq 0}$ e $(\iota(y_n))_{n\geq 0}$ convergen a x y a y. Estas dos últimas sucesiones en Y convergen, así que satisfacen la condición de Cauchy: como la función ι es una inyección isométrica, esto implica que las sucesiones $(x_n)_{n\geq 0}$ e $(y_n)_{n\geq 0}$ también satisfacen la condición de Cauchy. Si ε es un número positivo, hay entonces un entero positivo n_0 tal que $\|x_m - x_n\| < \varepsilon/2$ y $\|y_m - y_n\| < \varepsilon/2$ siempre que m y n son enteros mayores que n_0, y se sigue de esto que

$$\|(x_m + y_m) - (x_n + y_n)\| \leq \|x_m - x_n\| + \|y_m - y_n\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

siempre que m y n son dos enteros mayores que n_0. Vemos con esto que la sucesión $(x_n + y_n)_{n\geq 0}$ satisface la condición de Cauchy en Y, así que, como ι es una inyección isométrica, también la satisface la sucesión $(\iota(x_n + y_n))_{n\geq 0}$ en Y. Ahora bien, como el espacio Y es completo, podemos concluir que esta última sucesión, de hecho, converge en Y: escribamos s a su límite.

Sean, por otro lado, $(x'_n)_{n\geq 0}$ e $(y'_n)_{n\geq 0}$ otras dos sucesiones en X tales que las sucesiones $(\iota(x'_n))_{n\geq 0}$ e $(\iota(y'_n))_{n\geq 0}$ convergen a x y a y en Y. Lo que acabamos de hacer implica que la sucesión $(\iota(x_n+y_n))_{n\geq 0}$ converge en Y a un punto s'. Como $(\iota(x_n))_{n\geq 0}$ y $(\iota(x'_n))_{n\geq 0}$ convergen las dos a x, la sucesión $(d_Y(\iota(x_n), \iota(x'_n)))_{n\geq 0}$ converge a 0 en $\mathbb{R}$: como ι es una inyección isométrica, también tenemos que la sucesión $(\|x_n - x'_n\|)_{n\geq 0} = (d(x_n, x'_n))_{n\geq 0}$ converge a 0. De la misma forma, la sucesión $(\|y_n - y'_n\|)_{n\geq 0}$ converge a 0. Como para todo $n \in \mathbb{N}_0$ es

$$\begin{aligned} d(s, s') &\leq d(s, \iota(x_n + y_n)) + d(\iota(x_n + y_n), \iota(x'_n + y'_n)) + d(\iota(x'_n + y'_n), x') \\ &\leq d(s, \iota(x_n + y_n)) + d(x_n + y_n, x'_n + y'_n) + d(\iota(x'_n + y'_n), x') \\ &\leq d(s, \iota(x_n + y_n)) + \|(x_n + y_n) - (x'_n + y'_n)\| + d(\iota(x'_n + y'_n), x') \\ &\leq d(s, \iota(x_n + y_n)) + \|x_n - x'_n\| + \|y_n - y'_n\| + d(\iota(x'_n + y'_n), x'), \end{aligned}$$

podemos concluir que, de hecho, $s = s'$.

Todo esto prueba lo siguiente:

> *hay exactamente una función $\sigma : Y \times Y \to Y$ tal que siempre que x e y son dos elementos de Y y $(x_n)_{n\geq 0}$ e $(y_n)_{n\geq 0}$ son sucesiones en Y tales que las sucesiones $(\imath(x_n))_{n\geq 0}$ e $(\imath(y_n))_{n\geq 0}$ convergen en Y a x y a y, respectivamente, la sucesiones $(\imath(x_n + y_n))_{n\geq 0}$ converge a $\sigma(x, y)$.*

Mostremos ahora que si vemos a σ como una operación binaria sobre el conjunto Y, entonces Y es un grupo abeliano.

- Sean x, y, y z tres elementos de Y. Hay sucesiones $(x_n)_{n\geq 0}$, $(y_n)_{n\geq 0}$, y $(z_n)_{n\geq 0}$ en X tales que las sucesiones $(\imath(x_n))_{n\geq 0}$, $(\imath(y_n))_{n\geq 0}$, y $(\imath(z_n))_{n\geq 0}$ convergen en Y a x, a y, y a z, respectivamente, y entonces las sucesiones $(\imath(x_n + y_n))_{n\geq 0}$, $(\imath(y_n + x_n))_{n\geq 0}$, y $(\imath(y_n + z_n))_{n\geq 0}$ convergen en Y a $\sigma(x, y)$, a $\sigma(y, x)$ y a $\sigma(y, z)$. Como las sucesiones $(\imath(x_n + y_n))_{n\geq 0}$ y $(\imath(y_n + x_n))_{n\geq 0}$ coinciden, esto nos dice que $\sigma(x, y) = \sigma(y, x)$. Por otro lado, esto implica que las sucesiones $(\imath((x_n + y_n) + z_n))_{n\geq 0}$ y $(\imath(x_n + (y_n + z_n)))_{n\geq 0}$ convergen a $\sigma(\sigma(x, y), z)$ y a $\sigma(x, \sigma(y, z))$: como esas dos sucesiones coinciden, vemos que $\sigma(\sigma(x, y), z) = \sigma(x, \sigma(y, z))$. Esto prueba que la operación σ en Y es asociativa y conmutativa.
- Como la función $\imath$ es continua y la sucesión constante $(0)_{n\geq 0}$ converge a 0 en X, la sucesión $(\imath(0))_{n\geq 0}$ en Y converge a $\imath(0)$ en Y. Si x es un elemento cualquiera de X y $(x_n)_{n\geq 0}$ es una sucesión en X tal que la sucesión $(\imath(x_n))_{n\geq 0}$ converge a x en Y, entonces las sucesión $(\imath(0 + x_n))_{n\geq 0}$ e $(\imath(x_n + 0))_{n\geq 0}$ convergen a $\sigma(\imath(0), x)$ y a $\sigma(x, \imath(0))$ en Y: como esas dos sucesión son de hecho iguales a $(\imath(x_n))_{n\geq 0}$, esto nos dice que $\sigma(\imath(0), x) = x = \sigma(x, \imath(0))$ y, por lo tanto, que el elemento $\imath(0)$ de Y es un elemento neutro para la operación σ.
- Sea x un elemento de Y y sea $(x_n)_{n\geq 0}$ una sucesión en X tal que la sucesión $(\imath(x_n))_{n\geq 0}$ converge a x en Y. Esta última sucesión satisface entonces la condición de Cauchy y, como $\imath$ es una inyección isométrica, también lo hace la sucesión $(x_n)_{n\geq 0}$. Ahora bien, para toda elección de m y n en $\mathbb{N}_0$ es $\|(-x_m) - (-x_n)\| = \|x_m - x_n\|$, así que la sucesión $(-x_n)_{n\geq 0}$ satisface la condición de Cauchy y, como $\imath$ es una inyección isométrica, también lo hace $(\imath(-x_n))_{n\geq 0}$. Como el espacio Y es completo, esta sucesión converge, de hecho, a un elemento y de Y y la sucesión $(\imath(x_n + (-x_n)))_{n\geq 0}$ converge a $\sigma(x, y)$. Por supuesto $x_n + (-x_n) = 0$ para todo $n \in \mathbb{N}_0$, así que esta sucesión no es otra que la sucesión constante $(\imath(0))_{n\geq 0}$: vemos con esto que $\sigma(x, y) = \imath(0)$. Como la operación σ es conmutativa e $\imath(0)$ es un elemento neutro para σ, esto nos dice que el elemento y de Y es opuesto a x con respecto a σ.

Probemos que, más aún, con respecto a esta operación σ sobre Y la función $\imath : X \to Y$ es un homorfismo de grupos.

- Sean x e y dos elementos de X. Como las sucesiones constantes $(x)_{n\geq 0}$ e $(y)_{n\geq 0}$ son tales que las sucesiones $(\imath(x))_{n\geq 0}$ e $(\imath(y))_{n\geq 0}$ convergen a $\imath(x)$ y a $\imath(y)$, respectivamente, entonces la

sucesión $(\iota(x+y))_{n\geq 0}$ converge a $\sigma(\iota(x),\iota(y))$. Por supuesto, esta última sucesión converge a $\iota(x+y)$, ya que es constante, y esto nos dice quie $\iota(x+y)=\sigma(\iota(x),\sigma(y)$, de manera que la función ι es un homorfismo de grupos.

Segundo paso. Sea ahora r un escalar de $\Bbbk$ y sea x un elemento de Y, de manera que hay una sucesión $(x_n)_{n\geq 0}$ en X tal que la sucesión $(\iota(x_n))_{n\geq 0}$ converge en Y a x. En particular, esta última sucesión satisface la condición de Cauchy, así que, como la función ι es una inyección isométrica, también lo hace la sucesión $(x_n)_{n\geq 0}$. Si ε es un número positivo, entonces hay un entero positivo n_0 tal que $\|x_m-x_n\|<\varepsilon/(1+|r|)$ siempre que m y n son enteros mayores que n_0, y esto implica que

$$\|rx_m-rx_n\|=|r|\cdot\|x_m-x_n\|<|r|\cdot\frac{\varepsilon}{1+|r|}<\varepsilon$$

siempre que m y n son enteros mayores que n_0. Vemos con esto que la sucesión $(rx_n)_{n\geq 0}$ satisface la condición de Cauchy, así que, como ι es una inyección isométrica, también lo hace $(\iota(rx_n))_{n\geq 0}$ y, como el espacio Y es completo, esta última sucesión converge a un punto p de Y.

Sea ahora $(x'_n)_{n\geq 0}$ otra sucesión en X tal que la sucesión $(\iota(x'_n))_{n\geq 0}$ converge a x: lo que acabamos de hacer nos dice que la sucesión $(\iota(rx'_n))_{n\geq 0}$ converge a un punto p' de Y. Como $(\iota(x_n))_{n\geq 0}$ e $(\iota(x'_n))_{n\geq 0}$ convergen las dos a x, la sucesión

$$(\|x_n-x'_n\|)_{n\geq 0}=(d(x_n,x'_n))_{n\geq 0}=(d(\iota(x_n),\iota(x'_n)))_{n\geq 0}$$

converge a 0. Usando esto podemos concluir que $p=p'$, ya que para todo $n\in\mathbb{N}_0$ es

$$\begin{aligned}d(p,p')&\leq d(p,\iota(rx_n))+d(\iota(rx_n),\iota(rx'_n))+d(\iota(rx'_n),p')\\&\leq d(p,\iota(rx_n))+d(rx_n,rx'_n)+d(\iota(rx'_n),p')\\&\leq d(p,\iota(rx_n))+\|rx_n-rx'_n\|+d(\iota(rx'_n),p')\\&\leq d(p,\iota(rx_n))+|r|\cdot\|x_n-x'_n\|+d(\iota(rx'_n),p').\end{aligned}$$

Hemos probado que

> *hay exactamente una función $\mu:\Bbbk\times Y\to Y$ tal que cada vez que r es un escalar de $\Bbbk$, x un elemento de Y, y $(x_n)_{n\geq 0}$ una sucesión en X tal que la sucesión $(\iota(x_n))_{n\geq 0}$ converge a x se tiene que $(\iota(rx_n))_{n\geq 0}$ converge a $\mu(r,x)$.*

Probemos ahora que dotando al grupo abeliano (Y,s) de la acción del cuerpo $\Bbbk$ dada por la función μ hacemos de aquel un $\Bbbk$-espacio vectorial.

- Sean r y s dos escalares de $\Bbbk$, sea x un elemento de Y, y sea $(x_n)_{n\geq 0}$ una sucesión en X tal que $(\iota(x_n))_{n\geq 0}$ converge a x. Tenemos entonces que las sucesiones $(\iota(rx_n))_{n\geq 0}$ e $(\iota(sx_n))_{n\geq 0}$ convergen a $\mu(r,x)$ y a $\mu(s,x)$, respectivamente. Por un lado, esto nos dice que la sucesión $(\iota(rx_n+sx_n))_{n\geq 0}$ converge a $\sigma(\mu(r,x),\mu(s,x))$: como esta sucesión claramente

coincide con $(\iota((r+s)x_n))_{n\geq 0}$ y está connverge a $\mu(r+s,x)$, vemos que

$$\sigma(\mu(r,x),\mu(s,x)) = \mu(r+s,x).$$

Por otro lado, como $(\iota(sx_n))_{n\geq 0}$ converge a $\mu(s,x)$, la sucesión $(\iota(r(s(x_n))))_{n\geq 0}$ converge a $\mu(r,\mu(s,x))$. Esta última sucesión es igual a $(\iota((rs)(x_n)))_{n\geq 0}$, así que también converge a $\mu(rs,x)$: vemos así que

$$\mu(r,\mu(s,x)) = \mu(rs,x).$$

- Sea x un elemento de Y y sea $(x_n)_{n\geq 0}$ una sucesión en X tal que $(\iota(x_n))_{n\geq 0}$ converge a X. La sucesión $(\iota(1x_n))_{n\geq 0}$ converge entonces a $\mu(1,x)$ y, como coincide con $(\iota(x_n))_{n\geq 0}$, vemos que

$$\mu(1,x) = x.$$

- Sean finalmente r un escalar de $\mathbb{k}$ y x e y dos elementos de Y, y sean $(x_n)_{n\geq 0}$ e $(y_n)_{n\geq 0}$ dos sucesiones en X tales que $(\iota(x_n))_{n\geq 0}$ e $(\iota(y_n))_{n\geq 0}$ convergen a x y a y, respectivamente. Las sucesiones $(\iota(rx_n))_{n\geq 0}$ e $(\iota(ry_n))_{n\geq 0}$ convergen entonces a $\mu(r,x)$ y a $\mu(r,y)$ y, por lo tanto, la sucesión $(\iota(rx_n+ry_n))_{n\geq 0}$ converge a $\sigma(\mu(r,x),\mu(r,y))$. Esta última sucesión coincide con $(\iota(r(x_n+y_n)))_{n\geq 0}$, que converge a $\mu(r,\sigma(x,y)))$, ya que $(\iota(x_n+y_n))_{n\geq 0}$ converge a $\sigma(x,y)$, y podemos concluir entonces que

$$\sigma(\mu(r,x),\mu(r,y)) = \mu(r,\sigma(x,y)).$$

Con todo esto vemos que las funciones $\sigma : Y\times Y\to Y$ y $\mu : \mathbb{k}\times Y\to Y$ hacen de Y un $\mathbb{k}$-espacio vectorial. Desde ahora, y como hacemos con todos los espacios vectoriales, siempre que x e y sean elementos de Y y r sea un elemento de $\mathbb{k}$ escribiremos $x+y$ en lugar de $\sigma(x,y)$ y rx en lugar de $\mu(r,x)$, y escribiremos 0 al cero de Y.

Terminemos este segundo paso de la demostración de la proposición mostrando que con respecto a esta estructora de espacio vectorial sobre Y la función $\iota : X\to Y$ es lineal.

- Ya sabemos que $\iota(x+y) = \iota(x)+\iota(y)$ cualesquiera sean los elementos x e y de X. Sean, por otro lado, r un escalar de $\mathbb{k}$ y x un elemento de X. La sucesión constante $(x)_{n\geq 0}$ es ta; que $(\iota(x))_{n\geq 0}$ converge a $\iota(x)$, así que la sucesión $(\iota(rx))_{n\geq 0}$ converge a $\iota(rx)$ y tambien a $r\iota(x)$: vemos así que $\iota(rx) = r\iota(x)$, como queremos.

Tercer paso. Sea d_Y la métrica de Y, y consideremos la función $v : x\in Y\mapsto d_Y(0,x)\in\mathbb{R}$, que es continua. Queremos probar que se trata de una norma sobre el espacio vectorial Y.

- Es claro que $v(0) = d_Y(0,0) = 0$. Por otro lado, si x es un elemento de Y tal que $v(x) = 0$, entonces $d_Y(0,x) = 0$ y, por lo tanto, $x = 0$.

- Sea r un escalar de $\mathbb{k}$, sea x un elemento de Y, y sea $(x_n)_{n\geq 0}$ una sucesión en X tal que $(\imath(x_n))_{n\geq 0}$ converge a x. La sucesión $(\imath(rx_n))_{n\geq 0}$ converge entonces a rx en Y, así que la sucesión $(v(\imath(rx_n)))_{n\geq 0}$ converge a $v(rx)$ en $\mathbb{R}$. Ahora bien, para cada $n \in \mathbb{N}_0$ es

$$\begin{aligned} v(\imath(rx_n)) = d_Y(0, \imath(rx_n)) = d_Y(\imath(0), \imath(rx_n)) = d(0, rx_n) = \|rx_n\| = |r| \cdot \|x_n\| \\ = |r| \cdot d(0, x_n) = |r| \cdot d_Y(\imath(0), \imath(x_n)) = |r| \cdot v(\imath(x_n)) \end{aligned}$$

  y, como la sucesión $(\imath(x_n))_{n\geq 0}$ converge a x y v es continua, la sucesión $(v(\imath(x_n)))_{n\geq 0}$ converge a $v(x)$ y la sucesión $v(\imath(rx_n))_{n\geq 0} = (|r| \cdot v(\imath(x_n)))_{n\geq 0}$ converge a $|r| \cdot v(x)$. Vemos así que

$$v(rx) = |r| \cdot v(x).$$

- Sean finalmente x e y dos elementos de Y y sean $(x_n)_{n\geq 0}$ e $(y_n)_{n\geq 0}$ dos sucesiones en X tales que $(\imath(x_n))_{n\geq 0}$ e $(\imath(y_n))_{n\geq 0}$ convergen en Y a x y a y. La sucesión

$$(\|x_n\|)_{n\geq 0} = (d(0, x_n))_{n\geq 0} = (d(\imath(0), \imath(x_n)))_{n\geq 0} = (v(x_n))_{n\geq 0}$$

  converge a $v(x)$ y, de manera similar, la sucesión $(\|y_n\|)_{n\geq 0}$ converge a $v(y)$. Como además la sucesión $(\imath(x_n + y_n))_{n\geq 0}$ converge a $x + y$, la sucesión $(\|x_n + y_n\|)_{n\geq 0}$ converge a $v(x + y)$. Como $\|x_n + y_n\| \leq \|x_n\| + \|y_n\|$ para todo $n \in \mathbb{N}_0$, la sucesión $(\|x_n + y_n\|)_{n\geq 0}$ está acotada componente a componente por la sucesión $(\|x_n\| + \|y_n\|)_{n\geq 0}$, que claramente converge a $v(x) + v(y)$, así que podemos concluir que

$$v(x + y) \leq v(x) + v(y).$$

Si x e y son dos elementos de Y y $(x_n)_{n\geq 0}$ e $(y_n)_{n\geq 0}$ son sucesiones en X tales que $(\imath(x_n))_{n\geq 0}$ e $(\imath(y)_n))_{n\geq 0}$ convergen a x y a y, entonces sabemos que

$$(\|x_n - y_n\|)_{n\geq 0} = (d(x_n, y_n))_{n\geq 0} = (d_Y(\imath(x_n), \imath(y_n)))_{n\geq 0}$$

converge a $d_Y(x, y)$, y

$$(\|x_n - y_n\|)_{n\geq 0} = (d(0, x_n - y_n))_{n\geq 0} = (d_Y(\imath(0), \imath(x_n - y_n)))_{n\geq 0} = (v(\imath(x_n - y_n)))_{n\geq 0}$$

converge a $v(x - y)$, ya que $(\imath(x_n - y_n))_{n\geq 0}$ converge a $x - y$ y v es continua: esto nos dice que

$$d_Y(x, y) = v(x - y),$$

y, por lo tanto, que la métrica sobre Y determinada por la norma v coincide con la métrica d_Y original de ese espacio. En particular, como la métrica d_Y es completa, el espacio normado Y es completo, esto es, se trata de un espacio de Banach.

Cuarto paso. Hemos probado con todo lo que precede que es posible hacer del conjunto Y un espacio normado de manera tal que la función ι sea lineal y la métrica sobre Y correspondiente a su norma coincida con la métrica d_Y original de Y. Para completar la demostración de la proposición tenemos que probar que hay exactamente una forma de hacer eso.

Supongamos que $\sigma' : Y \times Y \to Y$, $\mu' : \Bbbk \times Y \to Y$ y $\nu' : Y \to \mathbb{R}$ son funciones tales que Y dotado de las operaciones de suma, producto por escalares, y norma dadas por σ', μ', y ν' es un espacio normado, que con respecto a esa estructura de espacio normado la función $\iota : X \to Y$ es lineal, y que la métrica sobre Y asociada a la norma ν' coincide con la métrica d_Y. Como la suma, el producto por escalares, y la norma de un espacio normado son funciones continuas con respecto a la métrica asociada a su norma, esto nos dice que las funciones μ, μ', σ, σ', ν, y ν' son todas continuas con respecto a las métricas en sus dominios y codominios determinadas por d_Y.

Si x e y son dos elementos de X, entonces

$$\sigma'(\iota(x), \iota(y)) = \iota(x + y) = \sigma(\iota(x), \iota(y)).$$

Esto nos dice que las funciones σ, $\sigma' : Y \times Y \to Y$, que son continuas con respecto a la métrica d_Y, coinciden sobre el subconjunto $\iota(X) \times \iota(X)$ de $Y \times Y$, que es denso: como sabemos, esto implica que $\sigma = \sigma'$. De manera similar, si r es un escalar de $\Bbbk$ y x es un elemento de X, entonces

$$\mu'(r, \iota(x)) = \iota(rx) = \mu(r, \iota(x),$$

y, por lo tanto, las funciones μ, $\mu' : \Bbbk \times Y \to Y$, que son continuas con respecto a la métrica d_Y, coinciden sobre el subconjunto $\Bbbk \times \iota(X)$ de $\Bbbk \times Y$, que es denso, así que también tenemos que $\mu = \mu'$. Finalmente, como las métricas asociadas a las normas ν y ν' coinciden con d_Y, para cada cada elemento x de X tenemos que

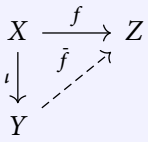
$$\nu(\iota(x)) = d_Y(\iota(0), \iota(x)) = \nu'(\iota(x)),$$

de manera que las funciones ν, $\nu' : Y \to \mathbb{R}$ coinciden sobre el subconjunto denso $\iota(X)$ de su dominio: como son continuas, vemos que $\nu = \nu'$. Como dijimos, esto termina la demostración de la proposición. □

Hemos mostrado que todo espacio normado admite completaciones y la siguiente proposición nos dice que podemos extender funciones lineales continuas a esa completación.

Proposición 12.4.3. *Sea X un espacio normado y sea $\iota : X \to Y$ una completación de X. Si Z es un espacio de Banach y $f : X \to Z$ es una función lineal continua, entonces hay exactamente una*

función lineal continua $\bar{f} : Y \to Z$ *tal que* $\bar{f} \circ \iota = f$.



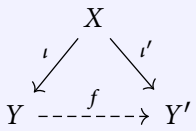
Demostración. Sea Z un espacio de Banach y sea $f : X \to Z$ una función lineal continua. De acuerdo a la Proposición 12.2.2, la función f es uniformemente continua y entonces la Proposición 8.5.8 nos dice que hay exactamente una función continua $\bar{f} : Y \to Z$ tal que $\bar{f} \circ \iota = f$. Para probar la proposición, entonces, es suficiente que mostremos que esta función $\bar{f}$ es lineal.

Sean x e y dos elementos de Y y sean a y b dos escalares de $\mathbb{k}$. Como la función ι tiene imagen densa, hay sucesiones $(x_n)_{n\geq 0}$ e $(y_n)_{n\geq 0}$ en X tales que las sucesiones $(\iota(x_n))_{n\geq 0}$ e $(\iota(y_n))_{n\geq 0}$ convergen a x y a y en Y. Como la función $\bar{f}$ es continua, sucesiones $(f(x_n))_{n\geq 0} = (\bar{f}(\iota(x_n)))_{n\geq 0}$ y $(f(y_n))_{n\geq 0} = (\bar{f}(\iota(y_n)))_{n\geq 0}$ convergen a $\bar{f}(x)$ y a $\bar{f}(y)$ en Z, y como Z es un espacio normado, esto implica que la sucesión $(af(x_n) + bf(y_n))_{n\geq 0}$ converge a $a\bar{f}(x) + b\bar{f}(x)$.

Por otro lado, como la función ι es lineal, es $(\iota(ax_n + by_n))_{n\geq 0} = (a\iota(x_n) + b\iota(y_n))_{n\geq 0}$, y como Y es un espacio normado, esta última sucesión converge a $ax + by$: como $\bar{f}$ es continua, esto implica que la sucesión $(af(x_n) + bf(y_n))_{n\geq 0} = (f(ax_n + by_n))_{n\geq 0} = (\bar{f}(\iota(ax_n + by_n)))_{n\geq 0}$ converge a $\bar{f}(ax + by)$. Podemos concluir entonces que $a\bar{f}(x) + b\bar{f}(x) = \bar{f}(ax + by)$. □

La Proposición 12.4.2 nos dice que toda completación de un espacio normado es, de manera única en el sentido apropiado, un espacio de Banach. Ahora bien, un espacio normado posee muchas completaciones: nuestro siguiente resultado nos dice que los espacios normados que obtenemos a partir de cada una de ellas son todos canónicamente isomorfos entre sí.

Proposición 12.4.4. *Sea X un espacio normado, y sean $\iota : X \to Y$ e $\iota' : X \to Y'$ dos completaciones de X. Hay exactamente una isometría lineal $f : Y \to Y'$ tal que $\iota \circ f = \iota'$.*



Demostración. De acuerdo a la Proposición 8.5.9, hay exactamente una función continua $\bar{f} : Y \to Y'$ tal que $f \circ \iota = \iota'$ y es una isometría. Más aún, como la función ι' es lineal e Y' es un espacio de Banach, la Proposición 12.4.3 nos dice que la función f es lineal. □

La demostración que dimos arriba de la Proposición 12.4.2 es larga y laboriosa. Probaremos ahora de una manera distinta que todo espacio normado admite una completación: este argumento

es considerablemente más eficiente, pero ahora depende de la teoría de dualidad que desarrollamos en la sección anterior.

Proposición 12.4.5. *Todo espacio normado posee una completación.*

Demostración. Sea X un espacio normado. Como $\mathbb{k}$ es un espacio de Banach, la Proposición 12.2.9 implica que X'' es un espacio de Banach. Por otro lado, la Proposición 12.3.9 nos dice que hay una inyección isométrica $\Phi : X \to X''$. Sea $Z \coloneqq \overline{\Phi(X)}$ que, de acuerdo a la Proposición 12.1.1, es un subespacio de X''. Como Z es cerrado en X'' y X'' es completo, Z es un espacio de Banach y es claro que la correstricción $\iota \coloneqq \Phi^Z : X \to Z$ es una completación del espacio métrico X que es lineal. Es claro, entonces, que ι es una completación de X en tanto espacio normado. □

12.5. El teorema de la función abierta y sus consecuencias

Definición 12.5.1. Sean X e Y dos espacios métricos. Una función $f : X \to Y$ es ***abierta*** si para todo subconjunto abierto U de X el conjunto $f(U)$ es abierto en Y.

Ejercicio 12.5.2. Pruebe que una función biyectiva $f : X \to Y$ entre dos espacios métricos es un homeomorfismo si y solamente si es continua y abierta.

De la misma forma en que para ver que una función lineal entre espacios normados es continua es suficiente con probar que es acotada sobre la bola unidad de su dominio, para ver que es abierta podemos usar el siguiente criterio.

Lema 12.5.3. *Sean X e Y dos espacios normados, sea $f : X \to Y$ una función lineal y continua, y consideremos las afirmaciones*

(*a*) *Hay números positivos r y s tales que $B_r^Y(0) \subseteq f(B_s^X(0))$.*
(*b*) *La función $f : X \to Y$ es abierta.*
(*c*) *Hay números positivos r y s tales que $B_r^Y(0) \subseteq \overline{f(B_s^X(0))}$.*

Valen las implicaciones (a) ⇒ (b) y (b) ⇒ (c), y si además el espacio X es completo, entonces también vale la implicación (c) ⇒ (a).

El resultado de que cuando el espacio X es completo vale la implicación $(c) \Rightarrow (a)$ es conocido como el *Lema de Schauder*, por Juliusz Schauder.

Demostración. $(a) \Rightarrow (b)$ Supongamos que hay números positivo r y $s$tal que $B_r^Y(0) \subseteq f(B_s^X(0))$ y sean U un subconjunto abierto de X e y un elemento de $f(U)$. Hay entonces un elemento x de U tal que $y = f(x)$ y, como U es abierto, un número positivo t tal que $B_t(x) \subseteq U$. Para cada elemento z de $B_{rt/s}^Y(y)$ tenemos que $s \cdot (z-y)/t \in B_r^Y(0)$, así que hay un vector w de $B_s^X(0)$ tal que $f(w) = s \cdot (z-y)/t$ y, por lo tanto,

$$z = y + \frac{t}{s} \cdot f(w) \in y + \frac{t}{s} \cdot f(B_s^X(0)) = f(x) + f\left(\frac{t}{s} \cdot B_s^X(0)\right) = f(B_t^X(x)) \subseteq f(U).$$

Vemos con esto que $B_{rt/2}^Y(y) \subseteq f(U)$ y, en definitiva, que el conjunto $f(U)$ es abierto. Esto prueba que la función f es abierta.

$(b) \Rightarrow (c)$ Si la función f es abierta, entonces el conjunto $f(B_1^X(0))$ es un abierto de Y que contiene a 0 y hay, por lo tanto, un número positivo r tal que $B_r^Y(0) \subseteq f(B_1(0)) \subseteq \overline{f(B_1^X(0))}$.

$(c) \Rightarrow (a)$ Supongamos ahora que el espacio X es completo y que hay números positivo r y s tales que $B_r^Y(0) \subseteq \overline{f(B_s^X(0))}$. Como el conjunto $\overline{f(B_s^X(0))}$ es cerrado, esto implica que, de hecho,

$$\overline{B}_r^Y(0) \subseteq \overline{f(B_s^X(0))}. \tag{12.6}$$

Esto implica que

> *si y es un elemento de Y y ε es un número positivo, entonces hay un vector x de X tal que $\|x\|_X < s\|y\|_Y/r$ y $\|y - f(x)\|_Y < \varepsilon$.* (12.7)

En efecto, esto es claro si y es 0. Sea, por otro lado, y un elemento no nulo de Y y sea ε un número positivo. Como $y' \coloneqq ry/\|y\|$ está en $\overline{B}_r^Y(0)$, la inclusión (12.6) nos dice que hay un elemento x' de $B_s^X(0)$ tal que $\|y' - f(x')\|_Y < r\varepsilon/\|y\|$, y entonces el vector $x \coloneqq \|y\|x'/r$ de X es tal que $\|y - f(x)\|_Y < \varepsilon$ y $\|x\| = \|x'\|\|y\|/r < s\|y\|/r$.

Sea ahora y un elemento de $B_r^Y(0)$, sea c un número positivo tal que $\|y\|_Y + sc < r$, y consideremos el conjunto $\mathscr{S}$ de todas las funciones $\phi : [\![0, n]\!] \to X$ definidas en un conjunto de la forma $[\![0, n]\!]$ para algún $n \in \mathbb{N}_0$ y tales que $\|\phi(0)\|_X < s\|y\|_Y/r$, $\|\phi(i)\|_X < sc/r2^i$ para cada $i \in [\![n]\!]$ y $\|y - \sum_{i=0}^n f(\phi(i))\|_Y < sc/2^{n+1}$.

- De acuerdo a (12.7), hay un vector x de X tal que $\|x\|_X < s\|y\|_Y/r$ y $\|y - f(x)\|_Y < sc/2$. Esto implica que la función $\phi : [\![0, 0]\!] \to X$ tal que $\phi(0) = x$ pertenece al conjunto $\mathscr{S}$.
- Sea $\phi : [\![0, n]\!] \to X$ un elemento de X y pongamos $y' \coloneqq y - \sum_{i=0}^n f(\phi(i))$. Tenemos entonces que $\|y'\|_Y < sc/2^{n+1}$ y, de acuerdo a (12.7), hay un elemento x de X tal que

  $$\|x\|_X < \frac{\|y'\|_Y}{r} < \frac{sc}{r2^{n+1}}, \qquad \|y' - f(x)\|_Y < \frac{sc}{2^{n+2}}.$$

  La función $\psi : [\![0, n+1]\!] \to X$ que en cada elemento i de $[\![0, n+1]\!]$ toma el valor

  $$\psi(i) = \begin{cases} \phi(i) & \text{si } i \leq n; \\ x & \text{si } i = n+1 \end{cases}$$

pertenece por lo tanto a $\mathscr{S}$ y es tal que $\psi|_{[\![0,n]\!]} = \phi$.

De acuerdo a la forma del *Principio de elección dependiente* que da la Proposición 2.1.9 hay una función $\phi : \mathbb{N}_0 \to X$ tal que $\phi|_{[\![0,n]\!]} \in \mathscr{S}$ cualquiera sea $n \in \mathbb{N}_0$. Si para cada $n \in \mathbb{N}_0$ ponemos $x_n \coloneqq \phi(n)$, entonces $\|x_0\|_X \leq \|y\|_Y / r$ y para cada entero positivo n es

$$\|x_n\| \leq \frac{sc}{r2^n}, \qquad \left\| y - \sum_{i=0}^{n} f(x_i) \right\| \leq \frac{sc}{2^{n+1}}.$$

Claramente esto implica que

$$\lim_{n\to\infty} \sum_{i=0}^{n} f(x_i) = y.$$

Por otro, como para cada $n \in \mathbb{N}_0$ tenemos que

$$\left\| \sum_{i=0}^{n+1} x_i - \sum_{i=0}^{n} x_i \right\| = \|x_{n+1}\| < \frac{sc}{r2^{n+1}}$$

y la serie $\sum_{n=0}^{\infty} sc/r2^{n+1}$ converge, la sucesión $(\sum_{i=0}^{n} x_i)_{n\geq 0}$ satisface la condición de Cauchy y, como el espacio X es completo, converge a un elemento x de X. Es

$$\|x\|_X = \lim_{n\to\infty} \left\| \sum_{i=0}^{n} x_i \right\|_X \leq \lim_{n\to\infty} \left(\|x_0\|_X + \sum_{i=1}^{n} \|x_i\|_X \right) \leq \frac{\|y\|_Y}{r} + \lim_{n\to\infty} \sum_{i=1}^{n} \frac{sc}{r2^i} = \frac{\|y\|_Y}{r} + \frac{sc}{r} < 1$$

y, como la función f es lineal y continua, tenemos que

$$y = \lim_{n\to\infty} \sum_{i=0}^{n} f(x_n) = \lim_{n\to\infty} f\left(\sum_{i=0}^{n} x_n \right) = f(x) \in f(B_1^X(0)).$$

Hemos probado que $B_r^Y(0) \subseteq f(B_1^X(0))$. □

Usaremos abajo el resultado del siguiente ejercicio.

Ejercicio 12.5.4. Pruebe que si G es un grupo abeliano dotado de una métrica invariante por translaciones y S y T son dos subconjuntos de G, entonces $\overline{S} + \overline{T} \subseteq \overline{S + T}$.

Apoyándonos en este criterio podemos probar el siguiente resultado, el *Teorema de la función abierta* o *Teorema de Banach–Schauder*, uno de los resultados fundamentales del análisis funcional.

Teorema 12.5.5. *Una función lineal y continua entre espacios de Banach es abierta si y solamente si es sobreyectiva.*

Demostración. Sean X e Y dos espacios de Banach y sea $f : X \to Y$ una función lineal y continua. Supongamos primero que la función f es abierta, de manera que, en particular, hay un número positivo r tal que $B_r^Y(0) \subseteq f(B_1(0))$. Si y es un elemento no nulo cualquiera de Y, entonces $y/2r\|y\|$ es un elemento de $B_r^Y(0)$, así que hay un vector x de X tal que $f(x) = y/2r\|y\|$ y, por lo tanto, $f(2r\|y\|x) = y$. Vemos así que la condición que da el teorema es necesaria para que la función f sea abierta.

Supongamos ahora que la función f es sobreyectiva. Como $X = \bigcup_{n=1}^{\infty} B_n^X(0)$, es

$$Y = f(X) = f\left(\bigcup_{n=1}^{\infty} B_n^X(0)\right) = \bigcup_{n=1}^{\infty} f(B_n^X(0)).$$

Como el espacio Y es completo, el *Teorema de Baire* 11.1.6 nos dice que hay un entero positivo n tal que el conjunto $f(B_n^X(0))$ no es nunca denso en Y y, por lo tanto, que hay un punto y_0 en Y y un número positivo r tal que

$$B_r(y_0) \subseteq \overline{f(B_n^X(0))}.$$

Para cada elemento y de Y tal que $\|y\|_Y < r$ es

$$\begin{aligned} y = (y - y_0) + y_0 \in B_r(y_0) + B_r(y_0) &\subseteq \overline{f(B_n^X(0))} + \overline{f(B_n^X(0))} \subseteq \overline{f(B_n^X(0)) + f(B_n^X(0))} \\ &\subseteq \overline{f(B_n^X(0) + B_n^X(0))} \subseteq \overline{f(B_{2n}^X(0))}, \end{aligned}$$

y esto nos dice que $B_r(0) \subseteq \overline{f(B_{2n}^X(0))}$. El *Lema de Schauder* 12.5.3 nos permite concluir, entonces, que la función f es abierta. □

Veamos ejemplos que muestran que la hipótesis de que el dominio y el codominio de la función lineal de la Proposición **??** sean completos es necesaria.

Ejemplo 12.5.6. Sea $\mathbb{k}$ o $\mathbb{R}$ o $\mathbb{C}$, y consideremos el espacio normado $X := \ell^1_{\mathbb{k}}$, que es completo. Escribamos Y al espacio normado que como espacio vectorial coincide con $\ell^1_{\mathbb{k}}$ y cuya norma es la restricción a $\ell^1_{\mathbb{k}}$ de la norma de $\ell^\infty_{\mathbb{k}}$, y consideremos la función $f := \mathrm{id}_{\ell^1_{\mathbb{k}}} : X \to Y$. Para cada elemento x de X es $\|x\|_\infty \leq \|x\|_1$, y esto implica inmediatamente que la función f, que evidentemente es sobreyectiva, es continua. Veamos que esta función f no es abierta. Para ello es suficiente que mostremos que para todo número positivo r la bola $B_r^Y(0)$ de Y no está contenida en la bola unidad $B_1^X(0)$ de X. Esto es fácil: si r es un número positivo, entonces hay un entero positivo n tal que $rn > 2$ y la sucesión

$$x := (r/2, r/2, \ldots, r/2, 0, 0, \ldots)$$

cuyas únicas componentes no nulas son las primeras n, que son todas iguales a $r/2$, es tal que $\|x\|_\infty = r/2$ y $\|x\|_1 = nr > 1$, así que $x \in B_r^Y(0)$ y $x \notin B_1^X(0)$. Por supuesto, el espacio normado Y

no es completo.

Ejemplo 12.5.7. Sea $\Bbbk$ o $\mathbb{R}$ o $\mathbb{C}$, y sea Y un espacio de Banach tal que hay una función lineal $\phi : Y \to \Bbbk$ que no es continua. Es fácil ver que la función

$$\|-\|_X : y \in Y \mapsto \|y\|_Y + |\phi(y)| \in \mathbb{R}$$

es una norma sobre el espacio vectorial Y, así que podemos considerar el espacio normado $(Y, \|-\|_X)$, al que escribiremos X. Consideremos la función $f = \mathrm{id}_Y : X \to Y$, que claramente es sobreyectiva. Como para cada vector x de X es

$$\|f(x)\|_Y \leq \|x\|_X,$$

la función f es continua. Si fuera abierta, habría, en particular, hay un número positivo r tal que $B_r^Y(0) \subseteq B_1^X(0)$, y entonces para todo elemento y de de Y tal que $\|y\|_Y < 1$ sería

$$|\phi(y)| = \frac{1}{r} \cdot |\phi(ry)| \leq \frac{1}{r} \cdot \big(\|ry\|_X + |\phi(ry)|\big) = \frac{1}{r} \cdot \|ry\|_X \leq \frac{1}{r}.$$

Esto es absurdo, porque la función lineal $\phi : Y \to \Bbbk$ no es continua. Es claro, en vista de la Proposición **??**, que el espacio normado X no es completo.

12.5.1. El teorema de la función inversa acotada

El *Teorema de Banach–Schauder* 12.5.5 tiene varios corolarios importantes. El primero es el llamado *Teorema de la función inversa acotada*.

Teorema 12.5.1.1. *La función inversa de una función lineal continua y biyectiva entre espacios de Banach es ella misma continua.*

Demostración. Sean X e Y dos espacios de Banach, y sea $f : X \to Y$ una función lineal continua y biyectiva. De acuerdo al *Teorema de Banach–Schauder* 12.5.5, la función f es abierta, así que su función inversa $f^{-1} : Y \to X$ es continua. □

Las funciones de los Ejemplos 12.5.6 y 12.5.7 son biyecciones continuas que no son abiertas, así que sus inversas no son continuas. Esto muestra que la hipótesis de que los espacios normados X e Y sean completos que aparece en el Teorema 12.5.1.1 es necesaria para llegar a su conclusión.

Corolario 12.5.1.2. *Sean X y Y dos espacios de Banach. Si $f : X \to Y$ es una función lineal continua cuya imagen* $\mathrm{img}(f)$ *es cerrada en Y, entonces hay un número positivo c tal que para cada elemento y de* $\mathrm{img}(f)$ *hay un elemento x de X tal que $f(x) = y$ y $\|x\|_X \leq c\|y\|_Y$.*

Demostración. Sea $f : X \to Y$ una función lineal y continua, sea $I \coloneqq \operatorname{img}(f)$ la imagen de f, y supongamos que I es un subespacio cerrado de Y.

□

12.5.2. El teorema del gráfico cerrado

Antes de presentar el segundo corolario del teorema necesitamos hacer una definición.

Definición 12.5.2.1. Sean X e Y dos conjuntos. El ***gráfico*** $f : X \to Y$ es el subconjunto

$$\Gamma(f) \coloneqq \{(x, f(x)) \in X \times Y : x \in X\}$$

del producto cartesiano $X \times Y$.

Notemos que, de hecho, el gráfico de una función *es* la función.

Si X e Y son dos espacios normados, D es un subespacio de X, y $f : D \to Y$ es una función lineal, entonces el inmediato verificar que el gráfico $\Gamma(f)$ de f es un subespacio vectorial del producto $D \times Y$ y, por lo tanto, de $X \times Y$. Si dotamos a $X \times Y$ de su norma producto usual, entonces tiene sentido decir que el gráfico de f es cerrado en $X \times Y$.

Lema 12.5.2.2. *Sean X e Y dos espacios de Banach, sea D un subespacio de X, y sea $f : D \to Y$ una función lineal. Las siguientes afirmaciones son equivalentes:*

(*a*) *El gráfico $\Gamma(f)$ de f es cerrado en $X \times Y$.*

(*b*) *Si $(x_n)_{n\geq 0}$ es una sucesión en D que converge a un punto x de X en X y tal que la sucesión $(f(x_n))_{n\geq 0}$ converge a un punto y de Y, entonces $x \in D$ y $f(x) = y$.*

(*c*) *La función $\|-\|_D : x \in D \mapsto \|x\|_X + \|f(x)\|_Y \in \mathbb{R}$ es una norma sobre D que hace de D un espacio de Banach.*

Demostración. (*a*) $\Rightarrow$ (*b*) Supongamos que el gráfico $\Gamma(f)$ es cerrado en $X \times Y$, y sea $(x_n)_{n\geq 0}$ una sucesión en D que converge en X a un punto x y tal que la sucesión $(f(x_n))_{n\geq)}$ converge en Y a un punto y. Se sigue de esto que la sucesión $((x_n, f(x_n)))_{n\geq 0}$, que toma valores en $\Gamma(f)$, converge en $X \times Y$ al par (x, y). Como $\Gamma(f)$ es un subconjunto cerrado de $X \times Y$, esto nos dice que (x, y) pertenece a $\Gamma(f)$, de manera que $x \in D$ e $y = f(x)$.

(*b*) $\Rightarrow$ (*c*) Supongamos que vale la afirmación (*b*) del lema. Es inmediato verificar que la función $\|-\|_D$ es una norma sobre D, ya que $\|-\|_X$ y $\|-\|_Y$ lo son y la función f es lineal. Probemos que D dotado de la métrica asociada a $\|-\|_D$, es completo. Sea $(x_n)_{n\geq 0}$ una sucesión en D que satisface la condición de Cauchy con respecto a la norma $\|-\|_D$. Es claro que entonces la sucesión $(x_n)_{n\geq 0}$ en X satisface la condición de Cauchy con respecto a la norma $\|-\|_X$ y que la sucesión $(f(x_n))_{n\geq 0}$ en Y satisface la condición de Cauchy con respecto a la norma $\|-\|_Y$.

Como X e Y son espacios de Banach, esto implica que la sucesión $(x_n)_{n\geq 0}$ converge en X a un punto x y que la sucesión $(f(x_n))_{n\geq 0}$ converge en Y a un punto Y. La hipótesis (b) nos permite entonces concluir que $x \in D$ y que $y = f(x)$. Esto nos dice que las sucesiones $(\|x_n - x\|_X)_{n\geq 0}$ y $(\|f(x_n - x)\|)_{n\geq 0}$ convergen a 0, así que también lo hace la sucesión $(\|x_n - x\|_D)_{n\geq 0}$. Vemos así que la sucesión $(x_n)_{n\geq 0}$ converge a x con respecto a la norma $\|-\|_D$.

$(c) \Rightarrow (a)$ Supongamos finalmente que la afirmaciones (c) del lema vale, y sea $((x_n, y_n))_{n\geq 0}$ una sucesión en $\Gamma(f)$ que converge en $X \times Y$ a un punto (x, y). Para cada $n \in \mathbb{N}_0$ el punto (x_n, y_n) pertenece a $\Gamma(f)$, así que $x_n \in D$ e $y_n = f(x_n)$. Por otro lado, como en $X \times Y$ estamos considerando la norma producto, sabemos que las sucesiones $(x_n)_{n\geq 0}$ e $(y_n)_{n\geq 0}$ convergen en X y en Y a x y a y, respectivamente. Eso implica que las sucesiones $(x_n)_{n\geq 0}$ y $(f(x_n))_{n\geq 0}$ satisfacen la condición de Cauchy con respecto a las normas $\|-\|_X$ y $\|-\|_Y$, así que claramente la sucesión $(x_n)_{n\geq 0}$, que toma valores en D, satisface la condición de Cauchy con respecto a la norma $\|-\|_D$. La hipótesis nos permite concluir que esa sucesión converge a un punto x' de D con respecto a esa norma: en vista de la definición de $\|-\|_D$, es claro que las sucesiones $(x_n)_{n\geq 0}$ y $(f(x_n))_{n\geq 0}$ convergen a x y a $f(x)$ con respecto a la normas $\|-\|_X$ y $\|-\|_Y$ y, por lo tanto, que $x = x' \in D$ y que $y = f(x)$. Vemos así que el punto (x, y) pertenece al gráfico $\Gamma(f)$ y, en definitiva, que este gráfico es cerrado en $X \times Y$. □

El ejemplo canónico de una función lineal con gráfico cerrado es el siguiente.

Ejemplo 12.5.2.3. Sea $C_b(0,1)$ el espacio vectorial real de las funciones continuas y acotadas $(0,1) \to \mathbb{R}$ dotado de la norma del supremo, con respecto a la que es un espacio de Banach, y sea D el subespacio de $C_b(0,1)$ de las funciones que son derivables en $(0,1)$ y cuyas derivadas pertenecen a $C_b(0,1)$. Hay entonces una función

$$f : \phi \in D \mapsto \phi' \in C_b(0,1).$$

Esta función tiene gráfico cerrado en $C_b(0,1) \times C_b(0,1)$. En efecto, supongamos que $(\phi_n)_{n\geq 0}$ es una sucesión en D que converge en $C_b(0,1)$ a una función ϕ y que la sucesión $(\phi_n')_{n\geq 0}$ converge en $C_b(0,1)$ a una función ψ. Para cada $t \in (0,1)$ tenemos que la sucesión $(\phi_n(t))_{n\geq 0}$ converge a $\phi(t)$. Por otro lado, si $t \in [\frac{1}{2}, 1)$, entonces

$$\left| \int_{1/2}^{t} \phi_n'(\tau)\,\mathrm{d}\tau - \int_{1/2}^{t} \psi(\tau)\,\mathrm{d}\tau \right| \leq \int_{1/2}^{t} |\phi_n'(\tau) - \psi(\tau)|\,\mathrm{d}\tau \leq \int_{1/2}^{t} \|\phi_n - \psi\|_\infty\,\mathrm{d}\tau = |t - \tfrac{1}{2}| \cdot \|\phi_n - \psi\|_\infty,$$

así que

$$\lim_{n\to\infty} \int_{1/2}^{t} \phi_n'(\tau)\,\mathrm{d}\tau = \int_{1/2}^{t} \psi(\tau)\,\mathrm{d}\tau,$$

y un cálculo similar muestra que esta igualdad también vale si $t \in (0, \frac{1}{2}]$. Se sigue de esto que para todo $t \in \mathbb{N}_0$ es

$$\phi(t) - \phi(\tfrac{1}{2}) = \lim_{n\to\infty} \big(\phi_n(t) - \phi_n(\tfrac{1}{2})\big) = \lim_{n\to\infty} \int_{1/2}^{t} \phi_n'(\tau)\,\mathrm{d}\tau = \int_{1/2}^{t} \psi(\tau)\,\mathrm{d}\tau$$

y, por lo tanto, que la función ϕ es derivable en $(0,1)$ y que $\phi'(t) = \psi(t)$ para todo $t \in (0,1)$. Esto implica que $\phi \in D$ y que $\psi = f(\psi)$. De acuerdo al Lema 12.5.2.2, el gráfico de f es cerrado en $C_b(0,1) \times C_b(0,1)$, como dijimos.

Veamos ahora que, si dotamos a D de la norma que se obtiene restringiendo a D la de $C_b(0,1)$, la función lineal f no es continua. En efecto, para cada $n \in \mathbb{N}$ la función

$$\phi_n : t \in (0,1) \mapsto \frac{\sin \pi n^2 t}{n} \in \mathbb{R}$$

claramente pertenece a D y la sucesión $(\phi_{n+1})_{n\geq 0}$ converge a 0 en D, pero la sucesión $(\phi'_{n+1})_{n\geq 0}$ no converge a 0 en $C_b(0,1)$. De hecho, como $\phi'_{n+1}(1/2) = (-1)^n 2(n+1)\pi$, la sucesión $(\|\phi'_{n+1}\|_\infty)_{n\geq 0}$ diverge a $+\infty$.

El siguiente es el *Teorema del gráfico cerrado*: nos dice que lo que vimos en este ejemplo no puede ocurrir cuando el dominio de las funciones que consideramos es un espacio de Banach.

Teorema 12.5.2.4. *Sean X e Y dos espacios de Banach. Una función lineal $f : X \to Y$ cuyo gráfico es cerrado en $X \times Y$ es continua.*

Demostración. Sea $f : X \to Y$ una función lineal cuyo gráfico $\Gamma(f)$ es cerrado en $X \times Y$, y consideremos a $\Gamma(f)$ como un espacio normado con respecto a la norma que se obtiene restringiendo la de $X \times Y$. Como $\Gamma(f)$ es cerrado en $X \times Y$ y este último es un espacio de Banach, vemos que $\Gamma(f)$ es, de esta forma, un espacio de Banach. Hay una función $g : \Gamma(f) \to X$ que en cada punto $p = (x, f(x))$ de $\Gamma(f)$ toma el valor $g(p) = x$. Esta función es claramente lineal y biyectiva, y además es continua. En efecto, para cada punto $p = (x, f(x))$ de $\Gamma(f)$ tenemos que $\|g(p)\|_X = \|x\|_X \leq \|x\|_X + \|f(x)\|_Y = \|p\|_{X\times Y}$. El *Teorema de la función inversa acotada* 12.5.1.1 nos dice que la función inversa $g^{-1} : X \to \Gamma(f)$ es continua.

Como la función $\pi : (x, y) \in X \times Y \mapsto y \in Y$ es continua, la restricción $\pi|_{\Gamma(f)} : \Gamma(f) \to Y$ también lo es y, en consecuencia, también es continua la composición $\pi|_{\Gamma(f)} \circ g^{-1} : X \to Y$. Como esta composición no es otra que la función f, esto prueba el Teorema. □

12.5.3. El teorema de acotación uniforme de Banach–Steinhaus

El tercer corolario del *Teorema de Banach–Schauder* que queremos presentar es el *Teorema de acotación uniforme de Banach–Steinhaus*, por Stefan Banach y Hugo Steinhaus. La demostración que damos viene del artículo [Ste1963] de Banach y Steinhaus, es la más conocida y depende directamente del *Teorema de Baire*. Las demostraciones originales de este teorema, debidas a Banach [Ban1922b] y a Hans Hahn [Hah1922], no dependen de ese teorema, como tampoco las que dieron más tarde Theophil Henry Hildebrandt en [Hil1923] y Felix Hausdorff en [Hau1932]. La demostración más elemental que el autor conoce es la que da Alan Sokal en [Sok2011].

Teorema 12.5.3.1. *Sea X un espacio de Banach, sea Y un espacio normado, y sea $\mathscr{F}$ una familia de funciones lineales continuas $X \to Y$. Si para todo elemento x de X el conjunto $\{f(x) : f \in \mathscr{F}\}$ es acotado en Y, entonces el conjunto $\mathscr{F}$ es acotado en el espacio normado $B(X, Y)$.*

En otras palabras, este teorema nos dice que un conjunto de funciones lineales continuas $X \to Y$ que es acotado puntualmente es uniformemente acotado.

Demostración. Supongamos que el conjunto $\{f(x) : f \in \mathscr{F}\}$ es acotado en Y para todo $x \in X$.

Sea n un entero positivo. Para cada elemento f de $\mathscr{F}$ la función $x \in X \mapsto \|f(x)\|_Y \in \mathbb{R}$ es continua, así que el conjunto $I_n(f) \coloneqq \{x \in X : \|f(x)\|_Y \leq n\}$ es cerrado en X, ya $(-\infty, n]$ lo es en $\mathbb{R}$, y esto implica que también es cerrado el conjunto

$$X_n = \bigcap_{f \in \mathscr{F}} I_n(f) = \{x \in X : \|f(x)\|_Y \leq n \text{ para todo } f \in \mathscr{F}\}.$$

Ahora bien, la hipótesis implica que $X = \bigcup_{n=1}^{\infty} X_n$, así que el *Teorema de Baire* 11.1.6 nos dice que hay un entero positivo m tal que el conjunto X_m tiene interior no vacío. Hay, por lo tanto, un vector x_0 de X y un número positivo r tal que $B_r(x_0) \subseteq X_m$. De hecho, como X_m es cerrado en X, tenemos también que $\overline{B}_r(x_0) \subseteq X_m$.

Si f es un elemento de $\mathscr{F}$, entonces para todo elemento x de $B_1^X(0)$ tenemos que

$$\|f(x)\|_Y = \frac{1}{r}\|f(rx - x_0) + f(x_0)\|_Y \leq \frac{1}{r}\big(\|f(rx - x_0)\|_Y + \|f(x_0)\|_Y\big) \leq \frac{1}{r}(m + m),$$

ya que los vectores $rx - x_0$ y x_0 están los dos en $\overline{B}_r(x_0)$, y, por lo tanto,

$$\|f\|_{B(X,Y)} = \sup_{x \in B_1^X(0)} \|f(x)\|_Y \leq \frac{2m}{r}.$$

Esto prueba que el conjunto $\mathscr{F}$ está contenido en la bola $\overline{B}_{2m/r}(0)$ de $B(X, Y)$, así que es acotado en este espacio. □

El siguiente corolario es una aplicación frecuente del teorema: el límite puntual de una sucesión de funciones lineales continuas entre espacios de Banach que converge puntualmente es continuo.

Corolario 12.5.3.2. *Sea X un espacio de Banach, sea Y un espacio normado, y sea $(f_n)_{n\geq 0}$ una sucesión de funciones lineales continuas $X \to Y$. Si para todo elemento x de X existe el límite $\lim_{n\to\infty} f_n(x)$, entonces la función*

$$f : x \in X \mapsto \lim_{n\to\infty} f_n(x) \in Y$$

es lineal y continua, y para todo subconjunto compacto de X la sucesión $(f_n|_K)_{n\geq 0}$ converge a $f|_K$ en el espacio $C(K, Y)$ de todas las funciones continuas $K \to Y$.

Demostración. HACER. □

Tercera parte

Capítulo 13

Cálculo diferencial e integral

13.1. Integración

13.1.1. Particiones

Definición 13.1.1.1. Sea n un elemento de $\mathbb{N}_0$. Una ***partición*** de un intervalo $[a, b]$ es una secuencia finita $(x_0, x_1, \ldots, x_n)$ tal que $x_0 = a$, $x_{i-1} < x_i$ para cada $i \in [\![n]\!]$, y $x_n = b$. Escribimos $\mathscr{P}[a, b]$ al conjunto de todas las particiones de $[a, b]$.

Si $P = (x_0, x_1, \ldots, x_n)$ es una partición de un intervalo $[a, b]$, llamamos a cada uno de los intervalos $[x_{i-1}, x_i]$ con $i \in [\![n]\!]$ un ***subintervalos*** determinado por P.

Hay una relación de orden sobre el conjunto $\mathscr{P}[a, b]$ que es fundamental para todo lo que hacemos. Para construirla empezamos haciendo la siguiente observación.

Lema 13.1.1.2. *Sea $[a, b]$ un intervalo de $\mathbb{R}$ y sean $P = (x_0, x_1, \ldots, x_n)$ y $Q = (y_0, y_1, \ldots, y_m)$ dos particiones de $[a, b]$. Hay como mucho una función $\pi : [\![0, m]\!] \to [\![0, n]\!]$ tal que $x_{\pi(i)} = y_i$ para cada $i \in [\![0, m]\!]$, y si la hay, entonces π es estrictamente creciente y es $\pi(0) = 0$ y $\pi(m) = n$.*

Demostración. Si $\pi, \pi' : [\![0, m]\!] \to [\![0, n]\!]$ son dos funciones tales que $x_{\pi(i)} = y_i$ y $x_{\pi'(i)} = y_i$ para cada $i \in [\![0, m]\!]$, entonces para cada $i \in [\![0, n]\!]$ es $x_{\pi(i)} = x_{\pi'(i)}$, así que $\pi(i) = \pi'(i)$: esto nos dice que $\pi = \pi'$ y prueba la primera afirmación del lema.

Supongamos ahora que $\pi : [\![0, m]\!] \to [\![0, n]\!]$ es una función tal que $x_{\pi(i)} = y_i$ para cada $i \in [\![0, m]\!]$. Si i e i' son elementos de $[\![0, m]\!]$ tales que $i < i'$, entonces $y_{\pi(i)} = x_i < x_{i'} = y_{\pi(i')}$, así

que, en particular, $\pi(i) < \pi(i')$: esto nos dice que la función π es estrictamente creciente. Como $x_{\pi(0)} = y_0 = a = x_0$ y $x_{\pi(m)} = y_m = b = x_n$, es claro que $\pi(0) = 0$ y $\pi(m) = n$. □

Es fácil ahora describir la relación que nos interesa y sus dos propiedades más importantes.

Lema 13.1.1.3. *Sea $[a, b]$ un intervalo cerrado de $\mathbb{R}$ y consideremos la relación binaria $\preceq$ sobre el conjunto $\mathscr{P}[a, b]$ tal que siempre que $P = (x_0, x_1, \ldots, x_n)$ y $Q = (y_0, y_1, \ldots, y_m)$ son elementos de $\mathscr{P}[a, b]$ se tiene que*

$$P \preceq Q \iff \text{hay una función } \pi : [\![0, m]\!] \to [\![0, n]\!] \text{ tal que } x_{\pi(i)} = y_i \text{ para cada } i \in [\![0, m]\!].$$

(i) La relación $\preceq$ es una relación de orden sobre $\mathscr{P}[a, b]$.

(ii) Si $P = (x_0, x_1, \ldots, x_n)$ y $Q = (y_0, y_1, \ldots, y_m)$ son dos elementos de $\mathscr{P}[a, b]$ y $P \preceq Q$, entonces hay exactamente una función $\phi : [\![n]\!] \to [\![m]\!]$ tal que $(x_{i-1}, x_i) \subseteq (y_{\phi(i)-1}, y_{\phi(i)})$ para cada $i \in [\![n]\!]$, esa función es no decreciente y, de hecho,

$$\phi(i) = \min\{j \in [\![0, m]\!] : x_i \leq y_j\}$$

cualquiera sea i en $[\![n]\!]$.

(iii) Si P y Q son dos elementos de $\mathscr{P}[a, b]$, entonces hay un elemento R de $\mathscr{P}[a, b]$ tal que $R \preceq P$ y $R \preceq Q$.

Cuando P y Q son dos elementos de $\mathscr{P}[a, b]$ tales que $P \preceq Q$, decimos que P ***refina*** a Q. La afirmación *(ii)* de este lema nos dice que si una partición P refina a una partición Q, entonces cada subintervalo determinado por P está contenido en uno de los subintervalos determinados por Q. Por otro lado, la afirmación *(iii)* nos dice que todo par de particiones de $[a, b]$ admite un refinamiento común — decimos por ello que el conjunto ordenado $\mathscr{P}[a, b]$ está ***dirigido***, como en el Ejercicio 7.1.28.

Demostración. *(i)* Veamos que $\preceq$ es una relación de orden sobre el conjunto $\mathscr{P}[a, b]$.

- Si $P = (x_0, x_1, \ldots, x_n)$ es un elemento de $\mathscr{P}[a, b]$, entonces $P \preceq P$ ya que la función $\pi \coloneqq \mathrm{id}_{[\![0,n]\!]} : [\![0, n]\!] \to [\![0, n]\!]$ es tal que $x_{\pi(i)} = x_i$ para cada $i \in [\![0, n]\!]$.
- Sean $P = (x_0, x_1, \ldots, x_n)$ y $Q = (y_0, y_1, \ldots, y_m)$ dos elementos de $\mathscr{P}[a, b]$ tales que $P \preceq Q$ y $Q \preceq P$, de manera que hay funciones $\pi : [\![0, m]\!] \to [\![0, n]\!]$ y $\sigma : [\![0, n]\!] \to [\![0, m]\!]$ tales que $x_{\pi(i)} = y_i$ para cada $i \in [\![0, m]\!]$ e $y_{\sigma(j)} = x_j$ para cada $j \in [\![0, n]\!]$. La parte *(i)* del lema nos dice que las funciones π y σ son estrictamente crecientes, así que son inyectivas. Podemos concluir que las funciones π y σ son, de hecho, biyectivas. Esto implica que $n = m$ y, como la función π es estrictamente creciente, que $\pi(i) = i$ para todo $i \in [\![0, n]\!]$, de manera que $P = Q$.

- Sean $P = (x_0, x_1, \ldots, x_n)$, $Q = (y_0, y_1, \ldots, y_m)$, y $R = (z_0, z_1, \ldots, z_p)$ tres elementos de $\mathscr{P}[a, b]$ tales que $P \preceq Q$ y $Q \preceq R$, de manera que hay funciones $\pi : [\![0, m]\!] \to [\![0, n]\!]$ y $\sigma : [\![0, p]\!] \to [\![0, m]\!]$ tales que $x_{\pi(i)} = y_i$ para cada $i \in [\![0, m]\!]$ e $y_{\sigma(j)} = z_j$ para cada $j \in [\![0, n]\!]$. Se sigue de esto que $P \preceq R$, ya que la función $\tau \coloneqq \pi \circ \sigma : [\![0, p]\!] \to [\![0, n]\!]$ es tal que $x_{\tau(k)} = x_{\pi(\sigma(k))} = y_{\sigma(k)} = z_k$ para cada $k \in [\![0, p]\!]$.

(*ii*) Sean $P = (x_0, x_1, \ldots, x_n)$ y $Q = (y_0, y_1, \ldots, y_m)$ dos elementos de $\mathscr{P}[a, b]$ y $P \preceq Q$, de manera que hay una función estrictamente creciente $\pi : [\![0, m]\!] \to [\![0, n]\!]$ tal que $x_{\pi(i)} = y_i$ para cada $i \in [\![0, m]\!]$. Para cada $i \in [\![n]\!]$ el conjunto $I_i \coloneqq \{j \in [\![0, m]\!] : x_i \leq y_j\}$ no es vacío, ya que contiene a m porque $x_i \leq b = y_m$, y no contiene a 0, ya que $y_0 = a < x_i$. Hay entonces una función

$$\phi : i \in [\![n]\!] \mapsto \min I_i \in [\![m]\!].$$

Si i e i' son dos elementos de $[\![n]\!]$ tales que $i < i'$, entonces $I_{i'} \subseteq I_i$ y $\phi(i) = \min I_i \leq \min I_{i'} = \phi(i')$: vemos así que la función ϕ es no decreciente. Además, si i es un elemento de $[\![n-1]\!]$, entonces $x_i > x_{i-1} \geq x_{\pi(i-1)} = y_{i-1}$, ya que π es estrictamente creciente, así que $i \in I_i$ y, por lo tanto, $\phi(i) \leq i$.

Sea i un elemento de $[\![n]\!]$. Como $\phi(i) \in I_i$, tenemos que $x_i \leq y_{\phi(i)}$. Por otro lado, como $\phi(i) - 1$ no está en I_i, porque es menor que $\phi(i) = \min I_i$, es $x_{\pi(\phi(i)-1)} = y_{\phi(i)-1} < x_i$ y, por lo tanto, $y_{\phi(i)-1} = x_{\pi(\phi(i)-1)} \leq x_{i-1}$: esto nos dice que $(x_{i-1}, x_i) \subseteq (y_{\phi(i)-1}, y_{\phi(i)})$.

Sea ahora $\phi' : [\![n]\!] \to [\![m]\!]$ otra función tal que $(x_{i-1}, x_i) \subseteq (y_{\phi'(i)-1}, y_{\phi'(i)})$ para cada $i \in [\![n]\!]$. Entonces para cada $i \in [\![n]\!]$ tenemos que $(x_{i-1}, x_i) \subseteq (y_{\phi(i)-1}, y_{\phi(i)}) \cap (y_{\phi'(i)-1}, y_{\phi'(i)})$, de manera que la intersección $(y_{\phi(i)-1}, y_{\phi(i)}) \cap (y_{\phi'(i)-1}, y_{\phi'(i)})$ no es vacía. Como la secuencia $(y_0, y_1, \ldots, y_m)$ es estrictamente creciente, esto implica que $\phi(i) = \phi'(i)$.

(*iii*) Sean $P = (x_0, x_1, \ldots, x_n)$ y $Q = (y_0, y_1, \ldots, y_m)$ dos elementos de $\mathscr{P}[a, b]$. El conjunto $X \coloneqq \{x_i : i \in [\![0, n]\!]\} \cup \{y_j \in [\![0, n]\!]\}$ es claramente finito y no vacío, así que hay un elemento k de $\mathbb{N}_0$ tal que $|X| = k + 1$. Para cada $x \in X$ el conjunto $X_x \coloneqq \{y \in X : y < x\}$ está contenido en $X \setminus \{x\}$, así que es finito y su cardinal es un elemento de $[\![0, k]\!]$. Hay, por lo tanto, una función

$$f : x \in X \mapsto |X_x| \in [\![0, k]\!].$$

Esta función es estrictamente creciente: en efecto, si x y x' son dos elementos de X tales que $x < x'$ entonces $X_x \subsetneq X_{x'}$ y, por lo tanto, $f(x) = |X_x| < |X_{x'}| = f(x')$. En particular, esta función f es inyectiva, y, como $|X| = |[\![0, k]\!]|$ y f es estrictamente creciente, biyectiva y su función inversa $f^{-1} : [\![0, k]\!] \to X$ es estrictamente creciente. Claramente $a = x_0 \in X$, $b = x_n \in X$ y, de hecho, a y B son el menor y el mayor elemento de X, respectivamente: como la función f^{-1} es estrictamente creciente, esto implica que $f^{-1}(0) = a$ y $f^{-1}(k) = b$, y, por lo tanto, la secuencia $R \coloneqq (z_0, z_1, \ldots, z_k)$ que tiene $z_r = f^{-1}(r)$ para cada $r \in [\![0, k]\!]$ es un elemento de $\mathscr{P}[a, b]$.

El conjunto

$$\pi \coloneqq \{(i, j) \in [\![0, n]\!] \times [\![0, k]\!] : j = f(x_i)\}$$

es una función $[\![0,n]\!] \to [\![0,k]\!]$. En efecto, si i es un elemento de $[\![0,n]\!]$, entonces $x_i \in X$ y $(i, f(x_i)) \in \pi$, y si $i \in [\![0,n]\!]$ y $j, j' \in [\![0,k]\!]$ son tales que $(i,j), (i,j') \in \pi$, entonces $j = f(x_i) = j'$. Como para cada elemento i de $[\![0,n]\!]$ tenemos que $(i, f(x_i)) \in \pi$, así que $z_{\pi(i)} = f^{-1}(f(x_i)) = x_i$, y vemos con esto que $R \preceq P$. De la misma forma podemos ver que el conjunto $\sigma \coloneqq \{(i,j) \in [\![0,m]\!] \times [\![0,k]\!] : j = f(y_i)\}$ es una función $[\![0,m]\!] \to [\![0,k]\!]$ tal que $z_{\sigma(j)} = y_j$ para cada $j \in [\![0,m]\!]$ y, en consecuencia, que $R \preceq Q$. □

Usaremos un par de veces la siguiente observación.

Lema 13.1.1.4. *Sea $[a,b]$ un intervalo de $\mathbb{R}$, sean $P = (x_0, x_1, \ldots, x_n)$ y $Q = (y_0, y_1, \ldots, y_m)$ dos particiones de $[a,b]$, y supongamos que $P \preceq Q$, de manera que hay*

- *una función estrictamente creciente $\pi : [\![0,m]\!] \to [\![0,n]\!]$ tal que $\pi(0) = 0$, $\pi(m) = n$, y $y_i = x_{\pi(i)}$ para cada $i \in [\![0,m]\!]$, y*
- *una función $\phi : [\![n]\!] \to [\![m]\!]$ tal que $(x_{i-1}, x_i) \subseteq (y_{\phi(i)-1}, y_{\phi(i)})$ para cada $i \in [\![n]\!]$.*

Si $i \in [\![n]\!]$ y $j \in [\![\pi(i-1)+1, \pi(i)]\!]$, entonces $\phi(j) = i$.

En el Lema 13.1.1.3 vimos que, de hecho, hay exactamente una función ϕ que satisface las condiciones del segundo punto del lema y la describimos completamente.

Demostración. Sea i un elemento de $[\![n]\!]$, y notemos que $1 \leq \pi(i-1)+1 \leq \pi(i) \leq m$. Si j es un elemento de $[\![\pi(i-1)+1, \pi(i)]\!] \subseteq [\![m]\!]$, entonces

$$x_{i-1} = y_{\pi(i-1)} < y_{\pi(i-1)+1} \leq y_j \leq y_{\pi(i)} = x_i,$$

así que la descripción de la función ϕ que dimos en el Lema 13.1.1.3 nos dice que $i - 1 < \phi(j) \leq i$, esto es, que $\phi(j) = i$ □

13.1.2. Funciones escalonadas

Usando las particiones de un intervalo podemos describir una clase útil de funciones.

Definición 13.1.2.1. Sea $[a,b]$ un intervalo cerrado de $\mathbb{R}$. Decimos que una función $f : [a,b] \to \mathbb{R}$ es ***escalonada con respecto a una partición*** $P = (t_0, t_1, \ldots, t_n)$ de $[a,b]$ si hay una función $\xi : [\![n]\!] \to \mathbb{R}$ tal que $f(x) = \xi(i)$ para cada elección de i en $[\![n]\!]$ y de x en (x_{i-1}, x_i), y decimos que es ***escalonada*** si es escalonada con respecto a alguna partición de $\mathscr{P}[a,b]$. Escribimos $\mathscr{E}[a,b]$ al conjunto de todas las funciones $[a,b] \to \mathbb{R}$ que son escalonadas.

Notemos que esta definición no impone ninguna condición sobre los valores de la función g en los puntos $t_0, t_1, \ldots, t_n$. Por otro lado, si sabemos que una función $f : [a,b] \to \mathbb{R}$ es escalonada con respecto a una partición de $[a,b]$, entonces hay exactamente una función ξ como en al definición.

Lema 13.1.2.2. *Sea $P = (x_0, x_1, \ldots, x_n)$ una partición de $[a, b]$, sea $f : [a, b] \to \mathbb{R}$ una función escalonada con respecto a P, y sea $\xi : [\![n]\!] \to \mathbb{R}$ una función tal que $f(x) = \xi(i)$ para cada elección de i en $[\![n]\!]$ y de x en (x_{i-1}, x_i). Para cada $i \in [\![n]\!]$ es*

$$\xi(i) = f\left(\frac{x_{i-1} + x_i}{2}\right).$$

Demostración. Para cada $i \in [\![n]\!]$ es $(x_{i-1} + x_i)/2 \in (x_{i-1}, x_i)$, así que $f((x_{i-1} + x_i)/2) = \xi(i)$. □

En general, una función que es escalonada lo es con respecto a muchas particiones. El siguiente lema nos da un ejemplo de esto.

Lema 13.1.2.3. *Sea $f : [a, b] \to \mathbb{R}$ una un función escalonada con respecto a una partición P de $[a, b]$. Si Q es una partición de $[a, b]$ que refina a P, entonces f también es escalonada con respecto a la partición Q.*

Demostración. Supongamos que $P = (x_0, x_1, \ldots, x_n)$ y que $Q = (y_0, y_1, \ldots, y_m)$ es una partición de $[a, b]$ que refina a P, de manera que hay funciones $\pi : [\![0, n]\!] \to [\![0, m]\!]$ y $\phi : [\![m]\!] \to [\![n]\!]$ tales que $y_{\pi(i)} = x_i$ para cada $i \in [\![0, n]\!]$ y $(y_{j-1}, y_j) \subseteq (x_{\phi(j)-1}, x_{\phi(j)})$ para cada $j \in [\![m]\!]$. Como f es escalonada con respecto a P, hay una función $\xi : [\![n]\!] \to \mathbb{R}$ tal que $f(x) = \xi(i)$ siempre que $i \in [\![n]\!]$ y $x \in (x_{i-1}, x_i)$. Consideremos la función $\zeta \coloneqq \xi \circ \phi : [\![m]\!] \to \mathbb{R}$ y observemos que si j es un elemento de $[\![m]\!]$ y x uno de (y_{j-1}, y_j), entonces $x \in (x_{\phi(j)-1}, x_{\phi(j)})$ y, por lo tanto, que $f(x) = \xi(\phi(j)) = \zeta(j)$. Esto nos dice que la función f es escalonada con respecto a la partición Q. □

Una consecuencia importante de esto es que el conjunto de las funciones escalonadas en un intervalo es un espacio vectorial.

Lema 13.1.2.4. *Sea $[a, b]$ un intervalo cerrado de $\mathbb{R}$.*

(*i*) *El conjunto $\mathscr{E}[a, b]$ es un subespacio vectorial del espacio $\mathbb{R}^{[a,b]}$ de todas las funciones $[a, b] \to \mathbb{R}$. Más aún, si f y g son dos elementos de $\mathscr{E}[a, b]$ que son funciones escalonadas con respecto a una partición P de $[a, b]$ y α y β son dos números reales, entonces $\alpha f + \beta g$ es una función escalonada con respecto a la misma partición P.*

(*ii*) *Si f y g son dos elementos de $\mathscr{E}[a, b]$, entonces las funciones* $\max\{f, g\}$, $\min\{f, g\}$, *y $|f|$ también son elementos de $\mathscr{E}[a, b]$.*

Demostración. Supongamos que $f, g : [a, b] \to \mathbb{R}$ son dos funciones escalonadas con respecto a particiones P y Q de $[a, b]$, respectivamente, sean α y β dos números reales, y consideremos la función $h \coloneqq \alpha f + \beta g : [a, b] \to \mathbb{R}$. De acuerdo a la tercera parte del Lema 13.1.1.3 hay una

partición $R = (x_0, x_1, \ldots, x_n)$ de $[a, b]$ que refina a P y a Q, y el Lema 13.1.2.3 nos dice que f y g son escalonadas con respecto a R, así que hay funciones $\xi, \zeta : [\![n]\!] \to \mathbb{R}$ tales que para todo $i \in [\![n]\!]$ y todo $x \in (x_{i-1}, x_i)$ es $f(x) = \xi(i)$ y $g(x) = \zeta(i)$. La función

$$\omega : i \in [\![n]\!] \to \alpha\xi(i) + \beta\zeta(i) \in \mathbb{R}$$

es entonces tal que para todo $i \in [\![n]\!]$ y todo $x \in (x_{i-1}, x_i)$ es

$$h(x) = \alpha f(x) + \beta g(x) = \alpha\xi(i) + \beta\zeta(i) = \omega(i),$$

y esto nos dice que la función h es escalonada con respecto a la partición R. En particular, la función h pertenece al conjunto $\mathscr{E}[a, b]$, y con esto y el hecho evidente de que la función nula es escalonada podems concluir que $\mathscr{E}[a, b]$ es un subespacio de $\mathbb{R}^{[a,b]}$. Por otro lado, las funciones

$$\alpha : i \in [\![n]\!] \mapsto \text{máx}\{\xi(i), \zeta(i)\} \in \mathbb{R}, \qquad \beta : i \in [\![n]\!] \mapsto \text{mín}\{\xi(i), \zeta(i)\} \in \mathbb{R}$$

son tales que para cada $i \in [\![n]\!]$ y cada $x \in (x_{i-1}, x_i)$ es

$$\text{máx}\{f, g\}(x) = \text{máx}\{f(x), g(x)\} = \text{máx}\{\xi(i), \zeta(i)\} = \alpha(i)$$

y

$$\text{mín}\{f, g\}(x) = \text{mín}\{f(x), g(x)\} = \text{mín}\{\xi(i), \zeta(i)\} = \beta(i),$$

así que las funciones $\text{máx}\{f, g\}$ y $\text{mín}\{f, g\}$ son escalonadas con respecto a la partición R. Finalmente, como $|f| = \text{máx}\{f, -f\}$, es claro que $|f|$ también es un elemento de $\mathscr{E}[a, b]$. □

Ahora queremos definir qué es la integral de una función escalonada. Si $f : [a, b] \to \mathbb{R}$ es una función y $P = (x_0, x_1, \ldots, x_n)$ es una partición de $[a, b]$, entonces ponemos

$$S_P(f) \coloneqq \sum_{i=1}^{n} f\left(\frac{x_{i-1} + x_i}{2}\right) \cdot (x_i - x_{i-1}).$$

Lema 13.1.2.5. *Sea $f : [a, b] \to \mathbb{R}$ una función. Si P y Q son dos particiones de $[a, b]$ con respecto a las que f es escalonada, entonces $S_P(f) = S_Q(f)$.*

Demostración. Sean P y Q dos particiones de $[a, b]$ y supongamos que f es escalonada con respecto a P y con respecto a Q. De acuerdo al Lema 13.1.1.3, hay una partición R de $[a, b]$ tal que $R \preceq P$ y $R \preceq Q$, y para mostrar que los números $S_P(f)$ y $S_P(g)$ son iguales es suficiente que probemos que ambos son iguales a $S_R(f)$.

Supongamos que $P = (x_0, x_1, \ldots, x_n)$ y que $R = (y_0, y_1, \ldots, y_m)$. Como R refina a P, hay una función $\pi : [\![0, n]\!] \to [\![0, m]\!]$ tal que $x_i = y_{\pi(i)}$ para cada $i \in [\![0, n]\!]$, y una función $\phi : [\![m]\!] \to [\![n]\!]$ tal que $(y_{j-1}, y_j) \subseteq (x_{\phi(j)-1}, x_{\phi(j)})$ para cada $j \in [\![m]\!]$. Recordemos que $\pi(0) = 0$ y $\pi(n) = m$, que

π es estrictamente creciente, y que ϕ es no decreciente.

Sea i un elemento de $[\![n]\!]$, de manera que, en particular, $1 \leq \pi(i-1)+1 \leq \pi(i) \leq m$. Si j es un elemento de $[\![\pi(i-1)+1, \pi(i)]\!] \subseteq [\![m]\!]$, entonces $(y_{j-1}+y_j)/2 \in (y_{j-1}, y_j) \subseteq (x_{\phi(j)-1}, x_{\phi(j)})$ y el Lema 13.1.1.4 nos dice que $\phi(j) = i$: esto implica, claro, que

$$f\left(\frac{y_{j-1}+y_j}{2}\right) = f\left(\frac{x_{\phi(j)-1}+x_{\phi(j)}}{2}\right) = f\left(\frac{x_{i-1}+x_i}{2}\right).$$

Vemos con esto que

$$\begin{aligned}\sum_{j=\pi(i-1)+1}^{\pi(i)} f\left(\frac{y_{j-1}+y_j}{2}\right)\cdot(y_j - y_{j-1}) &= f\left(\frac{x_{i-1}+x_i}{2}\right)\cdot \sum_{j=\pi(i-1)+1}^{\pi(i)} (y_j - y_{j-1}) \\ &= f\left(\frac{x_{i-1}+x_i}{2}\right)\cdot(y_{\pi(i)} - y_{pi(i-1)}) = f\left(\frac{x_{i-1}+x_i}{2}\right)\cdot(x_i - x_{i-1})\end{aligned}$$

y, por lo tanto, como la función π es estrictamente creciente y $\pi(0) = 0$ y $\pi(n) = m$, que

$$\begin{aligned}S_R(f) &= \sum_{j=1}^{m} f\left[\frac{y_{j-1}+y_j}{2}\right)\cdot(y_j - y_{j-1}) = \sum_{i=1}^{n}\left(\sum_{j=\pi(i-1)+1}^{\pi(i)} f\left(\frac{y_{j-1}+y_j}{2}\right)\cdot(y_j - y_{j-1})\right] \\ &= \sum_{i=1}^{n} f\left(\frac{x_{i-1}+x_i}{2}\right)\cdot(x_i - x_{i-1}) = S_P(f).\end{aligned}$$

De la misma forma, por supuesto, podemos ver que $S_R(f) = S_Q(f)$. □

Gracias a este lema podemos hacer la siguiente definición.

Definición 13.1.2.6. Si $f : [a,b] \to \mathbb{R}$ es una función escalonada y P es una partición de $[a,b]$ con respecto a la cual f es escalonada, ponemos

$$\int_a^b g \coloneqq S_P(f).$$

En efecto, el lema nos dice que el número $S_P(f)$ depende solamente de f y no de la partición P de $[a,b]$ con respecto a la cual f es escalonada.

Lema 13.1.2.7. *Sea $[a,b]$ un intervalo cerrado de $\mathbb{R}$.*

(i) La función $f \in \mathscr{E}[a,b] \mapsto \int_a^b f \in \mathbb{R}$ es lineal.

(ii) Si f y g son dos elementos de $\mathscr{E}[a,b]$ tales que $f \leq g$, entonces $\int_a^b f \leq \int_a^b g$.

(iii) Si f es un elemento de $\mathscr{E}[a,b]$, entonces $\left|\int_a^b f\right| \leq \int_a^b |f|$.

Demostración. Sea f y g dos elementos de $\mathscr{E}[a,b]$ y sean α y β dos números. Hay una partición $P = (x_0, x_1, \ldots, x_n)$ de $[a,b]$ con respecto a la cual las dos funciones son escalonadas, y vimos el Lema 13.1.2.4 nos dice que la función $\alpha f + \beta g$ también es escalonada con respecto a P. Se sigue de esto que

$$\begin{aligned}
\int_a^b (\alpha f + \beta g) &= \sum_{i=1}^n (\alpha f + \beta g)\left(\frac{x_{i-1}+x_i}{2}\right)(x_i - x_{i-1}) \\
&= \sum_{i=1}^n \left(\alpha f\left(\frac{x_{i-1}+x_i}{2}\right) + \beta g\left(\frac{x_{i-1}+x_i}{2}\right)\right)(x_i - x_{i-1}) \\
&= \alpha \sum_{i=1}^n f\left(\frac{x_{i-1}+x_i}{2}\right)(x_i - x_{i-1}) + \beta \sum_{i=1}^n g\left(\frac{x_{i-1}+x_i}{2}\right)(x_i - x_{i-1}) \\
&= \alpha \int_a^b f + \beta \int_a^b g,
\end{aligned}$$

y esto prueba la afirmación (*i*) del lema. Por otro lado, si $f \leq g$, entonces

$$\int_a^b f = \sum_{i=1}^n f\left(\frac{x_{i-1}+x_i}{2}\right)(x_i - x_{i-1}) \leq \sum_{i=1}^n f\left(\frac{x_{i-1}+x_i}{2}\right)(x_i - x_{i-1}) = \int_a^b g,$$

ya que para cada $i \in [\![n]\!]$ el número $x_i - x_{i-1}$ es positivo. Vemos así que la afirmación (*ii*) es cierta. Por otro lado, sabemos que la función $|f|$ es escalonada y que $-|f| \leq f \leq |f|$, así que lo que acabamos de probar nos dice que $-\int_a^b |f| \leq \int_a^b f \leq \int_a^b |f|$, de manera que $\left|\int_a^b f\right| \leq \int_a^b |f|$. □

Ejercicio 13.1.2.8. Sea $[a,b]$ un intervalo de $\mathbb{R}$, sea $P = (x_0, x_1, \ldots, x_n)$ una partición de $[a,b]$, y para cada $i \in [\![n]\!]$ sea $f_i : [x_{i-1}, x_i] \to \mathbb{R}$ una función escalonada sobre el intervalo $[x_{i-1}, x_i]$. Muestre que la función $f : [a,b] \to \mathbb{R}$ que en cada elemento x de $[a,b]$ toma el valor

$$f(x) = \begin{cases} f_i(x) & \text{si hay un elemento } i \text{ de } [\![n]\!] \text{ tal que } x \in [x_{i-1}, x_i); \\ f(x) & \text{si } x = b \end{cases}$$

es escalonanda en $[a,b]$ y su integral allí es

$$\int_a^b f = \sum_{i=1}^n \int_{x_{i-1}}^{x_i} f_i.$$

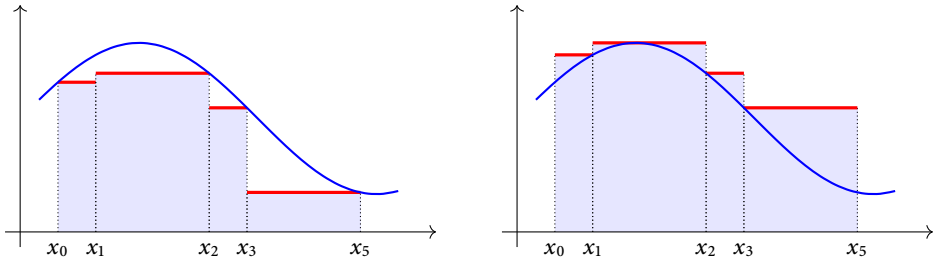


Figura 13.1. Una función escalonada inferior y una función escalonada superior.

13.1.3. La integral de Darboux

Podemos por fin definir cuándo una función es integrable y cuál es en ese caso su integral.

Definición 13.1.3.1. Sea $[a,b]$ un intervalo cerrado de $\mathbb{R}$, sea $f:[a,b]\to\mathbb{R}$ una función acotada, y consideremos los conjuntos

$$\underline{\mathscr{I}}(f) \coloneqq \left\{ \int_a^b g : g \in \mathscr{E}[a,b], g \le f \right\}, \qquad \overline{\mathscr{I}}(f) \coloneqq \left\{ \int_a^b h : h \in \mathscr{E}[a,b], h \ge f \right\}.$$

La ***integral inferior*** de f es el número

$$\underline{I}(f) \coloneqq \sup \underline{\mathscr{I}}(f)$$

y la ***integral superior*** de f es el número

$$\overline{I}(f) \coloneqq \text{ínf}\, \overline{\mathscr{I}}(f).$$

Decimos que la función f es ***integrable en el sentido de Darboux*** si $\underline{I}(f) = \overline{I}(f)$, y en ese caso la ***integral*** de f es

$$\int_a^b f \coloneqq \underline{I}(f).$$

Notemos que si $f:[a,b]\to\mathbb{R}$ es una función acotada, de manera que hay dos números m y M tales que $m \le f(x) \le M$ para todo $x \in [a,b]$, entonces las funciones constantes g y h de valor m y M, respectivamente, son elementos de $\mathscr{E}[a,b]$ tales que $g \le f \le h$: esto nos dice que los conjuntos $\underline{\mathscr{I}}(f)$ y $\overline{\mathscr{I}}(f)$ no son vacíos. Por otro lado, si u es un elemento cualquiera de $\mathscr{E}[a,b]$ tal que $u \le f$, entonces tenemos que $u \le h$ y, por lo tanto, que $\int_a^b u \le \int_a^b h$: vemos con esto que el número $\int_a^b h$ es una cota superior para el conjunto $\underline{\mathscr{I}}(f)$ y, por lo tanto, que su supremo es efectivamente un número real. De manera similar, el número $\int_a^b g$ es una cota inferior para el conjunto $\overline{\mathscr{I}}(f)$, así que el ínfimo de ese conjunto es también un número real.

Como dijimos en la definición, cuando $f : [a, b] \to \mathbb{R}$ es una función integrable, escribimos $\int_a^b f$ a su integral. Muchas veces escribimos también a este número en la forma

$$\int_a^b f(x)\,\mathrm{d}x.$$

Así, por ejemplo, si f es la función $x \in [0,1] \mapsto e^x \cos x \in \mathbb{R}$, entonces en lugar de $\int_0^1 f$ podemos escribir

$$\int_0^1 e^x \cos x\,\mathrm{d}x.$$

La ventaja de esto es que podemos escribir integrales de funciones sin ponerles nombre. Es importante observar que el signo dx que aparece aquí es solamente parte de la notación y no tiene ningún sentido propio. Más aún, la letra x que aparece en $f(x)$ o en dx en la notación $\int_a^b f(x)\,\mathrm{d}x$ es solo una variable, y la expresión no depende de ella. Así, si x e y son dos variables, entonces

$$\int_a^b e^x \cos x\,\mathrm{d}x = \int_a^b e^y \cos y\,\mathrm{d}y$$

simplemente porque las expresiones que aparecen a los dos lados del signo = significan precisamente lo mismo.

Una primera observación importante que tenemos que hacer es la siguiente.

Lema 13.1.3.2. *Si $f : [a, b] \to \mathbb{R}$ es una función acotada, entonces $\underline{I}(f) \leq \overline{I}(f)$.*

En vista de esto, para probar que una función $f : [a, b] \to \mathbb{R}$ es integrable es suficiente con mostrar que $\overline{I}(f) \leq \underline{I}(f)$.

Demostración. Si g y h son elementos de $\mathscr{E}[a, b]$ tales que $g \leq f \leq h$, entonces $\int_a^b g \leq \int_a^b h$. Esto nos dice que todo elemento del conjunto $\underline{\mathscr{I}}(f)$ es una cota inferior para el conjunto $\overline{\mathscr{I}}(f)$ y que todo elemento de $\overline{\mathscr{I}}(f)$ es una cota superior para el conjunto $\underline{\mathscr{I}}(f)$, de manera que $\underline{I}(f) = \sup \underline{\mathscr{I}}(f) \leq \inf \overline{\mathscr{I}}(f) = \overline{I}(f)$. □

Las propiedades de las funciones escalonadas y sus integrales se ven reflejadas en propiedades de las funciones integrables y sus integrales, como es de esperar. El siguiente es un ejemplo de ello.

Proposición 13.1.3.3. *Sea $[a, b]$ un intervalo cerrado de $\mathbb{R}$.*

(*i*) *El conjunto $\mathscr{D}[a, b]$ de todas las funciones $[a, b] \to \mathbb{R}$ que son integrables es un subespacio de $\mathbb{R}^{[a,b]}$.*

(*ii*) *La función $f \in \mathscr{D}[a, b] \mapsto \int_a^b f \in \mathbb{R}$ es lineal.*

Demostración. Sean $f, g : [a,b] \to \mathbb{R}$ dos elementos de $\mathscr{D}[a,b]$. Si u y v son dos elementos de $\mathscr{E}[a,b]$ tales que $u \leq f$ y $v \leq g$, entonces $u+v$ es una función escalonada en $[a,b]$ y $u+v \leq f+g$, así que

$$\int_a^b u + \int_a^b v = \int_a^b (u+v) \leq \underline{I}(f+g).$$

Esto nos dice que, en términos de la notación de la Definición 13.1.3.1, que $\underline{I}(f+g)$ es una cota superior para el conjunto

$$\underline{\mathscr{I}}(f) + \underline{\mathscr{I}}(g) \coloneqq \{s+t : s \in \underline{\mathscr{I}}(f), t \in \underline{\mathscr{I}}(g)\}.$$

Como los conjuntos $\underline{\mathscr{I}}(f)$ y $\underline{\mathscr{I}}(g)$ están acotados superiormente, es entonces

$$\underline{I}(f) + \underline{I}(g) = \sup \underline{\mathscr{I}}(f) + \sup \underline{\mathscr{I}}(g) = \sup(\underline{\mathscr{I}}(f) + \underline{\mathscr{I}}(g)) \leq \underline{I}(f+g).$$

Un argumento simétrico prueba que también $\overline{I}(f+g) \leq \overline{I}(f) + \overline{I}(g)$, así que

$$\underline{I}(f) + \underline{I}(g) \leq \underline{I}(f+g) \leq \overline{I}(f+g) \leq \overline{I}(f) + \overline{I}(g).$$

Como $\underline{I}(f) = \overline{I}(f)$ e $\underline{I}(g) = \overline{I}(g)$, esto nos permite concluir que $\underline{I}(f+g) = \overline{I}(f+g)$, esto es, que la función $f+g$ es integrable y, más aún, que su integral es

$$\int_a^b (f+g) = \int_a^b f + \int_a^b g.$$

Sea ahora $f : [a,b] \to \mathbb{R}$ un elemento de $\mathscr{D}[a,b]$, sea α un escalar, y supongamos primero que $\alpha \geq 0$. Si $u : [a,b] \to \mathbb{R}$ es una función escalonada tal que $u \leq f$, entonces también es $\alpha u \leq \alpha f$ y, por lo tanto,

$$\alpha \int_a^b u = \int_a^b \alpha u \leq \underline{I}(\alpha f).$$

Esto implica que $\alpha^{-1}\underline{I}(\alpha f)$ es una cota superior para el conjunto $\underline{\mathscr{I}}(f)$ y, en consecuencia, que $\underline{I}(f) \leq \alpha^{-1}\underline{I}(\alpha f)$ y $\alpha\underline{I}(f) \leq \underline{I}(\alpha f)$. De manera similar podemos ver $\overline{I}(\alpha f) \leq \alpha\overline{I}(f)$, así que

$$\alpha\underline{I}(f) \leq \underline{I}(\alpha f) \leq \overline{I}(\alpha f) \leq \alpha\overline{I}(f).$$

Como $\underline{I}(f) = \overline{I}(f)$, podemos concluir que $\underline{I}(\alpha f) = \overline{I}(\alpha f)$, de manera que la función αf es integrable, y que su integral es igual a $\alpha \int_a^b f$.

Supongamos ahora que $\alpha < 0$. Como $-\alpha > 0$, sabemos ya que $-\alpha f = (-\alpha) f$ es integrable y que su integral es $-\alpha \int_a^b f$. Para probar que αf es integrable y que su integral es $\alpha \int_a^b f$, entonces, es suficiente que mostremos esto cuando α es -1.

Ahora bien, si t es un elemento del conjunto $\underline{\mathscr{I}}(f)$, entonces hay una función escalonada g tal que $g \leq f$ y $t = \int_a^b g$, y esto implica que $-g$ es una función escalonada tal que $-f \leq -g$ y, por lo tanto, que $-t = \int_a^b (-g) \in \overline{\mathscr{I}}(-f)$. Vemos así que

$$\{-t : t \in \overline{\mathscr{I}}(f)\} \subseteq \underline{\mathscr{I}}(-f).$$

Esto es cierto cualquiera sea la función acotada $f : [a,b] \to \mathbb{R}$, así que, en particular, también tenemos que $\{-t : t \in \overline{\mathscr{I}}(-f)\} \subseteq \underline{\mathscr{I}}(f)$, así que $\underline{\mathscr{I}}(-f) \subseteq \{-t : t \in \overline{\mathscr{I}}(f)\}$. Juntando todo, vemos que, de hecho, $\underline{\mathscr{I}}(-f) = \{-t : t \in \overline{\mathscr{I}}(f)\}$ y, por lo tanto, que

$$\underline{I}(-f) = \sup \underline{\mathscr{I}}(-f) = \sup\{-t : t \in \overline{\mathscr{I}}(f)\} = -\inf \overline{\mathscr{I}}(f) = -\overline{I}(f).$$

Un argumento similar prueba que $\overline{I}(-f) = -\underline{I}(f)$ y, como f es integrable, que $-f$ lo es y que su integral es $-\int_a^b f$. □

Otra propiedad que hereda la integral de las funciones integrables de la de las funciones escalonadas es la monotonía.

Proposición 13.1.3.4. *Sea $[a,b]$ un intervalo cerrado de $\mathbb{R}$. Si f y g son dos elementos de $\mathscr{D}[a,b]$ tales que $f \leq g$, entonces $\int_a^b f \leq \int_a^b g$.*

Demostración. Sean $f, g : [a,b] \to \mathbb{R}$ dos funciones integrables en $[a,b]$ tales que $f \leq g$, y sean $u, v : [a,b] \to \mathbb{R}$ dos funciones escalonadas tales que $u \leq f$ y $g \leq v$. Tenemos entonces que $u \leq v$ y, por lo tanto, que $\int_a^b u \leq \int_a^b v$. Esto nos dice que $\int_a^b v$ es una cota superior para el conjunto $\underline{\mathscr{I}}(f)$, así que $\underline{I}(f) = \sup \underline{\mathscr{I}}(f) \leq \int_a^b v$, y esto, a su vez, nos dice que $\underline{I}(f)$ es una cota inferior para el conjunto $\overline{\mathscr{I}}(g)$, así que

$$\int_a^b f = \underline{I}(f) \leq \inf \overline{\mathscr{I}}(g) = \overline{I}(g) = \int_a^b g.$$ □

En este momento tenemos dos definiciones distintas de lo que queremos decir cuando hablamos de la integral de una función escalonada. De hecho, estas dos definiciones coinciden.

Lema 13.1.3.5. *Una función escalonada es integrable y su integral en el sentido de la Definición 13.1.3.1 coincide con su integral en el sentido de la Definición 13.1.2.6.*

Toda función escalonada es acotada, así que la afirmación del lema tiene sentido.

Demostración. Sea $f : [a,b] \to \mathbb{R}$ una función escalonada y sea P una partición de $[a,b]$ con respecto a la cual f es escalonada. Como f es escalonada y, por supuesto, $f \leq f$, es claro que $\overline{I}(f) \leq S_P(f) \leq \underline{I}(f)$, así que, de hecho, los tres números $\overline{I}(f)$, $S_P(f)$ y $\underline{I}(f)$ son iguales. □

Antes de continuar, calculemos una integral concreta usando la definición.

Ejemplo 13.1.3.6. Sea n un entero no negativo, y consideremos la función $f : x \in [0,1] \mapsto x^n \in \mathbb{R}$, que claramente es acotada. Sea N un entero positivo, para cada $i \in [\![N]\!]$ sea $x_i := i/N$, y consideremos la secuencia $P = (x_0, x_1, \ldots, x_N)$, que es una partición de $[0,1]$, y las funciones $g_N, h_N : [0,1] \to \mathbb{R}$ que en cada elemento x de $[0,1]$ toman valores

$$g_N(x) = \begin{cases} x_{i-1}^n & \text{si } x_{i-1} \le x < x_i \text{ para algún } i \in [\![N]\!]; \\ 1 & \text{si } x = 1; \end{cases}$$

y

$$h_N(x) = \begin{cases} 0 & \text{si } x = 0; \\ x_i^n & \text{si } x_{i-1} < x \le x_i \text{ para algún } i \in [\![N]\!]; \end{cases}$$

Es claro que se trata de funciones escalonadas con respecto a la partición P y que $g_N \le f \le h_N$. Es

$$\int_0^1 g_N = S_P(g_N) = \sum_{i=1}^N x_{i-1}^n (x_i - x_{i-1}) = \sum_{i=1}^N \frac{(i-1)^n}{N^n} \cdot \frac{1}{N} = \frac{1}{N^{n+1}} \sum_{i=0}^{N-1} i^n$$

y

$$\int_0^1 h_N = S_P(h_N) = \sum_{i=1}^N x_i^n (x_i - x_{i-1}) = \sum_{i=1}^N \frac{i^n}{N^n} \cdot \frac{1}{N} = \frac{1}{N^{n+1}} \sum_{i=1}^N i^n,$$

así que

$$0 \le \overline{I}(f) - \underline{I}(f) \le \int_0^1 h_N - \int_0^1 g_N = \frac{1}{N^{n+1}} \left(\sum_{i=1}^N i^n - \sum_{i=0}^{N-1} i^n \right) = \frac{N^n - 1}{N^{n+1}}.$$

Como

$$\lim_{N \to \infty} \frac{N^n - 1}{N^{n+1}} = 0,$$

podemos concluir que $\overline{I}(f) = \underline{I}(f)$ y, por lo tanto, que la función f es integrable.

Queremos determinar su integral. Para ello observemos primero que

la función $\Delta : p \in \mathbb{R}[X] \mapsto p(X) - p(X-1) \in \mathbb{R}[X]$ *es lineal y sobreyectiva, y que su núcleo es el subespacio de* $\mathbb{R}[X]$ *de los polinomios constantes.* (13.1)

Que es lineal es claro. Por otro lado, para cada $k \in \mathbb{N}$ es

$$\Delta(X^k) = X^k - (X-1)^k = X^k - \sum_{i=0}^k (-1)^{k-i} \binom{k}{i} X^i = \sum_{i=0}^{k-1} (-1)^{k-i+1} \binom{k}{i} X^i,$$

y esto es un polinomio de grado $n-1$. Vemos así que la imagen de Δ contiene polinomios de todos los grados, así que coincide con $\mathbb{R}[X]$. Más aún, esto prueba que

si p es un polinomio de grado d, entonces $\Delta(p)$ tiene grado $d-1$.

Por otro lado, si p es un elemento de $\mathbb{R}[X]$ tal que $\Delta(p) = 0$, entonces tenemos que $p(X) = p(X-1)$. Esto implica que para cada entero n es $p(n) = p(n-1)$ y, por lo tanto, que hay un número real ξ_n en el intervalo $(n-1, n)$ tal que $p'(\xi_n) = 0$. La función $n \in \mathbb{Z} \mapsto \xi_n \in \mathbb{R}$ es claramente inyectiva, así que el polinomio p' tiene infinitas raíces en $\mathbb{R}$: se trata entonces del polinomio nulo y, en consecuencia, el polinomio p es constante.

Como consecuencia de (13.1) podemos concluir que hay exactamente un polinomio p_n en $\mathbb{R}[X]$ tal que $\Delta(p_n) = X^n$ y $p_n(0) = 0$. Si N es un elemento cualquiera de $\mathbb{N}$, entonces es

$$p_n(N) = p_n(N) - p_n(0) = \sum_{i=1}^{N} (p_n(i) - p_n(i-1)) = \sum_{i=1}^{N} i^n.$$

Supongamos que p_n tiene grado d y que su coeficiente principal es α, de manera que $p_n = \alpha X^d + q$, con q un polinomio que o es nulo o no es nulo y tiene grado menor que d y, por lo tanto,

$$X^n = \Delta(p_n) = \alpha X^d - \alpha(X-1)^d + \Delta(q) = \sum_{i=0}^{d-1} (-1)^{d-i+1} \binom{d}{i} X^i + \Delta(q).$$

Este último polinomio tiene término principal $\alpha d X^{d-1}$, ya que $\Delta(q)$ tiene grado menor que $d-1$, y esto nos permite concluir que $d = n+1$ y que $\alpha = 1/(n+1)$. Hay entonces números reales $\alpha_0, \ldots, \alpha_n$ tales que

$$p_n(X) = \frac{1}{n+1} X^{n+1} + \alpha_n X^n + \cdots + \alpha_1 X + \alpha_0.$$

Todo esto implica que para cada $N \in \mathbb{N}$ es

$$\int_0^1 h_N = \frac{1}{N^{n+1}} \sum_{i=1}^{N} i^n = \frac{p_n(N)}{N^{n+1}} = \frac{1}{n+1} + \frac{\alpha_n}{N} + \cdots + \frac{\alpha_1}{N^n} + \frac{\alpha_0}{N^{n+1}}$$

y

$$\begin{aligned} \int_0^1 g_N &= \frac{1}{N^{n+1}} \sum_{i=1}^{N-1} i^n = \frac{(N-1)^{n+1}}{N^{n+1}} \cdot \frac{p_n(N-1)}{(N-1)^{n+1}} \\ &= \left(1 - \frac{1}{N}\right)^{n+1} \left(\frac{1}{n+1} + \frac{\alpha_n}{N-1} + \cdots + \frac{\alpha_1}{(N-1)^n} + \frac{\alpha_0}{(N-1)^{n+1}} \right), \end{aligned}$$

y es claro entonces que

$$\overline{I}(f) \leq \lim_{N\to\infty} \int_0^1 h_N = \frac{1}{n+1} \leq \lim_{N\to\infty} \int_0^1 g_N \leq \underline{I}(f).$$

Vemos con esto que

$$\int_0^1 f = \frac{1}{n+1}.$$

Veamos ahora un ejemplo de una función que no es integrable.

Ejemplo 13.1.3.7. Consideremos la llamada ***función de Dirichlet*** $f : [0,1] \to \mathbb{R}$, que en cada elemento x de $[0,1]$ toma el valor

$$f(x) = \begin{cases} 1 & \text{si } x \in \mathbb{Q}; \\ 0 & \text{en caso contrario.} \end{cases}$$

Sea $g : [0,1] \to \mathbb{R}$ una función escalonada con respecto a una partición $P = (x_0, x_1, \ldots, x_n)$ tal que $g \leq f$. Para cada $i \in [\![n]\!]$ hay un elemento x del intervalo (x_{i-1}, x_i) que es irracional y, por lo tanto, $g(x) \leq f(x) = 0$: como la función g es constante en el intervalo (x_{i-1}, x_i), esto nos dice que $g((x_{i-1} + x_i)/2) \leq 0$. Esto implica que

$$\int_0^1 g = S_P(g) = \sum_{i=1}^{n} g\left(\frac{x_{i-1} + x_i}{2}\right)(x_i - x_{i-1}) \leq 0.$$

Vemos así que 0 es una cota superior para el conjunto $\underline{\mathscr{I}}(f)$, así que $\underline{I}(f) \leq 0$.

Sea, por otro lado, $h : [0,1] \to \mathbb{R}$ una función escalonada con respecto a una partición $Q = (y_0, y_1, \ldots, y_m)$ tal que $f \leq h$. Para cada $j \in [\![m]\!]$ hay un elemento y del intervalo (y_{j-1}, y_j) que es racional y, por lo tanto, $1 = f(y) \leq h(y)$: como la función h es constante en el intervalo (y_{j-1}, y_j), esto nos dice que $h((y_{j-1} + y_j)/2) \geq 1$. Esto implica que

$$\int_0^1 h = S_Q(h) = \sum_{j=1}^{m} g\left(\frac{y_{j-1} + y_j}{2}\right)(y_j - y_{j-1}) \geq \sum_{j=1}^{m}(y_j - y_{j-1}) = y_m - y_0 = 1.$$

Vemos así que 1 es una cota inferior para el conjunto $\overline{\mathscr{I}}(f)$, así que $\overline{I}(f) \geq 1$.

Juntando todo vemos que

$$\overline{I}(f) - \underline{I}(f) \geq 1 - 0 = 1$$

y, por lo tanto, que $\overline{I}(f) \neq \underline{I}(f)$, de manera que la función f no es integrable.

Si $f : [a,b] \to \mathbb{R}$ es una función acotada y P es una partición del intervalo $[a,b]$, entre todas las funciones escalonadas g con respecto a P tales que $g \leq f$ hay una que es la más grande de todas. El siguiente lema la identifica explícitamente.

Lema 13.1.3.8. *Sea $f : [a,b] \to \mathbb{R}$ una función acotada, sea $P = (x_0, x_1, \ldots, x_n)$ una partición de $[a,b]$, para cada $i \in [\![n]\!]$ pongamos*

$$m_i \coloneqq \text{ínf}\{f(x) : x \in [x_{i-1}, x_i]\}, \qquad M_i \coloneqq \sup\{f(x) : x \in [x_{i-1}, x_i]\},$$

y consideremos las funciones f_P, $f^P : [a,b] \to \mathbb{R}$ *que en cada elemento* x *de* $[a,b]$ *toman el valor*

$$f_P(x) = \begin{cases} m_i & \text{si hay un elemento } i \in [\![n]\!] \text{ tal que } x \in (x_{i-1}, x_i); \\ f(x) & \text{si } x \in \{x_0, x_1, \ldots, x_n\}; \end{cases}$$

y

$$f^P(x) = \begin{cases} M_i & \text{si hay un elemento } i \in [\![n]\!] \text{ tal que } x \in (x_{i-1}, x_i). \\ f(x) & \text{si } x \in \{x_0, x_1, \ldots, x_n\} \end{cases}$$

(*i*) *Las funciones* f_P *y* f^P *son escalonadas con respecto a la partición* P*, y es* $f_P \le f \le f^P$.
(*ii*) *Si* $g : [a,b] \to \mathbb{R}$ *es una función escalonada con respecto a la partición* P *tal que* $g \le f$*, entonces* $g \le f_P$.
(*iii*) *Si* $h : [a,b] \to \mathbb{R}$ *es una función escalonada con respecto a la partición* P *tal que* $f \le g$*, entonces* $f^P \le h$.

Demostración. Todas las afirmaciones del lema son evidentes. □

13.1.4. El criterio de integrabilidad de Darboux

Analizando en detalle el argumento que usamos en el Ejemplo 13.1.3.6 para probar que la función es integrable podemos observar que depende solamente de que se trata de una función monótona. De hecho, vale el siguiente resultado.

Proposición 13.1.4.1. *Una función monótona definida en un intervalo cerrado y acotado de* $\mathbb{R}$ *es integrable.*

Notemos que una función monótona sobre un intervalo cerrado y acotado es automáticamente acotada, así que la afirmación de la proposición tiene sentido.

Demostración. Sea $f : [a,b] \to \mathbb{R}$ una función creciente — el caso de las funciones decrecientes es completamente similar. Fijemos un entero positivo n, para cada elemento i de $[\![0,n]\!]$ sea $x_i := a + i(b-a)/n$, y consideremos la secuencia $P_n := (x_0, x_1, \ldots, x_n)$, que es una partición del intervalo $[a,b]$. Para cada $i \in [\![n]\!]$ pongamos

$$m_i := \text{ínf}\{f(x) : x \in (x_{i-1}, x_i)\}, \qquad M_i := \sup\{f(x) : x \in (x_{i-1}, x_i)\}.$$

Por supuesto, $m_i \le M_i$ para todo $i \in [\![n]\!]$. Por otro lado, si $i \in [\![n-1]\!]$, entonces $f(x) \le f(y)$ siempre que $x \in (x_{i-1}, x_i)$ e $y \in (x_i, x_{i+1})$ porque f es creciente, y esto implica que $M_i \le m_{i+1}$. Usando esto vemos que

$$0 \le \sum_{i=1}^{n} (M_i - m_i) = \sum_{i=1}^{n} M_i - \sum_{i=1}^{n} m_i = M_n + \sum_{i=1}^{n-1} (M_i - m_{i+1}) - m_1 \le M_n - m_1$$

$$\leq f(b) - f(a).$$

Sean ahora $f_P, f^P : [a,b] \to \mathbb{R}$ las funciones descriptas en el Lema 13.1.3.8, que son escalonadas con respecto a la partición P de $[a,b]$ y tales que $f_P \leq f \leq f^P$. Tenemos entonces que

$$0 \leq \overline{I}(f) - \underline{I}(f) \leq S_P(f^P) - S_P(f_P) = \sum_{i=1}^{n} M_i(x_i - x_{i-1}) - \sum_{i=1}^{n} m_i(x_i - x_{i-1})$$
$$= \sum_{i=1}^{n}(M_i - m_i)(x_i - x_{i-1}) = \frac{1}{n}\sum_{i=1}^{n}(M_i - m_i) \leq \frac{f(b) - f(a)}{n}.$$

Como esto es cierto cualquiera sea el entero positivo n, nos permite concluir que $\overline{I}(f) = \underline{I}(f)$, esto es, que la función f es integrable. □

La razón por la que esta demostración funciona es que podemos usar el hecho de que la función f es monótona para acotar la variación en cada uno de los subintervalos de la partición de las funciones escalonadas con las que acotamos a f. Un argumento del mismo tipo nos permite mostrar que las funciones continuas son integrables, apoyándonos ahora no en la monotonía sino en la continuidad uniforme.

Proposición 13.1.4.2. *Una función continua definida en un intervalo cerrado y acotado es integrable.*

Demostración. Sea $f : [a,b] \to \mathbb{R}$ una función continua, y sea ε un número positivo. Como f es uniformemente continua en $[a,b]$, hay un número positivo δ tal que para todo par de puntos x y x' de $[a,b]$ vale $|x - x'| < \delta \implies |f(x) - f(x')| < \varepsilon$.

Sea n un entero positivo tal que $1/n < \varepsilon$. Para cada elemento i de $[\![0,n]\!]$ sea $x_i \coloneqq a + i(b-a)/n$, y consideremos la secuencia $P_n \coloneqq (x_0, x_1, \ldots, x_n)$, que es una partición del intervalo $[a,b]$. Para cada $i \in [\![n]\!]$ pongamos

$$m_i \coloneqq \text{ínf}\{f(x) : x \in (x_{i-1}, x_i)\}, \qquad M_i \coloneqq \sup\{f(x) : x \in (x_{i-1}, x_i)\}.$$

Sea i un elemento de $[\![n]\!]$. Es $|f(x) - f(x')| < \varepsilon$ si x y x' son dos elementos de (x_{i-1}, x_i), ya que el diámetro de este intervalo es menor que δ, y esto implica que $-\varepsilon < f(x) - f(x') < \varepsilon$. Vemos así que $f(x') - \varepsilon$ es una cota inferior para el conjunto $\{f(y) : y \in (x_{i-1}, x_i)\}$, así que $f(x') - \varepsilon \leq m_i$. Esto nos dice, a su vez, que $m_i + \varepsilon$ es una cota superior para $\{f(y) : y \in (x_{i-1}, x_i)\}$, de manera que $M_i \leq m_i + \varepsilon$. Todo esto nos permite concluir que para todo elemento i de $[\![n]\!]$ es $M_i - m_i \leq \varepsilon$.

Sean $f_P, f^P : [a,b] \to \mathbb{R}$ las funciones descriptas en el Lema 13.1.3.8, que son escalonadas con respecto a la partición P de $[a,b]$ y tales que $f_P \leq f \leq f^P$. Tenemos que

$$0 \leq \overline{I}(f) - \underline{I}(f) \leq S_P(f^P) - S_P(f_P) = \sum_{i=1}^{n} M_i(x_i - x_{i-1}) - \sum_{i=1}^{n} m_i(x_i - x_{i-1})$$

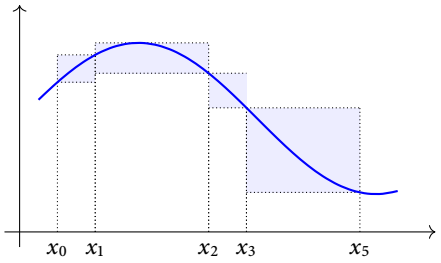


Figura 13.2. La oscilación de la función con respecto a la partición $(x_0, x_1, x_2, x_3, x_4)$ es el area pintada.

$$= \sum_{i=1}^{n} (M_i - m_i)(x_i - x_{i-1}) \leq \varepsilon \sum_{i=1}^{n} (x_i - x_{i-1}) = (b-a)\varepsilon,$$

ya que $x_i - x_{i-1} > 0$ para todo $i \in [\![n]\!]$. Como esto es así cualquiera sea el número positivo ε, vemos que $\overline{I}(f) = \underline{I}(f)$ y, en consecuencia, que la función f es integrable. □

Notemos que en las demostraciones de las Proposiciones 13.1.4.1 y 13.1.4.2 hicimos esencialmente lo mismo. Podemos abstraer la idea y obtener un criterio general de integrabilidad originalmente debido a Darboux. Empecemos, como siempre, con una definición.

Definición 13.1.4.3. Sea $f : [a,b] \to \mathbb{R}$ una función acotada. Si $P = (x_0, x_1, \ldots, x_n)$ es una partición del intervalo $[a,b]$, entonces la ***oscilación*** de f con respecto a P es el número

$$\omega_P(f) \coloneqq \sum_{i=1}^{n} \left(\sup_{x \in (x_{i-1}, x_i)} f(x) - \inf_{x \in (x_{i-1}, x_i)} f(x) \right) (x_i - x_{i-1}).$$

Lema 13.1.4.4. *Sea $f : [a,b] \to \mathbb{R}$ una función acotada, sea $P = (x_0, x_1, \ldots, x_n)$ una partición del intervalo $[a,b]$, y sean f_P, $f^P : [a,b] \to \mathbb{R}$ las funciones descriptas en el Lema 13.1.3.8, que son escalonadas con respecto a P. Es*

$$\omega_P(f) = S_P(f^P) - S_P(f_P),$$

Notemos que esta diferencia $S_P(f^P) - S_P(f_P)$ fue precisamente la que usamos al probar las Proposiciones 13.1.4.1 y 13.1.4.2.

Demostración. La igualdad de que afirma el lema es consecuencia directa de las definiciones de $\omega_P(f)$, f_P y f^P. □

El siguiente criterio de integrabilidad, debido a Darboux, nos dice que una función es integrable exactamente su oscilación puede hacerse arbitrariamente pequeña.

Proposición 13.1.4.5. *Una función acotada $f : [a, b] \to \mathbb{R}$ es integrable si y solamente si para todo número positivo ε hay una partición $P = (x_0, x_1, \ldots, x_n)$ del intervalo $[a, b]$ tal que $\omega_P(f) < \varepsilon$.*

Demostración. Sea $f : [a, b] \to \mathbb{R}$ una función acotada, supongamos que la condición de la proposición se satisface, y sea ε un número positivo. Hay una partición P de $[a, b]$ tal que $\omega_P(f) < \varepsilon$, y si $f_P, f^P : [a, b] \to \mathbb{R}$ son las funciones descriptas en el Lema 13.1.3.8, que son escalonadas con respecto a P y tales que $f_P \leq f \leq f^P$, entonces

$$0 \leq \overline{I}(f) - \underline{I}(f) \leq S_P(f^P) - S_P(f_P) = \omega_P(f) < \varepsilon.$$

Esto es así cualquiera sea el número positivo ε, así que $\overline{I}(f) = \underline{I}(f)$ y la función f es integrable. La condición de la proposición es, por lo tanto, suficiente.

Supongamos ahora que $f : [a, b] \to \mathbb{R}$ es una función acotada que es integrable, y sea ε un número positivo. Como

$$\sup \underline{\mathscr{I}}(f) = \underline{I}(f) = \overline{I}(f) = \text{ínf } \overline{\mathscr{I}}(f),$$

hay elementos s y t de $\underline{\mathscr{I}}(f)$ y de $\overline{\mathscr{I}}(f)$, respectivamente, tales que $t - s < \varepsilon$, y entonces hay funciones escalonadas $g, h : [a, b] \to \mathbb{R}$ con respecto a particiones P y Q de $[a, b]$ tales que $g \leq f \leq h$, $s = \int_a^b g$ y $t = \int_a^b h$. Sabemos que hay una partición R de $[a, b]$ que refina a P y a Q, y que g y h son escalonadas también con respecto a R.

Sean $f_R, f^R : [a, b] \to \mathbb{R}$ las funciones descriptas en el Lema 13.1.3.8, que son escalonadas con respecto a P y tales que $g \leq f_P \leq f \leq f^P \leq h$. Se sigue de esto que

$$\omega_R(f) = S_R(f^R) - S_R(f_R) \leq S_R(h) - S_R(g) = \int_a^b h - \int_a^b g = t - s < \varepsilon.$$

Esto demuestra que la condición de la proposición también es necesaria. □

El siguiente lema es útil al tratar de controlar cómo cambia la oscilación de una función con respecto a una partición cuando cambiamos la partición.

Lema 13.1.4.6. *Sea $f : [a, b] \to \mathbb{R}$ una función acotada. Si P y Q son dos particiones de $[a, b]$ y $P \leq Q$, entonces $\omega_P(f) \leq \omega_Q(f)$.*

Demostración. Sean $P = (x_0, x_1, \ldots, x_n)$ y $Q = (y_0, y_1, \ldots, y_m)$ dos particiones de $[a, b]$ y supongamos que $P \leq Q$. Sean $f_P, f_Q, f^P, f^Q : [a, b] \to \mathbb{R}$ como en el Lema 13.1.3.8. Como f_Q y f^Q

son escalonadas con respecto a Q y $P \preceq Q$, esas dos funciones son escalonadas también con respecto a P. Como $f_Q \leq f \leq f^Q$, ese lema nos dice que $f_Q \leq f_P$ y que $f^P \leq f^Q$. Se sigue de estas desigualdades que

$$\omega_P(f) = S_P(f^P) - S_P(f_P) \leq S_P(f^Q) - S_P(f_Q) = S_Q(f^Q) - S_Q(f_Q) = \omega_Q(f),$$

y esto es lo que afirma el lema. □

Vamos a dar varios ejemplos de cómo podemos usar el *Criterio de Darboux*. El primero es el siguiente resultado que nos dice que una función que es integrable en un intervalo $[a, b]$ también lo es en cualquier subintervalo de $[a, b]$.

Proposición 13.1.4.7. *Sea $f : [a, b] \to \mathbb{R}$ una función integrable. Si $[c, d]$ es un intervalo contenido en $[a, b]$, entonces la restricción $f|_{[c,d]} : [c, d] \to \mathbb{R}$ es integrable.*

Demostración. Sea $[c, d]$ un intervalo contenido en $[a, b]$, sea n el entero tal que el cardinal del conjunto $\{a, c, d, b\}$ es $n + 1$, y sean $x_0, x_1, \ldots, x_n$ los elementos de ese conjunto listados en orden creciente. La secuencia $P_0 \coloneqq (x_0, x_1, \ldots, x_n)$ es entonces una partición del intervalo $[a, b]$.

Sea ε un número positivo. Como f es integrable, hay una partición P de $[a, b]$ tal que $\omega_P(f) < \varepsilon$, y sabemos que hay una partición $P' = (y_0, y_1, \ldots, y_m)$ de $[a, b]$ que es un refinamiento común de P y de P_0. En particular, hay dos elementos i_1 e i_2 en $[\![0, m]\!]$ tales que $i_1 \leq i_2$, $y_{i_1} = c$ e $y_{i_2} = d$ y, en consecuencia, la secuencia $P'' \coloneqq (y_{i_1}, y_{i_1+1}, \ldots, y_{i_2})$ es una partición de $[c, d]$ y, gracias al Lema 13.1.4.6,

$$\begin{aligned}\omega_{P''}(f|_{[c,d]}) &= \sum_{i=i_1+1}^{i_2} \left(\sup_{x \in (y_{i-1}, y_i)} f(x) - \inf_{x \in (y_{i-1}, y_i)} f(x) \right) (y_i - y_{i-1}) \\ &\leq \sum_{i=1}^{m} \left(\sup_{x \in (y_{i-1}, y_i)} f(x) - \inf_{x \in (y_{i-1}, y_i)} f(x) \right) (y_i - y_{i-1}) = \omega_{P'}(f) \leq \omega_P(f) < \varepsilon.\end{aligned}$$

Esto prueba que la restricción $f|_{[c,d]} : [c, d] \to \mathbb{R}$ es integrable. □

El siguiente resultado va en la dirección opuesta al del de la Proposición 13.1.4.7: nos permite concluir la integrabilidad de una función en un intervalo si es integrable en *suficientes* subintervalos de este.

Proposición 13.1.4.8. *Sea $f : [a, b] \to \mathbb{R}$ una función acotada, sea $P = (x_0, x_1, \ldots, x_n)$ una partición de $[a, b]$, y supongamos que para cada $i \in [\![n]\!]$ la restricción $f|_{[x_{i-1}, x_i]} : [x_{i-1}, x_i] \to \mathbb{R}$ es*

integrable. La función f es integrable y su integral es

$$\int_a^b f = \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} f.$$

Demostración. Sea ε un número positivo. Para cada $i \in [\![n]\!]$ la restricción $f|_{[x_{i-1},x_i]}$ es integrable, así que hay funciones escalonadas $g_i, h_i : [x_{i-1}, x_i] \to \mathbb{R}$ tales que

$$g_i \leq f|_{[x_{i-1},x_i]} \leq h_i, \qquad \int_{x_{i-1}}^{x_i} g_i + \frac{\varepsilon}{n} \geq \int_{x_{i-1}}^{x_i} f \geq \int_{x_{i-1}}^{x_i} h_i - \frac{\varepsilon}{n}.$$

De acuerdo al Ejercicio 13.1.2.8, las funciones $g, h : [a, b] \to \mathbb{R}$ que en cada elemento x de $[a, b]$ toman los valores

$$g(x) = \begin{cases} g_i(x) & \text{si hay un elemento } i \text{ de } [\![n]\!] \text{ tal que } x \in [x_{i-1}, x_i); \\ g(x) & \text{si } x = b; \end{cases}$$

y

$$h(x) = \begin{cases} h_i(x) & \text{si hay un elemento } i \text{ de } [\![n]\!] \text{ tal que } x \in [x_{i-1}, x_i); \\ h(x) & \text{si } x = b \end{cases}$$

son escalonadas, y claramente tenemos que $g \leq f \leq h$. Más aún, ese ejercicio nos dice que

$$\overline{I}(f) \leq \int_a^b h = \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} h_i \leq \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} f + \varepsilon$$

y

$$\underline{I}(f) \geq \int_a^b g = \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} g_i \geq \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} f - \varepsilon,$$

así que

$$\sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} f - 2\varepsilon \leq \underline{I}(f) \leq \overline{I}(f) \leq \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} f + 2\varepsilon.$$

Esto es cierto cualquiera sea el número positivo ε, así que podemos concluir que

$$\sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} f \leq \underline{I}(f) \leq \overline{I}(f) \leq \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} f$$

Podemos concluir a partir de esto que $\underline{I}(f) = \overline{I}(f)$, así que la función f es integrable en $[a, b]$, y que su integral es

$$\int_a^b f = \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} f,$$

como afirma la proposición. □

En la Definición 13.1.3.1 dimos sentido a la expresión $\int_a^b f$ cuando $[a, b]$ es un intervalo cerrado y acotado de $\mathbb{R}$ y $f : [a, b] \to \mathbb{R}$ es una función integrable. Es conveniente extender esa notación de la siguiente manera.

Definición 13.1.4.9. Sea $[a, b]$ un intervalo cerrado y acotado de $\mathbb{R}$ y sea $f : [a, b] \to \mathbb{R}$ una función integrable. Ponemos

$$\int_b^a f \coloneqq -\int_a^b f.$$

Esta definición extiende el signo $\int_a^b$ al caso en que el límite inferior a es mayor que el límite superior b. De todas formas, es importante tener en cuenta que muchos de los resultados que presentamos sobre integrales $\int_a^b f$ dependen de que sea $a \leq b$. Muchos de ellos tienen extensiones más o menos obvias al caso en general en el que posiblemente sea $a > b$, pero es necesario prestar atención. Por ejemplo, si $f : [a, b] \to \mathbb{R}$ es una función integrable y $|f| \leq K$, entonces *no* es cierto en general cuando c y d son elementos cualesquiera de $[a, b]$ valga que

$$\left|\int_c^d f\right| \leq \int_c^d |f| \leq K \cdot (c - d).$$

Lo que sí vale es que

$$\left|\int_c^d f\right| = \left|\int_{\text{mín}\{c,d\}}^{\text{máx}\{c,d\}} f\right| \leq \int_{\text{mín}\{c,d\}}^{\text{máx}\{c,d\}} |f| \leq |\text{máx}\{c, d\} - \text{mín}\{c, d\}| \cdot |f| = |d - c| \cdot K.$$

Más allá de eso, usando esta extensión del significado del signo integral podemos enunciar el siguiente resultado extremadamente útil.

Proposición 13.1.4.10. *Sea $[a, b]$ un intervalo cerrado y acotado de $\mathbb{R}$ y sea $f : [a, b] \to \mathbb{R}$ una función integrable. Si c, d, y e son tres elementos de $[a, b]$ entonces las integrales $\int_c^d f$, $\int_c^e f$, y $\int_c^d f$ tienen sentido y*

$$\int_c^d f = \int_c^e f + \int_c^d f.$$

Demostración. Sean c, d, y e tres elementos de $[a, b]$. Los tres intervalos $[\text{mín}\{c, d\}, \text{máx}\{c, d\}]$, $[\text{mín}\{c, e\}, \text{máx}\{c, e\}]$, y $[\text{mín}\{e, d\}, \text{máx}\{e, d\}]$ están contenidos en $[a, b]$, así que la restricción de f a cualquiera de los tres es integrable. Esto nos dice que las tres integrales que aparecen en la proposición tienen sentido.

Si dos de los tres números c, d y e son iguales, entonces estamos en alguno de los siguientes tres casos.

- $c = d$: es $\int_c^d f = 0 = \int_c^e f + \int_e^c f = \int_c^e f + \int_e^d f$.

- $c = e$: es $\int_c^d f = 0 + \int_c^d f = \int_c^e f + \int_e^d f$.
- $e = d$: es $\int_c^d f = \int_c^d f + 0 = \int_c^c f + \int_e^d f$.

Si, en cambio, esos tres números son distintos dos a dos, entonces estamos en alguno de los siguientes seis casos:

- $c < e < d$: es $\int_c^d f = \int_c^e f + \int_e^d f$.
- $e < c < d$: es $\int_e^d f = \int_e^c f + \int_c^d f$, así que $\int_c^d = -\int_e^c f + \int_c^d f = \int_c^e f + \int_e^d f$.
- $c < d < e$: es $\int_c^e f = \int_c^d f + \int_d^e f$, así que $\int_c^d = \int_c^e f - \int_d^e f = \int_c^e f + \int_e^d f$.
- $e < d < c$: es $\int_e^c f = \int_e^d f + \int_d^c f$, así que $\int_c^d f = -\int_d^c = -\int_e^c f + \int_e^d f = \int_c^e f + \int_e^d f$.
- $d < e < c$: es $\int_d^c f = \int_d^e f + \int_e^c f$, así que $\int_c^d f = -\int_d^c = -\int_d^e f - \int_e^c f = \int_c^e f + \int_e^d f$.
- $d < c < e$: es $\int_d^e f = \int_d^c f + \int_c^e f$, así que $\int_c^d f = -\int_d^c = \int_c^e f - \int_d^e f = \int_c^e f + \int_e^d f$.

En cualquier caso, entonces, vale la igualdad de la proposición. □

Para la próxima aplición del *Criterio de Darboux* que queremos presentar necesitamos la siguiente observación.

Lema 13.1.4.11. *Sea $f : I \to \mathbb{R}$ es una función acotada definida sobre un intervalo cualquiera I, y sea $g : [c, d] \to \mathbb{R}$ es una función definida en un intervalo $[c, d]$ que contiene a la imagen de f. Si g satisface la condición de Lipschitz con constante L, entonces*

$$\sup_{x\in I} g(f(x)) - \inf_{x\in I} g(f(x)) \leq L \cdot \left(\sup_{x\in I} f(x) - \inf_{x\in I} f(x) \right).$$

Demostración. Supongamos que la función g satisface la condición de Lipschitz con constante L, y sea δ un número positivo. Hay elementos x_1 y x_2 de I tales que $g(f(x_1)) > \sup_{x\in I} g(f(x)) - \delta$ y $g(f(x_2)) < \inf_{x\in I} g(f(x)) + \delta$. Como

$$\inf_{x\in I} f(x) \leq f(x_1) \leq \sup_{x\in I} f(x), \qquad -\sup_{x\in I} f(x) \leq -f(x_2) \leq -\inf_{x\in I} f(x),$$

tenemos que

$$\inf_{x\in I} f(x) - \sup_{x\in I} f(x) \leq f(x_1) - f(x_2) \leq \sup_{x\in I} f(x) - \inf_{x\in I} f(x),$$

de manera que

$$|f(x_1) - f(x_2)| \leq \sup_{x\in I} f(x) - \inf_{x\in I} f(x).$$

Se sigue de esto que

$$\sup_{x\in I} g(f(x)) - \inf_{x\in I} g(f(x)) < g(f(x_1)) - g(f(x_2)) + 2\delta$$

$$
\begin{aligned}
&\leq |g(f(x_1)) - g(f(x_2))| + 2\delta \\
&\leq L \cdot |f(x_1) - f(x_2)| + 2\delta \\
&\leq L \cdot \left(\sup_{x \in I} f(x) - \inf_{x \in I} f(x) \right) + 2\delta.
\end{aligned}
$$

Esto es cierto cualquiera sea el número positivo δ, así que vale la desigualdad del lema. □

La siguiente es una aplición directa del lema.

Proposición 13.1.4.12. *Sea $f : [a, b] \to \mathbb{R}$ una función acotada y sea $g : [c, d] \to \mathbb{R}$ una función definida en un intervalo $[c, d]$ que contiene a la imagen de f y que satisface la condición de Lipschitz con constante L. Si f es integrable, entonces $g \circ f$ también.*

Demostración. Supongamos que f es integrable y sea ε un número positivo. De acuerdo al *Criterio de Darboux*, hay una partición $P = (x_0, x_1, \ldots, x_n)$ del intervalo $[a, b]$ tal que $\omega_P(f) < \varepsilon/L$. El Lema 13.1.4.11 nos dice que para cada elemento i de $[\![n]\!]$ es

$$
\sup_{x \in (x_{i-1}, x_i)} g(f(x)) - \inf_{x \in (x_{i-1}, x_i)} g(f(x)) \leq L \cdot \left(\sup_{x \in (x_{i-1}, x_i)} f(x) - \inf_{x \in (x_{i-1}, x_i)} f(x) \right),
$$

y esto implica que

$$
\begin{aligned}
\omega_P(g \circ f) &= \sum_{i=1}^{n} \left(\sup_{x \in (x_{i-1}, x_i)} g(f(x)) - \inf_{x \in (x_{i-1}, x_i)} g(f(x)) \right) (x_i - x_{i-1}) \\
&\leq L \cdot \left(\sum_{i=1}^{n} \left(\sup_{x \in (x_{i-1}, x_i)} f(x) - \inf_{x \in (x_{i-1}, x_i)} f(x) \right) (x_i - x_{i-1}) \right) \\
&= L \cdot \omega_P(f) < L \cdot \frac{\varepsilon}{L} = \varepsilon.
\end{aligned}
$$

El *Criterio de Darboux* nos permite entonces concluir que $g \circ f$ es integrable. □

Veamos algunas consecuencias de esto.

Corolario 13.1.4.13. *Sea $[a, b]$ un intervalo cerrado y acotado de $\mathbb{R}$. Si $f : [a, b] \to \mathbb{R}$ es una función integrable, entonces $|f| : [a, b] \to \mathbb{R}$ también es integrable y*

$$
\left| \int_a^b f \right| \leq \int_a^b |f|.
$$

Demostración. Sea $f : [a,b] \to \mathbb{R}$ una función integrable. La función $g : x \in \mathbb{R} \mapsto |x| \in \mathbb{R}$ satisface la condición de Lipschitz con constante 1, así que de acuerdo a la Proposición 13.1.4.12 la composición $g \circ f$ es integrable. Como esta composición es precisamente la función $|f|$, esto prueba la primera afirmación del corolario. Por otro lado, como $-|f| \leq f \leq |f|$, la Proposición 13.1.3.4 y la linealidad de la integral nos permiten concluir que $-\int_a^b |f| \leq \int_a^b f \leq \int_a^b |f|$, de manera que $|\int_a^b f| \leq \int_a^b |f|$. □

A partir de la integrabilidad del valor absoluto de una función integrable el siguiente resultado sigue inmediatamente.

Corolario 13.1.4.14. *Sea $[a,b]$ un intervalo cerrado y acotado de $\mathbb{R}$. Si $f, g : [a,b] \to \mathbb{R}$ son dos funciones integrables, entonces las funciones* $\max\{f,g\}$ *y* $\min\{f,g\}$ *son también integrables.*

Demostración. Esto es consecuencia del Corolario 13.1.4.13, de la linealidad de la integral, y de las igualdades

$$\max\{f,g\} = \frac{f+g}{2} + \frac{|f-g|}{2}, \qquad \min\{f,g\} = \frac{f+g}{2} - \frac{|f-g|}{2}.$$ □

Más interesante es el siguiente resultado.

Corolario 13.1.4.15. *Sea $[a,b]$ un intervalo cerrado y acotado de $\mathbb{R}$. Si $f, g : [a,b] \to \mathbb{R}$ son dos funciones integrables, entonces la función fg es integrable.*

Demostración. Sean $f, g : [a,b] \to \mathbb{R}$ dos funciones integrables. Como

$$fg = \frac{(f+g)^2 - (f-g)^2}{4},$$

en vista de lo que ya sabemos para probar que fg es integrable es suficiente que mostremos que

si $h : [a,b] \to \mathbb{R}$ es integrable, entonces también lo es h^2.

Sea para ello $h : [a,b] \to \mathbb{R}$ una función integrable, de manera que, en particular, es acotada y hay un número positivo c tal que $-c \leq h(x) \leq c$ para todo $x \in [a,b]$. La función

$$q : x \in [-c,c] \mapsto x^2 \in \mathbb{R}$$

satisface la condición de Lipschitz con constante $2c$. En efecto, es derivable en $(-c,c)$, así que si x y x' son dos elementos de $[-c,c]$ hay, de acuerdo al *Teorema del valor medio*, un elemento ξ de $(-c,c)$ tal que

$$h(x) - h(x') = h'(\xi) \cdot (x - x') = 2\xi(x - x')$$

y, por lo tanto,

$$|h(x) - h(x')| \leq 2|\xi| \cdot (x - x') \leq 2c \cdot |x - x'|.$$

La Proposición 13.1.4.12 nos dice entonces que la función $h^2 = q \circ h$ es integrable. □

13.1.5. Los teoremas fundamentales

El lema de separación

Empecemos con la siguiente observación sencilla pero importante de que podemos reconocer la función nula entre las funciones continuas por su integral.

Proposición 13.1.5.1. *Sea $f : [a, b] \to \mathbb{R}$ una función continua. Si $f \geq 0$, entonces $\int_a^b f \geq 0$ y vale la igualdad si y solamente si $f = 0$.*

Demostración. Supongamos que $f \geq 0$. La Proposición 13.1.3.4 nos dice que $\int_a^b f \geq \int_a^b 0 = 0$, y es claro, por la linealidad de la integral, que si $f = 0$ entonces $\int_a^b f = 0$.

Supongamos ahoras que $f \neq 0$, de manera que hay un elemento x_0 de $[a, b]$ tal que $f(x_0) > 0$. Como f es continua, hay entonces un número positivo ε tal que para todo $x \in (x_0 - \varepsilon, x_0 + \varepsilon) \cap [a, b]$ es $f(x) > f(x_0)/2$. Notemos que la intersección $(x_0 - \varepsilon, x_0 + \varepsilon) \cap (a, b)$ no es vacía, así que podemos suponer que el punto x_0 pertenece a (a, b). Más aún, podemos elegir ε de manera que el intervalo $(x_0 - \varepsilon, x_0 + \varepsilon)$ esté contenido en (a, b). Consideremos la función $g : [a, b] \to \mathbb{R}$ que en cada elemento x de $[a, b]$ toma el valor

$$g(x) = \begin{cases} 0 & \text{si } x \leq x_0 - \varepsilon; \\ \frac{f(x_0)}{2} & \text{si } x_0 - \varepsilon < x < x_0 + \varepsilon; \\ 0 & \text{si } x \geq x_0 + \varepsilon. \end{cases}$$

Esta función es claramente escalonada con respecto a la partición $P \coloneqq (a, x_0 - \varepsilon, x_0 + \varepsilon, b)$ y es $g \leq f$, así que

$$\int_a^b f \geq \int_a^b g = S_P(g) = \varepsilon f(x_0) > 0.$$

Esto prueba que al condición de la proposición es necesaria. □

El teorema del valor medio

En segundo lugar, veamos el llamado *Teorema del valor medio* para integrales.

Proposición 13.1.5.2. *Sea $f : [a, b] \to \mathbb{R}$ una función continua. Hay un elemento c de $[a, b]$ tal que*

$$\int_a^b f = f(c) \cdot (b - a).$$

Demostración. Si $a = b$ la afirmación es evidente, así que podemos suponer que $a < b$. Hay dos elementos x_1 y x_2 de $[a, b]$ tales que $f(x_1) \leq f(x) \leq f(x_2)$ para cada $x \in [a, b]$, así que

$$f(x_1) \cdot (b - a) = \int_a^b f(x_1) \leq \int_a^b f \leq \int_a^b f(x_2) = f(x_2) \cdot (b - a)$$

y, por lo tanto,

$$f(x_1) \leq \frac{1}{b - a} \int_a^b f \leq f(x_2).$$

El *Teorema de los valores intermedios*, entonces, nos dice que hay un número c entre x_1 y x_2 y, en particular, en $[a, b]$, tal que

$$\frac{1}{b - a} \int_a^b f = f(c).$$

Esto prueba la proposición. □

Este resultado tiene la siguiente generalización que es a veces útil.

Proposición 13.1.5.3. *Sean f, $g : [a, b] \to \mathbb{R}$ dos funciones continuas. Si g no cambia de signo en $[a, b]$, entonces hay un elemento c de $[a, b]$ tal que*

$$\int_a^b fg = f(c) \int_a^b g.$$

Demostración. Supongamos que, por ejemplo, g es no negativa en $[a, b]$. Si $\int_a^b g = 0$, entonces la Proposición 13.1.5.1 nos dice que $g = 0$ y la igualdad de la proposición es evidente. Podemos suponer entonces que $\int_a^b g > 0$. Hay dos elementos x_1 y x_2 de $[a, b]$ tales que $f(x_1) \leq f(x) \leq f(x_2)$ para cada $x \in [a, b]$, así que $f(x_1)g(x) \leq f(x)g(x) \leq f(x_2)g(x)$ para cada $x \in [a, b]$ y, por lo tanto,

$$f(x_1) \int_a^b g \leq \int_a^b fg \leq f(x_2) \int_a^b g$$

y, por lo tanto,

$$f(x_1) \leq \frac{\int_a^b fg}{\int_a^b g} \leq f(x_2).$$

El *Teorema de los valores intermedios* implica nos dice que hay un número c entre x_1 y x_2 y, en particular, en $[a, b]$, tal que

$$\frac{\int_a^b fg}{\int_a^b g} = f(c),$$

y esto prueba la proposición. □

El teorema fundamental del cálculo

Veamos ahora el célebre *Teorema fundamental del cálculo*, que establece una conexión entre el cálculo integral y el cálculo diferencial.

Proposición 13.1.5.4. *Sea I un intervalo de $\mathbb{R}$, sea $f : I \to \mathbb{R}$ una función que es integrable en todo intervalo cerrado y acotado contenido en I, sea a un elemento de I, y consideremos la función*

$$F : x \in I \mapsto \int_a^x f,$$

Si x es un elemento del interior de I tal que f es continua en x, entonces F es derivable en x y su derivada allí es $F'(x) = f(x)$.

Recordemos que solo definimos la derivabilidad de una función en los puntos interiores a su dominio: es por ello que aquí solamente consideramos puntos x que son interiores al intervalo I. Por otro lado, como I es un intervalo de $\mathbb{R}$ y $a \in I$, para todo elemento x de I el intervalo cerrado de extremos a y x — que es $[a, x]$ si $a \leq x$ o $[x, a]$ si $x < a$ — está contenido en I, así que la integral que aparece en el enunciado de la proposición tiene sentido.

Demostración. Sea x un elemento del interior de I tal que la función f es continua en x, y sea ε un número positivo. Hay un número positivo δ tal que para todo elemento x' de I vale

$$|x' - x| < \delta \implies |f(x') - f(x)| < \varepsilon.$$

Sea y un elemento de (a, b) tal que $0 < |y - x| < \delta$. Es

$$F(y) - F(x) - f(x)(x - y) = \int_c^y f - \int_c^x f - \int_x^y f(x) = \int_x^y (f - f(x)),$$

así que, como $|f(\xi) - f(x)| < \varepsilon$ para cualquier número ξ entre x y y,

$$\left|\frac{F(y) - F(x)}{y - x} - f(x)\right| \leq \frac{1}{|y - x|}\left|\int_x^y (f - f(x))\right| \leq \frac{1}{|y - x|} \cdot |y - x| \cdot \varepsilon = \varepsilon.$$

Esto nos dice que la función f es derivable en x y que su derivada allí es

$$F'(x) = \lim_{y \to x} \frac{F(y) - F(x)}{y - x} = f(x). \qquad \square$$

Nuestro siguiente resultado es usualmente considerado como una segunda parte del *Teorema fundamental del cálculo* y es la observación clave que nos permite calcular integrales en la práctica. Recordemos primero la siguiente definición.

Definición 13.1.5.5. Una ***primitiva*** de una función $f : (a, b) \to \mathbb{R}$ definida en un intervalo abierto de $\mathbb{R}$ es una función $F : (a, b) \to \mathbb{R}$ que es derivable en todo su dominio y tal que $F'(x) = f(x)$ para todo $x \in (a, b)$.

Una función f que posee una primitiva posee muchas: en efecto, si F es una primitiva de f, entonces para cualquier número c la función $F + c$ también es una primitiva de f. Es importante saber que, de hecho, de esta forma obtenemos *todas* las primitivas de f.

Lema 13.1.5.6. *Sea $f : (a, b) \to \mathbb{R}$ una función y sean $F_1, F_2 : (a, b) \to \mathbb{R}$ dos primitivas de f. La diferencia $F_1 - F_2$ es constante.*

Demostración. Sea c un elemento de (a, b). La función $G \coloneqq F_1 - F_2$ es derivable en (a, b) y su derivada es $G' = F_1' - F_2' = 0$, así que el *Teorema del valor medio* nos dice que para cada elemento x de (a, b) hay un número ξ entre c y x tal que $G(x) - G(c) = G'(x) \cdot (x - c) = 0$. Vemos así que la función G es constante de valor $G(c)$. $\square$

Un corolario del *Teorema fundamental del cálculo* es que podemos contruir primitivas para funciones continuas:

Corolario 13.1.5.7. *Sea $f : (a, b) \to \mathbb{R}$ una función continua y sea c un elemento de (a, b). La función*

$$F : x \in (a, b) \mapsto \int_c^x f \in \mathbb{R}$$

es una primitiva de f.

Demostración. La función f es integrable en cualquier subintervalo de (a,b), porque es continua, así que esto es consecuencia inmediata de la Proposición 13.1.5.4. □

La regla de Barrow

El Corolario 13.1.5.7 tiene como consecuencia inmediata la siguiente «segunda parte» del *Teorema fundamental del cálculo*, que nos dice que podemos calcular la integral de una función si conocemos una cualquiera de sus primitivas.

Proposición 13.1.5.8. *Sea $f : (a,b) \to \mathbb{R}$ una función continua y sea $P : (a,b) \to \mathbb{R}$ una primitiva de f. Para cada elección de c y d en (a,b) es*

$$\int_c^d f = P(d) - P(c).$$

Notemos que la integral que aparece aquí tiene sentido, ya que la función f es integrable en $[\text{mín}\{c,d\}, \text{máx}\{c,d\}]$ porque es continua allí. La igualdad de la proposición es conocida como la *Regla de Barrow*, por Isaac Barrow — que, entre muchas otras cosas, fue profesor de Isaac Newton.

Demostración. Sean c y d dos elementos de (a,b). Como la función f es continua en (a,b), el Corolario 13.1.5.7 nos dice que la función

$$F : x \in (a,b) \mapsto \int_c^x f \in \mathbb{R}$$

es una primitiva de f. Como la función P también es una primitiva de f, el Lema 13.1.5.6 implica que hay un número k tal que $F = P + k$. De hecho, como

$$0 = \int_c^c f = F(c) = P(c) + k,$$

es $k = -P(c)$ y, por lo tanto,

$$\int_c^d f = F(d) = P(d) - P(c),$$

que es lo que afirma la proposición. □

Usando este resultado podemos volver a calcular la integral del Ejemplo 13.1.3.6.

Ejemplo 13.1.5.9. Sea n un entero no negativo. La función $f : x \in \mathbb{R} \mapsto x^n \in \mathbb{R}$ tiene a la función

$$F : x \in \mathbb{R} \mapsto \frac{x^{n+1}}{n+1} \in \mathbb{R}$$

como primitiva, así que, de acuerdo a la Proposición 13.1.5.8 es

$$\int_0^1 f = F(1) - F(0) = \frac{1}{n+1}.$$

La regla de integración por sustitución

Una vez que tenemos a nuestra disposición el *Teorema fundamental del cálculo* es fácil establecer los resultados básicos que nos permiten calcular integrales en la práctica. Empecemos por la llamada *Regla de integración por sustitución*, que no es otra cosa que una manifestación en el cálculo integral de la *Regla de la cadena* del cálculo diferencial.

Proposición 13.1.5.10. *Sea $P : (a,b) \to \mathbb{R}$ una función derivable, sea $f \coloneqq P' : (a,b) \to \mathbb{R}$ su derivada, y sea $g : (c,d) \to \mathbb{R}$ una función derivable cuya imagen está contenida en (a,b). La función $P \circ g : (c,d) \to \mathbb{R}$ es una primitiva de la función*

$$x \in (c,d) \mapsto f(g(x)) \cdot g'(t) \to \mathbb{R}$$

y, por lo tanto,

$$\int_u^v f(g(x)) \cdot g'(x)\,\mathrm{d}x = P(g(v)) - P(g(u)) = \int_{g(u)}^{g(v)} f(x)\,\mathrm{d}x$$

para cualquier elección de u y b en (c,d).

Demostración. De acuerdo a la *Regla de la cadena* la función $Q \coloneqq P \circ g : (c,d) \to \mathbb{R}$ es derivable y para cada elemento x de (c,d) es

$$Q'(x) = P'(g(x)) \cdot g'(x) = f(g(x)) \cdot g'(x).$$

Esto nos dice que esa función Q es una primitiva para la función

$$h : x \in (c,d) \mapsto f(g(x)) \cdot g'(x) \in \mathbb{R}$$

y, de acuerdo a la *Regla de Barrow* entonces, es

$$\int_u^v f(g(x)) \cdot g'(x)\,\mathrm{d}x = \int_u^v h = Q(v) - Q(u) = P(g(v)) - P(g(u))$$

cualesquiera sean u y v en (c,d). Por otro lado, como la función P es una primitiva para la función f, la *Regla de Barrow* nos dice que también

$$\int_{g(u)}^{g(v)} f(x)\,\mathrm{d}x = P(g(v)) - P(g(u)).$$ □

Veamos un ejemplo de cómo podemos usar esta proposición para calcular integrales.

Ejemplo 13.1.5.11. Sea t un número positivo. Queremos calcular la integral

$$\int_0^t \frac{x}{\sqrt{1+x^2}}\,\mathrm{d}x.$$

Notemos que esto tiene sentido, ya que la función

$$x \in \mathbb{R} \mapsto \frac{x}{\sqrt{1+x^2}} \in \mathbb{R}$$

es continua. Consideremos las funciones

$$f : x \in (0,+\infty) \mapsto \frac{1}{2\sqrt{x}} \in \mathbb{R}, \qquad g : x \in \mathbb{R} \mapsto 1+x^2 \in \mathbb{R}.$$

La función g es derivable y su imagen está contenida en $(0,+\infty)$, y para cada $x \in \mathbb{R}$ es

$$\frac{x}{\sqrt{1+x^2}} = f(g(x)) \cdot g'(x).$$

Por otro lado, la función $P : x \in (0+\infty) \mapsto \sqrt{x} \in \mathbb{R}$ es una primitiva para f. De acuerdo a la Proposición 13.1.5.10, entonces, tenemos que

$$\int_0^t \frac{x}{\sqrt{1+x^2}}\,\mathrm{d}x = P(g(t)) - P(g(0)) = \sqrt{1+t^2} - 1.$$

La regla de integración por partes

En segundo lugar, veamos la *Regla de integración por partes*, que se deduce inmediatamente de la *Regla del producto* del cálculo diferencial.

Proposición 13.1.5.12. *Sean $u, v : (a,b) \to \mathbb{R}$ dos funciones derivables. Para cada elección de c y d en (a,b) es*

$$\int_c^d u'(x) \cdot v(x)\,\mathrm{d}x = u(d) \cdot v(d) - u(c) \cdot v(c) - \int_c^d u(x) \cdot v'(x)\,\mathrm{d}x.$$

Demostración. La *Regla del producto* nos dice que la función $w : x \in (a,b) \mapsto u(x) \cdot v(x) \in \mathbb{R}$ es derivable y que para cada elemento x de (a,b) es

$$w'(x) = u'(x) \cdot v(x) + u(x) \cdot v'(x).$$

Esto nos dice que la función w es una primitiva para la función

$$f : x \in (a,b) \mapsto u'(x) \cdot v(x) + u(x) \cdot v'(x) \in \mathbb{R}.$$

Esto implica, de acuerdo a la *Regla de Barrow*, para cada elección de c y d en (a, b) es

$$u(d)\cdot v(d) - u(c)\cdot v(c) = f(d) - f(c) = \int_c^d w = \int_c^d u'(x)\cdot v(x), \mathrm{d}x + \int_c^d u(x)\cdot v'(x)\,\mathrm{d}x.$$

La igualdad del enunciado es consecuencia inmediata de esta. □

Este resultado es extremadamente útil a la hora de calcular integrales.

Ejemplo 13.1.5.13. Sean a y b dos números reales. Queremos calcular la intergral

$$\int_a^b x\cos x\,\mathrm{d}x.$$

Las funciones $u : x \in \mathbb{R} \mapsto \operatorname{sen} x \in \mathbb{R}$ y $v : x \in \mathbb{R} \mapsto x \in \mathbb{R}$ son derivables, y la Proposición 13.1.5.12 nos dice que

$$\begin{aligned}\int_a^b x\cos x\,\mathrm{d}x &= \int_a^b u'(x)\cdot v(x)\,\mathrm{d}x \\ &= u(b)\cdot v(b) - u(a)\cdot v(a) - \int_a^b u(x)\cdot v'(x)\,\mathrm{d}x \\ &= b\operatorname{sen} b - a\operatorname{sen} a - \int_a^b \operatorname{sen} x\,\mathrm{d}x,\end{aligned}$$

ya que $v'(x) = 1$ para todo $x \in \mathbb{R}$, y como la función $-\cos$ es una primitiva de la función sen, la *Regla de Barrow* nos dice que esto es

$$= b\operatorname{sen} b - a\operatorname{sen} a + \cos b - \cos a.$$

Notemos que esto implica que la función

$$x \in \mathbb{R} \mapsto x\operatorname{sen} x + \cos x \in \mathbb{R}$$

es una primitiva para la función $x \in \mathbb{R} \mapsto x\cos x \in \mathbb{R}$.

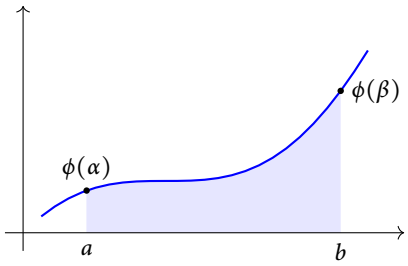
La *Regla de integración por partes* tiene una interpretación geométrica. Supongamos que $f, g : \mathbb{R} \to \mathbb{R}$ son dos funciones derivables tales que $f'(x) > 0$ y $g'(x) > 0$ para todo $x \in \mathbb{R}$, y consideremos la función

$$\phi : t \in \mathbb{R} \mapsto (f(t), g(t)) \in \mathbb{R}^2.$$

Sea $[\alpha, \beta]$ un intervalo cerrado y acotado de $\mathbb{R}$ y sean $a \coloneqq f(\alpha)$ y $b \coloneqq f(\beta)$. Como la función f es estrictamente creciente, podemos restringirla a una función $[\alpha, \beta] \to [a, b]$ y esta restricción es una biyección, así que podemos considerar su función inversa $u \coloneqq f^{-1} : [a, b] \to [\alpha, \beta]$. Esta función inversa es derivable y $u'(x) = 1/f'(u(x))$ para cada $x \in [a, b]$. Además, como para cada

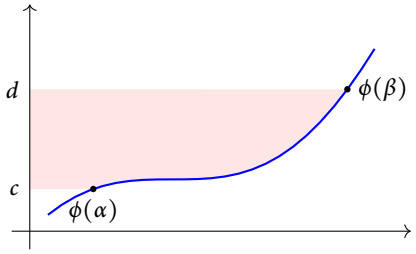
$x \in [a,b]$ es $\phi(u(x)) = (f(u(x), g(u(x)) = (x, g(u(x))$ y entonces el area de la región entre el eje x y el segmento de la curva ϕ entre los puntos $\phi(\alpha)$ y $\phi(\beta)$ es

$$\int_a^b g(u(x))\,\mathrm{d}x.$$



De manera similar, si ponemos $c \coloneqq g(\alpha)$ y $d \coloneqq g(\beta)$, entonces la función g se restringe a una función $[\alpha,\beta] \to [c,d]$ y esta restricción es una biyección, así que podemos considerar su función inversa $v \coloneqq g^{-1} : [c,d] \to [\alpha,\beta]$. Esta función inversa es derivable y $v'(y) = 1/g'(v(y))$ para cada $y \in [c,d]$, y, como para cada $y \in [c,d]$ es $\phi(v(y)) = (f(v(y), g(v(y)) = (f(v(y), y)$. el area de la región entre el eje y y el segmento de la curva ϕ entre los puntos $\phi(\alpha)$ y $\phi(\beta)$ es

$$\int_c^d f(v(y))\,\mathrm{d}y.$$



De acuerdo a la Proposición 13.1.5.10, tenemos que

$$\begin{aligned}\int_\alpha^\beta f'(t)\cdot g(t)\,\mathrm{d}t &= \int_{u(a)}^{u(b)} f'(t)\cdot g(t)\,\mathrm{d}t = \int_a^b f'(u(x))\cdot g(u(x))\cdot u'(x)\,\mathrm{d}x \\ &= \int_a^b f'(u(x))\cdot g(u(x))\cdot \frac{1}{f'(u(x))}\,\mathrm{d}x = \int_a^b g(u(x))\,\mathrm{d}x\end{aligned}$$

y

$$\begin{aligned}\int_\alpha^\beta f(t)\cdot g'(t)\,\mathrm{d}t &= \int_{v(c)}^{v(d)} f(t)\cdot g'(t)\,\mathrm{d}t = \int_c^d f(v(y))\cdot g(v(y))\cdot v'(y)\,\mathrm{d}y \\ &= \int_c^d f(v(y))\cdot g(v(y))\cdot \frac{1}{g'(v(y))}\,\mathrm{d}y = \int_c^d f(v(y))\,\mathrm{d}y,\end{aligned}$$

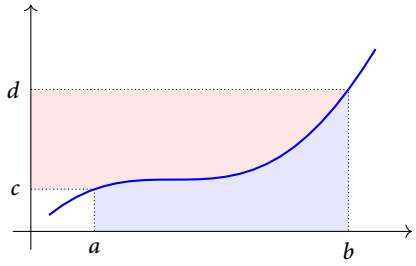
y la *Regla de integración por partes* nos dice que

$$\int_\alpha^\beta f'(t)\cdot g(t)\,\mathrm{d}t + \int_\alpha^\beta f(t)\cdot g'(t)\,\mathrm{d}t = f(\beta)\cdot g(\beta) - f(\alpha)\cdot g(\alpha).$$

Juntado todo podemos concluir que

$$\int_a^b g(u(x))\,\mathrm{d}x + \int_c^d f(v(y))\,\mathrm{d}y = bd - ac.$$

Vemos así que la *Regla de integración por partes* nos dice que la suma de las áreas de la región azul y de la región roja en el diagrama



es la diferencia entre el area del rectángulo $[0, b] \times [0, d]$ y el rectángulo $[0, a] \times [0, c]$. Notemos que este afirmación es intuitivamente evidente!

Para terminar, presentemos una aplicación de la fórmula de integración por partes, el llamado *Lemma de Riemann–Lebesgue* — o, para ser exactos, un caso especial de este.

Proposición 13.1.5.14. *Si $f : \mathbb{R} \to \mathbb{R}$ es una función periódica de periodo 2π que es derivable y que tiene derivada continua, entonces*

$$\lim_{n\to\infty} \int_0^{2\pi} f(x) \cos nx \, \mathrm{d}x = 0.$$

Notemos que el integrando de la intergral que aparece en el límite es continuo, así que es integrable en $[0, 2\pi]$ y que, por lo tanto, la afirmación de la proposición tiene sentido.

Demostración. Como la derivada f' es continua, hay un número positivo K tal que $|f'(x| \leq K$ para todo $x \in [0, 2\pi]$. Para cada entero positivo n es $\sin 0 = \sin 2\pi n = 0$, así que la *Fórmula de integración por partes* nos dice que

$$\int_0^{2\pi} f'(x) \sin nx \, \mathrm{d}x = -\int_0^{2\pi} f(x) n \cos nx \, \mathrm{d}x,$$

y, por lo tanto, es

$$\left| \int_0^{2\pi} f(x) \cos nx \, \mathrm{d}x \right| = \left| \frac{1}{n} \int_0^{2\pi} f'(x) \sin nx \, \mathrm{d}x \right| \leq \frac{1}{n} 2\pi K,$$

ya que $|f'(t) \sin nx| \leq K$ para todo $x \in [0, 2\pi]$. La afirmación de la proposición sigue inmediatamente de esto. □

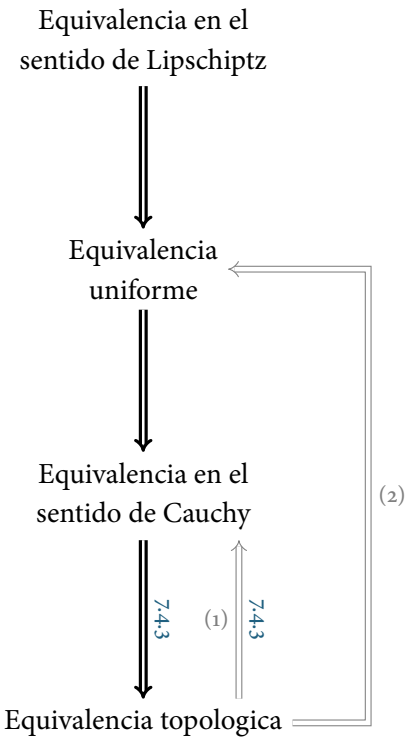


Figura 13.3. Las relaciones entre las distintas nociones de equivalencia. La implicación (1) vale si las dos métricas son completas; la implicación (2) si las dos métricas son compactas.

Bibliografía

[Ack1928] Wilhelm Ackermann, *Zum Hilbertschen Aufbau der reellen Zahlen.*, Math. Ann. **99** (1928), 118–133.

[Arz1885] Cesare Arzelà, *Un teorema intorno alle serie di funzioni*, Rom. Acc. L. Rend. (4) **1** (1885), 262–267 (French).

[Asc1884] Giulio Ascoli, *Le curve limite di una varietà data di curve*, Rom. Acc. L. Mem. (3) **18** (1884), 521–586 (Italian).

[AZ1999] Martin Aigner and Günter M. Ziegler, *Proofs from the book*, Springer-Verlag, Berlin, 1999. Including illustrations by Karl H. Hofmann, Corrected reprint of the 1998 original. MR1723092

[Bai1899] René-Louis Baire, *On functions of real variables.*, Annali di Mat. (3) **3** (1899), 1–123 (French).

[Ban1922a] Stefan Banach, *Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales*, Fund. Math. **3** (1922), 133–181. MR3949898

[Ban1922b] Stefan Banach, *Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales*, Fund. Math. **3** (1922), 133–181. MR3949898

[Ban1929a] Stefan Banach, *Sur les fonctionelles linéaires. I.*, Stud. Math. **1** (1929), 211–216 (French).

[Ban1929b] Stefan Banach, *Sur les fonctionelles linéaires. II.*, Stud. Math. **1** (1929), 223–239 (French).

[Ber1845] Joseph Bertrand, *Mémoire sur le nombre de valeurs que peut prendre une fonction quand on y permute les lettres qu'elle renferme*, Journal de l'École Royale Polytechnique **18** (1845), 123–140.

[Ber1905] Felix Bernstein, *Untersuchungen aus der Mengenlehre*, Math. Ann. **61** (1905), 117–155.

[Ber1963] Claude Berge, *Topological spaces*, Vol. XIII, Edinburgh-London: Oliver & Boyd, Ltd., 1963.

[Ber1997] George Bergman, *The Axiom of Choice, Zorn's Lemma, and all that*, Math 250B, 1997.

[Bou1949] Nicolas Bourbaki, *Sur le théorème de Zorn*, Arch. Math. (Basel) **2** (1949/50), 434–437 (1951). MR47739

[Bou2007a] Nicolas Bourbaki, *Éléments de mathématique. Topologie générale. Chapitres 1 à 4*, Reprint of the 1971 original, Berlin: Springer, 2007.

[Bou2007b] Nicolas Bourbaki, *Éléments de mathématique. Topologie générale. Chapitres 5 à 10*, Reprint of the 1974 original, Berlin: Springer, 2007 (French).

[Bro1912] Luitzen Egbertus Jan Brouwer, *Beweis der Invarianz des n-dimensionalen Gebiets.*, Math. Ann. **71** (1912), 305–313.

[Bro1970] Arlen Brown, *A remark on set theory*, Proc. Amer. Math. Soc. **25** (1970), 920–921. MR266749

[Bro2010] Kenneth Brown, *Zorn's lemma*, Mathematics 6310, Cornell University, 2010.
Available at https://pi.math.cornell.edu/~kbrown/6310/zorn.pdf

[Bum1965] Richard T. Bumby, *Modules which are isomorphic to submodules of each other*, Arch. Math. **16** (1965), 184–185.

[Cam1999] Peter J. Cameron, *Sets, logic and categories*, Springer Undergraduate Mathematics Series, Springer-Verlag London, Ltd., London, 1999. MR1700271.
Available at https://doi.org/10.1007/978-1-4471-0589-3

[Can1874] Georg Cantor, *Ueber eine Eigenschaft des Inbegriffs aller reellen algebraischen Zahlen*, J. Reine Angew. Math. **77** (1874), 258–262. MR1579605

[Can1877] Georg Cantor, *Ein Beitrag zur Mannigfaltigkeitslehre*, J. Reine Angew. Math. **84** (1877), 242–259.

[Can1882] Georg Cantor, *Ueber unendliche lineare Punktmannichfaltigkeiten*, Math. Ann. **20** (1882), 113–122.

[Can1887] Georg Cantor, *Mitteilungen zur Lehre vom Transfiniten. I. II*, Zeitschrift für Philosophie und philosophische Kritik **91, 92** (1887), 1: 81–125, 2: 240–265.

[Can1892] Georg Cantor, *Ueber eine elementare Frage der Mannigfaltigkeitslehre.*, Jahresber. Dtsch. Math.-Ver. **1** (1892), 75–78.

[Can1895] Georg Cantor, *Beitrage zur Begründung der transfiniten Mengenlehre. Art. I*, Math. Ann. **46** (1895), 481–512.

[Car1937] Henri Cartan, *Théorie des filtres*, Comptes rendus hebdomadaires des séances de l'Académie des sciences **205** (1937), 595–598.

[Cas1957] John William Scott Cassels, *An introduction to Diophantine approximation*, Cambridge Tracts in Mathematics and Mathematical Physics, vol. No. 45, Cambridge University Press, New York, 1957. MR87708

[CG1996] John H. Conway and Richard K. Guy, *The book of numbers*, Copernicus, New York, 1996. MR1411676.
Available at https://doi.org/10.1007/978-1-4612-4072-3

[CL1955] Earl A. Coddington and Norman Levinson, *Theory of ordinary differential equations*, McGraw-Hill Book Co., Inc., New York-Toronto-London, 1955. MR69338

[Coh1963] Paul Cohen, *The independence of the continuum hypothesis*, Proc. Nat. Acad. Sci. U.S.A. **50** (1963), 1143–1148. MR157890

[Dev1967] Vladimir Devidé, *A proof of Zermelo's theorem*, J. Symb. Log. **32** (1967), 366.

[Die1944] Jean Dieudonné, *Une généralisation des espaces compacts*, J. Math. Pures Appl. (9) **23** (1944), 65–76. MR13297

[Dug1966] James Dugundji, *Topology*, Allyn and Bacon, Inc., Boston, MA, 1966. MR193606

[Dys1947] Freeman J. Dyson, *The approximation to algebraic numbers by rationals*, Acta Math. **79** (1947), 225–240. MR23854

[EK1990] Paul Erdős and Péter Komjáth, *Countable decompositions of* $\mathbf{R}^2$ *and* $\mathbf{R}^3$, Discrete Comput. Geom. **5** (1990), no. 4, 325–331. MR1043714

[Erd1932] Pál Erdős, *Beweis eines Satzes von Tschebyschef.*, Acta Litt. Sci. Szeged **5** (1932), 194–198.

[Ewa1996] William Ewald (ed.), *From Kant to Hilbert: a source book in the foundations of mathematics. Vol. I, II*, Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1996. MR1465678

[Fer2010] Andrea Ferretti, *Slick proof?: A vector space has the same dimension as its dual if and only if it is finite dimensional*, MathOverflow, 2010. URL:https://mathoverflow.net/q/13372 (version: 2025-07-01).
Available at https://mathoverflow.net/q/13372

[FK1934] Grigoriĭ Mikhaĭlovich Fichtenholz and Leonid Vitalʹevich Kantorovitch, *Sur les opérations linéaires dans l'espace des fonctions bornees*, Stud. Math. **5** (1934), 69–98.

[Fré1906] René Maurice Fréchet, *Sur quelques points du calcul fonctionnel*, Rend. Circ. Mat. Palermo **22** (1906), 1–74 (French).

[Fri1952] Orrin Frink, *A proof of the maximal chain theorem*, Amer. J. Math. **74** (1952), 676–678. MR49866

[Göd1940] Kurt Gödel, *The Consistency of the Continuum Hypothesis*, Annals of Mathematics Studies, vol. No. 3, Princeton University Press, Princeton, NJ, 1940. MR2514

[Gra2007] Daniel R. Grayson, *Zorn's lemma*, University of Illinois at Urbana-Champaign, 2007. Available at `https://pages.uoregon.edu/apolish/Zorn.pdf`

[Gro1981] Mikhael Gromov, *Groups of polynomial growth and expanding maps*, Inst. Hautes Études Sci. Publ. Math. **53** (1981), 53–73. MR623534

[Gro1987] Mikhail Gromov, *Hyperbolic groups*, Essays in group theory, Publ., Math. Sci. Res. Inst. 8, 75-263 (1987)., 1987 (English).

[Hah1922] Hans Hahn, *über Folgen linearer Operationen*, Monatsh. Math. Phys. **32** (1922), no. 1, 3–88. MR1549169

[Hah1927] Hans Hahn, *über lineare Gleichungssysteme in linearen Räumen*, J. Reine Angew. Math. **157** (1927), 214–229. MR1581120

[Hal1960] Paul R. Halmos, *Naive set theory*, The University Series in Undergraduate Mathematics, D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto-London-New York, 1960. MR114756

[Ham1950] Richard Wesley Hamming, *Error detecting and error correcting codes*, Bell System Tech. J. **29** (1950), 147–160. MR35935

[Hat2002] Allen Hatcher, *Algebraic topology*, Cambridge University Press, Cambridge, 2002. MR1867354. Available at `http://www.math.cornell.edu/Ëœhatcher`

[Hau1914] Felix Hausdorff, *Grundzüge der Mengenlehre.*, 1914 (German).

[Hau1932] Felix Hausdorff, *Zur Theorie der linearen metrischen Räume*, J. Reine Angew. Math. **167** (1932), 294–311. MR1581343

[Hau1936] Felix Hausdorff, *Über zwei Sätze von G. Fichtenholz und L. Kantorovitch.*, Stud. Math. **6** (1936), 18–19.

[Her1873] Charles Hermite, *On the irrationality of the base of hyperbolic logarithms*, Rep. Britt. Ass. Messenger (2) III (1873), 98–100.

[Her2006] Horst Herrlich, *Axiom of choice*, Lecture Notes in Mathematics, vol. 1876, Springer-Verlag, Berlin, 2006. MR2243715

[Hil1890] David Hilbert, *Ueber die Theorie der algebraischen Formen*, Math. Ann. **36** (1890), no. 4, 473–534. MR1510634

[Hil1900] David Hilbert, *Mathematische Probleme. Vortrag, gehalten auf dem internationalen Mathematiker-Congress zu Paris 1900.*, Nachr. Ges. Wiss. Göttingen, Math.-Phys. Kl. **1900** (1900), 253–297.

[Hil1902] David Hilbert, *Mathematical problems*, Bull. Amer. Math. Soc. **8** (1902), no. 10, 437–479. MR1557926

[Hil1923] Theophil Henry Hildebrandt, *On uniform limitedness of sets of functional operations*, Bull. Amer. Math. Soc. **29** (1923), no. 7, 309–315. MR1560736

[Hil1925] David Hilbert, *Über das Unendliche.*, Math. Ann. **95** (1925), 161–190.

[Hin2013] Arie Hinkis, *Proofs of the Cantor-Bernstein theorem. A mathematical excursion*, Sci. Networks, Hist. Stud., vol. 45, Basel: Birkhäuser/Springer, 2013.

[Hur1891] Adolf Hurwitz, *On the approximation of irrational numbers by rational numbers*, Math. Ann. **39** (1891), 279–284 (German).

[IT2024] Guillermo L. Incatasciato and Pedro Sánchez Terraf, *Chain bounding, the leanest proof of zorn's lemma, and an illustration of computerized proof formalization*, 2024.
Available at https://arxiv.org/abs/2404.11638

[Jac1953] Nathan Jacobson, *Lectures in abstract algebra. Vol. II. Linear algebra*, D. Van Nostrand Co., Inc., Toronto-New York-London, 1953. MR53905

[Jai2012] Arjun Jain, *Zorn's lemma — an elementary proof under the axiom of choice*, 2012.
Available at https://arxiv.org/abs/1207.6698

[Jam1951] Robert C. James, *A non-reflexive Banach space isometric with its second conjugate space*, Proc. Nat. Acad. Sci. U.S.A. 37 (1951), 174–177. MR44024

[Jec1973] Thomas J. Jech, *The axiom of choice*, Studies in Logic and the Foundations of Mathematics, vol. Vol. 75, North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., Inc., New York, 1973. MR396271

[Jec2003] Thomas J. Jech, *Set theory*, millennium, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2003. MR1940513

[Kap1954] Irving Kaplansky, *Infinite abelian groups*, (University of Michigan Publications in Mathematics. No. 2.) Ann Arbor: University of Michigan Press V, 91 p. (1954)., 1954.

[Kec1995] Alexander S. Kechris, *Classical descriptive set theory*, Graduate Texts in Mathematics, vol. 156, Springer-Verlag, New York, 1995. MR1321597.
Available at https://doi.org/10.1007/978-1-4612-4190-4

[Kel1950] John L. Kelley, *Convergence in topology*, Duke Math. J. 17 (1950), 277–283. MR36982

[Kna1928] Bronisław Knaster, *Un théorème sur les fonctions d'ensembles*, Ann. Soc. Polon. Math. **6** (1928), 133–134.

[Kne1950] Hellmuth Kneser, *Eine direkte Ableitung des Zornschen Lemmas aus dem Auswahlaxiom*, Math. Z. **53** (1950), 110–113. MR37887

[Kom2001] Péter Komjáth, *Three clouds may cover the plane*, Annals of Pure and Applied Logic **109** (2001), no. 1, 71–75. Special Issue dedicated to Petr Vopenka.

[Kor1911] Alwin Reinhold Korselt, *Über einen Beweis des Äquivalenzsatzes*, Math. Ann. 70 (1911), no. 2, 294–296.

[Kur1922] Kazimierz Kuratowski, *Une méthode d'élimination des nombres transfinis des raisonnements mathématiques.*, Fundam. Math. 3 (1922), 76–108.

[Kut1998] Kenneth Kuttler, *Modern analysis*, Boca Raton, MA: CRC Press, 1998.

[Lam1768] Johann Heinrich Lambert, *Mémoire sur quelques propriétés remarquables des quantités transcendantes, circulaires et logarithmiques*, Mémoires de l'Académie Royale des Sciences de Berlin (1768), 265–322.

[Lan1966] Serge Lang, *Introduction to transcendental numbers*, Addison-Wesley Series in Mathematics, vol. VI, Addison-Wesley Publishing Company, Reading, Mass., 1966.

[Lan1993] Serge Lang, *Real and functional analysis*, Third, Graduate Texts in Mathematics, vol. 142, Springer-Verlag, New York, 1993. MR1216137.
Available at https://doi.org/10.1007/978-1-4612-0897-6

[Lan2002] Serge Lang, *Algebra*, third, Graduate Texts in Mathematics, vol. 211, Springer-Verlag, New York, 2002. MR1878556.
Available at https://doi.org/10.1007/978-1-4613-0041-0

[Lek2010] Lektor Lektorean, *Zorn's lemma*, 2010.
Available at https://people.math.aau.dk/~cornean/index.html/zornlemma.pdf

[Lew1991] Jonathan Lewin, *A simple proof of Zorn's lemma*, Amer. Math. Monthly **98** (1991), no. 4, 353–354. MR1103192

[Lio1851] Joseph Liouville, *Sur des classes très étendues de quantités dont la valeur n'est ni algébrique, ni même réductible à des irrationnelles algébriques*, J. Math. Pures Appl. **16** (1851), 133–142.

[Mil1978] John Milnor, *Analytic proofs of the "hairy ball theorem" and the Brouwer fixed-point theorem*, Amer. Math. Monthly **85** (1978), no. 7, 521–524. MR505523

[MS1922] Eliakim Hastings Moore and Herman L. Smith, *A General Theory of Limits*, Amer. J. Math. **44** (1922), no. 2, 102–121. MR1506463

[Mun2000] James R. Munkres, *Topology.*, 2nd ed., Upper Saddle River, NJ: Prentice Hall, 2000 (English).

[Nag1985] Jun-iti Nagata, *Modern general topology*, Second, North-Holland Mathematical Library, vol. 33, North-Holland Publishing Co., Amsterdam, 1985. MR831659

[Nui2024] Koji Nuida, *A Simple and Elementary Proof of Zorn's Lemma*, Discrete Mathematics Letters **13** (2024), 108–110.

[Ost1916] Alexander Ostrowski, *über einige Lösungen der Funktionalgleichung* $\psi(x) \cdot \psi(x) = \psi(xy)$, Acta Math. **41** (1916), no. 1, 271–284. MR1555153

[Pea1890] Giuseppe Peano, *Sur une courbe, qui remplit toute une aire plane*, Math. Ann. **36** (1890), 157–160.

[Pét1935] Rózsa Péter, *Konstruktion nichtrekursiver Funktionen.*, Math. Ann. **111** (1935), 42–60.

[Ram1919] Srinivasa Ramanujan, *A proof of Bertrand's postulate*, Journal of the Indian Mathematical Society **11** (1919), 181–182.

[Rot1955] Klaus Friedrich Roth, *Rational approximations to algebraic numbers*, Mathematika **2** (1955), 1–20; corrigendum, 168. MR72182

[RR1985] Herman Rubin and Jean E. Rubin, *Equivalents of the axiom of choice. II*, Studies in Logic and the Foundations of Mathematics, vol. 116, North-Holland Publishing Co., Amsterdam, 1985. MR798475

[Rud1969] Mary Ellen Rudin, *A new proof that metric spaces are paracompact*, Proc. Amer. Math. Soc. **20** (1969), 603. MR236876

[Rud1987] Walter Rudin, *Real and complex analysis*, McGraw-Hill Book Co., New York, 1987. MR924157

[Sch1898] Ernst Schröder, *Ueber zwei Definitionen der Endlichkeit und G. Cantor'sche Sätze*, Nova Acta Leop.-Carol. Akad. **71** (1898), 303–362.

[Shu2021] Jerry Shurman, *The Axiom of Choice, Zorn's Lemma, and the Well Ordering Principle*, 2021.
Available at `https://people.reed.edu/~jerry/332/23zorn.pdf`

[Sie1921] Carl Siegel, *Approximation algebraischer Zahlen*, Math. Z. **10** (1921), no. 3-4, 173–213. MR1544471

[Sie1934] Wacław Sierpiński, *Hypothèse du continu*, Monogr. Mat., Warszawa, vol. 4, PWN - Panstwowe Wydawnictwo Naukowe, Warszawa, 1934.
Available at `https://eudml.org/doc/219323`

[Sok2011] Alan D. Sokal, *A really simple elementary proof of the uniform boundedness theorem*, Amer. Math. Monthly **118** (2011), no. 5, 450–452. MR2805031

[Spr2017] Nico Spronk, *Axiom of choice, zorn's lemma and the well-ordering principle*, University of Waterloo, 2017.
Available at `https://www.math.uwaterloo.ca/~nspronk/math351/AofC.pdf`

[Sri1998] Shashi Mohan Srivastava, *A course on Borel sets*, Graduate Texts in Mathematics, vol. 180, Springer-Verlag, New York, 1998. MR1619545.
Available at `https://doi.org/10.1007/978-3-642-85473-6`

[Sta2015] Richard P. Stanley, *Catalan numbers*, Cambridge University Press, New York, 2015. MR3467982.
Available at `https://doi.org/10.1017/CBO9781139871495`

[Ste1963] Hugo Steinhaus, *Stefan Banach*, Studia Math. (Ser. Specjalna) Zeszyt **1** (1963), 7–15. MR175746

[Sto1937] Marshall H. Stone, *Applications of the theory of Boolean rings to general topology*, Trans. Amer. Math. Soc. **41** (1937), no. 3, 375–481. MR1501905

[Sto1948] Arthur Harold Stone, *Paracompactness and product spaces*, Bull. Amer. Math. Soc. **54** (1948), 977–982. MR26802

[Sze1950] Tibor Szele, *On Zorn's lemma*, Publ. Math. Debrecen **1** (1950), 254–256, erratum 257. MR39027

[Tar1955] Alfred Tarski, *A lattice-theoretical fixpoint theorem and its applications*, Pacific J. Math. **5** (1955), 285–309. MR74376

[Thu1909] Axel Thue, *über Annäherungswerte algebraischer Zahlen*, J. Reine Angew. Math. **135** (1909), 284–305. MR1580770

[Tie1914] Heinrich Tietze, *Über Funktionen, die auf einer abgeschlossenen Menge stetig sind.*, J. Reine Angew. Math. **145** (1914), 9–14 (German).

[Tyc1935] Andrey Nikolayevich Tychonoff, *über einen Funktionenraum*, Math. Ann. **111** (1935), no. 1, 762–766. MR1513030

[Ury1925] Paul Urysohn, *über die Mächtigkeit der zusammenhängenden Mengen*, Math. Ann. **94** (1925), no. 1, 262–295. MR1512258

[vL1882] Ferdinand von Lindemann, *Ueber die Zahl π*, Math. Ann. **20** (1882), no. 2, 213–225. MR1510165

[vN1923] John von Neumann, *Zur Einführung der transfiniten Zahlen.*, Acta Litt. Sci. Szeged **1** (1923), 199–208.

[Wes1957] Jeffrey Dennis Weston, *A short proof of Zorn's lemma*, Arch. Math. (Basel) **8** (1957), 279. MR92743

[Wit1951] Ernst Witt, *Beweisstudien zum Satz von M. Zorn*, Math. Nachr. **4** (1951), 434–438. MR39776

[ZD2019] Mohssin Zarouali-Darkaoui, *On the bourbaki's fixed point theorem and the axiom of choice*, arXiv, 2019. Available at `https://arxiv.org/abs/1905.09782`

[Zer1904] Ernst Zermelo, *Beweis, daß jede Menge wohlgeordnet werden kann.*, Math. Ann. **59** (1904), 514–516.

[Zer1908a] Ernst Zermelo, *Neuer Beweis für die Möglichkeit einer Wohlordnung.*, Math. Ann. **65** (1908), 107–128.

[Zer1908b] Ernst Zermelo, *Untersuchungen über die Grundlagen der Mengenlehre. I*, Math. Ann. **65** (1908), 261–281.

[Zor1935] Max Zorn, *A remark on method in transfinite algebra*, Bull. Amer. Math. Soc. **41** (1935), no. 10, 667–670. MR1563165

Notaciones

$|X|$ 32
$X \preceq Y$ 32
$X \prec Y$ 32
$[\![n]\!]$ 39
$\aleph_0$ 58
$\mathfrak{c}$ 59
$\mathfrak{a} + \mathfrak{b}$ 71
$\mathfrak{a} \cdot \mathfrak{b}$ 71
$\mathfrak{a}^{\mathfrak{b}}$ 71
Ord 99
$\alpha + 1$ 107
$\|-\|_\infty$ 124, 133
$\|-\|_p$ 126, 135
$\ell^\infty_{\mathbb{k}}$ 133
c 134
c_0 134
x_{00} 134
$\ell^p_{\mathbb{k}}$ 135
$\ell^\infty_{\mathbb{k}}(X \mod \mathscr{F})$ 144
S° 182
$\operatorname{int}(S)$ 182
$\operatorname{int}_X(S)$ 182
τ_X 190
$d \sim d'$ 200
$d \sim_{\mathrm{Lip}} d'$ 204
$\overline{B}_r(x)$ 211
S' 226
$\operatorname{cbr}(S)$ 231
$\lim_{n\to\infty} x_n$ 232
$\mathscr{C}(X)$ 265
$L(X, Y)$ 431
X' 431
$\mathscr{P}[a, b]$ 471
$P \preceq Q$ 472
$\mathscr{E}[a, b]$ 474
$S_P(f)$ 476

Lista de personas

Ackermann, Wilhelm 13
Aigner, Martin 300
Baire, René-Louis 411
Banach, Stefan 34, 252, 317, 439, 468
Barrow, Isaac 500
Bendixson, Ivar Otto 231
Berge, Claude 21
Bergman, George 21
Bernstein, Felix 34
Bertrand, Joseph 13
Bolzano, Bernard 347, 354
Borel, Émile 346
Bourbaki, Nicolas 21, 27, 242
Brouwer, Luitzen Egbertus Jan 76, 300
Brown, Arlen 21
Brown, Kenneth 21
Bumby, Richard 81
Bunyakovsky, Viktor 116
Cameron, Peter 21
Cantor, Georg 34, 38, 58, 69, 230, 231
Cartan, Henri 242
Cassels, John William Scott 424
Catalan, Eugène Charles 9
Cauchy, Augustin-Louis 116
Chebyshev, Pafnuty Lvovich 13
Cohen, Paul 69
Conway, John H. 424
Dedekind, Richard 34, 56, 58, 76
Devidé, Vladimir 87
Dieudonné, Jean 362
Dini, Ulisse 391
Dugundji, James 21
Dyson, Freeman 427
Erdős, Paul 13, 69
Euler, Leonard 68
Ferreti, Andrea 83
Fichtenholz, Grigoriĭ Mikhaĭlovich 83
Fraenkel, Abraham 69
Fréchet, René Maurice 112, 147
Gödel, Kurt 69
Grayson, Daniel 21
Gromov, Mikhail 156, 179, 209
Guy, Richard K. 424
Hahn, Hans 439, 468
Halmos, Paul 21
Hölder, Otto 128, 285

Hamming, Richard 157
Hartogs, Friedrich Moritz 34
Hastings Moore, Eliakim 241, 271
Hatcher, Allen 300
Hausdorff, Felix 29, 83, 87, 112, 411, 468
Heine, Eduard 346
Hermite, Charles 68
Herrlich, Horst 21
Hilbert, David 13, 20, 69, 70, 77, 252
Hildebrandt, Theophil Henry 468
Hinkis, Arie 34
Hurwitz, Adolf 423
Incatasciato, Guillermo 21
Jacobson, Nathan 83
Jain, Arjun 21
James, Robert 444
Jordan, Pascual 130
Kantorovitch, Leonid Vitaľevich 83
Kaplansky, Irving 81
Kechris, Alexander 333
Kelley, John L. 241
Knaster, Bronisław 34
Kneser, Hellmuth 21
Komjáth, Péter 69
König, Gyula 87
Korselt, Alwin Reinhold 34
Kronecker, Leopold 76
Kuratowski, Kazimierz 21, 28
Kuttler, Kenneth 21
Lagrange, Joseph-Louis 424
Lambert, Johann Heinrich 68
Lang, Serge 21, 68
Lebesgue, Henri Léon 358
Lejeune Dirichlet, Peter Gustav 42, 422
Lektorean, Lektor 21
Lewin, Jonathan 21
Lindelöf, Ernst 323
Lindelöf, Ernst Leonard 224
Liouville, Joseph 68, 421, 426
Lipschitz, Rudolf 285, 323
Markov, Andrei 423
Milnor, John 301
Minkowski, Hermann 126
Munkres, James 365
Nagata, Jin-iti 21
Newton, Isaac 500
Noether, Emmy 20
Nuida, Koji 21
Orrin, Frink 21
Osgood, William Fogg 271, 411
Ostrowski, Alexander 172
Péter, Rózsa 13
Peano, Giuseppe 77
Picard, Émile 323
Poincaré, Henri 76
Ramanujan, Srinivasa 13
Reed, Irving S. 163
Robinson, Raphael 13
Roth, Klaus Friedrich 421, 427
Rubin, Herman 21
Rubin, Jean 21
Rudin, Mary Ellen 365
Rudin, Walter 21
Sánchez Terraf, Pedro 21
Schauder, Juliusz 460
Schröder, Ernst 34
Schwarz, Hermann 116
Shurman, Jerry 21
Siegel, Carl Ludwig 427
Sierpiński, Wacław 69
Smith, Herman 241
Sokal, Alan 468
Solomon, Gustave 163
Spronk, Nico 21
Srivastava, Shashi Mohan 333
Stanley, Richard 9
Steinhaus, Hugo 468
Stone, Arthur Harold 365

Stone, Marshall 397
Sylvester, James Joseph 119
Szele, Tibor 21
Tarski, Alfred 34
Teichmüller, Oswald 322
Thue, Axel 427
Tietze, Heinrich 297
Tikhonov, Andrey Nikolayevich 350
Tomas, Jech 21
Urysohn, Pavel Samuilovich 296
von Lindemann, Ferdinand 68
von Neumann, John 99, 130
Weierstrass, Karl 347, 354, 397
Weston, Jeffrey Dennis 21
Weyl, Hermann 76
Witt, Ernst 27
Wittgenstein, Ludwig 76
Zarouali-Darkaoui, Mohssin 21
Zermelo, Ernest 69
Zermelo, Ernst 34, 87
Ziegler, Günther 301
Zorn, Max 21

Índice

Álgebra
de Banach . . . 394
normada . . . 393
Anticadena . . . 30
Approximate solution . . . 376
Automorfismo
de un conjunto ordenado . . . 94
Axioma
de elección . . . 15
Base
de entornos de un punto . . . 185
de una topología . . . 192
Bola
abierta . . . 176
cerrada . . . 211
Cadena . . . 19
Carácter de Teichmüller . . . 322
Cardinal . . . 32, 110
Clase
C_p^0 . . . 375
C_p^1 . . . 375
Clausura . . . 210
Completación
de un espacio métrico . . . 304
de un espacio normado . . . 452
Completitud . . . 252
Condición
de Hölder . . . 285
de Lipschitz . . . 285
Condición de Cauchy . . . 248
Conjunto
abierto . . . 182
acotado . . . 234
bien ordenado . . . 86
cerrado . . . 210
contable . . . 63
convexo . . . 123
de Cantor . . . 409
denso . . . 218
dirigido . . . 241, 472
F_σ . . . 332
finito . . . 39
G_δ . . . 332
inductivo . . . 2
infinito . . . 51
localmente cerrado . . . 216

numerable 58
subyacente 112
transitivo 99
Constante
de Lipschitz 285
de Markov 423
Continuidad
uniforme 285
Contracción 317
Cota superior 19
Cubrimiento 224
Desigualdad
de Cauchy–Bunyakovsky–Schwartz 116
de Hölder 128
de Minkowski 126
triangular 112
ultramétrica 164
Diagonal 211
Diagrama de Cayley 156
Distancia 112
Distancia mínima de un código 161
Elemento
máximo 19
maximal 19
Entorno 184
Epimorfismo 293
Equiconvergencia 270
Equipotencia 31
Equivalencia
Cauchy de métricas 265
de normas 205
Lipschitz de métricas 204
topológica de métricas 200
Espacio
compacto 338
de Banach 252
de Hilbert 252
localmente compacto 359
métrico 112
métrico completo 252
normado 121
paracompacto 362
regular 211
separable 219
totalmente acotado 341
ultramétrico 164
Espacio dual continuo 431
Espacio normado
reflexivo 444
Espacio vectorial
localmente convexo 188
Exponente de Hölder 285
Extensión lineal 26
Familia localmente finita 214, 362
Filtro 242
Función
abierta 460
compacta 373
continua 275
continua a trozos 375
continua en un punto 274
de Ackermann–Péter 13
de elección 15
derivable con continuidad a trozos . 375
escalonada 474
integrable 479
lineal a trozos 376
recursiva primitiva 12
transpuesta 448
uniforme continua 285
Gráfico de una función 278, 465
Grupo de homeomorfismos 280
Hipótesis del continuo 69
Homeomorfismo 280
Interior de un conjunto 182
Isomorfismo

de conjuntos ordenados 94
Jerarquía de Borel 333
Lema
de Riemann–Lebesgue 505
de Schauder 460
de Zorn 21
Ley del paralelogramo 129
Localización 180
Métrica 112
asociada a una norma 122
de Hamming 157
de las palabras 156
de los caminos 154
discreta 153
euclídea 114
invariante por translaciones 115
Métricas equivalentes
en el sentido de Cauchy 265
en el sentido de Lipschitz 204
topológicamente 200
Matriz definida positiva 119
Número
algebraico 68
de Catalan 9
de Lagrange 424
de Lebesgue 358
diofantino 427
trascendente 68
Norma 121
asociada a un producto interno 122
Normas equivalentes 205
Ordinal 99
Partición de un intervalo 471
Polinomio minimal 68
Polinomios de Laurent 400
Potencia del continuo 59
Pre-orden 34
Principio
de buena ordenación 2
de Dirichlet 42, 422
de inducción 2, 3
de los intervalos encajados 257
de recurrencia. 4
del palomar 42, 422
maximal de Hausdorff 29
Producto interno 116
Propiedad
de intersección finita 339
de intersección numerable 225
Punto
aislado 226
de acumulación 226
de adherencia 210
de condensación 228
interior 182
límite 245
Rango de Cantor–Bendixson 231
Razón áurea 423
Red 241
Refinamiento de particiones 472
Regla
de integración por partes 502
de integración por sustitución 501
Regla de Newton 500
Retracción 300
Retracto 300
Segmento 123
Semi-métrica 138
Semi-norma 140
Subbase 195
Subsucesión 243
Sucesión
acotada 234
de Cauchy 248
Sucesor 107
Teorema
de Arzelà–Ascoli 369
de Banach–Schauder 462

de Bolzano–Weierstrass 347, 354
de Dini 391
de extensión de Tietze 297
de hahn–Banach 439
de Heine–Borel 346
de Jordan y von Neumann 130
de la base de Hilbert 20
de la función abierta 462
de la función inversa acotada 464
de los valores extremos 354
de Moore–Osgood 271
de punto fijo de Banach 317
de punto fijo de Bourbaki–Witt 27
de punto fijo de Kuratowski 29
de Thue–Siegel–Roth 427
del gráfico cerrado 467
fundamental del álgebra 356
fundamental del cálculo 498
fundamental del cálculo, 2da parte . 500
Topología 190
discreta 191
Tricotomía 36
Ultramétrica 164
p-ádica 171
Valor absoluto 172
equivalente 174
Valor absoluto *p*-ádico 170
Valuación *p*-ádica 168