

# 1 Espacios normados, métricos y Topología

Durante todo este curso “espacio vectorial” querrá decir “espacio vectorial real”, es decir que el cuerpo de escalares será siempre  $\mathbb{R}$ .

Informalmente la palabra “espacio” quiere decir “conjunto con alguna estructura”. El vacío  $\emptyset$  sí que es un conjunto, pero no se le pone ninguna estructura. Por lo tanto, a todos los tipos de espacio que se definan se les exigirá ser no vacíos.

## 1.1 Normas euclídeas

Recordemos que un **producto escalar** en un espacio vectorial  $\mathbb{V}$  es una función de dos variables

$$\begin{aligned} \langle \cdot, \cdot \rangle : \mathbb{V} \times \mathbb{V} &\longrightarrow \mathbb{R} \\ (v, w) &\longmapsto \langle v, w \rangle \end{aligned}$$

cumpliendo las siguientes condiciones:

1. Bilineal: lineal en la variable  $v$  cuando se congela el valor de  $w$  y lineal en  $w$  cuando se congela el valor de  $v$ .
2. Simétrica:  $\langle v, w \rangle = \langle w, v \rangle$ .
3. Definida positiva:  $\langle v, v \rangle > 0$  para todo  $v \neq \mathbf{0}$ .

En el caso particular  $\mathbb{V} = \mathbb{R}^n$ , todos los productos escalares vienen dados por la siguiente fórmula (piénsese en  $x, y$  como vectores columna):

$$\langle x, y \rangle = x^t A y,$$

donde  $A$  es cualquier matriz  $n \times n$  simétrica y definida positiva (es decir, con todos los autovalores estrictamente positivos). El **producto escalar estándar** en  $\mathbb{R}^n$  corresponde a  $A = I_n$ , es decir que si  $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n)$  entonces:

$$x \cdot y = x^t y = x_1 y_1 + \dots + x_n y_n.$$

**Definición 1.** La longitud euclídea o norma euclídea asociada al producto escalar  $\langle \cdot, \cdot \rangle$  es la siguiente función:

$$\| \cdot \| : \mathbb{V} \rightarrow \mathbb{R}, \quad \|v\| = \sqrt{\langle v, v \rangle},$$

donde se toma la raíz cuadrada no negativa.

En particular la **norma euclídea estándar** en  $\mathbb{R}^n$  es  $\|(x_1, \dots, x_n)\| = \sqrt{x_1^2 + \dots + x_n^2}$ .

Para todo producto escalar las siguientes propiedades son obvias:

1.  $\|v\| \geq 0$ ,
2.  $\|v\| = 0 \iff v = \mathbf{0}$ ,
3.  $\|\lambda v\| = |\lambda| \|v\|$ ,

y se les añaden otras dos, no tan obvias:

$$\text{Desigualdad de Schwarz: } |\langle v, w \rangle| \leq \|v\| \|w\|.$$

$$\text{Desigualdad triangular: } \|v + w\| \leq \|v\| + \|w\|.$$

Para efectuar geoméricamente la suma de vectores  $v + w$  podemos trasladar paralelamente  $w$  hasta que su origen coincida con la punta de  $v$ , entonces  $v + w$  une el origen de  $v$  con la punta (trasladada) de  $w$ . Colocados así,  $v, w, v + w$  son los *lados de un triángulo*. La desigualdad triangular se llama así porque afirma que en cualquier triángulo la longitud de un lado no supera la suma de las longitudes de los otros dos lados.

La demostración algebraica de la desigualdad de Schwarz se basa en el concepto de **matriz de Gram** de una sucesión de vectores  $v_1, \dots, v_k$ , que es simplemente la “tabla de multiplicar”

$$G = [\langle v_i, v_j \rangle]_{1 \leq i, j \leq k},$$

obviamente simétrica  $k \times k$ . Si  $v_1, \dots, v_k$  son linealmente independientes entonces  $G$  es definida positiva y por lo tanto con determinante positivo. En el caso particular de dos vectores  $v, w$ :

$$G = \begin{bmatrix} \langle v, v \rangle & \langle v, w \rangle \\ \langle w, v \rangle & \langle w, w \rangle \end{bmatrix}, \quad \det G = \langle v, v \rangle \langle w, w \rangle - \langle v, w \rangle^2.$$

Por lo tanto, si  $v, w$  son linealmente independientes entonces  $\|v\|^2 \|w\|^2 - \langle v, w \rangle^2 > 0$  y en este caso la desigualdad de Schwarz es *estricta*:  $|\langle v, w \rangle| < \|v\| \|w\|$ . Es trivial ver que la desigualdad de Schwarz es una igualdad cuando  $v, w$  son linealmente dependientes.

Otra demostración del caso estricto de la desigualdad de Schwarz es la siguiente. Supuestos  $v, w$  linealmente independientes, les aplicamos el *proceso de Gram-Schmidt* y obtenemos dos vectores *ortonormales*  $v_1, v_2$  tales que  $v = a_1 v_1$ ,  $w = b v_1 + a_2 v_2$  para ciertos números  $a_1, b, a_2$  con  $a_1, a_2 > 0$ . Entonces:

$$|\langle v, w \rangle| = a_1 |b|, \quad \|v\| \|w\| = a_1 \sqrt{a_2^2 + b^2} > a_1 |b|.$$

Geoméricamente, en el plano vectorial generado por  $v, w$  hemos tomado unos *ejes ortogonales adaptados* al par  $v, w$ : el primer eje (el del vector  $v_1$ ) contiene al vector  $v$ .

Como corolario de la desigualdad de Schwarz se obtiene la demostración algebraica de la desigualdad triangular:

$$\langle v + w, v + w \rangle = \langle v, v \rangle + \langle w, w \rangle + 2 \langle v, w \rangle \leq \|v\|^2 + \|w\|^2 + 2 \|v\| \|w\| = (\|v\| + \|w\|)^2,$$

pero hay demostraciones más geométricas. Por ejemplo, empecemos por observar que si  $v_0$  es la *proyección ortogonal* de un vector  $v$  sobre cualquier subespacio vectorial entonces  $\|v_0\| \leq \|v\|$ , es decir que las proyecciones ortogonales acortan longitudes (o las dejan igual). La desigualdad triangular es obvia si  $v + w = \mathbf{0}$ . Supuesto  $v + w \neq \mathbf{0}$ , sean  $v_0, w_0$  las proyecciones ortogonales respectivas de  $v, w$  sobre la recta vectorial generada por  $v + w$ , con lo cual  $v + w = v_0 + w_0$ . Es fácil convencerse, con algunos dibujos, de que la longitud euclídea de  $v + w = v_0 + w_0$  es igual a uno de los tres números siguientes:

$$\|v_0\| + \|w_0\|, \quad \|v_0\| - \|w_0\|, \quad \|w_0\| - \|v_0\|,$$

en todo caso un valor menor o igual que  $\|v\| + \|w\|$ .

## 1.2 Normas en general

Antes de definir el concepto de norma, introducimos una noción más básica que tiene mucha utilidad.

**Definiciones 2.** Sea  $\mathbb{V}$  un espacio vectorial de dimensión no nula y sea  $f : \mathbb{V} \setminus \{\mathbf{0}\} \rightarrow \mathbb{R}$  una función escalar (que puede estar definida o no en el vector nulo).

Para un entero  $k$ , decimos que  $f$  es **homogénea de grado  $k$**  si

$$v \in \mathbb{V} \setminus \{\mathbf{0}\}, \lambda \neq 0 \implies f(\lambda v) = \lambda^k f(v).$$

Para un  $a \in \mathbb{R}$  (positivo, nulo o negativo) decimos que  $f$  es **positivamente homogénea de grado  $a$**  si

$$v \in \mathbb{V} \setminus \{\mathbf{0}\}, \lambda > 0 \implies f(\lambda v) = \lambda^a f(v).$$

Ejemplos. Todo polinomio homogéneo de grado  $k$  en  $n$  variables define una función homogénea de grado  $k$  en  $\mathbb{R}^n$ . La fórmula  $f(x, y) = 1/\sqrt{|x| + |y|}$  define una función en  $\mathbb{R}^2 \setminus \{(0, 0)\}$  que es positivamente homogénea de grado  $-0'5$ .

Cuando el grado es positivo (entero o fraccionario) extendemos esas funciones al origen poniendo  $f(\mathbf{0}) = 0$ . Entonces el **grafo** de  $f$  es el siguiente subconjunto de  $\mathbb{V} \times \mathbb{R}$ :

$$\text{grafo}(f) = \{(x, f(x)) : x \in \mathbb{V}\}.$$

Si  $\mathbb{V} = \mathbb{R}^n$ , entonces el grafo es un subconjunto de  $\mathbb{R}^{n+1}$ .

Que  $f$  sea positivamente homogénea de grado 1 significa que su grafo es un **cono**: una unión de semirrectas en  $\mathbb{V} \times \mathbb{R}$  que salen del origen  $(\mathbf{0}, 0)$ . Que  $f$  sea homogénea de grado 1 significa que el grafo es un cono simétrico respecto de dicho origen: una unión de rectas pasando por  $(\mathbf{0}, 0)$ .

**Definiciones 3.** Sea  $\mathbb{V}$  un espacio vectorial de dimensión no nula. Una **norma** en  $\mathbb{V}$  es cualquier función

$$\begin{aligned} \|\cdot\| &: \mathbb{V} \longrightarrow \mathbb{R} \\ v &\longmapsto \|v\| \end{aligned}$$

cumpliendo las siguientes condiciones:

1.  $\|\mathbf{0}\| = 0$  y  $\|v\| > 0$  cuando  $v \neq \mathbf{0}$ .
2. Positivamente homogénea de grado 1:  $\lambda > 0 \implies \|\lambda v\| = \lambda \|v\|$ .
3. Función par:  $\|-v\| = \|v\|$ .
4. Desigualdad triangular:  $\|v + w\| \leq \|v\| + \|w\|$ .

Un **espacio normado** es un par  $(\mathbb{V}, \|\cdot\|)$  formado por un espacio vectorial  $\mathbb{V}$  y una norma  $\|\cdot\|$  en  $\mathbb{V}$ .

Lo que esperamos de una norma es que sirva para *medir longitudes* de vectores. La propiedad 1. responde a dos ideas: (1) no queremos longitudes negativas, (2) un buen criterio para saber si un vector es nulo es mirar si su longitud es nula. Las propiedades 2. y 3. suelen juntarse en la siguiente fórmula:

$$\|\lambda v\| = |\lambda| \|v\|,$$

pero aquí las mantendremos separadas mientras analizamos su significado.

**Definiciones 4.** Sean  $(\mathbb{V}, \|\cdot\|)$  un espacio normado y  $r > 0$ . La **bola abierta de centro el origen y radio  $r$**  es el conjunto:

$$B(\mathbf{0}, r) = \{v \in \mathbb{V} : \|v\| < r\},$$

y la **bola cerrada**, del mismo centro y radio, es el conjunto:

$$\overline{B}(\mathbf{0}, r) = \{v \in \mathbb{V} : \|v\| \leq r\}.$$

La **bola unidad abierta** es  $B(\mathbf{0}, 1)$ . La **bola unidad cerrada** es  $\overline{B}(\mathbf{0}, 1)$ , cuya “cáscara” es el conjunto  $\{v \in \mathbb{V} : \|v\| = 1\}$  de los **vectores unitarios** para la norma  $\|\cdot\|$ .

Un subconjunto  $E \subset \mathbb{V}$  es **acotado** si está contenido en alguna bola, es decir si  $E \subseteq \overline{B}(\mathbf{0}, r)$  para algún  $r > 0$ .

Dado un espacio normado  $(\mathbb{V}, \|\cdot\|)$ , cada vector no nulo  $v \in \mathbb{V} \setminus \{\mathbf{0}\}$  tiene una “descomposición polar”, es decir una factorización  $v = \lambda \omega$  con  $\lambda$  escalar positivo y  $\omega$  vector unitario:

$$\mathbf{0} \neq v = \lambda \omega \quad \text{con} \quad \begin{cases} \lambda > 0 \\ \|\omega\| = 1 \end{cases}$$

Esta factorización resulta ser única, con  $\lambda = \|v\|$  y, por lo tanto, con  $\omega = v/\|v\|$ .

### 1.3 Las normas $p$ en $\mathbb{R}^n$

**Definiciones 5.** Sea  $1 < p < \infty$ . La **norma  $p$**  en  $\mathbb{R}^n$  es la función  $\|\cdot\|_p : \mathbb{R}^n \rightarrow \mathbb{R}$  definida de la siguiente manera:

$$\|(x_1, \dots, x_n)\|_p = \left[ \sum_{j=1}^n |x_j|^p \right]^{\frac{1}{p}}.$$

El **exponente conjugado** de  $p$  es el único número  $p'$  tal que

$$\frac{1}{p} + \frac{1}{p'} = 1. \quad (1)$$

Es trivial obtener la fórmula  $p' = \frac{p}{p-1}$ . Por la simetría de la ecuación (1), si  $q$  es el conjugado de  $p$  entonces  $p$  es el conjugado de  $q$ . Eso equivale, a su vez, a que el grafo de la conjugación

$$\left\{ (p, p') = \left( p, \frac{p}{p-1} \right) : 1 < p < \infty \right\},$$

sea simétrico respecto de la diagonal principal del plano  $pp'$ . La ecuación (1) se transforma fácilmente en  $(p-1)(p'-1) = 1$ , luego dicho grafo es el resultado de trasladar la rama de hipérbola  $\{xy = 1, x > 0, y > 0\}$  una unidad hacia la derecha y una unidad hacia arriba (y así se mantiene simétrico respecto de la diagonal principal). Otra consecuencia es que la conjugación es una biyección decreciente del intervalo  $(1, \infty)$  consigo mismo. Cuando  $p$  tiende a 1 su conjugado tiende a  $\infty$  y cuando  $p$  tiende a  $\infty$  su conjugado tiende a 1. El exponente  $p = 2$  es el único que es conjugado de sí mismo.

**Teorema 6.** Se cumplen las tres desigualdades siguientes:

$$\text{Desigualdad de Young: } a, b > 0 \implies ab \leq \frac{a^p}{p} + \frac{b^{p'}}{p'}.$$

$$\text{Desigualdad de Hölder: } x, y \in \mathbb{R}^n \implies |x \cdot y| \leq \|x\|_p \|y\|_{p'}.$$

$$\text{Desigualdad de Minkowski: } \|x + y\|_p \leq \|x\|_p + \|y\|_p.$$

Es fácil ver que la función  $\|\cdot\|_p$  cumple las propiedades 1., 2. y 3. de la definición de norma. La desigualdad de Minkowski nos dice que también cumple la propiedad 4., luego es efectivamente una norma en  $\mathbb{R}^n$ .

La desigualdad de Young con  $p = 2$  nos da  $\sqrt{a^2 b^2} \leq \frac{a^2}{2} + \frac{b^2}{2}$ , es decir la *desigualdad aritmético-geométrica*  $\sqrt{xy} \leq \frac{x+y}{2}$  con  $(x, y) = (a^2, b^2)$  números positivos cualesquiera.

La norma  $\|\cdot\|_2$  es la norma euclídea estándar en  $\mathbb{R}^n$  y la desigualdad de Hölder con  $p = 2$  es la desigualdad de Schwarz.

Introducimos a continuación unas nociones necesarias para demostrar el teorema 6. Recordemos que, dados un espacio vectorial  $\mathbb{V}$  y puntos  $x, y \in \mathbb{V}$ , el **segmento rectilíneo**  $[x, y]$  de extremos  $x, y$  admite la siguiente descripción:

$$[x, y] = \{ \lambda x + (1 - \lambda)y : 0 \leq \lambda \leq 1 \}.$$

**Definiciones 7.** Un subconjunto  $E \subseteq \mathbb{V}$  es **convexo** si siempre que  $x, y \in E$  se tiene  $[x, y] \subseteq E$ . Sean  $E \subseteq \mathbb{V}$  conjunto convexo no vacío y  $f : E \rightarrow \mathbb{R}$ . Decimos que  $f$  es **convexa** si cumple:

$$x, y \in E, \lambda \in [0, 1] \implies f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

Análogamente decimos que  $f$  es **cóncava** si cumple:

$$x, y \in E, \lambda \in [0, 1] \implies f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y).$$

Que  $f$  sea convexa quiere decir que el **supergrafo** de  $f$ :

$$\{(x, t) \in \mathbb{V} \times \mathbb{R} : x \in E, t \geq f(x)\},$$

es un subconjunto convexo de  $\mathbb{V} \times \mathbb{R}$ . Ejemplo:  $E = \mathbb{R}, f(x) = x^2$ .

Que  $f$  sea cóncava quiere decir que el **subgrafo** de  $f$ :

$$\{(x, t) \in \mathbb{V} \times \mathbb{R} : x \in E, t \leq f(x)\},$$

es un subconjunto convexo de  $\mathbb{V} \times \mathbb{R}$ . Ejemplo:  $E = (0, +\infty) \subset \mathbb{R}, f(x) = \log x$ .

*Demostración de la desigualdad de Young.* Hacemos  $\lambda = \frac{1}{p} \in [0, 1]$ , de donde  $1 - \lambda = \frac{1}{p'}$ . Por la concavidad del logaritmo neperiano:

$$\log\left(\frac{a^p}{p} + \frac{b^{p'}}{p'}\right) \geq \frac{1}{p} \log(a^p) + \frac{1}{p'} \log(b^{p'}) = \log a + \log b = \log(ab),$$

y tomando exponenciales sale la desigualdad de Young.

*Demostración de la desigualdad de Hölder.* Es obvia cuando  $x = \mathbf{0}$  o  $y = \mathbf{0}$ . Para  $x \neq \mathbf{0} \neq y$  tomamos las descomposiciones polares,  $x = \|x\|_p \alpha$  e  $y = \|y\|_{p'} \beta$  con  $\|\alpha\|_p = 1 = \|\beta\|_{p'}$ , lo que convierte la desigualdad de Hölder en  $\|x\|_p \|y\|_{p'} |\alpha \cdot \beta| \leq \|x\|_p \|y\|_{p'}$ . Basta, pues, con demostrar lo siguiente:

$$\|\alpha\|_p = 1 = \|\beta\|_{p'} \implies |\alpha \cdot \beta| \leq 1. \quad (2)$$

Si  $\alpha = (\alpha_1, \dots, \alpha_n), \beta = (\beta_1, \dots, \beta_n)$ , aplicamos la desigualdad de Young a cada producto  $|\alpha_j \beta_j|$  y obtenemos:

$$|\alpha \cdot \beta| = \left| \sum_{j=1}^n \alpha_j \beta_j \right| \leq \sum_{j=1}^n |\alpha_j \beta_j| \leq \frac{1}{p} \sum_{j=1}^n |\alpha_j|^p + \frac{1}{p'} \sum_{j=1}^n |\beta_j|^{p'} = \frac{1}{p} \cdot 1^p + \frac{1}{p'} \cdot 1^{p'} = 1,$$

lo que demuestra (2) y, por lo explicado, la desigualdad de Hölder en general.

*Demostración rápida de la desigualdad de Minkowski.*

Dados  $x = (x_1, \dots, x_n)$  e  $y = (y_1, \dots, y_n)$ , definimos los siguientes vectores:

$$x' = (|x_1|, \dots, |x_n|), \quad y' = (|y_1|, \dots, |y_n|), \quad z = (|x_1 + y_1|^{p-1}, \dots, |x_n + y_n|^{p-1}).$$

Empezamos con la siguiente estimación:

$$\|x + y\|_p^p = \sum_{j=1}^n |x_j + y_j|^p \leq \sum_{j=1}^n (|x_j| + |y_j|) |x_j + y_j|^{p-1} = x' \cdot z + y' \cdot z,$$

aplicamos la desigualdad de Hölder:

$$x' \cdot z + y' \cdot z \leq \|x'\|_p \|z\|_{p'} + \|y'\|_p \|z\|_{p'},$$

y comprobamos que  $\|z\|_{p'} = \|x + y\|_p^{p-1}$  (aquí se usa la fórmula  $p' = p/(p-1)$ ). Como es obvio que  $\|x'\|_p = \|x\|_p$  y que  $\|y'\|_p = \|y\|_p$ , llegamos a:

$$\|x + y\|_p^p \leq (\|x\|_p + \|y\|_p) \|x + y\|_p^{p-1}.$$

Si  $x + y \neq \mathbf{0}$ , dividimos por  $\|x + y\|_p^{p-1}$  y ya está. Si  $x + y = \mathbf{0}$  la desigualdad de Minkowski es obvia.  $\square$

Daremos también una demostración geométrica de la desigualdad de Minkowski. Con ese fin, haremos ahora un estudio válido para cualquier norma  $\|\cdot\|$  en cualquier espacio vectorial  $\mathbb{V}$ , no solamente para las normas  $p$  en  $\mathbb{R}^n$ .

Hermann Minkowski (1864-1909) se planteó cómo determinar la norma  $\|\cdot\|$  si sólo se conoce la bola unidad  $\overline{B} = \overline{B}(\mathbf{0}, 1)$  como subconjunto de  $\mathbb{V}$ . Por supuesto  $\|\mathbf{0}\| = 0$ , luego lo que hay que determinar es la norma de los vectores no nulos. Dados  $v \neq \mathbf{0}$  y  $\lambda > 0$ , observamos que:

$$\frac{v}{\lambda} \notin \overline{B} \text{ para } \lambda < \|v\| \quad \text{y} \quad \frac{v}{\lambda} \in \overline{B} \text{ para } \lambda \geq \|v\|.$$

Por lo tanto el siguiente conjunto de números:

$$\left\{ \lambda > 0 : \frac{v}{\lambda} \in \overline{B} \right\}, \quad (3)$$

es el intervalo  $[\|v\|, +\infty)$ . La norma  $\|v\|$  se determina como el ínfimo de este intervalo.

**Definición 8.** Dado un subconjunto  $\overline{B} \subset \mathbb{V}$  la **función de Minkowski** asociada a él es la función  $\varphi : \mathbb{V} \rightarrow \mathbb{R}$  definida como sigue. En primer lugar  $\varphi(\mathbf{0}) = 0$ . En segundo lugar:

$$v \neq \mathbf{0} \implies \left\{ \lambda > 0 : \frac{v}{\lambda} \in \overline{B} \right\} = [\varphi(v), +\infty).$$

Finalmente nos preguntamos cómo tiene que ser la geometría del conjunto  $\overline{B}$  para que su función de Minkowski sea una norma en  $\mathbb{V}$ . Primero pedimos que verdaderamente sea una función, es decir que para todo  $v \neq \mathbf{0}$  se determine un valor  $\varphi(v) > 0$ . Esto requiere que el conjunto (3) sea un intervalo de la forma  $[\lambda_0, +\infty)$  con  $\lambda_0 > 0$ . Poniendo  $a = 1/\lambda$ , es equivalente que el conjunto:

$$\{ av \in \overline{B} : a > 0 \},$$

sea de la forma  $(0, a_0]v$  para algún  $a_0$  finito y positivo. Como  $(0, +\infty)v = \{av : a > 0\}$  es una *semirrecta* que emana del origen en  $\mathbb{V}$ , y con el origen quitado, tenemos:

**Primer requisito:**  $\overline{B}$  contiene al  $\mathbf{0}$  e interseca cada semirrecta  $(0, +\infty)v$  en un segmento  $(0, a_0]v$  finito y no vacío.

Una vez que el conjunto  $\overline{B}$  cumple el primer requisito, la función de Minkowski asociada está definida en todo  $\mathbb{V}$  y cumple la condición 1. de la definición de norma. Es fácil ver que también cumple la condición 2, es decir que de hecho es positivamente homogénea de grado 1. *Ahora la podemos ver como la única función  $\varphi$  que tiene esa homogeneidad y cumple  $\varphi^{-1}([0, 1]) = \overline{B}$ .*

También es fácil ver que, cumplido el primer requisito, la función de Minkowski es par (condición 3. de la definición de norma) si y sólo si  $\overline{B}$  es *simétrico respecto del origen*:

**Segundo requisito:**  $v \in \overline{B} \iff -v \in \overline{B}$ .

**Proposición 9.** Si  $f : \mathbb{V} \rightarrow \mathbb{R}$  es convexa entonces cada **conjunto de subnivel**:

$$f^{-1}((-\infty, c]) = \{v \in \mathbb{V} : f(v) \leq c\},$$

es o vacío o convexo.

La demostración es inmediata y se deja como ejercicio. También queda como ejercicio ver que **toda norma es convexa**, luego por la proposición anterior la bola unidad de cualquier norma tiene que ser un conjunto convexo:

**Tercer requisito:**  $\overline{B}$  debe ser un subconjunto convexo de  $\mathbb{V}$ .

**Teorema 10.** Si un conjunto  $\overline{B} \subset \mathbb{V}$  cumple los tres requisitos entonces es la bola unidad de una (única) norma en  $\mathbb{V}$ .

*Demostración del teorema 10.* Ya sabemos que si se cumple el primer requisito entonces la función de Minkowski está definida en todo  $\mathbb{V}$  y cumple las propiedades 1. y 2. de la definición de norma. También hemos visto que los requisitos primero y segundo implican la propiedad 3. Habremos terminado si vemos que los requisitos primero y tercero implican la desigualdad triangular (propiedad 4.). Dicha desigualdad es obvia si uno de los vectores  $x, y$  es nulo, por lo que suponemos  $x \neq \mathbf{0} \neq y$ . Tomamos las descomposiciones polares:

$$x = a\alpha, \quad y = b\beta, \quad a = \varphi(x) > 0, \quad b = \varphi(y) > 0, \quad \varphi(\alpha) = \varphi(\beta) = 1,$$

y haciendo  $\lambda = a/(a+b) \in [0, 1]$  escribimos:

$$x + y = (a+b) \left( \frac{a}{a+b} \alpha + \frac{b}{a+b} \beta \right) = (a+b) (\lambda \alpha + (1-\lambda) \beta).$$

Como es  $\alpha, \beta \in \overline{B}$  y estamos suponiendo  $\overline{B}$  convexo, es  $\lambda \alpha + (1-\lambda) \beta \in \overline{B}$  y para la función de Minkowski  $\varphi$  se tiene  $\varphi(\lambda \alpha + (1-\lambda) \beta) \leq 1$ . Luego, como  $\varphi$  tiene la propiedad 2.:

$$\varphi(x+y) \leq (a+b) \cdot 1 = a+b = \varphi(x) + \varphi(y),$$

que es la desigualdad triangular. □

Ahora tenemos interpretaciones geométricas de las cuatro propiedades en la definición de norma. Dar una función cumpliendo 1. y 2. equivale a dar un conjunto  $\overline{B} \subset \mathbb{V}$  que contiene el origen  $\mathbf{0}$  y corta a cada semirrecta  $(0, +\infty)v$  en un segmento  $(0, a]v$  finito y no vacío. Cumplido eso, la propiedad 3. equivale a la simetría de  $\overline{B}$  respecto del origen y la 4. a la convexidad de  $\overline{B}$ .

*Demostración geométrica de la desigualdad de Minkowski.* La función  $\varphi(v) = \|v\|_p$  cumplirá la desigualdad triangular si el conjunto  $\overline{B} = \varphi^{-1}([0, 1])$  es convexo. Este conjunto tiene la siguiente descripción alternativa:

$$\overline{B} = \{x = (x_1, \dots, x_n) : |x_1|^p + \dots + |x_n|^p \leq 1\},$$

y, por la proposición 9, será convexo si la función  $x \mapsto \|x\|_p^p = \sum_{j=1}^n |x_j|^p$  es convexa.

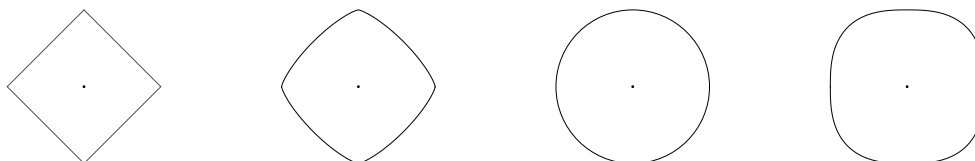
Fijado  $1 < p < \infty$  la función  $f : \mathbb{R} \rightarrow \mathbb{R}$  dada por  $f(t) = |t|^p$  es convexa. Se sigue que las  $n$  funciones  $f_j(x) = |x_j|^p$ ,  $1 \leq j \leq n$ , son todas convexas. Como la suma de funciones convexas es convexa,  $\sum_{j=1}^n |x_j|^p$  es función convexa de  $x = (x_1, \dots, x_n)$  y hemos terminado. □

Si en la fórmula que define las normas  $p$  en  $\mathbb{R}^n$  tomamos el valor extremo  $p = 1$ , resulta una función bien definida para todo  $x \in \mathbb{R}^n$ :

$$\|(x_1, \dots, x_n)\|_1 = |x_1| + \dots + |x_n|,$$

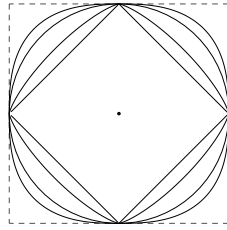
y es obvio que se trata de una norma.

Mostramos a continuación dibujos de la bola unidad  $\{v \in \mathbb{R}^2 : \|v\|_p \leq 1\}$ , para los valores (de izquierda a derecha)  $p = 1$ ,  $p = 1'3$ ,  $p = 2$  y  $p = 2'7$ :



Por supuesto, todas son subconjuntos convexos del plano.

A continuación mostramos esas mismas bolas superpuestas, y en línea de trazos la “forma límite” a la que tienden cuando  $p \rightarrow \infty$ :



Vemos que la bola aumenta en todas las direcciones a medida que  $p$  aumenta, excepto en las direcciones de los ejes coordenados en las que permanece igual. Esto se debe a una disminución monótona del valor de la norma:

**Proposición 11.** *Fijado un vector  $v_0 \in \mathbb{R}^n$  que no esté en los ejes coordenados, el número  $\|v_0\|_p$  es función estrictamente decreciente de  $p$ . Si por el contrario  $v_0$  está en uno de los ejes coordenados, entonces ese número no depende de  $p$ .*

En ambos casos  $p \mapsto \|v_0\|_p$  es una función monótona y tiene un límite finito (y no negativo) cuando  $p \rightarrow \infty$ . Es fácil demostrar que:

$$\lim_{p \rightarrow \infty} \|(x_1, \dots, x_n)\|_p = \max(|x_1|, \dots, |x_n|),$$

y el resultado define una norma en  $\mathbb{R}^n$ , que denotamos  $\|\cdot\|_\infty$  y de manera natural llamamos “norma infinito” en  $\mathbb{R}^n$ :

$$\|(x_1, \dots, x_n)\|_\infty = \max_{1 \leq j \leq n} |x_j|.$$

La bola unidad de  $\|\cdot\|_\infty$  es el producto cartesiano  $[-1, 1]^n$ . Cuando  $n = 2$  es el cuadrado límite que hemos mostrado más arriba.

**Advertencia.** La “cáscara”  $\{v : \|v\| = 1\}$ , que separa la bola unidad del espacio circundante, puede contener segmentos rectilíneos (ocurre, por ejemplo, con  $\|\cdot\|_1$  y con  $\|\cdot\|_\infty$ ). Repasando la demostración del teorema 10, vemos que si el segmento  $[\alpha, \beta]$  está todo él contenido en dicha cáscara entonces  $\|x + y\| = \|x\| + \|y\|$ , es decir que en este caso la desigualdad triangular no es estricta sino que es una igualdad para la pareja  $x, y$ , incluso si es una pareja linealmente independiente y el triángulo de lados  $x, y, x + y$  es no degenerado.

Para las normas  $p$  con  $1 < p < \infty$ , y triángulos no degenerados, la desigualdad triangular es estricta.

**Definición 12.** *Sea  $\mathbb{V}$  un espacio vectorial de dimensión finita  $n$ . Un **elipsoide** centrado en el origen es el conjunto imagen en  $\mathbb{V}$  de la bola estándar  $\{x \in \mathbb{R}^n : x \cdot x = 1\}$  por cualquier biyección lineal  $\mathbb{R}^n \rightarrow \mathbb{V}$ .*

*Definición equivalente:* es la bola unidad de alguna norma euclídea en  $\mathbb{V}$ .

Los dibujos que hemos mostrado en el plano pueden entenderse como las intersecciones con el plano  $x_1x_2$  de las bolas unidad de la norma  $p$  en  $\mathbb{R}^n$ . Es fácil demostrar que sólo para  $p = 2$  es una elipse en el plano  $x_1x_2$ , luego sólo la bola unidad de  $\|\cdot\|_2$  es un elipsoide en  $\mathbb{R}^n$ .

De todas las normas  $\|\cdot\|_p$  en  $\mathbb{R}^n$ , solamente la  $\|\cdot\|_2$  es euclídea.

La manera habitual de demostrar este último resultado es mediante la **identidad de paralelogramo**, que se satisface si y sólo si la norma es euclídea:

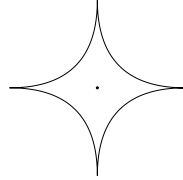
$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

Una vez que se satisface esta identidad, podemos recuperar el producto escalar (llamado **forma polar**), del que procede la norma, mediante la **identidad de polarización**:

$$4 \langle x, y \rangle = \|x + y\|^2 - \|x - y\|^2.$$



Cuando  $0 < p < 1$  la fórmula  $\varphi_p(x_1, \dots, x_n) = \left[ \sum_{j=1}^n |x_j|^p \right]^{1/p}$  todavía tiene sentido y la función  $\varphi_p(x)$  así definida tiene las propiedades 1., 2. y 3. de la definición de norma, pero no cumple la desigualdad triangular. Por ejemplo  $\varphi_{1/2}(x) = \left[ \sum_{j=1}^n \sqrt{|x_j|} \right]^2$  y el conjunto  $\{(x_1, x_2) : \varphi_{1/2}(x_1, x_2) \leq 1\}$  tiene la siguiente forma no convexa:



## 1.4 Distancias y continuidad

Sea  $(\mathbb{V}, \|\cdot\|)$  un espacio normado. Dados dos puntos  $x, y$ , el vector  $y - x$  puede entenderse como la flecha que sale de  $x$  y termina en  $y$ ; análogamente  $x - y$ . Es natural utilizar la longitud de este vector como *medida de lo apartados que están  $x, y$* . Por lo tanto, definimos la **distancia de  $x$  a  $y$**  (respecto de la norma  $\|\cdot\|$ ) como  $d(x, y) = \|x - y\|$  y tiene las siguientes propiedades:

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

**Definiciones 13.** Sea  $X$  un conjunto no vacío. Una **distancia** en  $X$  es cualquier función

$$\begin{aligned} d : X \times X &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto d(x, y) \end{aligned}$$

que cumpla las condiciones 1., 2. y 3., arriba indicadas, para cualesquiera  $x, y, z \in X$ .

Un **espacio métrico** es un par  $(X, d)$  formado por un conjunto  $X$  y una distancia  $d$  en  $X$ .

La propiedad 1. significa que no queremos distancias negativas y que un buen *criterio de igualdad* para dos puntos es ver si están a distancia nula. La propiedad 2. se llama *simetría*. La propiedad 3. se llama, por razones obvias, *desigualdad triangular*.

**Importante.** Escritas así, las condiciones 1., 2. y 3. tienen sentido en cualquier conjunto  $X$ , sin necesidad de que tenga estructura de espacio vectorial o de cualquier otro tipo. En consecuencia, hay muchos más espacios métricos que espacios normados. Incluso en el caso de que  $X$  es un espacio vectorial, pueden darse en él muchas distancias que no vengan de ninguna norma.

Como primer ejemplo, todo espacio normado tiene la distancia particular  $d(x, y) = \|x - y\|$ .

Segundo ejemplo. Si  $(X, d)$  es un espacio métrico y  $E \subseteq X$  es un subconjunto no vacío cualquiera, entonces la restricción

$$\begin{aligned} d_E : E \times E &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto d(x, y) \end{aligned}$$

es una distancia en  $E$  y así  $(E, d_E)$  también es un espacio métrico.

En particular, si  $\|\cdot\|$  es cualquier norma en  $\mathbb{R}^n$  y  $E \subseteq \mathbb{R}^n$  es *cualquier* subconjunto no vacío entonces  $E$  se convierte en un espacio métrico con la distancia  $\|x - y\|$  restringida a  $E \times E$ .

Tercer ejemplo. A partir de una distancia  $d_1$  en  $X$  construimos otra  $d_2$  poniendo:

$$d_2(x, y) = \min(d_1(x, y), 1). \quad (4)$$

**Definiciones 14.** Sea  $(X, d)$  un espacio métrico. La **bola abierta** con centro  $x_0 \in X$  y radio  $r > 0$  es el conjunto

$$B(x_0, r) = \{x \in X : d(x, x_0) < r\}.$$

La **bola cerrada** con centro  $x_0$  y radio  $r \geq 0$  (aquí permitimos el valor  $r = 0$ ) es el conjunto

$$\bar{B}(x_0, r) = \{x \in X : d(x, x_0) \leq r\}.$$

Un subconjunto  $E \subseteq X$  es **acotado** si está contenido en alguna bola, es decir si existen  $x_0 \in X$  y  $r \geq 0$  tales que  $E \subseteq \bar{B}(x_0, r)$ .

Sean  $(X, d_X), (Y, d_Y)$  espacios métricos y  $f : X \rightarrow Y$  una aplicación. Introducimos ahora el importante concepto de continuidad.

**Definiciones 15.** Decimos que  $f$  es **continua** en  $x_0 \in X$  si para todo  $\varepsilon > 0$  existe un  $\delta > 0$  tal que

$$f(B(x_0, \delta)) \subseteq B(f(x_0), \varepsilon).$$

Es decir, para toda bola  $B'$  centrada en  $f(x_0)$  hay una bola  $B$ , centrada en  $x_0$ , que es enviada por  $f$  dentro de  $B'$ .

Decimos que  $f$  es **continua** si es continua en todo punto de  $X$

**Importante.** Tomando cualquier  $\delta' < \delta$ , tenemos:

$$f(\bar{B}(x_0, \delta')) \subseteq f(B(x_0, \delta)) \subseteq B(f(x_0), \varepsilon) \subseteq \bar{B}(f(x_0), \varepsilon).$$

Por otra parte, si  $\bar{\varepsilon} < \varepsilon$  y  $\bar{\delta}$  es tal que  $f(\bar{B}(x_0, \bar{\delta})) \subseteq \bar{B}(f(x_0), \bar{\varepsilon})$  entonces:

$$f(B(x_0, \bar{\delta})) \subseteq \bar{B}(f(x_0), \bar{\varepsilon}) \subseteq B(f(x_0), \varepsilon).$$

Conclusión: **en la definición 15 podemos usar bolas abiertas o cerradas, a voluntad.**

**Proposición 16.** Sean espacios métricos  $(X, d_X), (Y, d_Y), (Z, d_Z)$  y aplicaciones  $X \xrightarrow{f} Y \xrightarrow{g} Z$ . Si  $f$  es continua en  $x_0 \in X$  y  $g$  es continua en  $f(x_0) \in Y$  entonces la compuesta  $g \circ f$  es continua en  $x_0$ .

Por lo tanto, si  $f, g$  son continuas entonces  $g \circ f$  es continua.

*Demostración.* Sea  $\varepsilon_3 > 0$ . Como  $g$  es continua en  $f(x_0)$ , existe  $\varepsilon_2 > 0$  tal que

$$g(B(f(x_0), \varepsilon_2)) \subseteq B(g(f(x_0)), \varepsilon_3) = B((g \circ f)(x_0), \varepsilon_3).$$

Como  $f$  es continua en  $x_0$ , dado  $\varepsilon_2$  existe  $\varepsilon_1 > 0$  tal que  $f(B(x_0, \varepsilon_1)) \subseteq B(f(x_0), \varepsilon_2)$  y así:

$$(g \circ f)(B(x_0, \varepsilon_1)) = g(f(B(x_0, \varepsilon_1))) \subseteq g(B(f(x_0), \varepsilon_2)) \subseteq B((g \circ f)(x_0), \varepsilon_3).$$

Como  $\varepsilon_3$  era arbitrario, queda probada la continuidad de  $g \circ f$  en  $x_0$ . □

## 1.5 Normas de operador

Sean  $(\mathbb{V}, \|\cdot\|_{\mathbb{V}}), (\mathbb{W}, \|\cdot\|_{\mathbb{W}})$  espacios normados y  $L : \mathbb{V} \rightarrow \mathbb{W}$  una aplicación lineal.

**Proposición 17.** Las condiciones siguientes son equivalentes:

1.  $L$  es continua en  $\mathbf{0}_{\mathbb{V}}$ .
2.  $L$  es continua en todo punto de  $\mathbb{V}$ .
3. La imagen  $L(\bar{B}(\mathbf{0}_{\mathbb{V}}, 1))$  es un subconjunto acotado en  $(\mathbb{W}, \|\cdot\|_{\mathbb{W}})$ , es decir que existe una "mayorante"  $M \geq 0$  tal que

$$\|v\|_{\mathbb{V}} \leq 1 \implies \|L(v)\|_{\mathbb{W}} \leq M.$$

**Definición 18.** A las  $L$  que cumplen esas condiciones se las llama **aplicaciones lineales acotadas**. El conjunto de todas ellas se denota por  $\mathcal{L}((\mathbb{V}, \|\cdot\|_{\mathbb{V}}), (\mathbb{W}, \|\cdot\|_{\mathbb{W}}))$  o, cuando no hay duda de cuales son las normas, por  $\mathcal{L}(\mathbb{V}, \mathbb{W})$ .

Recordemos que, fijado un vector  $v_0 \in \mathbb{V}$ , la **traslación por  $v_0$**  es la biyección:

$$T_{v_0} : \mathbb{V} \longrightarrow \mathbb{V} \quad , \quad v \longmapsto v + v_0 \quad ,$$

afín pero no lineal (excepto si  $v_0 = \mathbf{0}$ , en cuyo caso es la identidad).

Dado  $E \subseteq \mathbb{V}$ , el **trasladado**  $v_0 + E$  es la imagen  $\{v + v_0 : v \in E\}$  de  $E$  por  $T_{v_0}$ .

Recordemos asimismo que, dada una constante  $c > 0$ , la **homotecia de razón  $c$**  centrada en el origen es:

$$\mathbb{V} \longrightarrow \mathbb{V} \quad , \quad v \longmapsto cv \quad .$$

Dado  $E \subseteq \mathbb{V}$  denotamos por  $cE$  la imagen  $\{cv : v \in E\}$  de  $E$  por esa homotecia. Informalmente podemos entender  $cE$  como el mismo conjunto  $E$  visto a una escala diferente.

La identidad  $L(v + v_0) = L(v) + L(v_0)$  nos dice que  $L$  intercambia traslaciones según la siguiente fórmula:

$$L \circ T_{v_0} = T_{L(v_0)} \circ L \quad .$$

Análogamente  $L$  intercambia homotecias:  $L(cv) = cL(v)$ .

Las traslaciones y homotecias transforman bolas en bolas, según las siguientes fórmulas (recuerda que  $c > 0$ ), válidas también con bolas abiertas:

$$v_0 + \overline{B}(x_0, r) = \overline{B}(x_0 + v_0, r) \quad , \quad c\overline{B}(x_0, r) = \overline{B}(cx_0, cr) \quad .$$

*Demostración de la proposición 17.*

1.  $\iff$  2. Como  $L(\mathbf{0}_{\mathbb{V}}) = \mathbf{0}_{\mathbb{W}}$ , que  $L$  sea continua en  $\mathbf{0}_{\mathbb{V}}$  significa que para todo  $\varepsilon > 0$  existe un  $\delta > 0$  tal que

$$L(B(\mathbf{0}_{\mathbb{V}}, \delta)) \subseteq B(\mathbf{0}_{\mathbb{W}}, \varepsilon) \quad .$$

Aplicamos traslación por  $L(v_0)$  a ambos lados de esa inclusión y resulta:

$$L(v_0) + L(B(\mathbf{0}_{\mathbb{V}}, \delta)) \subseteq L(v_0) + B(\mathbf{0}_{\mathbb{W}}, \varepsilon) \quad ,$$

que, por las observaciones anteriores, se convierte en:

$$L(v_0 + B(\mathbf{0}_{\mathbb{V}}, \delta)) = L(B(v_0, \delta)) \subseteq B(L(v_0), \varepsilon) \quad ,$$

y  $L$  es continua en  $v_0$ , ya que  $\varepsilon$  era arbitrario.

1.  $\iff$  3. Para cualesquiera  $r_1, r_2 \geq 0$  y  $c > 0$  se tiene:

$$L(\overline{B}(\mathbf{0}_{\mathbb{V}}, r_1)) \subseteq \overline{B}(\mathbf{0}_{\mathbb{W}}, r_2) \implies L(\overline{B}(\mathbf{0}_{\mathbb{V}}, cr_1)) \subseteq \overline{B}(\mathbf{0}_{\mathbb{W}}, cr_2) \quad , \quad (5)$$

porque se pasa de la izquierda a la derecha aplicando homotecia de razón  $c$ .

Si  $L$  es continua en  $\mathbf{0}_{\mathbb{V}}$ , existe un  $\delta > 0$  con  $L(\overline{B}(\mathbf{0}_{\mathbb{V}}, \delta)) \subseteq \overline{B}(\mathbf{0}_{\mathbb{W}}, 1)$ . En la fórmula (5) hacemos  $(r_1, r_2, c) = (\delta, 1, 1/\delta)$  y nos da  $L(\overline{B}(\mathbf{0}_{\mathbb{V}}, 1)) \subseteq \overline{B}(\mathbf{0}_{\mathbb{W}}, 1/\delta)$ , o sea que se cumple 3. con  $M = 1/\delta$  como mayorante. Esto demuestra que 1.  $\implies$  3.

Si  $L(\overline{B}(\mathbf{0}_{\mathbb{V}}, 1)) \subseteq \overline{B}(\mathbf{0}_{\mathbb{W}}, M)$ , dado cualquier  $\varepsilon > 0$  en la fórmula (5) hacemos  $(r_1, r_2, c) = (1, M, \varepsilon/M)$  y resulta  $L(\overline{B}(\mathbf{0}_{\mathbb{V}}, \varepsilon/M)) \subseteq \overline{B}(\mathbf{0}_{\mathbb{W}}, \varepsilon)$ . Es decir que para todo  $\varepsilon > 0$  tenemos un  $\delta = \varepsilon/M$  tal que  $L$  envía la bola  $\overline{B}(\mathbf{0}_{\mathbb{V}}, \delta)$  dentro de la bola  $\overline{B}(\mathbf{0}_{\mathbb{W}}, \varepsilon)$ . Luego 3.  $\implies$  1.  $\square$

**Lema 19.** Sea  $L : \mathbb{V} \rightarrow \mathbb{W}$  lineal. Fijado  $M \geq 0$ , las condiciones siguientes son equivalentes:

1.  $\|v\| \leq 1 \implies \|L(v)\| \leq M$ , es decir que  $M$  es una mayorante para  $L(\overline{B}(\mathbf{0}, 1))$ .
2.  $\|\omega\| = 1 \implies \|L(\omega)\| \leq M$ .
3.  $\|L(v)\| \leq M \|v\|$  para todo  $v \in \mathbb{V}$ .

*Demostración.* Dado  $v \neq \mathbf{0}$ , tomamos su descomposición polar  $v = \|v\|\omega$  con  $\|\omega\| = 1$  y, suponiendo que se cumpla 2., deducimos (omitimos subíndices):

$$\|L(v)\| = \|L(\|v\|\omega)\| = \|\|v\|L(\omega)\| = \|v\|\|L(\omega)\| \leq \|v\|M.$$

Luego 2.  $\implies$  3., pues trivialmente  $\|L(\mathbf{0})\| = 0 = M\|\mathbf{0}\|$ . Como 1.  $\implies$  2. y 3.  $\implies$  1. son obvias, hemos terminado.  $\square$

El lema nos dice que si  $L$  es lineal acotada entonces los tres números siguientes son iguales:

- la mínima cota superior para  $\{\|L(v)\| : \|v\| \leq 1\}$ ,
- la mínima cota superior para  $\{\|L(\omega)\| : \|\omega\| = 1\}$ ,
- la mínima constante  $M \geq 0$  tal que  $\|L(v)\| \leq M\|v\|$  para todo  $v \in \mathbb{V}$ .

**Definición 20.** Dada  $L : (\mathbb{V}, \|\cdot\|_{\mathbb{V}}) \rightarrow (\mathbb{W}, \|\cdot\|_{\mathbb{W}})$  lineal acotada, su **norma de operador** (respecto de las normas de vectores  $\|\cdot\|_{\mathbb{V}}$  y  $\|\cdot\|_{\mathbb{W}}$ ) es el número:

$$\|L\| = \sup\{\|L(v)\|_{\mathbb{W}} : \|v\|_{\mathbb{V}} \leq 1\} = \sup\{\|L(\omega)\|_{\mathbb{W}} : \|\omega\|_{\mathbb{V}} = 1\}.$$

Se verifica (omitidos subíndices):

$$\|L(v)\| \leq \|L\| \|v\| \quad \text{para todo } v \in \mathbb{V}, \quad (6)$$

y no hay ninguna constante  $M$  menor que  $\|L\|$  que nos dé  $\|L(v)\| \leq M\|v\|$  para todo  $v \in \mathbb{V}$ .

**Proposición 21.** El conjunto  $\mathcal{L}(\mathbb{V}, \mathbb{W})$  es cerrado para la suma y producto por constante, por lo tanto un espacio vectorial. La norma de operador es, efectivamente, una norma en este espacio vectorial.

Dados tres espacios normados  $\mathbb{V}_1, \mathbb{V}_2, \mathbb{V}_3$  y aplicaciones  $\mathbb{V}_1 \xrightarrow{L_1} \mathbb{V}_2 \xrightarrow{L_2} \mathbb{V}_3$  lineales acotadas, la compuesta  $L_2 \circ L_1 : \mathbb{V}_1 \rightarrow \mathbb{V}_3$  es lineal acotada y se verifica:

$$\|L_2 \circ L_1\| \leq \|L_2\| \|L_1\|. \quad (7)$$

La demostración de esta proposición es fácil y se deja como ejercicio.

Supongamos elegidas normas en tres espacios numéricos:  $(\mathbb{R}^n, \|\cdot\|')$ ,  $(\mathbb{R}^m, \|\cdot\|'')$ ,  $(\mathbb{R}^k, \|\cdot\|''')$  y sean dadas matrices  $A \in M_{k \times m}(\mathbb{R})$ ,  $B \in M_{m \times n}(\mathbb{R})$ . Podemos verlas como operadores:

$$(\mathbb{R}^n, \|\cdot\|') \xrightarrow{B} (\mathbb{R}^m, \|\cdot\|'') \xrightarrow{A} (\mathbb{R}^k, \|\cdot\|'''),$$

con lo cual cada una tiene su norma de operador (respecto de las correspondientes normas de vectores). Además existe el producto  $AB$  y podemos verlo como un operador

$$AB : (\mathbb{R}^n, \|\cdot\|') \rightarrow (\mathbb{R}^k, \|\cdot\|''').$$

Entonces:

$$\|AB\| \leq \|A\| \|B\|. \quad (8)$$

Esta desigualdad es lo que hace útiles las normas de operador: la mayoría de las normas que podemos definir en los espacios de matrices no cumplen (8).

Si fijamos una norma  $\|\cdot\|$  en  $\mathbb{R}^n$  y consideramos las matrices  $n \times n$  como operadores

$$(\mathbb{R}^n, \|\cdot\|) \rightarrow (\mathbb{R}^n, \|\cdot\|),$$

utilizando la misma norma de vectores en salida y en llegada, queda definida una norma de operador para las matrices  $n \times n$  que cumple lo siguiente:

$$\|A_1 A_2 \cdots A_s\| \leq \|A_1\| \|A_2\| \cdots \|A_s\|,$$

y también cumple  $\|I_n\| = 1$ , lo que refuerza la idea de que es una norma especial.

Terminamos este apartado dando una fórmula explícita para la norma de operador  $\|A\|$  cuando  $A$  es una matriz *invertible*  $n \times n$  y ponemos la norma euclídea estándar en salida y en llegada:

$$A : (\mathbb{R}^n, \|\cdot\|_2) \rightarrow (\mathbb{R}^n, \|\cdot\|_2) .$$

Como  $A$  es invertible, la imagen  $A \cdot \overline{B}(\mathbf{0}, 1)$  de la bola unidad estándar es un elipsoide  $E \subset \mathbb{R}^n$ . La norma de operador  $\|A\|$  es el mínimo radio  $r$  tal que la bola estándar  $\overline{B}(\mathbf{0}, r)$  contiene a  $E$ . Este radio coincide con el valor del **semieje principal máximo** del elipsoide  $E$ . Hallemos, pues, los semiejes principales de  $E$ .

Como  $A$  es invertible existe la inversa  $B = A^{-1}$ . Entonces:

$$v \in E = A \cdot \overline{B}(\mathbf{0}, 1) \iff \|Bv\| \leq 1 \iff 1 \geq \|Bv\|^2 = (Bv)^t(Bv) = v^t(B^tB)v .$$

La desigualdad cuadrática  $v^t(B^tB)v \leq 1$  define al elipsoide  $E$ . La matriz  $B^tB$  es simétrica y definida positiva, por lo tanto existe una base ortonormal  $\{u_1, \dots, u_n\}$  de  $\mathbb{R}^n$  que diagonaliza  $B^tB$  con autovalores  $\mu_1, \dots, \mu_n$  reales positivos. Escribiendo el vector general  $v$  en esa base ortonormal, obtenemos:

$$v = y_1u_1 + \dots + y_nu_n \implies v^t(B^tB)v = \mu_1y_1^2 + \dots + \mu_ny_n^2 ,$$

de donde:

$$y_1u_1 + \dots + y_nu_n \in E \iff 1 \geq \mu_1y_1^2 + \dots + \mu_ny_n^2 = \left(\frac{y_1}{1/\sqrt{\mu_1}}\right)^2 + \dots + \left(\frac{y_n}{1/\sqrt{\mu_n}}\right)^2 .$$

Los semiejes principales de  $E$  son, pues,  $\sqrt{1/\mu_1}, \dots, \sqrt{1/\mu_n}$ .

Por otra parte  $B^tB$  es la inversa de la matriz  $AA^t$ , luego los autovalores de  $AA^t$  son:

$$\lambda_1 = 1/\mu_1, \dots, \lambda_n = 1/\mu_n ,$$

y los semiejes principales de  $E$  son  $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}$ . El mayor de ellos es la norma de operador  $\|A\|$ . Por último, la igualdad  $A^{-1}(AA^t)A = A^tA$  muestra que  $AA^t$  y  $A^tA$  tienen los mismos autovalores.

**Proposición 22.** *Si  $A$  es una matriz cuadrada invertible entonces su norma de operador, vista como  $A : (\mathbb{R}^n, \|\cdot\|_2) \rightarrow (\mathbb{R}^n, \|\cdot\|_2)$ , viene dada por  $\|A\| = \sqrt{\lambda}$  donde  $\lambda$  es, indistintamente, el máximo autovalor de  $AA^t$  o de  $A^tA$ .*

## 1.6 Abiertos y cerrados

**Definición 23.** *Sea  $(X, d)$  un espacio métrico. Un subconjunto  $U \subseteq X$  es un **abierto** (para la distancia  $d$ ) si se cumple:*

$$x \in U \implies \text{existe un } r > 0 \text{ tal que } B(x, r) \subseteq U .$$

*Dado un punto  $x_0 \in X$ , un **entorno** de  $x_0$  es cualquier abierto  $U$  tal que  $x_0 \in U$ .*

*Un subconjunto  $E \subseteq X$  es **cerrado** si su complemento  $X \setminus E$  es abierto. De manera equivalente, los cerrados son los complementos de los abiertos.*

Recordemos que en Matemáticas una implicación  $\mathcal{A} \implies \mathcal{B}$  se considera verdadera cuando la premisa  $\mathcal{A}$  es falsa. Por lo tanto al vacío  $\emptyset \subset X$  lo consideramos abierto.

**Proposición 24.** *La unión (finita o infinita) de abiertos produce un abierto.*

*La intersección finita de abiertos produce un abierto.*

*Como consecuencia, son cerrados: (1) la intersección finita o infinita de cerrados, (2) la unión finita de cerrados.*

*Demostración.* Sea  $(U_i)_{i \in I}$  una familia de abiertos de  $(X, d)$ , con el conjunto de índices  $I$  finito o infinito, y  $U = \bigcup_{i \in I} U_i$ . Por definición de unión, dado  $x \in U$  existe un  $i_0 \in I$  con  $x \in U_{i_0}$ . Al ser  $U_{i_0}$  un abierto, existe un  $r > 0$  tal que

$$B(x, r) \subseteq U_{i_0} \subseteq U,$$

y queda visto que  $U$  es un abierto.

Para demostrar que la intersección finita de abiertos es abierta, es suficiente demostrarlo para la intersección de dos. Sean, pues  $U_1, U_2$  abiertos en  $(X, d)$ . Si  $x \in U_1 \cap U_2$  entonces  $x \in U_1$  y  $x \in U_2$ , luego existen radios  $r_1, r_2 > 0$  con  $B(x, r_1) \subseteq U_1$  y  $B(x, r_2) \subseteq U_2$ . Si  $r = \min\{r_1, r_2\}$ , entonces:

$$r > 0, \quad B(x, r) \subseteq U_1, \quad B(x, r) \subseteq U_2,$$

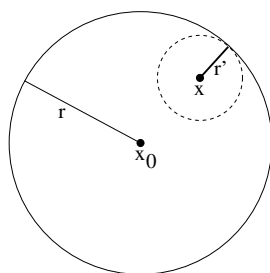
luego  $B(x, r) \subseteq U_1 \cap U_2$ . Esto prueba que  $U_1 \cap U_2$  es un abierto de  $(X, d)$  □

Como primer ejemplo, el vacío  $\emptyset$  y el total  $X$  son abiertos en todo espacio métrico  $(X, d)$ . También son cerrados, porque cada uno es el complementario del otro.

**Proposición 25.** *Las bolas abiertas  $B(x_0, r)$  son todos conjuntos abiertos.*

*En consecuencia, una bola abierta es entorno de todos sus puntos.*

*Demostración.* Dado  $x \in B(x_0, r)$ , se define un número positivo  $r' > 0$  por la igualdad  $r' = r - d(x, x_0)$ . El siguiente dibujo, que representa esa situación en el plano euclídeo, sirve para guiar nuestra intuición pero no es una demostración, ya que hay espacios métricos muy diferentes del plano euclídeo.



Para probar que, sea cual sea el espacio métrico, se cumple  $B(x, r') \subseteq B(x_0, r)$ , razonamos así:

$$y \in B(x, r') \iff d(x, y) < r' \implies d(x_0, y) \stackrel{*}{\leq} d(x_0, x) + d(x, y) < d(x_0, x) + r' = r,$$

donde se ha marcado con \* el lugar donde se usa la desigualdad triangular. Esto prueba que  $y \in B(x, r') \implies y \in B(x_0, r)$ , es decir  $B(x, r') \subseteq B(x_0, r)$ . Como  $r' > 0$  y  $x$  era un punto arbitrario de la bola  $B(x_0, r)$ , queda visto que dicha bola es un abierto.

Observa que la desigualdad triangular nos permite acotar distancias *por arriba*. □

**Proposición 26.** *Dado un espacio métrico  $(X, d)$ , un subconjunto  $U \subseteq X$  es un abierto si y sólo si es una unión de bolas abiertas.*

*Demostración.* Ya hemos visto que las bolas abiertas son abiertos y que la unión de abiertos es abierta. Sea ahora  $U$  un abierto cualquiera y para cada  $x \in U$  elijamos un  $r_x > 0$  tal que  $B(x, r_x) \subseteq U$ . Esto da lugar a una familia de bolas abiertas  $(B(x, r_x))_{x \in U}$  con  $U$  como conjunto de índices. Entonces:

$$U = \bigcup_{x \in U} \{x\} \subseteq \bigcup_{x \in U} B(x, r_x) \subseteq U,$$

donde la última inclusión se debe a que cada  $B(x, r_x)$  está contenida en  $U$ . Así  $U = \bigcup_{x \in U} B(x, r_x)$ . □

En la recta real el intervalo  $(a, b)$  es abierto porque es la bola abierta de centro  $(a + b)/2$  y radio  $(b - a)/2$ . Un intervalo  $(-\infty, b)$  es abierto porque es unión de intervalos del tipo anterior, lo mismo para  $(a, \infty)$ . El intervalo  $(-\infty, b]$  es cerrado porque su complemento  $(b, \infty)$  es abierto. Ese mismo intervalo  $E = (-\infty, b]$  no es abierto porque  $b \in E$  pero ninguna bola con centro en  $b$  está contenida en  $E$ . El intervalo  $[a, b]$  es cerrado porque su complemento  $(-\infty, a) \cup (b, \infty)$  es abierto. Siguiendo así, se llega a las siguientes conclusiones:

$(a, b), (-\infty, b), (a, \infty)$  abiertos, no cerrados.

$(-\infty, \infty) = \mathbb{R}$  abierto y cerrado.

$[a, a] = \{a\}, [a, b], (-\infty, b], [a, \infty)$  cerrados, no abiertos.

$(a, b), [a, b]$  ni abiertos ni cerrados.

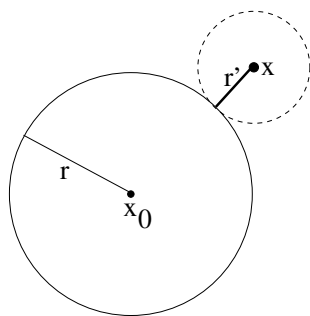
**Fíjate:** los subconjuntos de un espacio métrico no son “puertas”: pueden no ser ni abiertos ni cerrados. De hecho, *tanto los abiertos como los cerrados son subconjuntos muy especiales*: la inmensa mayoría de los subconjuntos no son de ninguno de esos dos tipos.

**Proposición 27.** Las bolas cerradas  $\overline{B}(x_0, r)$ , con  $r \geq 0$ , son subconjuntos cerrados.

*Demostración.* Hay que demostrar que el complemento

$$X \setminus \overline{B}(x_0, r) = \{x \in X : d(x_0, x) > r\},$$

es un abierto: para cada  $x$  con  $d(x, x_0) > r$  hay que hallar un  $r' > 0$  tal que la bola abierta  $B(x, r')$  esté toda ella contenida en  $X \setminus \overline{B}(x_0, r)$ . Dado un tal  $x$ , se define un número positivo  $r' > 0$  por la fórmula  $r' = d(x_0, x) - r$ . El siguiente dibujo muestra esa situación en el plano euclídeo:



y sugiere que, con la definición que hemos dado de  $r'$ , se tiene  $B(x, r') \subseteq X \setminus \overline{B}(x_0, r)$ . Insistimos en que el dibujo guía nuestra intuición pero no es una demostración.

Sea  $y \in B(x, r')$ . Razonamos de la manera siguiente, usando la simetría y la desigualdad triangular:

$$d(x, y) < r' \implies d(x_0, x) \leq d(x_0, y) + d(y, x) < d(x_0, y) + r' \implies r = d(x_0, x) - r' < d(x_0, y).$$

en definitiva:

$$d(x, y) < r' \implies d(x_0, y) > r,$$

es decir  $B(x, r') \subseteq X \setminus \overline{B}(x_0, r)$ , como se quería demostrar.

Observa que la desigualdad triangular nos permite acotar distancias *por abajo*. □

Sean  $(X, d)$  un espacio métrico e  $Y \subseteq X$  un subconjunto no vacío. La restricción  $d_Y$  es una distancia en  $Y$  que produce el nuevo espacio métrico  $(Y, d_Y)$ .

**Teorema-definición 28.** Un subconjunto  $V \subseteq Y$  es un **abierto relativo de  $Y$**  si es abierto en  $(Y, d_Y)$ . Esto equivale a  $V = Y \cap U$  para algún  $U$  abierto en  $(X, d)$ .

Un subconjunto  $F \subseteq Y$  es **cerrado relativo de  $Y$**  si es un cerrado de  $(Y, d_Y)$ . Esto equivale a  $F = Y \cap E$  para algún  $E$  cerrado en  $(X, d)$ .

La demostración se deja como ejercicio.

Por ejemplo, el conjunto  $A = (1, 2]$  no es abierto ni cerrado en  $\mathbb{R}$  pero sí es cerrado en  $Y_1 = (1, \infty)$  y sí es abierto en  $Y_2 = (-\infty, 2]$ .

Si conocemos la clase de los abiertos de  $X$  (y por lo tanto también los cerrados en  $X$ ) entonces ya conocemos los abiertos y cerrados relativos de  $Y$ .

## 1.7 Interior y cierre, caracterización mediante bolas

**Definiciones 29.** Sean  $(X, d)$  un espacio métrico y  $E \subseteq X$ .

Un punto  $x \in X$  es **interior a  $E$**  si existe un  $r > 0$  tal que  $B(x, r) \subseteq E$ . El conjunto de estos puntos se llama **interior de  $E$**  y se denota  $\text{int } E$ .

Un punto  $x \in X$  es **exterior a  $E$**  si es interior a  $X \setminus E$ : existe un  $r > 0$  con  $B(x, r) \cap E = \emptyset$ . Decimos que  $x \in X$  es un **punto adherente a  $E$**  si no es exterior a  $E$ , es decir si toda bola abierta centrada en  $x$  **corta a  $E$** : para todo  $r > 0$  tenemos  $B(x, r) \cap E \neq \emptyset$ . El conjunto de estos puntos se denota  $\overline{E}$  y se llama **adherencia o cierre de  $E$** .

Un punto  $x \in X$  es un **punto frontera de  $E$**  si no es ni interior ni exterior a  $E$ . El conjunto de estos puntos se llama **frontera topológica de  $E$**  y se denota  $\text{Fr } E$ .

Por supuesto, todo punto interior a  $E$  pertenece a  $E$ , es decir  $\text{int } E \subseteq E$ . También es obvio que todo  $x \in E$  es adherente a  $E$ , en definitiva:

$$\text{int } E \subseteq E \subseteq \overline{E},$$

y también  $\text{Fr } E = \overline{E} \setminus \text{int } E$  por la definición de frontera topológica.

Intuitivamente, que  $x$  sea exterior a  $E$  significa que no sólo no está en  $E$  sino que además está “un poco alejado” de  $E$ . Combinando las definiciones de exterior y adherencia, vemos inmediatamente que:

$$\overline{E} = X \setminus \text{int}(X \setminus E).$$

**Proposición 30.** El interior de  $E$  es el abierto más grande contenido en  $E$ . El cierre de  $E$  es el cerrado más pequeño que contiene a  $E$ .

Por lo tanto, un subconjunto  $E \subseteq X$  es cerrado si y sólo si  $\overline{E} = E$ .

La frontera topológica es siempre un cerrado.

*Demostración.* Empecemos por observar lo siguiente:

$$U \text{ abierto y } U \subseteq E \implies U \subseteq \text{int } E. \quad (9)$$

En efecto, para cada  $y \in U$  existe un  $r > 0$  tal que  $B(y, r) \subseteq U$ , luego  $B(y, r) \subseteq E$  y por lo tanto  $y \in \text{int } E$ . Como  $y$  era cualquier punto de  $U$ , queda visto que  $U \subseteq \text{int } E$ .

Dado  $x \in \text{int } E$ , se verifica  $B(x, r) \subseteq E$  para algún  $r > 0$ ; pero hemos visto que una bola abierta es un abierto, luego por (9) sabemos que  $B(x, r) \subseteq \text{int } E$ . Como  $x$  era cualquier punto de  $\text{int } E$ , hemos probado que  $\text{int } E$  es abierto. Pero (9) afirma que  $\text{int } E$  contiene a cualquier abierto contenido en  $E$ , luego es el mayor abierto que cabe dentro de  $E$ .

En particular  $\text{int}(X \setminus E)$  es un abierto, luego su complemento  $\overline{E}$  es un cerrado que, como ya hemos comentado, contiene a  $E$ .

Sea ahora  $C$  cualquier cerrado de  $(X, d)$  que contenga a  $E$ . El abierto  $X \setminus C$  está contenido en  $X \setminus E$ , luego contenido en  $\text{int}(X \setminus E)$ . Razonamos así:

$$X \setminus C \subseteq \text{int}(X \setminus E) \implies C \supseteq X \setminus \text{int}(X \setminus E) = \overline{E},$$

y así  $\overline{E} \subseteq C$ : el cierre  $\overline{E}$  es el cerrado más pequeño en el que cabe  $E$ .

La igualdad  $\text{Fr } E = \overline{E} \setminus \text{int } E = \overline{E} \cap (X \setminus \text{int } E)$  exhibe la frontera topológica como intersección de dos cerrados, luego  $\text{Fr } E$  es un cerrado sea cual sea el subconjunto  $E \subseteq X$ .  $\square$

**Corolario 31.** Una vez que conocemos la clase de los abiertos en  $(X, d)$  conocemos el interior y el cierre de cada subconjunto de  $X$ .

Estudiemos el caso de  $E = (a, b]$  en la recta real (con la distancia usual). Es fácil convencerse de que  $(a, b)$  es el abierto más grande de  $\mathbb{R}$  contenido en  $E$ , luego  $\text{int}(a, b] = (a, b)$ . También es fácil ver que  $[a, b]$  es el menor cerrado de  $\mathbb{R}$  que contiene a  $E$ , luego  $\overline{(a, b]} = [a, b]$ . Entonces la frontera topológica  $\text{Fr}(a, b] = [a, b] \setminus (a, b) = \{a\} \cup \{b\}$  está formada por los dos puntos que separan a  $E$  del resto de la recta real.



Sea  $(\mathbb{V}, \|\cdot\|)$  un espacio normado. Dejamos como ejercicio probar que para todo punto  $x \in \mathbb{V}$  y todo  $r > 0$  se tiene:

$$\text{int } B(x, r) = \text{int } \overline{B}(x, r) = \overline{B}(x, r) \quad , \quad \overline{B(x, r)} = \overline{\overline{B}(x, r)} = \overline{B}(x, r) ,$$

y por lo tanto la frontera topológica

$$\text{Fr } B(x, r) = \text{Fr } \overline{B}(x, r) = \{y : \|x - y\| = r\} ,$$

es la “cáscara” que separa la bola del resto del espacio.

**Comentarios.** 1. La frontera topológica es una “cáscara separadora” para *conjuntos muy sencillos en espacios sencillos*, como es una bola en un espacio normado. Pero para conjuntos o espacios métricos más complicados eso ya no es así. Un ejemplo es el conjunto  $\mathbb{Q}$  de los racionales en la recta real (con la distancia usual); enseguida se ve que:

$$\text{int } \mathbb{Q} = \emptyset \quad , \quad \overline{\mathbb{Q}} = \mathbb{R} ,$$

luego  $\text{Fr } \mathbb{Q} = \mathbb{R}$  sí que es un cerrado (tal como afirma la proposición 30) pero no tiene nada que ver con la idea de una cáscara separadora.

2. En la recta real  $\mathbb{R}$  consideramos dos distancias: la usual  $|x - y|$  y la  $d(x, y) = \min\{1, |x - y|\}$ . Utilizamos  $B$  o  $\overline{B}$  para denotar las bolas de la distancia usual y  $B_d$  o  $\overline{B}_d$  para las bolas de la distancia  $d$ . Se comprueba que ambas distancias definen los mismos conjuntos abiertos, y por lo tanto el mismo concepto de interior y el mismo concepto de cierre para subconjuntos de  $\mathbb{R}$ . Entonces el conjunto  $B(0, 1) = B_d(0, 1) = (-1, 1)$  tiene por cierre la bola  $\overline{B}(0, 1) = [-1, 1]$ , sin embargo  $\overline{B}_d(0, 1) = \mathbb{R}$  es mucho más grande que el cierre de  $B_d(0, 1)$ . Así el espacio métrico  $(\mathbb{R}, d)$  nos presenta un fenómeno que no ocurre ni en el plano euclídeo ni en ningún espacio normado. Nos enseña que los dibujos en el plano euclídeo, si bien pueden ser de gran ayuda (como en las demostraciones de las proposiciones 25 y 27), no son una “garantía”: si queremos mostrar que algo es verdad en *todos* los espacios métricos, es obligatorio elaborar una demostración basada sólo en las propiedades de las distancias.

Dado un espacio métrico  $(X, d)$  y un punto  $x \in X$ , el conjunto  $\{x\}$  es cerrado porque coincide con la bola cerrada  $\overline{B}(x, 0)$ . De hecho es un cerrado relativo de todo conjunto en el que esté  $x$ .

**Definición 32.** Sea  $E \subseteq X$  un subconjunto. Decimos que  $x \in E$  es un **punto aislado de  $E$**  si existe  $r_x > 0$  (dependiente de  $x$ ) tal que  $B(x, r_x) \cap E = \{x\}$ , es decir que el conjunto  $\{x\}$  es, además de cerrado, un abierto relativo de  $E$ .

Decimos que  $E$  es **discreto** si todos sus puntos son aislados.

El conjunto  $E = \{1/n : n = 1, 2, 3, \dots\}$  es discreto en la recta real. En cambio  $\overline{E} = E \cup \{0\}$  no es discreto porque 0 es un punto suyo que no es aislado en  $E$ . Vemos, de paso, que un conjunto discreto puede no ser cerrado.

Si  $(X, d)$  es un **espacio métrico discreto**, o sea que todos sus puntos son aislados en  $X$ , entonces cada conjunto de un solo elemento  $\{x\}$  es abierto y, como la unión arbitraria de abiertos es abierta, en este caso *todos* los subconjuntos de  $X$  son abiertos.

**Definición 33.** Un subconjunto  $E \subseteq X$  es **denso en  $X$**  si  $\overline{E} = X$ .

Por ejemplo, los conjuntos  $\mathbb{Q}$  y  $\mathbb{R} \setminus \mathbb{Q}$  son ambos densos en la recta real (con la distancia usual).

## 1.8 Sucesiones de puntos

**Definición 34.** Sea  $(X, d)$  un espacio métrico. Una sucesión  $(x_n)_{n=1}^{\infty} = (x_1, x_2, x_3, \dots)$  de puntos de  $X$  **converge al punto**  $x_0 \in X$ , lo cual indicamos escribiendo  $x_n \rightarrow x_0$ , si cada bola  $B(x_0, r)$  contiene una **cola**  $x_k, x_{k+1}, x_{k+2}, \dots$  de la sucesión.

Dicho con más precisión,  $(x_n)$  converge a  $x_0$  si y sólo si para cada  $r > 0$  existe un  $k$  (dependiente de  $r$  y la sucesión) tal que  $h \geq k \implies x_h \in B(x_0, r)$ .

**Proposición 35.** Si  $\overline{B}(x, r) \cap \overline{B}(x', r')$  no es vacía entonces  $d(x, x') \leq r + r'$ . Tenemos la desigualdad estricta  $d(x, x') < r + r'$  si una de las tres intersecciones siguientes es no vacía:

$$B(x, r) \cap B(x', r') \quad , \quad B(x, r) \cap \overline{B}(x', r') \quad , \quad \overline{B}(x, r) \cap B(x', r') .$$

La demostración se deja como ejercicio. Veamos, a partir de esta proposición, que **una sucesión no puede converger a dos puntos diferentes**  $x_0 \neq y_0$ . Hacemos  $r = d(x_0, y_0)/3 > 0$  y, como  $d(x_0, y_0) > r + r$ , se tiene  $B(x_0, r) \cap B(y_0, r) = \emptyset$ . Si cada bola contuviese una cola:

$$x_k, x_{k+1}, x_{k+2}, \dots \in B(x_0, r) \quad , \quad x_h, x_{h+1}, x_{h+2}, \dots \in B(y_0, r) ,$$

entonces para  $n = \max\{k, h\}$  el punto  $x_n$  tendría que estar en las dos bolas a la vez, imposible.

Concluimos que cada sucesión o bien no converge o lo hace a un único punto.

**Definición 36.** Una sucesión de puntos de  $(X, d)$  es **convergente en  $X$**  si existe un punto  $x_0 \in X$  al cual converge. Entonces  $x_0$  es único y se llama **límite** de la sucesión:  $\lim x_n = x_0$

La unicidad es, precisamente, lo que da importancia al concepto de límite. Muchos objetos importantes en Matemáticas (números, funciones, conjuntos) se construyen como límites y, gracias a la unicidad, quedan perfectamente definidos por tal construcción.

Es fácil ver que  $x_n \rightarrow x_0$  si y sólo si todo entorno de  $x_0$  contiene una cola de la sucesión  $(x_n)$ . Por lo tanto, la clase de los abiertos en  $(X, d)$  determina qué sucesiones convergen así como el límite de cada una de ellas.

Escribimos  $(x_n) \subset E$  para indicar que  $x_n \in E$  para todo  $n$ .

**Proposición 37.** Dado un subconjunto  $E \subseteq X$  la adherencia  $\overline{E}$  es el conjunto de los límites de sucesiones  $(x_n) \subset E$  convergentes en  $X$ . Por lo tanto  $E$  es cerrado si y sólo si contiene todos esos límites.

*Demostración.* Sea  $F$  el conjunto de esos límites. Dado  $x_0 \in F$ , hay una sucesión  $(x_n) \subset E$  convergente a  $x_0$ . Para cada  $r > 0$  la bola  $B(x_0, r)$  contiene una cola de la sucesión, luego contiene puntos de  $E$  y así  $B(x_0, r) \cap E \neq \emptyset$ . Esto prueba que  $x_0 \in \overline{E}$  y que  $F \subseteq \overline{E}$ .

Fijamos ahora un  $y_0 \in \overline{E}$ . Para cada entero positivo  $n$  tenemos  $B(y_0, 1/n) \cap E \neq \emptyset$  y elegimos un punto  $x_n$  en esta intersección, formando así una sucesión  $(x_n)_{n=1}^{\infty} \subset E$ . Dado ahora cualquier  $r > 0$ , hay un  $n$  tal que  $1/n < r$ , con lo cual:

$$m \geq n \implies d(x_m, y_0) < \frac{1}{m} \leq \frac{1}{n} < r \implies x_m \in B(y_0, r) ,$$

luego  $B(y_0, r)$  contiene la cola  $x_n, x_{n+1}, x_{n+2}, \dots$ . Así  $(x_n) \rightarrow y_0 \in F$ . Como  $y_0$  era un punto arbitrario de  $\overline{E}$ , tenemos  $\overline{E} \subseteq F$ .  $\square$

Podemos explicar ahora por qué los conjuntos cerrados se llaman así. Primero, cuando un conjunto tiene una operación (suma, producto, etc) se dice que un subconjunto  $E$  es *cerrado para esa operación* si al operar con elementos de  $E$  siempre resulta un elemento de  $E$ . Por ejemplo, en un grupo los subgrupos son los subconjuntos cerrados para la multiplicación  $(x, y) \mapsto xy$  y para la toma de inverso  $x \mapsto x^{-1}$ . Segundo, hay operaciones de más de dos argumentos, por ejemplo dados  $v_1, v_2, v_3 \in \mathbb{R}^3$  el determinante  $\det[v_1|v_2|v_3]$  es una operación que produce un número a partir de esos tres vectores. Tercero, hay operaciones que no se pueden efectuar con cualesquiera elementos, por ejemplo la división  $(x, y) \mapsto x/y$  sólo está definida en los pares  $(x, y)$  con  $y \neq 0$ . El *paso al límite* puede verse como una operación de infinitos argumentos

$$(x_1, x_2, x_3, \dots) \mapsto \lim x_n ,$$

que no está definida para todas las sucesiones (sólo para las convergentes) y decir que un subconjunto  $E$  es *cerrado para esta operación* es decir que  $E = \overline{E}$ . Por eso llamamos *subconjuntos cerrados* a los que cumplen  $E = \overline{E}$ . Coinciden con los complementarios de los abiertos.

## 1.9 Continuidad

El criterio más básico de si una aplicación  $f : (X, d_X) \rightarrow (Y, d_Y)$  es continua o no lo hemos dado en la definición 15 del apartado 1.4. Ahora daremos otros criterios que son de gran utilidad.

**Teorema 38.** *Para una aplicación  $f : (X, d_X) \rightarrow (Y, d_Y)$  son equivalentes:*

1.  $f$  es continua.
2. Para todo abierto  $V \subseteq Y$  la preimagen  $f^{-1}(V)$  es un abierto de  $X$ .
3. Para todo cerrado  $C \subseteq Y$  la preimagen  $f^{-1}(C)$  es un cerrado de  $X$ .
4. Para todo  $x_0 \in X$  y toda sucesión  $(x_n) \subset X$  con  $x_n \rightarrow x_0$  se tiene  $f(x_n) \rightarrow f(x_0)$ .

*Demostración.* Supongamos  $f$  continua y  $V \subseteq Y$  abierto. Dado  $x \in f^{-1}(V)$  es  $f(x) \in V$  y existe un  $\varepsilon > 0$  tal que  $B(f(x), \varepsilon) \subseteq V$ . Entonces existe un  $\delta > 0$  tal que

$$f(B(x, \delta)) \subseteq B(f(x), \varepsilon) \subseteq V,$$

luego  $B(x, \delta) \subseteq f^{-1}(V)$ . Esto prueba que  $f^{-1}(V)$  es abierto y que 1.  $\implies$  2.

Supongamos ahora que  $f$  satisface 2. Dados  $x \in X$  y  $\varepsilon > 0$ , el conjunto  $f^{-1}(B(f(x), \varepsilon))$  es un abierto de  $X$  al que pertenece  $x$ . Por lo tanto existe un  $\delta > 0$  tal que

$$B(x, \delta) \subseteq f^{-1}(B(f(x), \varepsilon)).$$

Aplicando  $f$  en ambos lados sale  $f(B(x, \delta)) \subseteq B(f(x), \varepsilon)$ . Luego  $f$  es continua y 2.  $\implies$  1.

Para todo  $A \subseteq Y$  se tiene  $f^{-1}(Y \setminus A) = X \setminus f^{-1}(A)$ . Si  $f$  satisface 2. y  $C$  es un cerrado de  $Y$ , entonces  $Y \setminus C$  es abierto de  $Y$  y su preimagen  $X \setminus f^{-1}(C)$  es abierto de  $X$ , luego  $f^{-1}(C)$  es un cerrado de  $X$ . Esto prueba que 2.  $\implies$  3. Análogamente 3.  $\implies$  2.

Sea ahora  $f$  continua y  $x_n \rightarrow x_0$  en  $X$ . Dado  $r > 0$  existe un  $\varepsilon > 0$  tal que  $f(B(x_0, \varepsilon)) \subseteq B(f(x_0), r)$ . A su vez la bola  $B(x_0, \varepsilon)$  contiene una cola  $x_n, x_{n+1}, x_{n+2}, \dots$  de la sucesión  $(x_n)$  y, aplicando  $f$ :

$$f(x_n), f(x_{n+1}), f(x_{n+2}), \dots \in f(B(x_0, \varepsilon)) \subseteq B(f(x_0), r),$$

es decir que para todo  $r > 0$  la bola  $B(f(x_0), r)$  contiene una cola de la sucesión  $(f(x_n))$ , lo cual equivale a  $f(x_n) \rightarrow f(x_0)$ . Esto prueba que 1.  $\implies$  4.

Supongamos, por último, que  $f$  satisface 4. pero hay un punto  $x_0 \in X$  en el que es discontinua. Entonces habrá un  $\varepsilon_0 > 0$  "malo", en el sentido de que para ningún  $\delta > 0$  estará la imagen  $f(B(x_0, \delta))$  contenida en  $B(f(x_0), \varepsilon_0)$ . En particular

$$f(B(x_0, 1/n)) \not\subseteq B(f(x_0), \varepsilon_0) \quad , \quad n = 1, 2, 3, \dots$$

Para cada entero positivo  $n$  podremos elegir un  $x_n \in B(x_0, 1/n)$  tal que  $f(x_n) \notin B(f(x_0), \varepsilon_0)$ . Estas elecciones formarán una sucesión  $(x_n) \subset X$  con  $x_n \rightarrow x_0$ , mientras que la bola  $B(f(x_0), \varepsilon_0)$  no contendrá ningún elemento de la sucesión  $(f(x_n))$ , con lo cual  $f(x_n) \not\rightarrow f(x_0)$  y  $f$  no satisfará 4., contradicción. Por reducción al absurdo, tal punto  $x_0$  no existe y  $f$  es continua. Esto prueba que 4.  $\implies$  1. □

**Corolario 39.** *Si conocemos la clase de los abiertos de  $X$  y la clase de los abiertos de  $Y$  entonces ya sabemos qué aplicaciones  $X \rightarrow Y$  son continuas y cuáles no.*

**Importante.** La condición 2. del teorema 38 dice que **una función continua es una máquina de generar abiertos por medio de desigualdades estrictas**. Concretamente, si  $f : X \rightarrow \mathbb{R}$  es continua entonces  $\{x : f(x) \neq 0\}$ ,  $\{x : f(x) > a\}$ ,  $\{x : f(x) < b\}$  y  $\{x : a < f(x) < b\}$  son abiertos de  $X$ . Análogamente la condición 3. dice que **una función continua es una máquina de generar cerrados por medio de ecuaciones o desigualdades no estrictas**: si  $f : X \rightarrow \mathbb{R}$  es continua entonces  $\{x : f(x) = 0\}$ ,  $\{x : f(x) \geq a\}$ ,  $\{x : f(x) \leq b\}$  y  $\{x : a \leq f(x) \leq b\}$  son cerrados de  $X$ .

**Proposición 40.** Dadas funciones escalares  $f_1, \dots, f_k : (X, d) \rightarrow \mathbb{R}$ , la correspondiente función vectorial

$$f : (X, d) \longrightarrow (\mathbb{R}^k, \|\cdot\|_\infty) \quad , \quad x \longmapsto (f_1(x), \dots, f_k(x)) \quad ,$$

es continua si y sólo si las  $f_1, \dots, f_k$  son todas continuas.

Una sucesión de vectores  $(x_n) \subset \mathbb{R}^2$  converge a  $x$  en  $(\mathbb{R}^k, \|\cdot\|_\infty)$  si y sólo si para  $i = 1, \dots, n$  la sucesión  $(x_n^i)_{n=1}^\infty$  de las  $i$ -ésimas coordenadas de los  $x_n$  converge a la  $i$ -ésima coordenada de  $x$ .

La demostración se deja como ejercicio.

Si en cada subconjunto de  $\mathbb{R}$  ponemos la distancia  $|x - y|$  y en cada subconjunto de  $\mathbb{R}^2$  ponemos la distancia  $\|x - y\|_\infty$ , entonces las siguientes funciones son todas continuas:

Suma:  $\mathbb{R}^2 \rightarrow \mathbb{R}$  ,  $(x, y) \mapsto x + y$ .

Multiplicación:  $\mathbb{R}^2 \rightarrow \mathbb{R}$  ,  $(x, y) \mapsto xy$ .

División:  $\mathbb{R} \times (\mathbb{R} \setminus \{0\}) \rightarrow \mathbb{R}$  ,  $(x, y) \mapsto \frac{x}{y}$ .

Directas:  $\mathbb{R} \rightarrow \mathbb{R}$  ,  $x \mapsto e^x = \exp(x)$ ,  $\text{sen } x$ ,  $\text{cos } x$ ,

Logaritmo:  $(0, +\infty) \rightarrow \mathbb{R}$  ,  $x \mapsto \log x$ .

Seno inversa:  $[-1, 1] \rightarrow [-\pi/2, \pi/2]$  ,  $x \mapsto \arcsen x$ .

Coseno inversa:  $[-1, 1] \rightarrow [0, \pi]$  ,  $x \mapsto \arccos x$ .

Raíz impar:  $\mathbb{R} \rightarrow \mathbb{R}$  ,  $x \mapsto \sqrt[n]{x}$ ,  $n = 3, 5, 7, \dots$

Raíz positiva:  $[0, +\infty) \rightarrow [0, +\infty)$  ,  $x \mapsto \sqrt[n]{x}$ ,  $n = 2, 3, 4, 5, \dots$

Valor absoluto:  $\mathbb{R} \rightarrow \mathbb{R}$  ,  $x \mapsto |x|$ .

A esa lista se podrían añadir muchas más. Combinando ahora la proposición 40 con la 16 del apartado 1.4 (la compuesta de continuas es continua), obtenemos:

Si  $f, g : (X, d) \rightarrow \mathbb{R}$  son continuas entonces  $f + g : (X, d) \rightarrow \mathbb{R}$  es continua.

Si  $f, g : (X, d) \rightarrow \mathbb{R}$  son continuas entonces  $fg : (X, d) \rightarrow \mathbb{R}$  es continua.

Si  $f, g : (X, d) \rightarrow \mathbb{R}$  son continuas y  $g$  nunca se anula entonces  $\frac{f}{g} : (X, d) \rightarrow \mathbb{R}$  es continua.

Si  $f : (X, d) \rightarrow \mathbb{R}$  es continua entonces  $e^f, \text{sen } f, \text{cos } f : (X, d) \rightarrow \mathbb{R}$  son continuas.

Si  $f : (X, d) \rightarrow \mathbb{R}$  es continua y siempre positiva entonces  $\log f : (X, d) \rightarrow \mathbb{R}$  es continua.

Si  $f, g : (X, d) \rightarrow \mathbb{R}$  son continuas y  $f$  siempre positiva entonces  $f^g = \exp(g \log f)$  es continua.

Si  $f : (X, d) \rightarrow [-1, 1]$  es continua entonces  $\arcsen f : (X, d) \rightarrow [-\pi/2, \pi/2]$  es continua.

Si  $f : (X, d) \rightarrow [-1, 1]$  es continua entonces  $\arccos f : (X, d) \rightarrow [0, \pi]$  es continua.

Si  $f : (X, d) \rightarrow \mathbb{R}$  es continua y  $n$  es impar entonces  $\sqrt[n]{f} : (X, d) \rightarrow \mathbb{R}$  es continua.

Si  $f : (X, d) \rightarrow [0, +\infty)$  es continua entonces  $\sqrt[n]{f} : (X, d) \rightarrow [0, +\infty)$  es continua.

Si  $f : (X, d) \rightarrow \mathbb{R}$  es continua entonces  $|f| : (X, d) \rightarrow \mathbb{R}$  es continua.

Haciendo combinaciones sucesivas de esos casos se consigue lo siguiente:

Sea  $\mathcal{A}(x)$  una *fórmula elemental* en las variables  $x = (x_1, \dots, x_n)$ , que en un conjunto  $E \subseteq \mathbb{R}^n$  no plantea ningún problema:

cuando  $x \in E$  entonces cada vez que en  $\mathcal{A}(x)$  hay un cociente el denominador es no nulo, cada vez que hay un logaritmo el logaritmando es estrictamente positivo, cada vez que hay una raíz de índice par el radicando es no negativo, etc.

Entonces  $x \mapsto \mathcal{A}(x)$  define una función  $E \rightarrow \mathbb{R}$  continua respecto de  $\|\cdot\|_\infty$ .

Lo mismo vale para una función vectorial  $E \rightarrow \mathbb{R}^m$  cada una de cuyas  $m$  componentes se define por una fórmula elemental sin problemas cuando  $x \in E$ .

**Aviso.** Aparte de las elementales, existen muchas más funciones continuas.

**Proposición 41. (Principio de prolongación de las identidades).** Sean  $f, g : (X, d_X) \rightarrow (Y, d_Y)$  aplicaciones continuas y  $E \subseteq X$  denso en  $X$ . Si  $f|_E \equiv g|_E$  entonces  $f \equiv g$ .

La prueba se deja como ejercicio.

## 1.10 Compacidad

**Definiciones 42.** Las subsucesiones de una sucesión  $(x_n)_{n=1}^{\infty}$  son el resultado de tomar una sucesión estrictamente creciente de enteros positivos  $n_1 < n_2 < n_3 < \dots$  y quedarse sólo con los correspondientes términos  $x_{n_1}, x_{n_2}, x_{n_3}, \dots$ . Obsérvese que  $n_k \geq k$  para todo  $k$ .

Dado un subconjunto  $K \subseteq X$  los **recubrimientos** de  $K$  son las familias  $(A_i)_{i \in I}$  de subconjuntos de  $X$  tales que  $K \subseteq \bigcup_{i \in I} A_i$ .

**Teorema-definición 43.** Sea  $(X, d)$  un espacio métrico. Dado un subconjunto  $K \subset X$ , consideremos las dos propiedades siguientes:

**Propiedad de sucesiones:** Toda sucesión  $(x_n) \subset K$  tiene una subsucesión  $(x_{n_k})$  convergente a algún punto de  $K$ .

**Propiedad de recubrimiento:** Todo recubrimiento  $(U_i)_{i \in I}$  de  $K$  por abiertos de  $X$  tiene una subfamilia finita  $U_{i_1}, \dots, U_{i_N}$  que también recubre  $K$ , es decir  $K \subseteq U_{i_1} \cup \dots \cup U_{i_N}$ .

Estas dos propiedades son equivalentes. Si se cumplen decimos que  $K$  es **compacto**.

En la demostración de este teorema necesitaremos el resultado siguiente.

**Lema 44.** Sea  $(U_i)_{i \in I}$  un recubrimiento de  $K$  por abiertos de  $X$ . Si  $K$  tiene la propiedad de sucesiones entonces existe un número  $\varepsilon > 0$  tal que para todo  $x \in K$  hay un índice  $i \in I$  (dependiente de  $x$ ) tal que  $B(x, \varepsilon) \subseteq U_i$ .

De un número  $\varepsilon > 0$  que cumpla eso decimos que es un **número de Lebesgue** del recubrimiento, propiedad que admite dos interpretaciones:

- Para cada “punto gordo”  $B(x, \varepsilon)$  con  $x \in K$  hay un abierto  $U_i$  que lo contiene.
- Los “abiertos adelgazados”  $U_i^\varepsilon = \{x \in U_i : B(x, \varepsilon) \subseteq U_i\}$  también recubren  $K$ .

*Demostración del lema.* Procedemos por reducción al absurdo. Supongamos que para cada entero positivo  $n$  el inverso  $1/n$  no es número de Lebesgue del recubrimiento dado, es decir que existe un  $x_n \in K$  tal que la bola  $B(x_n, 1/n)$  no está contenida en ninguno de los abiertos  $U_i$ . Esto genera una sucesión  $(x_n)$  de puntos de  $K$  con la siguiente propiedad:

$$\text{para ningún } n \text{ está la bola } B(x_n, 1/n) \text{ contenida en ninguno de los abiertos } U_i. \quad (10)$$

La propiedad de sucesiones dice que hay una subsucesión  $(x_{n_j})_{j=1}^{\infty}$  convergente a un punto  $x_0 \in K$ . Habrá un abierto  $U_{i_0}$  del recubrimiento con  $x_0 \in U_{i_0}$ . El abierto  $U_{i_0}$  contiene una bola  $B(x_0, r)$  con  $r > 0$ . Elegimos un  $n_0$  tal que  $\frac{1}{n_0} < \frac{r}{2}$  y un  $s$  tal que la “bola mitad”  $B(x_0, \frac{r}{2})$  contiene la cola  $(x_{n_j})_{j \geq s}$  de la subsucesión. Haciendo  $m = \max\{n_0, s\}$  tenemos:

$$\frac{1}{n_m} \leq \frac{1}{m} < \frac{r}{2}, \quad x_{n_m} \in B\left(x_0, \frac{r}{2}\right),$$

y entonces la desigualdad triangular nos permite concluir que

$$B\left(x_{n_m}, \frac{1}{n_m}\right) \subseteq B\left(x_0, \frac{r}{2} + \frac{r}{2}\right) = B(x_0, r) \subseteq U_{i_0},$$

en contradicción con la propiedad (10). Por reducción al absurdo, algún  $1/n$  tiene que ser número de Lebesgue del recubrimiento dado.  $\square$

*Demostración del teorema 43.* Primero partimos de un  $K$  con la propiedad de sucesiones y probamos que tiene la propiedad de recubrimiento. Lo hacemos en tres pasos.

**Primer paso.** Fijado un  $r > 0$  cualquiera, consideramos la siguiente propiedad para conjuntos finitos  $\{x_1, \dots, x_N\} \subseteq K$ :

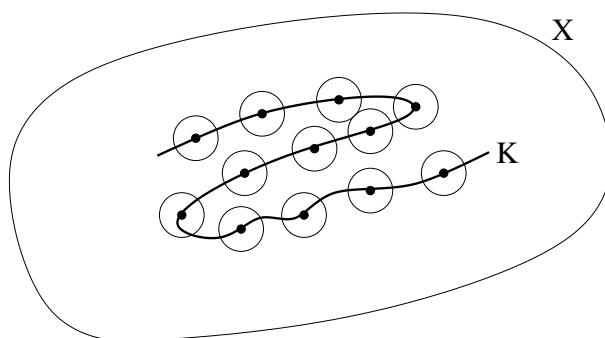
$\mathcal{P}$ : que para  $1 \leq i < j \leq N$  sea  $B(x_i, r) \cap B(x_j, r) = \emptyset$  (bolas en  $X$ ).

Afirmamos que hay uno maximal entre los conjuntos que tienen esta propiedad. Si no lo hubiera, podríamos construir una sucesión infinita  $\{x_1, x_2, x_3, \dots\}$  de puntos de  $K$  tal que las bolas  $B(x_n, r)$  fuesen disjuntas dos a dos en  $X$ . Esto implica  $d(x_n, x_m) \geq r$  siempre que  $n \neq m$  y, en consecuencia, *ninguna* subsucesión de  $(x_n)$  puede ser de Cauchy (ver el apartado 1.14) ni convergente. Pero hemos supuesto que  $K$  tiene la propiedad de sucesiones y, por reducción al absurdo, el conjunto maximal existe, es decir que hay un  $\{x_1, \dots, x_N\} \subseteq K$  tal que para todo  $x \in K$  la bola  $B(x, r)$  corta a alguna de las  $B(x_1, r), \dots, B(x_N, r)$ .

La idea intuitiva de la construcción del conjunto  $\{x_1, \dots, x_N\}$  es la siguiente. Al espacio  $X$  le vamos arracando bolas cumpliendo siempre dos reglas:

- Tienen que ser de radio  $r$  y estar centradas en puntos de  $K$ .
- No se permite arrancar un punto de  $X$  más de una vez, es decir que después de cada extracción de bolas el conjunto que queda debe contener *entera* la siguiente bola.

Hemos probado que este proceso no puede seguir indefinidamente; se detiene cuando ya no es posible arrancar una bola más cumpliendo esas dos reglas. En ese momento el conjunto  $Q = X \setminus (B(x_1, r) \cup \dots \cup B(x_N, r))$  es un “gruyère” con los agujeros tan juntos que en el queso situado entre ellos ya no cabe ninguna bola con centro en  $K$  y radio  $r$ .



**Advertencia.** Al haber hallado  $\{x_1, \dots, x_N\}$  por reducción al absurdo, que es un método no constructivo, no tenemos ningún control sobre el tamaño  $N$  de este conjunto. Más abajo damos un ejemplo de lo enorme que puede ser.

**Segundo paso.** Hacemos uso de un truco inventado por Giuseppe Vitali (1875-1932):

Las bolas de radio doble recubren todo  $K$ , es decir  $K \subseteq B(x_1, 2r) \cup \dots \cup B(x_N, 2r)$ .

Esto es así porque si hubiera un punto  $x_0 \in K$  fuera de todas esas bolas tendríamos  $d(x_0, x_i) \geq 2r$  para  $i = 1, \dots, N$  y, por la proposición 35 del apartado 1.8, la bola  $B(x_0, r)$  sería disjunta de cada una de las  $B(x_1, r), \dots, B(x_N, r)$  y el conjunto  $\{x_0, x_1, \dots, x_N\} \subseteq K$  tendría la propiedad  $\mathcal{P}$ , en contradicción con la maximalidad de  $\{x_1, \dots, x_N\}$ .

**Tercer paso.** Sea ahora un recubrimiento  $(U_j)_{j \in J}$  de  $K$  por abiertos de  $X$ . Tomamos un número de Lebesgue  $\varepsilon$  para él y hacemos  $r = \varepsilon/2$  en los pasos primero y segundo. Obtenemos un conjunto finito de bolas  $\{B(x_1, 2r), \dots, B(x_n, 2r)\} = \{B(x_1, \varepsilon), \dots, B(x_n, \varepsilon)\}$  centradas en puntos de  $K$  y que recubren todo  $K$ . Además, por ser  $\varepsilon$  un número de Lebesgue cada bola  $B(x_i, \varepsilon)$  está contenida en algún  $U_{j_i}$  del recubrimiento. Se deduce  $K \subseteq U_{j_1} \cup \dots \cup U_{j_n}$ . Luego  $K$  tiene la propiedad del recubrimiento.

Ahora partimos de un  $K$  que tiene la propiedad de recubrimiento y probamos que tiene también la de sucesiones. Llegaremos a una contradicción partiendo de suponer que hay una  $(x_n) \subset K$  sin subsucesiones convergentes en  $K$ .

**Paso auxiliar.** Fijada una tal  $(x_n)$ , demostremos (también por reducción al absurdo) que para cada punto  $y \in K$  existe un  $r_y > 0$  tal que el conjunto  $J_{y,r_y}$  es finito, siendo:

$$J_{y,r} \stackrel{\text{def}}{=} \{n \in \mathbb{N} : x_n \in B(y,r)\}.$$

Lo contrario sería que hubiese un  $y_0 \in K$  con el conjunto  $J_{y_0,r}$  infinito para todo  $r > 0$ . En particular serían infinitos  $J_{y_0,1}, J_{y_0,1/2}, J_{y_0,1/3}, \dots$  y estas infinidades nos permitirían elegir sucesivamente

$$n_1 \in J_{y_0,1} \quad , \quad n_2 \in J_{y_0,1/2} \quad \text{con} \quad n_2 > n_1 \quad , \quad n_3 \in J_{y_0,1/3} \quad \text{con} \quad n_3 > n_2 \quad , \quad \dots$$

y así  $n_1 < n_2 < n_3 < \dots$ . Entonces  $(x_{n_k})_{k=1}^\infty$  sería una subsucesión de  $(x_n)$  verificando  $x_{n_k} \in B(y_0, 1/k)$  para todo  $k$ , luego  $(x_{n_k}) \rightarrow y_0$  en contradicción con la hipótesis sobre la sucesión  $(x_n)$ . Por reducción al absurdo, para cada  $y \in K$  hay un radio  $r_y > 0$  con  $J_{y,r_y}$  finito.

Si existiera la sucesión  $(x_n) \subset K$  sin subsucesiones convergentes en  $K$ , el paso auxiliar produciría un recubrimiento  $(B(y,r_y))_{y \in K}$  de  $K$  por abiertos de  $X$ . Como estamos suponiendo que  $K$  tiene la propiedad del recubrimiento, habría una subfamilia finita  $B(y_1, r_{y_1}), \dots, B(y_N, r_{y_N})$  recubriendo todo  $K$  y por lo tanto recubriendo  $\{x_n, n \in \mathbb{N}\}$ . Esto implica que cada  $n \in \mathbb{N}$  pertenecería a algún  $J_{y_k, r_{y_k}}$ ,  $k = 1, \dots, N$ , y entonces  $\mathbb{N}$  se pondría como unión finita de conjuntos finitos, imposible. Por reducción al absurdo la sucesión  $(x_n)$  no puede existir y  $K$  tiene la propiedad de sucesiones.  $\square$

**El ejemplo.** Para  $(X, d) = (\mathbb{R}^n, \|\cdot\|_\infty)$  y  $K = \overline{B}(\mathbf{0}, 1)$ , el número máximo de bolas abiertas disjuntas dos a dos y de radio  $r = 1/k$  es  $N(r) = k^n = (1/r)^n$ , dependiente de  $r$  y de la dimensión. En  $\mathbb{R}^{333}$  es  $N(1/2) = 2^{333} > 10^{100}$  y  $N(1/4) > 10^{200}$ , o sea  $N(1/4)$  es como  $10^{100}$  veces  $N(1/2)$ .

Una consecuencia del teorema 43 es que una vez que conocemos los abiertos de  $(X, d)$  ya sabemos qué subconjuntos suyos son compactos y cuáles no.

**Teorema 45.** Si  $K \subseteq X$  es compacto entonces es cerrado y acotado en  $X$ .  
En  $(\mathbb{R}^n, \|\cdot\|_\infty)$  todo conjunto cerrado y acotado es compacto.

*Demostración.* Sean  $K \subseteq X$  compacto y  $(x_n) \subset K$  sucesión con  $x_n \rightarrow x \in X$ . Existe una subsucesión  $(x_{n_k})$  convergente a un punto  $y \in K$ . Como esta subsucesión también converge a  $x$ , tenemos  $x = y \in K$ . Es decir que  $K$  es cerrado para el paso al límite, luego es cerrado.

La familia de bolas  $(B(x, 1))_{x \in K}$  es un recubrimiento de  $K$  por abiertos de  $X$ . Tomamos un subrecubrimiento finito  $K \subseteq \overline{B}(x_1, 1) \cup \dots \cup \overline{B}(x_N, 1)$  y haciendo

$$r = 1 + \max\{d(x_1, x_2) \dots, d(x_1, x_N)\},$$

se llega a  $K \subseteq B(x_1, r)$ , luego  $K$  es acotado.

Sean ahora  $K \subset \mathbb{R}^n$ , cerrado y acotado, y  $(x_j)_{j=1}^\infty \subset K$ . Para  $i = 1, \dots, n$  la sucesión  $(x_j^i)_{j=1}^\infty$  de las  $i$ -ésimas coordenadas de los  $x_j$  es acotada en  $\mathbb{R}$ . La primera sucesión  $(x_j^1)$  tiene una subsucesión  $(x_{j_k}^1)$  convergente en  $\mathbb{R}$  a un número  $x^1$ . Dada ahora la sucesión  $(x_j^2)$  de las segundas coordenadas, consideramos la subsucesión  $(x_{j_k}^2)$  correspondiente a la misma sucesión de índices  $j_1 < j_2 < j_3 < \dots$  utilizada para definir  $(x_{j_k}^1)$ . Existe una sub-subsucesión  $(x_{j_{k_h}}^2)_{h=1}^\infty$  convergente en  $\mathbb{R}$  a un número  $x^2$  y la correspondiente sub-subsucesión  $(x_{j_{k_h}}^1)$  de las primeras coordenadas todavía converge a  $x^1$ . Siguiendo así, después de  $n$  pasos tenemos una sucesión creciente de índices  $m_1 < m_2 < m_3 < \dots$  y números  $x^1, \dots, x^n \in \mathbb{R}$  tales que para cada  $i = 1, \dots, n$  la sucesión de números  $(x_{m_s}^i)_{s=1}^\infty$  converge a  $x^i$ . Entonces la sucesión de vectores  $(x_{m_s})_{s=1}^\infty$  converge a  $x = (x^1, \dots, x^n)$  en  $(\mathbb{R}^n, \|\cdot\|_\infty)$  y además  $x \in K$  por ser  $K$  cerrado. Esto prueba que  $K$  tiene la propiedad de las sucesiones, luego es compacto.  $\square$

**Teorema 46.** Si  $f : X \rightarrow Y$  es continua y  $K \subseteq X$  es compacto, entonces  $f(K) \subseteq Y$  es compacto.

Si  $K \subseteq X$  es compacto y  $f : K \rightarrow \mathbb{R}$  es continua (para la distancia usual en  $\mathbb{R}$ ), entonces  $f$  alcanza su máximo y su mínimo. Es decir que existen  $x_1, x_2 \in K$  tales que  $f(x_1) \leq f(x) \leq f(x_2)$  para todo  $x \in K$ .

*Demostración.* La primera parte es por la caracterización de los compactos por sucesiones y la propiedad  $f(\lim x_n) = \lim f(x_n)$  de las aplicaciones continuas.

Para la segunda parte definimos el conjunto de números  $A = f(K) \subset \mathbb{R}$ , del que sabemos que es cerrado y acotado. Al ser acotado existen  $\inf A, \sup A$  y son valores finitos. Además son límites de sucesiones en  $A$  y, como  $A$  también es cerrado, resulta  $\inf A, \sup A \in A$ . Es decir que existen  $x_1, x_2 \in K$  tales que  $f(x_1) = \inf A$  y  $f(x_2) = \sup A$ , luego  $f(x_1) \leq f(x) \leq f(x_2)$  para todo  $x \in K$ .  $\square$

## 1.11 Normas equivalentes

**Teorema-definición 47.** Dos normas  $\|\cdot\|, \|\cdot\|'$  en un espacio vectorial  $\mathbb{V}$  son **equivalentes** si dan lugar a los mismos conjuntos abiertos en  $\mathbb{V}$ .

Esto resulta ser equivalente a la existencia de constantes  $c, C > 0$  tales que

$$c\|v\| \leq \|v\|' \leq C\|v\| \quad \text{para todo } v \in \mathbb{V}.$$

Por lo tanto, dos normas equivalentes también definen los mismos conjuntos acotados.

*Demostración.* Veamos que las dos condiciones siguientes son equivalentes:

1. Todo abierto para  $\|\cdot\|'$  es un abierto para  $\|\cdot\|$ .
2. Existe una  $C > 0$  tal que  $\|v\|' \leq C\|v\|$  para todo  $v \in \mathbb{V}$ .

Sea  $\text{id} : \mathbb{V} \rightarrow \mathbb{V}$  la identidad  $v \mapsto v$ . Expresando la condición 1. así:

Si  $U$  es un abierto en  $(\mathbb{V}, \|\cdot\|')$  entonces  $\text{id}^{-1}(U)$  es abierto en  $(\mathbb{V}, \|\cdot\|)$ , vemos que equivale a que  $\text{id}$  sea continua vista como  $(\mathbb{V}, \|\cdot\|) \rightarrow (\mathbb{V}, \|\cdot\|')$  y, por lo tanto, lineal acotada (véase el apartado 1.5) que es lo que expresa la condición 2.

Intercambiando  $\|\cdot\|$  con  $\|\cdot\|'$ , sabemos también que la existencia de una  $C' > 0$  tal que  $\|v\| \leq C'\|v\|'$  para todo  $v \in \mathbb{V}$  equivale a que todo abierto para  $\|\cdot\|$  sea un abierto para  $\|\cdot\|'$ . Poniendo  $c = 1/C'$ , esto último equivale a su vez a tener  $c\|v\| \leq \|v\|'$  para todo  $v \in \mathbb{V}$ .  $\square$

**Teorema-definición 48.** En  $\mathbb{R}^n$  todas las normas son equivalentes entre sí.

Llamamos **abiertos estándar** de  $\mathbb{R}^n$  a los definidos por cualquier norma.

Otra consecuencia útil es que las funciones y aplicaciones  $\mathbb{R}^n \supseteq E \xrightarrow{f} \mathbb{R}^m$  que hemos llamado *elementales* en el apartado 1.9 son continuas respecto de cualesquiera normas que pongamos en  $\mathbb{R}^n$  y en  $\mathbb{R}^m$ .

*Demostración del teorema 48.* Basta con probar que toda norma es equivalente a la  $\|\cdot\|_\infty$ . Denotaremos por  $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  la base estándar de  $\mathbb{R}^n$ . Para cualquier norma  $\|\cdot\|$  se tiene:

$$\|x\| = \|x_1\mathbf{e}_1 + \dots + x_n\mathbf{e}_n\| \leq |x_1|\|\mathbf{e}_1\| + \dots + |x_n|\|\mathbf{e}_n\| \leq C\|x\|_\infty,$$

donde hemos tomado  $C = \|\mathbf{e}_1\| + \dots + \|\mathbf{e}_n\| > 0$ . Esto implica que para cualesquiera  $x, y \in \mathbb{R}^n$  es:

$$\left| \|x\| - \|y\| \right| \leq \|x - y\| \leq C\|x - y\|_\infty,$$

luego si escribimos  $f(x) = \|x\|$  para todo  $x \in \mathbb{R}^n$  entonces:

$$\varepsilon > 0 \implies f(B(x, \varepsilon/C)) \subseteq (f(x) - \varepsilon, f(x) + \varepsilon),$$



donde  $B$  denota bolas abiertas respecto de  $\|\cdot\|_\infty$ . Esto nos dice que  $f \equiv \|\cdot\|$  es una función escalar continua respecto de  $\|\cdot\|_\infty$ .

La cáscara  $K = \{\omega : \|\omega\|_\infty = 1\}$  es un conjunto cerrado y acotado en  $(\mathbb{R}^n, \|\cdot\|_\infty)$ , luego compacto por el teorema 45. Por el teorema 46 existe un  $\omega_0 \in K$  tal que  $c = f(\omega_0)$  es el mínimo de  $f$  en  $K$ . Entonces  $c$  cumple las dos condiciones siguientes:

$$c > 0 \quad , \quad \|\omega\|_\infty = 1 \implies \|\omega\| \geq c .$$

Por el truco habitual de la descomposición polar, obtenemos  $\|x\| \geq c\|x\|_\infty$  para todo  $x \in \mathbb{R}^n$ . En definitiva  $c\|x\|_\infty \leq \|x\| \leq C\|x\|_\infty$  para todo  $x \in \mathbb{R}^n$ , como se quería demostrar.  $\square$

## 1.12 Conexión

**Definiciones 49.** Un camino en un espacio métrico  $(X, d)$  es cualquier aplicación continua  $\alpha : I \rightarrow X$  cuyo dominio  $I$  es un intervalo de la recta real.

Un subconjunto no vacío  $E \subseteq X$  es **conexo por caminos** o **conexo por arcos** si cada par de puntos  $p, q \in E$  se puede unir por un camino en  $E$ , es decir existe un camino  $\alpha(t) : [0, 1] \rightarrow E$  contenido en  $E$ , que empieza en  $p$  y termina en  $q$ :  $\alpha(0) = p, \alpha(1) = q$ .

**Proposición 50.** 1. Los subconjuntos conexos por caminos de la recta real son los intervalos y los de un elemento  $\{a\}$ .

2. La imagen de un conexo por caminos por una aplicación continua es conexas por caminos.

3. La unión no disjunta de dos conexos por caminos es conexas por caminos.

4. Si  $E$  es conexo por caminos entonces sólo  $\emptyset$  y  $E$  son a la vez abierto relativo y cerrado relativo de  $E$  (conceptos definidos al final del apartado 1.6).

Dado un punto  $p \in E$  el más grande subconjunto de  $E$  conexo por caminos y conteniendo a  $p$  es el conjunto:

$$\{q \in E : p \text{ se une con } q \text{ por un camino contenido en } E\} .$$

Estos conjuntos se llaman **componentes conexas por caminos** de  $E$ . La propiedad 3. de la proposición 50 hace que estas componentes formen una **partición**: su unión es todo  $E$  y dos cualesquiera que sean distintas son disjuntas.

**Proposición 51.** Si  $U$  es un abierto de  $\mathbb{R}^n$  entonces sus componentes conexas por caminos son abiertos y a lo más hay una cantidad numerable de ellas. De este modo todo abierto de  $\mathbb{R}^n$  tiene una única partición en abiertos conexos por caminos, en cantidad finita o numerable.

**Aviso.** Es importante no confundir las palabras “conexo” y “convexo”. Todo subconjunto convexo de  $\mathbb{R}^n$  es conexo por caminos (cualquier par de puntos suyos se une por un segmento rectilíneo que no se sale del conjunto) pero la mayoría de los conexos por caminos no son convexos (por ejemplo, un abierto del plano con la forma de la letra  $C$ ).

## 1.13 Continuidad uniforme

**Definición 52.** Una aplicación  $f : (X, d_X) \rightarrow (Y, d_Y)$  es **uniformemente continua** si

$$\text{Para cada } \varepsilon > 0 \text{ existe un } \delta > 0 \text{ tal que } d_X(x, x') \leq \delta \implies d_Y(f(x), f(x')) \leq \varepsilon .$$

Esta propiedad es estrictamente más fuerte que la continuidad. Por ejemplo, en el intervalo  $(0, 1]$  (acotado pero no compacto) la función escalar  $x \mapsto 1/x$  es continua pero no uniformemente continua: tómense  $x = 1/n$  y  $x' = 1/(n+1)$ .

**Teorema 53.** Si  $(X, d_X)$  es un espacio métrico compacto entonces toda aplicación continua  $f : (X, d_X) \rightarrow (Y, d_Y)$  es uniformemente continua.

*Demostración.* Fijamos un  $\varepsilon > 0$  y para cada  $x \in X$  definimos el conjunto:

$$U_x = \{x' \in X : d_Y(f(x), f(x')) < \varepsilon/2\}.$$

La familia que resulta  $(U_x)_{x \in X}$  es un recubrimiento de  $X$  por abiertos de  $X$ . Como  $X$  es compacto podemos aplicar el lema 44 y hay un número de Lebesgue  $\delta$  para ese recubrimiento. Sean ahora  $x_1, x_2$  puntos de  $X$  con  $d_X(x_1, x_2) < \delta$ . La bola  $B(x_1, \delta)$  contiene tanto a  $x_1$  como a  $x_2$  y, por ser  $\delta$  un número de Lebesgue, existe un  $x \in X$  tal que  $B(x_1, \delta) \subseteq U_x$ . En particular  $x_1, x_2 \in U_x$ , de donde:

$$d_Y(f(x), f(x_1)) < \frac{\varepsilon}{2}, \quad d_Y(f(x), f(x_2)) < \frac{\varepsilon}{2},$$

y un uso fácil de la simetría de  $d$  y la desigualdad triangular da:

$$d_Y(f(x_1), f(x_2)) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

## 1.14 Completitud

**Definición 54.** Sea  $(X, d)$  un espacio métrico. Decimos que  $(x_n) \subset X$  es una **sucesión de Cauchy** si para todo  $\varepsilon > 0$  existe  $N$  tal que  $n, m \geq N \implies d(x_m, x_n) < \varepsilon$ .

Es fácil ver que toda sucesión convergente es de Cauchy. El recíproco no es cierto en espacios métricos cualesquiera. Por ejemplo, es fácil dar sucesiones  $(x_n) \subset \mathbb{Q}$  convergentes a un número irracional, con lo cual son de Cauchy pero no son convergentes en  $\mathbb{Q}$ .

**Definiciones 55.** Un espacio métrico  $X$  es **completo** si toda sucesión  $(x_n) \subset X$  que sea de Cauchy es convergente en  $X$ . Un **espacio de Banach** es un espacio normado completo.

Todo cerrado  $C \subseteq \mathbb{R}^n$  es completo con la distancia inducida de  $\mathbb{R}^n$ . En cambio  $\mathbb{Q}$  no es completo.

## 1.15 Topologías

Sea  $d$  una distancia en un conjunto no vacío  $X$

La **topología** (con minúscula) **asociada a  $d$**  es la familia de los abiertos de  $(X, d)$ .

La **Topología** (con mayúscula) es una parte de las Matemáticas, estudia lo expuesto en este capítulo y más cosas.

**Definiciones 56.** Una **propiedad topológica** es la que puede decidirse cuando se conoce la topología del espacio involucrado, o las topologías de los espacios involucrados.

Una **construcción topológica** es la que se puede hacer una vez que se conoce la topología.

Son propiedades topológicas:

- ser cerrado en  $X$
- ser abierto o cerrado *relativo* en un subconjunto  $Y \subseteq X$ ,
- la continuidad de aplicaciones  $X \xrightarrow{f} Y$ ,
- la convergencia de sucesiones  $(x_j) \subset X$ ,
- la compacidad o no compacidad de cada subconjunto  $E \subseteq X$ ,
- ser conjunto discreto,
- ser conjunto denso.

Son construcciones topológicas

- el interior de cada  $E \subseteq X$ ,
- la adherencia de cada  $E \subseteq X$ ,
- el límite de cada sucesión convergente  $(x_j) \subset X$ .

Veamos, mediante un ejemplo, que las siguientes propiedades no son topológicas:

- ser conjunto acotado,
- continuidad uniforme,
- completitud.

Consideramos en  $X = (0, 1]$  las distancias  $d_{\text{estándar}}(x, y) = |x - y|$  y  $d(x, y) = |\log x - \log y|$ , que definen los mismos abiertos. El espacio métrico  $(X, d_{\text{estándar}})$  es acotado pero el  $(X, d)$  no. La función escalar  $x \mapsto \log x$  es uniformemente continua en  $(X, d)$  pero no lo es en  $(X, d_{\text{estándar}})$ . El espacio métrico  $(X, d)$  es completo mientras que  $(X, d_{\text{estándar}})$  no lo es.

Los tres ejemplos siguientes muestran propiedades que, en algún sentido que vamos a explicar, tienen menos calidad que la compacidad:

- La función  $f : (0, 1] \rightarrow \mathbb{R}$ ,  $f(x) = 1/x$ , con  $f((0, 1]) = [1, +\infty)$ , muestra que la imagen de un conjunto acotado por una función continua, puede que no sea acotada.
- La función  $g(x) = x^2$ , con  $g(\mathbb{R}) = [0, +\infty)$ , muestra que la imagen continua de un abierto puede que no sea abierta.
- La función  $h(x, y) = x$  y el cerrado  $F = \{(x, y) \in \mathbb{R}^2 : xy = 1\}$ , con  $h(F) = \mathbb{R} \setminus \{0\}$ , muestran que la imagen continua de un cerrado puede que no sea cerrada.

El teorema 46 dice que la compacidad es “robusta”: no desaparece al efectuar una transformación continua. En cambio acabamos de ver que la acotación, ser abierto y ser cerrado son propiedades que pueden desaparecer al efectuar una transformación continua y son, en este sentido, “de menos calidad” que la compacidad.