

Análisis Matemático, apuntes de clase

Universidad Autónoma de Madrid. Curso 2021-2022

Jesús Gonzalo Pérez

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 *no vacío* con alguna estructura”.

1.1 Espacios normados

La idea de un espacio normado es que se trata de un espacio vectorial en el que sabemos medir longitudes de vectores. Empezamos con el ejemplo básico.

1.1.1 Productos escalares

Definición 1. 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:

(Esc1). *Bilineal*: lineal en la variable v cuando se congela el valor de w y lineal en w cuando se congela el valor de v .

(Esc2). *Simétrica*: $\langle v, w \rangle = \langle w, v \rangle$.

(Esc3). *Definida positiva*: $\langle v, v \rangle > 0$ para todo $v \neq \mathbf{0}_{\mathbb{V}}$.

En el caso particular $\mathbb{V} = \mathbb{R}^n$, todos los productos escalares en \mathbb{R}^n vienen dados de la manera siguiente:

$$\langle x, y \rangle = x^t A y \quad , \quad \text{donde} \quad \begin{cases} x, y \text{ son vectores columna} \\ A \text{ es una matriz simétrica definida positiva} \end{cases}$$

Recordemos que una matriz simétrica A es **definida positiva** si cumple esto:

$$\text{para todo } x \in \mathbb{R}^n \setminus \{\mathbf{0}\} \quad x^t A x > 0 .$$

El **producto escalar estándar en \mathbb{R}^n** corresponde a la matriz identidad $A = I_n$, o sea

$$x \cdot y \stackrel{\text{def}}{=} x^t y = x_1 y_1 + \cdots + x_n y_n .$$

1.1.2 Normas euclídeas

Definición 2. La **longitud euclídea** o **norma euclídea** asociada al producto escalar $\langle \cdot, \cdot \rangle$ es la siguiente función:

$$\| \cdot \| : \mathbb{V} \rightarrow \mathbb{R} \quad , \quad \|v\| = \sqrt{\langle v, v \rangle} \geq 0 .$$

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

Para todo producto escalar las siguientes propiedades son obvias:

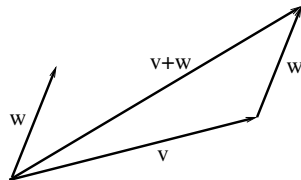
1. $\|v\| \geq 0$,
2. $\|v\| = 0 \iff v = \mathbf{0}$,
3. $\|\lambda v\| = |\lambda| \|v\|$,

y se les añaden otras dos, no tan obvias:

Desigualdad de Cauchy-Schwarz: $|\langle v, w \rangle| \leq \|v\| \|w\|$.

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, los tres vectores v , w y $v + w$ 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.



Demostración algebraica de la desigualdad de Cauchy-Schwarz

La **matriz de Gram** de una sucesión de vectores v_1, \dots, v_k es la “tabla de multiplicar”

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

obviamente simétrica $k \times k$. Veamos que tiene la siguiente propiedad:

$$v_1, \dots, v_k \text{ linealmente independientes} \implies G \text{ es definida positiva} \implies \det G > 0.$$

Definida positiva. Si $x \neq \mathbf{0}$, entonces $v = x_1 v_1 + \dots + x_k v_k \neq \mathbf{0}$, luego $0 < v^t v = x^t G x$.

Determinante positivo. Como G es real simétrica, hay una base ortonormal $\{u_1, \dots, u_k\}$ de \mathbb{R}^k formada por autovectores de G . Si $\lambda_1, \dots, \lambda_k$ son los correspondientes autovalores, entonces

$$\mathbb{R}^k \ni x = y_1 u_1 + \dots + y_k u_k \implies x^t G x = \lambda_1 y_1^2 + \dots + \lambda_k y_k^2,$$

y está claro que las desigualdades $x^t G x > 0$ para todo $x \neq \mathbf{0}$ fuerzan $\lambda_1, \dots, \lambda_k > 0$, de donde

$$\det G = \lambda_1 \cdots \lambda_k > 0.$$

Caso particular de dos vectores v, w :

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

Si v, w son linealmente independientes, es $\|v\|^2 \|w\|^2 - |\langle v, w \rangle|^2 > 0$ y $|\langle v, w \rangle| < \|v\| \|w\|$.

Si v, w son linealmente dependientes, es trivial ver que $|\langle v, w \rangle| = \|v\| \|w\|$.

Demostración geométrica de la desigualdad de Cauchy-Schwarz

Supuestos v, w linealmente independientes, les aplicamos el *proceso de Gram-Schmidt* y llegamos a una pareja ortonormal v_1, v_2 y números a_1, b, a_2 tales que

$$v = a_1 v_1, \quad w = b v_1 + a_2 v_2, \quad a_1, a_2 > 0, \quad b \text{ cualquiera}.$$

Entonces $|\langle v, w \rangle| = a_1 |b|$, y también:

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

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 .

Demostración algebraica de la desigualdad triangular

Como corolario de la desigualdad de Cauchy-Schwarz :

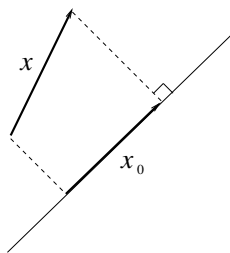
$$\langle v + w, v + w \rangle = \langle v, v \rangle + \langle w, w \rangle + 2\langle v, w \rangle \leq \|v\|^2 + \|w\|^2 + 2\|v\|\|w\| = (\|v\| + \|w\|)^2,$$

es decir $\|v + w\|^2 \leq (\|v\| + \|w\|)^2$. Tomando raíces cuadradas, sale $\|v + w\| \leq \|v\| + \|w\|$.

Demostración geométrica de la desigualdad triangular

Es obvia si $v + w = \mathbf{0}$. Supongamos $v + w \neq \mathbf{0}$.

Observación. Si x_0 es la *proyección ortogonal* de un vector x sobre cualquier subespacio vectorial, entonces $\|x_0\| \leq \|x\|$, es decir que *las proyecciones ortogonales acortan longitudes* (o las dejan igual).



Sean v_0, w_0 las proyecciones ortogonales respectivas de v, w sobre la recta vectorial $L = \langle v + w \rangle$, con lo cual $v + w = v_0 + w_0$.

Mirando sólo L , vemos que la longitud euclídea de $v_0 + w_0$ es uno de los tres números siguientes:

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

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

1.1.3 Homogeneidad de funciones

Definiciones 3. 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 $k \in \mathbb{Z}$, 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 homogéneos con $\mathbb{V} = \mathbb{R}^2$:

$$f(x, y) \equiv 5x^3y^2 - 9x^4y + y^5 \quad , \quad k = 5.$$

$$f(x, y) \equiv \frac{x}{x^4 + y^4} \quad , \quad k = -3.$$

$$f(x, y) = \sqrt[3]{x^3 + y^3} \quad , \quad k = 1.$$

Ejemplos positivamente homogéneos (pero no homogéneos):

$$f(x, y) = |x| + 3|y| \quad , \quad a = 1.$$

$$f(x, y) = \sqrt[6]{x^2 + 9y^2} \quad , \quad a = 1/3.$$

$$f(x, y) = |x| + \frac{y^3}{x^2 + y^2} \quad , \quad a = 1.$$

Cuando el grado es positivo (entero o fraccionario) extendemos estas funciones al origen poniendo $f(\mathbf{0}) = 0$. Entonces el **grafo** de f es el siguiente subconjunto de $\mathbb{V} \times \mathbb{R}$ (de \mathbb{R}^{n+1} cuando $\mathbb{V} = \mathbb{R}^n$):

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

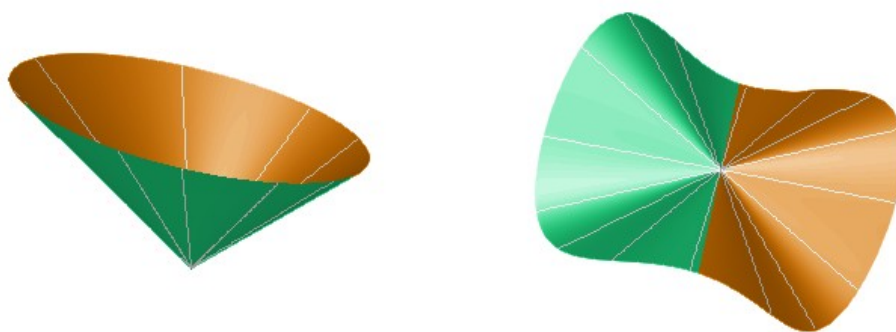
Que f sea positivamente homogénea de grado 1 significa que su grafo es un **cono**:

una unión de *semirrectas* en $\mathbb{V} \times \mathbb{R}$ que salen del origen $(\mathbf{0}, 0)$.

Si g es homogénea de grado 1, su grafo es un **cono simétrico respecto del origen**:

una unión de *rectas* pasando por $(\mathbf{0}, 0)$.

La siguiente figura muestra a la izquierda el grafo de $f(x, y) = 0.3x + \sqrt{x^2 + y^2}$ y a la derecha el grafo de $g(x, y) = (1/3)\sqrt[3]{x^3 - 3xy^2}$



1.1.4 Normas en general

Definiciones 4. 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:

(N1). $\|\mathbf{0}\| = 0$ y $\|v\| > 0$ cuando $v \neq \mathbf{0}$.

(N2.1). Positivamente homogénea de grado 1: $\lambda > 0 \implies \|\lambda v\| = \lambda \|v\|$.

(N2.2). Función par: $\|-v\| = \|v\|$.

(N3). Desigualdad triangular: $\|v + w\| \leq \|v\| + \|w\|$.

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

Lo que esperamos de una norma es que sirva para *medir longitudes* de vectores.

La propiedad (N1) responde a dos ideas:

- No queremos longitudes negativas.
- Un buen criterio para saber si un vector es nulo es mirar si su longitud es nula.

Las propiedades (N2.1) y (N2.2) suelen juntarse en la fórmula $\|\lambda v\| = |\lambda| \|v\|$, con lo que los axiomas para normas suelen darse así:

(N1). $\|\mathbf{0}\| = 0$ y $\|v\| > 0$ cuando $v \neq \mathbf{0}$.

(N2). $\|\lambda v\| = |\lambda| \|v\|$.

(N3). $\|v + w\| \leq \|v\| + \|w\|$.

1.1.5 Bolas y conjuntos acotados

Sea $(\mathbb{V}, \|\cdot\|)$ un espacio normado.

Dado $r > 0$ (desigualdad estricta), la **bola abierta de centro el origen y radio r** es el conjunto:

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

Dado $r \geq 0$, la **bola cerrada de centro el origen y radio r** es el conjunto:

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

En particular, la bola cerrada de radio nulo $\overline{B}(\mathbf{0}, 0)$ es el conjunto $\{\mathbf{0}\}$.

La **bola unidad abierta** es $B(\mathbf{0}, 1)$.

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

Definición 5. En un espacio normado \mathbb{V} , 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 $r \geq 0$.

1.1.6 Descomposición polar

Sea $N : \mathbb{B} \rightarrow \mathbb{R}$ una función que cumple los axiomas (N1) y (N2.1); en este apartado no entramos en si cumple o no (N2.2) o (N3). Cada vector no nulo $v \in \mathbb{V} \setminus \{\mathbf{0}\}$ tiene una factorización

$$v = \lambda a \omega \quad \text{con} \quad \begin{cases} a & \text{escalar positivo} \\ \omega & \text{vector “unitario”: } N(\omega) = 1 \end{cases}$$

Esta factorización resulta ser única, veámoslo. Se tiene que cumplir

$$N(v) = N(a\omega) = a N(\omega) = a \cdot 1,$$

luego $a = N(v)$ es el único valor posible para a . La única solución posible es, pues

$$a = N(v) \quad \text{y} \quad \omega = v/N(v),$$

que de hecho es solución, porque $N\left(\frac{v}{N(v)}\right) = N\left(\frac{1}{N(v)} v\right) = \frac{1}{N(v)} N(v) = 1$.

Observación. Una función $N(\cdot)$ que cumpla los axiomas (N1) y (N2.1) está completamente determinada por su “bola unidad” $\overline{B} = \{x : N(x) \leq 1\}$, veámoslo.

Sea $x \in \mathbb{V}$ un vector cualquiera. Si $x = \mathbf{0}$, ya sabemos que $N(x) = 0$. Si $x \neq \mathbf{0}$ y $x = a\omega$ es su descomposición polar, entonces para $\lambda > 0$ tenemos:

$$\lambda < a \implies \frac{x}{\lambda} = \frac{a}{\lambda} \omega \notin \overline{B}, \quad \text{pues } \frac{a}{\lambda} > 1,$$

$$\lambda \geq a \implies \frac{x}{\lambda} = \frac{a}{\lambda} \omega \in \overline{B}, \quad \text{pues } \frac{a}{\lambda} \leq 1,$$

es decir que el conjunto $\left\{\lambda > 0 : \frac{x}{\lambda} \in \overline{B}\right\}$ es el intervalo $[a, +\infty)$, quedando el número $N(x) = a$ determinado como el extremo inferior de este intervalo:

$$N(x) = \inf \left\{\lambda > 0 : x/\lambda \in \overline{B}\right\} = \min \left\{\lambda > 0 : x/\lambda \in \overline{B}\right\}.$$

Hay una manera más elegante (pero menos concreta) de decirlo:

$N(\cdot)$ es la *única* función positivamente homogénea de grado 1 con $x \in \overline{B} \iff N(x) \leq 1$.

1.1.7 Normas euclídeas y no euclídeas

Definiciones 6. Sea \mathbb{V} un espacio vectorial de dimensión finita. Un subconjunto $E \subset \mathbb{V}$ es un **elipsoide centrado en \mathbf{x}_0** si existen una forma cuadrática $Q : \mathbb{V} \rightarrow \mathbb{R}$ definida positiva y un $c > 0$ tales que

$$E = \{ y \in \mathbb{V} : Q(y - x_0) = c \} = x_0 + \{ x \in \mathbb{V} : Q(x) = c \} .$$

Un subconjunto $B \subset \mathbb{V}$ es un **elipsoide sólido centrado en \mathbf{x}_0** si existen una forma cuadrática $Q : \mathbb{V} \rightarrow \mathbb{R}$ definida positiva y un $c > 0$ tales que

$$B = \{ y \in \mathbb{V} : Q(y - x_0) \leq c \} = x_0 + \{ x \in \mathbb{V} : Q(x) \leq c \} .$$

Si $\|x\| \equiv \sqrt{\langle x, x \rangle}$ es una norma euclídea en un espacio vectorial \mathbb{V} , entonces las bolas cerradas de radio positivo

$$\overline{B}(\mathbf{0}, r) = \{ x \in \mathbb{V}, \langle x, x \rangle \leq r^2 \} ,$$

son elipsoides sólidos centrados en el origen $\mathbf{0}$, porque $Q(x) \equiv \langle x, x \rangle$ es una forma cuadrática definida positiva.

Recíproco: si la bola unidad es un elipsoide sólido (centrado en $\mathbf{0}$), entonces la norma es euclídea. Supongamos que la bola unidad de una norma $\|\cdot\|$ es un elipsoide sólido

$$\{ x : Q(x) \leq a \} = \left\{ x : \frac{1}{a} Q(x) \leq 1 \right\} .$$

Entonces la forma cuadrática $(1/a)Q$ procede de un producto escalar $\langle \cdot, \cdot \rangle$ en \mathbb{V} y la correspondiente norma euclídea $\|\cdot\|'$ tiene al elipsoide sólido $\{ x : Q(x) \leq a \}$ por bola unidad. Entonces $\|\cdot\| \equiv \|\cdot\|'$, porque la bola unidad determina la norma, y así $\|\cdot\|$ es euclídea.

Esta observación nos permite dar una infinidad de ejemplos de normas que no son euclídeas. En particular, es fácil comprobar que las siguientes fórmulas

$$\|(x, y)\|_1 \stackrel{\text{def}}{=} |x| + |y| \quad , \quad \|(x, y)\|_\infty \stackrel{\text{def}}{=} \max \{ |x|, |y| \} ,$$

definen normas en \mathbb{R}^2 cuyas bolas son cuadrados, no elipses; luego estas normas no son euclídeas.

Teorema 7. Una norma $\|\cdot\|$ es euclídea si y sólo si cumple la **regla del paralelogramo**:

$$2(\|x\|^2 + \|y\|^2) = \|x + y\|^2 + \|x - y\|^2 \quad , \quad \text{para cualesquiera } x, y \in \mathbb{V} .$$

□

Demostraremos el teorema 7 en un documento aparte.

Corolario 8. Una norma es euclídea si y sólo si el corte de la bola unidad con cada subespacio vectorial de dimensión 2 es una elipse sólida centrada en el origen.

Demostración. Si cumple esa condición de corte, entonces en cada subespacio bidimensional cumple la regla del paralelogramo. Pero entonces la cumple siempre, porque dicha regla involucra sólo cuatro vectores: $x, y, x+y, x-y$, que están metidos en algún subespacio con $\dim \leq 2$. □

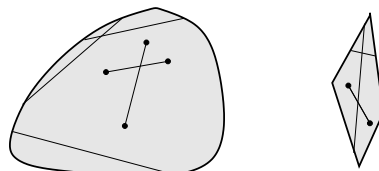
1.1.8 Conjuntos convexos

Definiciones 9. Dados $x, y \in \mathbb{V}$, el **segmento rectilíneo de extremos x, y** es el conjunto

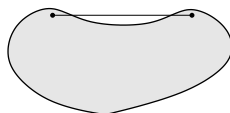
$$[x, y] \stackrel{\text{def}}{=} \{ (1 - \lambda)x + \lambda y : 0 \leq \lambda \leq 1 \} .$$

Un subconjunto no vacío $E \subseteq \mathbb{V}$ es **convexo** si siempre que $x, y \in E$ se tiene $[x, y] \subseteq E$.

La siguiente figura muestra dos conjuntos convexos en el plano \mathbb{R}^2 . También muestra los segmentos rectilíneos uniendo algunos pares de puntos.



Y ahora mostramos un subconjunto del plano que no es convexo, porque vemos que el segmento rectilíneo que une un par particular de puntos se sale del conjunto.



Intuitivamente, un conjunto convexo es el que no tiene ninguna “bahía”: visto desde afuera, es “saliente por todas partes”. Los subconjuntos convexos de la recta son los intervalos.

El concepto de conjunto convexo es muy importante en Análisis Matemático. En estos apuntes, aparece en los apartados 1.1.9, 1.1.10, 1.1.11, 1.1.12, 1.2.2, 1.2.9, 3.2, 3.3.2, 3.5 y 5.1.7.

1.1.9 Convexidad de funciones

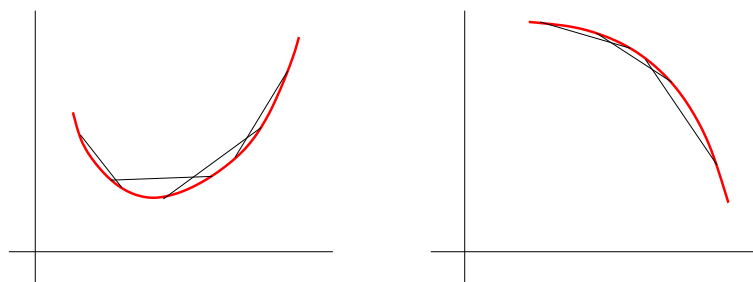
Definiciones 10. Sean $E \subseteq \mathbb{V}$ conjunto convexo y $f : E \rightarrow \mathbb{R}$. Decimos que f es una **función convexa** si cumple:

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

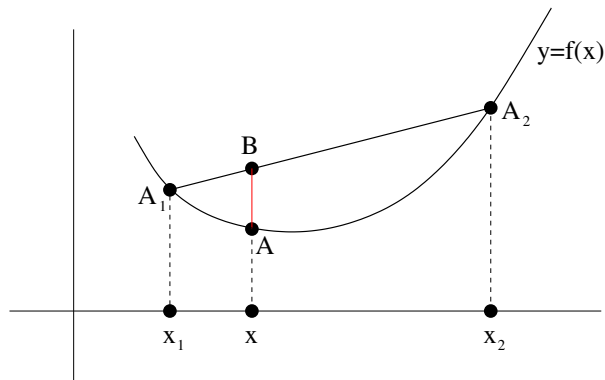
Análogamente decimos que f es una **función cóncava** si cumple:

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

Para una función convexa f de una variable, las **cuerdas** del grafo $\{(x, f(x)) : x \in I\}$ quedan por encima de él o forman parte de él. Para una función cóncava, las cuerdas del grafo quedan por debajo de él o forman parte de él.



Veamos, con una figura, que la condición de cuerdas por encima es equivalente a las desigualdades (1).



La función $f(x)$ está definida en un intervalo $I \subseteq \mathbb{R}$. Elegimos tres valores $x_1, x_2 \in I$ y $\lambda \in [0, 1]$. El valor $x = (1 - \lambda)x_1 + \lambda x_2$ es intermedio entre x_1 y x_2 . Los puntos correspondientes sobre el grafo son

$$A_1 = (x_1, f(x_1)) \quad , \quad A = (x, f(x)) \quad , \quad A_2 = (x_2, f(x_2)) \quad .$$

Con este mismo λ , definimos el punto $B = (1 - \lambda)A_1 + \lambda A_2 \in [A_1, A_2]$. El siguiente cálculo

$$B = (1 - \lambda)(x_1, f(x_1)) + \lambda(x_2, f(x_2)) = \left((1 - \lambda)x_1 + \lambda x_2, (1 - \lambda)f(x_1) + \lambda f(x_2) \right) ,$$

muestra que B tiene la misma abscisa que A :

$$B = (x, y(x)) \quad , \quad \text{con } y(x) = (1 - \lambda)f(x_1) + \lambda f(x_2) \quad .$$

Que el segmento rectilíneo $[A_1, A_2]$ esté por encima del grafo equivale a que para cada x , intermedio entre x_1 y x_2 , el punto $B = (x, y(x))$ esté por encima del punto $A = (x, f(x))$ (lo cual se indica en la figura por un segmento vertical rojo). Es decir, que para cualesquiera $x_1, x_2 \in I$ y $\lambda \in [0, 1]$ se tenga

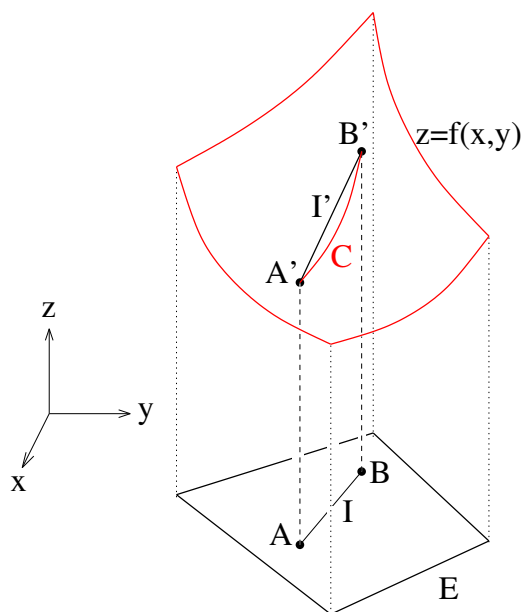
$$x = (1 - \lambda)x_1 + \lambda x_2 \implies f(x) \leq y(x) = (1 - \lambda)f(x_1) + \lambda f(x_2) ,$$

que es lo mismo que la desigualdad (1).

Teorema 11. Sean $I \subseteq \mathbb{R}$ un intervalo y $f : I \rightarrow \mathbb{R}$ una función derivable en todo punto de I . Si la derivada f' es monótona no decreciente, entonces f es convexa. Si f' es monótona no creciente, entonces f es cóncava. \square

Ejemplos: x^2 es convexa en $I = \mathbb{R}$, mientras que $\log x$ es cóncava en $I = (0, +\infty)$.

La siguiente figura muestra el grafo (en \mathbb{R}^3) de una función convexa $f(x, y)$ definida en una región convexa E del plano xy .



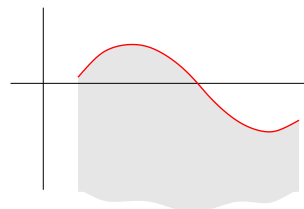
Sean $A, B \in E$, y sea $I \subseteq E$ el segmento rectilíneo $I = [A, B]$. Las verticales que pasan por I cortan a la superficie grafo $\{z = f(x, y)\}$ en una curva $C = \{(p, f(p)) : p \in I\}$. Si $A' = (A, f(A))$ y $B' = (B, f(B))$ son los extremos de C , entonces el segmento rectilíneo $I' = [A', B']$ queda por encima de la curva C , es decir que para todo $\lambda \in [0, 1]$ se tiene

$$(1 - \lambda)f(A) + \lambda f(B) \geq f((1 - \lambda)A + \lambda B) .$$

Por lo tanto, el segmento I' queda por encima de todo el grafo de f , ya que está en la vertical de la curva C . Podemos, pues, expresar la convexidad de $f(x, y)$ diciendo que, dado cualquier par A', B' de puntos del grafo, el segmento $[A', B']$ queda por encima de dicho grafo.

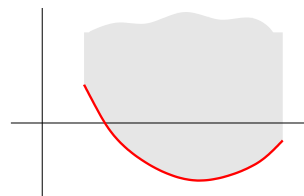
Definiciones 12. Sean un conjunto no vacío E y una función $f : E \rightarrow \mathbb{R}$. El **subgrafo** de f es el siguiente subconjunto de $E \times \mathbb{R}$:

$$\{(p, y) \in E \times \mathbb{R} : y \leq f(p)\}$$



El **supergrafo** de f es este otro conjunto:

$$\{(p, y) \in E \times \mathbb{R} : y \geq f(p)\}$$



Ejercicio 13. Sean $E \subseteq \mathbb{R}^n$ convexo y $f : E \rightarrow \mathbb{R}$. Demuestra que f es convexa si y sólo si el supergrafo de f es un conjunto convexo en \mathbb{R}^{n+1} . Demuestra que f es cóncava si y sólo si el subgrafo de f es un conjunto convexo en \mathbb{R}^{n+1} .

Ejercicio 14. Sean $E \subseteq \mathbb{R}^n$ convexo y $f : E \rightarrow \mathbb{R}$ convexa. Demuestra que para todo $c \in \mathbb{R}$ el subnivel $\{p \in E : f(p) \leq c\}$ es o vacío o un conjunto convexo en \mathbb{R}^n .

1.1.10 Teoría de Minkowski

Veamos ahora la estrecha relación que existe entre las normas y la convexidad.

Teorema 15. (Hermann Minkowski). Sea $N(\cdot)$ una función en \mathbb{V} que cumple los axiomas (N1) y (N2.1) de las normas. Entonces $N(\cdot)$ cumple la desigualdad triangular (N3) si y sólo si el conjunto $\overline{B} = \{x \in \mathbb{V} : N(x) \leq 1\}$ es convexo.

Demostración. Si $N(\cdot)$ cumple (N3), entonces es fácil comprobar que es una función convexa y, por el ejercicio 14, el conjunto \overline{B} es convexo.

Suponiendo ahora que $N(\cdot)$ cumple los axiomas (N1) y (N2.1) de las normas y que el conjunto \overline{B} es convexo, vamos a ver que

$$N(x + y) \leq N(x) + N(y) . \quad (2)$$

Esto último es obvio si $x = \mathbf{0}$ o si $y = \mathbf{0}$. Si $x \neq \mathbf{0} \neq y$, tomamos las descomposiciones polares

$$x = a\alpha \begin{cases} a = N(x) \\ N(\alpha) = 1 \end{cases} \quad y = b\beta \begin{cases} b = N(y) \\ N(\beta) = 1 \end{cases}$$

Entonces los números $\lambda = \frac{a}{a+b}$ y $\mu = \frac{b}{a+b}$ son positivos con $\lambda + \mu = 1$. Como α, β pertenecen al conjunto convexo \overline{B} , se tiene $\lambda\alpha + \mu\beta \in \overline{B}$, es decir,

$$1 \geq N(\lambda\alpha + \mu\beta) = N\left(\frac{a}{a+b}\alpha + \frac{b}{a+b}\beta\right) = N\left(\frac{a\alpha + b\beta}{a+b}\right) = N\left(\frac{x+y}{a+b}\right) = \frac{N(x+y)}{a+b} ,$$

en definitiva $N(x+y) \leq a+b$, que es la desigualdad (2). \square

¿Y el axioma (N2.2)? Éste equivale a que \overline{B} sea **simétrico respecto del origen**, es decir que se tenga $x \in \overline{B} \iff -x \in \overline{B}$.

Dar una norma en un espacio vectorial \mathbb{V} equivale a dar una convexo simétrico respecto del origen y que “contiene al origen en su interior”.

1.1.11 Las normas p en \mathbb{R}^n

Definición 16. Sea $1 \leq p < \infty$. La **norma p** en \mathbb{R}^n es la función $\|\cdot\|_p : \mathbb{R}^n \rightarrow \mathbb{R}$ definida de la siguiente manera:

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

Es evidente que estas funciones cumplen los axiomas (N1) y (N2). La desigualdad triangular para ellas

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p, \quad (3)$$

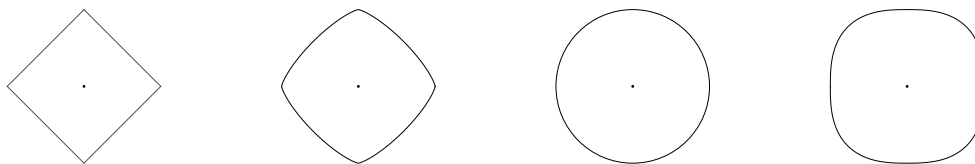
recibe el nombre de **desigualdad de Minkowski** y es consecuencia fácil del teorema 15. Como $p \geq 1$, la función $f(t) \equiv |t|^p$ es convexa. Se sigue (ejercicio) que la siguiente función también es convexa:

$$g(x_1, \dots, x_n) \equiv |x_1|^p + \dots + |x_n|^p \equiv \|x\|_p^p.$$

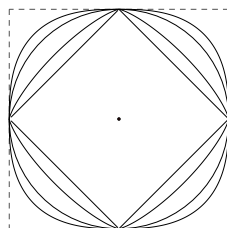
Entonces el conjunto $\{x \in \mathbb{R}^n : g(x) \leq 1\} = \{x \in \mathbb{R}^n : \|x\|_p \leq 1\}$ es convexo por el ejercicio 14 y, por el teorema 15, la función $\|\cdot\|_p$ cumple la desigualdad triangular.

En el apartado 1.1.14 puedes ver otra demostración de la desigualdad de Minkowski.

Mostramos a continuación dibujos de la bola unidad en el plano $\{v \in \mathbb{R}^2 : \|v\|_p \leq 1\}$ 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. Ahora mostramos esas mismas bolas superpuestas, y, en línea de trazos, la “forma límite” a la que tienden cuando $p \rightarrow \infty$:



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

Proposición 17. 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 $\|v_0\|_p$ 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:

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

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

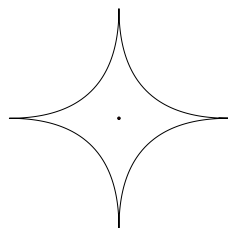
$$\|(x_1, \dots, x_n)\|_\infty \stackrel{\text{def}}{=} \max_{1 \leq j \leq n} |x_j|.$$

La bola unidad de $\|\cdot\|_\infty$ es $[-1, 1]^n \subset \mathbb{R}^n$. Cuando $n = 2$, es el cuadrado límite mostrado arriba.

Los dibujos que hemos mostrado en el plano son las intersecciones con el plano $x_1 x_2$ de las bolas unidad de la norma p en \mathbb{R}^n . Es fácil demostrar que no son elipses sólidas, excepto para $n = 2$, luego sólo la bola unidad de $\|\cdot\|_2$ es un elipsoide sólido en \mathbb{R}^n :

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

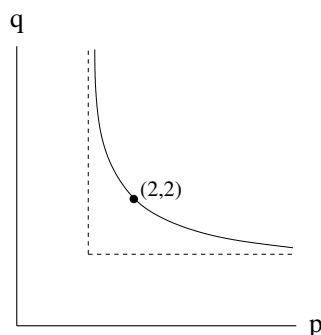
¿Por qué $p \geq 1$? Cuando $0 < p < 1$, la fórmula $\varphi_p(x_1, \dots, x_n) = \left[\sum_{j=1}^n |x_j|^p \right]^{1/p}$ todavía tiene sentido y la función $\varphi_p(x)$ así definida todavía cumple los axiomas (N1) y (N2) de las normas. Pero no cumple la desigualdad triangular. Por ejemplo, para $n = 2$ y $p = 1/2$ la “bola” $\{(x, y) : \varphi_{1/2}(x, y) \leq 1\}$ tiene la siguiente forma no convexa:



1.1.12 Desigualdades de Young y Hölder

Definición 18. Dos números $1 < p, q < \infty$ son **exponentes conjugados** si cumplen la ecuación $\frac{1}{p} + \frac{1}{q} = 1$.

La ecuación $(1/p) + (1/q) = 1$ se transforma fácilmente en $(p-1)(q-1) = 1$ y, por lo tanto, define una curva en el plano pq que es el resultado de trasladar una unidad hacia la derecha y una unidad hacia arriba la rama de hipérbola $\{pq = 1, p > 0, q > 0\}$.



La aplicación que lleva un exponente a su conjugado es una biyección decreciente del intervalo $(1, \infty)$ consigo mismo:

cuando p tiende a 1, su conjugado tiende a ∞ ; cuando p tiende a ∞ , su conjugado tiende a 1.

El exponente $p = 2$ es el único que es conjugado de sí mismo.

Teorema 19. *Dados exponentes conjugados p, q , se cumplen las desigualdades siguientes:*

$$\textbf{Desigualdad de Young: } a, b \geq 0 \implies ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

$$\textbf{Desigualdad de Hölder: } x, y \in \mathbb{R}^n \implies |x \cdot y| \leq \|x\|_p \|y\|_q.$$

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 Cauchy-Schwarz.

Demostración de la desigualdad de Young. Como $\frac{1}{p} + \frac{1}{q} = 1$ y $\log x$ es una función cóncava, para $a, b > 0$ tenemos:

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

y, tomando exponenciales, sale la desigualdad de Young. Es evidente para $a = 0$ o $b = 0$.

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 $y = \|y\|_q \beta$, con $\|\alpha\|_p = 1 = \|\beta\|_q$, lo que convierte la desigualdad de Hölder en $\|x\|_p \|y\|_q |\alpha \cdot \beta| \leq \|x\|_p \|y\|_q$. Basta, pues, con demostrar lo siguiente:

$$\|\alpha\|_p = 1 = \|\beta\|_q \implies |\alpha \cdot \beta| \leq 1. \quad (4)$$

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

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

lo que demuestra (4) y, por lo explicado, la desigualdad de Hölder en general. \square

1.1.13 Normas de operador

Cuando vemos \mathbb{R}^{mn} (matrices $m \times n$) como el espacio de las aplicaciones lineales de \mathbb{R}^n a \mathbb{R}^m , se distinguen en él unas normas especialmente útiles y que vamos a definir aquí.

Definición 20. Sean $(\mathbb{V}, \|\cdot\|_{\mathbb{V}}), (\mathbb{W}, \|\cdot\|_{\mathbb{W}})$ espacios normados. Para todo $r > 0$ definimos las bolas

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

Llamamos **aplicación lineal acotada de $(\mathbb{V}, \|\cdot\|_{\mathbb{V}})$ a $(\mathbb{W}, \|\cdot\|_{\mathbb{W}})$** a cualquier aplicación lineal $L : \mathbb{V} \rightarrow \mathbb{W}$ tal que el conjunto imagen $L(\overline{B}_{\mathbb{V}}(\mathbf{0}, 1))$ es acotado en $(\mathbb{W}, \|\cdot\|_{\mathbb{W}})$. Es decir, que existe una constante $M \geq 0$ tal que

$$L(\overline{B}_{\mathbb{V}}(\mathbf{0}, 1)) \subseteq \overline{B}_{\mathbb{W}}(\mathbf{0}, M). \quad (5)$$

El conjunto de las aplicaciones lineales acotadas 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 $\mathcal{L}(\mathbb{V}, \mathbb{W})$.

Consideremos la función:

$$\phi : \mathbb{V} \setminus \{\mathbf{0}\} \longrightarrow \mathbb{R}, \quad \phi(x) = \frac{\|L(x)\|_{\mathbb{W}}}{\|x\|_{\mathbb{V}}}, \quad (6)$$

que es homogénea de grado 0. Teniendo en cuenta esta homogeneidad, es fácil ver que la fórmula (5) equivale a que M sea una cota superior para ϕ :

$$\phi(x) \leq M \quad , \quad \text{para todo } x \in \mathbb{V} \setminus \{0\} \quad ,$$

y también equivale a que M sea una cota superior para $\phi(\omega)$ cuando $\|\omega\|_{\mathbb{V}} = 1$, vectores para los cuales $\phi(\omega) = \|L(\omega)\|_{\mathbb{W}}$.

Definición 21. 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 la mínima cota superior para el cociente ϕ definido en (6).

$$\|L\| \stackrel{\text{def}}{=} \sup_{x \neq 0} \frac{\|L(x)\|_{\mathbb{W}}}{\|x\|_{\mathbb{V}}} = \sup_{\|\omega\|_{\mathbb{V}}=1} \|L(\omega)\|_{\mathbb{W}} . \quad (7)$$

El número $\|L\|$ verifica (omitidos subíndices):

$$\|L(v)\| \leq \|L\| \|v\| \quad \text{para todo } v \in \mathbb{V} , \quad (8)$$

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

Proposición 22. El conjunto $\mathcal{L}(\mathbb{V}, \mathbb{W})$ es cerrado para la suma y producto por constante, por lo tanto es 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:

$$\|L_2 \circ L_1\| \leq \|L_2\| \|L_1\| . \quad (9)$$

Ejercicio 23. Demuestra esta proposición.

Definición 24. Sea $L : (\mathbb{V}, \|\cdot\|_{\mathbb{V}}) \rightarrow (\mathbb{W}, \|\cdot\|_{\mathbb{W}})$ una aplicación lineal entre espacios normados. Decimos que un vector $x_0 \in \mathbb{V} \setminus \{0\}$ es **maximal para L** si $\phi(x) \leq \phi(x_0)$ para todo $x \in \mathbb{V} \setminus \{0\}$, es decir que ϕ alcanza su máximo en x_0 .

Decimos que $x_0 \in \mathbb{V} \setminus \{0\}$ es **minimal para L** si $\phi(x) \geq \phi(x_0)$ para todo $x \in \mathbb{V} \setminus \{0\}$, es decir que ϕ alcanza su mínimo en x_0 .

El siguiente resultado afirma que en dimensión finita el supremo en la fórmula (7) siempre existe y es un máximo.

Lema 25. Si \mathbb{V} y \mathbb{W} tienen dimensiones finitas, entonces toda aplicación lineal $L : (\mathbb{V}, \|\cdot\|_{\mathbb{V}}) \rightarrow (\mathbb{W}, \|\cdot\|_{\mathbb{W}})$ tiene vectores maximales y vectores minimales. En este caso pues, toda aplicación lineal es acotada.

Demostraremos este lema en el apartado 1.2.8. De momento, lo vamos a utilizar.

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:

$$\|AB\| \leq \|A\| \|B\| . \quad (10)$$

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 (10).

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.

1.1.14 Norma dual de la norma p

El **espacio dual** \mathbb{V}^* de un espacio vectorial \mathbb{V} es el conjunto de las **formas lineales** (las funciones escalares lineales $\mathbb{V} \rightarrow \mathbb{R}$).

A cada vector $v \in \mathbb{R}^n$ le asociamos la forma lineal v^\flat (llamada “*v* bemol”) siguiente

$$v^\flat : \mathbb{R}^n \longrightarrow \mathbb{R} \quad , \quad v^\flat(x) = v^t x ,$$

de modo que si $v = (a_1, \dots, a_n)$ entonces $v^\flat(x_1, \dots, x_n) \equiv a_1 x_1 + \cdots + a_n x_n$. Es decir que las entradas a_j de v son los *coeficientes* de la forma lineal v^\flat . Esto define una biyección lineal

$$\mathbb{R}^n \longrightarrow (\mathbb{R}^n)^* \quad , \quad v \longmapsto v^\flat ,$$

que hace \mathbb{R}^n isomorfo a su dual.

En algún sentido, exponentes conjugados nos dan “normas duales”:

Proposición 26. *Fijamos $p > 1$. Cada forma lineal $\ell : \mathbb{R}^n \rightarrow \mathbb{R}$ es acotada como operador $\ell : (\mathbb{R}^n, \|\cdot\|_p) \rightarrow \mathbb{R}$. Más aún, si tomamos el único vector $v \in \mathbb{R}^n$ tal que $\ell \equiv v^\flat$ entonces*

$$\|v^\flat\| = \|v\|_q ,$$

siendo q el exponente conjugado de p .

Demostración. Es obvio cuando $v = \mathbf{0}$. Supongamos $v \neq \mathbf{0}$. La desigualdad de Hölder

$$|v^\flat(x)| \leq \|v\|_q \|x\|_p ,$$

nos dice que v^\flat es acotada como operador $(\mathbb{R}^n, \|\cdot\|_p) \rightarrow \mathbb{R}$, con norma de operador satisfaciendo la siguiente cota:

$$\|v^\flat\| \leq \|v\|_q . \tag{11}$$

Equivale a decir que $\|v\|_q$ es una cota superior para el cociente $\phi(x) = \frac{|v^\flat(x)|}{\|x\|_p}$ con $x \neq \mathbf{0}$.

Para reemplazar en (11) el signo de desigualdad por el de igualdad, vamos a construir explícitamente un vector $x_0 \neq \mathbf{0}$ tal que $\phi(x_0) = \|v\|_q$, con lo cual se prueba que $\|v\|_q$ es el máximo de ϕ y que x_0 es un vector maximal para v^\flat .

Definimos la **función signo** $\text{sig} : \mathbb{R} \rightarrow \mathbb{R}$ de la manera siguiente

$$\text{sig}(a) = \begin{cases} 1 & \text{si } a > 0 \\ 0 & \text{si } a = 0 \\ -1 & \text{si } a < 0 \end{cases}$$

de modo que $a = \text{sig}(a) |a|$ para todo $a \in \mathbb{R}$. Si $v = (a_1, \dots, a_n) \neq \mathbf{0}$, consideramos el vector

$$x_0 = (b_1, \dots, b_n) \quad , \quad \text{siendo } b_j = \text{sig}(a_j) |a_j|^{q-1} \quad \text{para } j = 1, \dots, n . \tag{12}$$

Entonces $a_j b_j = (\text{sig}(a_j))^2 |a_j| |a_j|^{q-1} = 1 \cdot |a_j|^q = |a_j|^q$, de donde

$$v^\flat(x_0) = \sum_{j=1}^n a_j b_j = \sum_{j=1}^n |a_j|^q = \|v\|_q^q = \|v\|_q (\|v\|_q)^{q-1} . \tag{13}$$

El que p, q sean conjugados equivale a la fórmula $p = \frac{q}{q-1}$, que con un cálculo fácil conduce a la igualdad $\|x_0\|_p = (\|v\|_q)^{q-1}$. Así (13) se convierte en $v^b(x_0) = \|v\|_q \|x_0\|_p$, de donde $\phi(x_0) = \|v\|_q = (\text{máximo de } \phi)$ y la fórmula (12) efectivamente define un vector maximal. \square

Otra demostración de la desigualdad de Minkowski. Dados $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_n)$, vamos a probar que $\|x + y\|_q \leq \|x\|_q + \|y\|_q$. Esto es trivial si $x + y = \mathbf{0}$, luego supondremos que $x + y \neq \mathbf{0}$. Definimos los siguientes vectores:

$$\begin{aligned} x' &= (|x_1|, \dots, |x_n|) & , & & y' &= (|y_1|, \dots, |y_n|) , \\ u &= (|x_1 + y_1|, \dots, |x_n + y_n|) & , & & z &= (|x_1 + y_1|^{q-1}, \dots, |x_n + y_n|^{q-1}) \end{aligned}$$

Observa que

$$\|x'\|_q = \|x\|_q \quad , \quad \|y'\|_q = \|y\|_q \quad , \quad \|u\|_q = \|x + y\|_q ,$$

y que z es vector maximal para el operador $u^b : (\mathbb{R}^n, \|\cdot\|_p) \rightarrow \mathbb{R}$.

Empezamos con la siguiente estimación:

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

aplicamos la desigualdad de Hölder:

$$x' \cdot z + y' \cdot z \leq \|x'\|_q \|z\|_p + \|y'\|_q \|z\|_p .$$

Por lo dicho sobre vectores definidos por la fórmula (12), es $\|z\|_p = \|u\|_q^{q-1} = \|x + y\|_q^{q-1}$. Llegamos a:

$$\|x + y\|_q^q \leq (\|x\|_q + \|y\|_q) \|x + y\|_q^{q-1} .$$

Dividimos por $\|x + y\|_q^{q-1}$ y ya está. \square

1.1.15 Elipsoides y norma de operador

Para A matriz invertible $n \times n$, y suponiendo que tanto en salida como en llegada tenemos la norma euclídea estándar

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

vamos a calcular la norma de operador $\|A\|$ explícitamente.

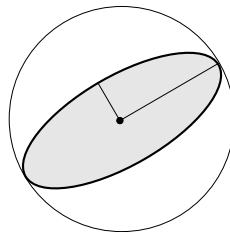
Lema 27. Cuando A es invertible, la imagen $A \cdot \overline{B}(\mathbf{0}, 1)$ de la bola unidad estándar es un elipsoide sólido (centrado en el origen).

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

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

Al ser B invertible, la matriz simétrica $B^t B$ es definida positiva. Por lo tanto, el conjunto $\{v : v^t(B^t B)v \leq 1\}$ es un elipsoide sólido, centrado en el origen, que coincide con $A \cdot \overline{B}(\mathbf{0}, 1)$. \square

La norma de operador $\|A\|$ es el mínimo radio M tal que la bola estándar $\overline{B}(\mathbf{0}, M)$ contiene al elipsoide sólido $E = A \cdot \overline{B}(\mathbf{0}, 1)$. Este radio coincide con el valor del **semieje principal máximo** de E .



Vamos a hallar los semiejes principales de E y a demostrar que, efectivamente, el mayor de ellos es igual a $\|A\|$.

La matriz $B^t B$ es simétrica y definida positiva, por lo tanto existe una base ortonormal $\{u_1, \dots, u_n\}$ de \mathbb{R}^n que diagonaliza $B^t B$ con autovalores reales positivos $0 < \mu_1 \leq \mu_2 \leq \dots \leq \mu_n$.

Escribiendo el vector general v en esa base ortonormal, obtenemos:

$$v = y_1 u_1 + \dots + y_n u_n \implies v^t (B^t B) v = \mu_1 y_1^2 + \dots + \mu_n y_n^2,$$

y la desigualdad $v^t B^t B v \leq 1$, que define al elipsoide E , puede escribirse como:

$$1 \geq \mu_1 y_1^2 + \dots + \mu_n y_n^2 = \left(\frac{y_1}{1/\sqrt{\mu_1}} \right)^2 + \dots + \left(\frac{y_n}{1/\sqrt{\mu_n}} \right)^2, \quad (14)$$

luego E es un elipsoide sólido cuyos semiejes principales miden

$$\sqrt{1/\mu_1} \geq \sqrt{1/\mu_2} \geq \dots \geq \sqrt{1/\mu_n},$$

con direcciones respectivas las de los vectores u_1, u_2, \dots, u_n . El semieje mayor mide $\sqrt{1/\mu_1}$ y el semieje menor mide $\sqrt{1/\mu_n}$.

La matriz $B^t B$ es la inversa de la matriz AA^t , luego los autovalores de AA^t son:

$$\lambda_1 = 1/\mu_1, \lambda_2 = 1/\mu_2, \dots, \lambda_n = 1/\mu_n,$$

con autovectores respectivos u_1, \dots, u_n . Así, pues, los semiejes principales de E tienen las medidas

$$\sqrt{\lambda_1} \geq \sqrt{\lambda_2} \geq \dots \geq \sqrt{\lambda_n}.$$

La matriz $A^t A$ tiene los mismos autovalores que la AA^t , aunque con autovectores diferentes:

$$A^t A (A^t u_j) = A^t (AA^t u_j) = A^t (\lambda_j u_j) = \lambda_j (A^t u_j),$$

pero los nuevos autovectores también son ortogonales dos a dos:

$$(A^t u_i) \cdot (A^t u_j) = (A^t u_i)^t A^t u_j = u_i^t AA^t u_j = u_i^t (\lambda_j u_j) = \lambda_j (u_i \cdot u_j) = 0 \quad \text{si } i \neq j,$$

y los vectores $v_j = (A^t u_j) / \|A^t u_j\|_2$ forman una base ortonormal $\{v_1, \dots, v_n\}$ de \mathbb{R}^n que diagonaliza $A^t A$ con autovalores respectivos $\lambda_1, \dots, \lambda_n$. Tenemos la siguiente fórmula:

$$v = z_1 v_1 + \dots + z_n v_n \implies \phi(v) = \frac{\|Av\|_2}{\|v\|_2} = \frac{\sqrt{v^t A^t A v}}{\sqrt{v^t v}} = \sqrt{\frac{\lambda_1 z_1^2 + \dots + \lambda_n z_n^2}{z_1^2 + \dots + z_n^2}},$$

que se simplifica a $\phi(\omega) = \sqrt{\lambda_1 z_1^2 + \dots + \lambda_n z_n^2}$ cuando $\|\omega\|_2 = 1$. De inmediato se deduce que v_1 es un vector maximal para A y que v_n es un vector minimal para A , siendo $\sqrt{\lambda_1}$ el máximo de la función ϕ y $\sqrt{\lambda_n}$ el mínimo, de donde:

$$\sqrt{\lambda_n} \|v\|_2 \leq \|Av\|_2 \leq \sqrt{\lambda_1} \|v\|_2, \quad \text{para todo } v \in \mathbb{R}^n. \quad (15)$$

Proposición 28. Si A es una matriz cuadrada invertible, entonces su norma como operador de $(\mathbb{R}^n, \|\cdot\|_2)$ a $(\mathbb{R}^n, \|\cdot\|_2)$ viene dada por $\|A\| = \sqrt{\lambda_1}$, donde λ_1 es, indistintamente, el máximo autovalor de AA^t o de $A^t A$. El valor $\sqrt{\lambda_1}$ es, asimismo, la longitud del semieje principal máximo del elipsoide sólido $E = A \cdot \overline{B}(\mathbf{0}, 1)$.

También se tiene:

$$\|Ax\|_2 \geq \sqrt{\lambda_n} \|x\|_2, \quad \text{para todo } x \in \mathbb{R}^n, \quad (16)$$

donde λ_n es, indistintamente, el mínimo autovalor de AA^t o de $A^t A$. Además $\sqrt{\lambda_n}$ es la longitud del semieje principal mínimo de E .

1.2 Espacios métricos y Topología

Las estructuras que vamos a definir aquí no sólo tienen sentido en espacios vectoriales, sino también en subconjuntos cualesquiera de un \mathbb{R}^n , que pueden no parecerse en nada a un espacio vectorial.

1.2.1 Funciones distancia

Definiciones 29. Sea X un conjunto no vacío. Una **función distancia** o **distancia** o **métrica** 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 siguientes condiciones para cualesquiera $x, y, z \in X$:

$$(\text{dist1}) \quad d(x, y) \geq 0 \quad y \quad d(x, y) = 0 \iff x = y.$$

$$(\text{dist2}) \quad d(x, y) = d(y, x).$$

$$(\text{dist3}) \quad d(x, z) \leq d(x, y) + d(y, z).$$

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

La propiedad (dist1) 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 (dist2) se llama *simetría*. La propiedad (dist3) se llama *desigualdad triangular (para distancias)*.

Decir que el espacio métrico es el par (X, d) significa, en particular, que si mantenemos el conjunto X pero cambiamos de una distancia a otra entonces tenemos *otro* espacio métrico.

Distancia procedente de una norma. Si $(\mathbb{V}, \|\cdot\|)$ es un espacio normado, podemos hallar la distancia entre dos puntos midiendo la longitud del vector que los une:

$$d(v, w) = \|v - w\| \quad , \quad \text{para cualesquiera } v, w \in \mathbb{V} . \quad (17)$$

Se comprueba de inmediato que esto, efectivamente, define una distancia. El axioma (dist1) es consecuencia del axioma (N1) para normas. La simetría (dist2) de la distancia es consecuencia de la simetría (N2.2) de la norma. La desigualdad triangular para la distancia (dist3) es consecuencia de la desigualdad triangular (N3) para la norma.

Observación. Los axiomas de distancia tienen sentido en *cualquier* conjunto (no vacío), con alguna estructura o sin ella. Esto lo hemos conseguido con la siguiente estrategia:

- Aunque el axioma (N3) para normas requiere que haya definida una operación de suma de vectores, al “convertirlo” en el (dist3) se elimina la necesidad de dicha suma.
- El axioma (N2.1), que requiere una operación de multiplicación de vector por número, sencillamente lo hemos “olvidado” por completo al definir el concepto de función distancia.

Por ejemplo, si $\|\cdot\|$ es una norma en \mathbb{R}^n y $E \subset \mathbb{R}^n$ es cualquier subconjunto no vacío (la mayoría de los cuales no se parecen en nada a un espacio vectorial) entonces tenemos una distancia d_E definida en E de la manera siguiente:

$$d_E(p, q) = \|p - q\| \quad , \quad \text{para } p, q \in E . \quad (18)$$

Si (X, d) es cualquier espacio métrico, podemos darle otra distancia d' definida así:

$$d'(x, y) = \min \{ d(x, y), 1 \} . \quad (19)$$

Incluso en el caso de un espacio vectorial \mathbb{V} , pueden darse en él muchas distancias que no vengan de ninguna norma. Damos un ejemplo de esto en el apartado 1.2.2.

1.2.2 Bolas métricas y conjuntos acotados

Definiciones 30. Sean (X, d) un espacio métrico y $x_0 \in X$ un punto suyo. La **bola abierta** con centro $x_0 \in X$ y radio $r > 0$ es el conjunto

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

La **bola cerrada** con centro x_0 y radio $r \geq 0$ es el conjunto

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

Cuando no hay duda de cuál es la función distancia, se omite el subíndice d .

Observación. Para la bola cerrada se permite el valor $r = 0$, siendo $\overline{B}(x_0, 0) = \{x_0\}$.

Definición 31. Un subconjunto $E \subseteq X$ es **acotado en (X, d)** si está contenido en alguna bola, es decir si $E \subseteq \overline{B}(x_0, r)$ para algún punto $x_0 \in X$ y algún radio $r \geq 0$.

A partir de la desigualdad triangular se deduce fácilmente que para todo $x \in X$ se tiene

$$\overline{B}(x_0, r) \subseteq \overline{B}(x, r + d(x, x_0)).$$

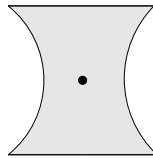
Esto implica que, una vez que el subconjunto E es acotado, podemos elegir *el punto que queramos* como centro de una bola conteniendo a E . Por supuesto, el radio de la bola contenedora dependerá del punto elegido como centro de dicha bola.

Supongamos que \mathbb{V} es un espacio vectorial y le damos una distancia d . Para que d venga de alguna norma según la fórmula (17), debe cumplir varias condiciones. En particular, las bolas (abiertas o cerradas) deben ser conjuntos convexos; veamos un ejemplo que no lo cumple.

Dada cualquier biyección $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ y cualquier norma $\|\cdot\|$ en \mathbb{R}^n , es trivial ver que la siguiente fórmula define una distancia

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

Tomando en el plano \mathbb{R}^2 la norma $\|\cdot\|_\infty$ y la biyección $\varphi(x, y) = (x/(1+y^2), y)$, queda definida una distancia cuyas bolas no son convexas sino que tienen la siguiente forma:



1.2.3 Convergencia de sucesiones

Definiciones 32. Dado un conjunto X , una **sucesión de elementos de X** es una aplicación

$$\{1, 2, 3, \dots\} \longrightarrow X, \quad n \longmapsto x_n,$$

del conjunto de los enteros positivos a X . Se la indica como una lista infinita x_1, x_2, x_3, \dots , también en la forma abreviada $\{x_n\}_{n=1}^\infty$ o simplemente $\{x_n\}$.

Dado un subconjunto $E \subseteq X$, escribimos $\{x_n\} \subset E$ para indicar que $x_n \in E$ para todo n .

Una **cola de la sucesión** es el resultado $x_k, x_{k+1}, x_{k+2}, \dots$ de suprimir los $k-1$ primeros términos para algún k .

La aplicación $\{1, 2, 3, \dots\} \rightarrow X$ no tiene por qué ser inyectiva, es decir que se permiten *repeticiones* en las sucesiones. Por ejemplo $1, 2, 1, 1, 3, 1, 1, 1, 4, 1, 1, 1, 1, 5, 1, 1, \dots$ es una sucesión de números perfectamente válida; algunos de sus valores son $x_3 = x_4 = 1, x_5 = 3, x_{12} = 1$, etc.

Definición 33. 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 en (X, d) 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 dicha 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 de la sucesión) tal que $m \geq k \implies x_m \in B(x_0, r)$.

Ejercicio 34. Si la intersección $B(x, r) \cap B(x', r')$ no es vacía, entonces $d(x, x') < r + r'$.

Proposición 35. Si una sucesión converge a dos puntos x_0, y_0 , entonces $x_0 = y_0$.

Demostración. Dado cualquier $\varepsilon > 0$ cada una de las bolas $B(x_0, \varepsilon)$ y $B(y_0, \varepsilon)$ contiene una cola de la sucesión, luego hay puntos en común a ambas bolas y, por el ejercicio 34, se tiene $d(x_0, y_0) < 2\varepsilon$. Al ser ε arbitrario, se deduce que $d(x_0, y_0) = 0$ y por lo tanto $x_0 = y_0$. \square

Definiciones 36. Una sucesión $\{x_n\}$, 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. Además de $\{x_n\} \rightarrow x_0$, también podemos escribir $\lim x_n = x_0$.

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.

Ejercicio 37. Para que una sucesión $\{x_n\}$ sea convergente, es necesario:

1. Que sea acotada.
2. Que sea una **sucesión de Cauchy**, es decir que para cada $\varepsilon > 0$ debe haber un k (dependiendo de ε) tal que $n, m \geq k \implies d(x_n, x_m) \leq \varepsilon$.

Definición 38. 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 .

Ejercicio 39. Si una sucesión $\{x_n\}_{n=1}^{\infty}$ converge al punto x_0 , entonces toda **subsucesión** suya $\{x_{n_j}\}_{j=1}^{\infty}$ converge también a x_0 .

Ejercicio 40. Una sucesión $\{x_j\} = \{(x_j^1, x_j^2, \dots, x_j^n)\}_{j=1}^{\infty} \subset \mathbb{R}^n$ converge en $(\mathbb{R}^n, \|\cdot\|_{\infty})$ al vector $x_0 \in \mathbb{R}^n$ si y sólo si para $i = 1, 2, \dots, n$ la sucesión de las i -ésimas coordenadas $\{x_j^i\}_{j=1}^{\infty}$ converge a la i -ésima coordenada de x_0 .

1.2.4 Abiertos, cerrados y topología

Definición 41. Sea (X, d) un espacio métrico. Un subconjunto $U \subseteq X$ es **abierto en (X, d)** si o bien $E = \emptyset$ o bien cumple lo siguiente

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

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

Dado un subconjunto $E \subseteq X$, un **entorno de E** es cualquier abierto U tal que $E \subseteq U$.

Debes pensar en los abiertos como *dominios en los que es cómodo estar*, porque si estamos en el punto $x \in U$ entonces podemos desplazarnos una distancia r en cualquier dirección sin salirnos de U . Esto nos facilitará el estudio de funciones definidas en dominios abiertos.

Definición 42. Un subconjunto $E \subseteq X$ es **cerrado en (X, d)** si o bien $E = \emptyset$ o bien para toda sucesión $\{x_n\} \subset E$ que converja a algún punto $x_0 \in X$ se tiene $x_0 \in E$.

Los subconjuntos cerrados se llaman así porque son cerrados para la *operación de paso al límite*, que lleva cada sucesión convergente a su punto límite

$$\{x_n\} \mapsto \lim_{n \rightarrow \infty} x_n.$$

En la recta real, con la distancia $d(x, y) = |x - y|$, el intervalo $(0, 1)$ es abierto y el intervalo $[0, 1]$ es cerrado. Por otra parte $[0, 1)$, \mathbb{Q} y $\mathbb{R} \setminus \mathbb{Q}$ no son ni abiertos ni cerrados. 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 ni abiertos ni cerrados.

Proposición 43. *Un subconjunto $C \subseteq X$ es cerrado si y sólo si su complemento $X \setminus C$ es abierto.*

Demostración. Lo que vamos a demostrar es que C no es cerrado si y sólo si $X \setminus C$ no es abierto.

Supongamos que $X \setminus C$ no es abierto. Por definición de abierto, eso significa que hay un $p \in X \setminus C$ tal que ninguna bola $B(p, \varepsilon)$ está contenida en $X \setminus C$. Esto último equivale a su vez a que toda bola $B(p, \varepsilon)$ contiene algún punto de C . Dando a ε los valores particulares $1, \frac{1}{2}, \frac{1}{3}, \dots$, obtenemos puntos $x_1, x_2, x_3, \dots \in C$ tales que para todo n es $d(p, x_n) < \frac{1}{n}$. Entonces $\{x_n\}_{n=1}^\infty$ es una sucesión contenida en C pero convergente a un punto p que no está en C , luego C no es cerrado.

Falta probar que si C no es cerrado entonces $X \setminus C$ no es abierto. La demostración consiste básicamente en recorrer el argumento anterior hacia atrás; los detalles se dejan como ejercicio. \square

Corolario 44. *Fijado un espacio métrico (X, d) , denotemos por \mathcal{A} la familia de los abiertos y por \mathcal{C} la de los cerrados. Entonces tenemos las siguientes biyecciones, inversas mutuas:*

$$\begin{array}{ccc} \mathcal{C} & \longrightarrow & \mathcal{A} \\ C & \longmapsto & X \setminus C \end{array} \qquad \begin{array}{ccc} \mathcal{A} & \longrightarrow & \mathcal{C} \\ U & \longmapsto & X \setminus U \end{array}$$

“Topología” (con mayúscula) es la parte de las Matemáticas que estudia las consecuencias de saber qué subconjuntos son los abiertos. Por su parte, “topología” (con minúscula) es la familia de los subconjuntos abiertos de un espacio.

1.2.5 Continuidad

Sean $(X, d), (X', d')$ dos espacios métricos cualesquiera y $f : X \rightarrow X'$ una aplicación.

Definiciones 45. *Dado un punto $x_0 \in X$, decimos que f es **continua en x_0** si para todo $\varepsilon > 0$ existe un $\delta > 0$ (dependiente de ε y de f) 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 .*

*Dado un subconjunto $Y \subseteq X'$, decimos que $f : X \rightarrow X'$ es **continua en Y** si es continua como aplicación $(Y, d_Y) \rightarrow (X', d')$, siendo $d_Y = d|_{Y \times Y}$. Dicho de otra manera, para $x_0 \in Y$ y B' bola centrada en $f(x_0)$ hay una bola B centrada en x_0 y tal que $f(B \cap Y) \subseteq B'$.*

Proposición 46. *Sean espacios métricos $(X, d), (X', d'), (X'', d'')$ y aplicaciones $X \xrightarrow{f} X' \xrightarrow{g} X''$. Si f es continua en $x_0 \in X$ y g es continua en $x'_0 = f(x_0) \in X'$, entonces la compuesta $g \circ f$ es continua en x_0 . Por lo tanto, si f y g son continuas entonces $g \circ f$ es continua.*

Demostración. Sea $\varepsilon_3 > 0$. Como g es continua en x'_0 , existe $\varepsilon_2 > 0$ tal que

$$g(B(x'_0, \varepsilon_2)) \subseteq B(g(x'_0), \varepsilon_3) = B((g \circ f)(x_0), \varepsilon_3).$$

Como f es continua en x_0 , a partir de ε_2 encontramos un $\varepsilon_1 > 0$ tal que $f(B(x_0, \varepsilon_1)) \subseteq B(x'_0, \varepsilon_2)$ y así:

$$(g \circ f)(B(x_0, \varepsilon_1)) = g(f(B(x_0, \varepsilon_1))) \subseteq g(B(x'_0, \varepsilon_2)) \subseteq B((g \circ f)(x_0), \varepsilon_3).$$

Como ε_3 era arbitrario, queda probada la continuidad de $g \circ f$ en x_0 . \square

Ejercicio 47. Sea $f : X \rightarrow X'$ continua. Demuestra que para todo subconjunto $Y \subseteq X$ la restricción $f|_Y : Y \rightarrow X'$ es continua.

Teorema 48. Para una aplicación $f : (X, d) \rightarrow (X', d')$, son equivalentes:

1. f es continua.
2. Para todo abierto $V \subseteq X'$ la preimagen $f^{-1}(V)$ es un abierto de X .
3. Para todo cerrado $C \subseteq X'$ 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 $V \subseteq X'$ 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 $A \subseteq X'$ se tiene $f^{-1}(X' \setminus A) = X \setminus f^{-1}(A)$. Si f satisface 2. y C es un cerrado de X' , entonces $X' \setminus C$ es abierto de X' 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 $\{f(x_n)\} \not\rightarrow f(x_0)$, en contradicción con 4. Por reducción al absurdo, tal punto x_0 no existe y f es continua. Esto prueba que 4. \implies 1. \square

Corolario 49. Si conocemos la clase de los abiertos de X y la clase de los abiertos de X' , entonces ya sabemos qué aplicaciones $X \rightarrow X'$ son continuas y cuáles no.

Ejercicio 50. 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.

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

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

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

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

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 la lista anterior con el hecho de que 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 $e^f, \operatorname{sen} f, \cos f : (X, d) \rightarrow \mathbb{R}$ 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 $\operatorname{arcsen} f : (X, d) \rightarrow [-\pi/2, \pi/2]$ es continua.

Si $f : (X, d) \rightarrow [-1, 1]$ es continua entonces $\operatorname{arccos} f : (X, d) \rightarrow [0, \pi]$ 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 $E \rightarrow \mathbb{R}^k$ cada una de cuyas k componentes se define por una fórmula elemental sin problemas cuando $x \in E$.

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

Importante. La condición 2. del teorema 48 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 . Las desigualdades $f(x) < g(x)$ se convierten en $(g - f)(x) > 0$ y se aplica lo dicho.

Análogamente, la condición 3. del teorema 48 dice que **una función continua es una máquina de generar cerrados por medio de ecuaciones o desigualdades no estrictas**. De este modo, 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 . Las desigualdades $f(x) \leq g(x)$ se convierten en $(g - f)(x) \geq 0$ y se aplica lo dicho.

Combinando eso con lo que hemos dicho sobre la continuidad de fórmulas elementales, ahora podemos decir de manera inmediata que un conjunto como el siguiente

$$\left\{ (x, y, z) \in \mathbb{R}^3 : \sin x + e^{3z} > 7, 5y < \frac{x^2 + 8xyz^4 - 10y^3}{x^2 + y^4 + e^z} < 3 \right\},$$

es intersección de dos abiertos y por lo tanto es abierto (ver el apartado 1.3.2), sin tener que pensarlo mucho. De los siguientes subconjuntos de \mathbb{R}^2

$$C = \{(x, y) : |x + y| \leq 4\}, \quad U = \{(x, y) : xy < 45\},$$

podemos decir de inmediato que C es cerrado y U es abierto.

1.2.6 Compacidad

Definición 51. 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 52. Sea (X, d) un espacio métrico. Dado un subconjunto $K \subseteq X$, las dos propiedades siguientes son equivalentes:

Propiedad de sucesiones, o de Bolzano-Weierstrass: Toda sucesión $\{x_n\} \subset K$ tiene una subsucesión $\{x_{n_k}\}$ convergente a algún punto de K .

Propiedad de recubrimiento, o de Heine-Borel: 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 que $K \subseteq U_{i_1} \cup \dots \cup U_{i_N}$.

Un **subconjunto compacto de X** es cualquier $K \subseteq X$ que tiene estas dos propiedades equivalentes.

Un **espacio métrico compacto** es cualquier espacio métrico (X, d) que es subconjunto compacto de sí mismo.

La equivalencia enunciada en este teorema se demuestra en otro documento, en el cual se define, además, el importante concepto de **número de Lebesgue de un recubrimiento**.

Una consecuencia del teorema 52 es que la clase de los abiertos en X , por sí sola, ya determina qué subconjuntos son compactos y cuáles no.

Teorema 53. Para que un subconjunto de un espacio métrico $K \subseteq X$ sea compacto es necesario que sea cerrado y acotado.

Si K es compacto, entonces todo cerrado contenido en K es también 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 $y \in K$. El ejercicio 39 nos dice que esta subsucesión también converge a x , luego $x = y \in K$ por la unicidad del límite (proposición 35). Esto prueba que K es cerrado.

Demostramos que K es acotado por reducción al absurdo: suponemos que no es acotado y deducimos una contradicción. Fijamos un punto $x_0 \in K$; como estamos suponiendo K no acotado, para cada entero positivo n hay un $x_n \in K$ tal que $d(x_0, x_n) > n$. Esto da lugar a una sucesión $\{x_n\}_{n=1}^\infty$ de puntos de K tal que toda subsucesión suya es no acotada, luego no convergente por el ejercicio 37. Esto contradice la compacidad de K .

Se deja como ejercicio la demostración de que, si K es compacto, entonces todo cerrado contenido en K es compacto. \square

En dimensión finita tenemos el recíproco de la primera parte del teorema 53.

Teorema 54. *En $(\mathbb{R}^n, \|\cdot\|_\infty)$ todo subconjunto cerrado y acotado es compacto.*

Demostración. Sea $K \subset \mathbb{R}^n$, cerrado y acotado, y $\{x_j\}_{j=1}^\infty \subset K$ una sucesión cualquiera de puntos de 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 sucesión $\{x_j^1\}$, de las primeras coordenadas, tiene una subsucesión $\{x_{j_k}^1\}$ convergente en \mathbb{R} a un número x_0^1 . Dada ahora la sucesión $\{x_j^2\}$ de las segundas coordenadas, consideramos la subsucesión $\{x_{j_k}^2\}$ con los mismos índices $j_1 < j_2 < j_3 < \dots$ utilizados 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_0^2 , y la correspondiente sub-subsucesión $\{x_{j_{k_h}}^1\}_{h=1}^\infty$ de las primeras coordenadas todavía converge a x_0^1 .

Siguiendo así, después de n pasos hay una sucesión creciente de índices $m_1 < m_2 < m_3 < \dots$ y números $x_0^1, \dots, x_0^n \in \mathbb{R}$ tales que

$$\begin{aligned} \{x_{m_s}^1\} &\rightarrow x_0^1 && \text{cuando } s \rightarrow \infty \\ \{x_{m_s}^2\} &\rightarrow x_0^2 && \text{cuando } s \rightarrow \infty \\ \{x_{m_s}^3\} &\rightarrow x_0^3 && \text{cuando } s \rightarrow \infty \\ &\vdots && \\ \{x_{m_s}^n\} &\rightarrow x_0^n && \text{cuando } s \rightarrow \infty \end{aligned}$$

Entonces la subsucesión de vectores $\{x_{m_s}\}_{s=1}^\infty$ converge al punto $x_0 = (x_0^1, \dots, x_0^n)$ en $(\mathbb{R}^n, \|\cdot\|_\infty)$ y además $x_0 \in K$ por ser K cerrado.

Esto prueba que K tiene la propiedad de las sucesiones, luego es compacto. \square

En un documento aparte demostramos que en dimensión infinita ocurre lo contrario:

Teorema 55. *Si $(V, \|\cdot\|)$ es un espacio normado de dimensión infinita, entonces ninguna de las bolas $\overline{B}(\mathbf{0}, r)$ es compacta.* \square

La propiedad de ser compacto es muy apreciada porque es *muy robusta*: no se pierde al aplicar transformaciones continuas.

Teorema 56. *Si $f : X \rightarrow Y$ es continua y $K \subseteq X$ es compacto, entonces $f(K) \subseteq Y$ es compacto.*

Demostración. Sea $\{y_n\} \subset f(K)$ cualquier sucesión de puntos de $f(K)$. Por definición del conjunto $f(K)$, para cada n existe un $x_n \in K$ con $y_n = f(x_n)$; esto da lugar a una sucesión $\{x_n\} \subset K$.

Como K es compacto, existe una subsucesión $\{x_{n_j}\}_{j=1}^\infty$ convergente a un punto $x_0 \in K$.

Como f es continua, se tiene $\{f(x_{n_j})\} \rightarrow f(x_0)$ cuando $j \rightarrow \infty$. Pero esto nos dice que la subsucesión $\{y_{n_j}\}$ converge al punto $y_0 = f(x_0) \in f(K)$. Luego $f(K)$ tiene la propiedad de las sucesiones y es compacto. \square

Por sí sola, la propiedad de ser cerrado no tiene esta robustez. Por ejemplo, la exponencial $e^x : \mathbb{R} \rightarrow \mathbb{R}$ es continua y lleva el cerrado \mathbb{R} al conjunto $(0, +\infty)$, que no es cerrado.

Tampoco la propiedad de ser acotado tiene esa robustez. Por ejemplo $\frac{1}{x} : (0, 1) \rightarrow \mathbb{R}$ es continua y lleva el conjunto acotado $(0, 1)$ al $(1, +\infty)$, que no es acotado.

Desde este punto de vista, es un excelente resultado juntar dos propiedades “no robustas” y obtener una propiedad que sí es robusta.

La compacidad también es muy apreciada porque nos da existencia de **puntos óptimos**: puntos donde una función toma su valor máximo o mínimo

Teorema 57. *Si $K \subseteq X$ es compacto y $f : K \rightarrow \mathbb{R}$ es continua, entonces f alcanza sus valores máximo y mínimo en K . 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. 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

1.2.7 Funciones Lipschitzianas

Definiciones 58. Dada una aplicación $f : (X_1, d_1) \rightarrow (X_2, d_2)$ entre espacios métricos, decimos que $M > 0$ es una **constante de Lipschitz para f** si se verifica:

$$d_2(f(p), f(q)) \leq M \cdot d_1(p, q) \quad , \quad \text{para cualesquiera } p, q \in X_1 .$$

Decimos que f es una **aplicación de Lipschitz**, o simplemente **lipschitziana**, si hay una constante de Lipschitz para ella.

De manera equivalente, la aplicación f es lipschitziana si el siguiente cociente es acotado

$$\frac{d_2(f(p), f(q))}{d_1(p, q)} \quad , \quad p \neq q ,$$

y las cotas superiores de este cociente son las constantes de Lipschitz para f .

Proposición 59. Toda aplicación de Lipschitz es continua.

Demostración. Sean $x_0 \in X_1$ y $\varepsilon > 0$ cualesquiera. Si M es una constante de Lipschitz para f , entonces

$$\begin{aligned} x \in B_{d_1}\left(x_0, \frac{\varepsilon}{M}\right) &\implies d_2(f(x), f(x_0)) \leq M \cdot d_1(x, x_0) < M \frac{\varepsilon}{M} = \varepsilon \implies \\ &\implies f(x) \in B_{d_2}(f(x_0), \varepsilon) , \end{aligned}$$

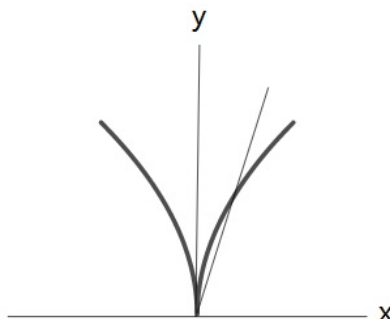
es decir que f envía $B_{d_1}(x_0, \varepsilon/M)$ dentro de $B_{d_2}(f(x_0), \varepsilon)$, luego es continua. \square

Dados un intervalo $I \subseteq \mathbb{R}$ y una función $f : I \rightarrow \mathbb{R}$, el cociente

$$\frac{|f(p) - f(q)|}{|p - q|} \quad , \quad p, q \in I \quad , \quad p \neq q ,$$

es el valor absoluto de la pendiente de la *recta secante a la gráfica de f* pasando por los puntos $(p, f(p))$ y $(q, f(q))$. En este caso, que f sea de Lipschitz significa que las pendientes de las secantes están acotadas (inferior y superiormente).

Entonces un ejemplo de una función continua pero no de Lipschitz es $f(x) = \sqrt{|x|} : \mathbb{R} \rightarrow \mathbb{R}$. El siguiente dibujo muestra en trazo grueso un trozo de su gráfica, y en trazo fino una secante pasando por el punto $(0, 0)$



Hay rectas secantes, pasando por $(0, 0)$, con pendiente tan grande como se quiera:

$$\frac{f(p) - f(0)}{p - 0} = \frac{\sqrt{|p|}}{p} = \frac{\pm 1}{\sqrt{|p|}} ,$$

luego f no es de Lipschitz, pese a ser continua.

Ser función de Lipschitz es estrictamente más fuerte que ser función continua.

Proposición 60. Dado un conjunto X , una función distancia d en X proporciona, para cada punto $p \in X$, una función escalar

$$f_p : X \longrightarrow \mathbb{R} \quad , \quad f_p(x) = d(p, x) \quad ,$$

y todas estas funciones tienen a 1 por constante de Lipschitz.

Demostración. Dados $p, q \in X$, aplicamos la desigualdad triangular de dos maneras diferentes:

$$\begin{aligned} d(p, x) &\leq d(p, y) + d(y, x) = d(p, y) + d(x, y) \implies d(p, x) - d(p, y) \leq d(x, y) \quad , \\ d(p, y) &\leq d(p, x) + d(x, y) \implies d(p, y) - d(p, x) \leq d(x, y) \quad , \end{aligned}$$

y deducimos $|d(p, x) - d(p, y)| \leq d(x, y)$, es decir $|f_p(x) - f_p(y)| \leq 1 \cdot d(x, y)$. □

Corolario 61. En cualquier espacio métrico las bolas abiertas son subconjuntos abiertos y las bolas cerradas son subconjuntos cerrados. También la “cáscara” $S(p, r) = \{q : d(p, q) = r\}$ es un subconjunto cerrado.

Esto es así porque $B(p, r) = f_p^{-1}((-\infty, r))$, $\overline{B}(p, r) = f_p^{-1}([0, r])$ y $S(p, r) = f_p^{-1}(\{r\})$.

Corolario 62. En cualquier espacio normado $(\mathbb{V}, \|\cdot\|)$ la norma $\|v\| = d(\mathbf{0}, v)$ es lipschitziana con respecto a sí misma, con 1 como constante de Lipschitz:

$$\left| \|x\| - \|y\| \right| \leq 1 \cdot \|x - y\| \quad . \quad (20)$$

Ejercicio 63. Sea $L : (\mathbb{V}, \|\cdot\|_{\mathbb{V}}) \rightarrow (\mathbb{W}, \|\cdot\|_{\mathbb{W}})$ una aplicación lineal entre espacios normados. Demuestra que son equivalentes:

1. L es acotada, según la definición 20 del apartado 1.1.13.
2. L es continua en el punto $\mathbf{0}_{\mathbb{V}}$.
3. L es continua.
4. L es de Lipschitz.

Y, en tal caso, la norma de operador $\|L\|$ es la mínima constante de Lipschitz para L .

1.2.8 Normas equivalentes

En el caso de \mathbb{R}^n , hemos definido los conceptos de convergencia, continuidad, compacidad, etc. mediante la norma $\|\cdot\|_{\infty}$ y la distancia asociada a ella. En este apartado vamos a ver que, mientras la dimensión sea finita, no importa mucho qué norma se utilice.

Definición 64. Dos normas $\|\cdot\|, \|\cdot\|'$ en un espacio vectorial \mathbb{V} son **equivalentes** si existen constantes $c, C > 0$ tales que

$$c \|v\| \leq \|v\|' \leq C \|v\| \quad \text{para todo } v \in \mathbb{V} \quad . \quad (21)$$

Dicho de otro modo, el cociente $\frac{\|v\|'}{\|v\|}$, $v \neq \mathbf{0}$, tiene cotas superior e inferior, ambas positivas.

Esto define, efectivamente, una relación de equivalencia entre las normas en \mathbb{V} .

Ejercicio 65. La cota $\|v\|' \leq C \cdot \|v\|$, para todo $v \in \mathbb{V}$, equivale a que cada conjunto abierto para $\|\cdot\|'$ sea también abierto para $\|\cdot\|$.

Por lo tanto, dos normas son equivalentes si y sólo si definen la misma clase de conjuntos abiertos.

Pero no queda ahí la cosa. Si dos normas en \mathbb{V} son equivalentes, entonces:

1. Definen la misma clase de subconjuntos acotados de \mathbb{V} .
2. Definen la misma clase de sucesiones convergentes en \mathbb{V} y, para cada una de éstas, definen el mismo punto límite.
3. Definen la misma clase de subconjuntos compactos de \mathbb{V} .
4. Para cada espacio métrico (X, d) , definen la misma clase de aplicaciones continuas $X \rightarrow \mathbb{V}$ y la misma clase de aplicaciones lipschitzianas $X \rightarrow \mathbb{V}$.
5. Para cada espacio métrico (X, d) , definen la misma clase de aplicaciones continuas $\mathbb{V} \rightarrow X$ y la misma clase de aplicaciones lipschitzianas $\mathbb{V} \rightarrow X$.
6. etc.

Teorema 66. *En \mathbb{R}^n todas las normas son equivalentes entre sí.*

Advertencia: ese enunciado es falso en espacios vectoriales de dimensión infinita.

Como consecuencia del teorema 66, quedan establecidas **nociones estándar en \mathbb{R}^n** de lo que es un subconjunto acotado, un subconjunto abierto, una sucesión convergente, etc, etc. Sencillamente, son esas nociones para *cualquier* norma que elijamos en \mathbb{R}^n .

Demostración del teorema 66. Basta con probar que toda norma es equivalente a la $\|\cdot\|_\infty$. Denotando 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 $C = \|\mathbf{e}_1\| + \dots + \|\mathbf{e}_n\| > 0$. Utilizando el corolario 62, deducimos que para cualesquiera $x, y \in \mathbb{R}^n$ es:

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

es decir que $\|\cdot\|$ es de Lipschitz respecto de la norma $\|\cdot\|_\infty$. A fortiori, es continua como función $\|\cdot\| : (\mathbb{R}^n, \|\cdot\|_\infty) \rightarrow \mathbb{R}$.

También en $(\mathbb{R}^n, \|\cdot\|_\infty)$ la cáscara $K = \{\omega : \|\omega\|_\infty = 1\}$ es un conjunto cerrado y acotado por el corolario 61 y compacto por el teorema 54. Entonces, al ser $\|\cdot\|$ continua en $(\mathbb{R}^n, \|\cdot\|_\infty)$, podemos aplicar el teorema 57 y existe un vector $\omega_0 \in K$ tal que $c = \|\omega_0\|$ es el mínimo de $\|\cdot\|$ en K . La constante c cumple las dos condiciones siguientes:

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

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

Una consecuencia del teorema 66 es que las funciones y aplicaciones $\mathbb{R}^n \supseteq E \xrightarrow{f} \mathbb{R}^k$ que hemos llamado *elementales* en el apartado 1.2.5 son continuas respecto de cualesquiera normas que pongamos en \mathbb{R}^n y en \mathbb{R}^k .

Demostración del lema 25. Ahora sabemos que el conjunto $K = \{v \in \mathbb{V} : \|v\|_{\mathbb{V}} = 1\}$ es compacto en $(\mathbb{V}, \|\cdot\|_{\mathbb{V}})$. Lo que dice el lema 25 es que la función

$$f : K \longrightarrow \mathbb{R}, \quad f(v) = \|L(v)\|_{\mathbb{W}},$$

alcanza sus valores máximo y mínimo en K . Esto quedará demostrado si vemos que f es continua. Como f es la compuesta de L y $\|\cdot\|_{\mathbb{W}}$, y por el corolario 62 la norma $\|\cdot\|_{\mathbb{W}}$ es continua en $(\mathbb{W}, \|\cdot\|_{\mathbb{W}})$, sólo nos falta demostrar que L es continua respecto de la norma $\|\cdot\|_{\mathbb{V}}$ en salida y la norma $\|\cdot\|_{\mathbb{W}}$ en llegada. El teorema 66 nos dice ahora que es suficiente demostrar la continuidad de L en el caso particular $(\mathbb{V}, \|\cdot\|_{\mathbb{V}}) = (\mathbb{R}^n, \|\cdot\|_\infty)$. Pero entonces el siguiente cálculo:

$$\|L(x)\|_{\mathbb{W}} = \|x_1L(\mathbf{e}_1) + \dots + x_nL(\mathbf{e}_n)\|_{\mathbb{W}} \leq \|x\|_\infty(\|L(\mathbf{e}_1)\|_{\mathbb{W}} + \dots + \|L(\mathbf{e}_n)\|_{\mathbb{W}}) = M\|x\|_\infty,$$

implica $\|L(x) - L(x')\|_{\mathbb{W}} \leq M\|x - x'\|_\infty$, luego L es de Lipschitz y, a fortiori, continua. \square

1.2.9 Conexión

Definiciones 67. 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**: existe un camino $\alpha(t) : [0, 1] \rightarrow E$ (atención: α contenido en E) que empieza en p y termina en q : $\alpha(0) = p, \alpha(1) = q$.

Proposición 68. 1. Los subconjuntos conexos por caminos de la recta real son los intervalos y los conjuntos de un elemento $\{a\}$.

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

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

El siguiente teorema viene a decir que un abierto conexo por caminos “es de una sola pieza”.

Teorema 69. Un abierto conexo por caminos nunca puede ser unión de dos abiertos disjuntos y no vacíos.

De hecho, si U es una unión $U_1 \cup U_2$, con U_1 y U_2 abiertos disjuntos, entonces ningún punto de U_1 puede unirse a ningún punto de U_2 por caminos contenidos en U .

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

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

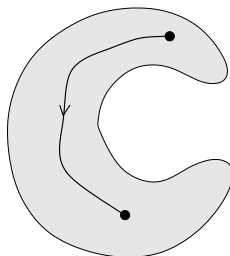
Estos conjuntos se llaman **componentes conexas por caminos** de E . La propiedad 3. de la proposición 68 hace que estas componentes formen una **partición**: su unión es todo E y dos cualesquiera que sean distintas son disjuntas.

Proposición 70. Si U es un abierto de \mathbb{R}^n entonces sus componentes conexas por caminos son también 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 $C \subseteq \mathbb{R}^n$ es conexo por caminos, pues si $p, q \in C$ entonces

$$\alpha : [0, 1] \longrightarrow C \quad , \quad \alpha(t) = (1 - t)p + tq \quad ,$$

es un camino (rectilíneo) en C que une p con q . Pero la mayoría de los conexos por caminos no son convexos. Un ejemplo muy sencillo es un abierto del plano con la forma de la letra C .



1.3 Alguna cosa más

En este apartado doy definiciones y resultados que no he explicado en clase, pero que pueden hacerte falta en algún momento del curso.

1.3.1 Sucesiones de Cauchy y convergencia

Ejercicio 71. Una sucesión de vectores $\{x_j\}_{j=1}^\infty \subset \mathbb{R}^n$ es de Cauchy en $(\mathbb{R}^n, \|\cdot\|_\infty)$ si y sólo si para cada $i = 1, \dots, n$ la sucesión $\{x_j^i\}_{j=1}^\infty$ de las i -ésimas coordenadas es una sucesión numérica de Cauchy.

Teorema 72. Sea $C \subseteq \mathbb{R}^n$ un subconjunto cerrado cualquiera. Toda sucesión $\{x_j\}_{j=1}^\infty \subset C$ que sea de Cauchy es convergente a algún punto de C .

Demostración. Para la norma $\|\cdot\|_\infty$, la sucesión es convergente en \mathbb{R}^n como consecuencia de los ejercicios 71 y 40. Para cualquier otra norma en \mathbb{R}^n , aplicar el teorema 66 tanto para la noción de sucesión de Cauchy como para la noción de sucesión convergente en \mathbb{R}^n .

El límite de la sucesión está en C porque C es cerrado. \square

1.3.2 Unión e intersección

Los abiertos de un espacio métrico tienen las siguientes propiedades:

1. La unión, finita o infinita, de abiertos es un abierto.
2. La intersección de una familia finita de abiertos es un abierto.

Puesto que los cerrados son los complementarios de los abiertos, tenemos los enunciados correspondientes:

1. La intersección, finita o infinita, de cerrados es un cerrado.
2. La unión de una familia finita de cerrados es un cerrado.

La intersección de infinitos abiertos puede dar un resultado que no sea abierto; por ejemplo:

$$\bigcap_{n=1}^{\infty} \left(-\frac{1}{n}, \frac{1}{n}\right) = \{0\}.$$

La unión de infinitos cerrados puede dar un resultado que no sea cerrado, por ejemplo:

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

1.3.3 Interior y cierre

Definición 73. 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 en (X, d)** y se denota $\text{int } E$.

Es trivial ver que $\text{int } E \subseteq E$ y que cualquier abierto contenido en E está contenido en $\text{int } E$.

Proposición 74. El conjunto $\text{int } E$ es abierto, luego es el abierto más grande contenido en E .

Definición 75. Dado un subconjunto $E \subseteq X$, el **cierre de E en (X, d)** o **adherencia de E en (X, d)** es el conjunto \overline{E} cuyos puntos son todos los límites de sucesiones contenidas en E y convergentes en (X, d) .

Se tiene $E \subseteq \overline{E}$, porque dado cualquier punto $p \in E$ la sucesión constante $\{p, p, p, \dots\}$ está contenida en E y converge a p . También es fácil ver que cualquier cerrado que contenga a E debe contener a \overline{E} .

Proposición 76. El cierre \overline{E} es un conjunto cerrado. Por lo tanto es el cerrado más pequeño que contiene a E .

Demostración. Sea $\{x_n\}_{n=1}^{\infty} \subset \overline{E}$ una sucesión contenida en el cierre \overline{E} y convergente a un punto $x_0 \in X$. Hay que demostrar que $x_0 \in \overline{E}$.

Para cada n hay una sucesión $\{x_{n,j}\}_{j=1}^{\infty} \subset E$ que converge al punto x_n . Esto produce una sucesión de sucesiones convergentes

$$\begin{array}{ccc} x_{1,1}, x_{1,2}, x_{1,3}, \dots & \rightarrow & x_1 \\ x_{2,1}, x_{2,2}, x_{2,3}, \dots & \rightarrow & x_2 \\ x_{3,1}, x_{3,2}, x_{3,3}, \dots & \rightarrow & x_3 \\ \vdots & & \vdots \end{array}$$

Sea k cualquier entero positivo. Primero, encontramos un n_k tal que $d(x_{n_k}, x_0) < 1/k$. Segundo, encontramos un j_k tal que $d(x_{n_k, j_k}, x_{n_k}) < 1/k$. La desigualdad triangular nos permite concluir que $d(x_{n_k, j_k}, x_0) < \frac{1}{k} + \frac{1}{k} = \frac{2}{k}$. Entonces $\{x_{n_k, j_k}\}_{k=1}^{\infty}$ es una sucesión contenida en E y convergente a x_0 , luego $x_0 \in \overline{E}$. \square

Corolario 77. Un subconjunto $E \subseteq X$ es cerrado si y sólo si $E = \overline{E}$.

Ejercicio 78. Demuestra que un punto $x \in X$ está en \overline{E} si y sólo si para todo $\varepsilon > 0$ la bola $B(x, \varepsilon)$ corta a E , es decir $B(x, \varepsilon) \cap E \neq \emptyset$.

1.3.4 Conjuntos discretos y conjuntos densos

Definiciones 79. Decimos que x es un **punto aislado de E** si existe un $\varepsilon > 0$ tal que $B(x, \varepsilon) \cap E = \{x\}$. Decimos que E es **discreto** si todos sus puntos son aislados.

Un conjunto discreto no está obligado a ser cerrado, por ejemplo $E = \{1/n : n = 1, 2, 3, \dots\}$.

Definición 80. 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).

Ejercicio 81. (Principio de prolongación de las identidades). Sean $f, g : (X, d) \rightarrow (X', d')$ aplicaciones continuas y $E \subseteq X$ denso en X . Si $f|_E \equiv g|_E$ entonces $f \equiv g$.

Este resultado se utiliza, por ejemplo, en el documento donde estudiamos la regla del paralelogramo.

1.3.5 Abiertos relativos y cerrados relativos

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

Teorema-definición 82. Los abiertos de (Y, d_Y) son las intersecciones $Y \cap U$, siendo U un abierto cualquiera de (X, d) . Estos conjuntos se llaman **abiertos relativos de Y** .

Los cerrados de (Y, d_Y) son las intersecciones $Y \cap C$, siendo C un cerrado cualquiera de (X, d) . Estos conjuntos se llaman **cerrados relativos de Y** . \square

Ejemplo. Sea \mathbb{R} con $d(x, y) = |x - y|$. 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, mediante la operación de intersección, los abiertos y cerrados relativos de Y .

2 Diferenciabilidad y extremos locales

Empezamos con una notación inventada por el físico ruso Lev Landau (1908-1968).

2.1 Oes de Landau

Tenemos dos espacios normados E, F y un punto $x_0 \in E$. En un entorno $U \subseteq E$ de x_0 , tenemos definidas una aplicación $f : U \rightarrow F$ y una función escalar $\varphi : U \rightarrow [0, +\infty)$ con $\varphi(x) > 0$ para $x \neq x_0$.

Definición 83. Decimos que **f es una o grande de Landau de φ** , y lo indicamos escribiendo $f(x) = O(\varphi(x))$ si existe una constante $C \geq 0$ tal que

$$\|f(x)\| \leq C \varphi(x) \quad , \quad \text{para todo } x \in U .$$

Decimos que **f es una o pequeña de Landau de φ** cuando **x tiende a x_0** , y se indica escribiendo $f(x) = o(\varphi(x))$ si se tiene

$$\lim_{\substack{x \rightarrow x_0 \\ x \neq x_0}} \frac{\|f(x)\|}{\varphi(x)} = 0 .$$

Normalmente no hace falta indicar el punto x_0 , porque es el único punto donde se anula φ .

Es obvio que si $f = o(\varphi)$ entonces $f = O(\varphi)$ cerca de x_0 , pero el recíproco suele ser falso.

Ejemplos con $E = F = \mathbb{R}$:

$$\sin x = O(|x|), \quad x_0 = 0, \quad U = \mathbb{R}.$$

$$x^3 \sqrt{x^2 + 10x^4} = O(|x|^4), \quad x_0 = 0, \quad U = (-3, 3).$$

$$e^x = O(1), \quad U = (a, b) \text{ intervalo finito.}$$

$$e^x - 1 = O(|x|), \quad x_0 = 0, \quad U = (-1, 1).$$

$$e^x - 1 - x - (x^2/2) - (x^3/6) = O(|x|^4), \quad x_0 = 0, \quad U = (-1, 1).$$

$$|x|^{1/2}x = o(|x|), \quad x_0 = 0.$$

$$|x|^{1/2} \sin x = o(|x|^{4/3}), \quad x_0 = 0.$$

Ejemplos con $E = \mathbb{R}^2$, $F = \mathbb{R}$ y $\|\cdot\|$ cualquier norma en \mathbb{R}^2 :

$$x^3 - 7x^5y + 4y^3 = O(\|(x, y)\|^3), \quad x_0 = (0, 0), \quad U = B(x_0, 1).$$

$$\sqrt[14]{x^2 + 3y^2} \log(x + 1) = o(\|(x, y)\|), \quad x_0 = (0, 0).$$

Advertencias. (1) Esta notación no sigue las pautas convencionales de uso del signo “=” . Si consideramos la **clase** de las funciones que son o grande de φ , entonces “ $O(\varphi)$ ” designa *el elemento general de dicha clase* y la fórmula “ $f = O(\varphi)$ ” significa que f es un elemento de esa clase. De hecho $O(\varphi)$ puede designar una función distinta cada vez que aparece, incluso en una misma fórmula. Por ejemplo, la fórmula

$$O(\varphi) + O(\varphi) = O(\varphi) , \tag{22}$$

significa que si f_1 es una o grande de φ y f_2 es otra o grande de φ entonces $f_1 + f_2$ también es una o grande de φ (lo cual es obvio).

Hoy día utilizaríamos “ $O(\varphi)$ ” para designar esa clase de funciones, y escribiríamos “ $f \in O(\varphi)$ ” para decir que f es una de las funciones en la clase; pero el símbolo “ \in ” era muy poco conocido cuando Landau inventó su notación, y usar el símbolo “ $=$ ” tiene la ventaja de permitir manipulaciones muy convenientes, como la fórmula (22) o las siguientes:

$$O(o(\varphi)) = o(\varphi) \quad , \quad O(\varphi)O(\psi) = O(\varphi\psi) \quad , \quad \text{etc} \quad ,$$

todas dictadas por el sentido común. Vale decir otro tanto de las notaciones “ $o(\varphi)$ ” y “ $f = o(\varphi)$ ”.

(2) Hay que tener en cuenta que la notación no es simétrica, porque establece una *acotación* de la función escrita a la izquierda por (múltiplos de) la función escrita a la derecha: es verdad que $o(|x|) = O(|x|)$, pero suele ser mentira que $O(|x|) = o(|x|)$; la primera fórmula dice que toda función que sea una o pequeña de $|x|$ es una o grande de $|x|$, mientras que la segunda fórmula afirma algo tan falso como que toda función que sea una o grande de $|x|$ es una o pequeña de $|x|$.

(3) También hay que recordar que no hay una “función que se llame O ”, con lo cual “ $f = O(\varphi)$ ” *no significa que f sea una función de φ* . Por ejemplo, llamando r, θ a las coordenadas polares en el plano, es verdad que $xy = O(x^2 + y^2)$, es decir $xy = O(r^2)$, pero el valor de r no determina el valor de xy (que también depende de θ). Análogamente, es verdad que $xy = o(r)$ pero el valor de r no determina el de xy .

Proposición 84. Sea k un entero positivo. Si $\phi : E \rightarrow F$ es positivamente homogénea de grado k y $\phi(x) = o(\|x\|^k)$, entonces $\phi \equiv 0$.

Demostración. Si v es cualquier vector no nulo de E , podemos definir la función

$$g(t) = \frac{\|\phi(tv)\|}{\|t^k v\|} \quad , \quad t > 0 \quad ,$$

que resulta ser constante. Además, la hipótesis $\phi(x) = o(\|x\|^k)$ implica que $g(t) \rightarrow 0$ cuando $t \rightarrow 0$. Luego $g(t)$ es idénticamente nula y tenemos $0 = g(1) = \phi(v)$. \square

2.2 Diferenciabilidad: derivadas de orden 1

Aquí introducimos el importante concepto de función diferenciable y explicamos su significado.

2.2.1 La definición

Definición 85. Sean: E, F espacios normados, un punto $x_0 \in E$ y un abierto $U \subseteq E$ entorno de x_0 . Decimos que una aplicación $f : U \rightarrow F$ es **diferenciable en x_0** si existe $L : E \rightarrow F$, lineal acotada, tal que:

$$R_1(x) \stackrel{\text{def}}{=} f(x) - [f(x_0) + L(x - x_0)] = o(\|x - x_0\|) \quad , \quad (23)$$

o, lo que es equivalente:

$$\mathcal{R}_1(h) \stackrel{\text{def}}{=} f(x_0 + h) - [f(x_0) + L(h)] = o(\|h\|) \quad . \quad (24)$$

La dos funciones $R_1(x)$ y $\mathcal{R}_1(h)$ reciben el nombre de **resto de Taylor de primer orden de f en x_0** .

Decimos que f es **diferenciable en U** si es diferenciable en cada punto de U .

Primeras propiedades:

a) L es única si existe, en cuyo caso se llama **diferencial de f en x_0** y se denota $(df)_{x_0}$.

En particular f es diferenciable en x_0 con diferencial nula $\iff f(x) = o(\|x - x_0\|)$

b) Para que f sea diferenciable en x_0 le es necesario (no suficiente) ser continua en x_0 .

c) Dadas funciones escalares $f_1, \dots, f_k : U \rightarrow \mathbb{R}$, la función vectorial

$$\mathbf{f} \equiv \begin{bmatrix} f_1 \\ \vdots \\ f_k \end{bmatrix} : U \longrightarrow \mathbb{R}^k ,$$

es diferenciable en x_0 si sólo si cada una de sus entradas f_j es diferenciable en x_0 , en cuyo caso:

$$(d\mathbf{f})_{x_0}(v) = \begin{bmatrix} (df_1)_{x_0}(v) \\ \vdots \\ (df_k)_{x_0}(v) \end{bmatrix} , \quad \text{para todo } v \in E . \quad (25)$$

d) Si $f, g : U \rightarrow F$ son diferenciables en x_0 y c constante, entonces $f + g$, $f - g$ y cf son diferenciables en x_0 y se tiene:

$$(d(f+g))_{x_0} = (df)_{x_0} + (dg)_{x_0} , \quad (d(f-g))_{x_0} = (df)_{x_0} - (dg)_{x_0} , \quad (d(cf))_{x_0} = c(df)_{x_0} .$$

La propiedad a) se debe a que si $f(x+h) - f(x_0) - L_1(h)$ y $f(x+h) - f(x_0) - L_2(h)$ son ambas oes pequeñas de $\|h\|$ entonces $(L_1 - L_2)(h) = o(\|h\|)$ y, por la proposición 84, es $L_1 - L_2 \equiv 0$.

La propiedad b) se deduce del siguiente cálculo:

$$\lim_{x \rightarrow x_0} f(x) = \lim_{h \rightarrow \mathbf{0}} f(x_0 + h) = \lim_{h \rightarrow \mathbf{0}} (f(x_0) + L(h) + o(\|h\|)) = f(x_0) + \mathbf{0} + \mathbf{0} = f(x_0) .$$

Por supuesto, no toda aplicación continua en x_0 es diferenciable en x_0 . Por ejemplo, una norma $\|\cdot\|$ siempre es continua y nunca es diferenciable en el origen $\mathbf{0}$.

Se deja como ejercicio la prueba de las propiedades c) y d).

Es útil estudiar el concepto de diferenciable para una función definida en un conjunto no abierto. La única razón por la que no lo hacemos aquí es para simplificar la exposición, dado que tenemos un límite de tiempo.

Si E y F tienen dimensiones finitas, entonces todas las normas en E son equivalentes, lo mismo ocurre en F y todas las aplicaciones lineales $L : E \rightarrow F$ son acotadas. El que la función $R_1(x)$ de la fórmula (23) sea un $o(\|x - x_0\|)$, o que $\mathcal{R}_1(h)$ de la fórmula (24) sea un $o(\|h\|)$, es algo que no depende de las normas elegidas en E y en F , luego tanto la diferenciableidad de f en x_0 como el valor de la diferencial $(df)_{x_0}$ son independientes de esas normas.

2.2.2 Significado analítico: regla de la cadena para caminos

Recordemos que un **camino** es una función vectorial de una variable independiente

$$\alpha(t) \equiv (x_1(t), \dots, x_n(t)) .$$

Una tal función es diferenciable en $t = t_0$ si y sólo si existe el vector derivada

$$\alpha'(t_0) = (x'_1(t_0), \dots, x'_n(t_0)) ,$$

que suele llamarse **velocidad en $t = t_0$** . La diferencial del camino en $t = t_0$ es la función

$$(d\alpha)_{t_0} : \mathbb{R} \longrightarrow \mathbb{R}^n , \quad \mathbb{R} \ni a \longmapsto a \alpha'(t_0) .$$

Definición 86. Dada una función $f(x_1, \dots, x_n) : U \rightarrow \mathbb{R}^k$, dado un punto $x_0 \in U$ y dado un vector cualquiera $v \in \mathbb{R}^n$, la **derivada direccional de \mathbf{f} con respecto a \mathbf{v} en el punto \mathbf{x}_0** es el siguiente elemento de \mathbb{R}^k

$$(D_v f)(x_0) \stackrel{\text{def}}{=} \left. \frac{d}{dt} \right|_{t=0} f(x_0 + tv) .$$

En particular, si $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ es la base estándar de \mathbb{R}^n entonces

$$(D_{\mathbf{e}_j} f)(x_0) = \frac{\partial f}{\partial x_j}(x_0) = f_{x_j}(x_0) \quad , \quad j = 1, \dots, n .$$

Es fácil ver que si f es diferenciable en x_0 entonces para todo $v \in \mathbb{R}^n$ existe la derivada $(D_v f)(x_0)$ y de hecho

$$v \mapsto (D_v f)(x_0) = (df)_{x_0}(v) ,$$

es una función lineal de v . Esta linealidad se traduce en la siguiente fórmula

$$v = (c_1, \dots, c_n) \implies \left. \frac{\partial}{\partial t} \right|_{t=0} f(x_0 + tv) = c_1 (D_{\mathbf{e}_1} f)(x_0) + \dots + c_n (D_{\mathbf{e}_n} f)(x_0) ,$$

es decir, si $x_0 = (a_1, \dots, a_n)$ entonces:

$$\left. \frac{\partial}{\partial t} \right|_{t=0} f(a_1 + c_1 t, \dots, a_n + c_n t) = c_1 f_{x_1}(a_1, \dots, a_n) + \dots + c_n f_{x_n}(a_1, \dots, a_n) , \quad (26)$$

para cualesquiera $c_1, \dots, c_n \in \mathbb{R}$.

Definición 87. La **regla de la cadena a lo largo de caminos cualesquiera** afirma que para todo camino diferenciable $\alpha(t) \equiv (x_1(t), \dots, x_n(t))$ se tiene

$$\boxed{\frac{d}{dt} f(x_1(t), \dots, x_n(t)) = x'_1(t) f_{x_1}(\alpha(t)) + \dots + x'_n(t) f_{x_n}(\alpha(t))} \quad (27)$$

Esta regla es una herramienta fundamental del cálculo en varias variables, por lo que nos interesan especialmente las funciones que la cumplen.

Observa que (26) es el caso particular de la regla cuando el camino es rectilíneo $\alpha(t) \equiv x_0 + tv$. La función f cumple esta “parte rectilínea” de la regla de la cadena si y sólo si $v \mapsto (D_v f)(x_0)$ es una función definida para todo v y lineal.

Teorema 88. La función $f(x_1, \dots, x_n)$ es diferenciable en el punto x_0 , según la definición 85, si y sólo si satisface la regla de la cadena (27) a lo largo de todos los caminos diferenciables cuando pasan por x_0 .

Por lo tanto f es diferenciable si y sólo si satisface la regla de la cadena a lo largo de todos los caminos diferenciables contenidos en su dominio. \square

Este teorema pone en valor la definición 85, por nuestro interés en que se cumpla la regla de la cadena. Dada una función, es imposible comprobar si cumple (27) a lo largo de todos los caminos diferenciables, porque hay una infinidad excesiva de tales caminos. En cambio, comprobar si cumple (23) es factible porque sabemos que, si existe, la función L que aparece en (23) sólo puede ser ésta:

$$L(c_1, \dots, c_n) = c_1 \frac{\partial f}{\partial x_1}(x_0) + \dots + c_n \frac{\partial f}{\partial x_n}(x_0) . \quad (28)$$

Una manera de ver si f es diferenciable en x_0 sería la siguiente:

1. Comprobar que existen las derivadas $\frac{\partial f}{\partial x_j}(x_0)$ y calcularlas.
2. Formar la función L mediante la fórmula (28).
3. Calcular la función $f(x) - f(x_0) - L(x - x_0)$ y ver si es una o pequeña de $\|x - x_0\|$.

Pero de hecho hay criterios más fáciles de utilizar que éste (ver el teorema 92 del apartado 2.3). El uso habitual de las fórmulas (23) o (24) es en demostraciones de teoremas que involucran el concepto de diferenciabilidad.

Ahora vamos a ver, con dos ejemplos especiales, el contenido profundo de la definición 85.

Primer ejemplo. Vamos a estudiar en $x_0 = (0, 0)$ la función $f(x, y) : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por:

$$f(x, y) = \begin{cases} \frac{xy^3}{x^2 + y^6} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases} \quad (29)$$

Empezamos viendo que la derivada direccional $(D_v f)(0, 0)$ existe para todo $v \in \mathbb{R}^2$ y depende linealmente de v . Es decir que f satisface la “parte rectilínea” de la regla de la cadena en $(0, 0)$.

Sea $v = (a, b)$. Si $a = 0$ o $b = 0$, se comprueba de inmediato que

$$f((0, 0) + tv) = f(at, bt) = 0 \quad \text{para todo } t,$$

luego $(D_v f)(0, 0)$ existe y es nula para $v = a\mathbf{e}_1$ y para $v = b\mathbf{e}_2$. Si $a \neq 0 \neq b$, entonces

$$f(at, bt) = \begin{cases} \frac{ab^3 t^4}{a^2 t^2 + b^6 t^6} = \frac{ab^3}{a^2 + b^6 t^4} t^2 & \text{si } t \neq 0 \\ 0 & \text{si } t = 0 \end{cases} = O(t^2) = o(|t|),$$

y en realidad es $(D_v f)(0, 0) = 0$ para todo v . La derivada direccional en $(0, 0)$ es una función lineal de v , porque es la función constante nula.

Veamos ahora que f es discontinua en $(0, 0)$. Esto, además de probar que f no es diferenciable en $(0, 0)$, pone de manifiesto lo débil que es la “parte rectilínea” de la regla de la cadena: ni siquiera obliga a la función a ser continua.

El camino $\alpha(t) \equiv (t^3, t)$ es continuo, pero la compuesta

$$(f \circ \alpha)(t) = f(t^3, t) = \begin{cases} \frac{t^6}{t^6 + t^6} = 1/2 & \text{si } t \neq 0 \\ 0 & \text{si } t = 0 \end{cases}$$

es claramente discontinua en $t = 0$, luego f es discontinua en $\alpha(0) = (0, 0)$.

Segundo ejemplo. Ahora consideramos la función $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por

$$g(x, y) = \begin{cases} \frac{xy^3}{x^2 + y^4} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases} \quad (30)$$

y de nuevo estudiamos el comportamiento en el punto $(0, 0)$. Volvemos a tener

$$g(t, 0) = g(0, t) \equiv 0,$$

luego $(D_v g)(0, 0) = 0$ para $v = a\mathbf{e}_1$ o $v = b\mathbf{e}_2$. Si $v = (a, b)$ con $a \neq 0 \neq b$, volvemos a obtener $g(at, bt) = O(t^2) = o(|t|)$. En definitiva $v \mapsto (D_v g)(0, 0)$ es la función constante nula, luego g también cumple la “parte rectilínea” de la regla de la cadena en $(0, 0)$.

Además g es continua en $(0, 0)$, porque la desigualdad

$$|xy^2| \leq \frac{x^2 + y^4}{2},$$

implica $|g(x, y)| \leq |y|/2$ y así $g(x, y) \rightarrow 0 = g(0, 0)$ cuando $(x, y) \rightarrow (0, 0)$.

A pesar de todo esto, tampoco g es diferenciable en $(0, 0)$, pues del siguiente cálculo

$$g(t^2, t) = \begin{cases} \frac{t^5}{t^4 + t^4} & \text{si } t \neq 0 \\ 0 & \text{si } t = 0 \end{cases} \equiv \frac{t}{2},$$

deducimos:

$$\frac{1}{2} = \frac{d}{dt} \Big|_{t=0} g(t^2, t) \neq 0 = \frac{d(t^2)}{dt} \Big|_{t=0} \cdot 0 + \frac{dt}{dt} \Big|_{t=0} \cdot 0 = \frac{d(t^2)}{dt} \Big|_{t=0} \cdot g_x(0, 0) + \frac{dt}{dt} \Big|_{t=0} \cdot g_y(0, 0).$$

Queda visto que g no cumple la regla de la cadena a lo largo del camino $\beta(t) \equiv (t^2, t)$ en $t = 0$. De ninguna manera vamos a considerar a g diferenciable en el punto $\beta(0) = (0, 0)$, ya que la regla de la cadena es lo que más nos interesa.

2.2.3 Significado geométrico

Empezamos con una función escalar de una variable independiente, es decir $f(x) : U \rightarrow \mathbb{R}$ siendo U un abierto de \mathbb{R} .

Cada función lineal $L : \mathbb{R} \rightarrow \mathbb{R}$ corresponde a una constante m mediante la fórmula $L(x) = mx$. Que f sea diferenciable en el punto $x_0 \in U$ significa que existe m tal que

$$f(x) - f(x_0) - m(x - x_0) = o(|x - x_0|),$$

es decir

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = m. \quad (31)$$

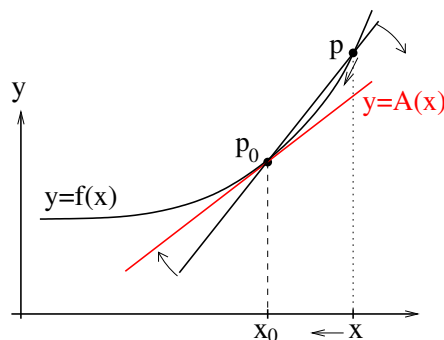
La primera interpretación geométrica consiste en ver que el **cociente incremental**

$$\frac{f(x) - f(x_0)}{x - x_0},$$

es la **pendiente** de la recta secante a la curva $\{y = f(x)\}$ pasando por los puntos

$$p_0 = (x_0, f(x_0)) \quad , \quad p(x) = (x, f(x)).$$

El límite de estas rectas secantes, cuando el punto $p(x)$ tiende al punto p_0 a lo largo de la curva, recibe el nombre de **recta afín tangente en el punto p_0** .



El número m es la pendiente de esta recta tangente, se lo denota $f'(x_0)$ y se llama **derivada de f en x_0** . En realidad, las siguientes condiciones son equivalentes:

- f es derivable en x_0 , con derivada $f'(x_0) = m$.
- f es diferenciable en x_0 , con diferencial $(df)_{x_0}(h) = mh$ para todo $h \in \mathbb{R}$.
- La curva $\{y = f(x)\}$ tiene recta tangente en el punto p_0 (límite de secantes), dada por $y = f(x_0) + m(x - x_0)$.

Es de destacar que la recta tangente, al tener una pendiente *finita* m , no es *vertical*.

Vamos a reescribir la fórmula (31) para darle otra interpretación visual. Consideramos la función afín

$$A(x) = f(x_0) + m(x - x_0),$$

cuyo grafo $\{y = A(x)\}$ es la recta tangente. También podemos definir $A(x)$ como la única función afín tal que $f(x) - A(x) = o(|x - x_0|)$. Para cada $\varepsilon > 0$ existe un $\delta > 0$ tal que

$$x \in (x_0 - \delta, x_0 + \delta) \implies \left| \frac{f(x) - A(x)}{x - x_0} \right| < \varepsilon.$$

Es suficiente dar a ε un conjunto discreto de valores: $1, \frac{1}{2}, \frac{1}{3}, \dots$ y obtener los correspondientes valores $\delta_1, \delta_2, \delta_3, \dots$ tales que

$$x \in (x_0 - \delta_j, x_0 + \delta_j) \implies \left| \frac{f(x) - A(x)}{x - x_0} \right| < \frac{1}{j}, \quad j = 1, 2, 3, \dots,$$

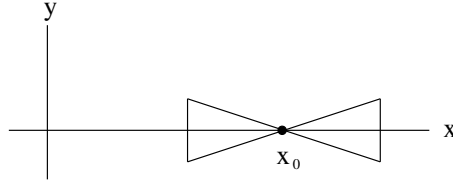
que a su vez equivale a

$$x \in (x_0 - \delta_j, x_0 + \delta_j) \implies A(x) - \frac{1}{j}|x - x_0| < f(x) < A(x) + \frac{1}{j}|x - x_0|, \quad (32)$$

El siguiente subconjunto del plano

$$\left\{ (x, y) : x \in (x_0 - \delta_j, x_0 + \delta_j), \quad -\frac{1}{j}|x - x_0| < y < \frac{1}{j}|x - x_0| \right\},$$

es una “cuña doble” formada por dos pequeños triángulos con un eje de simetría horizontal y con vértice en el punto $(x_0, 0)$

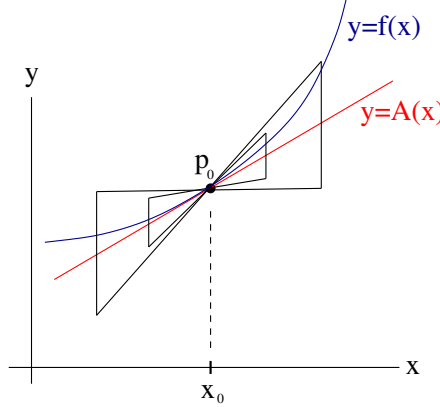


y este otro subconjunto

$$\left\{ (x, y) : x \in (x_0 - \delta_j, x_0 + \delta_j), \quad A(x) - \frac{1}{j}|x - x_0| < y < A(x) + \frac{1}{j}|x - x_0| \right\},$$

es también una cuña doble, posiblemente oblicua, con vértice en $(x_0, f(x_0)) = p_0$.

La fórmula (32) con $j = 1, 2, 3, \dots$, equivalente al límite (31), define unas cuñas oblicuas, cada vez más pequeñas y con el ángulo más cerrado, todas con vértice común p_0 y atravesadas por la recta afín $E = \{y = A(x)\}$. Que la curva $C = \{y = f(x)\}$ sea tangente a E en el punto p_0 significa que C se va metiendo en todas esas cuñas a medida que se acerca al punto p_0



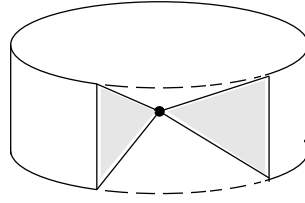
Veamos brevemente cómo es esa construcción para una función de dos variables $f(x) = f(x_1, x_2)$ en un punto $x_0 = (a, b)$. Ahora tenemos una función lineal de dos variables $L(h) = L(h_1, h_2)$ y la correspondiente función afín $A(x_1, x_2) = A(x) = f(x_0) + L(x - x_0)$. El análogo de la fórmula (32) es:

$$\|x - x_0\| < \delta_j \implies A(x) - \frac{1}{j}\|x - x_0\| < f(x) < A(x) + \frac{1}{j}\|x - x_0\|. \quad (33)$$

Si la norma utilizada en \mathbb{R}^2 es la euclídea, el conjunto

$$\left\{ (x, y) \in \mathbb{R}^2 \times \mathbb{R} : \|x - x_0\| < \delta_j, \quad -\frac{1}{j}\|x - x_0\| < y < \frac{1}{j}\|x - x_0\| \right\},$$

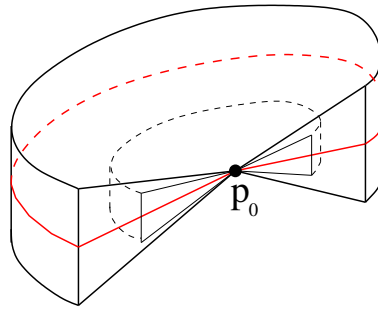
es un sólido de revolución engendrado al rotar alrededor de un eje vertical un triángulo que tiene un vértice en el punto $(x_0, 0) = (a, b, 0)$



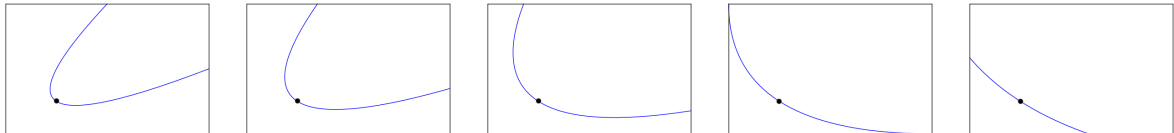
y el conjunto

$$\left\{ (x, y) \in \mathbb{R}^2 \times \mathbb{R} : \|x - x_0\| < \delta_j, \quad A(x) - \frac{1}{j} \|x - x_0\| < y < A(x) + \frac{1}{j} \|x - x_0\| \right\},$$

tiene una forma similar, aunque en general ya no será de revolución. Está “cortado en dos” por un disco en el plano afín $\{(x_1, x_2, y) \in \mathbb{R}^3 : y = A(x_1, x_2)\}$ que indicamos en color rojo en la siguiente figura. La fórmula (33), con $j = 1, 2, 3, \dots$, define una familia de estos conjuntos, cada uno más pequeño y más afilado, todos atravesados por el plano $\{y = A(x_1, x_2)\}$ y tales que la “superficie” $\{y = f(x_1, x_2)\}$ (no mostrada en la figura) va entrando en todos ellos a medida que se acerca al punto $p_0 = (x_0, f(x_0))$.



Existe otra construcción de la recta tangente, a veces muy útil, válida para una curva plana C aunque no sea un *grafo* (aunque y no esté puesta como función de x en C). El método consiste en mirar al punto $p_0 \in C$ y sus alrededores con microscopios cada vez más potentes, de modo que el punto p_0 siempre se vea en la misma posición del campo de visión, y ver a qué tiende la parte de la curva ampliada que va quedando dentro del campo de visión.



Si esa parte de la curva ampliada se va estirando a medida que aumenta la potencia del microscopio, de modo que tiende a un segmento de recta afín que pasa por p_0 , entonces esa recta límite es la **tangente afín a la curva** en el punto p_0 . En el caso particular $p_0 = (0, 0)$, estas ampliaciones las efectúan las transformaciones

$$H_\lambda(x, y) = (\lambda x, \lambda y) \quad , \quad \lambda \rightarrow +\infty,$$

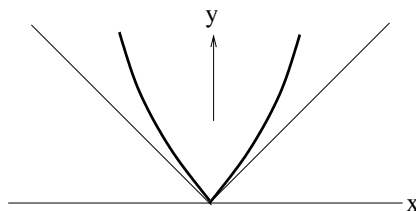
que son **homotecias con centro en $(0, 0)$** . En el caso general $p_0 = (a, b)$, las homotecias son

$$H_\lambda(x, y) = (a, b) + (\lambda(x - a), \lambda(y - b)) \quad , \quad \lambda \rightarrow +\infty.$$

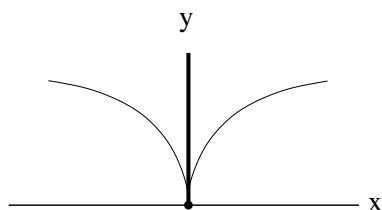
Lo que hacen estas transformaciones es expandir el plano con la misma intensidad en todas las direcciones, dejando fijo el punto p_0 .

El conjunto límite, de aquella parte de las curvas ampliadas $H_\lambda(C)$ que cae dentro del campo visual, es siempre una unión de semirrectas que empiezan en p_0 , es decir un **cono** con vértice en p_0 , pero no necesariamente una recta afín. Lo llamamos **cono tangente en p_0** .

Por ejemplo, para la curva $\{y = |x| + x^2\}$ (que es como el perfil de una bala de fusil) el cono tangente en el punto $(0, 0)$ no es una recta sino el conjunto $\{(x, y) : y = |x|\}$, unión de dos semirrectas que salen de $(0, 0)$ y forman un ángulo recto.

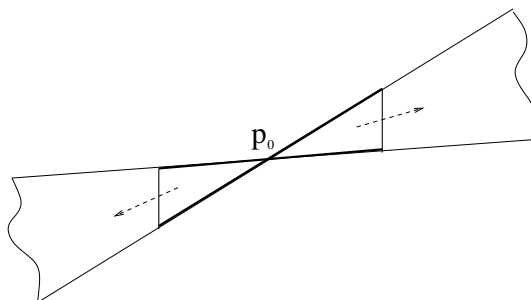


Para la curva $\{(x, y) : y = \sqrt{|x|}\}$ y el punto $(0, 0)$, el cono tangente es la semirrecta vertical que sube hacia arriba desde $(0, 0)$.

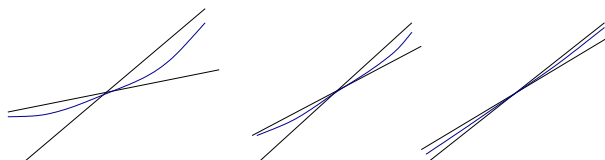


El cono tangente en $(0, 0)$ a la curva $\{y = \sqrt[3]{x}\}$ es una recta, pero es *vertical* y por lo tanto no es el grafo $\{y = A(x)\}$ para ninguna función afín $A(x)$: $\sqrt[3]{x}$ no es diferenciable en $x = 0$.

Las homotecias ampliadoras H_λ llevan cada cuña, de las antes descritas, dentro de la región angular subtendida por la cuña



luego, si la curva C es tangente en p_0 a una recta afín E , entonces la parte cercana a p_0 de las curvas ampliadas $H_\lambda(C)$ está dentro de ángulos cada vez más agudos que contienen a E



y el cono límite es la recta E . Este argumento también puede recorrerse hacia atrás, y así probar que si el cono tangente en el punto $(x_0, f(x_0))$ existe y es una recta *no vertical* entonces f es diferenciable en x_0 . En general:

Una función escalar de n variables $f(x_1, \dots, x_n)$ es diferenciable en el punto x_0 si y sólo si el grafo

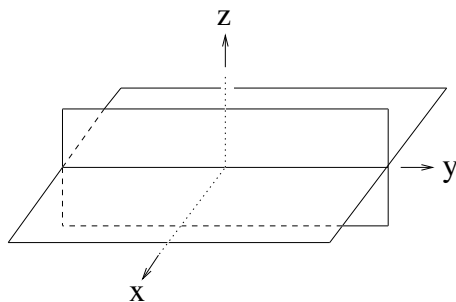
$$\{(x, y) : y = f(x)\} \subseteq \mathbb{R}^{n+1},$$

tiene cono tangente en el punto $p_0 = (x_0, f(x_0))$ y ese cono es un hiperplano afín no conteniendo ninguna recta afín vertical.

Éste es un *criterio visual* de diferenciability: conocido el cono tangente en p_0 , basta con mirar a ese cono y ver si es un hiperplano o no y, en su caso, si tiene parte vertical o no.

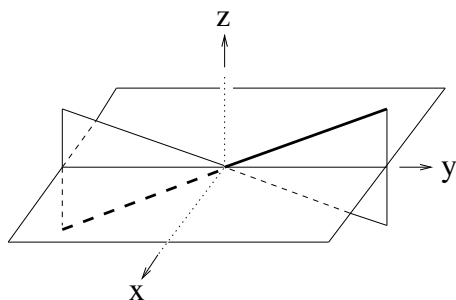
Observación. Si C es una curva espacial contenida en el grafo de f , y C pasa por el punto p_0 , entonces la recta tangente a C en p_0 es una parte del cono tangente al grafo en p_0 .

Para la función $f(x, y)$ definida por la fórmula (29) del apartado 2.2.2, el cono tangente al grafo en $(0, 0, 0)$ viene descrito por la siguiente figura



y claramente no es un plano. Así vemos que f no es diferenciable en $(0, 0)$.

Para la función $g(x, y)$ definida por la fórmula (30) del apartado 2.2.2, el cono tangente al grafo en $(0, 0, 0)$ viene descrito por la siguiente figura



La recta destacada en trazo grueso es la dirección del vector $(0, 1, 1/2)$ y está en el cono porque es la tangente en $t = 0$ al camino $\gamma(t) = (t^2, t, f(t^2, t))$, contenido en el grafo. El hecho de que esta recta sobresalga del plano $\{z = 0\}$ corresponde a que, como se vió en el apartado 2.2.2, la función g no satisface la regla de la cadena a lo largo del camino $\beta(t) = (t^2, t)$.

Ejercicio 89. Demuestra que las dos últimas figuras dicen la verdad: describen los conos tangentes correctamente.

2.2.4 Matriz jacobiana

Dados un abierto $U \subseteq \mathbb{R}^n$, un punto $x_0 \in U$ y una función $f : U \rightarrow \mathbb{R}^k$ diferenciable en x_0 , la diferencial es una aplicación lineal

$$(df)_{x_0} : \mathbb{R}^n \longrightarrow \mathbb{R}^k ,$$

y por lo tanto existe una única matriz $A \in M_{k \times n}(\mathbb{R})$ tal que

$$(df)_{x_0}(v) = Av \quad , \quad \text{para todo } v \in \mathbb{R}^n .$$

Llamamos a A la **matriz jacobiana de f en x_0** y la denotamos $(Df)_{x_0}$. También es frecuente la notación $(Df)(x_0)$, aunque aquí no la vamos a utilizar. Tenemos, pues:

$$(df)_{x_0}(v) = (D_v f)(x_0) = (Df)_{x_0} v \quad , \quad \text{para todo } v \in \mathbb{R}^n , \quad (34)$$

es decir:

$$\begin{aligned} & \text{imagen del vector por la diferencial} = \\ & = \text{derivada direccional en } x_0 = \\ & = \text{producto de la matriz jacobiana por el vector.} \end{aligned}$$

La regla de la cadena (27) del apartado 2.2.2 puede escribirse como un producto de matrices:

$$\boxed{\frac{d}{dt} f(\alpha(t)) = (Df)_{\alpha(t)} \alpha'(t)} \quad (35)$$

La jacobiana de un camino es una columna, coincidente con el vector velocidad:

$$\text{si } \alpha(t) \equiv \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}, \quad \text{entonces } (D\alpha)_t = \begin{bmatrix} x'_1(t) \\ \vdots \\ x'_n(t) \end{bmatrix} = \alpha'(t).$$

La jacobiana de una función escalar es una fila:

$$\text{si } f(x_1, \dots, x_n) : U \longrightarrow \mathbb{R}, \quad \text{entonces } Df = [f_{x_1} \ f_{x_2} \ \cdots \ f_{x_n}].$$

Las funciones escalares también tienen **vector gradiente**:

$$\nabla f \stackrel{\text{def}}{=} (f_{x_1}, f_{x_2}, \dots, f_{x_n}),$$

y se cumplen las siguientes igualdades, una con multiplicación matricial (fila por columna) y la otra con producto escalar (vector \cdot vector):

$$\frac{d}{dt} f(\alpha(t)) = (Df)_{\alpha(t)} \alpha'(t) = \nabla f_{\alpha(t)} \cdot \alpha'(t),$$

En el caso general $f : U \rightarrow \mathbb{R}^k$ hay dos maneras correctas de calcular la matriz jacobiana:

1. **Por columnas:** las columnas de A son $Ae_j = D_{e_j} f(x_0) = f_{x_j}(x_0)$, es decir las derivadas parciales de f en x_0 *puestas como columnas*:

$$(Df)_{x_0} = [f_{x_1}(x_0) \mid f_{x_2}(x_0) \mid \cdots \mid f_{x_n}(x_0)]_{k \times n}. \quad (36)$$

2. **Por filas:** si las funciones escalares $f_1(x), \dots, f_k(x) : U \rightarrow \mathbb{R}$ son las *entradas* del vector $f(x)$, es decir que

$$f(x) \equiv \begin{bmatrix} f_1(x) \\ \vdots \\ f_k(x) \end{bmatrix},$$

entonces las filas de la jacobiana son los gradientes de esas entradas:

$$(Df)_{x_0} = \begin{bmatrix} \overline{\nabla f_1(x_0)} \\ \overline{\nabla f_2(x_0)} \\ \vdots \\ \overline{\nabla f_k(x_0)} \end{bmatrix}_{k \times n}, \quad (37)$$

compara esta última fórmula con la (25).

2.2.5 Regla de la cadena para diferenciales y jacobianas

Teorema 90. Sea f diferenciable en x_0 . Sea $y_0 = f(x_0)$. Sea g diferenciable en y_0 .

La compuesta $g \circ f$ es diferenciable en x_0 y su diferencial es la compuesta de las diferenciales:

$$(d(g \circ f))_{x_0} = (dg)_{f(x_0)} \circ (df)_{x_0} \quad (38)$$

Si estas aplicaciones van entre dimensiones finitas, entonces tenemos sendas matrices jacobianas y la fórmula (38) equivale a esta:

$$(D(g \circ f))_{x_0} = (Dg)_{f(x_0)} (Df)_{x_0} \quad (39)$$

De la cual (35) es un caso particular.

Demostación. Esta demostración funciona incluso cuando alguna dimensión es infinita. Escribamos $L_1 = (df)_{x_0}$ y $L_2 = (dg)_{y_0}$. Tenemos los desarrollos:

$$\begin{aligned} f(x_0 + h) &= f(x_0) + L_1(h) + o(\|h\|) , \\ g(y_0 + k) &= g(y_0) + L_2(k) + o(\|k\|) . \end{aligned}$$

En la segunda de estas igualdades sustituimos y_0 por $f(x_0)$ y k por $L_1(h) + o(\|h\|)$. Obtenemos:

$$\begin{aligned} (g \circ f)(x_0 + h) &= g(f(x_0 + h)) = g(f(x_0) + L_1(h) + o(\|h\|)) = \\ &= g(y_0 + k) = g(y_0) + L_2(k) + o(\|k\|) = \\ &= (g \circ f)(x_0) + L_2(L_1(h) + o(\|h\|)) + o(\|L_1(h) + o(\|h\|)\|) \end{aligned}$$

Al ser L_1 lineal acotada, tenemos $\|L_1(h)\| = O(\|h\|)$, luego $\|L_1(h) + o(\|h\|)\| = O(\|h\|)$ y así:

$$(g \circ f)(x_0 + h) = (g \circ f)(x_0) + L_2(L_1(h)) + L_2(o(\|h\|)) + o(O(\|h\|)) .$$

Al ser L_2 lineal acotada, tenemos $L_2(o(\|h\|)) = o(\|h\|)$. Finalmente:

$$\begin{aligned} (g \circ f)(x_0 + h) &= (g \circ f)(x_0) + (L_2 \circ L_1)(h) + o(\|h\|) + o(O(\|h\|)) = \\ &= (g \circ f)(x_0) + (L_2 \circ L_1)(h) + o(\|h\|) + o(\|h\|) = \\ &= (g \circ f)(x_0) + (L_2 \circ L_1)(h) + o(\|h\|) . \end{aligned}$$

□

Veamos ahora que

La regla general de la cadena (39) consiste en varios casos particulares de la (27).

Sean $x_0 = (a_1, \dots, a_n)$ y $V \subseteq \mathbb{R}^k$ el abierto en el que g está definida. Para $\varepsilon > 0$ pequeño, están definidos los n caminos siguientes, posiblemente curvilíneos (el índice i va de 1 a n):

$$\alpha_i(t) : (-\varepsilon, \varepsilon) \longrightarrow V \quad , \quad \alpha_i(t) = f(x_0 + t \mathbf{e}_i) = f(a_1, \dots, a_{i-1}, a_i + t, a_{i+1}, \dots, a_n) .$$

Estos caminos satisfacen $\alpha_i(0) = y_0$ y $\alpha_i'(0) = f_{x_i}(x_0) = (Df)_{x_0} \mathbf{e}_i$. Pedir que g cumpla la regla de la cadena en $t = 0$ para estos n caminos:

$$\left. \frac{d}{dt} \right|_{t=0} g(\alpha_i(t)) = (Dg)_{y_0} \alpha_i'(0) \quad , \quad i = 1, \dots, n ,$$

es pedir que sea $(g \circ f)_{x_i}(x_0) = (Dg)_{y_0} f_{x_i}(x_0)$ para $i = 1, \dots, n$, y esto es exactamente lo mismo que pedir la siguiente igualdad matricial:

$$[(g \circ f)_{x_1} \mid (g \circ f)_{x_2} \mid \dots \mid (g \circ f)_{x_n}]_{x_0} = (Dg)_{y_0} [f_{x_1} \mid f_{x_2} \mid \dots \mid f_{x_n}]_{x_0} ,$$

que es la fórmula (39) cuando leemos $(D(g \circ f))_{x_0}$ columna a columna. La manera práctica de expresar esto es con la siguiente fórmula, en la que el índice \mathbf{i} toma cualquier valor entre 1 y n :

$$\left. \frac{\partial}{\partial x_{\mathbf{i}}} \right|_{x_0} g(f_1(x), \dots, f_k(x)) = g_{y_1}(f(x_0)) \frac{\partial f_1}{\partial x_{\mathbf{i}}}(x_0) + \dots + g_{y_k}(f(x_0)) \frac{\partial f_k}{\partial x_{\mathbf{i}}}(x_0) , \quad (40)$$

similar a la fórmula (27), con el operador $\frac{d}{dt}$ sustituido por $\left. \frac{\partial}{\partial x_{\mathbf{i}}} \right|_{x_0}$, para $\mathbf{i} = 1, \dots, n$.

Para hacer cálculos, todo lo que necesitas es (27). La fórmula (39) es útil en demostraciones.

2.3 Diferenciabilidad: funciones de clase \mathcal{C}^ℓ

Definición 91. Sea $U \subseteq \mathbb{R}^n$ un abierto. Se dice que $f : U \rightarrow \mathbb{R}^k$ es **diferenciable de clase \mathcal{C}^ℓ en U** si existen y son continuas en todo U las derivadas parciales de f de órdenes desde cero hasta ℓ (entendiendo que la derivada de orden cero es la propia f).

Decimos que f es **suave** o **de clase \mathcal{C}^∞** si es de clase \mathcal{C}^ℓ para todo ℓ .

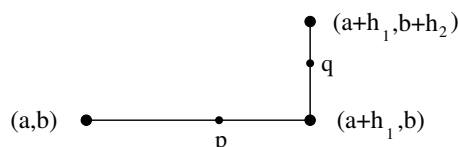
Decir $f \in \mathcal{C}^0$ es lo mismo que decir que f es continua. Es fácil ver que una función vectorial $\mathbf{f} \equiv (f_1, \dots, f_k)$ es \mathcal{C}^ℓ (incluido el caso $\ell = \infty$) si y sólo si sus entradas f_j son todas \mathcal{C}^ℓ .

Teorema 92. Sean U abierto de \mathbb{R}^n y $x_0 \in U$. Para que $f : U \rightarrow \mathbb{R}^k$ sea diferenciable en x_0 es suficiente (no necesario) que en un entorno de x_0 existan las funciones f_{x_1}, \dots, f_{x_n} y sean continuas en x_0 .

Como consecuencia, una función de clase \mathcal{C}^1 es diferenciable en todo punto de su dominio.

Demostración. Mirando las entradas de f una a una, vemos que basta demostrar el teorema cuando f es escalar. Vamos a probarlo para $n = 2$ y al final haremos un breve comentario sobre cómo se hace para n general.

Sea $f(x_1, x_2)$ función escalar definida en un abierto $U \subseteq \mathbb{R}^2$. Fijamos un punto $x_0 = (a, b) \in U$, tal que en una bola euclídea $B(x_0, r) \subseteq U$ existen las funciones f_{x_1}, f_{x_2} y son continuas en x_0 . Escribiendo un punto cualquiera $x \in B(x_0, r)$ como $x = (c, d) = (a + h_1, b + h_2)$, vemos que se puede pasar de x_0 a x por una línea poligonal de dos lados, uno horizontal y otro vertical.



La poligonal está contenida en la bola $\overline{B}(x_0, \|x - x_0\|_2) \subset B(x_0, r)$. Como $f(t, b)$ es una función *escalar* de una variable, le aplicamos el teorema de los incrementos finitos y resulta un punto $p = (\theta_1, b)$, en el segmento horizontal, tal que el incremento en dicho segmento $f(a + h_1, b) - f(a, b)$ es igual a $f_{x_1}(p) h_1$. De manera análoga, el incremento en el segmento vertical es $f(a + h_1, b + h_2) - f(a + h_1, b) = f_{x_2}(q) h_2$, con $q = (a + h_1, \theta_2)$ en dicho segmento. Definimos la función lineal $L(h_1, h_2) \equiv f_{x_1}(x_0) h_1 + f_{x_2}(x_0) h_2$ y calculamos:

$$f(a + h_1, b + h_2) = f(a, b) + f_x(p) h_1 + f_y(q) h_2 = f(a, b) + L(h_1, h_2) + \text{resto}.$$

siendo

$$\text{resto} = (f_{x_1}(p) - f_{x_1}(x_0)) h_1 + (f_{x_2}(q) - f_{x_2}(x_0)) h_2.$$

Como $|h_1|, |h_2| \leq \|(h_1, h_2)\|_2 = \|x - x_0\|_2$, tenemos

$$\begin{aligned} |\text{resto}| &\leq |f_{x_1}(p) - f_{x_1}(x_0)| |h_1| + |f_{x_2}(q) - f_{x_2}(x_0)| |h_2| \leq \\ &\leq (|f_{x_1}(p) - f_{x_1}(x_0)| + |f_{x_2}(q) - f_{x_2}(x_0)|) \|x - x_0\|_2, \end{aligned}$$

y tendremos $\text{resto} = o(\|x - x_0\|_2)$ si vemos que $f_{x_1}(p) - f_{x_1}(x_0)$ y $f_{x_2}(q) - f_{x_2}(x_0)$ tienden a cero cuando $x \rightarrow x_0$. Como los puntos p, q están en la bola $\overline{B}(x_0, \|x - x_0\|_2)$, se tiene $\|p - x_0\|_2, \|q - x_0\|_2 \leq \|x - x_0\|_2$ y así $p \rightarrow x_0$ y $q \rightarrow x_0$ cuando $x \rightarrow x_0$. Como f_{x_1} y f_{x_2} son continuas en el punto x_0 , se tiene

$$f_{x_1}(p) - f_{x_1}(x_0) \rightarrow 0 \quad \text{y} \quad f_{x_2}(q) - f_{x_2}(x_0) \rightarrow 0 \quad \text{cuando} \quad x \rightarrow x_0,$$

lo cual completa la demostración cuando $n = 2$.

Para n general utilizamos una poligonal de n lados, cada uno paralelo a un eje coordenado. \square

La condición suficiente de diferenciabilidad que proporciona el teorema 92 no es necesaria. Por ejemplo, la función $f : \mathbb{R} \rightarrow \mathbb{R}$ dada por

$$f(x) = \begin{cases} x^2 \sin \frac{1}{x^2} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

es diferenciable en $x = 0$, pero $f'(x)$ ni siquiera es acotada en el entorno de 0.

Teorema 93. (Lema de Schwarz). Sea $f(x_1, x_2)$ definida en un entorno de $x_0 = (a, b)$. Supongamos que las funciones f_{x_1}, f_{x_2} y $f_{x_1 x_2} = \frac{\partial}{\partial x_2} f_{x_1}$ existen cerca de x_0 , y que la función $f_{x_1 x_2}$ es continua en x_0 . Entonces existe $f_{x_2 x_1}(x_0) = \frac{\partial}{\partial x_1} \bigg|_{x_0} f_{x_2}$ y es igual a $f_{x_1 x_2}(x_0)$. Como consecuencia, si f es \mathcal{C}^2 entonces $f_{x_1 x_2}$ y $f_{x_2 x_1}$ son idénticas.

Demostración. Basta probarlo para f escalar. Para cada $h = (h_1, h_2)$ con $h_1 \neq 0 \neq h_2$, consideramos el siguiente rectángulo de lados paralelos a los ejes



Para h pequeño, todo el rectángulo está contenido en el entorno de x_0 donde existen $f_{x_1}, f_{x_2}, f_{x_1 x_2}$. Dando a los vértices los signos que se indican en la siguiente figura



definimos $\varphi(h) = [f(a, b) - f(a + h_1, b) - f(a, b + h_2) + f(a + h_1, b + h_2)] / (h_1 h_2)$.

Vamos a escribir $\varphi(h)$ de dos maneras muy parecidas. La primera es:

$$\varphi(h) = \frac{1}{h_2} \frac{[f(a + h_1, b + h_2) - f(a + h_1, b)] - [f(a, b + h_2) - f(a, b)]}{h_1}, \quad (41)$$

y la otra es:

$$\varphi(h) = \frac{1}{h_1} \cdot \left(\frac{f(a + h_1, b + h_2) - f(a + h_1, b)}{h_2} - \frac{f(a, b + h_2) - f(a, b)}{h_2} \right). \quad (42)$$

Definiendo la función $g(x_1) \equiv f(x_1, b + h_2) - f(x_1, b)$, convertimos la expresión (41) en:

$$\varphi(h) = \frac{1}{h_2} \cdot \frac{g(a + h_1) - g(a)}{h_1} = \frac{1}{h_2} \cdot g'(\theta_1) = \frac{1}{h_2} \cdot (f_{x_1}(\theta_1, b + h_2) - f_{x_1}(\theta_1, b)),$$

para cierto valor θ_1 situado entre a y $a + h_1$. Como $f_{x_1 x_2}$ existe en todo el rectángulo, hay un número θ_2 entre b y $b + h_2$, tal que $f_{x_1}(\theta_1, b + h_2) - f_{x_1}(\theta_1, b) = h_2 \cdot f_{x_1 x_2}(\theta_1, \theta_2)$, de donde $\varphi(h) = f_{x_1 x_2}(p)$, siendo $p = (\theta_1, \theta_2)$ un punto interior al rectángulo.

El rectángulo encoge cuando h tiende a $(0, 0)$, forzando $p \rightarrow x_0$ cuando $h \rightarrow (0, 0)$. Como $f_{x_1 x_2}$ es continua en x_0 , deducimos:

$$\lim_{\substack{h \rightarrow (0,0) \\ h_1 \neq 0 \\ h_2 \neq 0}} \varphi(h) = f_{x_1 x_2}(x_0). \quad (43)$$

Ahora vamos a utilizar la expresión (42) para $\varphi(h)$. Fijado un valor $h_1 \neq 0$, el siguiente límite está bien definido:

$$\begin{aligned} \lim_{\substack{h_2 \rightarrow 0 \\ h_2 \neq 0}} \varphi(h) &= \frac{1}{h_1} \cdot \lim_{h_2 \rightarrow 0} \left(\frac{f(a + h_1, b + h_2) - f(a + h_1, b)}{h_2} - \frac{f(a, b + h_2) - f(a, b)}{h_2} \right) = \\ &= \frac{1}{h_1} \cdot (f_{x_2}(a + h_1, b) - f_{x_2}(a, b)), \end{aligned}$$

lo cual junto con (43) implica que existe $\lim_{h_1 \rightarrow 0} \frac{f_{x_2}(a + h_1, b) - f_{x_2}(a, b)}{h_1}$ y es igual a $f_{x_1 x_2}(x_0)$.

Es decir que existe $\frac{\partial}{\partial x_1} \bigg|_{x_0} f_{x_2}$ y es igual a $f_{x_1 x_2}(x_0)$. \square

El lema de Schwarz implica que si f es \mathcal{C}^2 entonces se tiene $f_{x_i x_j} \equiv f_{x_j x_i}$ para i, j cualesquiera, es decir que para tal función el orden de derivación no importa en las derivadas segundas. Si f es \mathcal{C}^3 , aplicando el teorema de Schwarz varias veces deducimos identidades como las siguientes:

$$f_{xxy} = f_{xyx} = f_{yxx} \quad , \quad f_{xyz} = \begin{cases} f_{xzy} = f_{zxy} \\ f_{yxz} = f_{yzx} = f_{zyx} \end{cases}$$

y para una tal f el orden de derivación tampoco importa en las derivadas terceras. Ahora bien, la mayoría de las funciones \mathcal{C}^3 nos darán $f_{xxy} \neq f_{xyy}$, por ejemplo

$$f(x, y) = x^2 y \implies \begin{cases} f_{xxy} = 2 \\ f_{xyy} = 0 \end{cases}$$

Es, pues, muy importante cuántas veces se ha derivado respecto de cada variable independiente.

En general, en las derivadas de orden $\leq \ell$ de una función de clase \mathcal{C}^ℓ no importa el orden de derivación, pero sí importa (y mucho) cuántas veces se ha derivado respecto de cada variable independiente. Escribiendo $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ y $|\alpha| = \alpha_1 + \dots + \alpha_n$, se usa a veces la siguiente notación:

$$D^\alpha f \stackrel{\text{def}}{=} \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}} .$$

La n -upla de enteros no negativos α recibe el nombre de **multíndice** y codifica, precisamente, el número de veces que derivamos f respecto de cada x_i . A $|\alpha|$ se la llama **longitud de α** .

Claramente, la suma de funciones \mathcal{C}^ℓ es \mathcal{C}^ℓ . Lo mismo para la diferencia y para el producto de constante por función. Para el producto de dos funciones tenemos la **regla de Leibniz**, cuyo caso escalar dice así:

$$D(fg)_{x_0} = g(x_0) Df_{x_0} + f(x_0) Dg_{x_0} . \quad (44)$$

Demostración. La función producto $\text{prod} : \mathbb{R}^2 \rightarrow \mathbb{R}$, dada por $(x, y) \mapsto xy$, tiene jacobiana $D \text{prod} = [y \ x]$, claramente continua, luego prod es \mathcal{C}^1 y es diferenciable en todo punto de \mathbb{R}^2 . Dados un abierto $U \subseteq \mathbb{R}^n$ y $f, g : U \rightarrow \mathbb{R}$ diferenciables en $x_0 \in U$, el producto se describe como compuesta $fg \equiv \text{prod} \circ \begin{bmatrix} f \\ g \end{bmatrix}$ y podemos aplicar la regla de la cadena:

$$\begin{aligned} D(fg)_{x_0} &= D \text{prod}_{(x,y)=(f(x_0),g(x_0))} \cdot \begin{bmatrix} Df \\ Dg \end{bmatrix}_{x_0} = [y \ x]_{(x,y)=(f(x_0),g(x_0))} \cdot \begin{bmatrix} Df_{x_0} \\ Dg_{x_0} \end{bmatrix} = \\ &= g(x_0) Df_{x_0} + f(x_0) Dg_{x_0} . \end{aligned}$$

□

Proposición 94. Si las funciones escalares f, g son \mathcal{C}^ℓ en U , entonces fg también es \mathcal{C}^ℓ en U .

Demostración. Sabemos que es verdad para $\ell = 0$. Procedemos por inducción sobre ℓ . Sea $\ell > 0$ y supongamos cierta la proposición para funciones $\mathcal{C}^{\ell-1}$. Dadas f, g de clase \mathcal{C}^ℓ , la jacobiana del producto (vista como función matriz) está dada por

$$D(fg) \equiv g Df + f Dg ,$$

de donde se deduce, por la hipótesis de inducción, que $D(fg) \in \mathcal{C}^{\ell-1}$. Luego $fg \in \mathcal{C}^\ell$. □

De aquí se deduce el caso matricial de la misma proposición. Tenemos un abierto $U \subseteq \mathbb{R}^n$ y dos **funciones matriz**, es decir funciones con valores matrices:

$$A : U \longrightarrow M_{h \times k}(\mathbb{R}) \quad , \quad B : U \longrightarrow M_{k \times s}(\mathbb{R}) .$$

Si A, B son \mathcal{C}^ℓ , la función producto $AB : U \rightarrow M_{h \times s}(\mathbb{R})$ es también \mathcal{C}^ℓ porque sus entradas son sumas finitas de productos de entradas de A y B .

Proposición 95. *La compuesta de aplicaciones \mathcal{C}^ℓ es \mathcal{C}^ℓ .*

Demostración. Por inducción sobre ℓ . Es cierto para $\ell = 0$, en cuyo caso dice que la compuesta de continuas es continua. Sea ahora $\ell \geq 1$ y supongámoslo cierto para $\ell - 1$. Dadas f, g de clase \mathcal{C}^ℓ , en particular son \mathcal{C}^1 y diferenciables en todo punto. La regla de la cadena nos da entonces la siguiente identidad entre funciones matriz:

$$D(g \circ f) \equiv [(Dg) \circ f] Df.$$

El primer factor $(Dg) \circ f$ es la compuesta de una función \mathcal{C}^ℓ con una función matriz de clase $\mathcal{C}^{\ell-1}$, luego es $\mathcal{C}^{\ell-1}$ por la hipótesis de inducción. El segundo factor Df es otra función matriz de clase $\mathcal{C}^{\ell-1}$. Como el producto de funciones matriz de clase $\mathcal{C}^{\ell-1}$ es $\mathcal{C}^{\ell-1}$, tenemos $D(g \circ f) \in \mathcal{C}^{\ell-1}$. Luego $g \circ f \in \mathcal{C}^\ell$. \square

Una fórmula elemental define una aplicación \mathcal{C}^∞ en el abierto en el que no se anule ningún denominador, los radicandos y logaritmandos permanezcan positivos y las cantidades dentro de un valor absoluto o de la función sig no se anulen.

En particular, las dos funciones vistas la final del apartado 2.2.2 son \mathcal{C}^∞ en $\mathbb{R}^2 \setminus \{(0, 0)\}$, luego $(0, 0)$ es el único punto donde no son diferenciables.

La función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por

$$f(x, y) = \begin{cases} \frac{x^3 y}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases} \quad (45)$$

es \mathcal{C}^∞ en $\mathbb{R}^2 \setminus \{(0, 0)\}$, luego $f_{xy} = f_{yx}$ salvo quizás en $(x, y) = (0, 0)$.

Ejercicio 96. *Comprueba que para la función definida en (45) se tiene $f_{xy}(0, 0) \neq f_{yx}(0, 0)$. Por supuesto, para esta f las derivadas f_{xy} y f_{yx} son discontinuas en $(0, 0)$.*

El espacio $M_{n \times n}(\mathbb{R})$ de las matrices $n \times n$ puede identificarse con \mathbb{R}^{n^2} y entonces la función determinante $\det : M_{n \times n}(\mathbb{R}) \rightarrow \mathbb{R}$ es un polinomio (de grado n) y es \mathcal{C}^∞ .

El conjunto $\text{GL}(n, \mathbb{R})$ de las matrices invertibles $n \times n$ es la preimagen del abierto $\mathbb{R} \setminus \{0\}$ por la función determinante, luego es un abierto de \mathbb{R}^{n^2} . En este abierto la función $A \mapsto A^{-1}$, que lleva cada matriz a su inversa, es \mathcal{C}^∞ porque cada una de sus n^2 entradas es un cociente de dos polinomios y el denominador no se anula en el abierto.

2.4 Desarrollos de Taylor y extremos locales

Una función de una variable admite desarrollos de Taylor en cada punto. En este apartado vamos a hacer la misma construcción para funciones de varias variables.

En todo este apartado f es una función *escalar*.

2.4.1 Desarrollo cuadrático de Taylor

Definición 97. *Sean $U \subseteq \mathbb{R}^n$ un abierto y $a \in U$. Sea $f \in \mathcal{C}^2(U, \mathbb{R})$ una función escalar. La matriz hessiana de f en a es el cuadrado formado por la derivadas segundas de f en a :*

$$\text{Hess}(f)_a = [f_{x_i x_j}(a)]_{n \times n},$$

que, por el lema de Schwarz, es una matriz simétrica. La forma hessiana de f en a es la forma cuadrática correspondiente a esta matriz simétrica:

$$\mathbb{R}^n \ni v \mapsto v^t \text{Hess}(f)_a v = \sum_{1 \leq i, j \leq n} v_i v_j f_{x_i x_j}(a).$$

A partir de la función f , del punto a y de un vector cualquiera $h \in \mathbb{R}^n$, construimos la siguiente función escalar de una variable:

$$g(t) = f(a + th) , \quad (46)$$

que es de clase \mathcal{C}^2 en un intervalo abierto $I \subseteq \mathbb{R}$ con $0 \in I$. Aplicando la regla de la cadena a f , obtenemos:

$$g'(t) = (\nabla f)_{a+th} \cdot h = \sum_{1 \leq i \leq n} f_{x_i}(a + th) h_i , \quad (47)$$

y volviendo a aplicar la regla de la cadena, a cada una de las f_{x_i} , calculamos:

$$\begin{aligned} g''(t) &= \frac{d}{dt} g'(t) = \frac{d}{dt} \sum_{1 \leq i \leq n} h_i f_{x_i}(a + th) = \\ &= \sum_{1 \leq i \leq n} h_i \frac{d}{dt} f_{x_i}(a + th) = \sum_{1 \leq i, j \leq n} h_i h_j f_{x_i x_j}(a + th) , \end{aligned}$$

en definitiva:

$$g''(t) = \sum_{1 \leq i, j \leq n} f_{x_i x_j}(a + th) h_i h_j = h^t [\text{Hess}(f)_{a+th}] h . \quad (48)$$

Teorema 98. *En las condiciones de la definición anterior, tenemos un desarrollo:*

$$f(a + h) = f(a) + \frac{1}{1!} (\nabla f)_a \cdot h + \frac{1}{2!} h^t \text{Hess}(f)_a h + R_2(h) ,$$

donde el resto $R_2(h)$ es una $o(\|h\|^2)$, y es una $O(\|h\|^3)$ si f es \mathcal{C}^3 .

Demostración. Fijamos una bola euclídea $B(a, r)$ contenida en el dominio de f . Dado $x = a + h \in B(a, r)$, el segmento rectilíneo $[a, x]$ está contenido en $\overline{B}(a, \|h\|)$ y la función $g(t)$ dada por (46) está definida en un intervalo abierto I que contiene a $[0, 1]$. El teorema de Taylor para funciones de una variable nos dice que existe un valor intermedio $\theta \in (0, 1)$ tal que:

$$g(1) - g(0) = (1 - 0) g'(0) + \frac{1}{2!} (1 - 0)^2 g''(\theta) .$$

Escribiendo $p = a + \theta h$, tenemos:

$$\begin{aligned} f(a + h) - f(a) &= g'(0) + (1/2) g''(\theta) = (\nabla f)_a \cdot h + \frac{1}{2} h^t [\text{Hess}(f)_p] h = \\ &= (\nabla f)_a \cdot h + \frac{1}{2} h^t [\text{Hess}(f)_a] h + R_2 . \end{aligned}$$

donde $R_2 = (1/2) \sum_{1 \leq i, j \leq n} h_i h_j (f_{x_i x_j}(p) - f_{x_i x_j}(a))$.

Por otra parte, como para todo i es $|h_i| \leq \|h\|_2$, tenemos $|h_i h_j| \leq \|h\|_2^2$ para todo par i, j , luego:

$$|R_2| \leq \frac{\|h\|_2^2}{2} \sum_{1 \leq i, j \leq n} |f_{x_i x_j}(p) - f_{x_i x_j}(a)| . \quad (49)$$

Como $p \in \overline{B}(a, \|h\|)$, se tiene $p \rightarrow a$ cuando $h \rightarrow \mathbf{0}$ y, como cada función $f_{x_i x_j}$ es continua:

$$f_{x_i x_j}(p) - f_{x_i x_j}(a) \rightarrow 0 \quad \text{cuando } h \rightarrow \mathbf{0} , \quad \text{para todo par } i, j .$$

De esto y (49), deducimos que $R_2 = o(\|h\|^2)$.

Supongamos ahora que f no sólo es \mathcal{C}^2 sino que también es \mathcal{C}^3 , o sea que también existen las derivadas $f_{x_i x_j x_k}$ y son continuas.

Hemos obtenido la fórmula (48) aplicando la regla de la cadena dos veces. En general, si f es de clase \mathcal{C}^ℓ y $s \leq \ell$ podemos aplicar la regla de la cadena s veces y llegar a la siguiente fórmula:

$$\frac{d^s}{dt^s} f(a + th) = \sum_{1 \leq i_1, \dots, i_s \leq n} f_{x_{i_1} \dots x_{i_s}}(a + th) h_{i_1} \dots h_{i_s} . \quad (50)$$

Cuando f es de clase \mathcal{C}^3 , la función $g(t)$ dada por (46) es \mathcal{C}^3 y

$$g'''(t) = \sum_{1 \leq i, j, k \leq n} f_{x_i x_j x_k}(a + th) h_i h_j h_k .$$

Además, existe un valor $\theta' \in (0, 1)$ tal que

$$g(1) - g(0) = (1 - 0) g'(0) + \frac{1}{2} (1 - 0)^2 g''(0) + \frac{1}{3!} (1 - 0)^3 g'''(\theta') ,$$

lo que nos da una nueva fórmula para el resto R_2 de antes:

$$R_2 = \frac{1}{6} g'''(\theta') = \frac{1}{6} \sum_{1 \leq i, j, k \leq n} f_{x_i x_j x_k}(q) h_i h_j h_k ,$$

siendo $q = a + \theta' h \in \overline{B}(a, \|h\|_2)$. Entonces

$$|R_2| \leq \frac{1}{6} \left[\sum_{1 \leq i, j, k \leq n} |f_{x_i x_j x_k}(q)| \right] \|h\|_2^3 .$$

Pero las funciones $f_{x_i x_j x_k}$, siendo continuas, son acotadas en $\overline{B}(a, \|h\|_2)$; luego hay una constante K tal que $|R_2| \leq (1/6) K \|h\|_2^3$ y ahora el resto R_2 es una $O(\|h\|^3)$. \square

2.4.2 Extremos locales

Definición 99. Decimos que a es un **punto crítico de f** si se tiene $\nabla f_a = \mathbf{0}$, es decir

$$f_{x_1}(a) = f_{x_2}(a) = \cdots = f_{x_n}(a) = 0 .$$

Dada f , de clase \mathcal{C}^1 , para que a sea **máximo local** o **mínimo local** de f es *necesario* que sea punto crítico: si no es crítico, hay un vector h tal que la función g dada por (46) satisface $g'(0) \neq 0$ y existen valores t, t' , arbitrariamente pequeños, con $g(t) < g(0) < g(t')$; luego hay puntos p, q , arbitrariamente cerca de a , con $f(p) < f(a) < f(q)$.

Teorema 100. Sean f de clase \mathcal{C}^2 en un abierto $U \subseteq \mathbb{R}^n$ y $a \in U$, un punto crítico de f . Sea $A = \text{Hess}(f)_a$.

Si A es definida positiva, entonces a es un **mínimo local estricto** para f .

Si A es definida negativa, entonces a es un **máximo local estricto** para f .

Si A es indefinida (degenerada o no), entonces a no es ni **máximo local** ni **mínimo local** para f .

Demostración. Supongamos A definida positiva. Como A es simétrica, existen una base ortonormal $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ y números $0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ tales que

$$A \mathbf{u}_j = \lambda_j \mathbf{u}_j \quad , \quad j = 1, \dots, n .$$

Expresando el vector general $h \in \mathbb{R}^n$ en esa base $h = y_1 \mathbf{u}_1 + \cdots y_n \mathbf{u}_n$, tenemos:

$$\begin{aligned} h^t h &= y_1^2 + \cdots + y_n^2 , \\ h^t A h &= \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 , \end{aligned}$$

de donde

$$h^t A h \geq \lambda_1 h^t h = \lambda_1 \|h\|_2^2 . \quad (51)$$

Por otra parte, usando el teorema 98:

$$f(a + h) = f(a) + 0 + \frac{1}{2} h^t A h + o(\|h\|_2^2) \geq f(a) + \frac{\lambda_1}{2} \|h\|_2^2 + o(\|h\|_2^2) . \quad (52)$$

Tomemos un trocito de λ_1 , por ejemplo el uno por ciento. Cuando h es pequeño, la función $(0'01)\lambda_1 \|h\|_2^2$ supera a la $o(\|h\|_2^2)$ de (52). Esto implica que hay un $r > 0$ tal que

$$a + h \in B(a, r) \implies (0'01)\lambda_1 h^t A h + o(\|h\|_2^2) \geq 0 ,$$

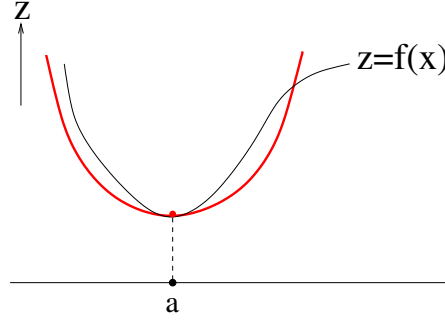
que junto con (52) nos da:

$$a + h \in B(a, r) \implies f(a + h) \geq f(a) + (0'49)\lambda_1 h^t h .$$

Para $x = a + h$ cerca de a , el grafo $\{(a + h, z) : z = f(a + h)\} \subset \mathbb{R}^{n+1}$ está por encima del grafo

$$\{(a + h, z) : z = f(a) + (0'49)\lambda_1 h^t h\} ,$$

que es un **paraboloide** con vértice en $(a, f(a))$ y faldas hacia arriba,



esto obliga a f a tener un mínimo local estricto en a .

Se procede igual cuando A es definida negativa.

Si A es indefinida, existen dos vectores $v, w \in \mathbb{R}^n$ tales que $v^t A v < 0$ y $w^t A w > 0$. Las funciones

$$g_1(t) = f(a + tv) \quad , \quad g_2(t) = f(a + tw) ,$$

satisfacen:

$$g_1'(0) = g_2'(0) = 0 \quad , \quad g_1''(0) < 0 \quad , \quad g_2''(0) > 0 ,$$

luego hay valores t , arbitrariamente pequeños, con $g_1(t) < g_1(0) = g_2(0) < g_2(t)$, o sea:

$$f(a + tv) < f(a) < f(a + tw) ,$$

y así hallamos puntos p, q , arbitrariamente cerca de a , con $f(p) < f(a) < f(q)$. \square

Advertencia. Si A es **semidefinida** (positiva o negativa) pero degenerada, la información que nos dan las derivadas segundas en el punto a es *insuficiente* para decidir qué clase de punto crítico es a ; hay que examinar la función en mayor profundidad. Por ejemplo, las tres funciones

$$f(x, y) = x^3 - y^2 \quad , \quad g(x, y) = -x^4 - y^2 \quad , \quad h(x, y) = -e^x y^2 ,$$

tienen un punto crítico en $(0, 0)$ con matriz hessiana $\begin{bmatrix} 0 & 0 \\ 0 & -2 \end{bmatrix}$. Este punto no es máximo local ni mínimo local de f , es máximo local estricto de g y es máximo local no estricto de h .

2.4.3 Desarrollo general de Taylor

Recordemos la notación con multíndices $D^\alpha f$ que hemos introducido en el apartado 2.3.

Teorema-definición 101. Sean un abierto $U \subseteq \mathbb{R}^n$, una función escalar $f(x) \in \mathcal{C}^\ell(U)$ y un punto $a \in U$. Existe un **único** polinomio $P_\ell(x)$ que cumple las siguientes condiciones:

1. $P_\ell(x)$ tiene grado $\leq \ell$.
2. Se tiene $D^\alpha f(a) = D^\alpha P_\ell(a)$ para todo multíndice $\alpha = (\alpha_1, \dots, \alpha_n)$ con $0 \leq |\alpha| \leq \ell$.

A este único polinomio $P_\ell(x)$ se le llama **polinomio de Taylor de orden ℓ de f en a** . \square

Veamos una fórmula para este polinomio. Fijados f y a , para cada $s = 0, 1, \dots, \ell$, hay un polinomio ϕ_s , homogéneo de grado s , definido por la siguiente fórmula:

$$\phi_s(h) = \left. \frac{d^s}{dt^s} \right|_{t=0} f(a + th) \quad , \quad \text{para todo } h \in \mathbb{R}^n .$$

La fórmula (47) implica que $\phi_1 = (df)_a$. La fórmula (48) implica que $\phi_2(h) = h^t \text{Hess}(f)_a h$. La fórmula (50) implica que, para $0 \leq s \leq \ell$, se tiene:

$$\phi_s(h) = \sum_{1 \leq i_1, \dots, i_s \leq n} f_{x_{i_1} \dots x_{i_s}}(a) h_{i_1} \dots h_{i_s} . \quad (53)$$

Entonces el polinomio de Taylor de orden ℓ viene dado por la siguiente igualdad:

$$P_\ell(x) = \phi_0 + \frac{1}{1!} \phi_1(x - a) + \frac{1}{2!} \phi_2(x - a) + \dots + \frac{1}{\ell!} \phi_\ell(x - a) , \quad (54)$$

equivalente a esta otra:

$$P_\ell(a + h) = f(a) + \nabla f_a \cdot h + \frac{1}{2} h^t \text{Hess}(f)_a h + \frac{1}{3!} \phi_3(h) + \dots + \frac{1}{\ell!} \phi_\ell(h) . \quad (55)$$

Podemos llamar a (53) la “fórmula con repeticiones” porque en ella un mismo monomio aparece varias veces. Por ejemplo, si $s = 5$ y escribimos $c = f_{x_1 x_1 x_2 x_2 x_2}(a)$, entonces el monomio $c h_1^2 h_2^3$ aparece $10 = \binom{5}{2} = \frac{5!}{2!3!}$ veces en la suma (53):

$$c h_1 \mathbf{h}_1 h_2 h_2 h_2 + c h_1 h_2 \mathbf{h}_1 h_2 h_2 + c h_1 h_2 h_2 \mathbf{h}_1 h_2 + \dots + c h_2 h_2 h_2 h_1 h_1 .$$

Si agrupamos esas repeticiones en un solo monomio, entonces la parte de $\frac{1}{5!} \phi_5(h)$ de grado 2 en h_1 y grado 3 en h_2 es:

$$\frac{1}{5!} [c h_1 h_1 h_2 h_2 h_2 + c h_1 h_2 h_1 h_2 h_2 + \dots + c h_2 h_2 h_2 h_1 h_1] = \frac{\frac{5!}{2!3!}}{5!} c h_1^2 h_2^3 = \frac{1}{2!3!} c h_1^2 h_2^3 .$$

Se demuestra que dividir por $s!$ y juntar monomios iguales en (53) da el siguiente resultado:

$$\frac{1}{s!} \phi_s(h) = \sum_{|\alpha|=s} \frac{1}{\alpha_1! \dots \alpha_n!} (D^\alpha f)_a h_1^{\alpha_1} \dots h_n^{\alpha_n} , \quad (56)$$

que puede llamarse la “fórmula sin repeticiones” para $(1/s!) \phi_s$. Mientras que (53) es una suma de n^s sumandos, la suma (56) sólo tiene $\binom{n+s-1}{s-1}$ sumandos. Es habitual escribir:

$$|\alpha| = \alpha_1! \dots \alpha_n! \quad , \quad h^\alpha = h_1^{\alpha_1} \dots h_n^{\alpha_n} .$$

La “fórmula sin repeticiones” para el polinomio de Taylor es, entonces

$$P_\ell(a + h) = \sum_{0 \leq |\alpha| \leq \ell} \frac{1}{\alpha!} (D^\alpha f)_a h^\alpha . \quad (57)$$

Por ejemplo, si $n = 2$ y $a = (a_1, a_2)$, la fórmula con repeticiones para el polinomio de orden 2 es ésta (7 sumandos):

$$P_2(x, y) = f(a) + f_x(a)(x - a_1) + f_y(a)(y - a_2) + \frac{1}{2} [f_{xx}(a)(x - a_1)^2 + f_{xy}(a)(x - a_1)(y - a_2) + f_{yx}(a)(y - a_2)(x - a_1) + f_{yy}(a)(y - a_2)^2] ,$$

mientras que la fórmula sin repeticiones es esta otra (mismo polinomio, 6 sumandos):

$$P_2(x, y) = f(a) + f_x(a)(x - a_1) + f_y(a)(y - a_2) + \frac{1}{2} f_{xx}(a)(x - a_1)^2 + f_{xy}(a)(x - a_1)(y - a_2) + \frac{1}{2} f_{yy}(a)(y - a_2)^2 .$$

Importante. Se utilizan indistintamente las fórmulas (54), (55) o (57) para el polinomio de Taylor. Esto puede causarte confusión. Ante cualquier duda, debes comprobar que el polinomio que estés manejando cumple las dos condiciones de la definición 101: si no cumple las dos, no es el polinomio de Taylor y hay que revisarlo.

Las fórmulas (54), (55) y (57) se usan casi siempre para demostrar teoremas que involucran polinomios de Taylor. Para calcular el polinomio de Taylor de una función concreta, lo que suele usarse es el siguiente resultado.

Teorema 102. Sean f de clase \mathcal{C}^ℓ en un abierto $U \subseteq \mathbb{R}^n$ y $a \in U$. Sea $P(x)$ un polinomio. Son equivalentes:

- (a) $P(x)$ es el polinomio de Taylor de orden ℓ de f en a .
- (b) $P(x)$ tiene grado $\leq \ell$ y la diferencia $f(x) - P(x)$ es una $o(\|x - a\|^\ell)$.

Además, si f es $\mathcal{C}^{\ell+1}$ y $P_\ell(x)$ es su polinomio de Taylor de orden ℓ en a , entonces

$$f(x) - P_\ell(x) = O(\|x - a\|^{\ell+1}).$$

En particular, si f es \mathcal{C}^∞ , entonces para cada ℓ el correspondiente polinomio de Taylor de f en a es el único polinomio $P_\ell(x)$ de grado $\leq \ell$ tal que $f(x) = P_\ell(x) + O(\|x - a\|^{\ell+1})$.

Atención. La hipótesis $f \in \mathcal{C}^\ell$ del teorema 102 es muy importante. Por ejemplo, la función

$$f(x) = \begin{cases} x^4 \operatorname{sen} \frac{1}{x^4} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

es derivable, pero su derivada no es acotada en el entorno de $x = 0$. Por lo tanto, no existe la segunda derivada, mucho menos el polinomio de Taylor de orden 3 en $x = 0$. Sin embargo, se tiene $f(x) = 0 + O(|x|^4)$.

El teorema 102 nos ahorra el cálculo de las derivadas de $f(x)$. Veamos un ejemplo.

Consideramos la función $f(x, y) = \frac{e^{xy}}{1 - y^2}$, que es \mathcal{C}^∞ en el entorno de $(0, 0)$.

Empezamos por recordar el desarrollo $e^t = 1 + t + (1/2)t^2 + O(|t|^3)$, que nos da:

$$e^{xy} = 1 + xy + \frac{1}{2}x^2y^2 + O(|xy|^3) = 1 + xy + \frac{1}{2}x^2y^2 + O(\|(x, y)\|^6).$$

Recordamos también que $\frac{1}{1-t} = 1 + t + t^2 + O(|t|^3)$, de donde:

$$\frac{1}{1-y^2} = 1 + y^2 + y^4 + O(|y|^6) = 1 + y^2 + y^4 + O(\|(x, y)\|^6).$$

Por lo tanto:

$$\begin{aligned} f(x, y) &= \left[1 + xy + \frac{1}{2}x^2y^2 + O(\|(x, y)\|^6) \right] \cdot \left[1 + y^2 + y^4 + O(\|(x, y)\|^6) \right] = \\ &= 1 + xy + y^2 + \frac{1}{2}x^2y^2 + xy^3 + y^4 + O(\|(x, y)\|^6). \end{aligned}$$

Ahora el teorema 102 nos dice que el polinomio de Taylor de orden 5 de f en $(0, 0)$ es el siguiente:

$$P_5(x, y) = 1 + xy + y^2 + \frac{1}{2}x^2y^2 + xy^3 + y^4,$$

y no hemos tenido que calcular ninguna derivada de f para obtener este resultado.

Las derivadas quintas de f en $(0, 0)$ son todas nulas, por eso P_5 tiene grado 4.

Demostración del teorema 102. Sea $P_\ell(x)$ el verdadero polinomio de Taylor de orden ℓ de f en a . Sea $P(x)$ un polinomio de grado $\leq \ell$.

Que $P_\ell(x)$ difiere de $f(x)$ en una $o(\|x - a\|^\ell)$, se prueba igual que el teorema 98. Por lo tanto, si $P(x)$ cumple la condición (b) del teorema 102 entonces:

$$P(x) - P_\ell(x) = o(\|x - a\|^\ell) . \quad (58)$$

Veamos que esto implica $P(x) - P_\ell(x) \equiv 0$. Tomando la descomposición en partes homogéneas:

$$P(a + h) - P_\ell(a + h) \equiv \varphi_0 + \varphi_1(h) + \varphi_2(h) + \cdots + \varphi_\ell(h) ,$$

de (58) deducimos primero que $\varphi_0 = 0$. Entonces tenemos:

$$\varphi_1(h) = -\varphi_2(h) - \cdots - \varphi_\ell(h) + o(\|h\|^\ell) = O(\|h\|^2) + o(\|h\|^\ell) = o(\|h\|^1) ,$$

que implica $\varphi_1(h) \equiv 0$ por la proposición 84. Esto, a su vez, permite deducir $\varphi_2(h) = o(\|h\|^2)$ y de nuevo $\varphi_2(h) \equiv 0$ por la proposición 84. Así hasta llegar a $\varphi_\ell(h) \equiv 0$. En definitiva $P(x) - P_\ell(x) \equiv 0$, es decir $P(x) \equiv P_\ell(x)$.

Si $f(x) \in \mathcal{C}^{\ell+1}(U)$, se obtiene:

$$\begin{aligned} f(x) &= P_{\ell+1}(x) + o(\|x - a\|^{\ell+1}) = P_\ell(x) + \frac{1}{(\ell+1)!} \phi_{\ell+1}(x - a) + o(\|x - a\|^{\ell+1}) = \\ &= P_\ell(x) + O(\|x - a\|^{\ell+1}) . \end{aligned}$$

□

3 Funciones inversas e implícitas

3.1 Aplicaciones contractivas y punto fijo

3.1.1 Espacios completos

Definiciones 103. Un espacio métrico (X, d) es **completo** si toda sucesión $\{x_n\}_{n=1}^\infty \subset X$ que sea de Cauchy para la distancia d es convergente para dicha distancia.

Si un espacio vectorial normado $(V, \|\cdot\|)$ es completo para la distancia $d(x, x') = \|x - x'\|$, decimos que es un **espacio de Banach**.

En un espacio completo tenemos la ventaja de poder decir si una sucesión tiene límite, sin necesidad de conocer explícitamente ese límite.

Ejemplo 1. El teorema 72 del apartado 1.3 nos dice que cada espacio numérico \mathbb{R}^n , dotado de cualquier norma, es completo con la distancia $d(x, x') = \|x - x'\|$; luego es un espacio de Banach. De hecho cualquier cerrado $C \subseteq \mathbb{R}^n$ es completo con esa distancia: un espacio métrico puede ser completo sin ser un espacio vectorial.

Ejemplo 2. Todo espacio métrico compacto es completo. El recíproco no es cierto porque, por ejemplo, hay muchos cerrados no compactos en \mathbb{R}^n y todos son completos.

Ejemplos de lo contrario: cualquier subconjunto $X \subset \mathbb{R}^n$ que no sea cerrado es incompleto con la distancia inducida por una norma cualquiera en \mathbb{R}^n . Ejemplos típicos: $\mathbb{Q}, (0, 1] \subset \mathbb{R}$.

3.1.2 Aplicaciones contractivas

Empezamos con una definición informal:

Decimos que una aplicación entre espacios métricos $f : (X, d_X) \rightarrow (Y, d_Y)$ es

Lipschitz-pequeña si es una aplicación Lipschitz con constante de Lipschitz menor que 1, es decir, si existe ε tal que:

$$\varepsilon < 1 \quad \text{y} \quad d_Y(f(x), f(x')) \leq \varepsilon d_X(x, x') \quad \text{para cualesquiera } x, x' \in X. \quad (59)$$

Definición 104. Una **aplicación contractiva** es una aplicación f que cumple las tres condiciones siguientes:

1. El dominio es igual al codominio: $f : X \rightarrow X$.
2. La función distancia es la misma en salida y en llegada: $f(X, d) \rightarrow (X, d)$.
3. f es Lipschitz-pequeña: $d(f(x), f(x')) \leq \varepsilon \cdot d(x, x')$, con $\varepsilon < 1$.

Veamos ejemplos de lo contrario, para que quede bien claro lo exigente que es esta definición.

Ej 1. La función $f : [3, +\infty) \rightarrow [0, 1/2]$, dada por $f(x) = 1/x$, satisface

$$|f(x) - f(x')| \leq \frac{1}{9} |x - x'| \quad , \quad \text{para cualesquiera } x, x' \in [3, +\infty) .$$

Por lo tanto, es Lipschitz-pequeña. Pero no es contractiva, porque su dominio es distinto de su codominio.

Ej 2. Si en $[1, +\infty)$ ponemos las distancias $d_1(x, x') = |x - x'|$ y $d_2(y, y') = |y - y'|/10$, entonces la aplicación

$$g : ([1, +\infty), d_1) \longrightarrow ([1, +\infty), d_2) \quad , \quad g(x) = 2x ,$$

es Lipschitz-pequeña pero no es contractiva, porque d_1 es distinta de d_2 . De hecho g dilata en vez de contraer.

Ej 3. La función $h : \mathbb{R} \rightarrow \mathbb{R}$, dada por $h(x) = 7 + \sqrt{x^2 + 1}$, satisface:

$$|h(x) - h(x')| < |x - x'| \quad \text{para cualesquiera } x, x' \in \mathbb{R} \text{ con } x \neq x' ,$$

pero no es Lipschitz-pequeña, porque su constante mínima de Lipschitz es exactamente 1.

3.1.3 Punto fijo

Teorema 105. (Teorema de la aplicación contractiva). Si $f : (X, d) \rightarrow (X, d)$ es contractiva y (X, d) es completo, entonces f tiene un **único punto fijo**, es decir que existe un único $p \in X$ tal que $f(p) = p$.

Demostración.

Existencia. Empecemos con cualquier punto $x_0 \in X$ y consideremos la sucesión $\{x_n\}_{n=0}^{\infty}$ que se define de la manera siguiente:

$$x_1 = f(x_0) \quad , \quad x_2 = f(x_1) \quad , \quad x_3 = f(x_2) \quad , \quad \dots$$

es decir $x_{n+1} = f(x_n)$ para $n = 0, 1, 2, 3, \dots$. Esto equivale a:

$$x_1 = f(x_0) \quad , \quad x_2 = f \circ f(x_0) = f^2(x_0) \quad , \quad x_3 = f \circ f \circ f(x_0) = f^3(x_0) \quad , \quad \dots$$

y en general:

$$x_n = \underbrace{f \circ \dots \circ f}_{n \text{ factores}}(x_0) = f^n(x_0) \quad \text{para } n \geq 1.$$

Observa lo importante que es, para poder definir las f^n , que el conjunto X de salida de f sea igual a su conjunto de llegada. Es decir, que el dominio y el codominio de f coincidan.

Sea $\varepsilon < 1$ una constante de Lipschitz para f . Se tiene:

$$d(x_1, x_2) = d(f(x_0), f(x_1)) \leq \varepsilon d(x_0, x_1) ,$$

y de igual modo:

$$d(x_2, x_3) \leq \varepsilon d(x_1, x_2) \leq \varepsilon^2 d(x_0, x_1) \quad , \quad d(x_3, x_4) \leq \varepsilon d(x_2, x_3) \leq \varepsilon^3 d(x_0, x_1) \quad , \quad \dots$$

y en general:

$$d(x_n, x_{n+1}) \leq \varepsilon^n d(x_0, x_1) \quad , \quad \text{para todo } n \geq 0.$$

Observa lo importante que es, para obtener estas desigualdades, que en la desigualdad general

$$d(f(x), f(x')) \leq \varepsilon d(x, x') ,$$

la “ d ” en el término de la izquierda sea la misma que en el término de la derecha. Es decir, que la función distancia d en el espacio métrico de salida de f sea la misma que en el espacio de llegada.

Estimemos la distancia entre dos términos de la sucesión no necesariamente consecutivos:

$$\begin{aligned} d(x_n, x_{n+s}) &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \dots + d(x_{n+s-1}, x_{n+s}) \leq \\ &\leq (\varepsilon^n + \varepsilon^{n+1} + \dots + \varepsilon^{n+s-1}) d(x_0, x_1) \leq d(x_0, x_1) \sum_{j=n}^{\infty} \varepsilon^j = d(x_0, x_1) \frac{\varepsilon^n}{1 - \varepsilon} . \end{aligned}$$

Vemos que $d(x_n, x_{n+s})$ desciende a cero como $\text{constante} \cdot \varepsilon^n$, luego $\{x_n\}_{n=0}^{\infty}$ es una sucesión de Cauchy y, como (X, d) es completo por hipótesis, hay un punto $p \in X$ con $\{x_n\} \rightarrow p$ cuando $n \rightarrow \infty$.

La aplicación f , siendo Lipschitz, es continua. Entonces de $\{x_n\} \rightarrow p$ se deduce

$$\{f(x_0), f(x_1), f(x_2), \dots\} \rightarrow f(p) ,$$

que es lo mismo que:

$$\{x_1, x_2, x_3, \dots\} \rightarrow f(p) .$$

La cola $\{x_1, x_2, x_3, \dots\}$ de $\{x_n\}$ converge tanto a p como a $f(p)$. Por la unicidad del límite, deducimos que $p = f(p)$ y así p es un punto fijo para f .

Unicidad. Sea $x \in X$ cualquier punto fijo, es decir $f(x) = x$. Razonamos así:

$$d(p, x) = d(f(p), f(x)) \leq \varepsilon d(p, x),$$

de donde $(1 - \varepsilon) d(p, x) \leq 0$ y, como es $1 - \varepsilon > 0$, se deduce $d(p, x) \leq 0$. Pero siempre es $d(p, x) \geq 0$, luego $d(p, x) = 0$ y, por lo tanto, es $p = x$. Esto prueba que p es el único punto fijo que tiene f . \square

Es interesante que esta demostración es *constructiva*: el punto fijo se aproxima, a velocidad exponencial, por los puntos $x_n = f^n(x_0)$.

Se llama **iterados del punto x_0** a los puntos $f^n(x_0)$. Se llama **iterados de f** a las funciones f^n .

El teorema funciona en cualquier espacio métrico completo. En los apartados siguientes lo vamos a aplicar al caso de una bola cerrada $X = \overline{B}(x_0, r) \subset \mathbb{R}^n$.

Repasamos ahora los ejemplos del apartado 3.1.2. Se verá lo esenciales que son todas las condiciones en la definición 104.

La función f del ejemplo 1 no tiene punto fijo en $[3, +\infty)$ (a pesar de que este espacio es completo). No puede tenerlo, porque esta función tiene dominio y codominio disjuntos.

La función g del ejemplo 2 no tiene punto fijo en $[1, +\infty)$ (a pesar de que este espacio es completo), es decir que $2x = x$ no tiene solución en $[1, +\infty)$. Esto no contradice al teorema 105, porque hemos usado métricas distintas en salida y en llegada.

Para la función h del ejemplo 3, se tiene $h(x) > 7 + x$ y por lo tanto $|h(x) - x| > 7$. A pesar de que h es “casi contractiva” (en el sentido de que $|h'(x)|$ sube hasta el valor 1 sólo cuando $x \rightarrow \infty$), resulta que el valor $h(x)$ *ni siquiera llega a acercarse a x* : $h(x)$ se mantiene a distancia mayor que 7 de x , para todo $x \in \mathbb{R}$.

3.2 Cómo obtener una constante de Lipschitz

En el caso de una función diferenciable, se obtienen constantes de Lipschitz a partir de una cota de las derivadas parciales primeras. Cuando esas derivadas sean lo bastante pequeñas, vamos a poder asegurar que la función es Lipschitz-pequeña.

Tenemos un abierto $U \subseteq \mathbb{R}^n$ y una función $f : U \rightarrow \mathbb{R}^k$ de clase \mathcal{C}^1 . Consideramos la restricción $f|_E$, donde $E \subseteq U$ es un subconjunto *convexo*.

Primer método: mirar cada entrada de f por separado.

El teorema de los incrementos finitos *sólo es válido para funciones escalares*. Esto estaba implícito en las demostraciones que hicimos de los teoremas 92 y 93 del apartado 2.3. Por lo tanto, damos nombre a las entradas de f : $f(x) \equiv (f_1(x), \dots, f_k(x))$ y estudiamos cada f_j individualmente.

Sea $j \in \{1, \dots, k\}$. Como E es convexo, dados $x, x' \in E$ se tiene $(1 - t)x + tx' \in E$ para todo $t \in [0, 1]$ y podemos definir la función:

$$g_j(t) = f_j((1 - t)x + tx') = f_j(x + t(x' - x)) \quad , \quad t \in [0, 1] .$$

Como g_j es escalar, le podemos aplicar el teorema de los incrementos finitos y existe $\theta_j \in (0, 1)$ tal que

$$f_j(x') - f_j(x) = g_j(1) - g_j(0) = (1 - 0) g'_j(\theta_j) = (x' - x) \cdot (\nabla f_j)_{p_j} ,$$

siendo $p_j = x + \theta_j(x' - x) \in E$. Así se hallan puntos $p_1, \dots, p_k \in E$ tales que:

$$|f_j(x') - f_j(x)| \leq \|x' - x\|_\infty \|(\nabla f_j)_{p_j}\|_1 \quad , \quad j = 1, \dots, k .$$

Si el siguiente número es finito

$$K_1 = \sup_{\substack{p \in E \\ 1 \leq j \leq k}} \|(\nabla f_j)_p\|_1 ,$$

entonces:

$$\|f(x') - f(x)\|_\infty \leq K_1 \cdot \|x' - x\|_\infty \quad \text{para cualesquiera } x, x' \in E,$$

y así K_1 es una constante de Lipschitz para $f : E \rightarrow \mathbb{R}^k$, cuando en salida y en llegada se usa la métrica $d(x, x') = \|x - x'\|_\infty$.

Segundo método: utilizar integrales vectoriales.

Dada una función vectorial de una sola variable $v(t) \equiv (v_1(t), \dots, v_k(t))$, su integral en el intervalo $[t_0, t_1]$ es el siguiente vector de \mathbb{R}^k

$$\int_{t_0}^{t_1} v(t) dt \stackrel{\text{def}}{=} \left(\int_{t_0}^{t_1} v_1(t) dt, \int_{t_0}^{t_1} v_2(t) dt, \dots, \int_{t_0}^{t_1} v_k(t) dt \right).$$

Proposición 106. Para cualquier norma $\|\cdot\|$ en \mathbb{R}^k , se cumple la siguiente desigualdad:

$$\left\| \int_{t_0}^{t_1} v(t) dt \right\| \leq \int_{t_0}^{t_1} \|v(t)\| dt.$$

Se demuestra fácilmente aproximando las integrales por sumas de Riemann. \square

Sean $x, x' \in E$ y pongamos $h = x' - x$. Al ser E convexo, cada punto $x + th$ con $t \in [0, 1]$ está en $E \subseteq U$ y tenemos la siguiente igualdad:

$$f(x') - f(x) = \int_0^1 \frac{d}{dt} f(x + th) dt = \int_0^1 (Df)_{x+th} h dt.$$

Sea $\|\cdot\|_0$ una norma en \mathbb{R}^n . Para cada matriz $A \in M_{k \times n}(\mathbb{R})$ denotamos por $\|A\|$ la norma como operador $A : (\mathbb{R}^n, \|\cdot\|_0) \rightarrow (\mathbb{R}^k, \|\cdot\|)$. Si el siguiente número es finito

$$K_0 = \sup_{p \in E} \|(Df)_p\|,$$

entonces aplicamos la proposición 106 y deducimos:

$$\begin{aligned} \|f(x') - f(x)\| &= \left\| \int_0^1 (Df)_{x+th} h dt \right\| \leq \int_0^1 \|(Df)_{x+th} h\| dt \leq \\ &\leq \int_0^1 \|(Df)_{x+th}\| \cdot \|h\|_0 dt \leq \|h\|_0 \int_0^1 K_0 dt = K_0 \|x' - x\|_0, \end{aligned}$$

y K_0 es una constante de Lipschitz para $f|_E : E \rightarrow \mathbb{R}^k$, cuando en salida se usa $d_0(x, x') = \|x - x'\|_0$ y en llegada se usa $d(y, y') = \|y - y'\|$.

El segundo método sirve para cualquier par de normas $\|\cdot\|_0$ y $\|\cdot\|$. Nos interesa especialmente el caso $k = n$ y las dos normas iguales a la euclídea estándar $\|\cdot\|_2$ en \mathbb{R}^n . Vimos en el apartado 1.1.15 que, para calcular la norma de $A_{n \times n}$ como operador $A : (\mathbb{R}^n, \|\cdot\|_2) \rightarrow (\mathbb{R}^n, \|\cdot\|_2)$, es preciso calcular autovalores de $A^t A$. Esto es difícil en cuanto $n \geq 3$. Lo que hacemos es hallar una cota superior, es decir un número K tal que

$$\|(Df)_p\| \leq K \quad \text{para todo } p \in E, \tag{60}$$

que equivale a $K \geq K_0$. Estaremos hallando para $f|_E$ una constante de Lipschitz que, por lo común, será mayor que la óptima. Con un poco de suerte, esta K todavía servirá para nuestro propósito y nos conformaremos con ella.

Haremos uso repetido de la desigualdad

$$|ab| \leq \frac{1}{2}(a^2 + b^2), \quad \text{o bien } |2ab| \leq a^2 + b^2, \tag{61}$$

o de la siguiente versión, más astuta, de (61):

$$|ab| = \left| \frac{a}{\sqrt{\alpha}}(\sqrt{\alpha} \cdot y) \right| \leq \frac{1}{2} \left(\frac{a^2}{\alpha} + \alpha b^2 \right), \quad \text{para todo } \alpha > 0. \tag{62}$$

Ejemplo de matriz constante. Hallemos una cota superior K para $\|A\|$ cuando $A = \begin{bmatrix} 1 & 2 \\ -7 & 4 \end{bmatrix}$. Buscamos un número $c > 0$ tal que para todo vector $h = (a, b) \in \mathbb{R}^2$ se tenga

$$\|Ah\|_2^2 \leq c \cdot \|h\|_2^2, \quad (63)$$

y así $\|Ah\|_2 \leq \sqrt{c} \cdot \|h\|_2$, con lo cual $K = \sqrt{c}$ será la cota que buscamos. Calculamos:

$$\|Ah\|_2^2 = (a + 2b)^2 + (-7a + 4b)^2 = 50a^2 - 52ab + 20b^2.$$

La desigualdad (61) nos permite escribir $|-52ab| \leq 26a^2 + 26b^2$, luego

$$\|Ah\|_2^2 \leq (50 + 26)a^2 + (20 + 26)b^2 = 76a^2 + 46b^2 \leq 76 \|h\|_2^2.$$

Finalmente $K = \sqrt{76} = 8'72$ es una cota superior para $\|A\|$, obtenida sin calcular autovalores. Para mejorar esta cota, probamos a aplicar la desigualdad (62) con $\alpha = 3/2$. Ahora tenemos:

$$\begin{aligned} |-52ab| &\leq \frac{1}{2} \cdot \left(\frac{52}{3/2} a^2 + (3/2) \cdot 52b^2 \right) \leq 18a^2 + 39b^2, \\ \|Ah\|_2^2 &\leq (50 + 18)a^2 + (20 + 39)b^2 = 68a^2 + 59b^2, \end{aligned}$$

luego $\|Ah\|_2^2 \leq 68 \|h\|_2^2$ y la cota mejorada es $K_{\text{mejorada}} = \sqrt{68} = 8'25$.

Veamos cuál es el valor óptimo. Los autovalores de $A^t A = \begin{bmatrix} 50 & -26 \\ -26 & 20 \end{bmatrix}$ son

$$\frac{70 \pm \sqrt{3604}}{2} \approx \begin{cases} 65 \\ 5 \end{cases}$$

luego $\|A\| \approx \sqrt{65} \approx 8'1$. Hemos perdido muy poco conformándonos con una cota superior.

Ejemplo de matriz variable. Hallemos una cota superior para $\|(Df)_p\|$, siendo

$$f(x, y) = \left(\frac{x^2 - 1}{10} + \frac{1}{3} \cos y, \frac{xy}{2} \right), \quad p \in E = \overline{B}((0, 0), 1) \quad (\text{bola euclídea}).$$

Observa que E es compacto y convexo.

La cota será $K = \sqrt{c}$, siendo c un número tal que para $p = (x, y) \in E$ y $h = (a, b) \in \mathbb{R}^2$ se tenga $\|(Df)_p h\|_2^2 \leq c \cdot \|h\|_2^2$. Primero calculamos:

$$Df = \begin{bmatrix} \frac{x}{5} & -\frac{1}{3} \sin y \\ \frac{y}{2} & \frac{x}{2} \end{bmatrix}, \quad (Df)h = \begin{pmatrix} \frac{x}{5}a - \frac{1}{3} \sin y \\ \frac{y}{2}a + \frac{x}{2}b \end{pmatrix},$$

y después:

$$\begin{aligned} \|(Df)h\|_2^2 &= \left(\frac{x}{5}a - \frac{1}{3} \sin y \right)^2 + \left(\frac{y}{2}a + \frac{x}{2}b \right)^2 = \\ &= \left(\frac{x^2}{25} + \frac{y^2}{4} \right) a^2 + \left(\frac{-x}{15} \sin y + \frac{xy}{4} \right) 2ab + \left(\frac{\sin^2 y}{9} + \frac{x^2}{4} \right) b^2 \leq \\ &\leq \frac{x^2 + y^2}{4} a^2 + \left(\frac{|x|}{15} + \frac{x^2 + y^2}{8} \right) (a^2 + b^2) + \left(\frac{1}{9} + \frac{1}{4} \right) b^2 \leq \\ &\leq \left(\frac{1}{4} + \frac{1}{15} + \frac{1}{8} \right) a^2 + \left(\frac{1}{15} + \frac{1}{8} + \frac{1}{9} + \frac{1}{4} \right) b^2 \leq \frac{199}{360} \|h\|_2^2 \leq (0'56) \|h\|_2^2. \end{aligned}$$

Finalmente $c = 0'56$ y $K = \sqrt{0'56} < 0'75$ es una constante de Lipschitz para $f : E \rightarrow \mathbb{R}^2$, cuando se usa la distancia euclídea $d(x, y) = \|x - y\|_2$ tanto en salida como en llegada. Hemos obtenido esta constante de Lipschitz sin hallar ningún autovalor.

Como $f(0,0) = (\frac{1}{3} - \frac{1}{10}, 0) = (0'2\bar{3}, 0)$, tenemos:

$$f(\bar{B}((0,0),1)) \subseteq \bar{B}((0'2\bar{3},0), 0'75) \subset \bar{B}((0,0), 0'2\bar{3} + 0'75) \subset \bar{B}((0,0), 0'984),$$

Luego $f : \bar{B}((0,0),1) \rightarrow \bar{B}((0,0),1)$ es una aplicación contractiva, cuando en salida y en llegada usamos la distancia euclídea estándar. Entonces el teorema 105 garantiza que en la bola $\bar{B}((0,0),1)$ hay una solución única al siguiente sistema de ecuaciones

$$\begin{cases} \frac{x^2-1}{10} + \frac{1}{3} \cos y = x \\ \frac{xy}{2} = y \end{cases}$$

No sólo sabemos que hay una solución, también sabemos dónde está: en la bola cerrada de centro $(0,0)$ y radio $0'984$, luego en la bola *abierta* $B((0,0),1)$. También sabemos cómo calcular esa solución con tantos decimales de precisión como queramos: evaluando el iterado $f^n(0,0)$ para un n adecuado.

Interesante: el truco con el que se pasa de (61) a (62) se adapta fácilmente a la desigualdad de Young.

3.3 Funciones inversas

Es sabido que si $f(x)$ es una función escalar, definida en un intervalo $I \subseteq \mathbb{R}$ y con derivada $f'(x)$ estrictamente positiva, entonces f es estrictamente creciente y de hecho es una biyección entre dos intervalos $f : I \rightarrow J$. Además, la función inversa $f^{-1} : J \rightarrow I$ es también derivable. En este apartado vamos a generalizar eso a una aplicación diferenciable entre dominios de \mathbb{R}^n . Para construir la inversa de f buscamos una solución x a la ecuación $f(x) = y$, basándonos en el estudio hecho en el apartado 3.1.3 de la ecuación de punto fijo.

3.3.1 Funciones coercivas

Definición 107. Una función entre espacios métricos $f : (X, d_X) \rightarrow (Y, d_Y)$ es **coerciva** si existe una constante $\lambda > 0$ tal que

$$d_Y(f(x'), f(x)) \geq \lambda \cdot d_X(x', x) \quad \text{para cualesquiera } x, x' \in X.$$

Las constantes λ que cumplen esto se llaman **constantes de coercividad** de f .

El primer significado que acude a nuestra mente es que f no permite que las imágenes $f(x), f(x')$ “se acerquen demasiado”. De hecho las aplicaciones coercivas son inyectivas, porque si $x \neq x'$ entonces $d(f(x), f(x')) \geq \lambda \cdot d(x, x') > 0$, luego $f(x) \neq f(x')$. Pero la coercividad es una propiedad *estrictamente más fuerte* que la inyectividad.

Ejercicio 108. (a) Demuestra que las condiciones siguientes son equivalentes:

1. f es coerciva, con constante de coercividad λ .
2. f es inyectiva y su **inversa desde el conjunto imagen** $f(X) \ni f(x) \mapsto x$ es Lipschitziana, admitiendo $1/\lambda$ como constante de Lipschitz.

(b) Demuestra que la compuesta de coercivas $f \circ g$ es coerciva, con constante de coercividad el producto de las constantes de coercividad de f y de g .

Ejercicio 109. Fijamos una norma $\|\cdot\|$ en \mathbb{R}^n y la utilizamos tanto en salida como en llegada, escribiendo $|||\cdot|||$ para la correspondiente norma de operador.

1. Dada A , matriz invertible $n \times n$, demuestra que $L(x) \equiv Ax$ es coerciva, con constante de coercividad $\lambda = 1/|||A^{-1}|||$.
2. Dadas $f, g : \mathbb{R}^n \rightarrow \mathbb{R}^n$, La primera coerciva con constante de coercividad λ , la segunda Lipschitziana con constante de Lipschitz K . Demuestra que si $K < \lambda$ entonces $f + g$ es coerciva, admitiendo $\lambda - K$ como constante de coercividad.

3.3.2 Teorema cualitativo de la función inversa

Teorema 110. Sean $U \subseteq \mathbb{R}^n$ un abierto y $f : U \rightarrow \mathbb{R}^n$ de clase \mathcal{C}^1 con $A = (Df)_{x_0}$ invertible. Existe un par de radios $r_1, r_2 > 0$ tales que f es coerciva en $B(x_0, r_1)$ y la imagen inyectiva $f(B(x_0, r_1))$ contiene a $B(f(x_0), r_2)$.

Llamamos “cualitativo” a este teorema porque no especificamos el valor de r_1 ni el de r_2 .

Demostración. Escribimos $y_0 = f(x_0)$ y $A = (Df)_{x_0}$. Tenemos

$$f(x) \equiv A(x - x_0) + \varphi_0(x) , \quad (64)$$

con φ_0 de clase \mathcal{C}^1 y tal que $\varphi_0(x_0) = y_0$ y $(D\varphi_0)_{x_0} = 0_{n \times n}$. Dado un vector y , vamos a convertir la siguiente ecuación, con incógnita x ,

$$f(x) = y , \quad (65)$$

en un problema de punto fijo. Transformamos (65) así:

$$\begin{aligned} A(x - x_0) + \varphi_0(x) &= y , \\ A(x - x_0) &= y - \varphi_0(x) , \\ x &= x_0 + A^{-1}y - A^{-1}\varphi_0(x) . \end{aligned}$$

Es decir que (65) equivale a la siguiente ecuación de punto fijo

$$x = F_y(x) , \quad (66)$$

siendo:

$$\begin{aligned} F_y(x) &\equiv x_0 + A^{-1}y + \varphi(x) , \\ \varphi(x) &\equiv -A^{-1}\varphi_0(x) . \end{aligned}$$

La función φ es de clase \mathcal{C}^1 y satisface:

$$\varphi(x_0) = -A^{-1}y_0 , \quad (D\varphi)_{x_0} = 0_{n \times n} .$$

Como $F_y \equiv (\text{vector constante}) + \varphi$, cualquier constante de Lipschitz para φ es también una constante de Lipschitz para F_y .

Consideramos valores de r para los que $\overline{B}(x_0, r) \subset U$. Como la bola $E_r = \overline{B}(x_0, r)$ es convexa, los resultados del apartado 3.2 nos dicen que el número

$$K_r = \sup_{p \in \overline{B}(x_0, r)} \|(D\varphi)_p\| ,$$

es una constante de Lipschitz para $\varphi|_{E_r}$ y para $F_y|_{E_r}$.

La función escalar $x \mapsto \|(D\varphi)_x\|$ es continua y nula en $x = x_0$. Se sigue que K_r tiende a cero cuando r tiende a cero.

Dado un valor $r_1 > 0$ con $K_{r_1} < 1$, sabemos que F_y es Lipschitz-pequeña en $E = \overline{B}(x_0, r_1)$. Veamos que entonces $f|_E$ es coerciva. La fórmula (64) implica esta otra:

$$f(x) = A(x - x_0 - \varphi(x)) = (L \circ \alpha)(x) ,$$

donde $\alpha(x) \equiv x - x_0 - \varphi(x)$ y $L(x) \equiv Ax$. Por el ejercicio 109, la función α es coerciva en la bola E , con constante de coercividad $1 - K_{r_1}$, mientras que L es coerciva en todo \mathbb{R}^n . El ejercicio 108 nos dice ahora que la compuesta $f|_E = L \circ (\alpha|_E)$ también es coerciva. Como constante de coercividad para $f|_E$ vale $(1 - K_{r_1})/\|A^{-1}\|$, pero no vamos a usar este dato.

Para ser una aplicación contractiva, a F_y sólo le falta llevar E dentro de E . Veamos cómo tiene que ser el vector $h = y - y_0$ para que se cumpla $F_y(E) \subseteq E$. Dado $x \in E$, se tiene:

$$\begin{aligned} d(x_0, F_y(x)) &\leq d(x_0, F_y(x_0)) + d(F_y(x_0), F_y(x)) = \\ &= \|A^{-1}h\| + d(F_y(x_0), F_y(x)) \leq \|A^{-1}h\| + K_{r_1} \cdot d(x_0, x) \leq \\ &\leq \|A^{-1}h\| + K_{r_1} \cdot r_1 . \end{aligned}$$

La condición $K_{r_1} < 1$ y la desigualdad estricta

$$\|A^{-1}h\| < (1 - K_{r_1}) \cdot r_1 , \quad (67)$$

nos aseguran esta inclusión:

$$\varphi_y(\overline{B}(x_0, r_1)) \subseteq B(x_0, r_1) . \quad (68)$$

Entonces φ_b es una aplicación de $\overline{B}(x_0, r_1)$ en sí misma que es contractiva y, como esta bola es un espacio completo, el teorema 105 nos dice que las ecuaciones equivalentes (65) y (66) tienen una solución $x \in \overline{B}(x_0, r_1)$ y que la igualdad $x = \varphi_b(x)$, junto con (68), implica $x \in B(x_0, r_1)$.

Se pueden elegir $r_1, r_2 > 0$ tales que

$$K_{r_1} < 1 \quad \text{y} \quad \|h\| < r_2 \implies \|A^{-1}h\| < (1 - K_{r_1}) \cdot r_1 .$$

Elegidos esos valores r_1, r_2 , para cada $y \in B(y_0, r_2)$ el vector $h = y - y_0$ satisface (67) y, por lo que acabamos de ver, hay un $x \in B(x_0, r_1)$ único tal que $y = f(x)$.

Fíjate. Como (67) implica $\|A^{-1}h\| < r_1$, cuanto más pequeño sea r_1 menor tiene que ser r_2 . \square

Al ser $f : B(x_0, r_1) \rightarrow \mathbb{R}^n$ coerciva, su inversa desde el conjunto imagen

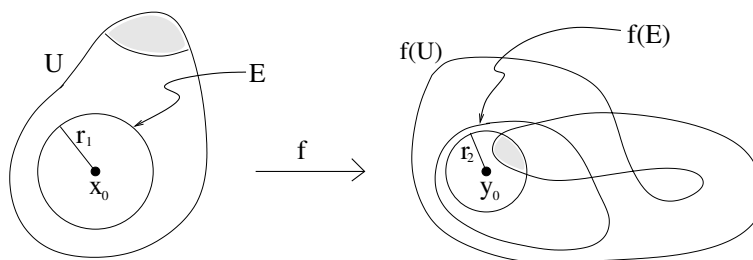
$$f^{-1} : f(B(x_0, r_1)) \longrightarrow B(x_0, r_1) ,$$

es Lipschitziana. Lo que hemos demostrado es que el dominio de esta f^{-1} contiene a la bola $B(y_0, r_2) = B(f(x_0), r_2)$. En particular $g = f^{-1}|_{B(y_0, r_2)}$ admite la siguiente definición:

$$\text{para cada } y \in B(f(x_0), r_2) \quad , \quad g(y) = \left(\text{el único } x \in B(x_0, r_1) \text{ tal que } f(x) = y \right) . \quad (69)$$

Esta g es la **inversa local de f definida en $B(f(x_0), r_2)$ y con valores en $B(x_0, r_1)$** .

Aviso. Dado $y \in B(y_0, r_2)$, sucede con frecuencia que fuera de la bola $E = \overline{B}(x_0, r_1)$ hay otros puntos $x' \in U$ que también cumplen $f(x') = y$. Dicho de otra manera: la función f , tomada en todo su dominio U , puede *cubrir varias veces* todo o parte de la bola $B(y_0, r_2)$



porque f puede no ser inyectiva en todo U , ver el apartado 3.5. En la definición (69) de g es, pues, importante indicar que la preimagen $g(y) = f^{-1}(y)$ sólo la buscamos en la bola $B(x_0, r_1)$. Otra manera de expresarlo es que la función $g(y)$ es la solución al siguiente sistema, formado por una identidad entre funciones y una restricción a dónde permitimos que g tome sus valores:

$$\left. \begin{aligned} f(g(y)) &\equiv y \\ g(y) &\in B(x_0, r_1) \end{aligned} \right\} \quad \text{para } y \in B(f(x_0), r_2) ,$$

siendo necesarias ambas condiciones para determinar g .

En el tercer comentario al final del apartado 3.4.1 damos otro método para determinar g .

Proposición 111. Sean $U \subseteq \mathbb{R}^n$ abierto y $f : U \rightarrow \mathbb{R}^n$ de clase \mathcal{C}^1 .

Dado cualquier abierto $V \subseteq U$ tal que $(Df)_x$ es invertible para todo $x \in V$, la imagen $f(V)$ es un conjunto abierto.

Demostración. Para cada $y \in f(V)$ existe al menos un $x \in V$ con $y = f(x)$. El teorema 110 proporciona radios $r_1, r_2 > 0$ tales que $B(x, r_1) \subseteq V$ y $f(B(x, r_1)) \supseteq B(y, r_2)$. A fortiori es $f(V) \supseteq B(y, r_2)$, es decir $B(y, r_2) \subseteq f(V)$. Como y era cualquier elemento de $f(V)$, queda demostrado que $f(V)$ es abierto. \square

3.3.3 Diferenciabilidad de la inversa local

Seguimos con las notaciones utilizadas en el apartado 3.3.2 .

Supongamos que g es diferenciable en $y_0 = f(x_0)$. Al diferenciar la identidad $f(g(y)) \equiv y$ se obtiene lo siguiente, después de aplicar la regla de la cadena

$$(Df)_{x_0} (Dg)_{y_0} = I_n .$$

La conclusión es:

Que g sea diferenciable en y_0 , obliga $\begin{cases} \text{a } (Df)_{x_0} \text{ a ser invertible ,} \\ \text{a } (Dg)_{y_0} \text{ a ser la inversa de } (Df)_{x_0} . \end{cases}$

Teorema 112. En las hipótesis del teorema 110, la inversa local g es diferenciable en y_0 .

Demostración. Tenemos

$$f(x) = f(x_0) + A(x - x_0) + R(x) \quad , \quad \text{con } R(x) = o(\|x - x_0\|) , \quad (70)$$

Sustituimos en (70) la variable x por $g(y)$, con y recorriendo $B(y_0, r_2)$. El resultado es:

$$y = y_0 + A(g(y) - g(y_0)) + R(g(y)) \quad , \quad \text{para todo } y \in B(y_0, r_2) ,$$

de donde se despeja:

$$g(y) = g(y_0) + A^{-1}(y - y_0) - A^{-1}R(g(y)) . \quad (71)$$

Al final del apartado 3.3.2 hemos señalado que g es una función Lipschitziana. Si C es la constante del Lipschitz de g , entonces

$$\|g(y) - g(y_0)\| = \|g(y) - g(y_0)\| \leq C \|y - y_0\| ,$$

luego $R(g(y)) = o(g(y) - g(y_0)) = o(y - y_0)$. También

$$\|A^{-1}R(g(y))\| \leq \|A^{-1}\| \|R(g(y))\| ,$$

implica que $A^{-1}R(g(y)) = o(\|y - y_0\|)$, luego (71) da lugar a:

$$g(y) = g(y_0) + A^{-1}(y - y_0) + o(\|y - y_0\|) \quad \text{para todo } y \in B(y_0, r_2) ,$$

y queda visto que g es diferenciable en y_0 , con jacobiana A^{-1} . \square

Teorema 113. Sean un abierto $U \subseteq \mathbb{R}^n$ y una función $f : U \rightarrow \mathbb{R}^n$ de clase \mathcal{C}^ℓ .

Si f es inyectiva y la jacobiana $(Df)_x$ es invertible para todo $x \in U$, entonces la inversa $f^{-1} : f(U) \rightarrow U$ es de clase \mathcal{C}^ℓ . Aquí se incluye el caso $\ell = \infty$.

Demostración. La proposición 111 nos da que $f(U)$ es un abierto. El teorema 112 dice que f^{-1} es diferenciable, y que en $f(U)$ tenemos la siguiente identidad entre funciones matriz:

$$D(f^{-1}) \equiv (Df)^{-1} \circ f^{-1} = F \circ Df \circ f^{-1} , \quad (72)$$

Siendo F la función que lleva cada matriz invertible $n \times n$ a su inversa. Al final del apartado 2.3 vimos que F es \mathcal{C}^∞ .

Si f es \mathcal{C}^1 , entonces Df es continua y los tres factores en el lado derecho de (72) son continuos, con lo cual la función $D(f^{-1})$ es continua y f^{-1} es \mathcal{C}^1 .

Sea ahora f de clase \mathcal{C}^2 . En particular es \mathcal{C}^1 y ya hemos visto que f^{-1} es también \mathcal{C}^1 . Pero entonces los tres factores en el lado derecho de (72) son \mathcal{C}^1 , luego la función $D(f^{-1})$ es \mathcal{C}^1 y por lo tanto f^{-1} es \mathcal{C}^2 .

Este proceso continúa indefinidamente y prueba que si f es \mathcal{C}^ℓ entonces f^{-1} es \mathcal{C}^ℓ , siendo válido el argumento para todo ℓ finito.

Si f es \mathcal{C}^∞ , en particular es \mathcal{C}^ℓ para todo ℓ finito. Entonces f^{-1} es \mathcal{C}^ℓ para todo ℓ finito, luego es \mathcal{C}^∞ . \square

El teorema 110 no prohíbe que la inversa local exista cuando $(Df)_{x_0}$ no es invertible, pero entonces esa inversa no es ni siquiera diferenciable en el punto $y_0 = f(x_0)$. Por ejemplo, la función $f(x) \equiv 5 + (x-2)^3$ es biyectiva de \mathbb{R} a \mathbb{R} y \mathcal{C}^∞ ; pero su inversa $f^{-1}(y) = 2 + \sqrt[3]{y-5}$ ni siquiera es diferenciable en $y_0 = 5 = f(2)$, lo cual se debe a que $f'(2)$ es nula.

Corolario 114. Sean un abierto $U \subseteq \mathbb{R}^n$ y una función $f : U \rightarrow \mathbb{R}^n$ de clase \mathcal{C}^ℓ e inyectiva, con $f(U)$ conjunto abierto.

Si la inversa $f^{-1} : f(U) \rightarrow U$ es diferenciable, entonces esta inversa es \mathcal{C}^ℓ .

Sugerencia: busca en la Internet el *teorema de invariancia de dominio*.

3.3.4 Algunos comentarios al teorema 110

(1) La inversa local g , definida en la fórmula (69), se puede calcular por la siguiente fórmula:

$$g(y) = \lim_{k \rightarrow \infty} \underbrace{F_y \circ \cdots \circ F_y}_{k \text{ factores}}(x_0).$$

Esta fórmula *no es finita*, lo cual era de esperar ya que g puede no ser elemental. Por ejemplo, la función $f : \mathbb{R} \rightarrow \mathbb{R}$ dada por $f(x) \equiv x + e^x$ es elemental y biyectiva, pero su inversa no es elemental.

(2) Una función en las condiciones del teorema 110 es inyectiva en una bola $B(x_0, r_1)$ suficientemente pequeña. Pero no está obligada a ser inyectiva en todo su dominio. Para más detalles ver el apartado 3.5.

(3) Tampoco está garantizado que una $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, con todas las jacobianas invertibles, cubra todo el espacio. Por ejemplo, la función $f(x) \equiv e^x$ tiene $f'(x) > 0$ para todo x , pero el recorrido de esta función es $f(\mathbb{R}) = (0, +\infty) \neq \mathbb{R}$.

(4) En el teorema 110 no se puede debilitar la hipótesis $f \in \mathcal{C}^1$. La siguiente función

$$f(x) = \begin{cases} x + x^2 \sin \frac{1}{x^2} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

es derivable (con derivada discontinua) y satisface $f'(0) = 1 \neq 0$; pero no es inyectiva en ningún entorno $B(0, r) = (-r, r)$, porque tiene infinitos máximos y mínimos locales estrictos cuando nos acercamos a $x = 0$.

3.3.5 Teorema cuantitativo de la función inversa

En este apartado vamos a concretar el valor de la pareja de radios r_1, r_2 , que se anuncia en el teorema 110, por un método mucho mejor que el que se extrae de la demostración dada en el apartado 3.3.2.

En todo este apartado la norma $\|\cdot\|$ es la euclídea estándar.

Teorema 115. Sean un abierto $U \subseteq \mathbb{R}^n$, un punto $x_0 \in U$ y $f : U \rightarrow \mathbb{R}^n$ de clase \mathcal{C}^1 . Si dos números $r, \lambda > 0$ y una matriz **ortogonal** $P \in O(n)$ verifican lo siguiente

$$\text{para todo } x \in \overline{B}(x_0, r) \text{ y todo } v \in \mathbb{R}^n \text{ se tiene } v^t P (Df)_x v \geq \lambda \|v\|^2, \quad (73)$$

entonces f es una función coerciva en $\overline{B}(x_0, r)$, con constante de coercividad λ , y además:

$$f(B(x_0, r)) \supseteq B(f(x_0), \lambda r),$$

luego hay una inversa local g de f , definida en $B(f(x_0), \lambda r)$ y con sus valores en $B(x_0, r)$.

Demostraremos este teorema entre los apartados 3.5 y 3.6.

Veamos un par de ejemplos, en los que vamos a utilizar la desigualdad (61) una y otra vez.

El primer ejemplo es una $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$:

$$f(x, y, z) \equiv (5x + 2x^2y - x^2z + 2xy^2, x + 7y + 2xy + 2x^2, 4x + y + 6z + 3y^2 - yz) \quad , \quad x_0 = (0, 0, 0),$$

Calculamos:

$$Df = \begin{pmatrix} 5 + 4xy - 2xz + y^2 & 2x^2 + 4xy & -x^2 \\ 1 + 2y + 4x & 7 + 2x & 0 \\ 4 & 1 + 6y - z & 6 - y \end{pmatrix}.$$

La matriz $(Df)_{x_0} = \begin{bmatrix} 5 & 0 & 0 \\ 1 & 7 & 0 \\ 4 & 1 & 6 \end{bmatrix}$, además de ser invertible, tiene las entradas de la diagonal más bien grandes, por lo que esperamos que en la forma cuadrática $v^t Df v$ los cuadrados dominen a los términos mixtos. Vamos a probar a ver si hay un $\lambda > 0$ tal que (73) se satisface para $P = I_3$ y $r = 0'1$. Escribimos $v = (a, b, c)$ y calculamos:

$$\begin{aligned} v^t Df v &= (5 + 4xy - 2xz + y^2) a^2 + (2x^2 + 4xy) ab - x^2 ac + \\ &\quad + (1 + 2y + 4x) ab + (7 + 2x) b^2 + 0 \cdot bc + \\ &\quad + 4ac + (1 + 6y - z) bc + (6 - y) c^2 \end{aligned}$$

Estimamos el primer coeficiente, teniendo en cuenta que $\|(x, y, z)\| \leq 0'1$:

$$\begin{aligned} |4y - 2z| &= |(x, y, z) \cdot (0, 4, -2)| \leq (0'1) \sqrt{20} < 0'45, \\ |4xy - 2xz| &= |x| \cdot |4y - 2z| \leq 0'045, \\ 5 + 4xy - 2xz + y^2 &\geq 5 - 0'045 + 0 = 4'955. \end{aligned}$$

El coeficiente de b^2 es $7 + 2x \geq 6'8$, y el de c^2 es $6 - y \geq 5'9$, luego:

$$\begin{aligned} v^t Df v &\geq (4'955) a^2 + (6'8) b^2 + (5'9) c^2 + \\ &\quad + \left(x^2 + 2xy + \frac{1}{2} + y + 2x \right) 2ab + \left(-\frac{x^2}{2} + 2 \right) 2ac + \left(0 + \frac{1}{2} + 3y - \frac{z}{2} \right) 2bc. \end{aligned}$$

Ahora estimamos los coeficientes de los términos mixtos, recordando que $\|(x, y, z)\| \leq r = 0'1$:

$$\begin{aligned} |x + 2y| &= |(x, y, z) \cdot (1, 2, 0)| \leq (0'1) \sqrt{5} < 0'23, \\ |x^2 + 2xy| &= |x(x + 2y)| \leq (0'1) 0'23 = 0'023, \\ |y + 2x| &= |(x, y, z) \cdot (2, 1, 0)| \leq (0'1) \sqrt{5} < 0'23, \\ \left| x^2 + 2xy + \frac{1}{2} + y + 2x \right| &\leq 0'023 + 0'5 + 0'23 = 0'753, \\ \left| -\frac{x^2}{2} + 2 \right| &\leq 2, \\ \left| 3y - \frac{z}{2} \right| &= \left| (x, y, z) \cdot \left(0, 3, \frac{1}{2} \right) \right| \leq (0, 1) \sqrt{9'25} < 0'305, \\ \left| \frac{1}{2} + 3y - \frac{z}{2} \right| &< 0'805, \end{aligned}$$

y deducimos que, para $\|(x, y, z)\| < 0'1$, se tiene:

$$\begin{aligned} v^t Df v &\geq (4'955) a^2 + (6'8) b^2 + (5'9) c^2 - \\ &\quad - 0'753 (a^2 + b^2) - 2 (a^2 + c^2) - 0'805 (b^2 + c^2) = \\ &= (4'955 - 0'753 - 2) a^2 + (6'8 - 0'753 - 0'805) b^2 + (5'9 - 2 - 0'805) c^2 = \\ &= (2'202) a^2 + (5'242) b^2 + (3'095) c^2, \end{aligned}$$

de donde:

$$v^t (Df)_p v \geq (2'2) \|v\|^2 \quad \text{para todo } v \in \mathbb{R}^3 \text{ y todo } p \in B((0, 0, 0), 0'1).$$

Concluimos que f es coerciva en $\overline{B}((0, 0, 0), 0'1)$, con constante de coercividad $2'2$, y que f tiene una inversa local $g : B((0, 0, 0), 0'22) \rightarrow B((0, 0, 0), 0'1)$ bien definida y C^∞ .

Nuestro segundo ejemplo es una $f : U \rightarrow \mathbb{R}^2$, siendo U el abierto $(-0'25, +\infty) \times \mathbb{R}$:

$$f(x, y) = \left(x^2 + e^{5y}, \frac{1}{1+4x} + \cos y \right), \quad x_0 = (0, 0), \quad f(x_0) = (1, 2).$$

Calculamos:

$$Df = \begin{pmatrix} 2x & 5e^{5y} \\ -4 & -\sin y \\ \frac{1}{(1+4x)^2} & -\sin y \end{pmatrix}, \quad (Df)_{x_0} = \begin{pmatrix} 0 & 5 \\ -4 & 0 \end{pmatrix}.$$

Dado $v = (a, b)$, los coeficientes de a^2 y b^2 en la forma cuadrática $v^t (Df)_{x_0} v$ son ambos nulos. Es necesario colocar en esas posiciones los números 5 y $|-4| = 4$, para lo cual intercambiamos las filas de la jacobiana, cambiando el signo de la segunda fila antes de moverla. Esto se consigue multiplicando Df desde la izquierda por la matriz $P = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, que es ortogonal:

$$P Df = \begin{pmatrix} 4 & \sin y \\ \frac{1}{(1+4x)^2} & \sin y \\ 2x & 5e^{5y} \end{pmatrix}.$$

Ahora tenemos:

$$\begin{aligned} v^t (P Df) v &= \frac{4}{(1+4x)^2} a^2 + \left(\frac{\sin y}{2} + x \right) 2ab + 5e^{5y} b^2 \geq \\ &\geq \left(\frac{4}{(1+4x)^2} - \left| \frac{\sin y}{2} + x \right| \right) a^2 + \left(5e^{5y} - \left| \frac{\sin y}{2} + x \right| \right) b^2. \end{aligned}$$

Vamos a probar con $r = 0'1$. Hacemos las siguientes estimaciones, para $\|(x, y)\| \leq 0'1$:

$$\begin{aligned} \frac{4}{(1+4x)^2} &\geq \frac{4}{(1'4)^2} > 2'04, \\ \left| \frac{\sin y}{2} + x \right| &\leq \frac{|y|}{2} + |x| \leq 0'05 + 0'1 = 0'15, \\ 5e^{5y} &\geq 5e^{-0'5} > 3, \end{aligned}$$

y deducimos:

$$v^t (P Df) v \geq (2'04 - 0'15) a^2 + (3 - 0'15) b^2 \geq (1'89) \|v\|^2.$$

Ahora sabemos que f es coerciva en $\overline{B}((0, 0), 0'1)$, con coeficiente de coercividad $1'89$, y que tiene una inversa local bien definida $g : B((1, 2), 0'189) \rightarrow B((0, 0), 0'1)$ que es C^∞ .

3.4 Funciones implícitas

En este apartado consideramos el espacio \mathbb{R}^{n+k} , donde n y k son enteros positivos. Para indicar el elemento general de este espacio, escribimos lo siguiente:

$$(x, y) = (x_1, \dots, x_n, y_1, \dots, y_k),$$

de modo que x denota un vector de \mathbb{R}^n e y denota un vector de \mathbb{R}^k . Correspondiendo a eso, una función de $n+k$ variables $F(x_1, \dots, x_n, y_1, \dots, y_k)$ puede ser vista como una función de dos variables vector: $F(x, y)$.

3.4.1 Teorema cualitativo de la función implícita

Sea W un abierto del espacio \mathbb{R}^{n+k} . Dados una función $F : W \rightarrow \mathbb{R}^k$ y un punto $b \in \mathbb{R}^k$, al **conjunto preimagen**

$$F^{-1}(\{b\}) = \{ (x, y) \in W : F(x, y) = b \},$$

lo llamamos **conjunto implícito definido por la ecuación $F(x, y) = b$** , porque es el conjunto de soluciones (x, y) a dicha ecuación.

“Resolver la ecuación $F(x, y) = b$ ” se refiere a describir el conjunto $F^{-1}(\{b\})$ por completo. “Resolver localmente” esa ecuación es describir *la parte cercana a un punto* del conjunto implícito. Más en concreto, elegimos un punto $p \in F^{-1}(\{b\})$ y damos una condición suficiente (no necesaria) para que haya un entorno B de p tal que la parte $B \cap F^{-1}(\{b\})$ del conjunto implícito tenga una descripción donde unas incógnitas quedan “libres de recorrer un abierto” mientras las otras incógnitas son funciones de las libres.

Teorema 116. Sean un abierto $W \subseteq \mathbb{R}^{n+k}$ y una función $F : W \rightarrow \mathbb{R}^k$ de clase \mathcal{C}^1 . Fijamos un punto $p = (x_0, y_0) \in W$ y su imagen $b = F(p)$.

Que las derivadas parciales $F_{y_1}(p), \dots, F_{y_k}(p)$ formen una base de \mathbb{R}^k es condición **suficiente** para que existan un abierto $U \subset \mathbb{R}^n$, con $x_0 \in U$, y otro abierto $B \subset \mathbb{R}^{n+k}$, con $p \in B$, que nos proporcionan las dos cosas siguientes:

1. Una **función implícita** $\varphi(x) : U \rightarrow \mathbb{R}^k$ que cumple la siguiente identidad:

$$F(x, \varphi(x)) \equiv b, \quad \text{para todo } x \in U. \quad (74)$$

Si F es de clase \mathcal{C}^ℓ , entonces φ es al menos \mathcal{C}^ℓ pero puede ser mejor.

2. Una **capa protectora** para el grafo de la función implícita:

$$B \cap F^{-1}(\{b\}) = (\text{grafo de } \varphi) = \{ (x, \varphi(x)) : x \in U \}, \quad (75)$$

es decir que en el abierto B no hay ninguna otra solución de la ecuación $F(x, y) = b$ aparte de los puntos del grafo de φ . En particular, es $\varphi(x_0) = y_0$.

La función $\varphi(x)$ puede ser escalar o vectorial, según sea $k = 1$ o $k \geq 2$. En el segundo caso las entradas de φ son k funciones escalares $\varphi_1, \dots, \varphi_k : U \rightarrow \mathbb{R}$ que reciben el nombre de **funciones implícitas**.

La ecuación vectorial $F(x, y) = b$ es, en realidad, un **sistema** de k ecuaciones escalares. El teorema 116 da una condición suficiente para poder resolverlo localmente, en un entorno adecuado B de p , dejando que n incógnitas x_i recorran cierto abierto $U \subset \mathbb{R}^n$ (que queden “libres”) y despejando las otras k incógnitas y_1, \dots, y_k en función de las x_i .

Una condición suficiente, para que cerca del punto p se puedan despejar tantas incógnitas como ecuaciones hay en el sistema, es que las derivadas de las ecuaciones **respecto de las incógnitas que queremos despejar** sean linealmente independientes en p .

Cuidado: nada dice el teorema sobre cómo son las posibles soluciones al sistema fuera del abierto B , o sea lejos del punto p .

Llamamos **capa protectora del grafo** a la parte $B \setminus (\text{grafo de } \varphi)$ porque es como una piel alrededor del grafo, y la fórmula (75) nos dice que el conjunto implícito $F^{-1}(\{b\})$ jamás entra en ella. Más sobre esto en el apartado 4.2.7 del capítulo 4.

Demostración del teorema 116.

Existencia. El truco estándar consiste en definir la siguiente función

$$G : W \longrightarrow \mathbb{R}^{n+k} \quad , \quad G(x, y) = (x, F(x, y)) \quad , \quad (76)$$

y observar que el conjunto implícito admite la siguiente descripción

$$F^{-1}(\{b\}) = G^{-1}(\mathbb{R}^n \times \{b\}) \quad . \quad (77)$$

La jacobiana de F se parte en dos submatrices $DF = [D_x F \mid D_y F]$ de la manera obvia:

$$D_x F = [F_{x_1} \mid F_{x_2} \mid \cdots \mid F_{x_n}]_{k \times n} \quad , \quad D_y F = [F_{y_1} \mid \cdots \mid F_{y_k}]_{k \times k} \quad .$$

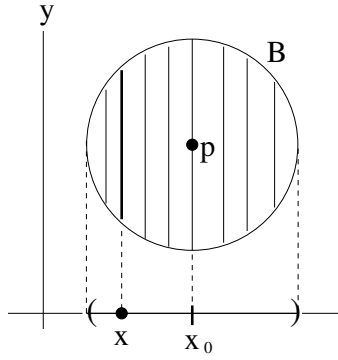
Entonces, la jacobiana de G

$$DG = \left[\begin{array}{c|c} I_n & 0_{n \times k} \\ \hline D_x F & D_y F \end{array} \right] \quad ,$$

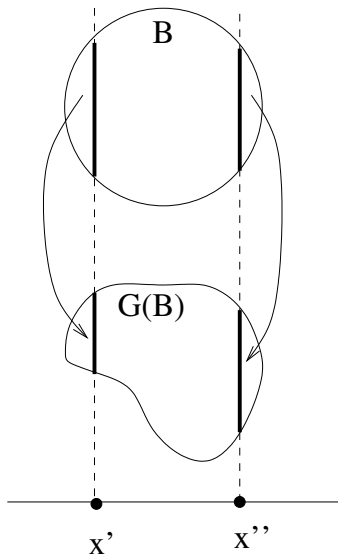
es triangular por cajas, luego invertible donde F_{y_1}, \dots, F_{y_k} sean linealmente independientes. Nuestra hipótesis equivale, pues, a que $(DG)_p$ sea una matriz invertible $(n+k) \times (n+k)$.

Al ser F de clase \mathcal{C}^1 , también lo es G y, como $(DG)_p$ es invertible, podemos aplicar a G el teorema 110 y la proposición 111. En particular, existe un radio $r > 0$ tal que G es inyectiva en la bola $B = B(p, r) \subseteq W$ y la imagen $G(B)$ es un abierto de \mathbb{R}^{n+k} que contiene al punto $G(p) = (x_0, b)$.

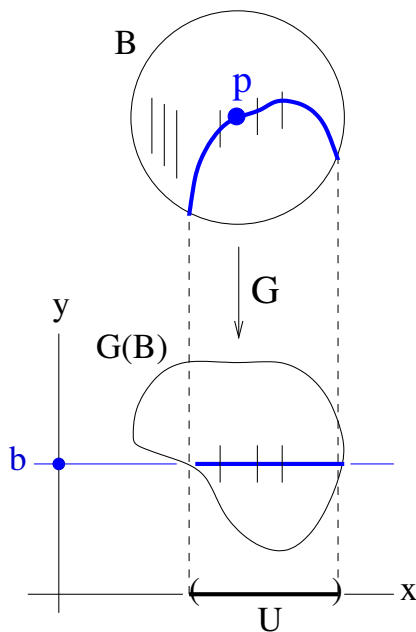
Para cada $x \in B(x_0, r)$, el subespacio afín $\{x\} \times \mathbb{R}^k$ corta a la bola B en un disco k -dimensional.



La definición (76) de G muestra claramente que G transforma cada disco $(\{x\} \times \mathbb{R}^k) \cap B$ sin sacarlo del correspondiente subespacio afín $\{x\} \times \mathbb{R}^k$



Por otra parte, cada subespacio afín $\{x\} \times \mathbb{R}^k$ corta a $\mathbb{R}^n \times \{b\}$ en un único punto (x, b) . Como $G : B \rightarrow G(B)$ es biyectiva, deducimos por (77) que cada disco $(\{x\} \times \mathbb{R}^k) \cap B$ contiene uno o ningún punto de $F^{-1}(\{b\})$, dependiendo de si (x, b) pertenece a $G(B)$ o no.



Entonces definimos

$$U = \{x \in B(x_0, r) : (x, b) \in G(B)\} \subset \mathbb{R}^n. \quad (78)$$

Resulta que $x_0 \in U$ y, para cada $x \in U$, el disco $(\{x\} \times \mathbb{R}^k) \cap B$ contiene exactamente un punto $q(x) \in F^{-1}(\{b\})$. Este punto es de la forma $q(x) = (x, \varphi(x))$ y así queda definida una función $\varphi(x) : U \rightarrow \mathbb{R}^k$. Como B es la unión de los discos $(\{x\} \times \mathbb{R}^n) \cap B$, con x recorriendo $B(x_0, r)$, deducimos que

$$B \cap F^{-1}(\{b\}) = \bigcup_{x \in U} \{q(x)\} = \{(x, \varphi(x)) : x \in U\} = \text{el grafo de } \varphi.$$

Como el punto $p = (x_0, y_0)$ pertenece a $F^{-1}(\{b\})$ y a B , está en el grafo de φ y forzosamente es $\varphi(x_0) = y_0$.

Dada $i_b(x) \equiv (x, b)$, vemos que la definición (78) equivale a $U = i_b^{-1}(G(B))$. Como i_b es continua y $G(B)$ es abierto, el conjunto U es un abierto de \mathbb{R}^n .

Diferenciabilidad. Si F es \mathcal{C}^ℓ entonces G y $H \equiv G^{-1} : G(B) \rightarrow B$ también son de clase \mathcal{C}^ℓ . Escribiendo $H \equiv (H_1, \dots, H_{n+k})$, para cada $x \in U$ se tiene

$$(x, \varphi(x)) = q(x) = G^{-1}(x, b) = H(x, b),$$

de donde se deduce la siguiente identidad, como funciones en U :

$$\varphi(x) \equiv (H_{n+1}(x, b), \dots, H_{n+k}(x, b)),$$

que nos muestra $\varphi(x)$ como una función que es al menos de clase \mathcal{C}^ℓ .

La razón por la que φ puede ser mejor que \mathcal{C}^ℓ es que puede suceder que la función H sea (solamente) \mathcal{C}^ℓ y que, al mismo tiempo, su restricción al subespacio afín $\mathbb{R}^n \times \{b\}$ sea mejor. Es lo que ocurre, por ejemplo, con $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por $F(x, y) \equiv y + |y|y$. Si despejamos y en función de x en la ecuación $F(x, y) = 0$, resulta la función implícita $\varphi(x) \equiv 0$ que es \mathcal{C}^∞ . Por otra parte F es \mathcal{C}^1 , pero no \mathcal{C}^2 , precisamente en los puntos $(x, y) = (x, 0)$ del grafo de φ , ya que $F_y \equiv 1 + 2|y|$. La explicación de la gran suavidad de φ está en la fórmula

$$H(x, b) \equiv G^{-1}(x, b) \equiv \left(x, \frac{2b}{1 + \sqrt{1 + 4|b|}} \right),$$

que nos muestra que, aunque H es \mathcal{C}^1 y no \mathcal{C}^2 en los puntos $(x, 0)$ del eje de abscisas, sin embargo su restricción a dicho eje es \mathcal{C}^∞ : $H(x, 0) \equiv (x, 0)$. \square

Terminamos con cuatro comentarios:

1. Al igual que dijimos para las funciones inversas, suele suceder que $\varphi(x)$ no es elemental aunque $F(x, y)$ sea una función elemental.
2. Sobre la **unicidad** de la función implícita. El asunto de la capa protectora, que ya hemos mencionado, implica que cualquier función $\psi(x)$ tal que $F(x, \psi(x)) \equiv b$ está obligada a coincidir con $\varphi(x)$ *mientras el grafo de ψ esté dentro de B* . Esto incluye a las $\psi(x)$ definidas en un entorno de x_0 , *continuas* y satisfaciendo el **dato particular** $\psi(x_0) = y_0$.
3. Las inversas locales de una función f son las funciones implícitas resultantes de resolver la ecuación $f(x) - y = 0$ despejando x en función de y . En particular, se les aplica la unicidad del comentario anterior: hay una única inversa local continua $g(y)$, definida en un entorno de y_0 y con el valor particular $g(y_0) = x_0$.
4. Los métodos del apartado 3.3.5 permiten dar **versiones cuantitativas** del teorema de la función implícita. Por ejemplo, se pueden hallar radios r_1, r_2 tales que $\varphi(x)$ está definida para $x \in B(x_0, r_1)$, toma sus valores en $B(y_0, r_2)$ y el conjunto implícito no tiene en $B(x_0, r_1) \times B(y_0, r_2)$ más puntos que los del grafo de φ .

3.4.2 Derivación implícita

La ecuación $F(x, y) = b$ produce la identidad

$$F(x, \varphi(x)) \equiv b, \tag{79}$$

que ha de ser satisfecha por la función implícita $\varphi(x)$. La derivación implícita es un procedimiento que consiste en derivar esa identidad, obteniendo nuevas identidades que han de ser satisfechas por $\varphi(x)$ y sus derivadas parciales. Si F es \mathcal{C}^2 podemos dar un segundo paso: las nuevas identidades se pueden volver a derivar y salen otras identidades, satisfechas por $\varphi(x)$ y sus derivadas primeras y segundas. El proceso continúa hasta ℓ pasos si F es de clase \mathcal{C}^ℓ .

Es interesante que el paso s -ésimo se puede combinar con los pasos que le preceden, resultando identidades que relacionan cada derivada de orden s de $\varphi(x)$ *solamente con* $\varphi(x)$, sin que

medien derivadas de órdenes entre 1 y $s - 1$. La consecuencia es que, en un punto x_0 en el que se conozca el valor $\varphi(x_0)$, se conoce el valor de todas las derivadas de φ .

Veamos la forma general del primer paso, vamos a utilizar la notación introducida al demostrar el teorema 116. Hallamos la matriz jacobiana de los dos miembros de la identidad (79), vistos como funciones de $x = (x_1, \dots, x_n)$, y aplicamos la regla de la cadena. El resultado es la siguiente identidad entre funciones matriz:

$$(D_x F)_{(x, \varphi(x))} + (D_y F)_{(x, \varphi(x))} (D\varphi)_x \equiv 0_{k \times n}.$$

Mientras la caja $D_y F$ se mantenga invertible, podemos despejar:

$$(D\varphi)_x \equiv -[(D_y F)_{(x, \varphi(x))}]^{-1} (D_x F)_{(x, \varphi(x))}.$$

Ahora veamos los pasos primero y segundo, pero en el caso particular $n = k = 1$. Tenemos una función escalar $F(x, y)$, de clase al menos \mathcal{C}^2 en un abierto de \mathbb{R}^2 , y una función $\varphi(x)$, definida en un intervalo $I \subseteq \mathbb{R}$, que satisface la identidad (79) en dicho intervalo y además $F_y(x, \varphi(x)) \neq 0$ para todo $x \in I$. Tomando derivada respecto de x en la identidad (79), se obtiene

$$F_x(x, \varphi(x)) + F_y(x, \varphi(x)) \varphi'(x) \equiv 0, \quad (80)$$

de donde

$$\varphi'(x) \equiv -\frac{F_x(x, \varphi(x))}{F_y(x, \varphi(x))}, \quad (81)$$

fórmula que nos da explícitamente el valor de $\varphi'(x)$ a partir de los valores de x y $\varphi(x)$. Para el segundo paso tenemos dos opciones: derivar la identidad (81) o derivar dos veces en (79). La segunda opción equivale a volver a derivar en (80), lo cual produce la siguiente identidad:

$$F_{xx}(x, \varphi(x)) + 2 F_{xy}(x, \varphi(x)) \varphi'(x) + F_{yy}(x, \varphi(x)) \varphi'(x)^2 + F_y(x, \varphi(x)) \varphi''(x) \equiv 0,$$

de donde primero despejamos $\varphi''(x)$ (porque $F_y \neq 0$) y después sustituimos el valor de $\varphi'(x)$ que hemos despejado de (80), o sea el valor dado por (81):

$$\begin{aligned} \varphi''(x) &\equiv \frac{-1}{F_y(x, \varphi(x))} \left[F_{xx} + 2 F_{xy} \varphi'(x) + F_{yy} \varphi'(x)^2 \right]_{y=\varphi(x)} = \\ &\equiv \frac{-1}{F_y(x, \varphi(x))} \left[F_{xx} + 2 F_{xy} \frac{-F_x}{F_y} + F_{yy} \left(\frac{F_x}{F_y} \right)^2 \right]_{y=\varphi(x)}. \end{aligned}$$

Conseguimos una fórmula que nos da explícitamente el valor de $\varphi''(x)$ a partir de los valores de x y $\varphi(x)$, sin que medie el valor de $\varphi'(x)$.

Observa que en todos los pasos se divide solamente por F_y . Cuando las dimensiones n, k son cualesquiera, en cada paso se multiplica por la matriz $(D_y F)^{-1}$.

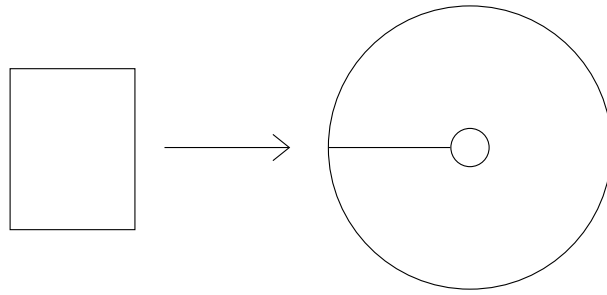
3.5 La cuestión de la inyectividad

Una función de clase \mathcal{C}^1 o mejor, con todas sus jacobianas invertibles, es inyectiva en abiertos pequeños pero puede no ser inyectiva en todo su dominio, *ni siquiera en el caso de que el dominio sea convexo*.

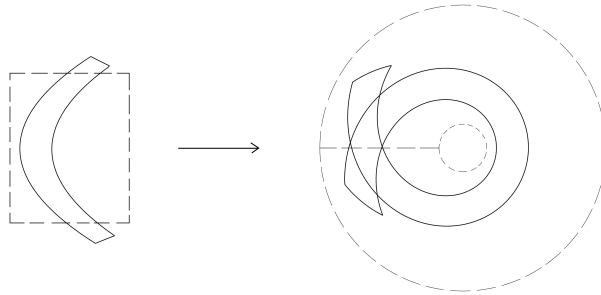
Un sencillo ejemplo es la siguiente función (el “cambio a polares”), cuyo dominio es un semiplano:

$$\text{CP} : (0, +\infty) \times \mathbb{R} \longrightarrow \mathbb{R}^2, \quad \text{CP}(r, \theta) \equiv (r \cos \theta, r \sin \theta).$$

Se tiene $\det D(\text{CP}) \equiv r$, positivo en todo punto del dominio de CP. La siguiente figura muestra el rectángulo $\{1 \leq r \leq 4, -\pi \leq \theta \leq \pi\}$ y su imagen por CP.



y esta otra figura muestra una región en forma de letra C (y de altura mayor que 2π) y cómo su imagen por CP “da la vuelta y se solapa consigo misma”.



Este comportamiento es muy frecuente en funciones con las jacobianas invertibles: si la dimensión es 2 o mayor, la imagen de una región puede dar la vuelta y volver a pasar por donde ya ha pasado.

Otro ejemplo del mismo problema, con un dominio más grande, es la siguiente función

$$g : \mathbb{R}^2 \setminus \{(0,0)\} \longrightarrow \mathbb{R}^2 \quad , \quad g(x,y) = \frac{(x^2 - y^2, 2xy)}{r} = \frac{(x^2 - y^2, 2xy)}{\sqrt{x^2 + y^2}} ,$$

Con ayuda de las coordenadas polares, se comprueba que

$$\text{para cualesquiera } p \in \mathbb{R}^2 \setminus \{(0,0)\} \text{ y } v \in \mathbb{R}^2 \quad , \quad \text{se tiene } \|(Dg)_p v\| \geq \|v\| ,$$

pero g no es inyectiva, ya que $g(-x, -y) = g(x, y)$.

La parte de coercividad del teorema 115 es consecuencia directa de la siguiente proposición, que nos da una condición suficiente para que la función sea coerciva (que es mejor que inyectiva) en ciertos dominios. Denotamos por $\|\cdot\|$ la norma euclídea estándar en \mathbb{R}^n .

Proposición 117. Sean un abierto $U \subseteq \mathbb{R}^n$, una función $f : U \rightarrow \mathbb{R}^n$ de clase \mathcal{C}^1 y un subconjunto convexo $E \subseteq U$. Si hay una constante $\lambda > 0$ y una matriz $P \in O(n)$ tales que

$$\text{para cualesquiera } p \in E, v \in \mathbb{R}^n \text{ se tiene } v^t P (Df)_p v \geq \lambda \|v\|^2 , \quad (82)$$

entonces $f|_E$ es coerciva con constante de coercividad λ .

Atención: la condición (82) es estrictamente más fuerte que la condición

$$\|(Df)_p v\| \geq \sqrt{\lambda} \|v\| , \quad \text{para } p \in E, v \in \mathbb{R}^n ,$$

que, por ejemplo, la cumple CP en el semiplano $\{r \geq 1\}$ sin que sea inyectiva en dicho semiplano.

Demostración de la proposición 117. Pongamos $g = Pf$. Como P es ortogonal, la función f es coerciva con constante de coercividad λ si y sólo si lo es g . La condición (82) afirma que para cualesquiera $p \in E$ y $v \in \mathbb{R}^n$,

$$v^t (Dg)_p v \geq \lambda \|v\|^2 .$$

Sean $p, q \in E$ con $p \neq q$ y pongamos $v = q - p$. Como $p + tv \in E$ para todo $t \in [0, 1]$, podemos definir la función $h(t) : [0, 1] \rightarrow \mathbb{R}$ dada por $h(t) = v \cdot g(p + tv)$. Al ser h una función escalar, se le puede aplicar el teorema de los incrementos finitos y deducir que existe $\theta \in (0, 1)$, tal que

$$v \cdot g(q) - v \cdot g(p) = h(1) - h(0) = (1 - 0) h'(\theta) = v^t (Dg)_{p+\theta v} v \geq \lambda \|v\|^2,$$

de donde:

$$\|v\| \cdot \|g(q) - g(p)\| \geq v \cdot (g(q) - g(p)) \geq \lambda \|v\|^2,$$

desigualdad en la que podemos dividir por $\|v\|$, ya que $v \neq \mathbf{0}$, para obtener:

$$\|g(q) - g(p)\| \geq \lambda \|v\| = \lambda \|q - p\|,$$

luego g y f son coercivas con constante de coercividad λ . □

Veamos, con un ejemplo, lo importante que es la hipótesis de convexidad del conjunto E . En el abierto $U = \mathbb{R}^2 \setminus ([0, +\infty) \times \{0\})$, que no es convexo, definimos la función

$$f : U \rightarrow \mathbb{R}^2, \quad f(x, y) = \begin{cases} (x, y + \sqrt{x^2 + 1} - 1) & \text{si } y < 0 \text{ y } x \geq 0 \\ (x, y) & \text{en el resto de } U \end{cases}$$

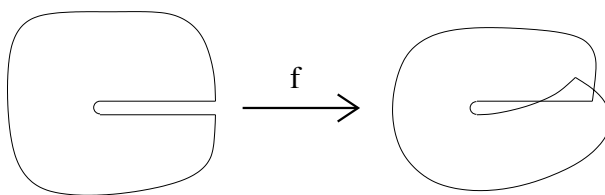
Se comprueba que f es de clase \mathcal{C}^1 y que su matriz jacobiana es la siguiente

$$Df = \begin{cases} \begin{pmatrix} 1 & 0 \\ \frac{x}{\sqrt{x^2+1}} & 1 \end{pmatrix} & \text{si } y < 0 \text{ y } x \geq 0 \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \text{en el resto de } U \end{cases}$$

Como siempre es $|x/\sqrt{x^2+1}| \leq 1$, vemos que para todo $(x, y) \in U$ y todo $v = (a, b)$ se tiene

$$v^t (Df)_{(x,y)} v \geq a^2 + b^2 - |ab| \geq \frac{a^2 + b^2}{2} = \frac{1}{2} \|v\|^2,$$

sin embargo f ni siquiera es inyectiva, por ejemplo $f(\sqrt{3}, -1/2) = (\sqrt{3}, 1/2) = f(\sqrt{3}, 1/2)$



3.6 La cuestión de la suprayectividad

La parte de suprayectividad del teorema 115 (que la imagen por f de $B(x_0, r)$ contiene totalmente a cierta bola con centro en $f(x_0)$) es consecuencia directa de las proposiciones 111 y 117 y del siguiente enunciado.

Proposición 118. Sean un abierto $U \subseteq \mathbb{R}^n$, un punto $x_0 \in U$ y $f : U \rightarrow \mathbb{R}^n$. Si hay números $r, \lambda > 0$ tales que f cumple las dos condiciones siguientes

1. La imagen $f(B(x_0, r))$ es un abierto de \mathbb{R}^n .
2. $f|_{B(x_0, r)}$ es continua y coerciva, con constante de coercividad λ .

entonces el abierto $f(B(x_0, r))$ contiene a la bola $B(f(x_0), \lambda r)$.

Observa que aquí no suponemos f diferenciable.

Demostración.

Pongamos $V = f(B(x_0, r))$. Falta ver que V contiene a $B(f(x_0), \lambda r)$.

Fijamos un vector unitario cualquiera ω , es decir que $\|\omega\| = 1$, y consideramos el siguiente conjunto de números

$$T_\omega = \{c \in [0, \lambda r] : f(x_0) + t\omega \in V \text{ para todo } t \in [0, c]\}.$$

Como $0 \in T_\omega$, este conjunto es no vacío y acotado, lo que nos permite definir

$$c_\omega \stackrel{\text{def}}{=} \sup T_\omega.$$

La función $\alpha_\omega(t) = f(x_0) + t\omega$ es continua y con $\alpha_\omega(0) \in V$. Como V es abierto, existe $\delta > 0$ tal que $\alpha_\omega([0, \delta]) \subset V$ y tenemos $0 < \delta \leq c_\omega \leq \lambda r$. En particular $c_\omega > 0$ y el intervalo $[0, c_\omega)$ es no vacío. De las definiciones de T_ω y c_ω se deduce

$$\alpha_\omega([0, c_\omega)) \subset V. \quad (83)$$

De suponer $c_\omega < \lambda r$ vamos a deducir una contradicción. Existe una sucesión estrictamente creciente $\{c_j\}_{j=1}^\infty \subset T_\omega$ que converge a c_ω . Por la definición de T_ω se verifica $\alpha_\omega(c_j) \in V$ para $j = 1, 2, 3, \dots$, luego existe una sucesión de puntos $\{p_j\}_{j=1}^\infty \subset B(x_0, r)$ tal que

$$f(p_j) = \alpha_\omega(c_j) \text{ para todo } j.$$

Como $f|_{B(x_0, r)}$ es coerciva, con constante de coercividad λ , para todo j se tiene

$$d(x_0, p_j) \leq \frac{1}{\lambda} d(f(x_0), f(p_j)) = \frac{1}{\lambda} d(\alpha_\omega(0), \alpha_\omega(c_j)) = \frac{c_j}{\lambda} \leq \frac{c_\omega}{\lambda},$$

y la sucesión $\{p_j\}$ está toda ella contenida en la bola cerrada $\overline{B}\left(x_0, \frac{c_\omega}{\lambda}\right)$. Como estamos suponiendo $c_\omega < \lambda r$, tenemos $\frac{c_\omega}{\lambda} < r$, es decir que $\{p_j\}$ está contenida en una bola cerrada a su vez contenida en la bola abierta $B(x_0, r)$.

Hay una subsucesión $\{p_{j_s}\}_{s=1}^\infty$ convergente a un punto $p_0 \in \overline{B}\left(x_0, \frac{c_\omega}{\lambda}\right)$. El punto p_0 está en la bola abierta $B(x_0, r)$ y, por la continuidad de f y $\alpha_\omega(t)$, tenemos $\alpha_\omega(c_\omega) = f(p_0) \in V$. Pero entonces la situación es ésta:

1. se tiene $\alpha_\omega([0, c_\omega)) \subset V$,
2. se tiene $\alpha_\omega(c_\omega) \in V$,
3. existe $\varepsilon > 0$ tal que $\alpha_\omega([c_\omega, c_\omega + \varepsilon)) \subset V$, porque V es abierto y $\alpha_\omega(t)$ es continua,

luego en realidad $\alpha_\omega([0, c_\omega + \varepsilon)) \subset V$. Como estamos suponiendo $c_\omega < \lambda r$, podemos elegir ε de modo que $c_\omega + \varepsilon \leq \lambda r$ y sería $c_\omega + \varepsilon \in T_\omega$. Esto contradice que c_ω sea el supremo de T_ω . Por reducción al absurdo, queda demostrado que $c_\omega = \lambda r$.

Ahora la fórmula (83) nos da $\alpha_\omega([0, \lambda r)) \subset V$. Al ser esto cierto para todos los vectores unitarios ω , toda la bola abierta

$$B(f(x_0), \lambda r) = \bigcup_{\|\omega\|=1} \alpha_\omega([0, \lambda r)),$$

está contenida en V . □

4 Subvariedades y extremos condicionados

Cada espacio \mathbb{R}^N tiene unos subconjuntos especiales llamados subvariedades. Intuitivamente, una subvariedad de dimensión geométrica n es un “espacio posiblemente curvilíneo de dimensión n ”. Decimos “posiblemente” porque a los subespacios afines (que son “rectilíneos”) también los consideramos subvariedades.

Veremos métodos para buscar los máximos y mínimos de una función en una subvariedad. Éstos son los máximos y mínimos condicionados.

4.1 Difeomorfismos

Definición 119. Sea ℓ positivo (aquí no admitimos el caso $\ell = 0$) y sean $W, W' \subseteq \mathbb{R}^N$ dos abiertos. Un **difeomorfismo entre W y W' de clase \mathcal{C}^ℓ** es una biyección $\sigma : W \rightarrow W'$ tal que ella y su inversa σ^{-1} son de clase \mathcal{C}^ℓ .

No debe confundirse este concepto con el de *biyección de clase \mathcal{C}^ℓ* . Por ejemplo, la función

$$f : \mathbb{R} \longrightarrow \mathbb{R} \quad , \quad f(x) = x^3 \quad ,$$

es una biyección de clase \mathcal{C}^∞ , pero su inversa $f^{-1}(y) = \sqrt[3]{y}$ no es diferenciable en $y = 0$. Por lo tanto esta f no es un difeomorfismo.

Para verificar que $\sigma : W \rightarrow W'$ es un difeomorfismo, tenemos que comprobar

1. Que σ es biyectiva de W a W' .
2. Que σ es \mathcal{C}^ℓ .
3. Que σ^{-1} también es \mathcal{C}^ℓ .

Veamos cómo hacer eso mismo con algo menos de esfuerzo. Dados un abierto $W \subseteq \mathbb{R}^N$ y una función vectorial $\sigma : W \rightarrow \mathbb{R}^N$ de clase \mathcal{C}^ℓ , con $\ell \geq 1$, si comprobamos

(a) que σ es inyectiva,

(b) que la matriz jacobiana $(D\sigma)_p$ es invertible en todo punto $p \in W$,

entonces σ es una biyección de W al conjunto imagen $W' = \sigma(W)$, el cual es un abierto de \mathbb{R}^N por la proposición 111 del apartado 3.3.2, y la inversa $\sigma^{-1} : W' \rightarrow W$ es también de clase \mathcal{C}^ℓ por el teorema 113 del apartado 3.3.3. Mientras que comprobar la condición (b) es fácil, no existe ningún método sistemático para comprobar la condición (a): es necesario hacer un estudio particular para cada función σ .

Por lo dicho al principio del apartado 3.3.3, la condición (b) es siempre necesaria para que σ pueda ser un difeomorfismo. Por eso $f(x) \equiv x^3$ no es un difeomorfismo, ya que hay un punto donde se anula su derivada.

Un difeomorfismo $\sigma = (\sigma_1, \dots, \sigma_N)$ puede entenderse como un **sistema de coordenadas curvilíneas** según explicamos a continuación. Las funciones $\sigma_1, \dots, \sigma_N : W \rightarrow \mathbb{R}$ son unas **coordenadas** en U en el sentido de que cada punto $p \in W$ está determinado por los números $\sigma_1(p), \dots, \sigma_N(p)$ y de hecho la función σ^{-1} , que reconstruye el punto p a partir de esos números, es de clase \mathcal{C}^ℓ .

4.2 Subvariedades de \mathbb{R}^N

4.2.1 La definición

Definición 120. Fijamos dos enteros $n, k \geq 0$ y escribimos $N = n + k$. Decimos que un subconjunto $X \subseteq \mathbb{R}^N$ es una **subvariedad de dimensión geométrica n** si X es no vacío y para todo punto $p \in X$ existen abiertos $W, W' \subseteq \mathbb{R}^N$ y un difeomorfismo $\sigma : W \rightarrow W'$ tales que

$$p \in W \quad \text{y} \quad \sigma(X \cap W) = W' \cap \mathbb{A} \quad , \quad (84)$$

para algún subespacio afín $\mathbb{A} \subseteq \mathbb{R}^N$ de dimensión n .

Primeros comentarios a esta definición:

1. Al vacío \emptyset no lo consideramos una subvariedad.
2. Las subvariedades de dimensión 0 son los conjuntos de puntos aislados, mientras que las de dimensión N son los abiertos de \mathbb{R}^N .
3. Se puede suponer, sin pérdida de generalidad, que \mathbb{A} es el subespacio afín estándar

$$\mathbb{R}^n \times \{\mathbf{0}\} = \{ (x_1, \dots, x_N) \in \mathbb{R}^N : x_{n+1} = x_{n+2} = \dots = x_N = 0 \} .$$

Para ello tomamos, por ejemplo, una biyección $\tau : \mathbb{R}^N \rightarrow \mathbb{R}^N$ que sea una rotación seguida de una traslación y tal que $\tau(\mathbb{A}) = \mathbb{R}^n \times \{\mathbf{0}\}$. Entonces $W'' = \tau(W')$ es un abierto, la compuesta $\tau \circ \sigma : W \rightarrow W''$ es un difeomorfismo y se tiene

$$\tau \circ \sigma(X \cap W) = W'' \cap (\mathbb{R}^n \times \{\mathbf{0}\}) .$$

4. La fórmula (84) equivale a esta otra:

$$p \in X \cap W = \sigma^{-1}(\mathbb{A}) , \quad (85)$$

que nos permite olvidarnos de W' .

5. Veremos en el apartado 4.2.5 por qué es necesario exigir que σ sea un difeomorfismo, no siendo adecuado contentarse con que sea una biyección de clase \mathcal{C}^ℓ para algún $\ell > 0$.

4.2.2 Unicidad de la dimensión geométrica en cada punto

Proposición 121. Sean $W_1, W'_1, W_2, W'_2 \subseteq \mathbb{R}^N$ cuatro abiertos, con dos difeomorfismos

$$\sigma_1 : W_1 \longrightarrow W'_1 , \quad \sigma_2 : W_2 \longrightarrow W'_2 ,$$

y dos subespacios afines $\mathbb{A}_1, \mathbb{A}_2 \subseteq \mathbb{R}^N$ tales que para $j = 1, 2$ se tiene

$$p \in W_j \quad y \quad \sigma_j(X \cap W_j) = W'_j \cap \mathbb{A}_j ,$$

entonces $\dim \mathbb{A}_1 = \dim \mathbb{A}_2$.

Demostración. Podemos suponer que

$$\mathbb{A}_1 = \mathbb{R}^{n_1} \times \{\mathbf{0}\} , \quad \mathbb{A}_2 = \mathbb{R}^{n_2} \times \{\mathbf{0}\} .$$

Consideremos el abierto $W = W_1 \cap W_2$. Tenemos $p \in W$ y las imágenes $\sigma_1(W), \sigma_2(W)$ son abiertos tales que para $j = 1, 2$ se tiene

$$\sigma_j(X \cap W) = \sigma_j(W) \cap (\mathbb{R}^{n_j} \times \{\mathbf{0}\}) . \quad (86)$$

El siguiente diagrama de biyecciones

$$\sigma_1(W) \xrightarrow{\sigma_1^{-1}} W \xrightarrow{\sigma_2} \sigma_2(W) ,$$

hace posible definir la biyección compuesta

$$\tau_{1,2} = \sigma_2 \circ \sigma_1^{-1} : \sigma_1(W) \longrightarrow \sigma_2(W) ,$$

que es de clase \mathcal{C}^ℓ . La compuesta análoga $\tau_{2,1} : \sigma_2(W) \rightarrow \sigma_1(W)$ es también \mathcal{C}^ℓ y es inversa de $\tau_{1,2}$, luego ambas son difeomorfismos. La fórmula (86) implica que

$$\tau_{1,2} \left(\sigma_1(W) \cap (\mathbb{R}^{n_1} \times \{\mathbf{0}\}) \right) = \sigma_2(W) \cap (\mathbb{R}^{n_2} \times \{\mathbf{0}\}) ,$$

y $\tau_{2,1}$ hace lo análogo. Consideremos $a = \sigma_1(p)$. Si escribimos $\tau_{1,2} \equiv (f_1, \dots, f_N)$, entonces cerca del punto a tenemos la siguiente identidad:

$$\tau_{1,2}(x_1, \dots, x_{n_1}, 0, \dots, 0) \equiv (f_1(x), \dots, f_{n_2}(x), 0, \dots, 0).$$

de la que se deduce que la diferencial $(d\tau_{1,2})_a$ lleva el subespacio vectorial $\mathbb{R}^{n_1} \times \{0\}$ dentro del subespacio vectorial $\mathbb{R}^{n_2} \times \{0\}$. Como las diferenciales de un difeomorfismo son aplicaciones lineales biyectivas, deducimos que $n_1 \leq n_2$.

Utilizando ahora $\tau_{2,1}$, deducimos de igual modo que $n_2 \leq n_1$. En definitiva $n_1 = n_2$. \square

La proposición 121 conduce a un sexto comentario a la definición 120: en dicha definición se fija el número n de antemano; por lo tanto la dimensión geométrica es, *por cómo se redacta la definición*, el mismo número n en todos los puntos de la subvariedad y, por la proposición 121, está totalmente determinado por la subvariedad.

Por ejemplo, al siguiente conjunto

$$\{(x, y, z) : z = 0\} \cup \{(x, y, z) : x = 0, z = 1\},$$

unión disjunta de un plano afín y un recta afín, no lo consideramos, *por definición*, como una subvariedad de \mathbb{R}^3 .

Vamos a ver que tenemos dos inclusiones estrictas:

$$\begin{aligned} \{\text{subvariedades de dimensión 1 de } \mathbb{R}^N\} &\subsetneq \{\text{curvas en } \mathbb{R}^N\}, \\ \{\text{subvariedades de dimensión 2 de } \mathbb{R}^N\} &\subsetneq \{\text{superficies en } \mathbb{R}^N\}. \end{aligned}$$

Concretamente, en los apartados 4.2.5 y 4.2.7 veremos curvas que no son subvariedades de dimensión 1, y en el apartado 4.2.6 veremos superficies que no son subvariedades de dimensión 2.

4.2.3 Caracterización por ecuaciones

Si $N = n + k$ y $X \subseteq \mathbb{R}^N$ es una subvariedad de dimensión n , al número $k = N - \dim X$ lo llamamos la **codimensión de X en \mathbb{R}^N** . Vamos a ver ahora que entonces X se define *localmente* por sistemas de k ecuaciones escalares.

Teorema 122. Sea $N = n + k$. Para un subconjunto no vacío $X \subseteq \mathbb{R}^N$, son equivalentes:

- (1) X es una subvariedad de \mathbb{R}^N con dimensión geométrica n .
- (2) Para cada $p \in X$ existen un abierto $A \subseteq \mathbb{R}^N$ con $p \in A$, un vector $b \in \mathbb{R}^k$ y una función $F \equiv (F_1, \dots, F_k) : A \rightarrow \mathbb{R}^k$, de clase al menos \mathcal{C}^1 , tales que
 - (a) Los gradientes $\nabla F_1, \dots, \nabla F_k$ son linealmente independientes en todo punto de A .
 - (b) Se tiene la igualdad conjuntista:

$$X \cap A = F^{-1}(\{b\}) = \{x \in A : F_1(x) = b_1, \dots, F_k(x) = b_k\}. \quad (87)$$

Demostración.

Primera parte: (1) \implies (2). Aquí suponemos que X es una subvariedad con $\dim X = n$. Dado $p \in X$ existen un abierto $W \subseteq \mathbb{R}^N$, con $p \in W$, y un difeomorfismo

$$\sigma \equiv (\sigma_1, \dots, \sigma_N) : W \rightarrow \sigma(W),$$

tales que

$$X \cap W = \sigma^{-1}(\mathbb{R}^n \times \{0\}) = \{x \in W : \sigma_{n+1}(x) = 0, \dots, \sigma_{n+k}(x) = 0\}.$$

Al ser la jacobiana $D\sigma$ invertible en todo punto de W , sus filas son linealmente independientes. En particular $\nabla \sigma_{n+1}, \dots, \nabla \sigma_{n+k}$ son las k últimas filas de dicha jacobiana, luego

son linealmente independientes en todo punto de W . Así, pues, se cumple (b) con $A = W$, $F \equiv (\sigma_{n+1}, \dots, \sigma_{n+k})$ y $b = \{0\}$.

Segunda parte: (2) \implies (1). Dado cualquier $p \in W$, ahora estamos suponiendo que existen A, b, F tales que $p \in A$ y se cumplen las condiciones (a) y (b). Vamos a hallar otro abierto $p \in B \subseteq A$ (en general, más pequeño que A) y un difeomorfismo $\sigma : B \rightarrow \sigma(B)$, tales que

$$X \cap B = (\sigma|_B)^{-1}(\mathbb{R}^n \times \{b\}). \quad (88)$$

La condición (a) implica que las jacobianas $(DF_1)_p, \dots, (DF_k)_p$ son linealmente independientes. Elijamos otras n filas $\mathbf{a}_1, \dots, \mathbf{a}_n \in M_{1 \times N}(\mathbb{R})$ tales que $\mathbf{a}_1, \dots, \mathbf{a}_n, (DF_1)_p, \dots, (DF_k)_p$ aún sean linealmente independientes. Hecho eso, definamos

$$\sigma : A \longrightarrow \mathbb{R}^N, \quad \sigma(x) \equiv (\mathbf{a}_1 x, \dots, \mathbf{a}_n x, F_1(x), \dots, F_k(x)), \quad (89)$$

y las filas de $(D\sigma)_p$ son $\mathbf{a}_1, \dots, \mathbf{a}_n, DF_1(p), \dots, DF_k(p)$, linealmente independientes por construcción, luego $(D\sigma)_p$ es una matriz $N \times N$ invertible.

Como σ es al menos de clase \mathcal{C}^1 , hay un abierto $p \in B \subseteq A$ tal que $\sigma : B \rightarrow \sigma(B)$ es un difeomorfismo. Al ser F_1, \dots, F_k las k últimas entradas del vector $F(x)$, para $x \in B$ se tiene

$$x \in X \iff F(x) = b \iff \sigma(x) \in \mathbb{R}^n \times \{b\},$$

luego $\sigma|_B$ es un difeomorfismo que cumple (88). \square

Corolario 123. Sean un abierto $W \subseteq \mathbb{R}^N$ y una función $F : W \rightarrow \mathbb{R}^k$ de clase al menos \mathcal{C}^1 . Fijado un vector $b \in \mathbb{R}^k$, consideramos el conjunto

$$X = F^{-1}(\{b\}) = \{x \in W : F(x) = b\}.$$

Si la jacobiana $(DF)_x$ tiene rango k en todo punto $x \in X$ (es decir, en los puntos en los que se satisface el sistema de ecuaciones $F(x) = b$) entonces X es o el vacío \emptyset o una subvariedad de dimensión $n = N - k$ de \mathbb{R}^N .

Atención: no toda subvariedad de \mathbb{R}^N se obtiene así, véase el apartado 4.5.

Llamamos **hipersuperficies de \mathbb{R}^N** a las subvariedades de codimensión 1, es decir de dimensión $N - 1$. Si $W \subseteq \mathbb{R}^N$ es un abierto y $F : W \rightarrow \mathbb{R}$ es una función escalar de clase al menos \mathcal{C}^1 , dado $b \in \mathbb{R}$ una condición suficiente para que el **conjunto de nivel** $X = \{x \in W : F(x) = b\}$ sea una hipersuperficie es $X \neq \emptyset$ y que el gradiente ∇F no se anule en ningún punto de X , aunque se permite que ∇F se anule fuera de X .

4.2.4 Grafos generalizados

En la prueba del teorema 122 se han elegido filas $\mathbf{a}_1, \dots, \mathbf{a}_n$ que junto con las jacobianas $(DF_1)_p, \dots, (DF_k)_p$ forman una base del espacio de matrices $1 \times N$. Si denotamos por $\{\mathbf{e}_1, \dots, \mathbf{e}_N\}$ la base estándar de \mathbb{R}^N , se pueden elegir las \mathbf{a}_j como traspuestas de vectores de esta base. Por ejemplo, si $n = 3$ y $k = 2$ (con lo cual $N = 5$) entonces las filas

$$\mathbf{e}_1^t, \mathbf{e}_2^t, \mathbf{e}_4^t, (DF_1)_p, (DF_2)_p,$$

son linealmente independientes si y sólo si en la jacobiana $(DF)_p$ las columnas 3ª y 5ª forman una matriz invertible 2×2 . En tal caso el difeomorfismo $\sigma : B \rightarrow \sigma(B)$, que se construye en la fórmula (89), viene dado por la siguiente fórmula

$$\sigma(x_1, x_2, x_3, x_4, x_5) \equiv (x_1, x_2, x_4, F_1(x), F_2(x)),$$

y se demuestra, imitando la prueba del teorema 116 del apartado 3.4, que $X \cap B$ es un **grafo** en el que las incógnitas x_1, x_2, x_4 son libres de recorrer un abierto $U \subseteq \mathbb{R}^3$ mientras que las x_3, x_5 son funciones al menos \mathcal{C}^1 de las libres:

$$\begin{aligned} X \cap B &= \{x \in B : (x_1, x_2, x_4) \in U, x_3 = \varphi(x_1, x_2, x_4), x_5 = \psi(x_1, x_2, x_4)\} = \\ &= \{(x_1, x_2, \varphi(x_1, x_2, x_4), x_4, \psi(x_1, x_2, x_4)) : (x_1, x_2, x_4) \in U\}. \end{aligned}$$

Observa que las variables x_3, x_5 , que se han puesto como funciones de las otras en $X \cap B$, son precisamente aquellas para las que la correspondiente submatriz de la jacobiana es invertible

$$(D_{x_3, x_5} F)_p = [F_{x_3}(p) \mid F_{x_5}(p)]_{2 \times 2}.$$

La pareja (φ, ψ) es una función que va de un abierto (representado por U) del espacio $\mathbb{R}_{x_1 x_2 x_4}^3$ al espacio $\mathbb{R}_{x_3 x_5}^2$ y $X \cap B$ es el grafo de esta función.

En general, para cada punto p de una subvariedad $X \subseteq \mathbb{R}^N$, con $\dim X = n$ y $0 < n < N$, se encuentra un abierto $B \subseteq \mathbb{R}^N$ con $p \in B$ y una partición de $\{1 \dots, N\}$ en dos conjuntos no vacíos

$$\{1, \dots, N\} = \{i_1, \dots, i_n\} \sqcup \{j_1, \dots, j_k\}, \quad (90)$$

tales que $X \cap B$ es el conjunto de los (x_1, \dots, x_N) con $(x_{i_1}, \dots, x_{i_n})$ recorriendo un cierto abierto $U \subseteq \mathbb{R}^n$ y al mismo tiempo $(x_{j_1}, \dots, x_{j_k}) = \varphi(x_{i_1}, \dots, x_{i_n})$ para cierta función $\varphi : U \rightarrow \mathbb{R}^k$ al menos de clase C^1 . En este sentido $X \cap B$ es el grafo de φ , aunque con las coordenadas puestas en un orden tal vez distinto del habitual.

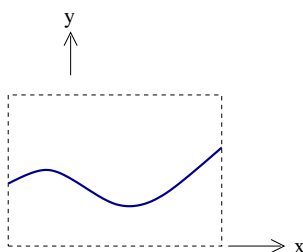
Puesto que las coordenadas $(x_{i_1}, \dots, x_{i_n})$ pueden moverse en un abierto de \mathbb{R}^n , el grafo $X \cap B$ tiene **n grados de libertad**.

La dimensión geométrica de una subvariedad es igual al número de sus grados de libertad.

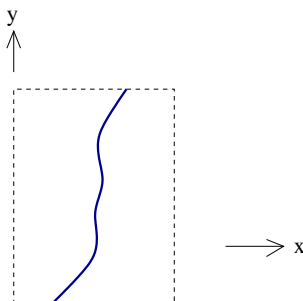
4.2.5 Tipos de grafo en el plano

Fijados N y n , el número total de particiones como en (90) es $\binom{N}{n}$. Así, pues, hay $\binom{N}{n}$ tipos diferentes de grafos en \mathbb{R}^N con n grados de libertad. En el plano \mathbb{R}_{xy}^2 tenemos dos tipos de grafo con 1 grado de libertad:

- Aquellos en los que y es una función, de clase C^1 o mejor, de la variable x la cual, a su vez, recorre un abierto de su eje



- Aquellos en los que x es una función C^1 de y , que a su vez recorre un abierto de eje y .



Esta descripción local como grafo permite reconocer si un subconjunto $X \subset \mathbb{R}^2$ es una subvariedad o no. Veamos dos ejemplos.

Ejemplo 1. Estudiamos la curva $X_1 = \{(x, y) : y = |x|\}$ en el entorno del punto $(0, 0)$. En dicho entorno no es un grafo en el que x sea función de y , porque un valor $y > 0$ corresponde a dos valores $x = \pm y$, además un valor $y < 0$ no corresponde a nada. Sí que es X_1 un grafo

con y función de x , pero la función en cuestión es $\varphi(x) \equiv |x|$ que no es diferenciable en el valor $x = 0$, correspondiente al punto $(0,0)$ que estamos estudiando. Luego X_1 no es una subvariedad de \mathbb{R}^2 .

Ejemplo 2. Estudiamos la curva

$$X_2 = \{(x, y) : y = 2x + |x|\} = \left\{ (x, y) : x = \frac{2}{3}y - \frac{1}{3}|y| \right\},$$

en el entorno de $(0,0)$. Es un grafo de las dos maneras posibles, de dos funciones no diferenciables. Luego tampoco es este conjunto una subvariedad de \mathbb{R}^2 .

Si quitamos el punto problemático, tanto $X_1 \setminus \{(0,0)\}$ como $X_2 \setminus \{(0,0)\}$ son subvariedades de \mathbb{R}^2 .

La lección que nos enseñan estos ejemplos es que en una subvariedad de dimensión 1 del plano están prohibidas las *esquinas*. Para conseguir esta prohibición, ha sido necesario exigir en la definición 120 que σ sea un difeomorfismo. Veamos con un ejemplo que, si sólo pedimos a σ que sea una biyección de clase \mathcal{C}^ℓ , las esquinas son posibles. La siguiente función

$$\sigma : \mathbb{R}^2 \longrightarrow \mathbb{R}^2, \quad \sigma(x, y) \equiv (x^3, y^3 - |x^3|),$$

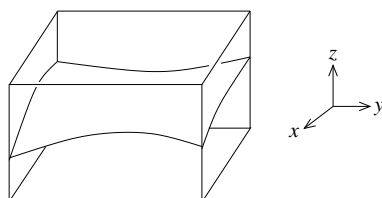
es biyectiva y de clase \mathcal{C}^2 , aunque no un difeomorfismo porque $\sigma^{-1}(a, b) = (\sqrt[3]{a}, \sqrt[3]{b + |a|})$ no es diferenciable. Se comprueba de inmediato que $\sigma(X_1) = \mathbb{R} \times \{0\}$.

Se puede “aplanar” una esquina mediante una biyección diferenciable, pero no mediante un difeomorfismo.

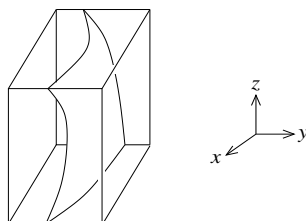
4.2.6 Tipos de grafo en \mathbb{R}^3

En el espacio \mathbb{R}^3 tenemos $\binom{3}{2} = 3$ tipos de grafo con 2 grados de libertad:

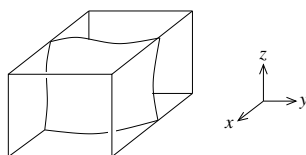
- z es función \mathcal{C}^1 de (x, y) , que a su vez recorre un abierto A_{xy} del plano xy .



- y es función \mathcal{C}^1 de (x, z) , que a su vez recorre un abierto A_{xz} del plano xz .



- x es función \mathcal{C}^1 de (y, z) , que a su vez recorre un abierto A_{yz} del plano yz .



Podemos usar esas descripciones locales para decidir si un subconjunto $X \subset \mathbb{R}^3$ es una subvariedad o no. Veamos tres ejemplos.

Ejemplo 1. Estudiamos $X_1 = \{(x, y, z) : z = |x - y|\}$ en el entorno de $(0, 0, 0)$. En dicho entorno X_1 sólo es un grafo del tipo $z = f(x, y)$, con f no diferenciable en el valor $(x, y) = (0, 0)$ correspondiente al punto $(0, 0, 0)$. Luego la superficie X_1 no es una subvariedad de \mathbb{R}^3 .

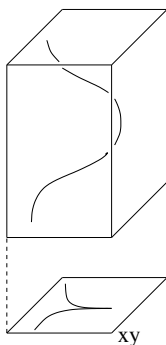
Ejemplo 2. Estudiamos $X_2 = \{(x, y, z) : z = \sqrt{x^2 + y^2}\}$ en el entorno de $(0, 0, 0)$. En dicho entorno sólo es un grafo del tipo $z = g(x, y)$ y de nuevo la función g no es diferenciable en $(x, y) = (0, 0)$, luego tampoco la superficie X_2 es una subvariedad de \mathbb{R}^3 .

Ejemplo 3. El conjunto $X_3 = \{(x, y, z) : yz = 0\}$, unión de dos planos que se cortan, es una superficie pero no una subvariedad de \mathbb{R}^3 , porque cerca de los puntos $(x, 0, 0)$ no es un grafo de ninguno de los tres tipos.

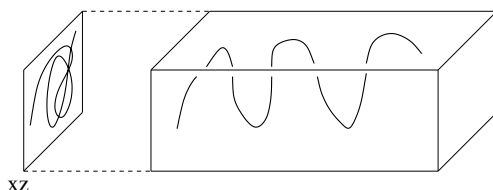
Estos ejemplos muestran tres de los muchos “accidentes” que están prohibidos en una subvariedad de dimensión 2 de \mathbb{R}^3 : esquinas, picos y cruces.

Tenemos $\binom{3}{1} = 3$ tipos de grafo en \mathbb{R}^3 con 1 grado de libertad:

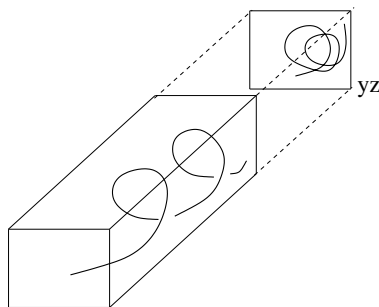
- (x, y) es una función \mathcal{C}^1 de z , que a su vez recorre un abierto A_z del eje z . Es decir, el grafo de una función vectorial $A_z \rightarrow \mathbb{R}_{xy}^2$.



- (x, z) es una función \mathcal{C}^1 de y , que a su vez recorre un abierto A_y del eje y . Es decir, el grafo de una función vectorial $A_y \rightarrow \mathbb{R}_{xz}^2$.



- (y, z) es una función \mathcal{C}^1 de x , que a su vez recorre un abierto A_x del eje x . Es decir, el grafo de una función vectorial $A_x \rightarrow \mathbb{R}_{yz}^2$.



4.2.7 La capa protectora

Aquí utilizamos las notaciones de la definición 120. El subespacio afín \mathbb{A} es un cerrado de \mathbb{R}^N , por lo tanto $W' \setminus \mathbb{A}$ es un abierto. Como σ es continua, la preimagen

$$W \setminus (X \cap W) = \sigma^{-1}(W' \setminus \mathbb{A}),$$

es también un abierto de \mathbb{R}^N al que llamamos **capa protectora del trozo $X \cap W$** . Le damos este nombre porque protege al trozo de subvariedad $X \cap W$ de la siguiente manera:

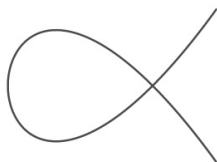
Como $W \setminus (X \cap W) = W \setminus X$, al conjunto X le está prohibido entrar en este conjunto. Dicho de otro modo, la única parte de X que entra en W es $\sigma^{-1}(\mathbb{A})$.

El papel que juega esta “protección” se entiende mejor con dos ejemplos.

Ejemplo 1. La *curva algebraica* $\Gamma = \{(x, y) : x^3 + x^2 - y^2 = 0\}$ es recorrida por el camino

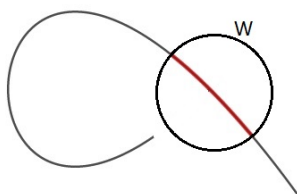
$$\alpha(t) : \mathbb{R} \rightarrow \mathbb{R}^2, \quad \alpha(t) \equiv (t^1 - 1, t^3 - t),$$

y es una **cúbica nodal**, lo que significa que tiene un **nodo**, o sea un punto en el que se cruza consigo misma

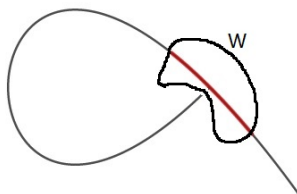


Fijamos el intervalo abierto $J_0 = (-1'5, -0'5)$, centrado en $t = -1$, y representamos el trozo $\alpha(J_0)$ en color rojo.

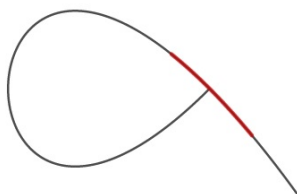
Si consideramos el conjunto $X_1 = \alpha(-\infty, 0'5)$, entonces encontramos con facilidad una capa protectora para el trozo $\alpha(J_0)$, es decir encontramos un abierto W que contiene a $\alpha(J_0)$ y no contiene ninguna otra parte de X_1



Para el conjunto $X_2 = \alpha(-\infty, 0'97)$, todavía encontramos una capa protectora para $\alpha(J_0)$, pero se va comprimiendo mucho por la parte de abajo



Para $X_3 = \alpha(-\infty, 1)$, la capa protectora de $\alpha(J_0)$ ya no existe: no importa lo pequeño que tomemos un entorno W del conjunto $\alpha(J_0)$, hay una parte de X_3 distinta de $\alpha(J_0)$ que entra en W



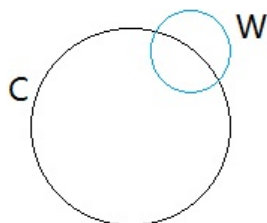
El conjunto $X_3 = \alpha(-\infty, 1)$ no es una subvariedad del plano, a pesar de que el camino $\alpha(t)$ *todavía es inyectivo* en el intervalo $(-\infty, 1)$.

Hay superficies en \mathbb{R}^3 , muy parecidas a este ejemplo, que no son subvariedades de dimensión 2. Tampoco es una subvariedad la curva nodal completa Γ a causa del nodo $(0, 0)$, que es el único punto de Γ en el que se anula el gradiente $\nabla(x^3 + x^2 - y^2)$. No se contradice, pues, el corolario 123.

Ejemplo 2. Consideramos la función $F(x, y) \equiv x^2 + y^2$, que es \mathcal{C}^∞ en todo el plano. La circunferencia unidad $C = \{(x, y) : F(x, y) = 1\}$ es trazada por el camino

$$\beta(t) : \mathbb{R} \longrightarrow \mathbb{R}^2, \quad \beta(t) \equiv (\cos t, \sin t),$$

que no es inyectivo (por el contrario, es periódico). Sin embargo C es una subvariedad de dimensión 1 del plano, porque el gradiente $\nabla F \equiv (2x, 2y)$ no se anula en ningún punto de C . Para cualquier arco $\alpha(J)$ de la circunferencia encontramos una capa protectora: un abierto W que contiene a ese arco y tal que, aunque el camino $\alpha(t)$ vuelve a entrar infinitas veces más en W , cada vez que lo hace es volviendo a pasar por ese mismo arco sin jamás entrar en la capa protectora.



4.3 Espacios tangente y normal

No indicamos aparte la definición del espacio tangente, porque dicha definición ocupa casi todo el apartado 4.3.1.

4.3.1 Espacio tangente

Sea $X \subseteq \mathbb{R}^N$ una subvariedad con $\dim X = n$. Dado un punto $p \in X$ elegimos un abierto $p \in W \subseteq \mathbb{R}^n$ y un difeomorfismo $\sigma : W \rightarrow \sigma(W)$ con $\sigma(X \cap W) = \sigma(W) \cap (\mathbb{R}^n \times \{\mathbf{0}\})$. Como $\mathbb{R}^n \times \{\mathbf{0}\}$ es un subespacio vectorial de \mathbb{R}^N y las diferenciales de σ^{-1} son biyectivas, el conjunto

$$\mathbb{V} = (d\sigma^{-1})_{\sigma(p)}(\mathbb{R}^n \times \{\mathbf{0}\}), \quad (91)$$

es un subespacio vectorial de \mathbb{R}^N , con dimensión lineal igual a la dimensión geométrica n de X . Veamos ahora que \mathbb{V} tiene la siguiente propiedad:

Dados cualquier abierto $U \subseteq \mathbb{R}^s$ y cualquier función $f : U \rightarrow \mathbb{R}^N$, de clase al menos \mathcal{C}^1 , con $f(U) \subseteq X$ y con $f(a) = p$, se tiene

$$(df)_a(\mathbb{R}^s) \subseteq \mathbb{V}. \quad (92)$$

El conjunto $U_0 = f^{-1}(W)$ es un abierto de \mathbb{R}^s con $a \in U_0 \subseteq U$ y

$$f(U_0) \subseteq X \cap W,$$

luego $(\sigma \circ f)(U_0) \subseteq \mathbb{R}^n \times \{\mathbf{0}\}$. Escribiendo $g = \sigma \circ (f|_{U_0})$, se tiene $g \equiv (g_1, \dots, g_n, 0, \dots, 0)$. Definimos $h \equiv (g_1, \dots, g_n)$ y llegamos a la identidad

$$f|_{U_0} \equiv \sigma^{-1} \circ (h, \mathbf{0}),$$

a la que aplicamos la regla de la cadena (38), para obtener:

$$(df)_a = (d\sigma^{-1})_{\sigma(p)} \circ (d(h, \mathbf{0}))_a ,$$

de donde

$$(df)_a(\mathbb{R}^s) = (d\sigma^{-1})_{\sigma(p)} \left((d(h, \mathbf{0}))_a(\mathbb{R}^s) \right) \subseteq (d\sigma^{-1})_{\sigma(p)}(\mathbb{R}^n \times \{\mathbf{0}\}) ,$$

o sea $(df)_a(\mathbb{R}^s) \subseteq \mathbb{V}$. □

Como primera consecuencia, veamos que se obtiene el mismo subespacio vectorial \mathbb{V} con cualquier “difeomorfismo aplanador” para la parte de X cercana a p .

Sean, pues, otro abierto $p \in W' \subseteq \mathbb{R}^N$ y otro difeomorfismo $\tau : W' \rightarrow \tau(W')$ tales que

$$\tau(X \cap W') = \tau(W') \cap (\mathbb{R}^n \times \{\mathbf{0}\}) .$$

Aplicando la fórmula (92) a la función $f = \tau^{-1}|_{\tau(W') \cap (\mathbb{R}^n \times \{\mathbf{0}\})}$ y el valor $a = \tau(p)$, deducimos de inmediato

$$(d\tau^{-1})_{\tau(p)}(\mathbb{R}^n \times \{\mathbf{0}\}) \subseteq \mathbb{V} ,$$

que es una inclusión entre dos espacios vectoriales de la misma dimensión n , luego en realidad es una igualdad

$$(d\tau^{-1})_{\tau(p)}(\mathbb{R}^n \times \{\mathbf{0}\}) = \mathbb{V} .$$

Tenemos, pues, un subespacio vectorial $\mathbb{V} \subseteq \mathbb{R}^N$ totalmente determinado por la subvariedad X y el punto $p \in X$. Le cambiamos la notación de \mathbb{V} a $T_p X$ y lo llamamos **espacio tangente a X en el punto p** . Llamamos **vectores tangentes a X en p** a los elementos de este espacio vectorial.

Cada subvariedad X viene con una colección de subespacios vectoriales $T_p X$, uno por cada punto $p \in X$, y se tiene $\dim T_p X = \dim X$ para todo $p \in X$.

La propiedad arriba enunciada puede ahora reescribirse así:

Proposición 124. *Dados cualesquiera abierto $U \subseteq \mathbb{R}^s$, función $f(u_1, \dots, u_s) : U \rightarrow \mathbb{R}^N$ de clase al menos \mathcal{C}^1 y con $f(U) \subseteq X$, y valor $a \in U$, se tiene*

$$(df)_a(\mathbb{R}^s) \subseteq T_{f(a)} X , \tag{93}$$

es decir $f_{u_1}(a), \dots, f_{u_s}(a) \in T_{f(a)} X$.

En particular, dados un intervalo $J \subseteq \mathbb{R}$ y un camino $\alpha(t) : J \rightarrow \mathbb{R}^N$ con $\alpha(J) \subseteq X$, se tiene

$$\alpha'(t) \in T_{\alpha(t)} X \quad , \quad \text{para todo } t \in J .$$

El espacio tangente $T_p X$ contiene las velocidades de todos los caminos contenidos en X cuando pasan por p . Recíprocamente, estas velocidades llenan por completo el espacio tangente porque cada vector

$$T_p X \ni v = (d\sigma^{-1})_{\sigma(p)}(c_1, \dots, c_n, 0, \dots, 0) ,$$

es la velocidad $\alpha'(0)$ del camino $\alpha(t) \equiv \sigma^{-1}(\sigma(p) + t(c_1, \dots, c_n, 0, \dots, 0))$, que está contenido en X y en el instante $t = 0$ pasa por $\alpha(0) = p$.

El subespacio $T_p X$ es “verdaderamente tangente a X ” en el punto p , porque es tangente a todos los caminos en X al pasar por p y no tiene ningún vector que no sea tangente en p a algunos de esos caminos.

Corolario 125. *Sea $n = \dim X$. Si $U \subseteq \mathbb{R}^n$ es un abierto y $\Phi(u_1, \dots, u_n) : U \rightarrow \mathbb{R}^N$ es al menos de clase \mathcal{C}^1 , con $\Phi(U) \subseteq X$, $\Phi(a) = p$ y $(d\Phi)_a$ inyectiva, entonces $\{\Phi_{u_1}(a), \dots, \Phi_{u_n}(a)\}$ es una base para $T_p X$.*

4.3.2 Espacio normal

Definición 126. Dada una subvariedad $X \subseteq \mathbb{R}^N$ y dado un punto $p \in X$, el **espacio normal a X en el punto p** es el complemento ortogonal del espacio tangente, o sea el conjunto

$$\{ w \in \mathbb{R}^N : w \cdot v = 0 \text{ para todo } v \in T_p X \}.$$

Decimos que un vector w es **normal a X en p** si pertenece al espacio normal.

Si $n = \dim X$ y $N = n + k$, entonces para cada punto de X el espacio normal es un subespacio vectorial de dimensión k .

La dimensión de cada espacio normal coincide con la codimensión de X en \mathbb{R}^N .

Proposición 127. Dados un abierto $W \subseteq \mathbb{R}^N$ y una función diferenciable $f : W \rightarrow \mathbb{R}$ tal que $f|_{X \cap W}$ es constante, para cada $p \in X \cap W$ el gradiente $\nabla f(p)$ es normal a X en p .

Demostración. Para cada vector $v \in T_p X$ existen caminos $\alpha(t) \subset X$ con $\alpha(0) = p$ y $\alpha'(0) = v$. Derivando la identidad

$$f \circ \alpha(t) \equiv \text{constante},$$

se obtiene

$$\nabla f(\alpha(0)) \cdot \alpha'(0) = 0,$$

es decir $\nabla f(p) \cdot v = 0$. Como v era cualquier vector tangente a X en p , queda probado que $\nabla f(p)$ es un vector normal. \square

Ahora nos situamos en el contexto del corolario 123. Tenemos un abierto $W \subseteq \mathbb{R}^N$ y una subvariedad X definida implícitamente por un sistema de ecuaciones

$$X = \{ x \in W : F_1(x) = b_1, \dots, F_k(x) = b_k \},$$

siendo b_1, \dots, b_k constantes y $F_1, \dots, F_k : W \rightarrow \mathbb{R}$ funciones al menos de clase \mathcal{C}^1 con los gradientes $\nabla F_1(p), \dots, \nabla F_k(p)$ linealmente independientes para todo punto $p \in X$. Estos gradientes están todos en el espacio normal a X en p , el cual tiene dimensión k porque ésta es la codimensión de la variedad implícita X . Luego para cada $p \in X$ la lista $\{ \nabla F_1(p), \dots, \nabla F_k(p) \}$ es una base del espacio normal a X en p .

Veamos el caso $k = 1$ con un poco más de detalle. Tenemos un abierto $W \subseteq \mathbb{R}^N$ y una función escalar $F : W \rightarrow \mathbb{R}$, al menos de clase \mathcal{C}^1 . Definimos otro abierto:

$$V = \{ x \in W : \nabla F(x) \neq \mathbf{0} \} \subseteq W.$$

Entonces para todo $c \in \mathbb{R}$ el conjunto de nivel

$$X_c = \{ x \in V : F(x) = c \},$$

es o vacío o una subvariedad de dimensión $N - 1$. El campo de vectores ∇F es normal en cada punto $p \in V$ al conjunto de nivel que pasa por p . Como este campo es continuo y no se anula en V , ahí podemos definir el campo $\nabla F / \|\nabla F\|$ que proporciona una **normal unitaria continua** a lo largo de cada conjunto de nivel en V .

4.4 Máximos y mínimos condicionados

Dados un abierto $W \subseteq \mathbb{R}^n$ y una función escalar $f : W \rightarrow \mathbb{R}$, buscamos el valor máximo o mínimo de f en un subconjunto $R \subset W$ definido por ciertas *condiciones*, que pueden ser ecuaciones o desigualdades.

4.4.1 Multiplicadores de Lagrange

Tenemos una subvariedad $X \subset \mathbb{R}^N$, un abierto $W \subseteq \mathbb{R}^N$ que contiene a X y una función diferenciable escalar $f : W \rightarrow \mathbb{R}$. Consideramos la restricción $f|_X : X \rightarrow \mathbb{R}$.

Proposición 128. *Una condición necesaria (no suficiente) para que un punto $p \in X$ sea un máximo local o mínimo local de $f|_X$ es que el gradiente $\nabla f(p)$ sea normal a X en p .*

Demostración. Supongamos que $\nabla f(p)$ no es normal a X en p y veamos que entonces p no es ni máximo local ni mínimo local para $f|_X$.

Elegimos un vector $v \in T_p X$ con $v \cdot \nabla f(p) \neq 0$. Elegimos un intervalo abierto $J \ni 0$ y un camino diferenciable $\alpha(t) : J \rightarrow \mathbb{R}^N$ con $\alpha(t) \subset X$, $\alpha(0) = p$ y $\alpha'(0) = v$. La función $\varphi(t) \equiv f \circ \alpha(t)$ tiene la siguiente derivada en $t = 0$:

$$\left. \frac{d}{dt} \right|_{t=0} \varphi(t) = \alpha'(0) \cdot \nabla f(\alpha(0)) = v \cdot \nabla f(p) \neq 0,$$

luego hay valores $t', t'' \in J$, arbitrariamente cercanos a 0, tales que

$$\varphi(t') < \varphi(0) \quad , \quad \varphi(t'') > \varphi(0) .$$

Entonces $p' = \alpha(t')$ y $p'' = \alpha(t'')$ son puntos de X , arbitrariamente cercanos a p , con

$$f(p') < f(p) < f(p'') .$$

□

Sea $X \subset \mathbb{R}^N$ una subvariedad construida según el corolario 123, por un sistema de ecuaciones $(F_1, \dots, F_k)(x) = b$. Como $\{\nabla F_1(p), \dots, \nabla F_k(p)\}$ es una base del espacio normal, el gradiente $\nabla f(p)$ es normal a X en p si y sólo si existen números $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ tales que

$$(\nabla f)(p) = \lambda_1 \nabla F_1(p) + \dots + \lambda_k \nabla F_k(p) . \quad (94)$$

Definición 129. Los números $\lambda_1, \dots, \lambda_k$, satisfaciendo (94), se llaman **multiplicadores de Lagrange en el punto p** . Están totalmente determinados por f , las F_j y el punto p .

La proposición 128 implica que todo punto de máximo local o de mínimo local para $f|_X$ da lugar a unos multiplicadores de Lagrange.

Veamos, con un ejemplo, que los multiplicadores de Lagrange pueden variar de un punto a otro. Consideramos $X = \{x^2 + y^2 + z^2 = 1\}$ y $f(x, y, z) = x^2 + y$. Queremos hallar los valores máximo y mínimo de $f|_X$. A menudo, eso se expresa diciendo que vamos a hallar el máximo o mínimo de $x^2 + y$ sujeto a la condición $x^2 + y^2 + z^2 = 1$.

Calculamos:

$$\nabla f = (2x, 1, 0) \quad , \quad \nabla (x^2 + y^2 + z^2) = (2x, 2y, 2z) .$$

En cualquier punto $p \in X$, es $T_p X = \left\{ \nabla (x^2 + y^2 + z^2)(p) \right\}^\perp = \{2p\}^\perp = \{p\}^\perp$, lo que conduce al siguiente planteamiento del problema de multiplicadores de Lagrange:

$$\left. \begin{aligned} x^2 + y^2 + z^2 &= 1 \\ 2x &= \lambda x \\ 1 &= \lambda y \\ 0 &= \lambda z \end{aligned} \right\}$$

Las soluciones son:

$$\begin{aligned} p^\pm &= \left(\pm \frac{\sqrt{3}}{2}, \frac{1}{2}, 0 \right) , \quad \lambda = 2 , \\ q^+ &= (0, 1, 0) , \quad \lambda = 1 , \\ q^- &= (0, -1, 0) , \quad \lambda = -1 . \end{aligned}$$

Los puntos p^+ y p^- comparten el multiplicador de Lagrange, pero los otros dos puntos lo tienen diferente.

Por lo explicado más arriba, sabemos que los puntos donde $f|_X$ alcanza sus valores máximo o mínimo tienen que estar entre los cuatro puntos p^\pm, q^+, q^- . Evaluamos:

$$f(p^\pm) = \frac{5}{4} , \quad f(q^+) = 1 , \quad f(q^-) = -1 ,$$

y concluimos que el valor máximo de $f|_X$ es $5/4$, que se alcanza en p^\pm , y el valor mínimo es -1 que se alcanza en q^- .

Ahora vamos a hallar un máximo y un mínimo en una *región* definida por una *desigualdad*. Buscamos el máximo y el mínimo de la función $f(x, y) = (x - 1)^2 + 4y^2$, así como los puntos donde se alcanzan, en la región $R = \{x^2 + y^2 \leq 3\}$. Esos puntos existen porque R es compacta.

Separamos la región R en su interior, que es la bola abierta $U = \{x^2 + y^2 < 3\}$, y su **frontera** (la curva que separa la región del resto del plano), que es la circunferencia $X = \{x^2 + y^2 = 3\}$.

Para resolver el problema buscamos puntos de máximo y mínimo en U (si es que los hay), puntos de máximo y mínimo en X (que seguro que los hay, porque X es compacta) y comparamos valores de f en los puntos resultantes.

Un máximo o un mínimo de f en el interior, si existe, debe ser un punto de anulación del gradiente $\nabla f = (2x - 2, 8y)$. El único punto cumpliendo eso es el $p_0 = (1, 0)$, que guardamos para luego.

Un punto de máximo o mínimo de $f|_X$ tiene que satisfacer:

$$x^2 + y^2 = 3 , \quad (2x - 2, 8y) = \lambda(2x, 2y) .$$

Puesto como sistema de ecuaciones escalares:

$$\left. \begin{aligned} x^2 + y^2 &= 3 \\ x - 1 &= \lambda x \\ 4y &= \lambda y \end{aligned} \right\}$$

Las soluciones con $y = 0$ son los puntos $(\pm\sqrt{3}, 0)$. Las soluciones con $y \neq 0$ tienen forzosamente $\lambda = 4$ y $x = -1/3$, luego son los puntos $(-1/3, \pm\sqrt{26}/3)$. Evaluamos:

$$f(p_0) = 0,$$

$$\text{en } p_1 = (\sqrt{3}, 0) \text{ es } \lambda = 1 - (1/\sqrt{3}) \text{ y } f(p_1) = (\sqrt{3} - 1)^2 \approx 0'53,$$

$$\text{en } p_2 = (-\sqrt{3}, 0) \text{ es } \lambda = 1 + (1/\sqrt{3}) \text{ y } f(p_2) = (\sqrt{3} + 1)^2 \approx 7'46,$$

$$\text{en } p^\pm = (-1/3, \pm\sqrt{26}/3) \text{ es } \lambda = 4 \text{ y } f(p^\pm) = 40/3 \approx 13'33.$$

De nuevo, el multiplicador de Lagrange varía de un punto a otro. El punto p_0 no tiene multiplicador de Lagrange, porque no está sobre la circunferencia sino en el interior.

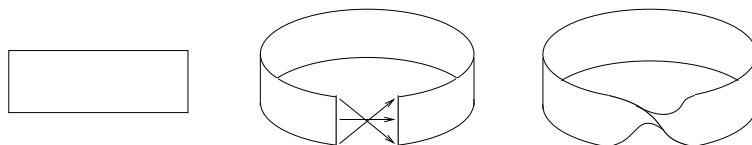
Vemos que 0 es el mínimo de f en toda la región y se alcanza solamente en el punto $p_0 = (1, 0)$, que está en el interior. El máximo de f en la región es $40/3$ y se alcanza en los puntos p^\pm , que están en la frontera. El mínimo de f en la frontera es $(\sqrt{3} - 1)^2$, mayor que el mínimo en el interior, y sólo se alcanza en el punto $p_1 = (\sqrt{3}, 0)$.

4.5 Un ejemplo de subvariedad no implícita

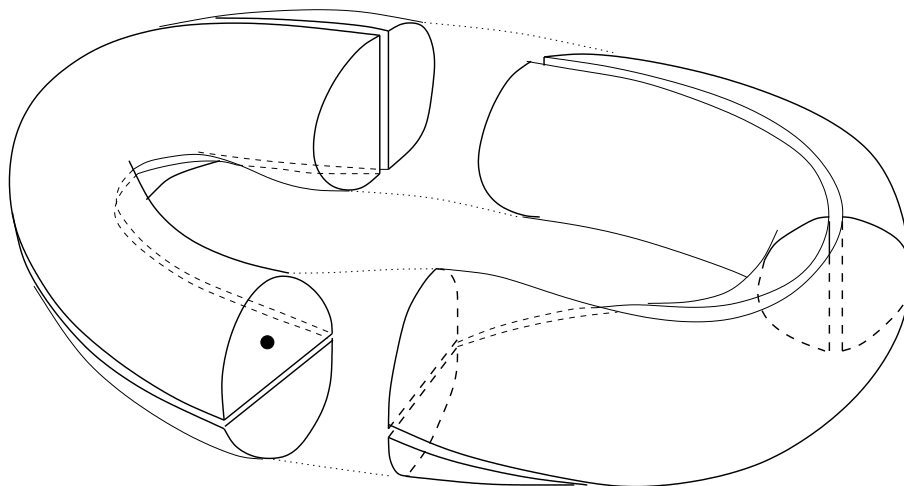
El corolario 123 nos proporciona una inmensa cantidad de ejemplos de subvariedades. Cabe preguntarse si toda subvariedad de \mathbb{R}^N puede obtenerse por ese procedimiento.

Proposición 130. *Para $N \geq 3$ hay subvariedades de \mathbb{R}^N que no se pueden obtener de la manera descrita en el corolario 123.*

El ejemplo estándar es la **banda de Möbius**, que es la subvariedad $X \subset \mathbb{R}^3$ de dimensión 2 que se obtiene doblando en el espacio un rectángulo flexible y pegando sus extremos, habiendo torcido primero uno de esos extremos por media vuelta.



Hay un abierto $W \subset \mathbb{R}^3$ en forma de anillo sólido, tal que la banda está adentro con su borde pegado a la “cáscara” que separa el anillo sólido del resto del espacio. Al quitar la banda del anillo sólido, éste “se corta por la mitad” y la capa protectora $W \setminus X$ que resulta da dos vueltas al anillo antes de cerrarse.



Si empezamos en el “túnel de arriba” (marcado en la figura con un punto gordo) y caminamos por $W \setminus X$, al cabo de la primera vuelta aparecemos en el “túnel de abajo” sin haber tocado la banda en ningún momento. Una hipotética función $f : W \rightarrow \mathbb{R}$ que se anulara en los puntos de X , con gradiente no nulo en esos puntos, tendría que ser positiva a un lado de la banda y negativa al otro, lo que nos impediría pasar del “túnel de arriba al de abajo” sin atravesar la banda. Por lo tanto, esa función f no existe.

Sí que existen tales funciones en cada una de las dos mitades que se insinúan en la figura (la mitad izquierda y la mitad derecha), pero es imposible pegarlas en los dos “empalmes” a la vez: si las funciones coinciden en un empalme, sus signos se contradicen en el otro empalme.

5 Campos

Dada una región $R \subseteq \mathbb{R}^n$, el tipo más sencillo de campo que podemos considerar en R es un **campo de escalares** o **campo escalar**, que es una función $R \rightarrow \mathbb{R}$.

El primer tipo de campo que vamos a estudiar son los **campos vectoriales**, que son las funciones $R \rightarrow \mathbb{R}^n$. Después, en apartados sucesivos, consideraremos otros tipos de campo.

5.1 Primer tipo de campos: vectoriales

Utilizaremos dos nombres y dos notaciones para estos objetos. Cuando dan lugar a flujos los llamamos *campos de velocidades* y normalmente los designamos por la letra V . Cuando los tratamos como funciones a las que vamos a derivar, los llamamos *campos de vectores* y normalmente los designamos por las letras F o G .

5.1.1 Flujos

Definición 131. *Dados un abierto no vacío $U \subseteq \mathbb{R}^n$ y una función $V(x) : U \rightarrow \mathbb{R}^n$, un camino integral de V viene dado por un intervalo $J \subseteq \mathbb{R}$ y una función $\alpha(t) : J \rightarrow U$ tal que*

$$\alpha'(t) = V(\alpha(t)) \quad \text{para todo } t \in J. \quad (95)$$

En cada instante t , el camino pasa por el correspondiente punto $\alpha(t)$ con la velocidad $V(\alpha(t))$ que V prescribe en ese punto. Dicho de otra manera, el punto $\alpha(t)$ se mueve dentro de U “obedeciendo en cada instante” a las velocidades que el campo V le va indicando.

Existen n funciones escalares $V_1, \dots, V_n : U \rightarrow \mathbb{R}$ con $V(x) \equiv (V_1(x), \dots, V_n(x))$. A su vez, un camino $\alpha(t) : J \rightarrow U$ viene dado por n funciones $x_1(t), \dots, x_n(t) : J \rightarrow \mathbb{R}$ tales que $\alpha(t) \equiv (x_1(t), \dots, x_n(t))$. Dicho esto, la ecuación vectorial (95) equivale al siguiente sistema de ecuaciones escalares

$$\left. \begin{aligned} x_1'(t) &= V_1(x_1(t), \dots, x_n(t)) \\ &\vdots \\ x_n'(t) &= V_n(x_1(t), \dots, x_n(t)) \end{aligned} \right\}$$

que es un sistema de **ecuaciones diferenciales** con funciones incógnitas las $x_j(t)$. Se demuestra que, si $V(x)$ es al menos de clase \mathcal{C}^1 , entonces para cada punto $x_0 \in U$ existe un único camino integral $\alpha(t) : J \rightarrow U$ satisfaciendo el **dato inicial** $\alpha(0) = x_0$ y con J lo más grande posible.

Definiciones 132. *Las trayectorias de V son los conjuntos imagen $\alpha(J)$ de los caminos integrales, con J lo más grande posible. Se llama **retrato de fase** al dibujo de U con las trayectorias.*

Las trayectorias forman una **partición** de U : dadas dos de ellas, o son la misma o son disjuntas.

Definiciones 133. *Dado un campo de velocidades $V : U \rightarrow \mathbb{R}^n$, al menos de clase \mathcal{C}^1 , las transformaciones de flujo de V son unas funciones $x \mapsto \varphi_t(x)$ definidas de la manera siguiente:*

Para cada $x \in U$ tomamos el (único) camino integral
 $\alpha(t)$ de V con $\alpha(0) = x$ y hacemos $\varphi_t(x) \stackrel{\text{def}}{=} \alpha(t)$.

Intuitivamente, dejamos que el punto x viaje por U , obedeciendo en cada instante a las velocidades prescritas por V , y vemos dónde acaba cuando han transcurrido t unidades de tiempo.

*Llamamos **flujo de V** al conjunto $\{\varphi_t\}$ de las transformaciones de flujo de V .*

Para cualquier campo de velocidades, se tiene $\varphi_0 = \text{id}_U$.

Dada una región $R \subseteq U$, tenemos las regiones transformadas $\varphi_t(R)$ para distintos valores de t . Si consideramos R como una masa de fluido en el instante $t = 0$, cada molécula de este fluido está viajando con las velocidades prescritas por V y $\varphi_t(R)$ es la posición de esa misma masa de fluido después de t unidades de tiempo.

Se demuestra que, si $V(x)$ es \mathcal{C}^ℓ con $\ell \geq 1$, entonces las transformaciones φ_t son difeomorfismos de clase \mathcal{C}^ℓ con determinante jacobiano positivo.

Ejemplo 1. Sea c una constante. En la recta \mathbb{R} consideramos el campo $V(x) = cx$, visto como campo de velocidades. Los caminos integrales son las funciones escalares $x(t)$ tales que

$$x'(t) = cx(t). \quad (96)$$

Dada una solución cualquiera $x(t)$ de esta ecuación diferencial, calculamos:

$$\frac{d}{dt}(e^{-ct}x(t)) = -ce^{-ct}x(t) + e^{-ct}x'(t) = -ce^{-ct}x(t) + e^{-ct}cx(t) \equiv 0,$$

luego $e^{-ct}x(t)$ es una función constante, de donde $x(t) \equiv \text{cte}e^{ct}$. Pero además

$$x(0) = \text{cte}e^{c0} = \text{cte} \cdot 1 = \text{cte},$$

y la **solución general** de la ecuación (96) queda escrita de la manera siguiente:

$$x(t) = x(0)e^{ct},$$

donde vemos claramente que hay un único camino integral para cada dato inicial $x(0)$.

Ejemplo 2. En el plano consideramos el campo de velocidades $V(x, y) \equiv (x, 2y)$. Un camino integral es una función $\alpha(t) \equiv (x(t), y(t))$ tal que

$$\left. \begin{aligned} x'(t) &= x(t) \\ y'(t) &= 2y(t) \end{aligned} \right\}$$

Por lo que acabamos de ver sobre las soluciones a la ecuación (96), deducimos:

$$x(t) = x(0)e^t, \quad y(t) = y(0)e^{2t}.$$

Volvemos a comprobar que cada dato inicial $\alpha(0) = (x(0), y(0))$ lo realiza un único camino integral $\alpha(t)$. También deducimos que cada punto inicial (x_0, y_0) es llevado por la transformación de flujo φ_t al punto

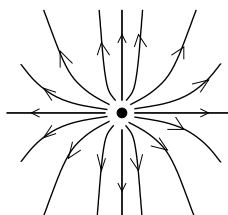
$$\varphi_t(x_0, y_0) = (x_0e^t, y_0e^{2t}),$$

y, como los valores x_0, y_0 son cualesquiera, tenemos la siguiente fórmula para las transformaciones de flujo:

$$\varphi_t(x, y) \equiv (xe^t, ye^{2t}),$$

donde se ve que las φ_t son difeomorfismos \mathcal{C}^∞ con jacobiano positivo, y que φ_0 es la identidad.

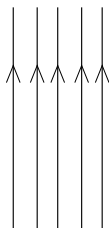
Veamos cómo son las trayectorias de este V . El camino constante $\alpha(t) \equiv (0, 0)$ es un camino integral de V , por lo tanto el conjunto $\{(0, 0)\}$, formado por un único punto, es una trayectoria. Hay otras cuatro trayectorias que son los semiejes coordenados quitado el origen; por ejemplo, la imagen del camino $\alpha(t) \equiv (e^t, 0)$ es el semieje x positivo. En las demás trayectorias nunca se anulan ni la abscisa ni la ordenada, y es evidente que la cantidad y/x^2 permanece constante a lo largo de cada una de ellas; luego cada una de esas trayectorias es una de las dos mitades que resultan de alguna parábola $\{y = \text{cte}x^2\}$ al quitarle el origen.



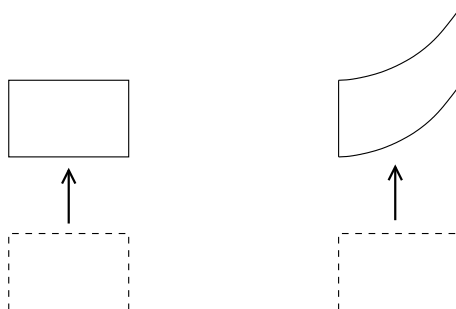
Ejemplo 3. Los campos de velocidades

$$V(x, y) = (0, 1) \quad , \quad W(x, y) = (0, e^x) \quad ,$$

tienen el mismo retrato de fase: líneas verticales que llenan el plano



pero tienen transformaciones de flujo que son diferentes. Las de V son traslaciones verticales $\varphi_t(x, y) = (x, y + t)$. Las de W no son traslaciones: $\psi_t(x, y) = (x, y + te^x)$



Las transformaciones de flujo no se determinan a partir solamente del retrato de fase. Para conocerlas es necesario calcular.

5.1.2 Divergencia

Definición 134. Dados un abierto $U \subseteq \mathbb{R}^n$ y un campo de velocidades $V \equiv (V_1, \dots, V_n)$, al menos de clase \mathcal{C}^1 en U , llamamos **divergencia de V** a la función escalar $\operatorname{div} V : U \rightarrow \mathbb{R}$ definida de la siguiente manera

$$\operatorname{div} V = \operatorname{traza}(DV) = \frac{\partial V_1}{\partial x_1} + \dots + \frac{\partial V_n}{\partial x_n} .$$

Para entender el significado de esta función escalar necesitamos demostrar primero un lema.

Lema 135. Sea $A(t)$ una función diferenciable de la variable t , con valores en matrices $n \times n$.

Si $A(t_0) = I_n$, entonces $\left. \frac{d}{dt} \right|_{t=t_0} \det A(t) = \operatorname{traza}(A'(t_0))$.

En general, se tiene $\left. \frac{d}{dt} \right|_{t=t_0} \det A(t) = \det A(t_0) \cdot \operatorname{traza}[A'(t_0) A(t_0)^{-1}]$.

Demostración. Escribiendo $A(t) = (a_{ij}(t))$, si $A(t_0) = I_n$ entonces las únicas entradas $a_{ij}(t_0)$ no nulas son las $a_{ii}(t_0) = 1$. Se deduce que, de todos los términos del desarrollo de la derivada

$$\left. \frac{d}{dt} \right|_{t=t_0} \det A = \left. \frac{d}{dt} \right|_{t=t_0} \sum_{\sigma \in S_n} (\operatorname{sig} \sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} ,$$

los únicos que pueden ser no nulos son los siguientes:

$$a'_{11}(t_0) \cdot 1 \cdots 1 + 1 \cdot a'_{22}(t_0) \cdot 1 \cdots 1 + \cdots + 1 \cdots 1 \cdot a'_{nn}(t_0) = \operatorname{traza} A'(t_0) .$$

En el caso general, razonamos así:

$$\left. \frac{d}{dt} \right|_{t=t_0} \det A = \left. \frac{d}{dt} \right|_{t=t_0} \det (A(t) A(t_0)^{-1} A(t_0)) = \det A(t_0) \cdot \left. \frac{d}{dt} \right|_{t=t_0} \det (A(t) A(t_0)^{-1}) ,$$

pero la función matriz $B(t) = A(t) A(t_0)^{-1}$ satisface $B(t_0) = I_n$, luego se le aplica el caso particular, o sea

$$\left. \frac{d}{dt} \right|_{t=t_0} \det A = \det A(t_0) \cdot \left. \frac{d}{dt} \right|_{t=t_0} \det B(t) = \det A(t_0) \cdot \text{traza } B'(t_0),$$

y como $B'(t_0) = A'(t_0) A(t_0)^{-1}$ la prueba está completa. \square

Sea ahora V un campo de velocidades, al menos \mathcal{C}^1 , definido en un abierto $U \subseteq \mathbb{R}^n$, y sean φ_t sus transformaciones de flujo.

Dada una región $R \subset U$, la **medida** de R es la integral múltiple $\int_R 1 \cdot dx_1 \cdots dx_n$. Para $n = 1$ la medida es la longitud de R , para $n = 2$ es el área y para $n = 3$ es el volumen; pero tiene sentido en cualquier dimensión, no sólo 1, 2 y 3.

Fijada R , consideramos las regiones transformadas $\varphi_t(R)$ y la función $m(t) = \text{medida}(\varphi_t(R))$. Como las φ_t tienen determinante jacobiano positivo, la fórmula de cambio de variables en integrales múltiples nos proporciona la fórmula

$$m(t) = \int_R \det(D\varphi_t)_x dx_1 \cdots dx_n. \quad (97)$$

Por otra parte, haciendo uso del lema 135,

$$\frac{\partial}{\partial t} \det(D\varphi_t)_x = \det(D\varphi_t)_x \cdot \text{traza} \left[\left(\frac{\partial}{\partial t} (D\varphi_t) \right)_x (D\varphi_t)_x^{-1} \right]. \quad (98)$$

Además:

$$\frac{\partial}{\partial t} (D\varphi_t)_x = D \left(\frac{\partial \varphi_t}{\partial t} \right)_x = D(V \circ \varphi_t(x)) = (DV)_{\varphi_t(x)} (D\varphi_t)_x,$$

de donde:

$$\left(\frac{\partial}{\partial t} (D\varphi_t) \right)_x (D\varphi_t)_x^{-1} = (DV)_{\varphi_t(x)} (D\varphi_t)_x (D\varphi_t)_x^{-1} = (DV)_{\varphi_t(x)}. \quad (99)$$

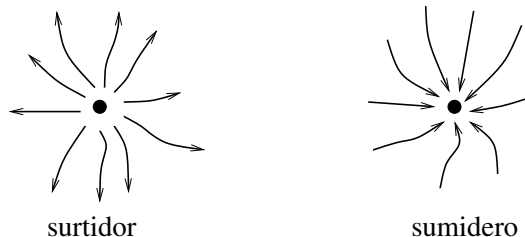
Juntando (97), (98) y (99), obtenemos:

$$m'(t) = \int_R \det(D\varphi_t)_x \cdot \text{traza} (DV)_{\varphi_t(x)} dx_1 \cdots dx_n = \int_{\varphi_t(R)} \text{traza} (DV) dx_1 \cdots dx_n,$$

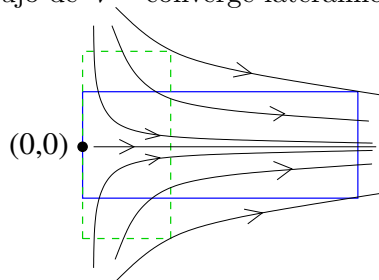
en definitiva:

$$\boxed{\frac{d}{dt} \text{medida}(\varphi_t(R)) = \int_{\varphi_t(R)} (\text{div } V) dx_1 \cdots dx_n} \quad (100)$$

Vamos a explicar por qué se le puso el nombre “divergencia” a la traza de DV . Sea p un **surtidor de V** , es decir un punto donde $V(p) = \mathbf{0}$ y cerca del cual los caminos integrales se alejan radialmente de p (o sea, *divergen*), entonces $\text{div } V$ es *positiva en promedio* en las cercanías de p : cada bola de radio pequeño centrada en p aumenta de medida al seguir el flujo de V , lo que junto con (100) implica que la integral de $\text{div } V$ es positiva en dichas bolas. Del mismo modo $\text{div } V$ es negativa en promedio cerca de los **sumideros**.



Esto puede dar lugar a confusión, ya que la divergencia existe en todo punto y no sólo en los surtidores y sumideros. Más aún, puede ser positiva en un lugar donde los caminos integrales se juntan en vez de separarse: por ejemplo, para el campo de velocidades $V(x, y) \equiv (3x, -y)$, en un tiempo positivo t el rectángulo $[0, 1] \times [-1, 1]$ es llevado por la transformación de flujo φ_t al rectángulo $[0, e^{3t}] \times [-e^{-t}, e^{-t}]$, de menor altura pero bastante más ancho (por lo que tiene más área). Habría que decir que el flujo de V “converge lateralmente pero diverge hacia adelante”.



5.1.3 Flujos lineales

En este apartado consideramos el caso particular de un campo de velocidades **lineal**. Empezamos considerando $V(x) \equiv Ax$, donde A es una matriz $n \times n$ constante.

Para un campo así los caminos integrales forman un espacio vectorial (de funciones). Concretamente, si $\alpha(t)$ y $\beta(t)$ cumplen

$$\alpha'(t) = A\alpha(t) \quad \text{y} \quad \beta'(t) = A\beta(t),$$

entonces para cualesquiera constantes c_1, c_2 se tiene

$$(c_1\alpha + c_2\beta)'(t) = A(c_1\alpha(t) + c_2\beta(t)),$$

y $c_1\alpha(t) + c_2\beta(t)$ también es un camino integral. Dados datos iniciales $v = \alpha(0)$ y $w = \beta(0)$, el camino integral $\gamma(t)$ con dato inicial $\gamma(0) = c_1v + c_2w$ es

$$\gamma(t) \equiv c_1\alpha(t) + c_2\beta(t),$$

de donde se deduce

$$\varphi_t(c_1v + c_2w) = c_1\varphi_t(v) + c_2\varphi_t(w) \quad \text{para cualesquiera } t, v, w, c_1, c_2,$$

lo que nos dice que las transformaciones de flujo φ_t son lineales. Como son difeomorfismos con jacobiano positivo, en realidad son aplicaciones lineales invertibles con determinante positivo para todo t .

Hemos comprobado esto en el ejemplo 2 del apartado 5.1.1, cuyas transformaciones de flujo son transformaciones lineales con determinante $e^{3t} > 0$:

$$\varphi_t \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} e^t & \\ & e^{2t} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Si A es simétrica entonces hay una base ortonormal $\{v_1, \dots, v_n\}$ de \mathbb{R}^n y números reales $\lambda_1, \dots, \lambda_n$ tales que $Av_j = \lambda_j v_j$ para $j = 1, \dots, n$. Si ahora escribimos el camino general en términos de esa base

$$\alpha(t) \equiv y_1(t)v_1 + \dots + y_n(t)v_n,$$

entonces la ecuación vectorial $\alpha'(t) = A\alpha(t)$ equivale al sistema

$$y_j'(t) = \lambda_j y_j(t), \quad j = 1, \dots, n,$$

y deducimos que las transformaciones de flujo son transformaciones lineales

$$\varphi_t(y_1 v_1 + \dots + y_n v_n) = (e^{\lambda_1 t} y_1) v_1 + \dots + (e^{\lambda_n t} y_n) v_n,$$

con matriz diagonal $\begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix}$ en esa base ortonormal.

Todavía más interesante es el caso en que la matriz A es antisimétrica, es decir $A^t = -A$. En tal caso, dados dos caminos integrales cualesquiera $\alpha(t), \beta(t)$ calculamos:

$$\begin{aligned} \frac{d}{dt} (\alpha(t) \cdot \beta(t)) &= \alpha'(t) \cdot \beta(t) + \alpha(t) \cdot \beta'(t) = (A \alpha(t)) \cdot \beta(t) + \alpha(t) \cdot (A \beta(t)) = \\ &= \alpha(t)^t A^t \beta(t) + \alpha(t)^t A \beta(t) = \alpha(t)^t (A^t + A) \beta(t) = \alpha(t)^t 0_{n \times n} \beta(t) \equiv 0, \end{aligned}$$

de donde:

$$\varphi_t(v) \cdot \varphi_t(w) = v \cdot w \quad \text{para cualesquiera } t, v, w,$$

es decir que cada φ_t es una transformación lineal que conserva el producto escalar. Esto quiere decir que las φ_t son transformaciones **ortogonales**. Como además tienen determinante positivo, en realidad tenemos $\det \varphi_t = 1$ para todo t , es decir

$$A^t = -A \implies \varphi_t \in \text{SO}(n) \quad \text{para todo } t.$$

Si A es $n \times n$ y antisimétrica, las transformaciones de flujo del campo $V(x) \equiv Ax$ son **rotaciones de \mathbb{R}^n** que fijan el origen.

Inmediatamente se deduce que, fijado un punto $\mathbf{a} \in \mathbb{R}^n$ y dado el campo $V(x) \equiv A(x - \mathbf{a})$, es decir

$$V(\mathbf{a} + h) = Ah \quad \text{para todo } h \in \mathbb{R}^n,$$

con A antisimétrica, las transformaciones de flujo de V son rotaciones de \mathbb{R}^n que fijan \mathbf{a} . Esto es, existen matrices $M_t \in \text{SO}(n)$ tales que $\varphi_t(\mathbf{a} + h) \equiv \mathbf{a} + M_t h$. Además $M_0 = I_n$.

5.1.4 Factorización del flujo y matriz rotacional

En este apartado V es un campo de velocidades definido en un abierto $U \subseteq \mathbb{R}^n$ y al menos de clase \mathcal{C}^2 . Fijamos un punto $\mathbf{a} \in U$ y damos una factorización de las transformaciones de flujo de V en un pequeño entorno de \mathbf{a} y en tiempos cercanos a 0. Para enunciar dicha factorización, establecemos varias notaciones.

Consideramos la matriz $A = (DV)_{\mathbf{a}}$. Hay una única descomposición $A = A_1 + A_2 + A_3$ cuyos sumandos son como sigue:

1. A_1 es simétrica con traza nula.
2. A_2 es antisimétrica.
3. A_3 es una **matriz escalar**: $A_3 = s I_n$, $s \in \mathbb{R}$.

Por ejemplo, para matrices 2×2 la descomposición es

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} (a-d)/2 & (b+c)/2 \\ (b+c)/2 & (d-a)/2 \end{bmatrix} + \begin{bmatrix} 0 & (b-c)/2 \\ (c-b)/2 & 0 \end{bmatrix} + (a+d) \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix},$$

y en general es:

$$A = \left(\frac{1}{2} (A + A^t) - (\text{traza } A) (I_n/n) \right) + \frac{1}{2} (A - A^t) + (\text{traza } A) (I_n/n).$$

Para $j = 1, 2, 3$ consideramos los campos $V_j(\mathbf{a} + h) \equiv A_j h$ y las respectivas transformaciones de flujo φ_t^j . Denotamos por φ_t las transformaciones de flujo de V .

Teorema 136. Dados el camino integral $\alpha(t) = \varphi_t(\mathbf{a})$ y el vector $v(t) = \alpha(t) - \mathbf{a}$, para cualquier permutación i, j, k de $1, 2, 3$ tenemos desarrollos:

$$\varphi_t(\mathbf{a} + h) = \varphi_t^k \circ \varphi_t^j \circ \varphi_t^i(\mathbf{a} + h) + v(t) + |t| O(\|(t, h)\|^2), \quad (101)$$

$$\varphi_t^k \circ \varphi_t^j \circ \varphi_t^i(\mathbf{a} + h) = \mathbf{a} + h + t(A_1 + A_2 + A_3)h + |t| O(\|(t, h)\|^2). \quad (102)$$

(Los restos de estos desarrollos dependen de la permutación elegida).

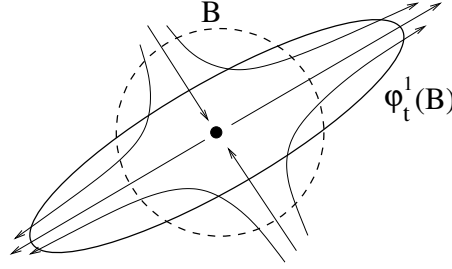
Este teorema se demuestra en documento aparte.

Elegimos, por ejemplo, el desarrollo:

$$\varphi_t(\mathbf{a} + h) = \varphi_t^3 \circ \varphi_t^2 \circ \varphi_t^1(\mathbf{a} + h) + v(t) + R(t, h) \quad , \quad \text{con } R(t, h) = |t| O(\|(t, h)\|^2),$$

y en base a él vamos describiendo cada transformación por turno.

1. Como A_1 es simétrica, existe una base ortonormal $\{v_1, \dots, v_n\}$ de \mathbb{R}^n formada por autovalores de A_1 : $A_1 v_i = \mu_i v_i$ ($i = 1, \dots, n$). Además $\mu_1 + \dots + \mu_n = 0$. En términos de esta base, tenemos $\varphi_t^1(\mathbf{a} + y_1 v_1 + \dots + y_n v_n) = \mathbf{a} + (e^{\mu_1 t} y_1) v_1 + \dots + (e^{\mu_n t} y_n) v_n$. Empezamos con una bola euclídea $B = B(\mathbf{a}, r)$, con r pequeño. La imagen $\varphi_t^1(B)$ es un *elipsoide* centrado en \mathbf{a} cuyos semiejes principales son paralelos a los v_i y tienen longitudes respectivas $e^{\mu_i t} r$.

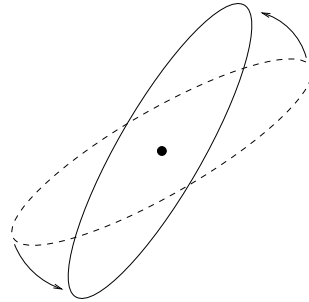


Denotando por m la *medida* n -dimensional (longitud si $n = 1$, área si $n = 2$, volumen si $n = 3$, etc) tenemos:

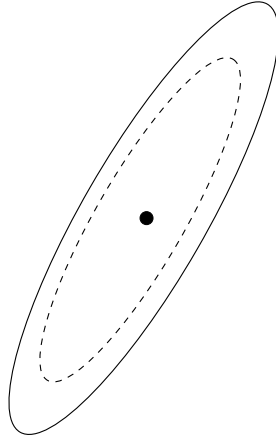
$$m(\varphi_t^1(B)) = e^{(\mu_1 + \dots + \mu_n)t} \cdot m(B) = e^0 \cdot m(B) = m(B),$$

es decir que φ_t^1 es una **deformación pura**, sin cambio ni en el volumen ni en las direcciones de los v_i ; se la puede aplicar a gases y líquidos. Se la puede aplicar a un sólido para t muy pequeño, lo que provocará una *reacción elástica* del sólido.

2. Como A_2 es antisimétrica, la transformación φ_t^2 es una **rotación de \mathbb{R}^n** que deja fijo el punto \mathbf{a} . La imagen $\varphi_t^2 \circ \varphi_t^1(B) = \varphi_t^2(\varphi_t^1(B))$ es un elipsoide centrado en \mathbf{a} , con las mismas longitudes de semiejes principales que $\varphi_t^1(B)$ pero con las direcciones de éstos cambiadas. Es decir que φ_t^2 no efectúa deformación alguna ni cambio de volumen, sólo cambia la posición angular. Se la puede aplicar a sólidos, líquidos o gases.



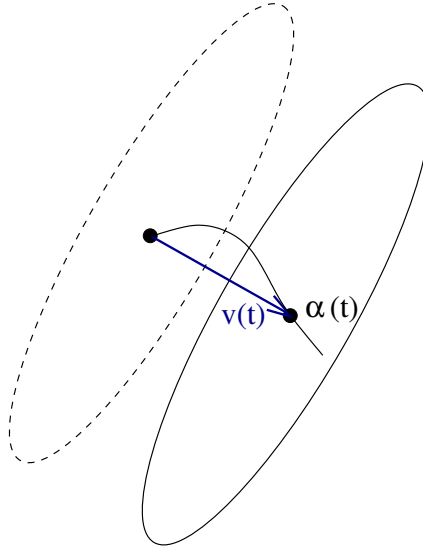
3. Para $A_3 = (\text{traza } A)(1/n)I_n$ las transformaciones $\varphi_t^3(\mathbf{a} + h) = \mathbf{a} + e^{(\text{traza } A)(1/n)t} h$ son **homotecias centradas en \mathbf{a}** . No cambian la forma ni la posición angular, pero sí el tamaño.



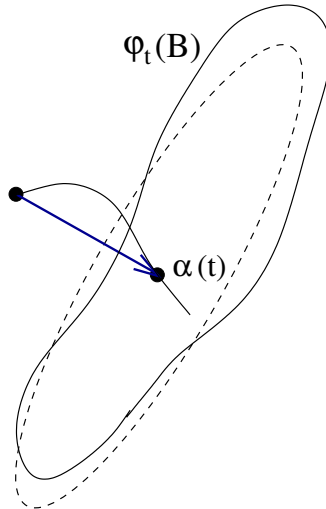
El cambio de la medida es

$$m(\varphi_t^3 \circ \varphi_t^2 \circ \varphi_t^1(B)) = e^{n(\text{traza } A)(1/n)t} m(B) = e^{(\text{traza } A)t} \cdot m(B). \quad (103)$$

4. Al sumar el vector $v(t)$ efectuamos una simple **traslación**. No se modifica la forma, la posición angular ni el tamaño, tan sólo cambiamos la posición del elipsoide.



5. Finalmente sumamos el resto cúbico $R(t, h)$. Cambia muy ligeramente la forma de la región, que ya no será exactamente un elipsoide. Hemos llegado a la imagen $\varphi_t(B)$.



En la factorización (101) del flujo de V , los factores $\varphi_t^1, \varphi_t^2, \varphi_t^3$ son únicos: si cambiásemos alguno de ellos, la alteración que se produce en el término $t(A_1 + A_2 + A_3)h$ de la fórmula (102)

es cuadrática en el vector (t, h) y no puede ser absorbida por el resto de dicha fórmula (que es cúbico en (t, h) y por ende mucho más pequeño) y la compuesta $\varphi_t^3 \circ \varphi_t^2 \circ \varphi_t^1$ cambiaría.

Definición 137. Llamamos a $[DV - (DV)^t]_{\mathbf{a}}$ **matriz rotacional de \mathbf{V} en \mathbf{a}** , porque es el doble de la matriz A_2 causante de las pequeñas rotaciones φ_t^2 que entran en la factorización única del flujo de V dada por la fórmula (101).

Al dejar variar el punto se construye la función matriz $DV - (DV)^t$, que es la **matriz rotacional de \mathbf{V}** .

Por su parte, la matriz A_3 , en la cual hemos puesto la traza de A , es la causante del aumento o disminución de la medida, tal como se ve en la fórmula (103). Esto es perfectamente consistente con la fórmula (100) del apartado 5.1.2, que nos dice que los cambios en medida al fluir son causados por la divergencia, o sea por la traza de DV .

5.1.5 Dimensión 3: vector rotacional

Las matrices antisimétricas $n \times n$ forman un espacio vectorial de dimensión $n(n-1)/2$, número que coincide con n solamente cuando $n = 3$. En tal caso, el **producto vectorial** nos da un isomorfismo lineal que lleva vectores de \mathbb{R}^3 a matrices antisimétricas 3×3 . En concreto:

$$(a_1, a_2, a_3) \times h = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} h, \quad \text{para todo } h \in \mathbb{R}^3. \quad (104)$$

Para el vector $c\mathbf{e}_3 = (0, 0, c)$, la matriz es $A = \begin{bmatrix} 0 & -c & 0 \\ c & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ y las transformaciones de flujo del campo $V(\mathbf{a} + h) \equiv Ah \equiv (c\mathbf{e}_3) \times h$ son

$$\varphi_t(\mathbf{a} + h) = \mathbf{a} + \begin{bmatrix} \cos(ct) & -\sin(ct) & 0 \\ \sin(ct) & \cos(ct) & 0 \\ 0 & 0 & 1 \end{bmatrix} h,$$

es decir rotaciones de eje la recta paralela al vector $c\mathbf{e}_3$ que pasa por \mathbf{a} . La **velocidad angular** con que se efectúan dichas rotaciones es la constante $c = \pm \|c\mathbf{e}_3\|$.

Para el vector general $v \in \mathbb{R}^3$, las transformaciones de flujo del campo $V(\mathbf{a} + h) \equiv v \times h$ son rotaciones cuyo eje es la recta paralela a v que pasa por el punto \mathbf{a} y la velocidad angular con que se efectúan dichas rotaciones es $\pm \|v\|$.

Definición 138. Si $V \equiv (V_1, V_2, V_3)$, es un campo de velocidades en un abierto $U \subseteq \mathbb{R}^3$, definimos el **vector rotacional de \mathbf{V}** como el siguiente campo de vectores:

$$\mathbf{rot} V : U \longrightarrow \mathbb{R}^3, \quad \mathbf{rot} V = (V_{3y} - V_{2z}, -V_{3x} + V_{1z}, V_{2x} - V_{1y}), \quad (105)$$

que cumple la siguiente identidad:

$$\text{para cualesquiera } \mathbf{a} \in U, h \in \mathbb{R}^3, \quad [DV - (DV)^t]_{\mathbf{a}} h = (\mathbf{rot} V)_{\mathbf{a}} \times h, \quad (106)$$

luego $\mathbf{rot} V$ es el vector que representa la matriz rotacional de V según la fórmula (104).

La parte antisimétrica de DV está representada, según (104), por $(1/2)\mathbf{rot} V$.

Por lo que acabamos de explicar, En la factorización (101) del flujo de V las pequeñas rotaciones φ_t^2 tienen por eje la recta que pasa por el punto \mathbf{a} y es paralela al vector $\mathbf{rot} V_{\mathbf{a}}$. Además, la velocidad angular con que se efectúan dichas rotaciones es $\pm(1/2)\|\mathbf{rot} V_{\mathbf{a}}\|$.

Hay un truco para recordar la fórmula (105), que es ver el vector rotacional como el producto vectorial $\nabla \times V$ del vector simbólico $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$ (llamado “nabla”) por V .

5.1.6 Dimensión 2: rotacional escalar

Cada matriz antisimétrica 2×2

$$\begin{bmatrix} 0 & -c \\ c & 0 \end{bmatrix},$$

la podemos representar por el escalar c .

Correspondiendo a eso, si V es un campo de velocidades en un abierto de \mathbb{R}^2 , su matriz rotacional

$$DV - (DV)^t = \begin{bmatrix} 0 & P_y - Q_x \\ Q_x - P_y & 0 \end{bmatrix},$$

la representamos por el escalar $Q_x - P_y$, que algunas personas llaman el **rotacional escalar** del campo. En la factorización (101) del flujo de V , las pequeñas rotaciones φ_t^2 se efectúan con velocidad angular $(1/2)(Q_x - P_y)(\mathbf{a})$, medida en el sentido antihorario.

El truco para acordarse de cuál es el minuendo y cuál el sustraendo es verlo como un “determinante simbólico”

$$Q_x - P_y = \det \begin{bmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \\ P & Q \end{bmatrix}.$$

También es útil relacionarlo con la divergencia: la divergencia de $F = (f, g)$ es el rotacional escalar de $(-g, f)$ que, por su parte, es el resultado de girar F un cuarto de vuelta en sentido antihorario.

5.1.7 Potenciales

Definiciones 139. *Dados un abierto $U \subseteq \mathbb{R}^n$ y un campo de vectores $F : U \rightarrow \mathbb{R}^n$, un **potencial escalar para F** es (si es que existe) una función $h : U \rightarrow \mathbb{R}$ tal que $F \equiv \nabla h$.*

*Dados un abierto $U \subseteq \mathbb{R}^3$ y un campo de vectores $F : U \rightarrow \mathbb{R}^3$, un **potencial vector para F** es (si es que existe) una función vectorial $A : U \rightarrow \mathbb{R}^3$ tal que $F = \mathbf{rot} A$.*

Supondremos que F es C^∞ . Si existe una función escalar h tal que $\nabla h = F$, entonces las derivadas parciales de h , iguales a las entradas de F , son C^∞ y h también lo es. Pero entonces el lema de Schwarz (teorema 93 del apartado 2.3) nos dice que el orden de derivación no importa en las derivadas segundas: $h_{x_j x_i} \equiv h_{x_i x_j}$.

para que un campo de vectores suave F sea un gradiente, es necesario que la matriz DF sea simétrica.

Por supuesto, esta condición necesaria equivale a que la matriz rotacional $DF - (DF)^t$ sea idénticamente nula. Si además estamos en dimensión 3, entonces la condición es $\mathbf{rot} F \equiv \mathbf{0}$.

Por ejemplo, el campo $F \equiv (1, x)$ no es un gradiente en ningún abierto $U \subset \mathbb{R}^2$, porque si en un tal abierto existiera h tal que $h_x = 1$ y $h_y = x$ entonces sería $h \in C^\infty(U)$ y se cumpliría la siguiente identidad en U :

$$0 \equiv \frac{\partial 1}{\partial y} \equiv h_{xy} \equiv h_{yx} \equiv \frac{\partial x}{\partial x} \equiv 1,$$

claramente imposible, excepto para $U = \emptyset$.

Los gradientes son unos campos de vectores sumamente especiales: la mayoría de los campos de vectores no son un gradiente en ningún abierto.

¿Es la condición necesaria $DF = (DF)^t$ una condición suficiente para que haya potencial escalar? Respuesta: sí en algunos abiertos, no en otros abiertos.

Veamos que la condición es suficiente si $U = \mathbb{R}^2$ o si $U = B(p, r)$ es una bola del plano. Tenemos $F \equiv (F_1, F_2) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ de clase C^∞ y con $F_{1y} \equiv F_{2x}$. Buscamos $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ tal que

$$h_x = F_1 \quad , \quad h_y = F_2 \quad .$$

Empezamos resolviendo la primera ecuación. Una solución particular es

$$h^1(x, y) = \int_0^x F_1(x, y) dx \quad ,$$

que ciertamente es C^∞ , y la solución general es $h^1(x, y) + C(y)$, donde $C(y)$ es *cualquier* función suave que sólo depende de la variable y . A esta solución general de la primera ecuación le imponemos la segunda ecuación:

$$\begin{aligned} F_2 &= \frac{\partial}{\partial y} (h^1(x, y) + C(y)) = C'(y) + \int_0^x \frac{\partial}{\partial y} F_1(x, y) dx = \\ &= C'(y) + \int_0^x F_{2x}(x, y) dx = F_2(x, y) + C'(y) - F_2(0, y) \quad . \end{aligned}$$

Al final nos queda la ecuación $C'(y) - F_2(0, y) = 0$, que ya no contiene a la variable x en ningún sitio. Esto nos permite dar la solución $C(y) = \int_0^y F_2(0, y) dy + C_0$, donde C_0 es cualquier constante, y el potencial escalar para F es

$$h = h^1 + C(y) = \int_0^x F_1(x, y) dx + \int_0^y F_2(0, y) dy + C_0 \quad .$$

Dos comentarios importantes:

1. En cada paso ha sido posible integrar desde 0 hasta el valor de la variable (x o y) gracias a que el dominio de F es todo \mathbb{R}^2 o una bola suya.
2. El potencial escalar de F , además de existir, es único *salvo sumarle una constante*.

Volveremos sobre el segundo comentario en la proposición 143 del apartado 5.2.2 y en el corolario 147 del apartado 5.2.6. Vamos ahora con el primer comentario.

Consideramos el abierto $U = \mathbb{R}^2 \setminus \{(0, 0)\}$ y el siguiente campo de vectores

$$F(x, y) = \frac{(y, -x)}{r^2} = \left(\frac{y}{x^2 + y^2}, \frac{-x}{x^2 + y^2} \right) \quad ,$$

que es C^∞ en todo U . Se comprueba de inmediato que cumple la condición $DF = (DF)^t$ y nos preguntamos si existe una función $h(x, y)$ que cumpla las dos ecuaciones siguientes

$$h_x = \frac{y}{x^2 + y^2} \quad , \quad h_y = \frac{-x}{x^2 + y^2} \quad . \quad (107)$$

Para resolver la primera ecuación integráramos $y/(x^2 + y^2)$ desde un valor x_0 hasta x , dejando fija y en un valor arbitrario. La dificultad aparece para el valor particular $y = 0$, para el cual la integral se ve “interrumpida” cuando $x = 0$.

Proposición 140. *No existe ninguna solución $h(x, y)$ a la ecuación*

$$h_x = \frac{y}{x^2 + y^2} \quad ,$$

que sea continua en todo el plano perforado $\mathbb{R}^2 \setminus \{(0, 0)\}$.

Demostración. Supongamos que existe una tal h y deduzcamos una contradicción. Para $\varepsilon \neq 0$ pequeño, positivo o negativo, consideramos el siguiente incremento de h

$$\begin{aligned} h(1, \varepsilon) - h(-1, \varepsilon) &= \int_{-1}^1 \frac{\varepsilon}{x^2 + \varepsilon^2} dx = \int_{x=-1}^{x=1} \frac{1}{\left(\frac{x}{\varepsilon}\right)^2 + 1} d\frac{x}{\varepsilon} = \\ &= \arctan \frac{1}{\varepsilon} - \arctan \frac{-1}{\varepsilon} \rightarrow \begin{cases} \pi & \text{cuando } \varepsilon \searrow 0 \\ -\pi & \text{cuando } \varepsilon \nearrow 0 \end{cases} \end{aligned}$$

Puesto que h es continua en el punto $(-1, 0)$, calculamos:

$$\begin{aligned} 2\pi &= \pi - (-\pi) = \lim_{\varepsilon \searrow 0} (h(1, \varepsilon) - h(-1, \varepsilon)) - \lim_{\varepsilon \nearrow 0} (h(1, \varepsilon) - h(-1, \varepsilon)) = \\ &= \left(\lim_{\varepsilon \searrow 0} h(1, \varepsilon) - \lim_{\varepsilon \nearrow 0} h(1, \varepsilon) \right) + \left(\lim_{\varepsilon \nearrow 0} h(-1, \varepsilon) - \lim_{\varepsilon \searrow 0} h(-1, \varepsilon) \right) = \\ &= \lim_{\varepsilon \searrow 0} h(1, \varepsilon) - \lim_{\varepsilon \nearrow 0} h(1, \varepsilon) + 0, \end{aligned}$$

y h tendría que ser discontinua en el punto $(1, 0)$. \square

En este ejemplo, pues, ni siquiera la primera de las dos ecuaciones que forman (107) puede resolverse en todo $\mathbb{R}^2 \setminus \{(0, 0)\}$.

El plano perforado $\mathbb{R}^2 \setminus \{(0, 0)\}$ es un ejemplo de abierto en el cual la condición $DF = (DF)^t$ no es suficiente para que F tenga un potencial escalar.

Ahora trataremos brevemente la cuestión del potencial vector. Si $A \equiv (A_1, A_2, A_3)$ es un campo suave en un abierto de \mathbb{R}^3 , el lema de Schwarz nos permite hacer el siguiente cálculo:

$$\begin{aligned} \operatorname{div}(\mathbf{rot} A) &= \operatorname{div}(A_{3y} - A_{2z}, -A_{3x} + A_{1z}, A_{2x} - A_{1y}) = \\ &= A_{3yx} - A_{2zx} - A_{3xy} + A_{1zy} + A_{2xz} - A_{1yz} = \\ &= (A_{3yx} - A_{3xy}) + (-A_{2zx} + A_{2xz}) + (A_{1zy} - A_{1yz}) = \\ &= 0 + 0 + 0 = 0, \end{aligned}$$

Para que un campo F pueda tener un potencial vector suave es necesario que F tenga divergencia nula.

Esta condición necesaria resulta ser suficiente si el abierto donde F está definido es convexo, por ejemplo todo \mathbb{R}^3 . No vamos a demostrar eso, pero vamos a hacer un ejemplo.

Es fácil comprobar que el campo $F = (x + e^y, z^2 - e^y, -z + ze^y + 7y)$ tiene divergencia nula. Si buscamos un potencial vector de tipo general $A = (f, g, h)$, las ecuaciones que resultan

$$\left. \begin{aligned} h_y - g_z &= x + e^y \\ -h_x + f_z &= z^2 - e^y \\ g_x - f_y &= -z + ze^y + 7y \end{aligned} \right\}$$

son **acopladas**, es decir que una misma función incógnita aparece en varias ecuaciones. En cambio, si buscamos un potencial vector con alguna entrada nula, por ejemplo $A = (f, g, 0)$, hay dos ecuaciones que se desacoplan

$$\left. \begin{aligned} -g_z &= x + e^y \\ f_z &= z^2 - e^y \\ g_x - f_y &= -z + ze^y + 7y \end{aligned} \right\}$$

De las dos ecuaciones desacopladas, tomamos una solución particular de una y la solución general de la otra. Por ejemplo

$$g = -xz - ze^y, \quad f = \frac{z^3}{3} - ze^y + c(x, y),$$

donde $c(x, y)$ es cualquier función suave que depende sólo de x, y . Las llevamos a la tercera ecuación

$$-z + ze^y + 7y = (-z + 0) - (0 - ze^y + c_y),$$

y al agrupar términos desaparece por completo la variable z , quedando $-c_y = 7y$. Una solución es $c = -(7/2)y^2$. Finalmente $A = (\frac{z^3}{3} - ze^y - \frac{7}{2}y^2, -xz - ze^y, 0)$ es un potencial vector para F .

Conviene advertir que, una vez que F tiene un potencial vector, tiene una gran infinidad de ellos. En efecto, si A_0 es tal que $\text{rot } A_0 = F$ entonces también $\text{rot } (A_0 + \nabla h) = F$ para toda función escalar suave h .

5.2 Segundo tipo de campos: campos de formas lineales

5.2.1 Formas lineales

Recordemos las siguientes nociones del Álgebra Lineal:

1. Una **forma lineal en \mathbb{R}^n** es cualquier función escalar $\ell : \mathbb{R}^n \rightarrow \mathbb{R}$ que sea *lineal*.
2. El **espacio dual $(\mathbb{R}^n)^*$** es el espacio vectorial formado por todas las formas lineales en \mathbb{R}^n .

Veamos que el espacio dual $(\mathbb{R}^n)^*$ también tiene dimensión n . Cada función lineal $\ell : \mathbb{R}^n \rightarrow \mathbb{R}$ tiene, respecto de la base estándar, una matriz que es una fila $1 \times n$

$$\ell \left(\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \right) \equiv [c_1 \ c_2 \ \cdots \ c_n] \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \equiv c_1 a_1 + c_2 a_2 + \cdots + c_n a_n. \quad (108)$$

Para $j = 1, \dots, n$ definamos $\ell_j : \mathbb{R}^n \rightarrow \mathbb{R}$ como la función lineal dada por la matriz fila cuyas entradas son todas cero excepto la j -ésima entrada, que es 1:

$$\ell_j(\vec{a}) \equiv [0 \ \cdots \ 0 \ 1 \ 0 \ \cdots \ 0] \vec{a} \equiv a_j,$$

es decir que $\ell_1, \dots, \ell_n : \mathbb{R}^n \rightarrow \mathbb{R}$ son las **coordenadas estándar en \mathbb{R}^n** , vistas como funciones lineales. Entonces la identidad (108) se puede escribir así:

$$\ell \equiv c_1 \ell_1 + \cdots + c_n \ell_n, \quad (109)$$

y los coeficientes c_j que satisfacen (109) son únicos para cada forma lineal ℓ , pues están obligados a ser las entradas de la matriz fila que representa a ℓ en la base estándar, es decir:

$$c_j = \ell(\mathbf{e}_j) \quad \text{para } j = 1, \dots, n. \quad (110)$$

En definitiva $\{\ell_1, \dots, \ell_n\}$ es una base para $(\mathbb{R}^n)^*$ y los valores $c_j = \ell(\mathbf{e}_j)$ son las coordenadas de ℓ en la base $\{\ell_1, \dots, \ell_n\}$.

5.2.2 Campo diferencial de una función

Definición 141. Sea $U \subseteq \mathbb{R}^n$ un abierto. Sea $f : U \rightarrow \mathbb{R}$ una función escalar, al menos de clase \mathcal{C}^1 . El **campo diferencial de f** se denota df y es la aplicación cuyo valor en cada punto $p \in U$ es la diferencial $(df)_p$.

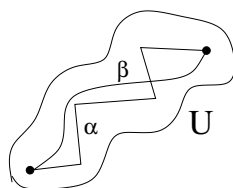
Como cada diferencial es una forma lineal $(df)_p : \mathbb{R}^n \rightarrow \mathbb{R}$, el campo diferencial es una aplicación

$$df : U \longrightarrow (\mathbb{R}^n)^*, \quad p \longmapsto (df)_p,$$

que asigna a cada punto $p \in U$ una forma lineal en \mathbb{R}^n .

Enunciamos ahora un resultado de **aproximación**: no importa cuán “estrecho” sea un abierto U conteniendo a un camino, hay un camino suave con los mismos extremos y que no se sale de U .

Lema 142. Sean $\alpha(t) : [t_0, t_1] \rightarrow \mathbb{R}^n$ un camino continuo. Para todo abierto $U \subseteq \mathbb{R}^n$ conteniendo a α , existe un camino suave $\beta(t) : [t_0, t_1] \rightarrow U$ con los mismos extremos que α , es decir $\beta(t_0) = \alpha(t_0)$ y $\beta(t_1) = \alpha(t_1)$. \square



Proposición 143. Dados un abierto $U \subseteq \mathbb{R}^n$ y una función diferenciable $f : U \rightarrow \mathbb{R}$, si $df \equiv 0$ entonces f es constante en cada componente conexa por caminos de U , valor constante que puede cambiar de una componente conexa por caminos a otra.

Demostración. Sea $V \subseteq U$ una componente conexa por caminos. Dados $p, q \in V$ cualesquiera, existe un camino continuo $\alpha(t) : [0, 1] \rightarrow V$ con $\alpha(0) = p$ y $\alpha(1) = q$.

Aplicamos el lema 142 y tenemos un camino suave $\beta(t) : [0, 1] \rightarrow V$ con $\beta(0) = p$ y $\beta(1) = q$. La regla de la cadena nos permite hacer el siguiente cálculo:

$$\begin{aligned} f(q) - f(p) &= f(\beta(1)) - f(\beta(0)) = \int_0^1 \frac{d}{dt} f(\beta(t)) dt = \\ &= \int_0^1 (df)_{\beta(t)}(\beta'(t)) dt = \int_0^1 0 dt = 0, \end{aligned}$$

y deducir que $f(p) = f(q)$. Como p y q eran puntos arbitrarios de V , queda visto que f es constante en V . \square

5.2.3 Formas de Pfaff

Definición 144. Sea $R \subseteq \mathbb{R}^n$ una región. Una **forma de Pfaff en R** es una aplicación ω que asigna a cada punto $p \in R$ una forma lineal $\omega_p : \mathbb{R}^n \rightarrow \mathbb{R}$. Es, pues, una aplicación

$$\omega : R \longrightarrow (\mathbb{R}^n)^* \quad , \quad R \ni p \longmapsto \omega_p \in (\mathbb{R}^n)^* ,$$

es decir que es un **campo de formas lineales** definido en R .

También se llama a ω **1-forma en R** o **forma diferencial de grado 1 en R** .

Los campos diferenciales df , definidos en el apartado anterior, son ejemplos especiales de formas de Pfaff.

La **suma de dos formas de Pfaff** se define sumando los valores en cada punto:

$$(\omega + \omega')_p \stackrel{\text{def}}{=} \omega_p + \omega'_p \quad , \quad \text{para cada } p \in R .$$

Dada una función escalar $h : R \rightarrow \mathbb{R}$, el **producto de la función por la forma de Pfaff** se define también punto a punto:

$$(h\omega)_p \stackrel{\text{def}}{=} h(p)\omega_p \quad , \quad \text{para cada } p \in R .$$

En particular, para las funciones constantes $h \equiv c$ queda definido el producto $c\omega$ que, junto con la suma, da al conjunto de las formas de Pfaff en R una estructura natural de espacio vectorial.

Si $F : R \rightarrow \mathbb{R}^n$ es un campo de vectores en R , la función escalar $\omega(F) : R \rightarrow \mathbb{R}$ se define evaluando cada forma lineal ω_p en el correspondiente vector $F(p)$:

$$\omega(F)(p) \stackrel{\text{def}}{=} \omega_p(F(p)) \quad , \quad \text{para cada } p \in R .$$

5.2.4 Expresión en coordenadas de una forma de Pfaff

Sean $x_1, \dots, x_n : \mathbb{R}^n \rightarrow \mathbb{R}$ las coordenadas estándar, vistas como funciones diferenciables. Para $j = 1, \dots, n$, el campo diferencial dx_j es la aplicación $dx_j : \mathbb{R}^n \rightarrow (\mathbb{R}^n)^*$ tal que para cada $p \in \mathbb{R}^n$ el valor $(dx_j)_p$ es la j -ésima coordenada estándar vista como función lineal en \mathbb{R}^n . O sea, cada dx_j es un *campo constante de formas lineales*.

Sean de nuevo $R \subseteq \mathbb{R}^n$ una región y ω una forma de Pfaff en R . Para cada punto $p \in R$ existen coeficientes únicos $c_{1p}, \dots, c_{np} \in \mathbb{R}$ tales que

$$\omega_p \equiv c_{1p}(dx_1)_p + \dots + c_{np}(dx_n)_p. \quad (111)$$

Para $j = 1, \dots, n$, sea $f_j : R \rightarrow \mathbb{R}$ la función escalar cuyo valor en cada punto $p \in R$ es c_{jp} . Entonces, utilizando las operaciones del apartado 5.2.3, las identidades puntuales (111) equivalen a la siguiente identidad entre campos de formas lineales:

$$\omega \equiv f_1 dx_1 + \dots + f_n dx_n, \quad (112)$$

Las funciones f_1, \dots, f_n que satisfacen (112) son únicas. De hecho, por (110) sabemos que en todo punto $p \in R$ se tiene $f_j(p) = \omega_p(\mathbf{e}_j)$. De nuevo usando las operaciones definidas en el apartado 5.2.3, esto último puede expresarse así:

$$f_j \equiv \omega(\mathbf{e}_j) \quad , \quad j = 1, \dots, n, \quad (113)$$

donde \mathbf{e}_j es visto como un *campo constante de vectores*.

En particular, para una función f diferenciable en algún abierto U se tiene:

$$\boxed{df \equiv f_{x_1} dx_1 + \dots + f_{x_n} dx_n} \quad (114)$$

pues para cada $j \in \{1, \dots, n\}$ y cada $p \in U$ tenemos $(df)_p(\mathbf{e}_j) = D_{\mathbf{e}_j} f(p) = f_{x_j}(p)$, es decir que $(df)(\mathbf{e}_j) \equiv f_{x_j}$ para $j = 1, \dots, n$.

Podemos considerar a ω como una función escalar de $2n$ variables, definida en el producto $R \times \mathbb{R}^n$ de la siguiente manera:

$$\omega : R \times \mathbb{R}^n \longrightarrow \mathbb{R} \quad , \quad \omega(p, v) = \omega_p(v),$$

es decir $\omega(x_1, \dots, x_n, y_1, \dots, y_n) \equiv f_1(x_1, \dots, x_n)y_1 + \dots + f_n(x_1, \dots, x_n)y_n$, fórmula que nos dice que las formas de Pfaff son **funciones lineales en y_1, \dots, y_n y arbitrarias en x_1, \dots, x_n** .

5.2.5 Isomorfismos musicales

De los tres isomorfismos musicales que utilizaremos, en este apartado definimos los dos primeros.

Dada una región $R \subseteq \mathbb{R}^n$, a cada campo de vectores $F : R \rightarrow \mathbb{R}^n$ le asociamos la forma de Pfaff F^\flat (se lee “efe bemol”) definida de la manera siguiente

$$\text{para cualesquiera } p \in R, v \in \mathbb{R}^n \quad , \quad F_p^\flat(v) = F(p) \cdot v.$$

En coordenadas estándar:

$$(F_1, \dots, F_n)^\flat = F_1 dx_1 + \dots + F_n dx_n.$$

La aplicación $F \longmapsto F^\flat$ es un isomorfismo lineal

$$\{\text{campos de vectores en } R\} \longrightarrow \{\text{formas de Pfaff en } R\},$$

cuyo inverso es el segundo isomorfismo musical

$$\begin{array}{ccc} \{\text{formas de Pfaff en } R\} & \longrightarrow & \{\text{campos de vectores en } R\} \\ \omega & \longmapsto & \omega^\# \end{array}.$$

La fórmula $\omega^\#$ se lee “omega sostenido”. En coordenadas estándar:

$$(f_1 dx_1 + \dots + f_n dx_n)^\# = (f_1, \dots, f_n).$$

5.2.6 Formas de Pfaff cerradas y exactas

Dada una función h diferenciable en un abierto $U \subseteq \mathbb{R}^n$, está claro que

$$dh = \nabla h^\flat, \quad \nabla h = dh^\sharp.$$

Definición 145. Decimos que una forma de Pfaff ω , definida en un abierto $U \subseteq \mathbb{R}^n$, es **exacta** o, con más precisión, **exacta en U** , si existe una función $h : U \rightarrow \mathbb{R}$ tal que $\omega = dh$.

Es evidente que ω es exacta si y sólo si el campo de vectores ω^\sharp es un gradiente en U . Por lo tanto, una condición *necesaria* para que $\omega = f_1 dx_1 + \cdots + f_n dx_n$ pueda ser exacta es que la matriz jacobiana $D(\omega^\sharp)$ sea simétrica, lo cual se traduce en las **condiciones de derivadas cruzadas**

$$\text{para cualesquiera } i, j \in \{1, \dots, n\} \text{ con } i < j, \text{ se tiene } \frac{\partial f_j}{\partial x_i} = \frac{\partial f_i}{\partial x_j}, \quad (115)$$

que es un total de $\binom{n}{2} = n(n-1)/2$ condiciones.

Definición 146. Decimos que la forma de Pfaff $\omega = f_1 dx_1 + \cdots + f_n dx_n$ es **cerrada en U** , o simplemente **cerrada**, si cumple las condiciones (115) en todo punto de U .

Los resultados del apartado 5.1.7 nos dicen que en el plano \mathbb{R}^2 , o en una bola del plano, toda forma de Pfaff cerrada es exacta, y que en el plano perforado $\mathbb{R}^2 \setminus \{(0,0)\}$ la siguiente forma de Pfaff es cerrada pero no exacta:

$$\omega = \left(\frac{(y, -x)}{r^2} \right)^\flat = \frac{y dx - x dy}{x^2 + y^2}.$$

Hay dominios en los que las propiedades cerrada y exacta son equivalentes, y dominios en los que no son equivalentes.

Ahora podemos enunciar la siguiente consecuencia de la proposición 143 del apartado 5.2.2.

Corolario 147. Si ω es una forma de Pfaff exacta en un abierto conexo por caminos, entonces la función h tal que $\omega = dh$ es única salvo sumarle una constante.

Si F es un gradiente en un abierto conexo por caminos, entonces el potencial escalar de F es único salvo sumarle una constante.

5.2.7 Integración de formas de Pfaff

Las formas de Pfaff son integrandos que pueden integrarse sobre caminos.

Definición 148. Sea $\alpha(t) : [t_0, t_1] \rightarrow \mathbb{R}^n$ un camino diferenciable. Si una forma de Pfaff ω está definida en el conjunto imagen $\alpha([t_0, t_1])$, entonces definimos la **integral de ω a lo largo de α** de la manera siguiente

$$\int_{\alpha} \omega = \int_{t_0}^{t_1} \omega_{\alpha(t)}(\alpha'(t)) dt. \quad (116)$$

Atención: este número depende del camino, no sólo del conjunto imagen $\alpha([t_0, t_1])$. Por ejemplo, los caminos

$$\begin{aligned} \alpha(t) &= (\cos t, \sin t) \quad , \quad 0 \leq t \leq 2\pi, \\ \beta(t) &= (\cos t, \sin t) \quad , \quad 0 \leq t \leq 5\pi, \end{aligned}$$

tienen ambos por imagen la circunferencia unidad, pero

$$\int_{\alpha} x dy = \pi \quad \text{mientras que} \quad \int_{\beta} x dy = \frac{5}{2} \pi.$$

Esta integral es **aditiva en el camino**, lo que significa que si $t_0 < t' < t_1$ entonces

$$\int_{\alpha} \omega = \int_{\alpha|_{[t_0, t']}} \omega + \int_{\alpha|_{[t', t_1]}} \omega ,$$

y eso nos sugiere definir la integral también cuando α es **diferenciable a tramos**:

se integra ω sobre cada tramo diferenciable y se suman los resultados.

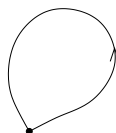
Sea $\varphi(s) : [s_0, s_1] \rightarrow [t_0, t_1]$ función diferenciable. Es fácil ver que

- Si $\varphi(s_0) = t_0$ y $\varphi(s_1) = t_1$, entonces $\int_{\alpha \circ \varphi} \omega = \int_{\alpha} \omega$.
- Si $\varphi(s_0) = t_1$ y $\varphi(s_1) = t_0$, entonces $\int_{\alpha \circ \varphi} \omega = - \int_{\alpha} \omega$.

Es decir que la integral de una forma de Pfaff es “sensible al sentido de recorrido”, de manera que si invertimos dicho sentido la integral se multiplica por -1 .

Definición 149. Dada una región $R \subseteq \mathbb{R}^n$, un **lazo en R** es un camino $\alpha(t) : [t_0, t_1] \rightarrow R$ que acaba en el mismo punto en el que empieza: $\alpha(t_1) = \alpha(t_0)$, que se llama **punto base**.

Conviene advertir que a un lazo diferenciable no se le pide que acabe con la misma velocidad con la que empieza, o sea que se permite $\alpha'(t_0) \neq \alpha'(t_1)$. Puede, pues, “tener una esquina” en el punto base.



Teorema 150. Sea ω una forma de Pfaff definida en un abierto $U \subseteq \mathbb{R}^n$. Son equivalentes:

- (1) ω es exacta en U .
- (2) Para los caminos $\alpha \subset U$, la integral $\int_{\alpha} \omega$ sólo depende de los extremos del camino.
- (3) La integral de ω a lo largo de los lazos contenidos en U es nula.

Demostración.

(1) \implies (3). Suponemos que existe $h : U \rightarrow \mathbb{R}$ tal que $\omega = dh$. Sea $\alpha(t) : [t_0, t_1] \rightarrow U$ un lazo diferenciable a trozos, con punto base $p = \alpha(t_0) = \alpha(t_1)$. Calculamos:

$$\int_{\alpha} \omega = \int_{t_0}^{t_1} (dh)_{\alpha(t)}(\alpha'(t)) dt = h(\alpha(t_1)) - h(\alpha(t_0)) = h(p) - h(p) = 0 .$$

(3) \implies (2). Elegimos dos puntos cualesquiera $p, q \in U$. Sean

$$\alpha(t) : [t_0, t_1] \longrightarrow U \quad , \quad \beta(t) : [t_2, t_3] \longrightarrow U ,$$

dos caminos que empiezan en p y terminan en q . Denotamos por β^{-1} un camino que haga el mismo recorrido que β pero en sentido inverso, con el parámetro ajustado para que empiece en $t = t_1$. La manera más sencilla de conseguir eso es definiendo:

$$\beta^{-1}(t) = \beta(t_1 + t_3 - t) \quad , \quad t_1 \leq t \leq t_1 + (t_3 - t_2) .$$

Sea $\gamma = \alpha\beta^{-1}$ el camino que primero hace el recorrido de α y después el de β^{-1}

$$\gamma(t) = \begin{cases} \alpha(t) & \text{para } t_0 \leq t \leq t_1 \\ \beta^{-1}(t) & \text{para } t_1 \leq t \leq t_1 + t_3 - t_2 \end{cases}$$

La definición de γ es válida porque $\alpha(t_1) = q = \beta^{-1}(t_1)$. Además γ es un lazo, porque empieza y termina en p . Por lo tanto

$$0 = \int_{\gamma} \omega = \int_{\alpha} \omega + \int_{\beta^{-1}} \omega = \int_{\alpha} \omega - \int_{\beta} \omega ,$$

es decir $\int_{\alpha} \omega = \int_{\beta} \omega$, igualdad que se cumple para cualesquiera dos caminos en U que tengan los mismos extremos.

(2) \implies (1). Para cada componente conexa por caminos $V \subseteq U$ vamos a construir una función $h^V : V \rightarrow \mathbb{R}$ tal que $\omega|_V = dh^V$. Como los V son abiertos disjuntos, las funciones h^V se juntan en una $h : U \rightarrow \mathbb{R}$ diferenciable y tal que $\omega = dh$.

Elegida V fijamos un punto $p_0 \in V$. Para todo $p \in V$ existen caminos $\alpha(t) : [t_0, t_1] \rightarrow V$ que unen p_0 con p , es decir $\alpha(t_0) = p_0$ y $\alpha(t_1) = p$. El lema 142 del apartado 5.2.2 nos dice que entre los caminos en V que unen p_0 con p los hay que son suaves. Definimos

$$h^V : V \longrightarrow \mathbb{R} \quad , \quad h^V(p) = \int_{\alpha} \omega ,$$

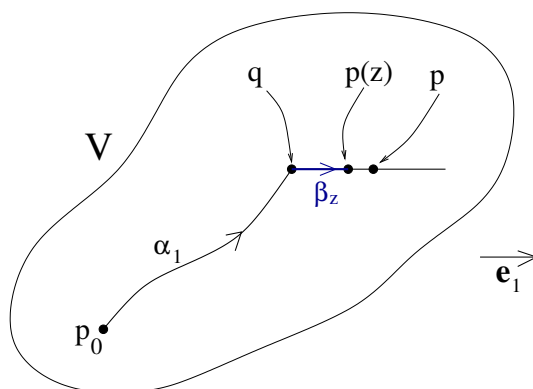
siendo $\alpha \subset V$ un camino suave que une p_0 con p . La hipótesis (2) garantiza que el resultado sólo depende de p , no del camino elegido para ir de p_0 a p sin salirse de V .

Sea $\omega = f_1 dx_1 + \dots + f_n dx_n$. Hay que demostrar que para cada $p \in V$ y cada $j \in \{1, 2, \dots, n\}$ la derivada $\left. \frac{\partial}{\partial x_j} \right|_p h^V$ existe y es igual a $f_j(p)$. Lo haremos para $j = 1$ y quedará claro cómo hacerlo para los otros j .

Fijamos $p = (a_1, \dots, a_n) \in V$. Como V es un abierto, existe un $r > 0$ (dependiente de p) tal que $B = B(p, 2r) \subseteq V$. El punto $q = (a_1 - r, a_2, \dots, a_n)$ está en B y fijamos un camino suave $\alpha_1(t) : [-1, 0] \rightarrow V$ con $\alpha_1(-1) = p_0$ y $\alpha_1(0) = q$. Para $a_1 - r < z < a_1 + r$, el punto q se une con el punto $p(z) = (z, a_2, \dots, a_n)$ mediante el camino

$$\beta_z(t) = (a_1 - r + t, a_2, \dots, a_n) \quad , \quad t \in [0, r - a_1 + z] ,$$

que es rectilíneo y paralelo al vector \mathbf{e}_1 .



La notación multiplicativa $\alpha_1 \beta_z$ indica el camino que primero hace el recorrido de α_1 y después el de β_z . Por la definición que hemos dado de h^V , y por ser $\beta'_z(t) \equiv \mathbf{e}_1$, tenemos

$$\begin{aligned} h^V(z, a_2, \dots, a_n) &= \int_{\alpha_1 \beta_z} \omega = \int_{\alpha_1} \omega + \int_{a_1 - r}^z \omega_{(s, a_2, \dots, a_n)}(\mathbf{e}_1) ds = \\ &= \text{cte}_1 + \int_{\text{cte}_2}^z f_1(s, a_2, \dots, a_n) ds , \end{aligned}$$

luego la derivada $\left. \frac{\partial}{\partial x_1} \right|_{(a_1, a_2, \dots, a_n)} h^V$ existe y es igual a $f_1(a_1, a_2, \dots, a_n)$: $h^V_{x_1}(p) = f_1(p)$.

Como p era un punto arbitrario de V , tenemos $h^V_{x_1} \equiv f_1|_V$. \square

La forma de Pfaff $\frac{ydx - xdy}{x^2 + y^2}$ no es exacta en el plano perforado porque su integral a lo largo del lazo $\alpha(t) = (\cos t, \sin t)$, $0 \leq t \leq 2\pi$, es igual a -2π en vez de ser nula. Esto significa que no existe en el plano perforado ninguna función h que satisfaga *las dos* ecuaciones de la fórmula (107) del apartado 5.1.7. Lo que hemos demostrado en dicho apartado es un poco más fuerte: no hay en el plano perforado ninguna h que cumpla *la primera* de las dos ecuaciones de la fórmula (107).

5.3 Tercer tipo de campos: campos de formas bilineales

5.3.1 Funciones alternadas de grado 2

Sea E un espacio vectorial. Recordemos que una función $\phi : E \times E \rightarrow \mathbb{R}$ es **bilineal** si cumple lo siguiente:

1. Para cada vector fijo $w_0 \in E$, la función $v \mapsto \phi(v, w_0)$ es lineal.
2. Para cada vector fijo $v_0 \in E$, la función $w \mapsto \phi(v_0, w)$ es lineal.

Definición 151. Una función bilineal $\phi : E \times E \rightarrow \mathbb{R}$ es **alternada** si cumple $\phi(v, v) = 0$ para todo $v \in E$. También se dice que ϕ es una **forma alternada de grado 2 en E** . El conjunto de todas estas funciones, denotado $A^2(E)$, es cerrado para la suma habitual de funciones y el producto de constante por función, luego es un espacio vectorial.

El cálculo estándar:

$$0 = \phi(v + w, v + w) = \phi(v, v) + \phi(v, w) + \phi(w, v) + \phi(w, w) = 0 + \phi(v, w) + \phi(w, v) + 0,$$

demuestra que $\phi(w, v) \equiv -\phi(v, w)$, es decir que el valor de una función de éstas se multiplica por -1 cuando se intercambian las variables vector. De esto se deduce que cada función bilineal alternada $\phi : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ viene dada por la fórmula

$$\phi(x, y) \equiv x^t A y, \quad (117)$$

siendo A una matriz antisimétrica $n \times n$.

5.3.2 Producto exterior de dos formas lineales

Definición 152. Dadas dos formas lineales $\ell, \ell' : E \rightarrow \mathbb{R}$, construimos una forma bilineal alternada $\ell \wedge \ell' : E \times E \rightarrow \mathbb{R}$ de la manera siguiente:

$$(\ell \wedge \ell')(v_1, v_2) \stackrel{\text{def}}{=} \ell(v_1)\ell'(v_2) - \ell(v_2)\ell'(v_1) = \det \begin{pmatrix} \ell(v_1) & \ell(v_2) \\ \ell'(v_1) & \ell'(v_2) \end{pmatrix}.$$

Llamamos a $\ell \wedge \ell'$ el **producto exterior de ℓ y ℓ'** .

Algunas propiedades obvias:

$$\begin{aligned} \ell \wedge (\ell' + \ell'') &= \ell \wedge \ell' + \ell \wedge \ell'', \\ (c\ell) \wedge \ell' &= c(\ell \wedge \ell') = \ell \wedge (c\ell'), \\ \ell \wedge \ell &= 0, \quad \ell' \wedge \ell = -\ell \wedge \ell'. \end{aligned}$$

Vamos a determinar el elemento general $\phi \in A^2(\mathbb{R}^n)$. Sean $\ell_1, \dots, \ell_n : \mathbb{R}^n \rightarrow \mathbb{R}$ las coordenadas estándar, vistas como funciones lineales. Partiendo de la fórmula (117), escribimos $A = (a_{ij})_{n \times n}$ y se cumplen las relaciones $a_{ii} = 0$ y $a_{ij} = -a_{ji}$. Entonces:

$$\phi(x, y) = \sum_{1 \leq i, j \leq n} x_i a_{ij} y_j = \sum_{i \neq j} x_i a_{ij} y_j = \sum_{i < j} x_i a_{ij} y_j + \sum_{i > j} x_i a_{ij} y_j.$$

Intercambiamos las letras i, j en la última suma y obtenemos:

$$\begin{aligned}\phi(x, y) &= \sum_{i < j} x_i a_{ij} y_j + \sum_{i < j} x_j a_{ji} y_i = \sum_{i < j} x_i a_{ij} y_j + \sum_{i < j} y_i (-a_{ij}) x_j = \\ &= \sum_{i < j} a_{ij} (x_i y_j - y_i x_j) = \sum_{i < j} a_{ij} (\ell_i \wedge \ell_j)(x, y)\end{aligned}$$

En definitiva:

$$\phi \equiv \sum_{i < j} a_{ij} \ell_i \wedge \ell_j, \quad (118)$$

que nos dice que $\{\ell_i \wedge \ell_j : i < j\}$ es un sistema de generadores para el espacio $A^2(\mathbb{R}^n)$. Pero de hecho es una base, porque dadas dos parejas $i < j$ e $i' < j'$ se comprueba de inmediato que

$$(\ell_i \wedge \ell_j)(\mathbf{e}_{i'}, \mathbf{e}_{j'}) = \begin{cases} 1 & \text{si } (i, j) = (i', j') \\ 0 & \text{si } (i, j) \neq (i', j') \end{cases}$$

de donde se deduce:

$$\phi \equiv \sum_{i < j} a_{ij} \ell_i \wedge \ell_j \implies \phi(\mathbf{e}_i, \mathbf{e}_j) = a_{ij} \text{ para cualesquiera } i < j, \quad (119)$$

y así los coeficientes en el desarrollo (118) son únicos. Una consecuencia es

$$\dim A^2(\mathbb{R}^n) = \text{cardinal} \{(i, j) : 1 \leq i < j \leq n\} = \binom{n}{2} = \frac{n(n-1)}{2}.$$

5.3.3 Formas diferenciales de grado 2

Definición 153. Sea $R \subseteq \mathbb{R}^n$ una región. Una **2-forma en R** o **forma diferencial de grado 2 en R** es una aplicación Ω que asigna a cada punto $p \in R$ una función bilineal alternada $\Omega_p : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$. Dicho de otra manera, es una aplicación

$$\Omega : R \longrightarrow A^2(\mathbb{R}^n).$$

o sea un **campo de formas bilineales alternadas** definido en R .

De nuevo definimos la suma:

$$(\Omega + \Omega')_p \stackrel{\text{def}}{=} \Omega_p + \Omega'_p \text{ para todo } p \in R,$$

y el producto por una función $h : R \rightarrow \mathbb{R}$:

$$(h\Omega)_p \stackrel{\text{def}}{=} h(p) \Omega_p \text{ para todo } p \in R.$$

Si $F_1, F_2 : R \rightarrow \mathbb{R}^n$ son dos campos de vectores, la función $\Omega(F_1, F_2) : U \rightarrow \mathbb{R}$ se define así:

$$\Omega(F_1, F_2)(p) \stackrel{\text{def}}{=} \Omega_p(F_1(p), F_2(p)) \text{ para todo } p \in R.$$

El **producto exterior de dos formas de Pfaff ω, ω' en R** es una 2-forma $\omega \wedge \omega'$ en R que se define multiplicando los valores en cada punto:

$$(\omega \wedge \omega')_p \stackrel{\text{def}}{=} \omega_p \wedge \omega'_p \text{ para todo } p \in R.$$

Sean $x_1, \dots, x_n : \mathbb{R}^n \rightarrow \mathbb{R}$ las coordenadas estándar, vistas como funciones diferenciables. Dada una 2-forma Ω en R , para cada punto $p \in R$ existen coeficientes únicos $\{c_{ijp}\}_{1 \leq i < j \leq n}$ tales que

$$\Omega_p \equiv \sum_{1 \leq i < j \leq n} c_{ijp} (dx_i)_p \wedge (dx_j)_p,$$

además cada coeficiente viene dado por $c_{ijp} = \Omega_p(\mathbf{e}_i, \mathbf{e}_j)$. Si $f_{ij} : R \rightarrow \mathbb{R}$ son las funciones tales que $f_{ij}(p) = c_{ijp}$ para cada $p \in R$, entonces

$$\Omega \equiv \sum_{1 \leq i < j \leq n} f_{ij} dx_i \wedge dx_j ,$$

y las funciones que cumplen esta identidad son únicas, estando dadas por las fórmulas

$$f_{ij} \equiv \Omega(\mathbf{e}_i, \mathbf{e}_j) ,$$

en las que $\mathbf{e}_i, \mathbf{e}_j$ son vistos como campos constantes de vectores.

Algunas propiedades evidentes, que nos ayudan con los cálculos:

$$\begin{aligned} \omega \wedge \omega &\equiv 0 , \\ \omega_2 \wedge \omega_1 &= -\omega_1 \wedge \omega_2 , \\ (f \omega_1) \wedge \omega_2 &= f \cdot (\omega_1 \wedge \omega_2) = \omega_1 \wedge (f \omega_2) , \\ (\omega_1 + \omega'_1) \wedge \omega_2 &= \omega_1 \wedge \omega_2 + \omega'_1 \wedge \omega_2 , \\ \omega_1 \wedge (\omega_2 + \omega'_2) &= \omega_1 \wedge \omega_2 + \omega_1 \wedge \omega'_2 . \end{aligned}$$

Un ejemplo sencillo en \mathbb{R}^4 :

$$\begin{aligned} &(dx_1 + x_1 x_3 dx_2 - 2e^{x_2} dx_4) \wedge (e^{x_1} dx_2 + x_3^2 dx_4) = \\ &= e^{x_1} dx_1 \wedge dx_2 + x_3^2 dx_1 \wedge dx_4 + 0 + x_1 x_3 x_3^2 dx_2 \wedge dx_4 - 2e^{x_2+x_1} \underbrace{dx_4 \wedge dx_2}_{=-dx_2 \wedge dx_4} + 0 = \\ &= e^{x_1} dx_1 \wedge dx_2 + x_3^2 dx_1 \wedge dx_4 + (x_1 x_3^3 + 2e^{x_1+x_2}) dx_2 \wedge dx_4 . \end{aligned}$$

5.3.4 Derivada exterior de una forma de Pfaff

Definición 154. Sea $\omega = f_1 dx_1 + \cdots + f_n dx_n$ una forma de Pfaff en un abierto $U \subseteq \mathbb{R}^n$. La **derivada exterior de ω** es una 2-forma en U que se denota $d\omega$ y viene dada por la siguiente fórmula

$$d\omega = df_1 \wedge dx_1 + \cdots + df_n \wedge dx_n .$$

También se la llama **diferencial exterior de ω** o **coborde de ω** .

Algunas propiedades:

$$d(\omega_1 + \omega_2) = d\omega_1 + d\omega_2 \quad , \quad d(f\omega) = df \wedge \omega + f d\omega .$$

Un ejemplo sencillo en \mathbb{R}^4 :

$$\begin{aligned} &d((x_1 + 2x_3 - x_4) dx_1 + x_1 x_2^5 x_3 dx_2 - e^{x_2} dx_3) = \\ &= (dx_1 + 2dx_3 - dx_4) \wedge dx_1 + (x_2^5 x_3 dx_1 + 5x_1 x_2^4 x_3 dx_2 + x_1 x_2^5 dx_3) \wedge dx_2 - (e^{x_2} dx_2) \wedge dx_3 = \\ &= 0 + 2dx_3 \wedge dx_1 - dx_4 \wedge dx_1 + x_2^5 x_3 dx_1 \wedge dx_2 + 0 + x_1 x_2^5 dx_3 \wedge dx_2 - e^{x_2} dx_2 \wedge dx_3 = \\ &= x_2^5 x_3 dx_1 \wedge dx_2 - 2dx_1 \wedge dx_3 + dx_1 \wedge dx_4 - (x_1 x_2^5 + e^{x_2}) dx_2 \wedge dx_3 . \end{aligned}$$

Es fácil comprobar que

$$d(f_1 dx_1 + \cdots + f_n dx_n) = \sum_{1 \leq i < j \leq n} \left(\frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j} \right) dx_i \wedge dx_j , \quad (120)$$

por lo tanto, la condiciones de derivadas cruzadas (115) del apartado 5.2.6 equivalen a $d\omega \equiv 0$.

Una forma de Pfaff ω es cerrada si y sólo si $d\omega \equiv 0$.

Puesto que vimos en el apartado 5.2.6 que toda forma de Pfaff exacta es cerrada, deducimos la siguiente identidad:

$d(df) \equiv 0$

(121)

válida para toda función escalar f que sea al menos de clase \mathcal{C}^2 .

5.3.5 Rotacional y derivada exterior

La derivada exterior de formas de Pfaff está relacionada con la matriz rotacional. Dado un campo de vectores F en un abierto $U \subseteq \mathbb{R}^n$, consideramos la forma de Pfaff F^\flat y su derivada exterior. Se comprueba que para cualesquiera $p \in U$ y $a, b \in \mathbb{R}^n$ se verifica:

$$d(F^\flat)_p(a, b) = a^t [(DF)^t - DF] b, \quad (122)$$

de hecho esta fórmula es esencialmente igual a la (120). Nótese que $(DF)^t - DF$ es la opuesta de la matriz rotacional.

En el caso particular $n = 3$ tenemos también el vector rotacional, y la identidad

$$[DF - (DF)^t]_p b = (\mathbf{rot} F)_p \times b. \quad (123)$$

Juntando (122) y (123), deducimos para $n = 3$:

$$d(F^\flat)_p(a, b) = -a \cdot ((\mathbf{rot} F)_p \times b) = -\det [a \mid \mathbf{rot} F \mid b] = \det [\mathbf{rot} F \mid a \mid b]. \quad (124)$$

Introducimos ahora el **tercer isomorfismo musical para $n = 3$** . Dada una región $R \subseteq \mathbb{R}^3$ y un campo $G : R \rightarrow \mathbb{R}^3$, definimos una 2-forma G^\sharp (se lee “ge becuadro”) que para $p \in R$ y $a, b \in \mathbb{R}^3$ tiene el siguiente efecto:

$$(G^\sharp)_p(a, b) = \det [G(p) \mid a \mid b]. \quad (125)$$

Entonces la identidad (124) la podemos escribir de la siguiente manera

$$\boxed{n = 3 \implies d(F^\flat) = (\mathbf{rot} F)^\sharp} \quad (126)$$

5.3.6 Variación de la integral al deformar el camino

En este apartado vamos a obtener una fórmula (la (127), escrita más abajo) que revela el verdadero significado de la derivada exterior.

Empezamos definiendo qué entendemos por deformar suavemente un camino.

Definición 155. En un abierto $U \subseteq \mathbb{R}^n$ tenemos un camino $\alpha(t) : [t_0, t_1] \rightarrow U$.

Una **deformación de α** viene dada por un intervalo $J \subset \mathbb{R}$, un valor particular $\lambda_0 \in J$ y una función de dos variables

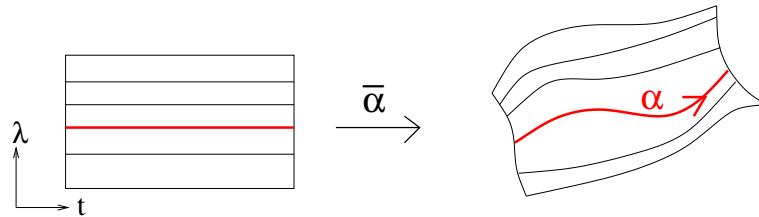
$$\bar{\alpha}(t, \lambda) : [t_0, t_1] \times J \longrightarrow U,$$

tal que $\bar{\alpha}(t, \lambda_0) = \alpha(t)$ para todo $t \in [t_0, t_1]$.

Para cada valor $\lambda \in J$ hay un camino $t \mapsto \bar{\alpha}(t, \lambda)$ que va cambiando con λ , por esto decimos que λ es el **parámetro de deformación**.

Esta definición la vamos a aplicar en dos situaciones

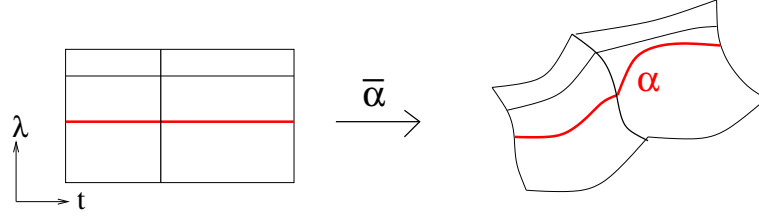
1. Cuando $\alpha(t)$ es suave y $\bar{\alpha}(t, \lambda)$ también lo es.



Entonces $\bar{\alpha}$ es una **deformación suave** del camino suave α .

2. Cuando hay una partición $t_0 < s_1 < s_2 < \dots < s_k < t_1$ y $\bar{\alpha}$ es continua en todo el rectángulo $[t_0, t_1] \times J$ y suave en cada parte:

$$[t_0, s_1] \times J \quad , \quad [s_j, s_{j+1}] \times J \quad \text{para } j = 1, \dots, k-1 \quad , \quad [s_k, t_1] \times J .$$

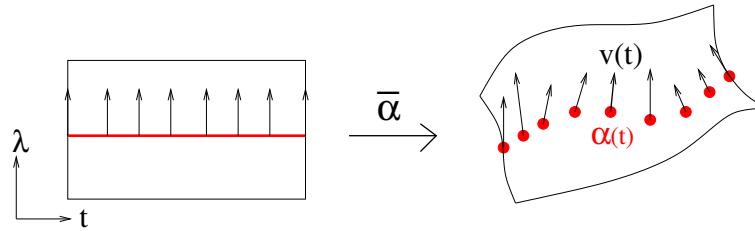


En este caso α es un camino suave a trozos y $\bar{\alpha}$ es una deformación suave a trozos de α .

En realidad, una deformación suave a trozos son varias deformaciones suaves pegadas una a continuación de la otra.

Fijado λ , el vector $\frac{\partial \bar{\alpha}}{\partial t}$ es tangente al correspondiente camino $t \mapsto \bar{\alpha}(t, \lambda)$, por lo que lo llamamos **velocidad tangencial**. Además $\frac{\partial \bar{\alpha}}{\partial t}(t, \lambda_0) = \alpha'(t)$.

Para definir la **velocidad lateral** conviene ver el camino α como un “collar de perlas”. Fijado un tiempo $t \in [t_0, t_1]$, a medida que cambiamos el valor de λ (o sea, deformamos el camino) la “perla” $\bar{\alpha}(t, \lambda)$ se desplaza lateralmente y cuando $\lambda = \lambda_0$ (cuando la perla está sobre el camino α) su velocidad de desplazamiento lateral es $\left. \frac{\partial \bar{\alpha}}{\partial \lambda} \right|_{\lambda=\lambda_0}$.



Llamamos **velocidad de deformación en $\lambda = \lambda_0$** a la función

$$v : [t_0, t_1] \longrightarrow \mathbb{R}^n \quad , \quad v(t) = \frac{\partial \bar{\alpha}}{\partial \lambda}(t, \lambda_0) ,$$

cuyos valores son las velocidades laterales. Este objeto es un campo de vectores a lo largo de α . La idea es que, como el “collar” se deforma por los desplazamientos laterales de infinitas perlas, la velocidad (en $\lambda = \lambda_0$) de esa deformación está hecha de infinitos vectores $v(t)$.

Teorema 156. Sean ω forma de Pfaff en U y $\bar{\alpha}(t, \lambda) \subset U$ deformación con $\bar{\alpha}(t, \lambda_0) = \alpha(t)$. Podemos definir una función escalar $I(\lambda) : J \rightarrow \mathbb{R}$ de la siguiente manera

$$I(\lambda) = \int_{\bar{\alpha}(\cdot, \lambda)} \omega ,$$

es decir, integramos ω a lo largo del λ -ésimo camino de la deformación. Entonces:

$$I'(\lambda_0) = \left[\omega_{\alpha(t)}(v(t)) \right]_{t_0}^{t_1} + \int_{t_0}^{t_1} (d\omega)_{\alpha(t)}(v(t), \alpha'(t)) dt \quad (127)$$

Atención: la velocidad lateral se pone a la izquierda de la coma en el paréntesis de la derivada exterior.

Llamamos a $\left[\omega_{\alpha(t)}(v(t)) \right]_{t_0}^{t_1}$ **término de evaluación** y a $\int_{t_0}^{t_1} (d\omega)_{\alpha(t)}(v(t), \alpha'(t)) dt$ **término integral**.

Demostración.

Es importante notar que en el término de evaluación sólo aparece la velocidad lateral, no la tangencial.

En una deformación suave a trozos, la velocidad tangencial tiene dos valores diferentes en los tiempos de transición s_j (una por la izquierda, la otra por la derecha) mientras que la velocidad lateral tiene un único valor, bien definido.

Supongamos cierto el teorema para deformaciones suaves y consideremos una deformación suave a trozos. Entonces la fórmula (127) es cierta para cada pieza suave de la deformación y podemos sumar esas igualdades miembro a miembro. En el miembro de la derecha se suman los términos de evaluación, pero en cada tiempo de transición s_j la velocidad lateral “desde la izquierda” es igual a la que “viene de la derecha”. Por lo tanto, en la suma de los términos de evaluación se cancelan dos a dos los valores intermedios $\omega(v(s_j))$:

$$\cdots + [\omega(v(s_j)) - \omega(v(s_{j-1}))] + [\omega(v(s_{j+1})) - \omega(v(s_j))] + \cdots$$

y lo único que sobrevive al simplificar esa suma es $\omega(v(t_1)) - \omega(v(t_0))$, que es el término de evaluación para el camino α sin cortarlo en tramos. También la suma de los términos integrales es el término integral para todo el camino α , y análogamente la suma de los términos de la izquierda. Luego (127) también será cierta para deformaciones suaves a trozos.

En el resto de la demostración vamos a suponer que la deformación es suave.

Los dos miembros de la igualdad (127) dependen linealmente de ω . En particular, si $\omega = \omega_1 + \omega_2$ entonces la integrales respectivas cumplen $I(\lambda) \equiv I_1(\lambda) + I_2(\lambda)$, luego $I'(\lambda_0) = I'_1(\lambda_0) + I'_2(\lambda_0)$. La aditividad respecto de ω del miembro de la derecha de (127) es aún más obvia. Como toda forma de Pfaff es una suma finita $f_1 dx_1 + \cdots + f_n dx_n$, es suficiente con demostrar el teorema para cada una de las formas $f_i dx_i$.

En el resto de la demostración fijamos un $j \in \{1, \dots, n\}$ y suponemos que $\omega = f dx_j$.

Para poder hacer cálculos, damos nombres a las entradas de las funciones:

$$\begin{aligned} \bar{\alpha}(t, \lambda) &\equiv (\bar{x}_1(t, \lambda), \dots, \bar{x}_n(t, \lambda)) , \\ \alpha(t) &\equiv (x_1(t), \dots, x_n(t)) , \\ v(t) &\equiv (v_1(t), \dots, v_n(t)) , \end{aligned}$$

y tenemos las siguientes identidades, para $i = 1, \dots, n$ y para todo t

$$\bar{x}_i(t, \lambda_0) = x_i(t) \quad , \quad \frac{\partial \bar{x}_i}{\partial t}(t, \lambda_0) = x'_i(t) \quad , \quad \frac{\partial \bar{x}_i}{\partial \lambda}(t, \lambda_0) = v_i(t) .$$

Comenzamos nuestro cálculo utilizando la fórmula para derivar una integral respecto de un parámetro y después las reglas del producto y de la cadena:

$$\begin{aligned} I'(\lambda_0) &= \left. \frac{d}{d\lambda} \right|_{\lambda=\lambda_0} \int_{\bar{\alpha}(\cdot, \lambda)} \omega = \left. \frac{d}{d\lambda} \right|_{\lambda=\lambda_0} \int_{t_0}^{t_1} \omega_{\bar{\alpha}(t, \lambda)} \left(\frac{\partial \bar{\alpha}}{\partial t} \right) dt = \\ &= \int_{t_0}^{t_1} \left. \frac{\partial}{\partial \lambda} \right|_{\lambda=\lambda_0} \left[f(\bar{\alpha}(t, \lambda)) \frac{\partial \bar{x}_j}{\partial t} \right] dt = \\ &= \int_{t_0}^{t_1} \left[(df)_{\alpha(t)}(v(t)) x'_j(t) + f(\alpha(t)) \frac{\partial^2 \bar{x}_j}{\partial t \partial \lambda}(t, \lambda_0) \right] dt . \end{aligned}$$

Como estamos suponiendo que la deformación es suave, se tiene

$$\frac{\partial^2 \bar{x}_j}{\partial t \partial \lambda}(t, \lambda_0) = \frac{\partial^2 \bar{x}_j}{\partial \lambda \partial t}(t, \lambda_0) = v'_j(t) ,$$

y así

$$I'(\lambda_0) = \int_{t_0}^{t_1} (df)_{\alpha(t)}(v(t)) x'_j(t) dt + \int_{t_0}^{t_1} f(\alpha(t)) v'_j(t) dt . \quad (128)$$

Integrando por partes:

$$\begin{aligned} \int_{t_0}^{t_1} f(\alpha(t)) v'_j(t) dt &= \left[f(\alpha(t)) v_j(t) \right]_{t_0}^{t_1} - \int_{t_0}^{t_1} \left(\frac{d}{dt} f(\alpha(t)) \right) v_j(t) dt = \\ &= \left[\omega_{\alpha(t)}(v(t)) \right]_{t_0}^{t_1} - \int_{t_0}^{t_1} (df)_{\alpha(t)}(\alpha'(t)) v_j(t) dt , \end{aligned}$$

que llevado a (128) nos da:

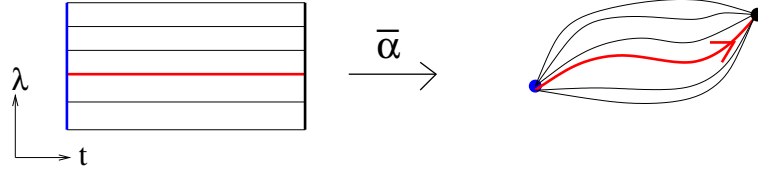
$$\begin{aligned} I'(\lambda_0) &= \left[\omega_{\alpha(t)}(v(t)) \right]_{t_0}^{t_1} + \int_{t_0}^{t_1} \left((df)_{\alpha(t)}(v(t)) x'_j(t) - (df)_{\alpha(t)}(\alpha'(t)) v_j(t) \right) dt = \\ &= \left[\omega_{\alpha(t)}(v(t)) \right]_{t_0}^{t_1} + \int_{t_0}^{t_1} (df \wedge dx_j)_{\alpha(t)}(v(t), \alpha'(t)) dt . \end{aligned}$$

Como $df \wedge dx_j = d\omega$, la demostración está completa. \square

Describimos a continuación dos situaciones en las que el término de evaluación es cero:

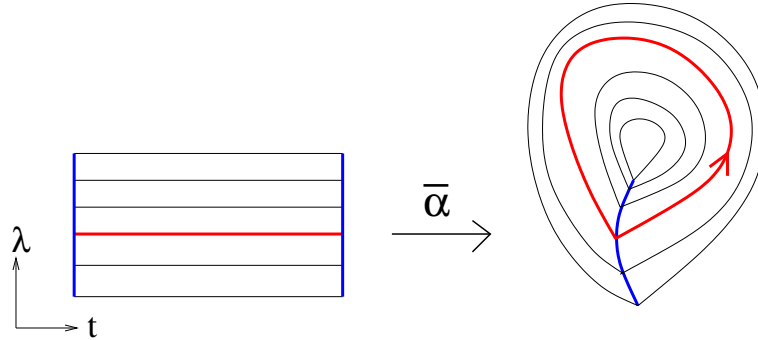
1. En una **deformación de extremos fijos**. Esto significa que para todo $\lambda \in J$ se tiene

$$\bar{\alpha}(t_0, \lambda) = \alpha(t_0) \quad \text{y} \quad \bar{\alpha}(t_1, \lambda) = \alpha(t_1) .$$



En este caso es $v(t_0) = v(t_1) = \mathbf{0}$, luego $\omega(v(t_0)) = \omega(v(t_1)) = 0$ y el término de evaluación es $0 - 0 = 0$.

2. Cuando α es un lazo y $\bar{\alpha}$ es una **deformación a través de lazos**. Se permite que el punto base de los lazos se vaya desplazando a medida que λ cambia, trazando un camino $\beta(\lambda)$ que indicamos en color azul



En este caso se tiene $v(t_0) = v(t_1) = \beta'(\lambda_0)$ y también $\alpha(t_0) = \alpha(t_1) = \beta(\lambda_0)$, con lo cual

$$\omega_{\alpha(t_1)}(v(t_1)) - \omega_{\alpha(t_0)}(v(t_0)) = \omega_{\beta(\lambda_0)}(\beta'(\lambda_0)) - \omega_{\beta(\lambda_0)}(\beta'(\lambda_0)) = 0 .$$

En esos dos casos particulares se cumple la siguiente fórmula, ya sólo con el término integral

$$I'(\lambda_0) = \int_{t_0}^{t_1} (d\omega)_{\alpha(t)}(v(t), \alpha'(t)) dt . \quad (129)$$

5.3.7 Integrales de 2-formas sobre funciones paramétricas

Definición 157. Tenemos un abierto $A \subseteq \mathbb{R}_{uv}^2$ del plano de los parámetros uv y $R \subset A$ región acotada. Tenemos también una función $\varphi(u, v) : R \rightarrow \mathbb{R}^n$, que es la restricción a R de una función diferenciable $A \rightarrow \mathbb{R}^n$, y una 2-forma Ω definida al menos en el conjunto imagen $\varphi(R) \subset \mathbb{R}^n$.

La **integral de Ω sobre φ** es el número definido por la siguiente fórmula

$$\int_{\varphi} \Omega = \iint_R \Omega_{\varphi(u,v)}(\varphi_u, \varphi_v) du dv. \quad (130)$$

Obsérvese el inmenso parecido con la fórmula (116) del apartado 5.2.7.

Si Ω es una 2-forma definida en una región $R \subset \mathbb{R}_{x_1 x_2}^2$, hacemos $\int_R \Omega \stackrel{\text{def}}{=} \int_{\text{id}_R} \Omega$, siendo id_R la aplicación identidad que lleva cada punto de R a sí mismo. En este caso, pues

$$\int_R \omega \stackrel{\text{def}}{=} \iint_R \Omega_{(x_1, x_2)}(\mathbf{e}_1, \mathbf{e}_2) dx_1 dx_2,$$

de modo que si tenemos escrita $\Omega = f dx_1 \wedge dx_2$ entonces $\int_R \Omega = \iint_R f dx_1 dx_2$, mientras que si tenemos $\Omega = g dx_2 \wedge dx_1$ entonces es $\int_R \Omega = \iint_R (-g) dx_1 dx_2$.

5.3.8 Cambio de parámetros en la integral de una 2-forma

La llamada **fórmula de cambio de parámetros en integrales dobles** puede enunciarse de la siguiente manera. Si tenemos una región $R_0 \subset \mathbb{R}_{s,t}^2$ y una aplicación $\tau(s, t) : R_0 \rightarrow \mathbb{R}_{uv}^2$ diferenciable y que cumple las siguientes condiciones

- (a) la imagen $\tau(R_0)$ coincide con R ,
- (b) las “preimágenes múltiples” (aquellos $(s, t) \in R_0$ para los que existe otro $(s', t') \in R_0$ distinto pero con $\tau(s', t')$ y $\tau(s, t)$ iguales) forman un conjunto de área nula,
- (c) $\det D\tau$ puede anularse, pero no cambia de signo en R_0 ,

entonces para toda $f : R \rightarrow \mathbb{R}$ se verifica la igualdad

$$\iint_{R_0} (f \circ \tau)(s, t) |\det D\tau| ds dt = \iint_R f(u, v) du dv.$$

Ejemplo. Eso se cumple para las coordenadas polares

$$\tau : [0, r_0] \times [0, 2\pi] \longrightarrow \mathbb{R}^2, \quad \tau(r, \theta) = (r \cos \theta, r \sin \theta),$$

cuyo conjunto imagen es el disco $R = \overline{B}((0, 0), r_0)$. Cada punto $(u, 0) \in R$ con $0 < u \leq r_0$ tiene dos preimágenes por θ , que son $(u, 0)$ y $(u, 2\pi)$. El origen $(0, 0)$ tiene infinitas preimágenes por θ : los $(0, \theta)$ con $\theta \in [0, 2\pi]$. En total, el conjunto de preimágenes múltiples es la unión de tres segmentos de recta

$$(\{0\} \times [0, 2\pi]) \cup ((0, r_0] \times \{0\}) \cup ((0, r_0] \times \{2\pi\}),$$

y por lo tanto tiene área nula. El jacobiano $\det D\tau$ es nulo en los puntos $(0, \theta)$, pero positivo en todos los demás. Se sigue que para toda $f(u, v) : R \rightarrow \mathbb{R}$ tenemos la igualdad

$$\iint_R f(u, v) du dv = \iint_{R_0} (f \circ \tau)(r, \theta) |\det D\tau| dr d\theta = \iint_{R_0} f(r \cos \theta, r \sin \theta) r dr d\theta.$$

Sea de nuevo cualquier $\tau(s, t) : R_0 \rightarrow \mathbb{R}^2$ diferenciable. Si en un punto (s_0, t_0) es $D\tau = \begin{bmatrix} a & b \\ a' & b' \end{bmatrix}$, entonces en dicho punto se tiene

$$(\varphi \circ \tau)_s = a \varphi_u \circ \tau + a' \varphi_v \circ \tau, \quad (\varphi \circ \tau)_t = b \varphi_u \circ \tau + b' \varphi_v \circ \tau,$$

luego para cualquier función bilineal alternada $\phi(\cdot, \cdot)$ se verifica

$$\phi((\varphi \circ \tau)_s, (\varphi \circ \tau)_t) = (ab' - a'b) \phi(\varphi_u, \varphi_v)_{\tau(s_0, t_0)} = \det D\tau \cdot \phi(\varphi_u, \varphi_v)_{\tau(s_0, t_0)}.$$

Al ser eso cierto para todo (s_0, t_0) , si τ satisface las condiciones (a), (b), (c) arriba indicadas deducimos

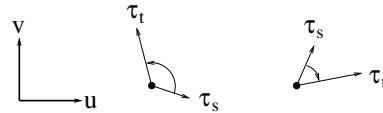
$$\begin{aligned} \int_{\varphi \circ \tau} \Omega &= \iint_{R_0} (\Omega(\varphi_u, \varphi_v))_{\tau(s, t)} \det D\tau \, ds dt = \\ &= (\text{sig } D\tau) \cdot \iint_{R_0} (\Omega(\varphi_u, \varphi_v))_{\tau(s, t)} |\det D\tau| \, ds dt, \end{aligned}$$

donde $\text{sig } D\tau$ es el signo del jacobiano allí donde éste no es cero, signo que permanece constante por la hipótesis 3. En definitiva:

$$\int_{\varphi \circ \tau} \Omega = \begin{cases} \int_{\varphi} \Omega & \text{si } \det D\tau \geq 0 \\ -\int_{\varphi} \Omega & \text{si } \det D\tau \leq 0 \end{cases} \quad (131)$$

Este resultado es análogo a lo observado en el apartado 5.2.7 de que la integral de una forma de Pfaff se multiplica por -1 si cambiamos el sentido de recorrido.

Tenemos $\det D\tau > 0$ si el sentido de giro del vector τ_s al vector τ_t , por el camino angular más corto, es el mismo que del semieje $u > 0$ al semieje $v > 0$ (sentido “antihorario”, con la colocación habitual de los ejes). Tenemos $\det D\tau < 0$ si el giro es del semieje $v > 0$ al semieje $u > 0$ (habitualmente “horario”).

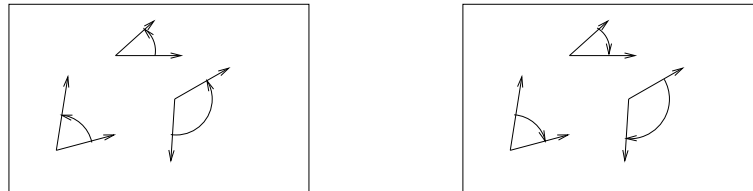


El sustituto en regiones planas del “sentido de recorrido” es el siguiente concepto

Definición 158. Una **orientación** en una curva es un sentido de recorrido para ella.

Una **orientación** en el plano \mathbb{R}^2 es un sentido de giro, por el camino angular más corto, común a todas las parejas ordenadas (v_1, v_2) , $v_1, v_2 \in \mathbb{R}^2$, que sean linealmente independientes.

El plano tiene exactamente dos orientaciones



lo cual permite reescribir la fórmula (131) de la siguiente manera

$$\int_{\varphi \circ \tau} \Omega = \begin{cases} \int_{\varphi} \Omega & \text{si } \tau \text{ conserva la orientación} \\ -\int_{\varphi} \Omega & \text{si } \tau \text{ invierte la orientación} \end{cases} \quad (132)$$

5.3.9 Fórmula paramétrica de Stokes

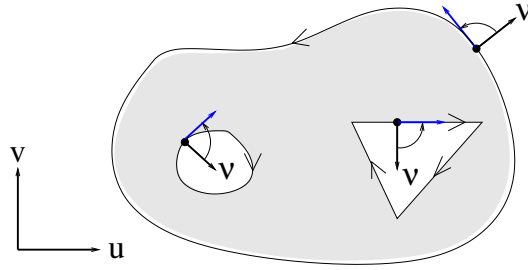
Definiciones 159. Sea $R \subset \mathbb{R}_{uv}^2$ una región acotada, limitada por unas curvas cerradas $\Gamma_1, \dots, \Gamma_k$ simples y disjuntas dos a dos. Llamamos **borde de R** al conjunto $\Gamma_1 \cup \dots \cup \Gamma_k$. La **normal unitaria exterior** es un campo de vectores a lo largo del borde

$$\nu : \Gamma_1 \cup \dots \cup \Gamma_k \longrightarrow \mathbb{R}^2,$$

tal que para cada j y cada punto $p \in \Gamma_j$ el vector $\nu(p)$ es normal a Γ_j en p , tiene longitud = 1 y apunta **saliendo de la región R** .

Para una región así, tenemos la **orientación del borde** que consiste en una orientación (un sentido de recorrido) para cada curva Γ_j . Damos dos descripciones:

1. Si el dibujo de la región es relativamente sencillo. Vemos que hay una **curva exterior**, que orientamos en el sentido del semieje $u > 0$ al semieje $v > 0$ (habitualmente, sentido “antihorario”) y todas las demás son **curvas interiores**, que orientamos en el sentido de giro del semieje $v > 0$ al semieje $u > 0$ (habitualmente, sentido “horario”).



2. Si estamos calculando en vez de recurrir al dibujo. Elegimos un punto $p_j \in \Gamma_j$, en el que sea fácil hallar el valor $\nu(p_j)$ de la normal exterior, y damos a Γ_j el sentido de recorrido que señala el vector tangente \mathbf{t}_j , resultado de rotar $\nu(p_j)$ un cuarto de vuelta en el sentido del semieje $u > 0$ al semieje $v > 0$. La orientación que resulta para Γ_j no depende de la elección de p_j , incluso si Γ_j sólo es diferenciable a tramos.

Recordamos aquí que la rotación alrededor del origen de un cuarto de vuelta, en el sentido de giro del semieje $u > 0$ al semieje $v > 0$, viene dada por la fórmula

$$(u, v) \longmapsto (v, -u).$$

Definición 160. Para cada j elegimos un lazo $\alpha_j(t)$ que dé a Γ_j exactamente una vuelta, en el sentido de la orientación del borde recién descrita.

Definimos el **borde de φ** como el sistema de lazos $\varphi \circ \alpha_1, \dots, \varphi \circ \alpha_k \subset \mathbb{R}^n$ y lo denotamos $\partial\varphi$.

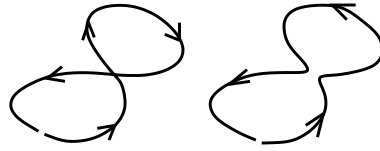
Teorema 161. (Fórmula de Stokes paramétrica). Sean $A \subseteq \mathbb{R}_{uv}^2$ un abierto y $R \subset A$ una región con borde orientado. Sea $\varphi(u, v) : R \rightarrow \mathbb{R}^n$ la restricción a R de una función diferenciable en A .

Para toda forma de Pfaff ω , diferenciable en algún abierto de \mathbb{R}^n que contenga la imagen $\varphi(R)$, se tiene:

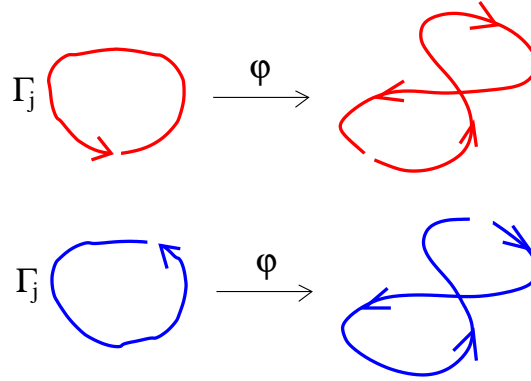
$$\boxed{\int_{\partial\varphi} \omega = \int_R d\omega} \quad (133)$$

Algunos comentarios:

1. Si la imagen $\varphi(\Gamma_j) \subset \mathbb{R}^n$ tiene autointersecciones, entonces hay muchos lazos en \mathbb{R}^n , que le dan exactamente una vuelta, de los cuales algunos son de la forma $\varphi \circ \alpha_j$, con α_j dando una vuelta a Γ_j , y otros no son en absoluto de esa forma.



Afortunadamente, gracias a las propiedades vistas en el apartado 5.2.7 la integral de ω es la misma en todos los lazos de la forma $\varphi \circ \alpha_j$.

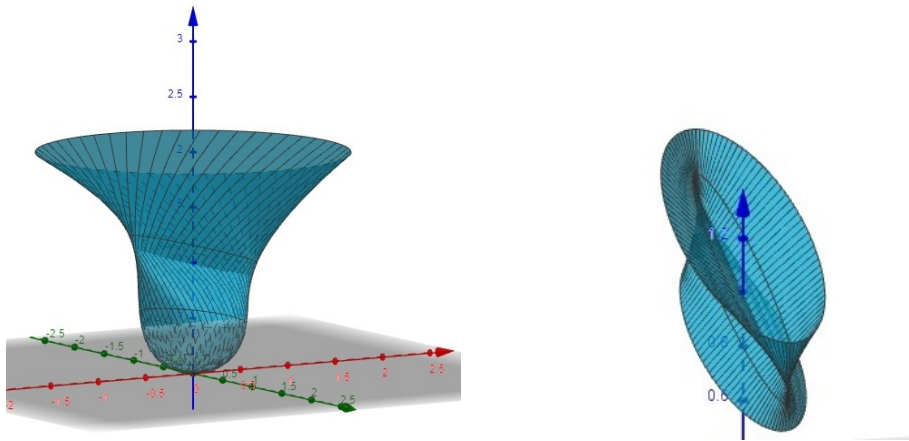


Luego la integral $\int_{\partial\varphi} \omega$ está bien definida.

2. El teorema 161 es cierto aunque la imagen $\varphi(R)$ sea una “superficie con autointersecciones”. Por ejemplo, vale para cualquier forma de Pfaff ω en \mathbb{R}^3 y la función paramétrica

$$\varphi(u, v) = ((1 - u^2 - v^2)u, v, u^2 + v^2) \quad , \quad R = \{(u, v) : u^2 + v^2 \leq 2\} \quad ,$$

cuya imagen se corta consigo misma a lo largo de un segmento rectilíneo, imagen de los valores paramétricos (u, v) con $u^2 + v^2 = 1$ (imagen derecha en el siguiente dibujo).



El borde $\partial\varphi$ de esta función es un lazo que da una vuelta a la circunferencia espacial

$$\{(x, y, z) : x^2 + y^2 = 2, z = 2\} \quad ,$$

pero se lo da en el sentido del semieje $y > 0$ al semieje $x > 0$. Se comprueba que:

$$\omega = xdy \implies \int_{\partial\varphi} \omega = -2\pi = \int_{\varphi} \omega \quad ,$$

también:

$$\omega = zdy \implies \int_{\partial\varphi} \omega = 0 = \int_{\varphi} \omega \quad ,$$

etc. etc.

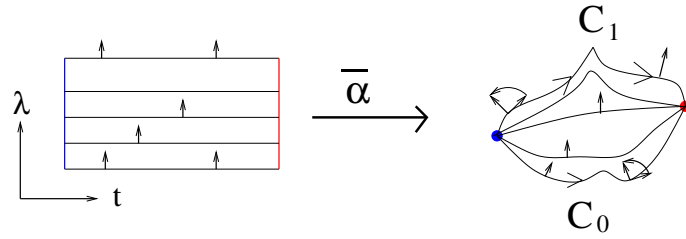
Demostración del teorema 161.

Supongamos primero que R es una *región sin agujeros*. En particular, su borde tiene sólo una curva Γ , que es la exterior. Elegir dos puntos $p, q \in \Gamma$ separa la curva en dos arcos C_0 y C_1 , ambos con p, q como extremos. Elijiendo adecuadamente a qué arco llamamos C_0 y a cuál C_1 , es posible dar una deformación de extremos fijos

$$\bar{\alpha}(t, \lambda) : [0, 1] \times [0, 1] \longrightarrow \mathbb{R}_{u,v}^2, \quad \bar{\alpha}(0, \lambda) \equiv p, \quad \bar{\alpha}(1, \lambda) \equiv q,$$

que cumpla las siguientes condiciones, donde $R_0 = [0, 1] \times [0, 1]$

- (1) $\bar{\alpha}(R_0) = R$.
- (2) El camino $\alpha_0(t) = \bar{\alpha}(t, 0)$ recorre C_0 en el sentido de la orientación del borde de R .
- (3) El camino $\alpha_1(t) = \bar{\alpha}(t, 1)$ recorre C_1 en el sentido opuesto a la orientación del borde.
- (4) El vector $\bar{\alpha}_\lambda$ es normal interior a R cuando $\lambda = 0$ y normal exterior cuando $\lambda = 1$.
- (5) $\bar{\alpha}$ es inyectiva y con jacobiano positivo para $(t, \lambda) \in (0, 1) \times [0, 1]$.



En primer lugar, la función $\bar{\alpha}$ cumple todos los requisitos que le hemos pedido a un cambio de parámetros en el apartado 5.3.8 para que se cumpla la siguiente igualdad

$$\int_{\varphi} d\omega = \int_{\varphi \circ \bar{\alpha}} d\omega. \quad (134)$$

En segundo lugar $\varphi \circ \bar{\alpha}$ es una deformación de extremos fijos que empieza en el camino $\varphi \circ \alpha_0$ y termina en el $\varphi \circ \alpha_1$. Aplicando la fórmula (129), obtenemos

$$\begin{aligned} \int_{\varphi \circ \bar{\alpha}} d\omega &= \iint_{0 \leq t, \lambda \leq 1} d\omega((\varphi \circ \bar{\alpha})_t, (\varphi \circ \bar{\alpha})_\lambda) dt d\lambda = \\ &= - \iint_{0 \leq t, \lambda \leq 1} d\omega((\varphi \circ \bar{\alpha})_\lambda, (\varphi \circ \bar{\alpha})_t) dt d\lambda = \\ &= - \int_{\lambda=0}^{\lambda=1} \frac{d}{d\lambda} \left(\int_{\varphi \circ \bar{\alpha}(\cdot, \lambda)} \omega \right) d\lambda = \\ &= - \int_{\varphi \circ \bar{\alpha}(\cdot, 1)} \omega + \int_{\varphi \circ \bar{\alpha}(\cdot, 0)} \omega = \\ &= - \int_{\varphi \circ \alpha_1} \omega + \int_{\varphi \circ \alpha_0} \omega = \int_{\varphi \circ (\alpha_0 \alpha_1^{-1})} \omega, \end{aligned} \quad (135)$$

donde $\alpha_0 \alpha_1^{-1}$ denota un camino que primero hace el recorrido de α_0 y después hace en sentido inverso el recorrido de α_1 .

En tercer lugar, las condiciones (2) y (3) nos dicen que $\alpha_0 \alpha_1^{-1}$ es un lazo que le da exactamente una vuelta al borde de R en el sentido de la orientación del borde, por lo tanto

$$\int_{\varphi \circ (\alpha_0 \alpha_1^{-1})} \omega = \int_{\partial \varphi} \omega,$$

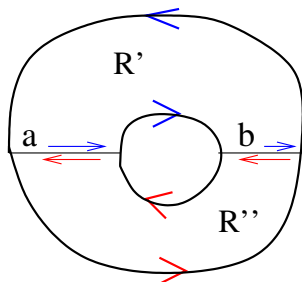
lo cual junto, con (134) y (135), nos conduce a la igualdad (133).

Importante: en el apartado 5.3.6 vimos que la fórmula (129) es válida para deformaciones suaves a trozos. Por lo tanto, la demostración que acabamos de hacer vale para una región sin agujeros R aunque su borde sea suave a trozos, con posibles esquinas.

Haremos ahora el caso de una región R con un agujero. Podemos hallar dos arcos $a, b \subset R$ que cortan la región en dos subregiones R', R'' que no tienen agujeros. Sabemos que la fórmula (133) es válida para estas subregiones y, sumando término a término las dos igualdades que resultan, obtenemos

$$\int_{\varphi} d\omega = \int_{\varphi|_{R'}} d\omega + \int_{\varphi|_{R''}} d\omega = \int_{\partial(\varphi|_{R'})} \omega + \int_{\partial(\varphi|_{R''})} \omega,$$

Los arcos a, b son parte de los bordes de R' y de R'' , recibiendo de R' orientaciones opuestas a las recibidas de R'' .



Luego las integrales de ω a lo largo de las partes de $\partial(\varphi|_{R'})$ y $\partial(\varphi|_{R''})$ procedentes de a o b se cancelan. Lo que único que queda después de esas cancelaciones es $\partial\varphi$ y tenemos

$$\int_{\partial(\varphi|_{R'})} \omega + \int_{\partial(\varphi|_{R''})} \omega = \int_{\partial\varphi} \omega.$$

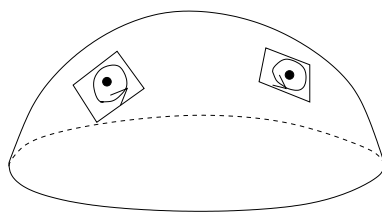
Queda hecho el caso de una región con un agujero.

El caso de una región con varios agujeros se hace de manera análoga. □

Observa que en la demostración recién hecha *todo el análisis se ha hecho en el plano de parámetros* \mathbb{R}_{uv}^2 . A los objetos φ y ω se les ha pedido que sean diferenciables y nada más.

5.3.10 Fórmula de Stokes clásica en \mathbb{R}^n

Definiciones 162. Sea $X \subset \mathbb{R}^n$ una subvariedad de dimensión 2. Una **orientación** es una estructura en X que consiste en definir un sentido de giro, en cada plano tangente de X , **con continuidad**: que no haya saltos bruscos en el sentido de giro al pasar de un punto de X a otro cercano.



Decimos que un “trozo de superficie” S es **orientable** si es parte de alguna subvariedad bidimensional que admite alguna orientación.

La **orientación estándar de \mathbb{R}^2** es la del sentido de giro de \mathbf{e}_1 a \mathbf{e}_2 por el camino angular más corto

Sea $X \subset \mathbb{R}^n$ una subvariedad bidimensional orientable y \mathcal{O} una orientación elegida en X . Dados un trozo acotado $S \subset X$ (una “superficie simple”) y una 2-forma Ω , definida al menos en S , definimos la **integral de Ω respecto de la orientación \mathcal{O}** , escrito $\int_{(S, \mathcal{O})} \Omega$, de la siguiente manera:

Elegimos cualquier región acotada $R \subset \mathbb{R}_{uv}^2$ y cualquier función diferenciable $\varphi(u, v) : R \rightarrow \mathbb{R}^n$ que tenga las siguientes propiedades (compáralas con las del principio del apartado 5.3.8)

- (a) la imagen $\varphi(R)$ coincide con S ,
- (b) las preimágenes múltiples (aquellos $(u, v) \in R$ para los que existe otro (u', v') distinto pero con $\varphi(u, v)$ y $\varphi(u', v')$ iguales) forman un conjunto de área nula,
- (c) φ es **compatible con la orientación \mathcal{O}** , que quiere decir que hay un conjunto de área nula $E \subset R$ tal que para todo $(u_0, v_0) \in R \setminus E$ el par ordenado $\{\varphi_u(u_0, v_0), \varphi_v(u_0, v_0)\}$ es una base de $T_{\varphi(u_0, v_0)}X$ que define el sentido de giro de la orientación \mathcal{O} ,

entonces definimos

$$\int_{(S, \mathcal{O})} \Omega \stackrel{\text{def}}{=} \int_{\varphi} \Omega. \quad (136)$$

En primer lugar, tal función φ existe (suele ser fácil hallarla en ejemplos concretos).

En segundo lugar, si $\psi : R_1 \rightarrow \mathbb{R}^n$ es otra función con esas mismas propiedades entonces hay conjuntos de área nula

$$E_0 \subset R \quad \text{y} \quad E_1 \subset R_1,$$

tales que $\psi|_{R_1 \setminus E_1} \equiv (\varphi|_{R \setminus E_0}) \circ \tau$, donde τ es un cambio de parámetros que cumple las condiciones indicadas al principio del apartado 5.3.8. Además, el signo constante de $\det D\tau$ es positivo, precisamente porque tanto φ como ψ son compatibles con la misma orientación \mathcal{O} . Luego

$$\int_{\varphi} \Omega = \int_{\psi} \Omega,$$

y (136) es una definición correcta de la integral de Ω sobre (S, \mathcal{O}) . Esta integral depende de la orientación de la siguiente manera:

1. Si \mathcal{O}' es la orientación en X que es inversa de \mathcal{O} en cada plano tangente $T_p X$, entonces

$$\int_{(S, \mathcal{O}')} \Omega = - \int_{(S, \mathcal{O})} \Omega.$$

2. Es imposible definir una integral de Ω sobre S si S es **no orientable**. Un ejemplo de superficie no orientable es la banda de Möbius, que hemos descrito en el apartado 4.5: sobre ella no se integran las formas de grado 2.

Teorema 163. (Fórmula clásica de Stokes en \mathbb{R}^n). Sea $X \subset \mathbb{R}^n$ una subvariedad bidimensional en la que se ha fijado una orientación \mathcal{O} . Sea $S \subseteq X$ un trozo de superficie, limitado por una cantidad finita de curvas cerradas simples $\Gamma_1, \dots, \Gamma_k \subset X$ disjuntas dos a dos. Llamamos **borde de S** al conjunto

$$\partial S = \Gamma_1 \cup \dots \cup \Gamma_k.$$

A partir de \mathcal{O} se define una cierta orientación en el borde ∂S tal que para **toda** forma de Pfaff ω , que sea diferenciable en algún abierto de \mathbb{R}^n conteniendo a S , se tiene

$$\boxed{\int_{(S, \mathcal{O})} d\omega = \int_{\partial S} \omega} \quad (137)$$

El teorema 161 nos dice cómo tenemos que definir la orientación de las curvas Γ_j para que se cumpla la fórmula (137):

Elegida $\varphi(u, v) : R \rightarrow S$, satisfaciendo (a), (b), (c), la utilizamos para trasladar a las curvas de ∂S la orientación como borde de las curvas de ∂R .

5.3.11 Fórmula de Stokes clásica en \mathbb{R}^3

Si $n = 3$, tenemos un método especial para definir las orientaciones de una subvariedad bidimensional $X \subset \mathbb{R}^3$:

corresponden biyectivamente a elecciones de **campo normal unitario**.

Es decir, corresponden a funciones $\nu : X \rightarrow \mathbb{R}^3$ con las siguientes propiedades:

1. ν es continua,
2. ν tiene longitud constante 1,
3. ν es normal a X : para cada punto $p \in X$, el vector $\nu(p)$ es normal al plano $T_p X$.

La correspondencia se describe de la siguiente manera:

1. Elegida una normal unitaria ν , el sentido de giro en cada plano tangente $T_p X$ es el de las bases $\{v_1, v_2\}$ de dicho plano tales que $0 < \det [\nu(p) \mid v_1 \mid v_2] = \nu(p) \cdot (v_1 \times v_2)$.
2. Elegida una orientación \mathcal{O} en X , en cada punto $p \in X$ tomamos $\nu(p) = \frac{v_1 \times v_2}{\|v_1 \times v_2\|}$, siendo $\{v_1, v_2\}$ cualquier base de $T_p X$ cuyo sentido de giro sea el de \mathcal{O} .

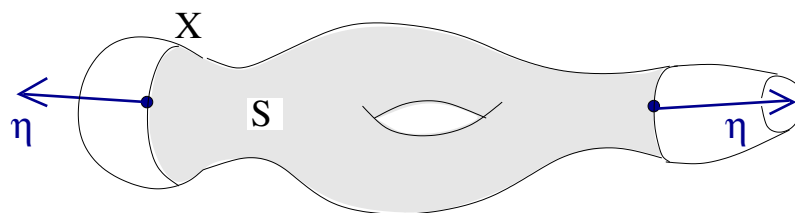
Una función $\varphi(u, v) : R \rightarrow X$ es compatible con la orientación asociada a la normal unitaria ν si y sólo si en todo $(u, v) \in R$ se verifica

$$\det [\nu(\varphi(u, v)) \mid \varphi_u \mid \varphi_v] \geq 0.$$

Pero eso no es todo, cuando $n = 3$ hay un método alternativo para hallar la orientación del borde ∂S que hace verdadera la fórmula de Stokes (137):

En cada una de las curvas Γ_j del borde ∂S , elegimos un punto $p_j \in \Gamma_j$ y determinamos una **conormal exterior** en dicho punto: un vector $\eta_{p_j} \neq \mathbf{0}$ que cumple las tres condiciones siguientes

1. $\eta_{p_j} \in T_{p_j} X$, es decir η_{p_j} es tangente a la superficie en el punto p_j ,
2. η_{p_j} es ortogonal a la recta tangente a Γ_j en p_j ,
3. η_{p_j} **apunta saliendo de S en p_j** .



Entonces el vector

$$\mathbf{t}_{p_j} = \nu(p_j) \times \eta_{p_j}.$$

es tangente a Γ_j en el punto p_j y da a Γ_j la orientación “correcta” para la fórmula de Stokes.

Es recomendable elegir, por cada curva Γ_j , un punto p_j en el que sea especialmente fácil determinar η_{p_j} . No es necesario que η_{p_j} sea un vector unitario, pero debe ser no nulo.

Utilizando el lenguaje del apartado 5.3.5, para $n = 3$ la fórmula (137) se escribe así:

$$\int_{(S, \mathcal{O})} (\mathbf{rot} F)^{\natural} = \int_{\partial S} F^{\flat},$$

lo que traducido al lenguaje del Cálculo Vectorial se convierte en el enunciado clásico de la fórmula clásica de Stokes:

$$\int_{(S, \mathcal{O})} \mathbf{rot} F \cdot d\mathbf{A} = \int_{\partial S} F \cdot d\mathbf{s}.$$

5.4 Cuarto tipo de campos: formas multilineales

5.4.1 Funciones alternadas de grado k

Sea E un espacio vectorial. Escribimos $E^k = \underbrace{E \times \cdots \times E}_{k \text{ factores}}$. Recordemos que una función de k variables vector

$$\phi : E^k \rightarrow \mathbb{R} \quad E^k \ni (v_1, \dots, v_k) \mapsto \phi(v_1, \dots, v_k),$$

es **multilineal** si dado $j \in \{1, \dots, n\}$ y fijados $k-1$ vectores cualesquiera

$$v_1^0, \dots, v_{j-1}^0, v_{j+1}^0, \dots, v_n^0 \in E,$$

la siguiente función es lineal:

$$E \rightarrow \mathbb{R} \quad , \quad E \ni v \mapsto \phi(v_1^0, \dots, v_{j-1}^0, v, v_{j+1}^0, \dots, v_n^0).$$

Definición 164. Sea k un entero positivo. Una función multilineal $\phi : E^k \rightarrow \mathbb{R}$ es **alternada** si se cumple $\phi(v_1, \dots, v_n) = 0$ siempre que en la lista v_1, \dots, v_n haya un vector repetido.

Estas funciones también se llaman **formas alternadas de grado k en E** . El conjunto de todas ellas, para un E dado, se denota $A^k(E)$ y es un espacio vectorial con las operaciones habituales de suma y producto por constante.

En particular, es $A^1(E) = E^*$ y definimos $A^0(E) \stackrel{\text{def}}{=} \mathbb{R}$.

Se deduce de inmediato que el valor de una función de éstas se multiplica por -1 si intercambiamos dos cualesquiera de sus argumentos, dejando los demás intactos:

$$\phi(\dots, w, \dots, v, \dots) = (-1) \phi(\dots, v, \dots, w, \dots).$$

Más en general, si $\sigma : \{1, \dots, k\} \rightarrow \{1, \dots, k\}$ es cualquier permutación de k índices entonces

$$\phi(v_{\sigma(1)}, \dots, v_{\sigma(k)}) = \text{sig } \sigma \cdot \phi(v_1, \dots, v_k).$$

Definición 165. El **producto exterior de varias formas lineales $m_1, \dots, m_k : E \rightarrow \mathbb{R}$** es la forma alternada $m_1 \wedge \cdots \wedge m_k \in A^k(E)$ cuyo valor para $v_1, \dots, v_k \in E$ cualesquiera es

$$(m_1 \wedge \cdots \wedge m_k)(v_1, \dots, v_k) \stackrel{\text{def}}{=} \det(m_i(v_j))_{1 \leq i, j \leq k}. \quad (138)$$

Si $k = 2$, esta definición coincide con la del apartado 5.3.2.

Proposición 166. Sean $\ell_1, \dots, \ell_n : \mathbb{R}^n \rightarrow \mathbb{R}$ las coordenadas estándar, vistas como funciones lineales. Para $k \leq n$ el conjunto $\{\ell_{i_1} \wedge \cdots \wedge \ell_{i_k}\}_{1 \leq i_1 < \cdots < i_k \leq n}$ es una base de $A^k(\mathbb{R}^n)$. Dada cualquier $\phi \in A^k(\mathbb{R}^n)$, los coeficientes únicos $\{c_{i_1 \dots i_k}\}_{1 \leq i_1 < \cdots < i_k \leq n}$ tales que

$$\phi \equiv \sum_{1 \leq i_1 < \cdots < i_k \leq n} c_{i_1 \dots i_k} \ell_{i_1} \wedge \cdots \wedge \ell_{i_k},$$

vienen dados por $c_{i_1 \dots i_k} = \phi(\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_k})$.

Corolario 167. Se tiene $\dim A^k(\mathbb{R}^n) = \begin{cases} \binom{n}{k} & \text{si } k \leq n \\ 0 & \text{si } k > n \end{cases}$

Demostración. La igualdad para $k \leq n$ es consecuencia inmediata de la proposición 166.

La igualdad para $k > n$ se debe a que, si una lista de vectores v_1, \dots, v_k tiene rango $< k$, entonces es $\phi(v_1, \dots, v_k) = 0$ para toda función multilineal alternada ϕ . Por ejemplo, sea ϕ trilineal alternada y supongamos que en la lista v_1, v_2, v_3 el tercer vector es combinación lineal de los dos primeros, entonces

$$\phi(v_1, v_2, v_3) = \phi(v_1, v_2, av_1 + bv_2) = a\phi(v_1, v_2, v_1) + b\phi(v_1, v_2, v_2) = a \cdot 0 + b \cdot 0 = 0.$$

Se sigue que, para $k > n$, el único elemento del espacio $A^k(\mathbb{R}^n)$ es la función idénticamente nula $0 : (\mathbb{R}^n)^k \rightarrow \mathbb{R}$. \square

5.4.2 Álgebra exterior

Teorema 168. Sea E un espacio vectorial. Existe una multiplicación $\cdot \wedge \cdot$ con las siguientes propiedades:

1. Si $\alpha \in A^k(E)$ y $\beta \in A^s(E)$, entonces $\alpha \wedge \beta \in A^{k+s}(E)$.
2. Bilineal:

$$\alpha \wedge (\beta + \beta') = \alpha \wedge \beta + \alpha \wedge \beta' \quad , \quad (\alpha + \alpha') \wedge \beta = \alpha \wedge \beta + \alpha' \wedge \beta \quad , \quad (c\alpha) \wedge \beta = c(\alpha \wedge \beta) = \alpha \wedge (c\beta) .$$
3. Asociativa: si $\alpha \in A^k(E)$, $\beta \in A^r(E)$ y $\gamma \in A^s(E)$, entonces $(\alpha \wedge \beta) \wedge \gamma$ y $\alpha \wedge (\beta \wedge \gamma)$ son el mismo elemento de $A^{k+r+s}(E)$.
4. El producto de formas lineales $m_1 \wedge \cdots \wedge m_k$, definido en el apartado 5.3.2 para $k = 2$ y en la fórmula (138) en general, es un caso particular de esta multiplicación.

Esta operación se llama **producto exterior** o **multiplicación exterior** y convierte en un anillo al espacio vectorial

$$A(E) = \bigoplus_{0 \leq k \leq \dim E} A^k(E) ,$$

el cual es llamado **álgebra exterior de E** . Nótese que $\dim A(E) = 2^{\dim E}$.

Queremos enfatizar que este teorema no es obvio. Omitimos la demostración por falta de tiempo.

La multiplicación exterior, además de existir, es única. Como es bilineal, queda determinada cuando sabemos cómo multiplicar elementos de la forma $m_1 \wedge \cdots \wedge m_r$, con las m_j lineales. Pero, por la asociatividad, sabemos que:

$$(m_1 \wedge \cdots \wedge m_r) \wedge (m_{r+1} \wedge \cdots \wedge m_{r+s}) = m_1 \wedge \cdots \wedge m_{r+s} ,$$

que no sólo demuestra la unicidad del producto exterior, sino que además nos dice cómo calcularlo.

Una consecuencia útil del corolario 167 es que, si $\alpha \in A^k(\mathbb{R}^n)$, $\beta \in A^s(\mathbb{R}^n)$ y $k + s > n$, entonces $\alpha \wedge \beta = 0$.

Los elementos de grado par conmutan con *todos* los elementos:

$$\alpha \in A^{2k}(E) \text{ y } \beta \in A^s(E) \implies \alpha \wedge \beta = \beta \wedge \alpha .$$

Dos elementos de grados impares anticonmutan

$$\alpha \in A^{2r+1}(E) \text{ y } \beta \in A^{2s+1}(E) \implies \alpha \wedge \beta = -\beta \wedge \alpha .$$

5.4.3 Formas diferenciales de grado k

Definición 169. Dada una región $R \subseteq \mathbb{R}^n$, una **k -forma en R** o **forma diferencial de grado k en R** es una aplicación η que asigna a cada punto $p \in R$ una forma alternada $\eta_p \in A^k(\mathbb{R}^n)$. Es, pues, una aplicación

$$\eta : R \longrightarrow A^k(\mathbb{R}^n) ,$$

o sea, un **campo de formas multilineales alternadas de grado k** definido en R . Las **formas de grado 0 en R** son las funciones escalares $R \rightarrow \mathbb{R}$.

La suma de k -formas se define de la manera obvia. También el producto de una función escalar por una k -forma.

El producto exterior de una k -forma η por una s -forma ξ es una $(k+s)$ -forma $\eta \wedge \xi$ en R que se define multiplicando los valores en cada punto:

$$(\eta \wedge \xi)_p = \eta_p \wedge \xi_p \quad \text{para todo } p \in R.$$

Si $k+s > n$, entonces $\eta \wedge \xi$ es idénticamente nula.

Dados campos de vectores $F_1, \dots, F_k : R \rightarrow \mathbb{R}^n$ y una k -forma η en R , tenemos una función escalar $\eta(F_1, \dots, F_k) : U \rightarrow \mathbb{R}$ definida así:

$$\eta(F_1, \dots, F_k)(p) = \eta_p(F_1(p), \dots, F_k(p)) \quad \text{para todo } p \in R.$$

Sean $x_1, \dots, x_n : \mathbb{R}^n \rightarrow \mathbb{R}$ las coordenadas estándar, vistas como funciones diferenciables. Dada una k -forma η en R , existen funciones únicas $\{f_{i_1 \dots i_k}\}_{1 \leq i_1 < \dots < i_k \leq n}$ tales que

$$\eta \equiv \sum_{1 \leq i_1 < \dots < i_k \leq n} f_{i_1 \dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k},$$

y están dadas por $f_{i_1 \dots i_k} \equiv \eta(\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_k})$, donde cada \mathbf{e}_i es visto como un campo constante de vectores.

Un ejemplo de producto exterior en \mathbb{R}^7 :

$$\begin{aligned} & (x_3 dx_1 \wedge dx_2 + e^{x_1} dx_3 \wedge dx_4) \wedge (2 \sin x_4 dx_4 \wedge dx_5 \wedge dx_6 - x_4^2 x_5^3 dx_1 \wedge dx_6 \wedge dx_7) = \\ &= x_3 \cdot 2 \cdot \sin x_4 dx_1 \wedge dx_2 \wedge dx_4 \wedge dx_5 \wedge dx_6 + 0 + 0 - e^{x_1} x_4^2 x_5^3 dx_3 \wedge dx_4 \wedge dx_1 \wedge dx_6 \wedge dx_7 = \\ &= 2x_3 \sin x_4 dx_1 \wedge dx_2 \wedge dx_4 \wedge dx_5 \wedge dx_6 - e^{x_1} x_4^2 x_5^3 dx_1 \wedge dx_3 \wedge dx_4 \wedge dx_6 \wedge dx_7. \end{aligned}$$

5.4.4 Derivada exterior

Las operaciones con formas diferenciales que hemos visto hasta aquí (suma, producto por función, producto exterior, evaluación en campos de vectores) sólo requieren multiplicar y sumar funciones. En este apartado y el siguiente vamos a definir operaciones que requieren derivar funciones.

Definición 170. Sea $U \subseteq \mathbb{R}^n$ un abierto. Dada una k -forma η en U , su **derivada exterior** es una $(k+1)$ -forma $d\eta$ en el mismo abierto U , definida por la siguiente fórmula:

$$d \left(\sum_{1 \leq i_1 < \dots < i_k \leq n} f_{i_1 \dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k} \right) = \sum_{1 \leq i_1 < \dots < i_k \leq n} df_{i_1 \dots i_k} \wedge dx_{i_1} \wedge \dots \wedge dx_{i_k}.$$

También se la llama **diferencial exterior de η** o **coborde de η** .

Si vemos una función escalar $f : U \rightarrow \mathbb{R}$ como una 0-forma, entonces la derivada exterior df coincide con el campo diferencial de la definición 141 del apartado 5.2.2. En el caso en que η es una forma de grado 1, la definición 170 coincide con la 154 del apartado 5.3.4.

Para $k = 0, 1, \dots, n-1$, tenemos un *operador*:

$$d : \{k\text{-formas en } U\} \longrightarrow \{(k+1)\text{-formas en } U\}.$$

Este operador es lineal:

$$d(\eta_1 + \eta_2) = d\eta_1 + d\eta_2 \quad , \quad d(c\eta) = c d\eta,$$

y tiene también la siguiente propiedad:

$$\text{grado}(\eta) = k \implies d(\eta \wedge \xi) = (d\eta) \wedge \xi + (-1)^k \eta \wedge d\xi.$$

En particular, para una función escalar f tenemos

$$d(f\eta) = df \wedge \eta + f d\eta.$$

Ejemplo de derivada exterior en \mathbb{R}^5 :

$$\begin{aligned} & d\left(e^{x_1 x_5} dx_2 \wedge dx_3 \wedge dx_4 + (x_2 - x_3 + x_4^2) dx_1 \wedge dx_3 \wedge dx_4\right) = \\ &= (x_5 e^{x_1 x_5} dx_1 + x_1 e^{x_1 x_5} dx_5) \wedge dx_2 \wedge dx_3 \wedge dx_4 + \\ &+ (dx_2 - dx_3 + 2x_4 dx_4) \wedge dx_1 \wedge dx_3 \wedge dx_4 = \\ &= x_5 e^{x_1 x_5} dx_1 \wedge dx_2 \wedge dx_3 \wedge dx_4 - x_1 e^{x_1 x_5} dx_2 \wedge dx_3 \wedge dx_4 \wedge dx_5 - \\ &- dx_1 \wedge dx_2 \wedge dx_3 \wedge dx_4 = \\ &= (x_5 e^{x_1 x_5} - 1) dx_1 \wedge dx_2 \wedge dx_3 \wedge dx_4 - x_1 e^{x_1 x_5} dx_2 \wedge dx_3 \wedge dx_4 \wedge dx_5. \end{aligned}$$

La derivada exterior satisface una identidad muy notable:

$$\boxed{d \circ d \equiv 0} \quad (139)$$

es decir, al hacer la derivada exterior dos veces se obtiene siempre cero:

$$d(d\eta) = 0 \quad \text{para cualquier forma diferencial } \eta \in \mathcal{C}^2.$$

El caso particular para 0-formas (funciones escalares $f : U \rightarrow \mathbb{R}$) es la fórmula (121) del apartado 5.3.4.

5.4.5 Formas diferenciales traídas (pullback)

En este apartado vamos a definir una operación que actúa sobre una pareja forma-aplicación y que requiere derivar la aplicación (no la forma diferencial).

Empezamos con un abierto $U \subseteq \mathbb{R}^n$, un subconjunto $E \subseteq \mathbb{R}^s$ y una aplicación diferenciable φ , con dominio U y valores en E , que representamos por una flecha que apunta de U a E :

$$U \xrightarrow{\varphi} E.$$

Entonces (atención al sentido de las flechas):

1. φ *empuja* los puntos de U a puntos de E : $U \ni p \mapsto \varphi(p) \in E$.
2. φ *empuja* los pares punto-vector de U a E : $U \times \mathbb{R}^n \ni (p, v) \mapsto (\varphi(p), (d\varphi)_p(v))$.
3. φ *trae* las funciones (0-formas) de E a U : $\boxed{\varphi^* h \stackrel{\text{def}}{=} h \circ \varphi : U \rightarrow \mathbb{R}} \longleftarrow \boxed{h : E \rightarrow \mathbb{R}}$

Es decir, definimos la **función traída** $\varphi^* h$ así:

$$(\varphi^* h)(p) \stackrel{\text{def}}{=} h(\varphi(p)) = h(\text{punto empujado}), \quad \text{para todo punto } p \in U.$$

4. φ *trae* las formas de Pfaff de E a U : $\varphi^* \omega \longleftarrow \omega$. Dada una forma de Pfaff ω definida en E , la **traída por** φ es una forma de Pfaff $\varphi^* \omega$ en U que se define así:

$$(\varphi^* \omega)_p(v) \stackrel{\text{def}}{=} \omega_{\varphi(p)}((d\varphi)_p v) = \omega_{\text{punto empujado}}(\text{vector empujado}),$$

para cualesquiera $p \in U$ y $v \in \mathbb{R}^n$.

5. φ *trae* las k -formas de E a U : $\varphi^* \eta \longleftarrow \eta$. Dada una k -forma η en E , la **forma traída por** φ es una k -forma $\varphi^* \eta$ en U que se define así:

$$\begin{aligned} (\varphi^* \eta)_p(v_1, \dots, v_k) & \stackrel{\text{def}}{=} \eta_{\varphi(p)}((d\varphi)_p(v_1), (d\varphi)_p(v_2), \dots, (d\varphi)_p(v_k)) = \\ &= \eta_{\text{punto empujado}}(\text{vectores empujados}), \end{aligned}$$

para cualesquiera $p \in U$ y $v_1, \dots, v_k \in \mathbb{R}^n$.

Observa que se usa $(d\varphi)_p$ en las construcciones 2., 4. y 5., luego estamos derivando φ .
Tenemos unos operadores

$$\{k\text{-formas en } U\} \xleftarrow{\varphi^*} \{k\text{-formas en } E\} \quad , \quad \text{para } k = 0, 1, 2, \dots$$

Estos operadores reciben, todos, el nombre de **pullback** (en inglés “traer hacia atrás”).

Las propiedades obvias del pullback son:

1. Conserva el grado: si η es una k -forma, entonces $\varphi^*\eta$ también es una k -forma.
2. Lineal: $\varphi^*(\eta_1 + \eta_2) = \varphi^*\eta_1 + \varphi^*\eta_2$, $\varphi^*(c\eta) = c\varphi^*\eta$.

Pero tiene, además, unas propiedades notables.

Primera. También conserva el producto exterior:

$$\varphi^*(\eta \wedge \xi) = \varphi^*\eta \wedge \varphi^*\xi .$$

En particular $\varphi^*(h\eta) = (h \circ \varphi) \varphi^*\eta$, cuando h es una función escalar (una 0-forma).

Segunda. Dados abiertos $U \subseteq \mathbb{R}^n$, $V \subseteq \mathbb{R}^r$, $W \subseteq \mathbb{R}^s$ y aplicaciones diferenciables

$$U \xrightarrow{\varphi} V \xrightarrow{\psi} W ,$$

se tiene:

$$(\psi \circ \varphi)^*\eta = \varphi^*\psi^*\eta \quad , \quad \text{para toda } k\text{-forma } \eta \text{ en } W ,$$

es decir que tenemos la siguiente identidad entre operadores:

$$\boxed{(\psi \circ \varphi)^* = \varphi^* \circ \psi^*} \quad (140)$$

Observa que *se invierte el orden de composición*. Esto no es nuevo, ya lo habíamos visto con el producto de matrices y la trasposición: $(AB)^t = B^t A^t$. En Matemáticas, las transformaciones entre aplicaciones que invierten el orden de composición se denotan siempre por **exponentes**, o sea símbolos puestos arriba y a la derecha del nombre de la aplicación a transformar.

Tercera. Si $\varphi : U \rightarrow V$ es al menos de clase \mathcal{C}^2 , entonces

$$\varphi^* d\eta = d(\varphi^*\eta) \quad , \quad \text{para toda forma } \eta \text{ en } V ,$$

es decir que se cumple la siguiente identidad entre operadores:

$$\boxed{\varphi^* \circ d = d \circ \varphi^*} . \quad (141)$$

Es fácil calcular explícitamente $\varphi^*\eta$ utilizando esas propiedades. Un ejemplo:

$$\begin{aligned} \varphi : \mathbb{R}^3 &\longrightarrow \mathbb{R}^4 \quad , \quad \varphi(u, v, w) \equiv (\varphi_1, \varphi_2, \varphi_3, \varphi_4) = \left(v + u^2, e^w, uv^3, \frac{1}{1+w^2} \right) , \\ \varphi^* \left((e^{x_1} + x_4) dx_1 \wedge dx_2 \wedge dx_3 - 7x_2^2 dx_1 \wedge dx_3 \wedge dx_4 \right) &= \\ = (e^{x_1 \circ \varphi} + x_4 \circ \varphi) \varphi^*(dx_1) \wedge \varphi^*(dx_2) \wedge \varphi^*(dx_3) - 7(x_2 \circ \varphi)^2 \varphi^*(dx_1) \wedge \varphi^*(dx_3) \wedge \varphi^*(dx_4) &= \\ = (e^{\varphi_1} + \varphi_4) d(x_1 \circ \varphi) \wedge d(x_2 \circ \varphi) \wedge d(x_3 \circ \varphi) - 7\varphi_2^2 d(x_1 \circ \varphi) \wedge d(x_3 \circ \varphi) \wedge d(x_4 \circ \varphi) &= \\ = \left(e^{v+u^2} + \frac{1}{1+w^2} \right) d(v+u^2) \wedge de^w \wedge d(uv^3) - 7(e^w)^2 d(v+u^2) \wedge d(uv^3) \wedge d\frac{1}{1+w^2} , \end{aligned}$$

el resto del cálculo se deja como ejercicio.

5.4.6 Pullback e integración de formas

Tenemos una región acotada $R \subset \mathbb{R}^k$ y una función $\varphi(u_1, \dots, u_k) : R \rightarrow \mathbb{R}^n$ que es restricción a R de alguna función diferenciable en un abierto de \mathbb{R}^k conteniendo a R .

Definición 171. Si η es cualquier k -forma definida al menos en el conjunto $\varphi(R)$, entonces la **integral de η sobre φ** es el número dado por la siguiente fórmula

$$\int_{\varphi} \eta = \int_{u \in R} \eta_{\varphi(u)}(\varphi_{u_1}, \dots, \varphi_{u_k}) du_1 \cdots du_k, \quad (142)$$

enteramente análoga a la (116) del apartado 5.2.7 y a la (130) del apartado 5.3.7.

Si η es una k -forma definida al menos en R , definimos

$$\int_R \eta = \int_{id_R} \eta = \int_{x \in R} \eta_x(\mathbf{e}_1, \dots, \mathbf{e}_k) dx_1 \cdots dx_k.$$

Por ejemplo, si $k = 3$ y η viene escrita como $\eta = f dx_1 \wedge dx_2 \wedge dx_3$, entonces

$$\int_R \eta = \iiint_R f dx_1 dx_2 dx_3.$$

Si η viene escrita como $\eta = g dx_2 \wedge dx_1 \wedge dx_3$, entonces $\eta(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = -g$ y tenemos

$$\int_R \eta = \iiint_R -g dx_1 dx_2 dx_3.$$

Es inmediato observar que (142) equivale a

$$\boxed{\int_{\varphi} \eta = \int_R \varphi^* \eta} \quad (143)$$

fórmula que en cálculos concretos es más fácil de utilizar que la (142).

Esto arroja nueva luz sobre la fórmula de Stokes paramétrica (133). Para una función de dos variables $\varphi(u, v) : R \rightarrow \mathbb{R}^n$ y una forma de Pfaff ω , se verifican:

$$\int_{\partial \varphi} \omega = \int_{\partial R} \varphi^* \omega, \quad \int_{\varphi} d\omega = \int_R \varphi^*(d\omega) = \int_R d\varphi^* \omega,$$

donde ∂R es el borde orientado de la región R . Ahora $\mu = \varphi^* \omega$ es una forma de Pfaff en el plano y la fórmula (133) resulta ser lo mismo que

$$\int_{\partial R} \mu = \int_R d\mu. \quad (144)$$

Pero escribiendo $\mu = P du + Q dv$ la fórmula (144) se traduce en:

$$\int_{\partial R} (P du + Q dv) = \iint_R (Q_u - P_v) dudv,$$

que no es otra cosa que la fórmula de Green en el plano.

Las propiedades (143) y (141) del pullback permiten reducir la fórmula de Stokes paramétrica (133) a la fórmula de Green en el plano.

Habrà a quién le guste más esta demostración de la fórmula (133) que la que hicimos en el apartado 5.3.9. Creemos que es más instructivo conocer ambas demostraciones, porque ya dijimos que las fórmulas (127) y (129), en las que se basa la primera demostración, exponen el verdadero significado de la derivada exterior.

5.4.7 Formas cerradas y formas exactas

Definición 172. Sea ω una k -forma en un abierto $U \subseteq \mathbb{R}^n$. Sea $V \subseteq U$ también abierto. Decimos que ω es **cerrada en V** si $d\omega \equiv 0$ en V . Decimos que es **cerrada** si $d\omega \equiv 0$ en U .

Una **primitiva exterior de ω** , o **antiderivada exterior de ω** , es (si es que existe) una $(k-1)$ -forma η tal que $\omega = d\eta$. Decimos que ω es **exacta en V** si tiene una primitiva exterior en V . Decimos que es **exacta** si tiene una primitiva exterior en U .

En abiertos de \mathbb{R}^n toda n -forma ω es cerrada, porque $d\omega$ es nula al ser de grado $n+1$, y de hecho una tal ω siempre es exacta (resultado no trivial, que no vamos a demostrar aquí).

Debido a la fórmula (139) del apartado 5.4.4, para que ω sea exacta en *necesario* que sea cerrada. En ciertos abiertos es también suficiente, pero en otros no lo es.

Veamos un ejemplo de un abierto con una 2-forma cerrada que no es exacta. El **radio esférico** es la función

$$\rho : \mathbb{R}^n \longrightarrow \mathbb{R} \quad , \quad \rho(x) = \|x\|_2 = \sqrt{x \cdot x} \quad ,$$

que es \mathcal{C}^∞ en $\mathbb{R}^n \setminus \{\mathbf{0}\}$. Nos centramos en el caso $n = 3$; en esta dimensión, es sabido que el **campo gravitatorio**

$$F = \frac{\nabla \rho}{\rho^2} = \frac{(x, y, z)}{\rho^3} \quad ,$$

es \mathcal{C}^∞ en $U = \mathbb{R}^3 \setminus \{\mathbf{0}\}$ y tiene divergencia idénticamente nula en U .

Es fácil comprobar que el tercer isomorfismo musical, que se definió en la fórmula (125) del apartado 5.3.5, cumple la siguiente identidad

$$d(G^\sharp) = (\operatorname{div} G) dx \wedge dy \wedge dz \quad . \quad (145)$$

Como el campo gravitatorio F tiene divergencia nula, la correspondiente 2-forma

$$\Omega = F^\sharp = \frac{x dy \wedge dz + y dz \wedge dx + z dx \wedge dy}{\rho^3} \quad ,$$

es cerrada en U . Además es exacta en los siguientes abiertos, más pequeños que U :

$$U_1 = \mathbb{R}^3 \setminus (\{(0,0)\} \times [0, +\infty)) \quad , \quad U_2 = \mathbb{R}^3 \setminus (\{(0,0)\} \times (-\infty, 0]) \quad ,$$

porque se comprueba que:

$$\Omega = d \frac{y dx - x dy}{\rho(\rho - z)} \quad \text{en } U_1 \quad \text{y} \quad \Omega = d \frac{-y dx + x dy}{\rho(\rho + z)} \quad \text{en } U_2 \quad .$$

Afirmamos que Ω no es exacta en U . Más aún: no es exacta en ningún abierto de \mathbb{R}^3 que contenga una esfera S centrada en el origen $(0,0,0)$. Esto es así porque en cualquiera de estas esferas, orientadas por la normal exterior, se tiene

$$\int_S \Omega = 4\pi \quad ,$$

integral que tendría que haber sido nula si Ω hubiese sido exacta en un entorno de S , según el siguiente resultado.

Lema 173. Sea η una forma de Pfaff diferenciable en un abierto de \mathbb{R}^3 que contiene una esfera S . Entonces $\int_S d\eta = 0$ (da igual la orientación de S).

Es decir que las 2-formas exactas en dimensión 3 tienen integral nula en toda esfera. Compárese con la condición (3) del teorema 150 del apartado 5.2.7.

Demostración. Utilizando coordenadas esféricas

$$u = \text{longitud geográfica} \quad , \quad v = \text{latitud geográfica} \quad ,$$

construimos una función

$$\varphi(u, v) : [0, 2\pi] \times \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \longrightarrow \mathbb{R}^3 \quad ,$$

cumpliendo las condiciones (a), (b) y (c) del apartado 5.3.10 para, por ejemplo, la orientación de S por la normal exterior. Así orientada S , sabemos que

$$\int_S d\eta = \int_{\varphi} d\eta = \int_{\partial\varphi} \eta \quad ,$$

donde $\partial\varphi$ (ver la definición 160 del apartado 5.3.9) es en este caso un lazo formado por cuatro caminos:

- un camino con valor constante el polo sur de S ,
- un camino que recorre el *meridiano cero* de S de abajo a arriba,
- un camino con valor constante el polo norte de S ,
- un camino que recorre el meridiano cero de S de arriba a abajo.

Dos de esos caminos contribuyen cero a la integral $\int_{\partial\varphi} \eta$, mientras que los otros dos producen contribuciones opuestas la una de la otra. La suma de las cuatro contribuciones es cero. \square

Corolario 174. *El campo gravitatorio F no tiene ningún potencial vector en $\mathbb{R}^3 \setminus \{\mathbf{0}\}$.*

Si tuviera un tal potencial A , entonces A^\flat sería una primitiva exterior para F^\flat .

5.4.8 Generalizaciones

El concepto de deformación suave, del apartado 5.3.6, se generaliza de caminos a funciones $\varphi : R \rightarrow \mathbb{R}^n$ con $R \subset \mathbb{R}^k$ región acotada, de manera que una deformación suave de φ viene dada por un intervalo $J \subseteq \mathbb{R}$ y una función

$$\bar{\varphi}(u_1, \dots, u_k, \lambda) : R \times J \longrightarrow \mathbb{R}^n \quad ,$$

tal que $\bar{\varphi}(u_1, \dots, u_k, \lambda_0) \equiv \varphi(u_1, \dots, u_k)$ para cierto valor $\lambda_0 \in J$. El concepto de deformaciones de extremos fijos, también del apartado 5.3.6, se generaliza al concepto de **deformaciones de borde fijo**, que son las que cumplen

$$\bar{\varphi}(u, \lambda) = \varphi(u) \quad , \quad \text{para todo } (u, \lambda) \in \partial R \times J \quad ,$$

y éstas cumplen el análogo de la fórmula (129):

$$\left. \frac{d}{d\lambda} \right|_{\lambda=\lambda_0} \int_{\bar{\varphi}(\cdot, \lambda)} \eta = \int_{u \in R} (d\eta)_{\varphi(u)}(v(u), \varphi_{u_1}, \dots, \varphi_{u_k}) du_1 \cdots du_k \quad ,$$

donde $v(u) : R \rightarrow \mathbb{R}^n$, la “velocidad lateral de deformación”, viene dada por $v(u) = \frac{\partial \bar{\varphi}}{\partial \lambda}(u, \lambda_0)$. Esto explica el significado de la derivada exterior de formas de cualquier grado:

$d\eta$ nos dice cuánto varía la integral de η cuando agitamos la “película”, sobre la cual estamos integrando η , sin mover el borde de esa película.

También resulta de la generalización de (129) una fórmula de Stokes idéntica a la (133), con básicamente la misma demostración que hicimos en el apartado 5.3.9, pero para formas de cualquier grado y no sólo formas de Pfaff.

Los conceptos de orientación y de subvariedad orientable, definidos para objetos bidimensionales en los apartados 5.3.10 y 5.3.11, existen para subvariedades de cualquier dimensión k y también son verdad las correspondientes fórmulas de Stokes para formas de grado $k - 1$.

También el lema 173 tiene su análogo para formas de grados mayores, de donde resulta que para grados $k = 1, 2, \dots, n - 1$ hay abiertos en \mathbb{R}^n con k -formas cerradas que no son exactas.

5.4.9 Propiedades locales y no locales

Una propiedad es **local** si cuando se satisface en pequeños dominios también se satisface en el dominio total. Ejemplos de propiedades locales:

1. Ser función continua.
2. Ser función de clase \mathcal{C}^ℓ .
3. Ser subconjunto abierto.
4. Tener jacobiana invertible.
5. Ser forma diferencial cerrada.

Comprobar cualquiera de estas propiedades es relativamente sencillo, porque es suficiente hacer la comprobación local (en pequeños dominios).

Mucho más difícil es comprobar si se cumple o no una **propiedad no local**, es decir una propiedad que, aunque se cumpla en pequeños dominios, puede no cumplirse en el dominio total. Ejemplos de propiedades no locales:

1. Ser función acotada.
2. Ser conjunto convexo.
3. Ser conjunto conexo por caminos.
4. Ser función inyectiva (esto nos dió bastante trabajo en el capítulo 3).
5. Ser forma diferencial exacta.

Proposición 175. *Una forma de cualquier grado positivo η , suave en un abierto $U \subseteq \mathbb{R}^n$, es cerrada si y sólo si es **localmente exacta**, que significa que cada punto $p \in U$ tiene un entorno $U^p \subseteq U$ tal que $\eta|_{U^p}$ es exacta.*

Que ser localmente exacta es suficiente para ser cerrada, se deduce inmediatamente de la identidad (139) del apartado 5.4.4.

Que ser localmente exacta es condición necesaria es lo que afirma el siguiente resultado, que previamente habíamos visto para formas de Pfaff en el apartado 5.2.6.

Teorema 176. (Lema de Poincaré). *Una forma cerrada de grado positivo, definida en todo \mathbb{R}^n o en una bola de \mathbb{R}^n , es exacta en dicho dominio.*

Demostramos el lema de Poincaré en el apartado 5.4.10.

En dimensión n la propiedad “ser forma exacta” es no local para grado $k = 1, 2, \dots, n-1$, porque hay abiertos en \mathbb{R}^n con alguna forma de ese grado que es cerrada (localmente exacta) pero no exacta.

5.4.10 Cálculo de una primitiva exterior

En este apartado demostramos el lema 176 de manera constructiva: dado U , que es una bola o todo \mathbb{R}^n , y $k = 1, 2, \dots, n$, vamos a describir un método para calcular explícitamente una primitiva exterior para cualquier k -forma ω que sea cerrada en U .

Primer paso. De todo el desarrollo

$$\omega \equiv \sum_{1 \leq i_1 < \dots < i_k \leq n} a_{i_1 \dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k},$$

separamos los términos que contienen dx_1 y, si hace falta, los reordenamos para que el factor dx_1 quede a la izquierda, de modo que ω quede escrita así:

$$\omega = dx_1 \wedge \left(\sum_{\mathbf{2} \leq i_2 < \dots < i_k \leq n} a_{1i_2 \dots i_k} dx_{i_2} \wedge \dots \wedge dx_{i_k} \right) + (\text{términos sin } dx_1) .$$

Entonces calculamos funciones *suaves* $\tilde{a}_{1i_2 \dots i_k}$ tales que:

$$\frac{\partial \tilde{a}_{1i_2 \dots i_k}}{\partial x_1} = a_{1i_2 \dots i_k} ,$$

lo cual es posible porque U es una bola o todo el espacio. Con estas funciones, formamos la $(k-1)$ -forma:

$$\eta_1 = \sum_{\mathbf{2} \leq i_2 < \dots < i_k \leq n} \tilde{a}_{1i_2 \dots i_k} dx_{i_2} \wedge dx_{i_3} \wedge \dots \wedge dx_{i_k} ,$$

y consideramos la diferencia $\omega_1 = \omega - d\eta_1$ que, siendo diferencia de una cerrada menos una exacta, todavía es cerrada. Los términos que contienen dx_1 son, por construcción, los mismos en ω y en $d\eta_1$; luego la diferencia no contiene dx_1 :

$$\omega_1 = \omega - d\eta_1 = \sum_{\mathbf{2} \leq i_1 < \dots < i_k \leq n} b_{i_1 \dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k} .$$

Segundo paso. Afirmamos que, gracias a que ω es cerrada y cómo es U , la forma ω_1 *no contiene x_1 en ningún sitio*. Es decir que sus funciones coeficiente no dependen de x_1 :

$$b_{i_1 \dots i_k}(x_2, \dots, x_n) . \quad (146)$$

Como $d\omega_1 = 0$, dados $\mathbf{2} \leq i_1 < \dots < i_k \leq n$, el término (función) $\cdot dx_1 \wedge dx_{i_1} \wedge \dots \wedge dx_{i_k}$ del desarrollo de $d\omega_1$ es nulo. Pero a este término sólo contribuyen las derivadas de los términos $b_{j_1 \dots j_k} dx_{j_1} \wedge \dots \wedge dx_{j_k}$ de ω_1 que cumplen $\{j_1, \dots, j_k\} \subset \{1, i_1, \dots, i_k\}$, que son:

1. Los términos función $\cdot dx_1 \wedge (dx_{i_1} \wedge \dots \wedge dx_{i_{h-1}} \wedge dx_{i_{h+1}} \wedge \dots \wedge dx_{i_k})$, todos ellos idénticamente nulos por cómo hemos construido ω_1 ,
2. y el $b_{i_1 \dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k}$.

Por lo tanto, sólo contribuye la derivada del término $b_{i_1 \dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k}$, de donde:

$$0 \equiv \frac{\partial b_{i_1 \dots i_k}}{\partial x_1} dx_1 \wedge dx_{i_1} \wedge \dots \wedge dx_{i_k} .$$

Esto prueba que $\frac{\partial b_{i_1 \dots i_k}}{\partial x_1} \equiv 0$ para cada función coeficiente $b_{i_1 \dots i_k}$ del desarrollo de ω_1 . Por cómo es el abierto U , esto implica que se escriben como en (146) y llegamos a que:

$$\omega_1 = \sum_{\mathbf{2} \leq i_1 < \dots < i_k \leq n} b_{i_1 \dots i_k}(x_2, \dots, x_n) dx_{i_1} \wedge \dots \wedge dx_{i_k} .$$

Tercer paso. Ahora tomamos los términos del desarrollo de ω_1 que contienen dx_2 , poniendo este factor a la izquierda en cada término, e integramos sus funciones coeficiente respecto de x_2 , *pero cuidando de que las integrales indefinidas no dependan de x_1* , y formamos con esas nuevas funciones una $(k-1)$ -forma η_2 en U . Esta forma diferencial η_2 no contiene x_1 en ningún sitio y, por lo tanto, tampoco lo hace $d\eta_2$. Además la diferencia

$$\omega_2 = \omega_1 - d\eta_2 = \omega - d\eta_1 - d\eta_2 ,$$

tampoco contiene x_2 en ningún sitio y es también cerrada.

Paso general. Partiendo de ω_r , cerrada y que no contiene x_1, \dots, x_r en ningún sitio, hallamos una $(k-1)$ -forma η_{r+1} tal que $\omega_{r+1} = \omega_r - d\eta_{r+1}$ no contiene x_1, \dots, x_r, x_{r+1} en ningún sitio y, siendo diferencia de una cerrada menos una exacta, también es cerrada.

Paso final. Cuando hayamos hecho $n-k$ pasos, tenemos ω_{n-k} que solamente contiene las k últimas coordenadas x_{n-k+1}, \dots, x_n y se desarrolla *con un solo término*:

$$\omega_{n-k} = \omega - d\eta_1 - d\eta_2 - \dots - d\eta_{n-k} = c(x_{n-k+1}, \dots, x_n) \cdot dx_{n-k+1} \wedge \dots \wedge dx_n.$$

Tomamos una integral indefinida de la función coeficiente c :

$$\frac{\partial \tilde{c}}{\partial x_{n-k+1}} = c,$$

cuidando que \tilde{c} no dependa de x_1, \dots, x_{n-k} . Hacemos

$$\eta_{n-k+1} = \tilde{c} \cdot dx_{n-k+2} \wedge \dots \wedge dx_n,$$

y conseguimos $\omega_{n-k} = d\eta_{n-k+1}$. Finalmente:

$$\omega = d(\eta_1 + \dots + \eta_{n-k} + \eta_{n-k+1}).$$

Ejemplo 1. Una 2-forma en \mathbb{R}^3 (comprueba que $d\omega \equiv 0$):

$$\omega \equiv (e^{yz} - x) dx \wedge dy + e^z dx \wedge dz + (2yz - xy e^{yz} + 1) dy \wedge dz.$$

Primer paso. Separamos los términos que contienen dx :

$$(e^{yz} - x) dx \wedge dy \quad \text{y} \quad e^z dx \wedge dz,$$

en cada uno quitamos el factor dx (asegurándonos primero de que esté a la izquierda) y reemplazamos la función coeficiente por una integral indefinida suya respecto de x :

$$\eta_1 = \left(x e^{yz} - \frac{x^2}{2} \right) dy + x e^z dz.$$

Calculamos la derivada exterior de η_1 :

$$d\eta_1 = (e^{yz} - x) dx \wedge dy + e^z dx \wedge dz - xy e^{yz} dy \wedge dz,$$

y la diferencia $\omega_1 = \omega - d\eta_1$ ya no contiene x en ningún sitio:

$$\omega_1 = \omega - d\eta_1 = (2yz + 1) dy \wedge dz.$$

Sabemos por la teoría que esto se debe a que ω es cerrada: si hacemos lo mismo con una 2-forma no cerrada, la diferencia también tendrá nulos los coeficientes de $dx \wedge dy$ y de $dx \wedge dz$, pero el coeficiente de $dy \wedge dz$ en dicha diferencia contendrá la variable x .

Segundo paso. Al ser ω_1 un monomio con sólo las variables y, z , es fácil ponerlo como una derivada exterior:

$$(2yz + 1) dy \wedge dz = d((y^2 z + y) dz),$$

y finalmente:

$$\omega = d\eta \quad \text{con} \quad \eta = \eta_1 + (y^2 z + y) dz = \left(x e^{yz} - \frac{x^2}{2} \right) dy + (x e^z + y^2 z + y) dz.$$

Ejemplo 2. Una 3-forma en \mathbb{R}^4 (Comprueba que $d\omega \equiv 0$):

$$\omega \equiv (x_2 + 5x_3 x_4) dx_1 \wedge dx_2 \wedge dx_3 + (1 + x_3^2) dx_1 \wedge dx_2 \wedge dx_4 + (3x_1 x_3 - e^{x_3}) dx_2 \wedge dx_3 \wedge dx_4.$$

Primer paso. Separamos los términos que contienen dx_1 y nos aseguramos de que el factor dx_1 esté a la izquierda:

$$(x_2 + 5x_3x_4) dx_1 \wedge dx_2 \wedge dx_3 \quad \text{y} \quad (1 + x_3^2) dx_1 \wedge dx_2 \wedge dx_4 .$$

En cada término quitamos el factor dx_1 e integramos la función coeficiente respecto de x_1 :

$$\eta_1 = (x_1x_2 + 5x_1x_3x_4) dx_2 \wedge dx_3 + (x_1 + x_1x_3^2) dx_2 \wedge dx_4 .$$

Hallamos $d\eta_1$ y la diferencia:

$$\begin{aligned} d\eta_1 &= (x_2 + 5x_3x_4) dx_1 \wedge dx_2 \wedge dx_3 + (1 + x_3^2) dx_1 \wedge dx_2 \wedge dx_4 + \\ &\quad + 5x_1x_3 dx_4 \wedge dx_2 \wedge dx_3 + 2x_1x_3 dx_3 \wedge dx_2 \wedge dx_4 = \\ &= (x_2 + 5x_3x_4) dx_1 \wedge dx_2 \wedge dx_3 + (1 + x_3^2) dx_1 \wedge dx_2 \wedge dx_4 + 3x_1x_3 dx_2 \wedge dx_3 \wedge dx_4 , \\ \omega - d\eta_1 &= -e^{x_3} dx_2 \wedge dx_3 \wedge dx_4 . \end{aligned}$$

Segundo paso. Como $\omega - d\eta_1$ es un monomio conteniendo solamente las variables x_2, x_3, x_4 , es fácil ponerlo como una derivada exterior:

$$-e^{x_3} dx_2 \wedge dx_3 \wedge dx_4 = d(e^{x_3} dx_2 \wedge dx_4) ,$$

luego $\omega = d(\eta_1 + e^{x_3} dx_2 \wedge dx_4)$, es decir:

$$\omega = d\eta \quad \text{con} \quad \eta = (x_1x_2 + 5x_1x_3x_4) dx_2 \wedge dx_3 + (x_1 + x_1x_3^2 + e^{x_3}) dx_2 \wedge dx_4 .$$

5.5 Teorema de la divergencia

5.5.1 Tercer isomorfismo musical

Este isomorfismo ya lo hemos definido para dimensión 3 en el apartado 5.3.5. Ahora lo vamos a definir para todas las dimensiones.

Definición 177. Dada una región $R \subseteq \mathbb{R}^n$ y en ella un campo de velocidades $V : R \rightarrow \mathbb{R}^n$, definimos una $(n-1)$ -forma diferencial V^\natural en R de la manera siguiente. Para todo punto $p \in R$ y cualesquiera vectores $v_1, \dots, v_{n-1} \in \mathbb{R}^n$, hacemos:

$$(V^\natural)_p(v_1, \dots, v_{n-1}) \stackrel{\text{def}}{=} \det [V(p) \mid v_1 \mid \dots \mid v_{n-1}] .$$

Sea $V = (V_1, \dots, V_n)$. Desarrollando el determinante por la primera columna, en coordenadas estándar se tiene:

$$V^\natural = \sum_{j=1}^n (-1)^{j-1} V_j dx_1 \wedge \dots \wedge \widehat{dx_j} \wedge \dots \wedge dx_n ,$$

donde el circunflejo encima de una expresión significa que dicha expresión ha sido suprimida. En el caso particular $n = 3$, es habitual escribir esa igualdad así:

$$V^\natural = V_1 dy \wedge dz + V_2 dz \wedge dx + V_3 dx \wedge dy .$$

Es fácil comprobar que

$$d(V^\natural) = (\text{div } V) dx_1 \wedge \dots \wedge dx_n . \quad (147)$$

5.5.2 Flujo a través del borde

Definición 178. Sea $R \subset \mathbb{R}^n$ una región compacta con borde ∂R . Dado un campo de velocidades V definido en cada punto de ∂R , el **flujo a través del borde de R** de dicho campo es el número

$$\int_{\partial R} V^\natural, \quad (148)$$

donde el borde ∂R ha sido orientado por la **normal exterior a R** , un campo de vectores no nulo que en cada punto $p \in \partial R$ es normal a ∂R y apunta hacia afuera de R .

Es importante no confundir el flujo a través del borde, que es un número, con el flujo definido en el apartado 5.1.1, que es una familia de difeomorfismos.

La integral (148) se define tomando una región acotada $R_0 \subset \mathbb{R}^{n-1}$ y una función $\varphi : R_0 \rightarrow \mathbb{R}^n$ cumpliendo las siguientes propiedades

- (a) la imagen $\varphi(R_0)$ coincide con ∂R ,
- (b) las preimágenes múltiples (aquellos $u \in R_0$ para los que existe otro $u' \in R_0$ distinto pero con $\varphi(u)$ y $\varphi(u')$ iguales) forman un conjunto de medida $(n-1)$ -dimensional nula,
- (c) φ es **compatible con la orientación de ∂R definida por la normal exterior ν** , que quiere decir que hay un conjunto $E \subset \mathbb{R}^{n-1}$, de medida $(n-1)$ -dimensional nula, tal que para todo $u \in R_0 \setminus E$ se tiene $\det [\nu(\varphi(u)) \mid \varphi_{u_1} \mid \cdots \mid \varphi_{u_{n-1}}] > 0$.

y entonces $\int_{\partial R} V^\natural = \int_{\varphi} V^\natural$, siendo el resultado independiente de la elección de φ mientras cumpla (a), (b) y (c).

5.5.3 Enunciado y demostración del teorema

Teorema 179. Sea V un campo suave de velocidades definido en un abierto $U \subseteq \mathbb{R}^n$. Si $R \subset U$ es una región compacta, con borde ∂R orientado por la normal exterior, se tiene:

$$\int_{\partial R} V^\natural = \int_R (\operatorname{div} V) dx_1 \cdots dx_n. \quad (149)$$

Es decir, el flujo a través del borde coincide con la integral de la divergencia en el interior.

Vamos a demostrar el teorema suponiendo que el borde ∂R es una subvariedad, haciendo un breve comentario al final sobre regiones más generales. En particular, suponemos la normal exterior ν bien definida en todo punto del borde. No vamos a dar todos los detalles técnicos, omitiendo los más pesados.

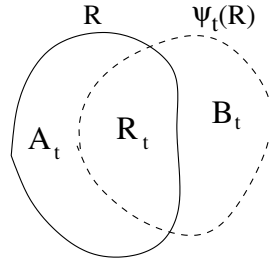
Escribiremos $m_n(A)$ para la medida n -dimensional de un conjunto $A \subset \mathbb{R}^n$ y $m_{n-1}(E)$ para la medida $(n-1)$ -dimensional de un conjunto $E \subset \mathbb{R}^{n-1}$.

Denotaremos por $\{\psi_s\}$ el flujo de V . Definimos dos subconjuntos del borde $\partial^- R, \partial^+ R \subset \partial R$ de la siguiente manera:

1. Un punto $p \in \partial R$ está en $\partial^- R$ si hay un $s_p > 0$ tal que $\psi_s(p)$ es un punto interior a R para $0 < s < s_0$. En estos puntos es $V(p) \cdot \nu(p) \leq 0$.
2. Un punto $p \in \partial R$ está en $\partial^+ R$ si hay un $s_p > 0$ tal que $\psi_s(p) \notin R$ para $0 < s < s_0$. En estos puntos es $V(p) \cdot \nu(p) \geq 0$.

Para cada $t > 0$ definimos las siguientes regiones

$$R_t = \psi_t(R) \cap R, \quad A_t = R \setminus \psi_t(R), \quad B_t = \psi_t(R) \setminus R.$$



Puesto que $\psi_t(R) = R_t \sqcup B_t$, $A_t \subseteq R$ y $R_t = R \setminus A_t$, tenemos

$$\begin{aligned} m_n(\psi_t(r)) &= m_n(R_t) + m_n(B_t) , \\ m_n(R_t) &= m_n(R) - m_n(A_t) , \end{aligned}$$

de donde

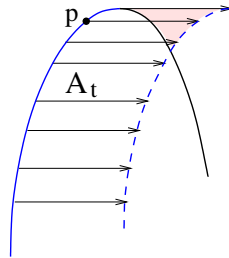
$$\begin{aligned} m_n(\psi_t(R)) &= m_n(R) - m_n(A_t) + m_n(B_t) , \\ m_n(\psi_t(R)) - m_n(R) &= -m_n(A_t) + m_n(B_t) . \end{aligned} \quad (150)$$

Fijados un valor $t > 0$ y un trozo de borde $Z \subset \partial R$, llamamos **región barrida por Z** al conjunto

$$\{ \psi_s(p) : p \in Z , 0 \leq s \leq t \} .$$

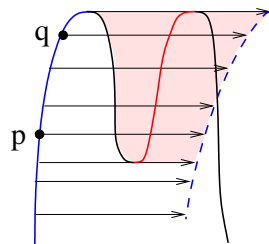
Cuando un trozo $Z \subseteq \partial^- R$ se desplaza por el flujo de V , primero entra dentro de R y la región que va barriendo ya no va a formar parte de $\psi_t(R)$. Esta parte barrida por Z contribuye, pues, a la región A_t que hay que quitar de R para obtener R_t . Pero *no toda la región barrida por Z debe quitarse de R*. Veamos dos ejemplos visuales, en los que utilizamos flechas de línea fina para representar los tramos de trayectoria $\{ \psi_s(p) : 0 \leq s \leq t \}$.

La primera figura muestra (en azul) un trozo $Z \subset \partial^- R$ y un punto particular $p \in Z$ que, al moverse por el flujo de V , primero entra en R y después sale y se aleja de R . La región barrida por Z es la formada por las flechas de línea fina y tiene una parte (en color rosa) formada por las posiciones de los puntos móviles cuando ya han salido de R . Esta parte rosa no contribuye al conjunto A_t .



Dicho de otra manera: un punto móvil que comience en $\partial^- R$ y sigue el flujo de V contribuye al conjunto A_t hasta la primera vez que sale de R , después ya no contribuye.

La segunda figura muestra otro trozo $Z \subset \partial^- R$ también en color azul. Vemos puntos $p \in Z$ que primero entran, salen y vuelven a entrar en R , y puntos $q \in Z$ que entran, salen, vuelven a entrar y vuelven a salir de R .



La parte rosa no debemos incluirla en A_t . Ni siquiera las posiciones cuando el punto móvil vuelve a entrar en R , porque de eso ya se encarga el trozo rojo de $\partial^- R$. Una vez más vemos que, en cuanto un punto móvil que arranca de $\partial^- R$ sale de R por primera vez, ya no debemos seguir usándolo en la construcción de A_t .

Fijado $t > 0$, para cada punto $p \in \partial^- R$ definimos un número $t_p \in (0, t]$ de la manera siguiente:

$$t_p = \max \{ \tau \in (0, t] : \psi_s(p) \text{ es interior a } R \text{ para todo } s \in (0, \tau) \} ,$$

es decir que t_p es el máximo tiempo (sin exceder t) que un punto móvil permanece sin salirse de R si empieza en p y se mueve por el flujo de V .

La discusión en los dos ejemplos visuales anteriores muestra que se tiene $\psi_s(p) \in A_t$ mientras sea $p \in \partial^- R$ y $s \in [0, t_p)$. De hecho se demuestra (no lo vamos a hacer aquí) que todo punto de A_t se construye así a partir de algún punto de $\partial^- R$, es decir que se cumple la siguiente igualdad conjuntista:

$$A_t = \{ \psi_s(p) : p \in \partial^- R, 0 \leq s < t_p \} . \quad (151)$$

Análogamente, para cada $p \in \partial^+ R$ definimos

$$t'_p = \max \{ \tau \in (0, t] : \psi_s(p) \notin R \text{ para todo } s \in (0, \tau) \} ,$$

es decir que t'_p es el máximo tiempo (sin exceder t) que un punto móvil permanece fuera de R si empieza en p y se mueve por el flujo de V . Se vuelve a cumplir:

$$B_t = \{ \psi_s(p) : p \in \partial^+ R, 0 < s \leq t'_p \} . \quad (152)$$

Los conjuntos $\partial^- R$ y $\partial^+ R$ no agotan ∂R . Para un punto $p \in \partial R \setminus (\partial^- R \cup \partial^+ R)$ hay cuatro posibilidades:

1. Hay un $\tau > 0$ tal que $\psi_s(p)$ permanece en ∂R para $0 \leq s \leq \tau$.
2. Hay una sucesión decreciente $\{t_j\}_{j=1}^\infty$, con límite 0, tal que $\psi_{t_j}(p) \in \partial R$ para j impar y $\psi_{t_j}(p)$ es interior a R para j par.
3. Hay una sucesión decreciente $\{t_j\}_{j=1}^\infty$, con límite 0, tal que $\psi_{t_j}(p) \in \partial R$ para j impar y $\psi_{t_j}(p) \notin R$ para j par.
4. Hay una sucesión decreciente $\{t_j\}_{j=1}^\infty$, con límite 0, tal que $\psi_{t_j}(p)$ es interior a R para j impar y $\psi_{t_j}(p) \notin R$ para j par.

En cualquiera de esos cuatro casos, se tiene $V(p) \cdot \nu(p) = 0$.

Suponemos elegidas una región $R_0 \subset \mathbb{R}^{n-1}$ y una función φ con un conjunto $E \subset R_0$ cumpliendo las condiciones (a), (b) y (c) del apartado 5.5.2. Entonces el flujo a través del borde viene dado por

$$\int_{\partial R} V^\natural = \int_{u \in R_0} \det [V(\varphi(u)) \mid \varphi_{u_1} \mid \cdots \mid \varphi_{u_{n-1}}] du_1 \cdots du_{n-1} . \quad (153)$$

Vamos a analizar el signo del integrando $g(u) = \det [V(\varphi(u)) \mid \varphi_{u_1} \mid \cdots \mid \varphi_{u_{n-1}}]$ en (153). Para cada $u \in R_0 \setminus E$ la siguiente lista de vectores es una base de \mathbb{R}^n

$$\{ \nu(\varphi(u)), \varphi_{u_1}(u), \dots, \varphi_{u_{n-1}}(u) \} ,$$

y nos permite escribir el vector $V(\varphi(u))$ como combinación lineal

$$V(\varphi(u)) = (V \cdot \nu)_{\varphi(u)} \nu(\varphi(u)) + a_1(u) \varphi_{u_1}(u) + \cdots + a_{n-1}(u) \varphi_{u_{n-1}}(u) ,$$

de donde $g(u) = (V \cdot \nu)_{\varphi(u)} \cdot \det [\nu(\varphi(u)) \mid \varphi_{u_1}(u) \mid \cdots \mid \varphi_{u_{n-1}}(u)]$ es un múltiplo positivo de $(V \cdot \nu)_{\varphi(u)}$. Deducimos que si $\varphi(u) \in \partial^- R$ entonces $g(u) \leq 0$, si $\varphi(u) \in \partial^+ R$ entonces $g(u) \geq 0$ y si $\varphi(u) \in \partial R \setminus (\partial^- R \cup \partial^+ R)$ entonces $g(u) = 0$.

Definiendo $R_0^- = \varphi^{-1}(\partial^- R)$, $R_0^+ = \varphi^{-1}(\partial^+ R)$ y la función $\Psi(u, s) = \psi_s(\varphi(u))$, las fórmulas (151) y (152) se pueden escribir así:

$$\begin{aligned} A_t &= \Psi(\{ (u, s) : u \in R_0^-, 0 \leq s < t_{\varphi(u)} \}) , \\ B_t &= \Psi(\{ (u, s) : u \in R_0^+, 0 < s \leq t'_{\varphi(u)} \}) . \end{aligned}$$

Para cada $u \in R_0$ tenemos

$$\begin{aligned} \Psi_{u_j}(u, 0) &= \varphi_{u_j}(u) \text{ para } j = 1, \dots, n-1 , \\ \Psi_s(u, 0) &= V(\varphi(u)) , \\ |\det(D\Psi)_{(u,0)}| &= \left| \det [\varphi_{u_1} \mid \dots \mid \varphi_{u_{n-1}} \mid V(\varphi(u))] \right| = |g(u)| , \end{aligned}$$

de donde, para $s \geq 0$ se cumple:

$$|\det(D\Psi)_{(u,s)}| = |g(u)| + O(s) . \quad (154)$$

Las fórmulas (150) y (154), junto con el estudio que hemos hecho del signo de $g(u)$, nos permiten hacer el siguiente cálculo para $t > 0$:

$$\begin{aligned} m_n(\psi_t(R)) - m_n(R) &= -m_n(A_t) + m_n(B_t) = \\ &= - \int_{u \in R_0^-} \left(\int_{0 \leq s < t_{\varphi(u)}} |\det(D\Psi)_{(u,s)}| ds \right) du_1 \cdots du_{n-1} + \\ &\quad + \int_{u \in R_0^+} \left(\int_{0 < s \leq t'_{\varphi(u)}} |\det(D\Psi)_{(u,s)}| ds \right) du_1 \cdots du_{n-1} = \\ &= \int_{u \in R_0^-} \left(\int_{0 \leq s < t_{\varphi(u)}} (-|g(u)| + O(s)) ds \right) du_1 \cdots du_{n-1} + \\ &\quad + \int_{u \in R_0^+} \left(\int_{0 < s \leq t'_{\varphi(u)}} (|g(u)| + O(s)) ds \right) du_1 \cdots du_{n-1} = \\ &= O(t^2) + \int_{u \in R_0^-} t_{\varphi(u)} g(u) du_1 \cdots du_{n-1} + \\ &\quad + \int_{u \in R_0^+} t'_{\varphi(u)} g(u) du_1 \cdots du_{n-1} = \\ &= t \cdot \int_{u \in R_0^- \cup R_0^+} g(u) du_1 \cdots du_{n-1} + a + O(t^2) . \end{aligned}$$

donde a es la siguiente cantidad:

$$a = - \int_{u \in R_0^-} (t - t_{\varphi(u)}) g(u) du_1 \cdots du_{n-1} - \int_{u \in R_0^+} (t - t'_{\varphi(u)}) g(u) du_1 \cdots du_{n-1} .$$

Como $g \equiv 0$ en $R_0 \setminus (R_0^- \cup R_0^+)$, y recordando la fórmula (153), llegamos a:

$$m_n(\psi_t(R)) - m_n(R) = t \cdot \int_{R_0} g(u) du_1 \cdots du_{n-1} + a + O(t^2) = t \cdot \int_{\partial R} V^\natural + a + O(t^2) . \quad (155)$$

Para acotar a , consideramos la constante $C = \max_{u \in R_0} |g(u)|$ y definimos los conjuntos

$$E(t) = \{ u \in R_0^- : t_{\varphi(u)} < t \} , \quad E'(t) = \{ u \in R_0^+ : t'_{\varphi(u)} < t \} ,$$

entonces:

$$|a| \leq C \cdot t \cdot \left[m_{n-1}(E(t)) + m_{n-1}(E'(t)) \right] . \quad (156)$$

El conjunto $E(t)$ encoge al disminuir t , es decir que si $t_1 < t_2$ entonces $E(t_1) \subseteq E(t_2)$, con el siguiente límite conjuntista

$$\bigcap_{t \searrow 0} E(t) = \emptyset ,$$

y el **teorema de la convergencia monótona**, que se estudia en un curso superior a éste, nos dice que

$$\lim_{t \searrow 0} m_{n-1}(E(t)) = 0 . \quad (157)$$

La imagen $\varphi(E(t))$ es el conjunto de los $p \in \partial^- R$ que, siguiendo el flujo de V , se salen de R antes de que transcurran t unidades de tiempo. Es decir, son los puntos “problemáticos” que hemos detectado antes con ayuda de dos dibujos. Entonces la fórmula (157) nos dice que, a medida que t desciende a cero, estos puntos tienen cada vez menos importancia sobre el cálculo de las integrales.

Del mismo modo se ve que $\lim_{t \searrow 0} m_{n-1}(E'(t)) = 0$. Ahora la fórmula (156) nos dice que $a = o(t)$, lo cual convierte la fórmula (155) en

$$t > 0 \implies m_n(\psi_t(R)) - m_n(R) = t \cdot \int_{\partial R} V^\natural + o(t) \quad (158)$$

El siguiente cálculo

$$\left. \frac{d}{dt} \right|_{t=0} m_n(\psi_t(r)) = \lim_{t \searrow 0} \frac{m_n(\psi_t(R)) - m_n(R)}{t} = \int_{\partial R} V^\natural + \lim_{t \searrow 0} \frac{o(t)}{t} ,$$

nos lleva finalmente a:

$$\left. \frac{d}{dt} \right|_{t=0} m_n(\psi_t(r)) = \int_{\partial R} V^\natural . \quad (159)$$

La fórmula (149) es consecuencia de combinar (159) con la fórmula (100) del apartado 5.1.2.

Esta demostración se extiende a otras regiones cuyo borde ∂R está formado por varias piezas, cada una contenida en una hipersuperficie de \mathbb{R}^n , como suelen aparecer en las colecciones de problemas o en las aplicaciones técnicas.

Por ejemplo, el borde del rectángulo $[x_1, x_2] \times [y_1, y_2]$ está formado por cuatro segmentos, cada uno contenido en un recta. La normal exterior está definida en todo el borde del rectángulo excepto en las cuatro *esquinas*, lo cual no afecta a las integrales que se han de considerar y todo vuelve a funcionar (teniendo un poco más de cuidado que en el argumento precedente).

Las fórmulas (150) y (158) nos dicen cuál es el verdadero significado del teorema de la divergencia:

Si la región está llena con un gas, entonces la variación del volumen que ocupa el gas es igual a lo que está saliendo de la región *a través del borde de la misma* menos lo que está entrando, también a través del borde.

Eso es igual a una integral en el interior de la región, porque otra manera de hallar la variación de volumen es sumar las contracciones o expansiones de pequeñas “celdas” en toda la región. Expansiones y contracciones que son medidas por la divergencia, tal como hemos explicado en los apartados 5.1.2 y 5.1.4.

5.5.4 Fórmula de Stokes y teorema de la divergencia

La fórmula (147) del apartado 5.5.1 permite reescribir el teorema de la divergencia así:

$$\int_{\partial R} V^\natural = \int_R d(V^\natural) ,$$

con lo cual es un caso particular de la fórmula de Stokes. Pero no hemos querido conformarnos con esta manera de demostrarlo, porque se habría perdido la interpretación del significado del teorema de la divergencia.

Veamos, con un ejemplo, cómo se puede deducir la fórmula de Stokes a partir del teorema de la divergencia. Empezamos con un abierto $U \subseteq \mathbb{R}^7$, y una 3-forma η definida en U . Tenemos también una región acotada $R \subset \mathbb{R}^4$ y una función $\varphi(u_1, u_2, u_3, u_4) : R \rightarrow U$ diferenciable. En principio, hay funciones $f_{i_1 i_2 i_3} : U \rightarrow \mathbb{R}$ tales que

$$\eta \equiv \sum_{1 \leq i_1 < i_2 < i_3 \leq 7} f_{i_1 i_2 i_3} dx_{i_1} \wedge dx_{i_2} \wedge dx_{i_3} .$$

En cada punto $x \in U$ la forma trilineal η_x es una suma de $\binom{7}{3} = 35$ sumandos. Tiene una enorme complejidad algebraica, equiparable a la de un polinomio de varias variables. Al hacer el pullback por φ se pierde esa complejidad, porque bastan cuatro funciones $g_1, g_2, g_3, g_4 : R \rightarrow \mathbb{R}$ para tener el desarrollo

$$\varphi^* \eta \equiv \sum_{j=1}^4 g_j(u_1, u_2, u_3, u_4) du_1 \wedge \cdots \widehat{du_j} \cdots \wedge du_4 ,$$

y definiendo el siguiente campo de velocidades en la región R

$$V = (g_1, -g_2, g_3, -g_4) ,$$

tenemos $\varphi^* \eta = V^\flat$. El campo V todavía tiene su complejidad, pero ésta es mucho menor que la complejidad original de η .

Utilizando la fórmula (143) del apartado 5.4.6 y la fórmula (141) del apartado 5.4.5, deducimos:

$$\begin{aligned} \int_{\varphi} d\eta &= \int_R \varphi^*(d\eta) = \int_R d\varphi^* \eta = \int_R d(V^\flat) = \\ &= \int_R (\operatorname{div} V) dx_1 dx_2 dx_3 dx_4 = \int_{\partial R} V^\flat = \\ &= \int_{\partial R} \varphi^* \eta = \int_{\partial \varphi} \eta . \end{aligned}$$

Hemos demostrado la fórmula de Stokes $\int_{\varphi} d\eta = \int_{\partial \varphi} \eta$, para 3-formas en dimensión 7, a partir del teorema de la divergencia en regiones de \mathbb{R}^4 .

Pero sería injusto decir que la fórmula de Stokes es una consecuencia directa del teorema de la divergencia: han intervenido las identidades (143) y (141). Lo que hemos puesto de manifiesto es lo inmensamente potentes que son estas dos identidades.

Esta demostración de la fórmula de Stokes no era posible en el Primer Curso, donde sólo se consideraban campos de vectores y éstos *no se pueden traer hacia atrás*: no tenemos operación de pullback para ellos.

5.6 Campos de Beltrami

Llamamos **campos de Beltrami** a los campos en abiertos de \mathbb{R}^3 que son **autofunciones** del operador rotacional, es decir campos F tales que $\mathbf{rot} F = \lambda \cdot F$ para alguna constante λ . Si $\lambda \neq 0$, eso fuerza $\operatorname{div} F \equiv 0$. En el caso $\lambda = 0$ se exige, además, que F tenga divergencia nula para considerarlo un campo de Beltrami. Son muy importantes en Física.

Toda constante es autovalor de \mathbf{rot} , por ejemplo $F = (\cos(\lambda z), -\cos(\lambda x) - \operatorname{sen}(\lambda z), \operatorname{sen}(\lambda x))$ es una autofunción con autovalor asociado λ .

La derivada exterior $d : \{\text{formas de Pfaff}\} \rightarrow \{2\text{-formas}\}$ no puede tener autofunciones, porque va de un espacio vectorial a otro muy distinto. Aunque en las páginas anteriores la hemos relacionado estrechamente con el rotacional, lo cierto es que estos dos operadores tienen su propio universo cada uno y, después del “encuentro” en este capítulo, se separan bastante.