Análisis Matemático

Miguel González
mgonzalez.contacto@gmail.com

Enero de 2020

$$\det(df)_{x_0} \neq 0 \implies \exists f^{-1} \in \mathcal{C}^1(B(f(x_0), \epsilon))$$

Revisado en 2022

Acerca de este documento

Estos apuntes son una versión revisada de los de la asignatura Análisis Matemático del grado en matemáticas, tomados en Enero de 2020 por Miguel González. A los apuntes originales se les ha añadido esta página, una imagen de portada, y breves párrafos explicativos en las zonas menos completas. Asimismo se han revisado las erratas y completado los contenidos faltantes.

Este documento es:

- Una recopilación ordenada y directa de las definiciones y resultados más importantes del tema en cuestión, al nivel de los estudios de grado.
- Una colección de demostraciones completas de dichos resultados (salvo en los casos más básicos).
- Una *guía* para revisar de manera rápida las ideas que se han adquirido previamente, o para consultar enunciados puntuales que puedan no haberse comprendido en su totalidad.

Este documento NO es:

- Un libro de texto de la asignatura.
- Una colección de ejercicios para practicar los conceptos adquiridos.
- Un listado de ejemplos para ilustrar las ideas tratadas. A pesar de ello, en ocasiones se incluyen ejemplos puntuales que puedan ser de especial interés o curiosidad, pero se intentan reducir al mínimo en virtud del primer punto de la lista anterior.

Sobre Análisis Matemático

En esta asignatura se presentan las herramientas avanzadas del análisis matemático de manera rigurosa. Se estudia, por un lado, el espacio euclídeo y las funciones sobre este, culminando en los teoremas de la función inversa y la función implícita. Después, se hace una introducción a la teoría de variedades diferenciables inmersas en el espacio euclídeo y se introducen las formas diferenciales, conectando el análisis tradicional con la geometría diferencial.

Requisitos previos

1. Familiaridad con la notación matemática básica.
2. Conocimientos de cálculo equivalentes a la asignatura de Cálculo I. (Sucesiones, funciones reales de una variable, límites, derivadas, integración).
3. Conocimientos de cálculo multivariable (asignatura de Cálculo II).
4. Conocimientos de álgebra lineal.

Índice

1. Preliminares

1.1. Espacios euclídeos

Definición 1. Una **forma** es una aplicación que va sobre el cuerpo de un espacio vectorial, es decir, si V es un $\mathbb{K}$-espacio, es una aplicación del tipo $\phi : V^n \mapsto \mathbb{K}$.

Definición 2. Una forma $\phi : V \times V \mapsto \mathbb{K}$ es **bilineal** si es lineal en cada componente, esto es, si $\forall v \in V$, las formas $\phi_v : V \mapsto \mathbb{K}$ dada por $\phi_v(w) = \phi(v, w)$; y $\phi'_v : V \mapsto \mathbb{K}$ dada por $\phi'_v(w) = \phi(w, v)$ son lineales.

Definición 3. Una forma bilineal $\phi : V \times V \mapsto \mathbb{K}$ es **simétrica** si $\forall u, v \in V$, se tiene $\phi(u, v) = \phi(v, u)$.

Definición 4. Una forma bilineal o sesquilineal $\phi : V \times V \mapsto \mathbb{K}$ es **definida positiva** si $\forall u \in V$, se tiene $\phi(u, u) \geq 0$, con $\phi(u, u) = 0 \iff u = 0$.

Definición 5. Un **espacio euclídeo** es un $\mathbb{R}$-espacio vectorial V junto con una forma bilineal simétrica definida positiva, el **producto escalar**, $\phi : V^2 \mapsto \mathbb{R}$. Se denota $E = (V, \phi)$. Se suele denotar $\phi(x, y) = \langle x, y \rangle$.

El producto escalar permite definir una aplicación sobre cada elemento del $\mathbb{R}$-espacio vectorial, que da una idea de ”tamaño”:

Definición 6. En un espacio euclídeo, la **norma inducida** por su producto escalar ϕ viene definida como $\|\|_\phi : V \mapsto \mathbb{R}$ dada por $\|v\|_\phi = \sqrt{\phi(v, v)}$.

Esta aplicación verifica las siguientes propiedades:

Proposición 1. *En un espacio euclídeo $(V, \langle \cdot, \cdot \rangle)$, con $\|\cdot\|$ la norma inducida, se tiene:*

1. $\forall x \in V$, $\|x\| \geq 0$ *y* $\|x\| = 0 \iff x = 0$.
2. $\forall x \in V, \lambda \in \mathbb{R}$, *se tiene* $\|\lambda x\| = |\lambda| \, \|x\|$.
3. $\forall x, y \in V$, *se tiene* $| \langle x, y \rangle | \leq \|x\| \, \|y\|$. *(Cauchy-Schwarz)*
4. $\forall x, y \in V$, *se tiene* $\|x + y\| \leq \|x\| + \|y\|$.

Demostración. 1 sigue de ser el producto escalar definido positivo. 2 es evidente dado que $(\langle \lambda x, \lambda x \rangle)^{\frac{1}{2}} = (\lambda^2 \langle x, x \rangle)^{\frac{1}{2}} = |\lambda| (\langle x, x \rangle)^{\frac{1}{2}}$, por bilinealidad. Para 3, basta con considerar $0 \leq \varphi(t) = \|x + ty\|^2 = \langle x, x \rangle + 2t \langle x, y \rangle + t^2 \langle y, y \rangle$. Para que este polinomio en t se mantenga no negativo, es necesaria la siguiente condición en su discriminante:

$$(2 \langle x, y \rangle)^2 - 4 \langle x, x \rangle \langle y, y \rangle \leq 0$$

Y reordenando se tiene la desigualdad deseada. Para 4, veamos que $\|x + y\|^2 = \langle x, x \rangle^2 + 2 \langle x, y \rangle + \langle y, y \rangle^2 \leq \langle x, x \rangle^2 + 2 \langle x, x \rangle \langle y, y \rangle + \langle y, y \rangle^2 = (\|x\| + \|y\|)^2$, y tomando raíces cuadradas se tiene. Se ha usado la desigualdad de Schwarz. □

1.2. Espacios normados

Con el fin de explorar aplicaciones de otros tipos que actúen con las mismas propiedades que la norma inducida por un producto escalar, se define:

Definición 7. Un espacio vectorial real E es **normado** si existe una función $\|\cdot\| : E \mapsto \mathbb{R}$, la **norma**, que verifique:

1. $\forall x \in E$, $\|x\| \geq 0$ y $\|x\| = 0 \iff x = 0$.
2. $\forall x \in E, \lambda \in \mathbb{R}$, se tiene $\|\lambda x\| = |\lambda| \, \|x\|$.
3. $\forall x, y \in E$, se tiene $\|x+y\| \leq \|x\| + \|y\|$.

Observación 1. Si relajamos la condición 1 pidiendo únicamente que $\|x\| \geq 0$, obtenemos lo que se conoce como una **seminorma**. Si consideramos $V = \{x \in E : \|x\| = 0\}$, y definimos en E/V la aplicación $\widetilde{\|\cdot\|}$, dada por $\|\widetilde{[x]}\| = \|x\|$, estará bien definida y mantendrá las propiedades de una norma completa. (Al partir por V colapsamos todos los puntos problemáticos en la clase del origen).

Definición 8. En $\mathbb{R}^n$, fijado un $p \geq 1$, se define la p-norma $\|\cdot\|_p : \mathbb{R}^n \mapsto \mathbb{R}$ por $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$. x_i, como es habitual, denota la i-ésima componente del vector x.

Asimismo, se define la ∞-norma por $\|x\|_\infty = \text{máx}_{1 \leq i \leq n}\{|x_i|\}$.

A continuación comprobaremos que esta norma verifica efectivamente las 3 propiedades requeridas.

Proposición 2. *Dado cualquier p con $1 \leq p \leq \infty$, la p-norma correspondiente es una norma.*

Demostración. Para 2, tenemos $\|\lambda x\|_p = (\sum_{i=1}^n |\lambda x_i|^p)^{\frac{1}{p}} = (|\lambda|^p \sum_{i=1}^n |x_i|^p)^{\frac{1}{p}} = |\lambda| \, \|x\|_p$, y además $\|\lambda x\|_\infty = \text{máx}_{1 \leq i \leq n}\{|\lambda x_i|\} = |\lambda| \, \text{máx}_{1 \leq i \leq n}\{|x_i|\}$. Para 1, es evidente que la p-norma nunca puede ser negativa y que $\|0\|_p = 0$ en cualquier caso. Ahora, si $\|x\|_p = 0$, y hay algún j tal que $x_j \neq 0$, entonces $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}} \geq (|x_j|^p)^{\frac{1}{p}} > 0$, contradiciendo el supuesto. Por tanto, $x = 0$. Lo mismo sucede si $\|x\|_\infty = 0$, dado que si $x_j \neq 0$, entonces $\|x\|_\infty \geq |x_j| > 0$.

Para 3, es fácil ver que en los casos $p = 1$ y $p = \infty$, la desigualdad triangular del valor absoluto garantiza la propiedad. En efecto: $\|x+y\|_1 = \sum_i |x_i + y_i| \leq \sum_i (|x_i| + |y_i|) = \sum_i |x_i| + \sum_i |y_i| = \|x\|_1 + \|y\|_1$, y $\|x+y\|_\infty = \text{máx}\{|x_i + y_i|\} \leq \text{máx}\{|x_i| + |y_i|\} = |x_{i_0}| + |y_{i_0}| \leq \text{máx}\{|x_i|\} + \text{máx}\{|y_i|\}$, como se quería.

Para demostrar la desigualdad con otros valores de p, es necesario el siguiente resultado:

Lema 1 (Young)**.** *Si $a, b > 0$ y $1 \leq p < \infty$, y definimos $p' = \frac{p}{p-1}$, de tal modo que $\frac{1}{p} + \frac{1}{p'} = 1$, entonces:*

$$ab \leq \frac{a^p}{p} + \frac{b^{p'}}{p'}$$

Demostración del lema. Como la función $t \mapsto \log(t)$ es cóncava, tenemos que $\log(\frac{1}{p}a^p + \frac{1}{p'}b^{p'}) \geq \frac{1}{p}\log a^p + \frac{1}{p'}\log b^{p'} = \log a + \log b = \log ab$, dado que tanto $\frac{1}{p}$ como $\frac{1}{p'}$ están entre 0 y 1. Como el logaritmo es creciente, se tiene lo que se quería. □

Ahora podemos continuar demostrando la desigualdad triangular. Para ello, necesitamos otro lema:

Lema 2 (Hölder)**.** *Si $x, y \in \mathbb{R}^n$ y $1 \leq p \leq \infty$, y definimos $p' = \frac{p}{p-1}$, de tal modo que también $1 \leq p' \leq \infty$ y $\frac{1}{p} + \frac{1}{p'} = 1$, entonces se tiene:*

$$|\langle x, y \rangle| \leq \|x\|_p \|y\|_{p'}$$

Donde $\langle x, y \rangle$ es el producto usual. Cabe destacar que si $p = 2$, se trata de la desigualdad de Schwarz.

Demostración del lema. En primer lugar, conviene descartar algunos casos triviales. Si $x = 0$ o $y = 0$, se tiene igualdad $0 = 0$. Asimismo, si $p = 1$ y $p' = \infty$, se tiene $|\langle x, y \rangle| = |\sum_1^n x_i y_i| \leq \sum_1^n |x_i y_i| \leq \sum_1^n |x_i| \, \|y\|_\infty = \|y\|_\infty \|x\|_1$, y se tiene lo que se buscaba.

Para probar el caso no trivial $1 < p < \infty$, con $x, y \neq 0$, aplicaremos la desigualdad de Young para ver que:

$$|\langle x, y\rangle| \leq \sum_{i=1}^{n} |x_i y_i| \leq \sum_{i=1}^{n} \frac{|x_i|^p}{p} + \sum_{i=1}^{n} \frac{|y_i|^{p'}}{p'} = \frac{1}{p}\|x\|_p^p + \frac{1}{p'}\|y\|_{p'}^{p'}$$

Ahora, como $\langle x, y\rangle = \left\langle tx, \frac{1}{t}y\right\rangle$, siempre que $t > 0$, podemos reescribir lo anterior como:

$$|\langle x, y\rangle| \leq \frac{1}{p}t^p\|x\|_p^p + \frac{1}{p'}t^{-p}\|y\|_{p'}^{p'} := \varphi(t)$$

Ahora, sea t_0 aquel tal que $t_0^p\|x\|_p^p = t_0^{-p'}\|y\|_{p'}^{p'} := \alpha$. Es inmediato encontrar un valor explícito para t_0 despejando, aunque por el momento no importa. En ese caso, tenemos:

$$|\langle x, y\rangle| \leq \frac{1}{p}\alpha + \frac{1}{p'}\alpha = \alpha$$

Ahora tan solo debemos obtener ese valor. Como $t_0 = \left(\frac{\|y\|_{p'}^{p'}}{\|x\|_p^p}\right)^{\frac{1}{p+p'}}$, entonces $\alpha = \left(\frac{\|y\|_{p'}^{p'}}{\|x\|_p^p}\right)^{\frac{p}{p+p'}} \cdot \|x\|_p^p = \frac{\|y\|_{p'}^{\frac{p'p}{p+p'}}}{\|x\|_p^{\frac{p^2}{p+p'}}}\|x\|_p^p$. La única observación que nos separa ahora del resultado es que $\frac{p'p}{p'+p} = \left(\frac{1}{p}+\frac{1}{p'}\right)^{-1} = 1$, así como que $\frac{p^2}{p+p'} = \frac{pp'}{p+p'}\frac{p}{p'} = p-1$. De este modo, $\alpha = \|x\|_p\|y\|_{p'}$ y se tiene lo deseado. □

Finalmente, con esta desigualdad podemos probar la triangular (También denominada *de Minkowski*). Una vez más, asumimos $x, y \neq 0$ dada la trivialidad del caso contrario. Tomamos $p \in (1, \infty)$. Comenzamos como sigue:

$$\|x+y\|_p^p = \sum_{i=1}^{n}|x_i+y_i|^p = \sum_{i=1}^{n}|x_i+y_i||x_i+y_i|^{p-1} \leq \sum_{i=1}^{n}|x_i||x_i+y_i|^{p-1} + \sum_{i=1}^{n}|y_i||x_i+y_i|^{p-1}$$

Ahora, observemos que cada suma puede interpretarse como un producto escalar entre dos vectores en particular. Por ejemplo, la primera, entre el vector $(|x_i|)_1^n$ y el $(|x_i+y_i|^{p-1})_1^n$. Con esto en mente, aplicamos la desigualdad de Hölder:

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

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

Ya solo queda encontrar desarrollar el segundo factor. Como $p' = \frac{p}{p-1}$, tenemos:

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

$$= \left[\left(\sum_{i=1}^{n}|x_i+y_i|^p\right)^{\frac{1}{p}}\right]^{\frac{p}{p'}} = \left(\|x+y\|_p\right)^{\frac{p'}{p}}$$

Y como $\frac{p'}{p} = p-1$ hemos terminado. (Basta con combinar los 3 pasos que hemos realizado, y dividir por $\|x+y\|_p^{p-1}$, no nula por hipótesis.) □

1.3. Convexidad

Definición 9. Si E es un $\mathbb{R}$-espacio, se dice que $\Omega \subset E$ es **convexo** si $\forall x, y \in \Omega$, se tiene que $[x,y] \subset \Omega$, donde $[x,y] = \{tx + (1-t)y : t \in [0,1]\}$ es el segmento que une los puntos x e y.

Definición 10 (Bola abierta)**.** Si $(E, \|\cdot\|)$ es un espacio vectorial normado, dado $x_0 \in E$ y $r \in \mathbb{R}$, $r > 0$, se llama **bola abierta** de centro x_0 y radio r, y se denota $B_r(x_0)$, al conjunto: $B_r(x_0) = \{x \in E : \|x - x_0\| < r\}$.

Definición 11 (Bola cerrada)**.** Si $(E, \|\cdot\|)$ es un espacio vectorial normado, dado $x_0 \in E$ y $r \in \mathbb{R}$, $r > 0$, se llama **bola cerrada** de centro x_0 y radio r, y se denota $\overline{B}_r(x_0)$, al conjunto: $\overline{B}_r(x_0) = \{x \in E : \|x - x_0\| \leq r\}$.

Proposición 3. *Cualquier bola abierta o cerrada es convexa.*

Demostración. Vamos a hacerlo para el caso cerrado, siendo análogo para el abierto. Si $y, z \in \overline{B}_r(x_0)$, entonces, si $t \in [0,1]$: $\|[(1-t)y + tz] - x\| = \|(1-t)(y-x) + t(z-x)\| \leq (1-t)\|y-x\| + t\|z-x\| \leq (1-t)r + tr = r$. □

Definición 12. Si E es un $\mathbb{R}$-espacio, se dice que $\Omega \subset E$ es **simétrico** si $\forall x \in \Omega$, $-x \in \Omega$.

Definición 13. Si E es un $\mathbb{R}$-espacio normado, se dice que $\Omega \subset E$ es **absorbente** si $\forall x \in E$, $\exists t > 0$ tal que $x \in t\Omega = \{tu : u \in \Omega\}$.

Es decir, es absorbente si escalándolo se puede abarcar cualquier punto deseado de E.

Proposición 4. *Se tiene que $B_1(0)$ y $\overline{B}_1(0)$ son absorbentes y simétricas en todo espacio normado.*

Demostración. Como $\forall x \in E$, se tiene $\|x\| = \|-x\|$, la simetría es evidente. Para la absorción, lo probaremos para $B_1(0)$, dado que $\overline{B}_1(0)$ lo contiene y por tanto abarca más aún. Si $x = 0$ hemos acabado dado que $0 \in B_1(0)$. En caso contrario, veamos que $\frac{x}{\|x\|} \in B_2(0)$ al ser unitario. Es decir, que $x \in \|x\| B_2(0) = 2\|x\| B_1(0)$ y por tanto es absorbente. □

A continuación describiremos una nueva norma en términos de conjuntos absorbentes simétricos. Antes, requeriremos del siguiente resultado intuitivo:

Lema 3. *Si $\exists t_0 > 0$ tal que $x \in t_0\Omega$, donde Ω es absorbente, simétrico y convexo, entonces $\forall t > t_0$ se tiene $x \in t\Omega$.*

Demostración. Por hipótesis, $x = t_0\omega$ con $\omega \in \Omega$. Por tanto, dado $t > t_0$, se tiene $x = t \cdot \left(\frac{t_0}{t}\omega\right)$. Por hipótesis se tiene que $\frac{t_0}{t}\omega \in [0, \omega]$, y $0 \in \Omega$ al ser convexo simétrico (dado que $0 \in [-\omega, \omega]$, por ejemplo). Por convexidad, $\frac{t_0}{t}\omega \in \Omega$ y entonces $x \in t\Omega$. □

La norma que vamos a describir se basa en ver cuánto hace falta estirar un conjunto absorbente para que alcance al punto deseado. Con esto, la nueva norma queda descrita como sigue:

Proposición 5. *Si E es un $\mathbb{R}$-espacio y $\Omega \subset E$ es **simétrico, conexo y absorbente**, podemos definir:*

$$\|x\|_\Omega = \operatorname{ínf}\{t > 0 : x \in t\Omega\}$$

Esta función:

1. *Está bien definida.*
2. *Es una seminorma.*
3. *Si $\nexists x \neq 0$ tal que $x \in t\Omega$ $\forall t > 0$, es una norma.*

4. Se tiene $B_1(0) \subset \Omega \subset \overline{B}_1(0)$ para esta norma.

Demostración. Para 1, veamos que el conjunto dado es no vacío al ser Ω absorbente, y está acotado inferiormente por definición (por 0, por ejemplo), de tal modo que tiene ínfimo. Para 2, está claro que es no negativa por definición (El conjunto está acotado inferiormente por 0).

Ahora, si $\lambda > 0$, veamos que $\|\lambda x\|_\Omega = \text{ínf}\{t > 0 : \lambda x \in t\Omega\} = \text{ínf}\,\lambda\{t > 0 : x \in t\Omega\} = \lambda\,\text{ínf}\{t > 0 : x \in t\Omega\} = \lambda \|x\|_\Omega$ dado que $\lambda > 0$. Si $\lambda = 0$, es claro que $\|\lambda x\|_\Omega = \|0\| = 0$, dado que $0 \in t\Omega$ $\forall t > 0$. Finalmente, como $\Omega = -\Omega$, se tiene que $x \in t\Omega \iff x \in -t\Omega \iff -x \in t\Omega$, de tal modo que $\|x\|_\Omega = \|-x\|_\Omega$, y por tanto si $\lambda < 0$, $\|\lambda x\|_\Omega = \|-|\lambda| x\|_\Omega = \||\lambda| x\|_\Omega = |\lambda| \|x\|_\Omega$.

Queda probar la desigualdad triangular. Fijemos $x, y \in E$. Tomamos $t > \|x\|_\Omega$ y $s > \|y\|_\Omega$. Está claro que entonces $x + y \in t\Omega + s\Omega$. Veamos que están también en $(t+s)\Omega$. Podemos poner $(x+y) = t\omega_1 + s\omega_2 = (t+s)\left(\frac{t}{t+s}\omega_1 + \frac{s}{t+s}\omega_2\right)$, luego $(x+y) \in (t+s)[\omega_1, \omega_2]$, y por convexidad $(x+y) \in (t+s)\Omega$. Por ello, $\|x+y\|_\Omega \leq t + s$. Como no importa qué $t > \|x\|_\Omega$ y $s > \|y\|_\Omega$ tomemos, podemos concluir que $\|x+y\|_\Omega \leq \|x\|_\Omega + \|y\|_\Omega$.

Para probar 3, veamos que se tiene que si $x \neq 0$, $\exists t_0 > 0$ con $x \in t_0\Omega$, de tal modo que $\|x\|_\Omega \geq t_0 > 0$.

Para probar 4, si $x \in B_1(0)$, entonces $\|x\|_\Omega < 1$, con lo que $x \in t\Omega$ $\forall t \geq 1$, y en particular $x \in \Omega$. Asimismo, si $x \in \Omega$, por el lema previo, $x \in t\Omega$ $\forall t \geq 1$, por ende $\|x_0\|_\Omega \leq 1$ y $x_0 \in \overline{B}_1(0)$. □

1.4. Acotación de aplicaciones lineales

Ahora introducimos la noción de norma en el espacio de aplicaciones lineales.

Definición 14. Sean E, F espacios normados, con $T : E \mapsto F$ lineal. Se dice que T está **acotada** si $\exists M < \infty$ tal que $\forall x \in E$, $\|Tx\|_F \leq M \|x\|_E$. Se denota $T \in \mathcal{L}(E, F)$.

Una característica de estos operadores es el menor valor M que satisface su definición. Por ello, definimos:

Definición 15. La **norma inducida** de $T \in \mathcal{L}(E, F)$ es:

$$\|T\| = \sup\{\frac{\|Tx\|_F}{\|x\|_E} : x \neq 0\}$$

Proposición 6. *Si $T \in \mathcal{L}(E, F)$, se tiene que $\|Tx\|_F \leq \|T\| \|x\|_E$ para cualquier x. Además, $\|T\|$ es la menor de las constantes que cumplen esto.*

Demostración. Tomamos $x \neq 0$ arbitrario para darnos cuenta de que, si M es una constante que satisface la definición anterior, $\|Tx\|_F \leq M \|x\|_E$, luego $\frac{\|Tx\|_F}{\|x\|_E} \leq M$. Como se cumple para todos los x, podemos comparar los supremos: $\sup\{\frac{\|Tx\|_F}{\|x\|_E} : x \neq 0\} \leq M$, luego $\|T\| \leq M$ si M satisface la definición. Por otro lado, por ser cota superior del conjunto, si $x \neq 0$, $\|T\| \geq \frac{\|Tx\|_F}{\|x\|_E}$ luego $\|T\| \|x\|_E \geq \|Tx\|_F$. Si $x = 0$ es inmediata la desigualdad. □

Además, el espacio de aplicaciones acotadas es un espacio vectorial y se mantiene la composición:

Proposición 7. *Si $T, S \in \mathcal{L}(E, F)$, entonces si $\alpha, \beta \in \mathbb{R}$, se tiene que $\alpha T + \beta S \in \mathcal{L}(E, F)$, y $\|\alpha T + \beta S\| \leq |\alpha| \|T\| + |\beta| \|S\|$.*

Asimismo, si $T \in \mathcal{L}(E, F)$, y $S \in \mathcal{L}(F, G)$, entonces $ST \in \mathcal{L}(E, G)$ y $\|ST\| \leq \|S\| \|T\|$.

Demostración. Para la primera, sea $x \neq 0$ arbitrario. $\|\alpha Tx + \beta Sx\|_F \leq |\alpha| \|Tx\|_F + |\beta| \|Sx\|_F \leq (|\alpha| \|T\| + |\beta| \|S\|) \|x\|_E$, luego está acotada. Por la proposición anterior se tiene la desigualdad. Para la segunda, sea $x \neq 0$ arbitrario. $\|STx\|_G \leq \|S\| \|Tx\|_F \leq \|S\| \|T\| \|x\|_E$, luego está acotada y por la proposición anterior se tiene la desigualdad. □

1.5. Espacios métricos

Definición 16. Sea $X \neq \emptyset$. Una **distancia** en X es una función $d : X \times X \mapsto X$ tal que:

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

La tupla (X, d) se denomina **espacio métrico**.

Definición 17. En el espacio métrico (X, d) se define la **bola abierta** de centro $x_0 \in X$ y radio $r > 0$ como $B_r(x_0) = \{x \in X : d(x_0, x) < r\}$, y la **bola cerrada** como $\overline{B}_r(x_0) = \{x \in X : d(x_0, x) \leq r\}$

Definición 18. En el espacio métrico (X, d), el conjunto $A \subset X$ es **abierto** si $\forall a \in A$, se tiene que $\exists r > 0$ con $B_r(a) \subset A$. El conjunto $B \subset X$ es **cerrado** si B^c es abierto.

Proposición 8. *Fijado $x \in X$ y $r > 0$, $B_r(x)$ es abierto.*

Demostración. Dado $y \in B_r(x)$, sea $\delta := r - d(x, y) > 0$. Veamos que $B_\delta(y) \subset B_r(x)$. Para ello, si $z \in B_\delta(y)$, entonces $d(z, y) < \delta = r - d(x, y)$, con lo que $r > d(z, y) + d(x, y) \geq d(z, x)$, luego $z \in B_r(x)$. $\square$

Proposición 9. *Fijado $x \in X$ y $r > 0$, $\overline{B}_r(x)$ es cerrado.*

Demostración. Dado $y \in \overline{B}_r(x)^c$, ponemos $\delta := d(x, y) - r > 0$. Veamos que $B_\delta(y) \subset \overline{B}_r(x)^c$. Sea $z \in B_\delta(y)$. Entonces $d(z, y) < d(x, y) - r$, luego $r < d(x, y) - d(z, y) \leq d(x, z) + d(z, y) - d(z, y) = d(x, z)$, como se quería. $\square$

Proposición 10. *Si (X, d) es espacio métrico, y $\{A_\lambda\}_{\lambda \in I}$, para cualquier I, es una colección de abiertos, entonces $\bigcup_{\lambda \in I} A_\lambda$ es abierto. Si Z, Y son abiertos, $Z \cap Y$ lo es también (y por inducción, toda intersección finita también). Si $\{B_\lambda\}_{\lambda \in I}$, para cualquier I, es una colección de cerrados, entonces $\bigcap_{\lambda \in I} B_\lambda$ es cerrado. Si T, W son cerrados, $T \cup W$ lo es también (y por inducción, toda intersección finita también).*

Demostración. Para la primera, veamos que si $a \in \bigcup_{\lambda \in I} A_\lambda$, entonces $a \in A_{\lambda_0}$, luego $B_r(a) \subset A_{\lambda_0} \subset \bigcup_{\lambda \in I} A_\lambda$ para algún r. Por otro lado, si $a \in Z \cap Y$, entonces $B_r(a) \subset Z$ y $B_s(a) \subset Y$ para r, s adecuados. Si $t := \min\{r, s\}$, entonces $B_t(a) \subset Z \cap Y$ al ser subconjunto de ambos. En el caso de los cerrados, veamos que $(\bigcap_{\lambda \in I} B_\lambda)^c = \bigcup_{\lambda \in I} B_\lambda^c$, que por la parte anterior es abierto. Asimismo, $(T \cup W)^c = T^c \cap W^c$, y por la parte anterior es abierto. $\square$

1.6. Acotación de aplicaciones lineales en espacios finitos.

Proposición 11. *Sea $A : \mathbb{R}^n \mapsto \mathbb{R}^m$ lineal, con la norma euclídea en ambos espacios. Entonces A es acotada, y si la matriz de A en bases canónicas es $M_A = (a_{i,j})_{i,j}$, se cumple que $\|A\| \leq \left(\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2\right)^{\frac{1}{2}}$. Este término de la derecha se denomina **norma de Fröbenius de** A.*

Demostración. Sea $x \in \mathbb{R}^n$. Vamos a escribir $M_A = \begin{pmatrix} F_1 \\ F_2 \\ \dots \\ F_m \end{pmatrix}$ donde cada F_k es la k-ésima fila (y por tanto un vector en $\mathbb{R}^n$). Podemos ver que el vector $Ax = \begin{pmatrix} F_1 x \\ F_2 x \\ \dots \\ F_m x \end{pmatrix}$. De este modo: $\|Ax\|^2 = \sum_{k=1}^m (F_k x)^2 \leq \sum_{k=1}^m \|F_k\|^2 \|x\|^2$ por la desigualdad de Schwarz. Por tanto, $\|Ax\|^2 \leq \|x\|^2 \sum_{k=1}^m \|F_k\|^2 = \|x\|^2 \sum_{k=1}^m \sum_{j=1}^n a_{kj}^2$, y tomando raíces se tiene la acotación de A. La desigualdad de la proposición se tiene recordando que $\|A\|$ es la menor constante que cumple esa desigualdad. $\square$

Proposición 12. *La norma de frobenius se puede expresar como $\|A\|_F = \sqrt{tr(A^tA)}$. Se tiene que si $A : \mathbb{R}^n \mapsto \mathbb{R}^m$ y $B : \mathbb{R}^m \mapsto \mathbb{R}^p$, entonces $\|BA\|_F \leq \|A\|_F \|B\|_F$.*

Demostración. Si $M_A = (a_{ij})$ y $M_A^t M_A = (b_{ij})$, entonces $b_{ii} = \sum_{k=1}^m a_{ik}^t a_{ki} = a_{ki}^2$, luego $tr(A^tA) = \sum_{i=1}^n b_{ii} = \sum_{i=1}^n \sum_{k=1}^m a_{ki}^2$. Para la segunda afirmación, $(BA)_{ij} = \sum_{k=1}^m b_{ik}a_{kj}$. Entonces $|(BA)_{ij}| = |\begin{pmatrix} b_{i1} & \dots & b_{im} \end{pmatrix} \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}| \leq (\sum_{k=1}^m b_{ik})^{\frac{1}{2}} (\sum_{k=1}^m a_{kj})^{\frac{1}{2}}$, por la desigualdad de Schwarz. Así, $\|BA\|_F^2 = \sum_{i=1}^p \sum_{j=1}^n |(BA)_{ij}|^2 \leq \sum_{i=1}^p \sum_{j=1}^n (\sum_{k=1}^m b_{ik})^{\frac{1}{2}} (\sum_{k=1}^m a_{kj})^{\frac{1}{2}} = (\sum_{i=1}^p \sum_{k=1}^m b_{ik}^2)(\sum_{j=1}^n \sum_{k=1}^m a_{kj}^2)$. □

Cabe observar que la estimación $\|A\|_F \geq \|A\|_2$ no es óptima pero sí fácil de calcular. Por ejemplo, $\|I_n\|_F = \sqrt{n}$ pero $\|I_n\|_2 = 1$. El siguiente resultado caracteriza explícitamente a $\|\cdot\|_2$.

Proposición 13. *Sea $A \in \mathcal{M}_{mn}(\mathbb{R})$. Se tiene que $\|A\|_2$, es decir, la norma matricial inducida por la euclídea, verifica: $\|A\|_2 = \max\{\sigma_j^{\frac{1}{2}}\}_1^n$, donde los σ_j son los distintos autovalores de A^tA.*

Demostración. Como A^tA es simétrica tiene todos sus autovalores reales. Asimismo, como es autoadjunta, observamos que para $x \in \mathbb{R}^n$ se tiene: $x^t(A^tA)x = \langle A^tAx, x\rangle = \langle Ax, Ax\rangle = \|Ax\|^2 \geq 0$, es decir, A^tA es semidefinida positiva. Por tanto, todos sus autovalores son no negativos. Sabemos que existe base ortonormal $\{v_1, \dots, v_n\}$ de autovectores tales que $A^tAv_j = \sigma_j v_j$.

Ahora, dado $x \in \mathbb{R}^n$, si sus coordenadas en la base anterior son $(x_1, \dots, x_n)$, sabemos que $\|x\|_2^2 = \sum_{j=1}^n x_j^2$ por ser la base ortonormal. Entonces, $\|Ax\|_2^2 = \langle A^tAx, x\rangle = \langle A^tA\sum_{i=1}^n x_iv_i, \sum_{k=1}^n x_kv_k\rangle = \sum_{j,k=1}^n x_jx_k \langle A^tAv_j, v_k\rangle = \sum_{j,k=1}^n x_jx_k\sigma_j \langle v_j, v_k\rangle = \sum_{j,k=1}^n x_jx_k\sigma_j\delta_{jk} = \sum_{k=1}^n x_k^2\sigma_k \leq \max\{\sigma_j\} \sum_{k=1}^n x_k^2 = \max\{\sigma_j\} \|x\|_2^2$. Esto prueba que $\|A\|_2 \leq \max\{\sigma_j^{\frac{1}{2}}\}$.

La implicación contraria no es difícil: si $\max\{\sigma_j\} = \sigma_{j_0}$, entonces $\|Av_{j_0}\|^2 = \langle A^tAv_{j_0}, v_{j_0}\rangle = \sigma_{j_0} \|v_{j_0}\|^2$, luego $\frac{\|Av_{j_0}\|}{\|v_{j_0}\|} = \sigma_{j_0}$, con lo que $\|A\|_2 \geq \sigma_{j_0}$. □

Definición 19. A las raíces de los autovalores de A^tA se los denomina **valores singulares** de A.

Observación 2. Dada $A \in \mathcal{M}_{mn}(\mathbb{R})$, como la traza es invariante al cambio de base, si $\{\sigma_1, \dots \sigma_n\}$ son sus valores singulares, se tiene que: $\|A\|_F = \sqrt{tr(A^tA)} = \sqrt{\sum_{i=1}^n \sigma_i^2}$.

Esto permite ver también, con facilidad, que $\|A\|_2 \leq \|A\|_F$.

A continuación vamos a caracterizar otras dos normas del operador A: las que inducen las normas 1 e ∞. Estas además serán mucho más fáciles de calcular explícitamente:

Proposición 14. *Sea $A : \mathbb{R}^n \mapsto \mathbb{R}^m$ lineal. Sean $\{C_i\}_1^n$ los vectores columna de A, y $\{F_j\}_1^n$ los vectores fila. Entonces:*

$$\|A\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^m |a_{ij}| = \max_{1\leq j\leq n} \|C_j\|_1$$

Y además:

$$\|A\|_\infty = \max_{1\leq i\leq m} \sum_{i=1}^n |a_{ij}| = \max_{1\leq i\leq m} \|F_i\|_1$$

Demostración. Sea $x = \sum_1^n x_ie_i$ un vector de $\mathbb{R}^n$, con las x_i las coordenadas canónicas. Entonces: $\|Ax\|_1 = \|A\sum_1^n x_ie_i\|_1 = \|\sum_1^n x_iAe_i\|_1 \leq \sum_1^n |x_i| \|Ae_i\|_1 \leq \max_{1\leq j\leq n} \|Ae_j\|_1 \sum_1^n |x_i| = \max_{1\leq j\leq n} \|C_j\|_1 \|x\|_1$, de tal modo que $\|A\|_1 \leq \max_{1\leq j\leq n} \|C_j\|_1$. Por otra parte, $\frac{\|Ae_j\|_1}{\|e_j\|_1} = \|Ae_j\|_1 = \|C_j\|_1$, luego, por ser supremo, $\|A_1\| \geq \|C_j\|$, con lo que $\|A\|_1 \geq \max_{1\leq j\leq n} \|C_j\|_1$ y se tiene lo que se quería.

Para la segunda afirmación, como $Ax = \begin{pmatrix} F_1x \\ \vdots \\ F_mx \end{pmatrix}$, tenemos que $\|Ax\|_\infty = \max_{1\leq i\leq m} |F_ix| \leq \max_{1\leq i\leq m} \|F_i\|_1 \|x\|_\infty$, por la desigualdad de Hölder, pero $\max_{1\leq i\leq m} \|F_i\|_1 \|x\|_\infty = \|x\|_\infty \max_{1\leq i\leq m} \|F_i\|_1$ luego $\|A\|_\infty \leq \max_{1\leq i\leq m} \|F_i\|_1$. Por otra parte, supongamos que $k = \arg\max_{1\leq i\leq m} \|F_i\|_1$. Si $F_k = (f_1, f_2, \ldots, f_n)$, vamos a poner $x = (sign(f_1), \ldots, sign(f_n))$. Cabe observar que $\|x\|_\infty \leq 1$. Ahora veamos que $\|Ax\|_\infty = \max_{1\leq i\leq m} |F_ix| \geq |F_kx| = |\sum_{i=1}^n f_i \cdot sign(f_i)| = |\sum_{i=1}^n |f_i|| = \|F_k\|_1$. Tenemos entonces que $\frac{\|Ax\|_\infty}{\|x\|_\infty} \geq \frac{\|F_k\|_1}{\|x\|_\infty} \geq \|F_k\|_1$, luego debe darse que $\|A\|_\infty \geq \max_{1\leq i\leq m} \|F_i\|_1$. □

A través de estas normas, se tiene un resultado que permite estimar de forma mucho más buena la cantidad $\|A\|_2$.

Proposición 15. *Se tiene que* $\|A\|_2 \leq \|A\|_1 \|A\|_\infty$.

Demostración. Observemos en primer lugar que $\|Ax\|_2^2 = \langle Ax, Ax\rangle = \langle A^tAx, x\rangle$. Ponemos $B = A^tA = (b_{ij})_{1\leq i,j\leq n}$. Entonces, si $x = (x_i)_1^n$, se tiene que $\|Ax\|_2^2 =< Bx, x >= \sum_{i,j} b_{ij}x_ix_j \leq \sum_{i,j} |b_{ij}||x_ix_j| \leq \sum_{i,j} |b_{ij}|\frac{x_i^2+x_j^2}{2}$ por la desigualdad de Hölder. Como $b_{ij} = b_{ji}$, vamos a agrupar de 2 en 2 estos términos, obteniendo así: $\sum_{i,j} |b_{ij}|\frac{x_i^2+x_j^2}{2} = \sum_{i,j} |b_{ij}|x_i^2 = \sum_i x_i^2 \sum_j |b_{ij}| \leq \sum_i x_i^2 \max_i \sum_j |b_{ij}| \leq \max_i(\sum_j |b_{ij}|) \|x\|_2^2$.

Por tanto, $\|A\|_2 \leq \max_i(\sum_j |b_{ij}|)$. Falta reordenar de esta manera: $\max_i(\sum_j |b_{ij}|) = \max_i \sum_j \sum_{l=1}^m |a_{li}a_{lj}| = \max_i \sum_{l=1}^m \sum_j |a_{li}a_{lj}| = \max_i \sum_{l=1}^m |a_{li}| \sum_j |a_{lj}| \leq \max_i \sum_{l=1}^m |a_{li}| \max_{1\leq l\leq m} \sum_j |a_{lj}| = \max_{1\leq l\leq m} \sum_j |a_{lj}| \max_i \sum_{l=1}^m |a_{li}| = \|A\|_\infty \|A\|_1$. □

1.7. Conceptos topológicos básicos

Definición 20. Dado $A \subset X$ con (X, d) métrico, definimos:

- $int(A) = \{x \in A : \exists r > 0 \,:\, B_r(x) \subset A\}$ el **interior** de A.
- $ext(A) = int(A^c)$ el **exterior** de A.
- $\overline{A} = \{x \in X : \forall r > 0, B_r(x) \cap A \neq \emptyset\}$ el **cierre** de A.
- $\partial A = \overline{A} \backslash int(A)$ el **borde** de A.

A continuación demostraremos algunas intuitivas propiedades de estos conjuntos.

Proposición 16. *Si $A \subset X$ y (X, d) es espacio métrico, entonces $int(A)$ y $ext(A)$ son abiertos. De hecho, $int(A)$ es el mayor abierto contenido en A. Si $A \subset B$, $int(A) \subset int(B)$.*

Demostración. Sea $x \in int(A)$. Sabemos entonces que $B_r(x) \subset A$ para cierto r. Como esa bola es abierta, si $y \in B_r(x)$, entonces $B_s(y) \subset B_r(x) \subset A$ para cierto s, de tal modo que $y \in int(A)$. Por tanto, $B_r(x) \subset int(A)$ y hemos acabado. Como $ext(A) = int(A^c)$ debe ser un abierto. Si $\Omega \subset A$ es abierto, entonces si $x \in \Omega$, $B_r(x) \subset \Omega \subset A$ para cierto r luego $x \in int(A)$ y por tanto $\Omega \subset int(A)$. Finalmente, si $x \in int(A)$, entonces $B_r(x) \subset A \subset B$ para cierto r luego $x \in int(B)$. □

Proposición 17. *Si $A \subset X$ y (X,d) es espacio métrico, entonces $\overline{A}$ y ∂A son cerrados. De hecho, $\overline{A}$ es el menor cerrado que contiene a A. Si $A \subset B$, $\overline{A} \subset \overline{B}$.*

Demostración. Sea $x \in \overline{A}^c$. Entonces, por definición, $B_r(x) \subset A^c$ para cierto r (si no, estaría en $\overline{A}$). Si $y \in B_r(x)$, entonces $B_s(y) \subset B_r(x) \subset A^c$ para cierto s, de tal modo que $y \in \overline{A}^c$. Por lo tanto, $B_r(x) \subset \overline{A}^c$ y $\overline{A}^c$ es entonces abierto. Como $\partial A = \overline{A} \cap int(A)^c$, es cerrado. Si $A \subset F$ y F es cerrado, entonces, si $x \in F^c$, se tiene $B_r(x) \subset F^c$ luego $B_r(x) \subset A^c$ dado que $F^c \subset A^c$, con lo que $x \neq \overline{A}$ y por tanto $x \in \overline{A}^c$. Hemos probado entonces que $F^c \subset \overline{A}^c$ y por tanto, $\overline{A} \subset F$. Finalmente, si $A \subset B$, entonces si $x \in \overline{A}$, $\forall r > 0$ se tiene $\emptyset \neq B_r(x) \cap A \subset B_r(x) \cap B$ luego $x \in \overline{B}$. □

1.7.1. Límites y continuidad en espacios métricos

Definición 21 (Límite de sucesión)**.** Si (X,d) es un espacio métrico y $(x_n)_{n\in\mathbb{N}}$ es una sucesión en X, se dice que $x_n \to x \in X$, o bien $\lim_{n\to\infty} x_n = x$, si $\forall \epsilon > 0$, $\exists N \in \mathbb{N}$ tal que $n \geq N \implies d(x_n, x) < \epsilon$.

Cabe observar que existen varias definiciones equivalentes, por ejemplo con $n < N$ en lugar de $n \leq N$, o $d(x_n, x) \leq \epsilon$ en lugar de la desigualdad estricta. La comprobación es inmediata.

Definición 22 (Sucesión de Cauchy)**.** La sucesión $(x_n)_{n\in\mathbb{N}}$ es de Cauchy si $\forall \epsilon > 0$, $\exists N \in \mathbb{N}$ tal que $n, m \geq N \implies d(x_n, x_m) < \epsilon$.

Proposición 18. *Supongamos que $x_n \to x$ y $x_n \to y$. Entonces $x = y$. Esto es, el límite de una sucesión es único.*

Demostración. Sea $\epsilon > 0$. Supongamos que N_x y N_y son los valores naturales que satisfacen la definición de límite para $\frac{\epsilon}{2}$, con x e y respectivamente. Entonces, si $n > \max\{N_x, N_y\}$, tenemos que $d(x,y) \leq d(x_n, x) + d(x_n, y) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$. Como además $d(x,y) \geq 0$, no queda otra que $d(x,y) = 0$. Por tanto $x = y$. □

Proposición 19. *Si $(x_n)_{n\in\mathbb{N}}$ tiene límite x, entonces es de Cauchy.*

Demostración. Dado $\epsilon > 0$, tomamos N el que satisface la definición de límite para $\frac{\epsilon}{2}$. Si $n, m \geq N$, entonces $d(x_n, x_m) \leq d(x_n, x) + d(x_m, x) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$. □

Podemos caracterizar el cierre de un conjunto:

Proposición 20. *Sea (X,d) un espacio métrico. Entonces $x \in \overline{A} \iff \exists (x_n)_{n\in\mathbb{N}} \subset A$ tal que $x_n \to x$.*

Demostración. Para $\impliedby$, veamos que si $\epsilon > 0$, entonces hay algún $x_n \in A$ tal que $d(x_n, x) \leq \epsilon$, es decir, $x_n \in B_\epsilon(x)$, luego $B_\epsilon(x) \cap A \neq \emptyset$ y por tanto $x \in \overline{A}$. Para $\implies$, vamos a definir x_n como un elemento de $B_{\frac{1}{n}}(x) \cap A$ para todo $n \in \mathbb{N}$. Sabemos que existe por ser x un punto de cierre. Tenemos así una sucesión $(x_n)_n \subset A$. Veamos que $x_n \to x$. Si $\epsilon > 0$, tomamos $N \in \mathbb{N}$ tal que $\epsilon > \frac{1}{N}$. Si $n \geq N$, tenemos entonces que $B_{\frac{1}{n}}(x) \subset B_{\frac{1}{N}}(x)$ y por tanto, $x_n \in B_{\frac{1}{N}}(x)$, luego $d(x_n, x) < \frac{1}{N} < \epsilon$. Así hemos probado que $x_n \to x$. □

Definición 23 (Continuidad local de una función)**.** Si (X, d_X) e (Y, d_Y) son espacios métricos, y $f : A \subset X \mapsto Y$, se dice que f es continua en $x_0 \in A$ si $\forall \epsilon > 0$, $\exists \delta > 0$ tal que si $x \in A$ y $0 < d_X(x, x_0) < \delta$, entonces $d_Y(f(x), f(x_0)) < \epsilon$.

Si la función es continua en todos los puntos de A, se dice que es continua en A.

A continuación veremos una caracterización de las funciones continuas por medio de sucesiones:

Proposición 21. *Si (X, d_X) e (Y, d_Y) son espacios métricos, y $f : A \subset X \mapsto Y$ es una función, se tiene que:*

$$f \text{ es continua} \iff \forall (x_n)_n \subset A \text{ tal que } x_n \to x_0 \text{ se tiene } f(x_n) \to f(x_0)$$

Demostración. Si f es continua y $(x_n)_n \subset A$ tiende a x_0, dado $\epsilon > 0$, sabemos que $\exists \delta > 0$ tal que $0 < d(x, x_0) < \delta \implies d(f(x), f(x_0)) < \epsilon$. Pero también $\exists N \in \mathbb{N}$ tal que $n \geq N \implies d(x_n, x_0) < \delta$. Por tanto, si $n \geq N$, se tiene $d(f(x_n), f(x_0)) < \epsilon$ y hemos acabado. Por otro lado, supongamos que f no fuese continua. Vamos a construir una x_n que tienda a x_0 pero $f(x_n)$ no tienda a $f(x_0)$. Como f no es continua, $\exists \epsilon > 0$ tal que $\forall \delta > 0$, $\exists x_\delta \in A$ tal que $d(x_\delta, x_0) < \delta$ pero $d(f(x_\delta), f(x_0)) \geq \epsilon$. Tomamos $(x_{\delta=\frac{1}{n}})_n$. Está claro que tiende a x_0 por construcción. No obstante, $f(x_n)$ no puede tender a $f(x_0)$ dado que $d(f(x_n), f(x_0)) \geq \epsilon$ siempre. □

Ahora vamos a ver una caracterización de las funciones continuas que puede ser de gran utilidad para determinar si un conjunto es abierto o cerrado:

Proposición 22. *Si (X, d_X) e (Y, d_Y) son espacios métricos y $f : X \mapsto Y$, son equivalentes:*

1. *f es continua en X.*
2. *$\forall U \subset Y$ abierto, $f^{-1}(U)$ es abierto.*
3. *$\forall R \subset Y$ cerrado, $f^{-1}(R)$ es cerrado.*

Demostración. Debemos darnos cuenta de que (2) $\iff$ (3) dado que $f^{-1}(A^c) = f^{-1}(A)^c$. Por tanto, dada una de las afirmaciones, basta con pasar al complementario y se tiene. Por tanto, probaremos solo que (1) $\iff$ (2). Para $\implies$, tomemos $x_0 \in f^{-1}(U)$. Como U es abierto, $B_\epsilon(f(x_0)) \subset U$ para cierto $\epsilon > 0$. Por continuidad, $\exists \delta > 0$ tal que $x \in B_\delta(x_0) \implies f(x) \in B_\epsilon(f(x_0))$, luego $B_\delta(x_0) \subset f^{-1}(B_\epsilon(f(x_0))) \subset f^{-1}(U)$ y tenemos por tanto que $f^{-1}(U)$ es abierto. Para $\impliedby$, sea $\epsilon > 0$. Como $B_\epsilon(f(x_0))$ es abierta, también lo es su preimagen por f, en la que está x_0. Es decir $B_\delta(x_0) \subset f^{-1}(B_\epsilon(f(x_0)))$ para cierto δ, pero esto es que si $d(x, x_0) < \delta$, entonces $d(f(x), f(x_0)) < \epsilon$, luego f es continua. □

A continuación vamos a ver que los subconjuntos de espacios métricos también lo son.

Definición 24. Sea (X, d) métrico y $A \subset X$, $A \neq \emptyset$. Entonces, (A, d_A), con $d_A = d|_{A \times A}$, es otro espacio métrico. d_A se denomina **métrica relativa a** A y los abiertos de A bajo esa métrica se denominan **abiertos relativos a** A.

Proposición 23. *Si (X, d) es métrico y $A \subset X$ es no vacío, entonces $U \subset A$ es abierto relativo a A $\iff \exists V \subset X$ abierto con $U = A \cap V$.*

Demostración. Debemos ver primero que, dado $x \in A$ y $r > 0$, $B_r^A(x) = B_r^X(x) \cap A$, donde $B_r^Y(x)$ es la bola de radio r centrada en x bajo la distancia del espacio métrico Y. En efecto, $y \in B_r^X(x) \cap A \iff d(x, y) < r$ y $y \in A \iff d_A(x, y) < r \iff y \in B_r^A(x)$. Sabiendo esto, veamos que si $U = A \cap V$ con V abierto de X, y $x_0 \in U$, como $x_0 \in V$, tenemos que $B_r^X(x_0) \subset V$ para cierto r, de tal modo que $B_r^X(x_0) \cap A \subset A \cap V = U$, es decir, $B_r^A(x_0) \subset U$ luego U es abierto de A. Por otro lado si $U \subset A$ es abierto de A, es fácil ver que $U = \bigcup_{x \in U} B_{r_x}^A(x)$ con r_x el radio que asegura que $B_{r_x}^A(x) \subset A$. Es decir, $U = \bigcup_{x \in U}(B_{r_x}^X(x) \cap A) = A \cap \bigcup_{x \in U} B_{r_x}^X(x)$, y esa unión es en sí un abierto de X. □

Definición 25. Si $X \neq \emptyset$, diremos que la familia $\{U_k\}_{k \in I}$ con I arbitrario, donde cada $U_k \subset X$, es un **recubrimiento** de X si $\bigcup_{k \in I} U_k = X$. Si cada U_k es abierto, es un **recubrimiento abierto** o **recubrimiento por abiertos**. Si $K \subset X$, diremos que es un recubrimiento (abierto) de K si $K \subset \bigcup_{k \in I} U_k$.

Definición 26 (Conjunto compacto)**.** Si (X, d) es espacio métrico, entonces $K \subset X$ es un **conjunto compacto** si $\forall \mathcal{R} = \{U_k\}_{k \in I}$ recubrimiento abierto de K, se tiene que $\exists \{k_1, \ldots, k_n\} \subset I$ tal que $\{U_{k_j}\}_{j=1}^n$ también es un recubrimiento abierto de K.

Obsérvese que lo que estamos imponiendo es que **todo** recubrimiento abierto pueda reducirse a un subrecubrimiento finito del mismo.

Proposición 24. *Supongamos que (X,d) es un espacio métrico y X es compacto. En ese caso, si $F \subset X$, F cerrado, entonces F es compacto.*

Demostración. Sea $\{U_k\}$ un recubrimiento abierto de F. Entonces, $\{U_k \cup F^c\}$ es un recubrimiento abierto de X, dado que los U_k cubren a F y F^c cubre el resto, y además son abiertos al ser F cerrado. De este modo, tenemos índices $\{k_1, \ldots, k_n\}$ tales que $X = \bigcup_{i=1}^n (U_{k_i} \cup F^c)$, luego $F = F \cap X = F \cap \bigcup_{i=1}^n (U_{k_i} \cup F^c) = F \cap \bigcup_{i=1}^n (U_{k_i})$, de tal modo que $F \subset \bigcup_{i=1}^n (U_{k_i})$. $\square$

En espacios métricos, se tiene que el recíproco, además, es cierto. Es decir, si un conjunto es compacto, entonces será cerrado.

Proposición 25. *Sea (X,d) un espacio métrico y $F \subset X$ un conjunto compacto. Se tiene que F es cerrado.*

Demostración. Supongamos que F no es cerrado. En ese caso, $\overline{F} \neq F$ y podemos tomar $p \in \overline{F} \backslash F$. Dado $x \in F$, sea $r_x = \frac{d(x,p)}{2} > 0$. Construimos así un recubrimiento abierto de F, $\{B_{r_x}(x)\}_{x \in F}$. Como F es compacto, tenemos un cierto $\{B_{r_j}(x_j)\}_{j=1}^n$ subconjunto finito, que también es recubrimiento.

Si ponemos $r := \min\{r_j\}$, entonces es inmediato ver que $B_r(p) \cap B_{r_j}(x_j) = \emptyset$. Efectivamente, si $x \in B_r(p)$, entonces $d(x_j,p) \leq d(x,x_j) + d(x,p)$ luego $d(x,x_j) \geq d(x_j,p) - d(x,p) > d(x_j,p) - r \geq d(x_j,p) - \frac{d(x_j,p)}{2} = r_j$, y por tanto $x \neq B_{r_j}(x_j)$.

Pero entonces $B_r(p) \cap F = \emptyset$, lo que contradice que $p \in \overline{F}$. $\square$

Definición 27. Sea (X,d) un espacio métrico y $A \subset X$. A es **acotado** si $\exists x_0 \in X$ y $r > 0$ tal que $A \subset B_r(x_0)$.

Proposición 26. *Sea (X,d) un espacio métrico y $A_1, \ldots, A_n \subset X$. Se tiene:*

1. *Si A_i es acotado $\forall i \in \{1, \ldots, n\}$, entonces $\bigcup_{i=1}^n A_i$ también lo es.*
2. *Si A_i es compacto $\forall i \in \{1, \ldots, n\}$, entonces $\bigcup_{i=1}^n A_i$ también lo es.*
3. *Si $\{F_\alpha\}_\alpha$ son compactos de X, entonces $\bigcap F_\alpha$ también lo es.*

Demostración. Para 1, vamos a ver que $A_1 \cup A_2$ es acotado, siguiendo el resto por inducción. Tenemos que $A_1 \subset B_r(x)$ y $A_2 \subset B_s(y)$ para ciertos puntos x, y y radios $r, s > 0$. Entonces $A_1 \cup A_2 \subset B_r(x) \cup B_s(y)$. Si $z \in B_r(x) \cup B_s(y)$, entonces $d(z,x) \leq d(z,y) + d(y,x) \leq s + d(y,x)$. Esto nos dice que $z \in B_{s+d(y,x)}(x)$ y por tanto $A_1 \cup A_2 \subset B_{s+d(y,x)}(x)$.

Para 2, si $\{U_\alpha\}_\alpha$ es recubrimiento abierto de $A_1 \cup A_2$, en particular lo es de A_1 y A_2. Por tanto $\{U_k\}_1^n$ y $\{U'_k\}_1^{n'}$ son sendos subrecubrimientos abiertos de A_1 y A_2. Es decir, $A_1 \cup A_2 \subset \bigcup_1^n U_k \cup \bigcup_1^{n'} U'_k$, con lo que $\{U_k\}_1^n \cup \{U'_k\}_1^{n'}$ es un subrecubrimiento abierto de la unión. Para el resto de conjuntos sigue por inducción.

Para 3, como cada F_α es cerrado por la proposición anterior, se tiene que $\bigcap_\alpha F_\alpha \subset F_{\alpha_0}$ es un subconjunto cerrado de un compacto y se tiene lo que se quería. $\square$

A continuación vamos a ver algunas definiciones que dan una idea similar a la compacidad, y veremos en un futuro las equivalencias entre ellas.

Definición 28. Si (X,d) es un espacio métrico, se dice que X es **secuencialmente compacto** si toda sucesión $(x_n)_n \subset X$ tiene una subsucesión $(x_{n_j})_j$ convergente.

Definición 29. El espacio métrico (X,d) es **completo** si toda sucesión de Cauchy $(x_n)_n$ es convergente.

Proposición 27. *Supongamos que* $(x_n)_{n\in\mathbb{N}}$ *es de Cauchy en* (X,d) *y tiene una subsucesión convergente,* $(x_{n_j})_{j\in\mathbb{N}}$ *de límite* x. *En ese caso* $x_n \to x$.

Demostración. Sea $\epsilon > 0$. Sabemos que $\exists J \in \mathbb{N}$ tal que $j \geq J$ implica que $d(x_{n_j}, x) < \frac{\epsilon}{2}$. Por otra parte, sabemos que existe $N \in \mathbb{N}$ tal que si $m, n \geq N$, se tiene $d(x_n, d_m) < \frac{\epsilon}{2}$. Ahora vamos a seleccionar n_j tal que $n_j \geq N$ y además $j \geq J$. Obsérvese que esto se puede hacer dado que hay infinitos $j \geq J$, pero si siempre se tuviese $n_j < N$, tendríamos finitos n_j dado que son naturales. Una vez seleccionado, basta ver que si $n \geq N$, se tiene $d(x, x_n) \leq d(x, x_{n_j}) + d(x_{n_j}, x_n) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$. □

Corolario. Se tiene que si (X,d) es secuencialmente compacto, entonces es completo.

Esto es así dado que si se tiene una sucesión de Cauchy, si es secuencialmente compacto, tiene una subsucesión convergente y por tanto converge.

Definición 30. Si (X,d) es espacio métrico y $x_0 \in X$, se dice que x_0 es **punto de acumulación de** A si $\forall r > 0$ se tiene que $B_r(x_0)\backslash\{x_0\} \cap A \neq \emptyset$. El conjunto de puntos de acumulación de A se denota A'.

Proposición 28. *Si* (X,d) *es métrico y* $A \subset X$, *entonces* A *es cerrado* $\iff A' \subset A$.

Demostración. Para $\implies$, veamos que si A es cerrado, entonces $\overline{A} = A$, y como trivialmente (atendiendo a las definiciones) se tiene $A' \subset \overline{A}$, hemos acabado. Para $\impliedby$, si $x \in A^c$, entonces por hipótesis $x \notin A'$, de modo que $B_r(x)\backslash\{x\} \subset A^c$ para cierto $r > 0$, pero entonces $B_r(x) \subset A^c$, con lo que finalmente A^c es abierto. □

Definición 31. El espacio (X,d) es **totalmente acotado** si $\forall\epsilon > 0$, se tiene $\exists\{x_i\}_1^n$ tales que $X = \bigcup_{i=1}^n B_\epsilon(x_i)$.

Es inmediato observar que si es totalmente acotado, entonces es acotado, y además si es compacto es totalmente acotado.

Definición 32. Sea (X,d) métrico. Se dice que tiene la **propiedad de Bolzano-Weierstrass** siempre que $\forall A \subset X$, A infinizto, se tenga $A' \neq \emptyset$.

Definición 33. Sea (X,d) métrico. Se dice que la familia $\{K_\alpha\}_{\alpha\in A}$ tiene la **propiedad de intersección finita** si $\forall I \subset A$, I finito, se tiene que $\bigcap_{\alpha\in I} K_\alpha \neq \emptyset$.

Proposición 29. *Sea* (X,d) *métrico compacto y* $\{K_\alpha\}_{\alpha\in A}$ *una familia de cerrados arbitraria que verifica la propiedad de intersección finita. Entonces, de hecho,* $\bigcap_{\alpha\in A} K_\alpha \neq \emptyset$.

Demostración. Supongamos que no fuese así. Entonces $X = (\bigcap_{\alpha\in A} K_\alpha)^c = \bigcup_{\alpha\in A} K_\alpha^c$. Como cada K_α es cerrado, se tiene que los $\{K_\alpha^c\}_{\alpha\in A}$ constituyen un recubrimiento abierto. Por compacidad, $\{K_i^c\}_1^n$ es un subrecubrimiento finito, de tal modo que $X = \bigcup_1^n F_i^c$, y por tanto $\emptyset = \bigcap_1^n F_i$, contradiciendo que verifiquen la propiedad de intersección finita. □

Proposición 30. *Sea* (X,d) *un espacio métrico compacto y* $\{K_i\}_1^\infty$ *una familia numerable de conjuntos cerrados, con propiedad de intersección finita, y tales que* $diam(K_n) \to 0$ *si* $n \to \infty$ *(donde* $diam(A) = \sup\{d(x,y) : x, y \in A\}$*). Entonces, además de la proposición anterior, se tiene que* $K := \bigcap_1^\infty K_i = \{c\}$ *para cierto* $c \in X$.

Demostración. Si $x, y \in K$ con $x \neq y$, entonces $x, y \in K_n$, de tal modo que $0 < d(x,y) \leq diam(K_n)$, impidiendo que $diam(K_n) \to 0$. □

Una vez presentadas todas estas nociones, que tratan de generalizar el concepto de *conjunto finito* a espacios métricos, y permiten extender propiedades locales a propiedades globales (como la proposición 29), vamos a ver las equivalencias que existen entre ellos. Para ello, necesitamos un resultado:

Teorema 1. *Sea (X, d) un espacio métrico. Son equivalentes:*

1. *X es compacto.*
2. *X tiene la propiedad de Bolzano-Weierstrass.*
3. *X es secuencialmente compacto.*
4. *X es totalmente acotado y completo.*

Demostración. Comenzamos por (1) $\implies$ (2). Supongamos que $A \subset X$ infinito tal que $A' = \emptyset$. Entonces, $\forall x \in X$, se tiene que x no es de acumulación de A, esto es, $\exists r_x > 0$ tal que $A \cap B_{r_x}(x) \subset \{x\}$. Consideramos el recubrimiento abierto $\{B_{r_x}(x)\}_{x \in X}$. Entonces hay un subrecubrimiento finito tal que $X = \bigcap_1^n B_{r_i}(x_i)$, y por tanto $A = X \cap A = \bigcap_1^n A \cap B_{r_i}(x_i) \subset \{x_1, x_2, \ldots, x_n\}$, lo que contradice que A sea infinito.

Seguimos por (2) $\implies$ (3). Sea $(x_n)_n$ una sucesión en X. Si toma un conjunto finito de valores, alguno se repite infinitas veces luego está claro que tiene una subsucesión convergente. Si no, por la propiedad de Bolzano-Weierstrass, el conjunto de valores tiene un x de acumulación. veamos que si $\epsilon_j = \frac{1}{j}$, la bola $B_{\epsilon_j}(x) \backslash \{x\} \cap \{x_n\}_n$ tiene infinitos elementos (si no los tuviese, el punto no sería de acumulación tomando el radio suficientemente pequeño). Tomamos el elemento x_{n_j} de tal modo que $n_j > n_{j-1}$. De esta manera, la sucesión $\{x_{n_j}\}_{j \in \mathbb{N}}$ así conseguida tiene límite x.

Ahora pasamos a (3) $\implies$ (4). Veremos en primer lugar que dado $r > 0$, existe un conjunto de puntos $\{x_1, \ldots, x_n\}$ finito, maximal, tal que $d(x_i, x_j) \geq r$ si $i \neq j$. Es decir, ningún superconjunto de este verifica la misma propiedad. Supongamos que no existiese. En ese caso, podríamos obtener K infinito numerable con esta propiedad. (A base de incrementar arbitrariamente el tamaño del conjunto al no tener maximal). Pongamos que $K = (x_n)_{n \in \mathbb{N}}$ es una numeración del conjunto. Entonces, como X es secuencialmente compacto, hay una subsucesión suya que converge, pero $d(x_i, x_j) \geq r$ si $i \neq j$, luego no es de Cauchy, lo que es una contradicción. Si ahora fijamos $\epsilon > 0$ y $\{x_1, \ldots, x_n\}$ es tal maximal para $r = \epsilon$, vemos que $X = \bigcup_{j=1}^n B_\epsilon(x_j)$. En efecto, si $x \notin \bigcup_{j=1}^n B_\epsilon(x_j)$, ocurre que $d(x, x_j) \geq \epsilon \; \forall j \in \{1, \ldots, n\}$, de tal manera que x verifica la propiedad y el conjunto no es maximal. Por lo tanto, está totalmente acotado. Por otro lado, si tenemos una sucesión de Cauchy, como hemos asumido (3), tiene una subsucesión convergente luego converge.

Finalmente veremos (4) $\implies$ (1). Para ello primero veremos que (4) $\implies$ (3). Sea $(x_n)_n$ sucesión en X. Si la sucesión toma finitos valores, como siempre, hemos terminado. Si no, como está totalmente acotado, tenemos que $X = \bigcup_{i=1}^n B_1(y_i)$ Alguna de esas bolas debe contener infinitos elementos de (x_n), la llamaremos B_1. Esos infinitos elementos forman la subsucesión $\{x_n^1\}_{n \in \mathbb{N}}$. Repitiendo el mismo argumento para $B_1 \cap X$, también totalmente acotada, pero con radios $\frac{1}{2}$, obtenemos otra subsucesión $\{x_n^2\}_n$ y otra bola B_2. Si seguimos así sucesivamente, obtenemos bolas B_j y sucesiones $\{x_n^j\}_n$, donde cada B_j tiene radio $\frac{1}{2^{j-1}}$ y cada subsucesión pertenece a todas las bolas anteriores. Definimos ahora la sucesión *diagonal* $y_j := x_j^j$. Como, si $l > k$, $d(y_k, y_l) \leq d(yl, y_{l-1}) + d(y_{l-1}, y_k) \leq \frac{1}{2^{l-1}} + d(y_{l-1}, y_k) \leq \cdots \leq 2 \sum_{n=k}^{l-1} \frac{1}{2^n} \leq 4 \cdot 2^{-k} \to 0$, luego esta sucesión es Cauchy y por completitud converge.

A continuación, veremos el **Lema de Lebesgue**, para probar que dado un recubrimiento abierto de X, $\{U_\alpha\}_{\alpha \in A}$, $\exists r > 0$, llamado número de Lebesgue, tal que $\forall x \in X$, $\exists \alpha \in A$ con $B_r(x) \subset U_\alpha$. Suponemos que tal numero no existe, y entonces podemos tomar x_n tal que $B_{\frac{1}{n}}(x_n) \not\subseteq U_\alpha$, para ningún $\alpha \in A$. Hemos probado que se tiene (3), luego la sucesión $\{x_n\}$ admite $\{y_n\}$ subsucesión convergente a y. Se tendrá que $B_{\frac{2}{n_1}}(x_0) \subset U_\alpha$ para ciertos n_1 y α al ser un abierto. Además, habrá un n_2 tal que si $n > n_2$, $d(y_n, y) < \frac{1}{n_1}$. Con ello, si $n > \max\{n_1, n_2\}$, tenemos que, dado $z \in B(y_n, \frac{1}{n})$, entonces $d(y, z) \leq d(y, x_n) + d(x_n, z) < \frac{1}{n_1} + \frac{1}{n} \leq \frac{2}{n_1}$. Pero entonces $z \in B_{\frac{2}{n_1}}(x_0) \subset U_\alpha$, luego $B(y_n, \frac{1}{n}) \subset U_\alpha$ lo que contradice la construcción de y_n.

Por tanto, finalmente, si $\{U_\alpha\}_\alpha$ recubre a X, y r es su número de Lebesgue, al ser totalmente acotado, $X = \bigcup_1^n B_r(x_i) \subset \bigcup_1^n U_{\alpha_i}$ para ciertos α_i, lo que termina las demostraciones. □

Proposición 31. *Sea $f : X \mapsto Y$ continua, donde (X, d_X) e (Y, d_Y) son métricos y X es compacto. Entonces $f(X)$ también lo es.*

Demostración. Supongamos que $\{U_\alpha\}_\alpha$ recubren a $f(X)$. Entonces, $X = f^{-1}(f(X)) \subset f^{-1}(\bigcup_\alpha U_\alpha) = \bigcup_\alpha f^{-1}(U_\alpha)$. Como f es continua, cada $f^{-1}(U_\alpha)$ es abierto, de tal modo que $X = \bigcup_1^n f^{-1}(U_i)$, luego $f(X) = f(\bigcup_1^n f^{-1}(U_i)) = \bigcup_1^n f(f^{-1}(U_i)) \subset \bigcup_1^n U_i$. □

Proposición 32 (Heine-Borel)**.** *Sea $(\mathbb{R}^n, d_\infty)$ el espacio métrico $\mathbb{R}^n$ con la distancia que induce la norma $\|\cdot\|_\infty$. Entonces $K \subset \mathbb{R}^n$ es compacto $\iff$ K es cerrado y acotado.*

Demostración. $\Longrightarrow$ es inmediata dado que hemos visto que, en general, debe ser cerrado y además totalmente acotado, de tal modo que es acotado. Para $\Longleftarrow$, comencemos con $n = 1$. Tenemos que $K \subset [-R, R]$, al ser acotado, y dado $\epsilon > 0$, lo vamos a trocear en N intervalos de la forma $I_j = [-R + \frac{2R}{N}j, -R + \frac{2R}{N}(j+1)]$ con cada $j \in \{0, \ldots, N-1\}$. Tomamos N para que $|I_j| = \frac{2R}{N} < \epsilon$. Si ahora cogemos $\tilde{I}_j$ como el abierto de mismo centro que I_j y norma 2ϵ, es claro que $K \subset \bigcup_j \tilde{I}_j$, luego es totalmente acotado. Ahora, dada $(x_n) \in \mathbb{N}$ de Cauchy en K, como $\mathbb{R}$ es completo, converge y dentro de K al ser cerrado. Así K es completo y hemos acabado. Si $n > 1$, la idea es similar. Habrá un $R > 0$ con $\|x\|_\infty \leq R \ \forall x \in \mathbb{R}^n$, y si $\epsilon > 0$, cubrimos el cubo $[-R, R]^n$ que acota a K mediante N^n sub-cubos de lado $\frac{2R}{N} \leq \epsilon$. Rodeamos a estos sub-cubos de bolas abiertas de radio 2ϵ, que cubren K por tanto, y es totalmente acotado. Por otro lado si $(x_n)_n \in K$ es Cauchy, sus sucesiones de coordenadas lo son, convergen todas ellas y por tanto $(x_n)_n$ también, dentro de K, por cierre. □

Proposición 33 (Weierstrass)**.** *Sea (X, d) métrico y $f : K \subset X \mapsto \mathbb{R}$ con K compacto, f continua y $K \neq \emptyset$. Entonces f alcanza máximo y mínimo en K.*

Demostración. Como $f(K)$ es compacto al ser f continua, está acotado y por tanto $\exists s = \sup f(X)$. Podemos construir la sucesión $\{s_n\}_n \subset f(K)$ con $s_n \to s$. Si ahora $a_n \in f^{-1}(s_n) \subset K$, entonces, por compacidad, hay una subsucesión $\{a_{n_j}\}_j$ tal que $a_{n_j} \to a \in K$. Por continuidad, $f(a_{n_j}) \to f(a)$, pero por otro lado $f(a_{n_j}) = s_{n_j} \to s$, luego $f(a) = s$. Análogo para el mínimo. □

El teorema de Heine-Borel aplica para cualquier distancia proveniente de normas. Eso lo vamos a ver introduciendo un nuevo concepto:

Definición 34. Si E es un $\mathbb{R}$-espacio y n_a, n_b dos normas, se dice que **son equivalentes** y se denota $n_a \sim n_b$ siempre que $\exists c, C$ con $0 < c \leq C$, finitas, tales que $cn_b(x) \leq n_a(x) \leq Cn_b(x)$, $\forall x \in X$.

Proposición 34. *La relación $\sim$ entre normas es de equivalencia. Asimismo, si $n_a \sim n_b$, entonces U es abierto bajo n_a $\iff$ U es abierto bajo n_b.*

Demostración. Está claro que $n_a \sim n_a$. Si se tiene $cn_a \leq n_b \leq Cn_a$, entonces $C^{-1}n_b \leq n_a \leq c^{-1}n_b$. Finalmente si $cn_b \leq n_a \leq Cn_b$ y $c'n_c \leq n_b \leq C'n_c$, entonces $cc'n_c \leq n_a \leq CC'n_c$. Por tanto es de equivalencia. Para la segunda afirmación supongamos que $n_a \sim n_b$. Sea $r > 0$, basta con probar que $B^a_{r'}(x) \subset B^b_r(x)$, para cierto r'. Tomamos $r' < cr$. Si $y \in B^a_{r'}(x)$, entonces $n_a(x-y) < r' < cr$, por tanto $n_b(x-y) < c^{-1}n_a(x-y) < r$. Como la relación es simétrica se tiene la otra implicación. □

Por lo tanto, si dos normas equivalen, el espacio métrico que inducen goza de la misma topología (mismos abiertos, pero esto conlleva mismos cerrados, compactos...). En $\mathbb{R}^n$ en particular se tiene que **todas** las normas equivalen, por tanto el teorema de Heine-Borel se verifica en general. Para probar esto, vamos a comenzar con unas observaciones:

Definición 35. Sean (X, d_X) e (Y, d_Y) espacios métricos. Se dice que $f : X \mapsto Y$ es una función **Lipschitz** si $\exists K > 0$ tal que $d_Y(f(x), f(y)) \leq Kd_X(x, y)$, $\forall x, y \in X$.

Es fácil probar que las funciones Lipschitz son de hecho uniformemente continuas y por tanto continuas. La norma es una de ellas:

Proposición 35. *La función $f : E \mapsto \mathbb{R}$ donde E es un $\mathbb{R}$-espacio normado, dada por $f(x) = \|x\|$, es Lipschitz para $K = 1$ en $(E, \|\cdot\|)$.*

Demostración. Lo único que hay que ver es que $|\,\|x\| - \|y\|\,| \leq \|x - y\|$. Pero $\|x\| = \|x + y - y\| \leq \|x - y\| + \|y\|$, de donde $\|x\| - \|y\| \leq \|x - y\|$. Revirtiendo la x y la y se llega a la otra desigualdad que prueba lo que queremos. $\square$

Con esto:

Proposición 36. *En $\mathbb{R}^n$ todas las normas son equivalentes.*

Demostración. Probaremos que todas equivalen a $\|\cdot\|_\infty$. Sea $\|\cdot\|$ una norma arbitraria de $\mathbb{R}^n$. Ya hemos visto que es Lipschitz en $(\mathbb{R}^n, \|\cdot\|)$. Vamos a ver que también es Lipschitz en $(\mathbb{R}^n, \|\cdot\|_\infty)$. Efectivamente $\|x\| = \|\sum x_j e_j\| \leq \sum |x_j|\, \|e_j\| \leq \|x\|_\infty \sum \|e_j\|$, de tal modo que $\|x\| \leq C\,\|x\|_\infty$. Pero entonces $|\,\|x\| - \|y\|\,| \leq \|x - y\| \leq C\,\|x - y\|_\infty$, de tal modo que tenemos lo que queremos.

Ahora consideremos $K = f^{-1}(\{1\})$ donde $f(x) = \|x\|_\infty$. En $(R^n, \|\cdot\|_\infty)$, tenemos que f es continua (Es Lipschitz), luego K es cerrado. Por definición, si $k \in K$, $\|x\|_\infty = 1$, luego allí es acotado también. Aplicando el Teorema de Heine-Borel, K es compacto ahí. Ahora sea $g(x) = \|x\|$ que hemos visto es Lipschitz en $(\mathbb{R}^n, \|\cdot\|_\infty)$. Por tanto, es continua en K y alcanza máximo y mínimo tales que $c_1 \leq \|x\| \leq c_2$, si $x \in K$.

Finalmente sea $x \in \mathbb{R}^n$ arbitrario. Entonces $x = \|x\|_\infty \frac{x}{\|x\|_\infty}$, donde el vector normalizado está en K, con lo que:

$$c_1\,\|x\|_\infty \leq \left\| \|x\|_\infty \frac{x}{\|x\|_\infty} \right\| \leq c_2\,\|x\|_\infty$$

Y por tanto hemos acabado. $\square$

Proposición 37. *Sean (X, d_x), (Y, d_y), (Z, d_z) métricos. Si $f : A \subset X \mapsto Y$ continua, $g : B \subset Y \mapsto Z$ continua y $f(A) \subset B$, entonces la composición $g \circ f : A \mapsto Z$ es continua.*

Demostración. Supongamos que U es un abierto de Z. Está claro que $(g \circ f)^{-1}(U) = f^{-1}(g^{-1}(U))$, que es abierto en A al ser g, f, continuas. $\square$

Proposición 38. *En el producto consigo mismo de (X, d) métrico definimos $d_2 : X^2 \times X^2 \mapsto \mathbb{R}$ la distancia $d_2((x, y), (x', y')) = d(x, x') + d(y, y')$. Es fácil comprobar que verifica las propiedades de distancia. En ese caso, la función d es Lipschitz entre (X^2, d_2) y $(\mathbb{R}, |\cdot|)$ y por tanto es continua en (X^2, d_2).*

Demostración. Como $d(x, y) \leq d(x, x') + d(x', y') + d(y, y')$, se tiene que $d(x, y) - d(x', y') \leq d(x, x') + d(y, y') = d_2((x, x'), (y, y'))$. Invirtiendo (x, y) e (x', y'), se tiene de hecho que $|d(x, y) - d(x', y')| \leq d_2((x, x'), (y, y'))$. $\square$

1.7.2. Continuidad uniforme

Definición 36. Sean (X, d_x) e (Y, d_y) métricos. En ese caso $f : A \subset X \mapsto Y$ es **uniformemente continua** en A si $\forall \epsilon > 0\ \exists \delta > 0$ tal que $\forall x, y \in A$ se tiene $d_x(x, y) < \delta \implies d_y(f(x), f(y)) < \epsilon$.

A diferencia de la continuidad puntual, donde podíamos escoger un δ para cada ϵ pero también para cada $x_0 \in A$, aquí el δ ha de ser global a A y garantizar que cualesquiera puntos cercanos más que δ hacen que la función está cercana más que ϵ. Esta observación implica que la continuidad uniforme es más fuerte que la puntual.

Una manera de ver la continuidad uniforme es a través de la existencia de una función que nos relaciona la distancia de las imágenes con la distancia de los puntos:

Definición 37. Si (X, d_x), (Y, d_y) son métricos y $f : A \subset X \mapsto Y$ es uniformemente continua en A, definimos su **módulo de continuidad** por $\varphi : \mathbb{R}^+ \mapsto \mathbb{R}^+$ dada por:

$$\varphi(t) := \sup\{d_y(f(x), f(y)) : d_x(x, y) \leq t\}$$

Proposición 39. *Si (X, d_x), (Y, d_y) son métricos y $f : A \subset X \mapsto Y$ es uniformemente continua en A, se tienen:*

1. $\forall x, y \in X$, $d_y(f(x), f(y)) \leq \varphi(d_x(x, y))$.
2. *φ es no decreciente.*
3. $\lim_{x \to 0^+} \varphi(x) = 0$.

Además, si fijada $f : A \subset X \mapsto Y$ se tiene una $\varphi : \mathbb{R}^+ \mapsto \mathbb{R}^+$ con estas características, entonces f es uniformemente continua.

Demostración. $d_Y(f(x), f(y)) \leq \sup\{d_y(f(u), f(v)) : d_x(u, v) \leq d_x(x, y)\}$ dado que pertenece a dicho conjunto, pero el elemento de la derecha es justo $\varphi(d_x(x, y))$. Por otro lado, si $r \leq t$, y $\varphi(x) := \sup A_x$, está claro que $A_r \subset A_t$, pero entonces $\sup A_r \leq \sup A_t$ y por tanto $\varphi(r) \leq \varphi(t)$. Para 3, si $\epsilon > 0$, sabemos que $\exists \delta$ tal que si $d(x, y) \leq \delta$, $d(f(x), f(y)) \leq \epsilon$. Es decir, $\varphi(\frac{\delta}{2}) \leq \epsilon$. Por ser no decreciente, si $0 \leq t \leq \frac{\delta}{2}$, entonces $\varphi(t) \leq \epsilon$.

Para la parte final (recíproco), veamos que si $\epsilon > 0$, entonces $\exists \delta > 0$ tal que $0 < t < \delta$ implica $\varphi(\delta) \leq \epsilon$. Si $d(x, y) \leq \delta$, por el no decrecimiento, entonces $\varphi(d(x, y)) \leq \epsilon$ y por la propiedad 1, inmediatamente $d_Y(f(x), f(y)) \leq \epsilon$. □

Proposición 40. *Si $(E, \|\cdot\|)$ es un espacio normado y $f : \Omega \subset E \mapsto X$ es continua, donde (X, d) es métrico y Ω es convexo, entonces φ, su módulo de continuidad, es subaditivo. Es decir: $\varphi(t + s) \leq \varphi(t) + \varphi(s)$, $\forall t, s \in [0, \infty)$.*

Demostración. Tomamos $u, v \in \Omega$ tales que $\|u - v\| \leq t + s$. Entonces, si definimos $\omega := \frac{s}{t+s}u + \frac{t}{t+s}v$, es inmediato que $\|u - \omega\| \leq t$ y $\|v - \omega\| \leq s$. Tenemos por convexidad que $\omega \in \Omega$. Pero entonces $d_X(f(u), f(v)) \leq d_X(f(u), f(\omega)) + d_X(f(v), f(\omega)) \leq \varphi(t) + \varphi(s)$, es decir, esta cantidad acota superiormente al conjunto del que $\varphi(t + s)$ es supremo, luego se tiene lo deseado. □

Definición 38. Sea (X, d) un espacio métrico y $A \subset X$. Se define $d(x, A) = \inf\{d(x, y) : y \in A\}$.

Proposición 41. *Se tiene que $d(\cdot, A)$ es Lipschitz.*

Demostración. Sea $a \in A$. Tenemos que $d(x, y) \geq d(x, a) - d(a, y) \geq d(x, A) - d(a, y)$, de tal modo que $d(a, y) \geq d(x, A) - d(x, y)$. Por definición de ínfimo, $d(y, A) \geq d(x, A) - d(x, y)$ luego $d(x, y) \geq d(x, A) - d(y, A)$. Si invertimos la x y la y se llega a la desigualdad con el signo cambiado, luego $|d(x, A) - d(y, A)| \leq d(x, y)$. □

Algunas propiedades de la distancia a conjunto son:

Proposición 42. *Sea (X, d) métrico, con $A \subset X$. Se cumple:*

1. $d(x, A) = 0 \iff x \in \overline{A}$.
2. *$\forall t > 0$, definimos $A_t = \{x \in X : d(x, A) < t\}$. Los A_t son abiertos y anidados: $\overline{A} \subset A_r \subset A_s$ si $r \leq s$.*
3. $\bigcap_{t>0} A_t = \overline{A}$.
4. *Si $x \notin \overline{A}$, $\exists U, V$ abiertos, con $\overline{A} \subset U$ y $x \in V$, tales que $U \cap V = \emptyset$.*

5. *Si $A, K \subset X$, con K compacto y $K \cap \overline{A} = \emptyset$, entonces $d(A,K) = \text{ínf}\{d(x,y) : x \in K, y \in A\}$ verifica $d(A,K) > 0$.*

Demostración. Supongamos que $d(x,A) = 0$. Entonces $d(x,A) < \epsilon$ si $\epsilon > 0$. Esto quiere decir que $d(x, y_\epsilon) < \epsilon$ para cierto y_ϵ de A. Pero entonces $y_\epsilon \in B_\epsilon(x) \cap A$, y por tanto $x \in \overline{A}$. A la inversa, si $x \in \overline{A}$, entonces $B_\epsilon(x) \cap A \neq \emptyset$, sea cual sea $\epsilon > 0$, luego si $y \in B_\epsilon(x) \cap A$, tenemos que $d(x,y) < \epsilon$ con $y \in A$, y por tanto $d(x,A) < \epsilon$, luego no queda otra que $d(x,A) = 0$.

Para (2), como d es Lipschitz, es continua, y de hecho $A_t = d^{-1}((-\infty, t))$, luego es abierto. La cadena de inclusiones viene de que $(-\infty, 0] \subset (-\infty, r) \subset (-\infty, s)$.

Para (3), tenemos que, claramente, $\overline{A} \subset \bigcap_{t>0} A_t$. Por otro lado, si $x \notin \overline{A}$, sigue que $d_A(x) > 0$ (por 1), y entonces $x \notin A_t$ si $t < d_A(x)$.

Para (4), si $x \notin \overline{A}$, entonces $\delta = d_A(x) > 0$. Ponemos $U := A_{\frac{\delta}{2}}$ y $V := B_{\frac{\delta}{2}}(x)$. Entonces, si $y \in U \cap V$, tenemos que $\delta \leq d(x,A) \leq d(y,A) + d(y,x) < \frac{\delta}{2} + \frac{\delta}{2} = \delta$, lo cual es una contradicción.

Para (5), $d(A,K) = \text{ínf}\{\text{ínf}\{d(x,y) : y \in A\} : x \in K\} = \text{ínf}\{d(x,A) : x \in K\}$. Como $\overline{A} \cap K = 0$, tenemos que $d(x,A) > 0$. Como d es continua en K alcanza un mínimo m, de donde sigue que $d(A,K) = d(m,A) > 0$. □

Proposición 43. *Se tiene que $d(x,A) = d(x,\overline{A})$.*

Demostración. Como $A \subset \overline{A}$, está claro que $d(x,\overline{A}) \leq d(x,A)$. Por otro lado, si $\epsilon > 0$, sabemos que $\exists y \in \overline{A}$ tal que $d(x,y) < d(x,\overline{A}) - \epsilon$. También, por ser la clausura, podemos tomar $y' \in A$ tal que $d(y,y') < \epsilon$. Por tanto, $d(x,y') \leq d(x,y) + d(y,y') \leq d(x,\overline{A}) + 2\epsilon$, pero como $d(x,A) \leq d(x,y')$, tenemos que $d(x,A) \leq d(x,\overline{A}) + 2\epsilon$ y por tanto hemos acabado. □

Teorema 2 (Heine, Cantor)**.** *Sean (X, d_x) e (Y, d_y) espacios métricos y sea $f : X \mapsto Y$ continua, con X compacto. Entonces f es uniformemente continua.*

Demostración. Supongamos que no fuese así. Entonces $\exists \epsilon_0 > 0$ tal que $\forall \delta > 0$, tenemos $x_\delta, y_\delta \in X$ con $d_x(x_\delta, y_\delta) < \delta$ pero $d_y(f(x_\delta), f(y_\delta)) \geq \epsilon_0$. Sea $x_n := x_{\delta = \frac{1}{n}}$ e $y_n := y_{\delta=\frac{1}{n}}$. Como X es compacto, la sucesión x_n admite la subsucesión convergente $\{x_{n_j}\}_j$ con $\lim\limits_{j\to\infty} x_{n_j} = p$. Se afirma que entonces $\exists \lim\limits_{j\to\infty} y_{n_j} = p$. Efectivamente, $d(y_{n_j}, p) \leq d(y_{n_j}, x_{n_j}) + d(x_{n_j}, p) \leq \frac{1}{n_j} + \epsilon$ para j suficientemente grande, luego $y_{n_j} \to p$.

Entonces, como f es continua, $f(x_{n_j}) \to f(p)$ y $f(y_{n_j}) \to f(p)$, pero esto contradice que $d(f(x_{n_j}), f(y_{n_j})) \geq \epsilon_0$. □

1.7.3. Conexión

Definición 39. Sea (X,d) métrico y $A \subset X$. Se dice que A es **conexo** si $\nexists U, V \subset X$ que verifiquen $U \cap V = \emptyset$, $A \subset U \cup V$ y $A \cap U$, $V \cap U \neq \emptyset$. Tales conjuntos, si existiesen, se denominan **separación** de A.

Definición 40. Sea (X,d) métrico y $A \subset X$. Se dice que A es **conexo por arcos o arco-conexo** si $\forall a, b \in A$, $\exists \gamma : [0,1] \mapsto X$ continua tal que $\gamma(0) = a$, $\gamma(1) = b$ y $\gamma(t) \in A$ $\forall t \in [0,1]$.

Definición 41. Sea $(E, \|\cdot\|_E)$ un espacio vectorial normado y $A \subset E$. Se dice que A es **conexo por poligonales** si $\forall a, b \in A$, $\exists p_0, \ldots, p_n$ con $p_0 = a$, $p_n = b$ y $[p_i, p_{i+1}] \subset A$ para todo $i \in \{0, \ldots, n-1\}$

Proposición 44.
1. *A conexo por arcos $\implies$ A conexo.*
2. *Si E es normado, $A \subset E$ conexo por poligonales $\implies$ A conexo por arcos.*
3. *Si $U \subset E$, U abierto, E normado, entonces $U \implies U$ conexo por poligonales.*

Demostración. Para 1, supongamos que A fuese conexo por arcos pero no conexo. Tomamos la separación U, V de A y definimos $F(x) = \begin{cases} 0 & \text{si } x \in U \\ 1 & \text{si } x \in V \end{cases}$ que es continua por ser constante en U y en V y ser estos abiertos disjuntos. Sea $x_0 \in A \cap U$ y $x_1 \in A \cap V$. Como es conexo por arcos, tenemos el arco $\gamma : [0,1] \mapsto A$ continuo que los une. Sea $\varphi : [0,1] \mapsto \mathbb{R}$ dada por $\varphi = F \circ \gamma$. Debe ser continua por ser composición de continuas, y además $\varphi(0) = 0$ y $\varphi(1) = 1$ luego $Im\varphi = \{0,1\}$, contradiciendo el teorema de valores intermedios.

Para 2, si $p, q \in A$, tomamos $p_0, \ldots, p_n$ los puntos de la poligonal que une a p con q y definimos un camino continuo que es exactamente la poligonal: $\gamma(t) = (1 - n(t - \frac{j}{n}))p_j + n(t - \frac{j}{n})p_{j+1}$ cuando $t \in [\frac{j}{n}, \frac{j+1}{n}]$, y entonces es conexo por arcos.

Para 3, tomamos $U \subset E$ abierto conexo. La relación en U dada por $p \sim q \iff$ existe una poligonal entre p y q es de equivalencia (es fácil comprobarlo). Ahora fijamos $x_0 \in U$ y escribimos $V = [x_0] = \{x \in U : x \sim x_0\}$. Lo que vamos a ver es que V es tanto abierto como cerrado en U, luego V^c, V sería una separación de U siempre que $V \neq U$, lo que sabemos que no puede ocurrir, con lo que concluiremos que $V = U$ y por tanto todos los puntos se conectan por poligonales.

V será abierto si y solo si $V = U \cap V'$ con V' abierto en E, es decir si y solo si es abierto en E (dado que U es abierto). Dado $x \in V$, como $x \in U$, tenemos la bola $B_r(x) \subset V$, y dado $x_1 \in B_r(x)$, como $x \sim x_1$ al ser $B_r(x)$ convexa, entonces tenemos que $x_1 \in V$ luego $B_r(x) \subset V$ y es abierto.

Para ver que es cerrado, damos $\{x_n\} \subset V$ con $x_n \to x \in U$. Tenemos que $B_r(x) \subset U$ por ser abierto, y para n grande sabemos que también $x_n \in B_r(x)$ luego $[x_n, x] \subset U$ con lo cual $x \sim x_n$ y entonces $x \in V$. □

Proposición 45. *Sea $\{C_\alpha\}_\alpha$ una familia arbitraria de conexos de (X,d) con $\bigcap C_\alpha \neq 0$. Entonces $C = \bigcup C_\alpha$ también es conexo en X.*

Demostración. Si no lo fuese, tomamos U, V una separación. Sea $p \in \bigcap C_\alpha$. Entonces $\forall \alpha \in A$, $p \in C_\alpha \cap U$, luego esa intersección es no vacía. Como $C \cap V \neq \emptyset$, y $C \cap V = \bigcup (C_\alpha \cap V)$, entonces $C_{\alpha_0} \cap V \neq \emptyset$ luego resulta que U, V también son una separación de C_{α_0} lo cual es contradictorio. □

Proposición 46. *En (X,d) métrico, la relación en $A \subset X$ dada por $x \sim y \iff x, y \in C$ para algún conexo C, es de equivalencia.*

Demostración. Viendo que $\{x\}$ es conexo para cualquier x, la reflexividad y la simetría son inmediatas. Para la transitividad: si $x \sim y$ y $y \sim z$, y C_1,C_2 son los conexos que los contienen, como y está en ambos, $C_1 \cup C_2$ es conexo y x, z pertenecen ahí luego $x \sim z$. □

Definición 42. Si (X,d) es métrico y $A \subset X$, las **componentes conexas** de A son las clases de equivalencia bajo la relación anterior. Si definimos otra relación, $x \sim y \iff \exists \gamma : [0,1] \mapsto A$ camino continuo de x a y, su conjunto cociente está constituido por las **componentes conexas por arcos**.

Proposición 47. *Si (X, d_x) es métrico y $A \subset X$ es conexo, y la función $f : A \mapsto Y$, donde (Y, d_y) es métrico, es continua, entonces $f(A)$ es conexo.*

Demostración. Supongamos que no lo fuese y sea U, V una separación. Como f es continua, $f^{-1}(U)$ y $f^{-1}(V)$ son abiertos, disjuntos al serlo U, V, como $f(A) \subset U \cup V$, tenemos que $A \subset f^{-1}(f(A)) \subset f^{-1}(U) \cup f^{-1}(V)$ e intersecan con A, contradiciendo que A sea conexo. □

Proposición 48. *$A \subset \mathbb{R}$ es conexo $\iff$ A es un intervalo.*

Demostración. $\Longleftarrow$ se tiene dado que intervalo equivale a convexo en $\mathbb{R}$. Por otro lado, para $\Longrightarrow$, supongamos $A \neq \emptyset$ y no es un intervalo. Como $\{x\}$ es un intervalo, debe ser que A tenga al menos 2 puntos, y como no es convexo, debe haber $x, y \in A$ con $[x,y] \nsubseteq A$. Sea $z \in [x,y] \backslash A$. Entonces $U = (-\infty, z)$ y $V = (z, \infty)$ es separación de A, en contradicción con que sea conexo. □

2. Diferenciabilidad

Definición 43. Sean $(E, \|\cdot\|_E)$ y $(F, \|\cdot\|_F)$ espacios vectoriales normados y $\Omega \subset E$ abierto. Entonces, se dice que $f : \Omega \mapsto F$ **es diferenciable** en $x_0 \in \Omega$ si $\exists T : E \mapsto F$ lineal, acotada, tal que:

$$\lim_{x \to x_0} \frac{\|f(x) - f(x_0) - T(x - x_0)\|_F}{\|x - x_0\|_E} = 0$$

O lo que es lo mismo, $f(x_0 + h) - f(x_0) - Th = o(h)$. En este caso, el **diferencial** de f en x_0 es la aplicación T y se denota por $(df)_{x_0}$.

Se debe pensar el diferencial como una buena estimación lineal de f en x_0, y como una aplicación que nos asigna vectores velocidad de paso por x_0 en vectores velocidad de paso por $f(x_0)$.

Proposición 49 (Unicidad)**.** *Si T, S verifican la condición de diferenciablidad para $f : E \mapsto F$, entonces $T = S$.*

Demostración. Tenemos que $\frac{\|Th-Sh\|_F}{\|h\|_E} \le \frac{\|f(x_0+h)-f(x_0)-Th\|_F}{\|h\|_E} + \frac{\|f(x_0+h)-f(x_0)-Sh\|_F}{\|h\|_E}$ por la desigualdad triangular, es decir, $\lim\limits_{h\to 0} \frac{\|Th-Sh\|_F}{\|h\|_E} = 0$. Si ahora tenemos en cuenta que T es lineal y fijamos $y \in E$, $y \neq 0$, tenemos que: $0 = \lim\limits_{x\to 0} \frac{\|Txy-Sxy\|_F}{\|xy\|_E} = \lim\limits_{x\to 0} \frac{\|Ty-Sy\|_F}{\|y\|_E} = \frac{\|Ty-Sy\|_F}{\|y\|_E}$, luego tenemos que $\|Ty - Sy\|_F = 0$ y por tanto $Ty = Sy$, $\forall y \in E$. □

Proposición 50. *Si $f : E \mapsto F$ es diferenciable en x_0, entonces también es continua en x_0.*

Demostración.
$\|f(x) - f(x_0)\|_F \le \|f(x) - f(x_0) - T(x - x_0)\|_F + \|T(x - x_0)\|_F = \frac{\|f(x)-f(x_0)-T(x-x_0)\|_F}{\|x-x_0\|_E} \|x - x_0\|_E + \|T(x - x_0)\|_F \le \frac{\|f(x)-f(x_0)-T(x-x_0)\|_F}{\|x-x_0\|_E} \|x - x_0\|_E + \|T\| \|(x - x_0)\|_E$. Como el lado de la derecha tiende a 0 cuando $x \to x_0$, hemos acabado. □

Ya sabemos que en espacios finitos, un operador lineal siempre es acotado. El siguiente resultado nos caracteriza la acotación y la continuidad de otra manera:

Proposición 51. *Sean $(E, \|\cdot\|_E)$ y $(F, \|\cdot\|_F)$ espacios vectoriales normados y $T : E \mapsto F$ lineal. Equivalen:*

1. *T es continua en E.*
2. *T es continua en algún $x_0 \in E$.*
3. *T es acotada.*

Demostración. (1) $\implies$ (2) es por definición. Para (2) $\implies$ (1), veamos sea $x \in E$. Tenemos que $\|T(x + h) - Tx\|_F = \|T(x_0 + h + (x - x_0)) - T(x_0 + x - x_0)\|_F = \|T(x_0 + h) + T(x - x_0) - Tx_0 - T(x - x_0)\|_F = \|T(x_0 + h) - Tx_0\|_F \to 0$ si $h \to 0$ y hemos acabado.

Para (1) $\implies$ (3), definimos $V = B_1^F(0)$ y entonces $T^{-1}(V)$ es abierto. Eso quiere decir que $\exists\delta > 0$ tal que $B_\delta(0) \in T^{-1}(V)$, dado que $0 \in T^{-1}(V)$ al tenerse que $\|T0\|_F = \|0\|_F = 0$. Pero entonces, fijado $v \in E \backslash \{0\}$, tenemos que $\frac{v}{2\|v\|_E}\delta \in B_\delta(0)$, lo que indica que $\left\| T \frac{v}{2\|v\|_E}\delta \right\|_F \le 1$, es decir, por linealidad, $\|Tv\|_F \le \frac{2\|v\|_e}{\delta}$. Finalmente, para (3) $\implies$ (1), veamos que como $x - y \in E$, se tiene $\|Tx - Ty\|_F = \|T(x - y)\|_F \le M \|x - y\|_E$ luego es Lipschitz. □

Definición 44. Dada $f : U \subset \mathbb{R}^n \mapsto \mathbb{R}^m$, U abierto, se define su **derivada parcial i-ésima** en $x_0 \in U$ como $\frac{\partial f}{\partial x_i}(x_0) = \lim\limits_{t\to 0} \frac{f(x_0+te_i)-f(x_0)}{t}$, si existe.

Definición 45. Sea $f : U \subset \mathbb{R}^n \mapsto \mathbb{R}^m$ con U abierto y $x_0 \in U$. Pongamos que $f_i : \mathbb{R}^n \mapsto \mathbb{R}$ es la i-ésima componente de f. Supongamos que existen todas las derivadas parciales de la forma $\frac{\partial f_j}{\partial x_i}(x_0)$ con $j \in \{1, \ldots, m\}$, $i \in \{1, \ldots, m\}$ en x_0. Se define la **matriz jacobiana** como $Df(x_0) = (\frac{\partial f_j(x_0)}{\partial x_i})_{j,i}$.

La importancia de la matriz jacobiana es que permite dar explícitamente el diferencial de las funciones que sean diferenciables:

Proposición 52. *Sea $f : U \subset \mathbb{R}^n \mapsto \mathbb{R}^m$ con U abierto, diferenciable en $x_0 \in U$. Entonces el operador lineal $(df)_{x_0} : \mathbb{R}^n \mapsto \mathbb{R}^m$ tiene matriz $Df(x_0)$ en las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$.*

Demostración. Por la definición de derivada parcial, tenemos que $f_j(x_0+te_i) = f_j(x_0)+\frac{\partial f_j(x_0)}{\partial x_i}t+g_{ij}(t)$ donde $g_{ij}(t) = o(t)$. Entonces, tenemos que:

$$f(x_0 + te_i) = (f_1(x_0), \ldots, f_m(x_0)) + t\left(\frac{\partial f_1(x_0)}{\partial x_i}, \ldots, \frac{\partial f_m(x_0)}{\partial x_i}\right) + (g_{i1}(t), \ldots, g_{im}(t))$$

Si ponemos $G(t) := (g_{i1}(t), \ldots, g_{im}(t))$ es fácil ver que $G(t) = o(t)$. En efecto, $\lim_{t\to 0} \frac{\|G(t)\|_1}{t} = \lim_{t\to 0} \frac{|g_{i1}(t)|+\cdots+|g_{im}(t)|}{t} = 0$.

Atendiendo a la expresión original del diferencial, y viendo que si $t \to 0$, entonces $te_i \to 0$, se observa que debe ser $(df)_{x_0} \cdot e_i = \left(\frac{\partial f_1(x_0)}{\partial x_i}, \ldots, \frac{\partial f_m(x_0)}{\partial x_i}\right)$, que es lo que queríamos ver. □

Proposición 53. *Sean $(E, \|\cdot\|_E), (F, \|\cdot\|_F), (G, \|\cdot\|_G)$ espacios normados, $f : U \subset E \mapsto F$ y $g : V \subset F \mapsto G$ con U, V abiertos, y sea $x_0 \in U$ con $f(x_0) \in V$ tal que f es diferenciable en x_0 y g es diferenciable en $f(x_0)$. Entonces $g \circ f$ es diferenciable en x_0, y $(dg \circ f)_{x_0} = (dg)_{f(x_0)} \circ (df)_{x_0}$.*

Demostración. Primero veamos que tiene sentido considerar el diferencial de $g \circ f$ en x_0, dado que por ser V abierto, $B_\epsilon(f(x_0)) \subset V$ para cierto $\epsilon > 0$, pero por continuidad de f en x_0 tenemos que $f(B_\delta(x_0)) \subset B_\epsilon(f(x_0)) \subset V$ para un $\delta > 0$ luego $g \circ f$ existe en ese entorno de x_0.

Ahora, si $\|h\| < \delta$, entonces $x_0 + h \in B_\delta(x_0)$ y por tanto $f(x_0 + h) \in B_\epsilon(f(x_0))$. Si ponemos $f(x_0+h) = f(x_0)+w$ entonces estamos diciendo que $\|w\| < \epsilon$. Tenemos las siguientes igualdades a causa de la diferenciablidad:

$$\begin{cases} f(x_0 + h) = f(x_0) + (df)_{x_0}h + o_1(h) \\ g(f(x_0) + w) = g(f(x_0)) + (dg)_{f(x_0)}w + o_2(w) \end{cases}$$

Pero entonces $w = (df)_{x_0}h + o_1(h)$, luego:

$$g(f(x_0+h)) = g(f(x_0)+w) = g(f(x_0))+(dg)_{f(x_0)}w+o_2(w) = g(f(x_0))+(dg)_{f(x_0)}(df)_{x_0}h+(dg)_{f(x_0)}o_1(h)+o_2(w)$$

Como $(dg)_{f(x_0)}(df)_{x_0}$ es acotado por ser composición de operadores acotados, lo único que basta ver es que $(dg)_{f(x_0)}o_1(h) + o_2(w) = o(h)$. Por un lado:

$$\left\|(dg)_{f(x_0)}o_1(h)\right\| \leq M \|o_1(h)\| = o(h)$$

.

Y por otro:

$$\lim_{h\to 0} \frac{\|o_2(w)\|}{\|h\|} = \lim_{h\to 0} \frac{\|o_2(w)\|}{\|w\|}\frac{\|w\|}{\|h\|} = \lim_{h\to 0} \frac{\|o_2(w)\|}{\|w\|}\frac{\|(df)_{x_0}h + o_1(h)\|}{\|h\|} \leq \lim_{h\to 0} \frac{\|o_2(w)\|}{\|w\|}\cdot\left(M_2 + \frac{\|o_1(h)\|}{\|h\|}\right) = 0$$

Dado que por continuidad si $h \to 0$ también $w \to 0$. □

Definición 46. Sea $f : U \subset E \mapsto F$ con E, F, espacios normados y U abierto. Si $x_0 \in U$ y $v \in E$, se define la **derivada direccional en dirección de v, en** x_0, como $D_v f(x_0) = \lim\limits_{t \to 0} \frac{f(x_0+tv)-f(x_0)}{t}$, si existe.

Proposición 54. *Si f con las características anteriores es diferenciable en x_0, entonces $\exists D_v f(x_0) = (df)_{x_0} v$, $\forall v \in E$.*

Demostración. Si $\gamma : \mathbb{R} \mapsto E$ está dada por $\gamma(t) = x_0 + tv$, entonces es inmediato comprobar que $D_v f(x_0) = \frac{d}{dt}_{t=0}(f \circ \gamma)(t)$. Como $\gamma(t+h) = \gamma(t) + vh + 0$, está claro que $(d\gamma)_t \xi = v\xi$ para cualquier $t \in E$. Por tanto, usando la regla de la cadena, $D_v f(x_0) = (df)_{x_0}(d\gamma)_0 = (df)_{x_0} v$. $\square$

Esto nos dice además que si no existe una relación lineal entre $D_v f(x_0)$ y v, no puede ser diferenciable f en x_0. El recíproco en general no es cierto: hay funciones que tienen todas las derivadas direccionales, incluso guardando una relación lineal con el vector dirección, sin ser diferenciables.

Proposición 55. *Si $f : U \subset E \mapsto \mathbb{R}^n$, con E normado y U abierto, y ponemos $f = (f_1, f_2, \ldots, f_n)$, con cada $f_i : U \mapsto \mathbb{R}$, entonces, dado $x_0 \in U$, se tiene que f es diferenciable en $x_0 \iff f_i$ es diferenciable en x_0 $\forall i \in \{1, \ldots, n\}$.*

Además, en ese caso, $(df)_{x_0} h = ((df_1)_{x_0} h, \ldots, (df_n)_{x_0} h)$.

Demostración. Para $\implies$, basta con definir el operador lineal $\pi_j : \mathbb{R}^n \mapsto R$ con $(x_1, \ldots, x_n) \mapsto x_j$. Entonces observamos que $f_j = \pi_j \circ f$. Basta con usar la regla de la cadena: $d(f_j)_{x_0} = (d\pi_j)_{f(x_0)} \circ (df)_{x_0} = \pi_j \circ (df)_{x_0}$, que además nos da la expresión que queríamos (es la j-ésima componente del diferencial de f).

Para $\impliedby$, si $h \in E$ verifica $x_0 + h \in U$, entonces:

$$f_j(x_0 + h) = f_j(x_0) + (df_j)_{x_0} h + o_j(h)$$

Y por lo tanto:

$$f(x_0 + h) = f(x_0) + ((df_1)_{x_0} h, \ldots, (df_n)_{x_0} h) + (o_1(h), \ldots, o_n(h))$$

Si denotamos $((df_1)_{x_0} h, \ldots, (df_n)_{x_0} h) = Th$, basta ver que T es acotada ya que es claramente lineal. $\|Th\|_1 = \sum |(df_j)_{x_0} h| \leq \|h\| \sum M_j$ al ser las diferenciales de cada componente acotadas. Ya solo queda ver que:
$\frac{\|(o_1(h), \ldots, o_n(h))\|_1}{\|h\|} = \frac{\sum |o_j(h)|}{\|h\|} \to 0$ luego el término de error es $o(h)$ y hemos acabado. $\square$

Proposición 56. *Sea $f : U \subset \mathbb{R}^n \mapsto \mathbb{R}^m$ con U abierto. Si $\forall x \in U$, $\exists \partial_i f_j(x)$ la i-ésima derivada parcial de la j-ésima componente en todos los i, j y son continuas en U, entonces f es diferenciable en U.*

Demostración. Por la proposición anterior, basta con suponer que $f : \mathbb{R}^n \mapsto \mathbb{R}$. Sea $x_0 = (a_1, a_2, \ldots, a_n) \in U$. Como U es abierto, el cuadrado $B_\delta(x_0)^\infty = x_0 + (-\delta, \delta)^n$ está en U para cierto δ. Sea $h = (h_1, h_2, \ldots, h_n) \in (-\delta, \delta)^n$. Entonces tiene sentido considerar $f(x_0 + h)$. Vamos a desplazarnos de x_0 a h en incrementos paralelos a los ejes, de la forma $\Delta_j = f(a_1 + h_1, \ldots, a_j + h_j, a_{j+1}, a_{j+2}, \ldots, a_n) - f(a_1 + h_1, \ldots, a_{j-1} + h_{j-1}, a_j, a_{j+1}, \ldots, a_n)$.

Entonces tenemos que $f(x_0 + h) - f(x_0) = \sum_{j=1}^n \Delta_j$. Como f es derivable coordenada a coordenada, podemos considerar Δ_j como una función de una variable (la j-ésima de f) y aplicar el teorema del valor medio para ver que:

$$f(x_0 + h) - f(x_0) = \sum_{j=1}^n \Delta_j = \sum_{j=1}^n \partial_j f(\xi_j) h_j$$

Donde $\xi_j \in [(a_1 + h_1, \ldots, a_j + h_j, a_{j+1}, \ldots, a_n), (a_1 + h_1, \ldots, a_{j-1} + h_{j-1}, a_j, \ldots, a_n)] \subset x_0 + (-\delta, \delta)^n$.

Pero entonces:

$$f(x_0+h)-f(x_0)-\sum_{j=1}^{n}\partial_j f(x_0)h_j=\sum_{j=1}^{n}(\partial_j f(\xi_j)-\partial_j f(x_0))h_j$$

Para concluir lo que se quería basta con ver que esta cantidad es $o(h)$. Pero se tiene que:

$$\frac{|\sum_{j=1}^{n}(\partial_j f(\xi_j)-\partial_j f(x_0))h_j|}{\|h\|_\infty}\leq \|h\|_\infty\frac{\left\|\sum_{j=1}^{n}(\partial_j f(\xi_j)-\partial_j f(x_0))\right\|_\infty}{\|h\|_\infty}=\left|\sum_{j=1}^{n}(\partial_j f(\xi_j)-\partial_j f(x_0))\right|$$

Y como cada $\partial_j f$ es continua, entonces, si $h\to 0$, tenemos $\xi_j\to x_0$ luego $\partial_j f(\xi_j)-\partial_j f(x_0)\to 0$ y hemos acabado. $\square$

Definición 47. Sea $f:U\subset\mathbb{R}^n\mapsto\mathbb{R}^m$, con U abierto. Si $\exists\partial_i f_j(x)$ en U para todos los i,j y son continuas, se dice que $f\in C^1(U)$.

Asimismo, si $k\in\mathbb{N}$, y existen y son continuas todas las $\partial_1^{\alpha_1}\dots\partial_n^{\alpha_n}f$ en U, con $\sum_1^n\alpha_j\leq k$, entonces se dice que $f\in C^k(U)$.

Finalmente, si $\forall k\in\mathbb{N}$, se tiene $f\in C^k(U)$, entonces se dice que $f\in C^\infty(U)$.

Proposición 57 (Schwarz)**.** *Si $f:U\subset\mathbb{R}^n\mapsto\mathbb{R}$, U abierto, y $f\in\mathcal{C}^2(U)$, entonces $\forall x\in U$ se verifica que $\partial^2_{ij}f(x)=\partial^2_{ji}f(x)$.*

Demostración. Sea $p\in U$ y $i<j$ (si son iguales es evidente). Como U es abierto, $B_\delta^\infty(p)\subset U$ para un $\delta>0$. Dado $r<\delta$ sea $C_r=p+te_i+se_j$ donde $|t|,|s|\leq r$. Es decir, son puntos donde las coordenadas distintas de i,j son las de p, y las otras verifican que $|x_i-p_i|,|x_j-p_j|\leq r$, o, lo que es lo mismo, es un cuadrado en $\mathbb{R}^n$ de lado $2r$. Entonces, como la derivada segunda es continua, aplicando Fubini:

$$\begin{aligned}&\iint_{C_r}\partial^2_{ij}f(x)dx_idx_j=\int_{p_j-r}^{p_j+r}\int_{p_i-r}^{p_i+r}\partial^2_{ij}f(x)dx_idx_j=\int_{-r}^{r}\partial_j f(p+te_i+se_j)|_{t=-r}^{r}ds=\\&=f(p+re_i+re_j)-f(p+re_i-re_j)-f(p-re_i+re_j)+f(p-re_i-re_j)\end{aligned}$$

Pero aplicándolo al revés llegamos a que $\iint_{C_r}\partial^2_{ji}f(x)dx_idx_j=\iint_{C_r}\partial^2_{ij}f(x)dx_idx_j$ y de hecho:

$\frac{1}{4r^2}\iint_{C_r}\partial^2_{ij}f(x)dx_idx_j=\frac{1}{4r^2}\iint_{C_r}\partial^2_{ji}f(x)dx_idx_j$. Ahora ya solo queda probar que $\lim_{r\to 0}\frac{1}{4r^2}\iint_{C_r}\partial^2_{ij}f(x)dx_idx_j=\partial^2_{ij}f(p)$ y, tomando límites a ambos lados, habremos acabado. Se tiene que:

$$\left|\frac{1}{4r^2}\iint_{C_r}\partial^2_{ij}f(x)dx_idx_j-\partial^2_{ij}f(p)\right|=\frac{1}{4r^2}\left|\iint_{C_r}\partial^2_{ij}f(x)-\partial^2_{ij}f(p)dx_idx_j\right|\leq\frac{1}{4r^2}\iint_{C_r}\left|\partial^2_{ij}f(x)-\partial^2_{ij}f(p)\right|dx_idx_j$$

Pero por continuidad de las derivadas parciales segundas, si r es pequeño, entonces ese término verifica que $\frac{1}{4r^2}\iint_{C_r}\left|\partial^2_{ij}f(x)-\partial^2_{ij}f(p)\right|dx_idx_j\leq\frac{1}{4r^2}\iint_{C_r}(\epsilon)dx_idx_j=\epsilon$ y hemos acabado. $\square$

Observación 3. Un argumento inductivo permite probar que si $f:U\subset\mathbb{R}^n\mapsto\mathbb{R}^m$ con U abierto y $f\in C^k(U)$, entonces todas las derivadas parciales k-ésimas (o de orden menor que k) no se ven afectadas por el orden en que se derive.

Definición 48. Si $f\in\mathcal{C}^2(U)$ y $x\in U$, se define el **hessiano** como $Hf(x)=\left(\partial^2_{ij}f(x)\right)_{i,j}$. Por el lema de Schwarz es simétrica. Induce una **aplicación bilineal hessiana**, $hess(f)_{x_0}$, dada por $h\to h^tHf(x_0)h$.

El hessiano, si existe, permite entre otras cosas refinar la aproximación incremental que hacía el diferencial:

Proposición 58. *Si $f : U \subset \mathbb{R}^n \mapsto \mathbb{R}$ con U abierto:*

1. *Si $f \in C^2(U)$ y $x_0 \in U$, y $h \in \mathbb{R}^n$ tal que $x_0 + h \in U$, entonces $f(x_0+h) = f(x_0) + (df)_{x_0}h + \frac{h^t Hf(x_0)h}{2} + o(\|h\|^2)$.*

2. *Si además $f \in C^3(U)$ entonces $f(x_0+h) = f(x_0) + (df)_{x_0}h + \frac{h^t Hf(x_0)h}{2} + O(\|h\|^3)$.*

Demostración. En primer lugar, $B_\delta(x_0) \subset U$. Si fijamos h con $\|h\| < \delta$ entonces $x_0 + th \in U$ para $t \in [0,1]$ y está bien definida $g(t) = f(x_0 + th)$ entre $[0,1]$ y $\mathbb{R}$. Como $g = f \circ \gamma$ donde $\gamma(t) = x_0 + th$, entonces $g \in \mathcal{C}^2([0,1])$ y por el teorema de Taylor:

$$g(1) = g(0) + g'(0) + \frac{g''(\xi)}{2} \qquad \xi \in [0,1]$$

Aplicando la regla de la cadena, $g'(0) = df_{\gamma(0)} \circ \gamma'(0) = \nabla f(x_0)h$. Aplicándola de nuevo, $g''(t) = \frac{d}{dt}\sum_{i=1}^n \partial_i f(x_0+th)h_i = \sum_{i=1}^n \nabla\partial_i f(x_0+th)hh_i = \sum_{i,j=1}^n \partial^2_{ji} f(x_0+th)h_j h_i = h^t Hf(x_0+th)h$.
Luego finalmente, tras sumar y restar $g''(0)$:

$$f(x_0+h) = f(x_0) + df_{x_0}h + \frac{1}{2}\left(h^t Hf(x_0+\xi h)h - h^t Hf(x_0)h\right) + \frac{1}{2}h^t Hf(x_0)h$$

Pero se tiene que:

$$\left|h^t\left(Hf(x_0+\xi h) - Hf(x_0)\right)h\right| \le \|h\|^2 \|Hf(x_0+\xi h) - Hf(x_0)\| = o(\|h\|^2)$$

Dado que si $h \to 0$, entonces $Hf(x_0+\xi h) \to Hf(x_0)$ al ser las segundas derivadas continuas.
Si además $f \in \mathcal{C}^3(U)$, entonces análogamente:

$$g(1) = g(0) + g'(0) + \frac{g''(0)}{2} + \frac{g'''(\xi)}{6} \qquad \xi \in [0,1]$$

Y solo hay que ver que $g'''(\xi) = \sum_{i,j,k=1}^n \partial^3_{ijk} f(x_0+th)h_i h_j h_k$. Luego, si $M = \max_{i,j,k}\sup_{z\in\overline{B}_{\|h\|}(x_0)} |\partial^3_{ijk}f(z)|$, que se alcanza al ser cada parcial continua y esa bola un compacto, entonces tenemos que $|g'''(\xi)| \le M\sum_{i,j,k=1}^n |h_i||h_j||h_k| = M\|h\|_1^3 \le MC^3\|h\|_2^3$, luego $g'''(\xi) = O(\|h\|^3)$. ☐

Proposición 59 (Valor Medio)**.** *Sea $f : U \subset E \mapsto \mathbb{R}$ con E un espacio vectorial normado y U abierto. Si, dados $p, q \in U$, se tiene que $[p,q] \subset U$, y f es diferenciable en $[p,q]$, entonces $f(p) - f(q) = \nabla f(r)(p-q)$ para un $r \in [p,q]$.*

Demostración. Damos $\gamma(t) = tp + (1-t)q$ en $[0,1] \mapsto U$. Entonces $h = f \circ \gamma$ es diferenciable, y $dh_t = \nabla f(\gamma(t))\gamma'(t)$. Si aplicamos el teorema de valor medio real, entonces $h(1) - h(0) = dh_\xi \cdot 1$ con $\xi \in [0,1]$, es decir, $f(p) - f(q) = \nabla f(\xi p + (1-\xi)q)(p-q)$, como se quería.

Proposición 60. *Si $f : U \subset E \mapsto \mathbb{R}^n$, E normado, $U \subset E$ abierto. Si $[p,q] \subset U$ y f es diferenciable en ese segmento, y tomamos un $\xi \in \mathbb{R}^n$, entonces, hay un $r \in [p,q]$ con :*

$$\langle \xi, f(p) - f(q)\rangle = \langle \xi, df_r(p-q)\rangle$$

Demostración. Si escribimos $h(x) = \langle \xi, f(x)\rangle$, es decir, $h = l \circ f$ donde $l(x) = \langle \xi, x\rangle$, entonces verifica las condiciones del teorema anterior, y lo único que hay que ver es que $\nabla h(r) = d(l)_{f(r)}\nabla f(r)$. Sin embargo, como l es lineal, entonces $d(l)_{f(r)} = l$. Por lo tanto tenemos que $\langle \xi, f(p)\rangle - \langle \xi, f(q)\rangle = l \circ \nabla f(r)(p-q) = \langle \xi, \nabla f(r)(p-q)\rangle$. ☐

Proposición 61. *Sea $f : U \subset E \mapsto \mathbb{R}^n$ con E normado y U abierto convexo. Si f es diferenciable en U, y $K = \sup\{\|Df(x)\|_2\, x \in U\}$, entonces f es Lipschitz de constante K.*

Demostración. Sean $p, q \in U$. Fijado $u \in \mathbb{R}^n$ ya sabemos que $\langle u, f(p) - f(q)\rangle = \langle u, Df(z)(p-q)\rangle$ con $z \in [p,q]$. Por otro lado sabemos que se puede escribir $\|v\|_2 = \max\limits_{\|u\|_2=1} |\langle u, v\rangle|$ (Cauchy-Schwarz) luego entonces $\|f(p)-f(q)\|_2 = \max\limits_{\|u\|_2=1} |\langle u, f(p)-f(q)\rangle| = \max\limits_{\|u\|_2=1} |\langle u, Df(z)(p-q)\rangle| \leq \|Df(z)(p-q)\|_2 = \|Df(z)\| \, \|p-q\|_E \leq K \|p-q\|_E$. □

Corolario. Si $Df(x) = 0$ para todo $x \in U$ en las hipótesis anteriores, entonces f es constante.

Definición 49. Sea $f : U \subset E \mapsto \mathbb{R}$ con E normado y U abierto. Se dice que f tiene un **punto crítico** en $x_0 \in U$ si $df_{x_0} = 0$.

Definición 50. Si $f : A \subset X \mapsto \mathbb{R}$ con (X, d) métrico y $x_0 \in A$, se dice que x_0 es **máximo local** (*resp. mínimo local*) si $\exists \delta > 0$ tal que si $x \in A$ se verifica $d(x, x_0) < \delta \implies f(x) \leq f(x_0)$ (*resp.* $f(x) \geq f(x_0)$). Si las desigualdades son estrictas, es un máximo local estricto (*resp. mínimo local estricto*).

Proposición 62. *Si $f : U \subset E \mapsto \mathbb{R}$ con U abierto en E normado, y f es diferenciable en U, entonces si $x_0 \in U$ es un extremo local, se tiene que x_0 es crítico.*

Demostración. Suponemos que x_0 es máximo siendo análogo para el mínimo. Fijamos $h \in E$. Se tiene $x_0 + th \in U$ si $\|th\| < \delta$ para cierto $\delta > 0$ por ser abierto, luego queda definida en un entorno del 0 la función: $\varphi(t) = f(x_0 + th) = f \circ \gamma(t)$ con $\gamma(t) = x_0 + th$. Como el 0 es máximo de φ por hipótesis, entonces $0 = \varphi'(0) = df_{x_0} h$. Como el h era arbitrario entonces $df_{x_0} = 0$. □

Proposición 63. *Sea $f : U \subset \mathbb{R}^n \mapsto \mathbb{R}$ con U abierto y $f \in \mathcal{C}^2(U)$.*

1. *Si x_0 es crítico no degenerado (es decir, $Hf(x_0)$ es no singular), entonces $\exists \delta > 0$ tal que si $x_1 \in U$ es otro punto crítico, se tiene $\|x_1 - x_0\| \geq \delta$.*

2. *Si x_0 es un máximo, se tiene que la aplicación Hessiana, $Hess(f)_{x_0}$ dada por $h \to h^t Hf(x_0) h$, verifica $Hess(f)_{x_0} h \leq 0$. Si es mínimo, se tiene $Hess(f)_{x_0} h \geq 0$.*

3. *Si x_0 es crítico y además $Hess(f) < 0$, entonces es máximo estricto. Si es crítico y $Hess(f) > 0$ entonces es mínimo estricto.*

Demostración. Como $Hf(x_0)$ es no singular, entonces $\|Hf(x_0)x\|_2 \geq c\|x\|_2$ $\forall x \in \mathbb{R}^n$ (es fácil comprobar que $c = \left\|Hf(x_0)^{-1} Hf(x_0)\right\|_2$ lo verifica). Tomamos $u \in \mathbb{R}^n$ con $\|u\|_2 = 1$, y consideramos $\varphi : [0,1] \mapsto \mathbb{R}$ dada por $t \mapsto \langle u, \nabla f(x_0 + th)\rangle$. Como $\nabla f \in C^1(U)$ entonces es diferenciable y $\varphi'(t) = \langle u, Hf(x_0 + th)h\rangle$. Como $\varphi(0) = 0$ usamos el teorema del valor medio y se tiene $\varphi(1) = \varphi'(\xi) \implies \langle u, \nabla f(x_0 + h)\rangle = \langle u, Hf(r)h\rangle$ con $r \in [x_0, x_0 + h]$. Para cierto u de esos se tiene $\langle u, Hf(x_0)h\rangle = \|Hf(x_0)h\|$, luego, finalmente, $\langle u, \nabla f(x_0 + h)\rangle = \|Hf(x_0)h\| + \langle u, (Hf(r) - Hf(x_0))h\rangle \geq c\|h\|_2 - |\langle u, (Hf(r) - Hf(x_0))h\rangle| \geq c\|h\|_2 - \|Hf(r) - Hf(x_0)\| \, \|h\|_2$. Por continuidad de Hf, para h pequeño, se tiene que $\langle u, \nabla f(x_0 + h)\rangle \geq \frac{c}{2}\|h\|_2 > 0$ luego $\nabla f(x_0 + h) \neq 0$.

Para la segunda afirmación, sea $\varphi(t) = f(x_0 + th)$ definida para un entorno del 0. Como 0 es máximo de φ, entonces $\varphi''(0) \leq 0 \implies h^t Hf(x_0) h \leq 0$. Análogo para el mínimo.

Para la tercera afirmación, supongamos que $\nabla f(x_0) = 0$ y que $Hf(x_0) > 0$. Entonces $h^t Hf(x_0) h \geq c\|h\|^2$ donde c es el mínimo autovalor de $Hf(x_0)$ que es positivo. Como $f(x_0 + h) = f(x_0) + \nabla f(x_0)h + \frac{h^t Hf(x_0)h}{2} + o(\|h\|^2) \geq f(x_0) + \frac{c}{2}\|h\|^2 + o(\|h\|^2)$, para un h cerca del 0, tenemos $|o(\|h\|^2)| \leq \frac{c}{4}\|h\|^2$ luego $f(x_0 + h) \geq f(x_0) + \frac{c}{4}\|h\|^2 > f(x_0)$ y análogo para el máximo. □

Proposición 64. *Sea $f : U \subset E \mapsto \mathbb{R}^n$ con E normado. Si U es abierto conexo y $\exists df(x) = 0$ en todo $x \in U$, entonces f es constante.*

Demostración. Dado $p \in U$ sea $V = \{x \in U : f(x) = f(p)\}$. Sea $x \in V$. Entonces $B_r(x) \subset U$ para cierto $r > 0$, y si $y \in B_r(x)$, como las funciones de diferencial nulo en un convexo son constantes, tenemos que $f(y) = f(x) = f(p)$ luego $B_r(x) \subset V$, luego V es abierto en X y por tanto en U al ser subconjunto, y U abierto. Por otro lado, si $\{x_n\} \subset V$ cumple $x_n \to x$, entonces $f(x_n) \to f(x)$ por continuidad de f, es decir, $f(p) \to f(x)$ luego $f(x) = f(p)$ y por tanto $x \in V$ y V es cerrado en X y por tanto en U. De esta manera, debe darse que $V = U$ por conexión (si no, $U \backslash V$ y V separarían U) y hemos acabado. $\square$

Definición 51. Sea $f : U \subset E \mapsto \mathbb{R}$ con E un $\mathbb{R}$-espacio vectorial cualquiera y U convexo. Se dice que f **es convexa** en U si $\forall x, y \in U$, $\forall t \in [0,1]$, se tiene $f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$.

Proposición 65. *Son equivalentes:*

1. *f es convexa en U.*
2. *$Epi(f) = \{(x,t) \in U \times \mathbb{R} : t \geq f(x)\}$ es convexo en $E \times \mathbb{R}$.*
3. *$\forall x_1, \dots x_n \in U$ y $t_1, \dots, t_n \geq 0$ con $\sum_i t_i = 1$, se tiene $f(\sum_i t_i x_i) \leq \sum_i t_i f(x_i)$.*

Definición 52. Sea E un $\mathbb{R}$-espacio y $U \subset E$. Entonces $Env(U)$, la **envoltura convexa** de U, es el mínimo convexo de E que contiene a U.

Proposición 66. *$Env(U) = \bigcap U'$ donde los U' son los convexos que contienen a U. También, $Env(U) = \{x \in E : x_1, \dots x_n \in U$ y $t_1, \dots, t_n \geq 0$ con $\sum_i t_i = 1$ tales que $x = \sum_i t_i x_i\}$.*

Proposición 67. *Sea $f : \Omega \subset \mathbb{R}^n \mapsto \mathbb{R}$ con Ω abierto convexo y $f \in C^2(\Omega)$. Entonces f es convexa en $\Omega \iff Hf(x) \geq 0\ \forall x \in \Omega$.*

Demostración. $\Longrightarrow$. Sea $x \in \Omega$ y $v \in \mathbb{R}^n$. Como en demostraciones anteriores, la aplicación $\varphi(t) = f(x + tv)$ está bien definida en cierto $(-\delta, \delta)$ y es de clase C^2. Además, es convexa: $\varphi((1-u)p + uq) = f((1-u)(x+pv) + u(x+qv)) \leq (1-u)f(x+pv) + uf(x+qv) = (1-u)\varphi(p) + u\varphi(q)$. Por tanto, tenemos $\varphi''(t) \geq 0$, es decir $v^t Hf(x+tv)v \geq 0$ y si $t = 0$, tenemos lo que queríamos dado que v y x eran arbitrarios.

Para $\Longleftarrow$, sean $x, y \in \Omega$ y $\varphi : [0,1] \mapsto \mathbb{R}$ dada por $\varphi(t) = f((1-t)x + ty)$. Como $\varphi''(t) = (x-y)^t Hf((1-t)x + ty)(x-y) \geq 0$ por hipótesis, entonces φ es convexa, y en particular $\varphi(t) \leq (1-t)\varphi(0) + t\varphi(1) = (1-t)f(x) + tf(y)$, que es lo que se buscaba. $\square$

De esto se sigue el siguiente resultado, que nos garantiza extremos incluso en puntos degenerados:

Proposición 68. *Si $f : U \subset \mathbb{R}^n \mapsto \mathbb{R}$, con U abierto, $f \in C^2(U)$ y $x_0 \in U$ es un punto crítico tal que $Hf(x) \geq 0$ en un W entorno de x_0 en U, entonces x_0 es mínimo local. (Análogo para el máximo).*

Demostración. Tomamos $W' = B_r(x_0) \subset W$. Este conjunto es abierto convexo y tal que $Hf(x) \geq 0$ en W' por hipótesis. Por tanto f es convexa en W'. Supongamos que en W' hubiera un x_1 con $f(x_1) < f(x_0)$. Por lo pronto sabemos que $f((1-t)x_0 + tx_1) \leq (1-t)f(x_0) + tf(x_1) = f(x_0) + t(f(x_1) - f(x_0))$. Como las funciones de ambos lados coinciden en 0, su derivada en 0 preserva la monotonía: $Df(x_0) \cdot (x_1 - x_0) \leq f(x_1) - f(x_0) < 0$, contradiciendo que $Df(x_0) = 0$. $\square$

Proposición 69. *Si $f : \Omega \subset \mathbb{R}^n \mapsto \mathbb{R}$ con Ω abierto convexo y f convexa, entonces f es continua.*

3. Teoremas de aplicación inversa e implícita

El objetivo de la sección es establecer el teorema de la función inversa, que permite invertir funciones localmente, y el de la función implícita, que permite parametrizar superficies de nivel de funciones o, lo que es lo mismo, *resolver* ecuaciones implícitas. En primer lugar definiremos un tipo de aplicaciones que acortan la distancia entre puntos:

Definición 53. Si (X, d) es un espacio métrico, la aplicación $f : X \mapsto X$ es **contractiva** si $\forall x, y \in X$, con $x \neq y$, se tiene $d(f(x), f(y)) < d(x, y)$. Se dice que es **estrictamente contractiva** si es de Lipschitz con constante $\Lambda < 1$.

Observación 4. Las funciones contractivas son de Lipschitz luego son continuas. Asimismo, las funciones estrictamente contractivas son en particular contractivas.

Teorema 3 (Del punto fijo de Banach)**.** *Si (X, d) es métrico y completo, y $f : X \mapsto X$ es estrictamente contractiva, entonces $\exists! x_0 \in X$ tal que $f(x_0) = x_0$.*

Demostración. La unicidad viene de ser contractiva: si $f(x) = x$ y $f(y) = y$ tenemos que $d(f(x), f(y)) = d(x, y)$ lo que solo puede darse si $x = y$. Vamos a presentar una construcción del punto fijo para ver su existencia. Sea $x_0 \in X$ arbitrario y $x_n = f^n(x_0)$, donde f^n es la composición n veces. Una primera observación es que $d(x_n, x_{n-1}) = d(f(x_{n-1}), f(x_{n-2})) \leq \Lambda d(x_{n-1}, x_{n-2}) \leq \cdots \leq \Lambda^{n-1} d(x_0, x_1)$.

Ahora, dados $n > m \geq 1$, tenemos que $d(x_n, x_m) \leq d(x_m, x_{m+1}) + d(x_{m+1}, x_{m+2}) + \cdots + d(x_{n-1}, x_n) \leq d(x_0, x_1) \sum_{j=m}^{n-1} \Lambda^j \leq d(x_0, x_1) \sum_{j=m}^{\infty} \Lambda^j = d(x_0, x_1) \frac{\Lambda^m}{1-\Lambda} \to 0$ cuando $m \to \infty$, luego esta sucesión es de Cauchy.

Como X es completo, sea x' aquel con $x_n \to x'$, como f es continua por ser Lipschitz, finalizamos con que $x_{n+1} = f(x_n) \to f(x')$, luego $x' = f(x')$. □

Definición 54. Se dice que el espacio normado E es **de Banach** si es completo con la distancia de la norma.

Proposición 70. *Sea $\gamma : [0, 1] \mapsto E$ con E de Banach, γ continua. Entonces, si $\{p_n\}_n$ es una sucesión de particiones $p_n = \{0, x_1, \ldots, x_{n-1}, 1\}$, tales que $diam(p_n) \to 0$, se tiene que la suma de Riemann inferior, $I(\gamma, p_n) = \sum_{j=0}^{n-1} \gamma(x_j)|x_{j+1} - x_j|$, verifica que $\exists \lim\limits_{n\to\infty} I(\gamma, p_n)$, no depende de la sucesión de particiones, y se denota $\int_0^1 \gamma(t)dt$.*

Asimismo, se tiene $\left\| \int_0^1 \gamma(t)dt \right\| \leq \int_0^1 \|\gamma(t)\| \, dt$, y, en caso de estar en $E = \mathbb{R}^n$, y de que $\gamma \in C^1([0, 1])$, se tiene que $\int_0^1 \gamma'(t)dt = \gamma(1) - \gamma(0)$.

Demostración. Como γ es uniformemente continua al estar en un compacto, si $\epsilon > 0$, hay un $\delta > 0$ tal que si $y > x$ e $y - x < \delta$, entonces $\|\gamma(y) - \gamma(x)\| \leq \epsilon$. Supongamos que P y Q fuesen 2 particiones de $[0, 1]$, de tamaños n y m, con $diam(P), diam(Q) < \delta$. Primero vamos a ver qué ocurre si a $P = (x_i)_0^n$ le agregamos tantos puntos como se desee, comprendidos entre x_j y x_{j+1}. Sea $R = (x'_i)_1^m$ esta nueva partición y supongamos que hemos agregado k puntos. Entonces, la mayoría de términos son comunes y se cancelan en: $I(\gamma, R) - I(\gamma, P) = \sum_{i=j}^{j+k} \gamma(x'_i)(x'_{i+1} - x'_i) - \gamma(x_j)(x_{j+1} - x_j) = \sum_{i=j}^{j+k} (\gamma(x'_i) - \gamma(x_j))(x'_{i+1} - x'_i)$. Esto nos da la estimación $\|I(\gamma, R) - I(\gamma, P)\| \leq \sum_{i=j}^{j+k} \|(\gamma(x'_i) - \gamma(x_j))\| \, (x'_{i+1} - x'_i) \leq \epsilon(x_{j+1} - x_j)$.

Supongamos ahora que $T = P \cup Q$, como lo podemos considerar como intercalar puntos en todos los tramos de P, sigue que: $\|I(\gamma, P) - I(\gamma, T)\| \leq \sum_{j=0}^{n-1} \epsilon(x_{j+1} - x_j) = \epsilon$.

Ahora si tomamos $P = p_n$ y $Q = p_m$ con n, m suficientemente grandes para que el diámetro sea pequeño, entonces la estimación anterior nos indica que $\{I(\gamma, p_n)\}_n$ es Cauchy, y por estar en un espacio de Banach, converge. Si tomamos $P = p_n$ y $Q = q_n$, donde $\{q_n\}$ es otra sucesión de particiones de diámetro convergente a 0, la desigualdad anterior nos indica que $\lim\limits_{n\to\infty} \|I(\gamma, p_n) - I(\gamma, q_n)\| = 0$ y por tanto dan lugar al mismo límite.

Ahora, tenemos que $\left\|\int_0^1 \gamma(t)dt\right\| = \left\|\lim_{n\to\infty} I(\gamma, p_n)\right\| = \lim_{n\to\infty} \|I(\gamma,p_n)\| \leq \lim_{n\to\infty} I(\|\gamma\|, p_n) = \int_0^1 \|\gamma(t)\|\, dt$.

Para la afirmación final, sea l_j la aplicación lineal y continua de proyección sobre la j-ésima coordenada: $l_j(x_1,\ldots,x_n) = x_j$. Por linealidad tenemos que $(l_j\circ\gamma)'(t) = l_j\circ\gamma'(t)$, y es una función de $[0,1]$ en $\mathbb{R}$. Por un lado, $\int_0^1 (l_j\circ\gamma)'(t)dt = \int_0^1 \gamma_j'(t)dt = \gamma_j(1)-\gamma_j(0)$. Pero por otro, como l_j es lineal y continua, se tiene $\int_0^1 l_j(\gamma'(t))dt = l_j(\int_0^1 \gamma'(t)dt) = [\int_0^1 \gamma'(t)dt]_j$, y se tiene lo que se quería. □

Con anterioridad hemos probado que las funciones $f: E \mapsto \mathbb{R}^n$ diferenciables son Lipschitz. Vamos a dar otro resultado similar pero siendo el espacio de llegada más general:

Proposición 71. *Sea $f: U\subset E\mapsto E$, con U abierto conexo y E de Banach. Si f es diferenciable en U, y $\Lambda = \sup\{\|(df)_x\| : x\in U\}$, entonces f es Lipschitz de constante Λ.*

Demostración. Sean $p,q\in U$. Sea $g:[0,1]\mapsto E$, dada por $g(t) = f(tp+(1-t)q)$. Vamos a abreviar $tp+(1-t)q \equiv r_t$. Entonces, $f(p)-f(q) = g(1)-g(0) = \int_0^1 g'(t)dt = \int_0^1 (df)_{r_t}(p-q)dt$. Por tanto, $\|f(p)-f(q)\| \leq \int_0^1 \|(df)_{r_t}(p-q)\|\, dt \leq \int_0^1 \Lambda \|p-q\|\, dt = \Lambda\|p-q\|$. □

Definición 55. Sean (X,d_x) y (Y,d_y) espacios métricos y $f: X\mapsto Y$ tal que $\exists\lambda>0$ con $\forall x,x'\in X$, se tiene $d_y(f(x),f(x'))\geq \lambda d_x(x,x')$. Se dice que f es **coercitiva** de constante λ.

Observación 5. Si f es coercitiva, entonces es inyectiva.

Proposición 72. *$f: X\mapsto Y$ es coercitiva de constante $\lambda \iff f$ es inyectiva y $f^{-1}: f(X)\mapsto X$ Lipschitz $\Lambda = \frac{1}{\lambda}$.*

Demostración $\implies$. Dados $y,y'\in f(X)$, si escribimos $f^{-1}(y) = x$ y $f^{-1}(y') = x'$ entonces por ser coercitiva, tenemos que $d(f(x),f(x'))\geq \lambda d(x,x')$, luego $d(y,y')\geq \lambda d(f^{-1}(y), f^{-1}(y'))$, luego es Lipschitz con parámetro λ^{-1}. $\impliedby$ es análogo: dados $x,x'\in X$, por ser Lipschitz, escribimos $d(f^{-1}(f(x)), f^{-1}(f(x)))\leq \frac{1}{\lambda}d(f(x),f(x'))$, de donde sigue lo que se quiere. □

Teorema 4 (De la Función Inversa). *Sea $f:\Omega\subset E\mapsto E$, con E de Banach y Ω abierto de E. Si $f\in C^1(\Omega)$ y $x_0\in\Omega$ es tal que $(df)_{x_0}$ es invertible, entonces $\exists\delta>0$ tal que:*

1. *$B_\delta(x_0)\subset\Omega$.*
2. *$f: B_\delta(x_0)\mapsto f(B_\delta(x_0))$ es invertible y abierta (es decir, si $A\subset B_\delta(x_0)$ es abierto de E, entonces $f(A)$ también).*
3. *Se tiene $f^{-1}: f(B_\delta(x_0))\mapsto B_\delta(x_0)$ verificando $f^{-1}\in C^1(f(B_\delta(x_0)))$ y además $(df^{-1})_{f(x_0)} = (df)_{x_0}^{-1}$.*

Demostración. En caso de que obtengamos f^{-1} como se describe en el teorema, la parte final es inmediata, dado que $f\circ f^{-1} = f^{-1}\circ f = Id_E$, de tal modo que $Id = d(f\circ f^{-1})_{f(x_0)} = (df)_{x_0}\circ (df^{-1})_{f(x_0)}$, así como $Id = d(f^{-1}\circ f)_{x_0} = (df^{-1})_{f(x_0)}\circ(df)_{x_0}$. Por tanto solo hay que probar la existencia y diferenciablidad de f^{-1}.

Vamos a comenzar con un caso *normalizado* más sencillo. Supongamos que $x_0 = f(x_0) = 0$ y $(df)_{x_0} = Id$. La definición de diferencial nos indica que para $x\in\Omega$, se tiene $f(x) = x+g(x)$ donde además $g(x) = o(\|x\|)$. Claramente $g(x) = f(x)-x$ y por tanto también $g\in C^1(\Omega)$. Es decir, $x\mapsto (dg)_x$ es continua y además $(dg)_0 = (df)_0 - Id = Id - Id = 0$, de tal manera que si $\epsilon>0$, tenemos $\delta>0$ con $x\in \overline{B}_\delta(0)\subset U \implies \|(dg)_x\|\leq\epsilon$.

Por tanto, como esa bola es conexa, la proposición 71 garantiza que $\forall x,y\in\overline{B}_\delta(0)$, tenemos $\|g(x)-g(y)\|\leq \epsilon\|x-y\|$. Si además suponemos que $\epsilon<1$ arbitrario, entonces $\|f(x)-f(y)\| = \|x+g(x)-y-g(y)\|\geq \|x-y\| - \|g(x)-g(y)\| \geq (1-\epsilon)\|x-y\|$. Por lo tanto f es coercitiva y sigue que f es inyectiva en $B_\delta(0)$.

Ahora vamos a ver que $B_{(1-\epsilon)\delta}(0) \subset f(B_\delta(0))$, es decir, que hay una bola que se alcanza en su totalidad partiendo desde $B_\delta(0)$. Esto nos permitirá fijarnos exclusivamente en los puntos que van a parar a esa bola, teniendo así sobreyectividad (además de inyectividad, ya garantizada por estar en $B_\delta(0)$). En primer lugar, sea $y \in B_{(1-\epsilon)\delta}(0)$. Veamos que $y = f(x)$ con $x \in B_\delta(0) \iff x+g(x) = y \iff x = y - g(x) \iff x = G_y(x)$, con $G_y(x) = y - g(x)$. Por tanto, solo hay que probar que $G_y(x)$ tiene un punto fijo en $B_\delta(0)$. Que tiene un punto fijo sigue del teorema de punto fijo, dado que $\|G_y(x) - G_y(x')\| = \|g(x) - g(x')\| < \epsilon \|x - x'\|$, como ya vimos, con $\epsilon < 1$, luego es estrictamente contractiva, y aplicando el teorema en el completo $\overline{B}_\delta(0)$ se tiene el punto fijo. Para ver que además está en $B_\delta(0)$, observemos que $\|G_y(x)\| = \|y - g(x)\| = \|y - g(0) + g(0) - g(x)\| \leq \|y\| + \|g(x) - g(0)\| < (1-\epsilon)\delta + \epsilon \|x\| \leq \delta$, es decir, que $G_y(\overline{B}_\delta(0)) \subset B_\delta(0)$, luego hemos acabado esta parte.

Así, por lo anterior, podemos tomar $V = B_{(1-\epsilon)\delta}(0)$ y $U = f^{-1}(V) \cap B_\delta(0)$, los dos abiertos, y por los pasos anteriores, $f : U \mapsto V$ es biyectiva (invertible), dado que como $U \subset B_\delta(0)$ es inyectiva, y además sabemos que $f(U) = f(f^{-1}(V)) \cap f(B_\delta(0)) = V \cap f(B_\delta(0)) = V$ por lo anterior, luego es sobreyectiva.

Como f es coercitiva en ese espacio, f^{-1} es Lipschitz y por tanto continua. Es decir, f entre esos espacios es un homeomorfismo (Aplicación invertible continua con inversa continua). Ahora veremos que además es diferenciable en 0. Sabemos que, de serlo, su diferencial debe ser la identidad, luego pongamos $y \in V \backslash \{0\}$ y notemos que $\frac{\|f^{-1}(y)-y\|}{\|y\|} = \frac{\|x-f(x)\|}{\|y\|} = \frac{\|x-f(x)\|}{\|x\|}\frac{\|x\|}{\|y\|}$. Pero por ser Lipschitz, $\|x\| = \|f^{-1}(y)\| \leq \frac{1}{1-\epsilon}\|y\|$, de tal manera que:

$$\frac{\|f^{-1}(y) - y\|}{\|y\|} \leq \frac{\|x - f(x)\|}{\|x\|}\frac{1}{1-\epsilon} \to 0 \qquad \text{si } x \to 0$$

El límite de la derecha se tiene dado que $(df)_0 = Id$. Pero si $y \to 0$, entonces $x \to 0$ al ser $x = f^{-1}(y)$, f^{-1} continua y $f(0) = 0$.

Ahora procedamos con el caso general. Vamos a normalizar f al caso previo, a través de funciones invertibles, de tal modo que una vez la normalización sea invertible local por lo visto anteriormente, lo será f deshaciendo las transformaciones. Sabemos que si $\|h\| \leq \delta$, entonces $x_0 + h \in \Omega$. Sea $G : B_\delta(0) \mapsto E$ con $G(h) = (df)_{x_0}^{-1}(f(x_0+h) - y_0)$. Como vemos, la aplicación G es composición de una afinidad (invertible) con f, con otra afinidad, con $(df)_{x_0}^{-1}$, todas invertibles. Por tanto, si G es invertible, f deberá serlo. Ahora bien, $G(0) = (df)_{x_0}^{-1}(f(x_0) - y_0) = 0$ y además $(dG)_0 = (df)_{x_0}^{-1} \circ (df)_{x_0} \circ Id = Id$, dado que la diferencial de la afinidad $h \to x_0 + h$ es la identidad. Por tanto, aplicamos el caso previo a G y tenemos lo que queríamos. Además, como G es diferenciable en 0, f^{-1} ha de serlo en y_0, dado que podemos escribir:

$$G(t - x_0) = (df)_{x_0}^{-1}(f(t) - y_0) \implies f(t) = df_{x_0}(G(t-x_0)) + y_0 \implies f^{-1}(t) = x_0 + G^{-1}(df_{x_0}^{-1}(t - y_0))$$

Y la regla de la cadena nos indica que f^{-1} es diferenciable en y_0.

Finalmente queremos poder afirmar que $f^{-1} \in C^1$ en un entorno de y_0, que esté en V. En concreto, a $V' = \{f(x) : x \in V, \exists (df)_x^{-1}\}$, que es donde hemos probado que es diferenciable. Así, restringiendo aún más f, tendremos todo lo que queremos. Para ello necesitamos el siguiente resultado:

Lema 4. *Si $A \in L(E)$ es invertible y $X \in L(E)$ con $\|X\| \leq \|A^{-1}\|^{-1}$, entonces $A + X$ también es invertible y $\|(A+X)^{-1} - A^{-1}\| \leq \frac{\|A^{-1}\|^2\|X\|}{1-\|A^{-1}\|\|X\|}$. En particular $X \mapsto X^{-1}$ es continua.*

Este lema surge de que si $\|A\| < 1$, entonces $(I + A)$ es invertible y su inversa es $\sum_{n=0}^{\infty}(-1)^n(A)^n$. Entonces, tomamos $A + X = A(I + A^{-1}X)$ y aplicamos el lema sobre el factor de la derecha.

Gracias a este resultado, tenemos que como f es C^1, entonces $x \to (df)_x$ es continua, luego $x \to (df)_x^{-1} = (df^{-1})_{f(x)}$ también lo es y tenemos lo que queríamos. □

A continuación vamos a ver que en $\mathbb{R}^n$ la regularidad, de hecho, se mantiene a niveles superiores, gracias al siguiente resultado:

Proposición 73. *Sea $j : GL(\mathbb{R}^n) \mapsto GL(\mathbb{R}^n)$ dada por $j(A) = A^{-1}$. Se tiene $j \in \mathbb{C}^\infty(GL(\mathbb{R}^n))$.*

Demostración. Es correcto hablar de diferenciales dado que $GL(\mathbb{R}^n) = det^{-1}(\mathbb{R}\backslash\{0\})$ y por tanto es un abierto. Como $A^{-1} = \frac{1}{det(A)} adj(A^T)$, entonces, cada componente es $(A^{-1})_{i,j} = (-1)^{i+j}\frac{\det A_{ji}}{\det A}$ que se trata de un polinomio en las componentes de A y por tanto es C^∞, siempre que $\det A \neq 0$ lo cual ocurre siempre al ser $A \in GL(\mathbb{R}^n)$. □

Proposición 74. *Si $f : U \subset \mathbb{R}^n \mapsto \mathbb{R}^n$ con U abierto, $f \in C^k(U)$ y $x_0 \in U$ con df_{x_0} invertible, entonces la inversa local f^{-1} del teorema de función inversa verifica que $f^{-1} \in C^k(f(B_\delta(x_0)))$.*

Demostración. Tenemos que $(df^{-1})_y = (df)^{-1}_{f^{-1}(y)} = (j \circ df \circ f^{-1})(y)$ si entendemos por j la inversión y por df la función $x \mapsto df_x$. Si $k = 1$ por tanto el resultado es inmediato del teorema de función inversa y válido en cualquier espacio de banach. Ahora procedemos por inducción. Supongamos que se da el resultado hasta k, y sea $f \in C^{k+1}(U)$. Hemos visto que df^{-1} es composición de j, que es C^∞, df, que es C^k por hipótesis, y f^{-1}, que por hipótesis de inducción es C^k dado que f es también C^k. Por tanto df^{-1} es C^k y esto indica que f^{-1} es C^{k+1}. □

El teorema de la función inversa nos da lugar al de la función implícita: si tenemos una función de $m+n$ variables y n componentes, aquellos puntos que la anulan (si queremos, soluciones de un sistema dado por las n componentes) verifican (bajo ciertas condiciones) que, en un entorno suyo que también anule la función, existe una función *implícita* que relaciona n de las variables con las otras m. Es decir, en un entorno de la solución dada, las demás soluciones se pueden saber conociendo únicamente m de las componentes: por ello, se dice que hay m *grados de libertad*: las otras n son función de ellas.

Teorema 5 (De la función implícita)**.** *Sea $f : U \subset \mathbb{R}^{m+n} \mapsto \mathbb{R}^n$ con U abierto y $f \in \mathcal{C}^k(U)$ con $k \geq 1$. Dado $x \in \mathbb{R}^{m+n}$ lo escribiremos como $x = (y, z)$ con $y \in \mathbb{R}^m$ y $z \in \mathbb{R}^n$. Sea $a = (y_0, z_0) \in U$ tal que:*

1. *$f(a) = 0$*
2. *Si escribimos $Df(a) = Df(y_0, z_0) = \left(\left[\frac{\partial f(a)}{\partial y}\right] \;\middle|\; \left[\frac{\partial f(a)}{\partial z}\right]\right)$, con $\left[\frac{\partial f(a)}{\partial y}\right] \in \mathcal{M}_{n\times m}$ la matriz de parciales respecto a las m primeras componentes, y $\left[\frac{\partial f(a)}{\partial z}\right] \in \mathcal{M}_{n\times n}$ aquella con respecto a las n últimas, se tiene que $\left[\frac{\partial f(a)}{\partial z}\right]$ es invertible.*

Entonces, $\exists V \subset \mathbb{R}^m$ y $W \subset \mathbb{R}^n$ abiertos, con $y_0 \in V$, $z_0 \in W$ tales que:

a. *$V \times W \subset U$*

b. *$\exists! \varphi : V \mapsto \mathbb{R}^n$, con $\varphi \in \mathcal{C}^k(V)$, $\varphi(V) \subset W$ y $(V \times W) \cap f^{-1}(0) = \{(y, \varphi(y)) : y \in V\}$. Esta es la que se conoce como implícita y determina el resto de soluciones a $f(x) = 0$ en ese entorno de a.*

Demostración. Vamos a denotar $D_1 f(x) = \left[\frac{\partial f(a)}{\partial y}\right]$ y $D_2 f(x) = \left[\frac{\partial f(a)}{\partial z}\right]$ para abreviar. Definimos $h : U \subset \mathbb{R}^{m+n} \mapsto \mathbb{R}^{m+n}$ dada por $h(y, z) = (y, f(y, z))$. $h \in \mathcal{C}^k(U)$ al serlo f, y es fácil calcular:

$$Dh(y, z) = \begin{bmatrix} I_m & 0 \\ D_1 f(y, z) & D_2 f(y, z) \end{bmatrix}$$

Que es invertible en $a = (y_0, z_0)$ por ser triangular inferior por bloques, y ser $D_2 f(y_0, z_0)$ invertible. Por el teorema de función inversa tenemos $U_1 \subset \mathbb{R}^{m+n}$ y $\Omega \subset \mathbb{R}^{m+n}$ con $a \in U_1$, $(y_0, 0) \in \Omega$, abiertos y tales que $h : U_1 \mapsto \Omega$ es un C^k-difeomorfismo. Sea $\pi_1 : \mathbb{R}^{m+n} \mapsto \mathbb{R}^m$ dada por $\pi_1(y, z) = y$ y $\pi_2 : \mathbb{R}^{m+n} \mapsto \mathbb{R}^n$ dada por $\pi_2(y, z) = z$. Son claramente lineales, C^∞ y además veremos que son abiertas. Basta con que la imagen de una bola sea abierta, dado que los abiertos son uniones de bolas. $(y', z') \in B_r^\infty(y, z) \iff$

$\|(y'-y,z'-z)\|_\infty < r \iff \text{máx}\{\|y-y'\|_\infty, \|z-z'\|_\infty\} < r \iff (y',z') \in B_r^\infty(y) \times B_r^\infty(z)$, luego se tiene $\pi_1(B_r^\infty((y,z))) = B_r^\infty(y)$ y $\pi_2(B_r^\infty((y,z))) = B_r^\infty(z)$.

Ahora, si $g_i = \pi_i \circ h^{-1}$, es decir, $h^{-1}(x) = (g_1(x), g_2(x))$, tenemos lo siguiente si $x = (u,v) \in \Omega$:

$$(u,v) = h(h^{-1}(u,v)) = h(g_1(u,v), g_2(u,v)) = (g_1(u,v), f(g_1(u,v), g_2(u,v)))$$

Luego $u = g_1(u,v)$ y $v = f(g_1(u,v), g_2(u,v))$, o lo que es lo mismo, $v = f(u, g_2(u,v))$ y $h^{-1}(u,v) = (u, g_2(u,v))$. Esto nos dice que $U_1 \cap f^{-1}(0) = \{(u, g_2(u,0)) : (u,0) \in \Omega\}$. La inclusión a derecha es porque $f(u, g_2(u,0)) = 0$ según la ecuación anterior, y porque $h^{-1}(u,0) = (u, g_2(u,0))$ luego está en U_1. La inclusión a izquierda es porque dado $(u,v) \in U_1$, $f(u,v) = 0$, por un lado tenemos $h^{-1}(u,v) \in \Omega$ y por otro $h^{-1}(u,v) = (u, g_2(u,v))$, pero $0 = f(u,v) = f(u, g_2(u,v)) = v$ luego de hecho $v = 0$.

Por lo tanto, podemos reescribir aquello como $U_1 \cap f^{-1}(0) = \{(u, \varphi(u)) : u \in V\}$, donde $V = \pi_1(U_1)$ abierto según discutimos antes, y $\varphi(u) = g_2(u,0)$, que es C^k por serlo g_2. Ahora, si $x \in U_1$ con $x = (y,z)$, $y \in V$, $z \in \pi_2(U_1) = W$, tenemos que $V \times W$ es entorno de a, por serlo U_1, además V es entorno de y_0 y W de z_0, y $(V \times W) \cap f^{-1}(\{0\}) = U_1 \cap f^{-1}(\{0\}) = \{(u, \varphi(u)) : u \in V\}$ como se quería. □

Proposición 75 (Derivación implícita)**.** *En las hipótesis del teorema anterior y con la misma notación, tenemos que $D\varphi(y) = -D_2 f(y, \varphi(y))^{-1} D_1 f(y, \varphi(y))$, siempre que $y \in V$.*

Demostración. Si llamamos $q(y) = (y, \varphi(y))$ con $q : V \mapsto \mathbb{R}^{m+n}$. La regla de la cadena nos indica que como $0 = f(y, \varphi(y))$, entonces $0 = Df(y, \varphi(y)) Dq(y)$. Es decir:

$$0 = \begin{bmatrix} D_1 f(y, \varphi(y)) & D_2 f(y, \varphi(y)) \end{bmatrix} \begin{bmatrix} I_m \\ D\varphi(y) \end{bmatrix}$$

Multiplicando y despejando se tiene lo deseado. □

4. Subvariedades de $\mathbb{R}^n$

Definición 56. Sean $U_1 \subset \mathbb{R}^n$ y $U_2 \subset \mathbb{R}^m$ abiertos no vacíos. Se dice que $\sigma : U_1 \mapsto U_2$ es un **C^s-difeomorfismo**, con $s \geq 1$, si $\sigma \in C^s(U_1)$, es biyectiva, y $\sigma^{-1} \in C^s(U_2)$.

Observación 6. Para que sea un difeomorfismo, es obligatorio que $m = n$, es decir, que las dimensiones de los espacios de partida y llegada coincidan. Esto es así porque como $\sigma\sigma^{-1} = Id_{U_2}$ y $\sigma^{-1}\sigma = Id_{U_1}$, la regla de la cadena nos dice que $D\sigma(x)D\sigma^{-1}(\sigma(x)) = Id = D\sigma^{-1}(\sigma(x))D\sigma(x)$, y para que esas matrices se puedan operar de ambos lados, debe darse que $m = n$.

Definición 57. Sean los enteros $n \geq 1$ y $0 \leq k \leq n$. Se dice que $X \subset \mathbb{R}^n$ es una **C^s-subvariedad de dimensión** k si, $\forall x \in X$, $\exists U, U' \subset \mathbb{R}^n$ abiertos con $x \in U$ y $\exists \sigma : U \mapsto U'$ un C^s difeomorfismo (denominado *aplanador*) tal que $\sigma(X \cap U) = U' \cap (\mathbb{R}^k \times \{0_{n-k}\})$.

Es decir, localmente, la variedad X es difeomorfa a un entorno en $\mathbb{R}^k$.

Observación 7. Si $k = 1$ se habla de *curvas*, y si $k = 2$ se habla de *superficies*. Asimismo, si $k = 0$ se trata de *conjuntos discretos*, y si $k = n - 1$ se trata de *hipersuperficies*.

Por ejemplo, en $\mathbb{R}^n$ si $U \subset \mathbb{R}^n$ es abierto, basta con tomar para todo $x \in U$, la aplicación $Id : U \mapsto U$ y vemos que se trata de una n-variedad.

Otro ejemplo: Si $A = x_0 + V$ un espacio afín de $\mathbb{R}^n$, con V siendo k-dimensional, y $U \subset A$ es abierto relativo, es decir, $U = A \cap \Omega$ con Ω abierto de $\mathbb{R}^n$, tenemos que U es una k-variedad. Tomamos $T : \mathbb{R}^n \mapsto \mathbb{R}^n$ ortogonal, lineal, que mueva V así: $T(V) = V' = \mathbb{R}^k \times \{0_{n-k}\}$. Definimos $\varphi(x) := T(x - x_0)$. Entonces como tanto T como $x - x_0$ son C^∞, e inyectivas, se trata de un difeomorfismo entre Ω y $\varphi(\Omega)$. Asimismo, $\varphi(\Omega \cap U) = \varphi(\Omega \cap A) = \varphi(\Omega) \cap \varphi(A) = \varphi(\Omega) \cap T(V) = \varphi(\Omega) \cap (\mathbb{R}^k \times \{0_{n-k}\})$, satisfaciendo la definición de k-variedad.

Proposición 76. *Sea $X \subset \mathbb{R}^n$ un conjunto de puntos aislados. Es decir, $\forall p \in X$, $\exists U \subset \mathbb{R}^n$ abierto con $p \in U$ y $U \cap X = \{p\}$. Se tiene que X es una 0-subvariedad C^∞.*

Demostración. Dado $p \in X$, tomamos el U de la definición de aislado y exponemos $\sigma : U \mapsto \mathbb{R}^n$ dada por $\sigma(x) = x - p$. Está claro que es un C^∞-difeomorfismo. Además, $\sigma(X \cap U) = \sigma(\{p\}) = \{0\}$. □

Proposición 77. *Si $\Phi : \Omega \subset \mathbb{R}^k \mapsto \mathbb{R}^{n-k}$ es $C^s(\Omega)$ con Ω abierto, entonces $X = Gr(\Phi) = \{(x, \Phi(x)), x \in \Omega\}$ es una k-subvariedad C^s en $\mathbb{R}^n$.*

Demostración. Sea $W = \Omega \times \mathbb{R}^{n-k}$, que es abierto por serlo Ω en $\mathbb{R}^k$. Entonces damos $\sigma : W \mapsto W$ dada por $\sigma(x, x') = (x, x' - \Phi(x))$ donde $x = (x_1, \ldots, x_k) \in \Omega$ y $x' = (x_{k+1}, \ldots, x_n)$. Es C^s por serlo Φ y además tiene inversa: $\sigma^{-1}(y, y') = (y, y' + \Phi(y))$ con la misma notación. Efectivamente: $\sigma^{-1}\sigma(x, x') = \sigma^{-1}(x, x' - \Phi(x)) = (x, x' - \Phi(x) + \Phi(x)) = (x, x')$ y análogo al revés. Esta inversa es C^s al serlo Φ luego σ es difeomorfismo. Finalmente $(y, y') \in \sigma(W \cap X) \iff (y, y') = \sigma((x, \Phi(x)))$ con $(x, \Phi(x)) \in W \iff (y, y') = (x, \Phi(x) - \Phi(x)) = (x, 0)$ con $(x, 0) \in W \iff (y, y') \in \Omega \times \{0\}$. □

Observación 8. Podemos aplicar Φ en cualesquiera k variables de $\mathbb{R}^n$ y entonces $Gr(\Phi)$ es subvariedad. Arriba se ha probado cuando Φ se aplica en las k primeras pero basta con reordenarlas si es preciso.

Ahora que sabemos que los grafos son variedades, veamos otra manera de obtener variedades en relación a funciones:

Proposición 78. *Sean m, n enteros con $1 \leq m \leq n$. Sea $\Omega \subset \mathbb{R}^n$ abierto y $f : \Omega \mapsto \mathbb{R}^m$ con $f \in C^s(\Omega)$, $s \geq 1$. Dado $b \in \mathbb{R}^m$ tal que $rg(Df(a)) = m$ (el máximo posible) en todo $a \in f^{-1}(b)$, entonces, o bien $f^{-1}(b) = \emptyset$, o se trata de una C^s-subvariedad de dimensión $n - m$.*

Demostración. Suponemos que $f^{-1}(b) \neq \emptyset$. Si $n = m$, dado $a \in f^{-1}(b)$, como $Df(a)$ es invertible (tiene rango m y es cuadrada), aplicamos el teorema de la función inversa y obtenemos $U_a \subset U$ abierto con f inyectiva en U_a, y $a \in U_a$. En ese caso $U_a \cap f^{-1}(b) = \{a\}$ a causa de la inyectividad. Pero entonces $f^{-1}(b)$ es un conjunto de puntos aislados, de tal manera que es una 0-variedad. Por otro lado, si $m < n$, dado $a \in f^{-1}(b)$ el teorema de la función implícita afirma que hay un abierto de $\mathbb{R}^n$, digamos $W \subset \mathbb{R}^n$, con $a \in W$ y $W \cap f^{-1}(b)$ es gráfica de cierta función $\Phi : U \mapsto \mathbb{R}^m$ con $U \subset \mathbb{R}^{n-m}$, con $\Phi \in C^s(U)$, en cierto abierto $U \in \mathbb{R}^{n-m}$, y por lo probado anteriormente es una $(n-m)$-subvariedad. □

Por ejemplo, la esfera $(n-1)$-dimensional, $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : \|x\|_2 = 1\} = g^{-1}(1)$, donde $g : \mathbb{R}^n \mapsto \mathbb{R}$ dada por $g(x) = \|x\|_2^2$, que es $C^\infty(\mathbb{R}^n)$, verifica por el resultado anterior que es una subvariedad $n-1$ dimensional de $\mathbb{R}^n$, dado que $Dg(a) = 2a$ y si $g(a) = 1$, se tiene que $2a$ es de rango 1 o superior.

Definición 58. Sea $U \subset \mathbb{R}^k$ abierto. Se dice que $f : U \mapsto \mathbb{R}^n$ es una **inmersión** si $f \in \mathcal{C}^s(U)$ y $\forall x \in U$ la aplicación $Df(x) : \mathbb{R}^k \mapsto \mathbb{R}^n$ es inyectiva.

Observemos que debe ser que $k \leq n$ porque si no es imposible que sea inyectiva. Por tanto, la condición de inyectividad equivale a decir que $rg(Df(x)) = k$ para cualquier $x \in U$.

Proposición 79. *Si $k < n$ y $f : U \subset \mathbb{R}^k \mapsto \mathbb{R}^n$ es una inmersión de clase C^s, y $U \neq \emptyset$, entonces $f(U)$ se puede expresar una unión de subvariedades C^s de dimensión k de $\mathbb{R}^n$.*

Demostración. Sea $x_0 \in U$. Si $f = (f_1, \dots, f_n)$, reordenamos si es preciso para que $Df(x_0) = \begin{pmatrix} \nabla f_1(x_0) \\ \vdots \\ \nabla f_n(x_0) \end{pmatrix}$ tenga las primeras k filas linealmente independientes. Entonces definimos $\sigma(x) = (f_1(x), \dots, f_k(x))$, como vemos $\sigma : U \mapsto \mathbb{R}^k$ y es C^s. Por construcción, $D\sigma(x_0)$ es invertible luego $\exists U_{x_0} \subset U$ con $\sigma : U_{x_0} \mapsto W_{x_0}$ difeomorfismo de clase C^s. Si $\pi : \mathbb{R}^n \mapsto \mathbb{R}^k$ y $\tilde{\pi} : \mathbb{R}^n \mapsto \mathbb{R}^{n-k}$ son las proyecciones dadas por $\pi(x_1, \dots, x_n) = (x_1, \dots, x_k)$, $\tilde{\pi}(x_1, \dots, x_n) = (x_{k+1}, \dots, x_n)$, tenemos que $f(U_{x_0}) = \{y \in \mathbb{R}^n : \pi(y) \in W_{x_0}, \tilde{\pi}(y) = \tilde{\pi}(f((\sigma^{-1}(\pi(y)))))\} = Gr(\tilde{\pi} \circ f \circ \sigma^{-1} \circ \pi, W_{x_0})$ y por ser gráfico de una función C^s en un abierto de $\mathbb{R}^k$ que va en $\mathbb{R}^{n-k}$, es una variedad k-dimensional. Aplicando esto a cada punto se tiene la unión deseada, dado que $U = \bigcup U_{x_0}$ luego $f(U) = \bigcup f(U_{x_0})$ que son variedades. □

Definición 59. Si X es una subvariedad C^s k-dimensional en $\mathbb{R}^n$, una **parametrización** de X es $\Phi : W \subset \mathbb{R}^k \mapsto \mathbb{R}^n$ con W abierto, $\Phi \in C^s(W)$ y $\Phi(W) \subset X$, además de que $\Phi(W)$ considerado en X debe tener interior no vacío (es decir, $\exists U \neq \emptyset$ abierto en X con $U \subset \Phi(W)$). Al conjunto $\Phi(W)$ se lo denomina **parte parametrizada**.

Si además es inyectiva e inmersión, se dice que es una **parametrización regular.**

Observación 9 (Parametrizaciones grafo). Un ejemplo de parametrización: fijado $U \subset \mathbb{R}^k$ abierto y $\varphi : U \mapsto \mathbb{R}^n$ una función C^s. tenemos la parametrización $\Phi : U \mapsto \mathbb{R}^{n+k}$ con $\Phi(u) = (u, \varphi(u))$ que es regular. Además es inyectiva y $D\Phi(u) = \begin{pmatrix} I_k \\ D\varphi(u) \end{pmatrix}$ que es de rango k. Se denomina **parametrización grafo.** y parametriza la variedad gráfica de φ en U. Consideraremos que si $P : \mathbb{R}^{n+k} \mapsto \mathbb{R}^{n+k}$ permuta las coordenadas $(P(x_1, \dots, x_{n+k}) = (x_{\sigma(1)}, \dots, x_{\sigma(n+k)})$ para $\sigma \in S_{n+k})$, la parametrización $P(\Phi(u))$ también es parametrización grafo.

Proposición 80. *Sea $U \subset \mathbb{R}^k$ un abierto y $\varphi : U \mapsto \mathbb{R}^n$ con $\varphi \in C^s(U)$. Sea $X = \{(x, \varphi(x)) : x \in U\}$ una variedad (su gráfica) k-dimensional de regularidad C^r en $\mathbb{R}^{n+k}$. Entonces se tiene $r = s$. (En el sentido de que X no supera en regularidad a φ, ya sabíamos que al menos era C^s).*

Definición 60. Sea X una subvariedad C^s k-dimensional de $\mathbb{R}^n$. Se define, dado $p \in X$, el **espacio tangente a X en p** como $T_pX = \{\gamma'(0) : \exists \delta > 0 \text{ y } \gamma : (-\delta, \delta) \mapsto X \text{ con } \gamma \in C^1((-\delta, \delta)), \gamma(0) = p\}$.

Es decir, son los posibles vectores velocidad de curvas que pasan por el punto a través de la variedad. Forman un espacio vectorial:

Proposición 81. *Si $X \subset \mathbb{R}^n$ es una subvariedad C^s k-dimensional, y $p \in X$, entonces T_pX es un subespacio vectorial de $\mathbb{R}^n$, de dimensión k.*

Demostración. Sea $p \in X$. Por ser variedad, es localmente gráfica de función (se verá posteriormente), es decir, $\exists U \subset \mathbb{R}^n$ abierto con $p \in X$ tal que $U \cap X = \{\Phi(u) : u \in V\}$ con $V \subset \mathbb{R}^k$ abierto y $\Phi(u) = P(u, \varphi(u))$ una parametrización grafo con $\varphi \in C^s$. Si definimos entonces $\pi : \mathbb{R}^n \mapsto \mathbb{R}^k$ por $\pi(x_1, \ldots, x_n) = (x_{\sigma(1)}, \ldots, x_{\sigma(k)})$ con σ la dada por P, estamos diciendo que $q = \Phi(\pi(q))$ siempre que $q \in U \cap X$.

Ahora, dada $\alpha : (-\delta, \delta) \mapsto X$ una función C^1 con $\alpha(0) = p$, como U es abierto, $\exists \delta_1 > 0$ tal que si $|t| < \delta_1$, $\alpha(t) \in U$. Pero $\alpha(t) = \Phi(\beta(t))$ si $\beta = \pi \circ \alpha$ luego $\beta \in C^1$. Por la regla de la cadena $\alpha'(t) = D\Phi_{\beta(t)}\beta'(t)$ luego $\beta'(0) = D\Phi_{\pi(p)}\beta'(0) \in Im(D\Phi_{\pi(p)})$. Así, $T_pX \subset Im(D\Phi_{\pi(p)})$.

Por otra parte, dado cualquier $v \in \mathbb{R}^k$, tenemos que si $\gamma(t) = \pi(p) + vt$, como $\pi(p) \in V$ abierto, tenemos que si t es pequeño, $\gamma(t) \in V$. Si $\alpha(t) = \Phi(\beta(t))$, entonces $\alpha'(0) = D\Phi_{\pi(p)}\beta'(0) = D\Phi_{\pi(p)}v$, luego como v era arbitrario, $Im(D\Phi_{\pi(p)}) \subset T_pX$, por lo que vale la igualdad y es un subespacio imagen. La dimensión es k porque como Φ es inmersión, $rg(D\Phi_{\pi(p)}) = k$. □

Proposición 82. *Sea $U \subset \mathbb{R}^k$ abierto y $f : U \mapsto X$ con X una subvariedad C^s k-dimensional de $\mathbb{R}^n$. Dado $a \in U$ se tiene que $Im(D_af) \subset T_{f(a)}X$, siempre que $f \in C^1(U)$.*

Demostración. Damos $\alpha : (-\delta, \delta) \mapsto U$ dada por $\alpha(t) = a + vt$ con $v \in \mathbb{R}^k$ arbitrario y δ para que la imagen esté en U. Sea $\beta = f \circ \alpha$. Entonces $\beta'(0) = Df_a\alpha'(0) = Df_av$, y $\beta'(0) \in T_{f(a)}X$ por definición, pero v era arbitrario luego $Im(D_af) \subset T_{f(a)}X$. □

Proposición 83. *Si $X \subset \mathbb{R}^n$ es C^s-subvariedad k-dimensional y $\Phi : U \subset \mathbb{R}^k \mapsto X$ con U abierto, $\Phi \in C^s(U)$ es una parametrización regular de X, entonces $\Phi(U)$ es abierto en X y $\Phi : U \mapsto \Phi(U)$ es un homeomorfismo (biyectiva e inversa continua).*

Demostración. Si $p \in \Phi(U)$, es decir, $p = \Phi(a)$, $a \in U$, entonces $\exists V \subset \mathbb{R}^n$ entorno de p con $\sigma : V \mapsto V'$ verifica $\sigma(V \cap X) = V' \cap (\mathbb{R}^k \times \{0\})$ y es C^s difeomorfismo. Sea $\pi : \mathbb{R}^n \mapsto \mathbb{R}^k$ la proyección sobre las primeras k coordenadas. Observemos que $\pi|_{\sigma(V \cap X)}$ es biyección al ser el valor de las últimas $n - k$ coordenadas fijo a 0. Como Φ es continua, $\exists U_p \subset U$ entorno de a con $\Phi(U_a) \subset V$. Definimos el morfismo $\Omega = \pi \circ \sigma \circ \Phi : U_a \mapsto \mathbb{R}^k$. Es inyectivo al serlo cada uno de los 3 de los que está compuesto (Φ en U_0 por ser parametrización regular, σ en $\Phi(U_0)$ por estar en V y en X, y π en $\sigma(V \cap X)$ como hemos discutido), y es C^s por serlo cada uno. Además, si $b \in U_a$, tenemos $D\Omega_b = \pi \circ D\sigma_{\Phi(b)} \circ D\Phi_b$, que tiene rango k (máximo) por tener rango máximo cada una de las 3 de las que está compuesto.

Por tanto, el teorema de la función inversa dice que Ω es localmente invertible en todo punto de U_a, y como Ω ya es invertible sobre $\Omega(U_a)$ por ser inyectiva, nos dice que $\Omega : U_a \mapsto \Omega(U_a)$ es un C^s-difeomorfismo, pero $\Phi^{-1} = \Omega^{-1} \circ \pi \circ \sigma$, luego es continua. Pero podemos expresar $U = \bigcup_{a \in U} U_a$, luego como Φ^{-1} es homeomorfismo en cada U_a, tenemos que $\Phi(U) = \bigcup_{a \in U} \Phi(U_a)$ es unión de abiertos y por tanto es abierta. El mismo razonamiento vale para cualquier subconjunto abierto $W \subset U$, luego Φ es homeomorfismo en U. □

Proposición 84. *Sea X una subvariedad C^s k-dimensional de $\mathbb{R}^n$ y $\Phi_i : U_i \subset \mathbb{R}^k \mapsto X$ unas parametrizaciones regulares C^s, con U_i abiertos, $i \in \{1, 2\}$ y $W = \Phi(U_1) \cap \Phi(U_2) \neq \emptyset$. Entonces $\psi : U_1' \mapsto U_2'$, dada por $\psi = \Phi_2^{-1} \circ \Phi_1$, donde $U_1' = \Phi_1^{-1}(W)$ y $U_2' = \Phi_2^{-1}(W)$ son abiertos, verifica que ψ es un C^s-difeomorfismo. Es decir, el cambio de parametrizaciones es difeomorfismo.*

Demostración. U_i' es abierto porque Φ_i es continua y W es abierto por ser cada Φ_i un homeomorfismo de acuerdo a la proposición previa. Sabemos que ψ es homeomorfismo por ser composición de homeomorfismos. Ahora, sea $p \in W$. Entonces, tenemos $\Phi_i^{-1} = \Omega_i^{-1} \circ \pi \circ \sigma$, con Ω_i un C^s-difeomorfismo y σ el aplanador, como en la demostración de la proposición previa, en un entorno de p. Pero entonces $\psi = (\Omega_2^{-1} \circ \sigma \circ \pi) \circ \pi^{-1} \circ \sigma^{-1} \circ \Omega_1 = \Omega_2^{-1} \circ \Omega_1$ en un entorno de $\Phi_1^{-1}(p)$, que es un C^s-difeomorfismo por ser composición de difeomorfismos. Como esto vale para todo $p \in W$, entonces ψ es un C^s-difeomorfismo en todo $\Phi_1^{-1}(W) = U_1'$. □

Definición 61 (Función diferenciable en una variedad)**.** Sea X una variedad C^s k-dimensional de $\mathbb{R}^n$, y $f : X \mapsto \mathbb{R}^m$ una función en la variedad. Se dice que f **es diferenciable en** $p \in X$ si $\exists \Phi : U \subset \mathbb{R}^k \mapsto X$ con U abierto y Φ parametrización regular tal que $p \in \Phi(U)$ y $f \circ \Phi : U \mapsto \mathbb{R}^m$ es diferenciable en $\Phi^{-1}(p)$.

Proposición 85. *Dadas $\Phi_i : U_i \mapsto X$ dos parametrizaciones regulares con $p \in X$ y $p \in \Phi_1(U_1) \cap \Phi_2(U_2)$, y $f : X \mapsto \mathbb{R}^n$, la diferenciablidad de $f \circ \Phi_1$ en $\Phi_1^{-1}(p)$ equivale a la de $f \circ \Phi_2$ en $\Phi_2^{-1}(p)$, luego si f es diferenciable en p,* ***todas las parametrizaciones regulares*** *que lleguen a p verifican la definición anterior.*

Demostración. Tenemos que $f \circ \Phi_2 = f \circ \Phi_1 \circ \psi^{-1}$ con ψ el cambio de coordenadas de la proposición previa, luego si $f \circ \Phi_1$ es diferenciable en $\Phi_1^{-1}(p)$, entonces $f \circ \Phi_2$ lo es en $\psi(\Phi_1^{-1}(p)) = \Phi_2^{-1}(p)$. Análogamente se tiene la otra implicación. □

Proposición 86. *Si X es una C^s-variedad k-dimensional en $\mathbb{R}^n$ y $p \in X$, entonces $\exists U \subset \mathbb{R}^n$ abierto con $p \in U$, tal que $U \cap X = \Phi(V)$ con $\Phi : V \subset \mathbb{R}^k \mapsto X$ una parametrización regular C^s y V abierto. Es decir, las variedades se pueden parametrizar regularmente de manera local.*

Demostración. Sabemos por definición que $\exists U \subset \mathbb{R}^n$ abierto con $p \in U$ y $\sigma : U \mapsto U'$ el difeomorfismo aplanador tal que $\sigma(X \cap U) = V$ con $V = U' \cap (\mathbb{R}^k \times \{0\})$ que se puede interpretar como un subconjunto de $\mathbb{R}^k$ (ignorando las $n-k$ últimas coordenadas), que de hecho es abierto dado que U' es abierto en $\mathbb{R}^n$. Entonces basta poner $\Phi = \sigma^{-1} : V \mapsto X \cap U$. Por lo que hemos visto antes, $\Phi(V) = \sigma^{-1}(V) = X \cap U$, como σ es C^s-difeomorfismo, entonces Φ es C^s, es inyectiva por serlo σ^{-1}, y finalmente $D\Phi(x) = D\sigma^{-1}(x)$ que es invertible porque σ es difeomorfismo. □

Proposición 87. *Sea X una C^s-subvariedad k-dimensional de $\mathbb{R}^n$ y $p \in X$. Entonces, $\exists U \subset \mathbb{R}^n$ abierto con $p \in U$ tal que $U \cap X = \{P(x, \varphi(x)) : x \in V\}$ donde $P(x, \varphi(x))$ es una parametrización grafo, $V \subset \mathbb{R}^k$ abierto y $\varphi : V \mapsto \mathbb{R}^{n-k}$ con $\varphi \in C^s(V)$.*

Demostración. La proposición previa nos dice que hay Φ una parametrización regular cuya parte parametrizada contiene a p. En la demostración de la proposición 79 vimos que toda inmersión verifica que su imagen es unión de grafos, dado que en un entorno de cada punto de su imagen, se puede expresar la parametrización como grafo. Por tanto, combinando ambos resultados, se tiene que en un entorno de p, la parte parametrizada es un grafo. □

Proposición 88. *Si X es una C^s-subvariedad k-dimensional de $\mathbb{R}^n$, y $U \subset \mathbb{R}^n$ es abierto con $U \cap X \neq \emptyset$ y $f : U \mapsto \mathbb{R}^m$ es $f \in C^r(U)$, entonces $\overline{f} = f|_{U \cap X}$, la restricción sobre la variedad, es diferenciable de regularidad $C^{\text{mín}\{r,s\}}$ en la variedad.*

Demostración. Dado $p \in X \cap U$, sabemos que una parametrización regular $\exists \Phi : V \mapsto X$ tal que $p \in \Phi(V)$, $V \subset \mathbb{R}^k$ abierto. Pero entonces la función $\overline{f} \circ \Phi : \mathbb{R}^k \mapsto \mathbb{R}^m$ es de regularidad $C^{\text{mín}\{r,s\}}$ en p, dado que es composición de Φ que es C^s y f que es C^r. Por definición, entonces, se tiene lo que se quería. □

4.1. Extremos condicionados. Multiplicadores de Lagrange.

La motivación es restringir funciones a variedades, es decir, establecer relaciones que deben verificar las variables de las funciones, y tras esa restricción, obtener los puntos donde las funciones alcanzan su valor máximo o mínimo.

Proposición 89. *Sea X una C^s-variedad de dimensión $n-m$ en $\mathbb{R}^n$ dada por $X = g^{-1}(a)$ donde $g : U \subset \mathbb{R}^n \mapsto \mathbb{R}^m$, con U abierto y $g \in C^s$, y para la cual $Dg(x)$ es de rango máximo (m) $\forall x \in X$.*

Sea ahora $f : V \subset \mathbb{R}^n \mapsto \mathbb{R}$ con V abierto. Si $V \cap X \neq \emptyset$, y ponemos $\overline{f} = f|_{V \cap X}$, se tiene que si $p \in V \cap X$ es un ***extremo local*** *de $\overline{f}$, entonces debe cumplirse lo siguiente:*

1. *$\exists l : \mathbb{R}^m \mapsto \mathbb{R}$ lineal tal que p es crítico en la función $f - l \circ g$.*

2. *$\exists \lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$, llamados **multiplicadores de Lagrange**, que verifican:*

$$\nabla f(p) = \lambda^T Dg(p) = \sum_{i=1}^{m} \lambda_i \nabla g_i(p)$$

Demostración. Denotamos $D_1 f(p) = [\frac{\partial f(p)}{\partial(x_1,\ldots,x_{n-m})}]$ y $D_2 f(p) = [\frac{\partial f(p)}{\partial(x_{n-m+1},\ldots,x_n)}]$ y lo mismo para g.

Reordenamos si hace falta las variables para que $\exists \Omega \subset \mathbb{R}^n$ abierto, con $p \in \Omega$, tal que $\Omega \cap X = Gr(\varphi, V)$ con $V \subset \mathbb{R}^{n-m}$ abierto y $\varphi : V \mapsto \mathbb{R}^m$ una función C^s. Recordamos de la demostración de la proposición 78 que esto es posible cuando $D_2 g(p)$ es invertible, pero como tiene rango m basta con reordenar convenientemente las variables.

Tenemos que la función $F : V \mapsto \mathbb{R}$ dada por $F(x) = \overline{f}(x, \varphi(x))$, tiene un extremo en x_0, donde $p = \Phi(x_0) = (x_0, \varphi(x_0))$, dado que $\overline{f}$ tenía un extremo en p. Por lo tanto, $0 = DF(x_0) = Df(p) \cdot D\Phi(x_0) = D_1 f(p) + D_2 f(p) D\varphi(x_0)$.

Por otra parte, como $g(p) = a$, entonces $g(x_0, \varphi(x_0)) = a$, de tal manera que, diferenciando, $0 = Dg(p)D\Phi(x_0) = D_1 g(p) + D_2 g(p) D\varphi(x_0)$. En ambos casos hemos usado que $D\Phi(x_0) = \begin{pmatrix} I_{(n-m)\times(n-m)} \\ D\varphi(x_0) \end{pmatrix}$.

Ahora tenemos entonces que $D\varphi(x_0) = -D_2 g(p)^{-1} D_1 g(p)$. Sustituyendo en la primera expresión se sigue que $0 = D_1 f(p) - D_2 f(p) D_2 g(p)^{-1} D_1 g(p) = D_1 f(p) - l \circ D_1 g(p)$. Donde la aplicación lineal $l = D_2 f(p) D_2 g(p)^{-1} : \mathbb{R}^m \mapsto \mathbb{R}$. Por otra parte, directamente de la definición de l se tiene que $0 = D_2 f(p) - l \circ D_2 g(p)$. Combinando las expresiones se llega a $0 = Df(p) - l \circ Dg(p)$, que muestra lo que se quería. Para terminar, como $l(y) = \sum_{i=1}^m \lambda_i y_i$, entonces se tiene que $0 = \frac{\partial f(p)}{\partial x_j} - \sum_{i=1}^m \lambda_i \frac{\partial g_i(p)}{\partial x_j}$. □

Proposición 90. *Sea X una C^s-subvariedad k-dimensional en $\mathbb{R}^n$. Se tiene que toda componente conexa de X es otra C^s-subvariedad k-dimensional, y además son conexas por arcos.*

Demostración. Si es conexa hemos acabado. Si no sea $p \in X$ y X_p la componente conexa de p. Dado $q \in X_p$, seleccionamos U_q entorno de q que verifica $(X \cap X_p) \backslash U_q = \emptyset$, es decir $U_q \cap X = U_q \cap X_p$. El $\sigma : V_q \mapsto V_q'$ aplanador que existe por definición de X, entonces $\sigma|_{V_q \cap U_q}$ sobre su imagen sigue siendo difeomorfismo C^s y por como se eligió el U_q, verifica que $\sigma(X_q \cap V_q \cap U_q) = \sigma(X \cap V_q \cap U_q)$ que sabemos que queda en $\mathbb{R}^k \times \{0\}$ como se quería.

Ahora, para ver que X_p es conexa por arcos, sea X_p' la componente arcoconexa de p. Lo que vamos a ver es que X_p' es abierta y cerrada a la vez en X_p, luego como es no vacía debe coincidir. Dada $\{q_n\}_n \subset X_p' \mapsto q \in X_p$, como $\exists U \subset \mathbb{R}^n$ abierto tal que $U \cap X = \{P(x, \psi(x)) : x \in V\}$, V abierto de $\mathbb{R}^k$, tomamos n suficientemente grande como para que $q_n \in U$ y por tanto $q_n = P(x_n, \psi(x_n))$. Asimismo escribamos $q = P(x_0, \psi(x_0))$. Como V es abierto, $\exists \delta > 0$ tal que $B_\delta(x_0) \subset V$. Si n es lo suficientemente grande como para que $q_n \in U$ y además $x_n \in B_\delta(x_0)$, entonces como $[x_n, x_0] \in B_\delta(x_0) \subset V$, podemos definir el arco continuo: $\gamma(t) = P((1-t)x_n + tx_0, \psi((1-t)x_n + tx_0))$ que conecta q_n con q, luego $q \in X_p'$.

Por otro lado, si $q \in X_p'$, tenemos que si $\Phi(x) = P(x, \psi(x))$, como es parametrización regular sobre la variedad X_p, $\Phi(V)$ es abierto de X. Tomamos $B = B_r(\Phi^{-1}(q)) \in V$, conexa por arcos, luego $\Phi(B)$ es abierto conexo por arcos y por tanto $q \in \Phi(B) \subset X_p'$. □

5. Formas diferenciales de $\mathbb{R}^n$

Definición 62. Sea $U \in \mathbb{R}^n$ abierto. Una **1-forma exterior** es una aplicación definida en U, dada por: $\omega : p \mapsto \omega_p$ con $\omega_p : \mathbb{R}^n_p \mapsto \mathbb{R}$ un elemento del dual de $\mathbb{R}^n_p$, el espacio tangente a p (isomorfo a $\mathbb{R}^n$).

Obsérvese que fijada una base $\{e_1, \ldots, e_n\}$ en $\mathbb{R}^n_p$, denotando a su dual por $\{(dx_1)_p, \ldots, (dx_n)_p\}$ (esta notación proviene de denominar x_i a las formas duales de la base, y como son lineales, su diferencial coincide con ellas), tenemos que $\omega(p) = \sum_{i=1}^n \omega_i(p)(dx_i)_p$, donde cada $\omega_i(p)$ es la coordenada i-ésima de $\omega(p)$ en la base dual. Es por esto que podemos escribir $\omega = \sum_{i=1}^n \omega_i dx_i$. Una 1-forma se dice **clase** $C^k(U)$ si todas las aplicaciones ω_i lo son.

Para definir las formas diferenciales tenemos que recordar aspectos de álgebra multilineal:

Definición 63. Sea E un espacio vectorial real. Se dice que la forma k-lineal $\varphi : E^k \mapsto \mathbb{R}$ es **alternada** si $\varphi(u_{\sigma(1)}, \ldots, u_{\sigma(k)}) = sgn(\sigma) \cdot \varphi(u_1, \ldots, u_k)$, para cualesquiera $u_1, \ldots u_k \in E$, y $\sigma \in S_k$.

El espacio de las formas k-lineales alternadas se denomina **potencia exterior k-ésima de** E y se denota $\Lambda^k(E)$. Forma un espacio vectorial real.

La definición es equivalente a que φ cambie de signo cuando se intercambian dos parámetros. Es por esto que si dos parámetros coinciden, la forma vale 0.

Definición 64. Dadas las formas 1-lineales alternadas $\varphi_1, \ldots, \varphi_k \in \lambda^1(E)$, definimos su **producto exterior** $\psi \equiv (\varphi_1 \wedge \cdots \wedge \varphi_k) : E^k \mapsto \mathbb{R}$ como la forma k-lineal alternada dada por:

$$\psi(x_1, \ldots, x_n) = \det(\varphi_i(x_j)) = \det \begin{pmatrix} \varphi_1(x_1) & \ldots & \varphi_1(x_k) \\ \vdots & \ddots & \vdots \\ \varphi_k(x_1) & \ldots & \varphi_k(x_k) \end{pmatrix}$$

Como vemos la operación convierte k formas 1-lineales (por tanto alternadas) en una forma k-lineal alternada.

Proposición 91. *Sea E un $\mathbb{R}$-espacio n-dimensional y $\varphi_1, \ldots, \varphi_k \in \Lambda^1(E)$. Entonces, si $\{dx_1, \ldots, dx_n\}$ es la base de $\Lambda^1(E) \simeq E^*$, se tiene que $\{dx_{i_1} \wedge \cdots \wedge dx_{i_k} : 1 \leq i_1 < \ldots < i_k \leq n\}$ es una base de $\Lambda^k(E)$, y por tanto $\dim \Lambda^k(E) = \binom{n}{k}$ si $k \leq n$ y 0 si $k > n$.*

Demostración. Sean constantes $\{a_{i_1 i_2 \ldots i_k}\}_{1 \leq i_1 < \ldots < i_k \leq n}$ tales que $0 = \sum_{1 \leq i_1 < \ldots < i_k \leq n} a_{i_1 \ldots i_k}(dx_{i_1} \wedge \cdots \wedge dx_{i_k})$. Puesto que, atendiendo a la definición, $(dx_{i_1} \wedge \cdots \wedge dx_{i_k})(e_{i_1}, \ldots, e_{i_k}) = \det I_k = 1$, y si $1 \leq j_1 < \ldots < j_k \leq n$ y algún $j_t \neq i_t$, se tiene (también de la definición) que $(dx_{i_1} \wedge \cdots \wedge dx_{i_k})(e_{j_1}, \ldots, e_{j_k}) = 0$, concluimos que $0 = \sum_{1 \leq i_1 < \ldots < i_k \leq n} a_{i_1 \ldots i_k}(dx_{i_1} \wedge \cdots \wedge dx_{i_k})(e_{j_1}, \ldots, e_{j_k}) = a_{j_1 \ldots j_k}$ y por tanto cada coeficiente es nulo (evaluando en el multivector correspondiente).

Para ver que además generan todo el espacio, sea $\psi \in \Lambda^k(E)$ y definimos $\tilde{\psi} = \sum_{1 \leq i_1 < \ldots < i_k \leq n} \psi(e_{i_1}, \ldots, e_{i_k})(dx_{i_1} \wedge \cdots \wedge dx_{i_k})$. Se afirma que $\tilde{\psi} = \psi$. En primer lugar $\tilde{\psi} \in \Lambda^k(E)$ claramente, por ser combinación lineal de elementos de ese espacio. En segundo lugar, según lo discutido anteriormente, $\tilde{\psi}(e_{j_1}, \ldots, e_{j_k}) = \psi(e_{j_1}, \ldots, e_{j_k})$, siempre que $1 \leq i_1 < \ldots < i_k \leq n$ dado que se anulan todos los sumandos salvo el conveniente. Obsérvese que en esta desigualdad podemos reordenar en ambos lados como queramos y sigue valiendo, dado que como mucho aparecen en ambos lados dos signos negativos que se cancelan. Por tanto ambas formas k-lineales coinciden en todo subconjunto de k elementos de la base, y eso basta para afirmar la igualdad: dado $v_1, \ldots, v_k \in E$, tenemos que $v_i = \sum_j \lambda_{ij} e_j$, de tal modo que $\psi(v_1, \ldots, v_k) = \psi(\sum_{j_1} \lambda_{1 j_1} e_{j_1}, \ldots, \sum_{j_k} \lambda_{k j_k} e_{j_k}) = \sum_{j_1} \ldots \sum_{j_n} \lambda_{1 j_1} \ldots \lambda_{k j_k} \psi(e_{j_1}, \ldots, e_{j_k})$, y como vemos el valor se reduce a una combinación en imágenes de subconjuntos de la base, que coinciden en ψ y $\tilde{\psi}$. $\square$

Observación 10 (Notación de multiíndices). Como vemos, los índices de la forma $(i_1, \ldots, i_k)$ tales que $1 \leq i_1 < \ldots < i_k \leq n$ aparecen con frecuencia. Por ello, introducimos la notación $I_k = (i_1, \ldots, i_k)$ con $1 \leq i_1 < \ldots < i_k \leq n$, y entonces $v_{I_k} = (v_{i_1}, \ldots, v_{i_k})$, por ejemplo.

Definición 65 (k-forma diferenciable)**.** Sea $U \subset \mathbb{R}^n$ abierto. Una **k-forma exterior** es una función $\omega : U \mapsto \Lambda^k(\mathbb{R}^n)$, donde cada $\omega(p) \in \Lambda^k(\mathbb{R}^n)$. Por tanto, podemos poner $\omega(p) = \sum_{I_k} a_{I_k}(p)(dx_{I_k})_p$, donde $(dx_{I_k})_p = (dx_{i_1})_p \wedge \cdots \wedge (dx_{i_k})_p$. Por tanto, como función de p, denotamos:

$$\omega = \sum_{I_k} a_{I_k} dx_{I_k}$$

Donde cada coordenada $a_{I_k} : U \mapsto \mathbb{R}$ es una función del punto de U. Si cada a_{I_k} es $C^s(U)$, entonces la k-forma se dice C^s. Asimismo, si $s \geq 1$, se denomina **k-forma diferenciable**.

Definición 66 (Producto exterior de k-formas)**.** Dadas una k-forma ω y una l-forma φ en U, definimos una nueva forma de la siguiente manera: si $\omega = \sum_{I_k} a_{I_k} dx_{I_k}$ y $\varphi = \sum_{J_k} b_{J_k} dx_{J_k}$, ponemos su **producto exterior**:

$$\omega \wedge \varphi = \sum_{I_k, J_k} a_{I_k} b_{J_k} (dx_{I_k} \wedge dx_{J_k})$$

Que como vemos es una $k{+}l$-forma, dado que, una vez evaluada en un punto, tenemos una combinación lineal de productos exteriores de $k+l$ aplicaciones 1-lineales, lo que se trata de una aplicación $k+l$-lineal.

Obsérvese que aunque a priori parece que haya $\binom{n}{k}\binom{n}{l}$ sumandos, muchos de los productos exteriores se van a cancelar en cuanto se repita alguna de las dx_{i_k}, dado que sin importar en que puntos evaluemos, el $k + l$ funcional resultante va a ser el nulo.

Observación 11. Un ejemplo de producto exterior. Damos en $\mathbb{R}^3$ la 1-forma $\omega = x_1 dx_1 + x_2 dx_2 + x_3 dx_3$ y la 2-forma $\varphi = x_1(dx_1 \wedge dx_2) + dx_1 \wedge dx_3$. Para tomar el producto tenemos que ir, por cada sumando, multiplicando las 'funciones coordenada' y tomando el exterior de cada dx_i involucrado. Esto, eliminando sumandos con productos exteriores nulos, da lugar a:

$\omega \wedge \varphi = x_2(dx_2 \wedge dx_1 \wedge dx_3) + x_3 x_1 (dx_3 \wedge dx_1 \wedge dx_2)$. Con el fin de agrupar todo en un único sumando (sabemos que solo hay uno porque es una 3-forma en $\mathbb{R}^3$ luego el funcional 3-lineal en todo punto está en $\Lambda^3(\mathbb{R}^3)$ que es de dimensión 1), podemos reordenar los funcionales 1-lineales que aparecen en cada producto exterior al evaluar en un punto, lo que nos da lugar a un cambio de signo por cada trasposición (recordemos se trataba de un determinante), de donde $\omega \wedge \varphi = (x_3 x_1 - x_2)(dx_1 \wedge dx_2 \wedge dx_3)$. La aplicación de coordenadas es $(x_1, x_2, x_3) \mapsto x_3 x_1 - x_2$ que es C^∞ luego así lo es la forma.

Proposición 92. *Sean ω una k-forma, φ una s-forma, y θ una r-forma sobre $U \subset \mathbb{R}^n$. Se tienen:*

1. $(\omega \wedge \varphi) \wedge \theta = \omega \wedge (\varphi \wedge \theta)$
2. *Si $r = s$, y $\alpha, \beta \in \mathbb{R}$, entonces $\omega \wedge (\alpha\varphi + \beta\theta) = \alpha\omega \wedge \varphi + \beta\omega \wedge \theta$.*
3. $\omega \wedge \varphi = (-1)^{ks} \varphi \wedge \omega$.

Demostración. El 1 y 2 son rutinarios, aplicando la definición 66. Para el 3, tenemos que $\omega \wedge \varphi = \sum_{I,J} a_I b_J dx_I \wedge dx_J = \sum_{I,J} a_I b_J (-1)^{ks} dx_J \wedge dx_I = (-1)^{ks} \varphi \wedge \omega$. Hemos usado que cada uno de los s elementos de dx_J los tenemos que mover al otro lado mediante k trasposiciones con cada uno de los elementos de dx_I, lo que da lugar a un total de ks trasposiciones (y por tanto, cambios de signo). □

Definición 67. Sea $X \subset \mathbb{R}^n$ una variedad C^s de dimensión k. Se dice que ω es una **r-forma diferencial exterior en la variedad** X si asocia cada $p \in X$ a $\omega(p) \in \Lambda^k(T_p X)$, con la aplicación $p \mapsto \omega(p)$ diferenciable en X.

Definición 68. Sea $X \subset \mathbb{R}^m$, $Y \subset \mathbb{R}^n$, variedades C^s, de dimensiones k y l, respectivamente. Sea $f : X \mapsto Y$, $f \in C^1$. Si ω es una r-forma diferencial exterior en Y, f induce una r-forma diferencial exterior en X, dada por:

$$f^*(\omega)_p(v_1,\dots,v_k)=\omega_{f(p)}(df_p(v_1),\dots,df_p(v_k)) \qquad v_1,\dots,v_k\in T_pX$$

Esta forma se denomina **pull-back de** ω, o **forma inducida en** X **por cambio de coordenadas**.

Esto es porque la aplicación $(df)_p : T_pX \mapsto T_{f(p)}Y$ relaciona ambos espacios tangentes. Si $f : X \mapsto Y$ es C^1 y $p\in X$, entonces $\exists\Phi : \Omega\subset\mathbb{R}^k\mapsto X$, con Ω abierto, y Φ coordenadas locales en un entorno $\Phi(\Omega)\subset X$. Que f sea C^1 indica que $f\circ\Phi$ es C^1 en Ω. Si $v\in T_pX$, sabemos que $v=\gamma'(0)$ para una $\gamma : (-\delta,\delta)\mapsto X$, $\gamma(0)=p$. Entonces $\tilde\gamma : f\circ\gamma : (-\delta,\delta)\mapsto Y$ es curva de Y y también C^1. Se tiene $\tilde\gamma(0)=f(p)=q\in Y$. Además $\tilde\gamma'(0)=df_p\circ\gamma'(0)=df_p(v)$.

Proposición 93. *Sea $f : X\mapsto Y$, función C^1 entre las variedades $X\subset\mathbb{R}^m$ e $Y\subset\mathbb{R}^n$, C^s, de dimensiones k y l. Si ω,φ son r-formas diferenciales en Y y g es una 0-forma en Y (función escalar), entonces:*

1. *$f^*(\alpha\omega\wedge\beta\varphi)=\alpha f^*(\omega)+\beta f^*(\varphi)$.*
2. *$f^*(g\omega)=f^*(g)f^*(\omega)$, entendiendo que $f^*(g)=g\circ f$.*
3. *Si $\varphi_1,\dots,\varphi_s$, son 1-formas diferenciales de Y, entonces $f^*(\varphi_1\wedge\cdots\wedge\varphi_s)=f^*(\varphi_1)\wedge\cdots\wedge f^*(\varphi_s)$.*

Demostración. $1,2$ son inmediatas de la definición. Para 3, sea $p\in X$ y sea $q=f(p)\in Y$, y $v_1,\dots,v_s\in T_pX$. Entonces $f^*(\varphi_1\wedge\cdots\wedge\varphi_s)(v_1,\dots,v_s)=(\varphi_1\wedge\cdots\wedge\varphi_s)(Df_p(v_1),\dots,Df_p(v_s))=\det(\varphi_i(Df_p(v_j)))=\det(f^*(\varphi_i)(v_j))=[f^*(\varphi_1)\wedge\cdots\wedge f^*(\varphi_s)](v_1,\dots,v_s)$. $\square$

Proposición 94. *En las condiciones de la proposición previa, si $(x_1,\dots,x_k)$ e $(y_1,\dots,y_l)$ son coordenadas locales en $p\in X$ y $q=f(p)\in Y$ respectivamente, $\omega=\sum_{I_r}a_{I_r}dy_{I_r}$ es una $r-forma$ diferencial en Y, con las $a_I : V\mapsto\mathbb{R}$, V entorno de q donde se tienen las coordenadas. Denotamos a $f : X\mapsto Y$ por: $f(x_1,\dots,x_k)=(f_1,\dots,f_l)(x_1,\dots,x_k)=(y_1,\dots,y_l)$ (expresión local de $p\mapsto f(p)$ en entornos de p y q donde están las coordenadas locales: realmente f iría de Ω a Ω' los entornos donde están definidas las coordenadas, es un pequeño abuso de notación).*

Se tiene que $f^*(\omega)=\sum_I a_I\circ f df_I$.

Demostración. Por la proposición previa tenemos que $f^*(\omega)=\sum_I f^*(a_I)(f^*(dy_{i_1})\wedge\cdots\wedge f^*(dy_{i_r}))$. Pero si $v\in T_pX$, tenemos que $f^*(dy_i)(v)=dy_i(df_p(v))=d(y_i\circ f)_p(v)=(df_i)_p(v)$. Entonces, $f^*(\omega)=\sum_I a_I\circ f(df_{i_1}\wedge\cdots\wedge df_{i_r})=\sum_I a_I\circ f df_I$. $\square$

Observación 12 (Ejemplo). Damos la forma $\omega=\frac{-y}{x^2+y^2}dx+\frac{x}{x^2+y^2}dy$, fuera del origen. Damos el cambio de coordenadas $f : U\mapsto\mathbb{R}^2$ dado por $f(r,\theta)=(r\cos\theta,r\sin\theta)$, con $U=\mathbb{R}^+\times(0,2\pi)$.

Si $x=r\cos\theta$, entonces $dx=d(r\cos\theta)=\cos\theta dr-r\sin\theta d\theta$. Del mismo modo como $y=d(r\sin\theta)=\sin\theta dr+r\cos\theta d\theta$. Por la proposición previa, entonces, $f^*(\omega)=-\frac{r\sin\theta}{r^2}(\cos\theta dr-r\sin\theta d\theta)+\frac{r\cos\theta}{r^2}(\sin\theta dr+r\cos\theta d\theta)=d\theta$, sería el pull-back en U.

Proposición 95. *Sean $X\subset\mathbb{R}^m$, $Y\subset\mathbb{R}^n$ y $Z\subset\mathbb{R}^a$ variedades C^s de dimensión k,l,r, respectivamente. Entonces si $f : X\mapsto Y$ y $g : Y\mapsto Z$ son diferenciables:*

1. *Si ω,φ son formas diferenciales en Y, entonces $f^*(\omega\wedge\varphi)=f^*(\omega)\wedge f^*(\varphi)$ (forma en X).*
2. *Si ω es una forma en Z, entonces se verifica la igualdad $(g\circ f)^*(\omega)=f^*(g^*(\omega))$*

Demostración. Para 1, dado $p\in X$, $(x_1,\dots,x_k)$ coordenadas locales en Ω entorno de p en X e $(y_1,\dots,y_l)$ coordenadas locales en un entorno Ω' de $q=f(p)$. Entonces $f(x)=(f_1,\dots,f_l)(x)$. Si $\omega=\sum_I a_Idy_I$, $\varphi=\sum_I b_Idy_I$, la proposición previa afirma que $f^*(\omega\wedge\varphi)=f^*(\sum_{I,J}a_Ib_Jdy_I\wedge dy_J)=\sum_{I,J}(a_I\circ f)(b_j\circ f)df_I\wedge df_J=(\sum_I(a_I\circ f)df_I)\wedge(\sum_J(a_J\circ f)df_J)=f^*(\omega)\wedge f^*(\varphi)$

Dando también coordenadas locales $(z_1, \ldots, z_r)$ en un entorno de $g(f(p))$, escribimos la forma $\omega = \sum_K c_K dz_K$. Entonces, $(g \circ f)^*(\omega) = \sum_K c_K(g \circ f)d(g \circ f)_K = \sum_K f^*(c_K \circ g)f^*(d_K g) = f^*(\sum_K c_K \circ g d_K g) = f^*(g^*(\omega))$. $\square$

Definición 69 (Diferencial exterior)**.** Sea X una C^s variedad k-dimensional de $\mathbb{R}^n$. Definimos el operador **diferencial exterior** como una manera de obtener $r+1$ formas a partir de r formas, de la siguiente manera:

1. Si $f : X \mapsto \mathbb{R}$ es una 0-forma diferenciable, entonces dado $p \in X$, en un entorno donde haya coordenadas $(x_1, \ldots, x_k)$, su diferencial exterior es simplemente la 1-forma dada por su diferencial:

$$df = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j} dx_j$$

2. Si ω es una r-forma en X, con $\omega = \sum_I a_I dx_I$, su diferencial exterior es:

$$d\omega = \sum_I da_I \wedge dx_I$$

Proposición 96. *Sean $X \subset \mathbb{R}^m$, $Y \subset \mathbb{R}^n$ variedades C^s de dimensiones k y l, y $f : X \mapsto Y$ diferenciable. Se tiene:*

1. *Si ω, φ son r-formas en X, entonces $d(a\omega + b\varphi) = ad\omega + bd\varphi$, para $a, b \in \mathbb{R}$.*
2. *Si ω es una r-forma de X y φ una r'-forma de X, entonces $d(\omega \wedge \varphi) = d\omega \wedge \varphi + (-1)^r \omega \wedge d\varphi$.*
3. *Si $\omega \in C^2$, entonces $d(d\omega) = 0$.*
4. *Si ω es r-forma de Y, entonces $d(f^*\omega) = f^*(d\omega)$.*

Demostración. 1 es rutinaria. Para 2, en coordenadas locales ponemos $\omega = \sum_I a_I dx_I$, $\varphi = \sum_J b_J dx_J$. Tenemos que $d(\omega \wedge \varphi) = \sum_{I,J} d(a_I b_I) \wedge dx_I \wedge dx_J = \sum_{I,J}(a_I db_J + b_J da_I) \wedge dx_I \wedge dx_J = \sum_{I,J} b_J da_I \wedge dx_I \wedge dx_J + \sum_{I,J} a_I db_J \wedge dx_I \wedge dx_J$.

Evaluamos cada sumando. Por un lado, $\sum_{I,J} b_J da_I \wedge dx_I \wedge dx_J = \sum_J b_J(\sum_I da_I \wedge dx_I) \wedge dx_J = \sum_J b_J d\omega \wedge dx_J = d\omega \wedge \sum_J b_J dx_J = d\omega \wedge \varphi$. Por el otro, $\sum_{I,J} a_I db_J \wedge dx_I \wedge dx_J = \sum_I \sum_J (-1)^r a_I dx_I \wedge db_J \wedge dx_J = (-1)^r \sum_I a_I dx_I \wedge (\sum_J db_J \wedge dx_J) = (-1)^r \omega \wedge d\varphi$.

Para 3, comenzaremos probándolo en caso de que ω sea 0-forma. Tenemos $d(d\omega) = d(\sum_i \frac{\partial f}{\partial x_i} dx_i) = \sum_i \sum_j \frac{\partial^2 f}{\partial x_j \partial x_i} dx_j \wedge dx_i = \sum_{i<j}(\frac{\partial^2 f}{\partial x_j \partial x_i} dx_j \wedge dx_i + \frac{\partial^2 f}{\partial x_i \partial x_j} dx_i \wedge dx_j) = \sum_{i<j}(\frac{\partial^2 f}{\partial x_j \partial x_i}(dx_j \wedge dx_i + dx_i \wedge dx_j)) = 0$ donde usamos el lema de Schwarz para la igualdad de las cruzadas. Ahora, para el caso general, por linealidad basta verlo para formas como $\omega = a_I dx_I$. Tenemos $d\omega = da_I \wedge dx_I$, luego $d(d\omega) = d(da_I) \wedge dx_I - da_I \wedge d(dx_I)$. Por un lado, $d(da_I) = 0$ al ser 0-forma. Por otro, $d(dx_I) = d(dx_{i_1} \wedge \cdots \wedge d_{i_r}) = d(dx_{i_1}) \wedge dx_{i_2} \wedge \cdots \wedge d_{i_r} - dx_{i_1} \wedge d(dx_{i_2} \wedge \cdots \wedge d_{i_r}) = 0 - dx_{i_1} \wedge d(dx_{i_2} \wedge \cdots \wedge d_{i_r})$, y un argumento inductivo permite ver que $d(dx_{i_2} \wedge \cdots \wedge d_{i_r}) = 0$ y hemos acabado.

Finalmente, para 4, escribimos $\omega = \sum_I a_I dy_I$, y entonces $d\omega = \sum_I da_I \wedge dy_I = \sum_{j=1}^{l} \sum_I \frac{\partial a_I}{\partial y_j} dy_j \wedge dy_I$. Por tanto, $f^*(d\omega) = \sum_{j=1}^{l} \sum_I \frac{\partial a_I}{\partial y_j} \circ f df_j \wedge df_I$. Por otra parte, $d(f^*(\omega)) = d(\sum_I a_I \circ f df_I) = \sum_I d(a_I \circ f) \wedge df_I = \sum_{j=1}^{l} \sum_I \frac{\partial a_I}{\partial y_j} \circ f df_j \wedge df_I$, como se quería. $\square$

Proposición 97. *Sea X una C^s variedad k-dimensional en $\mathbb{R}^n$. Si en $p \in X$ tenemos sistemas de coordenadas $(x_1, \ldots, x_k)$ en un entorno $\Omega \subset X$ y $(y_1, \ldots, y_k)$ en $\Omega' \subset X$, y ω una k-forma diferencial en X, es decir, $\omega = adx_1 \wedge \cdots \wedge dx_k = bdy_1 \wedge \cdots \wedge dy_k$. Si $\Phi : U \subset \mathbb{R}^k \mapsto \Omega$ y $\Psi : V \subset \mathbb{R}^k \mapsto \Omega'$ son las* *parametrizaciones de coordenadas, $W = \Omega \cap \Omega'$ y ponemos $U' = \Phi^{-1}(W)$, $V' = \Psi^{-1}(W)$, sabemos que tenemos $F = \Phi^{-1}\Psi : V' \mapsto U'$ el difeomorfismo de cambio de coordenadas.*

Se tiene que $b = (a \circ F)\det(dF)$.

Demostración. $bdy_1 \wedge \cdots \wedge dy_k = F^*(adx_1 \wedge \cdots \wedge dx_k) = (a \circ F)dF_1 \wedge \ldots dF_k = (a \circ F)\det(\partial_i F_j)dy_1 \wedge \cdots \wedge dy_k = (a \circ F)\det(dF)(dy_1 \wedge \cdots \wedge dy_k)$. □

Definición 70. Las bases $B_1 = \{v_1, \ldots, v_n\}$, $B_2 = \{v_1', \ldots, v_n'\}$ de $\mathbb{R}^n$ tienen la misma **orientación** si $P : \mathbb{R}^n \mapsto \mathbb{R}^n$ dada por $P(v_j) = v_j'$ verifica que $\det(P) > 0$.

Definición 71. Sea $X \subset \mathbb{R}^n$ una variedad C^s k-dimensional. Se dice que X es **orientable** si X admite un conjunto de coordenadas locales $\{\Phi_\alpha\}_\alpha$ tal que $\Phi_\alpha : U_\alpha \subset \mathbb{R}^k \mapsto X$, con $X \subset \bigcap_{\alpha\in I}\Phi_\alpha(U_\alpha)$, y tal que si $W_{\alpha,\beta} = \Phi_\alpha(U_\alpha) \cap \Phi_\beta(U_\beta)$ es no vacío, el $\Psi_{\alpha,\beta} : U_\beta' \mapsto U_\alpha'$ difeomorfismo del cambio satisface $\det(d\Psi_{\alpha,\beta}) > 0$ siempre.

Dicho conjunto $\{\Phi_\alpha\}_\alpha$ se denomina **atlas orientado**.

Definición 72. Sea X una C^s variedad k-dimensional en $\mathbb{R}^n$, y ω una r-forma en X. Se define el **soporte de** ω como $sop(\omega) = \overline{A}$, con $A = \{p \in X : \omega(p) \neq 0\}$, y la clausura es en X.

Se dice que ω tiene **soporte compacto** si su soporte es compacto en X.

Es decir, es el conjunto cerrado más pequeño tal que ω vale 0 fuera de él. Obsérvese (se utilizará en la siguiente definición) que ω podría ser una 0-forma, luego esta definición aplica también a funciones en una variedad.

Definición 73. Sea X una variedad C^s k-dimensional de $\mathbb{R}^n$. Sea $A \subset X$ compacto en X, y $\{\Phi_\alpha(U_\alpha)\}_{\alpha\in I}$ un recubrimiento abierto de A por entornos coordenados. Se dice que $\{\varphi_1, \ldots, \varphi_m\}$, con cada $\varphi_m : X \mapsto \mathbb{R}$, y $|I| = m$ una función C^s, es una **partición de la unidad en A subordinada al recubrimiento** $\{\Phi_\alpha(U_\alpha)\}$ si verifica:

1. $\sum_{i=1}^m \varphi_i = 1$ en todo punto de A.
2. $0 \leq \varphi_i \leq 1$ en todo punto de A.
3. Para todo $i \in \{1, \ldots, m\}$, hay un $\alpha_i \in I$ tal que $sop(\varphi_i) \subset \Phi_{\alpha_i}(U_{\alpha_i})$.

Proposición 98. *Si X es una variedad C^s de dimensión k en $\mathbb{R}^n$, $A \subset X$ es compacto, y $\{\Phi_\alpha(U_\alpha)\}_\alpha$ es un recubrimiento abierto de A por entornos coordenados, entonces existe una partición de la unidad.*

Demostración. Necesitaremos tres lemas previos:

Lema 5. *$\exists F : \mathbb{R} \mapsto \mathbb{R}$ tal que $F(x) = F(-x)$, $F(x) \in [0,1]$, $F(x) = 1$ si $|x| \leq 1$, $F(x) = 0$ si $|x| \geq 2$ y $F \in \mathcal{C}^\infty$.*

Demostración. Sea $g(x) = e^{-\frac{1}{1-x^2}}$ en $|x| < 1$, y 0 en $|x| \geq 1$. Sea ahora $h(x) = \int_{-\infty}^x g(y)dy$. Sea $f(x) = \frac{1}{h(1)}h(\frac{x-3}{2})$. Estamos normalizando h y desplazándola, de tal modo que f vale 0 si $x \leq -2$ y vale 1 si $x \geq -1$, y por como se ha construido, es C^∞. Basta con asignar $F(x) = f(-|x|)$ y se tiene todo lo deseado, dado que fuera de 0 sigue siendo C^∞, y en un entorno del 0 (en concreto en $(-1,1)$) es constante luego también C^∞. □

Lema 6. *$\exists f : \mathbb{R}^n \mapsto \mathbb{R}$ tal que $\forall \rho \in O(n)$ aplicación ortogonal de $\mathbb{R}^n$ en $\mathbb{R}^n$, se tiene $f(\rho x) = f(x)$, así como $f(x) \in [0,1]$, $f \in C^\infty$, $f(x) = 1$ si $\|x\| \leq 1$ y $f(x) = 0$ si $\|x\| \geq 2$.*

Demostración. Basta con tomar $f(x) = F(\|x\|)$ con F la dada por el lema anterior, y como $\|x\| = \|\rho x\|$, verifica todo lo deseado. □

Lema 7. *Dada X una variedad C^s k-dimensional de $\mathbb{R}^n$, y $p \in X$, existe una parametrización regular $\psi_P : B_3(0) \subset \mathbb{R}^k \mapsto X$ tal que $\psi_p(0) = p$. (Vale cualquier radio para la bola de partida, pero nos centramos en el radio 3 para uso posterior.)*

Demostración. Sabemos que $\exists \Phi : U \subset \mathbb{R}^k \mapsto X$ parametrización regular en torno a p. Si $x_0 = \Phi^{-1}(p)$, como U es abierto, $B_\delta(x_0) \subset U$. Si ahora definimos el difeomorfismo afín $F : B_3(0) \mapsto B_\delta(0)$, dado por $F(x) = x_0 + \frac{\delta}{3}x$, tenemos que $\psi_p = \Phi \circ F$ verifica lo deseado. $\square$

Procedemos ahora a la demostración de la proposición. Dado $p \in X$, tomamos $\psi_p : B_3(0) \mapsto X$ la garantizada por el lema previo. Sea $V_p = \psi_p(B_3(0))$ y sea $W_p = \psi_p(B_1(0))$. Está claro que $p \in W_p \subset V_p$. Como A es compacto, $\exists p_1, \ldots, p_N \in A \subset X$ tales que $A \subset \bigcup_1^N W_{p_j}$. Denotamos $\psi_i \equiv \psi_{p_i}$, así como $V_i \equiv V_{p_i}$. Sea F la función dada por el lema 6. Definimos las funciones en X:

$$\theta_i(p) = \begin{cases} F \circ \psi_i^{-1}(p) \text{ si } p \in V_i \\ 0 \text{ si } p \in X \backslash V_i \end{cases}$$

Definida así es inmediato comprobar que $sop(\theta_i) \subset \overline{V_i}$. Además, si denotamos $\tilde{V_i} = \psi_i(B_2(0))$, como F anula todos los puntos fuera de la bola de radio 2, podemos escribir:

$$\theta_i(p) = \begin{cases} F \circ \psi_i^{-1}(p) \text{ si } p \in V_i \\ 0 \text{ si } p \in X \backslash \tilde{V_i} \equiv U_i \end{cases}$$

Esta definición, aunque se solapan ambos trozos, es consistente con la original, por lo discutido anteriormente. Además observamos claramente que cada uno de los trozos es C^s, y como solapan (con interior no vacío), se sigue que θ_i es C^s en todo X. Para terminar de definir una partición de la unidad, solo hay que normalizar. Sea $\Theta(p) = \sum_{i=1}^N \theta_i(p)$. Basta con definir:

$$\varphi_i(p) = \begin{cases} \frac{\theta_i(p)}{\Theta(p)} \text{ si } p \in \bigcup_{i=1}^N V_i \\ 0 \text{ si } p \notin \bigcup_{i=1}^N V_i \end{cases}$$

Que son funciones C^s, y verifican la definición de partición de la unidad, subordinada al recubrimiento $\{\psi_{p_i}(B_3(0))\}$. $\square$

A continuación definiremos con estas herramientas la integral en una variedad.

Definición 74. Sea X variedad C^s k-dimensional en $\mathbb{R}^n$, orientable y ω una k-forma diferenciable en X. Si $sop(\omega)$ es compacto en X y damos $\{\Phi_i\}_{i=1}^N$ un atlas orientado con $\Phi_i : U_i \subset \mathbb{R}^k \mapsto X$ tal que $sop(\omega) \subset \bigcup_{i=1}^N \Phi_i(U_i)$, dada una partición de la unidad $\{\varphi_1, \ldots, \varphi_N\}$ subordinada a dicho recubrimiento en $sop(\omega)$, se define:

$$\int_X \omega = \sum_{j=1}^N \int_X \varphi_j \omega$$

Donde, si $\omega = a dx_1 \wedge \cdots \wedge dx^k$ con las coordenadas de $\Phi_{i'}$ en el abierto $U_{i'}$ tal que $sop(\varphi_i \omega) \subset \Phi_{i'}(U_{i'})$ (existe por estar la partición de la unidad subordinada al recubrimiento $\{\Phi_i(U_i)\}$), se define

$$\int_X \varphi_j \omega = \int_{U_{i'}} \varphi_i(x_1, \ldots, x_k) a(x_1, \ldots, x_k) dx_1 \ldots dx_k$$

Como una integral usual en $\mathbb{R}^k$.

La motivación de esta definición es que como $\{\Phi_i(U_i)\}$ recubre el soporte de ω, se tiene que $\omega = (\sum_{i=1^N} \varphi_i)\omega$ al sumar 1 la partición de la unidad en tal soporte.

Proposición 99. *Sea X una variedad C^s k-dimensional en $\mathbb{R}^n$ orientable y ω una k-forma diferenciable en X. Supongamos que tenemos un atlas orientado de X, $\{\Phi_i : U_i \subset \mathbb{R}^k \mapsto X\}_{i=1}^N$, que recubre $sop(\omega)$. Entonces:*

1. *Si $sop(\omega) \subset \Phi_i(U_i) \cap \Phi_j(U_j)$ (y por tanto basta solo una carta local para recubrirlo, ya sea Φ_i o Φ_j), entonces no importa que carta tomemos para dar las coordenadas, la integral valdrá lo mismo.*
2. *Si $\{\Psi_i : U_i \subset \mathbb{R}^k \mapsto X\}_{i=1}^M$ es otro atlas orientado que cubre a $sop(\omega)$, $\{\varphi_1, \dots, \varphi_N\}$ es partición de la unidad subordinada a $\{\Phi_i(U_i)\}$ y $\{\tilde{\varphi}_1, \dots, \tilde{\varphi}_M\}$ otra, subordinada a $\{\Psi_i(U_i)\}$, ambas de $sop(\omega)$, entonces $\sum_{i=1}^N \int_X \varphi_i\omega = \sum_{i=1}^M \int_X \tilde{\varphi}_i\omega$, es decir, la definición es independiente del atlas.*

Demostración. Para 1, sea $W = \Phi_i(U_i) \cap \Phi_j(U_j)$. Si $sop(\omega) \subset W$, entonces si $\omega = a dx_1 \wedge \cdots \wedge dx_k = b dy_1 \wedge \cdots \wedge dy_k$, cada expresión en cada una de las coordenadas, entonces $b = a \circ \psi \det D\psi$ con $\psi = \Phi_i^{-1} \circ \Phi_j$ el difeomorfismo del cambio. Además como el atlas es orientado, $\det D\psi > 0$. Como solo hace falta una carta para recubrir el soporte, la integral consta de un sumando y la única partición de la unidad para un recubrimiento que consta de un único conjunto es $\varphi_1 = 1$. Entonces, $\int_X \omega = \int_{U_j} b dy_1 \dots dy_k = \int_{u_j} a \circ \psi | \det D\psi | dy_1 \dots dy_k = \int_{U_i} a dx_1 \dots dx_k$, con un cambio de variables usual en $\mathbb{R}^k$, con ψ la aplicación de cambio.

Para 2, sea $\Omega_{ij} = \Phi_i(U_i) \cap \Phi_j(U_j)$, para $i \in \{1, \dots, N\}$ y $j \in \{1, \dots, M\}$. Como $\bigcup_{i,j} \Omega_{ij} = (\bigcup_i \Phi_i(U_i)) \cap (\bigcup_j \Phi_j(U_j))$, y el soporte de ω está contenido en ambas uniones, sigue que los abiertos Ω_{ij} cubren $sop(\omega)$. Además las funciones $\{\varphi_i\tilde{\varphi}_j\}_{i,j}$ que se obtienen como producto de una de cada partición de la unidad, son a su vez una partición subordinada a Ω_{ij} dado que $sop(\varphi_i\tilde{\varphi}_j) \subset sop(\varphi(i)) \cap sop(\tilde{\varphi}_j) \subset \Omega_{i'j'}$ y $\sum_{i,j} \varphi_i\tilde{\varphi}_j = \sum_i \varphi_i \sum_j \tilde{\varphi}_j = 1 \cdot 1 = 1$ en $sop(\omega)$, y verifican el resto de propiedades de manera inmediata. Pero entonces $\sum_i \int_X \varphi_i\omega = \sum_i \int_X \varphi_i(\sum_j \tilde{\varphi}_j)\omega = \sum_{i,j} \int_X \varphi_i\tilde{\varphi}_j\omega$ y eso es igual a $\sum_j \int_X \tilde{\varphi}_j\omega$ de manera análoga. □

A continuación daremos una definición más debil de variedad, que permite que el aplanado no tenga por qué ser abierto completamente:

Definición 75. Se dice que $X \subset \mathbb{R}^n$ es una **variedad con borde** de dimensión k y regularidad C^s si en todo punto p se tiene una de las siguientes:

1. $\exists U_p \subset \mathbb{R}^n$ abierto con $p \in U_p$ y $\sigma : U_p \mapsto \sigma(U_p)$ un C^s-difeomorfismo tal que $\sigma(X \cap U_p) = \sigma(U_p) \cap (\mathbb{R}^k \times \{0\})$.
2. $\exists U_p \subset \mathbb{R}^n$ abierto con $p \in U_p$ y $\sigma : U_p \mapsto \sigma(U_p)$ un C^s-difeomorfismo tal que $\sigma(X \cap U_p) = \sigma(U_p)_- \cap (\mathbb{R}^k \times \{0\})$, donde $A_- = A \cap \mathbb{R}^n_-$ y $\mathbb{R}^n_- = \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_1 \leq 0\}$.

Es decir, como vemos, todo punto es localmente difeomorfo a un abierto de $\mathbb{R}^k$, o bien es localmente difeomorfo a un conjunto en $\mathbb{R}^k$ que no es abierto a causa de un *borde* dado por los puntos de la forma $(0, x_2, \dots, x_n)$. Como vemos, las variedades son variedades con borde. Bajo esta interpretación, la siguiente definición es inmediata:

Definición 76. Si X es una variedad con borde, se dice que $p \in X$ es **punto de frontera** si algún $\sigma : U_p \mapsto \sigma(U_p)$ difeomorfismo local C^s de los de tipo 2 según la definición de variedad con borde verifica que $\sigma(p) \in \partial\mathbb{R}^n_- = \{(0, x_2, \dots, x_n)\}$.

Proposición 100. *Dada una variedad con borde, y $p \in \partial X$, todo difeomorfismo σ del tipo 2 según la definición de variedad con borde verifica que $\sigma(p) \in \partial\mathbb{R}^n_-$.*

Demostración. Supongamos $p \in X$ y σ_1, σ_2 difeomorfismos de ese tipo, que podemos asumirlos definidos en $U \subset \mathbb{R}^n$ abierto de p, común. Si se tuviese $\sigma_1(p) = (0, x_2, \ldots, x_n) = q_1$ pero $\sigma_2(p) = (x_1', x_2', \ldots, x_n') = q_2$ con $x_1' \neq 0$, pongamos $\sigma_1 \circ \sigma_2^{-1} \equiv \psi : \sigma_2(U) \mapsto \sigma_1(U)$. Se tiene $\psi(q_2) = q_1$. Dado $V \subset \sigma_2(U)$ abierto en torno a q_2 y $V' = \psi(V)$, hay puntos de V' cuya primera coordenada verifica $x_1 > 0$ (es un entorno de un punto q_1 con primera coordenada nula), y sin embargo $\psi(\sigma_2(U)) = \sigma_1(U) \subset \mathbb{R}^n_-$ lo cual es contradictorio. □

Proposición 101. *Si X es una variedad orientable C^s de dimensión k en $\mathbb{R}^n$ con borde ∂X, entonces ∂X es a su vez una variedad C^s orientable de dimensión $k-1$ y $\partial(\partial X) = \emptyset$. Además X induce en ∂X una orientación (denominada* ***por la normal exterior****).*

Teorema 6 (Stokes)**.** *Sea X una variedad C^s de dimensión k en $\mathbb{R}^n$ orientable y con frontera ∂X. Si ω es una $(k-1)$-forma en X de soporte compacto e $i : \partial X \mapsto X$ es la inclusión, y denotamos $\int_{\partial X} i^*(\omega) \equiv \int_{\partial X} \omega$, se tiene que:*

$$\int_{\partial X} \omega = \int_X d\omega$$

Demostración. Vamos a probar primero que el caso general sigue de probarlo para cuando *sop*$(\omega) \subset \Phi(U)$ un único entorno coordenado, y luego lo probaremos para ese caso.

Sean $\{\Phi_\alpha(U_\alpha)\}_{\alpha \in I}$ abiertos coordenados en X que cubren *sop*(ω), y $\varphi_1, \ldots, \varphi_N$ funciones C^s que dan una partición de la unidad de *sop*(ω) subordinada a esos abiertos. Entonces $\omega = (\sum_i \varphi_i)\omega$, al valer la suma de la partición de la unidad 1 en el soporte. Se tiene que $d\omega = \sum_i d(\varphi_i \omega) = \sum_i \varphi_i d\omega + d\varphi_i \wedge \omega = (\sum_i \varphi_i)\omega + (\sum_i d\varphi_i) \wedge \omega$. Como en *sop*$(\omega)$ $\sum_i \varphi_i = 1$, tenemos que $0 = d(\sum_i \varphi_i) = \sum_i d\varphi_i$, por lo que finalmente $d\omega = \sum_i \varphi_i d\omega$. Si denotamos $d\omega_i = \varphi_i d\omega$, tenemos que *sop*$(d\omega_i) \subset \Phi_j(U_j)$ la carta donde tiene el soporte φ_i, y que, por lo visto anteriormente, $d\omega = \sum_i d\omega_i$.

Ahora aplicamos la suposición que hemos realizado al principio para cada $d\omega_i$:

$$\int_X d\omega = \int_X \sum_{i=1}^N d\omega_i = \sum_{i=1}^N \int_X d\omega_i = \sum_{i=1}^N \int_{\partial X} i^*(\omega_i) = \int_{\partial X} \sum_{i=1}^N i^*(\omega_i) = \int_{\partial X} i^* \left(\sum_{i=1^N} \omega_j \right) = \int_{\partial X} i^*(\omega)$$

A continuación hay que probar la situación que se describe al comienzo de la demostración, y que hemos usado durante esta parte. Tenemos 2 casos, dependiendo de $K =$ *sop*(ω). Como hemos dicho, estamos ahora en el supuesto de que $K \subset \Phi(U)$ para un entorno coordenado.

Si $K \cap \partial X = \emptyset$, como en la frontera vale 0, tenemos que $\int_{\partial X} i^*(\omega) = \int_{\partial X} 0 = 0$. Ponemos en $\Phi(U)$ la expresión en coordenadas $\sum_{j=1}^k a_j dx_{\tilde{j}}$ donde $dx_{\sim j} = dx_1 \wedge \cdots \wedge dx_{j-1} \wedge dx_{j+1} \wedge \cdots \wedge dx_k$. La expresión es de ese tipo porque, recordemos, ω es una $k-1$ forma y la dimensión de X es k. Como $da_j = \sum_{i=1}^n \frac{\partial a_j}{\partial x_i} dx_i$, tenemos que:

$$d\omega = \sum_{j=1}^k da_j \wedge dx_{\sim j} = \sum_{j=1}^k \frac{\partial a_j}{\partial x_j} (-1)^{j-1} dx_1 \wedge \cdots \wedge dx_k$$

Dado que el único sumando de da_j que sobrevive es el que va con dx_j, y reordenarlo cuesta $j-1$ trasposiciones. Como K es compacto en U, y **no interseca la frontera**, podemos escoger $Q \subset U \subset \mathbb{R}^k$, siendo $Q = \times_{j=1}^j [b_j, c_j]$ un paralelepípedo que encierre a U y con $c_1 < 0$, es decir, que no toca la banda vertical $H_n = \{x \in \mathbb{R}^k : x_1 = 0\}$ que correspondería a ∂X por Φ. Pero entonces, la integral queda:

$$\int_X d\omega = \int_Q \left(\sum_{j=1}^k (-1)^{j-1} \frac{\partial a_j}{\partial x_j} \right) dx_1 \ldots dx_k =$$

$$= \sum_{j=1}^{k} \int_{b_1}^{c_1} \dots \int_{b_{j-1}}^{c_{j-1}} \int_{b_{j+1}}^{c_{j+1}} \dots \int_{b_k}^{c_k} \left(\int_{b_j}^{c_j} \frac{\partial a_j}{\partial x_j} dx_j \right) dx_1 \dots dx_{j-1} dx_{j+1} \dots dx_k$$

Pero $\int_{b_j}^{c_j} \frac{\partial a_j}{\partial x_j} dx_j = a_j(x_1, \dots, c_j, \dots, x_k) - a_j(x_1, \dots, b_j, \dots, x_k)$. Recordemos que hemos tomado el Q con $U \subset Q$ y además *por fuera* (con distancia positiva entre sus aristas y todo punto de U, podíamos hacerlo porque U es abierto salvo, como mucho, H_n si tocase a la frontera, cosa que no ocurre), es decir, las aristas de Q quedan fuera de U y por tanto $a_j(x_1, \dots, c_j, \dots, x_k) - a_j(x_1, \dots, b_j, \dots, x_k) = 0 - 0 = 0$ y hemos acabado con este caso.

Si, por el contrario, $K \cap \partial X \neq \emptyset$, a la hora de tomar el $Q \subset U \subset \mathbb{R}^k$ con $K \subset Q$ como antes, no podemos hacerlo sin que toque a la frontera, es decir, $c_1 = 0$ y el primer intervalo $[b_1, 0]$ toca a $H_k = \{x \in \mathbb{R}^k : x_1 = 0\}$, donde $\Phi^{-1}(\partial X \cap \Phi(U)) = H_k$, luego K no queda excluido completamente del borde de Q. Si $j \neq 1$, el mismo razonamiento de antes nos lleva a que $\int_Q (-1)^{j-1} \frac{\partial A_j}{\partial x_j} dx_1 \dots dx_k = 0$. Si denotamos $Q' = \times_{j=2}^{n} [b_j, c_j]$, tenemos que:

$$\int_Q \frac{\partial A_1}{\partial x_1} dx_1 \dots dx_k = \int_{Q'} \left(\int_{b_1}^{c_1} \frac{\partial A_1}{\partial x_1} dx_1 \right) dx_2 \dots dx_k = \int_{Q'} A(0, x_2, \dots, x_k) dx_2 \dots dx_k$$

En el último paso hemos usado el teorema fundamental del cálculo, que $c_1 = 0$ y que $A(b_1, x_2, \dots, x_k) = 0$ porque $b_1 \times \mathbb{R}^{k-1}$ ya queda fuera de $sop(\omega)$. Pero, por otro lado, $i^*(\omega) = A_1(0, x2, \dots, x_k) dx_2 \wedge \dots \wedge dx_k$ ya que el resto de términos irían acompañados de dx_1 pero $x_1 = 0$ en ∂X por definición, luego $dx_1 = 0$. Entonces:

$$\int_X i^*(\omega) = \int_{H_k} A_1(0, x_2, \dots, x_k) dx_2 \dots dx_k$$

Pero integrar en H_k y en Q' es lo mismo porque $sop(\omega) \subset Q' \subset H_k$. $\square$