Universidad de Granada

Apuntes de Análisis Matemático I

María D. Acosta
Camilo Aparicio
Antonio Moreno

Índice general

Capítulo I

Continuidad

Tema 1

Introducción al Análisis de una variable.

1.1. Resultados fundamentales en $\mathbb{R}$.

Los números reales constituyen la base sobre la cual se asienta el Análisis Matemático. Consecuentemente, la primera premisa para avanzar provechosamente en este área será establecer las propiedades del conjunto $\mathbb{R}$ de los números reales.

Codificamos a continuación las propiedades básicas de $\mathbb{R}$ que de hecho lo caracterizan. $\mathbb{R}$ es un conjunto provisto de una suma y un producto respecto de los cuales *es un cuerpo conmutativo*. Está dotado además de una *relación de orden total compatible con las operaciones del cuerpo*, esto es, una relación de orden $\leq$ que verifica las siguientes propiedades:

i) $a,b \in \mathbb{R} \Rightarrow a \leq b \text{ o } b \leq a$,

ii) $[a,b,c \in \mathbb{R},\ a \leq b] \Rightarrow a+c \leq b+c$, y

iii) $[a,b,c \in \mathbb{R},\ a \leq b,\ 0 \leq c] \Rightarrow ac \leq bc$.

Nótese, sin embargo, que el cuerpo $\mathbb{Q}$ de los números racionales también goza de las anteriores propiedades por lo que lógicamente no son éstas por sí solas las que caracterizan a $\mathbb{R}$. La propiedad fundamental de $\mathbb{R}$ (que ya lo distingue de $\mathbb{Q}$) es el *axioma del supremo*.

Axioma 1.1 (del supremo)**.** *Todo conjunto no vacío y mayorado de números reales tiene supremo, es decir, el conjunto de sus mayorante tiene mínimo.*

Es claro que para que un conjunto de números reales tenga supremo ha de ser no vacío y mayorado. El axioma anterior nos asegura que estas dos condiciones son también suficientes. El supremo de un conjunto A no vacío y mayorado de números reales se notará en lo sucesivo $\sup A$.

Es fácil comprobar a partir del axioma del supremo que todo conjunto A de números reales no vacío y minorado tiene ínfimo, que en lo sucesivo se notará $\inf A$. En efecto, si notamos $-A := \{-a : a \in A\}$, se tiene que

$$m \text{ es minorante de } A \Leftrightarrow -m \text{ es mayorante de } -A$$

De lo anterior se deduce que $\inf A = -\sup(-A)$.

Corolario 1.2. *Todo conjunto de números reales no vacío y acotado tiene supremo e ínfimo.*

Proposición 1.3 (Caracterización de supremo y de ínfimo)**.** *Sea A un conjunto no vacío de números reales y sea x un número real. Entonces*

i) $x = \sup A \Leftrightarrow \begin{cases} a \leq x, \ \forall a \in A \\ \forall \varepsilon > 0, \ \exists a \in A : \ x - \varepsilon < a \end{cases}$

ii) $x = \inf A \Leftrightarrow \begin{cases} x \leq a, \ \forall a \in A \\ \forall \varepsilon > 0, \ \exists a \in A : \ a < x + \varepsilon \end{cases}$

Demostración:

$i)$ $\boxed{\Rightarrow}$ Si $x = \sup A$, x es mayorante de A y dado $\varepsilon > 0$, $x - \varepsilon$ no puede ser mayorante de A, luego existe $a \in A$ tal que $x - \varepsilon < a$.

$\boxed{\Leftarrow}$ Claramente x es mayorante de A. Sea y un mayorante cualquiera de A. Si fuese $y < x$, aplicando la hipótesis al positivo $\varepsilon = x - y$, existiría $a \in A$ tal que $y = x + (y - x) = x - \varepsilon < a$, lo cual es absurdo pues y es un mayorante de A. Así pues $x \leq y$, lo que prueba que x es el mínimo de los mayorantes de A.

$ii)$ Análoga a $i)$. ∎

Comenzamos ya a exponer las primeras consecuencias del axioma del supremo que constituyen los pilares sobre los que se sustenta el Análisis Matemático.

Teorema 1.4. *Sea $\{x_n\}$ una sucesión monótona de números reales. Se verifican las siguientes afirmaciones:*

i) Si $\{x_n\}$ es creciente y mayorada, entonces $\{x_n\} \to \sup\{x_n : n \in \mathbb{N}\}$.

ii) Si $\{x_n\}$ es decreciente y minorada, entonces $\{x_n\} \to \inf\{x_n : n \in \mathbb{N}\}$.

Demostración:

$i)$ Es claro que $\{x_n : \ n \in \mathbb{N}\}$ es un conjunto no vacío y mayorado. Sea

$$L = \sup\{x_n : \ n \in \mathbb{N}\} \ .$$

Fijado $\varepsilon > 0$ existe $m \in N$ tal que $L - \varepsilon < x_m$, de donde se deduce que

$$L - \varepsilon < x_m \leq x_n \leq L < L + \varepsilon, \ \forall n \geq m \ ,$$

donde se ha utilizado que la sucesión $\{x_n\}$ es creciente.

$ii)$ Utilizando $i)$ se obtiene que

$$\lim x_n = -\lim(-x_n) = -\sup\{-x_n : n \in \mathbb{N}\} = \inf\{x_n : n \in \mathbb{N}\}.$$

∎

Definición 1.5. Sean $\{x_n\}$, $\{y_n\}$ dos sucesiones de números reales. Diremos que $\{y_n\}$ es una *subsucesión* o *sucesión parcial* de $\{x_n\}$ si existe una aplicación

$$\sigma : \mathbb{N} \longrightarrow \mathbb{N}$$

estrictamente creciente tal que

$$y_n = x_{\sigma(n)}, \ \forall n \in \mathbb{N} \ .$$

Teorema 1.6 (de los intervalos encajados)**.** *Sea $\{I_n\}$ una sucesión decreciente de intervalos cerrados no vacíos tal que $\{\ell(I_n)\} \to 0$ (donde $\ell(I_n)$ denota la longitud del intervalo I_n). Entonces existe un número real x tal que $\cap_{n=1}^{\infty} I_n = \{x\}$.*

Demostración:

$\cap_{n=1}^{\infty} I_n$ puede tener a lo sumo un punto. En efecto, si $x, y \in \cap_{n=1}^{\infty} I_n$, $|x-y| \leq \ell(I_n)$, $\forall n \in \mathbb{N}$, y en consecuencia $x = y$.

Veamos que dicho conjunto no es vacío. Si para cada natural n notamos $I_n = [a_n, b_n]$, entonces $\{a_n\}$ es creciente y mayorada. El teorema anterior nos asegura que

$$\{a_n\} \to x := \sup\{a_n : n \in \mathbb{N}\} \ .$$

En consecuencia para cada natural m la sucesión $\{a_{m+h}\}_{h\in\mathbb{N}}$, parcial de $\{a_n\}$, también converge a x, de donde deducimos, al ser

$$a_m \leq a_{m+h} \leq b_m, \ \forall h \in \mathbb{N} \ ,$$

que $x \in I_m$, para todo natural m, es decir,

$$x \in \cap_{n=1}^{\infty} I_n \ .$$

$\blacksquare$

Definición 1.7. Una sucesión $\{x_n\}$ de números reales es de *Cauchy* si

$$\forall \varepsilon > 0, \ \exists m \in \mathbb{N} : p, q \geq m \Rightarrow |x_p - x_q| < \varepsilon.$$

Teorema 1.8 (de complitud de $\mathbb{R}$)**.** *En $\mathbb{R}$, toda sucesión de Cauchy es convergente.*

Demostración:

Sea $\{x_n\}$ una sucesión de Cauchy de números reales. Veamos que $\{x_n\}$ está acotada. En efecto, por definición,

$$\exists m \in \mathbb{N} : n \geq m \Rightarrow -1 < x_n - x_m < 1, \text{ o sea } x_m - 1 < x_n < x_m + 1,$$

de donde se deduce que $\{x_n\}$ está acotada.

Para cada natural n definimos

$$a_n := \inf\{x_k : k \geq n\} \quad \text{y} \quad b_n := \sup\{x_k : k \geq n\}$$

(los conjuntos que aparecen en las definición son no vacíos y acotados). Es inmediato que $\{[a_n, b_n]\}$ es una sucesión decreciente de intervalos cerrados y acotados. Fijado $\varepsilon > 0$ existe m tal que

$$p, q \geq m \Rightarrow |x_p - x_q| < \varepsilon.$$

De la expresión $x_p < x_q + \varepsilon,\ \forall p, q \geq m$ se deduce que

$$b_n := \sup\{x_p : p \geq n\} \leq x_q + \varepsilon,\ \forall n, q \geq m\ ,$$

que podemos escribir en la forma $b_n - \varepsilon \leq x_q,\ \forall n, q \geq m$, de donde obtenemos

$$b_n - \varepsilon \leq \inf\{x_q : q \geq n\} := a_n\ .$$

Hemos probado que $n \geq m \Rightarrow b_n - a_n \leq \varepsilon$. El teorema anterior nos asegura que $\cap_{n=1}^{\infty}[a_n, b_n] = \{x\}$, para algún x real. Se tiene ahora para $n \geq m$ que

$$|x_n - x| \leq b_n - a_n \leq \varepsilon\ ,$$

es decir la sucesión $\{x_n\}$ converge a x. ∎

El procedimiento usado en la demostración del teorema de complitud de $\mathbb{R}$, que asigna a cada sucesión acotada $\{x_n\}$ dos sucesiones de números reales $\{a_n\}$ y $\{b_n\}$ dadas por

$$a_n := \inf\{x_k : k \geq n\} \quad \text{y} \quad b_n := \sup\{x_k : k \geq n\}\ ,$$

es especialmente útil. Cuando la sucesión $\{x_n\}$ no está acotada, al menos una de estas dos sucesiones no está definida (en $\mathbb{R}$). Esta dificultad desaparece considerando el siguiente conjunto:

Definición 1.9 (El conjunto $[-\infty, +\infty]$)**.** Sean $-\infty$, $+\infty$ dos objetos matemáticos distintos que no son números reales. En el conjunto $[-\infty, +\infty] = \mathbb{R} \cup \{-\infty, +\infty\}$ se extiende el orden usual de $\mathbb{R}$ definiendo $-\infty < x < +\infty,\ \forall x \in \mathbb{R}$.

A partir del Corolario 1.2 se obtiene el siguiente resultado:

Corolario 1.10. *Todo subconjunto no vacío de $[-\infty,+\infty]$ tiene supremo e ínfimo.*

Definición 1.11 (de límite en $[-\infty,+\infty]$)**.** Se dice que una sucesión $\{x_n\}$ de elementos de $[-\infty,+\infty]$ tiende hacia $x\in[-\infty,+\infty]$, lo que notaremos por $\{x_n\}\to x$, si se verifica alguna de las siguientes tres afirmaciones claramente excluyentes:

i) $\{x_n\}\to x\in\mathbb{R}$ si $\forall\varepsilon>0,\ \exists m\in\mathbb{N}:\quad n\geq m\Rightarrow\quad x-\varepsilon<x_n<x+\varepsilon$ ($\{x_n\}$ *converge a* x).

ii) $\{x_n\}\to+\infty$ si $\forall K\in\mathbb{R},\ \exists m\in\mathbb{N}:\quad n\geq m\Rightarrow K<x_n$ ($\{x_n\}$ *diverge positivamente*).

iii) $\{x_n\}\to-\infty$ si $\forall K\in\mathbb{R},\ \exists m\in\mathbb{N}:\quad n\geq m\Rightarrow x_n<K$ ($\{x_n\}$ *diverge negativamente*).

La siguiente proposición extiende al conjunto $[-\infty,+\infty]$ las siguientes conocidas propiedades en el caso de que las sucesiones sean de números reales.

Proposición 1.12. *Sean $\{x_n\},\{y_n\},\{z_n\}$ tres sucesiones de elementos de $[-\infty,+\infty]$ y sean $x,y\in[-\infty,+\infty]$. Entonces*

i) $[x_n\leq y_n,\ \forall n\in\mathbb{N},\ \{x_n\}\to x,\ \{y_n\}\to y]\Rightarrow x\leq y$.

ii) $[x_n\leq y_n\leq z_n,\ \forall n\in\mathbb{N},\ \{x_n\}\to x,\ \{z_n\}\to x]\Rightarrow\{y_n\}\to x.$

Obsérvese que de i) se deduce que

$$[\{x_n\}\to x,\{x_n\}\to y]\Rightarrow x=y.$$

Esto nos permite definir, en el caso de que $\{x_n\}$ tienda a un elemento de $[-\infty,+\infty]$, el $\{x_n\}$ como el único $x\in[-\infty,+\infty]$ tal que $\{x_n\}\to x$, en cuyo caso escribiremos $\lim x_n=x$.

Ahora del Teorema 1.4 se deduce el siguiente resultado.

Proposición 1.13. *Sea $\{x_n\}$ una sucesión monótona de elementos de $[-\infty,+\infty]$. Se verifican las siguiente afirmaciones:*

i) *Si $\{x_n\}$ es creciente, entonces* $\{x_n\}\to\sup\{x_n:\ n\in\mathbb{N}\}$.

ii) *Si $\{x_n\}$ es decreciente, entonces* $\{x_n\}\to\inf\{x_n:\ n\in\mathbb{N}\}$.

Definición 1.14 (de límite superior y límite inferior)**.** El *límite superior* de una sucesión $\{x_n\}$ en $[-\infty,+\infty]$ es el elemento de $[-\infty,+\infty]$ definido por

$$\limsup x_n:=\lim\ b_n$$

donde para cada n natural

$$b_n := \sup\{x_k :\ k \geq n\}$$

(la sucesión $\{b_n\}$ es una sucesión decreciente de elementos de $[-\infty, +\infty]$ que obligadamente tiene límite).

El *límite inferior* de una sucesión $\{x_n\}$ en $[-\infty, +\infty]$ es el elemento de $[-\infty, +\infty]$ definido por

$$\liminf x_n := \lim\ a_n$$

donde para cada n natural

$$a_n := \inf\{x_k :\ k \geq n\}$$

(la sucesión $\{a_n\}$ es una sucesión creciente de elementos de $[-\infty, +\infty]$ que obligadamente tiene límite).

Nótese que para cada $n \in \mathbb{N}$ se tiene que $a_n \leq b_n$ y por tanto $\lim\ a_n \leq \lim\ b_n$ o lo que es lo mismo

$$\liminf\ x_n \leq \limsup\ x_n.$$

Ahora de la definición de límite y de la Proposición 1.12 ii) se deduce fácilmente la siguiente caracterización de la existencia de límite en $[-\infty, +\infty]$.

Proposición 1.15. *Una sucesión $\{x_n\}$ en $[-\infty, +\infty]$ tiene límite si, y sólo si,*

$$\liminf\ x_n = \limsup\ x_n.$$

En consecuencia en el caso de que la sucesión sea de números reales, ésta es convergente si, y sólo si,

$$\liminf\ x_n = \limsup\ x_n \in \mathbb{R}.$$

Teorema 1.16 (Bolzano-Weierstrass)**.** *Toda sucesión acotada de números reales admite (al menos) una subsucesión convergente.*

Demostración:

Sea $\{x_n\}$ una sucesión acotada de números reales. Sea $L := \limsup x_n$. Es claro que $L \in \mathbb{R}$. Vamos a definir por inducción una aplicación $\sigma : \mathbb{N} \to \mathbb{N}$ estrictamente creciente tal que

$$\{x_{\sigma(n)}\} \ \to\ L\ .$$

Por definición de límite superior, la sucesión $\{b_n\}$ definida por

$$b_n := \sup\{x_k : k \geq n\},\ \forall n \in \mathbb{N}$$

decrece a L. Luego

$$\exists m_1 \in \mathbb{N} :\ L \leq b_{m_1} < L + 1.$$

Por definición de b_{m_1} y la caracterización del supremo

$$\exists \sigma(1) \geq m_1 :\ b_{m_1} - 1 < x_{\sigma(1)} \leq b_{m_1}.$$

Por tanto
$$L-1 < x_{\sigma(1)} < L+1.$$
De nuevo y análogamente
$$\exists m_2 > \sigma(1) : \ L \le b_{m_2} < L + \frac{1}{2}$$
y por definición de b_{m_2}
$$\exists \sigma(2) \ge m_2 : b_{m_2} - \frac{1}{2} < x_{\sigma(2)} \le b_{m_2}.$$
De ambas desigualdades se sigue que
$$L - \frac{1}{2} < x_{\sigma(2)} < L + \frac{1}{2}.$$

Supongamos definido $\sigma(n)$ con $n \ge 2$ tal que $\sigma(n) > \sigma(n-1)$ y
$$L - \frac{1}{n} \ < x_{\sigma(n)} < L + \frac{1}{n} \ .$$
Por definición de L
$$\exists m_{n+1} > \sigma(n) : \ L \le b_{m_{n+1}} < L + \frac{1}{n+1} \ .$$
y por definición de $b_{m_{n+1}}$ y la caracterización del supremo, razonando igual que antes, obtenemos
$$\exists \sigma(n+1) \ge m_{n+1} : \ L - \frac{1}{n+1} < x_{\sigma(n+1)} < L + \frac{1}{n+1}.$$

Queda así definida $\{x_{\sigma(n)}\}$ que verifica
$$|L - x_{\sigma(n)}| < \ \frac{1}{n} \ , \ \forall n \in \mathbb{N},$$
y en consecuencia
$$\{x_{\sigma(n)}\} \to L \ .$$

■

1.2. Numerabilidad.

Sabemos que existen conjuntos finitos e infinitos y que los primeros se clasifican atendiendo a su número de elementos. Para desarrollar la teoría del Análisis Matemático es importante clasificar los subconjuntos infinitos atendiendo también a su "tamaño". Los conjuntos numerables son los conjuntos infinitos "más pequeños".

Probaremos que $\mathbb{Q}$ es numerable (de hecho existe $f : \mathbb{N} \to \mathbb{Q}$ biyectiva con lo que $\mathbb{Q} = \{r_n : \ n \in \mathbb{N}\}$, donde $r_n = f(n)$, $\forall n \in \mathbb{N}$, es una *enumeración* de $\mathbb{Q}$) .

Sin embargo, $\mathbb{R}$ no es numerable.

Definición 1.17 (de conjunto numerable)**.** Un conjunto A se dice *numerable* si es vacío o si existe $f : A \to \mathbb{N}$ inyectiva.

Ejemplos 1.18 (de conjuntos numerables)**.**

a) Todo conjunto finito es numerable.

b) $\mathbb{N}$ es un conjunto infinito numerable.

c) Todo subconjunto de un conjunto numerable es numerable.

d) $\mathbb{Z}$ es numerable. En efecto, la aplicación $f : \mathbb{Z} \to \mathbb{N}$ definida por

$$f(0) = 1 \ , \ f(n) = 2n+1 \ , \ f(-n) = 2n \ , \forall n \in \mathbb{N}$$

es claramente biyectiva.

e) $\mathbb{N} \times \mathbb{N}$ es numerable. En efecto, la aplicación $f : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N}$ definida por

$$f(a,b) = 2^a 3^b \ , \forall (a,b) \in \mathbb{N} \times \mathbb{N}$$

es inyectiva.

f) $\mathbb{Q}$ es numerable. En efecto, la aplicación de $\mathbb{Q}$ en $\mathbb{Z} \times \mathbb{N}$ que a cada número racional r le hace corresponder $(p,n) \in \mathbb{Z} \times \mathbb{N}$, donde $\frac{p}{n}$ es la fracción irreducible de r con denominador positivo, es inyectiva. Sabemos que existe una aplicación inyectiva de $\mathbb{Z} \times \mathbb{N}$ en $\mathbb{N}$ (de hecho existe una biyección). Así podemos definir una aplicación inyectiva de $\mathbb{Q}$ en $\mathbb{N}$, esto es, $\mathbb{Q}$ es numerable.

Hemos descrito ya una biyección de $\mathbb{N}$ sobre $\mathbb{Z}$. En el ejercicio 1.6 se presenta una biyección usual de $\mathbb{N}$ sobre $\mathbb{N} \times \mathbb{N}$.

Teorema 1.19. *Sea $A \subset \mathbb{N}$ infinito. Entonces existe una aplicación $f : \mathbb{N} \to A$ biyectiva.*

Demostración:

Sea $f(1) = \min A$. Supuesto definidos $f(1), f(2), \ldots, f(n)$, sea

$$f(n+1) = \min\{a \in A \setminus \{f(1), f(2), \ldots, f(n)\}\}.$$

Claramente $f(n) < f(n+1)$, $\forall n \in \mathbb{N}$ de donde se deduce que f es inyectiva. Veamos que f es sobreyectiva. Sea $b \in A$ un elemento fijo y definimos

$$k := \max\{n \in \mathbb{N} : \; f(n) \leq b\}$$

(el conjunto anterior es no vacío y finito, de hecho, tiene b elementos a lo sumo). Es claro que $f(k) \leq b$. Si fuese $f(k) < b$, entonces se tendría

$$b \in A \setminus \{f(1), \ldots, f(k)\},$$

y, en consecuencia, se seguiría de la definición de f que $f(k+1) \leq b$, lo que contradice la definición de k. Hemos probado que $f(k) = b$. ∎

Corolario 1.20. *Existe una biyección de $\mathbb{N}$ sobre $\mathbb{Q}$.*

Demostración:

Sea $g : \mathbb{Q} \to \mathbb{N}$ inyectiva, entonces $\mathbb{Q}$ y $g(\mathbb{Q})$ son biyectivos. El resultado se deduce del teorema anterior. ∎

Nuestro próximo objetivo es probar que la unión numerable de conjuntos numerables es numerable y que $\mathbb{R}$ no es numerable. Para ello usaremos el siguiente resultado.

Lema 1.21. *Un conjunto A no vacío es numerable si, y sólo si, existe $g : \mathbb{N} \to A$ sobreyectiva.*

Demostración:

Supongamos que A es numerable y sea $f : A \to \mathbb{N}$ inyectiva. Sea $b \in A$ fijo. La aplicación $g : \mathbb{N} \to A$ definida por

$$\begin{aligned} g(f(a)) &= a, \quad \forall a \in A \\ g(n) &= b, \quad \forall n \in \mathbb{N} \setminus f(A) \;\; (\text{si } f(A) \neq \mathbb{N}) \end{aligned}$$

es claramente sobreyectiva.

Supongamos ahora que existe $g : \mathbb{N} \to A$ sobreyectiva. La aplicación $f : A \to \mathbb{N}$ definida por $f(a) = \min\{n \in \mathbb{N} : \; g(n) = a\}$ es inyectiva. ∎

Proposición 1.22. *La unión numerable de conjuntos numerables es numerable. Es decir si I es un conjunto numerable y A_i es un conjunto numerable para cada $i \in I$, entonces el conjunto $A = \cup_{i \in I} A_i$ es un conjunto numerable.*

Demostración:

Podemos suponer que $I \neq \emptyset \neq A_i$. Existen funciones

$$g : \mathbb{N} \longrightarrow I \ , \ f_i : \mathbb{N} \longrightarrow A_i, \ \forall i \in I$$

sobreyectivas.

Definimos $h : \mathbb{N} \times \mathbb{N} \longrightarrow A$ por

$$h(m,n) = f_{g(n)}(m) \ .$$

Veamos que h es sobreyectiva. Dado $a \in A$, elegimos $i \in I$ tal que $a \in A_i$, entonces elegimos $m,n \in \mathbb{N}$ tales que $g(n) = i$ y $f_i(m) = a$. Al ser $\mathbb{N} \times \mathbb{N}$ numerable se puede construir una aplicación de $\mathbb{N}$ sobre $\mathbb{N} \times \mathbb{N}$, y en consecuencia existe también una aplicación de $\mathbb{N}$ sobre A. En virtud del lema anterior hemos probado que A es numerable. ∎

Teorema 1.23. *El conjunto $\mathbb{R}$ de los números reales no es numerable.*

Demostración:

Basta probar que el intervalo $[0,1]$ no es numerable. Si lo fuese existiría una aplicación $f : \mathbb{N} \rightarrow [0,1]$ sobreyectiva. Dividimos $[0,1]$ en tres intervalos cerrados de igual longitud y al menos en uno de ellos, que notamos I_1, no está el punto $f(1)$. Dividimos I_1 en tres intervalos de igual longitud y al menos en uno de ellos, que notamos I_2, no está $f(2)$. En el n-ésimo paso I_n es un intervalo que no contiene al punto $f(n)$ y cuya longitud es un tercio de la del intervalo I_{n-1}.

La sucesión $\{I_n\}$ así definida es una sucesión decreciente de intervalos cerrados no vacíos de $\mathbb{R}$ verificado que $\ell(I_n) = \frac{1}{3^n}$, para todo natural n. El teorema de los intervalos encajados (Teorema 1.6) nos asegura que existe $x \in [0,1]$ tal que $\cap_{n=1}^{\infty} I_n = \{x\}$. Esto es absurdo pues al ser f sobreyectiva ha de existir $k \in \mathbb{N}$ tal que $f(k) = x$ con lo que $x \notin I_k$ y con mayor motivo $x \notin \cap_{n=1}^{\infty} I_n$. ∎

Para familiarizarse, de una manera "informal" y agradable, con el duro concepto de la numerabilidad recomendamos la lectura del Capítulo 2 de [Thi] titulado "Fábulas".

1.3. Notas

Nota 1.24. Existen diversos procedimientos para presentar el cuerpo de los números reales. En los métodos "constructivos", los axiomas de Peano definen los números naturales ("Los números naturales los hizo Dios y los demás los hizo el hombre", Kronecker).

Axiomas de Peano. Existe un conjunto $\mathbb{N}$, *un elemento* 1 *en* $\mathbb{N}$ *y una aplicación* $\sigma : \mathbb{N} \rightarrow \mathbb{N}$ *que a cada natural n le hace corresponder otro natural* $\sigma(n)$, *llamado su siguiente, verificando:*

i) σ *es inyectiva.*

ii) $1 \notin \sigma(\mathbb{N})$ *(*1 *no es el siguiente de ningún natural).*

iii) Principio de inducción. *Si* A *es un subconjunto de* $\mathbb{N}$ *que satisface:*

a) $1 \in A$,

b) $a \in A \Rightarrow \sigma(a) \in A$,

entonces $A = \mathbb{N}$.

A partir de los números naturales, de manera fácil, se obtienen los números enteros y racionales. Después, para la construcción de los reales a partir de los racionales, hay diversos procedimientos: el método de Cantor parte de familias de sucesiones de Cauchy de números racionales [Lin], en otros procedimientos se parte de sucesiones de intervalos encajados de números racionales [BaGa], que en esencia constituye una variante del método de pares de sucesiones monótonas convergentes de números racionales [Rey] y, por último, la construcción a partir de las cortaduras de Dedekind de números racionales [Gau]. Todos los métodos constructivos conducen a conjuntos que gozan de las mismas propiedades. Los métodos "axiomáticos" admiten la existencia de un conjunto que goza de algunas propiedades, elegidas de tal forma que de ellas puedan deducirse todas las demás [ApPa, Capítulo I]. Así los métodos que llevan a "ejemplificaciones" de $\mathbb{R}$ construyen estructuras matemáticamente idénticas, entendiendo por tal que, aunque los conjuntos resultantes sean diferentes, sin embargo, tienen exactamente las mismas propiedades. Por ello hablamos de *el cuerpo* $\mathbb{R}$ *de los números reales*.

Más formalmente, cuando decimos que dos ejemplificaciones cualesquiera $\mathbb{R}$ y $\mathbb{R}'$ de los números reales son matemáticamente idénticas estamos afirmando que existe una biyección $f : \mathbb{R} \to \mathbb{R}'$ que conserva la estructura de cuerpo ordenado, esto es,

i) $f(x+y) = f(x) + f(y)$,

ii) $f(xy) = f(x)f(y)$ y

iii) $x \leq y \Leftrightarrow f(x) \leq f(y)$.

En [McBo] puede consultarse una demostración detallada de la unicidad del cuerpo de los números reales.

Nota 1.25. Del estudio hecho en la sección segunda es fácil obtener el siguiente importante resultado:
Sea A *un conjunto numerable. Entonces* A *es finito* o equipotente *a* $\mathbb{N}$, *es decir, existe una biyección de* A *sobre* $\mathbb{N}$.
Demostración:

Si A es vacío entonces A es finito. En otro caso, sea $f : A \to \mathbb{N}$ inyectiva, es claro que A y $f(A)$ son equipotentes. Si $f(A)$ es finito, A también lo es, mientras que si $f(A)$ es infinito, aplicando el Teorema 1.19 tenemos que $f(A)$ es equipotente a $\mathbb{N}$ y por tanto A es equipotente a $\mathbb{N}$. ■

1.4. Referencias recomendadas.

[ApPa], [BaGa], [Gau], [Lin], [McBo], [Rey], [SoSi], [Thi].

1.5. Resumen de resultados del Tema 1.

El conjunto $\mathbb{R}$ de los números reales. $\mathbb{R}$ es el "único" cuerpo conmutativo con una relación de orden $\leq$ verificando:

i) $a,b \in \mathbb{R} \Rightarrow a \leq b$ o $b \leq a$ (orden total),

ii) $[a,b,c \in \mathbb{R},\ a \leq b] \Rightarrow\ a+c \leq b+c$ (compatibilidad del orden con la suma), y

iii) $[a,b,c \in \mathbb{R},\ a \leq b,\ 0 \leq c] \Rightarrow\ ac \leq bc$ (compatibilidad del orden con el producto),

que satisface el axioma del supremo, esto es, todo conjunto no vacío y mayorado de números reales tiene supremo (el conjunto de sus mayorante tiene mínimo).

Teorema de los intervalos encajados. *Sea $\{I_n\}$ una sucesión decreciente de intervalos cerrados no vacíos tal que $\{\ell(I_n)\} \to 0$ (donde $\ell(I_n)$ denota la longitud del intervalo I_n). Entonces $\cap_{n=1}^{\infty} I_n = \{x\}$ para algún $x \in \mathbb{R}$.*

Teorema de complitud de $\mathbb{R}$. *En $\mathbb{R}$, toda sucesión de Cauchy es convergente.*

Teorema de Bolzano-Weierstrass. *Toda sucesión acotada de números reales admite una sucesión parcial convergente.*

Conjunto numerable. Un conjunto A se dice *numerable* si es vacío o si existe $f : A \to \mathbb{N}$ inyectiva.

Es fácil deducir de la definición que el conjunto $\mathbb{Z}$ de los números enteros y el conjunto $\mathbb{Q}$ de los números racionales son numerables.

Teorema. *El conjunto $\mathbb{R}$ de los números reales no es numerable.*

Teorema. *Todo conjunto numerable o es finito o equipotente a $\mathbb{N}$.*

1.6. Ejercicios del Tema 1.

1.1 (**Propiedad arquimediana**)
Sean $x, y \in \mathbb{R}$ con $0 < x$. Probar que existe $n \in \mathbb{N}$ tal que $y < nx$. El enunciado equivale a la siguiente afirmación: $\mathbb{N}$ no está mayorado ($\Leftrightarrow \{\frac{1}{n}\} \to 0$).

Indicación: Si se supone, razonando por reducción al absurdo, que $\mathbb{N}$ está mayorado, tendría supremo. Sea $\alpha = \sup \mathbb{N}$. Entonces $n \leq \alpha$, $\forall n \in \mathbb{N}$ y, por la caracterización de supremo, ha de existir un natural m tal que $\alpha - 1 < m$, de donde $\alpha < m+1$ y esto contradice el hecho de que α es el supremo de los naturales y, por tanto, mayorante.

1.2 Sean $x, y \in \mathbb{R}$ con $x < y$. Probar que se verifican las siguientes afirmaciones:

i) (**Densidad de** $\mathbb{Q}$ **en** $\mathbb{R}$). Existe $a \in \mathbb{Q}$ tal que $x < a < y$.

ii) (**Densidad de** $\mathbb{R} \setminus \mathbb{Q}$ **en** $\mathbb{R}$). Existe $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ tal que $x < \alpha < y$.

Indicación: Para *i*): Sea $n \in \mathbb{N}$ tal que $\dfrac{1}{n} < y - x$. Comprobar que el racional

$$a := \frac{E(nx)+1}{n} \ ,$$

donde $E : \mathbb{R} \to \mathbb{Z}$ es la función parte entera definida por

$$E(x) := \max\{p \in \mathbb{Z} : p \leq x\} \ ,$$

verifica $x < a < y$.
Para *ii*). Sea β un irracional positivo. Elijamos en virtud de *i*) un racional no nulo a tal que $\frac{x}{\beta} < a < \frac{y}{\beta}$. Comprobar que el irracional $\alpha := a\beta$ verifica $x < \alpha < y$.

1.3 i) Probar que toda sucesión de números reales convergente es de Cauchy.

ii) Probar que toda sucesión de números reales de Cauchy que admite una subsucesión convergente es convergente.

iii) Deducir el teorema de complitud de $\mathbb{R}$ a partir del teorema de Bolzano-Weierstrass.

1.4 Sea $\{x_n\}$ una sucesión de números reales positivos. Probar que

$$\liminf \frac{x_{n+1}}{x_n} \leq \liminf \sqrt[n]{x_n} \leq \limsup \sqrt[n]{x_n} \leq \limsup \frac{x_{n+1}}{x_n}.$$

En particular $\left\{\dfrac{x_{n+1}}{x_n}\right\} \longmapsto x \Rightarrow \{\sqrt[n]{x_n}\} \longmapsto x$. Por ejemplo $\{\sqrt[n]{n}\} \longmapsto 1$ y $\{\sqrt[n]{n!}\} \longmapsto +\infty$.

Indicación: Para probar la primera desigualdad nótese para cada $n \in \mathbb{N}$, $a_n := \inf\left\{\frac{x_{k+1}}{x_k} : k \geq n\right\}$. Si $\lim a_n = 0$ no hay nada que probar. En otro caso, sea $\alpha \in \mathbb{R}$ tal que $0 < \alpha < \lim\, a_n$ y sea $m \in \mathbb{N}$ tal que $\alpha < a_m$. Para $n > m$ se tiene que

$$\alpha^{n-m} < a_m a_{m+1} \ldots a_{n-1} \leq \frac{x_{m+1}}{x_m}\frac{x_{m+2}}{x_{m+1}} \ldots \frac{x_n}{x_{n-1}} = \frac{x_n}{x_m}$$

es decir $\sqrt[n]{x_m}\alpha\sqrt[n]{\alpha^{-m}} < \sqrt[n]{x_n}$. Deducir que

$$\alpha = \lim \sqrt[n]{x_m}\alpha\sqrt[n]{\alpha^{-m}} \leq \liminf \sqrt[n]{x_n}$$

para concluir que

$$\liminf \frac{x_{n+1}}{x_n} = \lim a_n \leq \liminf \sqrt[n]{x_n}.$$

La segunda desigualdad es bien conocida y la tercera se demuestra de manera análoga a la primera.

1.5 Probar que la aplicación $f : \mathbb{R} \to]-1,1[$ definida por $f(x) := \dfrac{x}{1+|x|}$ es una biyección estrictamente creciente.

Calcular su inversa. ¿Es f^{-1} estrictamente monótona?. ¿Creciente o decreciente?. En general, ¿cómo es la función inversa de una biyección estrictamente creciente entre números reales?.

Extender, conservando el orden, f a una biyección $\hat{f} : [-\infty,+\infty] \to [-1,1]$.

1.6 Una forma natural de enumerar $\mathbb{N}^2$ (biyección $\sigma : \mathbb{N} \longrightarrow \mathbb{N}^2$) viene dada en el siguiente esquema

$$\begin{array}{cccccccc}
(1,1) & \rightarrow & (1,2) & & (1,3) & \cdots & \cdots & (1,m) \\
 & & \downarrow & & \downarrow & & & \downarrow \\
(2,1) & \leftarrow & (2,2) & & (2,3) & \cdots & \cdots & (2,m) \\
 & & & & & & & \\
\cdots & \cdots & & \cdots\cdots & \cdots & \cdots & \cdots & \cdots \\
 & & & & & & & \downarrow \\
 & & & & & & & \\
(m,1) & \leftarrow & (m,2) & \leftarrow & (m,3) & \leftarrow & \cdots & (m,m)
\end{array}$$

es decir, los pares asociados a los primeros naturales son

$$(1,1),(1,2),(2,2),(2,1),$$

$$(1,3),(2,3),(3,3),(3,2),(3,1),\cdots,$$

$$(1,m),(2,m),(3,m)\cdots,(m,m)\cdots,(m,3),(m,2),(m,1).$$

Con la enumeración descrita antes, calcular

a) $\sigma(11),\sigma(13),\sigma(15)$.

b) $\sigma(k)$.

c) $\sigma^{-1}(4,1),\sigma^{-1}(1,4)$.

d) $\sigma^{-1}(n,m)$.

Indicación para el apartado b): $(1,1) \longmapsto 1$, $(n,1) \longmapsto n^2$. Para numerar las

$$(n+1)^2 - n^2 = 2n+1$$

parejas que orlan el cuadrado anterior por una columna de $(n+1)-$parejas y una fila $(n+1)-$parejas, obsérvese que

$$(1,n+1) \longmapsto n^2+1 \ , \ (2,n+1) \longmapsto n^2+2 \ , \cdots \ (n+1,n+1) \longmapsto n^2+n+1$$

y

$$(n+1,1) \longmapsto (n+1)^2 \ , \ (n+1,2) \longmapsto (n+1)^2-1 \ , \cdots \ (n+1,n+1) \longmapsto (n+1)^2-n = n^2+n+1.$$

En [SoSi], Capítulo I puede encontrarse una amplia colección de ejercicios con soluciones.

1.7. Soluciones a los ejercicios del Tema 1.

1.1 El supremo de un conjunto no puede ser menor que uno de sus elementos. Para probar que $\mathbb{N}$ no es acotado tomar $x = 1$.

$$\left[\forall \varepsilon > 0, \exists m \in \mathbb{N} : n \geq m \Rightarrow \frac{1}{n} < \varepsilon\right] \Leftrightarrow [\forall K > 0, \exists m \in \mathbb{N} : n \geq m \Rightarrow K < n],$$

sin más que relacionar ε y K mediante la expresión $K\varepsilon = 1$.

1.2 i) $\frac{E(nx)}{n} \leq x < \frac{E(nx)+1}{n} \Rightarrow x < a \leq x + \frac{1}{n} < x + (y - x) = y.$

ii) El producto de un racional por un irracional es irracional.

1.3 i) Sea $\{x_n\} \to x$. Sea $\varepsilon > 0$

$$\exists m \in \mathbb{N} : n \geq m \Rightarrow |x - x_n| < \varepsilon.$$

$$p, q \geq m \Rightarrow |x_p - x_q| \leq |x_p - x| + |x - x_q| < 2\varepsilon.$$

ii) $|x - x_n| \leq |x - x_{\sigma(n)}| + |x_{\sigma(n)} - x_n| < 2\varepsilon$, pues $n \leq \sigma(n)$, $\forall n \in \mathbb{N}$.

iii) Como toda sucesión $\{x_n\}$ de Cauchy es acotada, el teorema de B-W nos asegura que existe $\{x_{\sigma(n)}\} \to x$. El resultado se sigue de *ii*).

1.4 Seguir las indicaciones

1.5 $f^{-1}(y) = \frac{y}{1-|y|}$, $\forall y \in \mathbb{R}$.

La función inversa de una biyección estrictamente creciente entre números reales es también estrictamente creciente.

1.6 a) $\sigma(11) = (2,4)$, $\sigma(13) = (4,4)$, $\sigma(15) = (4,2)$.

b) Si el natural k es un cuadrado perfecto, entonces $\sigma(k) = (\sqrt{k}, 1)$. En otro caso, existen dos naturales a, b tales que $a^2 < k < (a+1)^2$ y $k = a^2 + b$. Entonces

$$\sigma(k) = (b, a+1) \text{ si } b \leq a+1, \quad \sigma(k) = (a+1, 2a-b+2) \text{ si } a+1 \leq b \leq 2a.$$

c) $\sigma^{-1}(4,1) = 16$, $\sigma^{-1}(1,4) = 10$.

d)

$$\sigma^{-1}(n,m) = \begin{cases} (m-1)^2 + n & \text{si } n \leq m \\ (n-1)^2 + n + n - m = n^2 - m + 1 & \text{si } n \geq m \end{cases}.$$

Tema 2

Campos escalares y vectoriales continuos. Límite funcional.

En este tema, por abstracción de las propiedades de la norma y la distancia euclídea, se presentan las nociones de espacio normado y espacio métrico. La topología usual de $\mathbb{R}^N$ es la generada por la norma euclídea, esto es, los abiertos son uniones de bolas abiertas euclídeas. El Teorema de Hausdorff, resultado principal de este tema, afirma que dicha topología coincide con la topología asociada a cualquier norma en $\mathbb{R}^N$. Probamos también las extensiones a $\mathbb{R}^N$ de los Teoremas de complitud y de Bolzano-Weierstrass.

Definimos los compactos de $\mathbb{R}^N$ como los subconjuntos cerrados y acotados. Presentamos dos caracterizaciones de los compactos que son estupendas herramientas en las demostraciones por compacidad (aquellas cuyos enunciados están ligados a la noción de compacto). La primera afirma que toda sucesión en un compacto se acumula (Teorema 2.28) y la segunda que todo compacto verifica el axioma de Heine-Borel (Teorema 2.31). Introducimos también las nociones de convexidad y conexión en $\mathbb{R}^N$ que son las extensiones geométrica y topológica, respectivamente, de la noción de intervalo de $\mathbb{R}$.

Definimos la continuidad de funciones reales de varias variables reales y probamos que tales funciones conservan los compactos y los conexos. El Teorema de Dini da condiciones suficientes para que una sucesión de funciones que, en principio, converge sólo puntualmente converja uniformemente. El Teorema de Heine nos asegura que las funciones continuas definidas en compactos son de hecho uniformemente continuas.

Terminamos la lección estudiando el concepto de límite funcional (indispensable para definir el concepto de función derivable) y la relación que existe entre éste y la continuidad.

En lo sucesivo, para cada natural N, $\mathbb{R}^N$ denota el espacio vectorial real de las N-uplas de números reales, es decir,

$$\mathbb{R}^N := \{x = (x_1, \ldots, x_N) : x_1, \ldots, x_N \in \mathbb{R}\}$$

con las definiciones usuales de suma y producto por escalares

$$x + y := (x_1 + y_1, \ldots, x_N + y_N), \;\; \lambda x := (\lambda x_1, \ldots, \lambda x_N).$$

A las componentes $x_1, \ldots, x_N$ de la N-upla que define el vector x se les denominan <u>*coordenadas*</u> de dicho vector. Cuando haya lugar a confusión, y en especial cuando se esté trabajando con sucesiones en $\mathbb{R}^N$, denotaremos a las coordenadas de un vector $x \in \mathbb{R}^N$ en la forma $x = (x(1), ..., x(N))$. Así, por ejemplo, si $\{x_n\}$ es una sucesión de vectores en $\mathbb{R}^N$ a la coordenada k-ésima del término x_n se le denotará $x_n(k)$.

2.1. Normas y distancias.

Recordemos que la *norma euclídea* en $\mathbb{R}^N$, es decir, la aplicación $\|\cdot\|_2 : \mathbb{R}^N \longrightarrow \mathbb{R}_0^+$ definida por

$$\|x\|_2 : \sqrt{(x|x)} = \sqrt{x_1^2 + \cdots + x_N^2} \quad (x \in \mathbb{R}^N)$$

goza de las siguientes propiedades:

i) $\|x\|_2 = 0 \Leftrightarrow x = 0$.

ii) $\|\lambda x\|_2 = |\lambda|\ \|x\|_2,\ \forall \lambda \in \mathbb{R},\ \forall x \in \mathbb{R}^N$.

iii) $\|x+y\|_2 \leq \|x\|_2 + \|y\|_2,\ \forall x, y \in \mathbb{R}^N$.

A $(\mathbb{R}^N, \|.\|_2)$ se le llama *el espacio euclídeo* (de dimensión N).

Ello nos invita a dar la siguiente definición.

Definición 2.1. Si X es un espacio vectorial real, una *norma* en X es una función $\|\cdot\| : X \longrightarrow \mathbb{R}_0^+$ verificando

i) $\|x\| = 0 \Leftrightarrow x = 0$.

ii) $\|\lambda x\| = |\lambda|\ \|x\|,\ \forall \lambda \in \mathbb{R},\ \forall x \in X$ (homogeneidad).

iii) $\|x+y\| \leq \|x\| + \|y\|,\ \forall x, y \in X$ (desigualdad triangular) .

El par ordenado $(X, \|\cdot\|)$ se llama *espacio normado* .

Observaciones 2.2.

a) En i) basta exigir sólo la condición

$$\|x\| = 0 \Rightarrow x = 0,$$

ya que de ii) se deduce que $\|0\| = 0$.

b) De la definición se sigue que también se puede prescindir en la definición de que la norma toma valores no negativos, ya que

$$0 = \|x - x\| \leq \|x\| + \|x\| = 2\|x\| \Rightarrow \|x\| \geq 0.$$

c) De *ii*) y *iii*) se deduce fácilmente que

$$| \ \|x\| - \|y\| \ | \leq \|x \pm y\|, \ \forall x, y \in X.$$

d) Por último, de *iii*) se deduce fácilmente por inducción que

$$\|x_1 + x_2 + \cdots + x_n\| \leq \|x_1\| + \|x_2\| + \cdots + \|x_n\|, \ \forall x_1, \cdots, x_n \in X.$$

Ejemplos 2.3.

1. $(\mathbb{R}, |\cdot|)$ es un espacio normado. De hecho, en $\mathbb{R}$ todas las normas son producto del valor absoluto por una constante positiva.

2. Algunas normas en $\mathbb{R}^N$.
Además de la norma euclídea, en $\mathbb{R}^N$ consideraremos entre otras la <u>*norma de la suma*</u> y la <u>*norma del máximo*</u>, dadas, respectivamente, por

$$\begin{array}{r} \|x\|_1 := |x_1| + \cdots + |x_N|, \\ \|x\|_\infty := \max\{|x_1|, \cdots, |x_N|\} \end{array} \qquad (x \in \mathbb{R}^N)$$

No es difícil comprobar (hágase como ejercicio) que ambas son normas y que se verifica la desigualdad (véase el Ejemplo 2.13)

$$\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1 \leq N\|x\|_\infty \ \ (x \in \mathbb{R}^N).$$

Es conveniente dibujar la esfera unidad (elementos que tienen norma 1) en dimensión 2 para tener una idea de cómo se comporta la norma. Si consideramos la bola unidad cerrada asociada a una norma, esto es,

$$B := \{x \in \mathbb{R}^2 : \|x\| \leq 1\},$$

ocurre que a bolas menores corresponden mayores normas.

3. En el espacio vectorial $\mathscr{C}[a,b]$ de las funciones reales continuas definidas en el intervalo cerrado y acotado $[a,b]$, se puede definir, por la propiedad de compacidad, la norma dada por

$$\|f\|_\infty = \max \{|f(x)| : x \in [a,b]\} \qquad (f \in \mathscr{C}[a,b]).$$

Compruébese que de hecho es una norma.

Recordemos ahora que la <u>*distancia euclídea*</u> en $\mathbb{R}^N$, es decir, la aplicación

$$d_2 : \mathbb{R}^N \times \mathbb{R}^N \longrightarrow \mathbb{R}_0^+$$

definida por

$$d_2(x,y) = \|x - y\|_2 \qquad (x, y \in \mathbb{R}^N),$$

verifica las siguientes propiedades:

1. $d_2(x,y) = 0 \Leftrightarrow x = y$.
2. $d_2(x,y) = d_2(y,x),\ \forall x,y \in \mathbb{R}^N$.
3. $d_2(x,z) \leq d_2(x,y) + d_2(y,z),\ \forall x,y,z \in \mathbb{R}^N$.

Ello nos invita a dar la siguiente definición.

Definición 2.4. Una *distancia* (o *métrica*) definida en un conjunto no vacío E es una función $d : E \times E \longrightarrow \mathbb{R}_0^+$ que verifica:

1. $d(x,y) = 0 \Leftrightarrow x = y$.
2. $d(x,y) = d(y,x),\ \forall x,y \in E$ (propiedad simétrica).
3. $d(x,z) \leq d(x,y) + d(y,z),\ \forall x,y,z \in E$ (desigualdad triangular).

Al par ordenado (E,d) se le denomina *espacio métrico* .

Observaciones 2.5.

a) Una aplicación $d : E \times E \longrightarrow \mathbb{R}$ que verifique 1), 2) y 3) no toma valores negativos, ya que
$$0 = d(x,x) \leq d(x,y) + d(y,x) = 2d(x,y).$$

b) $|d(x,y) - d(y,z)| \leq d(x,z),\ \forall x,y,z \in E$.
En efecto, usando 2) y 3) se tiene
$$d(x,y) \leq d(x,z) + d(z,y) = d(x,z) + d(y,z) \Rightarrow$$
$$\Rightarrow d(x,y) - d(y,z) \leq d(x,z),$$
Intercambiando x por z y usando 2) se obtiene
$$d(y,z) - d(x,y) \leq d(x,z),$$
por tanto,
$$|d(x,y) - d(y,z)| \leq d(x,z),\ \forall x,y,z \in E.$$

c) De 3) se deduce por inducción la desigualdad
$$d(x_1,x_n) \leq d(x_1,x_2) + \cdots + d(x_{n-1},x_n),\ \forall x_1,\cdots,x_n \in E.$$

Ejemplos 2.6.

1. *Subespacio métrico.*
   Es claro que todo subconjunto A no vacío de un espacio métrico (E,d) también es un espacio métrico, sin más que considerar en A la restricción de la distancia de E. El conjunto A, dotado de esta métrica, es un subespacio métrico de (E,d).

2. *Distancia asociada a una norma.*
   Si $(X,\|\cdot\|)$ es un espacio normado, la *distancia en X asociada a la norma* viene dada por
   $$d(x,y) := \|x-y\| \qquad (x,y \in X).$$
   Compruébese que efectivamente es una distancia. En particular, la distancia usual en un subconjunto A de $\mathbb{R}$ viene dada por
   $$d(x,y) = |x-y|,\ \forall x,y \in A.$$

3. *Espacio métrico producto.*
   Dados n espacios métricos (E_1,d_1),(E_2,d_2),...,(E_n,d_n), podemos definir una distancia en el producto $E_1 \times E_2 \times ... \times E_n$ por
   $$d((x_1,...,x_n),(y_1,...,y_n)) := \max \{d_1(x_1,y_1),...,d_n(x_n,y_n)\}.$$

2.2. Topología de un espacio métrico.

Definición 2.7. Sea (E,d) un espacio métrico. Dados $a \in E$, $r \geq 0$, la *bola abierta* (resp. *cerrada*) de centro a y radio r son los conjuntos dados por
$$B(a,r) := \{x \in E : d(x,a) < r\},$$
$$\overline{B}(a,r) := \{x \in E : d(x,a) \leq r\}.$$
La *esfera* de centro a y radio r es, por definición, el conjunto
$$S(a,r) := \{x \in E : d(x,a) = r\}.$$

Un subconjunto O de un espacio métrico (E,d) se dice *abierto* si verifica la siguiente condición:
$$\forall a \in O,\ \exists r > 0 :\ B(a,r) \subset O\ .$$

Es fácil ver que una bola abierta es un conjunto abierto, que los conjuntos abiertos son aquellos que se pueden expresar como unión de bolas abiertas y, si notamos por $\Im$ a la familia de todos los conjuntos abiertos, se verifica:

i) $\emptyset, E \in \Im$.

ii) $\mathscr{A} \subset \Im \Rightarrow \bigcup_{O \in \mathscr{A}} O \in \Im$.

iii) $O_1, O_2 \in \Im \Rightarrow O_1 \cap O_2 \in \Im$.

Por tanto, $\Im$ es una *topología* en E.

Dado que todo espacio normado $(X, \|.\|)$ es un espacio métrico con la distancia definida por $d(x,y) := \|y - x\|$ $(x, y \in X)$, la *topología de un espacio normado* es la topología asociada a la métrica descrita.

Definición 2.8. Sea (E,d) un espacio métrico y sea $\Im$ la familia de sus conjuntos abiertos. Un subconjunto F de E es *cerrado* si su complementario es abierto, esto es, si $E \backslash F \in \Im$. Si notamos $\mathscr{F}$ a la familia de los conjuntos cerrados, es fácil comprobar que se verifica:

i) $\emptyset, E \in \mathscr{F}$.

ii) $\mathscr{B} \subset \mathscr{F} \Rightarrow \bigcap_{F \in \mathscr{B}} F \in \mathscr{F}$.

iii) $F_1, F_2 \in \mathscr{F} \Rightarrow F_1 \cup F_2 \in \mathscr{F}$.

Es fácil probar que las bolas cerradas y las esferas son conjuntos cerrados.

Sea A un subconjunto de E, un elemento $x \in E$ se dice que es *adherente* a A si para cada positivo r se verifica

$$B(x,r) \cap A \neq \emptyset.$$

Se llama *adherencia* o *cierre* de A al conjunto de todos los valores adherentes de A, que notaremos por $\overline{A}$. Es inmediato que $A \subset \overline{A}$.

Un elemento $x \in E$ se dice que es *interior* de A si se verifica que

$$\exists r > 0 : B(x,r) \subset A.$$

Notaremos por $\overset{\circ}{A}$ al conjunto de todos los puntos interiores de A, conjunto que claramente verifica $\overset{\circ}{A} \subset A$.

Por último, llamaremos *frontera* de A al conjunto $Fr(A) := \overline{A} \backslash \overset{\circ}{A}$.

Ejemplo 2.9. Pruébese que si $A \subset \mathbb{R}$ está mayorado (resp. minorado) y no es vacío entonces $\sup A \in \overline{A}$ (resp. $\inf A \in \overline{A}$). Hallar $\overline{\mathbb{Q}}, \overline{\mathbb{R} \setminus \mathbb{Q}}, \overline{\{\frac{1}{n} : n \in \mathbb{N}\}}$.

El siguiente resultado, cuya demostración se deja como ejercicio (¡también!), resume las primeras propiedades que relacionan los conceptos anteriores.

Proposición 2.10. *Sea (E,d) un espacio métrico y notemos por $\Im$ a la familia de sus abiertos y por $\mathscr{F}$ a la de sus cerrados. Cada subconjunto A de E verifica:*

i) $(O \in \Im, O \subset A) \Rightarrow O \subset \overset{\circ}{A}$.

ii) $\overset{\circ}{A} \in \Im$.

iii) $\overset{\circ}{A}$ *es el mayor abierto incluido en A, en consecuencia,* $\overset{\circ}{A}$ *es la unión de todos los abiertos incluidos en A.*

iv) $A \in \Im \Leftrightarrow A = \overset{\circ}{A}$.

v) $(F \in \mathscr{F}, A \subset F) \Rightarrow \overline{A} \subset F$.

vi) $\overline{A} \in \mathscr{F}$.

vii) $\overline{A}$ *es el menor cerrado que contiene a A y coincide con la intersección de todos los cerrados que contienen a A.*

viii) $A \in \mathscr{F} \Leftrightarrow A = \overline{A}$.

ix) $(E \overset{\circ}{\setminus} A) = E \setminus \overline{A}$, *equivalentemente* $\overset{\circ}{A} = E \setminus \overline{E \setminus A}$.

x) $E \setminus \overset{\circ}{A} = \overline{E \setminus A}$, *equivalentemente* $\overline{A} = E \setminus (E \overset{\circ}{\setminus} A)$.

De la igualdad $E \setminus \overset{\circ}{A} = \overline{E \setminus A}$, se deduce que $Fr(A) = \overline{A} \cap \overline{(E \setminus A)}$.
A continuación extendemos el concepto de convergencia en $\mathbb{R}$ a espacios métricos.

Definición 2.11 (Convergencia en espacios métricos)**.** Se dice que una sucesión $\{x_n\}$ de elementos de un espacio métrico (E,d) es <u>*convergente*</u> si existe un elemento $x \in E$ tal que

$$\forall \varepsilon > 0,\ \exists m \in \mathbb{N} : n \geq m \Rightarrow d(x_n, x) < \varepsilon,$$

equivalentemente, si $\{d(x_n, x)\} \to 0$. Si se verifica la condición anterior, diremos que $\{x_n\}$ converge a x y en tal caso escribiremos $\{x_n\} \to x$.
Es fácil comprobar (ejercicio) que el elemento x que verifica la condición de convergencia es único y se llama <u>*límite de la sucesión* $\{x_n\}$</u>, y entonces escribiremos $x = \lim\ x_n$.

El concepto de sucesión de Cauchy en un espacio métrico es también copia literal del dado para $\mathbb{R}$.

Definición 2.12 (sucesión de Cauchy)**.** Una sucesión $\{x_n\}$ de elementos de un espacio métrico (E,d) es de <u>*Cauchy*</u> si se verifica

$$\forall \varepsilon > 0,\ \exists m \in \mathbb{N} : p, q \geq m \Rightarrow d(x_p, x_q) \leq \varepsilon,$$

equivalentemente,

$$\forall \varepsilon > 0,\ \exists m \in \mathbb{N} :\ [n \geq m,\ h \in \mathbb{N}] \Rightarrow d(x_{n+h}, x_n) \leq \varepsilon.$$

Es inmediato comprobar que en un espacio métrico toda sucesión convergente es de Cauchy. $(\mathbb{Q}, |.|)$ es un ejemplo de que el recíproco de esta afirmación no es cierto. Aquellos espacios métricos que verifican que toda sucesión de Cauchy es convergente se llaman *completos*. Un espacio normado y completo para la métrica asociada a la norma es un *espacio de Banach*. El espacio normado $(\mathscr{C}[0,2], \|.\|_1)$, es decir, el espacio vectorial de las funciones reales continuas definidas en $[0,2]$ con la norma integral dada por

$$\|f\|_1 := \int_0^2 |f(x)|dx,$$

no es completo[1] (véase Ejercicio 2.3).

Un subconjunto A de un espacio métrico (E,d) se dice completo si el espacio métrico (A,d) es completo, es decir, si toda sucesión en A que sea de Cauchy converge a un elemento de A.

Ejemplo 2.13 (Convergencia en $(\mathbb{R}^N, \|\cdot\|_2)$**.** Si $\{x_n\}$ es una sucesión en $\mathbb{R}^N$ denotaremos por $x_n(k)$ a la coordenada k-ésima del término x_n. Probaremos que la convergencia en $\mathbb{R}^N$ se reduce a la convergencia coordenada a coordenada, esto es:

$$\{x_n\} \overset{\|\cdot\|_2}{\to} x \Leftrightarrow \{x_n(k)\} \to x(k),\ \forall k = 1,2,\cdots,N$$

donde $x = (x(1),...,x(N))$.

Demostración:

Probemos primeramente que para todo vector $x \in \mathbb{R}^N$ se verifica

$$\|x\|_\infty \leq \|x\|_2 \leq N\|x\|_\infty \tag{1}$$

Denotando por $x(k)$ a las componentes del vector x, se prueba fácilmente la primera desigualdad, ya que

$$\|x\|_\infty^2 := \max\ \{|x(k)| : k = 1,...,N\}^2 =$$

$$\max\ \{x(k)^2 : k = 1,...,N\} \leq \sum_{k=1}^N x(k)^2 = \|x\|_2^2.$$

Y por otro lado

$$\|x\|_2 := \sqrt{\sum_{k=1}^N x(k)^2} \leq \sqrt{\|x\|_\infty^2 + \overset{N}{\cdots} + \|x\|_\infty^2} = \sqrt{N}\|x\|_\infty \leq N\|x\|_\infty.$$

De la primera desigualdad de (1) se sigue que si una sucesión $\{x_n\}$ en $\mathbb{R}^N$ tiene límite x, entonces $\{x_n(k)\} \to x(k)$ para todo $k = 1,...,N$. Supongamos ahora que

$$\{x_n(k)\} \to x(k),\ \forall k = 1,...,N.$$

[1] La no complitud del espacio $(\mathscr{C}[0,2], \|.\|_1)$ no es consecuencia de la norma elegida, sino de que es necesario ampliar sensiblemente el conjunto de funciones integrables para conseguir la complitud y que en consecuencia las cosas marchen bien. Algo análogo ocurre con $\mathbb{Q}$ y su "completación" a $\mathbb{R}$

Entonces, dado $\varepsilon > 0$, para cada $k = 1, \cdots, N$ existe un natural m_k tal que si $n \geq m_k$ entonces $|x_n(k) - x(k)| < \frac{\varepsilon}{N}$. Luego, si tomamos $m := \max\ \{m_1, \cdots, m_N\}$, se obtiene para $n \geq m$ que

$$\|x_n - x\|_2 \leq N\|x_n - x\|_\infty = N\ \max\ \{|x_n(1) - x(1)|, ..., |x_n(N) - x(N)|\} < N\frac{\varepsilon}{N} = \varepsilon.$$

■

Es muy fácil probar que en $(\mathbb{R}^N, \|\cdot\|_2)$ se verifica

$$\{x_n\} \text{ es de Cauchy } \Leftrightarrow \{x_n(k)\} \text{ es de Cauchy } \forall k = 1, 2, \cdots, N.$$

Como consecuencia del Teorema de complitud de $\mathbb{R}$, se tiene que $(\mathbb{R}^N, \|\cdot\|_2)$ es un espacio de Banach.

Estúdiese la convergencia en $(\mathbb{R}^N, \|.\|_1)$ y en $(\mathbb{R}^N, \|.\|_\infty)$.

Ejemplo 2.14 (Convergencia en $(\mathscr{C}[a,b], \|\cdot\|_\infty)$)**.** De la definición de la norma $\|.\|_\infty$ se sigue que

$$\{f_n\} \overset{\|\cdot\|_\infty}{\to} f \Leftrightarrow (\forall \varepsilon > 0, \exists m \in \mathbb{N} : n \geq m \Rightarrow |f_n(x) - f(x)| \leq \varepsilon,\ \forall x \in [a,b]).$$

La convergencia en $(\mathscr{C}[a,b], \|.\|_\infty)$ es la *convergencia uniforme* , que implica la convergencia puntual pero el recíproco no es cierto (véanse los Ejercicios 2.2 y 2.4).

Proposición 2.15 (Caract. secuencial de la adherencia en esp. métricos)**.** *Sea A un subconjunto de un espacio métrico E y $x \in E$. Equivalen:*

i) x es un punto adherente a A.

ii) Existe una sucesión en A que converge a x.

En consecuencia, como $A \in \mathscr{F} \Leftrightarrow A \supset \overline{A}$, un subconjunto A de un espacio métrico es cerrado si, y sólo si, A contiene los límites de todas las sucesiones en A convergentes.

Demostración:

$i) \Rightarrow ii)$ Supongamos que x es un punto adherente a A, por tanto

$$B\left(x, \frac{1}{n}\right) \cap A \neq \emptyset,\ \forall n \in \mathbb{N}.$$

En consecuencia, para cada natural n, podemos elegir un elemento $a_n \in A$ que verifique $d(a_n, x) < \frac{1}{n}$. Es claro que la sucesión $\{a_n\}$, así construida, cuyos términos están en A, converge a x.

$ii) \Rightarrow i)$ Supongamos que $\{a_n\}$ es una sucesión en A convergente a x. Por tanto, para cada $\varepsilon > 0$, existe un natural m tal que

$$n \geq m \Rightarrow a_n \in B(x, \varepsilon),$$

y en consecuencia, $B(x,\varepsilon)\cap A\neq\emptyset$, para todo $\varepsilon>0$. ■

El siguiente resultado es útil para justificar que ciertos espacios métricos son completos. Su demostración hace uso de la anterior caracterización secuencial de la adherencia.

Proposición 2.16.

i) Todo subconjunto completo de un espacio métrico es cerrado.

ii) Todo subconjunto cerrado de un espacio métrico completo es completo.

Demostración:

$i)$ Supongamos que A es un subconjunto completo de un espacio métrico E. Sea $x\in\overline{A}$, por tanto, en vista de la Proposición 2.15, existe una sucesión $\{a_n\}$ de elementos de A convergente a x. La sucesión $\{a_n\}$ es de Cauchy y, por ser A completo, ha de converger a un elemento de A, por tanto, $x\in A$. Hemos probado que $\overline{A}\subset A$, y, por tanto, A es cerrado.

$ii)$ Sea A un subconjunto cerrado de un espacio métrico completo E. Fijamos una sucesión $\{a_n\}$ de Cauchy en A. Por ser E completo, existe un elemento $x\in E$ que es el límite de la sucesión $\{a_n\}$, por tanto, usando de nuevo la Proposición 2.15, x es adherente a A, y por ser A cerrado, concluimos que $x\in A$, luego A es completo. ■

Observación 2.17. Nótese que, como consecuencia de la caracterización secuencial de la adherencia (Proposición 2.15), en un espacio métrico es suficiente conocer las sucesiones convergentes y sus límites para conocer los conjuntos cerrados, y, por tanto, la topología.

Nota 2.18. El concepto de convergencia de una sucesión es topológico, esto es, depende de la topología del espacio métrico, pero no de la distancia concreta que se utilice. Esto significa simplemente que si dos distancias d y d^* generan la misma topología, entonces las sucesiones convergentes coinciden para ambas distancias. En efecto, supongamos que $\{x_n\}$ converge a x en la distancia d. Dado $\varepsilon>0$, puesto que $B_{d^*}(x,\varepsilon)$ es un abierto que contiene a x y d genera la misma topología que d^*, ha de existir $r>0$ tal que $B_d(x,r)\subset B_{d^*}(x,\varepsilon)$. Como estamos suponiendo que $\{x_n\}$ converge en la distancia d, ha de existir un natural m verificando que

$$n\geq m\Rightarrow d(x_n,x)<r.$$

En vista de la elección de r se tiene también que $d^*(x_n,x)<\varepsilon$ para $n\geq m$, y, por tanto $\{x_n\}$ converge también a x en la distancia d^*.

Probaremos que "*dos normas cualesquiera en* $\mathbb{R}^N$ *generan la misma topología*". Para preparar la prueba de este importantísimo resultado introducimos el siguiente concepto.

Definición 2.19. Dos normas $\|\cdot\|$ y $|||\cdot|||$ en un mismo espacio vectorial X se dicen <u>*equivalentes*</u> si existen constantes $m, M > 0$ verificando

$$m\|x\| \leq |||x||| \leq M\|x\|, \ \forall x \in X.$$

Es inmediato probar que la relación binaria que hemos definido entre normas es de equivalencia.

Proposición 2.20. *Sean $\|\cdot\|$ y $|||\cdot|||$ dos normas en el espacio vectorial X. Equivalen las siguientes condiciones:*

i) Las normas $\|\cdot\|$ y $|||\cdot|||$ son equivalentes.

ii) Ambas normas generan la misma topología.

Demostración:

i) $\Rightarrow$ *ii)* Sea O un abierto para la topología asociada a $\|\cdot\|$. Dado $a \in O$, existe $r > 0$ tal que

$$B_{\|\cdot\|}(a, r) \subset O,$$

pero, por ser, ambas normas equivalentes, existe una constante $m > 0$ tal que

$$m\|x\| \leq |||x|||, \ \forall x \in X.$$

Por tanto, se tiene

$$B_{|||\cdot|||}(a, mr) \subset B_{\|\cdot\|}(a, r) \subset O,$$

y O es abierto para la topología asociada a la norma $|||.|||$. La inclusión contraria es consecuencia de la otra desigualdad ente las normas.

ii) $\Rightarrow$ *i)* Por ser las bolas abiertas conjuntos abiertos, existe una constante $s > 0$ tal que

$$B_{|||\cdot|||}(0, s) \subset B_{\|\cdot\|}(0, 1).$$

Sea $x \in X$ un elemento no nulo, entonces, es claro que se verifica

$$\left|\left|\left| \frac{sx}{2\,|||x|||} \right|\right|\right| < s \Rightarrow \left\| \frac{sx}{2\,|||x|||} \right\| < 1 \Rightarrow \frac{s}{2}\|x\| \leq |||x|||$$

En vista de la hipótesis, ambas normas están en las mismas condiciones, luego también se puede probar que existe una constante M tal que $|||x||| \leq M\|x\|, \ \forall x \in X.$ ∎

Para probar el Teorema de Hausdorff, en primer lugar, generalizaremos al espacio euclídeo el Teorema de Bolzano-Weierstrass.

Definición 2.21 (conjunto acotado)**.** Un subconjunto A de un espacio métrico (E, d) se dice <u>*acotado*</u> si existen $M > 0$ y $x_0 \in E$ tales que $A \subset B(x_0, M)$. Así, un subconjunto A de un espacio normado $(X, \|\cdot\|)$, es acotado si existe $M > 0$ tal que

$$\|a\| < M, \ \forall a \in A \ .$$

Pruébese que toda sucesión convergente es acotada. De hecho, toda sucesión de Cauchy en un espacio métrico es también acotada.

Teorema 2.22 (Bolzano-Weierstrass en $(\mathbb{R}^N, \|\cdot\|)_2)$)**.** $(\mathbb{R}^N, \|\cdot\|_2)$ *verifica la propiedad de Bolzano-Weierstrass, es decir, toda sucesión acotada en* $(\mathbb{R}^N, \|\cdot\|_2)$ *admite una parcial convergente.*

Demostración:

Haremos la prueba por inducción sobre la dimensión del espacio. Para $N = 1$, se trata del Teorema de Bolzano-Weierstrass, que es conocido en $\mathbb{R}$. Supongamos que se verifica para $\mathbb{R}^N$. En $\mathbb{R}^{N+1}$ se tiene que

$$\|(x,y)\|_2 = \sqrt{x(1)^2 + \cdots + x(N)^2 + y^2},\ \forall (x,y) \in \mathbb{R}^N \times \mathbb{R}. \tag{$*$}$$

Fijamos una sucesión acotada en $\mathbb{R}^{N+1}$, que podemos suponer de la forma $\{(x_n, y_n)\}$, donde $x_n \in \mathbb{R}^N$, $y_n \in \mathbb{R}$, para cada natural n. En vista de $(*)$, las sucesiones $\{x_n\}$ e $\{y_n\}$ son acotadas. Por hipótesis de inducción, la primera admite una parcial convergente, que escribiremos $\{x_{\sigma(n)}\}$, ahora bien, por ser $\{y_{\sigma(n)}\}$ acotada (parcial de una acotada), el Teorema de Bolzano-Weierstrass nos asegura que admite una parcial convergente que escribiremos $\{y_{\sigma(\tau(n))}\}$[2]. Finalmente, la sucesión en $\mathbb{R}^{N+1}$ dada por $\{(x_{\sigma(\tau(n))}, y_{\sigma(\tau(n))})\}$ es una parcial convergente de $\{(x_n, y_n)\}$. ∎

Teorema 2.23 (Hausdorff)**.** *Todas las normas en* $\mathbb{R}^N$ *son equivalentes.*

Demostración:

Probaremos que si $\|\cdot\|$ es una norma cualquiera en $\mathbb{R}^N$, entonces equivale a la norma euclídea. Notamos por $\{e_1, e_2, \cdots, e_N\}$ a la base canónica de $\mathbb{R}^N$. Dado cualquier vector $x \in \mathbb{R}^N$ que se escriba de la forma $x = x(1)e_1 + \cdots + x(N)e_N$, se tiene

$$\begin{aligned}\|x\| = \|x(1)e_1 + \cdots + x(N)e_N\| \leq &\quad \text{(por la desigualdad triangular)}\\ &\leq |x(1)|\ \|e_1\| + \cdots + |x(N)|\ \|e_N\| \leq\\ &\leq (\|e_1\| + \cdots + \|e_N\|)\|x\|_\infty\\ &\leq (\|e_1\| + \cdots + \|e_N\|)\|x\|_2,\end{aligned}$$

luego, tomando $M = \|e_1\| + \cdots + \|e_N\|$ se tiene que

$$\|x\| \leq M\|x\|_2,\ \forall x \in \mathbb{R}^N.$$

Ahora definimos

$$m := \inf\{\|x\| : \|x\|_2 = 1\}.$$

[2]Obsérvese que si notamos $\{y_n\} = \{x_{\sigma(n)}\}$, entonces $\{y_{\tau(n)}\} = \{x_{\sigma(\tau(n))}\}$

Probaremos que m es un mínimo[3]. Sabemos que por ser m el ínfimo del conjunto anterior, existe una sucesión $\{x_n\}$ de elementos de $\mathbb{R}^N$ verificando

$$\|x_n\|_2 = 1, \qquad \{\|x_n\|\} \to m$$

(véanse el Ejemplo 2.9 y la Proposición 2.15).

Como $(\mathbb{R}^N, \|\cdot\|_2)$ verifica la propiedad de Bolzano-Weierstrass (Teorema 2.22), ha de existir un elemento $x_0 \in \mathbb{R}^N$ y una sucesión parcial $\{x_{\sigma(n)}\}$ de $\{x_n\}$ tal que

$$\{x_{\sigma(n)}\} \overset{\|\cdot\|_2}{\to} x_0,$$

con lo que se tiene

$$\|x_0\|_2 = \lim \{\|x_{\sigma(n)}\|_2\} = 1,$$

en particular $x_0 \neq 0$. Por la desigualdad ya probada entre las normas se tiene

$$|\ \|x_{\sigma(n)}\| - \|x_0\|\ | \leq \|x_{\sigma(n)} - x_0\| \leq M\|x_{\sigma(n)} - x_0\|_2, \ \forall n \in \mathbb{N},$$

y en virtud de la convergencia en la norma euclídea de $\{x_{\sigma(n)}\}$ a x_0, concluimos que $\{\|x_{\sigma(n)}\|\}$ converge a $\|x_0\|$, por tanto $m = \|x_0\| > 0$.

Queremos probar ahora que

$$m\|x\|_2 \leq \|x\|, \ \ \forall x \in \mathbb{R}^N,$$

desigualdad que es cierta si $x = 0$. Si $x \neq 0$, se tiene

$$\left\| \frac{x}{\|x\|_2} \right\|_2 = 1 \Rightarrow m \leq \left\| \frac{x}{\|x\|_2} \right\| \Rightarrow m\|x\|_2 \leq \|x\|. \qquad \blacksquare$$

Dado que "*no existe más espacio vectorial real de dimensión N que* $\mathbb{R}^N$", podemos decir que *en "cualquier" espacio vectorial real finito-dimensional todas las normas son equivalentes* (véase problema 2.8). En realidad, tal propiedad caracteriza la finito-dimensionalidad de un espacio vectorial.

Como corolario del teorema anterior y de la Proposición 2.20 se obtiene:

Corolario 2.24. *Existe una única topología en* $\mathbb{R}^N$ *que proceda de una norma a la que llamaremos la* <u>*topología de la norma*</u> .

En todo lo que sigue, se supondrá que $\mathbb{R}^N$ está dotado de la topología de la norma, cuyos abiertos no son más que uniones de bolas abiertas para alguna norma. En el caso de $\mathbb{R}$, los abiertos son uniones de intervalos abiertos.

La segunda consecuencia del Teorema de Hausdorff es que el concepto de sucesión de Cauchy en $\mathbb{R}^N$ es independiente de la norma. Este hecho, junto con el Teorema de complitud de $\mathbb{R}$ y el Ejemplo 2.13 nos prueban el siguiente resultado.

---

[3] De haber pospuesto la demostración de este teorema a la obtención de la propiedad de compacidad (Proposición 2.51), esto habría sido inmediato. En efecto, dado que la función norma $\|\cdot\|$ es continua en $(\mathbb{R}^N, \|\cdot\|_2)$ (nótese que $|\ \|x\| - \|y\|\ | \leq \|x - y\| \leq M\|x - y\|_2$) y que la esfera unidad para la norma euclídea $S_{\|\cdot\|_2}(0,1)$ es compacta, por la propiedad de compacidad, el conjunto $\{\|x\| : \|x\|_2 = 1\}$ tiene mínimo.

Teorema 2.25 (complitud)**.** *En $\mathbb{R}^N$, toda sucesión de Cauchy es convergente, esto es, $\mathbb{R}^N$ es un espacio de Banach con cualquier norma.*

La tercera consecuencia del Teorema de Hausdorff es que el concepto de acotación en $\mathbb{R}^N$ es independiente de la norma.

El Teorema 2.22 admite ahora el siguiente enunciado.

Teorema 2.26 (Bolzano-Weierstrass)**.** $\mathbb{R}^N$ *verifica la propiedad de Bolzano-Weierstrass, es decir, toda sucesión acotada admite una parcial convergente.*

2.3. Compactos, convexos y conexos.

Dedicamos esta sección a presentar tres tipos distinguidos de subconjuntos de $\mathbb{R}^N$: los compactos, los convexos y los conexos.

En la demostración del Teorema de Hausdorff sólo se ha tenido en cuenta que $S_{\|.\|_2}(0,1)$ es un conjunto acotado y cerrado. [4] Ello nos motiva a destacar estos subconjuntos de $\mathbb{R}^N$.

Definición 2.27. Un subconjunto K de $\mathbb{R}^N$ es <u>*compacto*</u> si es cerrado y acotado.

Decir cuáles de los siguientes conjuntos son compactos: $\mathbb{N}$, $\{\frac{1}{n} : n \in \mathbb{N}\}$, $\{a\}$, $B(a,r)$, $\overline{B}(a,r)$, $S(a,r)$, $[a,b] \cup [c,d]$, $\{(x,y) \in \mathbb{R}^2 : x > 0\}$.

Obsérvese que el punto crucial de la demostración del tan citado Teorema de Hausdorff consiste en probar que toda sucesión en la esfera $S_{\|.\|_2}(0,1)$ se acumula, es decir, admite una parcial convergente a un punto de dicha esfera. Esto nos motiva la siguiente caracterización de los compactos de $\mathbb{R}^N$ que será una herramienta básica en futuras demostraciones y que a su vez nos permite extender dicho concepto a espacios métricos cualesquiera.

Teorema 2.28 (Caracterización de los compactos de $\mathbb{R}^N$)**.** *Sea K un subconjunto de $\mathbb{R}^N$. Equivalen:*

i) K es compacto (cerrado y acotado de $\mathbb{R}^N$).

ii) Toda sucesión de puntos de K admite una sucesión parcial que converge a un punto de K.

Demostración:

$i) \Rightarrow ii)$ Sea $\{x_n\}$ una sucesión en K. Por hipótesis, $\{x_n\}$ es acotada y, por el Teorema de Bolzano-Weierstrass, tiene una subsucesión $\{x_{\sigma(n)}\}$ convergente a un vector x que necesariamente ha de pertenecer a K, por ser K un conjunto cerrado.

$ii) \Rightarrow i)$ K es cerrado: Sea $x \in \overline{K}$ y $\{x_n\}$ una sucesión en K convergente a x. Por hipótesis, existe $\{x_{\sigma(n)}\} \to y \in K$. Puesto que también $\{x_{\sigma(n)}\} \to x$, deducimos de la unicidad del límite que $x \in K$. Hemos probado que $\overline{K} \subset K$ y por tanto K es cerrado.

K es acotado: Supongamos que K no es acotado. Se tiene entonces que $K \backslash B(0,n) \neq \emptyset$, $\forall n \in \mathbb{N}$, luego podemos elegir para cada natural n, un elemento $x_n \in K \backslash B(0,n)$. La sucesión $\{x_n\}$ así elegida no puede tener ninguna parcial convergente ya que $\|x_n\| \geq n$, $\forall n \in \mathbb{N}$, y, por tanto, todas sus parciales son no acotadas. Hemos probado que si K es no acotado, entonces no se verifica $ii)$. ∎

[4]Del Teorema de Bolzano-Weierstrass se sigue que toda sucesión en un conjunto acotado admite una parcial convergente; si además el conjunto es cerrado, la caracterización secuencial de la adherencia (Proposición 2.15) nos asegura que el límite se queda en el conjunto.

Definición 2.29. Un subconjunto K de un espacio métrico es *compacto* si toda sucesión de puntos de K admite una sucesión parcial que converge a un punto de K.

Recordamos que en un espacio métrico (E,d) todo subconjunto $A \subset E$ es también un espacio métrico con la distancia inducida. Como todo espacio métrico tiene una topología que procede de la distancia, a la topología asociada a la restricción de d al subconjunto A se le llama *topología inducida* en A. En cualquier espacio métrico, los abiertos son uniones de bolas abiertas. Ahora bien, una bola abierta en A no es más que un conjunto del tipo $\{x \in A : d(x,a) < r\}$, donde $a \in A$ y $r \geq 0$, pero este conjunto no es otra cosa que $B(a,r) \cap A$. Por tanto, si G es un abierto de A, existe un abierto O de E tal que $G = O \cap A$, esto es, los *abiertos de A* o *abiertos relativos en A* no son otra cosa que las intersecciones con A de los abiertos de E. Recíprocamente, es muy fácil probar que los conjuntos que se escriben como intersección de un abierto de E con A son, de hecho, abiertos de A.

Los cerrados de A son los complementos en A de los abiertos de A. Por razones similares, también los cerrados de A se obtienen intersecando A con los cerrados de E.

Por ejemplo, $]\frac{1}{2},1]$ es abierto de $]0,1]$ y no es abierto. ¿Cuales son los abiertos y los cerrados de $\mathbb{Z}$?

Nota 2.30. Es interesante observar que en espacios métricos coinciden los subconjuntos compactos (anterior definición) y los subespacios compactos (con la topología inducida).

Es fácil probar que todo subconjunto compacto de un espacio métrico es cerrado y acotado[5], pero el recíproco no es cierto (véase el ejercicio 2.5).

El siguiente teorema caracteriza la compacidad en los espacios métricos en términos de su topología y permite definir dicho concepto en espacios topológicos generales.

Teorema 2.31 (Heine-Borel-Lebesgue)**.** *Sea K un subconjunto de un espacio métrico (E,d). Equivalen:*

i) K es compacto.

ii) K verifica el axioma de Heine-Borel: todo recubrimiento por abiertos de K admite un subrecubrimiento finito, esto es, si $\mathscr{U}$ es una familia de abiertos de E tales que $K \subset \bigcup_{U \in \mathscr{U}} U$, entonces existen $n \in \mathbb{N}$ y $U_1, \cdots, U_n \in \mathscr{U}$ tal que $K \subset \bigcup_{i=1}^{n} U_i$.

La demostración puede verse en el apéndice (véase también el ejercicio 2.27).

La propiedad arquimediana de $\mathbb{R}$ nos asegura que $\{]\frac{1}{n},1[: n \in \mathbb{N}\}$ (respectivamente $\{]-n,n[: n \in \mathbb{N}\}$) es un recubrimiento por abiertos del conjunto $]0,1[$ (resp. $\mathbb{R}$). Pruébese que en ninguno de los casos anteriores se puede extraer un subrecubrimiento finito de los recubrimientos ¿Contradice este hecho el teorema anterior?

[5] La verificación de este hecho es una simple adaptación de $ii) \Rightarrow i)$ del Teorema 2.28, entendiendo como $i)$ ser cerrado y acotado en un espacio métrico

El axioma de Heine-Borel (afirmación *ii*) del anterior teorema) se toma como definición de compacidad en un espacio topológico cualquiera y es una valiosa herramienta en muchas demostraciones (véase por ejemplo el Teorema de Dini, Teorema 2.52).

Nota 2.32. Obsérvese que en al axioma de Heine-Borel se pueden sustituir simultáneamente los abiertos por abiertos relativos y la inclusión por igualdad, es decir, la compacidad sólo depende de la topología del conjunto en cuestión. Con más precisión, a nivel de espacios topológicos también coinciden los subconjuntos compactos y los subespacios compactos.

Finalizamos esta sección presentando dos extensiones de la noción de intervalo de $\mathbb{R}$ a subconjuntos de $\mathbb{R}^N$, una de naturaleza geométrica y otra de naturaleza topológica.

Recordemos la siguiente caracterización geométrica de los intervalos de $\mathbb{R}$, un subconjunto I de $\mathbb{R}$ es un intervalo si, y sólo si, para cualesquiera $x, y \in I$ con $x \leq y$, se tiene que $[x, y] \subset I$ (es de resaltar que para probar esta caracterización se requiere el axioma del supremo). Como obviamente se tiene

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

entonces I es un intervalo si, y sólo si, para cualesquiera $x, y \in I$ y cualquier $t \in [0,1]$, se verifica $x + t(y-x) \in I$. En efecto, basta considerar la desigualdad

$$\min\{x,y\} \leq (1-t)x + ty \leq \max\{x,y\},\ \forall t \in [0,1].$$

La propiedad anterior que, como acabamos de recordar, caracteriza a los intervalos tiene perfecto sentido en $\mathbb{R}^N$ y, por tanto, nos invita a generalizar el concepto de intervalo de la siguiente forma:

Definición 2.33. Un subconjunto A de $\mathbb{R}^N$ es *convexo* si se verifica

$$a, b \in A \Rightarrow a + t(b-a) \in A,\ \forall t \in [0,1],$$

es decir, para cualesquiera dos elementos a, b en A, el segmento de extremos a y b está contenido en A.

Obviamente el anterior concepto tiene sentido en cualquier espacio vectorial. Es inmediato comprobar que en $\mathbb{R}^N$ las bolas abiertas son conjuntos convexos, e igual ocurre con las bolas cerradas. Como caso particular, por supuesto, se obtiene que los intervalos son conjuntos convexos. Al fin y al cabo, intentamos abstraer una propiedad que tienen (y que de hecho caracteriza a) los intervalos.

Para presentar la otra generalización de intervalo, la conexión, nos inspiraremos en esta otra caracterización topológica de los intervalos:

Proposición 2.34. *Sea $C \subset \mathbb{R}$. Equivalen:*

i) C es un intervalo.

ii) No existen particiones no triviales de C en abiertos relativos, esto es, si O_1, O_2 son abiertos en C tales que $O_1 \cap O_2 = \emptyset$ y $C = O_1 \cup O_2$, entonces $O_1 = \emptyset$ o bien $O_2 = \emptyset$.

Demostración:

$i) \Rightarrow ii)$ Sea C un intervalo y, razonando por reducción al absurdo, supongamos que C es la unión disjunta de dos abiertos O_1 y O_2 de C no vacíos. Sean entonces $a \in O_1, b \in O_2$. Podemos suponer sin perder de generalidad que $a < b$. Como C es un intervalo, se tiene que $[a,b] \subset C$. Definamos

$$c := \sup([a,b] \cap O_1).$$

Como $[a,b]$ es cerrado, es claro que $c \in [a,b] \subset C$. Si $c \in O_1$, entonces $c < b$, y al ser $[a,b]$ un intervalo y O_1 un abierto de C, se tiene que existe $\delta > 0$ tal que

$$\left.\begin{array}{r} [c, c+\delta[\subset [a,b] \\ \\ ]c-\delta, c+\delta[\cap C \subset O_1 \end{array}\right\} \Rightarrow \ [c, c+\delta[\ \subset [a,b] \cap O_1,$$

lo que contradice la definición de c.

Un razonamiento análogo (hágase) muestra que $c \notin O_2$, lo que lleva a contradicción. Hemos probado que un intervalo no admite particiones no triviales en abiertos relativos.

$ii) \Rightarrow i)$ Si C no fuese un intervalo, entonces existirían números reales $x, y \in C$ y un real $z \notin C$ tales que $x < z < y$. Entonces, la partición

$$C = (]-\infty, z[\cap C) \cup (]z, +\infty[\cap C)$$

contradice ii). ∎

Definición 2.35. Un conjunto C de un espacio métrico es <u>*conexo*</u> si verifica que la única partición de C en dos abiertos relativos es la trivial.

En la siguiente sección probaremos que, en los espacios normados, todo conjunto convexo es conexo, en particular, los segmentos y las bolas son conexos. Claramente un conjunto formado por dos elementos no es conexo.

Por supuesto, la definición de conexión puede darse en espacios topológicos.

2.4. Funciones continuas.

Recordemos que "la" topología de $\mathbb{R}^N$ es la topología de la norma.

Definición 2.36 (campo escalar y vectorial)**.** Sean $M,N \in \mathbb{N}$ y $A \subset \mathbb{R}^N$. Un *campo escalar* en A es una función de A en $\mathbb{R}$. Una función $f = (f_1,\dots,f_M) : A \to \mathbb{R}^M$ se llama *campo vectorial* en A, y a los campos escalares f_i, para $i = 1,\dots,M$, se les denomina *campos escalares* de f o funciones componentes de f.

Definición 2.37 (campo vectorial continuo)**.** Sean $M,N \in \mathbb{N}$, $A \subset \mathbb{R}^N$, $a \in A$ y $f : A \to \mathbb{R}^M$ un campo vectorial. Se dice que f es *continuo* en el punto a si para toda sucesión $\{a_n\}$ en A convergente a a, se tiene que la sucesión $\{f(a_n)\}$ converge a $f(a)$, es decir:

$$\big[\forall \{a_n\} \text{ en } A, \{a_n\} \to a\big] \Rightarrow \{f(a_n)\} \to f(a)$$

donde la convergencia de las sucesiones es relativa a las topologías de $\mathbb{R}^N$ y $\mathbb{R}^M$.

Se dice que la función f es continua en $B \subset A$ si lo es en todos los puntos de B. Se dice que la función f es continua si es continua en A.

Como consecuencia inmediata del Ejemplo 2.13, el estudio de la continuidad de los campos vectoriales se reduce al de sus campos escalares componentes como se recoge en el siguiente enunciado.

Proposición 2.38 (Reducción a campos escalares)**.** *Sean $M,N \in \mathbb{N}$, $A \subset \mathbb{R}^N$, $f = (f_1,\dots,f_M) : A \to \mathbb{R}^M$ un campo vectorial en A y a un punto de A. Entonces*

$$f \textit{ es continuo en } a \Leftrightarrow f_i \textit{ es continuo en } a, \ \forall i = 1,\dots,M.$$

El concepto de continuidad se extiende literalmente a funciones definidas entre espacios métricos:

Definición 2.39. Sean $(E,d),(F,\rho)$ espacios métricos, $A \subset E$, $f : A \to F$ y $a \in A$. Se dice que la función f es *continua* en el punto a si

$$\big[\forall \{a_n\} \text{ en } A, \{a_n\} \xrightarrow{d} a\big] \Rightarrow \{f(a_n)\} \xrightarrow{\rho} f(a)$$

Se dice que la función f es continua en $B \subset A$ si lo es en cada punto de B. Se dice que f es continua si es continua en A.

Como la convergencia de una sucesión es un concepto topológico, el concepto de continuidad también es topológico, es decir, depende de las topologías de los espacios métricos pero no de las métricas concretas que se utilicen.

Los siguientes resultados se demuestran rutinariamente y se dejan como ejercicios.

Proposición 2.40 (Regla de la cadena para la continuidad)**.** *Sean E_1, E_2, E_3 espacios métricos, $A \subset E_1$, $f : A \to E_2$, $B \subset E_2$, $g : B \to E_3$ y supongamos que $f(A) \subset B$. Si f es continua en un punto a de A y g es continua en $f(a)$, entonces la composición $g \circ f$ es continua en a. Como consecuencia, si f y g son continuas, entonces $g \circ f$ es continua.*

Pongamos ahora de manifiesto la buena convivencia entre el álgebra y la topología de un espacio normado $(X, \|.\|)$. Esto es,

1. La suma en $(X, \|.\|)$ es continua, es decir, la aplicación de $X \times X$ en X definida por $(x,y) \;\to\; x+y$. En efecto, si $\{(x_n, y_n)\} \to (x,y)$, entonces
$$\|(x+y)-(x_n+y_n)\| \leq \|x-x_n\| + \|y-y_n\| \leq 2 \ \max \{\|x-x_n\|, \|y-y_n\|\} \to 0.$$
2. El producto por escalares en X es continuo, es decir la aplicación de $\mathbb{R} \times X$ en X definida por $(\lambda, x) \;\to\; \lambda x$. En efecto, si $\{\lambda_n\} \to \lambda$, $\{x_n\} \to x$, se tiene
$$\|\lambda x - \lambda_n x_n\| \leq \|(\lambda - \lambda_n)x + \lambda_n(x-x_n)\| \leq |\lambda - \lambda_n| \ \|x\| + |\lambda_n| \ \|x-x_n\| \leq$$
$$(\|x\| + |\lambda_n|) \ \max \{|\lambda - \lambda_n|, \|x-x_n\|\} \to 0,$$
donde se ha tenido en cuenta que $\{\lambda_n\}$ es acotada.

   En particular, el producto en $\mathbb{R}$ es continuo.
3. La norma $\|.\|$ es continua, es decir la aplicación de X en $\mathbb{R}$ definida por $x \;\to\; \|x\|$. Basta tener en cuenta la desigualdad
$$|\,\|x\| - \|x_n\|\,| \leq \|x-x_n\|.$$

Ahora es inmediata la demostración del siguiente resultado:

Corolario 2.41. *Sean (E,d) un espacio métrico, $(X, \|.\|)$ un espacio normado y $a \in A \subset E$.*

i) Si $f, g : A \to X$ son continuas en a y $\lambda \in \mathbb{R}$, entonces $f+g$ y λf son continuas en a.

ii) Si $f : A \to \mathbb{R}$ y $g : A \to X$ son continuas en a, entonces fg es continua en a.

iii) Si $f : A \to \mathbb{R}$ es continua en a y $f(x) \neq 0, \ \forall x \in A$, entonces $\dfrac{1}{f}$ es continua en a.

iv) Si $f, g : A \to \mathbb{R}$ son continuas en a y $g(x) \neq 0, \ \forall x \in A$, entonces $\dfrac{f}{g}$ es continua en a.

Proposición 2.42 (Continuidad de la restricción)**.** *Sean E y F espacios métricos, $B \subset A \subset E$ y $f : A \to F$, y consideremos la aplicación restricción de f a B, $f_{|B} : B \to F$ definida por $f_{|B}(x) := f(x), \ \forall x \in B$. Entonces $f_{|B}$ es continua en todo punto de B en el que f sea continua. En consecuencia, si la función f es continua, entonces su restricción a B también lo es.*

Es importante destacar que la continuidad no se transfiere a extensiones (las funciones definidas en un sólo punto son continuas, luego si la continuidad se transfiriese a extensiones, concluiríamos que todas las funciones son continuas). Sin embargo, se tiene el siguiente resultado parcial.

Proposición 2.43 (Carácter local de la continuidad)**.** *Sean E, F espacios métricos, $A \subset E$, $f : A \to F$ y $b \in B \subset A$. Si B es un "entorno relativo" de b (existe $r > 0$ tal que $B(b,r) \cap A \subset B$) y $f_{|B}$ es continua en b, entonces f es continua en b. En particular, si $f : E \to F$, $B \subset E$ es abierto y si $f_{|B}$ es continua, entonces f es continua en B.*

Demostración:

Sea $r > 0$ tal que $B(b,r) \cap A \subset B$. Si $\{a_n\}$ es una sucesión en A convergente al punto b, entonces

$$\exists m \in \mathbb{N} : \ n \geq m \Rightarrow \ a_n \in B(b,r),$$

con lo que para $n \geq m$ se tiene que $a_n \in B$. Así $\{a_{m+n}\}_{n \in \mathbb{N}}$ es una sucesión en B que converge al punto b, ya que es parcial de la sucesión $\{a_n\}$. Por ser $f_{|B}$ continua en b se tiene que $\{f(a_{m+n})\}_{n \in \mathbb{N}} \to f(b)$ y, por tanto, también $\{f(a_n)\} \to f(b)$. ∎

Ejemplos 2.44 (funciones continuas)**.**

1. Las proyecciones de $\mathbb{R}^N$ en $\mathbb{R}$ son continuas, es decir las aplicaciones $\pi_k : \mathbb{R}^N \to \mathbb{R}$ $(k = 1, \ldots, N)$ definidas por $\pi_k(x_1, \ldots, x_N) = x_k$.
2. Toda función polinómica de $\mathbb{R}^N$ en $\mathbb{R}$ es continua.
3. Toda función racional $R = \dfrac{P}{Q}$ en N variables es continua en su conjunto de definición:
$$\{(x_1, \ldots, x_N) \in \mathbb{R}^N : \ Q(x_1, \ldots, x_N) \neq 0\}.$$
4. Las funciones elementales reales de variable real (exponencial, logaritmo, raíces, funciones trigonométricas) son continuas en su dominio de definición.

La siguiente caracterización proporcionará la manera satisfactoria de introducir en espacios topológicos el concepto de continuidad en un punto.

Proposición 2.45 (ε-δ-caracterización de la continuidad)**.** *Sean (E,d), (F,ρ) espacios métricos, $A \subset E$, $f : A \to F$ y $a \in A$. Equivalen las siguientes afirmaciones:*

i) *f es continua en a.*

ii) $\forall \varepsilon > 0, \ \exists \delta > 0 : \left\{ \begin{array}{c} x \in A \\ d(x,a) < \delta \end{array} \right\} \Rightarrow \ \rho(f(x), f(a)) < \varepsilon$, *equivalentemente*

$$\forall \varepsilon > 0, \ \exists \delta > 0 : \ f(B(a,\delta) \cap A) \subset B(f(a), \varepsilon).$$

Demostración:

$i) \Rightarrow ii)$ Supongamos que no se verifica $ii)$. Entonces existe un positivo ε_0 con la siguiente propiedad:

$$\forall \delta > 0,\ \exists a_\delta \in A : \begin{cases} d(a_\delta, a) < \delta \\ \rho(f(a_\delta), f(a)) \geq \varepsilon_0 \end{cases},$$

en particular

$$\forall n \in \mathbb{N},\ \exists a_n \in A : \begin{cases} d(a_n, a) < \frac{1}{n} \\ \rho(f(a_n), f(a)) \geq \varepsilon_0 \end{cases}.$$

Es inmediato que la sucesión $\{a_n\}$ en A así definida converge hacia a mientras que $\{f(a_n)\}$ no converge a $f(a)$.

$ii) \Rightarrow i)$ Sea $\{a_n\}$ una sucesión en A convergente hacia a. Fijemos $\varepsilon > 0$ y tomemos $\delta > 0$ tal que

$$[x \in A,\ d(x,a) < \delta] \Rightarrow \rho(f(x), f(a)) < \varepsilon.$$

Como $\{a_n\} \to a$, existe un natural m tal que si $n \geq m$ se verifica que $d(a_n, a) < \delta$, y por tanto

$$\rho(f(a_n), f(a)) < \varepsilon.$$

Hemos probado que f es continua en a. ■

Es importante observar que, en la ε-δ-caracterización de la continuidad, el número positivo δ depende tanto del número positivo ε elegido, como del punto a de A prefijado. En general, para un mismo ε positivo pueden aparecer distintos δ dependiendo del punto a de A del que se trate.

En el caso en que, para cada $\varepsilon > 0$ fijo pueda encontrarse un positivo δ común, válido para todos los puntos de A, la función gozará de propiedades adicionales que la diferencian de las funciones que son únicamente continuas.

Definición 2.46 (continuidad uniforme)**.** Sean (E,d) y (F,ρ) espacios métricos, $A \subset E$ y $f : A \to F$ una función. Se dice que f es *uniformemente continua* si

$$\forall \varepsilon > 0, \exists \delta > 0 : [x, y \in A, d(x,y) < \delta] \Rightarrow \rho(f(x), f(y)) < \varepsilon$$

No es difícil comprobar que una función $f : A \to F$ es uniformemente continua si, y sólo si, se verifica la condición

$$a_n, b_n \in A,\ \forall n \in \mathbb{N},\ \{d(a_n, b_n)\} \to 0 \Rightarrow \rho(f(a_n), f(b_n)) \to 0.$$

Es inmediato que si f es uniformemente continua, entonces f es continua. El recíproco no es cierto, de hecho existen funciones continuas muy sencillas que no son uniformemente continuas:

La función $f : \mathbb{R}^+ \to \mathbb{R}$ definida por $f(x) = \dfrac{1}{x}$ es continua y, sin embargo, no es uniformemente continua ya que para cada natural n se tiene que

$$\left| f\left(\frac{1}{2n}\right) - f\left(\frac{1}{n}\right) \right| = n.$$

(Como consecuencia $\frac{1}{f}$ puede no ser uniformemente continua aunque lo sea f).

La función $f:\mathbb{R}\to\mathbb{R}$ definida por $f(x)=x^2$ es continua y, sin embargo, no es uniformemente continua ya que para cada natural n se tiene que

$$\left|f\left(n+\frac{1}{n}\right)-f(n)\right|>2.$$

(Como consecuencia el producto de funciones uniformemente continuas no tiene por qué serlo).

En general, el producto por escalares en un espacio normado $(X,\|.\|)$ es continuo pero no es uniformemente continuo ya que para cada vector x no nulo y para cada natural n se tiene

$$d\left(\left(\frac{1}{n},nx\right),\left(\frac{1}{2n},nx\right)\right)=\frac{1}{2n} \quad \text{y} \quad \left\|f\left(\frac{1}{n},nx\right)-f\left(\frac{1}{2n},nx\right)\right\|=\left|1-\frac{1}{2}\right|\|x\|=\frac{1}{2}\|x\|.$$

Es claro que toda restricción de una función uniformemente continua también lo es.

Proposición 2.47 (Caracterización de la continuidad global)**.** *Sean (E,d), (F,ρ) espacios métricos, $A\subset E$ y $f:A\to F$. Denotemos por $\Im_F$ (resp. $\Im_A$) la familia de los abiertos de F (resp. abiertos relativos de A) y por $\mathscr{F}_F$ (resp. $\mathscr{F}_A$) la familia de los cerrados de F (resp. cerrados relativos de A). Equivalen las siguientes afirmaciones:*

i) f es continua, equivalentemente (Proposición 2.45):

$$\forall a\in A,\ \forall\varepsilon>0,\ \exists\delta_a>0:\left\{\begin{array}{c} x\in A\\ d(x,a)<\delta_a\end{array}\right\}\Rightarrow\ \rho(f(x),f(a))<\varepsilon.$$

ii) La imagen inversa por f de cualquier abierto de F es un abierto relativo de A:

$$O\in\Im_F\Rightarrow\ f^{-1}(O)\in\Im_A.$$

iii) La imagen inversa por f de cualquier cerrado de F es un cerrado relativo de A:

$$C\in\mathscr{F}_F\Rightarrow\ f^{-1}(C)\in\mathscr{F}_A.$$

iv) La función f aplica valores adherentes de cualquier subconjunto de A en valores adherentes de la imagen de dicho conjunto, esto es:

$$f(\overline{B}\cap A)\subset\overline{f(B)},\ \forall B\subset A.$$

Es importante destacar que en las sentencias *ii*) y *iii*) el calificativo relativo es obligado. Si no fuese así, entonces todos los conjuntos de un espacio métrico serían cerrados (cada subconjunto S se puede escribir de la forma $S=f^{-1}(\{0\})$ donde $f:S\to\mathbb{R}$ es la función constantemente cero, y por tanto S es la imagen inversa por una función continua de un cerrado), y en consecuencia también abiertos.

Si la función toma valores reales y está definida en todo el espacio métrico, se tiene el siguiente resultado que permite reconocer subconjuntos abiertos y cerrados de un espacio métrico de forma muy sencilla.

Corolario 2.48. *Sean (E,d) un espacio métrico y $f: E \to \mathbb{R}$ una aplicación continua. Denotemos por $\Im$ (resp. $\mathscr{F}$) la familia de los abiertos (resp. cerrados) de E. Entonces para todo $a \in \mathbb{R}$ se tiene que:*

i) $\{x \in E : f(x) < a\} \in \Im$.

ii) $\{x \in E : f(x) > a\} \in \Im$.

iii) $\{x \in E : f(x) \leq a\} \in \mathscr{F}$.

iv) $\{x \in E : f(x) \geq a\} \in \mathscr{F}$.

v) $\{x \in E : f(x) = a\} \in \mathscr{F}$.

Ejemplo 2.49. Los conjuntos

$$A = \left\{(x,y) \in \mathbb{R}^2 :\ 0 < x, y < \frac{1}{x}\right\}$$

y

$$B = \{(x,y) \in \mathbb{R}^2 :\ 0 < x^2 + y^2 - \text{sen }(xy) < 2,\ ye^x > 2\}$$

son abiertos de $\mathbb{R}^2$.
En efecto, puesto que $A = A_1 \cap A_2$, donde

$$A_1 = \{(x,y) \in \mathbb{R}^2 :\ 0 < x\} \text{ y } A_2 = \{(x,y) \in \mathbb{R}^2 :\ xy < 1\}$$

el resultado se sigue de la continuidad de la aplicación proyección primera $\pi_1 : \mathbb{R}^2 \to \mathbb{R}$ y de la aplicación polinómica $P : \mathbb{R}^2 \to \mathbb{R}$ definida por $P(x,y) = xy$, y del hecho de que

$$A_1 = \pi_1^{-1}(]0, +\infty[) \text{ y } A_2 = P^{-1}(]-\infty, 1[).$$

La prueba de que B es abierto es análoga a la anterior sin más que considerar las funciones continuas de $\mathbb{R}^2$ en $\mathbb{R}$

$$f(x,y) = x^2 + y^2 - \text{sen }(xy)\ ,\ g(x,y) := ye^x,$$

ya que

$$B = f^{-1}(]0,2[) \cap g^{-1}(]2, +\infty[).$$

Probaremos ahora que la compacidad y la conexión se conservan por funciones continuas.

Teorema 2.50 (Conservación de la compacidad por continuidad)**.** *Sean E, F espacios métricos, K un subconjunto compacto de E y $f : K \to F$ continua. Entonces $f(K)$ es compacto.*

Demostración:

Sea $\{y_n\}$ una sucesión en $f(K)$. Tomemos una sucesión $\{x_n\}$ en K tal que $f(x_n) = y_n$, $\forall n \in \mathbb{N}$. Al ser K compacto $\exists \{x_{\sigma(n)}\} \to x \in K$. La continuidad de f nos asegura que $\{f(x_{\sigma(n)})\} \to f(x)$, es decir:

$$\{y_{\sigma(n)}\} \to f(x) \in f(K).$$

Hemos probado que $f(K)$ es compacto. ∎

Corolario 2.51 (Propiedad de compacidad)**.** *Sean E un espacio métrico, K un subconjunto compacto de E y $f : K \to \mathbb{R}$ una función continua. Entonces f está acotada y alcanza su máximo y su mínimo absolutos.*

Demostración:

$f(K)$ es un compacto de $\mathbb{R}$, luego cerrado y acotado, de donde se deduce fácilmente que el conjunto $f(K)$ tiene máximo y mínimo. ∎

El siguiente resultado da condiciones suficientes para que una sucesión de funciones que, en principio, converge sólo puntualmente, converja de hecho uniformemente. La demostración que damos a continuación es un buen ejemplo de utilización del axioma de Heine-Borel.

Teorema 2.52 (Dini)**.** *Sean K un subconjunto compacto de un espacio métrico y $\{f_n\}$ una sucesión monótona de funciones continuas de K en $\mathbb{R}$ que converge puntualmente en K a una función continua f. Entonces la convergencia es uniforme en K.*

Demostración:

Sea $\varepsilon > 0$. Para cada natural n definimos

$$U_n := \{x \in K : |f_n(x) - f(x)| < \varepsilon\}.$$

La continuidad de las funciones f_n y f nos permite asegurar que U_n es un abierto relativo de K. La monotonía de la sucesión $\{f_n\}$ nos dice que $\{U_n\}$ es una sucesión creciente de abiertos relativos. La convergencia puntual implica que dicha familia recubre K. Por último, la compacidad de K (axioma de Heine-Borel y la Nota 2.32), nos permite afirmar que existe m natural tal que $K = U_m$, y por tanto:

$$n \geq m \Rightarrow |f_n(x) - f(x)| < \varepsilon, \ \forall x \in K.$$

∎

El siguiente resultado prueba que las funciones continuas definidas en compactos son de hecho uniformemente continuas.

Teorema 2.53 (Heine)**.** *Sean E y F espacios métricos, $K \subset E$ un subconjunto compacto y $f : K \to F$ una función continua. Entonces f es uniformemente continua.*

Demostración:

Notemos por d y ρ las distancias de E y F, respectivamente. Supongamos que f no es uniformemente continua. Entonces existe $\varepsilon_o > 0$ tal que

$$\forall \delta > 0,\ \exists x_\delta, y_\delta \in K:\ d(x_\delta, y_\delta) < \ \delta \ \text{ y } \ \rho(f(x_\delta), f(y_\delta)) \geq \varepsilon_o.$$

En consecuencia

$$\forall n \in \mathbb{N},\ \exists x_n, y_n \in K:\ d(x_n, y_n) < \frac{1}{n} \ \text{ y } \ \rho(f(x_n), f(y_n)) \geq \varepsilon_o.$$

Como K es compacto, $\exists \{x_{\sigma(n)}\} \to \ x \in K$. Se tiene, para cada natural n, que

$$d(y_{\sigma(n)}, x) \leq d(y_{\sigma(n)}, x_{\sigma(n)}) + d(x_{\sigma(n)}, x) < \frac{1}{\sigma(n)} + d(x_{\sigma(n)}, x),$$

de donde se deduce que $\{y_{\sigma(n)}\} \to x$. Por último, la continuidad de f nos asegura que

$$\{f(x_{\sigma(n)})\} \to f(x) \quad \text{y} \quad \{f(y_{\sigma(n)})\} \to f(x)$$

lo que contradice que $\rho(f(x_n), f(y_n)) \geq \varepsilon_o,\ \forall n \in \mathbb{N}.$ ■

En el Apéndice D del tema se da otra demostración del Teorema de Heine que usa el Teorema de Heine-Borel-Lebesgue.

Queda claro que, en virtud de la Nota 2.32, en los enunciados del Teorema de Dini y de Heine podemos suponer que K es un espacio métrico compacto (en lugar de un subconjunto compacto de un espacio métrico).

Recordemos que si X e Y son conjuntos, $A, B \subset Y$, y $f : X \to Y$, entonces

$$f(f^{-1}(A)) = A \cap f(X)$$

y

$$f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)\ , \qquad f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B).$$

Teorema 2.54 (Conservación de la conexión por la continuidad)**.** *Sean E, F dos espacios métricos, C un subconjunto conexo no vacío de E y $f : C \to F$ continua. Entonces $f(C)$ es conexo.*

Demostración:

Sea $f(C) = U \cup V$ una partición de $f(C)$ en abiertos relativos. Si O es un abierto de F tal que $U = O \cap f(C)$, entonces

$$f^{-1}(U) = f^{-1}(O \cap f(C)) = f^{-1}(O) \cap f^{-1}(f(C)) = f^{-1}(O) \cap C = f^{-1}(O),$$

lo que prueba que $f^{-1}(U)$ es un abierto relativo de C. Análogamente $f^{-1}(V)$ también es un abierto relativo de C. Así, en vista del comentario anterior al teorema, $\{f^{-1}(U), f^{-1}(V)\}$ es una partición de C en abiertos relativos, y al ser C conexo uno de los dos es vacío. Como en este caso, al ser $U,V \subset f(C)$, se verifica que $f(f^{-1}(U)) = U$, $f(f^{-1}(V)) = V$, concluimos que o bien $U = \emptyset$ o bien $V = \emptyset$. Así, la única partición de $f(C)$ en abiertos relativos es la trivial. ■

La anterior proposición generaliza el teorema del valor intermedio: "*Una función real continua definida en un intervalo si toma dos valores toma todos los intermedios*" ¿Por qué?

Corolario 2.55. *Los segmentos en $\mathbb{R}^N$ son conexos.*

Demostración:
El segmento $[x,y]$ es la imagen del intervalo $[0,1]$ por la función continua $f(t) = x + t(y - x)$. En general, los segmentos de extremos x e y ($[x,y], [x,y[,]x,y],]x,y[$) son conexos (hágase!). ■

A continuación mostramos que un conjunto C que verifique la propiedad que aparece en el Teorema 2.54, para cualquier función continua valuada en $\{0,1\}$, ha de ser conexo. A pesar de la sencillez de la prueba, esta caracterización resulta muy útil para probar ciertas propiedades de estabilidad de los conjuntos conexos.

Proposición 2.56 (Caracterización de la conexión)**.** *Sea C un subconjunto no vacío de un espacio métrico. Equivalen las siguientes afirmaciones:*

i) C es conexo.

ii) Toda función continua de C en $\{0,1\}$ es constante.

Demostración:
i) $\Rightarrow$ *ii)* Es consecuencia del teorema anterior y de que los únicos intervalos no vacíos contenidos en $\{0,1\}$ son $\{0\}$ y $\{1\}$.
ii) $\Rightarrow$ *i)* Supongamos, razonando por reducción al absurdo, que C no es conexo. Sea $C = U \cup V$ una partición no trivial de C en abiertos relativos. Entonces la función $f : C \to \{0,1\}$ definida por

$$f(x) := \begin{cases} 0 & \text{si } x \in U \\ 1 & \text{si } x \in V \end{cases}$$

es continua (por el carácter local de la continuidad), y verifica $f(C) = \{0,1\}$, y por tanto, *ii)* no es cierta. ■

Los siguientes resultados nos ayudan a probar la conexión de ciertos subconjuntos.

Proposición 2.57. *Todo conjunto comprendido entre un conexo y su cierre es conexo, es decir si C es conexo y $C \subset D \subset \overline{C}$, entonces D es conexo.*

Demostración:

Si $D = \emptyset$ es claro. Supongamos $D \neq \emptyset$ y sea $f : D \to \{0,1\}$ una aplicación continua. Por ser C conexo, en vista de la Proposición 2.56 ha de ser $f_{|C}$ constante. Como $D \subset \overline{C}$ y f es continua se tiene que

$$f(D) = f(\overline{C} \cap D) \subset \overline{f(C)}.$$

Por tanto, f es constante. ∎

Proposición 2.58. *Sea $\{C_i : \ i \in I\}$ una familia de conexos de un espacio métrico tal que dos cualesquiera tienen intersección no vacía. Entonces $\bigcup_{i \in I} C_i$ es conexo.*

Demostración:

Sea $f : \bigcup_{\in I} C_i \to \{0,1\}$ continua. Queremos probar que f es constante. Sean $x, y \in \bigcup_{i \in I} C_i$. Existen entonces $r, s \in I$ tales que $x \in C_r$ e $y \in C_s$. Como $f_{|C_r}$ y $f_{|C_s}$ son constantes y existe $c \in C_r \cap C_s$ concluimos que $f(x) = f(c) = f(y)$. ∎

Corolario 2.59. *Todo convexo de $\mathbb{R}^N$ es conexo.*

Demostración:

Sea C un subconjunto convexo de $\mathbb{R}^N$ y tomemos $c \in C$. Entonces $C = \cup_{x \in C} [c, x]$ es conexo por ser unión de conexos con al menos el punto c en común. ∎

2.5. Límite funcional.

Recordemos que "la" topología de $\mathbb{R}^N$ sigue siendo la topología de la norma.

Definición 2.60. Sean (E,d) un espacio métrico y A un subconjunto no vacío de E. Se dice que $\alpha \in E$ es un *punto de acumulación* de A si existe una sucesión $\{a_n\}$ de puntos de A distintos de α y convergente a α, equivalentemente, si

$$B(\alpha,\varepsilon)\cap(A\backslash\{\alpha\})\neq\emptyset,\ \forall\varepsilon>0.$$

Denotaremos por A' al conjunto de los puntos de acumulación de A. Se dice que un punto $a\in A$ es un *punto aislado* de A si existe $\varepsilon>0$ tal que $B(a,\varepsilon)\cap A=\{a\}$.

Es inmediato que todo punto adherente a A o es de acumulación de A o es aislado. En consecuencia los puntos de A o son de acumulación o son aislados.

Definición 2.61. Sean (E,d) y (F,ρ) espacios métricos, $A\subset E$, $f:A\to F$, α un punto de acumulación de A y $\ell\in F$. Se dice que f tiende a ℓ cuando x tiende a α, y se nota $f(x)\to\ell$ cuando $x\to\alpha$, si para toda sucesión de puntos de A distintos de α y convergente a α, se verifica que la sucesión imagen converge a ℓ, es decir:

$$[\ \forall\{a_n\} \text{ en } A\ ,\ a_n\neq\alpha\ ,\ \{a_n\}\xrightarrow{d}\alpha\ ]\Rightarrow\{f(a_n)\}\xrightarrow{\rho}\ell.$$

Supuesta la existencia de un tal ℓ, de la unicidad del límite secuencial se sigue que tal elemento es único, se llama *límite de la función f* en el punto α y se nota

$$\lim_{x\to\alpha} f(x)=\ell.$$

Como consecuencia inmediata del Ejemplo 2.13, el estudio de la existencia del límite de los campos vectoriales se reduce al de sus campos escalares componentes, tal como se recoge en el siguiente enunciado.

Proposición 2.62 (Reducción a campos escalares)**.** *Sean M,N naturales, $A\subset\mathbb{R}^N$, $f=(f_1,\ldots,f_M):A\to\mathbb{R}^M$ un campo vectorial en A y α un punto de acumulación de A. Entonces*

$$f \text{ tiene límite en } \alpha \Leftrightarrow f_i \text{ tiene límite en } \alpha,\ \forall i=1,\ldots,M.$$

En tal caso,

$$\lim_{x\to\alpha} f(x)=(\lim_{x\to\alpha} f_1(x),\ldots,\lim_{x\to\alpha} f_M(x)).$$

Proposición 2.63 (Álgebra de límites)**.** *Sean $A\subset\mathbb{R}^N$, α un punto de acumulación de A, $f,g:A\to\mathbb{R}$ y $\lambda\in\mathbb{R}$. Se verifica:*

i) Si f,g *tienen límite en* α*, entonces* $f+g$ *y* λf *tienen límite en* α *y*

$$\lim_{x\to\alpha}(f+g)(x)=\lim_{x\to\alpha}f(x)+\lim_{x\to\alpha}g(x),$$

$$\lim_{x\to\alpha}(\lambda f)(x)=\lambda\ \lim_{x\to\alpha}f(x).$$

ii) Si f,g *tienen límite en* α*, entonces* fg *tiene límite en* α *y*

$$\lim_{x\to\alpha}(fg)(x)=\lim_{x\to\alpha}f(x)\ \lim_{x\to\alpha}g(x).$$

iii) Si f,g *tienen límite en* α*,* $g(x)\neq 0\ \ \forall x\in A$*, y* $\lim_{x\to\alpha}g(x)\neq 0$*, entonces* $\dfrac{f}{g}$ *tiene límite en* α *y*

$$\lim_{x\to\alpha}\frac{f}{g}(x)=\frac{\lim_{x\to\alpha}f(x)}{\lim_{x\to\alpha}g(x)}.$$

Notas 2.64.
a) Puesto que en los espacios métricos la convergencia secuencial depende sólo de la topología, se tiene que, al igual que ocurría con la continuidad, también el concepto de límite funcional es de carácter topológico: depende de las topologías de los espacios métricos pero no de las métricas concretas que se utilicen.
b) Puede ocurrir que el punto α no pertenezca al conjunto A, pero que sea de acumulación. En el caso en que α pertenezca a A, el valor que tome la función f en α no afecta para nada a la existencia del límite, ni al valor de éste.
c) Nótese que, en la Definición 2.61, basta exigir la condición

$$[\,\forall\{a_n\}\text{ en }A\ ,\ a_n\neq\alpha\ ,\ \{a_n\}\xrightarrow{d}\alpha\,]\Rightarrow\{f(a_n)\}\text{ es convergente.}$$

En efecto, si $\{a_n\}$ y $\{a_n^{'}\}$ son dos sucesiones que verifican la hipótesis anterior, entonces la sucesión "mezcla" $\{a_1,a_1^{'},a_2,a_2^{'},\ldots,a_n,a_n^{'},\ldots\}$ está en las mismas circunstancias, y por tanto, la sucesión imagen

$$\{f(a_1),f(a_1^{'}),f(a_2),f(a_2^{'}),\ldots,f(a_n),f(a_n^{'}),\ldots\}$$

es convergente, lo que conlleva a que

$$\lim\ f(a_n)=\lim\ f(a_n^{'}).$$

La relación entre la continuidad de una función en un punto y la existencia de límite funcional en dicho punto se recoge en el siguiente resultado:

Proposición 2.65. *Sean* E,F *espacios métricos,* $A\subset E$*,* $f:A\to F$ *y* a *un punto de* A*.*

i) Si a *es un punto aislado de* A*, entonces* f *es continua en* a*.*

ii) Si a es un punto de acumulación de A, entonces f es continua en a si, y sólo si, $\lim_{x\to a} f(x) = f(a)$.

La demostración se deja como ejercicio.

El siguiente resultado recoge el utilísimo proceso de cambio a coordenadas polares a la hora de estudiar límites de funciones de dos variables reales. Claramente, vía el uso de traslaciones, no es restrictivo el llevar a cabo el estudio del límite en el punto $(0,0)$.

Recordemos que la función paso a coordenadas polares es la función de $\mathbb{R}_0^+ \times \mathbb{R}$ en $\mathbb{R}^2$ definida por

$$(\rho,\vartheta) \to (\rho\ cos\ \vartheta, \rho\ \text{sen}\ \vartheta),$$

que es una aplicación sobreyectiva, aplica el eje $\{0\}\times\mathbb{R}$ en $(0,0)$, y que es periódica de periodo 2π en la variable ϑ. En consecuencia, esta aplicación induce una biyección de la franja $\mathbb{R}^+\times]-\pi,\pi]$ sobre $\mathbb{R}^2\setminus\{(0,0)\}$. En efecto: dado $(x,y)\in\mathbb{R}^2\setminus\{(0,0)\}$ existe un único $(\rho,\vartheta)\in\mathbb{R}^+\times]-\pi,\pi]$ tal que

$$x = \rho\ cos\ \vartheta\ , y = \rho\ \text{sen}\ \vartheta.$$

De hecho, ρ y ϑ son el módulo y el "argumento principal" de (x,y), y (ρ,ϑ) se pueden obtener a a partir de (x,y) mediante las expresiones $\rho = \sqrt{x^2+y^2}$ y

$$\vartheta = \begin{cases} \pi & \text{si } x\in\mathbb{R}^-, y=0 \\ 2\ \arctan\dfrac{y}{\rho+x} & \text{en otro caso.} \end{cases}$$

Al par (ρ,ϑ) se le llama <u>*coordenadas polares*</u> del punto (x,y).

Proposición 2.66 (coordenadas polares)**.** *Sean* $f:\mathbb{R}^2\setminus\{(0,0)\}\to\mathbb{R}$ *una función y* ℓ *un número real. Equivalen las siguientes afirmaciones:*

i) $\lim_{(x,y)\to(0,0)} f(x,y) = \ell$.

ii) $\forall\varepsilon>0,\ \exists\delta>0:\ 0<\|(x,y)\|<\delta \Rightarrow |f(x,y)-\ell|<\varepsilon$.

iii) $\lim_{\rho\to 0} f(\rho\ cos\ \vartheta,\rho\ \text{sen}\ \vartheta) = \ell$ *uniformemente en* ϑ*, es decir:*

$$\forall\varepsilon>0,\ \exists\delta>0:\ [0<\rho<\delta,\ \vartheta\in\mathbb{R}\,] \Rightarrow |f(\rho\ cos\ \vartheta,\rho\ \text{sen}\ \vartheta)-\ell|<\varepsilon.$$

iv) Para cualquier sucesión $\{\rho_n\}$ *de positivos convergente a cero y cualquier sucesión de reales* $\{\vartheta_n\}$ *(si queremos* $|\vartheta_n|\le\pi$*), se verifica que*

$$\{f(\rho_n\ cos\ \vartheta_n,\rho_n\ \text{sen}\ \vartheta_n)\}\to\ell.$$

Demostración:

i) $\Rightarrow$ *ii)* Se deja como ejercicio (véase la $\varepsilon-\delta$-caracterización de la continuidad).

ii) $\Rightarrow$ *iii)*. Sea $\varepsilon > 0$ fijo. Por hipótesis

$$\exists \delta > 0 : 0 < \|(x,y)\|_2 < \delta \Rightarrow |f(x,y) - \ell| < \varepsilon.$$

Ahora, si ρ es tal que $0 < \rho < \delta$, entonces para todo real ϑ se verifica que

$$\|(\rho\ cos\ \vartheta, \rho\ \text{ sen }\ \vartheta)\|_2 = \rho < \delta,$$

luego se tiene que

$$|f(\rho\ cos\ \vartheta, \rho\ \text{ sen }\ \vartheta) - \ell| < \varepsilon.$$

iii) $\Rightarrow$ *iv)* Sean $\{\rho_n\}$ una sucesión de números reales positivos convergente a cero y $\{\vartheta_n\}$ una sucesión de números reales. Para $\varepsilon > 0$ fijo, por hipótesis

$$\exists \delta > 0 : (0 < \rho < \delta, \vartheta \in \mathbb{R}) \Rightarrow |f(\rho\ cos\ \vartheta, \rho\ \text{ sen }\ \vartheta) - \ell| < \varepsilon.$$

Como $\{\rho_n\} \to 0$, entonces

$$\exists m \in \mathbb{N}\ :\ n \geq m \Rightarrow 0 < \rho_n < \delta,$$

y por tanto, en vista de la hipótesis

$$n \geq m\ \Rightarrow\ |f(\rho_n\ cos\ \vartheta_n, \rho_n\ \text{ sen }\ \vartheta_n) - \ell| < \varepsilon.$$

En consecuencia,

$$\{f(\rho_n\ cos\ \vartheta_n, \rho_n\ \text{ sen }\ \vartheta_n)\} \to \ell.$$

iv) $\Rightarrow$ *i)* Sea $\{(x_n,y_n)\}$ una sucesión en $\mathbb{R}^2 \setminus \{(0,0)\}$ convergente a $(0,0)$. Para cada natural n tomemos $\rho_n := \|(x_n,y_n)\|_2$ y $\vartheta_n \in \mathbb{R}$ tal que

$$x_n = \rho_n\ cos\ \vartheta_n\ , y_n = \rho_n\ \text{ sen }\ \vartheta_n.$$

De la continuidad de la norma se sigue que $\{\rho_n\} \to 0$, y por tanto por *iv)*

$$\{f(x_n,y_n)\} = \{f(\rho_n\ cos\ \vartheta_n, \rho_n\ \text{ sen }\ \vartheta_n)\} \to \ell.$$

∎

Corolario 2.67 (Límites direccionales)**.** *Sea* $f : \mathbb{R}^2 \setminus \{(a,b)\} \to \mathbb{R}$. *Si existe* $\lim_{(x,y)\to(a,b)} f(x,y) = \ell$ *entonces, para todo* $\vartheta \in]-\pi,\pi]$ *existe*

$$\lim_{\rho\to 0} f(a + \rho\ cos\ \vartheta, b + \rho\ \text{ sen }\ \vartheta) = \ell.$$

Los anteriores límites se llaman <u>límites direccionales</u> en la dirección ϑ.

En consecuencia, de existir límite han de existir todos los límites direccionales y ser iguales al límite.

En la práctica es usual para calcular los límites direccionales de $f(x,y)$ en el punto (a,b), considerar las rectas $y = b + m(x-a)$ (de pendiente m y que pasa por (a,b)) y hallar el límite

$$\lim_{x \to a} f(x, b + m(x-a)).$$

Obsérvese que de no existir un límite direccional o en el caso de que dos límites direccionales sean distintos podemos afirmar que no hay límite. Supuesto que todos los límites direccionales existen y son iguales, éste es el candidato a límite, aunque puede que no haya límite (véase Ejercicio 2.27, g).

2.6. Apéndice.

2.6.1. A) Teorema de Heine-Borel-Lebesque.

Teorema de Heine-Borel-Lebesgue. *Sea K un subconjunto de un espacio métrico. Equivalen:*

i) K es compacto, esto es, toda sucesión de puntos de K admite una sucesión parcial convergente a un punto de K.

ii) K verifica el axioma de Heine-Borel : todo recubrimiento por abiertos de K admite un subrecubrimiento finito, esto es, si $\mathcal{U}$ es una familia de abiertos de E tales que $K \subset \bigcup_{U\in\mathcal{U}} U$, entonces existen $n \in \mathbb{N}$ y $U_1, \cdots, U_n \in \mathcal{U}$ tal que $K \subset \bigcup_{i=1}^{n} U_i$.

En la demostración del teorema anterior usaremos los dos siguientes resultados auxiliares, de interés en sí mismos:

Proposición 2.68. *Sea $\{x_n\}$ una sucesión de un espacio métrico E y $x \in E$. Son equivalentes:*

a) La sucesión $\{x_n\}$ admite una parcial convergente a x.

b) Para cada positivo ε, el conjunto $\{n \in \mathbb{N} : x_n \in B(x,\varepsilon)\}$ es infinito.

Demostración:
Por definición de convergencia es claro que $a) \Rightarrow b)$.
$b) \Rightarrow a)$ Supongamos $b)$ cierto y definamos $\sigma : \mathbb{N} \longrightarrow \mathbb{N}$ de la siguiente manera:

$$\sigma(1) = \min \{n \in \mathbb{N} : x_n \in B(x,1)\}.$$

Supuesto definido $\sigma(k)$, para $k \leq n$, definimos

$$\sigma(n+1) = \min \left\{ m \in \mathbb{N} : m > \sigma(n), x_m \in B\left(x, \frac{1}{n+1}\right) \right\}.$$

Así conseguimos una sucesión parcial $\{x_{\sigma(n)}\}$ que, por la forma de construirla, claramente converge a x. ■

Lema 2.69 (del número de Lebesgue)**.** *Sea K un subconjunto compacto de un espacio métrico. Entonces para todo recubrimiento por abiertos $\mathcal{U}$ de K existe $r > 0$ tal que*

$$x \in K \Rightarrow \exists U \in \mathcal{U} : B(x,r) \subset U.$$

En tal caso se dice que r es un número de Lebesgue asociado al recubrimiento $\mathcal{U}$.

Demostración:

Si no existiera un número de Lebesgue, entonces habría un recubrimiento por abiertos $\mathscr{U}$ de K tal que

$$n \in \mathbb{N} \Rightarrow \left(\exists x_n \in K : B\left(x_n, \frac{1}{n}\right) \subsetneq U,\ \forall U \in \mathscr{U}\right).$$

Por hipótesis $\{x_n\}$ admite una parcial convergente a un elemento x de K. Por tanto, ha de existir un conjunto $U \in \mathscr{U}$ tal que $x \in U$, y, por ser este conjunto abierto, para algún positivo ρ se verifica $B(x,\rho) \subset U$. Ahora, elegimos $\frac{1}{N} < \frac{\rho}{2}$, y sabemos que existe $n > N$ tal que $d(x_n,x) < \frac{\rho}{2}$. Entonces, si $y \in B(x_n, \frac{\rho}{2})$, se tiene

$$d(x,y) \leq d(x,x_n) + d(x_n,y) < \rho.$$

Así

$$B\left(x_n, \frac{\rho}{2}\right) \subset B(x,\rho) \subset U.$$

Como $\frac{1}{n} < \frac{1}{N} < \frac{\rho}{2}$, obtenemos que $B(x_n, \frac{1}{n}) \subset B(x_n, \frac{\rho}{2}) \subset U$, lo que contradice la elección de x_n. ∎

Demostración del Teorema de Heine-Borel-Lebesgue.

$i) \Rightarrow ii)$ Probamos primero que dado un positivo ε, existe un subconjunto finito $F \subset K$ tal que $K \subset \bigcup_{x\in F} B(x,\varepsilon)$. De no ser así, tomemos

$$x_1 \in K, \quad x_2 \in K \setminus B(x_1,\varepsilon),$$

y en general,

$$x_{n+1} \in K \setminus \cup_{i=1}^{n} B(x_i,\varepsilon).$$

Obtendríamos así una sucesión en K verificando que $d(x_n,x_m) \geq \varepsilon$ si $n \neq m$, o sea $\{x_n\}$ es una sucesión sin parciales convergentes (pues ni siquiera pueden ser de Cauchy), lo que contradice la hipótesis.

Ahora probamos que K verifica el axioma de Heine-Borel. Fijamos un recubrimiento por abiertos $\mathscr{U}$ de K. Sea r un número de Lebesgue asociado al recubrimiento. Sabemos que existe un conjunto finito $F \subset K$ tal que $K \subset \bigcup_{x\in F} B(x,r)$. Ahora bien, cada una de las bolas anteriores está contenida en algún abierto del recubrimiento por ser r un numero de Lebesgue, por tanto, éste admite un subrecubrimiento finito, como queríamos demostrar.

ii) $\Rightarrow$ *i*) Sea $\{x_n\}$ una sucesión en K y supongamos que no admite ninguna parcial convergente a ningún elemento de K. Por la Proposición 2.68, ha de ocurrir

$$y \in K \Rightarrow \exists r_y > 0 : \{n \in \mathbb{N} : x_n \in B(y, r_y)\} \text{ es finito.}$$

Como la familia de conjuntos

$$\{B(y, r_y) : y \in K\}$$

es un recubrimiento por abiertos de K, por hipótesis, deben de existir elementos $y_1, \cdots, y_m$ tales que

$$K \subset B(y_1, r_{y_1}) \cup \cdots \cup B(y_m, r_{y_m}),$$

lo cual es una contradicción, ya que en tal caso

$$\mathbb{N} = \{n \in \mathbb{N} : x_n \in K\} \subset \left\{ n \in \mathbb{N} : x_n \in \bigcup_{j=1}^{m} B(y_j, r_{y_j}) \right\}$$

sería finito, ya que cada una de las bolas $B(y_i, r_{y_i})$ cuya unión recubre K contiene a un número finito de términos de la sucesión.

$\blacksquare$

2.6.2. B) Desigualdad entre la media geométrica y aritmética.

Proposición. *Sea n un natural mayor o igual que 2. Entonces*

$$0 \leq a_1, a_2, ..., a_n \Rightarrow \sqrt[n]{a_1 a_2 ... a_n} \leq \frac{a_1 + a_2 + ... + a_n}{n}$$

y la igualdad ocurre si, y sólo si, $a_1 = a_2 = ... = a_n$.
Demostración:

Si alguno de los a_i es nulo la proposición es inmediata. También es inmediato que si todos los a_i son iguales se da la igualdad. Queda así reducida la demostración a probar la siguiente implicación:

$$[\, 0 < a_1, a_2, ..., a_n,\ a_1 < a_2 \,] \Rightarrow \sqrt[n]{a_1 a_2 ... a_n} < \frac{a_1 + a_2 + ... + a_n}{n}$$

que a su vez se deduce de probar

$$[\, 0 < x_1, x_2, ..., x_n,\ x_1 x_2 ... x_n = 1,\ x_1 < 1 < x_2 \,] \Rightarrow n < x_1 + x_2 + ... + x_n \qquad (*)$$

(basta aplicar (*) a los números $x_i = \frac{a_i}{\sqrt[n]{a_1 a_2 ... a_n}}, i = 1, 2, ..., n$). Vamos a demostrar (*) por inducción.

Comprobemos la implicación para $n = 2$.

$$x_1 < 1 < x_2 \ \Rightarrow\ (1 - x_1)(1 - x_2) < 0 \ \Leftrightarrow\ 1 + x_1 x_2 < x_1 + x_2 \ \Rightarrow\ 2 < x_1 + x_2.$$

Supongamos (*) cierta para n y probémosla para $n + 1$. Sean

$$0 < x_1, x_2, ..., x_n, x_{n+1},\ x_1 x_2 ... x_n x_{n+1} = 1,\ x_1 < 1 < x_2$$

y consideremos los siguientes n números $(x_1 x_2), x_3, ..., x_n, x_{n+1}$. Es claro que todos son positivos y su producto vale 1. Tanto si todos son iguales (en cuyo caso son todos iguales a 1) como si no (en cuyo caso aplicamos la hipótesis de inducción) tenemos que

$$n \leq x_1 x_2 + x_3 + ... + x_n + x_{n+1}$$

y por tanto (usando la desigualdad $1 + x_1 x_2 < x_1 + x_2$ vista en la prueba para $n = 2$) concluimos que

$$\begin{aligned} n + 1 &\leq x_1 x_2 + x_3 + ... + x_n + x_{n+1} + 1 = \\ &= (1 + x_1 x_2) + x_3 + ... + x_n + x_{n+1} < \\ &< x_1 + x_2 + x_3 + ... + x_n + x_{n+1}. \end{aligned}$$

■

2.6.3. C) Demostración de la caracterización de la continuidad global.

Demostración de la Proposición 2.47.

$i) \Rightarrow ii)$ Sea $O \subset F$ abierto. Si $f^{-1}(O)$ es vacío no hay nada que demostrar. En otro caso sea $a \in f^{-1}(O)$ fijo. Tomemos $\varepsilon > 0$ tal que $B(f(a), \varepsilon) \subset O$. La hipótesis nos asegura la existencia de un positivo δ_a tal que

$$f(B(a, \delta_a) \cap A) \subset B(f(a), \varepsilon),$$

y en consecuencia

$$B(a, \delta_a) \cap A \subset f^{-1}(O).$$

Hemos probado que $f^{-1}(O)$ contiene la intersección de A con una bola centrada en cada uno de sus puntos, luego $f^{-1}(O)$ es un abierto relativo de A.

$ii) \Rightarrow iii)$ Si C es un cerrado de F, entonces

$$A \backslash f^{-1}(C) = f^{-1}(F \backslash C)$$

es un abierto relativo de A, y por tanto $f^{-1}(C)$ es un cerrado relativo de A.

$iii) \Rightarrow iv)$ Sea $B \subset A$. Por hipótesis, $f^{-1}\left(\overline{f(B)}\right)$ es un cerrado relativo de A que, evidentemente, contiene a B, por tanto, ha de contener al cierre de B en la topología relativa a A, esto es,

$$\overline{B} \cap A \subset f^{-1}\left(\overline{f(B)}\right).$$

Aplicando f obtenemos $f(\overline{B} \cap A) \subset \overline{f(B)}$.

$iv) \Rightarrow i)$ Si no ocurre $i)$, entonces

$$\exists a \in A, \exists \varepsilon_0 > 0 : \ \forall \delta > 0, \exists x_\delta \in A \text{ verificando } \begin{cases} d(x_\delta, a) < \delta \\ \rho(f(x_\delta), f(a)) \geq \varepsilon_0 \end{cases}$$

y por tanto $B := \{x_\delta : \delta > 0\}$ es un subconjunto de A tal que $a \in \overline{B}$ y $f(a) \notin \overline{f(B)}$, y en consecuencia no se verifica $iv)$. ■

2.6.4. D) Otra demostración del Teorema de Heine.

Muchas veces la compacidad se usa para "uniformizar" una condición que a priori no parece ser uniforme. Como muestra de esta idea, daremos una nueva demostración del Teorema de Heine, que es directa.

Teorema de Heine. *Sea F un espacio métrico, K un espacio métrico compacto y $f : K \to F$ una función continua. Entonces f es uniformemente continua.*

Demostración. Notemos por d y ρ las distancias de K y F, respectivamente. Por la ε-δ-caracterización de continuidad, dado un positivo ε, para cada punto x de K existe un positivo $\delta(x)$ tal que

$$y \in K, d(y,x) < \delta(x) \Rightarrow \rho(f(y), f(x)) < \varepsilon.$$

Ahora, variando el punto x, recubrimos el compacto por una unión de abiertos, sin más que considerar

$$K \subset \bigcup_{x\in K} B\left(x, \frac{\delta(x)}{2}\right).$$

Por ser K compacto, se tiene que, en vista del Teorema de Heine-Borel-Lebesgue, podemos obtener un subrecubrimiento finito. Esto es, existen elementos $x_1, x_2, \cdots, x_n \in K$ tales que

$$(\star) \qquad K \subset \bigcup_{i=1}^{n} B\left(x_i, \frac{\delta_i}{2}\right),$$

donde hemos notado por δ_i a los positivos que verifican la condición de continuidad en el punto x_i.

Tomamos $\delta = \min\{\frac{\delta_i}{2} : 1 \leq i \leq n\}$ y si $x, y \in K$ verifican que $d(x,y) < \delta$, entonces, por $(\star)$, se verifica que, para algún i se tiene $x \in B\left(x_i, \frac{\delta_i}{2}\right)$, por tanto, como $\delta \leq \delta_i$ obtenemos que $x, y \in B(x_i, \delta_i)$ y usando la continuidad de f en x_i se tiene que

$$\rho(f(x), f(x_i)) < \varepsilon, \ \rho(f(y), f(x_i)) < \varepsilon.$$

Finalmente, usando la desigualdad triangular se deduce que

$$\rho(f(y), f(x)) < 2\varepsilon.$$

■

2.6.5. E) Fórmula para el argumento de un número complejo.

Proposición. *Para todo $z = x + iy \in \mathbb{C}^*$, si definimos ϑ por*

$$\begin{cases} \vartheta = 2\arctan \dfrac{y}{x+|z|} & \text{si} \quad z \in \mathbb{C}^*\backslash\mathbb{R}^- \\ \vartheta = \pi & \text{si} \quad z \in \mathbb{R}^- \end{cases}$$

entonces ϑ es el único número real en $]-\pi,\pi]$ que verifica que

$$z = |z|\ (\cos\vartheta + i \operatorname{sen}\vartheta).$$

Demostración. Supuesto que existan dos elementos ϑ, ϑ_0 que verifiquen lo anterior, entonces, se tendría que

$$\cos\vartheta = \cos\vartheta_0, \qquad \operatorname{sen}\vartheta = \operatorname{sen}\vartheta_0,$$

por tanto,

$$\cos(\vartheta - \vartheta_0) = \cos\vartheta\cos\vartheta_0 + \operatorname{sen}\vartheta\operatorname{sen}\vartheta_0 = 1,$$

de donde $\vartheta - \vartheta_0 \in 2\pi\mathbb{Z}$ y como $\vartheta, \vartheta_0 \in]-\pi,\pi]$, entonces $\vartheta = \vartheta_0$.

Comprobamos ahora la existencia y únicamente lo hacemos en el caso de que $z \in \mathbb{C}^*\backslash\{\mathbb{R}^-\}$. Tomamos

$$\vartheta = 2\arctan\frac{y}{x+|z|},$$

por tanto $\tan\frac{\vartheta}{2} = \frac{y}{x+|z|}$, de donde

$$\cos\vartheta = \cos^2\frac{\vartheta}{2} - \operatorname{sen}^2\frac{\vartheta}{2} = \frac{\cos^2\frac{\vartheta}{2} - \operatorname{sen}^2\frac{\vartheta}{2}}{\cos^2\frac{\vartheta}{2} + \operatorname{sen}^2\frac{\vartheta}{2}} = \frac{1-\tan^2\frac{\vartheta}{2}}{1+\tan^2\frac{\vartheta}{2}} =$$

$$= \frac{(x+|z|)^2 - y^2}{(x+|z|)^2 + y^2} = \frac{x^2 + x^2 + y^2 + 2|z|x - y^2}{x^2 + y^2 + |z|^2 + 2x|z|} = \frac{2x(x+|z|)}{2|z|(|z|+x)} = \frac{x}{|z|}.$$

Análogamente se tiene

$$\operatorname{sen}\vartheta = 2\operatorname{sen}\frac{\vartheta}{2}\cos\frac{\vartheta}{2} = \frac{2\operatorname{sen}\frac{\vartheta}{2}\cos\frac{\vartheta}{2}}{\operatorname{sen}^2\frac{\vartheta}{2} + \cos^2\frac{\vartheta}{2}} = \frac{2\tan\frac{\vartheta}{2}}{1+\tan^2\frac{\vartheta}{2}} =$$

$$= \frac{2\frac{y}{x+|z|}}{\frac{y^2}{(x+|z|)^2} + 1} = \frac{2y(x+|z|)}{y^2 + (x+|z|)^2} =$$

$$\frac{2y(x+|z|)}{y^2 + x^2 + x^2 + y^2 + 2x|z|} = \frac{2y(x+|z|)}{2|z|(|z|+x)} = \frac{y}{|z|}.$$

■

Dado $z \in \mathbb{C}^*$ se llama <u>*argumento*</u> de z a todo número real t que verifique la igualdad $z = |z|(\cos t + i \operatorname{sen} t)$.

2.7. Referencias recomendadas.

[Ber], [Bra], [Bri], [Cra], [Fe], [Jur], [MaHo], [SoSi] y [Stro].

2.8. Resumen de resultados del Tema 2.

Espacio normado. Si X es un espacio vectorial real, una *norma* en X es una función $\|\cdot\| : X \longrightarrow \mathbb{R}_0^+$ verificando

i) $\|x\| = 0 \Leftrightarrow x = 0$.

ii) $\|\lambda x\| = |\lambda| \, \|x\|, \ \forall \lambda \in \mathbb{R}, \ \forall x \in X$ (homogeneidad).

iii) $\|x+y\| \leq \|x\| + \|y\|, \ \forall x,y \in X$ (desigualdad triangular).

El par ordenado $(X, \|\cdot\|)$ se llama espacio normado.

Espacio métrico. Una distancia (o métrica) definida en un conjunto no vacío E es una función $d : E \times E \longrightarrow \mathbb{R}_0^+$ que verifica:

1. $d(x,y) = 0 \Leftrightarrow x = y$.
2. $d(x,y) = d(y,x), \ \forall x,y \in E$ (propiedad simétrica).
3. $d(x,z) \leq d(x,y) + d(y,z), \ \forall x,y,z \in E$ (desigualdad triangular).

Al par ordenado (E,d) se le denomina espacio métrico.

Topología de un espacio métrico. Sea (E,d) un espacio métrico.
Dado $a \in E, r \geq 0$, la bola abierta (resp. cerrada) de centro a y radio r son los conjuntos dados por

$$B(a,r) := \{x \in E : d(x,a) < r\}, \ \overline{B}(a,r) := \{x \in E : d(x,a) \leq r\}.$$

La *esfera* de centro a y radio r es el conjunto $S(a,r) := \{x \in E : d(x,a) = r\}$.
Un subconjunto O de un espacio métrico (E,d) se dice abierto si verifica:

$$\forall a \in O, \exists r > 0 : B(a,r) \subset O.$$

Así pues, un subconjunto es abierto si puede expresarse como unión de bolas abiertas.
Todo espacio normado es un espacio topológico con la topología asociada a la distancia definida por $d(x,y) := \|x-y\|$.

Convergencia. Una sucesión $\{x_n\}$ de elementos de un espacio métrico (E,d) es convergente a x si $\{d(x_n,x)\} \to 0$. El concepto de convergencia es topológico, esto es, si en un espacio E dos distancias d y d^* generan la misma topología, entonces las sucesiones convergentes en (E,d) y en (E,d^*) son las mismas (y además tienen los mismos límites).
La convergencia en el espacio normado $(\mathscr{C}[a,b], \|.\|_\infty)$ se llama convergencia uniforme, esto es,

$$\{f_n\} \xrightarrow{\|.\|_\infty} f \Leftrightarrow [\ \forall \varepsilon > 0, \ \exists m \in \mathbb{N} : \ n \geq m \Rightarrow |f_n(x) - f(x)| \leq \varepsilon, \ \forall x \in [a,b]\ ].$$

Una sucesión $\{x_n\}$ de elementos de un espacio métrico (E,d) es de Cauchy si se verifica

$$\forall \varepsilon > 0, \ \exists m \in \mathbb{N} : p,q \geq m \Rightarrow d(x_p,x_q) \leq \varepsilon,$$

equivalentemente,

$$\forall \varepsilon > 0, \ \exists m \in \mathbb{N} : \ [n \geq m, \ h \in \mathbb{N}] \Rightarrow d(x_{n+h}, x_n) \leq \varepsilon.$$

Toda sucesión convergente es de Cauchy, pero el recíproco no es cierto. Aquellos espacios métricos que verifican que toda sucesión de Cauchy es convergente se llaman completos. Un espacio normado y completo para la métrica asociada a la norma es un espacio de Banach.

Normas equivalentes. Dos normas $\|\cdot\|$ y $|||\cdot|||$ en un mismo espacio vectorial X se dicen equivalentes si existen constantes $m, M > 0$ verificando

$$m\|x\| \leq |||x||| \leq M\|x\|, \ \forall x \in X,$$

equivalentemente, (Proposición 2.20) si ambas normas generan la misma topología.

Teorema de Hausdorff. *Todas las normas en $\mathbb{R}^N$ son equivalentes.*

Como consecuencia, existe una única topología en $\mathbb{R}^N$ que proceda de una norma, a la que llamamos topología de la norma en $\mathbb{R}^N$. Los abiertos en $\mathbb{R}^N$ son las uniones de bolas abiertas para cualquier norma.

Teorema de complitud. *En $\mathbb{R}^N$, toda sucesión de Cauchy es convergente, esto es, $\mathbb{R}^N$ es un espacio de Banach con cualquier norma.*

Conjunto acotado. Un subconjunto A de un espacio métrico (E, d) se dice acotado si $A \subset B(x_0, M)$ para convenientes $x_0 \in E$, $M > 0$.

Toda sucesión de Cauchy es acotada.

Teorema de Bolzano-Weierstrass. *$\mathbb{R}^N$ verifica la propiedad de Bolzano-Weierstrass, es decir, toda sucesión acotada en $\mathbb{R}^N$ admite una parcial convergente.*

Compactos. En $\mathbb{R}^N$ los subconjuntos cerrados y acotados se llaman compactos.

Los compactos de $\mathbb{R}^N$ se caracterizan como aquellos subconjuntos K en los que toda sucesión de puntos de K admite una sucesión parcial convergente a un punto de K (Teorema 2.28). Esta caracterización es la que se toma como definición de compacto en un espacio métrico.

En un espacio métrico el **Teorema de Heine-Borel-Lebesgue** afirma que un subconjunto K es compacto si, y sólo si, verifica el axioma de Heine-Borel: todo recubrimiento por abiertos de K admite un subrecubrimiento finito, esto es, si $\mathscr{U}$ es una familia de abiertos de E tales que $K \subset \bigcup_{U \in \mathscr{U}} U$, entonces existen $n \in \mathbb{N}$ y $U_1, \cdots, U_n \in \mathscr{U}$ tal que $K \subset \bigcup_{i=1}^n U_i$. Esta caracterización es la que se toma como definición de compacto en un espacio topológico.

En general, si A es un subconjunto de un espacio métrico (E, d), los abiertos de A o abiertos relativos son las intersecciones de los abiertos de E con A. A dicha topología en A se le llama la topología inducida.

Nota: La compacidad de un subconjunto K de un espacio topológico es una propiedad intrínseca del espacio topológico $(K, \text{topología inducida})$.

Convexos. Un subconjunto A de $\mathbb{R}^N$ es convexo si

$$[a, b] := \{a + t(b - a) : t \in [0, 1]\} \subset A, \ \forall a, b \in A.$$

Conexos. Un subconjunto C de $\mathbb{R}^N$ es conexo si verifica que la única partición de C en dos abiertos relativos es la trivial. Todo convexo de $\mathbb{R}^N$ es conexo.

Función continua. Sean $(E,d),(F,\rho)$ espacios métricos, $A \subset E$, $f : A \to F$ y $a \in A$. Se dice que la función f es continua en el punto a si

$$\left[\forall \{a_n\} \text{ en } A, \{a_n\} \xrightarrow{d} a\right] \Rightarrow \{f(a_n)\} \xrightarrow{\rho} f(a)$$

El concepto de continuidad entre espacios métricos es topológico.

El estudio de la continuidad de los campos vectoriales se reduce al de sus campos escalares componentes: Si $f = (f_1, \dots, f_M) : A \subset \mathbb{R}^N \to \mathbb{R}^M$ es un campo vectorial y a es un punto de A, entonces

$$f \text{ es continuo en } a \Leftrightarrow f_i \text{ es continuo en } a, \ \forall i = 1,...,M.$$

Regla de la cadena para funciones continuas. *Sean E_1, E_2, E_3 espacios métricos, $A \subset E_1$, $f : A \to E_2$, $B \subset E_2$, $g : B \to E_3$ y supongamos que $f(A) \subset B$. Si f es continua en un punto a de A y g es continua en $f(a)$, entonces la composición $g \circ f$ es continua en a. Como consecuencia, si f y g son continuas, entonces $g \circ f$ es continua.*

Álgebra de las funciones continuas. *Sean (E,d) un espacio métrico, $(X, \|.\|)$ un espacio normado y $a \in A \subset E$.*

i) Si $f, g : A \to X$ son continuas en a y $\lambda \in \mathbb{R}$, entonces $f + g$ y λf son continuas en a.

ii) Si $f : A \to \mathbb{R}$ y $g : A \to X$ son continuas en a, entonces fg es continua en a.

iii) Si $f : A \to \mathbb{R}$ es continua en a y $f(x) \neq 0, \ \forall x \in A$, entonces $\dfrac{1}{f}$ es continua en a.

iv) Si $f, g : A \to \mathbb{R}$ son continuas en a y $g(x) \neq 0, \ \forall x \in A$, entonces $\dfrac{f}{g}$ es continua en a.

Carácter local de la continuidad. *Sean E, F espacios métricos, $A \subset E$, $f : A \to F$ y $b \in B \subset A$. Si B es un "entorno relativo" de b (existe $r > 0$ tal que $B(b,r) \cap A \subset B$) y $f_{|B}$ es continua en b, entonces f es continua en b. En particular, si $f : E \to F$, $B \subset E$ es abierto y si $f_{|B}$ es continua, entonces f es continua en B.*

$\varepsilon - \delta$-caracterización de la continuidad. *Sean (E,d), (F,ρ) espacios métricos, $A \subset E$, $f : A \to F$ y $a \in A$. Equivalen las siguientes afirmaciones:*

i) f es continua en a.

ii) $\forall \varepsilon > 0, \ \exists \delta > 0 : \left\{ \begin{array}{c} x \in A \\ d(x,a) < \delta \end{array} \right\} \Rightarrow \ \rho(f(x), f(a)) < \varepsilon$, *equivalentemente,*

$$\forall \varepsilon > 0, \ \exists \delta > 0 : \ f(B(a,\delta) \cap A) \subset B(f(a), \varepsilon).$$

Como consecuencia de la caracterización de la continuidad global (Proposición 2.47) se tiene:

Sean (E,d) un espacio métrico y $f: E \to \mathbb{R}$ una aplicación continua. Denotemos por $\mathfrak{I}$ (resp. $\mathscr{F}$) la familia de los abiertos (resp. cerrados) de E. Entonces para todo $a \in \mathbb{R}$ se tiene que:

i) $\{x \in E : f(x) < a\} \in \mathfrak{I}$.

ii) $\{x \in E : f(x) > a\} \in \mathfrak{I}$.

iii) $\{x \in E : f(x) \leq a\} \in \mathscr{F}$.

iv) $\{x \in E : f(x) \geq a\} \in \mathscr{F}$.

v) $\{x \in E : f(x) = a\} \in \mathscr{F}$.

Los compactos se conservan por funciones continuas (Teorema 2.50) y, como consecuencia las funciones continuas en un compacto valuadas en $\mathbb{R}$ alcanzan su mínimo y su máximo absolutos (Propiedad de compacidad: Corolario 2.51). Asimismo, los conexos se conservan por funciones continuas (Teorema 2.54).

Teorema de Dini. *Sean K un subconjunto compacto de un espacio métrico y $\{f_n\}$ una sucesión monótona de funciones continuas de K en $\mathbb{R}$ que converge puntualmente en K a una función continua f. Entonces la convergencia es uniforme en K.*

Continuidad uniforme. Sean (E,d) y (F,ρ) espacios métricos, $A \subset E$ y $f: A \to F$ una función. Se dice que f es <u>uniformemente continua</u> si

$$\forall \varepsilon > 0,\ \exists \delta > 0 :\ [x,y \in A,\ d(x,y) < \delta] \Rightarrow \rho(f(x), f(y)) < \varepsilon$$

Teorema de Heine. *Sean E y F espacios métricos, $K \subset E$ compacto y $f: K \to F$ una función continua. Entonces f es uniformemente continua.*

Límite funcional. Sean (E,d) y (F,ρ) espacios métricos, $A \subset E$, $f: A \to F$, α un punto de acumulación de A y $\ell \in F$. Se dice que ℓ es el límite de f en α (o que f tiende a ℓ cuando x tiende a α, y se nota $f(x) \to \ell$ cuando $x \to \alpha$), si

$$[\,\forall \{a_n\} \text{ en } A,\ a_n \neq \alpha,\ \{a_n\} \xrightarrow{d} \alpha\,] \Rightarrow \{f(a_n)\} \xrightarrow{\rho} \ell.$$

El concepto de límite funcional entre espacios métricos es también topológico.

Reducción a campos escalares. *Sean M,N naturales, $A \subset \mathbb{R}^N$, $f = (f_1, \ldots, f_M): A \to \mathbb{R}^M$ un campo vectorial en A y α un punto de acumulación de A. Entonces*

$$f \text{ tiene límite en } \alpha \Leftrightarrow f_i \text{ tiene límite en } \alpha,\ \forall i = 1, \ldots, M.$$

En tal caso,

$$\lim_{x\to\alpha} f(x) = (\lim_{x\to\alpha} f_1(x), \ldots, \lim_{x\to\alpha} f_M(x)).$$

Álgebra de límites. *Sean $A \subset \mathbb{R}^N$, α un punto de acumulación de A, $f, g : A \to \mathbb{R}$ y $\lambda \in \mathbb{R}$. Se verifica:*

i) Si f, g tienen límite en α, entonces $f+g$ y λf tienen límite en α y

$$\lim_{x\to\alpha}(f+g)(x) = \lim_{x\to\alpha} f(x) + \lim_{x\to\alpha} g(x),$$

$$\lim_{x\to\alpha}(\lambda f)(x) = \lambda \lim_{x\to\alpha} f(x).$$

ii) Si f, g tienen límite en α, entonces fg tiene límite en α y

$$\lim_{x\to\alpha}(fg)(x) = \lim_{x\to\alpha} f(x) \lim_{x\to\alpha} g(x).$$

iii) Si f, g tienen límite en α, $g(x) \neq 0\ \forall x \in A$, y $\lim_{x\to\alpha} g(x) \neq 0$, entonces $\frac{f}{g}$ tiene límite en α y

$$\lim_{x\to\alpha} \frac{f}{g}(x) = \frac{\lim_{x\to\alpha} f(x)}{\lim_{x\to\alpha} g(x)}.$$

Relación entre límite funcional y continuidad. *Sean E, F espacios métricos, $A \subset E$, $f : A \to F$ y a un punto de A.*

i) Si a es un punto aislado de A, entonces f es continua en a.

ii) Si a es un punto de acumulación de A, entonces f es continua en a si, y sólo si, $\lim_{x\to a} f(x) = f(a)$.

Coordenadas polares. *Existe una biyección de $\mathbb{R}^+ \times]-\pi,\pi]$ sobre $\mathbb{R}^2 \setminus \{(0,0)\}$. Esto es, dado $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$ existe un único $(\rho,\vartheta) \in \mathbb{R}^+ \times]-\pi,\pi]$ tal que*

$$x = \rho\ cos\ \vartheta\ , y = \rho\ \text{sen}\ \vartheta.$$

Al par (ρ,ϑ) se le llama coordenadas polares del punto (x,y).

Proposición (coordenadas polares). *Sean $f : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}$ una función y ℓ un número real. Equivalen las siguientes afirmaciones:*

i) $\lim_{(x,y)\to(0,0)} f(x,y) = \ell$.

ii) $\forall \varepsilon > 0, \exists \delta > 0 : 0 < \|(x,y)\| < \delta \Rightarrow |f(x,y) - \ell| < \varepsilon$.

iii) $\lim_{\rho\to 0} f(\rho\ cos\ \vartheta, \rho\ \text{sen}\ \vartheta) = \ell$ *uniformemente en ϑ, es decir:*

$$\forall \varepsilon > 0,\ \exists \delta > 0 :\ [\,0 < \rho < \delta,\ \vartheta \in \mathbb{R}\,] \Rightarrow |f(\rho\ cos\ \vartheta, \rho\ \text{sen}\ \vartheta) - \ell| < \varepsilon.$$

iv) *Para cualquier sucesión* $\{\rho_n\}$ *de positivos convergente a cero y cualquier sucesión de reales* $\{\vartheta_n\}$ *(si queremos* $|\vartheta_n| \leq \pi$*), se verifica que*

$$\{f(\rho_n\, cos\ \vartheta_n, \rho_n \text{ sen } \vartheta_n)\} \to \ell.$$

Límites direccionales. *Sea* $f : \mathbb{R}^2 \setminus \{(a,b)\} \to \mathbb{R}$. *Si existe*

$$\lim_{(x,y)\to(a,b)} f(x,y) = \ell,$$

entonces, para cada $\vartheta \in]-\pi,\pi]$, *existe*

$$\lim_{\rho\to 0} f(a+\rho\ cos\ \vartheta, b+\rho \text{ sen } \vartheta) = \ell.$$

Los anteriores límites se llaman <u>*límites direccionales*</u> *en la dirección* ϑ.

En consecuencia, de existir límite han de existir todos los límites direccionales y ser iguales al límite.

En la práctica es usual para calcular los límites direccionales de $f(x,y)$ *en el punto* (a,b), *considerar las rectas* $y = b + m(x-a)$ *(de pendiente* m *y que pasa por* (a,b)*) y hallar el límite*

$$\lim_{x\to a} f(x, b+m(x-a)).$$

2.9. Ejercicios del Tema 2.

Los problemas en los que aparezca el símbolo * se suponen conocidos. En clase nos dedicaremos a los demás.

Se recomienda (con las debidas precauciones) el uso de Mathematica (o programas similares) para dibujar las funciones que aparezcan en los problemas, especialmente en los que se pretende estudiar la convergencia puntual y uniforme, y en los problemas de cálculo de límites.

2.1 Probar que en el espacio vectorial $\mathscr{C}[a,b]$ de las funciones reales continuas definidas en el intervalo $[a,b]$ la aplicación

$$\|f\|_\infty := \max\ \{|f(x)| : x \in [a,b]\}$$

es una norma.

2.2 Probar que la sucesión de funciones $\{f_n\}$ definidas por

$$f_n(x) := x^n,\ \forall x \in [0,1],\ \forall n \in \mathbb{N}$$

converge puntualmente en $[0,1]$ pero no converge uniformemente en $[0,1[$. Pruébese que, de hecho, $\{f_n\}$ no admite parciales convergentes en $\mathscr{C}[0,1]$. Sin embargo, $\{f_n\}$ converge uniformemente en $[0,\alpha]$ con $0 < \alpha < 1$.

Indicación: Úsese la monotonía de $\{f_n\}$ para probar que no hay parciales convergentes.

2.3 Consideremos el espacio vectorial $\mathscr{C}[0,2]$ de las funciones continuas definidas en $[0,2]$. Probar que

i) $(\mathscr{C}[0,2], \|.\|_\infty)$ es un espacio de Banach.

ii) $(\mathscr{C}[0,2], \|.\|_1)$ es un espacio normado no completo, donde $\|f\|_1 := \int_0^2 |f(x)|dx$.

iii) ¿Son equivalentes las anteriores normas?

Indicación:

i) $\{f_n\}$ $\|.\|_\infty$-Cauchy $\Rightarrow$ $\{f_n(a)\}$ es de Cauchy para todo $a \in [0,2]$. El Teorema de complitud nos sugiere la siguiente candidata a límite:

$$f(a) := \lim\ f_n(a), \quad (a \in [0,2]).$$

Utilizando la condición de Cauchy "a tope" pruébese que f es acotada y que $\{f_n\}$ converge uniformemente a f. El Teorema de conservación de la continuidad por convergencia uniforme asegura que f es continua.

ii) Pruébese que si $g(a) \neq 0$, entonces, usando la continuidad de g en a, existe $\delta > 0$ tal que $\int_0^2 |g(x)|dx \geq \dfrac{|g(a)|}{2}\delta$.

Sea $\{f_n\}$ la sucesión decreciente de funciones continuas dada por

$$f_n(x) := \begin{cases} x^n & \text{si } x \in [0,1] \\ 1 & \text{si } x \in [1,2] \end{cases} \quad \forall n \in \mathbb{N}.$$

Probar que $\{f_n\}$ es $\|.\|_1$-Cauchy utilizando que

$$\int_0^2 |f_{n+h} - f_n| \leq \int_0^1 f_n = \frac{1}{n+1}.$$

Para demostrar que dicha sucesión no es convergente en $(\mathscr{C}[0,2], \|.\|_1)$, obsérvese que si f fuese la función límite, entonces para cada $0 < a < 1$ se tendría

$$0 \leq \int_0^a |f| \leq \int_0^a |f_n| + \int_0^a |f_n - f| \leq a f_n(a) + \|f_n - f\|_1 \to 0,$$

de donde se deduce que $f(a) = 0$ y por continuidad $f(1) = 0$.

Por otra parte

$$0 \leq \int_1^2 |1 - f| = \int_1^2 |f_n - f| \leq \|f_n - f\|_1 \to 0$$

obliga a que $f(1) = 1$.

iii) Para cada natural n basta considerar la función

$$f_n(x) := \begin{cases} -nx + 1 & \text{si } x \in [0, \frac{1}{n}] \\ 0 & \text{si } x \in [\frac{1}{n}, 2] \end{cases}$$

También se puede usar que la norma $\|\cdot\|_\infty$ es completa y $\|\cdot\|_1$ no lo es.

2.4 Probar que la sucesión de funciones

$$f_n(x) := \left(1 + \frac{x}{n}\right)^n, \ \forall x \in \mathbb{R}, \ \forall n \in \mathbb{N}$$

y

$$g_n(x) := n\left(\sqrt[n]{x} - 1\right), \ \forall x \in \mathbb{R}^+, \ \forall n \in \mathbb{N}$$

convergen puntualmente en su dominio pero no uniformemente.

Probar también que convergen uniformemente en todo intervalo compacto.
Indicación: Úsese la desigualdad entre la media geométrica y aritmética para probar la monotonía de la sucesión y aplíquese el Teorema de Dini. En el segundo caso puede calcularse $\|g_n\|_\infty$ (en acotados) derivando.

2.5 a) Constrúyase una sucesión $\{f_n\}$ de funciones en la esfera unidad de $\mathscr{C}[0,1]$ tal que $\|f_p - f_q\|_\infty = 1, \ \forall p \neq q$.

Indicación: Tómese, por ejemplo,

$$f_1(x) := \begin{cases} 1 & \text{si } x \in [0, \frac{1}{2}] \\ -2(x-1) & \text{si } x \in [\frac{1}{2}, 1] \end{cases}, f_2(x) := \begin{cases} 1 & \text{si } x \in [0, \frac{1}{2^2}] \\ -2^2(x - \frac{1}{2}) & \text{si } x \in [\frac{1}{2^2}, \frac{1}{2}] \\ 0 & \text{si } x \in [\frac{1}{2}, 1]. \end{cases}$$

b) ¿Es compacta la bola unidad cerrada de $\mathscr{C}[0,1]$?

2.6 Sea A un subconjunto de un espacio de Banach. Probar que equivalen las siguientes afirmaciones.

i) A es completo.

ii) A es cerrado.

2.7 Probar que el cierre de un subespacio de un espacio normado es un subespacio y que los subespacios de $\mathbb{R}^N$ son cerrados.

Indicación. Véase el Ejercicio 2.6.

2.8 Pruébese que no existe más espacio normado de dimensión N que $\mathbb{R}^N$ con una conveniente norma. Es decir, si $(X, \|.\|)$ es un espacio normado de dimensión N, entonces existe una norma $|||\,.\,|||$ en $\mathbb{R}^N$ y una biyección lineal

$$f : (X, \|.\|) \to (\mathbb{R}^N, |||\,.\,|||)$$

isométrica (i.e., $|||\,f(x)\,||| = \|x\|$ para todo x en X). Dedúzcase que en todo espacio finito-dimensional hay una única topología que proceda de una norma.

Indicación: Sea $f : X \to \mathbb{R}^N$ una biyección lineal (¡descríbase!). Basta comprobar que $|||\,f(x)\,||| := \|x\|$ es una norma en $\mathbb{R}^N$.

2.9 Sea N un número natural y sean $\alpha_0, \alpha_1, ..., \alpha_N$ números reales distintos. En el espacio vectorial X de todas las funciones polinómicas de grado menor o igual que N definimos:

$$\|f\| := \sum_{k=0}^{N} |f(\alpha_k)|, \quad (f \in X).$$

Probar que

i) $\|\cdot\|$ es una norma en X.

ii) La topología que genera esta norma no depende de la elección de los reales α_k.

iii) Una sucesión $\{p_n\}$ en X converge uniformemente en un intervalo $[a,b]$ si, y sólo si, existen $N+1$ números reales distintos del intervalo $[a,b]$, $\beta_0, \beta_1, ..., \beta_N$, tales que $\{p_n(\beta_k)\}$ converge para $k = 0, 1, ..., N$.

Indicación. Úsese que si N es un número natural, $x_0, ..., x_N$ son números reales distintos y $y_0, ..., y_N$ son números reales cualesquiera, existe un único polinomio p de grado menor o igual que N tal que

$$y_i := p(x_i) \text{ para } i = 0, 1, ..., N.$$

De hecho el polinomio p viene dado por la expresión

$$p(x) = \sum_{i=0}^{N} y_i \frac{\Pi_{r\neq i}(x-x_r)}{\Pi_{r\neq i}(x_i-x_r)}.$$

2.10 Sean M un subespacio vectorial de $(\mathbb{R}^N, \|\cdot\|)$ y $x_0 \in \mathbb{R}^N$. Probar que existe $m_0 \in M$ que materializa la distancia de x_0 a M, es decir,

$$\|x_0 - m_0\| = \text{dist}\,(x_0, M) := \inf\{\|x_0 - m\| : m \in M\}.$$

(Se dice que los subespacios de $\mathbb{R}^N$ son *proximinales*, por verificarse lo anterior).

Indicación: Justificar la existencia de una sucesión acotada $\{m_n\}$ en M tal que

$$\{\|x_0 - m_n\|\} \to \text{dist}\,(x_0, M).$$

2.11 Sea M un subespacio propio de $\mathbb{R}^N$. Probar que existe un vector x de la esfera unidad tal que $\text{dist}(x,M) = 1$.

Indicación: Sea $x_0 \in \mathbb{R}^N \setminus M$ y sea $m_0 \in M$ tal que $\|x_0 - m_0\| = dist\,(x_0, M)$. Considerar el vector normalizado de $x_0 - m_0$.

2.12 Probar que todo espacio normado X es *homeomorfo* a su bola abierta $B(0,1)$, es decir, existe una biyección bicontinua (continua con inversa continua) de X sobre $B(0,1)$.

Indicación: Véase el Ejercicio 1.5.

2.13 Sean E,F espacios métricos, $A \subset E$ y $f : A \to F$ una función. Probar que si f conserva las sucesiones convergentes, entonces f es continua.

2.14 Sea (E,d) un espacio métrico. Pruébese que

i) La distancia $d : E \times E \longrightarrow \mathbb{R}$ es continua.

ii) Si $A \subset E$ es un subconjunto no vacío, definimos

$$\text{dist}(x,A) = \inf\{d(x,a) :\in A\} \quad (x \in E).$$

Pruébese que la función $x \mapsto d(x,A)$ es continua en E. De hecho, la afirmación anterior es consecuencia de la siguiente desigualdad:

$$|\text{dist}(x,A) - \text{dist}(y,A)| \leq d(x,y),\ \forall x,y \in E.$$

2.15 Sea A un subconjunto no vacío de un espacio métrico (E,d). Se define el diámetro de A como

$$\text{diam }(A) = \sup\{d(x,y) :\ x,y \in A\} \in \mathbb{R}_0^+ \cup \{+\infty\}.$$

Pruébese que:

1. $\text{diam }(\overline{A}) = \text{ diam }(A)$.
2. $\overset{\circ}{A} = \{x \in E :\ \text{dist}(x, E \setminus A) > 0\}$ y $\overline{A} = \{x \in E :\ \text{dist}(x,A) = 0\}$. Deducir que en un espacio métrico todo conjunto abierto se puede expresar como una unión numerable de conjuntos cerrados y que cada conjunto cerrado se puede expresar como una intersección numerable de abiertos.
3. Si x es un vector de un espacio normado y $r > 0$, entonces $\overline{B(x,r)} = \overline{B}(x,r)$ y $\text{diam }(B(x,r)) = 2r$. ¿Son ciertas las anteriores igualdades para un espacio métrico cualquiera?

Indicación: Usar el Corolario 2.48 y el Ejercicio 2.14.

2.16 **Lema de Uryshon para espacios métricos.**

Sean A y B subconjuntos no vacíos, cerrados y disjuntos de un espacio métrico E. Probar que existe una aplicación continua $f : E \to \mathbb{R}$ verificando:

$$0 \le f \le 1 \ ,\ f(A) = \{0\} \ ,\ f(B) = \{1\}.$$

Deducir que existen subconjuntos abiertos U,V de E tales que:

$$A \subset U \ ,\ B \subset V \ ,\ U \cap V = \emptyset.$$

Indicación: Considérese la función

$$f(x) = \frac{\text{dist}(x,A)}{\text{dist}(x,A) + \text{dist}(x,B)} \qquad (x \in E).$$

Para probar la continuidad de esta función, úsese el Ejercicio 2.14.

2.17 Decimos que una familia de subconjuntos de un conjunto tiene la propiedad de la intersección finita si la intersección de cualquier subfamilia finita tiene intersección no vacía. Probar que un espacio topológico K es compacto si, y sólo si, toda familia $\{F_i\}_{i\in I}$ de cerrados de K tal que $\cap_{i\in I}(F_i) = \emptyset$ no verifica la propiedad de la intersección finita.

2.18 Pruébese que una función real $f \in \mathscr{C}^1[a,b]$ es uniformemente continua. De hecho, existe una constante $M \ge 0$ tal que

$$|f(y) - f(x)| \le M|y-x|,\ \forall x,y \in [a,b].$$

Indicación: Teorema del valor medio.

2.19 Sea A un conjunto no cerrado de un espacio métrico (E,d). Probar que existe una aplicación continua $f: A \to \mathbb{R}$ que no es uniformemente continua.

Indicación: Fijado $x_0 \in \overline{A} \setminus A$, considerar la función

$$f(a) := \frac{1}{d(a,x_0)}, \quad (a \in A)$$

y utilizar el Ejercicio 2.14.

2.20 Sea $f : \mathbb{R}^* \times \mathbb{R}^* \to \mathbb{R}$ la función definida por:

$$f(x,y) = (x+y) \operatorname{sen} \frac{\pi}{x} \operatorname{sen} \frac{\pi}{y}.$$

i) Estudiar la existencia de límite de la función f en los puntos $(0,0)$, $(0,1)$ y $(0,\pi)$.

ii) ¿Es f uniformemente continua?

2.21 **Límite a lo largo de una curva**.

Sean E,F espacios métricos, A un subconjunto de E, α un punto de acumulación de A y $f : A \to F$. Sea $\gamma : [0,1] \to E$ una aplicación continua verificando que $\gamma(]0,1[) \subset A \setminus \{\alpha\}$ y $\gamma(0) = \alpha$. Probar que si existe el límite $\lim_{x\to\alpha} f(x)$, entonces existe el límite $\lim_{t\to 0} f(\gamma(t))$ y además

$$\lim_{x\to\alpha} f(x) = \lim_{t\to 0} f(\gamma(t)).$$

2.22 **Límites iterados**.

Sean I,J intervalos de $\mathbb{R}$ que tienen a 0 como punto de acumulación, $A = I \times J \setminus \{(0,0)\}$ y $f : A \to \mathbb{R}$. Consideremos las siguientes afirmaciones:

$a_1)$ $[x \in I \setminus \{0\} \Rightarrow \exists \lim_{y\to 0} f(x,y) := f_1(x)$

$a_2)$ $[y \in J \setminus \{0\} \Rightarrow \exists \lim_{x\to 0} f(x,y) := f_2(y)$

$b)$ $\exists \lim_{(x,y)\to(0,0)} f(x,y) = \ell$.

Probar que $a_i)$ y $b)$ implican que f_i tiene límite ℓ en 0. Los límites

$$\lim_{x\to 0}(\lim_{y\to 0} f(x,y)), \quad \lim_{y\to 0}(\lim_{x\to 0} f(x,y)),$$

cuando existen, se llaman *límites iterados* de f en $(0,0)$.

2.23 * Estudiar la existencia de los límites iterados y del límite en $(0,0)$ de las siguientes funciones definidas en $\mathbb{R}^2 \setminus \{(0,0)\}$:

$$f(x,y) = \frac{x^2-y^2}{x^2+y^2},\ g(x,y) = \frac{y^3}{(x^2+y^2)^{\frac{3}{2}}};\ w(x,y) = \begin{cases} y \operatorname{sen} \frac{1}{x} & \text{si} \quad x \neq 0 \\ 0 & \text{si} \quad x = 0 \end{cases};$$

$$u(x,y) = \frac{x^2y^2}{x^2+y^2};\ v(x,y) = \begin{cases} x\frac{|y|}{y} & \text{si} \quad y \neq 0 \\ 0 & \text{si} \quad y = 0 \end{cases};\ h(x,y) = \frac{xy}{x^2+y^2}.$$

2.24 * Estudiar la existencia de límite en $(0,0)$ de las siguientes funciones reales definidas, en cada caso, en un subconjunto apropiado de $\mathbb{R}^2$:

$$f(x,y) = \frac{xy^2}{x^2+y^4};\ g(x,y) = \frac{x^2y}{(x^2+y^2)^{\frac{3}{2}}};\ h(x,y) = \left(\frac{1}{x^2}+\frac{1}{y^2}\right)\text{sen }(xy)$$

$$u(x,y) = \cot\ x\ \text{sen}\ y;\ \ v(x,y) = \frac{x^3+y^3}{x^2+y^2}.$$

2.25 * Estudiar la continuidad de la función $f : \mathbb{R}^2 \to \mathbb{R}$ dada por

$$f(x,y) = \begin{cases} x^2 \text{ si } |x| \le |y| \\ y^2 \text{ si } |x| > |y| \end{cases}.$$

Indicación. Obsérvese que min $\{x^2, y^2\} = \frac{1}{2}[x^2+y^2-|x^2-y^2|]$. También puede probarse que f es continua en los puntos (x,y) tales que $|x| = |y|$ y usar el carácter local de la continuidad.

2.26 * Estudiar la continuidad de la función $f : \mathbb{R}^2 \to \mathbb{R}^3$ definida por

$$f(x,y) = \left(e^{x^2-y^2}\ ,\ \frac{\log(1+|x|)}{\sqrt{1+x^2+y^4}}\ ,\ \text{sen}\ y\ \cos\ x\right).$$

2.27 * Estudiar la existencia de límite en $(0,0)$ de las siguientes funciones definidas en $\mathbb{R}^+ \times \mathbb{R}^+$

$$f(x,y) = \frac{\text{sen}\ x + \text{sen}\ y}{x+y}\ ;\ g(x,y) = \frac{x}{y}\ \text{sen}\ (x^2+y^2)\ ;$$

$$h(x,y) = \frac{xe^y+ye^x}{x+y}\ ;\ \phi(x,y) = x\ E\left(\frac{y}{x}\right).$$

2.28 * Estudiar para qué números positivos α la función

$$f(x,y) = \frac{\text{sen}\ (xy)}{(x^2+y^2)^\alpha},\ \forall (x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$$

tiene límite en $(0,0)$.

2.29 * Demostrar que

$$\lim_{(x,y)\to(0,b)} \frac{e^x-e^{yx}}{x} = 1-b\ ,\ \ \lim_{(x,y)\to(0,0)} \frac{1-cos\ x\ cos\ y}{\sqrt{x^2+y^2}} = 0.$$

2.30 * Estudiar la existencia de límite en $(0,0)$ de las siguientes funciones reales definidas, en cada caso, en un subconjunto apropiado de $\mathbb{R}^2$:

$$f(x,y) = \frac{x^4+3x^2y^2+2xy^3}{(x^2+y^2)^2}; \;\; g(x,y) = \frac{2x^5+2y^3(2x^2-y^2)}{(x^2+y^2)^2};$$

$$h(x,y) = \frac{y(x^2+y^2)}{x}; \;\; u(x,y) = \frac{x^2+y^2}{\sqrt{x^2+y^2+1}-1}.$$

2.31 * Sea $\mathbb{K}$ un cuerpo un cuerpo conmutativo con una relación de orden $\leq$ verificando

i) $a,b \in \mathbb{R} \Rightarrow a \leq b$ o $b \leq a$ (orden total),

ii) $[a,b,c \in \mathbb{R},\ a \leq b] \Rightarrow\ a+c \leq b+c$ (compatible con la suma), y

iii) $[a,b,c \in \mathbb{R},\ a \leq b,\ 0 \leq c] \Rightarrow\ ac \leq bc$ (compatible con el producto)

Se dota a $\mathbb{K}$ de la topología del orden (los abiertos de $\mathbb{K}$ son las uniones de intervalos abiertos). Probar que las siguientes afirmaciones son equivalentes

i) $\mathbb{K}$ verifica el axioma del supremo.

ii) $\mathbb{K}$ verifica el axioma de Bolzano (Toda función continua definida en un intervalo que toma valores positivos y negativos se anula).

iii) $\mathbb{K}$ es conexo.

Como consecuencia $\mathbb{R}$ es el único cuerpo ordenado que verifica el axioma de Bolzano, y también es el único cuerpo ordenado conexo

2.10. Soluciones a los ejercicios del Tema 2.

2.1 Lo único no rutinario es la propiedad triangular que se deduce de

$$|(f+g)(x)| \leq |f(x)| + |g(x)| \leq \|f\|_\infty + \|g\|_\infty, \ \forall x \in [a,b].$$

2.2 La sucesión $\{f_n\}$ converge puntualmente a la función

$$f:[0,1] \to \mathbb{R}, \ f(x) := \begin{cases} 0 & \text{si } x<1 \\ 1 & \text{si } x=1 \end{cases} .$$

$(f_n - f)(1-\frac{1}{n}) = (1-\frac{1}{n})^n \to e^{-1}$, luego no hay convergencia uniforme en $[0,1[$.
Por ser f_n monótona, si admitiese una parcial convergente, ella sería convergente en contra de lo probado.

Para probar la convergencia uniforme en $[0,\alpha]$, dado $\varepsilon > 0$ tómese $m = \dfrac{log\ \varepsilon}{log\ \alpha}$.

2.3 Seguir las indicaciones.

2.4 No hay convergencia uniforme en $\mathbb{R}$, pues

$$|f_n(n) - f(n)| = |2^n - e^n| \to +\infty, \quad |f_n(-2n) - f(-2n)| = |(-1)^n - e^{-2n}|.$$

Nótese que la sucesión $\{|(-1)^n - e^{-2n}|\}$ no tiene límite.
Además se tiene $|g_n(n^n) - g(n^n)| = |n(n-1-log\ n)| \to +\infty$,
$|g_n(n^{-n}) - g(n^{-n})| = |1 + n(\log(n)-1)| \to +\infty$.

Prueba de la monotonía: Dado $x \in [-a,a]$, para $n > a$ se tiene que $\left|\dfrac{x}{n}\right| < 1$, por tanto, usando la desigualdad de las medias para

$$a_1 = \ldots = a_n = 1 + \frac{x}{n}, \ \ a_{n+1} = 1,$$

se obtiene

$$\sqrt[n+1]{\left(1+\frac{x}{n}\right)^n} \leq \frac{n(1+\frac{x}{n})+1}{n+1} \leq 1 + \frac{x}{n+1},$$

esto es, para $n > a$, se tiene

$$f_n(x) \leq f_{n+1}(x), \ \ \forall x \in [-a,a].$$

Basta ahora aplicar el Teorema de Dini a la sucesión de funciones anterior. Para el caso de la segunda sucesión de funciones, se usa la misma técnica, obteniéndose para cada $x \in \mathbb{R}^+$

$$g_{n+1}(x) \leq g_n(x) \Leftrightarrow \sqrt[n+1]{x} \leq \frac{n\sqrt[n]{x}+1}{n+1},$$

expresiones que se deducen de la desigualdad entre la media geométrica y la aritmética sin más que observar que

$$\left(1+\frac{x}{n}\right)^n = \left(1+\frac{x}{n}\right)^n \cdot 1 \quad \text{y} \quad x = \left(\sqrt[n]{x}\right)^n \cdot 1 \ .$$

2.5 a)

$$f_n(x) := \begin{cases} 1 & \text{si } x \in [0, \frac{1}{2^n}] \\ -2^n(x - \frac{1}{2^{n-1}}) & \text{si } x \in [\frac{1}{2^n}, \frac{1}{2^{n-1}}] \\ 0 & \text{si } x \in [\frac{1}{2^{n-1}}, 1] \end{cases}$$

$\|f_n\|_\infty = f_n(0) = 1 \ \ , \ \ \|f_{n+h} - f_n\|_\infty = (f_n - f_{n+h})(\frac{1}{2^{n+h-1}}) = 1.$

2.6 Rutinario. Proposición 2.16.

2.7 Rutinario.

2.8 Seguir las indicaciones y aplicar el Teorema de Hausdorff.

2.9 $i)$ Rutinario.

$ii)$ y $iii)$ Por el ejercicio anterior, todas las normas en X son equivalentes ya que la dimensión de X es $N+1$.

2.10 Por definición de ínfimo, existe una sucesión $\{m_n\}$ en M tal que

$$\{\|x_0 - m_n\|\} \to dist\ (x_0, M).$$

La sucesión $\{m_n\}$ es acotada pues como

$$\exists m \in \mathbb{N} : n \geq m \Rightarrow \|x_0 - m_n\| < dist\ (x_0, M) + 1$$

se tiene que

$$n \geq m \Rightarrow \|m_n\| \leq \|m_n - x_0\| + \|x_0\| < dist\ (x_0, M) + 1 + \|x_0\|.$$

Ahora, por el Teorema de Bolzano-Weierstrass, $\exists\ \{m_{\sigma(n)}\} \to m_0 \in \mathbb{R}^N$. Como M es cerrado, por ser un subespacio finito-dimensional, se tiene que $m_0 \in M$. Finalmente

$$\left.\begin{array}{l} \{\|x_0 - m_{\sigma(n)}\|\} \to dist\ (x_0, M) \\ \\ \{\|x_0 - m_{\sigma(n)}\|\} \to \|x_0 - m_0\| \end{array}\right\} \Rightarrow \|x_0 - m_0\| = dist\ (x_0, M).$$

2.11 Nótese $x := \dfrac{x_0 - m_0}{\|x_0 - m_0\|}$. Se tiene que $dist(x,M) \leq \|x\| = 1$ y para todo $m \in M$ se verifica que

$$\|x - m\| = \left\| \frac{x_0 - m_0}{\|x_0 - m_0\|} - m \right\| = \frac{\|x_0 - (m_0 + \|x_0 - m_0\| m)\|}{\|x_0 - m_0\|} \geq \frac{dist(x_0,M)}{\|x_0 - m_0\|} = 1.$$

2.12 Considérense $f : X \to B(0,1)$ dada por $f(x) := \dfrac{x}{1 + \|x\|}$ y su inversa

$$f^{-1}(y) := \frac{y}{1 - \|y\|}.$$

2.13 Sea $\{a_n\}$ una sucesión en A convergente a un punto $a \in A$. La sucesión

$$\{a_1, a, a_2, a, a_3, a, \ldots\}$$

es claramente convergente hacia a, luego por hipótesis también es convergente la sucesión

$$\{f(a_1), f(a), f(a_2), f(a), f(a_3), f(a), \ldots\}.$$

Como la subsucesión de los términos pares converge a $f(a)$, igual le ocurre a la de los términos impares, es decir, $\{f(a_n)\} \to f(a)$.

2.14 *i)* La continuidad de la aplicación distancia se sigue de:

$$|d(x,y) - d(x_n,y_n)| \leq |d(x,y) - d(y,x_n)| + |d(x_n,y) - d(x_n,y_n)| \leq$$

$$d(x,x_n) + d(y,y_n) \leq 2\, d((x,y),(x_n,y_n)).$$

ii) Nótese que dados elementos $x, y \in E$ y $a \in A$ se verifica:

$$\mathrm{dist}(x,A) \leq d(x,a) \leq d(x,y) + d(y,a),$$

y reorganizando los términos de la desigualdad anterior tenemos

$$\mathrm{dist}(x,A) - d(x,y) \leq d(y,a), \ \forall a \in A.$$

Tomando ínfimo en la expresión anterior se obtiene

$$\mathrm{dist}(x,A) - d(x,y) \leq \mathrm{dist}(y,A),$$

esto es,

$$\mathrm{dist}(x,A) - \mathrm{dist}(y,A) \leq d(x,y).$$

Basta intercambiar los papeles de x e y para obtener

$$|\mathrm{dist}(x,A) - \mathrm{dist}(y,A)| \leq d(x,y),$$

de donde se deduce inmediatamente la continuidad de la función $x \mapsto \mathrm{dist}(x,A)$.

2.15 1. Suponemos que A es acotado, pues en otro caso queda la igualdad $\infty = \infty$. Es claro que $\operatorname{diam}(A) \leq \operatorname{diam}(\overline{A})$. Para probar la otra desigualdad, sean $x, y \in \overline{A}$ y tomemos dos sucesiones $\{x_n\}, \{y_n\}$ en A con $\{x_n\} \to x$, $\{y_n\} \to y$. Para cada natural n, se tiene

$$d(x,y) \leq d(x,x_n) + d(x_n,y_n) + d(y_n,y) \leq d(x,x_n) + \operatorname{diam}(A) + d(y_n,y)$$

y en consecuencia, tomando límites $d(x,y) \leq \operatorname{diam}(A), \forall x, y \in \overline{A}$, equivalentemente $\operatorname{diam}(\overline{A}) \leq \operatorname{diam}(A)$.

2.

$$x \in \overline{A} \iff [\exists \{a_n\} \to x \text{ con } a_n \in A] \iff \operatorname{dist}(x,A) = 0.$$

$$x \in \overset{\circ}{A} \iff x \notin \overline{E \setminus A} \iff \operatorname{dist}(x, E \setminus A) > 0.$$

En consecuencia,

$$\overset{\circ}{A} = \bigcup_{n \in \mathbb{N}} \left\{ x \in E : \operatorname{dist}(x, E \setminus A) \geq \frac{1}{n} \right\}.$$

y

$$\overline{A} = \bigcap_{n \in \mathbb{N}} \left\{ y \in E : \operatorname{dist}(y, A) < \frac{1}{n} \right\}.$$

3. Por ser el cierre de un conjunto el menor cerrado conteniéndolo, se tiene que

$$\overline{B(x,r)} \subset \overline{B}(x,r).$$

Para la otra inclusión, nótese que si $d(x,y) = r$, entonces $y \in \overline{B(x,r)}$ pues

$$\{x + \frac{n}{n+1}(y-x)\} \to y\ ,$$

y para cada natural n es

$$x + \frac{n}{n+1}(y-x) \in B(x,r).$$

Del primer apartado se sigue que

$$\operatorname{diam}(B(x,r)) = \operatorname{diam} \overline{B(x,r)} = \operatorname{diam}(\overline{B}(x,r)) = 2r,$$

ya que si $a, b \in \overline{B}(x,r)$,

$$d(a,b) \leq d(a,x) + d(x,b) \leq r + r = 2r,$$

y por tanto

$$\operatorname{diam}(\overline{B}(x,r)) \leq 2r.$$

La otra desigualdad se sigue de que, dado $s \in S_X$ se tiene que

$$x - rs, x + rs \in \overline{B}(x,r) \quad \text{y} \quad d(x+rs, x-rs) = \|2rs\| = 2r.$$

Con la métrica trivial se tiene que $\operatorname{diam}(B(x,1)) = 0$, y si además E tiene más de un punto $\overline{B(x,1)} = \{x\} \neq E = \overline{B}(x,1)$.

2.16 Dado un subconjunto $\emptyset \neq A$ de un espacio métrico, es fácil comprobar que la función

$$x \mapsto \text{dist}(x,A),$$

es continua (Ejercicio 2.14).

Definimos entonces la función $f : \mathbb{R} \longrightarrow \mathbb{R}$ por

$$f(x) := \frac{\text{dist}(x,A)}{\text{dist}(x,A)+\text{dist}(x,B)}, \ \forall x \in E,$$

que está bien definida pues A y B son cerrados disjuntos. De la continuidad de la aplicación distancia se deduce la continuidad de f. Es inmediato que esta aplicación verifica que $0 \leq f \leq 1$, $f(A) = \{0\}$ y $f(B) = \{1\}$. Por último, sean

$$U := \left\{x \in E : f(x) < \frac{1}{2}\right\}, \ \ V := \left\{x \in E : f(x) > \frac{1}{2}\right\}$$

que son abiertos por la continuidad de f y que claramente verifican $A \subset U$, $B \subset V$ y $U \cap V = \emptyset$.

2.17 Úsese el axioma de Heine-Borel y el paso a complementos.

2.18 Seguir la indicación.

2.19 Sea $x_0 \in \overline{A} \setminus A$. Es claro que $f(a) := \frac{1}{d(a,x_0)}$, $\forall a \in A$, es continua. Sea $\{a_n\}$ una sucesión en A tal que $d(x_0,a_n) < \frac{1}{n}$, $\forall n \in \mathbb{N}$. Se tiene entonces que la sucesión $\{f(a_n)\}$ no está acotada ($n < f(a_n)$, $\forall n \in \mathbb{N}$), lo que impide que f sea uniformemente continua. En efecto, si f fuese uniformemente continua, fijado $\varepsilon > 0$ existiría $\delta > 0$ tal que

$$[x,y \in A, d(x,y) < \delta] \Rightarrow |f(x) - f(y)| < \varepsilon,$$

de donde se deduce que la función f conservaría las sucesiones de Cauchy.

2.20 i) Como $|f(x,y)| \leq \|(x,y)\|_1$ se tiene que $\lim_{(x,y)\to(0,0)} f(x,y) = 0$.

Para $(x,y) \in \mathbb{R}^* \times \mathbb{R}^*$ cerca de $(0,1)$ se tiene

$$|f(x,y)| \leq 2\left|\text{sen } \frac{\pi}{y}\right|,$$

luego $\lim_{(x,y)\to(0,1)} f(x,y) = 0$.

Veamos que f no tiene límite en $(0,\pi)$. Sea $a_n = (\frac{2}{2n+1}, \pi)$, $\forall n \in \mathbb{N}$. Entonces

$$f(a_n) = (\frac{2}{2n+1} + \pi)(-1)^n \text{sen } 1 \to \begin{cases} \pi \text{ sen } 1 & \text{si } n \text{ es par} \\ -\pi \text{ sen } 1 & \text{si } n \text{ es impar} \end{cases}.$$

ii) Para probar que f no es uniformemente continua basta utilizar que hay dos sucesiones que convergen a $(0,\pi)$ cuyas sucesiones imágenes tienen límites distintos. En efecto

$$\{d(a_{2n}, a_{2n-1})\} \to 0 \text{ y } \{|f(a_{2n}) - f(a_{2n-1})|\} \to 2\pi \text{ sen } 1.$$

2.21 De existir límite tienen que existir los límites a lo largo de cualquier curva y ser iguales al límite. La demostración es inmediata por sucesiones:

$$[\{t_n\} \to 0, t_n \neq 0] \Rightarrow [\{\gamma(t_n)\} \to \alpha, \gamma(t_n) \neq \alpha] \Rightarrow \{f(\gamma(t_n)) \to \lim_{x\to\alpha} f(x).$$

2.22 La afirmación de b) equivale a la siguiente condición

$$\forall \varepsilon > 0, \exists \delta > 0 : \left\{ \begin{array}{c} 0 < \|(x,y)\|_\infty < \delta \\ (x,y) \in A \end{array} \right\} \Rightarrow |f(x,y) - \ell| \leq \varepsilon.$$

Supongamos que se verifica a_1. Fijemos $x \in I \backslash \{0\}$ con $|x| < \delta$. Al tomar límite cuando $y \to 0$ en la expresión anterior nos queda $|f_1(x) - \ell| \leq \varepsilon$ o, lo que es igual, $\lim_{x\to 0} f_1(x) = \ell$. Hemos probado que $\lim_{x\to 0}(\lim_{y\to 0} f(x,y)) = \ell$. La demostración del otro caso es análoga.

2.23 f: Como $\lim_{x\to 0}(\lim_{y\to 0} f(x,y)) = 1$ y $\lim_{y\to 0}(\lim_{x\to 0} f(x,y)) = -1$ no existe límite.

g: De $\lim_{x\to 0} g(x,y) = \left(\frac{y}{|y|}\right)^3$ se sigue que no existe $\lim_{y\to 0}(\lim_{x\to 0} g(x,y))$ y por tanto tampoco g tiene límite.

h: Los límites iterados valen cero, pero como $h(t,t) = \frac{1}{2}$ (si $t > 0$), concluimos que no existe límite.

u: Los límites iterados valen cero, luego de existir límite este es cero.

$$|u(x,y)| = \frac{|xy|^2}{x^2+y^2} \leq \frac{1}{4}(x^2+y^2) \to 0 \Rightarrow \lim_{(x,y)\to(0,0)} u(x,y) = 0.$$

v: No existe $\lim_{y\to 0} v(x,y)$. Como $\lim_{y\to 0}(\lim_{x\to 0} v(x,y)) = 0$, de existir límite este es cero. De $|v(x,y)| \leq |x|$ se sigue que $\lim_{(x,y)\to(0,0)} v(x,y) = 0$.

w: No existe $\lim_{x\to 0} w(x,y)$. Como $\lim_{x\to 0}(\lim_{y\to 0} w(x,y)) = 0$, de existir límite este es cero. De $|w(x,y)| \leq |y|$ se sigue que $\lim_{(x,y)\to(0,0)} w(x,y) = 0$.

Sumando las dos últimas funciones obtenemos una función que no tiene límites iterados y sin embargo tiene límite.

2.24 f: Como $f(0,t) = 0$ y $f(t^2,t) = \frac{1}{2}$ no existe límite.

g: Como $g(0,t) = 0$ y $g(t,t) = \frac{|t|}{t}\frac{1}{2\sqrt{2}}$ no existe límite.

h: Como $h(t,t) \to 2$ y $h(2t,t) \to \frac{5}{2}$ no existe límite.

u: Como $u(t,t) \to 1$ y $u(t,-t) \to -1$ no existe límite.

v: Como $v(t,0) = t$, de existir límite ha de ser cero.

$$|v(x,y)| \leq \frac{|x^3| + |y^3|}{x^2+y^2} = \frac{|x^3|}{x^2+y^2} + \frac{|y^3|}{x^2+y^2} \leq \frac{|x^3|}{x^2} + \frac{|y^3|}{y^2} = \|(x,y)\|_1,$$

luego $\lim_{(x,y)\to(0,0)} v(x,y) = 0$.

2.25 $f(x,y) = \min\{x^2, y^2\} = \frac{1}{2}[x^2 + y^2 - |x^2 - y^2|]$ y por tanto continua por la continuidad de los polinomios, del valor absoluto, la regla de la cadena, etc.

También se puede hacer directamente teniendo en cuenta el carácter local de la continuidad y el estudio de la función en las bisectrices de los cuadrantes.

2.26 La función f es continua por serlo cada unas de sus tres campos escalares componentes.

2.27 Para encontrar el candidato a límite o demostrar que no existe, estúdiense los límites direccionales pues de existir límite han de existir todos y ser iguales al límite. De haber candidato a límite, para probar si tal candidato es o no el límite hay que acudir a la definición y utilizar convenientes acotaciones.

f: $\lim_{(x,x)\to(0,0)} f(x,x) = 1$.

$$\left|\frac{\operatorname{sen} x + \operatorname{sen} y}{x+y} - 1\right| \leq \left|\frac{\operatorname{sen} x - x}{x}\right| + \left|\frac{\operatorname{sen} y - y}{y}\right| \to 0.$$

g: $g(x,y) = \frac{x(x^2+y^2)}{y}\frac{\operatorname{sen}(x^2+y^2)}{(x^2+y^2)}$. Como el segundo factor tiene límite 1, basta estudiar el primero, g_1, que tiene límite direccional cero en todas las direcciones. Pero g_1 no está acotada en un entorno de $(0,0)$ (ya que $g_1(\frac{1}{n}, \frac{1}{n^4}) = n + \frac{1}{n^5}$) por lo que no existe el límite.

h: $\lim_{(x,x)\to(0,0)} h(x,x) = 1$.

$$\left|\frac{xe^y + ye^x}{x+y} - 1\right| \leq |e^y - 1| + |e^x - 1| \to 0,$$

luego $\lim_{(x,y)\to(0,0)} h(x,y) = 1$.

Φ: Como $0 \leq E(\frac{y}{x}) \leq \frac{y}{x} \Rightarrow 0 \leq \Phi(x,y) \leq y$ se tiene que $\lim_{(x,y)\to(0,0)} \Phi(x,y) = 0$.

2.28 Acotaciones útiles:

$$\boxed{|xy| \leq \tfrac{1}{2}(x^2+y^2),\ \forall (x,y) \in \mathbb{R}^2, \quad |\operatorname{sen} x| \leq |x|, |\arctan x| \leq |x|,\ \forall x \in \mathbb{R}.}$$

Se tiene que $|\operatorname{sen}(xy)| \leq |xy| \leq \frac{1}{2}(x^2+y^2)$ con lo que

$$\alpha < 1 \Rightarrow |f(x,y)| \leq \frac{1}{2}(x^2+y^2)^{1-\alpha} \Rightarrow \lim_{(x,y)\to(0,0)} f(x,y) = 0.$$

Por otra parte, si $\alpha = 1$ no existe límite pues $f(x,0) = 0$ y $f(x,x) = \frac{\operatorname{sen}(x^2)}{2x^2} \to \frac{1}{2}$.

Y si $\alpha > 1$ tampoco existe el límite ya que $f(x,x) = \frac{\operatorname{sen}(x^2)}{x^2}\frac{x^{2(1-\alpha)}}{2^\alpha}$ no está acotada.

2.29
$$\frac{e^x - e^{yx}}{x} = \frac{e^x - 1}{x} + \frac{1 - e^{xy}}{x} = \frac{e^x - 1}{x} + y\frac{1 - e^{xy}}{yx} \to 1 - b,$$

pues $\frac{e^x-1}{x} \to 1$ si $x \to 0$ (derivada de la exponencial).

$$(x,y) \to (0,b) \Rightarrow xy \to 0 \Rightarrow \frac{1 - e^{xy}}{yx} \to -1.$$

$$\left|\frac{1 - cos\ x\ cos\ y}{\sqrt{x^2 + y^2}}\right| = \left|\frac{(cos\ x)(1 - cos\ y) + 1 - cos\ x}{\sqrt{x^2 + y^2}}\right| \leq \frac{|1 - cos\ y| + |1 - cos\ x|}{\sqrt{x^2 + y^2}} \leq$$

$$\left|\frac{1 - cos\ y}{y}\right| + \left|\frac{1 - cos\ x}{x}\right| \to 0 \text{ (derivada del coseno).}$$

2.30 f: $f(x,0) = 1, f(0,y) = 0$, luego no existe límite.

g: Como $g(x,0) = 2x$, de existir límite este es cero.

$$|g(\rho\ cos\ \vartheta, \rho\ \text{ sen } \vartheta)| = \rho|2cos^5\vartheta + 2\,\text{sen}^3\,\vartheta(2cos^2\vartheta - \text{sen}^2\,\vartheta)| \leq 8\rho,$$

luego $\lim_{(x,y)\to(0,0)} g(x,y) = 0$.

h: Como $h(x,0) = 0$, de existir límite este es cero.

$$|h(\rho\ cos\ \vartheta, \rho\ \text{ sen } \vartheta)| = \rho^2|tg\ \vartheta|.$$

Sea $\{\rho_n\} \to 0$. Para cada n, sea ϑ_n con $0 < \vartheta_n < \frac{\pi}{2}$ tal que $\rho_n^2 tg\ \vartheta_n = 1$. La proposición del cambio a coordenadas polares nos prueba que no existe límite de h pues

$$\sup\left\{\rho^2 tg\ \vartheta : 0 < \vartheta < \frac{\pi}{2}\right\} = +\infty,\ \forall \rho > 0.$$

u: Como $u(x,0) = \frac{x^2}{\sqrt{x^2+1}-1} = \sqrt{x^2+1} + 1 \to 2$, de existir este límite, valdría 2. De $|u(\rho cos\ \vartheta, \rho\,\text{sen}\ \vartheta) - 2| = |\sqrt{\rho^2+1} - 1| \leq \varepsilon$ si $0 < \rho < \varepsilon$, se concluye que

$$\lim_{(x,y)\to(0,0)} u(x,y) = 2.$$

2.31
$\boxed{\mathbf{i}) \Longrightarrow \mathbf{ii})}$

Es el conocido Teorema de los ceros de Bolzano. Sea $f : [a,b] \to \mathbb{K}$ una función continua verificando que

$$f(a) < 0 < f(b)\ .$$

Sea

$$C := \{x \in [a,b] :\ f(x) < 0\}$$

que es no vacío y mayorado. Sea

$$c := \sup\, C.$$

Probemos que $f(c) = 0$. Es claro que $c \in [a,b]$ y que $f(c) \leq 0$, y en consecuencia $c < b$. Supongamos, razonando por reducción al absurdo, que $f(c) < 0$. Por continuidad de la función f en c, existiría $\delta > 0$ tal que

$$f(c+\delta) < 0,$$

lo que contradice la elección de c.

$\boxed{\mathbf{ii)} \Longrightarrow \mathbf{iii)}}$

Razonando por reducción al absurdo, supongamos $\mathbb{K}$ es la unión disjunta de dos abiertos O_1 y O_2 no vacíos. Tendríamos entonces que la función $f : \mathbb{K} \to \mathbb{K}$ definida por

$$f(x) = \begin{cases} -1, \ \forall x \in O_1 \\ \ \ 1, \ \forall x \in O_2 \end{cases}$$

sería continua (carácter local de la continuidad) tomaría valores positivos y negativos y no se anularía. Hemos probado que si $\mathbb{K}$ verifica el axioma de Bolzano, entonces $\mathbb{K}$ es conexo.

$\boxed{\mathbf{iii)} \Longrightarrow \mathbf{i)}}$

Sea $A \subset \mathbb{K}$ no vacío y mayorado. Notamos $M(A)$ el conjunto de sus mayorantes. Veamos que $\mathbb{K} \setminus M(A)$ es abierto. En efecto si $m \in \mathbb{K} \setminus M(A)$, existe $a \in A$ tal que $m < a$, con lo cual

$$]\leftarrow, a[\subset \mathbb{K} \setminus M(A).$$

Si suponemos que $M(A)$ no tiene mínimo, sería también abierto. En efecto si $M \in M(A)$, existiría $M' \in M(A)$ con $M' < M$, y en consecuencia

$$]M', \rightarrow [\subset M(A).$$

Tendríamos entonces que $\{\mathbb{K} \setminus M(A), M(A)\}$ sería una desconexión de $\mathbb{K}$. Hemos probado que $\mathbb{K}$ es conexo, entonces $\mathbb{K}$ verifica el axioma del supremo.

2.11. Breve biografía de los matemáticos mencionados en los temas 1 y 2

Arquímedes de Siracusa, *(287-212 a.C.).*

Nació en Siracusa, Sicilia. Estudió con los sucesores de Euclides en Alejandría, con los que se mantuvo después en contacto y les comunicaba sus avances. Fue bastante conocido por sus contemporáneos como el inventor de algunas "eficaces" máquinas de guerra.

Es considerado como uno de los más grandes matemáticos de todos los tiempos. Usó el método de exhaucción, que fue el origen de la integración, y que le permitió calcular áreas, volúmenes y superficies. Dió una aproximación del número π, así como un método para aproximar raíces cuadradas. En mecánica, descubrió resultados fundamentales sobre el centro de gravedad de cuerpos. El famoso resultado que lleva su nombre, da el peso de un cuerpo inmerso en un líquido.

Muchos de sus trabajos han sobrevivido hasta la actualidad, como: Sobre equilibrios en el plano, Cuadratura de la parábola, Sobre la esfera y el cilindro, Sobre espirales, Medida de un círculo.

Arquímedes murió durante la toma de Siracusa en la Segunda Guerra Púnica.

Bolzano, Bernhard *(1781-1848).*

Filósofo, lógico y matemático checo de origen italiano. Además de sus importantes trabajos en el campo de los fundamentos de la lógica, anticipó importantes concepciones relativas a la teoría de conjuntos y creó la primera función continua no derivable en ningún punto.

Borel, Emile *(1871-1956)*

Matemático y político francés. Ocupó los cargos de diputado (1924) y ministro de Marina (1925). Trabajó con éxito en diferentes áreas de la Matemática, especialmente sobre funciones complejas, topología, variable real y teoría de la medida. Algunas de sus aportaciones fueron fundamentales para la moderna teoría de la integración.

También actuó como representante de la comunidad matemática al público general.

Trabajó como profesor en la universidades de Lille y la Sorbona y en la "Ecole Normale Supérieure". Publicó más de 300 trabajos, entre artículos y libros. También fue editor de una importante colección de libros, en la cual publicó su obra "Lecons sur la théorie des fonctions". En 1921 fue nombrado miembro de la Académie des Sciences y presidente de ésta en 1934; recibió la primera medalla de oro de la CNRS.

Como conferenciante impresionaba, con su aire de distinción y dignidad. Parece ser que la demostración del Teorema de Heine-Borel-Lebesgue (caracterización de la compacidad en $\mathbb{R}^N$) se debe únicamente a él.

Cantor, Georg *(1854-1918).*

Matemático alemán de origen ruso. Se le considera el creador de la llamada teoría de conjuntos y de la teoría de los números transfinitos. Su obra impulsó una revisión en profundidad de los fundamentos de las matemáticas.

Cauchy, barón Augustin *(1789-1857).*

Matemático francés. Autor de más de 700 memorias en diversos campos de la ciencia, introdujo métodos rigurosos en el campo del análisis y creó la llamada teoría de las funciones analíticas.

Dedekind, Richard *(1831-1916).*

Matemático alemán. Alumno de Gauss, e introductor en el campo del análisis de las nociones que permiten precisar el concepto de número inconmensurable. Se le deben trabajos relativos, entre otros, a las integrales eulerianas, a los números irracionales, a las ecuaciones y funciones algebraicas, etc.

Dini, Ulisse *(1845-1918)*

Matemático italiano nacido en Pisa. Se doctoró en Ciencias en 1864. Desde 1866 desempeñó varias cátedras en la Universidad de Pisa, entre otras la de Análisis Matemático y Física Matemática. Fue diputado del Parlamento italiano y senador. Escribió varios artículos y tres libros: *Analisi infinitesimali* (1878), *Fondamenti per la teorica delle funzioni di variabili reali* (1878) y *Sopra la serie de Fourier* (1880). Sus conocidos teorema y criterio aparecen, respectivamente, en el primer y tercer libro.

Euclides *(año 300 a. C.)*

Matemático griego, fundador de la escuela de Alejandría (año 322 a. C.). Además de sus aportaciones a otros campos del saber como la óptica, su principal obra fue la llamada *Elementos*, considerada la obra de geometría por excelencia, un completo tratado sobre la geometría clásica y la lógica, y que contiene el famoso postulado que lleva su nombre. Esta obra ha sido usada durante varios siglos, y en ella Euclides recogió el trabajo de Pitágoras e Hipócrates, entre otros.

Euclides estableció postulados (conjunto de hipótesis de trabajo) a partir de los cuales demostraba resultados. También fue autor de varias obras más sobre geometría plana, geometría esférica y perspectiva. Posiblemente ha sido el matemático más influyente en la historia de la geometría. Tal influencia llegó incluso hasta Galileo o Newton.

Hausdorff, Felix *(1868-1942)*

Matemático alemán que desarrolló nociones básicas como las de límite, continuidad, conexión y compacidad. Fundamentó la topología. En 1897 empezó a publicar sus trabajos sobre geometría, números complejos y probabilidad. También se interesó en el trabajo de Cantor sobre teoría de conjuntos y en el de Hilbert. Una de sus revolucionarias ideas, los espacios de dimensión no entera, juegan un importante papel en la teoría de los sistemas dinámicos y en la descripción de los populares fractales. Su interés no se limitaba a las matemáticas, sino que alcanzaba también, por ejemplo, al arte o la filosofía.

Heine, Heinrich Eduard *(1821-1881)*

Matemático alemán. Heine hizo sus principales contribuciones de las matemáticas en el campo del análisis (polinomios de Legendre, funciones de Bessel y Lamé, etc.). Formuló el concepto de continuidad uniforme y el Teorema de Heine-Borel-Lebesgue. Fue alumno de Gauss en la universidad y su tesis doctoral fue dirigida por Dirichlet. Activo miembro de la Academia Prusiana de Ciencias, recibió la Medalla de Gauss en 1877.

Kronecker, Leonard *(1823-1891)*.

Alumno en Berlín de Kummer, se opuso a las teorías de Cantor, Dedekind y Weierstrass. Fue uno de los más celebres algebristas del siglo XIX. Kronecker consideró la aritmética fundada en los números naturales como la única verdadera "creación divina". Estudió las propiedades de los números y la teoría de los cuerpos de números algebraicos. A él se debe el resultado que establece que un sistema de ecuaciones lineales no homogéneas es compatible si, y sólo si, la matriz de los coeficientes tiene el mismo rango que la matriz que se obtiene al adjuntar a la anterior la columna de los términos independientes. También introdujo la llamada función δ de Kronecker, que vale uno o cero, según dos elementos sean iguales o distintos, es decir,

$$\delta_{i,j} = \begin{cases} 1 & \text{si } i = j, \\ \\ 0 & \text{si } i \neq j \end{cases}$$

Lebesgue, Henri *(1875-1941)*

Matemático francés. Desarrolló durante la primera década del siglo XX la integral que lleva su nombre. Entró en 1894 en la "Ecole Normale Superieure"; Borel era en aquella época profesor y Lebesgue siguió las enseñanzas de Borel acerca de numerosas materias, en particular los trabajos de Cantor y la definición de medida. Lebesgue tomó conciencia de la falta de rigor del curso de Borel. Se cuestionó los métodos tradicionales de la Matemática y tras su graduación, estuvo trabajando de forma intensa durante dos años en la biblioteca.¿Cómo se desarrollaron sus primeros trabajos? Él cuenta "mi teoría de la integración data de la época en que yo tenía veintiuna horas de clase (y veintitrés años)". Dicha teoría permite integrar una clase más amplia de funciones que la integral de Riemann, y también con mejores propiedades de convergencia. Lebesgue no lidera ningún grupo de investigación. Sin embargo sus habilidades pedagógicas eran muy reconocidas.

Además, es autor de numerosos trabajos sobre teoría de funciones de variable real.

Peano, Giuseppe *(1858-1932)*.

Lógico y matemático italiano. Además de la exposición rigurosamente deductiva de diversos campos de las matemáticas, creó un sistema de símbolos para la descripción y enunciados de las proposiciones lógicas y matemáticas sin necesidad de recurrir al lenguaje ordinario.

Weierstrass, Karl *(1815-1897)*.

Matemático alemán. Desarrolló un trabajo de gran rigor en el campo del análisis y fue la cabeza de la escuela de analistas que acometió la revisión sistemática de las diferentes ramas del análisis matemático. Su nombre ha quedado indisolublemente unido a la teoría de funciones elípticas.

Capítulo II

Derivación

Tema 3

Campos escalares y vectoriales derivables. Reglas de derivación.

La manera natural de extender a campos vectoriales el concepto de función real de variable real derivable es introducir el concepto de derivada en el sentido de Fréchet:

Un campo vectorial f definido en $A \subset \mathbb{R}^N$ y con valores en $\mathbb{R}^M$ es derivable, en el sentido de Fréchet, en un punto a interior de A si existe $T \in \mathscr{L}(\mathbb{R}^N, \mathbb{R}^M)$, el espacio de los campos vectoriales de $\mathbb{R}^N$ en $\mathbb{R}^M$, verificando

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

en cuyo caso la aplicación T es única, se denomina la derivada de la función f en el punto a y se nota por $Df(a)$. Los conceptos de derivabilidad y derivada son algebraico-topológicos (no dependen de las normas elegidas en $\mathbb{R}^N$ y $\mathbb{R}^M$).

Así el estudio de la derivada de Fréchet para campos vectoriales requiere estar familiarizado con el espacio de Banach $\mathscr{L}(\mathbb{R}^N, \mathbb{R}^M)$ de los campos vectoriales lineales de $\mathbb{R}^N$ en $\mathbb{R}^M$. Empezamos estableciendo el isomorfismo que existe entre este espacio vectorial y el las matrices $\mathscr{M}_{M \times N}(\mathbb{R})$. Dicho isomorfismo hace corresponder a la composición de campos el producto de las correspondientes matrices.

Es claro que el único campo vectorial lineal acotado es el nulo y que dos campos vectoriales lineales que toman los mismos valores en la esfera unidad coinciden (!Hágase!). A cada $T \in \mathscr{L}(\mathbb{R}^N, \mathbb{R}^M)$ le asignamos el número real

$$\|T\| := \max\{\|T(x)\| : \|x\| = 1\}.$$

Probamos que tal aplicación es una norma en $\mathscr{L}(\mathbb{R}^N, \mathbb{R}^M)$, y que para $T \in \mathscr{L}(\mathbb{R}^N, \mathbb{R}^M)$ se tiene

$$\|T(x)\| \leq \|T\| \, \|x\|, \ \forall x \in \mathbb{R}^N$$

con lo que todo campo vectorial lineal es lipschitziano; probamos también que $\|T\|$ es la constante de Lipschitz de T (véase Definición 3.2).

Estudiamos también los isomorfismos topológicos en $\mathbb{R}^N$

$$\mathrm{Iso}\,(\mathbb{R}^N) = \{T \in \mathscr{L}(\mathbb{R}^N) : \det A_T \neq 0\}$$

que serán esenciales a la hora de establecer el Teorema de la función inversa. Probamos que $\mathrm{Iso}\,(\mathbb{R}^N)$ es un abierto de $\mathscr{L}(\mathbb{R}^N)$ y que la aplicación inversión $J : \mathrm{Iso}\,(\mathbb{R}^N) \to \mathscr{L}(\mathbb{R}^N)$ dada por

$$J(T) := T^{-1},\ \forall T \in \mathrm{Iso}\,(\mathbb{R}^N)$$

es continua.

3.1. El espacio de Banach $\mathscr{L}(\mathbb{R}^N,\mathbb{R}^M)$.

En lo sucesivo, para cada dos naturales N y M, $\mathscr{L}(\mathbb{R}^N,\mathbb{R}^M)$ denotará el espacio vectorial de los campos vectoriales lineales de $\mathbb{R}^N$ en $\mathbb{R}^M$, es decir, el conjunto de las aplicaciones $T:\mathbb{R}^N\to\mathbb{R}^M$ tales que

$$T(x+y)=T(x)+T(y),\ \ T(\lambda x)=\lambda T(x),\quad \forall x,y\in\mathbb{R}^N,\ \forall\lambda\in\mathbb{R},$$

con las operaciones usuales de suma y producto por escalares. En el caso de que $N=M$ escribiremos simplemente $\mathscr{L}(\mathbb{R}^N)$ en lugar de $\mathscr{L}(\mathbb{R}^N,\mathbb{R}^N)$.

Conviene dejar sentado desde el primer momento que el espacio vectorial $\mathscr{L}(\mathbb{R}^N,\mathbb{R}^M)$ y el espacio vectorial $\mathscr{M}_{M\times N}(\mathbb{R})$ de las matrices $M\times N$ de números reales son matemáticamente indistinguibles. Para cada $T\in\mathscr{L}(\mathbb{R}^N,\mathbb{R}^M)$ definimos $A_T\in\mathscr{M}_{M\times N}(\mathbb{R})$ por

$$A_T:=\Big(T_i(e_j)\Big)_{\substack{1\le i\le M\\ 1\le j\le N}}, \tag{3.1.1}$$

donde $T_1,\dots,T_M$ son los campos escalares componentes de T y $\{e_1,\dots,e_N\}$ es la base canónica de $\mathbb{R}^N$. La aplicación de $\mathscr{L}(\mathbb{R}^N,\mathbb{R}^M)$ en $\mathscr{M}_{M\times N}(\mathbb{R})$ definida por $T\to A_T$ es un isomorfismo de $\mathscr{L}(\mathbb{R}^N,\mathbb{R}^M)$ sobre $\mathscr{M}_{M\times N}(\mathbb{R})$, cuyo inverso es la aplicación de $\mathscr{M}_{M\times N}(\mathbb{R})$ sobre $\mathscr{L}(\mathbb{R}^N,\mathbb{R}^M)$ definida por $A\to T_A$ donde T_A es la aplicación lineal de $\mathbb{R}^N$ en $\mathbb{R}^M$ dada por

$$\Big(T_A(x)\Big)^t:=Ax^t,\ \forall x\in\mathbb{R}^N. \tag{3.1.2}$$

En particular, si $T\in\mathscr{L}(\mathbb{R}^N,\mathbb{R})$, entonces $A_T\in\mathbb{R}^N$ y si $A\in\mathbb{R}^N$, entonces la correspondiente aplicación T_A actúa de la forma siguiente

$$T_A(x)=(A|x),\ \forall x\in\mathbb{R}^N, \tag{3.1.3}$$

donde hemos notado por $(\cdot|\cdot)$ el producto escalar en $\mathbb{R}^N$.

Además esta identificación de $\mathscr{L}(\mathbb{R}^N,\mathbb{R}^M)$ con $\mathscr{M}_{M\times N}(\mathbb{R})$ tiene la importante siguiente propiedad. Si $T\in\mathscr{L}(\mathbb{R}^N,\mathbb{R}^M)$ y $S\in\mathscr{L}(\mathbb{R}^M,\mathbb{R}^P)$, entonces

$$A_{(ST)}=A_SA_T, \tag{3.1.4}$$

donde, como es usual, notamos por yuxtaposición la composición de los operadores S y T, es decir

$$(ST)(x)=S(T(x)),\forall x\in\mathbb{R}^N.$$

En efecto, si $x\in\mathbb{R}^N$ se tiene que

$$(A_SA_T)x^t=A_S(A_Tx^t)=A_S\Big(T(x)\Big)^t=(S(T(x)))^t=((ST)(x))^t=A_{(ST)}x^t.$$

Sabemos que un campo vectorial es continuo si, y sólo si, lo son sus campos escalares componentes. En el caso particular de campos vectoriales lineales esto es automático, pues en vista de 3.1.3, los campos escalares componentes son polinomios.

Es claro que si $M \neq N$, entonces no existen biyecciones lineales de $\mathbb{R}^N$ sobre $\mathbb{R}^M$. Además, las biyecciones lineales de $\mathbb{R}^N$ sobre $\mathbb{R}^N$ son automáticamente bicontinuas, es decir, son homeomorfismos lineales. Es usual la siguiente nomenclatura:

Definición 3.1 (Isomorfismo topológico)**.** Si X e Y son espacios normados, una aplicación $T : X \longrightarrow Y$ es un *isomorfismo topológico* si T es una aplicación biyectiva, lineal y continua cuya inversa también es continua.

Por el comentario anterior a la definición, toda biyección lineal $T : \mathbb{R}^N \to \mathbb{R}^N$ es un isomorfismo topológico. En lo que sigue, notaremos Iso $(\mathbb{R}^N)$ al conjunto de los isomorfismos topológicos de $\mathbb{R}^N$ sobre $\mathbb{R}^N$.

A partir de 3.1.4 se obtiene que

$$\text{Iso } (\mathbb{R}^N) = \{T \in \mathscr{L}(\mathbb{R}^N) : \det A_T \neq 0\}$$

y si $T \in \text{Iso } (\mathbb{R}^N)$, entonces

$$A_{(T^{-1})} = (A_T)^{-1}.$$

Veremos enseguida que los campos vectoriales lineales son de hecho algo más que uniformemente continuos.

Definición 3.2 (función lipschitziana)**.** Sean (E,d) y (F,ρ) espacios métricos. Una función $f : E \to F$ es *lipschitziana* si $\exists K \geq 0$ tal que

$$\rho(f(x), f(y)) \leq K d(x,y), \quad \forall x,y \in E. \tag{$*$}$$

A la menor de las constantes que verifican la condición $(*)$ se le llama la *constante de Lipschitz* de f. Evidentemente las funciones lipschitzianas con constante de Lipschitz 0 no son otra cosa que las funciones constantes.

Es inmediato probar que las funciones lipschitzianas son uniformemente continuas. En efecto, dado $\varepsilon > 0$, si f es lipschitziana de razón K, sea $\delta := \frac{\varepsilon}{K+1}$. Se tiene que

$$[d(x,y) < \delta] \Rightarrow \rho(f(x), f(y)) \leq K\, d(x,y) < \varepsilon.$$

La función $f : \mathbb{R} \to \mathbb{R}$ definida por

$$f(x) = \sqrt{|x|} \qquad (x \in \mathbb{R})$$

es uniformemente continua y, sin embargo, no es lipschitziana. En efecto, si $x,y \in \mathbb{R}$, se tiene que

$$\sqrt{|x|} \leq \sqrt{|x-y| + |y|} \leq \sqrt{|x-y|} + \sqrt{|y|},$$

es decir
$$\sqrt{|x|}-\sqrt{|y|}\leq\sqrt{|x-y|},$$
y en consecuencia, intercambiando los papeles de x e y, que
$$\left|\ \sqrt{|x|}-\sqrt{|y|}\ \right|\leq\sqrt{|x-y|}$$
de donde se deduce inmediatamente que es uniformemente continua (¡Hágase!).

Si fuese lipschitziana de razón $K>0$, tendríamos que
$$\left|\frac{1}{\sqrt{n}}-0\right|\leq K\ \frac{1}{n},\ \forall n\in\mathbb{N}$$
o lo que es lo mismo
$$n\leq K^2,\forall n\in\mathbb{N}$$
y los naturales estarían acotados.

Ejemplos 3.3 (funciones lipschitzianas)**.**

a) Sea $(X,\|\cdot\|)$ un espacio normado. Son lipschitzianas las siguientes aplicaciones:

- La norma en X:
$$|\,\|x\|-\|y\|\,|\leq\|x-y\|.$$
- La suma en X:
$$\|(x_1+y_1)-(x_2+y_2)\|\leq\|x_1-x_2\|+\|y_1-y_2\|\leq 2\|(x_1,y_1)-(x_2,y_2)\|_\infty.$$
- Sabemos que el producto por escalares ni siquiera es uniformemente continuo.

b) Sea (E,d) un espacio métrico. Son lipschitzianas las siguientes aplicaciones:

- La distancia de E:
$$|d(x_1,y_1)-d(x_2,y_2)|\leq|d(x_1,y_1)-d(y_1,x_2)|+|d(y_1,x_2)-d(x_2,y_2|\leq$$
$$d(x_1,x_2)+d(y_1,y_2)\leq 2d((x_1,y_1),(x_2,y_2)).$$
- La distancia a un subconjunto no vacío A de E:
$$|\text{dist}(x,A)-\text{dist}(y,A)|\leq d(x,y).$$
En efecto, para $x,y\in E$, se tiene
$$d(x,a)\leq d(x,y)+d(y,a),\quad\forall a\in A,$$
y en consecuencia
$$\text{dist}(x,A)\leq d(x,y)+\text{dist}(y,A),$$

o lo que es lo mismo

$$\operatorname{dist}(x,A)-\operatorname{dist}(y,A)\le d(x,y).$$

Ahora, intercambiando los papeles de x e y, se tiene la desigualdad

$$\operatorname{dist}(y,A)-\operatorname{dist}(x,A)\le d(x,y),$$

que unida a la anterior lleva a

$$\big|\operatorname{dist}(x,A)-\operatorname{dist}(y,A)\big|\le d(x,y).$$

En consecuencia, si (E,d) es un espacio métrico, A un subconjunto no vacío de E y a un número real, entonces los conjuntos

$$U:=\{x\in E:\operatorname{dist}(x,A)>a\}\quad \text{y}\quad V:=\{x\in E:\operatorname{dist}(x,A)<a\}.$$

son abiertos (ver caracterización topológica de la continuidad global).

La siguiente proposición prueba que una aplicación lineal de $\mathbb{R}^N$ en $\mathbb{R}^M$ es lipschitziana, y además da la constante de Lipschitz.

Proposición 3.4. *Sean $N,M\in\mathbb{N}$ y $T:\mathbb{R}^N\to\mathbb{R}^M$ lineal. Entonces se verifican las siguientes afirmaciones:*

i) $\|T\|$ alcanza el máximo en la esfera unidad y en consecuencia podemos definir

$$\|T\|:=\max\{\|T(x)\|:\|x\|=1\}.$$

ii) T es lipschitziana, de hecho se verifica que

$$\|T(x)\|\le\|T\|\|x\|,\forall x\in\mathbb{R}^N. \tag{3.1.5}$$

Además

$$\|T\|=\min\{K\ge 0:\ \|T(x)\|\le K\|x\|,\ \forall x\in\mathbb{R}^N\}=\max\{\|T(x)\|:\|x\|\le 1\}.$$

La primera igualdad nos asegura que $\|T\|$ es la constante de Lipschitz de T.

Demostración:

i) Como T es continua y la esfera unidad es compacta, $i)$ es consecuencia de la continuidad de la aplicación norma, de la regla de la cadena para funciones continuas y de la propiedad de compacidad.

ii) Es claro que

$$\left\|T\left(\frac{x}{\|x\|}\right)\right\|\le\|T\|,\ \forall x\in\mathbb{R}^N\backslash\{0\},$$

y usando que T es lineal, tenemos

$$\|T(x)\| \leq \|T\| \ \|x\|, \ \forall x \in \mathbb{R}^N \backslash \{0\}.$$

Dado que esta última expresión es también válida para cero, podemos escribirla equivalentemente en la forma

$$\|T(x)\| \leq \|T\| \|x\|, \ \forall x \in \mathbb{R}^N.$$

De la anterior desigualdad se deduce que T es lipschitziana, ya que si $x, y \in \mathbb{R}^N$, obtenemos

$$\|T(x) - T(y)\| = \|T(x-y)\| \leq \|T\| \ \|x-y\|,$$

es decir, T es lipschitziana.

Sea ahora $K \geq 0$ tal que

$$\|T(x)\| \leq K \ \|x\|, \ \ \forall x \in \mathbb{R}^N.$$

Se sigue que $\|T\| = \max\{\|T(x)\| : \ \|x\| = 1\} \leq K$, y por tanto

$$\|T\| = \min\{K \geq 0 : \|T(x)\| \leq K \ \|x\|, \ \forall x \in \mathbb{R}^N\},$$

es decir, $\|T\|$ es la constante de Lipschitz de T.

Finalmente, la propiedad de compacidad nos asegura que T alcanza el máximo en la bola unidad cerrada. Es claro que

$$\|T\| \leq \max\{\|T(x)\| : \ \|x\| \leq 1\}$$

y 3.1.5 nos asegura que también

$$\max\{\|T(x)\| : \ \|x\| \leq 1\} \leq \|T\|.$$

Hemos probado que

$$\|T\| = \max\{\|T(x)\| : \ \|x\| \leq 1\}.$$

∎

El teorema de Hausdorff nos asegura que todas las normas definibles en el espacio vectorial $\mathscr{L}(\mathbb{R}^N, \mathbb{R}^M)$ son equivalentes (es de dimensión $M \times N$). En el siguiente resultado presentamos la forma usual de normar este espacio.

Teorema 3.5 (El espacio de Banach $\mathscr{L}(\mathbb{R}^N, \mathbb{R}^M)$)**.** *La función que a cada aplicación lineal $T : \mathbb{R}^N \to \mathbb{R}^M$ le hace corresponder el número real*

$$\|T\| := \max\{\|T(x)\| : \|x\| = 1\}$$

es una norma en $\mathscr{L}(\mathbb{R}^N, \mathbb{R}^M)$, denominada <u>norma de operadores</u>. Además $\mathscr{L}(\mathbb{R}^N, \mathbb{R}^M)$ es un espacio de Banach.

Demostración:

Veamos que la función $\|\cdot\| : \mathscr{L}(\mathbb{R}^N,\mathbb{R}^M) \to \mathbb{R}$ es una norma. Para $T,S \in \mathscr{L}(\mathbb{R}^N,\mathbb{R}^M)$ y $\lambda \in \mathbb{R}$, se tiene:

i)
$$\|T\| = 0 \Rightarrow \left(\|T(x)\| \leq \|T\|\|x\|, \forall x \in \mathbb{R}^N\right) \Rightarrow T = 0.$$

ii)
$$\|\lambda T\| = \max\{\|(\lambda T)(x)\| : \|x\| = 1\} = \max\{|\lambda|\,\|T(x)\| : \|x\| = 1\} = |\lambda|\ \|T\|.$$

iii)
$$\|(T+S)(x)\| = \|T(x)+S(x)\| \leq \|T(x)\| + \|S(x)\| \leq \|T\|\ \|x\| + \|S\|\ \|x\|, \forall x \in \mathbb{R}^N,$$

donde se ha usado 3.1.5. Hemos probado que

$$\|(T+S)(x)\| \leq (\|T\| + \|S\|)\ \|x\|,\ \forall x \in \mathbb{R}^N$$

y en consecuencia, en vista de la proposición anterior, concluimos que

$$\|T+S\| \leq \|T\| + \|S\|.$$

Finalmente $\mathscr{L}(\mathbb{R}^N,\mathbb{R}^M)$, al ser de dimensión finita, es completo para cualquier norma, en particular, para la norma de operadores. ∎

Hállense la norma de operadores de $\mathscr{L}((\mathbb{R}^2,\|.\|_1),\mathbb{R})$, $\mathscr{L}((\mathbb{R}^2,\|.\|_2),\mathbb{R})$ (úsese en este caso la desigualdad de Cauchy-Schwarz, Apéndice A) y $\mathscr{L}((\mathbb{R}^2,\|.\|_\infty),\mathbb{R})$.

El siguiente resultado generaliza el hecho de que $\mathbb{R}^*$ sea abierto, ya que Iso $(\mathbb{R})$ se identifica con $\mathbb{R}^*$ (¿Por qué?).

Proposición 3.6. Iso $(\mathbb{R}^N)$ *es un abierto de* $\mathscr{L}(\mathbb{R}^N)$.

Demostración:

La aplicación de $\mathscr{L}(\mathbb{R}^N)$ en $\mathbb{R}$ definida por

$$T \to \det A_T \ \ (T \in \mathscr{L}(\mathbb{R}^N))$$

es continua (¡Hágase!). En consecuencia

$$\text{Iso }(\mathbb{R}^N) = \{T \in \mathscr{L}(\mathbb{R}^N) :\ \det A_T \neq 0\}$$

es un abierto de $\mathscr{L}(\mathbb{R}^N)$. ∎

Veamos para finalizar esta sección que, al igual que la aplicación de $\mathbb{R}^*$ en $\mathbb{R}$ dada por $x \to \frac{1}{x}$ es continua, también la aplicación de Iso $(\mathbb{R}^N)$ en $\mathscr{L}(\mathbb{R}^N)$ dada por $T \to T^{-1}$ es continua.

Proposición 3.7 (Continuidad de la aplicación inversión)**.** *La aplicación inversión* $J :$ Iso $(\mathbb{R}^N) \to \mathscr{L}(\mathbb{R}^N)$ *definida por*

$$J(T) := T^{-1}, \quad \forall T \in \text{Iso }(\mathbb{R}^N)$$

es continua.

Demostración:

Si $S,T \in \mathscr{L}(\mathbb{R}^N)$, es claro que $ST \in \mathscr{L}(\mathbb{R}^N)$. Veamos ahora la relación entre las normas de los tres operadores. Para $x \in \mathbb{R}^N$ se tiene que

$$\|(ST)(x)\| = \|S(T(x))\| \leq \|S\|\|T(x)\| \leq \|S\|\|T\|\|x\|,$$

y por tanto,

$$\|ST\| \leq \|S\|\|T\|. \tag{3.1.6}$$

Sean $T \in \text{Iso}\ (\mathbb{R}^N)$ y $\{T_n\}$ una sucesión en Iso $(\mathbb{R}^N)$ convergente a T. Para cada natural n se tiene que, en vista de la desigualdad anterior

$$\|J(T_n) - J(T)\| = \|T_n^{-1} - T^{-1}\| = \|T_n^{-1}(T - T_n)T^{-1}\| \leq \|T_n^{-1}\|\ \|T - T_n\|\ \|T^{-1}\|,$$

de donde se deduce la continuidad de J en T sin más que comprobar que la sucesión $\{\|T_n^{-1}\|\}$ está acotada. Para cada natural n tenemos, usando la desigualdad recién obtenida concluimos que

$$\begin{aligned}\left|\frac{1}{\|T^{-1}\|} - \frac{1}{\|T_n^{-1}\|}\right| &= \left|\frac{\|T_n^{-1}\| - \|T^{-1}\|}{\|T^{-1}\|\ \|T_n^{-1}\|}\right| \leq \\ &\leq \frac{\|T_n^{-1} - T^{-1}\|}{\|T^{-1}\|\ \|T_n^{-1}\|} \leq \\ &\leq \frac{\|T_n^{-1}\|\ \|T - T_n\|\ \|T^{-1}\|}{\|T^{-1}\|\ \|T_n^{-1}\|} = \|T - T_n\|.\end{aligned}$$

Por tanto

$$\lim \frac{1}{\|T_n^{-1}\|} = \frac{1}{\|T^{-1}\|},$$

en consecuencia

$$\lim\ \|T_n^{-1}\| = \|T^{-1}\|,$$

y en particular, la sucesión $\{\|T_n^{-1}\|\}$ está acotada. ∎

3.2. Concepto de derivada.

El concepto de derivada es de los que más han influido en el desarrollo de la matemática. Nuestro objetivo es el estudio de la derivabilidad, así como el cálculo de la derivada cuando ello proceda, de las funciones $f : A \rightarrow \mathbb{R}^M$, donde $A \subset \mathbb{R}^N$, siendo M y N dos naturales prefijados.

Esencialmente el estudio de la derivabilidad de un campo vectorial f en un punto a responde al problema de si f es aproximable por una función afín (continua) en el punto a, es decir, por una función g de la forma $g(x) = c + T(x),\ \forall x \in \mathbb{R}^N$, donde $c \in \mathbb{R}^M$ y $T \in \mathscr{L}(\mathbb{R}^N, \mathbb{R}^M)$. La parte lineal de la mejor aproximación afín de la función f en el punto a es la derivada de ésta en a y las propiedades de esta mejor aproximación (equivalentemente de la derivada) repercuten de manera natural sobre las propiedades locales de la función en el punto a. Así, el cálculo diferencial es una potente herramienta para estudiar el comportamiento local de funciones. El cálculo diferencial para aplicaciones entre espacios normados fue iniciado por Fréchet en 1.925, si bien la paternidad ha de ser compartida con muchos otros matemáticos como Stolz, Young, etc.

Empezaremos recordando la noción de derivada para funciones reales de variable real, así como su interpretación geométrica.

Definición 3.8 (derivada de funciones reales de variable real)**.** Sean $A \subset \mathbb{R}$, $a \in A \cap A'$ y $f : A \rightarrow \mathbb{R}$ una función. Se dice que f es derivable en el punto a si existe

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a},$$

en cuyo caso el valor de tal límite se nota $f'(a)$ y se denomina la <u>*derivada*</u> de la función f en el punto a.

La derivabilidad de una función tiene la siguiente magnífica caracterización en términos de existencia de una mejor aproximación afín, cuya interpretación es clara.

Proposición 3.9 (Caracterización de la derivabilidad)**.** *Sean $A \subset \mathbb{R}, a \in A \cap A'$ y $f : A \rightarrow \mathbb{R}$ una función. Equivalen:*

i) f es derivable en a.

ii) Existe una función afín (continua) $g : \mathbb{R} \rightarrow \mathbb{R}$ verificando que $g(a) = f(a)$ y que

$$\lim_{x \to a} \frac{f(x) - g(x)}{x - a} = 0.$$

En ese supuesto la función g es única y viene dada por

$$g(x) = f(a) + f'(a)(x - a),\ \forall x \in \mathbb{R}.$$

La Definición 3.8 y la Proposición 3.9 pueden extenderse literalmente a funciones vectoriales de una variable real.

Definición 3.10 (derivada elemental para funciones de $\mathbb{R}$ a $\mathbb{R}^M$)**.** Sean M un natural, $A \subset \mathbb{R}$, $a \in A \cap A'$ y $f: A \to \mathbb{R}^M$ una función. Se dice que f es *elementalmente derivable* en el punto a si existe

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a},$$

en cuyo caso el valor de tal límite (un vector de $\mathbb{R}^M$) se denomina la *derivada elemental* de f en el punto a y se nota $f'(a)$. Se dice que f es derivable elementalmente en un subconjunto $B \subset A$ si es derivable elementalmente en cada punto de B.

Sea $A_1 \subset A$ el conjunto de puntos donde f es derivable elementalmente. La aplicación $x \to f'(x)$ de A_1 en $\mathbb{R}^M$ se denomina la *aplicación derivada elemental* de f y se nota f'.

Proposición 3.11 (Caracterización de la derivabilidad elemental)**.** *Sean M un natural, $A \subset \mathbb{R}$, $a \in A \cap A'$ y $f = (f_1, ..., f_M) : A \to \mathbb{R}^M$ una función. Equivalen:*

i) f es derivable elementalmente en a.

ii) $f_1, ..., f_M$ son derivables en a.

iii) Existe una función afín (continua) $g : \mathbb{R} \to \mathbb{R}^M$ verificando que $g(a) = f(a)$ y que

$$\lim_{x \to a} \frac{f(x) - g(x)}{x - a} = 0.$$

En ese supuesto, $f'(a) = (f_1'(a), ..., f_M'(a))$, la función g es única y viene dada por

$$g(x) = f(a) + (x - a) f'(a), \ \forall x \in \mathbb{R}.$$

Demostración:

$i) \Leftrightarrow ii)$ Basta observar que para $x \in A \backslash \{a\}$ es

$$\frac{f(x) - f(a)}{x - a} = \left(\frac{f_1(x) - f_1(a)}{x - a}, \ldots, \frac{f_M(x) - f_M(a)}{x - a} \right)$$

y aplicar entonces la Proposición 2.62 sobre la reducción del límite a campos escalares.

$ii) \Leftrightarrow iii)$ Es consecuencia inmediata de las Proposiciones 3.9 y 2.62. ∎

El intento de llevar la Definición 3.10 al caso en que el dominio de la aplicación sea un subconjunto de un espacio de dimensión mayor que 1 tropieza con la imposibilidad de dar sentido al límite que en ella aparece. Sin embargo, la caracterización de la derivabilidad de una función en un punto por la existencia de una mejor aproximación afín a la función en el punto (afirmación *iii*) de la proposición anterior) es perfectamente trasladable al ámbito deseado si se cae en la cuenta de que

$$\lim_{x\to a}\frac{f(x)-g(x)}{x-a}=0 \iff \lim_{x\to a}\frac{f(x)-g(x)}{|x-a|}=0.$$

La anterior propiedad tiene sentido para funciones definidas en un subconjunto A de $\mathbb{R}^N$ y con valores en $\mathbb{R}^M$, sin más que sustituir el valor absoluto por una norma $\|\cdot\|$ en $\mathbb{R}^N$ y la función afín de $\mathbb{R}$ en $\mathbb{R}^M$ por una función afín de $\mathbb{R}^N$ en $\mathbb{R}^M$.

Parece pues que en el ambiente siguiente: $A\subset\mathbb{R}^N$, $a\in A\cap A'$, y $f:A\to\mathbb{R}^M$, la función f debe considerarse derivable en a cuando exista una función afín (continua) $g:\mathbb{R}^N\to\mathbb{R}^M$ verificando que $g(a)=f(a)$ y que

$$\lim_{x\to a}\frac{f(x)-g(x)}{\|x-a\|}=0.$$

Como quiera que una tal aplicación afín g ha de tener la forma

$$g(x)=f(a)+T(x-a),\ \forall x\in\mathbb{R}^N$$

para conveniente aplicación lineal $T:\mathbb{R}^N\to\mathbb{R}^M$, la condición anterior es equivalente a

$$\lim_{x\to a}\frac{f(x)-f(a)-T(x-a)}{\|x-a\|}=0.$$

Es importante observar que la condición anterior es topológica, es decir involucra solamente las topologías de $\mathbb{R}^N$ y $\mathbb{R}^M$ y no las concretas normas que se tomen. En efecto, fijada una norma en $\mathbb{R}^N$, al ser la noción de límite topológica concluimos que la condición es independiente de la norma elegida en $\mathbb{R}^M$. Ahora si $|||\cdot|||$ es otra norma en $\mathbb{R}^N$, entonces se tiene para $x\in A\backslash\{a\}$ que

$$\frac{f(x)-f(a)-T(x-a)}{\|x-a\|}=\frac{|||x-a|||}{\|x-a\|}\,\frac{f(x)-f(a)-T(x-a)}{|||x-a|||}$$

y como por la equivalencia de $\|\cdot\|$ y $|||\cdot|||$, existen $0<k,K$ tales que $k\leq\dfrac{|||x-a|||}{\|x-a\|}\leq K$, se sigue, teniendo en cuenta de nuevo que la noción de límite es topológica, que

$$\lim_{x\to a}\frac{f(x)-f(a)-T(x-a)}{\|x-a\|}=0 \iff \lim_{x\to a}\frac{f(x)-f(a)-T(x-a)}{|||x-a|||}=0.$$

Examinamos a continuación la repercusión que tiene la propiedad anterior sobre la continuidad de f en a. De ella deducimos que

$$\lim_{x\to a}f(x)-f(a)-T(x-a)=0,$$

lo que prueba que f es continua al ser T continua.

Para garantizar la unicidad de la aplicación lineal $T:\mathbb{R}^N\to\mathbb{R}^M$ verificando la condición requerida, necesitamos poder acercarnos al punto a en cualquier dirección, esto es, si para cada $x\in S(\mathbb{R}^N)$ notamos

$$A_x=\{t\in\mathbb{R}:\ a+tx\in A\},$$

deberá ser $0 \in A'_x$ para todo x en $S(\mathbb{R}^n)$ (véase Ejercicio 3.2). Es importante notar al respecto que la condición topológica más expeditiva para garantizar la anterior propiedad es que a sea un punto interior de A y esta condición será asumida en adelante. En este caso, sea $\delta > 0$ tal que $B(a,\delta) \subset A$. Para $x \in S(\mathbb{R}^N)$ y $t \in \mathbb{R}$ con $0 < |t| < \delta$, se tiene que $a + tx \in A$ y

$$\left\| \frac{f(a+tx) - f(a)}{t} - T(x) \right\| = \frac{\|f(a+tx) - f(a) - tT(x)\|}{|t|} =$$

$$\frac{\|f(a+tx) - f(a) - T(a+tx-a)\|}{\|a+tx-a\|},$$

de donde deducimos que

$$(*) \qquad \lim_{t\to 0} \frac{f(a+tx) - f(a)}{t} = T(x)$$

expresión que nos asegura la unicidad de T dado que una aplicación lineal está determinada por el comportamiento en la esfera unidad (¡Hágase!).

Estamos ya en condiciones de definir de manera coherente el concepto de derivada.

Definición 3.12 (función derivable Fréchet)**.** Sean $M, N \in \mathbb{N}, A \subset \mathbb{R}^N$, $a \in \overset{\circ}{A}$, y $f : A \to \mathbb{R}^M$ una función. Se dice que f es <u>derivable</u> (en el sentido de Fréchet) o diferenciable en el punto a si existe una aplicación lineal T de $\mathbb{R}^N$ en $\mathbb{R}^M$ verificando

$$(**) \qquad \lim_{x\to a} \frac{f(x) - f(a) - T(x-a)}{\|x-a\|} = 0$$

donde $\|\cdot\|$ es cualquier norma en $\mathbb{R}^N$. En tal caso la aplicación T es única, se denomina la <u>*derivada de una función en un punto*</u> o <u>*diferencial de f en a*</u> y se nota por $Df(a)$. Se dice que f es derivable en un subconjunto $B \subset A$ si es derivable en cada punto de B.

Sea $A_1 \subset A$ el conjunto de puntos donde f es derivable. La aplicación $x \mapsto Df(x)$ de A_1 en $\mathscr{L}(\mathbb{R}^N, \mathbb{R}^M)$ se denomina la <u>aplicación derivada</u> de f y se nota Df.

Se dice que f es de <u>*clase $\mathscr{C}^1$*</u> en a, y se nota $f \in C^1(a)$, si f es derivable en un entorno de a y la aplicación Df es continua en a. Se dice que f es de clase C^1 en un subconjunto $B \subset A$ si es de clase C^1 en cada punto de B.

Se dice que f es de clase C^1 cuando lo sea en todos los puntos de su conjunto de definición (que necesariamente será abierto). Notamos por $C^1(A)$ al conjunto de las funciones de clase C^1 en el abierto A.

Queda claro que si $N \geq 2$ se deriva en puntos interiores, lo que por comodidad, no siempre se resaltará en adelante.

Notas 3.13.

1. Resaltamos que los conceptos de derivabilidad y de derivada de una función en un punto involucran sólamente las topologías de $\mathbb{R}^N$ y $\mathbb{R}^M$ y no las normas concretas elegidas. Es decir, son conceptos algebraico-topológicos.

2. Obsérvese que la Proposición 3.11 tiene ahora la siguiente lectura para puntos interiores: Sean $A \subset \mathbb{R}$, $a \in \overset{\circ}{A}$, $M \in \mathbb{N}$ y $f = (f_1, \ldots, f_M) : A \to \mathbb{R}^M$ una función. Entonces las siguientes afirmaciones son equivalentes:

   i) f es derivable elementalmente en a.

   ii) $f_1, \ldots, f_M$ son derivables en a.

   iii) f es derivable en a.

   Además en el caso de que f sea derivable en a, la relación entre ambas derivadas viene dada, en virtud de (*), por

   $$Df(a)(x) = xf'(a), \ \forall x \in \mathbb{R},$$

   mientras que la función afín g viene dada por

   $$g(x) = (f(a) - af'(a)) + xf'(a)$$

   (nótese que $Df(a)$ es la aplicación lineal asociada a g). En particular, se obtiene que

   $$Df(a)(1) = f'(a) = (f_1'(a), ..., f_M'(a)).$$

3. Si f es derivable en el punto a, entonces f es continua en a.

De acuerdo con la definición, el estudio de la derivabilidad de una función en un punto conlleva dos problemas: el conocimiento de T dado por (*) y la verificación de (**).

Fijado un vector no nulo x, se dice que la función f es derivable en a en la dirección de x si existe

$$\lim_{t \to 0} \frac{f(a+tx) - f(a)}{t},$$

en cuyo caso el valor de tal límite (que es un vector de $\mathbb{R}^M$) se nota $f'(a;x)$ y se denomina la <u>*derivada de f en a en la dirección de x*</u>. Para $x = 0$ el anterior límite tiene sentido y vale cero, por lo que es natural definir también $f'(a;0) = 0$. La condición (*) se traduce ahora diciendo que la condición necesaria para que f sea derivable en a es la existencia de $f'(a;.)$, o sea deben existir todas las derivadas direccionales de f en a, siendo en tal caso la candidata a derivada de f en a la aplicación T dada por $T(x) = f'(a;x)$ para todo $x \in \mathbb{R}^N$ (supuesto que sea lineal). Es inmediato (¡Hágase!) que si la función f es derivable en a en las direcciones de los vectores de la esfera unidad, también lo es en la dirección de cualquier vector x, y en tal caso

$$f'(a;x) = \|x\| f'\left(a; \frac{x}{\|x\|}\right), \ \forall x \in \mathbb{R}^N \setminus \{0\}.$$

3.3. Campos escalares derivables. Vector gradiente.

Proposición 3.14 (Reducción de la derivabilidad a campos escalares)**.** *Sea $A \subset \mathbb{R}^N$, y $f = (f_1, \dots, f_M) : A \rightarrow \mathbb{R}^M$ un campo vectorial. Entonces f es derivable en $a \in \overset{\circ}{A}$ si, y sólo si, cada campo escalar componente es derivable en a, en cuyo caso*

$$Df(a)(x) = (Df_1(a)(x), ..., Df_M(a)(x)), \ \forall x \in \mathbb{R}^n.$$

Además $f \in C^1(a)$ si, y sólo si, $f_i \in C^1(a)$ para $i = 1, \dots, M$.

Demostración:

El enunciado es consecuencia de que para una aplicación lineal $T = (T_1, ..., T_M)$ de $\mathbb{R}^N$ en $\mathbb{R}^M$ se verifica para cada $x \in A \backslash \{a\}$

$$\frac{f(x) - f(a) - T(x-a)}{\|x-a\|} =$$

$$= \left(\frac{f_1(x) - f_1(a) - T_1(x-a)}{\|x-a\|}, ..., \frac{f_M(x) - f_M(a) - T_M(x-a)}{\|x-a\|} \right)$$

y de la Proposición 2.62. De la misma proposición se deduce también que la continuidad en a de la aplicación Df se reduce a la continuidad de las aplicaciones $Df_1, ..., Df_M$ en a, ya que son las componentes de la aplicación Df. ∎

En vista del resultado anterior, en el resto de la sección, estudiaremos los campos escalares derivables. Salvo mención expresa en contra, los resultados que se obtengan son válidos también para campos vectoriales.

Definición 3.15 (Derivadas parciales. Vector gradiente)**.** Sea $A \subset \mathbb{R}^N$, $a \in A$ y f un campo escalar en A. Para cada $1 \leq i \leq N$, supongamos que a_i es un punto de acumulación del conjunto

$$A_i := \{x_i \in \mathbb{R} : \ (a_1, \dots a_{i-1}, x_i, a_{i+1}, \dots, a_N) \in A\}.$$

Si la función de A_i en $\mathbb{R}$ definida por $x_i \mapsto f(a_1, \dots a_{i-1}, x_i, a_{i+1}, \dots, a_N)$ es derivable en a_i, se dice que f tiene <u>derivada parcial</u> respecto de la variable i-ésima en el punto a. En tal caso el valor del límite

$$\lim_{x_i \to a_i} \frac{f(a_1, \dots a_{i-1}, x_i, a_{i+1}, \dots, a_N) - f(a)}{x_i - a_i}$$

se denomina la derivada parcial de f respecto de la variable i-ésima en el punto a y se nota $D_i f(a)$ (o también $\frac{\partial f}{\partial x_i}(a)$).

Es consecuencia de la definición que el cálculo de la derivada parcial respecto de la variable i-ésima del campo escalar f en un punto genérico $x = (x_1, ..., x_N)$ se ha de llevar a cabo derivando la función real de variable real que resulta al considerar constantes las variables x_j $(j \neq i)$ y por tanto las reglas de derivación ordinarias se podrán utilizar.

Si $A^i \subset A$ es el conjunto de puntos donde f tiene derivada parcial respecto de la variable i-ésima, entonces el campo escalar en A^i definido por $x \ \mapsto \ D_i f(x)$ se denomina la *aplicación derivada parcial* de f respecto de la variable i-ésima y se nota $D_i f$ (o también $\frac{\partial f}{\partial x_i}$).

Se dice que el campo escalar f tiene *gradiente* en el punto a si admite derivadas parciales en a con respecto de todas las variables, en cuyo caso definimos el *vector gradiente* de f en a por:

$$\nabla f(a) := (D_1 f(a), ..., D_N f(a)) \in \mathbb{R}^N.$$

Si C es el conjunto de puntos de A donde f tiene gradiente, entonces el campo vectorial en C definido por $x \to \nabla f(x)$ se denomina *aplicación gradiente* de f y se nota ∇f.

Proposición 3.16 (Condición necesaria de deriv. y candidata a derivada)**.** *Sean $A \subset \mathbb{R}^N$ y f un campo escalar en A derivable en un punto a. Entonces f tiene gradiente en a con*

$$D_i f(a) = Df(a)(e_i), \ \forall i \in \{1, \ldots, N\},$$

donde $\{e_1, \ldots, e_N\}$ es la base canónica de $\mathbb{R}^N$.
En consecuencia,

$$Df(a)(x) = (\nabla f(a) \mid x), \ \forall x \in \mathbb{R}^N.$$

Demostración:

Para $i \in \{1, \ldots, N\}$ se tiene que

$$Df(a)(e_i) = \lim_{t \to 0} \frac{f(a + te_i) - f(a)}{t} = \lim_{x_i \to a_i} \frac{f(a + (x_i - a_i)e_i) - f(a)}{x_i - a_i} =$$

$$\lim_{x_i \to a_i} \frac{f(a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_N) - f(a)}{x_i - a_i} = D_i f(a),$$

donde se ha utilizado (*) de la sección anterior. El resto es consecuencia de la linealidad de $Df(a)$. ■

Notas 3.17.

1. La condición anterior no es suficiente. El campo escalar en $\mathbb{R}^2$

$$f(x,y) = \frac{x^6}{(y - x^2)^2 + x^6} \ \ (x,y) \neq (0,0), \quad f(0,0) = 0$$

   tiene gradiente en $(0,0)$ igual a $(0,0)$ y, sin embargo, no es ni tan siquiera continuo en $(0,0)$ (tómese $y = x^2$).

2. Supongamos que el campo escalar f es derivable en a y tiene derivada no nula. Puesto que

$$Df(a)(x) = (\nabla f(a) \mid x), \quad \forall x \in \mathbb{R}^N$$

se deduce de la desigualdad de Cauchy-Schwarz (Apéndice A) que

$$|Df(a)(x)| \leq \left\|\nabla f(a)\right\|_2 \|x\|_2$$

y que, para $x \in \mathbb{R}^N \backslash \{0\}$, la igualdad se alcanza si, y sólo si, existe $\lambda \in \mathbb{R}^*$ tal que $x = \lambda \nabla f(a)$, de donde se deduce que la derivada de f en el punto a es máxima en la dirección dada por el vector gradiente.

3. Obsérvese que las derivadas parciales no son otra cosa que las derivadas direccionales según los vectores de la base canónica. Es claro que pueden existir las derivadas parciales y no existir una derivada direccional (véase el Ejercicio 3.4).

Proposición 3.18 (Condición suficiente de derivabilidad)**.** *Sean $A \subset \mathbb{R}^N$, $a \in \overset{\circ}{A}$ y f un campo escalar en A. Supongamos que ∇f existe en un entorno de a y que ∇f es continuo en a. Entonces f es derivable en a.*

Demostración:

Dado $\varepsilon > 0$, por hipótesis, existe $\delta > 0$ tal que

$$\|x-a\|_1 < \delta \Rightarrow \begin{cases} x \in A \\ \exists \nabla f(x) \\ |D_i f(x) - D_i f(a)| \leq \varepsilon, \text{ para } i = 1, \ldots, N. \end{cases}$$

Para $x = (x_1, \ldots, x_N) \in \mathbb{R}^N$ tal que $0 < \|x-a\|_1 < \delta$ podemos escribir

$$f(x) - f(a) - (\nabla f(a) \, |x-a) =$$

$$\sum_{i=1}^{N} \big(f(a_1, \ldots, a_{i-1}, x_i, x_{i+1}, \ldots, x_N) - f(a_1, \ldots, a_{i-1}, a_i, x_{i+1}, \ldots, x_N) - D_i f(a)(x_i - a_i) \big)$$

luego, si probamos que para cada $i \in \{1, \ldots, N\}$ se verifica que

$$|f(a_1, \ldots, a_{i-1}, x_i, x_{i+1}, \ldots, x_N) - f(a_1, \ldots, a_{i-1}, a_i, x_{i+1}, \ldots, x_N) - D_i f(a)(x_i - a_i)| \leq$$
$$\leq \varepsilon |x_i - a_i|,$$

entonces tendremos que

$$|f(x) - f(a) - (\nabla f(a)|x-a)| \leq \sum_{i=1}^{N} \varepsilon |x_i - a_i| = \varepsilon \|x-a\|_1$$

y en consecuencia f será derivable en a.

Veamos que son ciertas las citadas desigualdades. Si $x_i = a_i$, la desigualdad es obvia. Supuesto $x_i \neq a_i$, podemos aplicar el Teorema del valor medio para encontrar y_i entre x_i y a_i tal que

$$f(a_1, \ldots, a_{i-1}, x_i, x_{i+1}, \ldots, x_N) - f(a_1, \ldots, a_{i-1}, a_i, x_{i+1}, \ldots, x_N) =$$
$$D_i f(a_1, \ldots, a_{i-1}, y_i, x_{i+1}, \ldots, x_N)(x_i - a_i),$$

de donde se deduce la desigualdad anunciada a partir de las condiciones impuestas a δ. ∎

Ejemplo 3.19. La condición suficiente del anterior resultado no es necesaria. En efecto, el campo escalar $f:\mathbb{R}^2 \to \mathbb{R}$ definido por

$$f(x,y) = (x^2+y^2)\ \operatorname{sen}\left(\frac{1}{\sqrt{x^2+y^2}}\right),\ \forall (x,y)\in\mathbb{R}^2\setminus\{(0,0)\}, \quad f(0,0)=0,$$

es derivable en $(0,0)$. En efecto, es inmediato que

$$D_1 f(0,0) = D_2 f(0,0) = 0$$

y que

$$\frac{f(x,y)}{\sqrt{x^2+y^2}} \to 0,\ \text{ cuando }\ (x,y)\to(0,0).$$

Sin embargo ∇f no es continuo en $(0,0)$. En efecto, como para $(x,y)\neq(0,0)$ se tiene

$$D_1 f(x,y) = 2x\ \operatorname{sen}\left(\frac{1}{\sqrt{x^2+y^2}}\right) - \frac{x}{\sqrt{x^2+y^2}}\ \cos\left(\frac{1}{\sqrt{x^2+y^2}}\right),$$

basta observar que

$$D_1 f\left(\frac{1}{n\pi},0\right) = \frac{2}{n\pi}\ \operatorname{sen}(n\pi) - \cos(n\pi) = (-1)^{n+1},\ \forall n\in\mathbb{N}$$

para concluir que $D_1 f$ no es continua en $(0,0)$. Como $f(x,y)=f(y,x)$, para cualquier (x,y), igual le ocurre a $D_2 f$.

Probaremos enseguida que un campo escalar es de clase C^1 si, y sólo si, tiene gradiente continuo. Sin embargo, no hemos obtenido ninguna caracterización de la derivabilidad de un campo escalar en términos del gradiente. A continuación resumimos toda la información para el estudio de la derivabilidad de campos escalares.

Nota 3.20 (Estudio de la derivabilidad de campos escalares)**.** Sean $A\subset\mathbb{R}^N$, $a\in\overset{\circ}{A}$ y f un campo escalar es A. Para el estudio de la derivabilidad de f en a se sugiere seguir los siguientes pasos:

1. Existencia de $\nabla f(a)$. Si no existe $\nabla f(a)$ entonces f no es derivable en a. En caso contrario, la candidata a derivada es la aplicación $x\mapsto(\nabla f(a)\mid x)$.

2. Continuidad de f en a. Si f no es continua en a, entonces no es derivable en a. De ser f continua en a, entonces f *puede ser no derivable* en a (véase la función g del Ejercicio 3.3, que es continua, con gradiente y no es derivable en $(0,0)$).

3. Continuidad de las derivadas parciales. Si ∇f es continuo en a entonces f es derivable en a. No se puede deducir de esto que f es de clase $\mathscr{C}^1$ en a, salvo que se globalice como se prueba en el teorema siguiente.

4. <u>Definición de derivabilidad</u>. Según que la expresión

$$\frac{f(x)-f(a)-(\nabla f(a)\mid x-a)}{\|x-a\|}$$

tienda a cero o no, cuando $x\to a$, f es derivable o no.

Teorema 3.21 (Caracterización de los campos escalares de clase $\mathscr{C}^1$)**.** *Sean A un abierto de $\mathbb{R}^N$ y f un campo escalar en A. Las siguientes afirmaciones son equivalentes:*

i) $f\in\mathscr{C}^1(A)$.

ii) $\nabla f\in\mathscr{C}(A)$, *esto es, f tiene gradiente en cada punto de A y el campo vectorial ∇f es continuo.*

Demostración:

i) $\Rightarrow$ ii) Puesto que f es derivable, por la Proposición 3.16 se sigue que f tiene gradiente en cada punto de A, y además para cada $1\le i\le N$, se tiene

$$D_i f(a)=Df(a)(e_i),\ \forall a\in A,$$

luego

$$D_i f=E_{e_i}\circ Df,$$

donde $E_{e_i}:\mathscr{L}(\mathbb{R}^N,\mathbb{R})\longrightarrow\mathbb{R}$ es la evaluación en el vector e_i, esto es,

$$E_{e_i}(T)=T(e_i),\ \forall T\in L(\mathbb{R}^N,\mathbb{R}),$$

aplicación que es lineal y (automáticamente) continua. En vista de la regla de la cadena para funciones continuas, obtenemos que $D_i f$ es continua. Finalmente, como las funciones $D_i f$ son las componentes de ∇f, concluimos que ∇f es continuo a.

ii) $\Rightarrow$ i) Por la Proposición 3.18, f es derivable. Además se tiene que

$$Df(a)=D_1 f(a)\pi_1+\cdots+D_N f(a)\pi_N,$$

donde π_k es la proyección k-ésima de $\mathbb{R}^N$, para $1\le k\le N$.
Equivalentemente,

$$Df=\alpha_1\circ D_1 f+\cdots+\alpha_N\circ D_N f,$$

donde, para cada $i=1,2,\cdots,N$, $\alpha_i:\mathbb{R}\longrightarrow\mathscr{L}(\mathbb{R}^N,\mathbb{R})$ es la aplicación lineal (continua) definida por $\alpha_i(t)=t\pi_i$. Basta, en consecuencia, utilizar la regla de la cadena para funciones continuas y que la suma de funciones continuas es continua para concluir que $Df:A\longrightarrow\mathscr{L}(\mathbb{R}^N,\mathbb{R})$ es continua. ∎

3.4. Campos vectoriales derivables. Matriz jacobiana.

Definición 3.22 (matriz jacobiana)**.** Sean $A \subset \mathbb{R}^N$, $a \in A$ y $f : A \to \mathbb{R}^M$ un campo vectorial. Se dice que f tiene matriz jacobiana en el punto a si cada uno de los campos escalares componentes tiene gradiente en dicho punto, en cuyo caso se define la *matriz jacobiana* de f en a por

$$J_f(a) := \begin{pmatrix} D_1 f_1(a) & \dots & D_j f_1(a) & \dots & D_N f_1(a) \\ \dots & \dots & \dots & \dots & \dots \\ D_1 f_i(a) & \dots & D_j f_i(a) & \dots & D_N f_i(a) \\ \dots & \dots & \dots & \dots & \dots \\ D_1 f_M(a) & \dots & D_j f_M(a) & \dots & D_N f_M(a) \end{pmatrix} = \big(D_j f_i(a)\big)_{\substack{1 \le i \le M \\ 1 \le j \le N}}.$$

En el caso $M = N$ al determinante de la matriz jacobiana se le denomina determinante jacobiano.

Obsérvese que los N números que componen la fila i-ésima de la matriz jacobiana son las componentes del vector $\nabla f_i(a)$, en consecuencia, en virtud de las Proposición 3.14 y 3.18, si el campo vectorial admite matriz jacobiana continua en el punto a, entonces es derivable en dicho punto y, en virtud de la Proposición 3.14 y del Teorema 3.21, si A es abierto y el campo escalar admite matriz jacobiana continua, entonces es de clase $\mathscr{C}^1$. La condición necesaria de derivabilidad y la candidata a derivada se codifican en el siguiente resultado.

Proposición 3.23. *Sean $A \subset \mathbb{R}^N$ y $f : A \to \mathbb{R}^M$ un campo vectorial. Si f es derivable en $a \in \overset{\circ}{A}$, entonces f tiene matriz jacobiana en a y*

$$(Df(a)(x))^t = J_f(a)x^t, \ \forall x \in \mathbb{R}^N.$$

Demostración:

Los campos escalares componentes de f son derivables en a. En consecuencia, en virtud de la Proposición 3.16, dichos campos escalares tienen gradiente en a, es decir f tiene matriz jacobiana en a. Se tiene además que

$$A_{Df(a)} = \big(Df_i(a)(e_j)\big)_{\substack{1 \le i \le M \\ 1 \le j \le N}} = \big(D_j f_i(a)\big)_{\substack{1 \le i \le M \\ 1 \le j \le N}},$$

donde se ha utilizado la Proposición 3.14. El resto es consecuencia de la igualdad (3.1.2) de la sección 3.1. ∎

Ejemplos 3.24.

1) Toda función constante de $\mathbb{R}^N$ en $\mathbb{R}^M$ es de clase $\mathscr{C}^1$ con

$$Df(a) = 0, \quad \forall a \in \mathbb{R}^N.$$

2) Toda aplicación lineal T de $\mathbb{R}^N$ en $\mathbb{R}^M$ es de clase $\mathscr{C}^1$ con

$$DT(a) = T, \quad \forall a \in \mathbb{R}^N.$$

3) Toda aplicación bilineal T de $\mathbb{R}^M \times \mathbb{R}^N$ en $\mathbb{R}^P$ (es decir, lineal en ambas variables) es de clase $\mathscr{C}^1$ con derivada dada por

$$DT(a,b)(x,y) = T(x,b) + T(a,y), \ \forall (a,b),(x,y) \in \mathbb{R}^M \times \mathbb{R}^N.$$

Calculemos en primer lugar las derivadas de T en las direcciones de la base canónica. Para ello si $\{e_1,\ldots,e_M\}$ es la base canónica de $\mathbb{R}^M$, y $\{f_1,\ldots,f_N\}$ es la base canónica de $\mathbb{R}^N$, entonces $\{(e_1,0),\ldots,(e_M,0),(0,f_1),\ldots,(0,f_N)\}$ es la base canónica de $\mathbb{R}^M \times \mathbb{R}^N$. Claramente se tiene

$$\frac{T((a,b)+t(e_i,0)) - T(a,b)}{t} = \frac{T((a+te_i,b)) - T(a,b)}{t} = \frac{T(a,b) + tT(e_i,b)) - T(a,b)}{t} = T(e_i,b) \quad (i = 1,\ldots,M),$$

es decir,

$$T'\left((a,b);(e_i,0)\right) = T(e_i,b) \quad (i = 1,\ldots,M)$$

Haciendo el mismo tipo de cálculo obtenemos que

$$T'\big((a,b);(0,f_j)\big) = T(a,f_j) \quad (j = 1,\ldots,N).$$

Hemos probado que

$$J_T(a,b) = \begin{pmatrix} T_1(e_1,b) & \ldots & T_1(e_M,b) & T_1(a,f_1) & \ldots & T_1(a,f_N) \\ T_2(e_1,b) & \ldots & T_2(e_M,b) & T_2(a,f_1) & \ldots & T_2(a,f_N) \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ T_P(e_1,b) & \ldots & T_P(e_M,b) & T_P(a,f_1) & \ldots & T_P(a,f_N) \end{pmatrix}.$$

Por tanto de ser T derivable en (a,b), su derivada en (a,b) habría de ser, en virtud de la Proposición 3.23,

$$(x,y) \longmapsto \left(J_T(a,b)(x,y)^t\right)^t = T(x,b) + T(a,y).$$

Como la matriz jacobiana es continua, basta tener en cuenta el comentario que sigue a la Definición 3.22, para concluir que $T \in \mathscr{C}^1(\mathbb{R}^M \times \mathbb{R}^N)$.

Otra forma:

Una aplicación bilineal T verifica que

$$\|T(x,y)\| \leq \|T\|\|x\|\|y\|, \ \forall (x,y) \in \mathbb{R}^M \times \mathbb{R}^N,$$

donde

$$\|T\| = \max\{\|T(x,y)\| : \ \|x\| = \|y\| = 1\}$$

(¡Pruébese!). Así

$$\frac{\|T(x,y) - T(a,b) - T(x-a,b) - T(a,y-b)\|}{\|(x-a,y-b)\|} = \frac{\|T(x-a,y-b)\|}{\|(x-a,y-b)\|} \leq$$
$$\frac{\|T\|\|x-a\|\|y-b\|}{\|(x-a,y-b)\|} \leq \frac{\|T\|\|(x-a,y-b)\|^2}{\|(x-a,y-b)\|} = \|T\|\|(x-a,y-b)\|,$$

donde se ha utilizado la desigualdad comentada. Basta tomar límite en (a,b) para concluir que T es derivable en (a,b) y su derivada es la aplicación que se ha enunciado antes. La aplicación $DT : \mathbb{R}^{M+N} \equiv \mathbb{R}^M \times \mathbb{R}^N \longrightarrow \mathscr{L}(\mathbb{R}^M \times \mathbb{R}^N, \mathbb{R}^P)$ es claramente lineal (por la bilinealidad de T), por tanto hemos probado que $T \in \mathscr{C}^1(\mathbb{R}^M \times \mathbb{R}^N)$.

4) La aplicación suma $\sigma : \mathbb{R}^N \times \mathbb{R}^N \longrightarrow \mathbb{R}^N$ dada por

$$\sigma(a,b) = a+b, \quad \forall a,b \in \mathbb{R}^N,$$

y la aplicación producto por escalares dada por

$$\pi(\lambda,a) = \lambda a, \ \ \forall \lambda \in \mathbb{R}, a \in \mathbb{R}^N,$$

son de clase C^1, por ser, en el primer caso lineal, y en el segundo una aplicación bilineal. Por tanto, se tiene

$$D\sigma(a,b) = \sigma, \ \forall a,b \in \mathbb{R}^N, \quad D\pi(\alpha,a)(\lambda,x) = \alpha x + \lambda a, \ \forall \alpha,\lambda \in \mathbb{R}, \ \forall a,x \in \mathbb{R}^N.$$

5) La aplicación inversión $J : \text{Iso}\ (\mathbb{R}^N) \longrightarrow \mathscr{L}(\mathbb{R}^N)$ dada por

$$J(T) = T^{-1}, \ \forall T \in \text{Iso}\ (\mathbb{R}^N)$$

es de clase $\mathscr{C}^1$ con derivada definida por

$$DJ(T)(S) = -T^{-1}ST^{-1}, \ \forall S \in \mathscr{L}(\mathbb{R}^N).$$

En efecto, es claro que la aplicación que se anuncia como $DJ(T)$ es una aplicación lineal. Para $S,T \in \text{Iso}\ (\mathbb{R}^N)$ se verifica que

$$S^{-1} - T^{-1} - [-T^{-1}(S-T)T^{-1}] = S^{-1} - T^{-1} + T^{-1}(S-T)T^{-1} =$$

$$= S^{-1}(T-S)T^{-1} + T^{-1}(S-T)T^{-1} = (T^{-1} - S^{-1})(S-T)T^{-1}.$$

Así,

$$\frac{\left\|S^{-1}-T^{-1}-\left(-T^{-1}(S-T)T^{-1}\right)\right\|}{\|S-T\|} \leq \|T^{-1}-S^{-1}\| \, \|T^{-1}\| \to 0 \quad (S\to T),$$

donde se ha utilizado la continuidad de J.

Finalmente, veamos que

$$DJ : \text{Iso}\,(\mathbb{R}^N) \longrightarrow \mathscr{L}(\mathscr{L}(\mathbb{R}^N),\mathscr{L}(\mathbb{R}^N))$$

es continua. Escribamos $DJ = \Phi\circ(J,J)$, donde $\Phi : \mathscr{L}(\mathbb{R}^N)\times\mathscr{L}(\mathbb{R}^N) \to \mathscr{L}(\mathscr{L}(\mathbb{R}^N))$ (se entiende $\mathscr{L}(\mathscr{L}(\mathbb{R}^N),\mathscr{L}(\mathbb{R}^N))$) está definida por

$$\Phi(F,G)(S) = -FSG,\ \forall F,G,S\in\mathscr{L}(\mathbb{R}^N).$$

En efecto, para $T\in \text{Iso}\,(\mathbb{R}^N)$ y $S\in\mathscr{L}(\mathbb{R}^N)$, tenemos que

$$(\Phi\circ(J,J))(T)(S) = \Phi(T^{-1},T^{-1})(S) = -T^{-1}ST^{-1} = DJ(T)(S).$$

Como Φ es una aplicación bilineal (entre espacios de dimensión finita), Φ es continua. De la continuidad de J se sigue la continuidad de la función vectorial (J,J), luego DJ es continua en virtud de la regla de la cadena para funciones continuas.

6) Cualquier norma $\|\cdot\| : \mathbb{R}^N \longrightarrow \mathbb{R}$ no es derivable en 0 (lo que generaliza la no derivabilidad del valor absoluto en 0). En efecto, si x es un vector no nulo, entonces si $t\in\mathbb{R}^*$ se tiene que

$$\frac{\|tx\|-0}{t} = \frac{|t|}{t}\|x\|,$$

función que no tiene límite cuando $t\to 0$. Luego no existe ninguna derivada direccional.

3.5. Reglas de derivación.

1) **Linealidad**.
Sean $A \subset \mathbb{R}^N$ y $f, g : A \longrightarrow \mathbb{R}^M$ funciones derivables en a y $\lambda \in \mathbb{R}$. Entonces $f+g$ y λf son derivables en a con

$$D(f+g)(a) = Df(a) + Dg(a)$$

$$D(\lambda f)(a) = \lambda Df(a).$$

Además, si $f, g \in \mathscr{C}^1(a)$, entonces $f+g, \lambda f \in \mathscr{C}^1(a)$.
La comprobación se deja como ejercicio.

2) **Regla de la cadena.**
Sean $A \subset \mathbb{R}^N, B \subset \mathbb{R}^M, f : A \longrightarrow \mathbb{R}^M$ tal que $f(A) \subset B$ y $g : B \longrightarrow \mathbb{R}^P$. Supongamos que f es derivable en a y que g es derivable en $f(a)$. Entonces la composición $h = g \circ f$ es derivable en a con

$$Dh(a) = Dg(f(a)) \circ Df(a)$$

y en consecuencia

$$J_h(a) = J_g(f(a)) J_f(a).$$

Además, si $f \in \mathscr{C}^1(a), g \in \mathscr{C}^1(f(a))$, entonces $h \in \mathscr{C}^1(a)$.

Demostración:

Notemos $b = f(a)$. Puesto que f es derivable en a y g es derivable en b, existen funciones

$$r : A \longrightarrow \mathbb{R}, \quad s : B \longrightarrow \mathbb{R},$$

tales que r es continua en a con $r(a) = 0$, s es continua en b con $s(b) = 0$, y

$$\begin{cases} \|f(x) - f(a) - Df(a)(x-a)\| = r(x)\|x-a\|, \quad \forall x \in A \\ \\ \|g(y) - g(b) - Dg(b)(y-b)\| = s(y)\|y-b\|, \quad \forall y \in B. \end{cases}$$

Para $x \in A$ se tiene

$$\begin{gathered} \|h(x) - h(a) - (Dg(b) \circ Df(a))(x-a)\| = \\ \|g(f(x)) - g(b) - Dg(b)(f(x) - b) + Dg(b)(f(x) - f(a) - Df(a)(x-a))\| \leq \\ \|g(f(x)) - g(b) - Dg(b)(f(x) - b)\| + \|Dg(b)(f(x) - f(a) - Df(a)(x-a)\| \leq \\ s(f(x))\ \|f(x) - f(a)\| + \|Dg(b)\|\ r(x)\ \|x-a\| = \\ s(f(x))\ \|f(x) - f(a) - Df(a)(x-a) + Df(a)(x-a)\| + \|Dg(b)\|\ r(x)\ \|x-a\| \leq \\ s(f(x)) \big[\|f(x) - f(a) - Df(a)(x-a)\| + \|Df(a)\| \|x-a\| \big] + \|Dg(b)\| r(x) \|x-a\| = \\ s(f(x)) \big[\ r(x)\ \|x-a\| +\ \|Df(a)\| \|x-a\| \big] + \|Dg(b)\|\ r(x)\ \|x-a\| \end{gathered}$$

de donde se sigue el resultado, pues $\lim_{x \to a} r(x) = 0$, y, como f es continua en a, también $\lim_{x \to a} s(f(x)) = 0$.

Supuesta la continuidad de Df en a y la de Dg en b, probaremos ahora la continuidad de Dh en a. Por ser g de clase $\mathscr{C}^1$ en b, existe $\varepsilon > 0$ tal que g es derivable en $B(b,\varepsilon) \subset B$; usando que $f \in \mathscr{C}^1$ en a, existe $\delta_1 > 0$ tal que f es derivable en $B(a,\delta_1) \subset A$. La continuidad de f en a nos asegura la existencia de $\delta_2 > 0$ tal que $f(B(a,\delta_2)) \subset B(b,\varepsilon)$. Tomando $\delta = \min\{\delta_1,\delta_2\}$ tenemos, aplicando la fórmula que ya hemos probado

$$Dh(x) = Dg(f(x)) \circ Df(x), \ \forall x \in B(a,\delta).$$

En consecuencia, $Dh = \Phi \circ (Dg \circ f, Df)$, donde Φ es la aplicación bilineal de $\mathscr{L}(\mathbb{R}^M \times \mathbb{R}^p) \times \mathscr{L}(\mathbb{R}^N,\mathbb{R}^M)$ en $\mathscr{L}(\mathbb{R}^N,\mathbb{R}^p)$ definida por

$$\Phi(T,S) = T \circ S, \ \ \forall (T,S) \in \mathscr{L}(\mathbb{R}^M,\mathbb{R}^p) \times \mathscr{L}(\mathbb{R}^N,\mathbb{R}^M).$$

Así pues, Dh es continua por la regla de la cadena para funciones continuas y la regla de continuidad de las funciones vectoriales ($Dg \circ f$ es continua en a pues f es continua en a y Dg es continua en $f(a)$ y Df es continua en a). ∎

Notas 3.25.

a) Es interesante observar la forma que adopta la regla de la cadena cuando se involucran derivadas elementales.

i) En el caso $A \subset \mathbb{R} \xrightarrow{f} B \subset \mathbb{R}^M \xrightarrow{g} \mathbb{R}^P$ se tiene que

$$(g \circ f)'(a) = Dg(f(a))(f'(a)).$$

En efecto,

$$(g \circ f)'(a) = D(g \circ f)(a)(1) = Dg(f(a))(Df(a)(1)) = Dg(f(a))(f'(a)),$$

donde se ha utilizado la Nota 3.13.2.

ii) En el caso

$$A \subset \mathbb{R} \xrightarrow{f} B \subset \mathbb{R} \xrightarrow{g} \mathbb{R}$$

se tiene que

$$(g \circ f)'(a) = f'(a)g'(f(a)),$$

y por tanto, la regla de la cadena recién enunciada generaliza la conocida para funciones de variable real. En efecto, por el caso anterior

$$(g \circ f)'(a) = Dg(f(a))(f'(a)) = f'(a)Dg(f(a))(1) = f'(a)g'(f(a)),$$

donde se ha vuelto a utilizar la Nota 3.13.2.

b) Veamos ahora la regla de la cadena para las derivadas parciales.

Las entradas de $J_h(a)$ vienen dadas por

$$D_j h_i(a) = \sum_{k=1}^{M} D_k g_i(f(a)) D_j f_k(a), \qquad (1 \le i \le P,\ 1 \le j \le N),$$

expresión que se conoce como *regla de la cadena para las derivadas parciales.*

Si consideramos el caso $P = 1$, con lo que g y h son campos escalares, la anterior expresión se escribe

$$D_j h(a) = \sum_{k=1}^{M} D_k g(f(a)) D_j f_k(a), \qquad (1 \le j \le N),$$

o bien, si notamos por $x_1, \cdots, x_N$ las coordenadas en $\mathbb{R}^N$, y por $y_1, \cdots, y_M$ las coordenadas en $\mathbb{R}^M$, se tendrá

$$\frac{\partial h}{\partial x_j}(a) = \sum_{k=1}^{M} \frac{\partial g}{\partial y_k}(f(a)) \frac{\partial f_k}{\partial x_j}(a), \qquad (1 \le j \le N). \qquad (*)$$

Si se identifican las variables con las funciones, es decir,

$$z \equiv z(y_1, \cdots, y_M),$$
$$w \equiv z(y_1(x_1, \cdots, x_N), \cdots, y_M(x_1, \cdots, x_N)) = w(x_1, \cdots, x_N)$$

queda finalmente

$$\frac{\partial w}{\partial x_j}(a) = \sum_{k=1}^{M} \frac{\partial z}{\partial y_k}(f(a)) \frac{\partial y_k}{\partial x_j}(a), \qquad (1 \le j \le N),$$

expresión en la que se debe saber reconocer $(*)$.

Ejemplo. Sea $z = z(x, y)$ un campo escalar derivable. Si cambiamos a coordenadas polares, esto es, hacemos

$$\begin{cases} x = \rho \cos \vartheta \\ y = \rho \operatorname{sen} \vartheta \end{cases},$$

entonces la función $w(\rho, \vartheta) := z(\rho \cos \vartheta, \rho \operatorname{sen} \vartheta)$ es derivable y sus derivadas parciales vienen dadas por

$$\begin{cases} \frac{\partial w}{\partial \rho} = \frac{\partial z}{\partial x}(\rho \cos \vartheta, \rho \operatorname{sen} \vartheta) \cos \vartheta + \frac{\partial z}{\partial y}(\rho \cos \vartheta, \rho \operatorname{sen} \vartheta) \operatorname{sen} \vartheta \\ \\ \frac{\partial w}{\partial \vartheta} = -\frac{\partial z}{\partial x}(\rho \cos \vartheta, \rho \operatorname{sen} \vartheta) \rho \operatorname{sen} \vartheta + \frac{\partial z}{\partial y}(\rho \cos \vartheta, \rho \operatorname{sen} \vartheta) \rho \cos \vartheta \end{cases}$$

Proposición 3.26 (Carácter local de la derivabilidad y de la derivada)**.** *Sea $A \subset \mathbb{R}^N, a \in \overset{\circ}{A}$ y $f : A \longrightarrow \mathbb{R}^M$. Las siguientes afirmaciones son equivalentes:*

i) f es derivable en a.

ii) $f_{|U}$ es derivable en a para algún entorno $U \subset A$ del punto a.

Además, en caso de que sean ciertas las afirmaciones anteriores, entonces las derivadas de ambas funciones (f y la restricción de f a U) coinciden.

Demostración:

Es consecuencia del carácter local del límite. ■

Nota. Supongamos que el conjunto A es unión disjunta de dos subconjuntos A_1 y A_2 y $f_1 : A_1 \longrightarrow \mathbb{R}^M, f_2 : A_2 \longrightarrow \mathbb{R}^M$ son campos vectoriales tales que

$$f(x) = \begin{cases} f_1(x) & \text{si} \quad x \in A_1 \\ f_2(x) & \text{si} \quad x \in A_2 \end{cases}$$

Por el resultado anterior, la derivabilidad de f en $\overset{\circ}{A_1}$ y en $\overset{\circ}{A_2}$ no es más que la derivabilidad de f_1 en $\overset{\circ}{A_1}$ y de f_2 en $\overset{\circ}{A_2}$, respectivamente (cuyo estudio posiblemente se puede llevar a cabo mediante el uso de las reglas de derivación). En los puntos de $(\mathrm{Fr}(A_1) \cup \mathrm{Fr}(A_2)) \cap A$, posiblemente haya que hacer un estudio particular.

Proposición 3.27 (Derivación de la función inversa)**.**

a) Sea $A \subset \mathbb{R}^N, a \in \overset{\circ}{A}$ y $f : A \longrightarrow \mathbb{R}^N$ un campo vectorial. Supongamos que f es inyectiva, derivable en a y que $f(a) \in f(\overset{\circ}{A})$, entonces

$$f^{-1} \text{ es derivable en } f(a) \Leftrightarrow \begin{cases} \det J_f(a) \neq 0 \\ f^{-1} \text{ es continua en } f(a) \end{cases}$$

Además, si se verifica lo anterior, se tiene

$$Df^{-1}(f(a)) = Df(a)^{-1},$$

y en consecuencia,

$$J_{f^{-1}}(f(a)) = J_f(a)^{-1}.$$

b) Sean A y B subconjuntos abiertos de $\mathbb{R}^N$ y f un homeomorfismo de A sobre B. Si $f \in \mathscr{C}^1(a)$ para algún $a \in A$ con $\det J_f(a) \neq 0$*, entonces $f^{-1} \in \mathscr{C}^1(f(a))$.*

Demostración:

a) $\Rightarrow$ f^{-1} es continua en $f(a)$ por ser derivable. Como

$$f \circ f^{-1} = Id_{f(A)}, \quad f^{-1} \circ f = Id_A,$$

la regla de la cadena nos da

$$Df(a) \circ Df^{-1}(f(a)) = Id_{\mathbb{R}^N}, \quad Df^{-1}(f(a)) \circ Df(a) = Id_{\mathbb{R}^N}$$

con lo que $Df(a) \in \mathrm{Iso}\ (\mathbb{R}^N)$ y $Df^{-1}(f(a)) = Df(a)^{-1}$, por tanto $\det J_f(a) \neq 0$.

$\Leftarrow$ Al ser f derivable en a, existe una función $r : A \longrightarrow \mathbb{R}$ continua en a con $r(a) = 0$ que verifica además

$$\|f(x) - f(a) - Df(a)(x-a)\| = r(x)\ \|x-a\|,\ \forall x \in A.$$

Se tiene que

$$
\begin{aligned}
x-a &= Df(a)^{-1}(Df(a)(x-a)) \\
&= Df(a)^{-1}\Big(Df(a)(x-a)-(f(x)-f(a))+(f(x)-f(a)\Big) \\
&= -Df(a)^{-1}\Big(f(x)-f(a)-Df(a)(x-a))\Big)+Df(a)^{-1}(f(x)-f(a).
\end{aligned} \tag{3.5.1}
$$

En consecuencia,

$$x-a-Df(a)^{-1}(f(x)-f(a)) = -Df(a)^{-1}(f(x)-f(a)-Df(a)(x-a))$$

y por tanto

$$\|x-a-Df(a)^{-1}(f(x)-f(a))\| \leq \|Df(a)^{-1}\|\ r(x)\ \|x-a\|. \tag{3.5.2}$$

También se sigue de 3.5.1 que

$$\|x-a\| \leq \|Df(a)^{-1}\|\ r(x)\ \|x-a\| + \|Df(a)^{-1}\|\ \|f(x)-f(a)\| \Rightarrow$$

$$\Rightarrow \Big(1-r(x)\ \|Df(a)^{-1}\|\Big)\ \|x-a\| \leq \|Df(a)^{-1}\|\ \|f(x)-f(a)\|$$

y cuando $r(x) < \dfrac{1}{\|Df(a)^{-1}\|}$ tendremos

$$\|x-a\| \leq \frac{\|Df(a)^{-1}\|}{1-r(x)\ \|Df(a)^{-1}\|}\ \|f(x)-f(a)\|$$

y por tanto, en vista de 3.5.2,

$$\|x-a-Df(a)^{-1}(f(x)-f(a))\| \leq r(x)\frac{\|Df(a)^{-1}\|^2}{1-r(x)\ \|Df(a)^{-1}\|}\ \|f(x)-f(a)\|$$

de donde se sigue el resultado, pues $\lim_{x\to a} r(x) = 0$ (¿Por qué?).

b) Como Iso $(\mathbb{R}^N)$ es un abierto por la Proposición 3.6, la continuidad de la aplicación $x \mapsto Df(x)$ en el punto a garantiza la existencia de un abierto U de $\mathbb{R}^N$ tal que

$$a \in U \subset A \quad \text{y} \quad Df(x) \in \text{Iso}\ (\mathbb{R}^N),\ \forall x \in U.$$

El apartado a) asegura que f^{-1} es derivable en cada punto de $V := f(U)$ (que es un abierto de $\mathbb{R}^N$ tal que $f(a) \in V \subset B$) y

$$Df^{-1}(y) = (Df(f^{-1}(y)))^{-1},\ \forall y \in V.$$

En consecuencia, $Df^{-1} = J \circ Df \circ f^{-1}$, lo que prueba la continuidad de Df^{-1} en $f(a)$ y, por tanto, $f^{-1} \in \mathscr{C}^1(f(a))$. ∎

Notas 3.28.

1. Este resultado generaliza para puntos interiores el visto para funciones reales de una variable real. En efecto, si $I \subset \mathbb{R}$ es un intervalo y $f : I \longrightarrow \mathbb{R}$ es una función (real de variable real), $a \in \overset{\circ}{I}$ y $b = f(a)$.

   Supongamos que f es continua, inyectiva y derivable en a. Entonces b es un punto interior de $f(I)$ y las siguientes afirmaciones son equivalentes:

   i) $f'(a) \neq 0$ y f^{-1} es continua en b.

   ii) f^{-1} es derivable en b.

   Además, en caso de que se cumplan i) y ii) se tiene: $(f^{-1})'(b) = \frac{1}{f'(a)}$. En efecto,

$$Df(a) \in \text{Iso }(\mathbb{R}) \Leftrightarrow f'(a) \neq 0 \ \text{ y } \ Df^{-1}(b) = Df(a)^{-1} \Leftrightarrow (f^{-1})'(b) = \frac{1}{f'(a)}.$$

2. Es bueno resaltar que la regla antes establecida no es el Teorema de la función inversa.

3.6. Interpretación geométrica del concepto de derivada. Hiperplano tangente.

Definición 3.29. Una *variedad afín* es el trasladado por un vector de un subespacio vectorial.

Sean $A \subset \mathbb{R}^N$ y $f : A \to \mathbb{R}^M$. Si f es derivable en el punto a, entonces la variedad afín de $\mathbb{R}^N \times \mathbb{R}^M$ que pasa por el punto $(a, f(a))$ con variedad de dirección Graf $(Df(a))$, esto es

$$(a, f(a)) + \text{Graf } (Df(a))$$

que coincide con la gráfica de la aproximación afín

$$g(x) = f(a) + Df(a)(x-a)$$

En efecto,

$$\begin{aligned}\{(x, g(x)) : x \in \mathbb{R}^N\} &= \{(a + (x-a), f(a) + Df(a)(x-a)) : x \in \mathbb{R}^N\} \\ &= (a, f(a)) + \text{Graf } (Df(a)).\end{aligned}$$

Dicha variedad afín de $\mathbb{R}^N \times \mathbb{R}^M$ es la "más próxima" a Graf (f) en el punto $(a, f(a))$. Se denomina *variedad afín tangente a la gráfica de f en el punto $(a, f(a))$*. Obsérvese que la aplicación $x \mapsto (x, Df(a)(x))$ de $\mathbb{R}^N$ en Graf $(Df(a))$ es un isomorfismo y en consecuencia la variedad afín tangente a la gráfica de f en el punto $(a, f(a))$ es afínmente isomorfa a $\mathbb{R}^N$.

Un subconjunto H de un espacio vectorial X es un *hiperplano vectorial* si es un subespacio maximal, es decir, si $H \neq X$ y el único subespacio vectorial que contiene estrictamente a H es X. El Teorema de extensión de la base nos permite afirmar que los hiperplanos vectoriales son los subespacios de codimensión uno. El trasladado A por un vector a de un hiperplano vectorial H se llama un *hiperplano* (afín), es decir, $A = a + H$. Para la caracterización de los hiperplanos vectoriales y de los hiperplanos (afines) véase el apéndice C.

Supuesto que un campo escalar f es derivable en a, la variedad afín tangente a la gráfica de f en el punto $(a, f(a))$ es el hiperplano de $\mathbb{R}^{N+1}$ dado por

$$\{(a, f(a)) + \big(x, Df(a)(x)\big) :\ x \in \mathbb{R}^N\}.$$

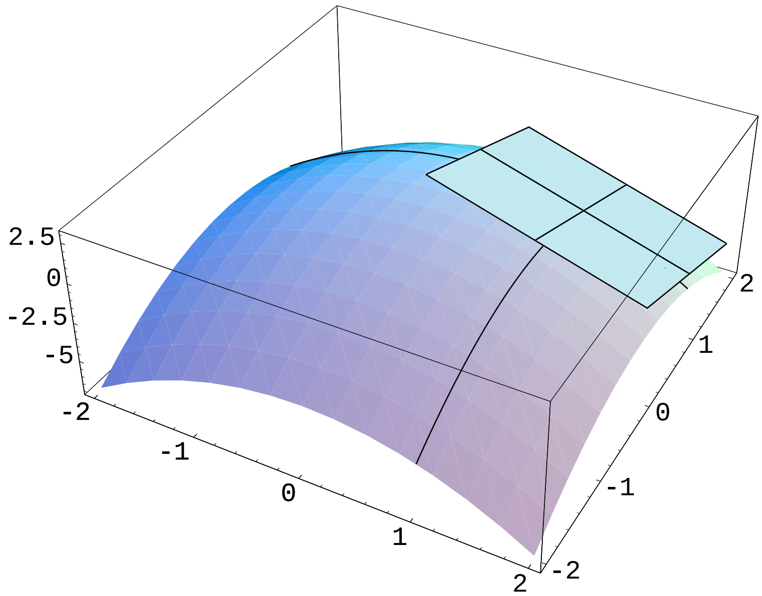
En virtud de la Proposición 3.16 este hiperplano se escribe ahora

$$\left\{(a, f(a)) + \big(x, (\nabla f(a)|x)\big) : x \in \mathbb{R}^N\right\} = \left\{\Big(x, f(a) + \big(\nabla f(a)|x-a\big)\Big) : x \in \mathbb{R}^N\right\}.$$

Por tanto, el *hiperplano tangente* a la gráfica del campo escalar f en el punto $(a, f(a)) = (a_1, \ldots, a_N, f(a_1, \ldots, a_N))$ es el hiperplano de ecuación

$$x_{N+1} = f(a) + D_1 f(a)(x_1 - a_1) + \ldots + D_N f(a)(x_N - a_N).$$

A continuación caracterizamos la variedad afín tangente a la gráfica de una función en un punto, en particular, los hiperplanos tangentes. Para ello recordamos, en primer lugar, que



una <u>*curva en* $\mathbb{R}^N$</u> es una aplicación continua γ de un intervalo I de $\mathbb{R}$ en $\mathbb{R}^N$. Normalmente la forma más cómoda de visualizar una curva es considerar su <u>*imagen*</u> (en $\mathbb{R}^N$, al menos cuando $N \leq 3$) en lugar de su <u>*gráfica*</u> en $\mathbb{R}^{N+1}$. Supongamos que γ es derivable en un punto $t_0 \in I$ con $\gamma'(t_0) \neq 0$, entonces la recta que pasa por el punto $\gamma(t_0)$ con dirección $\gamma'(t_0)$, esto es,

$$\{\gamma(t_0) + t\gamma'(t_0) : t \in \mathbb{R}\}$$

es la <u>*recta tangente a γ en el punto $\gamma(t_0)$.*</u>

Claramente $\operatorname{Im}(\gamma) = \pi_{\mathbb{R}^N}(\operatorname{Graf}(\gamma))$ (donde hemos identificado de forma natural $\mathbb{R}^{N+1} = \mathbb{R} \times \mathbb{R}^N$). Si la curva se visualiza en $\mathbb{R}^N$, es natural visualizar también en $\mathbb{R}^N$ la tangente en un punto como la proyección de la recta tangente a la gráfica de γ en el punto $(t_0, \gamma(t_0))$. Esta viene dada por

$$\{(t_0, \gamma(t_0)) + (t, t\gamma'(t_0)) : t \in \mathbb{R}\} = \{(t_0, \gamma(t_0)) + t(1, \gamma'(t_0)) : t \in \mathbb{R}\}$$

y su proyección es

$$\{\gamma(t_0) + t\gamma'(t_0) : t \in \mathbb{R}\},$$

que es la recta tangente a γ en $\gamma(t_0)$ cuando $\gamma'(t_0) \neq 0$.

Seguidamente mostramos que para un campo vectorial derivable f, la variedad afín tangente a la gráfica de f en el punto $(a, f(a))$ coincide con el conjunto de todas las rectas tangentes a las curvas cuya imagen está contenida en $\operatorname{Graf}(f)$ y que pasan por el punto $(a, f(a))$.

Proposición 3.30. *Sea $A \subset \mathbb{R}^N$ un abierto y $f : A \longrightarrow \mathbb{R}^M$ una función derivable. Si $a \in A$ y $u \in \mathbb{R}^N \times \mathbb{R}^M$, equivalen:*

i) $u \in \operatorname{Graf}(Df(a))$.

ii) Existen $\delta > 0$ y $\gamma : [-\delta, \delta] \longrightarrow \mathbb{R}^N \times \mathbb{R}^M$ continua tal que

$$\gamma([-\delta, \delta]) \subset \text{Graf } (f), \ \ \gamma(0) = (a, f(a)) \ \text{ y } \ \gamma'(0) = u.$$

Demostración:

i) $\Rightarrow$ ii) Notaremos por π_1 y π_2 a las proyecciones de $\mathbb{R}^N \times \mathbb{R}^M$ sobre $\mathbb{R}^N$ y $\mathbb{R}^M$, respectivamente. Si $u \in \text{Graf } (Df(a))$, sea $\delta > 0$ tal que

$$[a - \delta\pi_1(u), a + \delta\pi_1(u)] \subset A$$

y consideremos la curva $\gamma : [-\delta, \delta] \longrightarrow \mathbb{R}^N \times \mathbb{R}^M$ definida por

$$\gamma(t) = (a + t\pi_1(u), f(a + t\pi_1(u))).$$

Es claro que $\gamma([-\delta, \delta]) \subset \text{Graf } (f)$, $\gamma(0) = (a, f(a))$ y

$$\gamma'(0) = (\pi_1(u), Df(a)(\pi_1(u))),$$

donde se han utilizado las reglas elementales de derivación. Por otra parte, como $u \in \text{Graf } (Df(a))$, entonces $u = \big(\pi_1(u),\ Df(a)(\pi_1(u))\big)$ y en consecuencia $\gamma'(0) = u$.

ii) $\Rightarrow$ i) Supongamos ahora que γ es una curva verificando las condiciones requeridas en ii). Entonces, como $\text{Im } (\gamma) \subset \text{Graf } (f)$, se verifica que

$$f \circ \pi_1 \circ \gamma = \pi_2 \circ \gamma$$

y las reglas elementales de derivación nos aseguran que

$$Df(a)(\pi_1(\gamma'(0))) = \pi_2(\gamma'(0)),$$

esto es, $u = \gamma'(0) \in \text{Graf } (Df(a))$. ∎

3.7. Apéndice A) Desigualdad de Cauchy-Schwarz.

Proposición.

$$|(x|y)| \leq \|x\|_2 \ \|y\|_2, \ \forall x, y \in \mathbb{R}^N,$$

donde $(\cdot|\cdot)$ *denota el producto escalar en* $\mathbb{R}^N$*, esto es*

$$(x|y) := \sum_{k=1}^{N} x_k y_k,$$

y la igualdad ocurre si, y sólo si, los vectores son linealmente dependientes.

Demostración:

Sean $x, y \in \mathbb{R}^N$. Si $y = 0$ se da la igualdad y la condición (¿Por qué?). En caso contrario, la desigualdad de Cauchy-Schwarz equivale a probar

$$0 \leq \frac{(x|x)(y|y) - (x|y)^2}{(y|y)},$$

desigualdad que probamos a continuación. En efecto

$$\begin{aligned}
\frac{(x|x)(y|y) - (x|y)^2}{(y|y)} &= (x|x) - \frac{(x|y)}{(y|y)}(x|y) \\
&= (x|x) - 2\frac{(x|y)}{(y|y)}(x|y) + \frac{(x|y)}{(y|y)}(x|y) \\
&= (x|x) - 2\frac{(x|y)}{(y|y)}(x|y) + \left(\frac{(x|y)}{(y|y)}\right)^2 (y|y) \\
&= \left(x - \frac{(x|y)}{(y|y)}y \;\middle|\; x - \frac{(x|y)}{(y|y)}y\right) \geq 0
\end{aligned}$$

Finalmente, si por ejemplo $y = \alpha x$ para conveniente $\alpha \in \mathbb{R}$, entonces

$$(x|y)^2 = (x|\alpha x)^2 = \alpha^2 \|x\|_2^4 = \|x\|_2^2 \ \|\alpha x\|_2^2 = \|x\|_2^2 \ \|y\|_2^2.$$

■

3.8. Apéndice B) Normas duales.

Proposición. *Si consideramos en $\mathbb{R}^N$ la norma $\|.\|_1$ (resp. $\|.\|_2$, $\|.\|_\infty$), entonces la norma de operadores en $\mathscr{L}(\mathbb{R}^N,\mathbb{R}) \equiv \mathbb{R}^N$ es $\|.\|_\infty$ (resp. $\|.\|_2$, $\|.\|_1$).*

Demostración:

Sabemos que los espacios vectoriales $\mathscr{L}(\mathbb{R}^N,\mathbb{R})$ y $\mathbb{R}^N$ son matemáticamente indistinguibles vía la correspondencia que a cada $a \in \mathbb{R}^N$ le asigna la forma lineal

$$x \longmapsto (a|x).$$

1. El dual de $\left(\mathbb{R}^N, \|.\|_1\right)$ es isométricamente isomorfo a $\left(\mathbb{R}^N, \|.\|_\infty\right)$.

   Se tiene que
   $$|(a|x)| \leq \|a\|_\infty \, \|x\|_1 \;\; (a,x \in \mathbb{R}^N),$$
   y en consecuencia
   $$\|a\| \leq \|a\|_\infty.$$
   Por otra parte si $k \in \{1,...,N\}$ es tal que $\|a\|_\infty = |a_k|$, tomado $x_0 = e_k$ se tiene también que
   $$\|a\| \geq |(a|x_0)| = |a_k| = \|a\|_\infty.$$
   Hemos probado que $\|a\| = \|a\|_\infty$ y que la norma dual se alcanza en x_0.

2. El dual de $\left(\mathbb{R}^N, \|.\|_2\right)$ es isométricamente isomorfo a $\left(\mathbb{R}^N, \|.\|_2\right)$.

   La prueba es análoga al caso anterior. La primera desigualdad es en este caso la desigualdad de Cauchy-Schwarz y el punto donde se alcanza la norma dual es por ejemplo
   $$x_0 = \frac{a}{\|a\|_2}.$$

3. El dual de $\left(\mathbb{R}^N, \|.\|_\infty\right)$ es isométricamente isomorfo a $\left(\mathbb{R}^N, \|.\|_1\right)$.

   La prueba es también análoga al primer caso siendo ahora x_0 (vector donde se alcanza la norma dual) cualquier vector que verifique
   $$\left|x_0(k)\right| = 1, \quad a_k \, x_0(k) \geq 0 \; , \; k \in \{1,...,N\}.$$

$\blacksquare$

3.9. Apéndice C) Hiperplanos.

Proposición. *Sea H un subconjunto de $\mathbb{R}^N$. Equivalen las siguientes afirmaciones:*

i) H es un hiperplano vectorial.

ii) H es el núcleo de una forma lineal no nula.

Además, dos formas lineales no nulas determinan el mismo hiperplano vectorial, si y sólo si, son proporcionales.
Demostración:
$i) \Rightarrow ii)$ Sea $e \in \mathbb{R}^N \setminus H$. Se tiene que $\mathbb{R}^N = H \bigoplus < e >$, es decir:

$$\forall x \in \mathbb{R}^N, \exists! h \in H,\ \exists! \lambda \in \mathbb{R} :\ x = h + \lambda e.$$

La aplicación $f : \mathbb{R}^N \to \mathbb{R}$ definida por

$$f(x) = \lambda \ \text{ si } \ x = h + \lambda e \ \ (h \in H, \lambda \in \mathbb{R})$$

es un funcional lineal no nulo cuyo núcleo es H.
$ii) \Rightarrow i)$ Sea f un funcional no nulo tal que $H = Ker(f)$. Se tiene que H es un subespacio vectorial de $\mathbb{R}^N$. Probemos que H es maximal. Supongamos que S es un subespacio vectorial de $\mathbb{R}^N$ que contenga estrictamente a H. Fijado $y \in S \setminus H$, se tiene para cada vector $x \in \mathbb{R}^N$ que

$$x = \left(x - f(x)\frac{y}{f(y)}\right) + f(x)\frac{y}{f(y)} \ \in S$$

ya que $x - f(x)\frac{y}{f(y)} \ \in H$ y $f(x)\frac{y}{f(y)} \ \in S$. Hemos probado que $S = \mathbb{R}^N$.
Por último, sean f, g funcionales lineales no nulos. Si existe un real λ tal que $g = \lambda f$, es claro que $Ker(f) = Ker(g)$. Recíprocamente, si $Ker(f) = Ker(g)$, dado un vector $a \in \mathbb{R}^N \setminus Ker(f)$ se tiene que $g = \frac{g(a)}{f(a)} f$ ya que $x - \frac{f(x)}{f(a)} a \ \in Ker(f), \forall x \in \mathbb{R}^N$, y por tanto

$$0 = g(x) - \frac{f(x)}{f(a)} g(a),\ \forall x \in \mathbb{R}^N.$$

■

Corolario. *Un subconjunto A de $\mathbb{R}^N$ es un hiperplano, si y sólo si, existen un funcional no nulo f y un número real α tales que*

$$A = \{x \in \mathbb{R}^N :\ f(x) = \alpha\},$$

es decir, existen $A_1, ..., A_N, \alpha \in \mathbb{R}$ tales que

$$A = \{x \in \mathbb{R}^N :\ A_1 x_1 + ... + A_N x_N = \alpha\}.$$

Demostración:

Supongamos que $A = a + H$ para convenientes $a \in \mathbb{R}^N$ y H hiperplano vectorial. Si f es un funcional no nulo tal que $H = Ker(f)$, entonces se verifica que

$$A = \{x \in \mathbb{R}^N : \ f(x) = f(a)\}.$$

Recíprocamente, supongamos que $A = \{x \in \mathbb{R}^N : \ f(x) = \alpha\}$ para conveniente $\alpha \in \mathbb{R}$. Puesto que f es sobreyectiva (¡Hágase!), podemos elegir $a \in \mathbb{R}^N$ tal que $f(a) = \alpha$. Se tiene entonces que

$$A = f^{-1}(\alpha) = \{x \in \mathbb{R}^N : \ f(x) = \alpha\} = \{x \in \mathbb{R}^N : \ x - a \in Ker(f)\} = a + Ker(f).$$

■

Proposición. *Sean f un funcional lineal no nulo y x_0 un vector de $\mathbb{R}^N$, entonces*

$$\operatorname{dist}\Big(x_0, f^{-1}(\alpha)\Big) = \frac{|f(x_0) - \alpha|}{\|f\|}.$$

En consecuencia, si el funcional f viene dado por

$$f(x) = A_1x_1 + ... + A_Nx_N, \ \forall x \in \mathbb{R}^N, (A_1^2 + ... + A_N^2 > 0)$$

y se considera en $\mathbb{R}^N$ la norma euclídea, entonces la fórmula anterior es

$$\operatorname{dist}\big(x_0, A\big) = \frac{|A_1x_0(1) + ... + A_Nx_0(N) - \alpha|}{\sqrt{A_1^2 + ... + A_N^2}}.$$

Demostración:

Basta probar que $\operatorname{dist}\big(x_0, Ker(f)\big) = \frac{|f(x_0)|}{\|f\|}$. En efecto, sea a tal que $f(a) = \alpha$, entonces, en virtud de la linealidad de f, se tiene que

$$\operatorname{dist}\Big(x_0, f^{-1}(\alpha)\Big) = \operatorname{dist}\big(x_0 - a, Ker(f)\big).$$

Al ser f lipschitziana de razón $\|f\|$, se tiene para cada $x \in Ker(f)$ que

$$|f(x_0)| = |f(x_0 - x)| \leq \|f\| \|x_0 - x\| \qquad (*)$$

de donde se deduce que

$$\frac{|f(x_0|}{\|f\|} \leq \operatorname{dist}\big(x_0, Ker(f)\big).$$

Por definición de norma de un aplicación lineal, existe una sucesión $\{x_n\}$ en la esfera unidad de $\big(\mathbb{R}^N, \|\cdot\|\big)$ tal que $\|f\| < \frac{n+1}{n} \ |f(x_n)|, \forall n \in \mathbb{N}$. Al ser $\Big\{x_0 - f(x_0) \ \frac{x_n}{f(x_n)}\Big\}$ una sucesión en $Ker(f)$, se tiene para cada n natural que

$$\operatorname{dist}\big(x_0, Ker(f)\big) \leq \Big\|x_0 - \Big(x_0 - f(x_0) \ \frac{x_n}{f(x_n)}\Big)\Big\| = \Big|\frac{f(x_0)}{f(x_n)}\Big| < \ \frac{n+1}{n} \ \frac{|f(x_0)|}{\|f\|},$$

de donde deducimos tomando límites que $\operatorname{dist}\big(x_0, Ker(f)\big) \leq \frac{|f(x_0)|}{\|f\|}$. Hemos probado que

$$\operatorname{dist}\big(x_0, Ker(f)\big) = \frac{|f(x_0)|}{\|f\|}.$$

Nótese que también se puede razonar que la desigualdad (*) es, de hecho, una igualdad por alcanzarse la norma de operadores.

La última expresión se deduce de la anterior sin más que recordar que la norma euclídea es autodual. ∎

3.10. Referencias recomendadas.

[BRV], [Cra], [Fe], [FeSa], [MaHo], [Jur] y [MaHo].

3.11. Resumen del resultados del Tema 3

$\mathscr{L}(\mathbb{R}^N,\mathbb{R}^M) \equiv \mathscr{M}_{M\times N}(\mathbb{R})$.

El espacio vectorial $\mathscr{L}(\mathbb{R}^N,\mathbb{R}^M)$ y el espacio vectorial $\mathscr{M}_{M\times N}(\mathbb{R})$ de las matrices $M\times N$ de números reales son matemáticamente indistinguibles. Para cada T en $\mathscr{L}(\mathbb{R}^N,\mathbb{R}^M)$ definimos $A_T \in \mathscr{M}_{M\times N}(\mathbb{R})$ por

$$A_T := \Big(T_i(e_j)\Big)_{\substack{1\le i\le M\\ 1\le j\le N}},$$

donde $T_1,\cdots,T_M$ son los campos escalares componentes de T y $\{e_1,...,e_N\}$ es la base canónica de $\mathbb{R}^N$. La aplicación de $\mathscr{L}(\mathbb{R}^N,\mathbb{R}^M)$ en $\mathscr{M}_{M\times N}(\mathbb{R})$ definida por $T\to A_T$ es un isomorfismo de $\mathscr{L}(\mathbb{R}^N,\mathbb{R}^M)$ sobre $\mathscr{M}_{M\times N}(\mathbb{R})$, cuyo inverso es la aplicación de $\mathscr{M}_{M\times N}(\mathbb{R})$ sobre $\mathscr{L}(\mathbb{R}^N,\mathbb{R}^M)$ definida por $A\to T_A$ donde T_A es la aplicación lineal de $\mathbb{R}^N$ en $\mathbb{R}^M$ dada por

$$\Big(T_A(x)\Big)^t := Ax^t, \forall x\in\mathbb{R}^N.$$

Isomorfismo topológico. Si X e Y son espacios normados, una aplicación $T: X\longrightarrow Y$ es un <u>*isomorfismo topológico*</u> si T es una aplicación biyectiva, lineal y continua cuya inversa también es continua. Si $X=Y=\mathbb{R}^N$, toda aplicación lineal y biyectiva es un isomorfismo topológico. Notaremos $\mathrm{Iso}(\mathbb{R}^N)$ al conjunto de los isomorfismos de $\mathbb{R}^N$ en $\mathbb{R}^N$.

$$\mathrm{Iso}(\mathbb{R}^N) := \{T\in\mathscr{L}(\mathbb{R}^N) : \det A_T \neq 0\}.$$

Función lipschitziana. Sean (E,d) y (F,ρ) espacios métricos. Se dice que $f: E\to F$ es lipschitziana si existe $K\geq 0$ verificando

$$\rho(f(x),f(y)) \leq Kd(x,y), \quad \forall x,y\in E.$$

La menor constante que verifica esta desigualdad se denomina la constante de Lipschitz de T.

El espacio de Banach $\mathscr{L}(\mathbb{R}^N,\mathbb{R}^M)$. Norma del espacio $\mathscr{L}(\mathbb{R}^N,\mathbb{R}^M)$.

La función que a cada aplicación lineal $T:\mathbb{R}^N\to\mathbb{R}^M$ le hace corresponder el número real

$$\|T\| := \max\{\|T(x)\| : \|x\|=1\}$$

es una norma en $\mathscr{L}(\mathbb{R}^N,\mathbb{R}^M)$, denominada <u>norma de operadores</u>. $\mathscr{L}(\mathbb{R}^N,\mathbb{R}^M)$ es un espacio de Banach. Además se verifica que

$$\|T(x)\| \leq \|T\|\|x\|, \forall x\in\mathbb{R}^N$$

$$\|T\| = \min\{K\geq 0 : \ \|T(x)\|\leq K\|x\|, \ \forall x\in\mathbb{R}^N\} = \max\{\|T(x)\| : \|x\|\leq 1\}.$$

La primera igualdad nos asegura que $\|T\|$ es la constante de Lipschitz de T.

Proposición. Iso $(\mathbb{R}^N)$, *el conjunto de los isomorfismos de* $\mathbb{R}^N$ *en* $\mathbb{R}^N$*, es un abierto de* $\mathscr{L}(\mathbb{R}^N)$ *y la aplicación inversión* $J : \text{Iso }(\mathbb{R}^N) \to \mathscr{L}(\mathbb{R}^N)$ *definida por*

$$J(T) := T^{-1}, \quad \forall T \in \text{Iso }(\mathbb{R}^N)$$

es continua.

Función derivable Fréchet. Sean $M, N \in \mathbb{N}$, $A \subset \mathbb{R}^N$, $a \in \overset{\circ}{A}$, y $f : A \to \mathbb{R}^M$ una función. Se dice que f es derivable (en el sentido de Fréchet) o diferenciable en el punto a si existe una aplicación lineal T de $\mathbb{R}^N$ en $\mathbb{R}^M$ verificando

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

donde se consideran normas cualesquiera en $\mathbb{R}^N$ y $\mathbb{R}^M$. En tal caso la aplicación T es única, se denomina la derivada de la función f en el punto a o diferencial de f en a y se nota por $Df(a)$. Se dice que f es derivable en un subconjunto $B \subset A$ si es derivable en cada punto de B.

Sea $A_1 \subset A$ el conjunto de puntos donde f es derivable. La aplicación $x \to Df(x)$ de A_1 en $\mathscr{L}(\mathbb{R}^N, \mathbb{R}^M)$ se denomina la aplicación derivada de f o la diferencial de f y se nota Df.

Se dice que f es de clase C^1 en a, y se nota $f \in C^1(a)$, si f es derivable en un entorno de a y la aplicación Df es continua en a. Se dice que f es de clase C^1 en un subconjunto $B \subset A$ si es de clase C^1 en cada punto de B.

Se dice que f es de clase C^1 cuando lo sea en todos los puntos de su conjunto de definición (que necesariamente será abierto). Notamos por $C^1(A)$ al conjunto de las funciones de clase C^1 en el abierto A.

Reducción de la derivabilidad a campos escalares. *Sean* A *un subconjunto de* $\mathbb{R}^N$ *y* $f = (f_1, \ldots, f_M) : A \to \mathbb{R}^M$ *un campo vectorial. Entonces* f *es derivable en* $a \in \overset{\circ}{A}$ *si, y sólo si, cada campo escalar componente es derivable en* a*, en cuyo caso*

$$Df(a)(x) = (Df_1(a)(x), ..., Df_M(a)(x)), \ \forall x \in \mathbb{R}^n.$$

Además $f \in C^1(a)$ *si, y sólo si,* $f_i \in C^1(a)$ *para* $i = 1, \ldots, M$.

Derivadas parciales. Vector gradiente. Sea $A \subset \mathbb{R}^N$, $a \in A$ y f un campo escalar en A. Para cada $1 \leq i \leq N$, supongamos que a_i es un punto de acumulación del conjunto

$$A_i := \{x_i \in \mathbb{R} : \ (a_1, \ldots a_{i-1}, x_i, a_{i+1}, \ldots, a_N) \in A\}.$$

Si la función de A_i en $\mathbb{R}$ definida por $x_i \to f(a_1, \ldots a_{i-1}, x_i, a_{i+1}, \ldots, a_N)$ es derivable en a_i, se dice que f tiene derivada parcial respecto de la variable i-ésima en el punto a. En tal caso el valor del límite

$$\lim_{x_i \to a_i} \frac{f(a_1, \ldots a_{i-1}, x_i, a_{i+1}, \ldots, a_N) - f(a)}{x_i - a_i}$$

se denomina la derivada parcial de f respecto de la variable i-ésima en el punto a y se nota $D_i f(a)$ (o también $\frac{\partial f}{\partial x_i}(a)$).

Es consecuencia de la definición que el cálculo de la derivada parcial respecto de la variable i-ésima del campo escalar f en un punto genérico $x = (x_1, ..., x_N)$ se ha de llevar a cabo derivando la función real de variable real que resulta al considerar constantes las variables x_j $(j \neq i)$ y por tanto las reglas de derivación ordinarias se podrán utilizar.

Si $A^i \subset A$ es el conjunto de puntos donde f tiene derivada parcial respecto de la variable i-ésima, entonces el campo escalar en A^i definido por $x \mapsto D_i f(x)$ se denomina la aplicación derivada parcial de f respecto de la variable i-ésima y se nota $D_i f$ (o también $\frac{\partial f}{\partial x_i}$).

Se dice que el campo escalar f tiene gradiente en el punto a si admite derivadas parciales en a con respecto de todas las variables, en cuyo caso definimos el vector gradiente de f en a por:

$$\nabla f(a) := (D_1 f(a), ..., D_N f(a)) \in \mathbb{R}^N.$$

Si C es el conjunto de puntos de A donde f tiene gradiente, entonces el campo vectorial en C definido por $x \rightarrow \nabla f(x)$ se denomina aplicación gradiente de f y se nota ∇f.

Estudio de la derivabilidad de campos escalares.

Sean $A \subset \mathbb{R}^N$, $a \in \overset{\circ}{A}$ y f un campo escalar en A. Para el estudio de la derivabilidad de f en a se sugiere seguir los siguientes pasos:

1. Continuidad de f en a. Si f no es continua en a, entonces no es derivable en a. De ser f continua en a, entonces f *puede no ser derivable* en a (véase la función g del Ejercicio 3.4, que es continua, tiene gradiente y no es derivable en $(0,0)$).

2. Existencia de $\nabla f(a)$. Si no existe $\nabla f(a)$, entonces f no es derivable en a. En caso contrario, la candidata a derivada es la aplicación $x \rightarrow \Big(\nabla f(a) \mid x\Big)$.

3. Continuidad de las derivadas parciales. Si ∇f es continuo en a entonces f es derivable en a. No se puede deducir de esto que f es de clase $\mathscr{C}^1$ en a, salvo que se globalice.

4. Definición de derivabilidad. Según que la expresión

$$\frac{f(x) - f(a) - \Big(\nabla f(a) \mid x - a\Big)}{\|x - a\|}$$

tienda a cero o no, cuando $x \rightarrow a$, f es derivable o no.

Matriz jacobiana. Sean $A \subset \mathbb{R}^N$, $a \in A$ y $f : A \rightarrow \mathbb{R}^M$ un campo vectorial. Se dice que f tiene matriz jacobiana en el punto a si cada uno de los campos escalares componentes tiene gradiente en dicho punto, en cuyo caso se define la matriz jacobiana de f en a por

$$J_f(a) := \begin{pmatrix} D_1 f_1(a) & ... & D_j f_1(a) & ... & D_N f_1(a) \\ ... & ... & ... & ... & ... \\ D_1 f_i(a) & ... & D_j f_i(a) & ... & D_N f_i(a) \\ ... & ... & ... & ... & ... \\ D_1 f_M(a) & ... & D_j f_M(a) & ... & D_N f_M(a) \end{pmatrix} = \big(D_j f_i(a)\big)_{\substack{1 \leq i \leq M \\ 1 \leq j \leq N}}.$$

En el caso $M = N$ al determinante de la matriz jacobiana se le denomina determinante jacobiano.

Proposición. *Sean $A \subset \mathbb{R}^N$ y $f : A \to \mathbb{R}^M$ un campo vectorial. Si f es derivable en $a \in \overset{\circ}{A}$, entonces f tiene matriz jacobiana en a y*

$$(Df(a)(x))^t = J_f(a)x^t, \ \forall x \in \mathbb{R}^N.$$

Conviene recordar:

- Toda aplicación lineal T de $\mathbb{R}^N$ en $\mathbb{R}^M$ es de clase $\mathscr{C}^1$ con

$$DT(a) = T, \ \forall a \in \mathbb{R}^N.$$

- Toda aplicación bilineal T de $\mathbb{R}^M \times \mathbb{R}^N$ en $\mathbb{R}^P$ (es decir, lineal en ambas variables) es de clase $\mathscr{C}^1$ con derivada dada por

$$DT(a,b)(x,y) = T(x,b) + T(a,y), \ \forall (a,b),(x,y) \in \mathbb{R}^M \times \mathbb{R}^N.$$

- La aplicación inversión $J : \text{Iso}\ (\mathbb{R}^N) \longrightarrow \mathscr{L}(\mathbb{R}^N)$ dada por

$$J(T) = T^{-1}, \ \forall T \in \text{Iso}\ (\mathbb{R}^N)$$

  es de clase $\mathscr{C}^1$ con derivada definida por

$$DJ(T)(S) = -T^{-1}ST^{-1}, \ \forall S \in \mathscr{L}(\mathbb{R}^N).$$

- **Linealidad**.
  Sean $A \subset \mathbb{R}^N$ y $f,g : A \longrightarrow \mathbb{R}^M$ funciones derivables en a y $\lambda \in \mathbb{R}$. Entonces $f+g$ y λf son derivables en a con

$$D(f+g)(a) = Df(a) + Dg(a)$$

$$D(\lambda f)(a) = \lambda Df(a).$$

  Además, si $f,g \in \mathscr{C}^1(a)$, entonces $f+g, \lambda f \in \mathscr{C}^1(a)$.

- **Regla de la cadena.**
  Sean $A \subset \mathbb{R}^N, B \subset \mathbb{R}^M, f : A \longrightarrow \mathbb{R}^M$ tal que $f(A) \subset B$ y $g : B \longrightarrow \mathbb{R}^P$. Supongamos que f es derivable en a y que g es derivable en $f(a)$. Entonces la composición $h = g \circ f$ es derivable en a con

$$Dh(a) = Dg(f(a)) \circ Df(a)$$

  y en consecuencia

$$J_h(a) = J_g(f(a))J_f(a).$$

  Además, si $f \in \mathscr{C}^1(a), g \in \mathscr{C}^1(f(a))$, entonces $h \in \mathscr{C}^1(a)$.

- La derivabilidad es un concepto local.
- **Derivada de la función inversa.**

  Sea $A \subset \mathbb{R}^N, a \in \overset{\circ}{A}$ y $f : A \longrightarrow \mathbb{R}^N$ un campo vectorial. Supongamos que f es inyectiva, derivable en a y que $f(a) \in \overset{\circ}{f(A)}$, entonces

  $$f^{-1} \text{ es derivable en } f(a) \Leftrightarrow \begin{cases} \det J_f(a) \neq 0 \\ f^{-1} \text{ es continua en } f(a) \end{cases}$$

  Además, si se verifica lo anterior, se tiene

  $$Df^{-1}(f(a)) = Df(a)^{-1},$$

  y en consecuencia,

  $$J_{f^{-1}}(f(a)) = J_f(a)^{-1}.$$

  Si además A y B son subconjuntos abiertos de $\mathbb{R}^N$, f es un homeomorfismo de A sobre B y $f \in \mathscr{C}^1(a)$ para algún $a \in A$ con $\det J_f(a) \neq 0$, entonces $f^{-1} \in \mathscr{C}^1(f(a))$.

Interpretación geométrica. *Si un campo escalar f es derivable en un punto a, entonces la gráfica de f tiene un hiperplano tangente en el punto $(a, f(a))$ que viene dado por la expresión*

$$x_{N+1} = f(a) + D_1 f(a)(x_1 - a_1) + \ldots + D_N f(a)(x_N - a_N).$$

Dicho hiperplano coincide con el conjunto de todas las rectas tangentes en $(a, f(a))$ a las curvas cuya imagen está contenida en la gráfica de f y que pasan por el punto $(a, f(a))$.

3.12. Ejercicios del Tema 3

3.1 Sea $T \in \mathscr{L}(\mathbb{R}^2)$ y supongamos que

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

es la matriz asociada a T en términos de la base canónica. Pruébese que si consideramos las normas $\| \|_1$ y $\| \|_\infty$ en $\mathbb{R}^2$ se obtiene que la norma del operador T viene dada por

$$\max\Big\{|a|+|c|,|b|+|d|\Big\} \ \ (\text{para } \| \|_1)$$

$$\max\Big\{|a|+|b|,|c|+|d|\Big\} \ \ (\text{para } \| \|_\infty).$$

3.2 Probar que si en la definición de derivada en un punto el espacio normado $\mathbb{R}^N$ es de dimensión mayor que 1 y no exigimos que el punto a sea interior, en general, no se puede asegurar que la aplicación lineal que aparece en la definición sea única.

Indicación: Estudiar la derivabilidad en el punto $(0,0)$ de la aplicación

$$f : \mathbb{R} \times \{0\} \ \to \ \mathbb{R}, \ \ f(x,0) = x, \ \forall x \in \mathbb{R}.$$

3.3 Estudiar la continuidad, existencia de derivadas direccionales y derivabilidad en el origen de las siguientes funciones:

$$f(x,y) = \frac{xy}{x^2+y^4}, \ \forall (x,y) \in \mathbb{R}^2 \backslash \{(0,0)\}, \ f(0,0) = 0,$$

$$g(x,y) = \frac{x^2y}{x^2+y^4}, \ \forall (x,y) \in \mathbb{R}^2 \backslash \{(0,0)\}, \ g(0,0) = 0,$$

$$h(x,y) = \frac{x^2y^2}{x^2+y^4}, \ \forall (x,y) \in \mathbb{R}^2 \backslash \{(0,0)\}, \ h(0,0) = 0.$$

y

$$k(x,y) = \frac{x^6(x^2+y^2)}{(y-x^2)^2+x^6} \ , \ \forall (x,y) \in \mathbb{R}^2 \backslash \{(0,0)\}, \ k(0,0) = 0.$$

3.4 Justificar que una función racional de N variables es de clase $\mathscr{C}^1$ en su conjunto de definición.

3.5 Sea $f : \mathbb{R}^2 \ \to \ \mathbb{R}$ la función definida por:

$$f(x,y) = \frac{x^3+x^2y}{|y|+x^2} \ \text{ si } (x,y) \neq (0,0); \quad f(0,0) = 0.$$

Estudiar la continuidad y derivabilidad de f.

3.6 Para $\alpha \in \mathbb{R}^+$ sea $f_\alpha : \mathbb{R}^2 \to \mathbb{R}$ la función definida por:

$$f_\alpha(x,y) = \frac{|xy|^\alpha}{x^2 - xy + y^2} \text{ si } (x,y) \neq (0,0); \;\; f_\alpha(0,0) = 0.$$

Estudiar la continuidad y derivabilidad de f_α.

3.7 Estudiar la continuidad y derivabilidad en $(0,0)$ de la función $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ definida por:

$$f(x,y) = \frac{1 - \cos x \, \cos y}{\sqrt{x^2 + y^2}} \text{ si } (x,y) \neq (0,0); \;\; f(0,0) = 0.$$

3.8 Sean f un campo escalar definido en $A \subset \mathbb{R}^N$ y $a \in \overset{\circ}{A}$. Probar que equivalen:

i) f es derivable en a.

ii) Existen $h : A \to \mathbb{R}^N$ continua en a con $h(a) = 0$ y $b \in \mathbb{R}^N$:

$$f(x) - f(a) - (b|x-a) = (h(x)|x-a), \; \forall x \in A$$

donde $(.|.)$ denota el producto escalar usual de $\mathbb{R}^N$.

3.9 Sean A un abierto de $\mathbb{R}^N$ y f un campo escalar definido en A con derivadas parciales acotadas en A. Probar que f es continuo. Dar un ejemplo que muestre que un tal campo escalar puede no ser derivable.

Indicación: Considerar el campo escalar definido en $\mathbb{R}^2$ por

$$f(x,y) = \frac{xy}{\sqrt{x^2 + y^2}} \text{ si } (x,y) \neq (0,0); \;\; f(0,0) = 0.$$

3.10 *** (Condición suficiente de derivabilidad).**

Sean $A \subset \mathbb{R}^2$, $(a,b) \in \overset{\circ}{A}$ y $f : A \to \mathbb{R}$ un campo escalar. Supongamos que $D_1 f$ existe en un entorno de (a,b) y es continua en (a,b) y que existe $D_2 f(a,b)$. Probar que f es derivable en (a,b).

Estudiar un resultado análogo para campos escalares de N variables.

Indicación: Sea $\varepsilon > 0$, tómese $\delta > 0$ tal que

$$\|(x,y) - (a-b)\|_1 < \delta \Rightarrow \begin{cases} (x,y) \in A \\ \exists D_1 f(x,y) \\ |D_1 f(x,y) - D_1 f(a,b)| < \varepsilon \\ |f(a,y) - f(a,b) - D_2 f(a,b)(y-b)| \leq \varepsilon |y-b| \end{cases}$$

Para $(x,y) \in \mathbb{R}^2$ con $0 < \|(x,y)-(a,b)\|_1 < \delta$, se puede escribir

$$f(x,y) - f(a,b) - D_1 f(a,b)(x-a) - D_2 f(a,b)(y-b) =$$

$$[\, f(x,y) - f(a,y) - D_1 f(a,b)(x-a)\,] + [\, f(a,y) - f(a,b) - D_2 f(a,b)(y-b)\,]$$

y basta aplicar el teorema del valor medio al primer sumando (cuando $x \neq a$) para probar la derivabilidad de f en (a,b).

3.11 Sean $A \subset \mathbb{R}$ y $f, g : A \to \mathbb{R}^3$ funciones derivables en $a \in \overset{\circ}{A}$. Consideremos las funciones $h : A \to \mathbb{R}$ y $\varphi : A \to \mathbb{R}^3$ definidas por

$$h(x) := (f(x)|g(x)), \;\; \varphi(x) := f(x) \wedge g(x), \; \forall x \in A$$

(h es el producto escalar de f y g, φ es el producto vectorial de f por g). Calcular $h'(x)$ y $\varphi'(x)$.

3.12 Sea f un campo escalar derivable en $A = \mathbb{R}^+ \times \mathbb{R}^+$. Considérese el campo escalar h definido en $\Omega = \mathbb{R}^+ \times]0, \frac{\pi}{2}[$ por

$$h(\rho, \vartheta) := f(\rho \cos\, \vartheta, \rho \operatorname{sen}\, \vartheta), \; \forall (\rho, \vartheta) \in \Omega.$$

Probar que equivalen

i) $x \dfrac{\partial f}{\partial y}(x,y) - y \dfrac{\partial f}{\partial x}(x,y) = f(x,y), \; \forall (x,y) \in A.$

ii) $\dfrac{\partial h}{\partial \vartheta}(\rho, \vartheta) = h(\rho, \vartheta), \; \forall (\rho, \vartheta) \in \Omega.$

Deducir que las soluciones de i) son de la forma:

$$f(x,y) = \; \varphi\left(\sqrt{x^2+y^2}\right) \exp\left(\arctan \frac{y}{x}\right),$$

donde $\varphi : \mathbb{R}^+ \to \mathbb{R}$ es una función derivable.

3.13 Sea $f : \mathbb{R}^N \;\to\; \mathbb{R}^M$ derivable en cero y homogénea de grado 1, es decir:

$$f(tx) = t f(x), \; \forall x \in \mathbb{R}^N \setminus \{0\}, \; \forall t \in \mathbb{R}^+.$$

Probar que f es lineal.

3.14 **Teorema de Euler**.
Sea p un número real. Una función $f : \mathbb{R}^N \backslash \{0\} \;\to\; \mathbb{R}$ se llama *homogénea de grado p* si verifica:

$$f(tx) = t^p f(x), \; \forall x \in \mathbb{R}^N \setminus \{0\}, \; \forall t \in \mathbb{R}^+.$$

Demostrar que para una función derivable $f : \mathbb{R}^N \backslash \{0\} \;\to\; \mathbb{R}$ equivalen:

i) f es homogénea de grado p.

ii) $[Df(x)](x) = pf(x),\ \forall x \in \mathbb{R}^N \backslash \{0\}$.

3.15 * La identificación canónica de $(\mathbb{C}, |.|)$ con $(\mathbb{R}^2, \|.\|_2)$ nos permite tener una visión real de una función $f : \mathbb{C} \to \mathbb{C}$. A saber, la función $f : \mathbb{C} \to \mathbb{C}$ dada por

$$x + iy = z \ \mapsto f(z) = u + iv,$$

puede verse como $f : \mathbb{R}^2 \to \mathbb{R}^2$ dada por

$$(x, y) \mapsto f(x, y) = (u(x, y), v(x, y)).$$

Probar que para $z_0 = a + ib$ equivalen las siguientes afirmaciones:

i) f es derivable en (a, b) con $J_f(a, b) = \begin{pmatrix} A & -B \\ B & A \end{pmatrix}$.

ii) $\lim_{z \to z_0} \dfrac{f(z) - f(z_0)}{z - z_0} = A + iB$.

3.16 * Sean A un abierto convexo de $\mathbb{R}^N$ y $f : A \to \mathbb{R}^N$ una función derivable en A verificando que:

$$\sum_{i,j=1}^{N} D_i f_j(x) h_i h_j > 0,\ \forall x \in A,\ \forall h \in \mathbb{R}^N \backslash \{0\}.$$

Probar que f es inyectiva.

3.17 *

a) Probar que $\|\cdot\|_1$ es derivable sólo en los puntos del conjunto

$$\mathbb{R}^2 \backslash \{(x, y) \in \mathbb{R}^2 : xy = 0\}$$

con

$$\nabla \|.\|_1(x, y) = \left(\frac{x}{|x|}, \frac{y}{|y|} \right).$$

Estudiar la derivabilidad de $\|.\|_1$ en $\mathbb{R}^N$.

b) Probar que $\|.\|_2$ es derivable en los puntos del conjunto $\mathbb{R}^2 \backslash \{(0, 0)\}$ con

$$\nabla \|.\|_2(x, y) = \left(\frac{x}{\sqrt{x^2 + y^2}}, \frac{y}{\sqrt{x^2 + y^2}} \right).$$

Estudiar la derivabilidad de $\|.\|_2$ en $\mathbb{R}^N$.

c) Probar que $\|.\|_\infty$ es derivable sólo en los puntos del conjunto

$$\mathbb{R}^2\backslash\{(x,y)\in\mathbb{R}^2 : |x| = |y|\}$$

con

$$\nabla\|.\|_\infty(x,y) = \begin{cases} \left(0, \frac{y}{|y|}\right) & \text{si } |y| > |x| \\ \left(\frac{x}{|x|}, 0\right) & \text{si } |y| < |x| \end{cases}.$$

Estudiar la derivabilidad de $\|.\|_\infty$ en $\mathbb{R}^N$.

3.18 * Sean $A \subset \mathbb{R}^N$, $a \in \overset{\circ}{A}$ y $f : A \to \mathbb{R}^M$ una función. Probar que equivalen las siguientes afirmaciones:

1. f es derivable en a.
2. Existe $f'(a;\cdot)$, dicha aplicación de $\mathbb{R}^N$ en $\mathbb{R}^M$ es lineal y

$$\lim_{t\to 0}\frac{f(a+tx)-f(a)}{t} = f'(a;x)$$

es uniforme en x variando en cualquier subconjunto acotado de $\mathbb{R}^N$.

3.19 * Probar que si

$$f(x,y) = \frac{x^6\sqrt{x^2+y^2}}{(y-x^2)^2+x^6}\ ,\ \forall (x,y)\in\mathbb{R}^2\backslash\{(0,0)\},\ f(0,0) = 0$$

y $a = (0,0)$, entonces $\lim_{t\to 0}\dfrac{f(a+tu)-f(a)}{t} = 0$ no es uniforme cuando u varía en la esfera unidad euclídea de $\mathbb{R}^2$.

3.13. Soluciones a los ejercicios del Tema 3.

3.1 Sea $(x,y) \in \mathbb{R}^2$, entonces

$$T(x,y) = (ax+by, cx+dy).$$

Por tanto

$$\|T(x,y)\|_1 = |ax+by| + |cx+dy| \leq \Big(|a|+|c|\Big)|x| + \Big(|b|+|d|\Big)|y| \leq$$

$$\leq \Big(|x|+|y|\Big)\max\Big\{|a|+|c|, |b|+|d|\Big\}.$$

Asimismo, se tiene, para $(x,y) \in B_\infty(0,1)$

$$\|T(x,y)\|_\infty = \max\Big\{|ax+by|, |cx+dy|\Big\} \leq \max\Big\{|a|+|b|, |c|+|d|\Big\}.$$

Es muy fácil comprobar que las dos desigualdades que hemos obtenido para la norma del operador son de hecho igualdades. En el primer caso, basta evaluar el operador en los vectores de la base canónica, en el segundo en los vectores $(1,1), (1,-1)$.

3.2 Sea $f : \mathbb{R} \times \{0\} \to \mathbb{R}$ definida por $f(x,0) = x$, $\forall x \in \mathbb{R}$. Compruébese que f admite por "derivada" en $(0,0)$ cualquier T de la forma $(1,b)$ con $b \in \mathbb{R}$.

3.3
- f no es continua en $a = (0,0)$, pues $f(y^2,y) = \dfrac{1}{2y}, \forall y \in \mathbb{R}^*$. f no tiene derivada en a según el vector $u = (1,1)$, pues

  $$\frac{f(a+tu)-f(a)}{t} = \frac{f(t,t)}{t} = \frac{1}{t+t^3}, \ \forall t \in \mathbb{R}^*.$$

  En consecuencia f no es derivable en a.

- g es continua en $a = (0,0)$, pues $|g(xy)| \leq |y|$, $\forall (x,y) \in \mathbb{R}^2$. g admite en a todas las derivadas direccionales. En efecto, fijado $u \in \mathbb{R}^2 \backslash \{(0,0)\}$, tenemos que si $u = (x,y)$, entonces

  $$\frac{g(a+tu)-g(a)}{t} = \frac{g(tx,ty)}{t} = \frac{x^2y}{x^2+t^2y^4}\ ,$$

  y tomando límite cuando $t \to 0$ (distinguiendo los casos $x = 0$ y $x \neq 0$) se obtiene que

  (*) $$g'(a;(x,y)) = \begin{cases} y & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

  g no es derivable en a pues la aplicación de $\mathbb{R}^2$ en $\mathbb{R}$ $(x,y) \to g'(a;(x,y))$ no es lineal (si lo fuese sería de la forma $Ax+By$, lo que no permite (*)).

- h es continua en a, pues $|h(x,y)| \leq y^2$, $\forall (x,y) \in \mathbb{R}^2$. h admite en a todas las derivadas direccionales. En efecto, fijado $u \in \mathbb{R}^2 \backslash \{(0,0)\}$, tenemos que si $u = (x,y)$, entonces

$$\frac{h(a+tu)-h(a)}{t} = \frac{h(tx,ty)}{t} = \frac{tx^2y^2}{x^2+t^2y^4} \ ,$$

y tomando límite cuando $t \to 0$ (distinguiendo los casos $x = 0$ y $x \neq 0$) se obtiene que $h'(a;\cdot) = 0$, luego hay candidato a derivada. h es derivable en a. En efecto para $u \neq 0$ se verifica

$$\frac{h(u)-h(a)-h'(a;u-a)}{\|u-a\|} = \frac{h(u)}{\|u\|} = \frac{x^2y^2}{(x^2+y^4)\sqrt{x^2+y^2}} =: \varphi(x,y)$$

con

$$|\varphi(x,y)| \leq \frac{y^2}{\sqrt{x^2+y^2}} \leq |y|.$$

- k es continua en a, pues $|k(x,y)| \leq x^2+y^2$, $\forall (x,y) \in \mathbb{R}^2$. k admite en a todas las derivadas direccionales. En efecto, fijado $u \in \mathbb{R}^2 \backslash \{(0,0)\}$, tenemos que si $u = (x,y)$, entonces

$$\frac{k(a+tu)-k(a)}{t} = \frac{k(tx,ty)}{t} = \frac{t^5x^6(x^2+y^2)}{(y-tx^2)^2+t^4x^6},$$

y tomando límite cuando $t \to 0$ (distinguiendo los casos $y = 0$ e $y \neq 0$) se obtiene que $k'(a;\cdot) = 0$. Y tiene candidata a derivada pues $k'(a;\cdot)$ es lineal. k es derivable en a. En efecto para $u \neq 0$ se verifica

$$\frac{k(u)-k(a)-k'(a;u-a)}{\|u\|} = \frac{k(u)}{\|u\|} = \frac{x^6\sqrt{x^2+y^2}}{(y-x^2)^2+x^6} =: \varphi(x,y)$$

y $|\varphi(x,y)| \leq \sqrt{x^2+y^2}$.

3.4 Sabemos que una función racional de N variables está, de forma natural, definida en un conjunto abierto de $\mathbb{R}^N$ y es continua. Como las derivadas parciales de una función racional son también funciones racionales definidas en el mismo conjunto, concluimos que las derivadas parciales de una función racional son continuas y por tanto en virtud del Teorema 3.21 la función es de clase $\mathscr{C}^1$ en su conjunto de definición.

3.5 **Estudio en** $(0,0)$.

Continuidad: Como

$$|f(x,y)| \leq \left|\frac{x^3+x^2y}{x^2}\right| \leq |x|+|y| = \|(x,y)\|_1, \ \forall (x,y) \in \mathbb{R}^2$$

deducimos que f es continua en $(0,0)$.

<u>Derivabilidad</u>: Es inmediato que $D_1 f(0,0) = 1$ y $D_2 f(0,0) = 0$, con lo que si

$$g(x,y) := \frac{f(x,y) - x}{\|(x,y)\|_2} = \frac{x(xy - |y|)}{(|y| + x^2)\sqrt{x^2 + y^2}}$$

se tiene que

$$g(x,x) = \frac{x - \frac{|x|}{x}}{\sqrt{2}(1 + |x|)}$$

que no tiene límite cuando $x \to 0$. Hemos probado que la función f no es derivable en $(0,0)$.

Estudio en el eje de abscisas salvo el punto $(0,0)$.

Consideremos el punto $(a,0)$ con $a \neq 0$.

<u>Continuidad</u>: Como cociente de continuas con denominador distinto de cero f es continua en $(a,0)$.

<u>Derivabilidad</u>:

$$\frac{f(a,y) - f(a,0)}{y} = a \frac{a - \frac{|y|}{y}}{a^2 + |y|}$$

que no tiene límite cuando $y \to 0$. Así no existe $D_2 f(a,0)$ y en consecuencia f no es derivable en $(a,0)$.

Estudio fuera del eje de abscisas.

La función es de clase $\mathscr{C}^1$ pues tiene derivadas parciales continuas.

3.6 De

$$\frac{-1}{2}(x^2 + y^2) \leq -xy \leq \frac{1}{2}(x^2 + y^2) \Longrightarrow \frac{1}{2}(x^2 + y^2) \leq x^2 - xy + y^2 \leq \frac{3}{2}(x^2 + y^2)$$

se tiene que

$$0 \leq \frac{2|xy|^\alpha}{3(x^2 + y^2)} \leq f(x,y) \leq \frac{2|xy|^\alpha}{x^2 + y^2} \leq \left(\frac{x^2 + y^2}{2}\right)^{\alpha - 1}$$

Estudio en $(0,0)$.

<u>Continuidad</u>: f es continua $\Longleftrightarrow 1 < \alpha$

$1 < \alpha \Longrightarrow |f(x,y)| \leq \left(\frac{x^2+y^2}{2}\right)^{\alpha - 1} \to 0$ luego continua.

$0 < \alpha \leq 1 \Rightarrow 1 \leq f(x,x), \forall x \in]0,1[$ y por tanto no continua.

<u>Derivabilidad</u>: f es derivable $\Longleftrightarrow \dfrac{3}{2} < \alpha$

$\frac{3}{2} < \alpha \Longrightarrow \frac{|f(x,y)|}{\sqrt{x^2+y^2}} \leq \frac{(x^2+y^2)^{\alpha-\frac{3}{2}}}{2^{\alpha-1}} \to 0$, luego derivable.

$1 < \alpha \leq \frac{3}{2} \Longrightarrow \frac{1}{\sqrt{2}} \leq \frac{f(x,x)}{\sqrt{2}|x|}, \forall x \in]0,1[$ y por tanto no derivable.

También puede estudiarse la continuidad y derivabilidad mediante el paso a coordenadas polares. ¡Hágase!.

Estudio en los ejes coordenados salvo el punto $(0,0)$**.**

Continuidad: f es continua para cualquier $\alpha > 0$.

El numerador es continuo, el denominador también y no se anula.

Derivabilidad: f es derivable $\Leftrightarrow 1 < \alpha$.

En efecto, por razones de simetría, basta estudiar la función en un punto de la forma $(a,0)$ con $a \neq 0$. Es claro que $D_1 f(a,0) = 0$. Por otra parte

$$\frac{f(a,y)-f(a,0)}{y} = \frac{|a|^\alpha}{a^2-ay+y^2}\frac{|y|^\alpha}{y} \tag{$*$}$$

que tiende a cero cuando $y \to 0$ pues $1 < \alpha$. Así $D_2 f(a,0) = 0$. Se tiene que

$$0 \leq \left|\frac{f(a+x,y)}{\sqrt{x^2+y^2}}\right| \leq \frac{|a+x|^\alpha}{|(a+x)^2-(a+x)y+y^2|}\ |y|^{\alpha-1} \to 0.$$

Para $0 < \alpha \leq 1$ la expresión $(*)$ no tiene límite, es decir no existe $D_2 f(a,0)$ y por tanto la función f no es derivable en $(a,0)$.

Estudio fuera de los ejes coordenados.

La función es de clase $\mathscr{C}^1$ pues tiene derivadas parciales continuas.

3.7 Continuidad: f es continua en $(0,0)$.

La función $g(x,y) = 1-\cos x \cos y, \forall x,y \in \mathbb{R}$ es claramente de clase $\mathscr{C}^1$ en $\mathbb{R}^2$. Como $D_1 g(0,0) = D_2 g(0,0) = 0$, se tiene

$$f \text{ es continua en } (0,0) \Leftrightarrow g \text{ es derivable en } (0,0),$$

y por tanto f es continua en $(0,0)$.

Derivabilidad: f no es derivable en $(0,0)$.

$$\frac{f(x,0)}{x} = \frac{1-\cos x}{x|x|} = \frac{1-\cos x}{x^2}\frac{x}{|x|}$$

que no tiene límite cuando $x \to 0$ ya que el primer factor tiende a $\frac{1}{2}$ y el segundo no tiene límite. Así no existe $D_1 f(0,0)$ y por tanto f no es derivable en $(0,0)$.

3.8 Es inmediato que para funciones reales de variable real equivalen las siguientes afirmaciones:

i) f es derivable en a.

ii) Existen $h : A \to \mathbb{R}$ continua en a con $h(a) = 0$ y $b \in \mathbb{R}$:

$$f(x) - f(a) - b(x-a) = h(x)(x-a), \forall x \in A.$$

Se trata de probar que esta caracterización sigue siendo cierta para campos escalares *sustituyendo el producto de números reales por el producto escalar de vectores*:

i) f es derivable en a.

ii) Existen $h : A \to \mathbb{R}^n$ continua en a con $h(a) = 0$ y $b \in \mathbb{R}^n$:

$$f(x) - f(a) - (b|x-a) = (h(x)|x-a), \forall x \in A.$$

$i) \Rightarrow ii)$ Por hipótesis existe $r : A \to \mathbb{R}$ continua en a con $r(a) = 0$ tal que:

$$f(x) - f(a) - (\nabla f(a)|x-a) = r(x)\|x-a\|_2, \ \forall x \in A.$$

Es fácil probar que $b = \nabla f(a)$ y la función $h : A \to \mathbb{R}^n$ definida por

$$h(x) := \begin{cases} \dfrac{r(x)(x-a)}{\|x-a\|_2} & \text{si } x \neq a \\ 0 & \text{si } x = a \end{cases}$$

cumple la afirmación $ii)$, pues para todo $x \in A$

$$(h(x)|x-a) = r(x)\|x-a\|_2, \ \|h(x)\|_2 = |r(x)|.$$

$ii) \Rightarrow i)$ La desigualdad de Cauchy-Schwarz nos asegura

$$|f(x) - f(a) - (b|x-a)| = |(h(x)|x-a)| \leq \|h(x)\|_2 \|x-a\|_2$$

expresión que nos dice que la función f es derivable en a y que $\nabla f(a) = b$.

3.9 Sea $a \in A$. Consideremos $r > 0$ tal que $\|x-a\|_1 < r \Rightarrow \ x \in A$. Para $x \in \mathbb{R}^N$ con $\|x-a\|_1 < r$ se tiene:

$$f(x) - f(a) = f(x_1, x_2, \ldots, x_N) - f(x_1, x_2, \ldots, x_{N-1}, a_N) +$$

$$f(x_1, x_2, \ldots, x_{N-1}, a_N) - f(x_1, x_2, \ldots, x_{N-2}, a_{N-1}, a_N) +$$

$$\ldots\ldots\ldots\ldots \quad +$$

$$f(x_1, a_2, \ldots, a_N) - f(a_1, \ldots, a_N).$$

Si para algún $i \in \{1,\dots,N\}$ es $x_i = a_i$, entonces el correspondiente sumando es nulo. En otro caso aplicando el teorema del valor medio podemos encontrar y_i entre x_i y a_i tal que

$$f(a_1,\dots,a_{i-1},x_i,\dots,x_N) - f(a_1,\dots,a_i,x_{i+1},\dots,x_N) =$$
$$D_i f(a_1,\dots a_{i-1},y_i,x_{i+1},\dots,x_N)(x_i - a_i).$$

En consecuencia

$$|f(x) - f(a)| \le M[|x_1 - a_1| + \dots + |x_N - a_N|] = M\|x-a\|_1,$$

donde M es una cota de las derivadas parciales de f.

Nota: Una tal función f es más que continua en a, sin embargo, no tiene por qué ser derivable en a. Por ejemplo, el campo escalar definido en $\mathbb{R}^2$ por

$$f(x,y) = \frac{xy}{\sqrt{x^2+y^2}} \ ; \ f(0,0) = 0$$

tiene derivadas parciales acotadas y no es derivable en $\mathbb{R}^2$.

$$\frac{\partial f}{\partial x}(x,y) = \frac{y}{\sqrt{x^2+y^2}} - \frac{x^2 y}{(x^2+y^2)\sqrt{x^2+y^2}} = \frac{y^3}{(x^2+y^2)\sqrt{x^2+y^2}}$$

y en consecuencia

$$\left|\frac{\partial f}{\partial x}(x,y)\right| \le \frac{y^2}{(x^2+y^2)} \frac{|y|}{\sqrt{x^2+y^2}} \le 1.$$

La simetría de f nos da también que $\left|\frac{\partial f}{\partial y}(x,y)\right| \le 1$. Por último como no existe límite de $\frac{xy}{x^2+y^2}$ cuando $(x,y) \to (0,0)$, la función f no es derivable en $(0,0)$.

3.10 **(Condición suficiente de derivabilidad).**

Sea f un campo escalar definido en $A \subset \mathbb{R}^2$ tal que $D_1 f$ es continua en $(a,b) \in \overset{\circ}{A}$ y existe $D_2 f(a,b)$. Sea $\varepsilon > 0$ y tomemos $\delta > 0$ tal que

$$\|(x,y)-(a,b)\|_1 < \delta \Rightarrow \begin{cases} (x,y) \in A \\ \exists D_1 f(x,y) \\ |D_1 f(x,y) - D_1 f(a,b)| < \varepsilon \\ |f(a,y) - f(a,b) - D_2 f(a,b)(y-b)| \le \varepsilon |y-b| \end{cases}.$$

Para $(x,y) \in \mathbb{R}^2$ con $0 < \|(x,y)-(a,b)\|_1 < \delta$, se tiene que

$$f(x,y) - f(a,b) - D_1 f(a,b)(x-a) - D_2 f(a,b)(y-b) =$$
$$\big[f(x,y) - f(a,y) - D_1 f(a,b)(x-a)\big] + \big[f(a,y) - f(a,b) - D_2 f(a,b)(y-b)\big].$$

Si $x = a$ el primer sumando es nulo. En otro caso consideremos la función del intervalo I de extremos a y x en $\mathbb{R}$ definida por

$$\sigma : t \to f(t,y) - D_1 f(a,b) t.$$

El teorema del valor medio nos asegura que existe $c \in \overset{\circ}{I}$ tal que

$$\sigma(x) - \sigma(a) = \sigma'(c)(b-a) = [D_1 f(c,y) - D_1 f(a,b)](x-a),$$

y por tanto

$$|f(x,y) - f(a,y) - D_1 f(a,b)(x-a)| = |D_1 f(c,y) - D_1 f(a,b)||x-a| < \varepsilon |x-a|,$$

donde se ha utilizado una de las condiciones impuestas a δ.
Utilizando la otra condición impuesta a δ se tiene que

$$|f(x,y) - f(a,b) - D_1 f(a,b)(x-a) - D_2 f(a,b)(y-b)| \leq$$

$$\varepsilon |x-a| + \varepsilon |y-b| = \varepsilon \|(x,y) - (a,b)\|_1,$$

lo que prueba la derivabilidad del campo escalar f en el punto (a,b).

Caso $\mathbb{R}^N$.

Sea f un campo escalar definido en $A \subset \mathbb{R}^N$ tal que $D_1 f, \ldots, D_{N-1} f$ sean continuas en $a \in \overset{\circ}{A}$ y exista $D_N f(a)$. Entonces f es derivable en a.

Demostración:

Sea $\varepsilon > 0$. Tomemos $\delta > 0$ tal que

$$\|x-a\|_1 < \delta \Rightarrow \begin{cases} x \in A \\ \exists D_i f(x), \text{ para } i = 1, \ldots, N-1 \\ |D_i f(x) - D_i f(a)| < \varepsilon \text{ para } i = 1, \ldots, N-1 \\ |f(a_1, \ldots, a_{N-1}, x_N - f(a) - D_N f(a)(x_N - a_N)| \leq \varepsilon |x_N - a_N| \end{cases}$$

Sea $x \in \mathbb{R}^N$ tal que $0 < \|x-a\|_1 < \delta$. Escribiendo

$$f(x) - f(a) - (\nabla f(a)|x-a) =$$

$$\sum_{i=1}^{N} [f(a_1, \ldots, a_{i-1}, x_i, x_{i+1}, \ldots, x_N) - f(a_1, \ldots, a_{i-1}, a_i, x_{i+1}, \ldots, x_N) -$$

$$-D_i f(a)(x_i - a_i)],$$

basta probar que para cada $i \in \{1, \ldots, N\}$, se verifica

$$|f(a_1, \ldots, a_{i-1}, x_i, x_{i+1}, \ldots, x_N) - f(a_1, \ldots, a_{i-1}, a_i, x_{i+1}, \ldots, x_N) - D_i f(a)(x_i - a_i)| \leq$$

$$\leq \varepsilon |x_i - a_i|.$$

En efecto, para $i = N$ es la última condición impuesta a δ. Sea $i = 1, \ldots, N-1$. Si $x_i = a_i$, la desigualdad es claramente cierta. Supuesto $x_i \neq a_i$, podemos aplicar el teorema del valor medio para encontrar y_i entre x_i y a_i tal que

$$f(a_1, \ldots, a_{i-1}, x_i, x_{i+1}, \ldots, x_N) - f(a_1, \ldots, a_{i-1}, a_i, x_{i+1}, \ldots, x_N) =$$

$$D_i f(a_1, \ldots, a_{i-1}, y_i, x_{i+1}, \ldots, x_N)(x_i - a_i),$$

de donde se deduce la desigualdad anunciada en sentido estricto a partir de las condiciones impuestas a δ.

Nota: Es claro que el papel jugado por N lo puede jugar cualquier otro natural m con $1 \leq m \leq N-1$.

3.11 Como

$$h(x) := (f(x)|g(x)) = \sum_{i=1}^{3} f_i(x) g_i(x)$$

se tiene que

$$h'(x) = \sum_{i=1}^{3} [f_i'(x) g_i(x) + f_i(x) g_i'(x)] = (f'(x)|g(x)) + (f(x)|g'(x)).$$

Como

$$\varphi(x) := f(x) \wedge g(x) = \begin{vmatrix} i & j & k \\ f_1(x) & f_2(x) & f_3(x) \\ g_1(x) & g_2(x) & g_3(x) \end{vmatrix}$$

$$= (f_2(x)g_3(x) - f_3(x)g_2(x), -f_1(x)g_3(x) + f_3(x)g_1(x), f_1(x)g_2(x) - f_2(x)g_1(x))$$

se tiene que

$$\varphi'(x) = (f_2'(x)g_3(x) + f_2(x)g_3'(x) - f_3'(x)g_2(x) - f_3(x)g_2'(x),\ \ldots.\ ,\ \ldots.\ )$$

$$= (f_2'(x)g_3(x) - f_3'(x)g_2(x), -f_1'(x)g_3(x) + f_3'(x)g_1(x), f_1'(x)g_2(x) - f_2'(x)g_1(x))$$
$$+ (f_2(x)g_3'(x) - f_3(x)g_2'(x), -f_1(x)g_3'(x) + f_3(x)g_1'(x), f_1(x)g_2'(x) - f_2(x)g_1'(x))$$
$$= \begin{vmatrix} i & j & k \\ f_1'(x) & f_2'(x) & f_3'(x) \\ g_1(x) & g_2(x) & g_3(x) \end{vmatrix} + \begin{vmatrix} i & j & k \\ f_1(x) & f_2(x) & f_3(x) \\ g_1'(x) & g_2'(x) & g_3'(x) \end{vmatrix} = f'(x) \wedge g(x) + f(x) \wedge g'(x).$$

3.12 Equivalen:

i) $x \frac{\partial f}{\partial y}(x,y) - y \frac{\partial f}{\partial x}(x,y) = f(x,y),\ \forall (x,y) \in A.$

ii) $\frac{\partial h}{\partial \vartheta}(\rho, \vartheta) = h(\rho, \vartheta),\ \forall (\rho, \vartheta) \in \Omega.$

La regla de la cadena para las derivadas parciales nos da

$$\frac{\partial h}{\partial \vartheta}(\rho, \vartheta) = \frac{\partial f}{\partial x}(\rho \cos \vartheta, \rho \operatorname{sen} \vartheta)(-\rho \operatorname{sen} \vartheta) + \frac{\partial f}{\partial y}(\rho \cos \vartheta, \rho \operatorname{sen} \vartheta)(\rho \cos \vartheta)$$

es decir

$$\frac{\partial h}{\partial \vartheta}(\rho, \vartheta) = -y \frac{\partial f}{\partial x}(x,y) + x \frac{\partial f}{\partial y}(x,y),\ \text{si } (x,y) = (\rho \cos \vartheta, \rho \operatorname{sen} \vartheta)$$

de donde se obtiene $i) \Leftrightarrow ii)$.

Es fácil comprobar que las soluciones de $ii)$ son de la forma

$$h(\rho,\vartheta) = \varphi(\rho)e^{\vartheta}, \text{ con } \varphi : \mathbb{R}^+ \to \mathbb{R} \text{ derivable}$$

con lo que las soluciones de $i)$ son de la forma

$$f(x,y) = f(\rho\cos\ \vartheta, \rho\,\mathrm{sen}\ \vartheta) = h(\rho,\vartheta) = \varphi(\rho)e^{\vartheta}$$

y al estar en el primer cuadrante, concluimos que

$$f(x,y) = \varphi(\sqrt{x^2+y^2})exp\left(\arctan\frac{y}{x}\right), \text{ con } \varphi : \mathbb{R}^+ \to \mathbb{R} \text{ derivable}$$

3.13 Las funciones homogéneas de grado uno derivables en 0 son lineales, y por tanto de clase $\mathscr{C}^1$.

Como f es derivable en 0, es continua en 0. Sea x un vector no nulo, se tiene

$$f(0) = \lim_{t\to 0, t>0} f(tx) = \lim_{t\to 0, t>0} tf(x) = 0$$

con lo que, por ser derivable en 0, ha de ser

$$Df(0)(x) = f'(0;x) = \lim_{t\to 0^+}\frac{f(tx)-f(0)}{t} = \lim_{t\to 0^+}\frac{f(tx)}{t} = \lim_{t\to 0^+}\frac{tf(x)}{t} = f(x).$$

Hemos probado que $f = Df(0)$.

3.14 **Teorema de Euler.**

$$\exists p\in\mathbb{R} : f(tx) = t^p f(x), \forall x\in\mathbb{R}^N\setminus\{0\}, \forall t\in\mathbb{R}^+ \ \Leftrightarrow$$
$$\Leftrightarrow\ Df(x)(x) = pf(x), \forall x\in\mathbb{R}^N\setminus\{0\}$$

$\Rightarrow]$ Para $x\in\mathbb{R}^N\backslash\{0\}$ se tiene que

$$\frac{f(x+tx)-f(x)}{t} = \frac{(1+t)^p-1}{t}f(x)\ \to\ pf(x), \text{ cuando } t\to\ 0$$

(donde se ha usado la expresión de la derivada en el punto 1 de la función potencial), es decir

$$Df(x)(x) = f'(x;x) = pf(x), \forall x\in\mathbb{R}^N\backslash\{0\}.$$

$\Leftarrow]$ Fijado $x\in\mathbb{R}^N\backslash\{0\}$, definimos $g:\mathbb{R}^+\ \to\ \mathbb{R}$ por $g(t) := \frac{f(tx)}{t^p}$, $\forall t\in\mathbb{R}^+$. La función g es derivable pues, la regla de la cadena, nos asegura que la función real de variable real $\varphi : t\to f(tx)$ es derivable con

$$\varphi'(t) = Df(tx)(x), \forall t\in\mathbb{R}^+$$

por lo que, para cualquier t de $\mathbb{R}^+$, tenemos

$$g'(t) = \frac{t^p Df(tx)(x) - pt^{p-1}f(tx)}{t^{2p}} = \frac{t^{p-1}(Df(tx)(tx) - pf(tx))}{t^{2p}} = 0$$

y al ser $\mathbb{R}^+$ un intervalo deducimos que g es constante, así

$$g(t) = g(1) = f(x), \forall t \in \mathbb{R}^+ \Leftrightarrow f(tx) = t^p f(x), \forall t \in \mathbb{R}^+.$$

La arbitrariedad de x concluye la demostración.

3.15 Basta observar la igualdad de los siguientes pasos

$$\lim_{z\to z_0} \frac{f(z)-f(z_0)}{z-z_0} = A+iB$$

$$\lim_{z\to z_0} \frac{f(z)-f(z_0)-(A+iB)(z-z_0)}{z-z_0} = 0$$

$$\lim_{z\to z_0} \frac{|f(z)-f(z_0)-(A+iB)(z-z_0)|}{|z-z_0|} = 0$$

$$\lim_{x+iy\to a+ib} \frac{|f(x+iy)-f(a+ib)-(A+iB)((x-a)+i(y-b))|}{|x-a+i(y-b)|} = 0$$

$$\lim_{(x,y)\to(a,b)} \frac{\|f(x,y)-f(a,b)-(A(x-a)-B(y-b),A(y-b)+B(x-a))\|_2}{\|(x-a,y-b)\|_2} = 0$$

$$\lim_{(x,y)\to(a,b)} \frac{\left\| f(x,y)-f(a,b)-\left[\begin{pmatrix} A & -B \\ B & A \end{pmatrix}\begin{pmatrix} x-a \\ y-b \end{pmatrix}\right]^T \right\|_2}{\|(x-a,y-b)\|_2} = 0$$

$$\lim_{(x,y)\to(a,b)} \frac{f(x,y)-f(a,b)-Df(a,b)(x-a,y-b)}{\|(x-a,y-b)\|_2} = 0$$

3.16

$$\left[f \text{ derivable }, \sum_{i,j=1}^{n} D_i f_j(x) h_i h_j > 0, \forall x \in A, \forall h \in \mathbb{R}^n \backslash \{0\}\right] \Rightarrow f \text{ es inyectiva.}$$

Para $a, b \in A$, definimos la función $g : [0,1] \to \mathbb{R}$ por

$$g(t) = (f(a+t(b-a))|b-a), \forall t \in [0,1].$$

Razonando, por reducción al absurdo, si $f(a) = f(b)$ para algunos $b \neq a$, se tendría

$$g(1) - g(0) = (f(b) - f(a)|b-a) = 0$$

y el teorema del valor medio aseguraría que existe $\vartheta \in]0,1[$:

$$0 = g'(\vartheta) = (Df(a+\vartheta(b-a))(b-a)|b-a)$$

lo que contradice la hipótesis, ya que

$$(Df(x)(h)|h) = \sum_{i,j=1}^{n} D_i f_j(x) h_i h_j > 0.$$

3.17 Sabemos que cualquier norma no es derivable en el origen.

a) $\|\cdot\|_1$

Estudio en los ejes coordenados.

$\|.\|_1$ no es derivable en los ejes coordenados.

Consideremos un punto $(a,0)$ con $a \neq 0$. Se tiene que

$$\frac{\|(a,y)\|_1 - \|(a,0)\|_1}{y} = \frac{|y|}{y}$$

que no tiene límite cuando $y \to 0$. Es decir no existe $D_2\|\cdot\|_1(a,0)$ y en consecuencia $\|\cdot\|_1$ no es derivable en $(a,0)$. Análogamente $\|\cdot\|_1$ no es derivable en $(0,b)$ con $b \neq 0$.

Estudio fuera de los ejes coordenados.

Consideremos un punto (a,b) con $ab \neq 0$. Se tiene que

$$\frac{\|(x,b)\|_1 - \|(a,b)\|_1}{x-a} = \frac{|x|-|a|}{x-a} \to \frac{a}{|a|} \text{ cuando } x \to a.$$

Es decir $D_1\|\cdot\|_1(a,b) = \frac{a}{|a|}$. Análogamente $D_2\|\cdot\|_1(a,b) = \frac{b}{|b|}$.

Resumiendo

$$\nabla\|\cdot\|_1(x,y) = \left(\frac{x}{|x|}, \frac{y}{|y|}\right),\ \forall (x,y) \in \mathbb{R}^2 \setminus \{(x,y) \in \mathbb{R}^2 :\ xy = 0\}.$$

Al ser el gradiente continuo la función $\|\cdot\|_1$ es de clase $\mathscr{C}^1$ en el conjunto considerado.

Caso $\mathbb{R}^N$.

Consideremos un punto $a = (a_1, \ldots, a_N)$

Si $\exists k \in \{1,...,N\} : a_k = 0$ entonces $\|.\|_1$ no es derivable en a. En caso contrario $D\|.\|_1(a)(x) = \sum_{k=1}^{N} \frac{a_k}{|a_k|}$. Además la función $\|\cdot\|_1$ es de clase $\mathscr{C}^1$ en el conjunto en cuestión.

b) $\|\cdot\|_2$

Consideremos un punto $(x,y) \neq (0,0)$. Utilizando las reglas de derivación se tiene que

$$\nabla\|.\|_2(x,y) = \left(\frac{x}{\sqrt{x^2+y^2}}, \frac{y}{\sqrt{x^2+y^2}}\right),\ \forall (x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}.$$

Al ser el gradiente continuo la función $\|\cdot\|_1$ es de clase $\mathscr{C}^1$ en el conjunto considerado.

Caso $\mathbb{R}^N$.

Consideremos un punto $a = (a_1, \dots, a_N) \neq (0, \dots, 0)$. Se tiene que

$$D\|.\|_2(a)(x) = \frac{(a|x)}{\|a\|_2},$$

que es una generalización de la derivada de $|\cdot|$. Además la función $\|\cdot\|_2$ es de clase $\mathscr{C}^1$ en el conjunto en cuestión.

c) $\|\cdot\|_\infty$

Estudio en las diagonales de los cuadrantes.

Consideremos un punto (a,b) con $0 < |a| = |b|$. Para $x \in \mathbb{R}$ con $0 < |x| < |a|$ se tiene que

$$\frac{\|(a+x,b)\|_\infty - \|(a,b)\|_\infty}{x} = \begin{cases} \frac{|a+x|-|a|}{x} & \text{si } ax > 0 \\ 0 & \text{si } ax < 0 \end{cases}.$$

Como $\frac{|a+x|-|a|}{x} \to \frac{a}{|a|}$ cuando $x \to 0$, concluimos que no existe $D_1\|\cdot\|_\infty(a,b)$. Análogamente no existe $D_2\|\cdot\|_\infty(a,b)$. En consecuencia

$\|.\|_\infty$ no es derivable en las diagonales de los cuadrantes.

Estudio fuera de las diagonales de los cuadrantes.

Consideremos un punto (a,b) con $0 < |a| < |b|$. Para $x \in \mathbb{R}$ con $0 < |x| < |b| - |a|$, se tiene que

$$\frac{\|(a+x,b)\|_\infty - \|(a,b)\|_\infty}{x} = \frac{|b| - |b|}{x} = 0$$

pues $|a+x| \leq |a| + |x| < |b|$, y para $y \in \mathbb{R}$ con $0 < |y| < |b| - |a|$ se tiene que

$$\frac{\|(a,b+y)\|_\infty - \|(a,b)\|_\infty}{y} = \frac{|b+y| - |b|}{y} \to \frac{b}{|b|}$$

pues $|b+y| \geq |b| - |-y| > |a|$. En consecuencia el gradiente de la función $\|\cdot\|_\infty$ es el que se anuncia. Análogamente para puntos de la forma $0 < |b| < |a|$. Por último como el gradiente es continuo la función $\|\cdot\|_\infty$ es de clase $\mathscr{C}^1$ en el conjunto que se considera.

Caso $\mathbb{R}^N$.

Consideremos un punto $a = (a_1, \dots, a_N)$. Si $\exists k, q \in \{1, ..., N\} : |a_k| = |a_q| = |a\|_\infty$, entonces $\|.\|_\infty$ no es derivable en a. Si $\exists_1 k \in \{1, ..., N\} : |a_k| = \|a\|_\infty$ entonces $D\|.\|_\infty(a)(x) = \frac{a_k}{|a_k|} x_k$. Además la función $\|\cdot\|_\infty$ es de clase $\mathscr{C}^1$ en el conjunto en cuestión.

3.18 $i) \Rightarrow ii)$ Por ser f derivable en a, se tiene que

$$f'(a;x) = Df(a)(x), \ \forall x \in \mathbb{R}^N$$

con lo que existe $f'(a;\cdot)$ y es lineal.

Sean $B \subset \mathbb{R}^N$ acotado y $M \in \mathbb{R}$ tal que $\|x\| < M, \ \forall x \in B$. Por ser f derivable en a, existe $\delta > 0$ tal que

$$\|y-a\| < \delta \Rightarrow \|f(y) - f(a) - f'(a;y-a)\| < \frac{\varepsilon}{M}\|y-a\|.$$

Sea $x \in B\backslash\{0\}$ y tomemos $|t| < \frac{\delta}{M}$. Entonces el vector $y = a + tx$ verifica $\|y-a\| = |t|\|x\| < \frac{\delta}{M}M$, luego, en virtud de la desigualdad anterior se tiene

$$\|f(a+tx) - f(a) - f'(a;tx)\| < \frac{\varepsilon}{M}|t|\|x\| < \varepsilon|t|,$$

esto es

$$\left\|\frac{f(a+tx) - f(a) - tf'(a;x)}{t}\right\| < \varepsilon,$$

y por tanto

$$\left\|\frac{f(a+tx) - f(a)}{t} - f'(a;x)\right\| < \varepsilon, \ \forall x \in B.$$

ii) $\Rightarrow$ *i)* Sea $T \in \mathscr{L}(\mathbb{R}^N, \mathbb{R}^M)$ la función definida por $T(x) = f'(a;x), \ \forall x \in \mathbb{R}^N$. Sea $x \in A\backslash\{a\}$. Notamos $h := x - a$ con lo que $h \neq 0$, y se tiene que

$$\frac{f(x) - f(a) - T(x-a)}{\|x-a\|} = \frac{f(a+h) - f(a) - T(h)}{\|h\|} =$$

$$\frac{f\left(a + \|h\|\frac{h}{\|h\|}\right) - f(a) - \|h\|T\left(\frac{h}{\|h\|}\right)}{\|h\|} = \frac{f\left(a + \|h\|\frac{h}{\|h\|}\right) - f(a)}{\|h\|} - T\left(\frac{h}{\|h\|}\right) =$$

$$\frac{f\left(a + \|h\|\frac{h}{\|h\|}\right) - f(a)}{\|h\|} - f'\left(a; \frac{h}{\|h\|}\right) \to 0 \text{ cuando } h \to 0$$

pues el límite es uniforme en $S(\mathbb{R}^N)$. Hemos probado que

$$\frac{f(x) - f(a) - T(x-a)}{\|x-a\|} \to 0 \text{ cuando } x \to a$$

y por tanto f es derivable en a con $Df(a) = T$.

3.19 En efecto, dada $\{t_n\}$ sucesión de $]0,1[$ convergente a cero, tómese la sucesión $\{u_n\}$ en la esfera unidad euclídea de $\mathbb{R}^2$ dada por $u_n = (\sqrt{1-t_n^2}, t_n)$, y nótese que

$$\frac{f(a+t_n u_n) - f(a)}{t_n} = \frac{t_n^4(1-t_n^2)^3}{t_n^6 + t_n^4(1-t_n^2)^3} = \frac{(1-t_n^2)^3}{t_n^2 + (1-t_n^2)^3} \to 1.$$

Alternativamente: Si $u \in S_{\mathbb{R}^2}$, entonces existe $\vartheta \in]-\pi, \pi]$ con $u = (\cos\vartheta, \operatorname{sen}\vartheta)$. Sea

$$g(t,\vartheta) := \frac{f(t(\cos\vartheta, \operatorname{sen}\vartheta)) - f(0,0)}{t} = \frac{t^3|t|\cos^6\vartheta}{(\operatorname{sen}\vartheta - t\cos^2\vartheta)^2 + t^4\cos^6\vartheta}.$$

Si el límite en 0 de $t \mapsto \dfrac{f(tu)-f(0)}{t}$ fuese uniforme en $u \in S_{\mathbb{R}^2}$, entonces,

$$\forall \varepsilon > 0, \exists \delta > 0 : |t| < \delta \Rightarrow |g(t,\vartheta)| \leq \varepsilon, \forall\ \vartheta \in]-\pi,\pi].$$

Por tanto, si $\{t_n\} \to 0$, se tendría que $g(t_n,\vartheta_n) \to 0$ para cualquier sucesión $\{\vartheta_n\}$.

En este caso se tiene que si $\{\vartheta_n\} \to 0$ y $0 < \vartheta_n < \frac{\pi}{2}$ para cada n, entonces tomamos $t_n = \dfrac{\operatorname{sen}\vartheta_n}{\cos^2\vartheta_n} \to 0$ y $g(t_n,\vartheta_n) = \dfrac{t_n^4\cos^6\vartheta}{0+t_n^4\cos^6\vartheta_n} = 1$, para cada natural n. Luego el límite anterior no es uniforme en $u \in S_{\mathbb{R}^2}$.

Tema 4

Teorema del valor medio. Teoremas del punto fijo de Banach y de Schauder. Teorema de Picard-Lindelöf.

En la primera sección obtenemos una generalización del Teorema del valor medio real:

$$f(b)-f(a)=f'(c)(b-a) \text{ para conveniente } c\in]a,b[$$

a campos escalares

$$f(b)-f(a)=(\nabla f(c)|b-a) \text{ para conveniente } c\in]a,b[.$$

Fijadas dos normas en $\mathbb{R}^N$ y $\mathbb{R}^M$, probamos el importante resultado conocido como Teorema del valor medio o Desigualdad de Lagrange para campos vectoriales

$$\|f(a)-f(b)\|\leq\|b-a\|\sup\{\|Df(x)\|: x\in]a,b[\},$$

donde se considera en $\mathscr{L}(\mathbb{R}^N,\mathbb{R}^M)$ la norma de operadores.

Finalmente para las normas euclídeas probamos también una práctica desigualdad para campos vectoriales

$$\|f(b)-f(a)\|_2\leq\|b-a\|_2\,\sup\left\{\sqrt{\sum_{i,j}D_jf_i(x)^2}:\ x\in]a,b[\right\},$$

que se puede recordar como que la norma de operadores se mayora por la "norma euclídea" de la correspondiente matriz.

En la segunda sección definimos las funciones contractivas como las funciones lipschitzianas de constante de Lipschitz menor que 1. Obtenemos el importantísimo resultado teórico-práctico conocido como Teorema del punto fijo de Banach para espacios métricos completos (4.12) y deducimos de él el Teorema de Schauder para campos vectoriales de $\mathbb{R}^N$ en $\mathbb{R}^N$ (4.15).

Los Teoremas del valor medio y de Schauder serán herramientas fundamentales en la demostración del Teorema de la función inversa.

Dedicamos la última sección a demostrar otra importante consecuencia del Teorema del punto fijo de Banach, el Teorema de Picard-Lindelöf sobre la existencia de soluciones de ecuaciones diferenciales (4.18).

4.1. Teorema del valor medio.

Es conocido el siguiente importante resultado:

Teorema 4.1 (valor medio real)**.** *Sean $a,b \in \mathbb{R}$ con $a \neq b$ y sea I el intervalo cerrado de extremos a y b. Si $f : I \to \mathbb{R}$ es una función continua en I y derivable en $\overset{\circ}{I}$, entonces existe $c \in \overset{\circ}{I}$ tal que*

$$f(b) - f(a) = f'(c)(b-a).$$

Este teorema admite una generalización inmediata al caso de campos escalares definidos en $\mathbb{R}^N$. Recordemos que si a y b son dos puntos distintos de $\mathbb{R}^N$, entonces los segmentos cerrado y abierto de extremos a y b son respectivamente los conjuntos

$$[a,b] = \{a + t(b-a) : \ t \in [0,1]\} \ \text{ y } \ ]a,b[= \{a + t(b-a) : \ t \in]0,1[\}.$$

Teorema 4.2 (valor medio para campos escalares)**.** *Sean $a,b \in \mathbb{R}^N$ verificando que $a \neq b$ y que $[a,b] \subset A \subset \mathbb{R}^N$. Si $f : A \to \mathbb{R}$ es un campo escalar continuo en $[a,b]$ y derivable en $]a,b[$, entonces existe $c \in]a,b[$ tal que*

$$f(b) - f(a) = \big(\nabla f(c)|b-a\big).$$

Demostración:

La prueba de este teorema consiste en aplicar el Teorema del valor medio real a la función auxiliar $\sigma : [0,1] \to \mathbb{R}$ definida por

$$\sigma(t) = f(a + t(b-a)).$$

Puesto que σ es la composición con f de la función de $[0,1]$ en A definida por

$$t \mapsto a + t(b-a),$$

la regla de la cadena para funciones continuas nos permite afirmar que σ es continua en $[0,1]$. La nota 3.25 a) nos garantiza que σ es derivable en $]0,1[$ con

$$\sigma'(t) = D\sigma(t)(1) = Df(a + t(b-a))(b-a) = \big(\nabla f(a + t(b-a))|b-a\big).$$

Así, la función σ cumple las hipótesis del Teorema del valor medio real y, por tanto, existe $\vartheta \in]0,1[$ tal que

$$f(b) - f(a) = \sigma(1) - \sigma(0) = \big(\nabla f(a + \theta(b-a))|b-a\big).$$

∎

Es usual que el punto c que aparece en el enunciado del teorema anterior no se sepa calcular. El siguiente corolario nos da una mayoración de la distancia entre los valores tomados por un campo escalar en los extremos del segmento obtenida usando la desigualdad 3.1.5. No se debe olvidar que tal estimación es la responsable de la mayoría de las aplicaciones del Teorema del valor medio real.

Corolario 4.3. *Sean $a,b \in \mathbb{R}^N$ con $a \neq b$ tales que $[a,b] \subset A \subset \mathbb{R}^N$. Si $f : A \to \mathbb{R}$ es un campo escalar continuo en $[a,b]$ y derivable en $]a,b[$, entonces*

$$|f(b) - f(a)| \leq \|b-a\| \sup\{\|Df(x)\| : x \in]a,b[\}.$$

donde $\|\cdot\|$ es una norma en $\mathbb{R}^N$ y el supremo está tomado en $[-\infty,+\infty]$.

Simples ejemplos de funciones con valores en $\mathbb{R}^2$ prueban que no es posible la extensión literal del Teorema del valor medio real a campos vectoriales. He aquí uno de ellos.

Ejemplo 4.4. Sea $f : \mathbb{R} \to \mathbb{R}^2$ el campo vectorial dado por

$$f(x) = (\cos x, \operatorname{sen} x).$$

Es claro que f es continuo en $[0,2\pi]$ y derivable en $]0,2\pi[$ con

$$f'(x) = (-\operatorname{sen} x, \cos x) \neq (0,0), \ \forall x \in]0,2\pi[.$$

Se tiene pues para $x \in]0,2\pi[$ que

$$(0,0) = f(2\pi) - f(0) \neq 2\pi f'(x).$$

Nuestro próximo objetivo es probar que el corolario anterior sí puede extenderse a campos vectoriales. Esta extensión es la que se conoce como Teorema del valor medio, y también como Fórmula del incremento finito o Desigualdad de Lagrange.

Teorema 4.5 (valor medio)**.** *Sean $a,b \in \mathbb{R}^N$ verificando que $a \neq b$ y que $[a,b] \subset A \subset \mathbb{R}^N$. Si $f : A \to \mathbb{R}^M$ es un campo vectorial continuo en $[a,b]$ y derivable en $]a,b[$, entonces*

$$\|f(b) - f(a)\| \leq \|b-a\| \sup\{\|Df(x)\| : x \in]a,b[\}.$$

Demostración:

Si el conjunto

$$\{\|Df(x)\| : x \in]a,b[\}$$

no está mayorado, no hay nada que probar. En otro caso sea

$$M := \sup\{\|Df(x)\| : x \in]a,b[\}.$$

Consideremos, como en la prueba del Teorema 4.2, la función auxiliar $\sigma : [0,1] \to \mathbb{R}^M$ definida por

$$\sigma(t) := f(a + t(b-a)).$$

Sabemos que σ es continua en $[0,1]$ y derivable en $]0,1[$ con (ver Nota 3.25 a))

$$\sigma'(t) = Df(a + t(b-a))(b-a), \ \forall t \in]0,1[,$$

y por tanto

$$\|\sigma'(t)\| \leq \|Df(a+t(b-a))\|\|b-a\| \leq M\|b-a\|, \; \forall t \in]0,1[.$$

Puesto que

$$\|f(b)-f(a)\| = \|\sigma(1)-\sigma(0)\|,$$

el teorema estará probado si comprobamos que

$$\|\sigma(1)-\sigma(0)\| \leq M\|b-a\|.$$

La prueba de dicha desigualdad es consecuencia inmediata del siguiente resultado técnico que tiene importancia en sí mismo.

Proposición 4.6. *Sean* $\sigma : [0,1] \to \mathbb{R}^M$ *y* $g : [0,1] \to \mathbb{R}$ *funciones continuas en* $[0,1]$*, derivables en* $]0,1[$ *y que verifican*

$$\|\sigma'(t)\| \leq g'(t), \; \forall t \in]0,1[.$$

Entonces

$$\|\sigma(1)-\sigma(0)\| \leq g(1)-g(0).$$

Demostración:

Probaremos que para cada $\varepsilon > 0$ se verifica

$$\|\sigma(1)-\sigma(0)\| \leq g(1)-g(0)+2\varepsilon,$$

resultado del que se sigue el enunciado dada la arbitrariedad de ε.

Fijemos $\varepsilon > 0$ y consideremos el conjunto

$$A := \{t \in [0,1] : \; \|\sigma(t)-\sigma(0)\| < g(t)-g(0)+\varepsilon t+\varepsilon\}.$$

Es claro que $0 \in A$. Sea $a := \sup(A)$. De la continuidad de σ y g se deduce inmediatamente que

$$\|\sigma(a)-\sigma(0)\| \leq g(a)-g(0)+\varepsilon a+\varepsilon, \tag{4.1.1}$$

y que A es un abierto relativo de $[0,1]$ y por tanto $0 < a$. Veamos que $a = 1$ y en consecuencia

$$\|\sigma(1)-\sigma(0)\| \leq g(1)-g(0)+2\varepsilon,$$

como se pretende demostrar.

La prueba de que $a = 1$ la haremos por reducción al absurdo. Supongamos por tanto que $a < 1$ y veamos que entonces existe $s > 0$ tal que $a+s \in A$, en contra del hecho de ser a el supremo de A. En efecto, de la derivabilidad de σ y g en a, podemos tomar $s > 0$ tal que $a+s < 1$ y

$$\frac{\|\sigma(a+s)-\sigma(a)-s\sigma'(a)\|}{s} < \frac{\varepsilon}{2} \tag{4.1.2}$$

y

$$\frac{|g(a+s)-g(a)-sg'(a)|}{s} < \frac{\varepsilon}{2}. \tag{4.1.3}$$

Se tiene que

$$\|\sigma(a+s)-\sigma(0)\| = \|\sigma(a+s)-\sigma(a)-s\sigma'(a)+s\sigma'(a)+\sigma(a)-\sigma(0)\| \leq$$
$$\|\sigma(a+s)-\sigma(a)-s\sigma'(a)\|+s\|\sigma'(a)\|+\|\sigma(a)-\sigma(0)\| <$$
$$\frac{\varepsilon}{2}s+sg'(a)+g(a)-g(0)+\varepsilon a+\varepsilon =$$

(donde hemos utilizado 4.1.2, la desigualdad de la hipótesis y 4.1.1)

$$= \frac{\varepsilon}{2}s+\Big(sg'(a)-g(a+s)+g(a)\Big)+g(a+s)-g(0)+\varepsilon a+\varepsilon <$$
$$\frac{\varepsilon}{2}s+\frac{\varepsilon}{2}s+g(a+s)-g(0)+\varepsilon a+\varepsilon = g(a+s)-g(0)+\varepsilon(a+s)+\varepsilon.$$

(donde se ha utilizado 4.1.3)

Así $a+s\in A$ lo que concluye la demostración. ∎

Fin de la demostración del teorema:

Cosideremos la función $g:[0,1]\to\mathbb{R}$ definida por

$$g(t) := M\|b-a\|t,$$

la Proposición 4.6 nos asegura que $\|\sigma(1)-\sigma(0)\| \leq g(1)-g(0)$, o equivalentemente $\|f(b)-f(a)\| \leq M\|b-a\|$, es decir

$$\|f(b)-f(a)\| \leq \|b-a\| \sup\{\|Df(x)\| : x\in]a,b[\}.$$

∎

El Teorema del valor medio será aplicado reiteradas veces a lo largo de este curso. Basta como botón de muestra los siguientes resultados que generalizan los correspondientes para funciones reales de variable real (hemos comentado que la desigualdad del Corolario 4.3 es la responsable de la mayoría de las aplicaciones).

Corolario 4.7. *Sean A un abierto convexo de $\mathbb{R}^N$ y $f:A\to\mathbb{R}^M$ un campo vectorial derivable verificando que existe un número real M tal que*

$$\|Df(x)\| \leq M, \ \forall x\in A,$$

entonces f es lipschitziano con constante de Lipschitz $\leq M$.

Demostración:

Dados $x,y \in A$, veamos que $\|f(y)-f(x)\| \leq M\|y-x\|$. Ello es claro si $x=y$. En otro caso el Teorema del valor medio aplicado a f en el segmento $[x,y] \subset A$ nos da que

$$\|f(y)-f(x)\| \leq \|y-x\| \sup\{\|Df(z)\| : \ z \in]x,y[\}$$

y el resultado se sigue sin más que aplicar la hipótesis. ∎

Corolario 4.8. *Sean A un abierto conexo de* $\mathbb{R}^N$ *y* $f : A \to \mathbb{R}^M$ *un campo vectorial derivable verificando que*

$$Df(x) = 0, \ \forall x \in A,$$

entonces f es constante.

Demostración:

Sea a un punto de A y definamos

$$B := \{x \in A : \ f(x) = f(a)\}.$$

B es un cerrado relativo (f es continua) y no es vacío ($a \in B$). Veamos que B es abierto. Fijemos $b \in B$ y tomemos $\rho > 0$ tal que $B(b,\rho) \subset A$. Al ser $B(b,\rho)$ un convexo el corolario anterior nos dice que la restricción de f a $B(b,\rho)$ es constante, es decir $B(b,\rho) \subset B$. Hemos probado que B es un conjunto abierto. Al ser A conexo no queda más remedio que ser $B = A$, es decir, f es constante. ∎

Aunque el Teorema de valor medio es, como se acaba de comprobar, una valiosa herramienta teórica, si recordamos que fijadas dos normas en $\mathbb{R}^N$ y $\mathbb{R}^M$ es

$$\|Df(a)\| := \max\{\|Df(a)(x)\| : \ \|x\| = 1\},$$

concluimos que la aplicación práctica de dicho teorema no es fácil (piénsese que en el Teorema 4.5 intervienen tres normas: una en $\mathbb{R}^N$, otra en $\mathbb{R}^M$ y la correspondiente en $\mathscr{L}(\mathbb{R}^N,\mathbb{R}^M)$)

Terminamos esta sección obteniendo una versión práctica del Teorema del valor medio cuando se consideran las normas euclídeas en $\mathbb{R}^N$ y $\mathbb{R}^M$ cuya prueba consiste en obtener una mayoración calculable de la norma de operadores (la "norma euclídea" de la matriz asociada).

Corolario 4.9 (Teorema práctico del valor medio)**.** *Sean* $a,b \in \mathbb{R}^N$ *con* $a \neq b$ *tales que* $[a,b] \subset A \subset \mathbb{R}^N$. *Si* $f : A \to \mathbb{R}^M$ *es un campo vectorial continuo en* $[a,b]$ *y derivable en* $]a,b[$, *entonces*

$$\|f(b)-f(a)\|_2 \leq \|b-a\|_2 \ \sup\left\{ \sqrt{\sum_{i,j} D_j f_i(x)^2} : \ x \in]a,b[\right\}.$$

Demostración:

Para cada $x \in]a,b[$ se tiene

$$\|Df(x)\| := \max \{\|Df(x)(y)\|_2 : \|y\|_2 = 1\} =$$

$$\max \left\{ \sqrt{\sum_{i=1}^{M} \big(\nabla f_i(x)|y\big)^2} : \|y\|_2 = 1 \right\} \leq \sqrt{\sum_{i=1}^{M} \|\nabla f_i(x)\|_2^2} = \sqrt{\sum_{i,j} D_j f_i(x)^2} \,,$$

donde se ha utilizado la desigualdad de Cauchy-Schwarz.

El corolario es ahora una consecuencia inmediata del Teorema 4.5. ∎

Usamos en un ejemplo concreto el resultado que acabamos de probar.

Ejemplo 4.10. Probar que el campo vectorial $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definido por

$$f(x,y) = \left(\operatorname{sen}(x+y), \sqrt{1+x^2+y^2}\right) \quad ((x,y) \in \mathbb{R}^2)$$

es lipschitziana con constante de Lipschitz menor o igual que $\sqrt{3}$.

Para cada $(x,y) \in \mathbb{R}^2$ se tiene que

$$J_f(x,y) = \begin{pmatrix} \cos(x+y) & \cos(x+y) \\ \frac{x}{\sqrt{1+x^2+y^2}} & \frac{y}{\sqrt{1+x^2+y^2}} \end{pmatrix},$$

y, en consecuencia,

$$\sum_{i,j} D_j f_i(x)^2 = 2\, \cos^2(x+y) + \frac{x^2+y^2}{1+x^2+y^2} \leq 3.$$

El enunciado se deduce del Corolario 4.9.

4.2. Teoremas del punto fijo de Banach y de Schauder.

El principio de contracción de Banach está basado en un proceso iterativo que permite obtener aproximaciones hacia un punto fijo cada vez mejores, incluso se puede saber el número de iteraciones suficientes para obtener una aproximación del punto fijo con un cierto margen de error.

Definición 4.11 (función contractiva)**.** Sean (E,d) y (F,ρ) espacios métricos. Una función $f : E \to F$ es contractiva si $\exists \alpha \in [0,1[$ tal que

$$\rho(f(x), f(y)) \leq \alpha\, d(x,y),\ \forall x,y \in E.$$

Es decir, las funciones contractivas son las funciones lipschitzianas cuya constante de Lipschitz es menor que 1.

Teorema 4.12 (punto fijo de Banach)**.** *Sea (E,d) un espacio métrico completo y sea $f : E \to E$ una función contractiva. Entonces existe un único punto a de E fijo por f, es decir tal que $f(a) = a$.*

De hecho, dado a_0 en E, la sucesión $\{a_n\}$ definida mediante la expresión

$$a_{n+1} := f(a_n),\ \forall n \in \mathbb{N} \cup \{0\},$$

verifica

$$\{d(a,a_n)\} \searrow 0.$$

Además si $0 \leq \alpha < 1$ es tal que

$$d(f(x), f(y)) \leq \alpha\, d(x,y),\ \forall x,y \in E, \tag{4.2.1}$$

entonces para cada $n \in \mathbb{N}$ se verifica la desigualdad:

$$d(a_n, a) \leq \frac{\alpha^n}{1-\alpha}\, d(a_1, a_0).$$

Demostración:

Unicidad: Supongamos que $a,b \in E$ son tales que $f(a) = a$ y $f(b) = b$, se tiene que $d(a,b) = d(f(a), f(b)) \leq \alpha\, d(a,b)$ y por tanto $0 \leq (1-\alpha)d(a,b) \leq 0$. De donde se sigue que $a = b$. Hemos probado que no puede haber más de un punto fijo.

Existencia: Sea $a_0 \in E$. Para $n \geq 0$, definimos $a_{n+1} := f(a_n)$. La sucesión $\{a_n\}$ así obtenida es de Cauchy. En efecto, veamos que para cada natural n, se verifica la siguiente desigualdad:

$$d(a_{n+1}, a_n) \leq \alpha^n\, d(a_1, a_0).$$

La desigualdad es obvia para $n = 1$ por ser f contractiva. Supongamos que se verifica para n y probémosla para $n+1$. Se tiene

$$d(a_{n+2}, a_{n+1}) = d(f(a_{n+1}), f(a_n)) \leq \alpha\, d(a_{n+1}, a_n) \leq \alpha^{n+1}\, d(a_1, a_0)$$

donde se ha tenido en cuenta la definición de la sucesión $\{a_n\}$ y se ha utilizado que la función f es contractiva verificando 4.2.1.

Ahora para $n,h \in \mathbb{N}$ se tiene:

$$d(a_{n+h},a_n) \leq d(a_{n+h},a_{n+h-1}) + d(a_{n+h-1},a_{n+h-2}) + \cdots + d(a_{n+1},a_n) \leq$$

$$\left(\alpha^{n+h-1} + \alpha^{n+h-2} + \cdots + \alpha^n\right) d(a_1,a_0) = \frac{\alpha^n - \alpha^{n+h}}{1-\alpha}\, d(a_1,a_0) \leq \frac{\alpha^n}{1-\alpha}\, d(a_1,a_0).$$

De la anterior expresión se sigue que la sucesión $\{a_n\}$ es de Cauchy, y por tanto, en virtud de la complitud de E, convergente. Sea a el límite de $\{a_n\}$. La continuidad de la función f nos permite deducir que

$$a = \lim\, a_n = \lim\, a_{n+1} = \lim\, f(a_n) = f(\lim\, a_n) = f(a).$$

Tomando límite en h en la expresión

$$d(a_{n+h},a_n) \leq \frac{\alpha^n}{1-\alpha}\ d(a_1,a_0),$$

la continuidad de la aplicación distancia nos permite obtener que

$$d(a,a_n) \leq \frac{\alpha^n}{1-\alpha}\, d(a_1,a_0),\ \forall n \in \mathbb{N}.$$

Por último para todo natural n se tiene que

$$d(a,a_{n+1}) = d(f(a),f(a_n)) \leq \alpha\, d(a,a_n).$$

■

Corolario 4.13. *Sean E un espacio métrico completo y sea $f : E \to E$ una aplicación. Supongamos que existe $n \in \mathbb{N}$ tal que f^n es contractiva. Entonces f tiene un único punto fijo.*

Demostración:

El Teorema del punto fijo de Banach (4.12) nos asegura que f^n tiene un único punto fijo $a \in E$. Veamos que a es también punto fijo de f. En efecto

$$f^n(a) = a \Rightarrow f^n\big(f(a)\big) = f\big(f^n(a)\big) = f(a) \Rightarrow f(a) = a,$$

donde se ha utilizado que la función f^n tiene un único punto fijo.

Como cada punto fijo de f, también es punto fijo de f^n, entonces, a es el único punto fijo de f. ■

Ejemplo 4.14. Sea $f : \mathbb{R}^2 \to \mathbb{R}^2$ el campo vectorial definido por

$$f(x,y) = \left(\operatorname{sen}(x+y), \sqrt{1+x^2+y^2}\right).$$

Probar que existe un único $p \in \mathbb{R}^2$ tal que $f(p) = 2p$.

El ejemplo 4.10 nos asegura que la función $\dfrac{f}{2}$ es contractiva. Para obtener el resultado basta aplicar el Teorema 4.12 a dicha función.

El siguiente resultado, consecuencia del Teorema 4.12, será una pieza fundamental en la demostración del teorema de la función inversa.

Teorema 4.15 (Schauder)**.** *Sean $U \subset \mathbb{R}^N$ abierto y $\varphi : U \to \mathbb{R}^N$ una función contractiva. Si $f := i_U - \varphi$ (donde i_U es la aplicación inclusión de U en $\mathbb{R}^N$), entonces se verifica:*

i) $f(U)$ es un abierto de $\mathbb{R}^N$.

ii) f es un homeomorfismo (biyección bicontinua) de U sobre $f(U)$ (de hecho, f y f^{-1} son lipschitzianas).

iii) Si además $U = \mathbb{R}^N$, entonces f es un homeomorfismo de $\mathbb{R}^N$ sobre sí mismo.

La afirmación primera, que es la verdadera aportación de Schauder, es un punto crucial en la demostración del Teorema de la función inversa. Por razones didácticas conviene empezar demostrando la siguiente normalización de 4.15.

Lema 4.16 (versión normalizada del Teorema de Schauder)**.** *Sean D la bola unidad abierta de $\mathbb{R}^N$, $\varphi : D \to \mathbb{R}^N$ una función contractiva y $\alpha \in [0,1[$ tal que*

$$\|\varphi(y) - \varphi(x)\| \leq \alpha\, \|y-x\|,\ \forall x,y \in D.$$

Supongamos además que $\varphi(0) = 0$. Entonces si $f := i_D - \varphi$ (donde i_D es la aplicación inclusión de D en $\mathbb{R}^N$) se verifica:

$$B(0,1-\alpha) \subset f(D).$$

Demostración:

Sea $y \in B(0,1-\alpha)$. Consideremos la función auxiliar $g : D \to \mathbb{R}^N$ definida por

$$g(x) := y + \varphi(x).$$

Fijemos ρ tal que $\dfrac{\|y\|}{1-\alpha} < \rho < 1$, y veamos que $\overline{B}(0,\rho)$ es invariante por g. Para cada z de $\overline{B}(0,\rho)$ se tiene

$$\|g(z)\| \leq \|g(z)-g(0)\| + \|g(0)\| \leq \alpha\|z-0\| + \|y\| \leq \alpha\rho + (1-\alpha)\rho = \rho.$$

Aplicando el Teorema del punto fijo de Banach (4.12) a la función contractiva g en el espacio métrico completo $\overline{B}(0,\rho)$ (cerrado de un completo) obtenemos

$$\exists! x \in \overline{B}(0,\rho) \text{ tal que } x = g(x) \text{ , es decir } x = y + \varphi(x).$$

Equivalentemente $y = x - \varphi(x)$, es decir $y = f(x)$. Como $\overline{B}(0,\rho) \subset D$, se sigue que $y \in f(D)$. ∎

Nota: El anterior lema no se puede mejorar, es decir el radio de la bola de centro 0 contenida en $f(D)$ es óptimo. En efecto: la aplicación $\varphi : D \to \mathbb{R}^N$ definida por $\varphi(x) := \alpha x$ es contractiva de razón α y la función f asociada es tal que

$$f(D) = \{x - \varphi(x) : \ x \in D\} = \{(1-\alpha)x : \ x \in D\} = B(0, 1-\alpha).$$

Demostración del Teorema de Schauder:

i) Dado $a \in U$, tomemos $r > 0$ tal que $B(a,r) \subset U$. Probemos que si α es la constante de Lipschitz de φ se tiene que

$$B(f(a),(1-\alpha)r) \subset f(B(a,r)).$$

Consideremos la función $\overline{\varphi} : D \longrightarrow \mathbb{R}^N$ definida por

$$\overline{\varphi}(x) := \frac{1}{r}\,\big(\varphi(a+rx) - \varphi(a)\big).$$

Claramente $\overline{\varphi}(0) = 0$. Veamos que la aplicación $\overline{\varphi}$ es también contractiva con constante de Lipschitz menor o igual que α. Para cualesquiera $x_1, x_2 \in D$ se tiene

$$\|\overline{\varphi}(x_1) - \overline{\varphi}(x_2)\| = \frac{1}{r}\|\varphi(a+rx_1) - \varphi(a+rx_2)\| \leq$$

$$\frac{\alpha}{r}\|(a+rx_1) - (a+rx_2)\| = \alpha\|x_1 - x_2\|.$$

Aplicando el Lema 4.16, si $\overline{f} := i_D - \overline{\varphi}$, entonces

$$B(0, 1-\alpha) \subset \overline{f}(D). \tag{4.2.2}$$

Puesto que es inmediato comprobar que

$$\overline{f}(x) = \frac{1}{r}\big(f(a+rx) - f(a)\big), \ \forall x \in D,$$

se sigue que

$$\overline{f}(D) = \frac{1}{r}\big(f(a+rD) - f(a)\big) = \frac{1}{r}\big(f(a+B(0,r)) - f(a)\big) =$$

$$\frac{1}{r}\big(f(B(a,r)) - f(a)\big),$$

y por tanto, teniendo en cuenta la inclusión anterior (4.2.2), se tiene que

$$B(0,1-\alpha) \subset \frac{1}{r}\big(f(B(a,r)) - f(a)\big),$$

lo que implica $f(a) + r\,B(0,1-\alpha) \subset f(B(a,r))$, es decir

$$B(f(a),(1-\alpha)r) \subset f(B(a,r)).$$

Hemos probado que si $B(a,r) \subset U$, entonces $B(f(a),(1-\alpha)r) \subset f(U)$ y, por tanto, $f(U)$ es abierto.

ii) Comencemos comprobando que f es lipschitziana. Para $u,v \in U$ se tiene

$$\|f(u) - f(v)\| = \|u - \varphi(u) - (v - \varphi(v))\| = \|u - v - (\varphi(u) - \varphi(v))\| \leq$$
$$\|u-v\| + \|\varphi(u) - \varphi(v)\| \leq (1+\alpha)\|u-v\|.$$

También para $u,v \in U$ se tiene

$$\|f(u) - f(v)\| = \|u - \varphi(u) - (v - \varphi(v))\| \geq$$
$$\|u-v\| - \|\varphi(u) - \varphi(v)\| \geq (1-\alpha)\|u-v\|,$$

lo que prueba que f es inyectiva y también que $f^{-1} : f(U) \to \mathbb{R}^N$ es lipschitziana:

$$\|f^{-1}(f(u)) - f^{-1}(f(v))\| = \|u-v\| \leq \frac{1}{1-\alpha}\,\|f(u) - f(v)\|,\ \forall u,v \in U.$$

En consecuencia f es un homeomorfismo de U sobre $f(U)$.

iii) Sólo falta probar que f es sobreyectiva, lo que es una consecuencia inmediata de lo probado en i) pues al ser $B(0,r) \subset \mathbb{R}^N$, $\forall r > 0$ se sigue que

$$B(f(0),(1-\alpha)r) \subset f(\mathbb{R}^N),\ \forall r > 0$$

y, por tanto, $f(\mathbb{R}^N) = \mathbb{R}^N$.

■

Notas:

Aunque $ii) \Rightarrow i)$ en virtud del siguiente resultado:

Teorema 4.17 (Brouwer de invarianza del dominio)**.** *Si U es un abierto de $\mathbb{R}^N$ y $f : U \to \mathbb{R}^N$ es un homeomorfismo de U sobre $f(U)$, entonces $f(U)$ es abierto en $\mathbb{R}^N$.*

La demostración de éste, que puede verse en [HoYo, Teorema 6-53], es más complicada que la prueba de i).

De i) y ii) se deduce que la aplicación f es *abierta* (aplica abiertos en abiertos). En efecto, si $O \subset U$ es abierto, entonces $f(O) = (f^{-1})^{-1}(O)$ es un abierto relativo de $f(U)$, luego abierto de $\mathbb{R}^N$.

Por otra parte en iii) se puede probar la sobreyectividad de f directamente del Teorema 4.12. En efecto: dado $y \in \mathbb{R}^N$, consideremos la aplicación auxiliar $g : \mathbb{R}^N \to \mathbb{R}^N$ definida por $g(x) := y + \varphi(x)$, que es también contractiva. Aplicando el Teorema 4.12 a g obtenemos que existe un (único) $x \in \mathbb{R}^N$ tal que $x = g(x)$, es decir, $y = f(x)$.

4.3. Teorema de Picard-Lindelöf.

Teorema 4.18 (Picard-Lindelöf)**.** *Sean a,b números reales con $a<b$ y sea $f:[a,b]\times\mathbb{R}\to\mathbb{R}$ una función continua. Supongamos que además $\exists M\geq 0$ tal que*

$$|f(t,x)-f(t,y)|\leq\ M|x-y|,\ \forall t\in[a,b],\ \forall x,y\in\mathbb{R}$$

(f es uniformemente lipschitziana en segunda variable). Entonces existe una única $\Phi\in\mathscr{C}^1[a,b]$ tal que

$$\Phi(a)=0\ \ y\ \ \Phi'(t)=f(t,\Phi(t)),\ \forall t\in[a,b].$$

Demostración:

Sea $F:\mathscr{C}[a,b]\to\mathscr{C}[a,b]$ el operador dado por

$$F(\varphi)(t):=\int_a^t f(s,\varphi(s))ds,\ \forall\varphi\in\mathscr{C}[a,b],\ \forall t\in[a,b]. \tag{4.3.1}$$

F está bien definido y probaremos por inducción que se verifica para cada $n\in\mathbb{N}$ la siguiente desigualdad:

$$|F^n(\varphi)(t)-F^n(\psi)(t)|\leq\frac{M^n}{n!}(t-a)^n\|\varphi-\psi\|_\infty,\ \forall\varphi,\psi\in\mathscr{C}[a,b],\ \forall t\in[a,b]. \tag{4.3.2}$$

Para $n=1$ se tiene para cada $\varphi,\psi\in\mathscr{C}[a,b]$ y para cada $t\in[a,b]$ que

$$|F(\varphi)(t)-F(\psi)(t)|=\left|\int_a^t f(s,\varphi(s))ds-\int_a^t f(s,\psi(s))ds\right|\leq$$

$$\int_a^t\Big|f(s,\varphi(s))-f(s,\psi(s))\Big|ds\leq\int_a^t M|\varphi(s)-\psi(s)|ds\leq M\|\varphi-\psi\|_\infty(t-a),$$

y la desigualdad 4.3.2 es cierta para $n=1$.

Supongamos dicha desigualdad cierta para n, entonces para cada $\varphi,\psi\in\mathscr{C}[a,b]$ y para cada $t\in[a,b]$ se tiene que

$$|F^{n+1}(\varphi)(t)-F^{n+1}(\psi)(t)|=|F(F^n(\varphi))(t))-F(F^n(\psi))(t))|=$$

$$\left|\int_a^t f(s,F^n(\varphi)(s))ds-\int_a^t f(s,F^n(\psi)(s))ds\right|\leq$$

$$\int_a^t|f(s,F^n(\varphi)(s))-f(s,F^n(\psi)(s))|ds\leq\int_a^t M|F^n(\varphi)(s)-F^n(\psi)(s)|ds$$

donde hemos utilizado la desigualdad de la hipótesis y la monotonía de la integral. Así, en virtud de la hipótesis de inducción, tenemos

$$|F^{n+1}(\varphi)(t)-F^{n+1}(\psi)(t)|\leq\int_a^t M\frac{1}{n!}\,M^n\|\varphi-\psi\|_\infty(s-a)^n ds=$$

$$\frac{M^{n+1}}{n!}\|\varphi-\psi\|_\infty\frac{(t-a)^{n+1}}{n+1}=\frac{M^{n+1}}{(n+1)!}\|\varphi-\psi\|_\infty(t-a)^{n+1},$$

esto es, se verifica la desigualdad (4.3.2) para $n+1$.

Ahora bien, (4.3.2) nos asegura que

$$\|F^n(\varphi)-F^n(\psi)\|_\infty \leq \frac{(M(b-a))^n}{n!}\|\varphi-\psi\|_\infty,\ \forall\varphi,\psi\in\mathscr{C}[a,b],\ \forall n\in\mathbb{N}$$

y teniendo en cuenta que $\left(\mathscr{C}[a,b],\|\cdot\|_\infty\right)$ es un espacio de Banach (véase Ejercicio 2.3), que la sucesión $\left\{\frac{(M(b-a))^n}{n!}\right\}$ converge a cero y el Corolario 4.13 obtenemos que existe una única función $\Phi\in\mathscr{C}[a,b]$ tal que $F(\Phi)=\Phi$. Luego teniendo en cuenta la definición (4.3.1) concluimos que

$$\Phi(t)=\int_a^t f(s,\Phi(s))ds,\ \forall t\in[a,b].$$

Así $\Phi\in\mathscr{C}^1[a,b]$ y se verifica

$$\Phi(a)=0 \ \text{ y } \ \Phi'(t)=f(t,\Phi(t)),\ \forall t\in[a,b],$$

donde se ha utilizado el Teorema fundamental del cálculo para la integral de Riemann. ∎

Para más información sobre esta sección se puede consultar [Guz, Capítulo 4].

4.4. Referencias recomendadas.

[Guz], [MaHo], [HoYo].

4.5. Resumen del resultados del Tema 4

Teorema del valor medio para campos escalares. *Sean $a,b \in \mathbb{R}^N$ verificando que $a \neq b$ y que $[a,b] \subset A \subset \mathbb{R}^N$. Si $f : A \to \mathbb{R}$ es un campo escalar continuo en $[a,b]$ y derivable en $]a,b[$, entonces existe $c \in]a,b[$ tal que*

$$f(b) - f(a) = \big(\nabla f(c)|b-a\big).$$

Proposición. *Sean $\sigma : [0,1] \to \mathbb{R}^M$ y $g : [0,1] \to \mathbb{R}$ funciones continuas en $[0,1]$, derivables en $]0,1[$ y que verifican*

$$\|\sigma'(t)\| \leq g'(t), \ \forall t \in]0,1[.$$

Entonces

$$\|\sigma(1) - \sigma(0)\| \leq g(1) - g(0).$$

Teorema del valor medio. *Sean $a,b \in \mathbb{R}^N$ tal que $a \neq b$ y que $[a,b] \subset A \subset \mathbb{R}^N$. Si $f : A \to \mathbb{R}^M$ es un campo vectorial continuo en $[a,b]$ y derivable en $]a,b[$, entonces*

$$\|f(b) - f(a)\| \leq \|b-a\| \sup\{\|Df(x)\| : x \in]a,b[\}.$$

Teorema práctico del valor medio. *Sean $a,b \in \mathbb{R}^N$ verificando que $a \neq b$ y que $[a,b] \subset A \subset \mathbb{R}^N$. Si $f : A \to \mathbb{R}^M$ es un campo vectorial continuo en $[a,b]$ y derivable en $]a,b[$, entonces*

$$\|f(b) - f(a)\|_2 \leq \|b-a\|_2 \ \sup\left\{\sqrt{\sum_{i,j} D_j f_i(x)^2} : \ x \in]a,b[\right\}.$$

Definición. *Sean (E,d) y (F,ρ) espacios métricos. Una función $f : E \to F$ es contractiva si $\exists \alpha \in [0,1[$ tal que*

$$\rho(f(x), f(y)) \leq \alpha \, d(x,y), \ \forall x,y \in E.$$

Teorema del punto fijo de Banach. *Sea (E,d) un espacio métrico completo y sea $f : E \to E$ una función contractiva. Entonces existe un único punto a de E fijo por f, es decir tal que $f(a) = a$. De hecho, dado a_0 en E, la sucesión $\{a_n\}$ definida mediante la expresión*

$$a_{n+1} := f(a_n), \ \forall n \in \mathbb{N} \cup \{0\},$$

verifica

$$\{d(a,a_n)\} \searrow 0.$$

Además si $0 \leq \alpha < 1$ *es tal que*

$$d(f(x), f(y)) \leq \alpha\, d(x,y),\ \forall x, y \in E,$$

entonces para cada $n \in \mathbb{N}$ *se verifica la desigualdad:*

$$d(a_n, a) \leq \frac{\alpha^n}{1-\alpha}\, d(a_1, a_0).$$

Teorema de Schauder. *Sean* $U \subset \mathbb{R}^N$ *abierto y* $\varphi : U \to \mathbb{R}^N$ *una función contractiva. Si* $f := i_U - \varphi$ *(donde* i_U *es la aplicación inclusión de* U *en* $\mathbb{R}^N$*), entonces se verifica:*

i) $f(U)$ *es un abierto de* $\mathbb{R}^N$.

ii) f *es un homeomorfismo (biyección bicontinua) de* U *sobre* $f(U)$ *(de hecho* f *es bilipschitziana).*

iii) Si además $U = \mathbb{R}^N$*, entonces* f *es un homeomorfismo de* $\mathbb{R}^N$ *sobre sí mismo.*

Teorema de Picard-Lindelöf. *Sean* $a, b \in \mathbb{R}$ *con* $a < b$ *y sea* $f : [a,b] \times \mathbb{R} \to \mathbb{R}$ *una función continua. Supongamos que además* $\exists M \geq 0$ *tal que*

$$|f(t,x) - f(t,y)| \leq\ M|x-y|,\ \forall t \in [a,b],\ \forall x, y \in \mathbb{R}$$

(f *es uniformemente lipschitziana en segunda variable). Entonces existe una única función* $\Phi \in \mathscr{C}^1[a,b]$ *tal que*

$$\Phi(a) = 0\ \text{ y }\ \Phi'(t) = f(t, \Phi(t)),\ \forall t \in [a,b].$$

4.6. Ejercicios del Tema 4

4.1 Sean $A \subset \mathbb{R}^N$ abierto conexo y $f : A \longrightarrow \mathbb{R}^M$ un campo vectorial con derivada constante. Probar que f es la restricción a A de una función afín de $\mathbb{R}^N$ en $\mathbb{R}^M$.
Indicación: Considerar el campo vectorial $g : A \longrightarrow \mathbb{R}^M$ definido por

$$g(x) := f(x) - Df(a)(x) \text{ para conveniente } a \in A.$$

4.2 Probar que la ecuación $x = \operatorname{arc\,tg} x + 1$ tiene una única solución y aproximarla hasta las milésimas.
Indicación: Considerar la función

$$f(x) := \operatorname{arc\,tg} x + 1, \ \forall x \in [1, +\infty[\ .$$

4.3 Sea $A \subset \mathbb{R}^N$ un abierto convexo y sea $f : \overline{A} \longrightarrow \overline{A}$ una aplicación continua. Supongamos que f es derivable en A y que

$$\operatorname{Sup}\left\{ \sum_{i,j} D_i f_j(x)^2 : \ x \in A \right\} < \ 1.$$

Probar que f tiene un único punto fijo en $\overline{A}$.
Indicación: Utilizar el Corolario 4.9.

4.4 Probar que $(0,0)$ es la única solución en el conjunto C del sistema de ecuaciones que se indica, donde C es el conjunto dado por

$$C := \{(x,y) \in \mathbb{R}^2 : |x|, |y| \leq 1\}$$

y el sistema de ecuaciones viene dado por

$$\begin{cases} \operatorname{sen} x \ \operatorname{sen} y = 2x \\ x^2 + y^2 = 4y \end{cases}$$

Indicación: Aplicar el Teorema del punto fijo de Banach (4.12) al campo vectorial $f : C \longrightarrow \mathbb{R}^2$ dado por

$$f(x,y) = \left(\frac{\operatorname{sen} x \ \operatorname{sen} y}{2}, \frac{x^2 + y^2}{4} \right),$$

considerando en $\mathbb{R}^2$ la norma euclídea.

4.5 Sea $f:\mathbb{R}^2 \longrightarrow \mathbb{R}^2$ el campo vectorial definido por

$$f(x,y) = (1+\operatorname{sen} x, \operatorname{arc\,tg} y).$$

Probar que existe un único $p \in \mathbb{R}^2$ tal que $f(p) = 2p$.
Indicación: Ver Ejemplo 4.13.

4.6 Sea $f:\overline{B}(0,1) \subset \mathbb{R}^N \longrightarrow \mathbb{R}^M$ un campo vectorial continuo. Supongamos que $f(0)=0$ y que f es derivable en $B(0,1)$ con $\|Df(x)\| \le \|f(x)\|$, $\forall x \in B(0,1)$. Probar que

$$f(x) = 0,\ \forall x \in \overline{B}(0,1).$$

Indicación: Utilizar el Teorema del valor medio para probar que

$$\|f(x)\| \le \|x\| \ \mathrm{Sup}\ \{\|f(y)\| :\ y \in [0,x]\}.$$

Utilizar ahora la propiedad de compacidad para probar la existencia de $y_1 \in [0,x]$ tal que

$$\|f(x)\| \le \|x\|\ \|f(y_1)\|.$$

Iterar el proceso.

4.7 * Sea $f:A \subset \mathbb{R}^N \longrightarrow \mathbb{R}^N$ un campo vectorial de clase $\mathscr{C}^1$ en $a \in A$. Probar que:

i)

$$\forall \varepsilon > 0,\ \exists \delta > 0 : \begin{cases} B(a,\delta) \subset A \text{ y } f \text{ es derivable en } B(a,\delta) \\ x,y \in B(a,\delta) \Rightarrow \|f(x)-f(y)-Df(a)(x-y)\| \le \varepsilon\|x-y\| \end{cases}$$

ii) Si además, suponemos que $Df(a)$ es un isomorfismo topológico en $\mathbb{R}^N$, pruébese que existe $\rho > 0$ tal que $B(a,\rho) \subset A$ y f es un homeomorfismo de $B(a,\rho)$ sobre su imagen.

4.8 * Sean $A \subset \mathbb{R}^N$ abierto, $f:A \to \mathbb{R}^M$ un campo vectorial de clase $\mathscr{C}^1$ y $K \subset A$ un compacto. Probar que:

$$\forall \varepsilon > 0,\ \exists \delta > 0 : \begin{cases} \{x \in \mathbb{R}^N : \mathrm{dist(x,K)} < \delta\} \subset \mathrm{A} \\ \|h\| < \delta \Rightarrow \|f(x+h)-f(x)-Df(x)(h)\| \le \varepsilon\|h\|,\ \forall x \in K \end{cases}$$

(f es uniformemente derivable en K).

Indicación: Úsese el axioma de Heine-Borel para conseguir la primera condición. Después, el Teorema de Heine, para conseguir continuidad uniforme de la diferencial en un compacto de A que contenga a K. Finalmente, basta usar el Teorema del valor medio como en la parte i) del Ejercicio 4.7.

4.9 * Sea (E,d) un espacio métrico completo y sea $f : E \longrightarrow E$ una aplicación. Supongamos que existe una serie convergente de números reales $\sum_{n\geq 1} \alpha_n$ tal que

$$d\big(f^n(x), f^n(y)\big) \leq \alpha_n\, d(x,y),\ \forall x,y \in E,\ \forall n \in \mathbb{N}.$$

Entonces f tiene un único punto fijo. De hecho si x_0 es cualquier punto de E, la sucesión $\{f^n(x_0)\}$ converge a dicho punto.
Indicación: Utilizar la misma argumentación que en la demostración del Teorema del punto fijo de Banach cambiando la sucesión $\{\alpha^n\}$ que allí aparece por la sucesión $\{\alpha_n\}$.

4.7. Soluciones a los ejercicios del Tema 4.

4.1 Fijemos $a \in A$. El campo vectorial $g : A \longrightarrow \mathbb{R}^M$ dado por

$$g(x) := f(x) - Df(a)(x)$$

es derivable con derivada $Dg(x) = Df(x) - Df(a) = 0$. El Corolario 4.8 nos asegura que g es constante. Así

$$f = b + Df(a)_{|A} \text{ para conveniente } b \in \mathbb{R}^M,$$

es decir,

$$f(x) = f(a) + Df(a)(x-a), \ \forall x \in A.$$

4.2 El espacio métrico $\big([1,+\infty[, |\cdot|\big)$ es completo (cerrado de un completo). Es inmediato que $f([1,+\infty[) \subset ([1,+\infty[)$. Como

$$f'(x) = \frac{1}{1+x^2}, \ \forall x \in ([1,+\infty[,$$

se sigue del Teorema del valor medio real que

$$\big[x,y \in [1,+\infty[\big] \Rightarrow |f(x)-f(y)| = |f'(\theta)||x-y| = \frac{1}{1+\theta^2}|x-y|$$

para conveniente $\theta > 1$ y, por tanto,

$$x,y \in [1,+\infty[\big] \Rightarrow |f(x)-f(y)| \leq \frac{1}{2}\,|x-y|,$$

esto es, f es contractiva. El Teorema del punto fijo nos asegura que existe un único punto $a \in [1,+\infty[$ tal que $a = \operatorname{arc\,tg} a + 1$. Como la función de $\mathbb{R}$ en $\mathbb{R}$ dada por $x \mapsto \operatorname{arc\,tg} x + 1 - x$ tiene derivada negativa, deducimos que la ecuación $x = \operatorname{arc\,tg} x + 1$ tiene una única solución.

El Teorema del punto fijo nos asegura además que para cada natural n es

$$|a - a_n| \leq \frac{|a_1 - a_0|}{2^{n-1}}.$$

En particular si tomamos $a_0 = 1$, obtenemos que $|a - a_n| \leq \frac{\pi}{2^{n+1}}$, que nos permite conocer el punto a con la aproximación deseada, en este caso $a_{11} = 2'13222\ldots$, esto es, $a = 2'132\ldots$.

4.3 Sean $x, y \in A$ con $x \neq y$. Aplicamos el Corolario 4.9 a la función f en el segmento $[x,y]$ para obtener que

$$\|f(y) - f(x)\|_2 \leq \|y - x\|_2 \operatorname{Sup}\left\{\sqrt{\sum_{i,j} D_j f_i(x)^2} : \ x \in A\right\}.$$

Así, notando $\alpha := \text{Sup}\left\{\sqrt{\sum_{i,j} D_j f_i(x)^2} :\ x \in A\right\} < 1$, se tiene que

$$\|f(y) - f(x)\|_2 \leq \alpha \ \|y - x\|_2, \forall x, y \in A.$$

Por continuidad, también

$$\|f(y) - f(x)\|_2 \leq \alpha \ \|y - x\|_2, \forall x, y \in \overline{A}.$$

El Teorema del punto fijo de Banach nos asegura la existencia del punto fijo de f en $\overline{A}$.

4.4 El conjunto $C := \{(x,y) \in \mathbb{R}^2 :\ |x| \leq 1,\ |y| \leq 1\}$ es completo (cerrado de $\mathbb{R}^2$). Es inmediato que la aplicación $f : C \longrightarrow \mathbb{R}^2$ dada por

$$f(x,y) = \left(\frac{\operatorname{sen} x \ \operatorname{sen} y}{2}, \frac{x^2+y^2}{4}\right),$$

verifica $f(C) \subset C$.

f es diferenciable y la matriz jacobiana viene dada por

$$\frac{1}{2}\begin{pmatrix} \cos x \operatorname{sen} y & \operatorname{sen} x \cos y \\ \\ x & y \end{pmatrix}$$

Usando la versión práctica del Teorema del Valor medio (Corolario 4.9) se tiene que

$$\|Df(x,y)\|^2 \leq \frac{1}{4}\left(\cos^2 x \operatorname{sen}^2 y + \operatorname{sen}^2 x \cos^2 y + x^2 + y^2\right) \leq$$

$$\frac{1}{4}\left(\operatorname{sen}^2 y + \operatorname{sen}^2 x + x^2 + y^2\right)$$

$$\leq \frac{1}{4}\left(2\operatorname{sen}^2 1 + x^2 + y^2\right) \leq$$

$$\leq \frac{1}{4}\left(2\operatorname{sen}^2 1 + 2\right) \leq$$

$$\leq \frac{1}{2}\left(\operatorname{sen}^2 1 + 1\right) < 1,$$

por tanto, f es lipschitziana con constante de Lipschitz menor que 1, esto es, contractiva. Como C es completo, por el Teorema del punto fijo de Banach, f tiene un único punto fijo.

4.5 Para $x, y \in \mathbb{R}^2$ se tiene que

$$J_f(x,y) = \begin{pmatrix} \cos x & 0 \\ 0 & \frac{1}{1+y^2} \end{pmatrix},$$

y en consecuencia,

$$\sum_{i,j} D_j f_i(x)^2 = \cos^2 x + \frac{1}{(1+y^2)^2} \leq 2.$$

El Corolario (4.9) nos asegura que f es lipschitziana con constante de Lipschitz menor o igual que $\sqrt{2}$. Así $\frac{f}{2}$ es contractiva. El resultado se sigue del Teorema del punto fijo.

4.6 Sea $x \in B(0,1) \setminus \{0\}$. Se tiene que

$$\|f(x)\| \leq \|x\| \ \mathrm{Sup}\{\|Df(y)\| : \ y \in]0,x[\} \leq$$

$$\|x\| \ \mathrm{Sup}\{\|f(y)\| : \ y \in]0,x[\} \leq \|x\| \ \mathrm{Sup}\{\|f(y)\| : \ y \in [0,x]\},$$

donde se ha utilizado el Teorema del valor medio y la desigualdad de la hipótesis. La propiedad de compacidad nos asegura la existencia de $y_1 \in [0,x]$ tal que

$$\|f(x)\| \leq \|x\| \ \|f(y_1)\|.$$

Repitiendo el proceso, existe $y_2 \in [0,y_1] \subset [0,x]$ tal que $\|f(y_1)\| \leq \|y_1\| \ \|f(y_2)\|$, y en consecuencia

$$\|f(x)\| \leq \|x\|^2 \ \|f(y_2)\|.$$

Se prueba ahora por inducción que para cada n natural se tiene

$$\|f(x)\| \leq \|x\|^n \ \|f(y_n)\| \text{ para conveniente } y_n \in [0,x],$$

y por tanto para cada natural n se tiene que

$$\|f(x)\| \leq \|x\|^n \ \mathrm{Sup}\{\|f(y)\| : \ y \in [0,x]\} = \|x\|^n \ \max\{\|f(y)\| : \ y \in [0,x]\},$$

donde se ha vuelto a utilizar la propiedad de compacidad.

Como $\|x\| < 1$, se sigue que $\{\|x\|^n\} \to 0$ y por tanto $f(x) = 0$. La arbitrariedad de x nos asegura que f se anula en $B(0,1)$. Por último la continuidad de f nos asegura que dicha función se anula también en $\overline{B}(0,1)$.

4.7 *

i) Sea $\varepsilon > 0$ fijo. Tomemos $\delta > 0$ tal que si $\|x-a\| < \delta$, entonces

$$a \in A \quad f \text{ es derivable en } x \text{ y } \ \|Df(x) - Df(a)\| \leq \varepsilon.$$

La función $g : B(a,\delta) \to \mathbb{R}^N$ definida por

$$g(z) := f(z) - Df(a)(z)$$

es derivable con derivada $Dg(z) = Df(z) - Df(a), \forall z \in B(a,\delta)$.

Para cada $x,y \in B(a,\delta)$, el Teorema del valor medio(4.5) aplicado a la función g en el segmento $[x,y]$, nos asegura que

$$\|f(x) - f(y) - Df(a)(x-y)\| = \|g(x) - g(y)\| \leq \|x-y\| \ Sup\{\|Dg(z)\| : z \in]x,y[\}$$

$$= \|x-y\| Sup\{\|Df(z)-Df(a)\| : z \in]x,y[\} \leq \varepsilon \|x-y\|.$$

ii) Sea $\varepsilon > 0$. Tomemos $\delta > 0$ como en i). Sean $x,y \in B(a,\delta)$, se tiene que

$$\|Df(a)(x-y)\| - \|f(x)-f(y)\| \leq \|Df(a)(x-y) - f(x) - f(y)\| \leq \varepsilon \|x-y\|,$$

donde se han usado la propiedad triangular y el apartado i). Así

$$\|Df(a)(x-y)\| - \varepsilon \|x-y\| \leq \|f(x)-f(y)\|,$$

pero al ser $Df(a)$ un isomorfismo, se tiene también que

$$\|x-y\| \leq \|Df(a)^{-1}\| \| \, Df(a)(x-y)\|$$

y por tanto

$$\Big(\frac{1}{\|Df(a)^{-1}\|} \ -\varepsilon\Big)\|x-y\| \leq \|f(x)-f(y)\|.$$

Sea ahora $\varepsilon \in \mathbb{R}^+$ con $\varepsilon < \frac{1}{\|Df(a)^{-1}\|}$. Si notamos ρ a su correspondiente δ, se tiene que

$$\Big(\frac{1}{\|Df(a)^{-1}\|} \ -\varepsilon\Big)\|x-y\| \leq \|f(x)-f(y)\|, \ \forall x,y \in B(a,\rho),$$

expresión que nos da la inyectividad de f y la continuidad de su inversa.

Este ejercicio hace buena la expresión "f hereda localmente las propiedades de $Df(a)$". Conviene resaltar que este ejercicio no es el Teorema de la función inversa que nos asegurará además que f es localmente una función abierta (la demostración de dicho teorema requiere también el uso de otra herramienta fundamental, el Teorema de Schauder (4.15)).

4.8 * Consideramos la función $g : K \longrightarrow \mathbb{R}$ dada por

$$g(x) = \text{dist}(x, \mathbb{R}^N \backslash A) \quad (x \in K).$$

Como A es una abierto que contiene a K, entonces $\mathbb{R}^N \backslash A$ es un cerrado en $\mathbb{R}^N$ cuya intersección con K es vacía. Por tanto, $\text{dist}(x, \mathbb{R}^N \backslash A) > 0$ para todo $x \in K$. Como la función distancia a un conjunto es continua, entonces por la propiedad de compacidad, g alcanza el mínimo en K. Si $r := \min g(K)$, entonces se tiene que

$$x \in \mathbb{R}^N, \text{dist}(x,K) < r \Rightarrow x \notin \mathbb{R}^N \backslash A \Rightarrow x \in A.$$

Sea $\varepsilon > 0$ fijo. Para conseguir la otra condición, aplicamos el Teorema de Heine a la diferencial de f y al compacto

$$K' := \big\{y \in R^N : \text{dist}(y,K) \leq \frac{r}{2}\big\},$$

conjunto que contiene a K. Dado un positivo ε, por la continuidad uniforme de Df en K' existe un positivo $\delta < r$ que verifica

$$x,y\in K', \|x-y\|<\delta \Rightarrow \|Df(x)-Df(y)\|<\varepsilon. \tag{4.7.1}$$

Ahora bien, por la elección de $\delta < \frac{r}{2}$ es inmediato que se verifica

$$y\in\mathbb{R}^N, \operatorname{dist}(y,K)<\delta \Rightarrow y\in A.$$

Además, si $x\in K$, se tiene que un elemento de la forma $x+h$ con $\|h\|<\delta$ verifica que $x+h\in K'$, por tanto en vista de 4.7.1 tenemos

$$\|Df(x)-Df(y)\|<\varepsilon. \tag{4.7.2}$$

La desigualdad del enunciado es consecuencia del Teorema del valor medio. Consideremos la función $g: B(a,\delta_a)\to\mathbb{R}^M$ dada por $g(z)=f(z)-Df(a)(z)$. Dados $x,y\in B(a,\delta_a)$ con $x\neq y$, se tiene

$$\|f(y)-f(x)-Df(a)(x-y)\| = \|g(y)-g(x)\| \leq$$

$$\|y-x\|\ Sup\{\|Dg(z)\|:\ z\in]x,y[\} =$$

$$\|y-x\| Sup\{\|Df(z)-Df(a)\|:\ z\in]x,y[\} \leq \varepsilon\ \|y-x\|,$$

donde se ha vuelto a utilizar la desigualdad de (4.7.2).

4.9 * Aunque la existencia y unicidad del punto fijo de f se puede deducir del Corolario (4.13), haremos una demostración independiente con el fin de que el ejercicio generalice el Teorema del punto fijo (tómese $\alpha_n=\alpha^n$, $\forall n\in\mathbb{N}$).
<u>Unicidad:</u> Supongamos que $a,b\in E$ son tales que $f(a)=a$ y $f(b)=b$, se tiene para cada n natural que $d(a,b)=d(f^n(a),f^n(b))\leq\alpha_n\ d(a,b)$ de donde se deduce, por converger a cero la sucesión $\{\alpha_n\}$, que $a=b$.
<u>Existencia:</u> La sucesión $\{x_n\}=\{f^n(x_0)\}$ es de Cauchy. En efecto para $n,h\in\mathbb{N}$, se tiene

$$d(x_{n+h},x_n)\leq d(x_{n+h},x_{n+h-1})+\cdots+d(x_{n+1},x_n)=$$

$$d\big(f^{n+h-1}(x_1),f^{n+h-1}(x_0)\big)+\cdots+d\big(f^n(x_1),f^n(x_0)\big)\leq$$

$$\alpha_{n+h-1}\ d(x_1,x_0)+\cdots+\alpha_n\ d(x_1,x_0)\leq d(x_1,x_0)\ \sum_{k=n}^{\infty}\alpha_k.$$

Como la serie $\sum_{n\geq1}\alpha_n$ es convergente, la sucesión $\left\{\sum_{k=n}^{\infty}\alpha_k\right\}$ converge a cero y se puede concluir que $\{x_n\}$ es una sucesión de Cauchy, y por tanto convergente.

Sea $a:=\lim\ x_n$, se tiene entonces que

$$d(f(a),a)\leq d(f(a),x_{n+1})+d(x_{n+1},a)=d(f(a),f(x_n))+d(x_{n+1},a)\leq$$

$$\alpha_1 d(a,x_n)+d(x_{n+1},a)\to 0,$$

de donde se deduce que $f(a)=a$.

Tema 5

Derivada segunda. Matriz hessiana.

Recordemos que la derivabilidad de un campo vectorial f en un punto a garantiza la existencia de una mejor aproximación afín de la función f en dicho punto, esto es, la existencia de una aplicación lineal T de $\mathbb{R}^N$ en $\mathbb{R}^M$ tal que

$$\lim_{x\to a} \frac{f(x)-\big(f(a)+T(x-a)\big)}{\|x-a\|} = 0 \,.$$

Para una función f de variable real, si f' es de nuevo derivable, es conocido que se verifica

$$\lim_{x\to a} \frac{f(x)-\big[f(a)+f'(a)(x-a)+\frac{1}{2}f''(a)(x-a)^2\big]}{|x-a|^2} = 0 \,.$$

Ahora bien, si $f : A \subset \mathbb{R}^N \longrightarrow \mathbb{R}^M$ es un campo vectorial y la aplicación Df es derivable en un punto a, probaremos que existe una aplicación lineal S de $\mathbb{R}^N$ en $\mathscr{L}(\mathbb{R}^N,\mathbb{R}^M)$ tal que

$$\lim_{x\to a} \frac{f(x)-\big[f(a)+T(x-a)+\frac{1}{2}S(x-a)(x-a)\big]}{\|x-a\|^2} = 0 \,,$$

donde $T = Df(a)$ y a la aplicación S la llamaremos diferencial segunda de f en a, ($D^2f(a)$). Este resultado se conoce como Fórmula infinitesimal del resto (véase el Teorema 5.8).

También mejoraremos la información que obtuvimos en el Teorema del valor medio, esto es, la desigualdad

$$\|f(x)-f(a)\| \le \|x-a\| \sup\{\|Df(z)\| : z \in]a,x[\}.$$

Si el campo vectorial f es dos veces derivable, obtendremos el siguiente resultado que se conoce como Fórmula de Taylor (véase el Teorema 5.10), esto es,

$$\|f(x)-[f(a)+Df(a)(x-a)]\| \le \|x-a\|^2 \sup\{\|D^2f(z)\| : z \in]a,x[\}.$$

Los resultados conocidos y los que obtendremos en este tema se resumen en el siguiente y sugestivo esquema:

$$\left\{\begin{array}{ccc} \text{Definición de derivada} & \longrightarrow & \text{Fórmula infinitesimal del resto} \\ \text{Teorema del valor medio} & \longrightarrow & \text{Fórmula de Taylor} \end{array}\right\}.$$

Como la derivada segunda es un operador lineal con valores en el espacio de las aplicaciones lineales, empezaremos identificando este espacio de operadores con las aplicaciones bilineales.

Bajo esta identificación, probaremos que la diferencial segunda es simétrica, resultado que se conoce come Teorema de Schwarz (Teorema 5.6).

5.1. Aplicaciones bilineales.

Para cada $T \in \mathscr{L}(\mathbb{R}^N, \mathscr{L}(\mathbb{R}^M, \mathbb{R}^P))$, si definimos

$$\Phi(x,y) = T(x)(y) \qquad (x \in \mathbb{R}^N, y \in \mathbb{R}^M) \tag{$*$}$$

es fácil comprobar que $\Phi : \mathbb{R}^N \times \mathbb{R}^M \longrightarrow \mathbb{R}^P$ es una aplicación bilineal, esto es, Φ es lineal en cada variable. Recíprocamente, si $\Phi : \mathbb{R}^N \times \mathbb{R}^M \longrightarrow \mathbb{R}^P$ es una aplicación bilineal, la igualdad $(*)$ define una aplicación lineal $T : \mathbb{R}^N \longrightarrow \mathscr{L}(\mathbb{R}^M, \mathbb{R}^P)$. Como el espacio $\mathscr{L}(\mathbb{R}^N, \mathscr{L}(\mathbb{R}^M, \mathbb{R}^P))$ es un espacio normado si se considera la norma de operadores, y la anterior identificación entre operadores y formas bilineales es lineal, entonces la norma del espacio de operadores induce otra en el espacio $\mathscr{L}^2(\mathbb{R}^N \times \mathbb{R}^M, \mathbb{R}^P)$ de las aplicaciones bilineales de $\mathbb{R}^N \times \mathbb{R}^M$ en $\mathbb{R}^P$. Para $T \in \mathscr{L}(\mathbb{R}^N, \mathscr{L}(\mathbb{R}^M, \mathbb{R}^P))$ se tiene entonces

$$\|T\| = \max\left\{\|T(x)\| : x \in \mathbb{R}^N, \|x\| = 1\right\} =$$

$$= \max\left\{\|T(x)(y)\| : x \in \mathbb{R}^N, y \in \mathbb{R}^M, \|x\| = 1, \|y\| = 1\right\},$$

escribiendo la igualdad anterior en términos de Φ, la aplicación bilineal asociada a T, se tiene que

$$\|\Phi\| = \max\left\{\|\Phi(x,y)\| : \ x \in \mathbb{R}^N, y \in \mathbb{R}^M, \|x\| = 1, \|y\| = 1\right\}.$$

Si $N = M$, notaremos simplemente por $\mathscr{L}^2(\mathbb{R}^N, \mathbb{R}^P)$ al conjunto de las aplicaciones bilineales de $\mathbb{R}^N \times \mathbb{R}^N$ en $\mathbb{R}^P$. Cuando $P = 1$ y $N = M$, las aplicaciones bilineales de $\mathbb{R}^N \times \mathbb{R}^N$ en $\mathbb{R}$ se identifican con las matrices cuadradas de orden N de números reales. Si $A \in \mathscr{M}_{N \times N}(\mathbb{R})$, la forma bilineal asociada Φ_A viene dada por

$$\Phi_A(x,y) = T_A(x)(y) = \left(xA^t \mid y\right) = xA^t y^t \qquad (x, y \in \mathbb{R}^N),$$

donde estamos usando las identificaciones

$$\mathscr{L}^2(\mathbb{R}^N, \mathbb{R}) \equiv \mathscr{L}(\mathbb{R}^N, \mathscr{L}(\mathbb{R}^N, \mathbb{R})) \equiv \mathscr{L}(\mathbb{R}^N, \mathbb{R}^N) \equiv \mathscr{M}_{N \times N}.$$

Si suponemos que $A = \left(a_{ij}\right)$, como

$$\Phi_A(e_i, e_j) = \left(e_i A^t \mid e_j\right) = \left((a_{ni}) \mid e_j\right) = a_{ji},$$

entonces, es claro que la forma bilineal Φ_A es simétrica si, y sólo si, la matriz A lo es, en cuyo caso

$$\Phi_A(x,y) = xAy^t.$$

Proposición 5.1.

i) La aplicación $\| \ \| : \mathscr{L}^2(\mathbb{R}^N \times \mathbb{R}^M, \mathbb{R}^P) \longrightarrow \mathbb{R}$ *dada por*

$$\|\Phi\| = \max\left\{\|\Phi(x,y)\| : x \in \mathbb{R}^N, y \in \mathbb{R}^M, \|x\| = 1, \|y\| = 1\right\}$$

es una norma en $\mathscr{L}^2(\mathbb{R}^N \times \mathbb{R}^M, \mathbb{R}^P)$.

ii) Se verifica

$$\|\Phi\| = \max\left\{\|\Phi(x,y)\| : x \in \mathbb{R}^N, y \in \mathbb{R}^M, \|x\| \le 1, \|y\| \le 1\right\} =$$

$$= \min\left\{K \ge 0 : \|\Phi(x,y)\| \le K\|x\| \ \|y\|, \ \ \forall x \in \mathbb{R}^N, \forall y \in \mathbb{R}^M\right\}$$

iii) Φ *es continua.*

Demostración:
i) y ii) Basta usar que la identificación de $\mathscr{L}^2(\mathbb{R}^N \times \mathbb{R}^M, \mathbb{R}^P)$ en $\mathscr{L}(\mathbb{R}^N, \mathscr{L}(\mathbb{R}^M, \mathbb{R}^P))$ es una isometría y que las igualdades análogas se verifican en $\mathscr{L}(\mathbb{R}^N, \mathscr{L}(\mathbb{R}^M, \mathbb{R}^P))$.
iii) Como consecuencia de la igualdad

$$\Phi(x,y) - \Phi(a,b) = \Phi(x-a, y-b) + \Phi(x-a, b) + \Phi(a, y-b),$$

y en vista de ii) se tiene

$$\|\Phi(x,y) - \Phi(a,b)\| \le \|\Phi\| \ \Big(\|x-a\| \ \|y-b\| + \|x-a\| \ \|b\| + \|a\| \ \|y-b\|\Big),$$

de donde se deduce la continuidad de Φ. ∎

5.2. Derivada segunda.

Recordemos que si $A \subset \mathbb{R}$ y $f : A \longrightarrow \mathbb{R}$ es derivable, se dice que f es dos veces derivable si la función $f' : A \longrightarrow \mathbb{R}$ es derivable. Para campos vectoriales, se puede trasladar la misma definición sin más que tener en cuenta que la aplicación derivada toma valores en $\mathscr{L}(\mathbb{R}^N, \mathbb{R}^M)$, esto es, en el espacio vectorial de las matrices de orden $M \times N$ que se identifica con $\mathbb{R}^{M\times N}$, en el que, en virtud del Teorema de Hausdorff, todas las normas son equivalentes. Usualmente en $\mathscr{L}(\mathbb{R}^N, \mathbb{R}^M)$ consideraremos la norma de operadores.

Definición 5.2 (Campos vectoriales dos veces derivables)**.** Sean $A \subset \mathbb{R}^N$, $f : A \longrightarrow \mathbb{R}^M$ y sea $A_1 \subset A$ el conjunto de puntos donde f es derivable. Se dice que f es <u>*dos veces derivable*</u> en $a \in \overset{\circ}{A_1}$ si la función $Df : A_1 \longrightarrow \mathscr{L}(\mathbb{R}^N, \mathbb{R}^M) \equiv \mathbb{R}^{M\times N}$ es derivable en a, en cuyo caso la aplicación

$$D(Df)(a) \in \mathscr{L}(\mathbb{R}^N, \mathscr{L}(\mathbb{R}^N, \mathbb{R}^M)) \equiv \mathscr{L}^2(\mathbb{R}^N, \mathbb{R}^M)$$

se denomina la <u>*derivada segunda*</u> de f en a o <u>*diferencial segunda*</u> de f en a y se nota $D^2 f(a)$. La aplicación cuadrática de $\mathbb{R}^N$ en $\mathbb{R}^M$ asociada a $D^2 f(a)$ se nota $d^2 f(a)$. Esto es,

$$d^2 f(a)(x) := D^2 f(a)(x,x), \quad \forall x \in \mathbb{R}^N.$$

Diremos que f es <u>*dos veces derivable en un subconjunto B de A*</u> si es dos veces derivable en cada punto de B.

Sea $A_2 \subset A$ el conjunto de puntos de A donde f es dos veces derivable. La aplicación $D^2 f : A_2 \longrightarrow \mathscr{L}(\mathbb{R}^N, \mathscr{L}(\mathbb{R}^N, \mathbb{R}^M)) \equiv \mathscr{L}^2(\mathbb{R}^N, \mathbb{R}^M)$ definida por

$$x \to D^2 f(x)$$

se llama <u>*derivada segunda*</u> de f o diferencial segunda de f. Se dice que f es de <u>*clase $\mathscr{C}^2$ en a*</u> si es dos veces derivable en un entorno de a y la función $D^2 f$ es continua en a.

Se dice que f es <u>*dos veces derivable*</u> (resp. <u>*de clase $\mathscr{C}^2$*</u>) en A cuando lo sea en todos los puntos de su conjunto de definición (que necesariamente será abierto). Notamos por $\mathscr{C}^2(A)$ al conjunto de las funciones de clase $\mathscr{C}^2$ en el abierto A.

Nota 5.3. Sean $A \subset \mathbb{R}, f : A \longrightarrow \mathbb{R}^M$. Si f es derivable en un punto $a \in \overset{\circ}{A}$, sabemos que

$$Df(a)(x) = xf'(a), \quad \forall x \in \mathbb{R},$$

equivalentemente,

$$f'(a) = Df(a)(1).$$

Supongamos ahora que f es derivable en un entorno U de a. Las siguientes afirmaciones son equivalentes:

i) f es dos veces derivable en a (en el sentido de la definición anterior).

ii) f' es derivable en a.

Además, si son ciertas las afirmaciones anteriores, entonces se verifica

$$D^2 f(a)(s,t) = st f''(a), \quad \forall s,t \in \mathbb{R},$$

equivalentemente,

$$D^2 f(a)(1,1) = f''(a).$$

Demostración:

En efecto, sabemos que la aplicación $E: \mathscr{L}(\mathbb{R},\mathbb{R}^M) \longrightarrow \mathbb{R}^M$ definida por

$$E(T) = T(1), \quad \forall T \in \mathscr{L}(\mathbb{R},\mathbb{R}^M)$$

es un isomorfismo de $\mathscr{L}(\mathbb{R},\mathbb{R}^M)$ sobre $\mathbb{R}^M$ y la relación entre f' y Df viene determinada por este isomorfismo. Por tanto, la composición

$$U \xrightarrow{Df} \mathscr{L}(\mathbb{R},\mathbb{R}^M) \xrightarrow{E} \mathbb{R}^M$$

no es más que f'. Como la segunda función es derivable por ser lineal y su inversa también, entonces, gracias a la regla de la cadena, Df es derivable en a si, y sólo si, f admite segunda derivada en a. Además, si esto ocurre, usando la regla de la cadena, y teniendo en cuenta que, por ser E lineal $DE(T) = E, \ \forall T \in \mathscr{L}(\mathbb{R},\mathbb{R}^M)$, obtenemos

$$f''(a) = D(f')(a)(1) = D(E \circ Df)(a)(1) =$$

$$= \Big(DE(Df(a)) \circ D(Df)(a)\Big)(1) =$$

$$= E\Big(D(Df)(a)(1)\Big) = \Big(D(Df)(a)(1)\Big)(1) = D^2 f(a)(1,1).$$

Como la aplicación $(s,t) \mapsto D^2 f(a)(s,t)$ es bilineal, entonces

$$D^2 f(a)(s,t) = st D^2 f(a)(1,1) = st f''(a) \ \ (s,t \in \mathbb{R}).$$

■

Ejemplos 5.4 (Funciones dos veces derivables)**.**

1. Toda función constante $f: \mathbb{R}^N \longrightarrow \mathbb{R}^M$ es de clase $\mathscr{C}^2$ con

$$D^2 f(a) = 0, \quad \forall a \in \mathbb{R}^N.$$

2. Toda aplicación lineal $T: \mathbb{R}^N \longrightarrow \mathbb{R}^M$ es de clase $\mathscr{C}^2$ con

$$D^2 T(a) = 0, \quad \forall a \in \mathbb{R}^N.$$

3. Toda aplicación bilineal $T: \mathbb{R}^N \times \mathbb{R}^M \longrightarrow \mathbb{R}^P$ es de clase $\mathscr{C}^2$ con derivada segunda en $(a,b) \in \mathbb{R}^N \times \mathbb{R}^M$ definida por

$$D^2 T(a,b)\Big((x_1,y_1),(x_2,y_2)\Big) = T(x_1,y_2) + T(x_2,y_1), \quad \forall (x_1,y_1),(x_2,y_2) \in \mathbb{R}^N \times \mathbb{R}^M.$$

En efecto, como

$$DT(a,b)(x,y) = T(a,y) + T(x,b), \quad \forall (a,b),(x,y) \in \mathbb{R}^N \times \mathbb{R}^M,$$

DT es una aplicación lineal de $\mathbb{R}^N \times \mathbb{R}^M$ en $\mathscr{L}(\mathbb{R}^N \times \mathbb{R}^M, \mathbb{R}^P)$, luego DT es de clase $\mathscr{C}^1$, es decir, T es de clase $\mathscr{C}^2$. Calculemos ahora la derivada segunda de T en $(a,b) \in \mathbb{R}^N \times \mathbb{R}^M$. Se tiene

$$D^2T(a,b)\big((x_1,y_1),(x_2,y_2)\big) = \Big[D(DT)(a,b)(x_1,y_1)\Big]\big(x_2,y_2\big) =$$

$$\big(DT(x_1,y_1)\big)(x_2,y_2) = T(x_1,y_2) + T(x_2,y_1), \quad \forall (x_1,y_1),(x_2,y_2) \in \mathbb{R}^N \times \mathbb{R}^M.$$

Nótese la simetría de $D^2T(a,b)$.

4. En vista de los ejemplos 2 y 3, la aplicación suma $\sigma : \mathbb{R}^N \times \mathbb{R}^N \longrightarrow \mathbb{R}^N$ dada por

$$\sigma(a,b) = a+b, \quad \forall a,b \in \mathbb{R}^N$$

y la aplicación producto por escalares $\pi : \mathbb{R} \times \mathbb{R}^N \longrightarrow \mathbb{R}^N$ dada por

$$\pi(\alpha,a) = \alpha a, \quad \forall \alpha \in \mathbb{R},\ a \in \mathbb{R}^N$$

son de clase $\mathscr{C}^2$ con

$$D^2\sigma(a,b) = 0, \quad \forall a,b \in \mathbb{R}^N$$

y

$$D^2\pi(\alpha,a)\Big((\lambda_1,x_1),(\lambda_2,x_2)\Big) = \lambda_1 x_2 + \lambda_2 x_1, \quad \forall \alpha,\lambda_1,\lambda_2 \in \mathbb{R},\ a,x_1,x_2 \in \mathbb{R}^N.$$

5. La aplicación inversión $J : \text{Iso}\ (\mathbb{R}^N) \longrightarrow \mathscr{L}(\mathbb{R}^N)$ dada por

$$J(T) = T^{-1}, \quad \forall T \in \text{Iso}\ (\mathbb{R}^N)$$

es de clase $\mathscr{C}^2$ con derivada segunda en $T \in \text{Iso}\ (\mathbb{R}^N)$ definida por

$$D^2J(T)(R,S) = T^{-1}ST^{-1}RT^{-1} + T^{-1}RT^{-1}ST^{-1}, \qquad R,S \in \mathscr{L}(\mathbb{R}^N).$$

En efecto, sabemos que J es derivable con derivada

$$DJ = \Phi \circ (J,J)$$

de Iso $(\mathbb{R}^N)$ en $\mathscr{L}(\mathscr{L}(\mathbb{R}^N))$, donde Φ es la aplicación bilineal de $\mathscr{L}(\mathbb{R}^N) \times \mathscr{L}(\mathbb{R}^N)$ en $\mathscr{L}(\mathscr{L}(\mathbb{R}^N))$ definida por

$$\Phi(F,G)(S) = -FSG, \quad \forall F,G,S \in \mathscr{L}(\mathbb{R}^N).$$

Las reglas de derivación y los ejemplos de funciones derivables garantizan que DJ es de clase $\mathscr{C}^1$, y en consecuencia J es de clase $\mathscr{C}^2$, con derivada segunda dada por

$$\begin{gathered}
D^2J(T)(R,S) = D(DJ)(T)(R)(S) = D(\Phi \circ (J,J))(T)(R)(S) = \\
\Big[D\Phi(T^{-1},T^{-1}) \circ D(J,J)(T)\Big](R)(S) = \\
D\Phi(T^{-1},T^{-1})(DJ(T)(R),DJ(T)(R))(S) = \\
D\Phi(T^{-1},T^{-1})(-T^{-1}RT^{-1},-T^{-1}RT^{-1})(S) = \\
\Big(\Phi(T^{-1},-T^{-1}RT^{-1}) + \Phi(-T^{-1}RT^{-1},T^{-1})\Big)(S) = \\
-T^{-1}S(-T^{-1}RT^{-1}) - (-T^{-1}RT^{-1})ST^{-1} = \\
= T^{-1}ST^{-1}RT^{-1} + T^{-1}RT^{-1}ST^{-1}, \quad \forall R,S \in \mathscr{L}(\mathbb{R}^N).
\end{gathered}$$

5.3. Reglas de derivación.

i) **Linealidad.** Sean $A \subset \mathbb{R}^N$, $a \in A$ y $f, g : A \longrightarrow \mathbb{R}^M$ funciones dos veces derivables en a y sea $\lambda \in \mathbb{R}$. Entonces $f+g$ y λf son dos veces derivables en a con

$$D^2(f+g)(a) = D^2 f(a) + D^2 g(a), \quad D^2(\lambda f)(a) = \lambda D^2 f(a).$$

Además, si $f, g \in \mathscr{C}^2(a)$, entonces $f+g \in \mathscr{C}^2(a)$ y $\lambda f \in \mathscr{C}^2(a)$.
Demostración:

La comprobación de esta regla es rutinaria.

ii) **Regla de la cadena.** Sean $A \subset \mathbb{R}^N, B \subset \mathbb{R}^M, f : A \longrightarrow \mathbb{R}^M$ tal que $f(A) \subset B$ y $g : B \longrightarrow \mathbb{R}^P$. Supongamos que f es dos veces derivable en $a \in A$ y que g es dos veces derivable en $f(a)$. Entonces la composición $h = g \circ f$ es dos veces derivable en a con

$$D^2 h(a)(x_1, x_2) =$$

$$= D^2 g(f(a))\Big(Df(a)(x_1), Df(a)(x_2)\Big) + Dg(f(a))\Big(D^2 f(a)(x_1, x_2)\Big)$$

para cualesquiera $x_1, x_2 \in \mathbb{R}^N$.

Además si $f \in \mathscr{C}^2(a)$ y $g \in \mathscr{C}^2(f(a))$, entonces $h \in \mathscr{C}^2(a)$.
Demostración:

Teniendo en cuenta las hipótesis, podemos tomar U y V entornos de a y $f(a)$, respectivamente, tales que f es derivable en U, g es derivable en V, y $f(U) \subset V$. La regla de la cadena para la derivada primera garantiza que la función h es derivable en U, y además que

$$Dh(x) = Dg(f(x)) \circ Df(x), \quad \forall x \in U.$$

Por tanto

$$Dh = \Psi \circ \Phi,$$

donde Φ es la aplicación de U en $\mathscr{L}(\mathbb{R}^M, \mathbb{R}^P) \times \mathscr{L}(\mathbb{R}^N, \mathbb{R}^M)$ dada por

$$\Phi(x) = \Big((Dg \circ f)(x), Df(x)\Big), \quad \forall x \in U,$$

y Ψ es la aplicación de $\mathscr{L}(\mathbb{R}^M, \mathbb{R}^P) \times \mathscr{L}(\mathbb{R}^N, \mathbb{R}^M)$ en $\mathscr{L}(\mathbb{R}^N, \mathbb{R}^P)$ definida por

$$\Psi(R, S) = R \circ S.$$

Como g es dos veces derivable (resp. de clase $\mathscr{C}^2$) en $f(a)$ podemos asegurar que Dg es derivable (resp. de clase $\mathscr{C}^1$) en $f(a)$. En consecuencia, la regla de la cadena para la derivada primera nos dice que $Dg \circ f$ es derivable (resp. de clase $\mathscr{C}^1$) en a, luego la función componente primera de Φ es derivable (resp. de clase $\mathscr{C}^1$) en a. Como f es dos veces derivable (resp. de clase $\mathscr{C}^2$) en a, la función componente segunda de Φ , esto es, Df, es derivable (resp. de clase $\mathscr{C}^1$) en a. En consecuencia, Φ es derivable (resp. de clase $\mathscr{C}^1$) en a.

Por otra parte, la función Ψ es una aplicación bilineal y, por tanto, de clase $\mathscr{C}^1$ en $\mathscr{L}(\mathbb{R}^M,\mathbb{R}^P)\times\mathscr{L}(\mathbb{R}^N,\mathbb{R}^P)$.

Al ser $Dh=\Psi\circ\Phi$, el resultado es consecuencia de la regla de la cadena para la derivada primera.

Probemos ahora la fórmula. Para $x_1,x_2\in\mathbb{R}^N$, se tiene que

$$D^2h(a)(x_1,x_2)=D(Dh)(a)(x_1)(x_2)=D(\Psi\circ\Phi)(a)(x_1)(x_2)=$$

$$\Big(D\Psi(\Phi(a))\circ D\Phi(a)\Big)(x_1)(x_2)=D\Psi(\Phi(a))\Big(D\Phi(a)(x_1)\Big)(x_2)=$$

$$D\Psi\Big(Dg(f(a)),Df(a)\Big)\Big(D(Dg\circ f)(a)(x_1),D(Df)(a)(x_1)\Big)(x_2)=$$

$$D\Psi\Big(Dg(f(a)),Df(a)\Big)\Big(D(Dg)(f(a))\big(Df(a)(x_1)\big),D(Df)(a)(x_1)\Big)(x_2)=$$

$$Dg(f(a))\Big(D(Df)(a)(x_1)(x_2)\Big)+D(Dg)(f(a))\Big(Df(a)(x_1)\Big)\Big(Df(a)(x_2)\Big)=$$

$$Dg(f(a))\Big(D^2f(a)(x_1,x_2)\Big)+D^2g(f(a))\Big(Df(a)(x_1),Df(a)(x_2)\Big).$$

∎

Nota 5.5. Es interesante observar la forma que adopta la regla de la cadena cuando se involucran derivadas elementales.

En el caso de que $A\subset\mathbb{R}\xrightarrow{f}B\subset\mathbb{R}^M\xrightarrow{g}\mathbb{R}^P$ se tiene que

$$(g\circ f)''(a)=D^2(g\circ f)(a))(1,1)=D^2g(f(a))\,(Df(a)(1),Df(a)(1))+$$

$$Dg(f(a))\,\big(D^2f(a)(1,1)\big)=D^2g(f(a))\,\big(f'(a),f'(a)\big)+Dg(f(a))\,\big(f''(a)\big),$$

donde se han utilizado la Nota 5.3 y la regla de la cadena.

En el caso especial de que $M=1$ se obtiene

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

El carácter local de la derivabilidad y de la derivada permiten ahora obtener la siguiente regla.

iii) **Carácter local de la derivada segunda.** Sean $A\subset\mathbb{R}^N$, $a\in A$ y $f:A\longrightarrow\mathbb{R}^M$ un campo vectorial. Las siguientes afirmaciones son equivalentes:

i) f es dos veces derivable en a.

ii) Existe algún entorno U de a incluido en A tal que $f_{|U}$ es dos veces derivable en a.

Además, en caso de que sean ciertas las anteriores afirmaciones, las derivadas segundas en a de ambas funciones coinciden.

iv) **Derivación de la función inversa.** Sean A y B subconjuntos abiertos de $\mathbb{R}^N$ y f un homeomorfismo de A sobre B. Si f es dos veces derivable en a (resp. de clase $\mathscr{C}^2$ en a) para algún $a \in A$ y $Df(a) \in \text{Iso}\,(\mathbb{R}^N)$, entonces f^{-1} es dos veces derivable en $f(a)$ (resp. de clase $\mathscr{C}^2$ en $f(a)$).
Demostración:

La continuidad de la aplicación $x \mapsto Df(x)$ en el punto a garantiza la existencia de un abierto U de $\mathbb{R}^N$ tal que

$$a \in U \subset A,\ \ Df(x) \in \text{Iso}\,(\mathbb{R}^N), \quad \forall x \in U.$$

La regla de derivación de la función inversa (Proposición 3.27) asegura entonces que f^{-1} es derivable en $f(U)$ con

$$Df^{-1}(y) = \Big(Df(f^{-1}(y))\Big)^{-1}, \quad \forall y \in f(U).$$

En consecuencia,

$$Df^{-1} = J \circ Df \circ f^{-1},$$

lo que prueba que Df^{-1} es derivable en $f(a)$, es decir f^{-1} es dos veces derivable en $f(a)$. Es claro también que si f es de clase $\mathscr{C}^2$ en a, entonces Df^{-1} es de clase $\mathscr{C}^1$ en $f(a)$, es decir, f^{-1} es de clase $\mathscr{C}^2$ en $f(a)$. ■

v) Sea $A \subset \mathbb{R}^N$ y $f : A \longrightarrow \mathbb{R}^M$ una función dos veces derivable en $a \in A$ y $T \in \mathscr{L}(\mathbb{R}^M, \mathbb{R}^P)$, entonces $T \circ f$ es dos veces derivable en a con

$$D^2(T \circ f)(a) = T \circ D^2 f(a).$$

Como consecuencia, si $f \in \mathscr{C}^2(a)$, entonces $T \circ f \in \mathscr{C}^2(a)$.
Demostración:

Como T es de clase $\mathscr{C}^2$, la regla de la cadena garantiza que $T \circ f$ es dos veces derivable (resp. de clase $\mathscr{C}^2$) en a con

$$\begin{gathered} D^2(T \circ f)(a)(x_1, x_2) = \\ D^2T(f(a))\Big(Df(a)(x_1), Df(a)(x_2)\Big) + DT(f(a))\Big(D^2 f(a)(x_1, x_2)\Big) = \\ 0 + DT(f(a))\Big(D^2 f(a)(x_1, x_2)\Big) = \\ T\Big(D^2 f(a))(x_1, x_2)\Big) = \\ \Big(T \circ D^2 f(a)\Big)(x_1, x_2), \end{gathered}$$

donde se ha empleado que $DT(y) = T$, $\forall y \in \mathbb{R}^M$, y, por tanto, $D^2T = 0$. ■

vi) **Reducción a campos escalares.** Sean $A \subset \mathbb{R}^N$ y $f : A \longrightarrow \mathbb{R}^M$ un campo vectorial. Si $f = (f_1, \ldots, f_M)$, entonces f es dos veces derivable en $a \in A$ si, y sólo si, cada f_k es dos veces derivable en a para $k = 1, \ldots, M$, en cuyo caso

$$D^2 f(a)(x_1, x_2) = \Big(D^2 f_1(a)(x_1, x_2), \ldots, D^2 f_M(a)(x_1, x_2)\Big)$$

Además $f \in \mathscr{C}^2(a)$ si, y sólo si, $f_k \in \mathscr{C}^2(a)$ para $k = 1, \ldots, M$.
Demostración:

Para $k = 1, \ldots, M$, notaremos por π_k a la proyección k-ésima de $\mathbb{R}^M$ y por I_k a la inyección k-ésima de $\mathbb{R}$ en $\mathbb{R}^M$. Como

$$f_1 = \pi_1 \circ f, \ldots, f_M = \pi_M \circ f \quad \text{y} \quad f = I_1 \circ f_1 + \ldots + I_M \circ f_M,$$

la linealidad de la derivada segunda y la regla de la cadena aseguran que f es dos veces derivable en a si, y sólo si las funciones componentes lo son. El mismo argumento es válido en caso de que $f \in \mathscr{C}^2$. Además, el anterior apartado asegura que

$$D^2 f_k(a) = \pi_k \circ D^2 f(a), \quad k = 1, \ldots, M$$

y

$$D^2 f(a) = I_1 \circ D^2 f_1(a) + \ldots + I_M \circ D^2 f_M(a),$$

es decir,

$$D^2 f(a)(x_1, x_2) = \Big(D^2 f_1(a)(x_1, x_2), \ldots, D^2 f_M(a)(x_1, x_2)\Big), \quad \forall x_1, x_2 \in \mathbb{R}^N.$$

■

5.4. Teorema de Schwarz.

Teorema 5.6 (Schwarz)**.** *Sean $A \subset \mathbb{R}^N$ y $f : A \to \mathbb{R}^M$ un campo vectorial. Supongamos que f es dos veces derivable en un punto $a \in A$. Entonces la aplicación bilineal $D^2 f(a)$ es simétrica, esto es,*

$$D^2 f(a)(x_1, x_2) = D^2 f(a)(x_2, x_1), \quad \forall x_1, x_2 \in \mathbb{R}^N.$$

Demostración:

Dado $\varepsilon > 0$, elegimos $\delta > 0$ tal que si $x \in B(a, 2\delta)$ entonces se verifica

(5.4.1) $\quad x \in A, \ f$ es derivable en x y $\|Df(x) - Df(a) - D(Df)(a)(x-a)\| \leq \dfrac{\varepsilon}{2}\|x - a\|$

Sean $h, k \in \mathbb{R}^N$ con $0 < \|h\|, \|k\| < \delta$ fijos. Consideremos la función

$$\varphi : B(0, \delta) \subset \mathbb{R}^N \longrightarrow \mathbb{R}^M$$

definida por

$$\varphi(x) = f(a + h + x) - f(a + x) - D(Df)(a)(h)(x), \quad (x \in B(0, \delta)).$$

Es claro que φ es derivable en $B(0,\delta)$ con

$$D\varphi(x) = Df(a+h+x) - Df(a+x) - D(Df)(a)(h), \quad \forall x \in B(0,\delta),$$

donde se ha usado que $D(Df)(a)(h) \in \mathscr{L}(\mathbb{R}^N, \mathbb{R}^M)$ y en consecuencia $D(D(Df)(a)(h))(x) = D(Df)(a)(h)$. Aplicando a φ el Teorema del valor medio (Teorema 4.5) en el segmento $[0,k]$ obtenemos

$$\|\varphi(k) - \varphi(0)\| \leq \|k\| \ \sup\{\|D\varphi(x)\| : x \in]0,k[\}.$$

Para cada $x \in [0,k]$ se tiene

$$\|D\varphi(x)\| = \|Df(a+h+x) - Df(a+x) - D(Df)(a)(h)\| \leq$$

$$\|Df(a+h+x) - Df(a) - D(Df)(a)(h+x)\| + \|-Df(a+x) + Df(a) + D(Df)(a)(x)\| \leq$$

$$\leq \frac{\varepsilon}{2}\|h+x\| + \frac{\varepsilon}{2}\|x\| \leq \varepsilon\Big(\|h\| + \|x\|\Big) \leq \varepsilon\Big(\|h\| + \|k\|\Big),$$

donde se ha utilizado la desigualdad 5.4.1. En consecuencia,

$$\|\varphi(k) - \varphi(0)\| \leq \|k\|\varepsilon\Big(\|h\| + \|k\|\Big) \leq \varepsilon\Big(\|h\| + \|k\|\Big)^2.$$

Hemos probado que

$$\|f(a+h+k) - f(a+k) - f(a+h) + f(a) - D(Df)(a)(h)(k)\| \leq \varepsilon\Big(\|h\| + \|k\|\Big)^2.$$

De la simetría de la función

$$(h,k) \longrightarrow f(a+h+k) - f(a+h) - f(a+k) + f(a)$$

y de la propiedad triangular se deduce

$$\|D(Df)(a)(h)(k) - D(Df)(a)(k)(h)\| \leq 2\varepsilon\big(\|h\| + \|k\|\big)^2.$$

Sean ahora $x_1, x_2 \in \mathbb{R}^N$ vectores no nulos (en otro caso la igualdad del enunciado es inmediata); se tiene para cualquier $t > 0$ que

$$\frac{\|D^2f(a)(x_1,x_2) - D^2f(a)(x_2,x_1)\|}{\Big(\|x_1\| + \|x_2\|\Big)^2} =$$

$$\frac{\|D^2f(a)(tx_1,tx_2) - D^2f(a)(tx_2,tx_1)\|}{\Big(\|tx_1\| + \|tx_2\|\Big)^2} =$$

$$\frac{\|D(Df)(a)(tx_1)(tx_2) - D(Df)(a)(tx_2)(tx_1)\|}{\Big(\|tx_1\| + \|tx_2\|\Big)^2},$$

y por tanto, tomando $t > 0$ de manera que $0 < \|tx_1\|, \|tx_2\| < \delta$, concluimos que

$$\frac{\|D^2f(a)(x_1,x_2) - D^2f(a)(x_2,x_1)\|}{\Big(\|x_1\| + \|x_2\|\Big)^2} \leq 2\varepsilon$$

de donde se deduce el enunciado dada la arbitrariedad de ε. ∎

5.5. Fórmula de Taylor.

Definición 5.7 (Polinomio de Taylor de orden dos)**.** Sea $A \subset \mathbb{R}^N$ y $f : A \longrightarrow \mathbb{R}^M$ una función dos veces derivable en un punto $a \in A$. La función $P_2 : \mathbb{R}^N \longrightarrow \mathbb{R}^M$ definida por

$$P_2(x) = f(a) + df(a)(x-a) + \frac{1}{2}d^2 f(a)(x-a)$$

se denomina *polinomio de Taylor de orden 2 de f en a*, donde hemos escrito $df(a)$ en lugar de $Df(a)$ y

$$d^2 f(a)(x) = Df^2(a)(x,x), \ \forall x \in \mathbb{R}^N.$$

Teorema 5.8 (Fórmula infinitesimal del resto)**.** *Sean $A \subset \mathbb{R}^N$ y $f : A \longrightarrow \mathbb{R}^M$ un campo vectorial dos veces derivable en un punto $a \in A$. Entonces se verifica*

$$\lim_{x \to a} \frac{f(x) - P_2(x)}{\|x-a\|^2} = 0.$$

Demostración:

Por hipótesis, dado $\varepsilon > 0$, elegimos $\delta > 0$ verificando

$$\|x-a\| < \delta \Rightarrow \begin{cases} x \in A \\ f \text{ es derivable en } x \\ \|Df(x) - Df(a) - D(Df)(a)(x-a)\| \leq \varepsilon \|x-a\| \end{cases} \tag{5.5.1}$$

La función $g : B(a,\delta) \longrightarrow \mathbb{R}^M$ definida por

$$g(x) = f(x) - f(a) - df(a)(x-a) - \frac{1}{2}d^2 f(a)(x-a), \quad \forall x \in B(a,\delta)$$

es derivable con

$$Dg(x) = Df(x) - Df(a) - D^2 f(a)(x-a), \quad \forall x \in B(a,\delta).$$

En efecto, la única parte que merece ser comentada es la derivación de la función $Q : \mathbb{R}^N \longrightarrow \mathbb{R}^M$ definida por

$$Q(x) = \frac{1}{2}d^2 f(a)(x-a) = \frac{1}{2}D^2 f(a)(x-a,x-a), \quad \forall x \in \mathbb{R}^N.$$

De la derivación de una aplicación bilineal y de la regla de la cadena deducimos que Q es derivable en $x \in \mathbb{R}^N$ con

$$DQ(x)(z) = \frac{1}{2}\Big(D^2 f(a)(x-a,z) + D^2 f(a)(z,x-a)\Big), \quad \forall z \in \mathbb{R}^N.$$

Finalmente el Teorema de Schwarz (Teorema 5.6) garantiza que

$$DQ(x) = D(Df)(a)(x-a).$$

Sea ahora $x \in \mathbb{R}^N$ con $0 < \|x-a\| < \delta$. Se tiene que $[a,x] \subset B(a,\delta)$ y la función g es continua en $[a,x]$ y derivable en $]a,x[$. Aplicando el Teorema del valor medio al campo vectorial g en el segmento $[a,x]$ obtenemos que

$$\|g(x)-g(a)\| \leq \|x-a\| \ \sup\Big\{\|Df(z)-Df(a)-D(Df)(a)(z-a)\| : z \in]a,x[\Big\} \leq$$

$$\leq \|x-a\| \sup\Big\{\varepsilon\|z-a\| : z \in]a,x[\Big\} = \varepsilon\|x-a\|^2,$$

donde se ha utilizado 5.5.1. Por otra parte, por ser

$$\|g(x)-g(a)\| = \Big\|f(x)-\big(f(a)+df(a)(x-a)+\frac{1}{2}d^2f(a)(x-a)\big)\Big\|,$$

hemos concluido la demostración. ∎

Teorema 5.9 (Fórmula de Taylor con resto de Lagrange para campos esc.)**.** *Sean $A \subset \mathbb{R}^N$ y $a,x \in A$ con $a \neq x$, tales que el segmento $[a,x]$ está incluido en A. Si $f : A \longrightarrow \mathbb{R}$ es una función de clase $\mathscr{C}^1$ en $[a,x]$ y dos veces derivable en $]a,x[$, entonces existe $c \in]a,x[$ tal que*

$$f(x) = f(a)+df(a)(x-a)+\frac{1}{2}d^2f(c)(x-a).$$

Demostración:

La demostración de este resultado consiste en aplicar la Fórmula de Taylor clásica a la función auxiliar $\sigma : [0,1] \longrightarrow \mathbb{R}$ definida por

$$\sigma(t) = f(a+t(x-a)).$$

La regla de la cadena asegura que σ es de clase $\mathscr{C}^1$ en $[0,1]$ y dos veces derivable en $]0,1[$ con

$$\sigma'(t) = df(a+t(x-a))(x-a), \quad \forall t \in [0,1].$$

Usando la Nota 5.5, obtenemos

$$\sigma''(t) = d^2f\Big(a+t(x-a)\Big)(x-a), \quad \forall t \in [0,1].$$

Así, la función σ cumple las hipótesis de la Fórmula de Taylor y, por tanto, existe $t_0 \in]0,1[$ tal que

$$\sigma(1) = \sigma(0)+\sigma'(0)+\frac{1}{2}\sigma''(t_0),$$

es decir,

$$f(x) = f(a)+df(a)(x-a)+\frac{1}{2}d^2f(a+t_0(x-a))(x-a).$$

∎

Teorema 5.10 (Fórmula de Taylor)**.** *Sean $A \subset \mathbb{R}^N$ y $a,x \in A$ con $a \neq x$, tales que el segmento $[a,x]$ está incluido en A. Si $f : A \longrightarrow \mathbb{R}^M$ es un campo vectorial de clase $\mathscr{C}^1$ en $[a,x]$ y dos veces derivable en $]a,x[$, entonces*

$$\|f(x)-f(a)-df(a)(x-a)\| \leq \frac{1}{2}\|x-a\|^2 \sup\Big\{\|D^2f(z)\| : z \in]a,x[\Big\}.$$

Demostración:

Si el conjunto

$$\left\{\|D^2 f(z)\| : z \in]a,x[\right\}$$

no está mayorado, no hay nada que probar. En otro caso, sea

$$K := \sup\left\{\|D^2 f(z)\| : z \in]a,x[\right\}.$$

La prueba de este teorema consiste en aplicar la Proposición 4.6 a las funciones $\sigma : [0,1] \longrightarrow \mathbb{R}^M$ y $g : [0,1] \longrightarrow \mathbb{R}$ definidas por

$$\sigma(t) = f\Big(a+t(x-a)\Big) + (1-t)Df\Big(a+t(x-a)\Big)(x-a)$$

y

$$g(t) = \frac{-1}{2}K\|x-a\|^2(1-t)^2.$$

σ y g son continuas en $[0,1]$ y derivables en $]0,1[$ con

$$\sigma'(t) = Df\Big(a+t(x-a)\Big)(x-a) - Df\Big(a+t(x-a)\Big)(x-a)+$$

$$+(1-t)D^2 f\Big(a+t(x-a)\Big)(x-a,x-a) =$$

$$= (1-t)D^2 f\Big(a+t(x-a)\Big)(x-a,x-a)$$

y

$$g'(t) = K\|x-a\|^2(1-t).$$

En consecuencia,

$$\left\|\sigma'(t)\right\| \leq g'(t), \quad \forall t \in]0,1[.$$

Por tanto,

$$\|\sigma(1)-\sigma(0)\| \leq g(1)-g(0),$$

es decir,

$$\|f(x)-f(a)-df(a)(x-a)\| \leq \frac{1}{2}\|x-a\|^2 K.$$

■

En vista de las reglas de derivación (Sección 5.4, apartado vi)), para conocer la derivada segunda de un campo vectorial, es suficiente conocer las derivadas segundas de los campos escalares componentes. Por esta razón nos limitamos al estudio de la derivada segunda de campos escalares.

5.6. Campos escalares.

Definición 5.11 (Derivadas parciales segundas. Matriz hessiana)**.** Sean $A \subset \mathbb{R}^N$, $a \in A$ y f un campo escalar en A. Sea $j \in \{1,\ldots,N\}$ y $A^j \subset A$ el conjunto de puntos donde f tiene derivada parcial respecto de la variable j-ésima. Si $i \in \{1,\ldots,N\}$ y el campo escalar $D_j f$ en A^j tiene derivada parcial respecto de la variable i-ésima en el punto a, entonces se dice que f tiene *derivada parcial segunda* (o de segundo orden) respecto de las variables j e i en el punto a. En tal caso notaremos $D_{(i,j)}f(a)$ o bien $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(a)$ a esta derivada parcial de segundo orden. En caso de que $i = j$, escribiremos simplemente $\dfrac{\partial^2 f}{\partial x_i^2}(a)$.
Si $A^{(i,j)} \subset A$ es el conjunto de puntos donde f tiene derivada parcial segunda respecto de las variables j e i, entonces el campo escalar en $A^{(i,j)}$ definido por

$$x \mapsto D_{(i,j)}f(x)$$

es *la aplicación derivada parcial segunda* de f respecto de las variables j e i y se nota por $D_{(i,j)}f$ ó también $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$.

Se dice que el campo escalar f tiene matriz hessiana en el punto a si f admite cualquier derivada parcial de segundo orden en a. En tal caso se define la *matriz hessiana* de f en a por

$$H_f(a) = \begin{pmatrix} D_{(1,1)}f(a) & \ldots & D_{(N,1)}f(a) \\ \ldots\ldots & \ldots & \ldots\ldots \\ D_{(1,N)}f(a) & \ldots & D_{(N,N)}f(a) \end{pmatrix} = \Big(D_{(i,j)}f(a)\Big)_{1\le i,j\le N}.$$

Si C es el conjunto de puntos de A donde f tiene matriz hessiana, entonces la aplicación de C en $\mathscr{M}_{N\times N}(\mathbb{R})\big(\equiv \mathbb{R}^{N^2}\big)$ dada por

$$x \mapsto H_f(x)$$

es *la aplicación matriz hessiana* de f y se nota por H_f.

El siguiente resultado es una caracterización de la existencia de derivada segunda para campos escalares en función del gradiente. Recoge también que si un campo escalar tiene derivada segunda, entonces las derivadas parciales segundas cruzadas son iguales.

Teorema 5.12. *Sean $A \subset \mathbb{R}^N$, $a \in A$ y f un campo escalar en A. Las siguientes afirmaciones son equivalentes:*

i) f es dos veces derivable en a.

ii) f es derivable en un entorno de a y el gradiente de f es derivable en a.

En el caso de que se verifiquen las anteriores condiciones, se tiene que

$$D^2 f(a)(e_i, e_j) = D_{(i,j)}f(a) \quad (1 \le i,j \le N).$$

En consecuencia, la matriz $H_f(a)$ es simétrica y se verifica

$$D^2 f(a)(x,y) = xH_f(a)y^t, \qquad \forall x,y \in \mathbb{R}^N.$$

Además

$$f \in \mathscr{C}^2(a) \Leftrightarrow \nabla f \in \mathscr{C}^1(a).$$

Demostración:
i) ⇒ ii)
Sea U un entorno abierto de a incluido en A tal que f es derivable en U. Para cada $j \in \{1,\ldots,n\}$ existe $D_j f$ en U y

$$D_j f(x) = Df(x)(e_j), \quad \forall x \in U,$$

y en consecuencia,

$$D_j f = E_j \circ Df$$

donde E_j es el funcional de evaluación en e_j, esto es, la aplicación lineal de $\mathscr{L}(\mathbb{R}^N, \mathbb{R})$ en $\mathbb{R}$ definida por

$$E_j(T) = T(e_j).$$

La regla de la cadena garantiza ahora que $D_j f$ es derivable en a, y en el caso de que $f \in \mathscr{C}^2(a)$ se tiene que $D_j f \in \mathscr{C}^1(a)$. Además las derivadas parciales segundas de f en a se pueden calcular de la siguiente manera:

$$D_{(i,j)} f(a) = D_i(D_j f)(a) = D_i(E_j \circ Df)(a) =$$

$$D\Big(E_j \circ Df\Big)(a)(e_i) = \Big(E_j \circ D(Df)(a)\Big)(e_i) = E_j\Big(D(Df)(a)(e_i)\Big) =$$

$$= D(Df)(a)(e_i)(e_j) = D^2 f(a)(e_i, e_j),$$

para cualquier $i \in \{1,..,N\}$. El Teorema de Schwarz garantiza que $D^2 f(a)$ es simétrica, por tanto $H_f(a)$ también.

ii) ⇒ i)
Sea U un entorno abierto de a incluido en A tal que f es derivable en U. Se tiene que

$$Df(x)(y) = \Big(\nabla f(x) \mid y\Big), \quad \forall x \in U, \quad \forall y \in \mathbb{R}^N.$$

Si notamos por ϕ a la identificación usual de $\mathbb{R}^N$ con $\mathscr{L}(\mathbb{R}^N, \mathbb{R})$ dada por

$$\phi(x)(y) = (x \mid y), \quad \forall x, y \in \mathbb{R}^N,$$

entonces $Df = \phi \circ \nabla f$.

Como ∇f es derivable en a y ϕ es lineal, la regla de la cadena asegura que Df es derivable en a, esto es, f es dos veces derivable en a, y en el caso de que $\nabla f \in \mathscr{C}^1(a)$ se tiene que $Df \in \mathscr{C}^1(a)$, esto es, $f \in \mathscr{C}^2(a)$. ■

El resultado anterior junto con el teorema de caracterización de los campos escalares de clase $\mathscr{C}^1$ (Teorema 3.21) prueban el siguiente resultado:

Corolario 5.13. *Sean $A \subset \mathbb{R}^N$ un abierto y f un campo escalar en A. Las siguientes afirmaciones son equivalentes:*

i) f es dos veces derivable en A.

ii) ∇f *es derivable en A.*

Además,

$$f \in \mathscr{C}^2(A) \Leftrightarrow H_f \in \mathscr{C}(A).$$

A pesar de que el Teorema anterior afirma que si f es dos veces derivable en a, las derivadas parciales cruzadas coinciden, hay ejemplos sencillos de campos escalares donde esto no ocurre. Incluimos uno de ellos:

Ejemplo 5.14. Consideremos el campo escalar en $\mathbb{R}^2$ dado por

$$f(x,y) = \frac{xy(x^2-y^2)}{x^2+y^2} \quad \text{si} \ \ (x,y) \neq (0,0); \ f(0,0) = 0.$$

Se tiene que

$$D_1 f(0,0) = 0$$

$$D_1 f(x,y) = \frac{x^4 y + 4x^2 y^3 - y^5}{(x^2+y^2)^2} \ \text{si} \ \ (x,y) \neq (0,0)$$

por tanto,

$$D_{(2,1)} f(0,0) = \lim_{y \to 0} \frac{D_1 f(0,y) - D_1 f(0,0)}{y} = \lim_{y \to 0} \frac{-y^5}{y^5} = -1$$

$$D_2 f(0,0) = 0, \quad D_2 f(x,y) = \frac{x^5 - 4x^3 y^2 - xy^4}{(x^2+y^2)^2} \quad \text{si} \ \ (x,y) \neq (0,0),$$

de donde

$$D_{(1,2)} f(0,0) = \lim_{x \to 0} \frac{D_2 f(x,0) - D_2 f(0,0)}{x} = \lim_{x \to 0} \frac{x^5}{x^5} = 1.$$

■

La regla de la cadena para las derivadas parciales segundas se deduce a partir de la fórmula obtenida en la regla de la cadena para la diferencial segunda, si bien conviene codificar que se obtiene simplemente derivando parcialmente en la expresión obtenida para las derivadas parciales primeras.

Ejemplo 5.15 (Laplaciano en polares)**.** Sea $z = z(x,y)$ un campo escalar en $\mathbb{R}^2$ dos veces derivable. Si cambiamos a coordenadas polares, esto es, hacemos

$$x = \rho \cos \vartheta, \quad y = \rho \operatorname{sen} \vartheta$$

y notamos

$$w(\rho, \vartheta) = z(\rho \cos \vartheta, \rho \operatorname{sen} \vartheta),$$

entonces derivando parcialmente en las expresiones

$$\frac{\partial w}{\partial \rho}(\rho, \vartheta) = \frac{\partial z}{\partial x}(\rho \cos \vartheta, \rho \operatorname{sen} \vartheta) \cos \vartheta + \frac{\partial z}{\partial y}(\rho \cos \vartheta, \rho \operatorname{sen} \vartheta) \operatorname{sen} \vartheta$$

$$\frac{\partial w}{\partial \vartheta}(\rho,\vartheta)=\frac{\partial z}{\partial x}(\rho\cos\vartheta,\rho\operatorname{sen}\vartheta)(-\rho\operatorname{sen}\vartheta)+\frac{\partial z}{\partial y}(\rho\cos\vartheta,\rho\operatorname{sen}\vartheta)\rho\cos\vartheta$$

obtenemos

$$\frac{\partial^2 w}{\partial \rho^2}=\frac{\partial}{\partial\rho}\left(\frac{\partial w}{\partial\rho}\right)=\frac{\partial}{\partial\rho}\left(\frac{\partial z}{\partial x}\cos\vartheta+\frac{\partial z}{\partial y}\operatorname{sen}\vartheta\right)=$$

$$\frac{\partial}{\partial\rho}\left(\frac{\partial z}{\partial x}\right)\cos\vartheta+\frac{\partial z}{\partial x}\frac{\partial}{\partial\rho}(\cos\vartheta)+\frac{\partial}{\partial\rho}\left(\frac{\partial z}{\partial y}\right)\operatorname{sen}\vartheta+\frac{\partial z}{\partial y}\frac{\partial}{\partial\rho}(\operatorname{sen}\vartheta)=$$

$$\left(\frac{\partial^2 z}{\partial x^2}\cos\vartheta+\frac{\partial^2 z}{\partial y\partial x}\operatorname{sen}\vartheta\right)\cos\vartheta+0+\left(\frac{\partial^2 z}{\partial x\partial y}\cos\vartheta+\frac{\partial^2 z}{\partial y^2}\operatorname{sen}\vartheta\right)\operatorname{sen}\vartheta+0=$$

$$\frac{\partial^2 z}{\partial x^2}\cos^2\vartheta+\frac{\partial^2 z}{\partial y^2}\operatorname{sen}^2\vartheta+2\operatorname{sen}\vartheta\cos\vartheta\frac{\partial^2 z}{\partial y\partial x},$$

$$\frac{\partial^2 w}{\partial\vartheta\partial\rho}=\frac{\partial}{\partial\vartheta}\left(\frac{\partial w}{\partial\rho}\right)=\frac{\partial}{\partial\vartheta}\left(\frac{\partial z}{\partial x}\cos\vartheta+\frac{\partial z}{\partial y}\operatorname{sen}\vartheta\right)=$$

$$\frac{\partial}{\partial\vartheta}\left(\frac{\partial z}{\partial x}\right)\cos\vartheta+\frac{\partial z}{\partial x}\frac{\partial}{\partial\vartheta}(\cos\vartheta)+\frac{\partial}{\partial\vartheta}\left(\frac{\partial z}{\partial y}\right)\operatorname{sen}\vartheta+\frac{\partial z}{\partial y}\frac{\partial}{\partial\vartheta}(\operatorname{sen}\vartheta)=$$

$$\left(\frac{\partial^2 z}{\partial x^2}(-\rho\operatorname{sen}\vartheta)+\frac{\partial^2 z}{\partial y\partial x}(\rho\cos\vartheta)\right)\cos\vartheta+\frac{\partial z}{\partial x}(-\operatorname{sen}\vartheta)+$$

$$\left(\frac{\partial^2 z}{\partial x\partial y}(-\rho\operatorname{sen}\vartheta)+\frac{\partial^2 z}{\partial y^2}(\rho\cos\vartheta)\right)\operatorname{sen}\vartheta+\frac{\partial z}{\partial y}\cos\vartheta$$

y

$$\frac{\partial^2 w}{\partial\vartheta^2}=\frac{\partial}{\partial\vartheta}\left(\frac{\partial w}{\partial\vartheta}\right)=$$

$$=\frac{\partial}{\partial\vartheta}\left(\frac{\partial z}{\partial x}(-\rho\operatorname{sen}\vartheta)+\frac{\partial z}{\partial y}(\rho\cos\vartheta)\right)=$$

$$\frac{\partial}{\partial\vartheta}\left(\frac{\partial z}{\partial x}\right)(-\rho\operatorname{sen}\vartheta)+\frac{\partial z}{\partial x}\frac{\partial}{\partial\vartheta}(-\rho\operatorname{sen}\vartheta)+\frac{\partial}{\partial\vartheta}\left(\frac{\partial z}{\partial y}\right)(\rho\cos\vartheta)+\frac{\partial z}{\partial y}\frac{\partial}{\partial\vartheta}(\rho\cos\vartheta)=$$

$$\left(\frac{\partial^2 z}{\partial x^2}(-\rho\operatorname{sen}\vartheta)+\frac{\partial^2 z}{\partial y\partial x}(\rho\cos\vartheta)\right)(-\rho\operatorname{sen}\vartheta)+\frac{\partial z}{\partial x}(-\rho\cos\vartheta)+$$

$$\left(\frac{\partial^2 z}{\partial x\partial y}(-\rho\operatorname{sen}\vartheta)+\frac{\partial^2 z}{\partial y^2}(\rho\cos\vartheta)\right)(\rho\cos\vartheta)+\frac{\partial z}{\partial y}(-\rho\operatorname{sen}\vartheta)=$$

$$-\rho\left(\frac{\partial z}{\partial x}\cos\vartheta+\frac{\partial z}{\partial y}\operatorname{sen}\vartheta\right)+\rho^2\left(\frac{\partial^2 z}{\partial x^2}\operatorname{sen}^2\vartheta+\frac{\partial^2 z}{\partial y^2}\cos^2\vartheta-2\frac{\partial^2 z}{\partial y\partial x}\operatorname{sen}\vartheta\cos\vartheta\right).$$

En consecuencia, el laplaciano de z, definido por

$$\Delta z=D_{(1,1)}z+D_{(2,2)}z$$

se expresa en coordenadas polares de la siguiente manera

$$\Delta z(\rho\cos\vartheta, \rho\,\mathrm{sen}\,\vartheta) = \frac{\partial^2 w}{\partial \rho^2}(\rho, \vartheta) + \frac{1}{\rho^2}\frac{\partial^2 w}{\partial \vartheta^2}(\rho, \vartheta) + \frac{1}{\rho}\frac{\partial w}{\partial \rho}(\rho, \vartheta).$$

■

5.7. Referencias recomendadas.

[MaHo].

5.8. Resumen del resultados del Tema 5

Aplicaciones bilineales.

Se verifica que $\mathscr{L}(\mathbb{R}^N, \mathscr{L}(\mathbb{R}, \mathbb{R}^P)) \equiv \mathscr{L}^2(\mathbb{R}^N \times \mathbb{R}^M, \mathbb{R}^P))$ (conjunto de aplicaciones bilineales de $\mathbb{R}^N \times \mathbb{R}^M$ en $\mathbb{R}^P$). La identificación viene dada por

$$T \mapsto \Phi, \quad \text{donde} \quad \Phi(x,y) = T(x)(y) \quad (x \in \mathbb{R}^N, y \in \mathbb{R}^M).$$

La expresión de la norma inducida por esta identificación en el espacio de las aplicaciones bilineales es

$$\|\Phi\| = \max\left\{\|\Phi(x,y)\| : \ x \in \mathbb{R}^N, y \in \mathbb{R}^M, \|x\| = 1, \|y\| = 1\right\} =$$

$$= \max\left\{\|\Phi(x,y)\| : \ x \in \mathbb{R}^N, y \in \mathbb{R}^M, \|x\| \leq 1, \|y\| \leq 1\right\}.$$

Cuando $P = 1$ y $N = M$, las aplicaciones bilineales de $\mathbb{R}^N \times \mathbb{R}^N$ en $\mathbb{R}$ se identifican con $\mathscr{M}_{N\times N}(\mathbb{R})$; si $A \in \mathscr{M}_{N\times N}(\mathbb{R})$, la forma bilineal asociada Φ_A viene dada por

$$\Phi_A(x,y) = \left(xA^t \mid y\right) = xA^t y^t \quad (x, y \in \mathbb{R}^N).$$

Además Φ_A es simétrica si, y sólo si, A es simétrica, en cuyo caso

$$\Phi_A(x,y) = xAy^t.$$

Campos vectoriales 2 veces derivables.
Sean $A \subset \mathbb{R}^N, f : A \longrightarrow \mathbb{R}^M$ y sea $A_1 \subset A$ el conjunto de puntos donde f es derivable. Se dice que f es dos veces derivable en $a \in A_1$ si la función $Df : A_1 \longrightarrow \mathscr{L}(\mathbb{R}^N, \mathbb{R}^M)$ es derivable en a, en cuyo caso la aplicación

$$D(Df)(a) \in \mathscr{L}(\mathbb{R}^N, \mathscr{L}(\mathbb{R}^N, \mathbb{R}^M)) \equiv \mathscr{L}^2(\mathbb{R}^N \times \mathbb{R}^N, \mathbb{R}^M)$$

se denomina la derivada segunda de f en a y se nota $D^2 f(a)$.

Si $A \subset \mathbb{R}$ y $f : A \longrightarrow \mathbb{R}^M$ es un campo vectorial 2 veces derivable en un punto $a \in A$, entonces

$$D^2 f(a)(1,1) = f''(a),$$

Propiedades de la derivada segunda y de las funciones 2 veces derivables.

i) **Linealidad.** Sean $A \subset \mathbb{R}^N$, $a \in A$ y $f, g : A \longrightarrow \mathbb{R}^M$ funciones 2 veces derivables en a y sea $\lambda \in \mathbb{R}$. Entonces $f + g$ y λf son 2 veces derivables en a con

$$D^2(f+g)(a) = D^2 f(a) + D^2 g(a), \quad D^2(\lambda f)(a) = \lambda D^2 f(a).$$

Además, si $f, g \in \mathscr{C}^2(a)$, entonces $f + g \in \mathscr{C}^2(a)$ y $\lambda f \in \mathscr{C}^2(a)$.

ii) **Regla de la cadena.** Sean $A \subset \mathbb{R}^N, B \subset \mathbb{R}^M, f : A \longrightarrow \mathbb{R}^M$ tal que $f(A) \subset B$ y $g : B \longrightarrow \mathbb{R}^P$. Supongamos que f es 2 veces derivable en $a \in A$ y que g es 2 veces derivable en $f(a)$. Entonces la composición $h = g \circ f$ es 2 veces derivable en a. Además se tiene que

$$D^2h(a)(x_1,x_2) =$$

$$= D^2g(f(a))\Big(Df(a)(x_1), Df(a)(x_2)\Big) + Dg(f(a))\Big(D^2f(a)(x_1,x_2)\Big)$$

para cualesquiera $x_1, x_2 \in \mathbb{R}^N$.

En el caso particular de que $A \subset \mathbb{R}$ se obtiene

$$(g \circ f)''(a) = D^2g(f(a))\Big(f'(a), f'(a)\Big) + Dg(f(a))(f''(a)).$$

iii) **Carácter local de la derivada segunda.** Sean $A \subset \mathbb{R}^N$, $a \in A$ y $f : A \longrightarrow \mathbb{R}^M$ un campo vectorial. Entonces f es 2 veces derivable en a si, y sólo si, existe algún entorno $U \subset A$ de a tal que $f_{|U}$ es 2 veces derivable en a, en cuyo caso las derivadas segundas en a de ambas funciones coinciden.

iv) **Derivación de la función inversa.** Sean A y B subconjuntos abiertos de $\mathbb{R}^N$ y f un homeomorfismo de A sobre B. Si f es 2 veces derivable en a (resp. de clase $\mathscr{C}^k$ en a) para algún $a \in A$ y $Df(a) \in$ Iso $(\mathbb{R}^N)$, entonces f^{-1} es 2 veces derivable en $f(a)$ (resp. de clase $\mathscr{C}^2$ en $f(a)$).

vi) **Reducción a campos escalares.** Sean $A \subset \mathbb{R}^N$ y $f : A \longrightarrow \mathbb{R}^M$ un campo vectorial. Si $f = (f_1, \ldots, f_M)$, entonces f es 2 veces derivable en $a \in A$ si, y sólo si, cada f_i es 2 veces derivable en a para $i = 1, \ldots, M$, en cuyo caso

$$D^2f(a) = \Big(D^2f_1(a), \ldots, D^2f_M(a)\Big)$$

Por tanto, una campo vectorial es de clase 2 en un punto cuando todas sus componentes lo sean también.

Las aplicaciones lineales son de clase 2 y su diferencial segunda es nula. Toda aplicación bilineal $T : \mathbb{R}^N \times \mathbb{R}^M \longrightarrow \mathbb{R}^P$ es de clase $\mathscr{C}^2$ con derivada segunda en $(a,b) \in \mathbb{R}^N \times \mathbb{R}^M$ definida por

$$D^2T(a,b)\Big((x_1,y_1),(x_2,y_2)\Big) = T(x_1,y_2) + T(x_2,y_1), \quad \forall (x_1,y_1),(x_2,y_2) \in \mathbb{R}^N \times \mathbb{R}^M,$$

La aplicación inversión $J :$ Iso $(\mathbb{R}^N) \longrightarrow \mathscr{L}(\mathbb{R}^N)$ es de clase $\mathscr{C}^2$.

Teorema de Schwarz para la derivada segunda.
Sea $A \subset \mathbb{R}^N$, $f : A \longrightarrow \mathbb{R}^M$ un campo vectorial. Supongamos que f es 2 veces derivable en un punto $a \in A$. Entonces la aplicación bilineal $D^2f(a)$ es simétrica, esto es,

$$D^2f(a)\big(x,y\big) = D^2f(a)\big(y,x\big)$$

para cualesquiera $x,y \in \mathbb{R}^N$.

Polinomio de Taylor de orden 2.
Sea $A \subset \mathbb{R}^N$ y $f : A \longrightarrow \mathbb{R}^M$ una función. Supongamos que f es 2 veces derivable en un punto $a \in A$, entonces la función $P_2 : \mathbb{R}^N \longrightarrow \mathbb{R}^M$ definida por

$$P_2(x) = f(a) + df(a)(x-a) + \frac{1}{2}d^2 f(a)(x-a)$$

se llama polinomio de Taylor de orden 2 de f en a. El polinomio de Taylor de orden 2 de una función en un punto a verifica

$$P_2(a) = f(a), \quad DP_2(a) = Df(a), \quad D^2 P_2(a) = D^2 f(a).$$

Fórmula infinitesimal del resto.
Sea $A \subset \mathbb{R}^N$ y $f : A \longrightarrow \mathbb{R}^M$ un campo vectorial. Supongamos que f es 2 veces derivable en un punto $a \in A$, entonces se verifica

$$\lim_{x \to a} \frac{f(x) - P_2(x)}{\|x-a\|^2} = 0.$$

Fórmula de Taylor con resto de Lagrange para campos escalares.
Sean $A \subset \mathbb{R}^N$ y $a,x \in A$ con $a \neq x$, tales que el segmento $[a,x]$ está incluido en A. Si $f : A \longrightarrow \mathbb{R}$ es una función de clase $\mathscr{C}^1$ en $[a,x]$ y dos veces derivable en $]a,x[$, entonces existe $c \in]a,x[$ tal que

$$f(x) = f(a) + df(a)(x-a) + \frac{1}{2}d^2 f(c)(x-a).$$

Fórmula de Taylor.
Sean $A \subset \mathbb{R}^N$ y $a,x \in A$ con $a \neq x$, tales que el segmento $[a,x]$ está incluido en A. Si $f : A \longrightarrow \mathbb{R}^M$ es un campo vectorial de clase $\mathscr{C}^1$ en $[a,x]$ y dos veces derivable en $]a,x[$, entonces

$$\|f(x) - f(a) - df(a)(x-a)\| \leq \frac{1}{2}\|x-a\|^2 \sup\left\{ \|D^2 f(z)| : z \in]a,x[\right\}.$$

RESULTADOS PARA CAMPOS ESCALARES.

Derivadas parciales sucesivas. Matriz hessiana.
Sean $A \subset \mathbb{R}^N$, $a \in A$ y f un campo escalar en A. Sea $j \in \{1,\ldots,N\}$ y $A^j \subset A$ el conjunto de puntos donde está definida $D_j(f)$. Si $i \in \{1,\ldots,N\}$ y el campo escalar $D_j f$ en A^j tiene derivada parcial respecto de la variable i-ésima en el punto a, entonces se dice que f tiene derivada parcial segunda (o de segundo orden) respecto de las variables j e i en el punto a.

En tal caso notaremos $D_{(i,j)}f(a)$ o bien $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(a)$ a ésta derivada parcial de segundo orden.

En caso de que $i = j$ escribiremos simplemente $\dfrac{\partial^2 f}{\partial^2 x_i}(a)$. La aplicación

$$x \;\mapsto\; D_{(i,j)}f(x)$$

definida en el conjunto donde f tiene derivas parciales respecto de las variables j e i, es la aplicación derivada parcial segunda de f respecto de estas variables y se nota por $D_{(i,j)}f$ ó también $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$.

Se dice que el campo escalar f tiene matriz hessiana en el punto a si f admite cualquier derivada parcial de segundo orden en a. En tal caso se define la matriz hessiana de f en a por

$$H_f(a) = \begin{pmatrix} D_{(1,1)}f(a) & \dots & D_{(N,1)}f(a) \\ \dots\dots & \dots\dots & \dots\dots \\ D_{(1,N)}f(a) & \dots & D_{(N,N)}f(a) \end{pmatrix} = \Big(D_{(i,j)}f(a)\Big)_{1 \le i,j \le N}.$$

Si C es el conjunto de puntos de A donde f tiene matriz hessiana, entonces la aplicación de C en $\mathscr{M}_{N\times N}(\mathbb{R})\left(\equiv \mathbb{R}^{N^2}\right)$ dada por

$$x \mapsto H_f(x)$$

es la aplicación matriz hessiana de f y se nota por H_f.

El Teorema de Schwarz afirma que si f es dos veces derivable en a, la matriz hessiana es simétrica.

Caracterización de campos escalares 2 veces derivables.
Sean $A \subset \mathbb{R}^N$, $a \in A$ y f un campo escalar en A. Entonces f es 2 veces derivable en a si, y sólo si, f es derivable en un entorno de a y gradiente de f es derivable en a.

En el caso de que f sea dos veces diferenciable en a, se tiene que

$$D^2 f(a)(x,y) = xH_f(a)y^t, \qquad \forall x,y \in \mathbb{R}^N,$$

equivalentemente,

$$D_{(i,j)}f(a) = D^2 f(a)(e_i, e_j) \quad (1 \le i,j \le N).$$

Caracterización de campos escalares 2 veces derivables en términos del vector gradiente.
Sean $A \subset \mathbb{R}^N$ un abierto y f un campo escalar en A. Las siguientes afirmaciones son equivalentes:

i) *f es dos veces derivable en A.*

ii) *∇f es derivable en A.*

Además,

$$f \in \mathscr{C}^2(A) \Leftrightarrow H_f \in \mathscr{C}(A).$$

5.9. Ejercicios del Tema 5

5.1 Sea $f:\mathbb{R}^2 \longrightarrow \mathbb{R}$ la función definida por:

$$f(x,y)=\frac{\arctan(x)\operatorname{sen}y-xy}{x^2+y^2} \qquad \text{si} \quad (x,y)\neq(0,0), \quad f(0,0)=0.$$

Probar que f es de clase $\mathscr{C}^1$ en $\mathbb{R}^2$. Calcular $D_{(1,2)}f(0,0)$ y $D_{(2,1)}f(0,0)$. ¿Es f dos veces derivable en $(0,0)$?

5.2 Sea p un número real y $f:\mathbb{R}^N\backslash\{0\}\longrightarrow\mathbb{R}$ una función dos veces derivable y homogénea de grado p. Probar que

$$D^2f(x)(x,x)=p(p-1)f(x), \quad \forall x\in\mathbb{R}^N\backslash\{0\}.$$

Probar también que las derivadas parciales de primer orden D_if son funciones homogéneas de grado $p-1$ y deducir, cuando $p=1$ que el determinante Hessiano de f es nulo en todo punto.
Indicación: El Teorema de Euler y la relación entre diferencial y derivada direccional permiten comprobar la fórmula para la derivada segunda. Directamente, la definición de las derivadas parciales permite comprobar que son funciones $(p-1)$-homogéneas. Por último, expresar el Teorema de Euler en términos de las derivadas parciales. Derivando en esta expresión de nuevo parcialmente, se puede comprobar para $p=1$ que un sistema de ecuaciones lineales matriz de coeficientes la matriz hessiana tiene solución no trivial, de donde se concluye la condición sobre el determinante del hessiano.

5.3 Calcular el polinomio de Taylor de orden 2 de las siguientes funciones en el punto que se indica:

$$f(x,y)=\operatorname{sen}(x^2+3xy) \qquad \text{en } (0,0),$$
$$g(x,y)=\frac{\arctan(xy)}{1+x^2+y^2} \qquad \text{en } (0,0)$$
$$h(x,y)=\log(x^2+y^2) \qquad \text{en } (1,1).$$

5.4 Utilizar la fórmula infinitesimal del resto para probar que

$$\lim_{(x,y)\to(0,0)}\frac{e^{x\operatorname{sen}y}-(1+xy)}{x^2+y^2}=0.$$

5.5 Sea $A\subset\mathbb{R}^N$ un abierto conexo y $f:A\longrightarrow\mathbb{R}^M$ una función dos veces derivable tal que D^2f es constante. Probar que existen $b\in\mathbb{R}^M, T\in\mathscr{L}(\mathbb{R}^N,\mathbb{R}^M)$ y $S\in\mathscr{L}^2(\mathbb{R}^N,\mathbb{R}^M)$ tales que

$$f(x)=b+T(x)+S(x,x), \quad \forall x\in A.$$

5.6 Sea T una isometría lineal de $\left(\mathbb{R}^N, \|\cdot\|_2\right)$ en sí mismo y f un campo escalar de clase $\mathscr{C}^2$ en un abierto Ω de $\mathbb{R}^N$. Justificar que:

$$\Delta(f \circ T)(x) = \Delta f\big(T(x)\big), \quad \forall x \in T^{-1}(\Omega)$$

donde $\Delta f(x) := \sum_{i=1}^N \dfrac{\partial^2 f}{\partial x_i^2}(x)$.

Indicación: a) Pruébese que T conserva el producto escalar; para ello basta desarrollar $\|x+y\|_2^2$ con el fin de expresar el producto escalar en función de la norma euclídea.
b) Calcular las derivadas parciales que aparecen en le definición del operador laplaciano (Δ) (regla de la cadena para las derivadas parciales segundas). Traducir a) en términos de la matriz asociada a T para obtener el enunciado.

5.7 Obtener las funciones $f : \mathbb{R}^+ \longrightarrow \mathbb{R}$ de clase $\mathscr{C}^2$ tales que:

$$\Delta\big(f \circ \|\cdot\|_2\big)(x) = 0, \quad \forall x \in \mathbb{R}^N \backslash \{0\}.$$

Indicación: Calcular las derivadas parciales de segundo orden que intervienen en el laplaciano y resolver la ecuación diferencial en la que se traduce la hipótesis.

5.10. Soluciones a los ejercicios del Tema 5

5.1 Es inmediato que $D_1 f(0,0) = D_2 f(0,0) = 0$. Por comodidad de notación definimos

$$g(x,y) := \arctan(x)\operatorname{sen}(y) - xy, \quad \forall x,y \in \mathbb{R}.$$

Derivabilidad de f:
Es claro que

$$g(x,y) = (\arctan(x) - x)\operatorname{sen} y + x(\operatorname{sen}(y) - y),$$

por tanto,

$$\frac{|g(x,y)|}{\left(x^2+y^2\right)^{3/2}} \leq \frac{|\arctan(x) - x|}{x^2+y^2}\,\frac{|\operatorname{sen} y|}{(x^2+y^2)^{\frac{1}{2}}} + \frac{|x|}{(x^2+y^2)^{\frac{1}{2}}}\,\frac{|\operatorname{sen}(y) - y|}{x^2+y^2} \leq$$

$$\leq \frac{|\arctan(x) - x|}{x^2}\,\frac{|\operatorname{sen} y|}{|y|} + \frac{|\operatorname{sen}(y) - y|}{y^2}.$$

Es claro que el límite de la función anterior es cero, ya que

$$\lim_{x \to 0} \frac{\arctan(x) - x}{x^2} = 0 = \lim_{y \to 0} \frac{\operatorname{sen}(y) - y}{y^2}.$$

Para comprobar que el límite en $(0,0)$ de la función $\dfrac{|g(x,y)|}{\left(x^2+y^2\right)^{3/2}}$ vale cero, también puede usarse la fórmula infinitesimal del resto. En tal caso basta tener en cuenta que el polinomio de Taylor de orden 3 de g en el punto $(0,0)$ es nulo y que el denominador no es mas que $\|(x,y)\|_2^3$.

Continuidad de las derivadas parciales:

Las derivadas parciales fuera del origen son:

$$D_1 f(x,y) = \frac{D_1 g(x,y)}{x^2+y^2} - \frac{2x}{\left(x^2+y^2\right)^{1/2}}\,\frac{g(x,y)}{\left(x^2+y^2\right)^{3/2}}$$

$$D_2 f(x,y) = \frac{D_2 g(x,y)}{x^2+y^2} - \frac{2y}{\left(x^2+y^2\right)^{1/2}}\,\frac{g(x,y)}{\left(x^2+y^2\right)^{3/2}}$$

Queremos probar que ambas tienden a cero. Es inmediato de lo ya demostrado que los segundos sumandos tienden a cero. Probaremos que igual le ocurre a los primeros.

En efecto, por ser

$$D_1 g(x,y) = \frac{1}{1+x^2}\operatorname{sen}(y) - y,$$

$$\frac{|D_1 g(x,y)|}{x^2+y^2} = \frac{1}{1+x^2}\,\frac{|\operatorname{sen}(y) - y - x^2 y|}{x^2+y^2} \leq$$

$$\leq \frac{1}{1+x^2}\left(\frac{|\operatorname{sen}(y)-y|}{y^2}+\frac{x^2|y|}{x^2}\right) \longrightarrow 0.$$

Por otra parte,

$$D_2g(x,y) = \arctan(x)\cos(y) - x,$$

por tanto,

$$\frac{|D_2g(x,y)|}{x^2+y^2}| \leq \frac{|(\arctan(x)-x)\ \cos y|}{x^2+y^2} + \frac{|x\ (\cos(y)-1)|}{x^2+y^2} \leq$$

$$\leq \frac{|(\arctan(x)-x)\ \cos y|}{x^2} + \frac{|x\ (\cos(y)-1)|}{y^2} \to 0.$$

Como es claro que $f \in \mathscr{C}^1(\mathbb{R}^2\backslash\{(0,0)\})$, acabamos de probar que $f \in \mathscr{C}^1(\mathbb{R}^2)$.

Cálculo de las derivadas parciales segundas en el origen:

$$\frac{D_1f(0,y)-D_1f(0,0)}{y} = \frac{\operatorname{sen}(y)-y}{y^3} \longrightarrow \frac{-1}{6} =: D_{(2,1)}f(0,0)$$

$$\frac{D_2f(x,0)-D_2f(0,0)}{x} = \frac{\arctan(x)-x}{x^3} \longrightarrow \frac{-1}{3} =: D_{(1,2)}f(0,0).$$

Al no ser iguales las derivadas parciales segundas cruzadas, el Teorema 5.6 (la diferencial segunda es simétrica, por tanto la matriz hessiana también si f es dos veces derivable) nos permite concluir que f no es dos veces derivable en $(0,0)$.

5.2 El Teorema de Euler nos asegura que :

$$Df(x)(x) = pf(x), \quad \forall x \in \mathbb{R}^N\backslash\{0\}.$$

Dado un vector no nulo x de $\mathbb{R}^N$, definimos la función

$$g(t) = f(tx) \ \ (t > 0),$$

y usando la regla de la cadena para la diferencial segunda en el caso de que la primera función función sea de variable real, se obtiene que

$$g''(t) = D^2f(tx)(x,x).$$

Teniendo en cuenta que $g(t) = t^pf(x)$, se tiene también que

$$g''(t) = p(p-1)t^{p-2}f(x)$$

y basta evaluar en 1.

Para comprobar que las derivadas parciales de primer orden son homogéneas de grado $p-1$ se puede usar la definición de derivada parcial y, por supuesto, la homogeneidad de f. En vista de la igualdad, para $1 \leq i \leq N, x \in \mathbb{R}^N\backslash\{0\}$, $t \in \mathbb{R}^*, s > 0$

$$\frac{f(sx+te_i)-f(sx)}{t} = s^p\frac{f\left(x+\dfrac{t}{s}e_i\right)-f(x)}{t} = s^{p-1}\frac{f\left(x+\dfrac{t}{s}e_i\right)-f(x)}{\dfrac{t}{s}}$$

tomando límite $(t \to 0)$ se obtiene

$$D_i f(sx) = s^{p-1} D_i f(x),$$

esto es, $D_i f$ es homogénea de grado $p-1$.

La homogeneidad de las derivadas parciales de primer orden se puede obtener también del teorema de Euler, ya que

$$pf(x) = Df(x)(x) = (\nabla f(x) \mid x) = \sum_{j=1}^{N} x_j D_j f(x).$$

Derivando respecto de x_i $(i = 1, \ldots, N)$, obtenemos

$$pD_i f(x) = D_i f(x) + \sum_{j=1}^{N} x_j D_i(D_j f)(x)$$

y al ser f dos veces derivable, las derivadas parciales segundas cruzadas son iguales y, por tanto,

$$(p-1)D_i f(x) = \sum_{j=1}^{N} x_j D_j(D_i f)(x) = D\big(D_i f\big)(x)(x) \tag{5.10.1}$$

lo que nos asegura, en virtud del Teorema de Euler, que las derivadas parciales primeras son homogéneas de grado $p-1$.

Si $p = 1$, como, para $x \in \mathbb{R}^N \backslash \{0\}$ se tiene

$$H_f(x) = \begin{pmatrix} D_1(D_1 f)(x) & \ldots & D_N(D_1 f)(x) \\ \ldots & \ldots & \ldots \\ D_1(D_N f)(x) & \ldots & D_N(D_N f)(x) \end{pmatrix},$$

y la expresión 5.10.1 nos asegura que el sistema de ecuaciones lineales con matriz de coeficientes $H_f(x)$ tiene una solución no trivial, por ser $x \neq 0$, por tanto, el determinante del hessiano de f en x es nulo.

También puede usarse el Problema 3.14. Como hemos probado que $D_i f$ es 0-homogénea $(p = 1)$, sabemos que

$$D(D_i f)(x)(x) = 0, \quad \forall x \in \mathbb{R}^N \backslash \{0\}.$$

Es decir,

$$0 =< \nabla(D_i f)(x), x >= \sum_{j=1}^{N} D_{(j,i)} f(x) x_j, \quad \forall x \in \mathbb{R}^N \backslash \{0\}, \forall i \in \{1, \ldots, N\},$$

esto es,

$$H_f(x) x^t = 0, \quad \forall x \in \mathbb{R}^N \backslash \{0\}.$$

Por tanto, $\det H_f(x) = 0$, para todo vector no nulo x de $\mathbb{R}^N$.

Otra forma de obtener la condición sobre la diferencial segunda:

Llamamos $g \equiv Df$, con lo que, por ser f dos veces derivable en $\mathbb{R}^N \backslash \{0\}$ tenemos

$$g : \mathbb{R}^N \backslash \{0\} \longrightarrow \mathscr{L}(\mathbb{R}^N, \mathbb{R}), \quad Dg : \mathbb{R}^N \backslash \{0\} \longrightarrow \mathscr{L}(\mathbb{R}^N, \mathscr{L}(\mathbb{R}^N, \mathbb{R})) \equiv \mathscr{L}^2(\mathbb{R}^N, \mathbb{R}),$$

y se tiene para $a \in \mathbb{R}^N \backslash \{0\}$

$$Dg(a)(x) = \lim_{t \to 0} \frac{g(a+tx) - g(a)}{t}$$

y como la evaluación en un punto es una función continua en el espacio de operadores, tenemos

$$Dg(a)(x,x) = \lim_{t \to 0} \frac{g(a+tx)(x) - g(a)(x)}{t},$$

por tanto,

$$Dg(x)(x,x) = \lim_{t \to 0} \frac{g(x+tx)(x) - g(x)(x)}{t}. \tag{5.10.2}$$

En cada punto, g es lineal, por ser la diferencial de f, luego usando el Teorema de Euler, obtenemos la cadena de igualdades

$$\frac{g(x+tx)(x) - g(x)(x)}{t} = \frac{\dfrac{g(x+tx)((1+t)x)}{1+t} - g(x)(x)}{t} =$$

$$= \frac{\dfrac{pf((1+t)x)}{1+t} - pf(x)}{t} = pf(x)\frac{(1+t)^{p-1} - 1}{t},$$

tomando límite en cero, y usando la igualdad 5.10.2 concluimos que

$$Dg(x)(x,x) = p(p-1)f(x).$$

5.3 Para calcular las derivadas sucesivas en los puntos en cuestión se puede derivar y sustituir ó utilizar la definición de derivada parcial. Derivando con respecto a ambas variables se obtiene

$$D_1 f(x,y) = (2x+3y)\cos(x^2+3xy), \quad D_2 f(x,y) = 3x \cos(x^2+3xy),$$

por tanto

$$D_{(1,1)} f(x,y) = 2\cos(x^2+3xy) - (2x+3y)^2 \operatorname{sen}(x^2+3xy),$$

$$D_{(1,2)} f(x,y) = 3\cos(x^2+3xy) - (3x)(2x+3y)\operatorname{sen}(x^2+3xy),$$

$$D_{(2,2)} f(x,y) = -9x^2 \operatorname{sen}(x^2+3xy).$$

Evaluando en $(0,0)$ se tiene

$$H_f(0,0) = \begin{pmatrix} 2 & 3 \\ 3 & 0 \end{pmatrix},$$

por tanto,

$$P_2(x,y) = f(0,0) + \big(\nabla f(0,0)|(x,y)\big) + \frac{1}{2}(x,y)H_f(0,0)\begin{pmatrix} x \\ y \end{pmatrix} =$$

$$= x^2 + 3xy$$

Ahora calculamos las derivadas parciales de g.

$$D_1g(x,y) = \frac{y\frac{1}{(1+x^2y^2)} - 2x\arctan(xy)}{(1+x^2+y^2)^2}, \quad D_2g(x,y) = \frac{x\frac{1}{(1+x^2y^2)} - 2y\arctan(xy)}{(1+x^2+y^2)^2},$$

Evaluando en 0, se tiene $D_1g(0,0) = D_2g(0,0) = 0$. Para calcular las derivadas parciales de segundo orde usamos la definición

$$D_{(1,1)}g(0,0) = \lim_{t\to 0} \frac{D_1g(t,0) - D_1g(0,0)}{t} = 0,$$

$$D_{(1,2)}g(0,0) = \lim_{t\to 0} \frac{D_1g(0,t) - D_1g(0,0)}{t} = 1,$$

$$D_{(2,2)}g(0,0) = \lim_{t\to 0} \frac{D_2g(0,t) - D_2g(0,0)}{t} = 0.$$

El Hessiano de g en $(0,0)$ es

$$Hg(0,0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

y el polinomio de Taylor de orden 2

$$Q_2(x,y) = g(0,0) + \big(\nabla g(0,0)|(x,y)\big) + \frac{1}{2}(x,y)H_g(0,0)\begin{pmatrix} x \\ y \end{pmatrix} = xy$$

Hacemos los cálculos para h. Primero evaluamos la función, $h(1,1) = \log 2$.

$$D_1h(x,y) = \frac{2x}{x^2+y^2}, \quad D_2h(x,y) = \frac{2y}{x^2+y^2}.$$

Por tanto $\nabla h(1,1) = (1,1)$.

$$D_{(1,1)}h(1,1) = \lim_{t\to 0} \frac{D_1h(1+t,1) - D_1(1,1)}{t} = \lim_{t\to 0} \frac{\frac{2(1+t)}{(1+t)^2+1} - 1}{t} = 0.$$

$$D_{(2,1)}h(1,1) = \lim_{t\to 0} \frac{D_1h(1,1+t) - D_1(1,1)}{t} = \lim_{t\to 0} \frac{\frac{2}{(1+t)^2+1} - 1}{t} = -1.$$

Como $h(x,y) = h(y,x)$, entonces $D_{(2,2)}h(1,1) = D_{(1,1)}h(1,1) = 0$, así la matriz hessiana en $(1,1)$ es

$$H_h(1,1) = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$$

y el polinomio de Taylor de orden 2

$$R_2(x,y) = h(1,1) + \big(\nabla h(1,1)|(x-1,y-1)\big) + \frac{1}{2}(x-1,y-1)H_h(1,1)\begin{pmatrix} x-1 \\ y-1 \end{pmatrix} =$$

$$= \log 2 + (x-1) + (y-1) - (x-1)(y-1).$$

5.4 Sabemos que

$$\lim_{(x,y)\longrightarrow(0,0)} \frac{f(x,y) - P_2(x,y)}{\|(x,y)\|^2} = 0,$$

siendo P_2 el polinomio de Taylor de orden 2 de f en $(0,0)$ y

$$f(x,y) = e^{x \operatorname{sen} y}.$$

Basta pues calcular las derivadas parciales de orden menor o igual que 2 y evaluarlas en $(0,0)$. Se tiene que

$$D_1 f(x,y) = \operatorname{sen} y\, e^{x \operatorname{sen} y}, \quad D_2 f(x,y) = x \cos y\, e^{x \operatorname{sen} y},$$

$$D_{(1,1)} f(x,y) = e^{x \operatorname{sen} y} \operatorname{sen}^2 y, \quad D_{(1,2)} f(x,y) = e^{x \operatorname{sen} y} \cos y\big(1 + x \operatorname{sen} y\big),$$

$$D_{(2,2)} f(x,y) = x\, e^{x \operatorname{sen} y}\big(-\operatorname{sen} y + x \cos^2 y\big),$$

Evaluando en $(0,0)$ la función y las derivadas parciales de orden menor o igual que dos, se obtienen los siguientes valores

f	D_1	D_2	$D_{(1,1)}$	$D_{(1,2)}$	$D_{(2,2)}$
1	0	0	0	1	0

Por tanto,

$$P_2(x,y) = 1 + xy.$$

El enunciado es entonces consecuencia del Teorema 5.8.

5.5 Por hipótesis, sabemos que $D^2 f$ es constante en A, la llamamos B, por tanto $B \in \mathscr{L}^2(\mathbb{R}^N, \mathbb{R}^M)$. Fijamos un elemento $a \in A$, llamamos $L = Df(a)$ y definimos el campo vectorial en A dado por

$$g(x) = f(x) - \Big(f(a) + L(x-a) + \frac{1}{2}B(x-a, x-a)\Big) \quad (x \in A)$$

que toma valores en $\mathbb{R}^M$. Como f es diferenciable, L lineal y B bilineal, entonces g es diferenciable en A. Además, derivando y usando que $B \equiv D^2 f(a)$ es simétrica (Teorema de Schwarz) tenemos

$$Dg(x) = Df(x) - L - B(x-a) \quad (x \in A)$$

Como Dg es derivable, usando que el segundo sumando es constante y el último es afín, tenemos

$$D^2g(x) = D^2f(x) - B = 0.$$

En vista de que A es abierto y conexo, el Corolario 4.8 nos asegura que Dg es constante y, por ser $Dg(a) = 0$, sabemos que Dg es idénticamente cero. Aplicando de nuevo el Corolario 4.8 obtenemos que g es constante y basta evaluar en a para terminar el problema. Obtenemos entonces que se verifica el enunciado para

$$b = f(a) - Df(a)(a) + \frac{1}{2}D^2f(a)(a,a),$$

$$T(x) = Df(a)(x) - D^2f(a)(a,x),$$

$$S(x,x) = \frac{1}{2}D^2f(a)(x,x),$$

se obtiene el enunciado.

5.6 Como para cualesquiera $x,y \in \mathbb{R}^N$ se tiene que

$$\|x+y\|_2^2 = \big(x+y \mid x+y\big) = \|x\|_2^2 + \|y\|_2^2 + 2\big(x \mid y\big),$$

si T es un operador lineal que conserva la norma, también conserva el producto escalar, es decir,

$$\big(T(x) \mid T(y)\big) = \big(x \mid y\big), \quad \forall x,y \in \mathbb{R}^N. \qquad (*)$$

Si escribimos $T = (T_1,\ldots,T_N)$, $y_j = T_j(x)$, la regla de la cadena para las derivadas parciales nos da para $i = 1,\ldots,N$

$$\frac{\partial(f\circ T)}{\partial x_i}(x) = \sum_{j=1}^{N} \frac{\partial f}{\partial y_j}(T(x))\frac{\partial T_j}{\partial x_i}(x).$$

Si $\{e_1,\ldots,e_N\}$ es la base canónica de $\mathbb{R}^N$, se tiene que

$$\frac{\partial T_j}{\partial x_i}(x) = D_iT_j(x) = DT_j(x)(e_i) = T_j(e_i),$$

donde se ha utilizado que T_j es lineal.

Derivando otra vez se tiene

$$\frac{\partial^2(f\circ T)}{\partial x_i^2}(x) = \sum_{j=1}^{N}\sum_{k=1}^{N} \frac{\partial^2 f}{\partial y_k \partial y_j}(T(x))\frac{\partial T_k}{\partial x_i}(x)T_j(e_i)$$

es decir,

$$\frac{\partial^2(f\circ T)}{\partial x_i^2}(x) = \sum_{j=1}^{N}\sum_{k=1}^{N} \frac{\partial^2 f}{\partial y_k \partial y_j}(T(x))T_k(e_i)T_j(e_i),$$

y por tanto,

$$\sum_{i=1}^{N}\frac{\partial^2(f\circ T)}{\partial x_i^2}(x) = \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N} \frac{\partial^2 f}{\partial y_k \partial y_j}(T(x))T_k(e_i)T_j(e_i) =$$

$$(5.10.3) \qquad = \sum_{j=1}^{N}\sum_{k=1}^{N}\frac{\partial^2 f}{\partial y_k \partial y_j}(T(x))\sum_{i=1}^{N}T_k(e_i)T_j(e_i).$$

Hemos probado antes que T conserva el producto escalar, luego

$$\big(T(e_k) \mid T(e_j)\big) = \big(e_k \mid e_j\big) = \delta_{kj},$$

esto es, si M es la matriz asociada a T, la igualdad anterior nos dice que el producto de la fila k-ésima de M^t por la columna j-ésima de M es δ_{kj}, luego

$$M^t M = I \quad \Rightarrow MM^t = I.$$

Traduciendo de nuevo esta última condición tenemos

$$\sum_{i=1}^{N}T_k(e_i)T_j(e_i) = \big((T_k(e_1),\ldots,T_k(e_N))|(T_j(e_1),\ldots,T_j(e_N))\big) = \delta_{kj}.$$

Sustituyendo en 5.10.3 tenemos

$$\sum_{i=1}^{N}\frac{\partial^2(f\circ T)}{\partial x_i^2}(x) = \sum_{j=1}^{N}\frac{\partial^2 f}{\partial y_j^2}(Tx).$$

Otra demostración alternativa usando la regla de la cadena para la derivada segunda:

$$D^2(f\circ T)(x)(u,v) =$$

$$D^2 f(T(x))\big(DT(x)(u),DT(x)(v)\big) + Df(x)\big(D^2T(x)(u,v)\big) =$$

$$D^2 f(T(x))\big(T(u),T(v)\big) + 0 = D^2 f(T(x))\big(T(u),T(v)\big),$$

y en consecuencia

$$D_{(i,i)}(f\circ T)(x) = D^2(f\circ T)(x)(e_i,e_i) = D^2 f(T(x))\big(T(e_i),T(e_i)\big) =$$

$$T(e_i)H_f(T(x))T(e_i)^t = e_i\big(M^t H_f(T(x))M\big)e_i^t,$$

donde, como antes, M es la matriz asociada a T, esto es

$$T(x)^t = Mx^t.$$

Así

$$\Delta(f\circ T)(x) = \sum_{i=1}^{N}D_{(i,i)}(f\circ T)(x) = \text{ traza } \big(M^t H_f(T(x))M\big).$$

Análogamente

$$\Delta f(T(x)) = \sum_{i=1}^{N}D_{(i,i)}f(T(x)) = \sum_{i=1}^{N}D^2 f(T(x))(e_i,e_i) =$$

$$\sum_{i=1}^{N} e_i H_f(x) e_i^t = \text{ traza } \big(H_f(T(x))\big).$$

Al ser T una isometría, hemos probado que

$$MM^t = I.$$

Finalmente,

$$\Delta(f \circ T)(x) = \text{ traza } \big(M^t H_f(T(x))M\big) = \text{ traza } \big(H_f(T(x))\big) = \Delta f(T(x)),$$

donde se ha utilizado que

$$A, B \in \mathscr{M}_{N\times N}(\mathbb{R}) \Rightarrow \text{ traza } (AB) = \text{ traza } (BA).$$

5.7 Notamos, por comodidad,

$$g(x) = f\big(\|x\|_2\big), \quad \forall x \in \mathbb{R}^N \backslash \{0\}.$$

Se tiene para $i = 1, \ldots, N$

$$D_i g(x) = f'\Big(\|x\|_2\Big) \frac{x_i}{\|x\|_2},$$

por tanto,

$$D_{(i,i)} g(x) = f''\Big(\|x\|_2\Big) \frac{x_i^2}{\|x\|_2^2} + f'\Big(\|x\|_2\Big) \; \frac{\|x\|_2^2 - x_i^2}{\|x\|_2^3}$$

de donde se deduce que

$$\Delta g(x) = f''\Big(\|x\|_2\Big) + f'\Big(\|x\|_2\Big) \frac{(N-1)\|x\|_2^2}{\|x\|_2^3} =$$

$$f''\Big(\|x\|_2\Big) + f'\Big(\|x\|_2\Big) \frac{(N-1)}{\|x\|_2}.$$

Por tanto, las soluciones de $\Delta g(x) = 0$ verifican la ecuación diferencial

$$t\ f''(t) + (N-1) f'(t) = 0.$$

Supuesto que $f'(t) > 0$ para un cierto valor de $t > 0$ se tiene, que la función

$$h(t) = \log f'(t)$$

verifica $h'(t) = \frac{1-N}{t}$. Integrando, obtenemos

$$h(t) = (1-N)\log(t) + A.$$

Por tanto,

$$f'(t) = Bt^{1-N},$$

para cierta constante B y así, si $N = 2$, entonces

$$f(t) = B\log(t) + K,$$

mientras que si $N > 2$, se tiene

$$f(t) = Ct^{2-N} + K,$$

donde C y K son constantes arbitrarias.

Tema 6

Derivadas sucesivas.

Sabemos que la derivada de una función es una aplicación lineal, la derivada segunda es bilineal y ocurre que la derivada k-ésima es una aplicación k-lineal. Antes de nada, presentaremos la notación que usaremos para estas aplicaciones.

Definición 6.1. Si k es un natural, notaremos por

$$\mathscr{L}^k(\mathbb{R}^N,\mathbb{R}^M) \equiv \mathscr{L}^k(\mathbb{R}^N \times \overset{k}{\cdots} \times \mathbb{R}^N,\mathbb{R}^M)$$

al conjunto de las aplicaciones de $\mathbb{R}^N \times \overset{k}{\cdots} \times \mathbb{R}^N$ en $\mathbb{R}^M$ que son k-lineales, esto es, lineales en cada variable. Este conjunto es un espacio vectorial definiendo las operaciones de forma puntual y es fácil probar que se puede identificar

$$\mathscr{L}^{m+n}(\mathbb{R}^N,\mathbb{R}^M) \equiv \mathscr{L}^m(\mathbb{R}^N,\mathscr{L}^n(\mathbb{R}^N,\mathbb{R}^M))$$

de la siguiente forma natural

$$T\left(x_1,\dots,x_{m+n}\right) = T\left(x_1,\dots,x_m\right)\left(x_{m+1},\dots,x_{m+n}\right) \quad \forall x_1,\dots,x_{m+n} \in \mathbb{R}^N.$$

En particular, si $k > 1$, se tiene

$$\mathscr{L}^k(\mathbb{R}^N,\mathbb{R}^M) \equiv \mathscr{L}(\mathbb{R}^N,\mathscr{L}^{k-1}(\mathbb{R}^N,\mathbb{R}^M)).$$

El espacio $\mathscr{L}^k(\mathbb{R}^N,\mathbb{R}^M)$ se puede normar si definimos

$$\|T\| = \max\{\|T(x_1,x_k,\dots,x_k)\| : x_i \in \mathbb{R}^N, \|x_i\| = 1, 1 \le i \le k\} \qquad (T \in \mathscr{L}^k(\mathbb{R}^N,\mathbb{R}^M)).$$

La norma verifica propiedades análogas a las que aparecen en la Proposición 5.1.

Definición 6.2 (Función k veces derivable)**.** Sea $A \subset \mathbb{R}^N$ y $f : A \longrightarrow \mathbb{R}^M$ una función y k un natural mayor que 1. Sea $A_{k-1} \subset A$ el conjunto de puntos donde f es $k-1$ veces derivable y sea $D^{k-1}f : A_{k-1} \longrightarrow \mathscr{L}^{k-1}(\mathbb{R}^N, \mathbb{R}^M)$ la aplicación derivada $(k-1)$-ésima de f. Se dice que f es <u>k veces derivable</u> en $a \in \overset{\circ}{A}_{k-1}$ si la función $D^{k-1}f$ es derivable en dicho punto, en cuyo caso a la aplicación

$$D(D^{k-1}f)(a) \in \mathscr{L}(\mathbb{R}^N, \mathscr{L}^{k-1}(\mathbb{R}^N, \mathbb{R}^M)) \equiv \mathscr{L}^k(\mathbb{R}^N, \mathbb{R}^M)$$

se llama <u>*derivada k-ésima*</u> de f en a y se nota $D^k f(a)$. A $D^k f(a)$ se le asocia la aplicación $d^k f(a) : \mathbb{R}^N \longrightarrow \mathbb{R}^M$ definida por:

$$d^k f(a)(x) := D^k f(a)(x, \overset{k}{\ldots}, x), \quad \forall x \in \mathbb{R}^N.$$

Se dice que f es k veces derivable en un subconjunto $B \subset A$ si es k veces derivable en cada punto de B.

Sea $A_k \subset A$ el conjunto de puntos de A donde f es k veces derivable. La aplicación

$$x \longrightarrow D^k f(x)$$

de A_k en $\mathscr{L}^k(\mathbb{R}^N, \mathbb{R}^M)$ se denomina la <u>*aplicación derivada k-ésima*</u> de f y se nota $D^k f$.

Se dice que f es de <u>*clase $\mathscr{C}^k$*</u> en a, y se nota $f \in \mathscr{C}^k(a)$, si es k veces derivable en un entorno de a y la función $D^k f$ es continua en a. Se dice que f es de clase $\mathscr{C}^k$ en un subconjunto $B \subset A$ si es de clase $\mathscr{C}^k$ en cada punto de B.

Se dice que f es k veces derivable (resp. $\mathscr{C}^k$) cuando lo sea en todos los puntos de su conjunto de definición (que necesariamente será abierto). Notamos por $\mathscr{C}^k(A)$ al conjunto de las funciones de clase $\mathscr{C}^k$ en el abierto A.

Finalmente se dice que f es de clase $\mathscr{C}^\infty$ en $a \in A$ (resp. $B \subset A$) si $f \in \mathscr{C}^k(a)$ (resp. $f \in \mathscr{C}^k(B)$), $\forall k \in \mathbb{N}$.

Si $m, n \in \mathbb{N}$, entonces la identificación de los espacios $\mathscr{L}^{m+n}(\mathbb{R}^N, \mathbb{R}^M)$ y $\mathscr{L}^m(\mathbb{R}^N, \mathscr{L}^n(\mathbb{R}^N, \mathbb{R}^M))$ nos permite comprobar (mediante una sencilla inducción) que f es $m+n$ veces derivable en un punto a si, y sólo si, es n veces derivable en un entorno de a y la función $D^n f$ es m veces derivable en a, en cuyo caso

$$D^{m+n} f(a) \equiv D^m(D^n f)(a).$$

A partir de ahora presentamos para la derivada k-ésima los resultados obtenidos en el tema anterior para la derivada segunda. Conviene tener presente que el paso de la derivada $(k-1)$-ésima a la derivada k-ésima es idéntico al paso de la derivada primera a la derivada segunda.

Nota 6.3. Sea $A \subset \mathbb{R}$, $f : A \longrightarrow \mathbb{R}^M$ un campo vectorial y $k > 1$. Supongamos que f es k veces derivable en un punto $a \in A$. Entonces, la relación entre la derivada de f y la derivada elemental viene dada por

$$D^k f(a)(s_1, \ldots, s_k) = s_1 \ldots s_k f^{(k)}(a), \quad \forall s_1, \ldots, s_k \in \mathbb{R},$$

equivalentemente

$$f^{(k)}(a) = d^k f(a)(1).$$

Sabemos que la diferencial segunda es una aplicación bilineal simétrica (Teorema 5.6). También es cierto el resultado análogo para la diferencial k-ésima:

Teorema 6.4 (Schwarz)**.** *Sea $A \subset \mathbb{R}^N$, $f : A \longrightarrow \mathbb{R}^M$ un campo vectorial y $k > 1$. Supongamos que f es k veces derivable en un punto $a \in A$. Entonces la aplicación k-lineal $D^k f(a)$ es simétrica, esto es,*

$$D^k f(a)\big(x_1, \ldots, x_k\big) = D^k f(a)\big(x_{\sigma(1)}, \ldots, x_{\sigma(k)}\big)$$

para cualesquiera $x_1, \ldots, x_k \in \mathbb{R}^N$ y cualquier permutación σ del conjunto $\{1, \ldots, k\}$.

Demostración:

Sabemos que para $k = 2$ es cierto (Teorema 5.6). Para probar el caso general, basta razonar por inducción sobre k. Sea $k > 2$ y supongamos cierto el teorema de Schwarz para funciones $k-1$ veces derivables en un punto. Sean $x_1, \ldots, x_k \in \mathbb{R}$, por hipótesis de inducción sabemos que

$$D^2(D^{k-2} f)(a)\big(x_1, x_2\big) = D^2(D^{k-2} f)(a)\big(x_2, x_1\big)$$

y en consecuencia

$$D^k f(a)\big(x_1, x_2, x_3, \ldots, x_k\big) = D^k f(a)\big(x_2, x_1, x_3, \ldots, x_k\big).$$

Para probar que si $2 \le i < j \le k$, entonces

$$D^k f(a)\big(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_k\big) = D^k f(a)\big(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_k\big),$$

basta tener en cuenta que si x está en un cierto entorno de a se verifica

$$D^{k-1} f(x)\big(x_2, \ldots, x_i, \ldots, x_j, \ldots, x_k\big) =$$

$$D^{k-1} f(x)\big(x_2, \ldots, x_j, \ldots, x_i, \ldots, x_k\big), \quad \forall x \in U,$$

esto es,

$$D^{k-1} f = E_{(i,j)} \circ D^{k-1} f,$$

donde $E_{(i,j)}$ es la aplicación lineal en $\mathbb{R}^N$ que intercambia las componentes i-ésima y j-ésima. Basta usar la regla de la cadena y el hecho de que una permutación es composición de trasposiciones para terminar la demostración. ∎

Definición 6.5 (Polinomio de Taylor de orden k)**.** Sea $A \subset \mathbb{R}^N$ y $f : A \longrightarrow \mathbb{R}^M$ una función. Supongamos que f es k veces derivable en un punto $a \in A$, entonces la función $P_k : \mathbb{R}^N \longrightarrow \mathbb{R}^M$ definida por

$$P_k(x) = f(a) + df(a)(x-a) + \frac{1}{2} d^2 f(a)(x-a) + \cdots + \frac{1}{k!} d^k f(a)(x-a)$$

se llama *polinomio de Taylor de orden k* de f en a.

El Problema 6.6 afirma que si f es un polinomio de grado k, entonces, el polinomio de Taylor de orden k en cualquier punto coincide con f. Primero aprenderemos a derivar polinomios homogéneos.

Lema 6.6. *Sea $T \in \mathscr{L}^k(\mathbb{R}^N, \mathbb{R}^M)$ una aplicación simétrica. Entonces la aplicación $P : \mathbb{R}^N \longrightarrow \mathbb{R}^M$ definida por:*

$$P(x) := T(x, \overset{k}{\cdots}, x), \quad \forall x \in \mathbb{R}^N,$$

es derivable. Además, para cada $a \in \mathbb{R}^N$ se tiene

$$DP(a)(x) = kT(a, \overset{k-1}{\ldots}, a, x), \quad \forall x \in \mathbb{R}^N.$$

Demostración:
Fijemos $a \in \mathbb{R}^N$, entonces, para cada $x \in \mathbb{R}^N \backslash \{a\}$ se verifica que

$$P(x) = T(x, \overset{k}{\ldots}, x) = T(a + (x-a), \overset{k}{\ldots}, a + (x-a)) =$$

$$= \sum_{i=0}^{k} \binom{k}{i} T(a, \overset{k-i}{\ldots}, a, x-a, \overset{i}{\ldots}, x-a) =$$

$$= P(a) + kT(a, \overset{k-1}{\ldots}, a, x-a) + \sum_{i=2}^{k} \binom{k}{i} T(a, \overset{k-i}{\ldots}, a, x-a, \overset{i}{\ldots}, x-a),$$

y por tanto

$$\frac{P(x) - P(a) - kT(a, \overset{k-1}{\ldots}, a, x-a)}{\|x-a\|} = \sum_{i=2}^{k} \binom{k}{i} \frac{T(a, \overset{k-i}{\ldots}, a, x-a, \overset{i}{\ldots}, x-a)}{\|x-a\|}.$$

Como para todo $i \in \{2, \ldots, k\}$ se tiene que

$$\|T(a, \overset{k-i}{\ldots}, a, x-a, \overset{i}{\ldots}, x-a)\| \leq \|T\| \, \|a\|^{k-i} \|x-a\|^i,$$

se sigue que

$$\lim_{x \to a} \frac{T(a, \overset{k-i}{\ldots}, a, x-a, \overset{i}{\ldots}, x-a)}{\|x-a\|} = 0,$$

lo que termina la demostración. ■

Comprobaremos ahora que el polinomio de Taylor de orden k de una función f en un punto a tiene sus k primeras derivadas en a iguales a las de la función.

Proposición 6.7. *Sea $A \subset \mathbb{R}^N$ y $f : A \longrightarrow \mathbb{R}^M$ un campo vectorial. Supongamos que f es k veces derivable en un punto $a \in A$ y que $P_k : \mathbb{R}^N \longrightarrow \mathbb{R}^M$ es el polinomio de Taylor de orden k de f en a, esto es,*

$$P_k(x) = f(a) + df(a)(x-a) + \frac{1}{2} d^2 f(a)(x-a) + \ldots + \frac{1}{k!} d^k f(a)(x-a).$$

Entonces P_k es de clase $\mathscr{C}^\infty$ y se verifica

$$DP_k(x) = Df(a) + d(Df)(a)(x-a) + \frac{1}{2} d^2(Df)(a)(x-a) + \ldots +$$

$$+ \frac{1}{(k-1)!} d^{k-1}(Df)(a)(x-a).$$

Además,

$$D^n P_k(a) = D^n f(a) \quad (n \leq k) \quad \text{y} \quad D^n P_k(x) = 0, \quad \forall n > k.$$

Demostración:
Basta probar que DP_k es el polinomio de Taylor de la función Df en el punto a. Para cada $m \in \{2, \ldots, k\}$, la función $\varphi : \mathbb{R}^N \longrightarrow \mathbb{R}^M$ dada por

$$\varphi(x) = d^m f(a)(x-a)$$

es composición de la función $g : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ dada por $g(x) = x - a$ y de la función $d^m f(a) : \mathbb{R}^N \longrightarrow \mathbb{R}^M$. Luego, el lema anterior y la regla de la cadena, permiten afirmar que φ es derivable y que para cada $x \in \mathbb{R}^N$ se verifica

$$D\varphi(x) = D(d^m f(a) \circ g)(x) = D(d^m f(a))(g(x)) \circ Dg(x) =$$
$$= D(d^m f(a))(x-a) \circ Id_{\mathbb{R}^N} = D(d^m f(a))(x-a) =$$
$$= mD^m f(a)(x-a, \overset{m-1}{\ldots}, x-a, \cdot) =$$
$$= mD^{m-1}(Df)(a)(x-a, \overset{m-1}{\ldots}, x-a)(\cdot) = md^{m-1}(Df)(a)(x-a),$$

donde se ha utilizado la expresión de la derivada que da el lema anterior. De aquí se sigue inmediatamente que P_k es derivable y su derivada viene dada por

$$DP_k(x) = Df(a) + d(Df)(a)(x-a) + \frac{1}{2}d^2(Df)(a)(x-a) + \ldots + \frac{1}{(k-1)!}d^{k-1}(Df)(a)(x-a),$$

que es el polinomio de Taylor de Df en a de orden $k-1$. En particular, $DP_k(a) = Df(a)$. El resultado se sigue aplicando el mismo argumento. ∎

Teorema 6.8 (Fórmula infinitesimal del resto)**.** *Sea $A \subset \mathbb{R}^N$ y $f : A \longrightarrow \mathbb{R}^M$ un campo vectorial. Supongamos que f es k veces derivable en un punto $a \in A$, entonces se verifica*

$$\lim_{x \to a} \frac{f(x) - P_k(x)}{\|x-a\|^k} = 0.$$

Demostración:
Razonamos por inducción sobre k, por lo que supondremos conocida la fórmula infinitesimal del resto para funciones $k-1$ veces derivables en un punto.

Dado un positivo ε, como f es k veces derivable en a, existe $\delta > 0$ tal que

$$\|x-a\| < \delta \Rightarrow \begin{cases} x \in A \\ f \text{ es } k-1 \text{ veces derivable en } x \\ \|Df(x) - Q_{k-1}(x)\| \leq \varepsilon \|x-a\|^{k-1} \end{cases} \tag{6.0.1}$$

donde Q_{k-1} es el polinomio de Taylor de orden $k-1$ de Df en a, esto es

$$Q_{k-1}(x) = Df(a) + d(Df)(a)(x-a) + \ldots + \frac{1}{(k-1)!}d^{k-1}(Df)(a)(x-a).$$

La demostración se concluye probando que para cada $x \in B(a, \delta) \backslash \{a\}$ se verifica que

$$\|f(x) - P_k(x)\| \leq \varepsilon \|x-a\|^k,$$

lo cual será consecuencia del Teorema del valor medio aplicado a la función $g : B(a,\delta) \longrightarrow \mathbb{R}^M$ definida por

$$g(z) = f(z) - P_k(z), \quad \forall z \in B(a,\delta)$$

en el segmento $[a,x]$. En efecto, por el Teorema 4.5 y la proposición anterior tenemos

$$\|g(x) - g(a)\| \leq \sup\{\|Dg(z)\| : z \in]a,x[\}\|x-a\| \leq$$
$$= \sup\{\|Df(z) - Q_{k-1}(z)\| : z \in]a,x[\}\ \|x-a\| =$$
$$= \sup\{\varepsilon\|z-a\|^{k-1} : z \in]a,x[\} = \varepsilon\|x-a\|^k.$$

■

Ejemplos 6.9.

1. Toda función constante f de $\mathbb{R}^N$ en $\mathbb{R}^M$ es de clase $\mathscr{C}^\infty$ con $D^k f(a) = 0, \ \ \forall a \in \mathbb{R}^N, \forall k \in \mathbb{N}$.

2. Toda aplicación lineal T de $\mathbb{R}^N$ en $\mathbb{R}^M$ es de clase $\mathscr{C}^\infty$ con
$$D^k T(a) = 0, \quad \forall a \in \mathbb{R}^N, \quad \forall k \geq 2.$$

3. Toda aplicación bilineal T de $\mathbb{R}^N \times \mathbb{R}^M$ en en $\mathbb{R}^P$ es de clase $\mathscr{C}^\infty$ con
$$D^k T(a,b) = 0, \quad \forall (a,b) \in \mathbb{R}^N \times \mathbb{R}^M, \quad \forall k \geq 3.$$

4. La aplicación suma $\sigma : \mathbb{R}^N \times \mathbb{R}^N \longrightarrow \mathbb{R}^N$ y la aplicación producto por escalares $\pi : \mathbb{R} \times \mathbb{R}^N \longrightarrow \mathbb{R}^N$ son de clase $\mathscr{C}^\infty$.

5. La aplicación inversión $J : \text{Iso}\ (\mathbb{R}^N) \longrightarrow \mathscr{L}(\mathbb{R}^N)$ dada por
$$J(T) = T^{-1}, \quad T \in \text{Iso}\ (\mathbb{R}^N)$$
es de clase $\mathscr{C}^\infty$.

   Escribiendo la aplicación derivada primera de J como ya hicimos en el Ejemplo 5.4.5, $DJ = \phi \circ (J,J)$, el resultado se sigue de la regla de la cadena (véase Sección 5.8.2) y del Ejemplo 6.9.3.

6.1. Reglas de derivación.

Por la forma de definir la diferencial k-ésima, ésta verifica análogas propiedades a las de la derivada segunda.

1. **Linealidad.** El conjunto de las funciones k veces derivables en un punto a es un espacio vectorial en el que la aplicación
$$f \mapsto D^k f(a)$$
es lineal.

2. **Regla de la cadena.** La composición de funciones de clase k es una función de clase k. La afirmación análoga es cierta para las funciones k veces derivables.

3. **Carácter local de la derivada k-ésima.** Un campo vectorial es k veces derivable en un punto a si la restricción de f a un entorno del punto a es k veces derivable, y en tal caso, las derivadas de orden k de ambas funciones coinciden.

4. **Derivación de la función inversa.** Sean A y B subconjuntos abiertos de $\mathbb{R}^N$ y f un homeomorfismo de A sobre B. Si f es de clase $\mathscr{C}^k$ en $a \in A$ y $Df(a) \in \text{Iso}\,(\mathbb{R}^N)$, entonces f^{-1} es de clase $\mathscr{C}^k$ en $f(a)$. Para comprobar esta afirmación, basta escribir la aplicación derivada primera de f^{-1} como en la Sección 5.4, apartado iv), esto es, $Df^{-1} = J \circ Df \circ f^{-1}$, y el resultado se sigue entonces usando inducción y la regla de la cadena (apartado 2).

5. Sean $A \subset \mathbb{R}^N, f : A \longrightarrow \mathbb{R}^M$ un campo vectorial y $T \in \mathscr{L}(\mathbb{R}^M, \mathbb{R}^P)$. Si f es k veces derivable en un punto $a \in A$, entonces $T \circ f$ es k veces derivable en a con
$$D^k(T \circ f)(a) = T \circ D^k f(a)$$
Por tanto, si $f \in \mathscr{C}^k(a)$, entonces $T \circ f \in \mathscr{C}^k(a)$.

6. **Reducción del estudio de la derivabilidad a campos escalares.** Sea $A \subset \mathbb{R}^N$ y $f : A \longrightarrow \mathbb{R}^M$ un campo vectorial. Entonces f es k veces derivable en $a \in A$ si, y sólo si, cada función componente f_i es k veces derivable en a para $i = 1, \ldots, M$, en cuyo caso
$$D^k f(a) = (D^k f_1(a), \ldots, D^k f_M(a)).$$

Teorema 6.10 (Fórmula de Taylor con resto de Lagrange para campos escalares)**.** *Sea $A \subset \mathbb{R}^N$, $a, x \in A$ con $a \neq x$, tales que el segmento $[a,x]$ está incluido en A. Si $f : A \longrightarrow \mathbb{R}$ es una función de clase $\mathscr{C}^k$ en $[a,x]$ y $k+1$ veces derivable en $]a,x[$, entonces existe $c \in]a,x[$ tal que*
$$f(x) = P_k(x) + \frac{1}{(k+1)!} d^{k+1} f(c)(x-a),$$
donde $P_k(x)$ es el polinomio de Taylor de orden k de f en a, esto es,
$$P_k(x) = f(a) + df(a)(x-a) + \frac{1}{2} d^2 f(a)(x-a) + \ldots + \frac{1}{k!} d^k f(a)(x-a).$$

Demostración:
Para demostrar este resultado aplicaremos la fórmula de Taylor clásica a la función auxiliar $\sigma : [0,1] \longrightarrow \mathbb{R}$ definida por
$$\sigma(t) = f(a + t(x-a)).$$
Sabemos que
$$\sigma'(t) = df(a + t(x-a))(x-a), \quad \forall t \in [0,1],$$

y usando la regla de al cadena es fácil probar que

$$\sigma^{(m)}(t) = d^m f(a+t(x-a))(x-a).$$

Finalmente basta tener en cuenta que existe $t_0 \in]0,1[$ tal que

$$\sigma(1) = \sigma(0) + \frac{1}{2}\sigma'(0) + \ldots + \frac{1}{k!}\sigma^{(k)}(0) + \frac{1}{(k+1)!}\sigma^{(k+1)}(t_0),$$

de donde se deduce el resultado. ∎

Teorema 6.11 (Fórmula de Taylor)**.** *Sean $A \subset \mathbb{R}^N$ y $a, x \in A$ con $a \neq x$, tales que el segmento $[a,x]$ está incluido en A. Si $f : A \longrightarrow \mathbb{R}^M$ es una función de clase $\mathscr{C}^k$ en $[a,x]$ y $k+1$ veces derivable en $]a,x[$, entonces*

$$\|f(x) - P_k(x)\| \leq \frac{\|x-a\|^{k+1}}{(k+1)!} \sup\{\|D^{k+1} f(z)\| : z \in]a,x[\},$$

donde $P_k(x)$ es el polinomio de Taylor de orden k de f en a, es decir,

$$P_k(x) = f(a) + df(a)(x-a) + \frac{1}{2}d^2 f(a)(x-a) + \ldots + \frac{1}{k!}d^k f(a)(x-a).$$

Demostración:
Supongamos que el conjunto

$$\{\|D^{k+1} f(z)\| : z \in]a,x[\}$$

está mayorado, ya que en otro caso, no hay nada que demostrar. Sea K el supremo de este conjunto. Para probar este resultado, basta aplicar la Proposición 4.6 a las funciones $\sigma : [0,1] \longrightarrow \mathbb{R}^M$ y $g : [0,1] \longrightarrow \mathbb{R}$ definidas por

$$\sigma(t) = f(a+t(x-a)) + (1-t)df(a+t(x-a))(x-a) + \ldots +$$

$$+ \frac{1}{k!}(1-t)^k d^k f(a+t(x-a))(x-a)$$

y

$$g(t) = \frac{-1}{(k+1)!} K \|x-a\|^{k+1} (1-t)^{k+1}.$$

∎

6.2. Derivadas de orden superior de campos escalares.

Definición 6.12 (Derivadas parciales sucesivas)**.** Sea $A \subset \mathbb{R}^N, a \in A$ y f un campo escalar en A. Sean $i_2, \ldots, i_k \in \{1, \ldots, N\}$ y $A^{(i_2,\ldots,i_k)} \subset A$ el conjunto de puntos donde f tiene derivada parcial de orden $k-1$ respecto de las variables $i_k, \ldots, i_2$. Si $i_1 \in \{1, \ldots, N\}$ y el campo escalar

$D_{(i_2,\ldots,i_k)}f$ en $A^{(i_2,\ldots,i_k)}$ tiene derivada parcial respecto de la variable i_1 en el punto a, entonces se dice que f tiene derivada parcial de orden k respecto de las variables $i_k,\ldots,i_1$ en el punto a y la derivada $D_{i_1}\big(D_{(i_2,\ldots,i_k)}f\big)(a)$ se nota $D_{(i_1,\ldots,i_k)}f(a)$, esto es

$$D_{(i_1,\ldots,i_k)}f(a) = D_{i_1}D_{i_2}\ldots D_{i_k}f(a).$$

Quedan así definidas inductivamente las N^k derivadas parciales k-ésimas del campo escalar f en el punto a.

Si $A^{(i_1,\ldots,i_k)} \subset A$ es el conjunto de puntos donde f tiene derivada parcial de orden k respecto de las variables $i_k,\ldots,i_1$, entonces el campo escalar en $A^{(i_1,\ldots,i_k)}$ definido por

$$x \longrightarrow D_{(i_1,\ldots,i_k)}f(x)$$

se denomina la aplicación derivada parcial de orden k de f respecto de las variables $i_k,\ldots,i_1$ y se nota $D_{(i_1,\ldots,i_k)}f$.

Teorema 6.13. *Sean $A \subset \mathbb{R}^N, a \in A$ y f un campo escalar en A y $k \geq 2$. Las siguientes afirmaciones son equivalentes:*

i) f es k veces derivable en a.

ii) f es $k-1$ veces derivable en un entorno de a y todas las derivadas parciales de orden $k-1$ son derivables en a.

En el caso de que se verifiquen las anteriores condiciones, se tiene que

$$D_{(i_1,\ldots,i_k)}f(a) = D^k f(a)\big(e_{i_1},\ldots,e_{i_k}\big)$$

para cualesquiera $i_1,\ldots,i_k \in \{1,\ldots,N\}$, donde $\{e_1,\ldots,e_N\}$ es la base canónica de $\mathbb{R}^N$. En consecuencia, a partir del Teorema de Schwarz, se verifica que

$$D_{(i_{\sigma(1)},\ldots,i_{\sigma(k)})}f(a) = D_{(i_1,\ldots,i_k)}f(a)$$

para cualesquiera $i_1,\ldots,i_k \in \{1,\ldots,N\}$ y cualquier permutación σ del conjunto $\{1,\ldots,k\}$.

Se tiene, por tanto,

$$D^k f(a)\big(x_1,\ldots,x_k\big) = \sum_{i_1,\ldots,i_k=1}^{N} D_{(i_1,\ldots,i_k)}f(a)\pi_{i_1}(x_1)\ldots\pi_{i_k}(x_k)$$

para cualesquiera $x_1,\ldots,x_k \in \mathbb{R}^N$, donde, para cada $i \in \{1,\ldots,N\}, \pi_i$ denota la i-ésima proyección de $\mathbb{R}^N$.

Además

$$f \in \mathscr{C}^k(a) \Leftrightarrow D_{(i_1,\ldots,i_{k-1})}f \in \mathscr{C}^1(a), \quad \forall i_1,\ldots,i_{k-1} \in \{1,\ldots,N\}.$$

Este resultado para $k=2$ es el Teorema 5.12. Ahora, un argumento inductivo permite probar el resultado para $k>2$.

Corolario 6.14. *Sean $A \subset \mathbb{R}^N$ un abierto, f un campo escalar en A y $k > 1$. Las siguientes afirmaciones son equivalentes:*

i) f es k veces derivable en A.

ii) Todas las derivadas parciales de orden $k-1$ son derivables en A.

Además,

$$f \in \mathscr{C}^k(A) \Leftrightarrow D_{(i_1,\ldots,i_{k-1})} f \in \mathscr{C}^1(A), \quad \forall i_1,\ldots,i_{k-1} \in \{1,\ldots,N\}.$$

Notación 6.15. Sean $A \subset \mathbb{R}^N$ y f un campo escalar en A que es k veces derivable en un punto $a \in A$.

La simetría que garantiza el Teorema 6.13 asegura que el orden en el que se efectúen las derivadas parciales sucesivas del campo escalar f en a es irrelevante. En consecuencia, las derivadas parciales de orden k se pueden reorganizar con el fin de simplificar la notación. Este hecho permite definir, para $k_1,\ldots,k_N \in \mathbb{N} \cup \{0\}$ tales que $k_1 + \cdots + k_N = k$

$$D^{(k_1,\ldots,k_N)} f(a) := D_1^{k_1} \ldots D_N^{k_N} f(a) = D_{(1,\overset{k_1}{\ldots},1,\ldots,N,\overset{k_N}{\ldots},N)} f(a).$$

Conviene observar que el número de derivadas parciales de orden k eventualmente distintas ha quedado reducido de N^k a

$$\frac{(N+k-1)!}{k!(N-1)!} \quad \left(RC_N^k = \binom{N+k-1}{k} \right).$$

Pongamos $a = (a_1,\ldots,a_N)$, entonces el <u>*polinomio de Taylor de orden k*</u> de f en a es el polinomio $P_k : \mathbb{R}^N \longrightarrow \mathbb{R}$ definido por

$$P_k(x) = f(a) + \big(\nabla f(a)|x-a\big) + \frac{1}{2}(x-a)H_f(a)(x-a)^t + \frac{1}{3!}d^3 f(a)(x-a) + \ldots +$$

$$+\frac{1}{k!}d^k f(a)(x-a),$$

que, de nuevo en virtud del Teorema 6.13, adquiere ahora el siguiente aspecto

$$f(a) + \sum_{r=1}^{k} \sum_{r_1+\ldots+r_N=r} \frac{1}{r_1!\ldots r_N!} D^{(r_1,\ldots,r_N)} f(a)(x_1-a_1)^{r_1} \ldots (x_N-a_N)^{r_N}$$

para cualquier $x = (x_1,\ldots,x_N) \in \mathbb{R}^N$, ya que el sumando

$$D^{(r_1,\ldots,r_N)} f(a)(x_1-a_1)^{r_1} \ldots (x_N-a_N)^{r_N}$$

se repite

$$\frac{r!}{r_1!\ldots r_N!} \quad \left(RP_r^{r_1,\ldots,r_N} \right)$$

veces.

Nota 6.16. Para obtener la fórmula de Taylor (6.10) para campos escalares, hay que sustituir los correspondientes polinomios de Taylor por la fórmula que acabamos de presentar y también $\frac{1}{(k+1)!}d^{k+1}f(c)(x-a)$ por

$$\sum_{k_1+\ldots+k_N=k+1} \frac{1}{k_1!\ldots k_N!} D^{(k_1,\ldots,k_N)} f(c)(x_1-a_1)^{k_1}\ldots(x_N-a_N)^{k_N},$$

expresión que sabemos tiene $\frac{(N+k)!}{(k+1)!(N-1)!}$ sumandos.

6.3. Referencias recomendadas.

[MaHo].

6.4. Resumen del resultados del Tema 6

Aplicaciones multilineales.

El conjunto $\mathscr{L}^{m+n}(\mathbb{R}^N,\mathbb{R}^M)$ *de las aplicaciones* $(m+n)$*-lineales de* $\mathbb{R}^N\times \overset{m+n}{\cdots} \times\mathbb{R}^N$ *en* $\mathbb{R}^M$ *se identifica con* $\mathscr{L}^m(\mathbb{R}^N,\mathscr{L}^n(\mathbb{R}^N,\mathbb{R}^M))$ *y para* $T\in\mathscr{L}^{m+n}(\mathbb{R}^N,\mathbb{R}^M)$*, la identificación viene dada por*

$$T\big(x_1,\dots,x_{m+n}\big)=T\big(x_1,\dots,x_m\big)\big(x_{m+1},\dots,x_{m+n}\big)\quad \forall x_1,\dots,x_{m+n}\in\mathbb{R}^N.$$

En particular, si $k>1$*, se tiene*

$$\mathscr{L}^k(\mathbb{R}^N,\mathbb{R}^M)\equiv\mathscr{L}(\mathbb{R}^N,\mathscr{L}^{k-1}(\mathbb{R}^N,R^M)).$$

La norma de $\mathscr{L}^k(\mathbb{R}^N,\mathbb{R}^M)$ *viene dada por*

$$\|T\|=\max\{\|T(x_1,x_k,\dots,x_k)\|:x_i\in\mathbb{R}^N,\|x_i\|=1,1\le i\le k\}\qquad (T\in\mathscr{L}^k(\mathbb{R}^N,\mathbb{R}^M)).$$

Campos vectoriales k veces derivables.
Sean $A\subset\mathbb{R}^N, f:A\longrightarrow\mathbb{R}^M$ y sea $A_1\subset A$ el conjunto de puntos donde f es derivable. Se dice que f es dos veces derivable en $a\in A_1$ si la función $Df:A_1\longrightarrow\mathscr{L}(\mathbb{R}^N,\mathbb{R}^M)$ es derivable en a, en cuyo caso la aplicación

$$D(Df)(a)\in\mathscr{L}(\mathbb{R}^N,\mathscr{L}(\mathbb{R}^N,\mathbb{R}^M))\equiv\mathscr{L}^2(\mathbb{R}^N\times\mathbb{R}^N,\mathbb{R}^M)$$

se denomina la derivada segunda de f en a y se nota $D^2f(a)$.

Si $k\ge 2$, sea $A_{k-1}\subset A$ el conjunto de puntos donde f es $k-1$ veces derivable y sea $D^{k-1}f:A_{k-1}\longrightarrow\mathscr{L}^{k-1}(\mathbb{R}^N,\mathbb{R}^M)$ la aplicación derivada $(k-1)$-ésima de f. Se dice que f es k veces derivable en $a\in A_{k-1}$ si la función $D^{k-1}f$ es derivable en dicho punto, en cuyo caso a la aplicación

$$D(D^{k-1}f)(a)\in\mathscr{L}(\mathbb{R}^N,\mathscr{L}^{k-1}(\mathbb{R}^N,\mathbb{R}^M))\equiv\mathscr{L}^k(\mathbb{R}^N,\mathbb{R}^M)$$

se llama derivada k-ésima de f en a y se nota $D^kf(a)$. A $D^kf(a)$ se le asocia la aplicación $d^kf(a):\mathbb{R}^N\longrightarrow\mathbb{R}^M$ definida por:

$$d^kf(a)(x):=D^kf(a)(x,\overset{k}{\dots},x),\qquad \forall x\in\mathbb{R}^N.$$

Sea $A_k\subset A$ el conjunto de puntos de A donde f es k veces derivable. La aplicación

$$x\longrightarrow D^kf(x)$$

de A_k en $\mathscr{L}^k(\mathbb{R}^N,\mathbb{R}^M)$ se denomina la aplicación derivada k-ésima de f y se nota D^kf.

Se dice que f es de clase $\mathscr{C}^k$ en a, y se nota $f\in\mathscr{C}^k(a)$, si es k veces derivable en un entorno de a y la función D^kf es continua en a. Se dice que f es k veces derivable (resp. $\mathscr{C}^k$) cuando lo sea en todos los puntos de su conjunto de definición (que necesariamente será abierto). $\mathscr{C}^k(A)$ es el conjunto de las funciones de clase $\mathscr{C}^k$ en el abierto A.

Finalmente se dice que f es de clase $\mathscr{C}^\infty$ en $a \in A$ si $f \in \mathscr{C}^k(a)$ para todo k.

Si $A \subset \mathbb{R}$ y $f : A \longrightarrow \mathbb{R}^M$ es un campo vectorial k veces derivable en un punto $a \in A$, entonces

$$f^{(k)}(a) = D^k f(a)(1, \overset{k}{\ldots}, 1) = d^k f(a)(1).$$

Propiedades de la derivada k-ésima y de las funciones k veces derivables.

i) **Linealidad.** Sean $A \subset \mathbb{R}^N$, $a \in A$ y $f, g : A \longrightarrow \mathbb{R}^M$ funciones k veces derivables en a y sea $\lambda \in \mathbb{R}$. Entonces $f + g$ y λf son k veces derivables en a con

$$D^k(f+g)(a) = D^k f(a) + D^k g(a), \quad D^k(\lambda f)(a) = \lambda D^k f(a).$$

Además, si $f, g \in \mathscr{C}^k(a)$, entonces $f + g \in \mathscr{C}^k(a)$ y $\lambda f \in \mathscr{C}^k(a)$.

ii) **Regla de la cadena.** Sean $A \subset \mathbb{R}^N, B \subset \mathbb{R}^M, f : A \longrightarrow \mathbb{R}^M$ tal que $f(A) \subset B$ y $g : B \longrightarrow \mathbb{R}^P$. Supongamos que f es k veces derivable en $a \in A$ y que g es k veces derivable en $f(a)$. Entonces la composición $h = g \circ f$ es k veces derivable en a.

Además si $f \in \mathscr{C}^k(a)$ y $g \in \mathscr{C}^k(f(a))$, entonces $h \in \mathscr{C}^k(a)$.

iii) **Carácter local de la derivada k-ésima.** Sean $A \subset \mathbb{R}^N$, $a \in A$ y $f : A \longrightarrow \mathbb{R}^M$ un campo vectorial. Entonces f es k veces derivable en a si, y sólo si, existe algún entorno $U \subset A$ de a tal que $f_{|U}$ es k veces derivable en a, en cuyo caso las derivadas k-ésimas en a de ambas funciones coinciden.

iv) **Derivación de la función inversa.** Sean A y B subconjuntos abiertos de $\mathbb{R}^N$ y f un homeomorfismo de A sobre B. Si f es k veces derivable en a (resp. de clase $\mathscr{C}^k$ en a) para algún $a \in A$ y $Df(a) \in$ Iso $(\mathbb{R}^N)$, entonces f^{-1} es k veces derivable en $f(a)$ (resp. de clase $\mathscr{C}^k$ en $f(a)$).

vi) **Reducción a campos escalares.** Sean $A \subset \mathbb{R}^N$ y $f : A \longrightarrow \mathbb{R}^M$ un campo vectorial. Si $f = (f_1, \ldots, f_M)$, entonces f es k veces derivable en $a \in A$ si, y sólo si, cada f_i es k veces derivable en a para $i = 1, \ldots, M$, en cuyo caso

$$D^k f(a) = \Big(D^k f_1(a), \ldots, D^k f_M(a)\Big)$$

Por tanto, una campo vectorial es de clase k en un punto cuando todas sus componentes lo sean también.

Las aplicaciones lineales son de clase $\mathscr{C}^\infty$ y su diferencial segunda es nula. Toda aplicación bilineal $T : \mathbb{R}^N \times \mathbb{R}^M \longrightarrow \mathbb{R}^P$ es de clase $\mathscr{C}^\infty$ con derivada segunda en $(a, b) \in \mathbb{R}^N \times \mathbb{R}^M$ definida por

$$D^2 T(a,b)\Big((x_1,y_1),(x_2,y_2)\Big) = T(x_1,y_2) + T(x_2,y_1), \quad \forall (x_1,y_1),(x_2,y_2) \in \mathbb{R}^N \times \mathbb{R}^M,$$

por tanto $D^3 T \equiv 0$, por tanto, $D^k T \equiv 0$ para $k \geq 3$.
La aplicación inversión J : Iso $(\mathbb{R}^N) \longrightarrow \mathscr{L}(\mathbb{R}^N)$ es de clase $\mathscr{C}^\infty$.

Teorema de Schwarz para la derivada k-ésima.
Sea $A \subset \mathbb{R}^N$, $f : A \longrightarrow \mathbb{R}^M$ un campo vectorial y $k > 1$. Supongamos que f es k veces derivable en un punto $a \in A$. Entonces la aplicación k-lineal $D^k f(a)$ es simétrica, esto es,

$$D^k f(a)\big(x_1, \ldots, x_k\big) = D^k f(a)\big(x_{\sigma(1)}, \ldots, x_{\sigma(k)}\big)$$

para cualesquiera $x_1, \ldots, x_k \in \mathbb{R}^N$ y cualquier permutación σ del conjunto $\{1, \ldots, k\}$.

Polinomio de Taylor de orden k.
Sea $A \subset \mathbb{R}^N$ y $f : A \longrightarrow \mathbb{R}^M$ una función. Supongamos que f es k veces derivable en un punto $a \in A$, entonces la función $P_k : \mathbb{R}^N \longrightarrow \mathbb{R}^M$ definida por

$$P_k(x) = f(a) + df(a)(x-a) + \frac{1}{2}d^2 f(a)(x-a) + \ldots + \frac{1}{k!}d^k f(a)(x-a)$$

se llama polinomio de Taylor de orden k de f en a. El polinomio de Taylor de orden k de una función en un punto a verifica

$$P_k(a) = f(a), \quad D^n P_k(a) = D^n f(a) \quad (n \le k) \quad \text{y} \quad D^n P_k(x) = 0, \quad \forall n > k.$$

Fórmula infinitesimal del resto.
Sea $A \subset \mathbb{R}^N$ y $f : A \longrightarrow \mathbb{R}^M$ un campo vectorial. Supongamos que f es k veces derivable en un punto $a \in A$, entonces se verifica

$$\lim_{x \to a} \frac{f(x) - P_k(x)}{\|x-a\|^k} = 0.$$

Fórmula de Taylor con resto de Lagrange para campos escalares.
Sea $A \subset \mathbb{R}^N$, $a, x \in A$ con $a \neq x$, tales que el segmento $[a,x]$ está incluido en A. Si $f : A \longrightarrow \mathbb{R}$ es una función de clase $\mathscr{C}^k$ en $[a,x]$ y $k+1$ veces derivable en $]a,x[$, entonces existe $c \in]a,x[$ tal que

$$f(x) = P_k(x) + \frac{1}{(k+1)!}d^{k+1} f(c)(x-a),$$

donde $P_k(x)$ es el polinomio de Taylor de orden k de f en a.

Fórmula de Taylor.
Sean $A \subset \mathbb{R}^N$ y $a, x \in A$ con $a \neq x$, tales que el segmento $[a,x]$ está incluido en A. Si $f : A \longrightarrow \mathbb{R}^M$ es una función de clase $\mathscr{C}^k$ en $[a,x]$ y $k+1$ veces derivable en $]a,x[$, entonces

$$\|f(x) - P_k(x)\| \le \frac{\|x-a\|^{k+1}}{(k+1)!} \sup\{\|D^{k+1} f(z)\| : z \in]a,x[\},$$

donde $P_k(x)$ es el polinomio de Taylor de orden k de f en a.

RESULTADOS PARA CAMPOS ESCALARES.

Derivadas parciales sucesivas.
Sean $A \subset \mathbb{R}^N$, $a \in A$ y f un campo escalar en A.

Sean $i_2, \ldots, i_k \in \{1, \ldots, N\}$ y $A^{(i_2,\ldots,i_k)} \subset A$ el conjunto de puntos donde f tiene derivada parcial de orden $k-1$ respecto de las variables $i_k, \ldots, i_2$. Si $i_1 \in \{1, \ldots, N\}$ y el campo escalar $D_{(i_2,\ldots,i_k)} f$ en $A^{(i_2,\ldots,i_k)}$ tiene derivada parcial respecto de la variable i_1 en el punto a, entonces se dice que f tiene derivada parcial de orden k respecto de las variables $i_k, \ldots, i_1$ en el punto a y la derivada $D_{i_1}(D_{(i_2,\ldots,i_k)} f)(a)$ se nota $D_{(i_1,\ldots,i_k)} f(a)$, esto es

$$D_{(i_1,\ldots,i_k)} f(a) = D_{i_1} D_{i_2} \ldots D_{i_k} f(a).$$

Si $A^{(i_1,\ldots,i_k)} \subset A$ es el conjunto de puntos donde f tiene derivada parcial de orden k respecto de las variables $i_k, \ldots, i_1$, entonces el campo escalar en $A^{(i_1,\ldots,i_k)}$ definido por

$$x \longrightarrow D_{(i_1,\ldots,i_k)} f(x)$$

se denomina la aplicación derivada parcial de orden k de f respecto de las variables $i_k, \ldots, i_1$ y se nota $D_{(i_1,\ldots,i_k)} f$.

Caracterización de campos escalares k veces derivables ($k \geq 2$).
Sean $A \subset \mathbb{R}^N$, $a \in A$ y f un campo escalar en A. Entonces f es k veces derivable en a si, y sólo si, f es $k-1$ veces derivable en un entorno de a y todas las derivadas parciales de orden $k-1$ son derivables en a

Si f es k veces derivable en a, se tiene que

$$D^k f(a)\big(e_{i_1}, \ldots, e_{i_k}\big) = D_{(i_1,\ldots,i_k)} f(a),$$

equivalentemente,

$$D^k f(a)\big(x_1, \ldots, x_k\big) = \sum_{i_1,\ldots,i_k=1}^{N} D_{(i_1,\ldots,i_k)} f(a) \pi_{i_1}(x_1) \ldots \pi_{i_k}(x_k)$$

para cualesquiera $x_1, \ldots, x_k \in \mathbb{R}^N$, donde, para cada $i \in \{1, \ldots, N\}, \pi_i$ denota la i-ésima proyección de $\mathbb{R}^N$.

Además

$$f \in \mathscr{C}^k(a) \;\Leftrightarrow\; D_{(i_1,\ldots,i_{k-1})} f \in \mathscr{C}^1(a) \;\; \forall i_1, \ldots, i_{k-1} \in \{1, \ldots, N\}.$$

Para $k_1, \ldots, k_N \in \mathbb{N} \cup \{0\}$ tales que $k_1 + \cdots + k_N = k$ se define

$$D^{(k_1,\ldots,k_N)} f(a) = D_1^{k_1} \ldots D_N^{k_N} f(a) = D_{(1,\overset{k_1}{\cdots},1,\cdots,N,\overset{k_N}{\cdots},N)} f(a).$$

El Teorema de simetría de Schwarz permite reducir el número de derivadas parciales distintas. Teniendo en cuenta este resultado se obtiene la siguiente expresión del polinomio de Taylor de orden k de un campo escalar en el punto a.

$$P_k(x) = f(a) + (\nabla f(a) \mid x-a) + \frac{1}{2}(x-a) H_f(a)(x-a)^t +$$

$$\sum_{r=3}^{k} \sum_{r_1+\cdots+r_N=r} \frac{1}{r_1! \cdots r_N!} D^{(r_1,\ldots,r_N)} f(a)(x_1-a_1)^{r_1} \ldots (x_N-a_N)^{r_N}$$

para cualquier $x = (x_1, \ldots, x_N) \in \mathbb{R}^N$.

Caracterización de la existencia de derivada de orden k.
Sean $A \subset \mathbb{R}^N$ un abierto, f un campo escalar en A y $k > 1$. Las siguientes afirmaciones son equivalentes:

i) *f es k veces derivable en A.*

ii) *Todas las derivadas parciales de orden $k-1$ son derivables en A.*

Además,

$$f \in \mathscr{C}^k(A) \Leftrightarrow D_{(i_1,\ldots,i_{k-1})} f \in \mathscr{C}^1(A), \quad \forall i_1, \ldots, i_{k-1} \in \{1, \ldots, N\}.$$

6.5. Ejercicios del Tema 6

6.1 Ordenar el polinomio $2x^3y + x^2y^2 + 3x + y + 1$ en potencias de $x-1$ e $y-2$.

6.2 Sea f el campo escalar en $\mathbb{R}^2$ definido por

$$f(x,y) = e^{x \operatorname{sen} y}.$$

¿Cuál es la Fórmula de Taylor de orden 2 con resto de Lagrange en el segmento $\big[(0,0),(x,y)\big]$?

6.3 Sea A un subconjunto abierto de $\mathbb{R}^N$ y f un campo vectorial de clase $\mathscr{C}^\infty$ de A en $\mathbb{R}^M$. Supongamos que existe $C > 0$ tal que

$$\|D^k f(x)\| \leq C^k, \quad \forall x \in A, k \in \mathbb{N}.$$

Probar que si $a \in A$ y $r > 0$ son tales que $B(a,r) \subset A$, entonces la serie

$$\sum_{k \geq 1} \frac{1}{k!} d^k f(a)(x-a)$$

converge uniformemente en $B(a,r)$ y

$$f(x) = f(a) + \sum_{k=1}^{\infty} \frac{1}{k!} d^k f(a)(x-a), \quad \forall x \in B(a,r).$$

6.4 * Demostrar por inducción sobre k la regla de derivación 5) de la Sección 6.1. *Indicación*: Supóngase el resultado cierto para k, y aplíquese la regla de la cadena.

6.5 * Sea $k > 1$ y T una aplicación k-lineal de $\left(\mathbb{R}^N\right)^k$ en $\mathbb{R}^M$.

i) Probar que T es derivable en a con

$$DT(a_1,\ldots,a_k)(x_1,\ldots,x_k) =$$
$$= T(x_1,a_2,\ldots,a_k) + T(a_1,x_2,a_3,\ldots,a_k) + \ldots + T(a_1,\ldots,a_{k-1},x_k)$$

para cualesquiera $a_1,\ldots,a_k,x_1,\ldots,x_k \in \mathbb{R}^N$.

ii) Deducir de i) el Lema 6.6.

6.6 * Sea P un polinomio de grado k en $\mathbb{R}^N$. Probar que P es de clase $\mathscr{C}^\infty$ y que para cada natural $r \geq k$, el polinomio de Taylor de orden r de P en cualquier punto $a \in \mathbb{R}^N$ coincide con P. Deducir que si existe $a \in \mathbb{R}^N$ tal que

$$P(a) = 0 \;\text{ y }\; D^{(i_1,\ldots,i_N)}P(a) = 0, \;\; i_1,\ldots,i_N \in \mathbb{N} \cup \{0\}, \;\; 1 \leq i_1 + \ldots + i_N \leq k,$$

entonces $P(x) = 0, \quad \forall x \in \mathbb{R}^N$.

6.6. Soluciones a los ejercicios del Tema 6.

6.1 Como f es un polinomio de grado 4, entonces f coincide con su polinomio de Taylor de orden 4 en cualquier punto, luego

$$f(x,y) = f(1,2) + \sum_{r=1}^{4} \sum_{r_1+r_2=r} \frac{1}{r_1!r_2!} D^{(r_1,r_2)} f(1,2)\ (x-1)^{r_1}(y-2)^{r_2}.$$

A continuación, calculamos las derivadas parciales de orden menor o igual que cuatro de f y escribiremos en un cuadro la evaluación de éstas en $(1,2)$.

$$\begin{aligned}
D_1 f(x,y) &= 6x^2y + 2xy^2 + 3,\\
D_2 f(x,y) &= 2x^3 + 2x^2y + 1,\\
D^{(2,0)} f(x,y) &= 12xy + 2y^2,\\
D^{(1,1)} f(x,y) &= 6x^2 + 4xy,\\
D^{(0,2)} f(x,y) &= 2x^2,\\
D^{(3,0)} f(x,y) &= 12y,\\
D^{(2,1)} f(x,y) &= 12x + 4y,\\
D^{(1,2)} f(x,y) &= 4x,\\
D^{(0,3)} f(x,y) &= 0,\\
D^{(4,0)} f(x,y) &= 0,\\
D^{(3,1)} f(x,y) &= 12,\\
D^{(2,2)} f(x,y) &= 4,\\
D^{(1,3)} f(x,y) &= 0,\\
D^{(0,4)} f(x,y) &= 0,
\end{aligned}$$

Valores de f y de $D^{(i,j)}f$ en $(1,2)$ $(i+j \leq 4)$

f	D_1	D_2	$D^{(2,0)}$	$D^{(1,1)}$	$D^{(0,2)}$	$D^{(3,0)}$	$D^{(2,1)}$	$D^{(1,2)}$	$D^{(0,3)}$
14	23	7	32	14	2	24	20	4	0

$D^{(4,0)}$	$D^{(3,1)}$	$D^{(2,2)}$	$D^{(1,3)}$	$D^{(0,4)}$
0	12	4	0	0

con lo que, calculando las derivadas parciales y evaluando en $(1,2)$ se tiene

$$\begin{aligned}
f(x,y) = 14 &+ \big(23(x-1) + 7(y-2)\big) + \big(16(x-1)^2 + 14(x-1)(y-2) + (y-2)^2\big) +\\
&\big(4(x-1)^3 + 10(x-1)^2(y-2) + 2(x-1)(y-2)^2\big) +\\
&\big(2(x-1)^3(y-2) + (x-1)^2(y-2)^2\big)
\end{aligned}$$

6.2 Usaremos el Teorema 6.10, para $a = (0,0)$, por lo que existe $t_0 \in]0,1[$ tal que

$$f(x,y) = P_2(x,y) + \sum_{r_1+r_2=3} \frac{1}{r_1!r_2!} D^{(r_1,r_2)} f(t_0x, t_0y) x^{r_1} y^{r_2},$$

donde $P_2(x,y)$ es el polinomio de Taylor de orden 2 en el punto $(0,0)$. Por el Ejercicio 5.4 sabemos que

$$P_2(x,y) = 1 + xy,$$

y además las derivadas parciales de orden 2 son

$$D^{(2,0)} f(x,y) = e^{x \operatorname{sen} y} \operatorname{sen}^2 y,$$

$$D^{(1,1)} f(x,y) = e^{x \operatorname{sen} y} \cos y \big(1 + x \operatorname{sen} y\big),$$

$$D^{(0,2)} f(x,y) = x\, e^{x \operatorname{sen} y} \big(-\operatorname{sen} y + x \cos^2 y\big),$$

de donde

$$D^{(3,0)} f(x,y) = e^{x \operatorname{sen} y} \operatorname{sen}^3 y$$

$$D^{(2,1)} f(x,y) = e^{x \operatorname{sen} y} \operatorname{sen} y \cos y \big(2 + x \operatorname{sen} y\big),$$

$$D^{(1,2)} f(x,y) = e^{x \operatorname{sen} y} \big(-\operatorname{sen} y - x \operatorname{sen}^2 y + 2x \cos^2 y + x^2 \operatorname{sen} y \cos y\big),$$

$$D^{(0,3)} f(x,y) = x\, e^{x \operatorname{sen} y} \cos y \big(-1 - 3x \operatorname{sen} y + x^2 \cos^2 y\big).$$

Por tanto, en vista de la forma que tiene el polinomio de orden 2 en $(0,0)$ de esta función y del Teorema 6.10 se tiene

$$f(x,y) = 1 + xy + \frac{1}{3!} D^{(3,0)} f(t_0x, t_0y) x^3 + \frac{1}{2!} D^{(2,1)} f(t_0x, t_0y) x^2 y +$$

$$\frac{1}{2!} D^{(1,2)} f(t_0x, t_0y) x y^2 + \frac{1}{3!} D^{(0,3)} f(t_0x, t_0y) y^3$$

6.3 La sucesión de sumas parciales de la serie

$$f(a) + \sum_{k \geq 1} \frac{1}{k!} d^k f(a)(x - a)$$

es $P_k(x)$, donde P_k es el polinomio de Taylor de orden k de f en a.

La fórmula de Taylor garantiza que si $0 < \|x - a\| < r$, entonces

$$\|f(x) - P_k(x)\| \leq \frac{1}{(k+1)!} \|x - a\|^{k+1} \sup\{\|D^{k+1} f(z)\| : z \in]a, x[\} \leq$$

$$\leq \frac{1}{(k+1)!} \|x - a\|^{k+1} C^{k+1} \leq \frac{(rC)^{k+1}}{(k+1)!},$$

luego

$$\|f(x) - P_k(x)\| \leq \frac{(rC)^{k+1}}{(k+1)!}, \ \forall x \in B(a,r),$$

y en consecuencia

$$\sup\{\|f(x)-P_k(x)\| : x\in B(a,r)\} \le \frac{(rC)^{k+1}}{(k+1)!} \to 0.$$

Hemos probado que $f-P_k$ converge a cero uniformemente en $B(a,r)$, por tanto

$$f(x) = f(a) + \sum_{k=1}^{\infty} \frac{1}{k!} d^k f(a)(x-a),\ \forall x\in B(a,r).$$

6.4 Sabemos que es cierta para $k=1$. Supongámosla cierta para k. Como f es $k+1$ veces derivable en a, sea U un entorno abierto de a incluido en A donde f es k veces derivable. Por la hipótesis de inducción

$$D^k(T\circ f)(x) = T\circ D^k f(x), \quad \forall x\in U,$$

es decir,

$$D^k(T\circ f) = T\circ D^k f.$$

La regla de la cadena nos dice que $D^k(T\circ f)$ es derivable en a, es decir, $T\circ f$ es $k+1$ veces derivable en a y se tiene que

$$D(D^k(T\circ f)) = D(T\circ D^k f) = T\circ D(D^k f) = T\circ D^{k+1} f$$

y en consecuencia, para cualesquiera $x_1,\ldots,x_{k+1}\in\mathbb{R}^N$, se tiene

$$D^{k+1}(T\circ f)(a)(x_1,\ldots,x_{k+1}) = T\Big(D^{k+1}f(a)(x_1,\ldots,x_{k+1})\Big).$$

6.5 i) Si $T\in\mathscr{L}^k(\mathbb{R}^N,\mathbb{R}^M)$, probaremos que

$$DT\big(a_1,\ldots,a_k\big)(x_1,\ldots,x_k) =$$

$$= T\big(x_1,a_2,\ldots,a_k\big) + T\big(a_1,x_2,a_3,\ldots,a_k\big) + \ldots + T\big(a_1,\ldots,a_{k-1},x_k\big)$$

para cualesquiera $a_1,\ldots,a_k,x_1,\ldots,x_k\in\mathbb{R}^N$. Nótese que la candidata a diferencial en $(a_1,\ldots,a_k)$ es una aplicación lineal.

Probaremos la afirmación para $k=3$. En efecto,

$$T\big(x_1,x_2,x_3\big) = T\big(a_1+(x_1-a_1),a_2+(x_2-a_2),a_3+(x_3-a_3)\big) =$$

$$= T\big(a_1,a_2,a_3\big) + T\big(x_1-a_1,a_2,a_3\big) + T\big(a_1,x_2-a_2,a_3\big) + T\big(a_1,a_2,x_3-a_3\big) +$$
$$+T\big(x_1-a_1,x_2-a_2,a_3\big) + T\big(x_1-a_1,a_2,x_3-a_3\big) + T\big(a_1,x_2-a_2,x_3-a_3\big) +$$
$$+T\big(x_1-a_1,x_2-a_2,x_3-a_3\big),$$

por tanto

$$\Big\|T\big(x_1,x_2,x_3\big) - T\big(a_1,a_2,a_3\big) - \Big(T\big(x_1-a_1,a_2,a_3\big) + T\big(a_1,x_2-a_2,a_3\big) + T\big(a_1,a_2,x_3-a_3\big)\Big)\Big\| =$$

$$= \left\|T\big(x_1-a_1,x_2-a_2,a_3\big)+T\big(x_1-a_1,a_2,x_3-a_3\big)+\right.$$
$$\left.+T\big(a_1,x_2-a_2,x_3-a_3\big)+T\big(x_1-a_1,x_2-a_2,x_3-a_3\big)\right\| \leq$$
$$\leq \|T\| \Big(\|x_1-a_1\|\ \|x_2-a_2\|\ \|a_3\|+\|x_1-a_1\|\ \|a_2\|\ \|x_3-a_3\|+$$
$$+\|a_1\|\ \|x_2-a_2\|\ \|x_3-a_3\|+\|x_1-a_1\|\ \|x_2-a_2\|\ \|x_3-a_3\|\Big).$$

Finalmente, es claro que dividiendo la expresión anterior por $\|(x_1-a_1,x_2,-a_2,x_3-a_3)\|$ se obtiene una función que tiende al cero en el punto (a_1,a_2,a_3).

ii) Para probar esta afirmación, basta usar el apartado anterior y la regla de la cadena para la composición
$$x \mapsto (x,\overset{k}{\cdots},x) \mapsto T(x,\overset{k}{\cdots},x).$$
Si llamamos Φ a la primera aplicación, por ser ésta lineal, se tiene
$$DP(a) = D(T\circ\Phi)(a) = DT(\Phi(a))\circ\Phi.$$
Por tanto,
$$DP(a)(x) = DT(a,\overset{k}{\cdots},a)(x,\overset{k}{\cdots},x) = kT(x,a\overset{k-1}{\cdots},a).$$
donde hemos usado la simetría de T.

6.6 En vista del problema anterior, la diferencial de P es de nuevo un polinomio de grado $k-1$, por tanto, P es de clase $\mathscr{C}^\infty$. Además, como D^k es constante, entonces $D^{k+1}\equiv 0$, por tanto, el Teorema 6.11 nos asegura que P coincide con su polinomio de Taylor de orden m en cada punto, para $m \geq k$. En consecuencia, si $P(a)=0$ y las derivadas parciales de orden menor o igual que k se anulan, entonces $P\equiv 0$.

Tema 7

Extremos relativos.

Dedicamos esta lección a generalizar, en lo posible, a campos escalares el estudio de extremos relativos de funciones reales de variable real. A este respecto recordamos los siguientes resultados, condiciones necesarias y suficientes para la existencia de extremos relativos de funciones reales de variable real.

7.1. Condiciones necesarias y suficientes de extremo relativo

Proposición (Condición necesaria de extremo relativo). *Sea I un intervalo abierto, $f : I \to \mathbb{R}$ una función real de variable real y supongamos que f alcanza un extremo relativo en un punto a de I en el que f es derivable. Entonces $f'(a) = 0$.*

Proposición (Condiciones necesaria y suficiente de extremo relativo). *Sea I un intervalo abierto, a un punto de I, k un número natural con $k \geq 2$ y $f : I \to \mathbb{R}$ una función $k-1$ veces derivable en I y k veces derivable en el punto a. Supongamos además que*

$$f'(a) = f''(a) = \ldots = f^{(k-1)}(a) = 0 \text{ y } f^{(k)}(a) \neq 0.$$

Entonces:

a) Si k es impar, entonces f no alcanza un extremo relativo en el punto a.

b) Si k es par, se tiene

i) Si $f^{(k)}(a) > 0$ entonces f alcanza en a un mínimo relativo.
ii) Si $f^{(k)}(a) < 0$ entonces f alcanza en a un máximo relativo.

Se incluye también la definición de forma cuadrática de N variables y la clasificación de Sylvester de las mismas lo que nos ayudará a conseguir un cálculo práctico de los extremos relativos.

Recordemos en primer lugar la siguiente

Definición 7.1 (extremo relativo)**.** Sea f un campo escalar definido en $A \subset \mathbb{R}^N$. Se dice que f alcanza en un punto $a \in A$ un *máximo relativo* si existe un entorno abierto U de a tal que $U \subset A$ y

$$f(x) \leq f(a),\ \forall x \in U.$$

Se dice que f alcanza en el punto $a \in A$ un máximo relativo estricto si existe un entorno abierto U de a tal que $U \subset A$ y

$$f(x) < f(a),\ \forall x \in U \backslash \{a\}.$$

Las definiciones para mínimo se obtienen cambiando el sentido de las desigualdades. La expresión *extremo relativo* significa o bien máximo relativo, o bien mínimo relativo.

Proposición 7.2 (Condición necesaria de existencia de extremo relativo)**.** *Sea f un campo escalar definido en $A \subset \mathbb{R}^N$. Si f alcanza un extremo relativo en un punto $a \in A$ y f tiene gradiente en a, entonces $\nabla f(a) = 0$.*

Demostración:

Sean $i \in \{1, \ldots, N\}$ y sea g_i la función real de variable real definida en un entorno de cero por $g_i(t) = f(a + te_i)$, donde como es habitual $\{e_1, \ldots, e_N\}$ es la base canónica de $\mathbb{R}^N$. g_i alcanza un extremo relativo en cero y además es derivable en 0. En consecuencia, por el resultado real de variable real, $0 = g_i'(0) = D_i f(a)$. ∎

Definición 7.3 (punto crítico)**.** Sea f un campo escalar definido en $A \subset \mathbb{R}^N$. Los puntos de A donde f tiene gradiente nulo se llaman *puntos críticos* de f. Es decir son las soluciones del sistema de ecuaciones (no necesariamente lineal) $\nabla f(x) = 0$ definido en los puntos en los que f tiene gradiente.

Nota 7.4. El mismo argumento utilizado en la demostración anterior asegura que si f alcanza un extremo relativo en un punto $a \in A$ y f tiene derivada direccional en a según el vector u, entonces $f'(a; u) = 0$.

Los siguientes ejemplos muestran que la condición no es suficiente y que no se tiene información en los puntos en que f no tiene gradiente (si bien, la nota anterior podría sernos de ayuda en tal caso).

Ejemplos 7.5.

a) La función $f(x,y) = x^2 - y^2, \forall (x,y) \in \mathbb{R}^2$, es derivable en $(0,0)$ y $\nabla f(0,0) = (0,0)$. Sin embargo, f no alcanza en $(0,0)$ un extremo relativo:

$$\alpha \neq 0 \Rightarrow\ f(\alpha, 0) > 0,\ f(0, \alpha) < 0.$$

b) La función $f(x,y) = |x| + |y|, \forall (x,y) \in \mathbb{R}^2$, no tiene gradiente en $(0,0)$ y alcanza en $(0,0)$ un mínimo estricto, ya que $f(0,0) = 0$ y $f(x,y) > 0,\ \forall (x,y) \in \mathbb{R}^2 \backslash \{(0,0)\}$.

c) La función $f(x,y) = |x| + y,\ \forall (x,y) \in \mathbb{R}^2$, no tiene gradiente en $(0,0)$ y no alcanza en $(0,0)$ extremo relativo: $\alpha \neq 0 \Rightarrow\ f(0,\alpha) = \alpha$ (lo que también se deduce de la nota anterior ya que $f'((0,0);(0,1)) = 1$).

Teorema 7.6 (Condic. neces. y sufic. de existencia de extremo relativo)**.** *Sean f un campo escalar definido en $A \subset \mathbb{R}^N$ y $a \in A$ un punto crítico de f. Supongamos que f es dos veces derivable en a y que $D^2 f(a)$ es no nula.*

i) (Condiciones suficientes)

Si $d^2 f(a)(x) < 0,\ \forall x \neq 0$, entonces f alcanza en a un máximo relativo estricto.

Si $d^2 f(a)(x) > 0,\ \forall x \neq 0$, entonces f alcanza en a un mínimo relativo estricto.

ii) (Condiciones necesarias)

Si f alcanza en a un máximo relativo, entonces $d^2 f(a)(x) \leq 0,\ \forall x \in \mathbb{R}^N$.

Si f alcanza en a un mínimo relativo, entonces $d^2 f(a)(x) \geq 0,\ \forall x \in \mathbb{R}^N$.

Demostración:

Obsérvese que el polinomio de Taylor de segundo orden de f en a es

$$f(a) + \frac{1}{2} d^2 f(a)(x-a),\ \forall x \in \mathbb{R}^N.$$

i) Sea $\|\cdot\|$ una norma cualquiera de $\mathbb{R}^N$ y denotemos por $S(\mathbb{R}^N)$ la esfera unidad para dicha norma. Si

$$d^2 f(a)(x) < 0,\ \forall x \in \mathbb{R}^N \backslash \{0\},$$

la propiedad de compacidad nos permite asegurar que

$$\exists c > 0 :\ d^2 f(a)(u) \leq -c,\ \forall u \in S(\mathbb{R}^N).$$

Desnormalizando obtenemos

$$d^2 f(a)(x) \leq -c\|x\|^2,\ \forall x \in \mathbb{R}^N.$$

La fórmula infinitesimal del resto nos asegura que existe $\delta > 0$ tal que $B(a,\delta) \subset A$ y que para $x \in \mathbb{R}^N$ con $0 < \|x-a\| < \delta$ se verifica

$$\left| f(x) - f(a) - \frac{1}{2} d^2 f(a)(x-a) \right| < \frac{c}{2}\|x-a\|^2,$$

de donde se deduce que

$$f(x) - f(a) = \left[f(x) - f(a) - \frac{1}{2} d^2 f(a)(x-a) \right] + \frac{1}{2} d^2 f(a)(x-a) < \frac{c}{2}\|x-a\|^2 - \frac{c}{2}\|x-a\|^2 = 0.$$

Hemos probado que f alcanza un máximo relativo estricto en a.

ii) Por definición de máximo relativo existe $\delta > 0$ tal que $B(a,\delta) \subset A$ y para $x \in B(a,\delta)$ se verifica $f(x) \le f(a)$. Si $u \in S(\mathbb{R}^N)$ y t es un real con $0 < |t| < \delta$, se tiene entonces que

$$0 \ge \frac{f(a+tu)-f(a)}{t^2} =$$

$$\frac{f(a+tu)-f(a)-\frac{1}{2}d^2f(a)(tu)}{t^2} + \frac{1}{2}\frac{d^2f(a)(tu)}{t^2} =$$

$$\frac{f(a+tu)-f(a)-\frac{1}{2}d^2f(a)(tu)}{\|tu\|^2} + \frac{1}{2}\, d^2f(a)(u)$$

y basta, teniendo en cuenta la fórmula infinitesimal del resto, tomar límite cuando $t \to 0$ para obtener $d^2f(a)(u) \le 0$. Si ahora x es un vector no nulo, se tiene que

$$d^2f(a)(x) = \|x\|^2 d^2f(a)\left(\frac{x}{\|x\|}\right) \le 0.$$

Hemos probado que $d^2f(a)(x) \le 0,\ \forall x \in \mathbb{R}^N$, ya que esta desigualdad es claramente válida para $x = 0$.

Los resultados para mínimos se obtienen considerando la función $-f$. ∎

Nota 7.7. En la utilización práctica del teorema convendrá tener presente la siguiente consecuencia del apartado *ii*):

Si $d^2f(a)(x) \le 0,\ \forall x \in \mathbb{R}^N$, entonces de tener f un extremo relativo en a, éste ha de ser máximo.

Análogamente, si $d^2f(a)(x) \ge 0,\ \forall x \in \mathbb{R}^N$, entonces de tener f un extremo relativo en a, éste ha de ser mínimo.

En efecto, si f alcanzase en a un mínimo relativo, entonces tendríamos para todo $x \in \mathbb{R}^N$ que $d^2f(a)(x) \ge 0$, y en consecuencia $d^2f(a) = 0$, luego también $D^2f(a) = 0$ en contra de la hipótesis [nótese que $D^2(d^2f(a))(x) = 2D^2f(a)$ pues para $1 \le i,j \le N$ se tiene $D_{(i,j)}(d^2f(a))(x) = 2D_{(i,j)}f(a)$].

Ejemplos 7.8. El teorema no se puede mejorar como nuestran los siguientes ejemplos:

a) La función $f(x,y) = x^2 - y^4, \forall (x,y) \in \mathbb{R}^2$, verifica que

$$d^2f(0,0)(x,y) = 2x^2 \ge 0\ ,\ \forall (x,y) \in \mathbb{R}^2,$$

y sin embargo no tiene ningún extremo en $(0,0)$.

b) La función $f(x,y) = x^2 + y^4, \forall (x,y) \in \mathbb{R}^2$, alcanza un mínimo relativo estricto en $(0,0)$ y sin embargo

$$d^2f(0,0)(x,y) = 2x^2,\ \forall (x,y) \in \mathbb{R}^2,$$

se anula en $\{(0,y):\ y \in \mathbb{R}\}$.

Las condiciones necesarias y suficientes de existencia de extremo relativo dadas en el Teorema 7.6 ponen de relieve el interés que tiene para nosotros el conocimiento de criterios que permitan decidir el "signo" del polinomio $d^2f(a)$.

Definición 7.9 (forma cuadrática)**.** Una *forma cuadrática* en $\mathbb{R}^N$ es una aplicación $Q:\mathbb{R}^N \to \mathbb{R}$ de la forma $Q(x) = xH_Qx^t$, $\forall x \in \mathbb{R}^N$, donde $H_Q \in \mathscr{M}_{N\times N}(\mathbb{R})$ es una matriz simétrica.

Fijada una base en $\mathbb{R}^N$, la matriz asociada a una forma cuadrática es única (véase la Proposición 7.21 y la Definición 7.22). Concretamente,

$$d^2f(a)(x) = \sum_{i,j=1}^{N} x_iD_{(i,j)}f(a)x_j = xH_f(a)x^t\,,$$

esto es, la matriz hessiana $H_f(a)$ es la matriz asociada a la forma cuadrática $d^2f(a)$ cuando se considera la base canónica.

Definición 7.10 (Clasificación de las formas cuadráticas)**.** Una forma cuadrática Q en $\mathbb{R}^N$ se dice que es
definida positiva si $Q(x) > 0$, $\forall x \in \mathbb{R}^N\backslash\{0\}$,
(semidefinida) positiva si $Q(x) \geq 0$, $\forall x \in \mathbb{R}^N$,
Análogamente se definen las formas definidas negativas y (semidefinidas) negativas.
indefinida si toma valores positivos y negativos, es decir ni es (semidefinida) positiva ni es (semidefinida) negativa.

En dimensión 1 las formas cuadráticas vienen dadas por $Q(x) = ax^2$ $(a \in \mathbb{R})$, luego son definidas positivas (si $a > 0$), definidas negativas (si $a < 0$) o nulas (si $a = 0$).

Sin embargo, en dimensión mayor que 1 se dan todas las posibilidades enumeradas en la definición. Así por ejemplo

$$Q(x,y) = x^2+y^2,\ \forall(x,y)\in\mathbb{R}^2 \text{ es definida positiva,}$$
$$Q(x,y) = x^2,\ \forall(x,y)\in\mathbb{R}^2 \text{ es positiva (no definida),}$$
$$Q(x,y) = x^2-y^2,\ \forall(x,y)\in\mathbb{R}^2 \text{ es indefinida.}$$

Como ya se ha comentado, es interesante disponer de un criterio útil y cómodo de clasificación de las formas cuadráticas. A ello nos dedicamos a continuación.

Definición 7.11 (submatrices principales)**.** Dados N,k naturales, con $k \leq N$ y $H = (a_{ij})_{1\leq i,j\leq N}$ una matriz cuadrada de orden N, se dice que una matriz K es una *submatriz principal* de H de orden k si existe una aplicación estrictamente creciente

$$\sigma : \{1,\ldots,k\} \to \{1,\ldots,N\}$$

tal que $K = \left(a_{\sigma(i)\sigma(j)}\right)_{1\leq i,j\leq k}$.

El número total de submatrices principales es $\binom{N}{1}+\ldots+\binom{N}{N-1}+\binom{N}{N} = 2^N-1$. De ellas destacaremos las N siguientes (las esquinas izquierdas superiores)

$$H_k = \left(a_{ij}\right)_{1\leq i,j\leq k},\ k = 1,\ldots,N.$$

La demostración del siguiente criterio puede verse en el apéndice de este tema.

Proposición 7.12 (Criterio de Sylvester). *Sea Q una forma cuadrática en $\mathbb{R}^N$, y $H = \left(a_{ij}\right)_{1\le i,j\le N}$ la matriz asociada a Q. Entonces se tiene:*

$$Q \textit{ es definida } \left\{ \begin{array}{c} \textit{positiva} \\ \textit{negativa} \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{c} \det\left(H_k\right) > 0 \\ (-1)^k \det\left(H_k\right) > 0 \end{array} \right\} k = 1, \ldots, N$$

$$Q \textit{ es } \left\{ \begin{array}{c} \textit{positiva} \\ \textit{negativa} \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{c} \det\left(K\right) \ge 0 \\ (-1)^{orden(K)} \det\left(K\right) \ge 0 \end{array} \right\} \forall K \textit{ submatriz principal de } H$$

$$Q \textit{ es indefinida } \Leftrightarrow \exists K, L \textit{ subm. princ. de } H : \det\left(K\right) < 0 \textit{ y } (-1)^{orden(L)} \det\left(L\right) < 0$$

En consecuencia,
Q es indefinida si $\det(K) < 0$ *para alguna submatriz principal K de orden par.*

Ejemplos 7.13.

a) Clasificar las formas cuadráticas de dos variables, es decir

$$Q(x,y) = ax^2 + 2bxy + cy^2, \forall (x,y) \in \mathbb{R}^2 \text{ donde } a,b,c \in \mathbb{R}.$$

Como $H = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$, y el determinante de H es $D = ac - b^2$, se tiene

$$Q \text{ es definida } \left\{ \begin{array}{c} \text{positiva} \\ \text{negativa} \end{array} \right\} \Leftrightarrow D > 0 \text{ y } \left\{ \begin{array}{c} a > 0 \\ a < 0 \end{array} \right\}$$

$$Q \text{ es } \left\{ \begin{array}{c} \text{positiva} \\ \text{negativa} \end{array} \right\} \Leftrightarrow D \ge 0 \text{ y } \left\{ \begin{array}{c} a,c \ge 0 \\ a,c \le 0 \end{array} \right\}$$

$$Q \text{ es indefinida } \Leftrightarrow D < 0$$

b) Clasificar la forma cuadrática asociada a la matriz $\begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$.

La forma cuadrática es indefinida pues $\det \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} < 0$.

Los resultados Teorema 7.6, Nota 7.7 y Proposición 7.12 nos permiten enunciar el siguiente resultado práctico.

Proposición 7.14 (Condic. nec. y suf. de existencia de extremo relativo)**.** *Sea f un campo escalar definido en $A \subset \mathbb{R}^N$. Supongamos que f es dos veces derivable en un punto crítico $a \in A$ y que*

$$H_f(a) = \begin{pmatrix} D_{(1,1)}f(a) & \dots & D_{(N,1)}f(a) \\ \dots & \dots & \dots \\ D_{(1,N)}f(a) & \dots & D_{(N,N)}f(a) \end{pmatrix} \neq 0.$$

Para $k = 1, \dots, N$, sean $H_k = \left(D_{(i,j)}f(a)\right)_{1 \le i,j \le k}$ las submatrices principales "esquinas".

Se verifican entonces las siguientes afirmaciones:

i) Si $\det(H_k) > 0,\ k = 1, \dots, N$, entonces f alcanza en a un mínimo relativo estricto.

ii) Si $(-1)^k \det(H_k) > 0,\ k = 1, \dots, n$, entonces f alcanza en a un máximo relativo estricto.

iii) Si $\det(K) \ge 0$, para cualquier submatriz principal K de $H_f(a)$, entonces de alcanzar f en a un extremo relativo ha de ser mínimo.

iv) Si $(-1)^{orden(K)} \det(K) \ge 0$, para cualquier submatriz principal K de $H_f(a)$, entonces de alcanzar f en a un extremo relativo, ha de ser máximo.

v) Si existen K y L submatrices principales de $H_f(a)$ tales que $\det(K) < 0$ y $(-1)^{orden(L)} \det(L) < 0$ (lo que en particular ocurre si el determinante de alguna submatriz principal de orden par de $H_f(a)$ es negativo), entonces f no alcanza en a un extremo relativo.

El resultado anterior para el caso más simple y habitual.

Proposición 7.15 (Condiciones de extremo relativo para $f : \mathbb{R}^2 \to \mathbb{R}$)**.** *Sea f un campo escalar definido en $A \subset \mathbb{R}^2$. Supongamos que f es dos veces derivable en un punto crítico $(a,b) \in A$ y que $H_f(a,b) \neq 0$. Se verifican entonces las siguientes afirmaciones:*

i) Si $\det(H_f(a,b)) > 0$ y $D_{(1,1)}f(a,b) > 0$ entonces f alcanza en (a,b) un mínimo relativo estricto.

ii) Si $\det(H_f(a,b)) > 0$ y $D_{(1,1)}f(a,b) < 0$ entonces f alcanza en (a,b) un máximo relativo estricto.

iii) Si $\det(H_f(a,b)) \ge 0$, $D_{(1,1)}f(a,b) \ge 0$ y $D_{(2,2)}f(a,b) \ge 0$, entonces de alcanzar f en (a,b) un extremo relativo ha de ser mínimo.

iv) Si $\det(H_f(a,b)) \ge 0$, $D_{(1,1)}f(a,b) \le 0$ y $D_{(2,2)}f(a,b) \le 0$, entonces de alcanzar f en (a,b) un extremo relativo ha de ser máximo.

v) Si $\det(H_f(a,b)) < 0$, entonces f no alcanza en (a,b) un extremo relativo.

Terminamos esta lección dando un teorema para la existencia de extremos relativos involucrando derivadas de orden superior a dos que generaliza el análogo en dimensión uno.

Proposición 7.16 (Condic. nec. y suf. de existencia de extremo relativo)**.** *Sean f un campo escalar definido en $A \subset \mathbb{R}^N$ y k un natural mayor que 2. Supongamos que f es k veces derivable en un punto crítico $a \in A$, que todas las derivadas parciales hasta las de orden $k-1$ son nulas y que alguna derivada parcial de orden k no es nula.*

Se verifican las siguientes afirmaciones:

a) Si k es impar, entonces f no alcanza extremo relativo en a.

b) Si k es par, se tiene:

i) Si $d^k f(a)(x) > 0,\ \forall x \neq 0$, entonces f alcanza en a un mínimo relativo estricto.

ii) Si $d^k f(a)(x) < 0,\ \forall x \neq 0$, entonces f alcanza en a un máximo relativo estricto.

iii) Si $d^k f(a)(x) \geq 0,\ \forall x \in \mathbb{R}^N$, entonces de alcanzar f en a un extremo relativo, ha de ser mínimo.

iv) Si $d^k f(a)(x) \leq 0,\ \forall x \in \mathbb{R}^N$, entonces de alcanzar f en a un extremo relativo, ha de ser máximo.

v) Si $[\exists x, y \in \mathbb{R}^N : d^k f(a)(x) < 0 < d^k f(a)(y)]$, entonces f no alcanza ningún extremo relativo en a.

Notas 7.17.

a) En el caso particular del resultado anterior para subconjuntos de $\mathbb{R}$, para $i = 1, \ldots, k$, se tiene

$$d^i f(a)(x) = f^{(i)}(a)x^i,\ \forall x \in \mathbb{R},$$

por lo que recaemos en los resultados conocidos para funciones reales de variable real (recordados en la introducción):

- Si k es impar, f no alcanza en a un extremo relativo.
- Si k es par se verifica: $f^{(k)}(a) \left\{ \begin{array}{c} <0 \\ >0 \end{array} \right\}$ si, y sólo si, f alcanza en a un $\left\{ \begin{array}{c} \text{máximo} \\ \text{mínimo} \end{array} \right\}$ relativo estricto.

b) Al no contar con criterios que faciliten la clasificación de formas de grado $k \geq 4$, las hipótesis de las afirmaciones que aparecen en la parte $b)$ son sólo comprobables en casos particulares.

7.2. Apéndice: Clasificación de formas cuadráticas de N variables

Las condiciones necesarias y suficientes de existencia de extremo relativo dadas en el Teorema 7.6, y su generalización a N variables, ponen de relieve el interés que tiene para nosotros el conocimiento de criterios que permitan decidir el "signo" del polinomio $d^2 f(a)$, y en general de $d^k f(a)$ para k par. Puesto que para $k \geq 4$ no parece que existan criterios con utilidad práctica, nos centramos en el caso $k = 2$.

En lo que sigue N es un número natural, $\{e_1, \ldots, e_N\}$ es la base canónica de $\mathbb{R}^N$ y $(\cdot|\cdot)$ es el producto escalar euclídeo

$$(x|y) := x_1 y_1 + \ldots + x_N y_N.$$

Es fácil comprobar que el producto escalar euclídeo goza de las siguientes propiedades:

1. $(\lambda x + y|z) = \lambda(x|z) + (y|z),\ \forall x, y, z \in \mathbb{R}^N,\ \forall \lambda \in \mathbb{R}.$
2. $(x|\lambda y + z) = \lambda(x|y) + (x|z),\ \forall x, y, z \in \mathbb{R}^N,\ \forall \lambda \in \mathbb{R}.$
3. $(x|y) = (y|x),\ \forall x, y \in \mathbb{R}^N.$

Es decir, es una forma bilineal y simétrica en el sentido que se recoge en la siguiente definición.

Definición 7.18 (forma bilineal)**.** Una *forma bilineal* en $\mathbb{R}^N$ es una aplicación $B : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ lineal en ambas variables, es decir:

i) $B(\lambda x + y, z) = \lambda B(x, z) + B(y, z),\ \forall x, y, z \in \mathbb{R}^N,\ \forall \lambda \in \mathbb{R}.$

ii) $B(x, \lambda y + z) = \lambda B(x, y) + B(x, z),\ \forall x, y, z \in \mathbb{R}^N,\ \forall \lambda \in \mathbb{R}.$

Si además verifica:

iii) $B(x, y) = B(y, x),\ \forall x, y \in \mathbb{R}^N,$

se dice entonces que B es una *forma bilineal simétrica* .

Definición 7.19 (forma cuadrática)**.** Una *forma cuadrática* en $\mathbb{R}^N$ es una aplicación $Q : \mathbb{R}^N \to \mathbb{R}$ tal que $Q(x) = B(x, x)$, $\forall x \in \mathbb{R}^N$ para alguna forma bilineal B.

Definición 7.20 (operador autoadjunto)**.** Una aplicación lineal $T : \mathbb{R}^N \to \mathbb{R}^N$ se dice que es un *operador autoadjunto* si verifica

$$(T(x)|y) = (x|T(y)),\ \forall x, y \in \mathbb{R}^N.$$

Es inmediato que si T es un operador autoadjunto en $\mathbb{R}^N$, entonces la aplicación

$$(x,y) \ \rightarrow \ (x|T(y))$$

es una forma bilineal simétrica. De hecho toda forma bilineal y simétrica responde a una expresión de este tipo:

Proposición 7.21. *Sea Q una forma cuadrática en $\mathbb{R}^N$. Entonces*

i) Existe una única forma bilineal simétrica B tal que $Q(x) = B(x,x)$, $\forall x \in \mathbb{R}^N$. De hecho se tiene que $B(x,y) = \frac{Q(x+y)-Q(x-y)}{4}$, $\forall x,y \in \mathbb{R}^N$. En consecuencia, si Q no es nula, entonces Q es un polinomio homogéneo de grado dos en N variables.

ii) Existe un único operador autoadjunto T tal que $B(x,y) = (x|T(y))$, $\forall x,y \in \mathbb{R}^N$. De hecho se tiene la siguiente (equivalente) expresión $Q(x) = (x|T(x))$, $\forall x \in \mathbb{R}^N$.

Demostración:

i) Supuesta la existencia veamos la unicidad: Si B es una forma bilineal simétrica en $\mathbb{R}^N$ tal que $Q(x) = B(x,x)$, $\forall x \in \mathbb{R}^N$, entonces para cualesquiera $x,y \in \mathbb{R}^N$ se tiene

$$Q(x+y) = B(x+y,x+y) = B(x,x) + 2B(x,y) + B(y,y)$$

$$Q(x-y) = B(x-y,x-y) = B(x,x) - 2B(x,y) + B(y,y),$$

y en consecuencia

$$B(x,y) = \frac{Q(x+y)-Q(x-y)}{4},$$

luego B está determinada por Q.

Existencia: Es rutina convencerse de que la aplicación $B : \mathbb{R}^N \times \mathbb{R}^N \ \rightarrow \ \mathbb{R}$ definida por

$$B(x,y) := \frac{Q(x+y)-Q(x-y)}{4}$$

es una aplicación bilineal simétrica tal que $B(x,x) = Q(x)$, $\forall x \in \mathbb{R}^N$.

ii) Supuesta la existencia, veamos la unicidad: Si T es un operador (lineal autoadjunto) en $\mathbb{R}^N$ tal que

$$B(x,y) = (x|T(y)), \ \forall x,y \in \mathbb{R}^N,$$

entonces para todo $x \in \mathbb{R}^N$ se verifica que

$$T(x) = \sum_{k=1}^{N} (e_k|T(x))e_k = \sum_{k=1}^{N} B(e_k,x)e_k,$$

y por tanto el operador T está determinado por B.

Veamos ahora la existencia: Es claro que la expresión

$$T(x) = \sum_{k=1}^{N} B(e_k,x)e_k, \ \forall x \in \mathbb{R}^N$$

define un operador lineal en $\mathbb{R}^N$. Además

$$\begin{aligned} B(x,y) &= B\left(\sum_{k=1}^{N}(x\,|\,e_k)e_k, y\right) \\ &= \sum_{k=1}^{N}(x|e_k)B(e_k,y) \\ &= \left(x|\sum_{k=1}^{N}B(e_k,y)e_k\right) \\ &= (x|T(y)), \quad \forall x,y \in \mathbb{R}^N. \end{aligned}$$

La simetría de B nos prueba que T es autoadjunto.

Finalmente, es claro que la condición $B(x,y) = (x|T(y))$, $\forall x,y \in \mathbb{R}^N$ implica $Q(x) = (x|T(x))$, $\forall x \in \mathbb{R}^N$.
El recíproco también es cierto. Supongamos que $Q(x) = (x|T(x))$, $\forall x \in \mathbb{R}^N$. Entonces $(x,y) \to B(x,y)$ y $(x,y) \to (x|T(y))$ son formas bilineales simétricas que determinan la misma forma cuadrática, luego coinciden. ∎

El producto escalar euclídeo verifica también la propiedad

4. $(x|x) > 0$, $\forall x \in \mathbb{R}^N \backslash \{0\}$,

es decir, es una forma bilineal simétrica que es definida positiva en el sentido de la Definición 7.10.

Definición 7.22 (Matriz de una forma cuadrática. Autovectores. Autovalores)**.** Sea Q una forma cuadrática en $\mathbb{R}^N$. La matriz simétrica $H = (B(e_i,e_j))_{1\leq i,j\leq n}$, donde B es la forma bilineal simétrica asociada a Q, se denomina *matriz asociada* a Q con respecto a la base canónica.

Es inmediato que $B(x,y) = xHy^t$, $\forall x,y \in \mathbb{R}^N$ (donde $x \equiv (x_1,\dots,x_N)$), o equivalentemente $Q(x) = xHx^t$, $\forall x \in \mathbb{R}^N$.
Resaltemos que H es también la matriz asociada, con respecto a la base canónica, al único operador lineal autoadjunto T asociado a Q. En efecto, de

$$(T(x)|y) = (x|T(y)) = B(x,y) = xHy^t = (xH|y),\ \forall x,y \in \mathbb{R}^N,$$

se deduce que $T(x) = xH$, $\forall x \in \mathbb{R}^N$, y por tanto ($H = H^t$)

$$(T(x))^t = Hx^t,\ \forall x \in \mathbb{R}^N.$$

Un *autovector* de Q es un vector u no nulo tal que $T(u) = \lambda u$ para algún real λ. Al número real λ, que es claramente único, se le llama su *autovalor* . Equivalentemente (Proposición 7.21):

$$B(x,u) = \lambda(x|u),\ \forall x \in \mathbb{R}^N.$$

El producto escalar euclídeo verifica también la desigualdad de Cauchy-Schwarz (Apéndice A de la lección 3):

5. $(x|y)^2 \leq (x|x)(y|y)\ \forall x, y \in \mathbb{R}^N$,

que como pone de manifiesto el siguiente enunciado, es consecuencia de las anteriores.

Proposición 7.23 (Desigualdad de Cauchy-Schwarz)**.** *Toda forma bilineal simétrica positiva B verifica:*

$$B(x,y)^2 \leq B(x,x)B(y,y),\ \forall x, y \in \mathbb{R}^N.$$

Demostración:

Sean $x, y \in \mathbb{R}^N$. Se tiene:

$$0 \leq B(x+\lambda y, x+\lambda y) = B(y,y)\lambda^2 + 2B(x,y)\lambda + B(x,x),\ \forall \lambda \in \mathbb{R},$$

con lo que

$$0 \geq \Delta = 4B(x,y)^2 - 4B(x,x)B(y,y),$$

y por tanto

$$B(x,y)^2 \leq B(x,x)B(y,y).$$

■

Sabemos que en dimensión finita las formas bilineales son continuas y por tanto las formas cuadráticas también lo son. La propiedad de compacidad y la, en apariencia ingenua, desigualdad anterior nos van a permitir demostrar la existencia de autovectores.

Lema 7.24 (Existencia de autovectores)**.** *Toda forma cuadrática tiene un autovector.*

Demostración:

Sea Q una forma cuadrática y denotemos por B a su forma bilineal simétrica asociada. La propiedad de compacidad asegura la existencia de un vector u de la esfera unidad euclídea $S_2(\mathbb{R}^N)$ tal que

$$\lambda := B(u,u) \geq B(x,x),\ \forall x \in S_2(\mathbb{R}^N).$$

Para probar que u es un autovector y λ su autovalor, definimos

$$F(x,y) := \lambda(x|y) - B(x,y),\ \forall x, y \in \mathbb{R}^N.$$

Es claro que F es una forma bilineal simétrica positiva por lo que verifica la desigualdad de Cauchy-Schwarz:

$$F(x,y)^2 \leq F(x,x)F(y,y),\ \forall x, y \in \mathbb{R}^N$$

y como $F(u,u) = 0$, se sigue que $F(x,u) = 0$, $\forall x \in \mathbb{R}^N$, es decir

$$B(x,u) = \lambda(x|u),\ \forall x \in \mathbb{R}^N.$$

■

Recordemos que una aplicación lineal $F : \mathbb{R}^N \to \mathbb{R}^N$ es una <u>*transformación ortogonal*</u> si es una isometría para la norma euclídea:

$$\|F(x)\|_2 = \|x\|_2,\ \forall x \in \mathbb{R}^N,$$

equivalentemente (en virtud de la fórmula de polarización en la Proposición 7.21 *i*)) si conserva el producto escalar:

$$(F(x)|F(y)) = (x|y),\ \forall x, y \in \mathbb{R}^N.$$

El siguiente resultado nos permite conocer mejor estas aplicaciones

Proposición 7.25. *Sean $F : \mathbb{R}^N \to \mathbb{R}^N$ una aplicación lineal, y A la matriz asociada a F con respecto a la base canónica:*

$$(F(x))^t = Ax^t,\ \forall x \in \mathbb{R}^N.$$

Equivalen las siguientes afirmaciones:

i) F es ortogonal.

ii) $A^tA = I$. En consecuencia A (y por tanto F) es inversible y $A^{-1} = A^t$.

iii) F transforma bases ortonormales en bases ortonormales.

iv) F es la aplicación cambio entre dos bases ortonormales.

Demostración:

i) $\Rightarrow$ *ii*) El elemento (i, j) de la matriz A^tA es:

$$e_i(A^tA)e_j^t = (e_iA^t)(Ae_j^t) = (F(e_i))(F(e_j))^t = (F(e_i)|F(e_j)) = (e_i|e_j) = \delta_{ij}.$$

ii) $\Rightarrow$ *iii*) Sea $\{u_1, \ldots, u_N\}$ una base ortonormal. Se tiene

$$(F(u_i)|F(u_j)) = (F(u_i))(F(u_j))^t = (u_iA^t)(Au_j^t) = u_i(A^tA)u_j^t = u_iu_j^t = (u_i|u_j) = \delta_{ij}.$$

iii) $\Rightarrow$ *iv*) Por hipótesis $\{F(e_1), \ldots, F(e_N)\}$ es una base ortonormal. Del hecho de que F es el operador de cambio de esta base a la canónica se sigue inmediatamente el siguiente desarrollo:

$$F(x_1, \ldots, x_N) = (y_1, \ldots, y_N) \Leftrightarrow F\left(\sum_{i=1}^N x_ie_i\right) = \sum_{i=1}^N y_ie_i \Leftrightarrow \sum_{i=1}^N x_iF(e_i) = \sum_{i=1}^N y_ie_i.$$

iv) $\Rightarrow$ *i*) Sean $\{u_1, \ldots, u_N\}, \{v_1, \ldots, v_N\}$ bases ortonormales tales que

$$F(x_1, \ldots, x_N) = (y_1, \ldots, y_N) \Leftrightarrow \sum_{i=1}^N x_iu_i = \sum_{i=1}^N y_iv_i.$$

Entonces

$$\sum_{i=1}^N x_i^2 = \left\|\sum_{i=1}^N x_iu_i\right\|_2^2 = \left\|\sum_{i=1}^N y_iv_i\right\|_2^2 = \sum_{i=1}^N y_i^2 \Rightarrow \|x\|_2 = \|F(x)\|_2,\ \forall x \in \mathbb{R}^N.$$

■

Teorema 7.26 (principal)**.** *Para cada forma cuadrática Q en $\mathbb{R}^N$ existe una base ortonormal $\{u_1,\ldots,u_N\}$ de $\mathbb{R}^N$ compuesta de autovectores de Q. Además si $\{\lambda_1,\ldots,\lambda_N\}$ son sus correspondientes autovalores, entonces*

$$Q(x)=\lambda_1 a_1^2+\ldots+\lambda_N a_N^2,\ \forall x=a_1u_1+\ldots+a_Nu_N\in\mathbb{R}^N.$$

En términos de matrices, $H=A^{-1}\Delta A$ donde H es la matriz asociada a Q con respecto a la base canónica,

$$\Delta=\begin{pmatrix}\lambda_1 & 0 & \ldots & 0\\ 0 & \lambda_2 & \ldots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \ldots & \lambda_N\end{pmatrix}$$

y A es la matriz inversible asociada a la función cambio de la base canónica a la base $\{u_1,\ldots,u_N\}$. En consecuencia,

$$\det(H)=\lambda_1\ldots\lambda_N.$$

Demostración:

Sean B,T respectivamente la forma bilineal simétrica y el operador autoadjunto asociados a Q. La demostración es por inducción sobre la dimensión del espacio.

Para dimensión uno se tiene para $\lambda=B(1,1)$ que $Q(x)=\lambda x^2,\ \forall x\in\mathbb{R}$ con lo que

$$T(x)=\lambda x,\ \forall x\in\mathbb{R}$$

y el vector 1 es un autovector, y por tanto en este caso es la base canónica la que "diagonaliza", siendo innecesario cambiar de base.

Consideremos ahora un natural $N>1$. El lema anterior nos asegura que Q tiene un autovector u_1 que supondremos de norma uno, es decir, para conveniente λ_1 real

$$B(x,u_1)=\lambda_1(x|u_1),\ \forall x\in\mathbb{R}^N.$$

Sea

$$V=\{u_1\}^\perp:=\{y\in\mathbb{R}^N:(u_1|y)=0\}.$$

V es un subespacio vectorial de $\mathbb{R}^N$ de dimensión $N-1$ y la restricción de Q a V es, obviamente, una forma cuadrática. Si $\{v_1,\ldots,v_{N-1}\}$ es una base ortonormal fijada de V, entonces es claro que el isomorfismo $\varphi:\mathbb{R}^{N-1}\to V$ que aplica la base canónica de $\mathbb{R}^{N-1}$ en la base de V fijada

$$\varphi(x_1,\ldots,x_{N-1}):=x_1v_1+\ldots+x_{N-1}v_{N-1}$$

preserva el producto escalar

$$(x|y)=(\varphi(x)|\varphi(y)),\ \forall x,y\in\mathbb{R}^{N-1}.$$

Vía φ, podemos considerar la forma cuadrática restricción de Q a V como una forma cuadrática en $\mathbb{R}^{N-1}$, esto es, podemos definir

$$Q_\varphi(x):=Q(\varphi(x)),\ \forall x\in\mathbb{R}^{N-1}$$

para obtener una forma cuadrática en $\mathbb{R}^{N-1}$ cuya forma bilineal simétrica asociada es:

$$B_\varphi(x,y) := B(\varphi(x),\varphi(y)),\ \forall x,y \in \mathbb{R}^{N-1}.$$

La hipótesis de inducción aplicada a B_φ nos asegura que $\mathbb{R}^{N-1}$ tiene una base ortonormal $\{w_2,\ldots,w_N\}$ de autovectores de Q_φ. Es decir, para cada $i = 2,\ldots,N$ hay un conveniente número real λ_i tal que

$$B_\varphi(y,w_i) = \lambda_i(y|w_i),\ \forall y \in \mathbb{R}^{N-1}.$$

Veamos ahora que $\varphi(w_2),\ldots,\varphi(w_n)$ son también autovectores de Q. Ello es consecuencia de que la anterior igualdad puede escribirse en la forma

$$B(v,\varphi(w_i)) = \lambda_i(v|\varphi(w_i)),\ \forall v \in V,$$

y de que para cada $x \in \mathbb{R}^N$, existen α real y $v \in V$, tales que $x = \alpha u_1 + v$, con lo que para $i = 2,\ldots,N$ se tiene:

$$B(x,\varphi(w_i)) = \alpha B(u_1,\varphi(w_i)) + B(v,\varphi(w_i)) =$$

$$\alpha\lambda_1(u_1|\varphi(w_i)) + \lambda_i(v|\varphi(w_i)) = \lambda_i(v|\varphi(w_i)) =$$

$$\lambda_i(\alpha u_1|\varphi(w_i)) + \lambda_i(v|\varphi(w_i)) = \lambda_i(x|\varphi(w_i)),$$

donde se ha utilizado que u_1 es un autovector de B y que $\varphi(w_i) \in V$. Consideremos para $i = 2,\ldots,N$, $u_i = \varphi(w_i)$; y veamos que entonces $\{u_1,\ldots,u_N\}$ es la base buscada. En efecto, dicha base es ortonormal y además para todo $x = a_1u_1 + \ldots + a_Nu_N$ se tiene que

$$Q(x) = B(x,x) = B(a_1u_1 + \ldots + a_Nu_N,x) = a_1B(u_1,x) + \ldots + a_NB(u_N,x) =$$

$$\lambda_1a_1(u_1|x) + \ldots + \lambda_Na_N(u_N|x) = \lambda_1a_1^2 + \ldots + \lambda_Na_N^2.$$

En consecuencia, si F es la transformación ortogonal en $\mathbb{R}^N$ determinada por el cambio de la base canónica a la base ortonormal $\{u_1,\ldots,u_N\}$, y A es su matriz asociada, se sigue que

$$Q(x) = (F(x))\Delta(F(x))^t = xA^t\Delta Ax^t,$$

con lo que $H = A^t\Delta A$, y basta recordar que $A^t = A^{-1}$. ■

Es claro que las formas cuadráticas diagonales

$$Q(x) = \alpha_1x_1^2 + \ldots + \alpha_Nx_N^2$$

se clasifican a simple vista (según el signo de los coeficientes). Como el teorema principal asegura que toda forma cuadrática "es" diagonal sin más que cambiar de base, se sigue inmediatamente el siguiente criterio de clasificación.

Proposición 7.27 (Primer criterio de clasificación de formas cuadráticas)**.** *Sean Q una forma cuadrática en $\mathbb{R}^N$ y $\lambda_1,\ldots,\lambda_N$ sus autovalores. Entonces se tiene:*

$$Q \text{ es definida} \left\{ \begin{array}{c} \textit{positiva} \\ \textit{negativa} \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{c} \lambda_k > 0 \\ \lambda_k < 0 \end{array} \right\} k = 1,\ldots,N$$

$$Q \text{ es} \left\{ \begin{array}{c} \textit{positiva} \\ \textit{negativa} \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{c} \lambda_k \geq 0 \\ \lambda_k \leq 0 \end{array} \right\} k = 1,\ldots,N$$

$$Q \text{ es indefinida } \Leftrightarrow \exists p,q \in \{1,\ldots,N\} : \lambda_p < 0 < \lambda_q$$

Volviendo a la definición de autovectores y autovalores de una forma cuadrática Q en $\mathbb{R}^N$ (Definición 7.22), es inmediato que, si H es la matriz asociada a Q con respecto a la base canónica, I es la matriz identidad y λ es un número real, entonces u es un autovector de Q con autovalor λ si, y sólo si, sus coordenadas son solución no nula del sistema homogéneo

$$(\lambda I - H) \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} = 0.$$

La existencia de solución no nula para tal sistema sabemos que equivale, por la regla de Cramer, a que $\det(\lambda I - H) = 0$. Así, como consecuencia del Teorema principal, los N autovalores de Q son las N raíces (reales) del polinomio de grado N, $\det(\lambda I - H)$.

Definición 7.28 (polinomio característico)**.** Sea Q una forma cuadrática en $\mathbb{R}^N$. El polinomio de grado N con coeficientes reales $P(\lambda) = \det(\lambda I - H)$, donde I es la matriz identidad y H es la matriz asociada a Q con respecto a la base canónica, se denomina *polinomio característico* de Q.

La dificultad de aplicación del primer criterio estriba en conocer el signo de las raíces del polinomio característico. Empezaremos probando un resultado de polinomios con coeficientes y raíces reales que nos permita salvar esta dificultad.

Proposición 7.29. *Sea $P(\lambda) = \lambda^N - A_1\lambda^{N-1} + A_2\lambda^{N-2} + \ldots + (-1)^N A_N$ un polinomio mónico con coeficientes reales y raíces reales $\lambda_1,\ldots,\lambda_N$. Equivalen las siguientes afirmaciones:*

i) $\lambda_k > 0,\ \forall k \in \{1,\ldots,N\}$.

ii) $A_k > 0,\ \forall k \in \{1,\ldots,N\}$.

Asimismo equivalen también:

iii) $\lambda_k \geq 0,\ \forall k \in \{1,\ldots,N\}$.

iv) $A_k \geq 0,\ \forall k \in \{1,\ldots,N\}$.

Demostración:

Como $P(\lambda) = (\lambda - \lambda_1)\dots(\lambda - \lambda_N)$, al desarrollar e igualar coeficientes se tiene que A_k es la suma de todos los productos de k elementos del conjunto $\{\lambda_1,\dots,\lambda_N\}$, $k = 1,\dots,N$, de donde se deduce la implicación $i) \Rightarrow ii)$. Para probar la implicación contraria basta observar que $\lambda \leq 0 \Rightarrow P(\lambda) \neq 0$. En efecto: $P(0) = (-1)^N A_N \neq 0$, y si $\lambda < 0$ y N es par (resp. impar) todos los sumandos de $P(\lambda)$ son positivos (resp. negativos).

La equivalencia $iii) \Leftrightarrow iv)$ se prueba análogamente. ∎

El siguiente criterio nos permite clasificar una forma cuadrática si conocemos el signo de los coeficientes del polinomio característico (¡no hace falta conocer el signo de las raíces de dicho polinomio!)

Proposición 7.30 (Segundo criterio de clasificación de formas cuadráticas)**.** *Sean Q una forma cuadrática en $\mathbb{R}^N$ y*

$$P(\lambda) = \lambda^N - A_1\lambda^{N-1} + A_2\lambda^{N-2} + \dots + (-1)^N A_N$$

su polinomio característico. Entonces se tiene:

$$Q \text{ es definida } \left\{ \begin{array}{c} \text{positiva} \\ \text{negativa} \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{c} A_k > 0 \\ (-1)^k A_k > 0 \end{array} \right\} k = 1,\dots,N$$

$$Q \text{ es } \left\{ \begin{array}{c} \text{positiva} \\ \text{negativa} \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{c} A_k \geq 0 \\ (-1)^k A_k \geq 0 \end{array} \right\} k = 1,\dots,N$$

$$Q \text{ es indefinida } \Leftrightarrow \exists p,q \in \{1,\dots,N\} : A_p, (-1)^q A_q < 0$$

(En consecuencia, Q es indefinida si para un k par A_k es negativo).

Demostración:

El primer criterio de clasificación de formas cuadráticas (Proposición 7.27) nos asegura que Q es definida positiva si, y sólo si, $\lambda_k > 0$, $k = 1,\dots,N$, lo que equivale a su vez, en virtud de la Proposición 7.29, a que $A_k > 0$, $k = 1,\dots,N$.

Para obtener ahora el criterio para formas definidas negativas nótese que Q es definida negativa si, y sólo si, $-Q$ es definida positiva, y que si H denota la matriz asociada a Q con respecto a la base canónica, entonces el polinomio característico de $-Q$ es

$$\det\,(\lambda I - (-H)) = \det\,((-1)(-\lambda I - H)) = (-1)^N \det\,(-\lambda I - H) =$$

$$(-1)^N P(-\lambda) = \lambda^N + A_1\lambda^{N-1} + A_2\lambda^{N-2} + \dots + A_N.$$

Finalmente en los casos en que la forma cuadrática es positiva o negativa la demostración es análoga. ∎

Es bien conocido que si $H = \left(a_{ij}\right)_{1\leq i,j\leq N}$ es la matriz asociada con respecto a la base canónica de la forma cuadrática Q en $\mathbb{R}^N$, entonces el polinomio característico de Q

$$P(\lambda) = \lambda^N - A_1\lambda^{N-1} + A_2\lambda^{N-2} + \ldots + (-1)^N A_N$$

es tal que A_k $(1 \leq k \leq N)$ es la suma de todos los menores principales de H de orden k. En consecuencia, la aplicación del criterio anterior conlleva el cálculo de un elevado número de determinantes. El siguiente criterio es más fácil de aplicar (requiere calcular menos determinantes). Recuérdese el concepto de submatrices principales dado en la Definición 7.11.

Proposición 7.31 (Condición de Sylvester)**.** *Sea Q una forma cuadrática en $\mathbb{R}^N$, y $H = \left(a_{ij}\right)_{1\leq i,j\leq N}$ la matriz de Q. Entonces se tiene:*

$$Q \textit{ es definida} \left\{ \begin{array}{c} \textit{positiva} \\ \textit{negativa} \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{c} \det\,(H_k) > 0 \\ (-1)^k \det\,(H_k) > 0 \end{array} \right\} k = 1, \ldots, N$$

$$Q \textit{ es} \left\{ \begin{array}{c} \textit{positiva} \\ \textit{negativa} \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{c} \det\,(K) \geq 0 \\ (-1)^{orden(K)} \det\,(K) \geq 0 \end{array} \right\} \forall K \textit{ submatriz principal de } H$$

$$Q \textit{ es indefinida } \Leftrightarrow \exists K, L \textit{ subm. princ. de } H : \det\,(K) < 0 \textit{ y } (-1)^{orden(L)} \det\,(L) < 0$$

En consecuencia,
Q es indefinida si $\det\,(K) < 0$ *para alguna submatriz principal K de orden par.*

Demostración:
$\Rightarrow$] Basta tener en cuenta que si $Q(x)$ es definida (resp. semidefinida) positiva/negativa entonces las formas cuadráticas que resultan de hacer nulas algunas coordenadas del vector genérico x mantienen el carácter definido (resp. semidefinido) positivo/negativo. Aunque la nueva forma cuadrática cambie de autovalores, en virtud del Teorema 7.26, los nuevos autovalores siguen conservando el signo y su producto es igual al determinante de la matriz asociada, que es una submatriz principal de H.

$\Leftarrow$] La demostración de esta implicación puede verse en [Stro, §6.2, §6.3]. Utiliza el método de eliminación de Gauss (procedimiento de factorización triangular). ∎

7.3. Referencias recomendadas.

[Stra], [MaHo].

7.4. Resumen de resultados del Tema 7.

Extremo relativo. Sea f un campo escalar definido en $A \subset \mathbb{R}^N$. Se dice que f alcanza en un punto $a \in A$ un máximo relativo si existe un entorno abierto U de a tal que $U \subset A$ y

$$f(x) \leq f(a),\ \forall x \in U.$$

Se dice que f alcanza en el punto $a \in A$ un máximo relativo estricto si existe un entorno abierto U de a tal que $U \subset A$ y

$$f(x) < f(a),\ \forall x \in U \backslash \{a\}.$$

Análogas definiciones se hacen para mínimos. La expresión <u>extremo relativo</u> significa o bien máximo relativo, o bien mínimo relativo.

Condición necesaria de existencia de extremo relativo. Puntos críticos. *Sea f un campo escalar definido en $A \subset \mathbb{R}^N$. Si f alcanza un extremo relativo en un punto $a \in A$ y f tiene gradiente en a, entonces $\nabla f(a) = 0$. Más aún, si f alcanza un extremo relativo en un punto $a \in A$ y f tiene derivada direccional en a según el vector u, entonces $f'(a;u) = 0$.*

Los puntos de A donde f tiene gradiente nulo se llaman <u>puntos críticos</u> de f. Es decir son las soluciones del sistema de ecuaciones (no necesariamente lineal) $\nabla f(x) = 0$ definido en los puntos en los que f tiene gradiente.

Condiciones necesarias y suficientes de existencia de extremo relativo. *Sean f un campo escalar definido en $A \subset \mathbb{R}^N$ y k un natural mayor o igual que 2. Supongamos que f es k veces derivable en un punto crítico $a \in A$, que todas las derivadas parciales hasta las de orden $k-1$ son nulas y que alguna derivada parcial de orden k no es nula.*

Se verifican las siguientes afirmaciones:

a) *Si k es impar, entonces f no alcanza extremo relativo en a.*

b) *Si k es par, se tiene:*

   i) *Si $d^k f(a)(x) > 0,\ \forall x \neq 0$, entonces f alcanza en a un mínimo relativo estricto.*

   ii) *Si $d^k f(a)(x) < 0,\ \forall x \neq 0$, entonces f alcanza en a un máximo relativo estricto.*

   iii) *Si $d^k f(a)(x) \geq 0,\ \forall x \in \mathbb{R}^N$, entonces de alcanzar f en a un extremo relativo, ha de ser mínimo.*

   iv) *Si $d^k f(a)(x) \leq 0,\ \forall x \in \mathbb{R}^N$, entonces de alcanzar f en a un extremo relativo, ha de ser máximo.*

   v) *Si $[\exists x, y \in \mathbb{R}^N : d^k f(a)(x) < 0 < d^k f(a)(y)]$, entonces f no alcanza ningún extremo relativo en a.*

Las condiciones necesarias y suficientes de existencia de extremo relativo ponen de relieve el interés que tiene para nosotros el conocimiento de criterios que permitan decidir el "signo" del polinomio $d^k f(a)$.

Formas cuadráticas y su clasificación. Submatrices principales. Una forma cuadrática en $\mathbb{R}^N$ es una aplicación $Q : \mathbb{R}^N \to \mathbb{R}$ de la forma $Q(x) = xH_Qx^t$, $\forall x \in \mathbb{R}^N$, donde $H_Q \in \mathscr{M}_{N\times N}(\mathbb{R})$ es una matriz simétrica (que se prueba ser única).

Una forma cuadrática Q en $\mathbb{R}^N$ se dice que es
definida positiva si $Q(x) > 0$, $\forall x \in \mathbb{R}^N \backslash \{0\}$,
(semidefinida) positiva si $Q(x) \geq 0$, $\forall x \in \mathbb{R}^N$,

Análogas definiciones se dan para forma cuadrática definida negativa y (semidefinida) negativa.
indefinida si toma valores positivos y negativos, es decir ni es (semidefinida) positiva ni es (semidefinida) negativa.

Dados N, k naturales, con $k \leq N$ y $H = (a_{ij})_{1\leq i,j\leq N}$ una matriz cuadrada de orden N, se dice que una matriz K es una submatriz principal de H de orden k si existe una aplicación estrictamente creciente $\sigma : \{1,\ldots,k\} \to \{1,\ldots,N\}$ tal que $K = \left(a_{\sigma(i)\sigma(j)}\right)_{1\leq i,j\leq k}$.

Como consecuencia del criterio de Sylvester (Proposición 7.12), y de las condiciones necesarias y suficientes de existencia de extremo relativo, se obtiene el siguiente resultado práctico.

Condiciones necesarias y suficientes de existencia de extremo relativo. *Sea f un campo escalar definido en $A \subset \mathbb{R}^N$. Supongamos que f es dos veces derivable en un punto crítico $a \in A$ y que*

$$H_f(a) = \begin{pmatrix} D_{(1,1)}f(a) & \ldots & D_{(N,1)}f(a) \\ \ldots & \ldots & \ldots \\ D_{(1,N)}f(a) & \ldots & D_{(N,N)}f(a) \end{pmatrix} \neq 0.$$

Para $k = 1,\ldots,N$, sean $H_k = \left(D_{(i,j)}f(a)\right)_{1\leq i,j\leq k}$ las submatrices principales "esquinas".

Se verifican entonces las siguientes afirmaciones:

i) Si $\det(H_k) > 0$, $k = 1,\ldots,N$, entonces f alcanza en a un mínimo relativo estricto.

ii) Si $(-1)^k \det(H_k) > 0$, $k = 1,\ldots,N$, entonces f alcanza en a un máximo relativo estricto.

iii) Si $\det(K) \geq 0$, para cualquier submatriz principal K de $H_f(a)$, entonces de alcanzar f en a un extremo relativo ha de ser mínimo.

iv) Si $(-1)^{orden(K)} \det(K) \geq 0$, para cualquier submatriz principal K de $H_f(a)$, entonces de alcanzar f en a un extremo relativo, ha de ser máximo.

v) Si existen K y L submatrices principales de $H_f(a)$ tales que $\det(K) < 0$ y $(-1)^{orden(L)} \det(L) < 0$ (lo que en particular ocurre si el determinante de alguna submatriz principal de orden par de $H_f(a)$ es negativo), entonces f no alcanza en a un extremo relativo.

7.5. Ejercicios del Tema 7

7.1 Una función real f definida en un abierto y convexo A de $\mathbb{R}^N$ se dice <u>*convexa*</u> si verifica:

$$f((1-\alpha)x+\alpha y) \leq (1-\alpha)f(x)+\alpha f(y),\ \forall x,y \in A,\ \forall \alpha \in [0,1].$$

i) Se supone que f es derivable en A. Probar que:

$$[f \text{ es convexa}] \Leftrightarrow [f(x) \geq f(a)+Df(a)(x-a),\ \forall x,a \in A].$$

Deducir que si f es convexa y $Df(a)=0$, entonces f alcanza en a un mínimo absoluto.

ii) Se supone que f es dos veces derivable en A. Probar que:

$$[f \text{ es convexa}] \Leftrightarrow [D^2 f(a)(x,x) \geq 0,\ \forall a \in A,\ \forall x \in \mathbb{R}^N].$$

iii) Probar que la función $f :]-1,1[\times]-1,1[\to \mathbb{R}$ definida por

$$f(x,y) = -\sqrt{1-x^2}\sqrt{1-y^2}$$

es convexa. Estudiar sus extremos.

7.2 (*) Estudiar los extremos relativos y absolutos de las siguientes funciones:

$$f_1(x,y) = \frac{x-y}{1+x^2+y^2}\ ;\ f_2(x,y) = \operatorname{sen} x \cos y\ ;\ f_3(x,y) = \operatorname{sen}(xy)\ ;$$

$$f_4(x,y) = x^3+y^3-3axy\ (a \in \mathbb{R})\ ;\ f_5(x,y) = x^4+y^4-4a^2xy\ (a \in \mathbb{R}^*)\ ;$$

$$f_6(x,y) = (x^2+2y^2)e^{-(x^2+y^2)}\ ;\ f_7(x,y) = \frac{xy}{(1+x)(1+y)(1+xy)}\ \ (x,y>0)\ ;$$

$$f_8(x,y,z) = x^2+y^2+3z^2-xy+2xz+yz\ ;\ f_9(x,y,z) = xy+xz+yz\ .$$

7.3 Estudiar los extremos relativos de:

i) Dado $a>0$, la función $f : \mathbb{R}^+ \times \mathbb{R}^+ \times \overset{N}{\ldots} \times \mathbb{R}^+ \to \mathbb{R}$ definida por:

$$f(x_1,x_2,\ldots,x_N) = x_1x_2\ldots x_N + a^{N+1}\sum_{i=1}^{N}\frac{1}{x_i}\ .$$

ii) Dados $a_1,a_2,\ldots,a_N \in \mathbb{R}$, la función $f : \mathbb{R}^N \to \mathbb{R}$ definida por:

$$f(x_1,x_2,\ldots,x_N) = \left(\sum_{i=1}^{N} a_ix_i\right)e^{-\|x\|_2^2}\ .$$

(Si quiere simplificarse algo, hágase para $N=2$ o $N=3$.)

7.4 Dados n puntos $(a_i, b_i) \in \mathbb{R}^2$, determinar las condiciones que han de cumplir α y β para que la recta γ de ecuación $y = \alpha x + \beta$ sea tal que la cantidad

$$S(\alpha,\beta) = \sum_{i=1}^{N} (b_i - \alpha a_i - \beta)^2$$

sea mínima.

7.5 Dados m puntos $a_i \in \mathbb{R}^N$ calcular el menor valor de la función

$$f(x) := \sum_{i=1}^{m} \|x - a_i\|_2^2.$$

7.6 **Teorema de Rolle generalizado.**

Sean A un subconjunto abierto acotado de $\mathbb{R}^N$ y $f : \overline{A} \to \mathbb{R}$ una función continua en $\overline{A}$, constante en $\mathrm{Fr}(A)$ y derivable en A. Probar que $Df(x) = 0$ para algún x de A.

7.7 Sea $f : \mathbb{R}^2 \to \mathbb{R}$ definida por $f(x,y) = x^2y - 4x^2 - 2y^2$. Justificar que para todo subconjunto compacto $K \subset \mathbb{R}^2$ la función f alcanza su valor mínimo en K en un punto de la frontera de K, es decir:

$$\min f(K) = \min f(Fr(K)).$$

7.8 (*) Sean f un campo escalar continuo definido en un abierto A de $\mathbb{R}^N$ y a un punto de A. Supongamos que f es derivable en $A\backslash\{a\}$, que existe $\delta > 0$ tal que $B(a,\delta) \subset A$ y que se verifica

$$(x - a|\nabla f(x)) < 0 \ \ (\text{resp. } > 0),\ \forall x \in B(a,\delta)\backslash\{a\}\ .$$

Entonces f alcanza un máximo (resp. mínimo) relativo estricto en el punto a.

7.9 Sea $f : \mathbb{R}^2 \to \mathbb{R}$ la función dada por $f(x,y) := x^2 + (1-x)^3y^2$. Demostrar que dicha función tiene un único punto crítico en el que alcanza un mínimo estricto y que no alcanza ningún extremo absoluto. ¿Puede presentarse tal situación para una función de $\mathbb{R}$ en $\mathbb{R}$?

7.10 Sean $A = \{(x,y,z) \in \mathbb{R}^3 : x+y+z = 1\}$ y $f : A \to \mathbb{R}$ la función dada por $f(x,y,z) = xyz$. Estudiar los extremos "locales" de f.

Indicación: Puede usarse la Nota 7.4. Mejor es reformular el problema en los siguientes términos: de la condición $x+y+z = 1$, despejamos $z = 1-x-y$, con lo que se nos pide es hallar los extremos de la función $f : \mathbb{R}^2 \to \mathbb{R}$ definida por $f(x,y) := xy(1-x-y)$.

7.6. Soluciones a los ejercicios del Tema 7.

7.1 *i)* $\boxed{\Rightarrow}$ Sean $a,x \in A$. Si $a = x$ no hay nada que demostrar. En otro caso para $0 < \alpha \leq 1$, se tiene

$$\frac{f(a+\alpha(x-a))-f(a)}{\alpha} \leq \frac{(1-\alpha)f(a)+\alpha f(x)-f(a)}{\alpha} = f(x)-f(a)$$

y basta tomar límite cuando $\alpha \to 0^+$ para obtener

$$Df(a)(x-a) \leq f(x)-f(a).$$

$\boxed{\Leftarrow}$ Sean $x,y \in A$. Queremos probar que para $0 \leq \alpha \leq 1$, se verifica

$$f((1-\alpha)x+\alpha y) \leq (1-\alpha)f(x)+\alpha f(y).$$

Notemos $a = (1-\alpha)x+\alpha y$. Por hipótesis

$$\left.\begin{array}{l} f(x) \geq f(a)+Df(a)(x-a) \\ f(y) \geq f(a)+Df(a)(y-a) \end{array}\right\} \Rightarrow$$

$$(1-\alpha)f(x)+\alpha f(y) \geq f(a)+Df(a)((1-\alpha)(x-a)+\alpha(y-a)) = f(a) = f((1-\alpha)x+\alpha y)$$

En consecuencia, si f es convexa y $Df(a) = 0$, se tiene

$$Df(a) = 0 \Rightarrow \ f(a) \leq f(x), \ \forall x \in A,$$

es decir f alcanza en a un mínimo absoluto.

ii) $\boxed{\Rightarrow}$ Sean $a \in A, x \in S_{\mathbb{R}^N}$. Existe $\delta > 0$ tal que $0 < t < \delta \Rightarrow \ a+tx \in A$. Por i):

$$0 \leq \frac{f(a+tx)-f(a)-Df(a)(tx)}{t^2} =$$

$$\frac{f(a+tx)-f(a)-Df(a)(tx)-\frac{1}{2}D^2f(a)(tx,tx)}{t^2} + \frac{1}{2}D^2f(a)(x,x)$$

y al tomar límite cuando $t \to 0$, la fórmula infinitesimal del resto nos asegura que $0 \leq D^2f(a)(x,x)$.

$\boxed{\Leftarrow}$ La fórmula de Taylor con resto de Lagrange para funciones valuadas en $\mathbb{R}$, nos asegura que si $a,x \in A$, con $a \neq x$, entonces $\exists c \in]a,b[$ tal que

$$f(x)-f(a)-Df(a)(x-a) = \frac{1}{2}D^2f(c)(x-a,x-a)$$

con lo que por hipótesis $f(x)-f(a)-Df(a)(x-a) \geq 0$, que nos dice que f es convexa por lo demostrado en i).

iii) Es inmediato comprobar que el único punto crítico es el $(0,0)$. Calculemos la matriz hessiana

$$H_f(x,y) = \begin{pmatrix} \frac{1}{1-x^2}\frac{\sqrt{1-y^2}}{\sqrt{1-x^2}} & \frac{-xy}{\sqrt{1-x^2}\sqrt{1-y^2}} \\ \frac{-xy}{\sqrt{1-x^2}\sqrt{1-y^2}} & \frac{1}{1-y^2}\frac{\sqrt{1-x^2}}{\sqrt{1-y^2}} \end{pmatrix}.$$

En consecuencia para todo $(x,y) \in]-1,1[\times]-1,1[$, se tiene

$$\det H_1 > 0,\ \det H_2 = \frac{1-x^2y^2}{(1-x^2)(1-y^2)} > 0,$$

de donde deducimos por el criterio de Sylvester que $d^2 f(x,y)$ es definida positiva para todo $(x,y) \in]-1,1[\times]-1,1[$. El apartado $ii)$ nos dice que f es convexa. Claramente f alcanza en $(0,0)$ un mínimo absoluto estricto.

7.2 $\boxed{f_1}$

$$\left.\begin{array}{l} \frac{\partial f}{\partial x} = 0 \\ \frac{\partial f}{\partial y} = 0 \end{array}\right\} \Leftrightarrow \left\{\begin{array}{l} x^2 = y^2 \Leftrightarrow \ |x| = |y| \\ 2xy+1 = 0 \Rightarrow \ xy < 0 \end{array}\right.$$

Puntos críticos: $A = (\frac{\sqrt{2}}{2}, \frac{-\sqrt{2}}{2})$; $B = (\frac{-\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$.

Matriz hessiana: Haciendo uso de que las derivadas parciales primeras son nulas en los puntos críticos, basta trabajar con la matriz

$$\Delta(x,y) = \begin{pmatrix} -x+y & x+y \\ -x-y & y-x \end{pmatrix}; \det\ (\Delta(x,y)) = 2(x^2+y^2)$$

(aunque no es simétrica en los puntos en cuestión si lo es)

Solución:

f alcanza en A un máximo relativo estricto.
f alcanza en B un mínimo relativo estricto.

Extremos absolutos:

$$|f(\rho\cos\vartheta, \rho\,\mathrm{sen}\,\vartheta)| = \left|\frac{\rho}{1+\rho^2}(\cos\vartheta - \mathrm{sen}\,\vartheta)\right| \le \frac{2\rho}{1+\rho^2} \ \to\ 0,\ (\rho \to\ +\infty),$$

y por tanto $\lim_{\|(x,y)\|\to+\infty} f(x,y) = 0$.

Si una función alcanza valores mayores (resp. menores) que el límite en ∞, entonces alcanza su máximo (resp mínimo) absoluto que claramente son también relativos.

f alcanza en A su máximo absoluto igual a $\frac{\sqrt{2}}{2}$
f alcanza en B su mínimo absoluto igual a $\frac{-\sqrt{2}}{2}$

En efecto, sea $R > 1$ tal que $\|(x,y)\|_2 > R \Rightarrow |f(x,y)| < \frac{1}{2}$. La continuidad de f en $\overline{B}((0,0),R)$ nos asegura que f tiene máximo y mínimo absoluto en $\overline{B}((0,0),R)$, y en consecuencia f tiene máximo y mínimo absoluto (que se alcanza en $\overline{B}((0,0),R)$).

$\boxed{f_2}$

$$\left.\begin{array}{l} \frac{\partial f}{\partial x} = 0 \\ \frac{\partial f}{\partial y} = 0 \end{array}\right\} \Leftrightarrow \left\{\begin{array}{l} \cos x \cos y = 0 \\ \operatorname{sen} x \operatorname{sen} y = 0 \end{array}\right\} \Leftrightarrow \left\{\begin{array}{l} \cos x = \operatorname{sen} y = 0 \\ \text{o} \\ \operatorname{sen} x = \cos y = 0 \end{array}\right.$$

<u>Puntos críticos</u>: $A = \left((2p+1)\frac{\pi}{2}, q\pi\right); B = \left(p\pi, (2q+1)\frac{\pi}{2}\right),\ p,q \in \mathbb{Z}$.

<u>Matriz hessiana</u>:

$$H_f(x,y) = \begin{pmatrix} -\operatorname{sen} x \cos y & -\cos x \operatorname{sen} y \\ -\cos x \operatorname{sen} y & -\operatorname{sen} x \cos y \end{pmatrix} \text{ ; det } (H_f(x,y)) = \operatorname{sen}^2 x - \operatorname{sen}^2 y$$

<u>Solución</u>: Puntos A: Si $p+q \in 2\mathbb{Z}$ (p,q tienen la misma paridad), entonces hay un máximo relativo estricto. En caso contrario hay un mínimo relativo estricto. En los puntos B no hay extremo pues det $(H_f) < 0$. La superficie es como un cartón de huevos.

<u>Extremos absolutos</u>: f está acotada por 1 y toma valores 1 y -1, luego tiene máximo y mínimo absolutos iguales a 1 y -1, que son alcanzados en infinitos puntos.

$\boxed{f_3}$

$$\left.\begin{array}{l} \frac{\partial f}{\partial x} = 0 \\ \frac{\partial f}{\partial y} = 0 \end{array}\right\} \Leftrightarrow \left\{\begin{array}{l} y \cos xy = 0 \\ x \cos xy = 0 \end{array}\right.$$

<u>Puntos críticos</u>: $A = (0,0); B = (a,b)$ con $ab = \frac{(2k+1)\pi}{2}, k \in \mathbb{Z}$

<u>Matriz hessiana</u>:

$$H_f(x,y) = \begin{pmatrix} -y^2 \operatorname{sen}(xy) & \cos(xy) - xy \operatorname{sen}(xy) \\ \cos(xy) - xy \operatorname{sen}(xy) & -x^2 \operatorname{sen}(xy) \end{pmatrix}$$

<u>Solución</u>: En A no alcanza extremo. En los puntos B, si $k \in 2\mathbb{Z}$ hay un posible máximo relativo, y si $k \notin 2\mathbb{Z}$ hay un posible mínimo relativo.

En los puntos B hay máximo y mínimo pues $|\operatorname{sen}(xy)| \leq 1$.

<u>Extremos absolutos</u>: f está acotada por 1 y toma valores 1 y -1, luego tiene máximo y mínimo absolutos iguales a 1 y -1, que son alcanzados en infinitos puntos.

$\boxed{f_4}$

$$\left.\begin{array}{l} \frac{\partial f}{\partial x} = 0 \\ \frac{\partial f}{\partial y} = 0 \end{array}\right\} \Leftrightarrow \left\{\begin{array}{l} 3(x^2 - ay) = 0 \\ 3(y^2 - ax) = 0 \end{array}\right.$$

<u>Puntos críticos</u>: $A = (0,0); B = (a,a)$

<u>Matriz hessiana</u>:

$$H_f(x,y) = \begin{pmatrix} 6x & -3a \\ -3a & 6y \end{pmatrix}$$

<u>Solución</u>: En el punto A no hay extremo pues si $a = 0$ se anulan la primera y segunda derivada y no la tercera. En el punto B, si $a > 0$ hay un mínimo relativo, y si $a < 0$ hay un máximo relativo.

<u>Extremos absolutos</u>: f no está ni mayorada ni minorada ($f(x,0) = x^3,\ \forall x \in \mathbb{R}$), luego no tiene ni máximo ni mínimo absolutos.

$\boxed{f_5}$

$$\left.\begin{array}{l} \frac{\partial f}{\partial x} = 0 \\ \frac{\partial f}{\partial y} = 0 \end{array}\right\} \Leftrightarrow \left\{\begin{array}{l} 4(x^3 - a^2 y) = 0 \\ 4(y^3 - a^2 x) = 0 \end{array}\right.$$

<u>Puntos críticos</u>:$A = (0,0); B = (a,a); C = (-a,-a)$

<u>Matriz hessiana</u>:

$$H_f(x,y) = \begin{pmatrix} 12x^2 & -4a^2 \\ -4a^2 & 12y^2 \end{pmatrix}$$

<u>Solución</u>: f no alcanza en A un extremo y alcanza en B y C un mínimo relativo estricto.

<u>Extremos absolutos</u>: f no está mayorada ($f(x,0) = x^4,\ \forall x \in \mathbb{R}$), luego no tiene máximo absoluto (también por no tener máximo relativo).

$$f(\rho\cos\vartheta, \rho \operatorname{sen}\vartheta) = \rho^4(\cos^4\vartheta + \operatorname{sen}^4\vartheta) - 4a^2\rho^2\cos\vartheta\operatorname{sen}\vartheta \geq$$

$$\rho^4(\cos^4\vartheta + \operatorname{sen}^4\vartheta) - 4a^2\rho^2 \geq \rho^2(\rho^2\alpha - 4a^2) \ \rightarrow\ +\infty, \text{ cuando } \rho \ \rightarrow\ +\infty,$$

siendo $\alpha = \operatorname{Min}\{\cos^4\vartheta + \operatorname{sen}^4\vartheta :\ \vartheta \in [0,2\pi]\} > 0$.

Sea $R > 0$ tal que $\|(x,y)\| > R \Rightarrow\ f(x,y) > 1$. La continuidad de f en $\overline{B}((0,0),R)$ nos asegura que f tiene mínimo absoluto en $\overline{B}((0,0),R)$ que será $\leq f(0,0) = 0$, y en consecuencia f tiene mínimo absoluto (que se alcanza en $\overline{B}((0,0),R)$, de hecho lo alcanza en los puntos B y C).

$\boxed{f_6}$

$$\left.\begin{array}{l} \frac{\partial f}{\partial x} = 0 \\ \frac{\partial f}{\partial y} = 0 \end{array}\right\} \Leftrightarrow \left\{\begin{array}{l} 2(x - x^3 - 2xy^2)e^{-(x^2+y^2)} = 0 \\ 2(2y - x^2 y - 2y^3)e^{-(x^2+y^2)} = 0 \end{array}\right.$$

<u>Puntos críticos</u>: $A = (0,0); B = (0,1); C = (0,-1); D = (1,0); E = (-1,0)$

<u>Matriz hessiana</u>: Derivando con "picardía" sabiendo lo que se anula y suprimiendo factores que no influyen en el signo:

$$H_f(x,y) \simeq \begin{pmatrix} 1-3x^2-2y^2 & -4xy \\ -4xy & 2-x^2-6y^2 \end{pmatrix}$$

(aunque no es simétrica, en los puntos en cuestión si lo es).

<u>Solución</u>: En A hay un mínimo relativo estricto, en B y C un máximo relativo estricto y en D y E no hay extremo.

<u>Extremos absolutos</u>: $0 \leq f(x,y),\ \forall (x,y) \in \mathbb{R}^2$, de hecho $f(x,y) > 0$ si $(x,y) \neq (0,0)$, luego tiene mínimo absoluto (que lo alcanza en el punto A). Es claro que $\lim_{\|(x,y)\|\to+\infty} f(x,y) = 0$, de donde, como se vió para f_1, deducimos que f tiene máximo absoluto (que lo alcanza en los puntos B y C).

$\boxed{f_7}$

$$\left.\begin{array}{l} \frac{\partial f}{\partial x} = 0 \\ \frac{\partial f}{\partial y} = 0 \end{array}\right\} \Leftrightarrow \left\{\begin{array}{l} \frac{y(1+y)(1-x^2y)}{d^2} = 0 \\ \frac{x(1+x)(1-xy^2)}{d^2} = 0 \end{array}\right.$$

<u>Puntos críticos</u>: $A = (1,1)$

<u>Matriz hessiana</u>:

$$H_f(x,y) \cong \begin{pmatrix} -2xy & 1-x^2y(2-y) \\ 1-xy^2(2-x) & -2xy \end{pmatrix}$$

<u>Solución</u>: En A hay un máximo relativo estricto.

<u>Extremos absolutos</u>: Podemos extender de manera continua f a $\mathbb{R}_0^+ \times \mathbb{R}_0^+$ definiéndola como 0 en la frontera. Además

$$0 \leq f(x,y) < \frac{1}{(1+x)(1+y)} \ \to\ 0, \text{ cuando } \|(x,y)\| \to\ +\infty$$

y por tanto $\lim_{\|(x,y)\|\to+\infty} f(x,y) = 0$. Así f tiene máximo absoluto en $\mathbb{R}_0^+ \times \mathbb{R}_0^+$, que al no ser 0 estará en $\mathbb{R}^+ \times \mathbb{R}^+$ (de hecho lo alcanza en A).

Como f no alcanza mínimo relativo en $\mathbb{R}^+ \times \mathbb{R}^+$, tampoco tiene mínimo absoluto en $\mathbb{R}^+ \times \mathbb{R}^+$.

$\boxed{f_8}$

$$\left.\begin{array}{l} \frac{\partial f}{\partial x} = 0 \\ \frac{\partial f}{\partial y} = 0 \end{array}\right\} \Leftrightarrow \left\{\begin{array}{l} 2x-y+2z = 0 \\ -x+2y+z = 0 \\ 2x+y+6z = 0 \end{array}\right.$$

Puntos críticos: $A(0,0,0)$

Matriz hessiana:

$$H_f(x,y,z) = \begin{pmatrix} 2 & -1 & 2 \\ -1 & 2 & 1 \\ 2 & 1 & 6 \end{pmatrix} \text{ ; det } (H_f(x,y,z)) = 4 \text{ ; etc.}$$

Solución: En A hay un mínimo relativo estricto.

Extremos absolutos: f coincide con su polinomio de Taylor de orden 2 en cualquier punto (Ejercicio 6.6), es decir: $2f = (x\ y\ z)H(x\ y\ z)^t$ que es una forma cuadrática definida positiva. En consecuencia

$$0 = f(0,0,0) < f(x,y,z), \ \forall (x,y,z) \neq (0,0,0)$$

luego f tiene mínimo absoluto (que lo alcanza en $(0,0,0)$).

Como $f(x,0,0) = x^2$, f no está mayorada y no tiene máximo absoluto (si queremos también se deduce de no tener ningún máximo relativo).

$\boxed{f_9}$

$$\left.\begin{array}{l} \frac{\partial f}{\partial x} = 0 \\ \frac{\partial f}{\partial y} = 0 \end{array}\right\} \Leftrightarrow \left\{\begin{array}{l} y+z=0 \\ x+z=0 \\ x+y=0 \end{array}\right.$$

Puntos críticos: $A = (0,0,0)$

Matriz hessiana:

$$H_f(x,y,z) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \text{ ; det } (H_f(x,y)) = 2 \text{ ; etc.}$$

Solución: En A no se alcanza extremo (det $(H_2) = -1$).

Extremos absolutos: f no está ni mayorada ni minorada ($f(x,1,0) = x$, $\forall x \in \mathbb{R}$), luego no tiene ni máximo ni mínimo absolutos (también se puede razonar que al no tener extremos relativos no puede tener extremos absolutos).

7.3 $\boxed{i)}$

$$D_i f(x) = 0 \Leftrightarrow \frac{a^{N+1}}{x_i} = x_1 \dots x_N \text{ , } i = 1, \dots, N.$$

Punto crítico: $A = (a, \dots, a)$.

$$[D_{ii} f(A) = 2a^{N-2} \text{ , } D_{ij} f(A) = a^{N-2}] \Rightarrow H_f(A) = a^{N-2} H,$$

donde $H = \begin{pmatrix} 2 & 1 & 1 & \dots & 1 \\ 1 & 2 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & 2 \end{pmatrix}$ y

$$\det(H) = \begin{vmatrix} 2 & 1 & 1 & \dots & 1 \\ 1 & 2 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & 2 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & \dots & 1 \\ 0 & 1 & 0 & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ -1 & -1 & -1 & \dots & 2 \end{vmatrix} =$$

$$\begin{vmatrix} 1 & 0 & 0 & \dots & 0 & 1 \\ 0 & 1 & 0 & \dots & 0 & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & -1 & -1 & \dots & -1 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & \dots & 0 & 1 \\ 0 & 1 & 0 & \dots & 0 & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & -1 & \dots & -1 & 4 \end{vmatrix} =$$

$$\begin{vmatrix} 1 & 0 & 0 & \dots & 0 & 1 \\ 0 & 1 & 0 & \dots & 0 & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -1 & N \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & \dots & 0 & 1 \\ 0 & 1 & 0 & \dots & 0 & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & N+1 \end{vmatrix} = N+1$$

(en el primer paso se resta la última columna de las demás, en el segundo se le suma a la última fila la primera fila, en el tercero se le suma a la última fila la segunda fila, ..., con lo que se triangula la matriz). Repitiendo el proceso con las otras submatrices principales esquinas izquierda obtenemos que sus determinantes valen respectivamente $2,3,\dots,N+1$. El criterio de Sylvester nos asegura que

<u>Solución</u>: f alcanza en A un mínimo relativo estricto. Con un razonamiento topológico consistente en considerar cuadrados muy grandes del primer cuadrante (en dimensión dos) se concluye que también es absoluto.

$\boxed{ii)}$ Suponemos que algún a_i es no nulo (pues en otro caso f es la función nula). Se tiene

$$D_i f(x) = 0 \Leftrightarrow a_i = 2(a|x)x_i \ , \ i = 1,\dots,n$$

luego (multiplicando por a_i y sumando las n expresiones)

$$\|a\|_2^2 = 2(a|x)^2 \Leftrightarrow \|a\|_2 = \sqrt{2}|(a|x)|$$

es decir

$$2(a|x) = \sqrt{2}\|a\|_2 \text{ o } 2(a|x) = -\sqrt{2}\|a\|_2$$

y por tanto:

<u>Puntos críticos</u>: $A = \frac{a}{\sqrt{2}\|a\|_2}; B = \frac{-a}{\sqrt{2}\|a\|_2}$

(conviene comprobar las soluciones en el sistema original).

Calculando las derivadas parciales segundas, se obtiene:

$$H_f(A) = \frac{-\sqrt{2}}{\|a\|_2} e^{-\frac{1}{2}} H_0, \ \ H_f(B) = \frac{\sqrt{2}}{\|a\|_2} e^{-\frac{1}{2}} H_0,$$

donde

$$H_0 = \|a\|_2^2 I + \begin{pmatrix} a_1^2 & a_1a_2 & \dots & a_1a_n \\ \dots & \dots & \dots & \\ a_1a_n & a_2a_n & \dots & a_n^2 \end{pmatrix}$$

con lo que $xH_0x^t = \|a\|_2^2\ \|x\|_2^2 + (a|x)^2$.

Solución: A: Definida negativa, luego máximo relativo estricto. B: Definida positiva, luego mínimo relativo estricto.

7.4 "La suma de los cuadrados de las diferencias entre las ordenadas de los puntos y las correspondientes ordenadas de la recta sea mínima" (recta de mínimos cuadrados).

$$\left.\begin{array}{l} \frac{\partial S}{\partial \alpha} = 0 \\ \frac{\partial S}{\partial \beta} = 0 \end{array}\right\} \Leftrightarrow \left\{ \begin{array}{l} \alpha\|a\|_2^2 + \beta\sum_{i=1}^n a_i = (a|b) \\ \alpha\sum_{i=1}^n a_i + n\beta = \sum_{i=1}^n b_i \end{array}\right. \qquad (*)$$

donde $a = (a_1, .., a_n)$, $b = (b_1, .., b_n)$.

El sistema es lineal (cosa rara) y su determinante es $n\|a\|_2^2 - (\sum_{i=1}^n a_i)^2$. La desigualdad de Cauchy-Schwarz nos dice

$$\left(\sum_{i=1}^n a_i\right)^2 = (a|\mathbf{1})^2 \le \|a\|_2^2\,\|\mathbf{1}\|_2^2 = n\|a\|_2^2 \qquad (**)$$

con lo que el determinante del sistema es no negativo y es nulo si, y sólo si, los a_i son constante ($a = C\mathbf{1}$).

En este caso los n puntos están en una recta paralela al eje OY y hay infinitas rectas solución de $(*)$, todas las rectas que pasan por $(C, \overline{b})$, con $\overline{b} = \frac{1}{n}\sum_{i=1}^n b_i$ (media de ordenadas), ya que $\overline{b}$ es el punto en que la función $x \to \sum_{i=1}^n (b_i - x)^2$ toma su mínimo absoluto. Es decir, las rectas $y = \alpha x + \beta$ con $\overline{b} = \alpha C + \beta$.

Supongamos ahora que al menos hay dos valores a_i distintos. En este caso hay una única solución de $(*)$:

$$\alpha_0 = \frac{(a \mid b) - n\overline{a}\overline{b}}{\|a\|_2^2 - n\overline{a}^2}, \beta_0 = \frac{\|a\|_2^2\overline{b} - (a \mid b)\overline{a}}{\|a\|_2^2 - n\overline{a}^2}$$

Ahora

$$H_S(\alpha, \beta) = 2\begin{pmatrix} \sum_{i=1}^n a_i^2 & \sum_{i=1}^n a_i \\ \sum_{i=1}^n a_i & n \end{pmatrix}$$

define una forma cuadrática definida positiva (por $(**)$), luego S tiene un mínimo relativo estricto y absoluto en (α_0, β_0).

También, de

$$S(\alpha,\beta) = (\sum_{i=1}^n a_i^2)\alpha^2 + n\beta^2 + 2(\sum_{i=1}^n a_i)\alpha\beta - 2(\sum_{i=1}^n a_ib_i)\alpha - 2(\sum_{i=1}^n b_i)\beta + \sum_{i=1}^n b_i^2,$$

se sigue que $S(\alpha,\beta) \to +\infty$ cuando $\|(\alpha,\beta)\|_\infty \to +\infty$, y, por tanto, podemos asegurar que

$$\exists K > 0 : \ \|(\alpha,\beta)\|_\infty > K \Rightarrow \ S(\alpha,\beta) > S(\alpha_0,\beta_0)$$

Así, al alcanzar por compacidad la restricción de la función S al compacto

$$\{(\alpha,\beta) \in \mathbb{R}^2 : \|(\alpha,\beta)\|_\infty \leq K\}$$

un mínimo absoluto, podemos asegura que la función S alcanza un mínimo absoluto que es también relativo por lo que a la fuerza ha de alcanzarse en (α_0,β_0).

7.5

$$D_k f(x) = 2\sum_{i=1}^{m}(x(k) - a_i(k)) = 0 \Leftrightarrow \ x(k) = \frac{1}{m}\sum_{i=1}^{m} a_i(k) \ , \ k = 1,\ldots,n.$$

Punto crítico: $u = \frac{1}{m}\sum_{i=1}^{m} a_i$ (baricentro de los m puntos).

Un razonamiento topológico nos asegura que la función f alcanza en el punto u un mínimo absoluto. De $f(x) \geq \|x - a_1\|_2^2 \geq (\|x\|_2 - \|a_1\|_2)^2$, se sigue que $\lim_{\|x\|_2 \to +\infty} f(x) = +\infty$ y por tanto f tiene un mínimo absoluto. También se puede razonar directamente como sigue:

$$\begin{aligned}
f(x) &= \sum_{i=1}^{m} \|x - a_i\|_2^2 = \sum_{i=1}^{m} \|(x-u) + (u-a_i)\|_2^2 \\
&= \sum_{i=1}^{m} ((x-u) + (u-a_i) \ \mid \ (x-u) + (u-a_i)) \\
&= \sum_{i=1}^{m} \|x-u\|_2^2 + \sum_{i=1}^{m} \|u-a_i\|_2^2 + 2\sum_{i=1}^{m} (x-u \mid u-a_i) \\
&= m\|x-u\|_2^2 + \sum_{i=1}^{m} \|u-a_i\|_2^2 + 2\left(x-u \mid \sum_{i=1}^{m}(u-a_i)\right) \\
&= m\|x-u\|_2^2 + \sum_{i=1}^{m} \|u-a_i\|_2^2 \\
&\geq \sum_{i=1}^{m} \|u-a_i\|_2^2 = f(u), \ \forall x \in \mathbb{R}^n.
\end{aligned}$$

7.6 **Teorema de Rolle generalizado.**

$\overline{A}$ es compacto, luego la función f alcanza el máximo y el mínimo. Razonamos por disyunción de casos:

- Si la función f alcanza ambos en la frontera, son iguales el valor máximo y el mínimo de f que, por tanto, es constante y la derivada se anula en cualquier punto de A.

- Si alguno no lo alcanza en la frontera, al alcanzarse en un punto interior la condición necesaria de extremo relativo nos asegura que $\exists a \in A : \ Df(a) = 0$.

7.7 Puntos críticos:

$$\left.\begin{array}{r} 2xy-8x=0 \\ x^2-4y=0 \end{array}\right\}$$

$A=(0,0); B=(4,4); C=(-4,4)$.

Matriz hessiana:

$$H_f(0,0)=\begin{pmatrix} -8 & 0 \\ 0 & -4 \end{pmatrix}; H_f(4,4)=\begin{pmatrix} 0 & 8 \\ 8 & -4 \end{pmatrix}; H_f(-4,4)=\begin{pmatrix} 0 & -8 \\ -8 & -4 \end{pmatrix}.$$

Solución: f alcanza en A un máximo relativo estricto y no alcanza extremo en B y C.

Así el mínimo no se puede alcanzar en el interior de K luego se tiene que alcanzar en la frontera.

7.8 Este ejercicio generaliza el bien conocido hecho para funciones reales de variable real:

$$f'(x)(x-a)\left\{\begin{array}{c} <0 \\ >0 \end{array}\right\} \Rightarrow f \text{ tiene en } a \text{ un } \left\{\begin{array}{c} \text{máximo} \\ \text{mínimo} \end{array}\right\} \text{ relativo estricto.}$$

Queremos probar que

$$x\in B(a,\delta)\backslash\{a\} \Rightarrow\ f(x)<f(a)\ (\text{resp. } f(x)>f(a)).$$

Sea x un tal punto. Aplicamos el T.V.M. a la función f en el segmento $[a,x]$ para obtener $c=a+t(x-a)\in]a,x[$ (con $0<t<1$) tal que

$$f(x)-f(a)=(\nabla f(c)\mid x-a)=\left(\nabla f(c)\Big|\frac{c-a}{t}\right)=\frac{1}{t}(\nabla f(c)|c-a)<0\ (\text{resp. } >0).$$

7.9 Los puntos críticos son solución del sistema

$$\left.\begin{array}{r} 2x-3(1-x)^2y^2=0 \\ 2(1-x)^3y^2=0 \end{array}\right\} \overset{x\neq 1}{\Leftrightarrow} x=y=0.$$

Veamos que $x\neq 1$: si $x=1\Rightarrow\ x=0$ absurdo.

Punto crítico: $A=(0,0)$.

$$H_f(0,0)=\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Solución: f alcanza en $(0,0)$ un mínimo relativo estricto.

Por otra parte al verificarse que

$$f(0,n)=n^2\ \to\ +\infty \quad\text{y}\quad f(2,n)=4-n^2\ \to\ -\infty,$$

concluimos que f no tiene extremos absolutos.

En funciones de $\mathbb{R}$ en $\mathbb{R}$ no se puede presentar la anterior situación. En efecto si f alcanza en un punto a un mínimo relativo estricto y suponemos que no es mínimo absoluto, entonces toma valores mayores y menores que $f(a)$ en una de las semirrectas $]-\infty, a[$ o $]a, +\infty[$. El teorema del valor intermedio nos asegura que también toma el valor $f(a)$. Finalmente concluimos por el teorema de Rolle que f tiene un punto crítico distinto de a.

7.10 La solución más fácil es plantearse hallar los extremos de $f : \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x,y) := xy(1-x-y)$.

Puntos críticos:

$$\left.\begin{array}{r} y(1-2x-y)=0 \\ x(1-x-2y)=0 \end{array}\right\}$$

$a=(1,0)$, $b=(0,1)$, $c=(0,0)$, $d=\left(\frac{1}{3},\frac{1}{3}\right)$.

Matriz hessiana:

$$H_f(x,y) = \begin{pmatrix} -2y & 1-2x-2y \\ 1-2x-2y & -2x \end{pmatrix}$$

con lo que

$$H_f(1,0) = \begin{pmatrix} 0 & -1 \\ -1 & -2 \end{pmatrix} \;;\; H_f(0,1) = \begin{pmatrix} -2 & -1 \\ -1 & 0 \end{pmatrix};$$

$$H_f(0,0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \;;\; H_f\left(\frac{1}{3},\frac{1}{3}\right) = \begin{pmatrix} \frac{-2}{3} & \frac{-1}{3} \\ \frac{-1}{3} & \frac{-2}{3} \end{pmatrix}.$$

Solución: f no alcanza ningún extremo en los puntos a, b y c pues $\det(H_2) < 0$, y f alcanza un máximo relativo y absoluto en d.

Otra solución: Decimos que a es un máximo local de $f : A \subset \mathbb{R}^3 \to \mathbb{R}$ si existe un entorno abierto U de a (en $\mathbb{R}^3$) tal que

$$f(x) \le f(a), \; \forall x \in U \cap A.$$

Análogamente se define mínimo local y el adjetivo estricto tiene el significado usual. Un extremo local es un máximo o un mínimo local.

Las direcciones admisibles en el conjunto A vienen definidas por los vectores (u,v,w) no nulos tales que existe $\delta > 0$ verificando

$$x+tu+y+tv+z+tw = 1, \text{ si } |t| < \delta$$

equivalentemente

$$t(u+v+w) = 0, \text{ si } |t| < \delta \Leftrightarrow u+v+w = 0.$$

Consideramos f definida en todo $\mathbb{R}^3$. Se tiene para $(x,y,z) \in A$

$$Df(x,y,z)(u,v,w) = uyz + xvz + xyw, \forall (u,v,w) \in \mathbb{R}^3,$$

y si nos limitamos a las direcciones admisibles

$$0 = f'((x,y,z);(u,v,w)) = uyz + xvz + xyw,\ \forall (u,v,w) \in \mathbb{R}^3 : \begin{cases} u^2+v^2+w^2 > 0 \\ u+v+w = 0 \end{cases}$$

En consecuencia los posibles extremos locales son, por tanto, las soluciones del sistema

$$\begin{cases} uyz + xvz + xyw = 0 \\ x+y+z = 1 \end{cases} \quad \text{donde} \begin{cases} u^2+v^2+w^2 > 0 \\ u+v+w = 0 \end{cases}$$

Eligiendo como direcciones admisibles $(1,1,-2)$ y $(1,-2,1)$, se tiene que

$$\left.\begin{cases} yz + xz - 2xy = 0 \\ yz - 2xz + xy = 0 \end{cases}\right\} \Rightarrow\ xy = xz = yz,$$

de donde se deduce que el punto (x,y,z) puede tener dos coordenadas nulas o ninguna.

Posibles extremos locales: $a = (1,0,0)$, $b = (0,1,0)$, $c = (0,0,1)$, $d = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$.

Como la restricción de f al compacto

$$K := \{(x,y,z) \in \mathbb{R}_0^+ \times \mathbb{R}_0^+ \times \mathbb{R}_0^+ :\ x+y+z = 1\}$$

alcanza un máximo y un mínimo absoluto, concluimos que f alcanza en d un máximo local.

Veamos que f no alcanza en a, b y c un extremo. En efecto basta considerar por ejemplo las sucesiones $\{(1+\frac{2}{n}, \frac{-1}{n}, \frac{-1}{n})\}$ y $\{(1+\frac{1}{n}, \frac{1}{n}, \frac{-2}{n})\}$, ya que f toma, respectivamente, valores positivos y negativos.

Tema 8

Teoremas de la función inversa y de la función implícita.

Dedicamos esta lección al estudio de dos de los resultados más sobresalientes del cálculo diferencial: los Teoremas de la función inversa (8.5) y de la función implícita (8.12) para campos vectoriales, los cuales son formulaciones equivalentes del mismo resultado (véanse la demostración del Teorema 8.12 y el Apéndice). Ambos teoremas responden, como veremos, al "principio general", ya aludido, según el cual *una función derivable hereda localmente las propiedades de su derivada*.

De lo dicho se deduce que es una cuestión de gusto personal decidir por cual de ellos se ha de empezar. Nos inclinamos por el Teorema de la función inversa que se puede motivar a partir de los resultados conocidos para funciones reales de variable real.

La primera sección está dedicada al Teorema de la función inversa y la segunda al Teorema de la función implícita. Ambos resultado se motivan con ejemplos.

También se incluye un apéndice en el que se prueba el Teorema de la función inversa a partir del Teorema de la función implícita.

8.1. Teorema de la función inversa.

Empezamos recordando el siguiente resultado que es consecuencia del Teorema del valor medio real (4.1):

Sea I un intervalo y $f : I \to \mathbb{R}$ una función derivable con $f'(x) \neq 0,\ \forall x \in I$. Entonces f es estrictamente monótona y ocurre una de las dos posibilidades siguientes:

$$f'(x) > 0,\ \forall x \in I \ \ o \ \ f'(x) < 0,\ \forall x \in I.$$

Si además hacemos uso de la regla de derivación de la función inversa (Proposición 3.27) obtenemos

Teorema global de la función inversa para funciones reales. *Sean $I \subset \mathbb{R}$ un intervalo abierto y $f : I \to \mathbb{R}$ una función derivable con derivada distinta de cero en todo punto. Entonces, f es inyectiva (de hecho es estrictamente monótona) y su función inversa $f^{-1} : f(I) \to \mathbb{R}$ es derivable, siendo*

$$(f^{-1})'(f(x)) = \frac{1}{f'(x)},\ \forall x \in I.$$

Puesto que la imagen por una función continua y estrictamente monótona de un intervalo abierto es un intervalo abierto, se sigue que, en las condiciones del enunciado del teorema anterior, $f(I)$ es un intervalo abierto de $\mathbb{R}$ y la función biyectiva $f : I \to f(I)$ es mucho más que un homeomorfismo, ya que es derivable (por hipótesis) y su inversa también es derivable (por tesis). Formulamos a continuación este hecho dando una definición adecuada.

Definición 8.1 (difeomorfismo)**.** Sean A, B abiertos de $\mathbb{R}^N$.

Se dice que $f : A \to B$ es un *difeomorfismo* si f es biyectiva derivable y f^{-1} también es derivable. En tal caso escribiremos $f \in \mathrm{Dif}(A,B)$.

Se dice que $f : A \to \mathbb{R}^N$ es un *difeomorfismo en A* si $f(A)$ es un abierto de $\mathbb{R}^N$ y $f \in \mathrm{Dif}(A, f(A))$.

Se dice que $f : A \to \mathbb{R}^N$ es un *difeomorfismo local* si para cada $a \in A$ existe un entorno abierto U de a contenido en A tal que f es un difeomorfismo en U.

En caso de que $f \in \mathrm{Dif}(A,B)$, y que tanto f como f^{-1} sean de clase $\mathscr{C}^k$, diremos que f es un difeomorfismo de clase $\mathscr{C}^k$. Análogamente se definen los difeomorfismos en A de clase $\mathscr{C}^k$ y los difeomorfismos locales de clase $\mathscr{C}^k$.

Por ejemplo, la función exponencial es un difeomorfismo en $\mathbb{R}$ de clase $\mathscr{C}^\infty$.

El siguiente ejemplo muestra que en dimensión mayor que 1 no se puede esperar un resultado similar al resultado para funciones reales que acabamos de recordar.

Ejemplo 8.2. La función $f : \mathbb{R}^2 \to \mathbb{R}^2$ (determinada por la exponencial compleja) dada por

$$f(x,y) = (e^x \cos y \,,\, e^x \operatorname{sen} y) \quad \big((x,y) \in \mathbb{R}^2\big),$$

es una función derivable que tiene derivada inversible en todo punto

$$\det \big(J_f(x,y)\big) = e^{2x} \neq 0, \ \forall (x,y) \in \mathbb{R}^2,$$

y no es (globalmente) inyectiva

$$f(x,y) = f(x, y+2\pi), \ \forall (x,y) \in \mathbb{R}^2.$$

Por localización del teorema recordado obtenemos el siguiente resultado:

Teorema 8.3 (local de la función inversa para funciones reales)**.** *Sea $A \subset \mathbb{R}$ y sea $f : A \to \mathbb{R}$. Supongamos que f es de clase $\mathscr{C}^1$ en un punto $a \in \overset{\circ}{A}$ y que $f'(a) \neq 0$. Entonces existe un entorno abierto U de a contenido en A tal que $V := f(U)$ es un abierto de $\mathbb{R}$, f es un difeomorfismo de U sobre V, $f^{-1} : V \to U$ es una función de clase $\mathscr{C}^1$ en $f(a)$ y*

$$(f^{-1})'(f(x)) = \frac{1}{f'(x)}, \ \forall x \in U.$$

Demostración:

Por ser $f \in \mathscr{C}^1(a)$ y $f'(a) \neq 0$, existe un intervalo abierto U centrado en el punto a, contenido en A y tal que $f'(x) \neq 0$, $\forall x \in U$. Aplicando el Teorema global a la función $f_{|U}$, y teniendo en cuenta el comentario que sigue a dicho teorema, podemos asegurar que f es un difeomorfismo de U sobre V tal que

$$(f^{-1})'(f(x)) = \frac{1}{f'(x)}, \ \forall x \in U.$$

Finalmente, puesto que la anterior igualdad puede escribirse en la forma

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

donde por g denotamos a la función real inversión dada por $g(x) = \frac{1}{x}$, $\forall x \in \mathbb{R}^*$, se sigue que $f^{-1} \in \mathscr{C}^1(f(a))$ ya que $f \in \mathscr{C}^1(a)$. ∎

El siguiente ejemplo pone de manifiesto que es esencial exigir que f sea de clase $\mathscr{C}^1$ en a para obtener el resultado anterior.

Ejemplo 8.4. La función $f : \mathbb{R} \to \mathbb{R}$ definida por

$$f(x) = x + 2\,x^2 \operatorname{sen}\left(\frac{1}{x}\right) \ ; \ f(0) = 0$$

es derivable en todo punto con

$$f'(0)=1\ ,\ \ f'(x)=1+4\,x\ \operatorname{sen}\left(\frac{1}{x}\right)-2\ \cos\left(\frac{1}{x}\right)\ \ (x\neq 0)$$

(f' no es continua en 0). Veamos que la función f ni siquiera admite una inversa local en 0. En efecto, si existiese un intervalo abierto I, con $0\in I$, tal que la función f fuese inyectiva en I, entonces f sería (estrictamente) monótona en I, y en consecuencia f' sería de signo constante en I, pero esto es imposible pues $f'\left(\dfrac{1}{2n\pi}\right)=-1,\ \forall n\in\mathbb{N}$ y $f'(0)=1$.

El Teorema local (8.3) sí se generaliza para campos vectoriales. Habida cuenta que los ejemplos de campos escalares de clase $\mathscr{C}^1$ en un punto que no sean de clase $\mathscr{C}^1$ en un entorno de dicho punto no son frecuentes, enunciaremos el resultado para campos vectoriales de clase $\mathscr{C}^1$ en un abierto.

Teorema 8.5 (función inversa)**.** *Sean $A\subset\mathbb{R}^N$ un abierto y $f:A\to\mathbb{R}^N$ un campo vectorial de clase $\mathscr{C}^1$ (o lo que es igual sus campos escalares componentes tienen gradiente continuo). Supongamos que existe $a\in A$ tal que*

$$\det\big(J_f(a)\big)=\det\begin{pmatrix} D_1f_1(a) & \dots & D_Nf_1(a)\\ \dots & \dots & \dots\\ D_1f_N(a) & \dots & D_Nf_N(a)\end{pmatrix}\neq 0.$$

Entonces existe un entorno abierto U de a contenido en A tal que $V:=f(U)$ es un abierto de $\mathbb{R}^N$, f es un difeomorfismo de clase $\mathscr{C}^1$ de U sobre V, y para cada $x\in U$ se verifica

$$J_{f^{-1}}\big(f(x)\big)=\begin{pmatrix} D_1f_1(x) & \dots & D_Nf_1(x)\\ \dots & \dots & \dots\\ D_1f_N(x) & \dots & D_Nf_N(x)\end{pmatrix}^{-1}.$$

Además, si para k natural mayor que 1 se verifica que f es k veces derivable (resp. f es de clase $\mathscr{C}^k$) en a, entonces f^{-1} es k veces derivable (resp. f^{-1} es de clase $\mathscr{C}^k$) en $f(a)$.

Demostración:

Notemos $S=Df(a)^{-1}$, y consideremos el campo vectorial $\varphi:A\to\mathbb{R}^N$ dado por

$$\varphi:=i_A-S\circ f.$$

La linealidad de la derivada y la regla de la cadena nos aseguran que φ en de clase $\mathscr{C}^1$, siendo para cada $x\in A$

$$D\varphi(x)=Id_{\mathbb{R}^N}-S\circ Df(x).$$

Al ser $D\varphi(a)=0$ y $\det\big(J_f(a)\big)\neq 0$, se sigue que existe una bola abierta U centrada en a y contenida en A, tal que

$$\|D\varphi(x)\|\leq\frac{1}{2}\quad\text{y}\quad\det\big(J_f(x)\big)\neq 0,\ \forall x\in U.$$

El Corolario 4.7 nos asegura que φ es contractiva en U. Aplicando ahora el Teorema de Schauder (4.15, apartados i) y ii)) a la función $i_U-\varphi_{|U}=S\circ f_{|U}$, obtenemos que

i) $G := S(f(U))$ es un abierto de $\mathbb{R}^N$.

ii) $S \circ f$ es un homeomorfismo de U sobre G.

Puesto que $S \in \mathrm{Iso}(\mathbb{R}^N)$, se sigue finalmente que:

a) $V := f(U) = S^{-1}(G)$ es un abierto de $\mathbb{R}^N$.

b) $f = S^{-1} \circ (S \circ f)$ es un homeomorfismo de U sobre V.

La demostración del Teorema se concluye utilizando la regla de derivación de la inversa (Proposición 3.27 y el apartado 4 de la Sección 6.1). ■

Nota 8.6. La condición det $\big(J_f(a)\big) \neq 0$ equivale a afirmar que el sistema lineal

$$Df(a)(x) = y$$

tiene solución global única:

$$x = Df(a)^{-1}(y),\ \forall y \in \mathbb{R}^N.$$

El Teorema de la función inversa asegura la existencia de un entorno abierto U de a tal que $f(U)$ es un abierto de $\mathbb{R}^N$ y para cada $y = (y_1, y_2, \ldots, y_N) \in f(U)$ el sistema de N ecuaciones con N incógnitas

$$\begin{cases} f_1(x_1, x_2, \ldots, x_N) & = & y_1 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ f_N(x_1, x_2, \ldots, x_N) & = & y_N \end{cases}$$

tiene una única solución $x = (x_1, x_2, \ldots, x_N) \in U$. Además la función inversa

$$f^{-1} = \Big((f^{-1})_1, (f^{-1})_2, \ldots, (f^{-1})_N\Big)$$

así definida es de clase $\mathscr{C}^1$ en $f(U)$ (equivalentemente sus campos escalares componentes $(f^{-1})_1, (f^{-1})_2, \ldots, (f^{-1})_N$ tienen gradiente continuo).

Queda justificada la frase, tantas veces citada, f hereda localmente propiedades de su derivada.

Sabemos (Ejemplo 8.2) que para dimensión mayor que uno no es cierta la versión global del Teorema de la función inversa, sin embargo, si imponemos la inyectividad, sí probaremos una versión global, que requiere la siguiente:

Proposición 8.7. *Sean $A \subset \mathbb{R}^N$ un abierto y $f : A \to \mathbb{R}^N$ un difeomorfismo local, entonces f es una aplicación abierta.*

Demostración:

Sea $O \subset A$ abierto. Veamos que $f(O)$ es abierto. Por hipótesis para cada $x \in A$ existe U_x entorno abierto de x contenido en A tal que $V_x := f(U_x)$ es un abierto de $\mathbb{R}^N$ y f es un difeomorfismo de U_x sobre V_x. Se tiene entonces que $f(O \cap U_x)$ es un abierto relativo de V_x y por tanto abierto de $\mathbb{R}^N$. Por último al ser

$$f(O) = f\Big(\cup_{x \in O}\big(O \cap U_x\big)\Big) = \cup_{x \in O} f\big(O \cap U_x\big),$$

concluimos que $f(O)$ es abierto. ∎

Corolario 8.8 (Teorema "global" de la función inversa)**.** *Sean $A \subset \mathbb{R}^N$ un abierto, k un natural y $f : A \to \mathbb{R}^N$ un campo vectorial. Supongamos que f es de clase $\mathscr{C}^k$ en A y que*

$$\det\,\big(J_f(x)\big) \neq 0,\ \forall x \in A.$$

Entonces f es un difeomorfismo local de clase $\mathscr{C}^k$ en A, y por tanto una aplicación abierta. Si además f es inyectiva, entonces se tiene que f es un difeomorfismo de clase $\mathscr{C}^k$ en A.

Demostración:

Aplicando el Teorema de la función inversa (8.5) en cada punto de A se obtiene que f es un difeomorfismo local de clase $\mathscr{C}^k$ en A y por tanto, en virtud de la Proposición 8.7, $f(A)$ es abierto.

Si además f es inyectiva, en vista del carácter local de la derivabilidad, al ser f un difeomorfismo local de clase $\mathscr{C}^k$ concluimos que f es un difeomorfismo de clase $\mathscr{C}^k$ en A. ∎

Nota 8.9. Que f sea un difeomorfismo en un abierto $A \subset \mathbb{R}^N$ suele interpretarse como que dicha función define nuevas coordenadas en $f(A)$. De hecho es usual denominar *cambios de coordenadas* a los difeomorfismos: *las coordenadas según f de un punto y de $f(A)$ son las coordenadas cartesianas del punto x de A tal que $f(x) = y$.*

Aunque, en general cuando se aplica el Teorema de la función inversa, no se conoce la función f^{-1}, el enunciado del Teorema nos permite calcular su matriz jacobiana. Ilustramos este hecho con un ejemplo.

Ejemplo 8.10 (coordenadas polares)**.** Recordemos que $f : A = \mathbb{R}^+ \times]-\pi, \pi[\to \mathbb{R}^2$ definida por

$$f(\rho, \vartheta) = (\rho\ \cos\vartheta, \rho\ \operatorname{sen}\vartheta)$$

es la aplicación que da el cambio a coordenadas polares, que sabemos es inyectiva.

f es de clase $\mathscr{C}^\infty$ verificando

$$\det\,\big(J_f(\rho, \vartheta)\big) = \rho > 0,\ \forall (\rho, \vartheta) \in A.$$

El Corolario (8.8) garantiza que dicha aplicación es un difeomorfismo de clase $\mathscr{C}^\infty$ de A sobre $f(A) = \mathbb{R}^2 \setminus \{(x,0) :\ x \leq 0\}$. Además, para cada $(x,y) = (\rho\ \cos\vartheta, \rho\ \operatorname{sen}\vartheta)$ se verifica que

$$J_{f^{-1}}(x,y) = J_f(\rho,\vartheta)^{-1} = \begin{pmatrix} \cos\vartheta & -\rho\ \operatorname{sen}\vartheta \\ \operatorname{sen}\vartheta & \rho\ \cos\vartheta \end{pmatrix}^{-1} = \frac{1}{\rho}\begin{pmatrix} \rho\ \cos\vartheta & \rho\ \operatorname{sen}\vartheta \\ -\operatorname{sen}\vartheta & \cos\vartheta \end{pmatrix}$$

$$= \begin{pmatrix} \cos\vartheta & \operatorname{sen}\vartheta \\ \dfrac{-\operatorname{sen}\vartheta}{\rho} & \dfrac{\cos\vartheta}{\rho} \end{pmatrix} = \begin{pmatrix} \dfrac{x}{\sqrt{x^2+y^2}} & \dfrac{y}{\sqrt{x^2+y^2}} \\ \dfrac{-y}{x^2+y^2} & \dfrac{x}{x^2+y^2} \end{pmatrix}.$$

En el ejemplo anterior, de manera excepcional, se conoce la función inversa:

$$f^{-1}(x,y) = \left(\sqrt{x^2+y^2}\ ,\ 2\ \arctan\left(\frac{y}{x+\sqrt{x^2+y^2}}\right)\right),\ \forall (x,y) \in f(A),$$

es decir, *las coordenadas polares del punto* (x,y) .

¡Un buen ejercicio es calcular directamente $J_{f^{-1}}(x,y)$!

8.2. Teorema de la función implícita.

Nos ocupamos ahora de establecer el Teorema de la función implícita a partir del de la función inversa. Empezamos analizando dos ejemplos que nos van a motivar dicho teorema.

Ejemplos 8.11.

a) Sean $a,b,c \in \mathbb{R}$, con $b \neq 0$ y $f : \mathbb{R}^2 \to \mathbb{R}$ el campo escalar dado por

$$f(x,y) := ax + by + c.$$

Es inmediato que el conjunto

$$\{(x,y) \in \mathbb{R}^2 : \ f(x,y) = 0\}$$

es una recta no vertical, y por tanto para cada número real x existe un único número real y tal que $f(x,y) = 0$, es decir, dicha ecuación define implícitamente a y como función de x. De hecho, tal función es calculable en este caso:

$$y = \varphi(x) := -\frac{a}{b}\,x - \frac{c}{b}.$$

b) Sea $f : \mathbb{R}^2 \to \mathbb{R}$ el campo escalar dado por

$$f(x,y) = x^2 + y^2 - 1.$$

En este caso el conjunto

$$\{(x,y) \in \mathbb{R}^2 : \ f(x,y) = 0\}$$

es la circunferencia unidad, y por tanto la ecuación $f(x,y) = 0$ no define implícitamente a y como función de x, ya que si

$$\left.\begin{array}{lcc} |x| > 1 & \Rightarrow & \text{no existe ningún real } y \\ |x| = 1 & \Rightarrow & y = 0 \ \text{ es el único real} \\ |x| < 1 & \Rightarrow & \text{existen dos reales } (y = \sqrt{1-x^2}\ ,\ y = -\sqrt{1-x^2}) \end{array}\right\} \ : \ f(x,y) = 0.$$

Este ejemplo pone de manifiesto que tenemos que localizar para abordar con éxito los problemas de existencia y unicidad de funciones determinadas implícitamente, en el siguiente sentido. Sea $(a,b) \in \mathbb{R}^2$ con $f(a,b) = 0$.

i) Problema de existencia: ¿Existe un entorno abierto U de a y una función (continua, derivable,...) $\varphi : U \to \mathbb{R}$ verificando que

$$b = \varphi(a) \ \text{ y } \ f(x,\varphi(x)) = 0,\ \forall x \in U\ ?$$

ii) Problema de unicidad: ¿Es esta función única en el sentido de existir un entorno abierto Ω de (a,b) con la propiedad

$$\{(x,y)\in\Omega:\ f(x,y)=0\}=\{(x,\varphi(x)):\ x\in U\}?$$

La respuesta a ambas preguntas en este ejemplo es la siguiente:

- Si $|a|\geq 1$ no existe una tal φ.
- Si $|a|<1$

$$U=]-1,1[\text{ y }\begin{cases}\varphi(x)=\sqrt{1-x^2} & \text{con}\quad \Omega=\mathbb{R}\times\mathbb{R}^+ \text{ si}\quad b>0\\ \varphi(x)=-\sqrt{1-x^2} & \text{con}\quad \Omega=\mathbb{R}\times\mathbb{R}^- \text{ si}\quad b<0\end{cases}.$$

En este sencillo ejemplo vale el mismo entorno U para cualquier a con $|a|<1$. En general los entornos U y Ω dependen del punto (a,b) en cuestión.

El Teorema de la función implícita da condiciones suficientes para dar respuesta positiva a los problemas de existencia y unicidad de determinaciones implícitas.

Teorema 8.12 (función implícita)**.** *Sean $G\subset\mathbb{R}^N\times\mathbb{R}^M$ un conjunto abierto y $f:G\to\mathbb{R}^M$ un campo vectorial de clase $\mathscr{C}^1$ (o lo que es igual sus campos escalares componentes tienen gradiente continuo). Supongamos que existe $(a,b)\in G$ tal que $f(a,b)=0$ y que*

$$\det\begin{pmatrix} D_{N+1}f_1(a,b) & \dots & D_{N+M}f_1(a,b)\\ \dots & \dots & \dots\\ D_{N+1}f_M(a,b) & \dots & D_{N+M}f_M(a,b)\end{pmatrix}\neq 0.$$

Entonces existen un entorno abierto Ω de (a,b) contenido en G, un entorno abierto U de a y una función $\varphi:U\to\mathbb{R}^M$ de clase $\mathscr{C}^1$ tal que

$$\{(x,y)\in\Omega:\ f(x,y)=0\}=\{(x,\varphi(x)):\ x\in U\}.$$

Se tiene también que para todo $x\in U$ se verifica que

$$\det\begin{pmatrix} D_{N+1}f_1(x,\varphi(x)) & \dots & D_{N+M}f_1(x,\varphi(x))\\ \dots & \dots & \dots\\ D_{N+1}f_M(x,\varphi(x)) & \dots & D_{N+M}f_M(x,\varphi(x))\end{pmatrix}\neq 0.$$

y que

$$J_\varphi(x)=-\begin{pmatrix} D_{N+1}f_1(x,y) & \dots & D_{N+M}f_1(x,y)\\ \dots & \dots & \dots\\ D_{N+1}f_M(x,y) & \dots & D_{N+M}f_M(x,y)\end{pmatrix}^{-1}\begin{pmatrix} D_1f_1(x,y) & \dots & D_Nf_1(x,y)\\ \dots & \dots & \dots\\ D_1f_M(x,y) & \dots & D_Nf_M(x,y)\end{pmatrix}$$

donde $y=\varphi(x)$.

Además si para k natural mayor que 1 se verifica que f es k veces derivable (resp. f es clase $\mathscr{C}^k$) en (a,b) , entonces φ es k veces derivable (resp. φ es de clase $\mathscr{C}^k$) en a.

Demostración:

Consiste en aplicar el Teorema de la función inversa (8.5) en el punto (a,b) al campo vectorial $F : G \to \mathbb{R}^N \times \mathbb{R}^M$ definido por

$$F(x,y) := (x, f(x,y)),\ \forall (x,y) \in G.$$

F es de clase $\mathscr{C}^1$ con matriz jacobiana en cada $(x,y) \in G$ dada por

$$J_F(x,y) = \begin{pmatrix} & I & & & 0 & \\ D_1 f_1(x,y) & \dots & D_N f_1(x,y) & D_{N+1} f_1(x,y) & \dots & D_{N+M} f_1(x,y) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ D_1 f_M(x,y) & \dots & D_N f_M(x,y) & D_{N+1} f_M(x,y) & \dots & D_{N+M} f_M(x,y) \end{pmatrix}$$

Por tanto

$$\text{(8.2.1)} \qquad \det\big(J_F(x,y)\big) = \det \begin{pmatrix} D_{N+1} f_1(x,y) & \dots & D_{N+M} f_1(x,y) \\ \dots & \dots & \dots \\ D_{N+1} f_M(x,y) & \dots & D_{N+M} f_M(x,y) \end{pmatrix},\ \forall (x,y) \in G,$$

y en particular

$$\det\big(J_F(a,b)\big) \neq 0.$$

El Teorema de la función inversa nos garantiza la existencia de un entorno abierto Ω de (a,b) en $\mathbb{R}^N \times \mathbb{R}^M$ contenido en G tal que $F : \Omega \to F(\Omega)$ es un difeomorfismo de clase $\mathscr{C}^1$.

No quita generalidad suponer que $F(\Omega)$ es de la forma $U \times W$, donde U es un entorno abierto de a en $\mathbb{R}^N$ y W es un entorno abierto de cero en $\mathbb{R}^M$. En efecto, $(a,0) = F(a,b) \in F(\Omega)$ y si $r > 0$ es tal que $B_{\mathbb{R}^N \times \mathbb{R}^M}((a,0),r) \subset F(\Omega)$, entonces existe $s > 0$ tal que $B_{\mathbb{R}^N}(a,s) \times B_{\mathbb{R}^M}(0,s) \subset B_{\mathbb{R}^N \times \mathbb{R}^M}((a,0),r)$. Se tiene que

$$U := B_{\mathbb{R}^N}(a,s) \ ,\ \ W := B_{\mathbb{R}^M}(0,s) \ \text{ y } \ \Omega := F^{-1}(U \times W)$$

cumplen las condiciones requeridas.

Nótese que forzosamente el difeomorfismo inverso F^{-1} tendrá la forma

$$F^{-1}(x,z) = (x, h(x,z)),\ \forall (x,z) \in U \times W$$

para conveniente campo vectorial $h : U \times W \to \mathbb{R}^M$ de clase $\mathscr{C}^1$, pues F^{-1} es de clase $\mathscr{C}^1$. Definimos ahora el campo vectorial $\varphi : U \to \mathbb{R}^M$ por

$$\varphi(x) := h(x,0),\ \forall x \in U.$$

La función φ es de clase $\mathscr{C}^1$, y, como $(a,b) = F^{-1}(a,0) = (a, h(a,0)) = (a, \varphi(a))$, se verifica que $\varphi(a) = b$. Además para cada $x \in U$ se tiene que

$$(x, \varphi(x)) = (x, h(x,0)) = F^{-1}(x,0) \in \Omega \subset G$$

y también que

$$(x, f(x, \varphi(x))) = F(x, \varphi(x)) = F(x, h(x,0) = F\big(F^{-1}(x,0)\big) = (x,0),$$

en consecuencia

$$f(x,\varphi(x)) = 0, \ \forall x \in U.$$

Hemos probado que φ está determinada implícitamente por f.

Los cálculos que acabamos de hacer muestran la inclusión

$$\{(x,\varphi(x)) : \ x \in U\} \subset \{(x,y) \in \Omega : \ f(x,y) = 0\}.$$

La unicidad de φ consiste en justificar que la inclusión anterior es de hecho una igualdad:

$$\left.\begin{array}{r}(x,y) \in \Omega \\ f(x,y) = 0\end{array}\right\} \Rightarrow (x,0) = (x, f(x,y)) = F(x,y) \in U \times W,$$

con lo cual $x \in U$ y $(x,y) = F^{-1}(x,0) = (x,h(x,0)) = (x,\varphi(x))$ y en consecuencia $y = \varphi(x)$.

Ahora (8.2.1) y el Teorema de la función inversa nos aseguran que

$$\det \begin{pmatrix} D_{N+1}f_1(x,\varphi(x)) & \dots & D_{N+M}f_1(x,\varphi(x)) \\ \dots & \dots & \dots \\ D_{N+1}f_M(x,\varphi(x)) & \dots & D_{N+M}f_M(x,\varphi(x)) \end{pmatrix} = \det\big(J_F(x,\varphi(x))\big) \neq 0, \ \forall x \in U.$$

Calculemos finalmente la matriz jacobiana de φ. Para cada $x \in U$ y para cada $i = 1,2,\dots,M$ se tiene que

$$f_i(x,\varphi(x)) = 0.$$

Así, para cada $j = 1,2,\dots,N$, la regla de la cadena para las derivadas parciales nos asegura

$$D_j f_i(x,\varphi(x)) + \sum_{k=1}^{M} D_{N+k} f_i(x,\varphi(x)) D_j \varphi_k(x) = 0,$$

(el elemento (i,j) de la matriz nula). En consecuencia para cada $x \in U$, si notamos $y = \varphi(x)$, se tiene:

$$\begin{pmatrix} D_1 f_1(x,y) & \dots & D_N f_1(x,y) \\ \dots & \dots & \dots \\ D_1 f_M(x,y) & \dots & D_N f_M(x,y) \end{pmatrix} + \begin{pmatrix} D_{N+1} f_1(x,y) & \dots & D_{N+M} f_1(x,y) \\ \dots & \dots & \dots \\ D_{N+1} f_M(x,y) & \dots & D_{N+M} f_M(x,y) \end{pmatrix} J_\varphi(x) = 0,$$

de donde se puede despejar $J_\varphi(x)$ para obtener la expresión del enunciado.

Finalmente las perfecciones adicionales de la función f se transfieren a la función implícita φ. Ello es consecuencia de que una tal perfección la tiene F y por tanto F^{-1} y h. Luego, la función implícita $\varphi = h \circ g$ (donde $g(x) = (x,0), \ \forall x \in U$) tiene también dicha perfección. ∎

Nota 8.13. La condición

$$\det \begin{pmatrix} D_{N+1}f_1(a,b) & \dots & D_{N+M}f_1(a,b) \\ \dots & \dots & \dots \\ D_{N+1}f_M(a,b) & \dots & D_{N+M}f_M(a,b) \end{pmatrix} \neq 0$$

equivale a afirmar que en el sistema lineal

$$Df(a,b)(x,y) = 0$$

se puede despejar y de manera global en función de $x \in \mathbb{R}^N$:

$$y^t = -\begin{pmatrix} D_{N+1}f_1(a,b) & \dots & D_{N+M}f_1(a,b) \\ \dots & \dots & \dots \\ D_{N+1}f_M(a,b) & \dots & D_{N+M}f_M(a,b) \end{pmatrix}^{-1} \begin{pmatrix} D_1f_1(a,b) & \dots & D_Nf_1(a,b) \\ \dots & \dots & \dots \\ D_1f_M(a,b) & \dots & D_Nf_M(a,b) \end{pmatrix} x^t.$$

El Teorema de la función implícita asegura la existencia de entornos U de a y Ω de (a,b) tales que para cada $x = (x_1, x_2, \dots, x_N) \in U$ el sistema de M ecuaciones con M incógnitas

$$\begin{cases} f_1(x_1,x_2,\dots,x_N,y_1,y_2,\dots,y_M) & = & 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ f_M(x_1,x_2,\dots,x_N,y_1,y_2,\dots,y_M) & = & 0 \end{cases}$$

tiene una única solución $y = (y_1, y_2, \dots, y_M)$ tal que $(x,y) \in \Omega$. Además la función implícita

$$\varphi = (\varphi_1, \varphi_2, \dots, \varphi_M)$$

así definida es de clase $\mathscr{C}^1$ en U (equivalentemente sus campos escalares componentes $\varphi_1, \varphi_2, \dots, \varphi_M$ tienen gradiente continuo).

Queda una vez más justificada la frase f hereda localmente propiedades de su derivada.

Acabamos este tema presentando la manera práctica de calcular las derivadas de la función implícita. Aunque, en general no se conoce de manera "explícita" la función φ (esto sólo es posible en casos muy particulares como los ejemplos 8.11), siempre podemos calcular su matriz jacobiana. En muchos casos no interesa calcular toda la matriz jacobiana sino solamente algunas derivadas (parciales) de los campos escalares componentes de la función implícita. Procediendo como se ha hecho en la demostración del Teorema de la función implícita para el cálculo de la matriz jacobiana de φ, fijado $j \in \{1,2,\dots,N\}$, se calcula la derivada parcial j-ésima de cada una de las ecuaciones del sistema:

$$f_i(x, \varphi(x)) = 0 \quad (i = 1,2,\dots,M),$$

para obtener el sistema

$$\frac{\partial f_i}{\partial x_j} + \sum_{k=1}^{M} \frac{\partial f_i}{\partial y_k} \frac{\partial y_k}{\partial x_j} = 0, \ \ i = 1,2,\dots,M. \tag{8.2.2}$$

de M ecuaciones lineales con M incógnitas, donde

$$\frac{\partial y_k}{\partial x_j} \equiv \frac{\partial \varphi_k}{\partial x_j}, \ k = 1,2,\dots,M.$$

Para cada $x \in U$ el determinante del sistema es

$$\det \begin{pmatrix} D_{N+1}f_1(x,\varphi(x)) & \dots & D_{N+M}f_1(x,\varphi(x)) \\ \dots & \dots & \dots \\ D_{N+1}f_M(x,\varphi(x)) & \dots & D_{N+M}f_M(x,\varphi(x)) \end{pmatrix} \neq 0,$$

lo que permite despejar

$$\frac{\partial \varphi_k}{\partial x_j},\ k = 1, 2, \ldots, M.$$

Sabemos que la igualdad matricial que nos da $J_\varphi(x)$ es la solución de todos los sistemas 8.2.2 para $j = 1, 2, \ldots, N$. La ventaja de esta expresión es que se puede volver a derivar, por la regla de la cadena para las derivadas parciales, para obtener nuevos sistemas de ecuaciones lineales que permiten despejar (todos tienen el mismo determinante distinto de cero) las derivadas parciales de orden superior de los campos escalares componentes de la función implícita φ.

En el ejemplo (8.11) **b)**, se tiene que $D_2 f(x,y) = 2y \neq 0$ si $y \neq 0$, con lo que $f(x,y) = 0$ define localmente a y como función implícita de x en los puntos tales que $f(x,y) = 0$ con $y \neq 0$. Se tiene

$$2x + 2yy' = 0 \iff y' = -\frac{x}{y}\ .$$

Para obtener la derivada segunda de y se vuelve a derivar en la ecuación anterior

$$1 + (y')^2 + yy'' = 0\ ,$$

y basta sustituir y' para despejar y''.

Análogamente se obtienen las derivadas sucesivas.

8.3. Apéndice: El Teorema de la función inversa se deduce del Teorema de la función implícita.

Supongamos fijado el ambiente del enunciado del Teorema de la función inversa (8.5) y definamos el campo vectorial $F:\mathbb{R}^N\times A\longrightarrow\mathbb{R}^N$ por

$$F(y,x)=f(x)-y,\ \forall (y,x)\in\mathbb{R}^N\times A$$

(nótese el intercambio de los papeles de las variables x e y).

La demostración consiste en aplicar el Teorema de la función implícita (8.12) al campo vectorial F en el punto $(f(a),a)$.

Claramente $F(f(a),a)=0$ y F es de clase $\mathscr{C}^1$ con

$$\det\begin{pmatrix} D_{N+1}F_1(f(a),a) & \ldots & D_{N+N}F_N(f(a),a)\\ \ldots & \ldots & \ldots \\ D_{N+1}F_N(f(a),a) & \ldots & D_{N+N}F_N(f(a),a)\end{pmatrix}=\det\big(J_f(a)\big)\neq 0.$$

El Teorema de la función implícita nos asegura que existen

$$\begin{cases} V \ \text{ entorno abierto de } f(a) \text{ en } \mathbb{R}^N\\ \varphi:V\to\mathbb{R}^N \text{ de clase } \mathscr{C}^1\\ \Omega \text{ entorno abierto de } (f(a),a) \text{ en } \mathbb{R}^N\times A\end{cases}$$

tales que

$$\begin{cases} (y,\varphi(y))\in\mathbb{R}^N\times A\ ,\ F(y,\varphi(y))=0\ ,\ \forall y\in V\\ \{(y,x)\in\Omega:\ F(y,x)=0\}=\{(y,\varphi(y)):\ y\in V\}\end{cases}.$$

Teniendo en cuenta la concreta definición de F, se sigue en nuestro caso que

$$\begin{cases} (1) & \varphi(V)\subset A \quad \text{y} \quad f(\varphi(y))=y\ ,\ \forall y\in V\\ (2) & \{(y,x)\in\Omega:\ y=f(x)\}=\{(y,\varphi(y)):\ y\in V\}\end{cases}.$$

De (1) se deduce que φ es inyectiva y tiene a $f_{|\varphi(V)}$ como inversa por la izquierda. En consecuencia, $\varphi:V\to\varphi(V)$ es biyectiva con inversa $f_{|\varphi(V)}$. Veamos que

$$\varphi(V)=\{x\in A:\ (f(x),x)\in\Omega\}.$$

$\boxed{\subseteq}$ Si $x\in\varphi(V)$, por (1) se tiene que $x\in A$, $f(x)\in V$ y $\varphi(f(x))=x$. Ahora, por (2), podemos concluir que $(f(x),x)\in\Omega$.

$\boxed{\supseteq}$ Si $x\in A$ es tal que $(f(x),x)\in\Omega$, entonces por (2) se obtiene que $f(x)\in V$ y $x=\varphi(f(x))$, de donde se desprende que $x\in\varphi(V)$.

Si notamos $U := \varphi(V)$, veamos que se verifica la tesis de Teorema de la función inversa. En efecto de la igualdad anterior se sigue que U es un abierto de $\mathbb{R}^N$ ya que $U = (f, Id_A)^{-1}(\Omega)$ es la imagen inversa del abierto Ω por la función continua de A (abierto) en $\mathbb{R}^N \times \mathbb{R}^N$ definida por

$$x \longmapsto (f(x), x).$$

Además, $a \in U$ ya que $(f(a), a) \in \Omega$.

También hemos probado que el campo vectorial f es un difeomorfismo de clase $\mathscr{C}^1$ de U sobre V.

A partir del jacobiano de la función implícita, para $x \in U$, si notamos $y = f(x)$ $(x = \varphi(y))$, tenemos que

$$J_{f^{-1}}\big(f(x)\big) = J_\varphi(y) =$$

$$-\begin{pmatrix} D_{N+1}F_1(y,x) & \dots & D_{N+N}F_1(y,x) \\ \dots & \dots & \dots \\ D_{N+1}F_N(y,x) & \dots & D_{N+N}F_N(y,x) \end{pmatrix}^{-1} \begin{pmatrix} D_1F_1(y,x) & \dots & D_NF_1(y,x) \\ \dots & \dots & \dots \\ D_1F_N(y,x) & \dots & D_NF_N(y,x) \end{pmatrix} =$$

$$-\begin{pmatrix} D_1f_1(x) & \dots & D_Nf_1(x) \\ \dots & \dots & \dots \\ D_1f_N(x) & \dots & D_Nf_N(x) \end{pmatrix}^{-1} \Big(-I\Big) = \Big(J_f(x)\Big)^{-1}.$$

Finalmente las perfecciones adicionales de la función f se transfieren a su inversa local. Ello es consecuencia de que una tal perfección la tiene F y por tanto φ, es decir la inversa local de f. ∎

8.4. Referencias recomendadas.

[Cra], [MaHo], [Jur].

8.5. Resumen de resultados del Tema 8.

Definición (difeomorfismo). *Sean A,B abiertos de $\mathbb{R}^N$.*

Se dice que $f : A \to B$ es un <u>*difeomorfismo*</u> *si f es biyectiva derivable y f^{-1} también es derivable. En tal caso escribiremos $f \in \mathrm{Dif}(A,B)$.*

Se dice que $f : A \to \mathbb{R}^N$ es un <u>*difeomorfismo en A*</u> *si $f(A)$ es un abierto de $\mathbb{R}^N$ y $f \in \mathrm{Dif}(A, f(A))$.*

Se dice que $f : A \to \mathbb{R}^N$ es un <u>*difeomorfismo local*</u> *si para cada $a \in A$ existe un entorno abierto U de a contenido en A tal que f es un difeomorfismo en U.*

En caso de que $f \in \mathrm{Dif}(A,B)$, y que tanto f como f^{-1} sean de clase $\mathscr{C}^k$, diremos que f es un difeomorfismo de clase $\mathscr{C}^k$. Análogamente se definen los difeomorfismos en A de clase $\mathscr{C}^k$ y los difeomorfismos locales de clase $\mathscr{C}^k$.

Por ejemplo, la función exponencial es un difeomorfismo en $\mathbb{R}$ de clase $\mathscr{C}^\infty$.

Teorema de la función inversa. *Sean $A \subset \mathbb{R}^N$ un abierto y $f : A \to \mathbb{R}^N$ un campo vectorial de clase $\mathscr{C}^1$ (o lo que es igual sus campos escalares componentes tienen gradiente continuo). Supongamos que existe $a \in A$ tal que*

$$\det\big(J_f(a)\big) = \det \begin{pmatrix} D_1 f_1(a) & \dots & D_N f_1(a) \\ \dots & \dots & \dots \\ D_1 f_N(a) & \dots & D_N f_N(a) \end{pmatrix} \neq 0.$$

Entonces existe un entorno abierto U de a contenido en A tal que $V := f(U)$ es un abierto de $\mathbb{R}^N$, f es un difeomorfismo de clase $\mathscr{C}^1$ de U sobre V, y para cada $x \in U$ se verifica

$$J_{f^{-1}}\big(f(x)\big) = \begin{pmatrix} D_1 f_1(x) & \dots & D_N f_1(x) \\ \dots & \dots & \dots \\ D_1 f_N(x) & \dots & D_N f_N(x) \end{pmatrix}^{-1}.$$

Además, si para k natural mayor que 1 se verifica que f es k veces derivable (resp. f es de clase $\mathscr{C}^k$), entonces f^{-1} es k veces derivable (resp. f^{-1} es de clase $\mathscr{C}^k$).

Proposición. *Sean $A \subset \mathbb{R}^N$ un abierto y $f : A \to \mathbb{R}^N$ un difeomorfismo local, entonces f es una aplicación abierta.*

Teorema "global" de la función inversa. *Sean $A \subset \mathbb{R}^N$ un abierto, k un natural y $f : A \to \mathbb{R}^N$ un campo vectorial. Supongamos que f es de clase $\mathscr{C}^k$ en A y que*

$$\det\big(J_f(x)\big) \neq 0,\ \forall x \in A.$$

Entonces f es un difeomorfismo local de clase $\mathscr{C}^k$ en A, y por tanto una aplicación abierta. Si además f es inyectiva, entonces se tiene que f es un difeomorfismo de clase $\mathscr{C}^k$ en A.

Teorema de la función implícita. *Sean $G \subset \mathbb{R}^N \times \mathbb{R}^M$ un conjunto abierto y $f : G \to \mathbb{R}^M$ un campo vectorial de clase $\mathscr{C}^1$ (o lo que es igual sus campos escalares componentes tienen gradiente continuo). Supongamos que existe $(a,b) \in G$ tal que $f(a,b) = 0$ y que*

$$\det \begin{pmatrix} D_{N+1} f_1(a,b) & \dots & D_{N+M} f_1(a,b) \\ \dots & \dots & \dots \\ D_{N+1} f_M(a,b) & \dots & D_{N+M} f_M(a,b) \end{pmatrix} \neq 0.$$

Entonces existen un entorno abierto Ω de (a,b) contenido en G, un entorno abierto U de a y una función $\varphi : U \to \mathbb{R}^M$ de clase $\mathscr{C}^1$ tal que

$$\{(x,y) \in \Omega : \ f(x,y) = 0\} = \{(x,\varphi(x)) : \ x \in U\}.$$

Se tiene también que para todo $x \in U$ se verifica que

$$\det \begin{pmatrix} D_{N+1}f_1(x,\varphi(x)) & \dots & D_{N+M}f_1(x,\varphi(x)) \\ \dots & \dots & \dots \\ D_{N+1}f_M(x,\varphi(x)) & \dots & D_{N+M}f_M(x,\varphi(x)) \end{pmatrix} \neq 0.$$

y que

$$J_\varphi(x) = - \begin{pmatrix} D_{N+1}f_1(x,y) & \dots & D_{N+M}f_1(x,y) \\ \dots & \dots & \dots \\ D_{N+1}f_M(x,y) & \dots & D_{N+M}f_M(x,y) \end{pmatrix}^{-1} \begin{pmatrix} D_1f_1(x,y) & \dots & D_Nf_1(x,y) \\ \dots & \dots & \dots \\ D_1f_M(x,y) & \dots & D_Nf_M(x,y) \end{pmatrix}$$

donde $y = \varphi(x)$.

Además si para k natural mayor que 1 se verifica que f es k veces derivable (resp. f es clase $\mathscr{C}^k$), entonces φ es k veces derivable (resp. φ es de clase $\mathscr{C}^k$).

8.6. Ejercicios del Tema 8.

8.1 Sea $f:\mathbb{R}^N \to \mathbb{R}^N$ un campo vectorial de clase $\mathscr{C}^1$ verificando

$$\|x-y\| \leq \|f(x)-f(y)\|, \ \forall x,y \in \mathbb{R}^N.$$

Probar que f es un difeomorfismo de $\mathbb{R}^N$ sobre $\mathbb{R}^N$.
Indicación: Deducir de la desigualdad que f verifica las hipótesis del Teorema global de la función inversa, Corolario (8.8). Probar que $f(\mathbb{R}^N)$ es cerrado y deducir por conexión que $f(\mathbb{R}^N)=\mathbb{R}^N$.

8.2 Sea $F:\mathbb{R}^2 \to \mathbb{R}^2$ el campo vectorial dado por

$$F(x,y) = (x+f(y), y+f(x)),$$

donde $f:\mathbb{R}\to\mathbb{R}$ es una función derivable con $|f'(t)| \leq \alpha < 1, \ \forall t \in \mathbb{R}$. Probar que F es un difeomorfismo de $\mathbb{R}^2$ sobre $\mathbb{R}^2$.
Indicación: Utilizar el Teorema de Schauder con

$$\varphi(x,y) := (-f(y), -f(x)), \ \forall (x,y) \in \mathbb{R}^2.$$

8.3 Justificar que el campo vectorial

$$f(x) = \frac{x}{\sqrt{1+\|x\|_2^2}}, \ \forall x \in \mathbb{R}^N$$

es un difeomorfismo de clase $\mathscr{C}^\infty$ de $\mathbb{R}^N$ sobre la bola unidad euclídea $B_{\|\cdot\|_2}(0,1)$.

8.4 Justificar que cada una de las siguientes funciones es un difeomorfismo de clase $\mathscr{C}^\infty$ sobre su imagen

(**Coordenadas esféricas**)

$$f(\rho,\vartheta,\varphi) = (\rho\ \cos\vartheta\ \cos\varphi, \rho\ \operatorname{sen}\vartheta\ \cos\varphi, \rho\ \operatorname{sen}\varphi), \forall(\rho,\vartheta,\varphi) \in \mathbb{R}^+ \times]-\pi,\pi[\times]\frac{-\pi}{2},\frac{\pi}{2}[$$

$$g(x,y) = (x-y, xy), \ \forall (x,y) \in \{(x,y)\in\mathbb{R}^2 : \ x+y>0\}$$

$$h(x,y,z) = (x-xy, xy-xyz, xyz), \ \forall (x,y,z) \in \{(x,y,z)\in\mathbb{R}^3 : \ xy \neq 0\}$$

Indicación: Utilizar en los tres casos el Teorema global de la función inversa (Corolario 8.8).

8.5 Justificar que el campo vectorial $f:\mathbb{R}^2\to\mathbb{R}^2$ dado por

$$f(x,y)=(\cos x+\cos y,\ \operatorname{sen} x+\operatorname{sen} y)$$

es un difeomorfismo local en todo punto de $\{(x,y)\in\mathbb{R}^2:\ x-y\notin\pi\mathbb{Z}\}$ ¿Qué ocurre en los puntos de la forma $(a,a+p\pi)$ con $p\in\mathbb{Z}$?
Indicación: Utilizar el Teorema de la función inversa en los puntos del conjunto $\{(x,y)\in\mathbb{R}^2:\ x-y\notin\pi\mathbb{Z}\}$. Comprobar que en cualquier entorno del resto de puntos de $\mathbb{R}^2$ la función f no es inyectiva.

8.6 **(*)** Una función de un abierto A de $\mathbb{R}^2$ en $\mathbb{R}^2$ definida por

$$f(x,y)=(u(x,y),v(x,y))$$

se dice que verifica las ecuaciones de Cauchy-Riemann si

$$\frac{\partial u}{\partial x}=\frac{\partial v}{\partial y}\ ,\ \ \frac{\partial u}{\partial y}=-\ \frac{\partial v}{\partial x}\ .$$

Probar que si f es de clase $\mathscr{C}^1$, con derivada no nula en cada punto de A y se verifican las ecuaciones de Cauchy-Riemann, entonces es un difeomorfismo local en todo punto de A. Si además es inyectiva, entonces su inversa también verifica las ecuaciones de Cauchy-Riemann.
Indicación: Utilizar el Teorema 8.5 y el Corolario 8.8 .

8.7 Justificar la existencia de un campo vectorial $f=(f_1,f_2)$ de clase $\mathscr{C}^\infty$ de un abierto U de $\mathbb{R}^2$ en $\mathbb{R}^2$ verificando

$$f(1,-1)=(1,1)\ \text{ y }\ \left\{\begin{matrix} x^2+y^2-2f_1(x,y)f_2(x,y)=0\\ x^3+y^3+f_1(x,y)^3-f_2(x,y)^3=0\end{matrix}\right\},\ \forall (x,y)\in U.$$

Calcular $Df(1,-1)$ y $\dfrac{\partial^2 f_1}{\partial x^2}(1,-1)$.
Indicación: Utilizar el Teorema de la función implícita para un conveniente campo vectorial de $\mathbb{R}^2\times\mathbb{R}^2\to\mathbb{R}^2$.

8.8 Justificar la existencia de una función φ de clase $\mathscr{C}^\infty$ definida en un entorno U de cero en $\mathbb{R}$ con valores en $\mathbb{R}$ tal que

$$1-\varphi(x)\mathrm{e}^{\,x}+x\mathrm{e}^{\,\varphi(x)}=0,\ \forall x\in U$$

Obtener el polinomio de Taylor de segundo orden de φ en $x=0$.
Indicación: Utilizar el Teorema de la función implícita.

8.9 Probar que las ecuaciones

$$\begin{cases} xy^5 + yu^5 + zv^5 = 1 \\ x^5 y + y^5 u + z^5 v = 1 \end{cases}$$

definen a u y a v como funciones implícitas de (x,y,z) en un entorno del punto $(0,1,1,1,0)$. Calcular $\dfrac{\partial u}{\partial y}(0,1,1)$ y $\dfrac{\partial^2 u}{\partial y^2}(0,1,1)$.
Indicación: Utilizar el Teorema de la función implícita para un conveniente campo vectorial.

8.10 Justificar la existencia de dos abiertos difeomorfos A y B de $\mathbb{R}^2$ verificando que $(1,1) \in A$, $(0,1) \in B$ y que para cada $(x,y) \in A$ existe un único $(u,v) \in B$ tal que:

$$\begin{cases} x\, e^u + y\, e^v = 1 + e \\ \quad u\, e^x + v\, e^y = e \end{cases}$$

Calcular la matriz jacobiana en $(1,1)$ (resp. $(0,1)$) de

$$(x,y) \longmapsto \big(u(x,y), v(x,y)\big) \ \Big(\text{resp. } (u,v) \longmapsto \big(x(u,v), y(u,v)\big)\Big).$$

Indicación: Utilizar los Teoremas de la función implícita y de la función inversa.

8.11 Sean P, N naturales, A un abierto de $\mathbb{R}^P$ y $f : A \to \mathbb{R}^N$ un campo vectorial de clase $\mathscr{C}^1$ en un punto $a \in A$. Probar que:

i) Si $P < N$ y $Df(a)$ es inyectiva, entonces existen entornos abiertos U de a en $\mathbb{R}^P$, V de 0 en $\mathbb{R}^{N-P}$, W de $f(a)$ en $\mathbb{R}^N$ y $g \in \mathrm{Dif}(W, U \times V)$ tal que

$$(g \circ f)(x) = (x_1, \ldots, x_P, 0, \ldots, 0),\ \forall x = (x_1, \ldots, x_P) \in U.$$

ii) Si $N < P$ y $Df(a)$ es sobreyectiva, entonces existen abierto U, V en $\mathbb{R}^P$ con $a \in V$ y $g \in \mathrm{Dif}(U, V)$ tal que

$$(f \circ g)(x) = (x_1, \ldots, x_N),\ \forall x = (x_1, \ldots, x_P) \in U.$$

Indicación: Aplicar en ambos supuestos el Teorema de la función inversa a conveniente campo vectorial de $\mathbb{R}^N$ (resp. $\mathbb{R}^P$) en $\mathbb{R}^N$ (resp. $\mathbb{R}^P$).

8.7. Soluciones a los ejercicios del Tema 8.

8.1 La desigualdad nos asegura que f es inyectiva. Sea $x \in \mathbb{R}^N$. Se tiene que

$$\|u\| \leq \left\| \frac{f(x+tu)-f(x)}{t} \right\| , \ \forall u \in \mathbb{R}^N \setminus \{0\}, \ \forall t \in \mathbb{R}^*$$

donde hemos vuelto a utilizar la hipótesis, esto es, $Df(x)$ es inyectiva, y por tanto biyectiva (aplica vectores linealmente independientes en vectores linealmente independientes). Hemos probado que

$$Df(x) \in \mathrm{Iso}(\mathbb{R}^N), \ \forall x \in \mathbb{R}^N .$$

El Teorema global de la función inversa (Corolario 8.8) nos asegura que

$$f \ \text{ es abierta y que } \ f \in \mathrm{Dif}(\mathbb{R}^N, f(\mathbb{R}^N)).$$

Finalmente, como $f(\mathbb{R}^N)$ es abierto, bastará probar, por razones de conexión, que $f(\mathbb{R}^N)$ es también cerrado. En efecto, sea $y \in \overline{f(\mathbb{R}^N)}$ y sea $\{f(x_n)\} \longrightarrow y$. La desigualdad de la hipótesis nos asegura, al ser $\{f(x_n)\}$ de Cauchy, que $\{x_n\}$ es también de Cauchy, luego convergente. Si $\{x_n\} \longrightarrow x$, la continuidad de f nos asegura que $\{f(x_n)\} \longrightarrow f(x)$. Hemos probado que $f(\mathbb{R}^N)$ es cerrado, y por tanto $f(\mathbb{R}^N) = \mathbb{R}^N$.

Conviene resaltar que hay que utilizar el Teorema global de la función inversa (o el Teorema de Schauder 4.15 si se quiere) para asegurar que $f(\mathbb{R}^N)$ es abierto, pues aunque la desigualdad del enunciado da la continuidad de la función inversa eso no garantiza que dicho conjunto sea abierto.

8.2 Consideremos la función auxiliar

$$\varphi(x,y) := (-f(y), -f(x)), \ \forall x,y \in \mathbb{R}^2.$$

Se tiene que

$$\|\varphi(x,y)-\varphi(u,v)\|_1 = |f(x)-f(u)| + |f(y)-f(v)| =$$

$$|x-u||f'(c)| + |y-v||f'(d)| \leq \alpha(|x-u|+|y-v|) = \alpha\|(x,y)-(u,v)\|_1,$$

esto es, φ es contractiva. Ahora el Teorema de Schauder (4.15) nos asegura que $F \equiv \mathrm{Id}_{\mathbb{R}^2} - \varphi$ es un homeomorfismo de $\mathbb{R}^2$ sobre $\mathbb{R}^2$. Por otra parte

$$\det\big(J_F(x,y)\big) = 1 - f'(x)f'(y) \neq 0,$$

así $DF(x,y) \in \mathrm{Iso}(\mathbb{R}^2)$. El ejercicio se concluye utilizando la regla de derivación de la función inversa (Proposición 3.27 y el apartado 4 de la Sección 5.8).

Obsérvese que no se puede utilizar directamente el Teorema de la función inversa (falta la hipótesis de ser de clase $\mathscr{C}^1$, y claro está tampoco obtenemos que F^{-1} sea de clase $\mathscr{C}^1$).

8.3 Es fácil comprobar (utilizando un procedimiento similar al empleado en el Ejercicio 2.12) que f es una biyección de $\mathbb{R}^N$ sobre $B_2(0,1)$. Se tiene que

$$f(x) = \left(\frac{x_1}{\sqrt{1+x_1^2+\cdots+x_N^2}}, \ldots, \frac{x_N}{\sqrt{1+x_1^2+\cdots+x_N^2}} \right)$$

y que

$$f^{-1}(y) = \left(\frac{y_1}{\sqrt{1-y_1^2-\cdots-y_N^2}}, \ldots, \frac{y_N}{\sqrt{1-y_1^2-\cdots-y_N^2}} \right)$$

y basta aplicar, puesto que los campos escalares son de clase $\mathscr{C}^\infty$, el Corolario (6.14).

8.4 En todos los casos se utiliza el Teorema global de la función inversa (Corolario 8.8). Las funciones son claramente de clase $\mathscr{C}^\infty$ en sus respectivos abiertos de definición. Tan sólo hay que probar que

i) $\det\big(J_f(x)\big) \neq 0$.

ii) f es inyectiva.

Coordenadas esféricas

$$f(\rho,\vartheta,\varphi) = (\rho\,\cos\vartheta\,\cos\varphi, \rho\,\operatorname{sen}\vartheta\,\cos\varphi, \rho\,\operatorname{sen}\varphi), \forall(\rho,\vartheta,\varphi)\in\mathbb{R}^+\times]-\pi,\pi[\times]\frac{-\pi}{2},\frac{\pi}{2}[$$

i) $\det\big(J_f(x)\big) = \rho^2\,cos\varphi > 0$, pues $\varphi \in\,]\frac{-\pi}{2},\frac{\pi}{2}[$.

ii) Veamos que f es una biyección de

$$A := \mathbb{R}^+\times]-\pi,\pi[\times]\frac{-\pi}{2},\frac{\pi}{2}[\quad \text{sobre} \quad B := \mathbb{R}^3 \setminus \{(x,0,z):\ x\leq 0\},$$

esto es,

$$\forall (x,y,z)\in B,\ \exists_1(\rho,\vartheta,\varphi)\in A:\ f(\rho,\vartheta,\varphi) = (x,y,z)$$

a dicha terna se denomina <u>las coordenadas esféricas del punto (x,y,z).</u>

<u>Inyectividad:</u> En efecto

$$\left.\begin{array}{r}\rho\,\cos\vartheta\,\cos\,\varphi = r\,\cos\alpha\,\cos\beta\\ \rho\,\operatorname{sen}\vartheta\,\cos\,\varphi = r\,\operatorname{sen}\alpha\,\cos\,\beta\\ \rho\,\operatorname{sen}\,\varphi = r\,\operatorname{sen}\,\beta\end{array}\right\}\Rightarrow\left\{\begin{array}{l}\qquad\|f(\rho,\vartheta,\varphi)\|_2^2 = \rho^2 = r^2 \Rightarrow \rho = r\\ \text{así } \operatorname{sen}\varphi = \operatorname{sen}\beta, \varphi,\beta\in]-\frac{\pi}{2},\frac{\pi}{2}[\Rightarrow \varphi=\beta\,.\\ \text{por tanto } \exists k\in\mathbb{Z}:\ \vartheta = \alpha + 2k\pi \Rightarrow \vartheta = \alpha\end{array}\right.$$

<u>Sobreyectividad:</u> Sea $(x,y,z)\in B$, se tiene

$$\left[\rho = \sqrt{x^2+y^2+z^2},\ x^2+y^2>0\right] \Rightarrow -1 < \frac{z}{\rho} < 1 \Rightarrow \exists\varphi\in\,]\frac{-\pi}{2},\frac{\pi}{2}[:\ z = \rho\ \operatorname{sen}\varphi.$$

Como
$$(x,y) \in \mathbb{R}^2 \setminus \{(x,0):\ x \leq 0\},$$
utilizando las coordenadas polares (8.10), sabemos que
$$\exists_1 (r,\vartheta) \in \mathbb{R}^+ \times]-\pi,\pi[:\ \begin{cases} x = r\ \cos\vartheta \\ y = r\ \operatorname{sen}\vartheta \end{cases},$$
y en consecuencia, teniendo en cuenta que $\frac{-\pi}{2} < \varphi < \frac{\pi}{2}$, concluimos que
$$r = \sqrt{x^2+y^2} = \sqrt{\rho^2 - z^2} = \rho\ \cos\varphi.$$

$$g(x,y) = (x-y,xy),\ \forall (x,y) \in \{(x,y) \in \mathbb{R}^2\, x+y>0\}$$

i) $\det\big(J_g(x,y)\big) = x+y > 0$.

ii)
$$\left.\begin{array}{r} x-y = a-b \\ xy = ab \end{array}\right\} \Rightarrow \left.\begin{array}{c} x-a = y-b \\ (x-a)y = a(b-y) \end{array}\right\} \Rightarrow\ (x-a)y = -a(x-a).$$
Si $x=a$, entonces $y=b$ y hemos terminado; en otro caso $x-a \neq 0$ lo que obligaría a que $y=-a$ y en consecuencia $x=-b$ lo cual es absurdo pues sería $x+y = -(a+b)$ en contra de que $x+y,\ a+b \in \mathbb{R}^+$.

$$h(x,y,z) = (x-xy, xy-xyz, xyz),\ \forall (x,y,z) \in \{(x,y,z) \in \mathbb{R}^3:\ xy \neq 0\}$$

i) $\det\big(J_h(x,y,z)\big) = x^2y \neq 0$.

ii)
$$\left.\begin{array}{c} x-xy = a-ab \\ xy-xyz = ab-abc \\ xyz = abc \end{array}\right\} \Rightarrow \text{(sumando las tres igualdades)}\quad x = a\ .$$
Como $a \neq 0$, de la primera ecuación se deduce que $y=b$. Por último como $ab \neq 0$, de la tercera ecuación concluimos que $z=c$.

8.5 Como $\det\big(J_f(x,y)\big) = \operatorname{sen}(y-x),\ \forall (x,y) \in \mathbb{R}^2$, se verifica que
$$\det\big(J_f(x,y)\big) = 0 \Longleftrightarrow x-y \in \pi\mathbb{Z}\ .$$

El Teorema de la función inversa nos asegura que la función f es un difeomorfismo local en todo punto (a,b) tal que $a-b \in \mathbb{R} \setminus \pi\mathbb{Z}$.

Sea $p \in \mathbb{Z}$. Veamos que f no es inyectiva en ningún entorno del punto $(a,a+p\pi)$. En efecto, para todo $\delta > 0$, se verifica

1. Si p es par:

$$f(a+\delta,a+p\pi-\delta)=f(a+\delta,a-\delta)=f(a-\delta,a+\delta)=f(a-\delta,a+p\pi+\delta).$$

2. Si p es impar:

$$f(a+\delta,a+p\pi+\delta)=f(a+\delta,a+\pi+\delta)=(0,0)=f(a,a+\pi)=f(a,a+p\pi),$$

donde se ha utilizado que:

$$\cos x+\cos(x+\pi)=0 \text{ y } \operatorname{sen} x+\operatorname{sen}(x+\pi)=0\ ,\ \forall x\in\mathbb{R}.$$

8.6 Como

$$\det\Big(J_f(x,y)\Big)=\begin{vmatrix}\dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y}\\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y}\end{vmatrix}=\Big(\frac{\partial u}{\partial x}\Big)^2+\Big(\frac{\partial u}{\partial y}\Big)^2=\Big(\frac{\partial v}{\partial x}\Big)^2+\Big(\frac{\partial v}{\partial y}\Big)^2=:\alpha>0,$$

el Teorema de la función inversa nos asegura que la función f es un difeomorfismo local en todo punto de A.

Si además f es inyectiva, el Teorema global de la función inversa nos asegura que $f\in\operatorname{Dif}(A,f(A))$. Por último

$$J_{f^{-1}}(u,v)=J_f(x,y)^{-1}=\frac{1}{\alpha^2}\begin{pmatrix}\dfrac{\partial v}{\partial y} & -\dfrac{\partial v}{\partial x}\\[2ex] -\dfrac{\partial u}{\partial y} & \dfrac{\partial u}{\partial x}\end{pmatrix}=\begin{pmatrix}A & B\\ -B & A\end{pmatrix},$$

esto es, f^{-1} verifica las condiciones de Cauchy-Riemann.

8.7 Se trata de definir un campo vectorial g que verifique en el punto $(1,-1,1,1)$ las hipótesis del Teorema de la función implícita y que nos permita justificar la existencia de los campos escalares f_1 y f_2 verificando las condiciones del enunciado. Este campo vectorial no puede ser otro que

$$g(x,y,u,v):=\big(x^2+y^2-2uv\ ,\ x^3+y^3+u^3-v^3\big),\ \forall(x,y,u,v)\in\mathbb{R}^4.$$

g es de clase $\mathscr{C}^\infty$, verifica que $g(1,-1,1,1)=(0,0)$ y como

$$J_g(1,-1,1,1)=\begin{pmatrix}2 & -2 & -2 & -2\\ 3 & 3 & 3 & -3\end{pmatrix},$$

al ser $\begin{vmatrix}-2 & -2\\ 3 & -3\end{vmatrix}=12\neq 0$, existen un entorno U de $(1,-1)$ y un campo vectorial $f: U\to\mathbb{R}^2$ de clase $\mathscr{C}^\infty$ tal que

$$g\big(x,y,f_1(x,y),f_2(x,y)\big)=(0,0),\ \forall(x,y)\in U.$$

El Teorema de la función implícita nos asegura además que

$$J_f(1,-1) = -\begin{pmatrix} -2 & -2 \\ 3 & -3 \end{pmatrix}^{-1} \begin{pmatrix} 2 & -2 \\ 3 & 3 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Derivando dos veces en el sistema respecto de la variable x y sustituyendo los valores conocidos, nos queda

$$\left.\begin{array}{rcl} \dfrac{\partial^2 f_1}{\partial x^2} + \dfrac{\partial^2 f_2}{\partial x^2} & = & 1 \\[2ex] \dfrac{\partial^2 f_1}{\partial x^2} - \dfrac{\partial^2 f_2}{\partial x^2} & = & 0 \end{array}\right\} \Longrightarrow \frac{\partial^2 f_1}{\partial x^2}(1,-1) = \frac{1}{2}\,.$$

8.8 El campo escalar definido en $\mathbb{R}^2$ por

$$f(x,y) = 1 - ye^x + xe^y,$$

verifica las hipótesis del Teorema de la función implícita en el punto $(0,1)$, pues f es de clase $\mathscr{C}^\infty$, $f(0,1) = 0$ y $D_2 f(0,1) = -1 \neq 0$. Así existen un entorno U de 0 y una función $\varphi : U \to \mathbb{R}$ de clase $\mathscr{C}^\infty$ tales que $f(x,\varphi(x)) = 0$, $\forall x \in \mathbb{R}$. Derivando dos veces y sustituyendo los valores de φ y de su derivada en 0, obtenemos

$$\varphi'(0) = e - 1\ ,\ \varphi''(0) = 2e^2 - 4e + 1\ ,$$

lo que nos permite calcular el polinomio de Taylor

$$P_2 = 1 + \frac{\varphi'(0)}{1!}\,x + \frac{\varphi''(0)}{2!}\,x^2\ .$$

8.9 El campo escalar $f : \mathbb{R}^3 \times \mathbb{R}^2 \to \mathbb{R}^2$ dado por

$$f(x,y,z,u,v) = \left(xy^5 + yu^5 + zv^5 - 1\ ,\ x^5 y + y^5 u + z^5 v - 1\right)\ ,$$

cumple las hipótesis del Teorema de la función implícita en el punto $(0,1,1,1,0)$, luego existen los campos escalares u, v cumpliendo las condiciones del enunciado. Los cálculos habituales nos dan

$$J_{(u,v)}(0,1,1) = \begin{pmatrix} \dfrac{-1}{5} & \dfrac{-1}{5} & 0 \\[2ex] \dfrac{1}{5} & \dfrac{-24}{5} & 0 \end{pmatrix}\ ;\ \frac{\partial^2 u}{\partial y^2}(0,1,1) = \frac{6}{25}\,.$$

8.10 El campo vectorial $f:\mathbb{R}^2\times\mathbb{R}^2\to\mathbb{R}^2$ dado por

$$f(x,y,u,v)=\left(xe^u+ye^v-1-e\ ,\ ue^x+ve^y-e\right)\ ,$$

cumple las hipótesis del Teorema de la función implícita en el punto $(1,1,0,1)$, luego existen un entorno abierto U de $(1,1)$ y un campo vectorial $\varphi:U\to\mathbb{R}^2$ de clase $\mathscr{C}^\infty$ verificando

$$\varphi(1,1)=(0,1)\ \text{ y }\ f(x,y,\varphi_1(x,y),\varphi_2(x,y))=(0,0),\ \forall(x,y)\in U.$$

Los cálculos habituales dan

$$J_\varphi(1,1)=-\begin{pmatrix}1 & e\\ e & e\end{pmatrix}^{-1}\begin{pmatrix}1 & e\\ 0 & e\end{pmatrix}=\frac{1}{e-1}\begin{pmatrix}1 & 0\\ -1 & 1-e\end{pmatrix}$$

y en consecuencia

$$\det\left(J_\varphi(1,1)\right)=\frac{1}{1-e}\neq 0.$$

Así el campo vectorial φ cumple las condiciones del Teorema de la función inversa, lo que nos asegura la existencia de una entorno abierto A de $(1,1)$ contenido en U y un entorno abierto B de $(0,1)$ tales que $\varphi\in\mathrm{Dif}(A,B)$. Por último, el Teorema de la función inversa nos asegura también que

$$J_{\varphi^{-1}}(0,1)=J_\varphi(1,1)^{-1}=\begin{pmatrix}e-1 & 0\\ -1 & -1\end{pmatrix}\ .$$

8.11

i) $P<N$ (Hay más campos escalares que variables). Como $Df(a)$ es inyectiva, el rango de la matriz $J_f(a)$ es máximo y hay P vectores linealmente independientes en

$$\{\nabla f_1(a),\dots,\nabla f_N(a)\},$$

que podemos suponer, sin mengua de generalidad, son los primeros. En este supuesto, definimos el campo vectorial $\Phi:A\times\mathbb{R}^{N-P}\longrightarrow\mathbb{R}^N$ dado por

$$\Phi(x_1,\dots,x_P,y_{P+1},\dots,y_N):=$$

$$\left(f_1(x_1,\dots,x_P),\dots,f_P(x_1,\dots,x_P),f_{P+1}(x_1,\dots,x_P)+y_{P+1},\dots,f_N(x_1,\dots,x_P)+y_N\right),$$

esto es, $\Phi(x,y):=f(x)+(0,y)$. Es claro que Φ es de clase $\mathscr{C}^1$ en $(a,0)$ y que

$$J_\Phi(a,0)=\begin{pmatrix}\nabla f_1(a) & \\ \dots & 0\\ \nabla f_P(a) & \\ \nabla f_{P+1}(a) & \\ \dots & I\\ \nabla f_N(a) & \end{pmatrix}$$

y en consecuencia det $\big(J_\Phi(a,0)\big) \neq 0$. El Teorema de la función inversa nos asegura la existencia de entornos abiertos U de a en $\mathbb{R}^P$, V de 0 en $\mathbb{R}^{N-P}$, W de $\Phi(a,0)(=f(a))$ en $\mathbb{R}^N$ tales que $\Phi \in \mathrm{Dif}(U \times V, W)$. Si notamos $g = \Phi^{-1}$, se tiene que $g \in \mathrm{Dif}(W, U \times V)$. Para cada $x \in U$ se tiene que $(x,0) \in U \times V$, y en consecuencia $(g \circ \Phi)(x,0) = (x,0)$. Hemos probado que

$$g(f(x)) = g(\Phi(x,0)) = (x,0),\ \forall x \in U,$$

esto es,

$$(g \circ f)(x) = (x_1, \ldots, x_P, 0, \ldots, 0),\ \forall x = (x_1, \ldots, x_P) \in U\ .$$

ii) $N < P$ (Hay más variables que campos escalares). Como $Df(a)$ es sobreyectiva, el rango de la matriz $J_f(a)$ es máximo y hay N vectores columna linealmente independientes, que podemos suponer, sin mengua de generalidad, son los primeros. En este supuesto, definimos el campo vectorial $\Phi : A \longrightarrow \mathbb{R}^P$ dado por

$$\Phi(x_1, \ldots, x_P) := \big(f_1(x_1, \ldots, x_P), \ldots, f_N(x_1, \ldots, x_P), x_{N+1}, \ldots, x_P\big).$$

Es claro que Φ es de clase $\mathscr{C}^1$ en a y que

$$J_\Phi(a) = \begin{pmatrix} D_1 f_1(a) & \ldots & D_N f_1(a) & D_{N+1} f_1(a) & \ldots & D_P f_1(a) \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ D_1 f_N(a) & \ldots & D_N f_N(a) & D_{N+1} f_N(a) & \ldots & D_P f_N(a) \\ & & & & & \\ & 0 & & & I & \\ & & & & & \end{pmatrix}$$

y en consecuencia det $\big(J_\Phi(a,0)\big) \neq 0$. El Teorema de la función inversa nos asegura la existencia de abiertos U, V en $\mathbb{R}^P$ con $a \in V$ tales que $\Phi \in \mathrm{Dif}(V, U)$. Si notamos $g = \Phi^{-1}$, se tiene que $g \in \mathrm{Dif}(U, V)$. Para cada $x \in U$ se tiene que

$$x = \Phi(g(x)) = \big(f_1(g(x)), \ldots, f_N(g(x)), g_{N+1}(x), \ldots, g_P(x)\big).$$

Hemos probado que

$$(f \circ g)(x) = (x_1, \ldots, x_N),\ \forall x = (x_1, \ldots, x_P) \in U.$$

Tema 9

Variedades. Extremos condicionados.

Es bien conocido que para k, N naturales con $k < N$, las variedades afines de dimensión k en $\mathbb{R}^N$ son los conjuntos M de la forma

$$M = \{(x_1, \ldots, x_N) \in \mathbb{R}^N : \ a_{i1}x_1 + \cdots + a_{iN}x_N + b_i = 0 \ \text{ para } \ i = 1, \ldots, N-k\}$$

donde la matriz de números reales $A = (a_{ij})$ tiene rango máximo, esto es, $N-k$, y $b_i \in \mathbb{R}$ para $i = 1, \ldots, N-k$.

El Teorema de Cramer permite despejar $N-k$ variables del sistema de ecuaciones lineales que define M en función de las otras k, obteniéndose las conocidas ecuaciones paramétricas de M.

Es totalmente natural considerar también los subconjuntos M de $\mathbb{R}^N$ asociados a un sistema de ecuaciones no necesariamente lineal:

$$M = \{(x_1, \ldots, x_N) \in \mathbb{R}^N : \ g_i(x_1, \ldots, x_N) = 0\ , \ \text{ para } \ i = 1, \ldots, N-k\}.$$

Inspirados en el Teorema de la función implícita 8.12, y con el propósito de garantizar que $N-k$ de las variables puedan (teóricamente) despejarse en función de las restantes, es razonable considerar la situación de que el campo vectorial $g = (g_1, \ldots, g_{N-k})$ sea tal que sus campos escalares componentes tengan gradiente continuo y que la matriz jacobiana tenga rango máximo, esto es

$$\text{rango}\big(J_g(x)\big) = N-k.$$

Mediante adecuada reordenación de las variables podemos entonces aplicar el citado Teorema de la función implícita para obtener representaciones paramétricas locales de M, y es esperable que la "variedad tangente" en cada punto a de M sea la variedad afín determinada por el sistema de ecuaciones

$$J_g(a)\, x^t = 0.$$

Dedicamos la primera sección de este tema a dar la definición de variedad diferenciable que acabamos de motivar. Probamos un resultado importante (Teorema 9.2) en el que caracterizamos las variedades, y cuya demostración usa esencialmente los Teoremas de la función inversa (8.5) y de la función implícita para campos vectoriales (8.12).

En la sección segunda se introducen las nociones de espacios tangente y normal a una variedad que serán útiles a la hora de estudiar, en la última sección, los extremos de campos escalares condicionados por una variedad diferenciable.

9.1. Variedades diferenciables.

Definición 9.1 (implícita de variedad diferenciable)**.** Sean N natural y k entero con $0 \leq k \leq N$. Se dice que un subconjunto no vacío M de $\mathbb{R}^N$ es una *variedad* (diferenciable) en $\mathbb{R}^N$ de dimensión k si para cada $a \in M$ existen,

$$\begin{cases} G \text{ entorno abierto de } a \text{ en } \mathbb{R}^N \\ g : G \to \mathbb{R}^{N-k} \text{ de clase } \mathscr{C}^1 \text{ verificando } \operatorname{rango}\big(J_g(x)\big) = N-k, \ \ \forall x \in G \end{cases}$$

tales que $M \cap G = \{x \in G : \ g(x) = 0\}$.

En el caso de que se pueda conseguir una función g como antes que determine la totalidad de M, diremos que la variedad M está globalmente determinada por g.

Hablando coloquialmente la anterior definición presenta las variedades diferenciables en $\mathbb{R}^N$ como los conjuntos de ceros de una función g de clase $\mathscr{C}^1$, esto es, como los conjuntos formados por los puntos que se someten a la ligazón impuesta por una función g de clase $\mathscr{C}^1$ que expresa la dependencia entre las variables de los puntos que componen la variedad. La noción de dimensión responde al grado de libertad de las variables y se justifica vía el "principio" de inherencia local de propiedades globales de la derivada. Nótese que en cada punto a de la variedad el conjunto de puntos ligados por $Dg(a)$, esto es $\operatorname{Ker} Dg(a)$, tiene dimensión k puesto que

$$\dim\big(\operatorname{Ker} Dg(a)\big) = N - \dim\big(\operatorname{Im} Dg(a)\big) = N - \operatorname{rango}\big(J_g(a)\big).$$

No es inmediato que el entero k que aparece en la definición de variedad esté determinado de forma única, con lo que la frase, M es una variedad de dimensión k, queda en el aire. El Corolario 9.4 justificará el concepto de dimensión de una variedad en el caso $0 < k < N$.

Los dos casos extremos, las variedades de dimensión 0 y de dimensión N en $\mathbb{R}^N$, son fáciles de describir:

$M \subset \mathbb{R}^N$ es una variedad de dimensión 0 si, y sólo si, M es un conjunto de puntos aislados y por tanto numerable. Si además M es compacta, el conjunto es finito.

En efecto, supongamos que M es una variedad de dimensión 0. Fijemos $a \in M$ y consideremos G y g como en la definición. La función g verifica que

$$\operatorname{rango}\big(J_g(x)\big) = N, \ \forall x \in G,$$

por lo que el Teorema de la función inversa nos asegura la inyectividad de g en un entorno abierto U de a contenido en G, y por tanto

$$M \cap U = \{x \in U : \ g(x) = 0\} = \{a\}.$$

De donde se concluye que a es un punto aislado de M. Si además M es compacta el conjunto es finito.

Recíprocamente, si M es un conjunto de puntos aislados de $\mathbb{R}^N$ nos basta considerar para a en M un entorno abierto G de a tal que $M \cap G = \{a\}$ y como $g : G \longrightarrow \mathbb{R}^N$ la función definida por $g(x) = x - a$.

$M \subset \mathbb{R}^N$ es una variedad de dimensión N si, y sólo si, M es un abierto de $\mathbb{R}^N$.

En efecto, sea M una variedad de dimensión N en $\mathbb{R}^N$. Si para $a \in M$ consideramos G_a y $g : G_a \longrightarrow \mathbb{R}^0 \equiv \{0\}$ (espacio vectorial de dimensión cero), como en la definición, se tiene que $M \cap G_a = g^{-1}(0)$ es un abierto de G_a y por tanto abierto de $\mathbb{R}^N$. Así

$$M = \cup_{a \in M}(M \cap G_a)$$

es un abierto.

Recíprocamente, si M es un abierto, basta considerar la función $g : M \longrightarrow \{0\}$ dada por $g(x) = 0$, la cual determina globalmente M.

Nos concentramos por tanto en el estudio de las variedades de dimensión k de $\mathbb{R}^N$ con $0 < k < N$.

Teorema 9.2 (Definiciones usuales equivalentes)**.** *Sean k, m, N números naturales con $k + m = N$. Para $M \subset \mathbb{R}^N$ no vacío equivalen:*

i) *M es una variedad de dimensión k en $\mathbb{R}^N$. Es decir, para cada $a \in M$, existen*

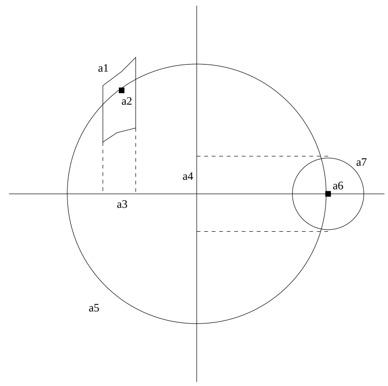
$$\begin{cases} G \text{ entorno abierto de } a \text{ en } \mathbb{R}^N \\ g : G \to \mathbb{R}^m \text{ de clase } \mathscr{C}^1 \text{ verificando } \operatorname{rango}\big(J_g(x)\big) = m, \ \forall x \in G \end{cases}$$

tales que $M \cap G = \{x \in G : \ g(x) = 0\}$.

ii) **(Definición explícita)** *M es localmente la gráfica de funciones de clase $\mathscr{C}^1$ definidas en abiertos de $\mathbb{R}^k$ y con valores en $\mathbb{R}^m$. Es decir, para cada $a \in M$, previa una reordenación de las variables en $\mathbb{R}^N$, existen*

$$\begin{cases} \Omega \text{ entorno abierto de } a \text{ en } \mathbb{R}^N \\ \varphi : U \to \mathbb{R}^m \text{ de clase } \mathscr{C}^1 \text{, donde } U = \pi(\Omega) \text{ siendo} \\ \pi : \mathbb{R}^N \to \mathbb{R}^k \text{ la proyección } (x_1, \dots, x_N) \longmapsto (x_1, \dots, x_k) \end{cases}$$

tales que $M \cap \Omega = \{(y, \varphi(y)) : \ y \in U\}$.

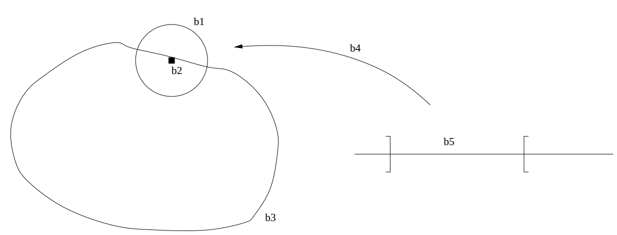


iii) **(Definición por parametrizaciones cartesianas locales)** *Para cada $a \in M$, existen*

$$\begin{cases} \Omega \text{ entorno abierto de } a \text{ en } \mathbb{R}^N \\ U \text{ abierto de } \mathbb{R}^k \\ p : U \to \mathbb{R}^N \text{ homeomorfismo de clase } \mathscr{C}^1 \text{ de } U \text{ sobre } p(U) \end{cases}$$

tales que $\begin{cases} p(U) = M \cap \Omega \\ \text{rango}\big(J_p(y)\big) = k, \ \ \forall y \in U \end{cases}$.

La terna (Ω, U, p) se denomina una parametrización cartesiana local de M en a.



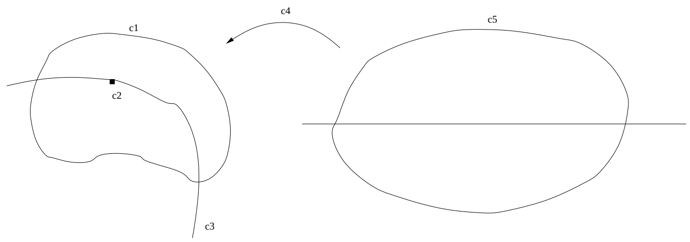
iv) **(Definición por difeomorfismos espaciales)** *Para cada $a \in M$, existen*

$$\begin{cases} W \text{ abierto de } \mathbb{R}^N \\ G \text{ entorno abierto de } a \text{ en } \mathbb{R}^N \\ \Phi : W \to G \text{ difeomorfismo de clase } \mathscr{C}^1 \end{cases}$$

tales que $\Phi\big(W \cap (\mathbb{R}^k \times \{0\})\big) = M \cap G$.

De hecho si (Ω, U, p) es una parametrización cartesiana local de M en a, entonces existe Φ satisfaciendo la condición anterior y además:

$$W \cap (\mathbb{R}^k \times \{0\}) \subset U \times \{0\} \quad y \quad \Phi(x) = p(x_1, \ldots, x_k), \ \ \forall x \in W \cap (\mathbb{R}^k \times \{0\}).$$



Demostración:

Sea $a \in M$.

i) $\Rightarrow$ *ii*) Sea $g : G \longrightarrow \mathbb{R}^m$ como en *i*). Como el rango de la matriz $J_g(a)$ es máximo hay m vectores columna linealmente independientes, que podemos suponer, sin mengua de generalidad, son los últimos, esto es,

$$\det \begin{pmatrix} D_{k+1}g_1(a) & \dots & D_N g_1(a) \\ \dots & \dots & \dots \\ D_{k+1}g_m(a) & \dots & D_N g_m(a) \end{pmatrix} \neq 0.$$

Bajo esta hipótesis adicional consideremos la proyección $\pi : \mathbb{R}^N \longrightarrow \mathbb{R}^k$ dada por $(x_1,\dots,x_N) \longmapsto (x_1,\dots,x_k)$ y denotemos $b := \pi(a)$. El Teorema de la función implícita aplicado a la función g en el punto a, nos asegura que existen

$$\begin{cases} U \text{ entorno abierto de } b \text{ en } \mathbb{R}^k \\ \varphi : U \to \mathbb{R}^m \text{ de clase } \mathscr{C}^1 \\ \Omega \text{ entorno abierto de } a \text{ incluido en } G \end{cases}$$

tales que $\begin{cases} 1)\ (y,\varphi(y)) \in G \quad \text{y} \quad g(y,\varphi(y)) = 0,\ \forall y \in U \\ 2)\ \{x \in \Omega :\ g(x) = 0\} = \{(y,\varphi(y)) :\ y \in U\} \end{cases}$.

Es claro que la igualdad conjuntista 2) puede escribirse en la forma

$$M \cap \Omega = \{(y,\varphi(y)) :\ y \in U\}\ .$$

Finalmente, cambiando Ω por $\Omega \cap \pi^{-1}(U)$ conseguimos que $U = \pi(\Omega)$.

ii) $\Rightarrow$ *iii*) Sea (Ω, U, φ) como en ii). La aplicación $p : U \longrightarrow \mathbb{R}^N$ definida por

$$p(y) = (y, \varphi(y)),$$

es inyectiva (lo es su primera componente), luego biyectiva de U sobre su imagen que es $M \cap \Omega$ y por tanto se cumple que

$$p(U) = M \cap \Omega.$$

p es de clase $\mathscr{C}^1$ ya que lo son sus componentes. Además $p^{-1} : M \cap \Omega \longrightarrow U$ es la restricción a $M \cap \Omega$ de π, luego es claramente continua.

Hemos probado que p es un homeomorfismo de clase $\mathscr{C}^1$ de U sobre $M \cap \Omega$.

Finalmente es claro que

$$J_p(y) = \begin{pmatrix} I \\ \ldots\ldots \\ J_\varphi(y) \end{pmatrix}, \ \forall y \ \in U$$

luego rango $\big(J_p(y)\big) = k, \ \forall y \in U$.

iii) $\Rightarrow$ *iv)* Sea (Ω, U, p) una parametrización cartesiana local de M en a. Sea b el único punto de U tal que $p(b) = a$. Como el rango de la matriz $J_p(b)$ es máximo hay k vectores linealmente independientes en $\{\nabla p_1(b), \ldots, \nabla p_N(b)\}$, que podemos suponer, sin mengua de generalidad, son los primeros. En este supuesto definimos el campo vectorial $\Phi : U \times \mathbb{R}^m \longrightarrow \mathbb{R}^N$ dado por $\Phi(y,z) = p(y) + (0,z)$, esto es,

$$\Phi(y_1, \ldots, y_k, z_{k+1}, \ldots, z_N) =$$

$$\Big(p_1(y_1,\ldots,y_k), \ldots, p_k(y_1,\ldots,y_k), p_{k+1}(y_1,\ldots,y_k) + z_{k+1}, \ldots, p_N(y_1,\ldots,y_k) + z_N\Big).$$

Es claro que ϕ es de clase $\mathscr{C}^1$ en $(b,0)$ con

$$J_\Phi(b,0) = \begin{pmatrix} \nabla p_1(b) & \\ \ldots\ldots\ldots & 0 \\ \nabla p_k(b) & \\ \nabla p_{k+1}(b) & \\ \ldots\ldots\ldots & I \\ \nabla p_N(b) & \end{pmatrix}$$

y en consecuencia det $\big(J_\Phi(b,0)\big) \neq 0$. El Teorema de la función inversa aplicado a la función Φ en el punto $(b,0)$ nos asegura que existen

$$\begin{cases} U^* \ \text{ entorno abierto de } \ b \ \text{ contenido en } \ U \\ V \text{ entorno abierto de } \ 0 \ \text{ en } \ \mathbb{R}^m \\ \Omega^* \ \text{ entorno abierto de } \ a \ \text{ contenido en } \ \Omega \end{cases}$$

tales que Φ es un difeomorfismo de clase $\mathscr{C}^1$ de $U^* \times V$ sobre Ω^*. Es inmediato que la restricción de Φ a $U^* \times \{0\}$ coincide con la restricción de la parametrización p a U^*, con lo que tenemos sendas inclusiones

$$M \supset p(U^*) = \Phi(U^* \times \{0\}) \subset \Omega^*,$$

luego

$$p(U^*) \subset M \cap \Omega^*. \tag{9.1.1}$$

Como p es un homeomorfismo de U sobre $M \cap \Omega$, y U^* es un abierto de U se sigue que $p(U^*)$ es un abierto de $M \cap \Omega$. Luego, teniendo en cuenta la definición de topología relativa, existe un abierto O de $\mathbb{R}^N$ tal que $p(U^*) = (M \cap \Omega) \cap O$ y teniendo en cuenta 9.1.1 concluimos que

$$p(U^*) = (M \cap \Omega^*) \cap O = M \cap (\Omega^* \cap O).$$

Ahora, si notamos $G := \Omega^* \cap O$, tenemos que G es un entorno abierto de a tal que $p(U^*) = M \cap G$ y, como hemos resaltado antes, el difeomorfismo ϕ extiende a la aplicación $p : U^* \longrightarrow M \cap G$.

Finalmente tómese $W := \Phi^{-1}(G)$.

iv) $\Rightarrow$ *i)* Sea (W, G, Φ) como en iv). Consideremos el campo vectorial $g : G \longrightarrow \mathbb{R}^m$ dado por

$$g(x) = (\pi \circ \Phi^{-1})(x),$$

donde $\pi : \mathbb{R}^N \longrightarrow \mathbb{R}^m$ es la proyección dada por $(x_1, \ldots, x_N) \longmapsto (x_{k+1}, \ldots, x_N)$. Se tiene que g es de clase $\mathscr{C}^1$ con

$$\text{rango}\,\big(J_g(x)\big) = m,\ \forall x \in G$$

ya que rango $\big(J_{\Phi^{-1}}(x)\big) = N,\ \forall x \in G$. Comprobemos finalmente que

$$M \cap G = \{x \in G :\ g(x) = 0\}.$$

En efecto,

$$\begin{aligned} M \cap G &= \Phi\big(W \cap (\mathbb{R}^k \times \{0\})\big) = \Phi\big(\Phi^{-1}(G) \cap (\mathbb{R}^k \times \{0\})\big) = G \cap \Phi(\mathbb{R}^k \times \{0\}) \\ &= \{x \in G :\ \Phi^{-1}(x) \in \mathbb{R}^k \times \{0\}\} = \{x \in G :\ \big(\pi \circ \Phi^{-1}\big)(x) = 0\} \\ &= \{x \in G :\ g(x) = 0\}\ . \end{aligned}$$

■

Nota 9.3.

a) Perfecciones adicionales de alguna de las aplicaciones g, φ, p, Φ inciden en las correspondientes perfecciones en el resto. Ello es consecuencia de que el procedimiento seguido para la “obtención” de cada una involucra a los Teoremas de la inversa y de la función implícita y a construcciones analíticas magníficas (de clase $\mathscr{C}^\infty$).

b) Las variedades en $\mathbb{R}^N$ de dimensión k con $k < N$ tienen interior vacío. En efecto, en otro caso, la definición de variedad por difeomorfismos espaciales nos llevaría a la existencia de un difeomorfismo de un abierto de $\mathbb{R}^k$ sobre un abierto de $\mathbb{R}^N$, y en consecuencia ambos espacios serían topológicamente isomorfos, con lo que se concluiría $k = N$.

Corolario 9.4 (Dimensión de una variedad)**.** *Sean N un natural y M un subconjunto de $\mathbb{R}^N$ que admite en cada punto parametrizaciones cartesianas locales. Sean (Ω_i, U_i, p_i) para $i =$ 1, 2 dos de estas parametrizaciones tales que el conjunto*

$$C := M \cap \Omega_1 \cap \Omega_2$$

es no vacío. Entonces la aplicación

$$p_2^{-1} \circ p_1 : p_1^{-1}(C) \longrightarrow p_2^{-1}(C)$$

es un difeomorfismo de clase $\mathscr{C}^1$ y en consecuencia el número k asociado a ambas parametrizaciones coincide.

Demostración:

Para $i = 1,2$, se tiene que $p_i^{-1}(C) = p_i^{-1}(\Omega_1 \cap \Omega_2)$ es un abierto relativo del abierto U_i, y por tanto abierto. Es inmediato que $p_i : p_i^{-1}(C) \longrightarrow C$ es un homeomorfismo y por tanto la aplicación

$$p_2^{-1} \circ p_1 : p_1^{-1}(C) \longrightarrow p_2^{-1}(C)$$

es un homeomorfismo.

Dado $a \in C$, veamos que $p_2^{-1} \circ p_1$ es de clase $\mathscr{C}^1$ en un entorno abierto de $p_1^{-1}(a)$ contenido en $p_1^{-1}(C)$, con lo que el carácter local de la continuidad y derivabilidad nos asegura que $p_2^{-1} \circ p_1$ es de case $\mathscr{C}^1$ en $p_1^{-1}(C)$. En efecto, al ser $\big(\Omega_1 \cap \Omega_2, p_2^{-1}(C), p_2\big)$ una parametrización cartesiana local de M en a , sabemos que existen

$$\begin{cases} W \text{ abierto de } \mathbb{R}^N \\ G \text{ entorno abierto de } a \text{ en } \mathbb{R}^N \\ \Phi : W \to G \text{ difeomorfismo de clase } \mathscr{C}^1 \end{cases}$$

tales que

$$\begin{cases} W \cap (\mathbb{R}^k \times \{0\}) \subset p_2^{-1}(C) \times \{0\} \\ \Phi\big(W \cap (\mathbb{R}^k \times \{0\})\big) = M \cap G \\ p_2 \circ \pi_{|W \cap (\mathbb{R}^k \times \{0\})} = \Phi_{|W \cap (\mathbb{R}^k \times \{0\})} \end{cases}$$

donde π denota la proyección en las primeras k coordenadas.

Finalmente nótese que

$$p_2^{-1} \circ p_{1_{|p_1^{-1}(G)}} = \big(\pi \circ \Phi^{-1} \circ p_1\big)_{|p_1^{-1}(G)},$$

igualdad que prueba que $p_2^{-1} \circ p_1$ es de clase $\mathscr{C}^1$ en $p_1^{-1}(G)$. Sin más que intercambiar los papeles $p_1^{-1} \circ p_2$ es también de clase $\mathscr{C}^1$. Luego $p_2^{-1} \circ p_1$ es un difeomorfismo de clase $\mathscr{C}^1$ como se quería demostrar. Ahora, la unicidad de k se justifica razonando como en la Nota b) de (9.3) . ■

Ejemplos 9.5.

a) Sean $k, N \in \mathbb{N}$ con $k < N$. Las variedades afines de dimensión k en $\mathbb{R}^N$ son variedades diferenciables en $\mathbb{R}^N$ de dimensión k (véase el Ejercicio 9.1).

b) La esfera unidad euclídea en $\mathbb{R}^N$

$$S_{\|\cdot\|_2}(0,1) := \{x \in \mathbb{R}^N : \ \|x\|_2 = 1\}$$

es una variedad de dimensión $N-1$.

La aplicación $g : \mathbb{R}^N \longrightarrow \mathbb{R}$ definida por

$$g(x) = \|x\|_2 - 1$$

es de clase $\mathscr{C}^\infty$ en $G := \mathbb{R}^N \setminus \{0\}$, y para cada $a \in G$, el Ejercicio 3.17 nos da

$$Dg(a)(x) = \frac{(a|x)}{\|a\|_2},$$

esto es,

$$J_g(a) = \frac{(a_1, \ldots, a_N)}{\|a\|_2},$$

y por tanto rango $\big(J_g(a)\big) = 1$.

La variedad determinada por g es

$$\{x \in G : \ \|x\|_2 - 1 = 0\} = S_{\|\cdot\|_2}(0,1) \ .$$

c) Sea $N > 1$. La esfera unidad

$$S_{\|\cdot\|_1}(0,1) := \{x \in \mathbb{R}^N : \ \|x\|_1 = 1\}$$

no es una variedad en $\mathbb{R}^N$, pero está muy cerca de serlo.

El conjunto

$$M = \{x \in S_{\|\cdot\|_1}(0,1) \ : \ x_i \neq 0 \ , \text{ para } \ i = 1, \ldots, N\}$$

es una variedad en $\mathbb{R}^N$ de dimensión $N-1$ globalmente determinada por la aplicación $g : G \longrightarrow \mathbb{R}$ de clase $\mathscr{C}^\infty$ definida por

$$g(x) = \|x\|_1 - 1,$$

donde

$$G = \{x \in \mathbb{R}^N \ : \ x_i \neq 0 \ , \text{ para } \ i = 1, \ldots, N\} \ .$$

En efecto, para cada $a \in G$, en vista del Ejercicio 3.17, se tiene

$$Dg(a)(x) = \sum_{i=1}^N \frac{a_i}{|a_i|} \, x_i, \ \forall x \in \mathbb{R}^N \ ,$$

es decir,

$$J_g(a) = \left(\frac{a_1}{|a_1|}, \ldots, \frac{a_N}{|a_N|} \right) \ ,$$

y por tanto rango $\big(J_g(a)\big) = 1, \ \forall a \in G$.

El motivo por el cual $S_{\|\cdot\|_1}(0,1)$ no es una variedad diferenciable es totalmente intuible: <u>los puntos que hemos excluido de $S_{\|\cdot\|_1}(0,1)$ para obtener M son puntos “esquinas” de $S_{\|\cdot\|_1}(0,1)$.</u> Comprobemos seguidamente que la intuición es correcta. Por comodidad trabajaremos en $\mathbb{R}^2$. Si $S_{\|\cdot\|_1}(0,1)$ fuese una variedad, existiría un entorno abierto U de $(1,0)$ en $\mathbb{R}^2$ y un campo escalar $h : U \longrightarrow \mathbb{R}$ de clase $\mathscr{C}^1$ tales que

$$\left\{ \begin{array}{l} \text{rango } \big(J_h(x,y)\big) = 1, \ \forall (x,y) \in U \\ U \cap S_{\|\cdot\|_1}(0,1) = \{(x,y) \in U : \ h(x,y) = 0\} \end{array} \right. .$$

Si consideramos $u^+ = (1,1),\ u^- = (1,-1)$, se tiene que

$$(1,0)+tu^{\pm} \in S_{\|\cdot\|_1}(0,1),\ \forall t \in]-1,0[\ ,$$

y en consecuencia

$$Dh(1,0)(u^{\pm}) = \lim_{t\to 0^-} \frac{h((1,0)+tu^{\pm})-h(1,0)}{t} = 0.$$

Puesto que $\{u^+,u^-\}$ genera $\mathbb{R}^2$, se concluye que $Dh(1,0)=0$, lo que contradice el hecho de que rango$\big(J_h(1,0)\big) = 1$.

Pese a que $S_{\|\cdot\|_1}(0,1)$ no es una variedad, sin embargo es unión disjunta de dos variedades: M (de dimensión 1) y $\{(1,0),(-1,0),(0,1),(0,-1)\}$ (de dimensión 0).

9.2. Espacios tangente y normal.

Puesto que las variedades son localmente gráficas de funciones derivables, la Sección 3.6 justifica las siguientes definiciones.

Definición 9.6 (espacios tangente y normal)**.** Sean $k,N \in \mathbb{N}$ con $k<N, M$ una variedad de $\mathbb{R}^N$ de dimensión k y a un punto de M.

Se dice que $u\in\mathbb{R}^N$ es un vector tangente a M en a si existe una curva contenida en M que pasa por el punto a y que tiene tangente en a con dirección u, esto es,

$$\exists\delta>0, \exists\gamma:]-\delta,\delta[\longrightarrow\mathbb{R}^N \text{ continua tal que } \gamma\big(]-\delta,\delta[\subset M \text{ con } \begin{cases}\gamma(0)=a\\ \gamma'(0)=u\end{cases}.$$

Se define el espacio tangente $T_M(a)$ a la variedad M en a como el conjunto de vectores tangentes a M en el punto a. Veremos enseguida que $T_M(a)$ es un subespacio vectorial de $\mathbb{R}^N$.

El complemento ortogonal en $\mathbb{R}^N$ del espacio tangente en a, que notaremos $T_M(a)^\perp$, se llama espacio normal a M en a, esto es,

$$T_M(a)^\perp = \{x\in\mathbb{R}^N:\ (x|u)=0,\ \forall u\in T_M(a)\}.$$

A la variedad afín $a+T_M(a)$ (resp. $a+T_M(a)^\perp$) se le llama la variedad afín tangente (resp. normal) a la variedad M en el punto a.

Si $k=1$ (resp. $2, N-1$) la variedad afín tangente recibe el nombre de recta (resp. plano, hiperplano) tangente a M en el punto a.

El siguiente resultado describe estos espacios a través de las funciones "determinación implícita" g y "parametrización cartesiana" p .

Proposición 9.7. *Sean $k,m,N \in \mathbb{N}$ con $m+k=N$, M una variedad en $\mathbb{R}^N$ de dimensión k, a un punto de M y $T_M(a)$ el espacio tangente a M en a .*

i) *Si g es una función que determina (localmente) la variedad M en el punto a, entonces*

$$T_M(a) = \text{Ker } Dg(a),$$

con lo que $T_M(a)$ es un subespacio vectorial de dimensión k, y

$$a + T_M(a) = \{x \in \mathbb{R}^N : \ Dg(a)(x-a) = 0\}.$$

Además

$$\{\nabla g_1(a), \ldots, \nabla g_m(a)\}$$

es una base del espacio vectorial normal $T_M(a)^{\perp}$.

ii) *Si (Ω, U, p) es una parametrización cartesiana local de M en a y b es el único punto de U tal que $p(b) = a$, entonces*

$$T_M(a) = \text{Im } \ Dp(b).$$

Los vectores columna de la matriz $J_p(b)$ son una base de $T_M(a)$ y por tanto

$$a + T_M(a)^{\perp} = \{x \in \mathbb{R}^N : \ \big(D_j p(b)|x-a\big) = 0, \ \forall j = 1, \ldots, k\}.$$

Demostración:

Probaremos en primer lugar que

$$\text{Im } \ Dp(b) \subseteq T_M(a) \subseteq \text{Ker } \ Dg(a). \tag{9.2.1}$$

Empezamos probando la inclusión Im $Dp(b) \subseteq T_M(a)$. Sea u un vector de Im $Dp(b)$. Tomemos $v \in R^k$ tal que $Dp(b)(v) = u$. Puesto que U es abierto podemos tomar $\delta > 0$ tal que $]b - \delta v, b + \delta v[\subset U$. Consideremos la función $\gamma :]-\delta, \delta[\longrightarrow \mathbb{R}^N$ definida por

$$\gamma(t) = p(b + tv).$$

Puesto que p es derivable y con imagen contenida en M, se sigue que γ es una curva contenida en M y con tangente en todo punto. Además

$$\left\{ \begin{array}{l} \gamma(0) = p(b) = a \\ \gamma'(0) = D\gamma(0)(1) = Dp(b)(v) = u \end{array} \right. .$$

Probamos ahora la inclusión $T_M(a) \subseteq \text{Ker } Dg(a)$. Sea $u \in T_M(a)$ y tomemos $\delta > 0$ y $\gamma :]-\delta, \delta[\longrightarrow M$ continua con $\left\{ \begin{array}{l} \gamma(0) = a \\ \gamma'(0) = u \end{array} \right.$. Puesto que $\gamma\big(]-\delta, \delta[\big) \subset M$, se tiene que la función $g \circ \gamma$ está definida y es nula en un entorno de 0, luego el carácter local de la derivabilidad nos asegura que

$$0 = D(g \circ \gamma)(0) = Dg(a) \circ D\gamma(0),$$

y por tanto

$$0 = (Dg(a) \circ D\gamma(0))(1) = Dg(a)\big(D\gamma(0)(1)\big) = Dg(a)\big(\gamma'(0)\big) = Dg(a)(u).$$

Ahora, para que se de la igualdad en 9.2.1 nos basta observar que los espacios extremos tienen la misma dimensión:

$$\dim\big(\text{Ker } Dg(a)\big) = N - \dim\big(\text{Im } Dg(a)\big) = N - m = k = \text{rango}\big(J_p(b)\big) = \dim\big(\text{Im } Dp(b)\big).$$

Hemos probado que

$$\text{Im } \ Dp(b) = T_M(a) = \text{Ker } \ Dg(a).$$

Determinaremos ahora bases de los espacios tangente $T_M(a)$ y normal $T_M(a)^{\perp}$ a la variedad M en el punto a. Si $\{e_1,\dots,e_k\}$ es la base canónica de $\mathbb{R}^k$ se tiene que $\{Dp(b)(e_1),\dots,Dp(b)(e_k)\}$ generan $T_M(a)$. Además, son linealmente independientes por ser los vectores columna de la matriz $J_p(b)$ que tiene rango k, y por tanto constituyen una base de $T_M(a)$.

Como $Dg(a)(x) = 0,\ \forall x \in T_M(a)$, o lo que es lo mismo,

$$\big(\nabla g_j(a)|x\big) = 0,\ \forall x \in T_M(a) \quad (j = 1,\dots,m),$$

se tiene que

$$\{\nabla g_1(a),\dots,\nabla g_m(a)\}$$

son vectores de $T_M(a)^{\perp}$. Además son linealmente independientes por ser los vectores fila de la matriz $J_g(a)$ que tiene rango m, luego constituyen una base de $T_M(a)^{\perp}$, ya que

$$\dim\big(T_M(a)^{\perp}\big) = N - \dim\big(T_M(a)\big) = N - k = m.$$

■

Nota 9.8. Puesto que usualmente las variedades nos vendrán dadas (globalmente) por "determinaciones implícitas" g, y el cálculo de parametrizaciones cartesianas locales puede ser complicado, en la práctica suele ser más útil la descripción del espacio tangente dado por $T_M(a) = \text{Ker } Dg(a)$. Para la determinación de una base de $T_M(a)$ se puede proceder como sigue:

Calculada $J_g(a)$, las ecuaciones implícitas de $T_M(a)$ vienen dadas por

$$J_g(a)x^t = 0,$$

y el teorema de Cramer nos permite pasar a las ecuaciones paramétricas de $T_M(a)$. Los vectores que se obtienen al ir haciendo cada uno de los k parámetros iguales a 1 y el resto iguales a 0, determinan una base de $T_M(a)$.

Por ejemplo, si $T_M(a)$ es de dimensión $N-1$ (hiperplano), y

$$T_M(a) = \{x \in \mathbb{R}^N :\ a_1x_1 + \cdots + a_Nx_N = 0\},$$

entonces, supuesto $a_1 \neq 0$, los vectores

$$\big\{(-a_2,a_1,0,\dots,0),(-a_3,0,a_1,0,\dots,0),\dots,(-a_N,,0,\dots,0,a_1)\big\}$$

son una base de $T_M(a)$ (son vectores de $T_M(a)$ linealmente independientes).

Ejemplo 9.9. Variedades afín tangente y afín normal a la esfera unidad euclídea $M = S_{\|\cdot\|_2}(0,1)$ en un punto a.

La función $g : \mathbb{R}^N \setminus \{0\} \longrightarrow \mathbb{R}$ definida por

$$g(x) = \|x\|_2 - 1$$

determina globalmente la variedad M . Para cada $a \in M$, sabemos que

$$\nabla g(a) = (a_1, \ldots, a_N).$$

La variedad afín tangente a la variedad M en el punto a es el hiperplano de $\mathbb{R}^N$ de ecuación

$$\begin{aligned} a + T_M(a) &= \{x \in \mathbb{R}^N : \ (\nabla g(a)|x - a) = 0\} \\ &= \{x \in \mathbb{R}^N : \ a_1(x_1 - a_1) + \cdots + a_N(x_N - a_N) = 0\} \\ &= \{x \in \mathbb{R}^N : \ a_1 x_1 + \cdots + a_N x_N = 1\}. \end{aligned}$$

La variedad afín normal a M en a es la recta

$$a + T_M(a)^\perp = \{a + t\nabla g(a) \ : \ t \in \mathbb{R}\},$$

es decir, la recta de ecuaciones paramétricas dadas (cambiando $1+t$ por t) por

$$\begin{cases} x_1 = a_1 t \\ \ldots\ldots \\ x_N = a_N t \end{cases} \quad (t \in \mathbb{R}).$$

9.3. Extremos condicionados.

Dedicamos esta sección al estudio de los extremos de campos escalares definidos en subconjuntos abiertos de $\mathbb{R}^N$ condicionados por una variedad diferenciable. Conviene tener presente que en el Tema 7 se trató este problema en el caso particular de variedades de dimensión N, es decir abiertos de $\mathbb{R}^N$. Ahora vamos a desarrollar la teoría para variedades de dimensión k con $0 < k < N$.

Definición 9.10 (Extremo condicionado)**.** Sean X un espacio topológico, $f : X \to \mathbb{R}$ una función y C un subconjunto de X.

Se dice que f alcanza en un punto $a \in C$ un *máximo condicionado* por C si existe un entorno abierto U de a tal que

$$f(x) \leq f(a), \quad \forall x \in U \cap \mathrm{C}.$$

Se dice que f alcanza en un punto $a \in C$ un *máximo condicionado por C estricto* si existe un entorno abierto U de a tal que

$$f(x) < f(a), \quad \forall x \in U \cap C \setminus \{a\}.$$

Si se cambian las desigualdades en las definiciones anteriores, aparecen los conceptos de *mínimo condicionado* y *mínimo condicionado estricto* . Un *extremo* es o bien un máximo condicionado, o bien un mínimo condicionado.

El estudio de los extremos de campos escalares condicionados por conjuntos que admiten una parametrización resulta sencillo, ya que se reduce al estudio de un problema de extremos relativos.

Proposición 9.11 (Relación entre extremos condicionados y relativos)**.** *Sean X un espacio topológico, $f : X \longrightarrow \mathbb{R}$ una función, C un subconjunto de X y a un punto de C. Si U es un espacio topológico, $p : U \longrightarrow C$ es un homeomorfismo de U sobre $p(U)$ con $a \in p(U)$ y b es el único punto de U tal que $p(b) = a$, entonces equivalen:*

i) f alcanza en a un máximo (resp. mínimo) condicionado por C.

ii) $f \circ p$ alcanza en b un máximo (resp. mínimo) relativo .

También se da la equivalencia para extremos condicionados estrictos y extremos relativos estrictos.

La anterior proposición reduce el problema de la determinación de los extremos de un campo escalar f condicionado por una variedad M, al estudio de los extremos relativos de la composición de f con parametrizaciones cartesianas locales de M. A este respecto resaltamos la importancia de la implicación i) $\Rightarrow$ iii) del Teorema 9.2.

¿Cómo proceder habida cuenta de que las parametrizaciones cartesianas locales de una variedad no suelen ser conocidas? La respuesta a esta cuestión fundamental nos la da el siguiente resultado debido a Lagrange que proporciona una condición necesaria de extremo condicionado que es similar a la condición necesaria de existencia de extremo relativo.

Teorema 9.12 (Lagrange (Condición necesaria de extremo condicionado))**.** *Sean $k,m,N \in \mathbb{N}$ con $k+m=N$ y M una variedad en $\mathbb{R}^N$ de dimensión k, G un abierto de $\mathbb{R}^N$ que contiene a M y $f : G \longrightarrow \mathbb{R}$ un campo escalar derivable. Si la función f alcanza en un punto a de la variedad M un extremo condicionado por dicha variedad y g es una función que determina localmente M en a, entonces existen únicos números reales $\alpha_1,\ldots,\alpha_m \in \mathbb{R}$ tales que*

$$D(f+\alpha_1 g_1+\ldots+\alpha_m g_m)(a)=0.$$

Demostración:

Sea (Ω, U, p) una parametrización cartesiana local de M en a y sea $b \in U$ tal que $p(b)$=a. La Proposición 9.11 y la condición necesaria de extremo relativo nos aseguran que

$$D(f \circ p)(b)=0.$$

La regla de la cadena nos dice en consecuencia que

$$Df(a) \circ Dp(b)=0,$$

expresión que se escribe equivalentemente

$$Df(a)(\operatorname{Im} Dp(b))=\{0\}.$$

Al ser, en virtud de la Proposición 9.7, Im $Dp(b) = T_M(a)$, concluimos que

$$Df(a)(T_M(a)) = 0, \text{ o equivalentemente } \nabla f(a) \in T_M(a)^{\perp}.$$

La citada Proposición 9.7 nos dice también que $\{\nabla g_1(a), \ldots, \nabla g_m(a)\}$ es una base de $T_M(a)^{\perp}$, de donde se deduce la existencia y unicidad de escalares $\alpha_1, \ldots, \alpha_m$ tales que

$$\nabla f(a) + \alpha_1 \nabla g_1(a) + \ldots + \alpha_m \nabla g_m(a) = 0,$$

o equivalentemente

$$Df(a) + \alpha_1 Dg_1(a) + \ldots + \alpha_m Dg_m(a) = 0.$$

De la linealidad de la derivada se sigue finalmente el enunciado. ■

A la vista del teorema anterior parece conveniente hacer un esfuerzo para codificar el proceso que se ha de seguir para determinar los candidatos a extremos condicionados. Aunque el Teorema de Lagrange se ha establecido para variedades no necesariamente globalmente determinadas, las siguientes dos razones justifican que la versión práctica de los resultados obtenidos se haga para variedades globalmente determinadas: por una parte, nuestro estudio es local y vía localización podemos limitarnos a variedades globalmente determinadas y, por otra, nuestro principal interés es la determinación de extremos de campos escalares condicionados por conjuntos que se puedan descomponer en unión finita de variedades globalmente determinadas.

9.4. Cálculo práctico de puntos críticos condicionados. Función de Lagrange, sistema de Lagrange y multiplicadores de Lagrange.

Sean $k, m, N \in \mathbb{N}$ con $k + m = N$, $G \subset \mathbb{R}^N$ un conjunto abierto y $g : G \longrightarrow \mathbb{R}^m$ una función de clase $\mathscr{C}^1$ tal que rango $(J_g(x)) = m, \quad \forall x \in G$.

Sea $M = \{x \in G : g(x) = 0\}$ la variedad de dimensión k determinada por g y supongamos que $f : G \longrightarrow \mathbb{R}$ es un campo escalar derivable. En este ambiente, la *función de Lagrange* $L : G \times \mathbb{R}^m \longrightarrow \mathbb{R}$ viene definida por

$$L(x, \lambda) = f(x) + \lambda_1 g_1(x) + \ldots + \lambda_m g_m(x) \quad (x \in G, \lambda \in \mathbb{R}^m)$$

donde $\lambda = (\lambda_1, \ldots, \lambda_m)$ y $g = (g_1, \ldots, g_m)$.

Es claro que el gradiente de L en un punto (x, λ) de $G \times \mathbb{R}^m$ viene dado por

$$\nabla L(x, \lambda) = \Big(D_1 L(x, \lambda), \ldots, D_N L(x, \lambda), g_1(x), \ldots, g_m(x) \Big).$$

El Teorema de Lagrange (9.12) tiene ahora la siguiente lectura:

Los extremos de f condicionados por M proceden de puntos críticos de la función de Lagrange L.

Esto es, se obtienen de las soluciones del sistema

$$\begin{cases} \nabla L(x,\lambda) = 0 \\ x \in G \end{cases}$$

Reescribiendo el sistema anterior se tiene

$$\begin{cases} D_i f(x) + \lambda_1 D_i g_1(x) + \ldots + \lambda_m D_i g_m(x) = 0 \quad (1 \leq i \leq N) \\ g_j(x) = 0 \quad (1 \leq j \leq m) \\ x \in G \end{cases}$$

que es un sistema (posiblemente no lineal) de $N+m$ ecuaciones con $N+m$ incógnitas, llamado *sistema de Lagrange* . Si

$$a = (a_1, \ldots, a_N), \quad \alpha = (\alpha_1, \ldots, \alpha_m)$$

es una solución del sistema de Lagrange, entonces en vista del Teorema de Lagrange (9.12) α está determinado de forma única y a sus coordenadas se les llama los *multiplicadores de Lagrange* para el punto a.

Para justificar una primera aplicación vistosa del Teorema de Lagrange empecemos recordando el procedimiento ya codificado para funciones de una variable. La propiedad de compacidad asegura que si I es un intervalo compacto de $\mathbb{R}$ y $f : I \longrightarrow \mathbb{R}$ es una función continua, entonces f tiene extremos absolutos. Es claro que para detectar los puntos de I en los que la función f alcanza sus valores máximo y mínimo absolutos basta localizar los puntos de $\overset{\circ}{I}$ en los que alcanza extremos relativos y contrastar los valores extremos relativos tomados con los valores de la función en los extremos de I. Si suponemos además que f es derivable en $\overset{\circ}{I}$, entonces los únicos puntos en los que f puede alcanzar el máximo o el mínimo son los extremos del intervalo y aquellos puntos de $\overset{\circ}{I}$ en los que se anula f'.

El estudio se complica si la función f está definida en un compacto de $\mathbb{R}^N$ con $N \geq 2$, pues en este caso la frontera del compacto puede ser un conjunto infinito.

9.5. Aplicación del Teorema de Lagrange al cálculo de extremos absolutos.

La teoría de extremos relativos y el Teorema de Lagrange nos permiten llevar a cabo el estudio de los extremos absolutos de un campo escalar f derivable sobre un subconjunto de $\mathbb{R}^N$ que se pueda expresar como unión finita de variedades. Es claro que si el campo escalar f alcanza en un punto un extremo absoluto, entonces f alcanza en dicho punto un extremo condicionado por la variedad a la que pertenece el punto. Por tanto

- Si la variedad es de dimensión N, esto es, un abierto de $\mathbb{R}^N$, entonces los posibles puntos candidatos a que f alcance un extremo absoluto tienen que ser puntos críticos de f.

- Si la variedad es de dimensión 0, esto es, un conjunto de puntos aislados de $\mathbb{R}^N$, entonces todos los puntos de la variedad son candidatos ya que f alcanza en ellos un extremo condicionado por la variedad.
- Finalmente si la variedad es de dimensión k con $0 < k < N$, entonces los posibles puntos en los que f alcance un extremo absoluto tienen que ser soluciones del correspondiente sistema de Lagrange.

Así, el procedimiento a seguir es el siguiente:

1. Detectar los puntos críticos de la función sobre cada una de las variedades en las que se ha descompuesto el conjunto.
2. Evaluar la función en todos estos puntos y finalmente comparar estos valores entre sí.

Ejemplos 9.13.

a) Calcular los extremos absolutos de la función definida por

$$f(x,y) = 2xy, \quad ((x,y) \in \mathbb{R}^2)$$

condicionados por el conjunto

$$K = \Big\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\Big\}.$$

La propiedad de compacidad nos asegura que f alcanza el máximo y el mínimo absolutos en K. Si el punto donde f alcanza un extremo absoluto pertenece al interior de K, entonces f alcanza en él un extremo relativo y en consecuencia las derivadas parciales de f en dicho punto son nulas, esto es, el punto es solución del sistema

$$\nabla f(x,y) = 0 \Leftrightarrow \left\{ \begin{array}{c} D_1 f(x,y) = 0 \\ D_2 f(x,y) = 0 \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{c} 2y = 0 \\ 2x = 0 \end{array} \right\} \Leftrightarrow (x,y) = (0,0).$$

En otro caso el punto pertenece a la frontera de K, esto es, al conjunto

$$M = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}.$$

Estudiamos ahora los extremos condicionados por M. Obsérvese que M es la variedad en $\mathbb{R}^2$ de dimensión uno determinada globalmente por la función

$$g(x,y) = x^2 + y^2 - 1, \quad ((x,y) \in \mathbb{R}^2 \backslash \{(0,0)\}).$$

Función de Lagrange: $L : \mathbb{R}^2 \backslash \{(0,0)\} \times \mathbb{R} \longrightarrow \mathbb{R}$ definida por

$$L(x,y,\lambda) = 2xy + \lambda(x^2 + y^2 - 1).$$

Sistema de Lagrange:

$$\left\{ \begin{array}{c} 2y + 2\lambda x = 0 \\ 2x + 2\lambda y = 0 \\ x^2 + y^2 = 1 \end{array} \right\} \Rightarrow x(1 - \lambda^2) = 0 \Rightarrow \lambda = \pm 1.$$

Solución:

$$\left(\frac{\sqrt{2}}{2},\frac{-\sqrt{2}}{2}\right),\quad \left(\frac{-\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)\quad (\lambda=1),$$
$$\left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right),\quad \left(\frac{-\sqrt{2}}{2},\frac{-\sqrt{2}}{2}\right)\quad (\lambda=-1),$$

Sólo falta comparar los valores que toma f en cada uno de estos puntos:

$$f(0,0)=0,\ f\left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)=f\left(\frac{-\sqrt{2}}{2},\frac{-\sqrt{2}}{2}\right)=1,\ f\left(\frac{-\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)=f\left(\frac{\sqrt{2}}{2},\frac{-\sqrt{2}}{2}\right)=-1$$

En consecuencia el máximo de f en K vale 1 y se alcanza en los puntos $\left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)$ y $\left(\frac{-\sqrt{2}}{2},\frac{-\sqrt{2}}{2}\right)$. El mínimo vale -1 y se alcanza en $\left(\frac{-\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)$ y $\left(\frac{\sqrt{2}}{2},\frac{-\sqrt{2}}{2}\right)$.

b) Determinar las constantes de equivalencia óptimas de las normas $\|\cdot\|_1$ y $\|\cdot\|_p$ en $\mathbb{R}^3$, donde $(1<p<\infty)$.

El problema consiste en determinar la mayor constante α y la menor constante β tales que
$$\alpha\|x\|_1\le\|x\|_p\le\beta\|x\|_1,\quad \forall x\in\mathbb{R}^3,$$
o equivalentemente,
$$\alpha\le\|x\|_p\le\beta,\quad \forall x\in\mathbb{R}^3:\|x\|_1=1,$$
y por tanto, se trata de determinar los valores máximo y mínimo absolutos de la función $\|\cdot\|_p$ sobre la esfera unidad para la norma $\|\cdot\|_1$.

Por razones de simetría, podemos restringir nuestra atención a la intersección de la esfera con el primer octante, es decir, al conjunto
$$K:=\Big\{(x,y,z)\in\mathbb{R}^3: x+y+z=1, 0\le x,y,z\Big\}.$$

Asimismo, debido al crecimiento de la función de $\mathbb{R}_0^+$ en $\mathbb{R}$ dada por $t\longrightarrow t^{1/p}$, podemos plantearnos el problema en los siguientes términos:

Determinar los extremos absolutos de la función f definida por
$$f(x,y,z)=x^p+y^p+z^p\quad (x,y,z\in\mathbb{R}_0^+)$$
en el conjunto
$$K:=\Big\{(x,y,z)\in\mathbb{R}^3: x+y+z=1, 0\le x,y,z\Big\}.$$

La continuidad de f y la compacidad de K nos garantizan la existencia de los extremos absolutos.

El conjunto K es una unión disjunta de
$$\binom{3}{1}+\binom{3}{2}+\binom{3}{3}=7$$
variedades diferenciables. En efecto:

- **3 variedades de dimensión** 0:

$$M_1 = \{(1,0,0)\}, \quad M_2 = \{(0,1,0)\}, \quad M_3 = \{(0,0,1)\}$$

- **3 variedades de dimensión** 1:

$$M_4 = \{(x,y,0) \in \mathbb{R}^3 : x+y=1, 0<x, 0<y\}$$

que está globalmente determinada por la función definida por

$$g(x,y,z) = (x+y+z-1, z), \quad (x,y,z) \in G := \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}$$

y análogamente las variedades

$$M_5 = \{(x,0,z) \in \mathbb{R}^3 : x+z=1, 0<x, 0<z\},$$

$$M_6 = \{(0,y,z) \in \mathbb{R}^3 : y+z=1, 0<y, 0<z\}.$$

- **Una variedad de dimensión** 2:

$$M_7 = \Big\{(x,y,z) \in \mathbb{R}^3 : x+y+z=1, 0<x,y,z\Big\},$$

que está globalmente determinada por la función definida por

$$g(x,y,z) = x+y+z-1, \quad (x,y,z) \in G = \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+.$$

Calculemos los posibles extremos de la función f condicionados por M_4. Puesto que f es derivable en $\mathbb{R}^3$ y la variedad M_4 puede verse como la variedad globalmente determinada por la función $g : G = \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R} \longrightarrow \mathbb{R}^2$ definida por

$$g(x,y,z) = (x+y+z-1, z)$$

se puede aplicar el Teorema de Lagrange.

Función de Lagrange: $L : G \times \mathbb{R}^2 \longrightarrow \mathbb{R}$ definida por

$$L(x,y,z,\lambda,\mu) = x^p + y^p + |z|^p + \lambda(x+y+z-1) + \mu z.$$

Sistema de Lagrange:

$$\left\{ \begin{array}{c} px^{p-1} + \lambda = 0 \\ py^{p-1} + \lambda = 0 \\ p \text{ signo } (z)|z|^{p-1} + \lambda + \mu = 0 \\ x+y+z = 1 \\ z = 0 \\ x,y > 0 \end{array} \right.$$

Solución:

$$a = \left(\frac{1}{2}, \frac{1}{2}, 0\right), \quad (\alpha, \beta) = \left(\frac{-p}{2^{p-1}}, \frac{p}{2^{p-1}}\right).$$

Este punto es el único en M_4 en el que f puede alcanzar un extremo absoluto en K.

El mismo procedimiento se sigue para las variedades M_5 y M_6, que nos proporcionan las soluciones $\left(\frac{1}{2},0,\frac{1}{2}\right)$ y $\left(0,\frac{1}{2},\frac{1}{2}\right)$, respectivamente.

De manera análoga podemos calcular los posibles extremos de la función f condicionados por M_7 aplicando el Teorema de Lagrange.

Función de Lagrange: $L : G\times\mathbb{R} \longrightarrow \mathbb{R}$ definida por

$$L(x,y,z,\lambda) = x^p + y^p + z^p + \lambda(x+y+z-1).$$

Sistema de Lagrange:

$$\begin{cases} px^{p-1}+\lambda = 0 \\ py^{p-1}+\lambda = 0 \\ pz^{p-1}+\lambda = 0 \\ x+y+z = 1 \\ x,y,z > 0 \end{cases}$$

Solución:

$$a = \left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right), \quad \alpha = \frac{-p}{3^{p-1}}.$$

Este punto es el único de M_7 donde f puede alcanzar un extremo absoluto en K.

Por consiguiente cada variedad aporta un único punto, por tanto hay siete candidatos. Basta evaluar la función en ellos para concluir que el máximo de f en K es 1 y este valor lo alcanza en los puntos que aportan las variedades de dimensión 0 y el valor mínimo es $\frac{1}{3^{p-1}}$ y lo alcanza en el punto que aporta la variedad de dimensión 2.

Como consecuencia de nuestros resultados se sigue que, si $\frac{1}{p}+\frac{1}{q}=1$, entonces

$$\left(\frac{1}{3^{p-1}}\right)^{\frac{1}{p}} = \frac{1}{3^{\frac{1}{q}}} \le \|x\|_p \le 1, \quad \forall x\in\mathbb{R}^3, \|x\|_1 = 1,$$

y las cotas obtenidas son inmejorables. En consecuencia, las constantes de equivalencia óptimas de las normas $\|\cdot\|_1$ y $\|\cdot\|_p$ en $\mathbb{R}^3$ son

$$\frac{1}{3^{\frac{1}{q}}}\|x\|_1 \le \|x\|_p \le \|x\|_1, \quad \forall x\in\mathbb{R}^3.$$

Nota: El estudio en $\mathbb{R}^N$ es análogo. El conjunto

$$K = \left\{x\in\mathbb{R}^N : x_1+\ldots+x_N = 1, 0\le x_1,\ldots,x_N\right\}$$

es unión disjunta de

$$\binom{N}{1}+\binom{N}{2}+\ldots+\binom{N}{N} = 2^N-1$$

variedades. Cada una aporta un único posible extremo absoluto de f. Al igual que ocurre en dimensión 3, el máximo se alcanza en los puntos críticos de las variedades de dimensión 0 (puntos) y el mínimo en el punto $\left(\frac{1}{N},\ldots,\frac{1}{N}\right)$ que es el que aporta la variedad de dimensión $N-1$. Se obtiene, como consecuencia, el siguiente resultado:

Sean $N\in\mathbb{N}$ y $p\in\mathbb{R}$ con $1<p<\infty$. Las constantes de equivalencia óptimas de las normas $\|\cdot\|_1$ y $\|\cdot\|_p$ en $\mathbb{R}^N$ son:

$$\frac{1}{N^{\frac{1}{q}}}\|x\|_1\le\|x\|_p\le\|x\|_1,\quad\forall x\in\mathbb{R}^N,\ \textit{donde}\ \frac{1}{p}+\frac{1}{q}=1.$$

c) Dado $N\in\mathbb{N}$, calcular el mayor valor de la función f definida por

$$f\big(x_1,\ldots,x_N\big)=x_1\ldots x_N,\quad\forall x=(x_1,\ldots,x_N)\in\mathbb{R}^N$$

en el conjunto

$$K=\Big\{x\in\mathbb{R}^N:\|x\|_2=1,0\le x_i,i=1,\ldots,N\Big\}.$$

La continuidad de f y la compacidad de K garantizan la existencia de máximo absoluto. Nótese que, como en el ejemplo anterior, K es unión disjunta de 2^N-1 variedades y f se anula en todas salvo en

$$M=\Big\{x\in\mathbb{R}^N:\|x\|_2=1,0<x_i,i=1,\ldots,N\Big\}.$$

Es claro entonces que los puntos de K en los que f alcanza su máximo absoluto han de ser puntos de M. Vamos por tanto a calcular los candidatos a extremos de f condicionados por M, la variedad globalmente determinada por la función $g:G=\mathbb{R}^+\times\overset{N}{\cdots}\times\mathbb{R}^+\longrightarrow\mathbb{R}$ definida por

$$g\big(x_1,\cdots,x_N\big)=x_1^2+\ldots+x_N^2-1.$$

Función de Lagrange: $L:G\times\mathbb{R}\longrightarrow\mathbb{R}$ definida por

$$L\big(x_1,\ldots,x_N,\lambda\big)=x_1\ldots x_N+\lambda(x_1^2+\ldots+x_N^2-1).$$

Sistema de Lagrange:

$$\begin{cases} x_2\ldots x_N+2\lambda x_1=0\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ x_1\ldots x_{N-1}+2\lambda x_N=0\\ x_1^2+\cdots+x_N^2=1\\ x_1,\ldots,x_N>0\end{cases}$$

Solución:

$$a=\Big(\frac{1}{\sqrt{N}},\ldots,\frac{1}{\sqrt{N}}\Big),\quad\alpha=\frac{-N}{2\sqrt{N^N}}.$$

Ya que este punto es el único candidato se sigue que en él la función f alcanza el máximo absoluto, que necesariamente (al ser único) es estricto. Se tiene pues que:

$$x_1x_2\cdots x_N\le\Big(\frac{1}{N}\Big)^{N/2},\quad\forall x\in K,$$

y se da la igualdad si, y sólo si,

$$x_1=x_2=\cdots=x_N.$$

Nota: La solución del ejercicio anterior es equivalente a la importante desigualdad entre las medias aritmética y geométrica. Demuéstrese (Ejercicio 9.5).

En la lectura que hicimos del Teorema de Lagrange hemos señalado que es condición necesaria para que f alcance en a un extremo condicionado por la variedad M el hecho de que exista $\alpha \in \mathbb{R}^m$ tal que (a,α) sea un punto crítico de la función de Lagrange: $\nabla L(a,\alpha) = 0$. Si denotamos por L_α a la función parcial de la función de Lagrange obtenida fijando α, esto es, $L_\alpha : G \longrightarrow \mathbb{R}$ definida por

$$L_\alpha(x) = L(x,\alpha), \quad (x \in G),$$

(nótese que L_α es una función sólo de N variables). Es claro que entonces se verifica que

$$DL_\alpha(a) = 0.$$

El siguiente resultado traduce en el contexto de extremos condicionados, las condiciones necesarias y suficientes de existencia de extremos relativos cuando la primera derivada no nula es la segunda.

Teorema 9.14 (Condiciones neces. y suf. de extremo condicionado)**.** *Sean $k,m,N \in \mathbb{N}$ con $k+m=N$, y M una variedad en $\mathbb{R}^N$ de dimensión k, G un abierto de $\mathbb{R}^N$ que contiene a M y $f : G \longrightarrow \mathbb{R}$ un campo escalar derivable. Sean $a \in M$, g una función que determina localmente M en a y $\alpha \in \mathbb{R}^m$ tales que (a,α) es un punto crítico de la función de Lagrange L.*

Si las funciones f y g son dos veces derivables en a y

$$d^2L_\alpha(a)_{|T_M(a)}$$

es no nula se verifican las siguientes afirmaciones:

i) (Condiciones suficientes)
Si
$$d^2L_\alpha(a)(u) < 0,\ \forall u \in T_M(a)\backslash\{0\},$$
entonces f alcanza en a un máximo condicionado por M estricto.

Si
$$d^2L_\alpha(a)(u) > 0,\ \forall u \in T_M(a)\backslash\{0\},$$
entonces f alcanza en a un mínimo condicionado por M estricto.

ii) (Condiciones necesarias)
Si f alcanza en a un máximo condicionado por M, entonces
$$d^2L_\alpha(a)(u) \leq 0,\ \forall u \in T_M(a)\backslash\{0\}.$$

Si f alcanza en a un mínimo condicionado por M, entonces
$$d^2L_\alpha(a)(u) \geq 0,\ \forall u \in T_M(a)\backslash\{0\}.$$

Demostración:

Puesto que g es dos veces derivable en a, podemos considerar una parametrización cartesiana local (Ω, U, p) de M tal que p es dos veces derivable en $b = p^{-1}(a)$. La estrategia de la demostración pasa por la Proposición 9.11 y consiste en aplicar a la función $f \circ p$ en el punto b las condiciones necesarias y suficientes de existencia de extremos relativos.

Como p es dos veces derivable en b y f es dos veces derivable en a, se sigue que $f \circ p$ es dos veces derivable en b.

Puesto que $g \circ p = 0$, se tiene que $g_j \circ p = 0$, $\forall j \in \{1, \ldots, m\}$, luego

$$(\alpha_1 g_1 + \ldots + \alpha_m g_m) \circ p = 0,$$

y en consecuencia, como $L_\alpha = f + (\alpha_1 g_1 + \cdots + \alpha_m g_m)$, tenemos que

$$f \circ p = L_\alpha \circ p.$$

Comprobemos que b es un punto crítico de $f \circ p$, esto es $D(f \circ p)(b) = 0$. En efecto,

$$D(f \circ p)(b) = D(L_\alpha \circ p)(b) = DL_\alpha(a) \circ Dp(b) = 0$$

pues $DL_\alpha(a) = 0$ por hipótesis.

Además, sabemos que para $y \in \mathbb{R}^k$ se verifica (ver Sección 5.3, *ii*)

$$D^2(f \circ p)(b)(y,y) = D^2(L_\alpha \circ p)(b)(y,y) =$$

$$D^2 L_\alpha(a)\Big(Dp(b)(y), Dp(b)(y)\Big) + DL_\alpha(a)\Big(D^2 p(b)(y,y)\Big) =$$

$$= D^2 L_\alpha(a)\Big(Dp(b)(y), Dp(b)(y)\Big).$$

Finalmente, puesto que en virtud de la Proposición 9.7, $Dp(b)$ es una biyección de $\mathbb{R}^k$ sobre $T_M(a)$ podemos cambiar en el enunciado $d^2 L_\alpha(a)(u)$ por $d^2(f \circ p)(b)(y)$ y bastará tener en cuenta la Proposición 9.11 para concluir la demostración del teorema. ∎

Nota: En la utilización práctica del teorema convendrá tener presente la siguiente consecuencia del apartado ii):

Si

$$d^2 L_\alpha(a)(u) \leq 0, \; \forall u \in T_M(a),$$

entonces si f tiene un extremo condicionado en a por M, éste ha de ser máximo. Análogamente, si

$$d^2 L_\alpha(a)(u) \geq 0, \; \forall u \in T_M(a),$$

entonces de alcanzar f en a un extremo condicionado por M, éste ha de ser mínimo.

A la aplicación práctica del resultado anterior dedicaremos el resto del tema.

Estudio práctico de las condiciones de extremo en los puntos críticos.

Nos interesa estudiar la forma cuadrática $d^2L_\alpha(a)$ restringida a $T_M(a)$. Sea $H(a)$ la matriz hessiana de L_α en a, esto es, la matriz asociada con respecto a la base canónica de $\mathbb{R}^N$ de dicha forma cuadrática

$$H(a) = \big(D_{(i,j)}L_\alpha(a)\big)_{1\leq i,j\leq N}.$$

Elegida una base $\{v_1,\ldots,v_k\}$ de $T_M(a)$, consideremos el isomorfismo de $\mathbb{R}^k$ sobre $T_M(a)$ que aplica la base canónica de $\mathbb{R}^k$ sobre $\{v_1,\ldots,v_k\}$ y que está dado por

$$\Phi(y) = \Phi(y_1,\ldots,y_k) := y_1v_1 + \ldots + y_kv_k.$$

Este isomorfismo se puede expresa en términos de una cierta matriz en la forma

$$\Phi(y) = yK(a),\ \forall y \in \mathbb{R}^k.$$

Es claro que $K(a)$ es la matriz cuyos vectores fila son los vectores básicos de $T_M(a)$ elegidos. La forma cuadrática Q en $\mathbb{R}^k$ obtenida restringiendo $d^2L_\alpha(a)$ a $T_M(a)$ previa composición con Φ viene dada por

$$Q(y) := d^2L_\alpha(a)(\Phi(y)),\ \forall y \in \mathbb{R}^k$$

y queda entonces determinada en términos de una matriz como sigue:

$$Q(y) = \Phi(y)H(a)\Phi(y)^t = yK(a)H(a)(yK(a))^t = yK(a)H(a)K(a)^t y^t.$$

Luego la matriz asociada a la forma cuadrática Q respecto de la base canónica de $\mathbb{R}^k$ es

$$H = K(a)H(a)K(a)^t.$$

La información que ahora se puede dar sobre el punto en cuestión requiere aplicar la teoría de extremos relativos, obteniéndose el siguiente criterio

i) Si $\det(H_i) > 0, i = 1,\ldots,k$, entonces f alcanza en a un mínimo condicionado por M estricto, donde H_i es la submatriz principal esquina izquierda superior de orden i de la matriz H.

ii) Si $(-1)^i\det(H_i) > 0, i = 1,\ldots,k$, entonces f alcanza en a un máximo condicionado por M estricto, donde H_i es la submatriz principal esquina izquierda superior de orden i de la matriz H .

iii) Si $\det(E) \geq 0$, para cualquier submatriz principal E de H, entonces de alcanzar f en a un extremo condicionado por M, ha de ser mínimo.

iv) Si $(-1)^{\text{orden }(E)}\det(E) \geq 0$, para cualquier submatriz principal E de H, entonces de alcanzar f en a un extremo condicionado por M, ha de ser máximo.

v) Si existen E y F submatrices principales de H tales que $(-1)^{\text{orden }(E)}\det(E) > 0$ y $(-1)^{\text{orden }(F)}\det(F) < 0$, (lo que en particular ocurre si el determinante de alguna submatriz principal de orden par de H es negativo), entonces f no alcanza en a un extremo condicionado por M.

Ejemplo 9.15. Determinar las dimensiones del ortoedro en $\mathbb{R}^3$ de volumen fijo una unidad y superficie lateral mínima.

La condición a la que están sujetas las variables es la nulidad de la función

$$g : G = \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+ \longrightarrow \mathbb{R} \text{ dada por } g(x,y,z) = xyz - 1.$$

La variedad $M = \{(x,y,z) \in G : xyz = 1\}$ (en $\mathbb{R}^3$ de dimensión 2 que define g) no es compacta. La función a minimizar en M es $f : G \longrightarrow \mathbb{R}$ definida por

$$f(x,y,z) = 2(xy + xz + yz).$$

Función de Lagrange: $L : G \times \mathbb{R} \longrightarrow \mathbb{R}$ dada por

$$L(x,y,z,\lambda) = 2(xy + xz + yz) + \lambda(xyz - 1).$$

Sistema de Lagrange:

$$\begin{cases} 2(y+z) + \lambda yz = 0 \\ 2(x+z) + \lambda xz = 0 \\ 2(x+y) + \lambda xy = 0 \\ xyz = 1 \\ x,y,z > 0 \end{cases}$$

Solución:

$$a = (1,1,1), \quad \alpha = -4.$$

Se tiene que la aplicación de Lagrange parcial $L_\alpha : G \longrightarrow \mathbb{R}$ está definida por

$$L_\alpha(x,y,z) = 2(xy + xz + yz) - 4(xyz - 1),$$

luego

$$H(a) = \left(D_{(i,j)} L_\alpha(a)\right)_{1 \le i,j \le 3} = \begin{pmatrix} 0 & -2 & -2 \\ -2 & 0 & -2 \\ -2 & -2 & 0 \end{pmatrix}.$$

Por otra parte $J_g(a) = (1,1,1)$, luego

$$T_M(a) = \{(x,y,z) \in \mathbb{R}^3 : x + y + z = 0\},$$

y por tanto dos vectores linealmente independientes de dicho espacio son, por ejemplo, $(1,-1,0)$ y $(1,0,-1)$. Luego

$$K(a) = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix},$$

con lo que

$$H = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & -2 & -2 \\ -2 & 0 & -2 \\ -2 & -2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix}.$$

Puesto que $\det(H) = 12$ y $\det(H_1) = 4$ concluimos que f alcanza en $a = (1,1,1)$ un mínimo condicionado por M estricto.

Aunque a sea mínimo estricto y además sea el único punto crítico, la justificación de que es mínimo absoluto no es fácil. Usaremos la desigualdad entre las medias geométrica y aritmética:

En nuestro caso si $0 < x, y, z$ se tiene que

$$\sqrt[3]{x^2y^2z^2} \leq \frac{xy + xz + yz}{3}$$

y se da la igualdad si, y sólo si, $xy = xz = yz$ si, y sólo si $x = y = z$. Como en nuestro caso ha de ser $xyz = 1$, se concluye que

$$3 \leq xy + xz + yz$$

y se da la igualdad si, y sólo si, $x = y = z = 1$. Hemos obtenido que

f alcanza en $(1,1,1)$ un mínimo condicionado por M absoluto estricto, esto es, el ortoedro en $\mathbb{R}^3$ de volumen 1 y superficie lateral mínima es el cubo.

Nota: Para más información, véase el Ejercicio 9.6.

9.6. Referencias recomendadas.

[Cra].

9.7. Resumen de resultados.

Definiciones de variedad. Sean k,m,N números naturales con $k+m=N$. Para $M\subset\mathbb{R}^N$ no vacío equivalen:

a) (**Definición implícita**) M es una variedad de dimensión k en $\mathbb{R}^N$. Es decir, para cada $a\in M$, existen

$$\begin{cases} G \text{ entorno abierto de } a \text{ en } \mathbb{R}^N \\ g: G\to\mathbb{R}^m \text{ de clase } \mathscr{C}^1 \text{ verificando } \operatorname{rango}\big(J_g(x)\big)=m,\ \ \forall x\in G \end{cases}$$

tales que $M\cap G=\{x\in G:\ g(x)=0\}$.

En el caso de que se pueda conseguir una función g como antes que determine la totalidad de M, diremos que la variedad M está <u>globalmente determinada por g</u>.

b) (**Definición por parametrizaciones cartesianas locales**) Para cada $a\in M$, existen

$$\begin{cases} \Omega \text{ entorno abierto de } a \text{ en } \mathbb{R}^N \\ U \text{ abierto de } \mathbb{R}^k \\ p: U\to\mathbb{R}^N \text{ homeomorfismo de clase } \mathscr{C}^1 \text{ de } U \text{ sobre } p(U) \end{cases}$$

tales que $\begin{cases} p(U)=M\cap\Omega \\ \operatorname{rango}\big(J_p(y)\big)=k,\ \ \forall y\in U \end{cases}$.

La terna (Ω,U,p) se denomina <u>una parametrización cartesiana local de M en a</u>.

Proposición. *Sean $k,m,N\in\mathbb{N}$ con $m+k=N$, M una variedad en $\mathbb{R}^N$ de dimensión k, a un punto de M y $T_M(a)$ el espacio tangente a M en a.*

i) *Si g es una función que determina (localmente) la variedad M en el punto a, entonces*

$$T_M(a)=\operatorname{Ker} Dg(a),$$

con lo que $T_M(a)$ es un subespacio vectorial de dimensión k, y

$$a+T_M(a)=\{x\in\mathbb{R}^M:\ Dg(a)(x-a)=0\}.$$

Además

$$\{\nabla g_1(a),\dots,\nabla g_m(a)\}$$

es una base del espacio vectorial normal $T_M(a)^\perp$.

ii) *Si (Ω,U,p) es una parametrización cartesiana local de M en a y b es el único punto de U tal que $p(b)=a$, entonces*

$$T_M(a)=\operatorname{Im}\ Dp(b).$$

Los vectores columna de la matriz $J_p(b)$ son una base de $T_M(a)$ y por tanto

$$a+T_M(a)^\perp=\{x\in\mathbb{R}^N:\ \big(D_jp(b)|x-a\big)=0,\ \forall j=1,\dots,k\}.$$

Extremo condicionado. Sean X un espacio topológico, $f : X \to \mathbb{R}$ una función y C un subconjunto de X.

Se dice que f alcanza en un punto $a \in C$ un máximo condicionado por C si existe un entorno abierto U de a tal que

$$f(x) \leq f(a), \quad \forall x \in U \cap \mathrm{C}.$$

Se dice que f alcanza en un punto $a \in C$ un máximo condicionado por C estricto si existe un entorno abierto U de a tal que

$$f(x) < f(a), \quad \forall x \in U \cap C \backslash \{a\}.$$

Si se cambian las desigualdades en las definiciones anteriores, aparecen los conceptos de mínimo condicionado y mínimo condicionado estricto. Un extremo condicionado es o bien un máximo condicionado, o bien un mínimo condicionado.

Teorema de Lagrange (Condición necesaria de extremo condicionado). *Sean $k, m, N \in \mathbb{N}$ con $k + m = N$ y M una variedad en $\mathbb{R}^N$ de dimensión k, G un abierto de $\mathbb{R}^N$ que contiene a M y $f : G \longrightarrow \mathbb{R}$ un campo escalar derivable. Si la función f alcanza en un punto a de la variedad M un extremo condicionado por dicha variedad y g es una función que determina localmente M en a, entonces existen únicos números reales $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ tales que*

$$D(f + \alpha_1 g_1 + \ldots + \alpha_m g_m)(a) = 0.$$

Teorema (Condiciones neces. y sufic. de extremo condicionado). *Sean $k, m, N \in \mathbb{N}$ con $k + m = N$, y M una variedad en $\mathbb{R}^N$ de dimensión k, G un abierto de $\mathbb{R}^N$ que contiene a M y $f : G \longrightarrow \mathbb{R}$ un campo escalar derivable. Sean $a \in M$, g una función que determina localmente M en a y $\alpha \in \mathbb{R}^m$ tales que (a, α) es un punto crítico de la función de Lagrange L.*

Si las funciones f y g son dos veces derivables en a y

$$d^2 L_\alpha(a)_{|T_M(a)}$$

es no nula se verifican las siguientes afirmaciones:

i) (Condiciones suficientes)
Si

$$d^2 L_\alpha(a)(u) < 0, \quad \forall u \in T_M(a) \backslash \{0\},$$

entonces f alcanza en a un máximo condicionado por M estricto.

Si

$$d^2 L_\alpha(a)(u) > 0, \quad \forall u \in T_M(a) \backslash \{0\},$$

entonces f alcanza en a un mínimo condicionado por M estricto.

ii) (Condiciones necesarias)
Si f alcanza en a un máximo condicionado por M, entonces

$$d^2 L_\alpha(a)(u) \leq 0, \quad \forall u \in T_M(a) \backslash \{0\}.$$

Si f alcanza en a un mínimo condicionado por M, entonces

$$d^2 L_\alpha(a)(u) \geq 0, \quad \forall u \in T_M(a) \backslash \{0\}.$$

Véanse además los apartados sobre:

- Aplicación del Teorema de Lagrange al cálculo de extremos absolutos.
- Estudio práctico de las condiciones de extremo en los puntos críticos.

9.8. Ejercicios del Tema 9

9.1 Sean $k,N \in \mathbb{N}$ con $k < N$. Probar que una variedad afín de dimensión k en $\mathbb{R}^N$ es una variedad (diferenciable) de dimensión k en $\mathbb{R}^N$.

9.2 **(*)** Sea N un número natural mayor que uno. Probar que no existe una parametrización cartesiana global de $S_{\|\cdot\|_2}(0,1)$. Calcular parametrizaciones cartesianas locales de dicha variedad.

Calcula una parametrización cartesiana global de la variedad M del Ejemplo 9.5.c).

9.3 Sea N un número natural mayor que uno. Probar que el conjunto

$$M = \{x \in \mathbb{R}^N : \ \|x\|_\infty = 1 \text{ y } \exists_1 i \in \{1,\ldots,N\} \text{ tal que } |x_i| = 1\}$$

es una variedad en $\mathbb{R}^N$ de dimensión $N-1$ que está globalmente determinada.

9.4 Justificar que cada uno de los subconjuntos de $\mathbb{R}^3$ es una variedad y calcular los espacios tangente y normal en el punto (a,b,c) que se indica:

1. (Elipsoide)

$$M = \left\{(x,y,z) \in \mathbb{R}^3 : \ \frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} + \frac{z^2}{\gamma^2} = 3\right\} \ (\alpha,\beta,\gamma > 0) \ ; \ (a,b,c) = (\alpha,\beta,-\gamma).$$

2. (Paraboloide hiperbólico)

$$M = \left\{(x,y,z) \in \mathbb{R}^3 : \ z = x^2 - y^2\right\} \ ; \ (a,b,c) = (0,0,0).$$

3. (Hélice)

$$M = \left\{(\cos t \ , \operatorname{sen} t \ , \ t) : \ t \in \mathbb{R}\right\} \ ; \ (a,b,c) = (-1,0,\pi).$$

4. (Borde de la bóveda de Viviani)

$$M = \left\{(x,y,z) \in \mathbb{R}^2 \times \mathbb{R}^+ : x^2+y^2+z^2 = 1, \ \left(x - \frac{1}{2}\right)^2 + y^2 = \frac{1}{4}\right\} ; \ (a,b,c) = (0,0,1).$$

5. (Toro)

$$M = \left\{(x,y,z) \in \mathbb{R}^3 : \ \left(\sqrt{x^2+y^2}-2\right)^2 + z^2 = 1\right\} \ ; \ (a,b,c) = (3,0,0).$$

9.5 Probar que para $a = (a_1, \ldots, a_N)$ con $a_i \geq 0 \quad (i = 1, \ldots, N)$ son equivalentes

i) $a_1 \cdots a_N \leq \left(\frac{1}{N}\right)^{\frac{N}{2}} \quad \left(\|a\|_2 = 1\right)$.

ii) $\sqrt[N]{a_1 \cdots a_N} \leq \dfrac{a_1 + \cdots + a_N}{N}$.

Además en ambos casos se da la igualdad si, y sólo si, $a_1 = \cdots = a_N$.

9.6

i) Determinar las dimensiones del ortoedro en $\mathbb{R}^3$ de superficie lateral seis unidades y volumen máximo ¿Existe un ortoedro de volumen mínimo con dicha superficie lateral?

ii) ¿Existe un ortoedro de superficie lateral máxima con volumen fijo una unidad?

9.7 Determinar las dimensiones del ortoedro de mayor volumen que se puede inscribir en el elipsoide

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \quad (a,b,c > 0).$$

9.8 Calcular la distancia mínima de un punto a un plano en $\mathbb{R}^3$.

9.9 Determinar los extremos absolutos de las siguientes funciones en el conjunto K indicado:

$$f(x,y) = 2x^2 - 3y^2 - 2x, \quad K := \left\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 5\right\}$$

$$g(x,y) = 10xy - x^2y - xy^2, \quad K := \left\{(x,y) \in \mathbb{R}^2 : 0 \leq x, 0 \leq y, x+y \leq 10\right\}$$

$$h(x,y) = x^2 + y^2 - xy - x - y, \quad K := \left\{(x,y) \in \mathbb{R}^2 : 0 \leq x, 0 \leq y, x+y \leq 3\right\}$$

$$u(x,y) = x^2 - 2xy + y^2, \quad K := \left\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 2x\right\}.$$

$$v(x,y) = y - x, \quad K := \left\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 - 2x = 0\right\}.$$

9.10 Estudiar los extremos la función

$$f(x,y,z) = xyz, \quad \forall (x,y,z) \in \mathbb{R}^3$$

condicionados por el conjunto

$$\left\{(x,y,z) \in \mathbb{R}^3 : x + y + z = 1\right\}.$$

9.9. Soluciones de los ejercicios del Tema 9.

9.1 Si la variedad afín viene dada como en la motivación, esto es,

$$M = \{x \in \mathbb{R}^N : \ Ax^t + b = 0\},$$

donde la matriz de números reales $A = \left(a_{ij}\right)_{\substack{1\le i\le N-k \\ 1\le j\le N}}$ tiene rango máximo, es decir, $N-k$, y $b \in \mathbb{R}^{N-k}$, entonces es inmediato que M es una variedad globalmente determinada por la función $g : \mathbb{R}^N \longrightarrow \mathbb{R}^{N-k}$ definida por

$$g(x) = Ax^t + b$$

que es de clase $\mathscr{C}^\infty$ y, como $J_g(x) = A$, $\forall x \in \mathbb{R}^N$, se verifica también que rango $\left(J_g(x)\right) = N-k$, $\forall x \in \mathbb{R}^N$.

Es claro que el espacio tangente a M en cualquier punto es $\{x \in \mathbb{R}^N : \ Ax^t = 0\}$.

Si por el contrario la variedad viene dada por $M = a + S$, con $a \in \mathbb{R}^N$ y S un subespacio de dimensión k de $\mathbb{R}^N$. Consideramos el espacio cociente

$$\mathbb{R}^N/S \cong \mathbb{R}^m,$$

con $m = N-k$ y la aplicación proyección canónica $T : \mathbb{R}^N \longrightarrow \mathbb{R}^m$ dada por

$$T(x) = x + S \equiv (x_{k+1}, \ldots, x_N)$$

(Sea $\{u_1, \ldots, u_k\}$ una base del subespacio S. Dicha base se puede prolongar de manera que

$$\{u_1, \ldots, u_k, u_{k+1}, \ldots, u_N\}$$

sea una base de $\mathbb{R}^N$. Sea ahora $x_1u_1 + \cdots + x_Nu_N$ un representante de la clase, entonces

$$(x_1u_1 + \cdots + x_nu_N) + S \longmapsto (x_{k+1}, \ldots, x_N).$$

Es claro que T es una aplicación lineal y sobreyectiva, luego el rango de la matriz asociada a T es m, y S es el núcleo de T. La aplicación $g : \mathbb{R}^N \longrightarrow \mathbb{R}^m$ definida por

$$g(x) = T(x-a) = T(x) - T(a)$$

es derivable en todo punto con derivada $Dg(x) = T$, $\forall x \in \mathbb{R}^N$, luego g es de clase $\mathscr{C}^\infty$. La variedad diferenciable determinada globalmente por g es

$$\{x \in \mathbb{R}^N : g(x) = 0\} = \{x \in \mathbb{R}^N : T(x-a) = 0\} =$$

$$\{x \in \mathbb{R}^N : x - a \in S\} = \{x \in \mathbb{R}^N : x \in a + S\} = M.$$

9.2 $S_{\|\cdot\|_2}(0,1)$ es un compacto de $\mathbb{R}^N$ luego no puede ser homeomorfo a un abierto de $\mathbb{R}^{N-1}$. En consecuencia $S_{\|\cdot\|_2}(0,1)$ no admite una parametrización cartesiana global. Sin embargo $S_{\|\cdot\|_2}(0,1)$ sí tiene parametrizaciones locales sencillas. Por ejemplo si $a_N \neq 0$, tómese U la bola unidad euclídea abierta de $\mathbb{R}^{N-1}$ y

$$p(y) = \begin{cases} \left(y, \sqrt{1-\|y\|_2^2}\right) & \text{si } \ a_N > 0, \ \text{ con } \ \Omega = U \times \mathbb{R}^+ \\ \left(y, -\sqrt{1-\|y\|_2^2}\right) & \text{si } \ a_N < 0, \ \text{ con } \ \Omega = U \times \mathbb{R}^- \end{cases}.$$

Veamos en primer lugar el caso particular $N = 2$. Sean $\Omega^{\pm} = \mathbb{R} \times \mathbb{R}^{\pm}$ los semiplanos abiertos, y consideremos el abierto de $\mathbb{R}$

$$U := \{x \in \mathbb{R} : \ 0 < |x| < 1\} =]-1,0[\cup]0,1[.$$

Consideremos por último las parametrizaciones

$$p_{\pm}(x) := \big(x, \pm(1-|x|)\big), \ \forall x \in U.$$

Entonces

- (Ω^+, U, p_+) es una parametrización cartesiana local para todo punto de M con ordenada positiva.
- (Ω^-, U, p_-) es una parametrización cartesiana local para todo punto de M con ordenada negativa.

Una parametrización cartesiana global se consigue separando el dominio de definición de ambas: Si notamos $\begin{cases} U^- = U - 1 =]-2,-1[\cup]-1,0[\\ U^+ = U + 1 =]0,1]\cup]1,2[\end{cases}$, entonces la aplicación $p : U^- \cup U^+ \longrightarrow \mathbb{R}^2$ dada por

$$p(x) = \begin{cases} p_-(x+1) & \text{si } \ x \in U^- \\ p_+(x-1) & \text{si } \ x \in U^+ \end{cases}$$

es una parametrización cartesiana global de M.

Veamos ahora el caso general. Sean

$$\Omega^+ = \{x \in \mathbb{R}^N : \ 0 < x_N\} \quad \text{y} \quad \Omega^- = \{x \in \mathbb{R}^N : \ x_N < 0\}$$

los dos semiespacios abiertos de $\mathbb{R}^N$. Consideremos el abierto de $\mathbb{R}^{N-1}$ dado por

$$U = \{(x_1, \ldots, x_{N-1}) \in \mathbb{R}^{N-1} : \ \ x_1, \ldots, x_{N-1} \neq 0, \ \ |x_1| + \cdots + |x_{N-1}| < 1\}.$$

Las aplicaciones $p_{\pm} : U \longrightarrow \mathbb{R}^N$ dadas por

$$p_{\pm}(x_1, \ldots, x_{N-1}) = \big(x_1, \ldots, x_{N-1}, \pm(1-|x_1| - \cdots - |x_{N-1}|)\big)$$

son respectivamente parametrizaciones cartesianas locales de $M \cap \Omega^{\pm}$. En consecuencia, si consideramos los abiertos

$$U^- = U - (1,\ldots,1) \quad \text{y} \quad U^+ = U + (1,\ldots,1),$$

entonces $(\mathbb{R}^N, U^+ \cup U^-, p)$, donde

$$p(x_1,\ldots,x_{N-1}) = \begin{cases} p_-\big((x_1,\ldots,x_{N-1}) + (1,\ldots,1)\big) & \text{si } (x_1,\ldots,x_{N-1}) \in U^- \\ p_+\big((x_1,\ldots,x_{N-1}) + (1,\ldots,1)\big) & \text{si } (x_1,\ldots,x_{N-1}) \in U^+ \end{cases},$$

es una parametrización cartesiana global de M .

9.3 La función

$$g(x) = \|x\|_\infty - 1, \;\; \forall x \in G = \{x \in R^N : \exists_1 i \in \{1,\ldots,N\} \text{ tal que } |x_i| = \|x\|_\infty\}$$

es de clase $\mathscr{C}^\infty$ verificando $J_g(a) = \left(0,\ldots,0,\dfrac{a_i}{|a_i|},0,\ldots,0\right)$, para cada $a \in G$ con $|a_i| = \|a\|_\infty$, y en consecuencia

$$\text{rango } \big(J_g(a)\big) = 1, \; \forall a \in G.$$

Así M es una variedad en $\mathbb{R}^N$ de dimensión $N-1$.

9.4 Todos los casos son variedades propias en $\mathbb{R}^3$ y en consecuencia de dimensión 2 (resp. 1). Se trata pues de describir el plano (resp. la recta) tangente y la recta (resp. el plano) normal a la variedad M en el punto (a,b,c).

Conviene recordar que si $A^2+B^2+C^2, \alpha^2+\beta^2+\gamma^2 > 0$, entonces se tiene que

$$\left.\begin{array}{l} \pi \equiv A(x-a)+B(y-b)+C(z-c)=0 \\ r \equiv \left\{\begin{array}{ll} x = & a+\alpha t \\ y = & b+\beta t \\ z = & c+\gamma t \end{array}\right\} \qquad (t \in \mathbb{R}) \end{array}\right\} \Longrightarrow$$

$$\begin{array}{l} \pi^\perp \equiv \left\{\begin{array}{l} x = a+At \\ y = b+Bt \\ z = c+Ct \end{array}\right\} \qquad (t \in \mathbb{R}) \\ r^\perp \equiv \alpha(x-a)+\beta(y-b)+\gamma(z-c)=0 \end{array}$$

1. (Elipsoide) La función

$$g(x,y,z) = \frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} + \frac{z^2}{\gamma^2} - 3, \;\; \forall (x,y,z) \in G = \mathbb{R}^3 \setminus \{(0,0,0)\}$$

es de clase $\mathscr{C}^\infty$ con $J_g(x,y,z) = \left(\dfrac{2x}{\alpha^2}, \dfrac{2y}{\beta^2}, \dfrac{2z}{\gamma^2}\right)$, y en consecuencia

$$\text{rango } \big(J_g(x,y,z)\big) = 1, \; \forall (x,y,z) \in G.$$

Así M es una variedad en $\mathbb{R}^3$ de dimensión 2.

Plano tangente: Como $J_g(\alpha,\beta,-\gamma) = \left(\frac{2}{\alpha},\ \frac{2}{\beta},\ -\frac{2}{\gamma}\right)$, la ecuación del plano tangente es

$$\frac{1}{\alpha}(x-\alpha)+\frac{1}{\beta}(y-\beta)-\frac{1}{\gamma}(z-\gamma)=0.$$

Recta normal:

$$\begin{cases} x=\alpha+\dfrac{t}{\alpha} \\ y=\beta+\dfrac{t}{\beta} \\ z=-\gamma-\dfrac{t}{\gamma} \end{cases} \quad (t\in\mathbb{R}).$$

2. (Paraboloide hiperbólico) Es la gráfica de la función de clase $\mathscr{C}^\infty$ definida por

$$\varphi(x,y)=x^2-y^2,\ \forall (x,y)\in\mathbb{R}^2$$

Así M es una variedad en $\mathbb{R}^3$ de dimensión 2.

Plano tangente: Como $\frac{\partial\varphi}{\partial x}(0,0)=\frac{\partial\varphi}{\partial y}(0,0)=0$, la ecuación del plano tangente es $z=0$.

Recta normal:

$$\begin{cases} x=0 \\ y=0 \\ z=t \end{cases} \quad (t\in\mathbb{R}).$$

3. (Hélice) La función

$$p(t)=(\cos t,\operatorname{sen} t,t),\ \forall t\in\mathbb{R},$$

es de clase $\mathscr{C}^\infty$ con $p'(t)=(-\operatorname{sen} t,\cos t,1)$, y en consecuencia

$$\operatorname{rango}\big(p'(t)\big)=1,\ \forall t\in\mathbb{R}.$$

Así M es una variedad en $\mathbb{R}^3$ de dimensión 1.

Recta tangente: Como $p'(\pi)=(0,-1,1)$, la ecuación de la recta tangente es

$$\begin{cases} x= & -1 \\ y= & -t \\ z= & \pi+t \end{cases} \quad (t\in\mathbb{R}).$$

Plano normal:

$$-y+(z-\pi)=0.$$

4. (Borde de la bóveda de Viviani) La función

$$g : G = \mathbb{R} \times \mathbb{R} \times \mathbb{R}^{+} \setminus \left\{ \left(\frac{1}{2}, 0, z\right) : z \in \mathbb{R} \right\} \longrightarrow \mathbb{R}^2$$

dada por

$$g(x,y,z) = (x^2 + y^2 + z^2 - 1, x^2 - x + y^2)$$

es de clase $\mathscr{C}^\infty$ con $J_g(x,y,z) = \begin{pmatrix} 2x & 2y & 2z \\ 2x-1 & 2y & 0 \end{pmatrix}$, y en consecuencia

$$\text{rango } \big(J_g(x,y,z)\big) = 2, \ \forall (x,y,z) \in G$$

(se han suprimido los puntos de la recta perpendicular al plano OXY que pasa por el cetro de la circunferencia $\left(x - \frac{1}{2}\right)^2 + y^2 = \frac{1}{4}$)
Así M es una variedad en $\mathbb{R}^3$ de dimensión 1.

Recta tangente: Como $J_g(0,0,1) = \begin{pmatrix} 0 & 0 & 2 \\ -1 & 0 & 0 \end{pmatrix}$, un vector de dirección de la recta tangente es por ejemplo $(0,1,0)$, y en consecuencia la ecuación de la recta tangente es

$$\begin{cases} x = & 0 \\ y = & t \\ z = & 1 \end{cases} \quad (t \in \mathbb{R}).$$

Plano normal: $y = 0$.

5. (Toro) La función

$$g : G = \mathbb{R}^3 \setminus \{(0,0,z) : z \in \mathbb{R}\} \cup \{(x,y,0) : x^2 + y^2 = 4\} \longrightarrow \mathbb{R}$$

dada por

$$g(x,y,z) = \left(\sqrt{x^2+y^2} - 2\right)^2 + z^2 - 1$$

es de clase $\mathscr{C}^\infty$ con

$$J_g(x,y,z) = \left(\frac{2(\sqrt{x^2+y^2}-2)x}{\sqrt{x^2+y^2}}, \ \frac{2(\sqrt{x^2+y^2}-2)y}{\sqrt{x^2+y^2}}, \ 2z \right),$$

y en consecuencia

$$\text{rango } \big(J_g(x,y,z)\big) = 1, \ \forall (x,y,z) \in G$$

(hay que quitar los puntos del eje OZ y la circunferencia eje interior del toro, esto es, el conjunto de centros de las circunferencias que describen el toro).

Así M es una variedad en $\mathbb{R}^3$ de dimensión 2.

Plano tangente: Como $J_g(3,0,0) = (2,0,0)$, la ecuación del plano tangente es $x - 3 = 0$.

Recta normal:

$$\begin{cases} x = & 3+t \\ y = & 0 \\ z = & 0 \end{cases} \quad (t \in \mathbb{R}).$$

9.5 $i) \Rightarrow ii)$ Si $a = 0$ no hay nada que probar, en otro caso aplicando $i)$ al vector $\frac{a}{\|a\|_2}$, se tiene que

$$\frac{a_1 \dots a_N}{\big(\|a\|_2\big)^N} \leq \big(\frac{1}{N}\big)^{\frac{N}{2}},$$

o lo que es lo mismo

$$\sqrt[N]{a_1^2 \cdots a_N^2} \leq \frac{a_1^2 + \cdots + a_N^2}{N}$$

de donde se deduce ii) sin más que observar que tan arbitrarios son los a_i como los a_i^2.

$ii) \Rightarrow i)$ Se tiene que

$$\sqrt[N]{a_1^2 \cdots a_N^2} \leq \frac{a_1^2 + \cdots + a_n^2}{N} \ ,$$

lo que en el caso de que $\|a\|_2 = 1$ equivale a

$$a_1 \cdots a_N \leq \Big(\frac{1}{N}\Big)^{\frac{N}{2}}.$$

Sabemos que en ambos casos se da la igualdad si, y sólo si, $a_1 = \cdots = a_N$.

9.6 *i)* La condición viene dada por la función $g : \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+ \longrightarrow \mathbb{R}$ dada por

$$g(x,y,z) = xy + xz + yz - 3.$$

La variedad M en $\mathbb{R}^3$ de dimensión 2 que define

$$M = \{(x,y,z) \in \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+ : xy + xz + yz = 3\}$$

no es compacta. La función a minimizar sobre M es $f : \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+ \longrightarrow \mathbb{R}$ definida por

$$f(x,y,z) = xyz.$$

Función de Lagrange: $L : \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R} \longrightarrow \mathbb{R}$ definida por

$$L(x,y,z,\lambda) = xyz + \lambda(xy + xz + yz - 3)$$

Sistema de Lagrange:

$$\begin{cases} yz + \lambda(y+z) = 0 \\ xz + \lambda(x+z) = 0 \\ xy + \lambda(x+y) = 0 \\ xy + xz + yz = 3 \\ x,y,z > 0 \end{cases}$$

Solución:

$$a=(1,1,1), \quad \alpha=\frac{-1}{2}.$$

Haciendo un estudio paralelo al realizado en el Ejemplo 9.15 obtendríamos

$$H=\begin{pmatrix}1 & -1 & 0\\ 1 & 0 & -1\end{pmatrix}\begin{pmatrix}0 & \frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & 0 & \frac{1}{2}\\ \frac{1}{2} & \frac{1}{2} & 0\end{pmatrix}\begin{pmatrix}1 & 1\\ -1 & 0\\ 0 & -1\end{pmatrix}=\begin{pmatrix}-1 & -\frac{1}{2}\\ -\frac{1}{2} & -1\end{pmatrix}.$$

y concluiríamos que f alcanza en $a=(1,1,1)$ un máximo estricto condicionado. La justificación topológica de que es máximo absoluto ahora no es difícil considerando el compacto

$$K=\{(x,y,z)\in\mathbb{R}_0^+\times\mathbb{R}_0^+\times\mathbb{R}_0^+ : xy+xz+yz=3\}.$$

Utilizando la desigualdad de las medias tenemos:

$$\sqrt[3]{x^2y^2z^2}\leq\frac{xy+xz+yz}{3},$$

además se da la igualdad si, y sólo si, $xy=xz=yz$, esto es, $x=y=z$.

Al haber un sólo punto crítico, concluimos que no existe un ortoedro de volumen mínimo con dicha superficie lateral. Una prueba alternativa de este hecho es la siguiente: Dado un natural n, tomando los valores

$$x=\frac{1}{n}, y=1, z=\frac{3n-1}{n+1} \;\Rightarrow\; xyz=\frac{3n-1}{n(n+1)}.$$

ii) No existe un ortoedro de superficie lateral máximo con volumen una unidad, pues hay un sólo punto crítico. Una prueba alternativa es la siguiente: dado un natural n, si tomamos

$$x=n, y=\frac{1}{n}, z=1 \;\Rightarrow\; xy+xz+yz\to+\infty.$$

9.7 Se puede hacer un estudio utilizando la función de Lagrange

$$L:\mathbb{R}^+\times\mathbb{R}^+\times\mathbb{R}^+\times\mathbb{R}\longrightarrow\mathbb{R}$$

definida por

$$L(x,y,z,\lambda)=8xyz+\lambda\left(\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}-1\right),$$

pero es mejor proceder usando la desigualdad de las medias:

$$\sqrt[3]{\frac{x^2y^2z^2}{a^2b^2c^2}}\leq\frac{\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}}{3}=\frac{1}{3} \;\Leftrightarrow\; 8xyz\leq\frac{8\sqrt{3}}{9}abc,$$

y se da la igualdad si, y sólo si,

$$\frac{x^2}{a^2} = \frac{y^2}{b^2} = \frac{z^2}{c^2},$$

y como se verifica

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1,$$

concluimos que la *solución* es

$$x = \frac{a\sqrt{3}}{3}, \quad y = \frac{b\sqrt{3}}{3}, \quad z = \frac{c\sqrt{3}}{3},$$

que corresponde al valor del parámetro $\alpha = \dfrac{-4\sqrt{3}abc}{9}$.

9.8 Sea

$$\pi \equiv Ax + By + Cz = D \text{ con } A^2 + B^2 + C^2 > 0$$

y el punto (a,b,c). Queremos hacer mínimo

$$d = \sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2}, \quad ((x,y,z) \in \pi),$$

ó lo que es igual, hacer mínimo d^2.

Función de Lagrange: $L : \mathbb{R}^4 \longrightarrow \mathbb{R}$ dada por

$$L(x,y,z,\lambda) = (x-a)^2 + (y-b)^2 + (z-c)^2 + \lambda(Ax + By + Cz - D).$$

Sistema de Lagrange:

$$\left\{ \begin{array}{l} 2(x-a) + \lambda A = 0 \\ 2(y-b) + \lambda B = 0 \\ 2(z-c) + \lambda C = 0 \\ Ax + By + Cz = D \end{array} \right. \Rightarrow \; 2\Big(A(x-a) + B(y-b) + C(z-c)\Big) = -\lambda(A^2 + B^2 + C^2),$$

es decir,

$$2\Big(D - (Aa + Bb + Cc)\Big) = -\lambda(A^2 + B^2 + C^2),$$

equivalentemente

$$\lambda = \frac{2(Aa + Bb + Cc - D)}{A^2 + B^2 + C^2}.$$

Solución:

$$(x,y,z), \quad \alpha = \frac{2(Aa + Bb + Cc - D)}{A^2 + B^2 + C^2},$$

con lo que

$$d^2 = (x-a)^2 + (y-b)^2 + (z-c)^2 = \frac{\alpha^2\Big(A^2 + B^2 + C^2\Big)}{4} =$$

$$= \frac{(Aa+Bb+Cc-D)^2}{A^2+B^2+C^2},$$

es decir,

$$d = \frac{|Aa+Bb+Cc-D|}{\sqrt{A^2+B^2+C^2}}.$$

La justificación de que es mínimo absoluto es fácil (ver Ejercicio 2.10 que utiliza la propiedad de Bolzano-Weierstrass.)

9.9 Los conjuntos K son compactos y las funciones son continuas luego alcanzan el máximo y el mínimo absolutos. Es claro que los extremos absolutos son también extremos condicionados (relativos en el caso de abiertos). Si el punto donde f alcanza un extremo absoluto pertenece al interior de K, entonces f alcanza en él un extremo relativo y en consecuencia las derivadas parciales de f en dicho punto son nulas. En otro caso, el punto pertenece a la frontera de K y se estudia como extremos condicionados (la frontera de K ó es una variedad ó es unión finita de variedades disjuntas).

$\boxed{\text{f}}$

$$f(x) = 2x^2 - 3y^2 - 2x. \quad K := \left\{(x,y) \in \mathbb{R}^2 : x^2+y^2 \leq 5\right\}.$$

En este caso K es un círculo de centro el origen y radio $\sqrt{5}$.

Estudio en el interior:

$$\left\{ \begin{array}{c} \frac{\partial f}{\partial x}(x,y) = 4x-2 \\ \frac{\partial f}{\partial y}(x,y) = -6y \end{array} \right\}.$$

El único punto crítico es $a = \left(\frac{1}{2}, 0\right)$.

Estudio en la frontera:

Función de Lagrange: $L : (\mathbb{R}^2 \backslash \{(0,0)\}) \times \mathbb{R} \longrightarrow \mathbb{R}$ definida por

$$L(x,y,\lambda) = 2x^2 - 3y^2 - 2x + \lambda(x^2+y^2-5).$$

Sistema de Lagrange:

$$\left\{ \begin{array}{c} (2+\lambda)x = 1 \\ (\lambda-3)y = 0 \\ x^2+y^2 = 5 \end{array} \right\} \Rightarrow \left\{ \begin{array}{c} y = 0 \Rightarrow x = \pm\sqrt{5}, \lambda = \frac{\pm 1}{\sqrt{5}} - 2 \\ \lambda = 3, x = \frac{1}{5}, y = \pm\frac{\sqrt{124}}{5} \end{array} \right\}$$

Solución:

$$b = \left(\sqrt{5}, 0\right), \alpha = \frac{1}{\sqrt{5}} - 2; \quad c = \left(-\sqrt{5}, 0\right), \alpha = -\frac{1}{\sqrt{5}} - 2;$$

$$d = \left(\frac{1}{5}, \frac{\sqrt{124}}{5}\right), \beta = 3; \quad e = \left(\frac{1}{5}, -\frac{\sqrt{124}}{5}\right), \beta = 3.$$

En consecuencia, f alcanza el máximo en c y vale $10+2\sqrt{5}$, mientras que el mínimo vale $-\frac{380}{25}=-15{,}2$ y se alcanza en d y e.

$\boxed{\text{g}}$

$$g(x,y)=10xy-x^2y-xy^2, \quad K:=\Big\{(x,y)\in\mathbb{R}^2: 0\le x, 0\le y, x+y\le 10\Big\}$$

K es el triángulo de vértices $(0,0),(0,10),(10,0)$.

Estudio en el interior:
$$\left\{\begin{array}{l} \frac{\partial g}{\partial x}(x,y)=10y-2xy-y^2 \\ \frac{\partial g}{\partial y}(x,y)=10x-x^2-2xy \end{array}\right\}.$$

El único punto crítico es $a=\left(\dfrac{10}{3},\dfrac{10}{3}\right)$.

Estudio en la frontera:
La frontera es unión disjunta de 4 variedades globalmente determinadas:

$$M_1=\{(0,0),(0,10),(10,0)\} \quad \text{(de dimensión cero)},$$

$$M_2=\{(x,y)\in\mathbb{R}^2: x=0, 0<y<10\},$$
$$M_3=\{(x,y)\in\mathbb{R}^2: y=0, 0<x<10\},$$
$$M_4=\{(x,y)\in\mathbb{R}^2: 0<x, 0<y,\ x+y=10\},$$

de dimensión uno.

La función $g(x,y)=(10-x-y)xy$ se anula en todos los puntos de la frontera y en el interior de K toma valores positivos, luego alcanza el mínimo en la frontera y éste vale cero y el máximo lo alcanza en el punto crítico del interior y vale $\left(\dfrac{10}{3}\right)^3$.

$\boxed{\text{h}}$

$$h(x,y)=x^2+y^2-xy-x-y, \quad K:=\Big\{(x,y)\in\mathbb{R}^2: 0\le x, 0\le y, x+y\le 3\Big\}$$

K es el triángulo de vértices $(0,0),(0,3),(3,0)$.

Estudio en el interior:
$$\left\{\begin{array}{l} \frac{\partial h}{\partial x}(x,y)=2x-y-1 \\ \frac{\partial h}{\partial y}(x,y)=2y-x-1 \end{array}\right\}.$$

El único punto crítico es $a=(1,1)$.

Estudio en la frontera:
La frontera es unión disjunta de 4 variedades globalmente determinadas:

$$M_1=\{(0,0),(0,3),(3,0)\} \quad \text{(de dimensión cero)},$$

$$M_2=\{(x,y)\in\mathbb{R}^2: x=0, 0<y<3\},$$

$$M_3 = \{(x,y) \in \mathbb{R}^2 : y = 0, 0 < x < 3\},$$

$$M_4 = \{(x,y) \in \mathbb{R}^2 : 0 < x, 0 < y,\ x + y = 3\},$$

de dimensión uno.

En las variedades M_2 y M_3 los puntos críticos son

$$c = \left(0, \frac{1}{2}\right), \quad \beta = \frac{-1}{2} \quad \text{y} \qquad d = \left(\frac{1}{2}, 0\right), \quad \gamma = \frac{-1}{2}.$$

Estudiamos los extremos condicionados de h en M_4.

Función de Lagrange: $L : \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R} \longrightarrow \mathbb{R}$ definida por

$$L(x,y,\lambda) = x^2 + y^2 - xy - x - y + \lambda(x + y - 3).$$

Sistema de Lagrange:

$$\left\{ \begin{array}{c} 2x - y - 1 + \lambda = 0 \\ 2y - x - 1 + \lambda = 0 \\ x + y = 3 \end{array} \right.$$

Solución:

$$b = \left(\frac{3}{2}, \frac{3}{2}\right), \quad \alpha = \frac{-1}{2}.$$

Evaluando se obtiene que h alcanza el máximo en $(0,3)$ y $(3,0)$ y éste vale 6; el mínimo vale -1 y se alcanza en $a = (1,1)$.

$\boxed{\text{u}}$

$$u(x,y) = x^2 - 2xy + y^2, \quad K := \left\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 2x\right\}.$$

K es el círculo de centro $(1,0)$ y radio 1. La propiedad de compacidad nos asegura que f alcanza el máximo y el mínimo absolutos en K. Si el punto donde f alcanza un extremo absoluto pertenece al interior de K, esto es, al conjunto

$$\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 < 2x\},$$

entonces f alcanza en él un extremo relativo y en consecuencia las derivadas parciales de f en dicho punto son nulas, esto es, es solución del sistema

$$\nabla f(x,y) = 0 \ \Leftrightarrow \ \left\{ \begin{array}{c} D_1 f(x,y) = 0 \\ D_2 f(x,y) = 0 \end{array} \right\} \Leftrightarrow x = y.$$

Puntos críticos: $\{(x,x) : 0 < x < 1\}$.

Estudio en la frontera:

$$M = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 - 2x = 0\}.$$

Estudiamos ahora los extremos condicionados por M. Obsérvese que M es la variedad en $\mathbb{R}^2$ de dimensión uno determinada globalmente por la función de clase $\mathscr{C}^\infty$

$$g(x,y) = x^2 + y^2 - 2x, \quad \forall (x,y) \in G = \mathbb{R}^2 \backslash \{(1,0)\},$$

pues

$$\text{rango}\,\big(J_g(x,y)\big) = 1, \quad \forall (x,y) \in G.$$

Función de Lagrange: $L : G \times \mathbb{R} \longrightarrow \mathbb{R}$ definida por

$$L(x,y,\lambda) = x^2 - 2xy + y^2 + \lambda (x^2 + y^2 - 2x).$$

Sistema de Lagrange:

$$\left\{ \begin{array}{c} x - y = \lambda (1 - x) \\ x - y = \lambda y \\ x^2 + y^2 = 2x \end{array} \right\} \Rightarrow \lambda (1 - x - y) = 0 \Rightarrow$$

$$\Rightarrow \left\{ \begin{array}{c} b = (0,0), \alpha = 0; \quad c = (1,1), \alpha = 0 \\ d = \left(\frac{2+\sqrt{2}}{2}, \frac{-\sqrt{2}}{2}\right), \beta = -2 - \sqrt{2}; \quad e = \left(\frac{2-\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right), \gamma = -2 + \sqrt{2} \\ y - x = \lambda y \end{array} \right.$$

Finalmente, evaluando u se obtiene que

$$u(x,x) = 0, \quad u(d) = 3 + 2\sqrt{2}, \quad u(e) = 3 - 2\sqrt{2}.$$

En consecuencia, el máximo se alcanza en d y el mínimo vale cero y se alcanza en los puntos del conjunto $\{(x,x) : 0 \leq x \leq < 1\}$.

$\boxed{\text{v}}$

$$v(x,y) = y - x, \quad K := \Big\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 - 2x = 0\Big\}.$$

La función

$$g(x,y) = x^2 + y^2 - 2x, \quad \forall (x,y) \in G = \mathbb{R}^2 \backslash \{(1,0)\}$$

es de clase $\mathscr{C}^\infty$ y verifica

$$\text{rango}\,\big(J_g(x,y)\big) = 1, \quad \forall (x,y) \in G$$

por lo que M es la variedad de dimensión uno en $\mathbb{R}^2$ globalmente determinada por g. De hecho M es la circunferencia $C((1,0),1)$, luego es compacta, lo que asegura la existencia de extremos absolutos.

En este caso, el interior del conjunto es vacío.

Estudio en la frontera:

Función de Lagrange: $L : G \times \mathbb{R} \longrightarrow \mathbb{R}$ definida por

$$L(x,y,\lambda) = y - x + \lambda(x^2 + y^2 - 2x).$$

Sistema de Lagrange:

$$\left\{\begin{array}{c} -1 + 2\lambda x - 2\lambda = 0 \\ 1 + 2\lambda y = 0 \\ x^2 + y^2 - 2x = 0 \end{array}\right. \Leftrightarrow \left\{\begin{array}{c} 1 = 2\lambda(x-1) \\ -1 = 2\lambda y \\ x^2 + y^2 = 2x \end{array}\right. \Leftrightarrow (\lambda \neq 0) \left\{\begin{array}{c} y = 1 - x \\ x^2 + y^2 = 2x \end{array}\right. \Leftrightarrow 2x^2 - 4x + 1 = 0.$$

Solución:

$$a = \Big(\frac{2+\sqrt{2}}{2}, \frac{-\sqrt{2}}{2}\Big), \alpha = \frac{\sqrt{2}}{2}; \quad b = \Big(\frac{2-\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\Big), \beta = \frac{-\sqrt{2}}{2}$$

Concluimos que v alcanza el máximo en b y en a el mínimo, y estos valen $-1 + \sqrt{2}$ y $-1 - \sqrt{2}$, respectivamente.

9.10

$$f(x,y,z) = xyz, \quad \forall (x,y,z) \in M = \{(x,y,z) \in \mathbb{R}^3 : x + y + z = 1\}.$$

Función de Lagrange: $L : \mathbb{R}^4 \longrightarrow \mathbb{R}$ definida por

$$L(x,y,z,\lambda) = xyz + \lambda(x + y + z - 1).$$

Sistema de Lagrange:

$$\left\{\begin{array}{c} yz + \lambda = 0 \\ xz + \lambda = 0 \\ xy + \lambda = 0 \\ x + y + z = 1 \end{array}\right. \Rightarrow yz = xz = xy$$

Soluciones:

$$a = \Big(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\Big), \quad \alpha = \frac{-1}{9}; \qquad b = (1,0,0), \;\; \beta = 0,$$
$$c = (0,1,0), \;\; \beta = 0; \qquad d = (0,0,1), \;\; \beta = 0$$

$$H_{L_\alpha}(x,y,z) = H_{L_\beta}(x,y,z) = \begin{pmatrix} 0 & z & y \\ z & 0 & x \\ y & x & 0 \end{pmatrix}$$

$$H_a = \frac{1}{3}\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \quad H_b = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad H_c = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad H_d = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

$$J_g = (1,1,1) \;\Rightarrow\; K(x,y,z) = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$$

f alcanza en a un máximo condicionado estricto.

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix}$$

f no alcanza en b un extremo condicionado.

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$$

f no alcanza en c un extremo condicionado.

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ 1 & 0 \end{pmatrix}$$

f no alcanza en d un extremo condicionado.

Una prueba alternativa de que en b no se alcanza un extremo condicionado es la siguiente: para cada natural n se tiene

$$f\left(1, \frac{1}{n}, \frac{-1}{n}\right) < 0 < f\left(1, \frac{1}{n}, \frac{1}{n}\right).$$

El mismo argumento se puede usar también para los puntos c y d.

Capítulo III

Integración

Tema 10

Medida de Lebesgue en $\mathbb{R}^N$.

Desde la antigüedad el hombre ha tenido que enfrentarse al problema de medir longitudes, áreas y volúmenes; si bien de hecho la historia está llena de acontecimientos que así lo avalan, hay que constatar que la formalización de la teoría de la medida es reciente.

La medida de Lebesgue en $\mathbb{R}^N$ es una extensión de la noción elemental de volumen de un intervalo acotado N-dimensional a una muy amplia clase de subconjuntos de $\mathbb{R}^N$. Puesto que la medida de Lebesgue tiene diferente significado en espacios de distinta dimensión, es usual fijar la dimensión $N \in \mathbb{N}$, lo que haremos desde este momento, y en consecuencia no será necesario indicarla explícitamente en la notación.

El procedimiento que seguiremos para construir la medida de Lebesgue en $\mathbb{R}^N$ consistirá en definir directamente la medida exterior de Lebesgue a partir del volumen de un intervalo acotado y restringir dicha medida exterior a la amplia clase de los conjuntos medibles, clase que describiremos utilizando la topología (Teorema 10.15), aunque se probará también que no hay que disponer de la topología para conocer los conjuntos de dicha clase (Teorema 10.16).

En términos generales el problema de medir en $\mathbb{R}^N$ consiste en asignar a cada subconjunto A de $\mathbb{R}^N$ su medida $\mu(A)$, con la que se pretende cuantificar su tamaño. Naturalmente esta asignación ha de poseer ciertas cualidades que nos dicta la razón, a saber

$$\mu(A) \geq 0 \ \text{ y } \ \mu(A_1 \cup \ldots \cup A_n) = \mu(A_1) + \ldots + \mu(A_n)$$

para cualesquiera conjuntos $A_1, \ldots, A_n$ disjuntos entre sí. No obstante, el desarrollo exitoso de la teoría exige que la anterior condición de aditividad sea válida para cualquier sucesión $\{A_n\}$ de conjuntos disjuntos entre sí, esto es

$$\mu\left(\cup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

Es también natural requerir que la medida asignada a los intervalos N-dimensionales venga dada por la fórmula tradicional. No todos los subconjuntos de $\mathbb{R}^N$ se pueden someter a estas reglas. Aquellos que sí lo hacen constituyen una familia, los conjuntos medibles, que goza de las propiedades que seguidamente codificamos. Conviene resaltar que el trabajo con medidas involucra inexorablemente el conjunto $[0, \infty]$ (véase el Apéndice A en el que se recogen las nociones usadas habitualmente en dicho conjunto).

10.1. σ-álgebras y medidas.

Definición 10.1 (σ-álgebra y medida)**.** Una *σ-álgebra* en un conjunto no vacío Ω es una familia $\mathscr{A} \subset \mathscr{P}(\Omega)$ que verifica las siguientes propiedades:

i) $\Omega \in \mathscr{A}$.

ii) $A \in \mathscr{A} \Rightarrow A^C := \Omega \backslash A \in \mathscr{A}$.

iii) $[A_n \in \mathscr{A}, \forall n \in \mathbb{N}] \Rightarrow \cup_{n=1}^{\infty} A_n \in \mathscr{A}$.

Los elementos de $\mathscr{A}$ se llaman *conjuntos medibles* y el par $(\Omega, \mathscr{A})$ se llama *espacio medible*.

Dado un espacio medible $(\Omega, \mathscr{A})$, una *medida* sobre $\mathscr{A}$ es una aplicación $\mu : \mathscr{A} \to [0, \infty]$ *σ-aditiva*, es decir:

$$\left[A_n \in \mathscr{A}, \forall n \in \mathbb{N}, A_i \cap A_j = \emptyset \ (i \neq j)\right] \Rightarrow \mu\left(\cup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

Es natural que $\mu(\emptyset) = 0$ lo que ocurre, como veremos, si existe $A \in \mathscr{A}$ tal que $\mu(A) < \infty$. Por ello, se supondrá siempre que existe un conjunto de medida finita.

La terna $(\Omega, \mathscr{A}, \mu)$ se llama un *espacio de medida*.

Se dice que un espacio de medida $(\Omega, \mathscr{A}, \mu)$ es *completo* si cualquier subconjunto de un conjunto de medida cero es medible, esto es

$$[A \in \mathscr{A}, \mu(A) = 0, B \subset A] \Rightarrow B \in \mathscr{A},$$

de lo cual se deducirá enseguida que, en tal caso, $\mu(B) = 0$.

Proposición 10.2 (Propiedades de las σ-álgebras y de las medidas)**.** *Sea $(\Omega, \mathscr{A})$ un espacio medible.*

1) Se verifican las siguientes afirmaciones:

a) $\emptyset \in \mathscr{A}$.

b) $A, B \in \mathscr{A} \Rightarrow A \cup B \in \mathscr{A}$.

c) $[A_n \in \mathscr{A}, \forall n \in \mathbb{N}] \Rightarrow \cap_{n=1}^{\infty} A_n \in \mathscr{A}$.

d) $A, B \in \mathscr{A} \Rightarrow A \cap B \in \mathscr{A}$.

e) $A, B \in \mathscr{A} \Rightarrow A \backslash B \in \mathscr{A}$.

2) Si μ es una medida sobre $\mathscr{A}$, entonces:

i) $\mu(\emptyset)=0$.

ii) $[A,B\in\mathscr{A}, A\cap B=\emptyset]\Rightarrow \mu(A\cup B)=\mu(A)+\mu(B)$ *(aditividad).*

iii) $[A,B\in\mathscr{A}, A\subset B]\Rightarrow \mu(A)\leq\mu(B)$ *(monotonía).*

iv) $[A_n\in\mathscr{A}, A_n\subset A_{n+1}, \forall n\in\mathbb{N}]\Rightarrow \mu(\cup_{n=1}^{\infty}A_n)=\lim\mu(A_n)$*(crecimiento continuo).*

v) $[A_n\in\mathscr{A}, A_n\supset A_{n+1}, \forall n\in\mathbb{N}, \mu(A_1)<\infty]\Rightarrow \mu(\cap_{n=1}^{\infty}A_n)=\lim\mu(A_n)$ *(decrecimiento continuo).*

vi) $[A,B\in\mathscr{A}]\Rightarrow \mu(A\cup B)\leq\mu(A)+\mu(B)$ *(subaditividad).*

vii) $[A_n\in\mathscr{A}, \forall n\in\mathbb{N}]\Rightarrow \mu(\cup_{n=1}^{\infty}A_n)\leq\sum_{n=1}^{\infty}\mu(A_n)$ *(σ-subaditividad).*

Demostración:

$\boxed{1}$

a) Es consecuencia de i) y ii) de la definición ya que $\Omega^C=\emptyset$.
b) Basta hacer en iii) de la definición $A=A_1, B=A_2=A_3=...$
c) Es consecuencia de ii) y iii) de la definición ya que $\cap_{n=1}^{\infty}A_n=\left(\cup_{n=1}^{\infty}A_n^C\right)^C$.
d) Se prueba como b) pero utilizando c).
e) Es consecuencia de la igualdad $A\backslash B=A\cap B^C$ usando d) y el axioma ii) de la definición.

$\boxed{2}$

i) Tomemos $A\in\mathscr{A}$ tal que $\mu(A)<\infty$, entonces

$$\mu(A)=\mu(A\cup\emptyset\cup\emptyset\ldots)=\mu(A)+\mu(\emptyset)+\mu(\emptyset)+\ldots,$$

luego $\mu(\emptyset)=0$.
ii) $\mu(A\cup B)=\mu(A\cup B\cup\emptyset\cup\emptyset\cup\ldots)=\mu(A)+\mu(B)+0+0+\ldots=\mu(A)+\mu(B)$.
iii) $\mu(A)\leq\mu(A)+\mu(B\backslash A)=\mu(A\cup(B\backslash A))=\mu(B)$.
iv) Pongamos $B_1=A_1$ y $B_{n+1}=A_{n+1}\backslash A_n$, $\forall n\in\mathbb{N}$. Se tiene que:

$$A_n=\cup_{k=1}^{n}B_k, \forall n\in\mathbb{N} \quad \text{y, en consecuencia,} \quad \cup_{n=1}^{\infty}A_n=\cup_{k=1}^{\infty}B_k$$

con lo que

$$\mu(\cup_{n=1}^{\infty}A_n)=\mu(\cup_{k=1}^{\infty}B_k)=\sum_{k=1}^{\infty}\mu(B_k)=$$

$$\lim\left(\mu(B_1)+\ldots+\mu(B_n)\right)=\lim\mu\left(\cup_{k=1}^{n}B_k\right)=\lim\mu(A_n),$$

donde se ha utilizado el axioma de σ-aditividad, el concepto de serie y la propiedad de aditividad finita.
v) La propiedad de aditividad finita nos asegura para $n\in\mathbb{N}$ que

$$\mu(A_1)=\mu(A_1\backslash A_n)+\mu(A_n),$$

y también que

$$\mu(A_1)=\mu\left(\cap_{n=1}^{\infty}A_n\right)+\mu\left(A_1\backslash\cap_{n=1}^{\infty}A_n\right),$$

con lo que al ser $\mu(A_1)<\infty$, utilizando además iv), se tiene

$$\mu(A_1)-\mu\left(\cap_{n=1}^{\infty}A_n\right)=\mu\left(A_1\backslash\cap_{n=1}^{\infty}A_n\right)=\mu\left(\cup_{n=1}^{\infty}(A_1\backslash A_n)\right)=$$

$$\lim \mu(A_1 \backslash A_n) = \lim(\mu(A_1) - \mu(A_n)) = \mu(A_1) - \lim \mu(A_n)$$

de donde obviamente se deduce la propiedad v).

vi) $\mu(A \cup B) = \mu(A \cup (B \backslash A)) = \mu(A) + \mu(B \backslash A) \leq \mu(A) + \mu(B)$.

vii) Pongamos $B_n = \cup_{k=1}^n A_k, \forall n \in \mathbb{N}$. Se tiene que

$$\mu\left(\cup_{n=1}^\infty A_n\right) = \mu\left(\cup_{n=1}^\infty B_n\right) = \lim \mu(B_n) \leq \lim \sum_{k=1}^n \mu(A_k) = \sum_{n=1}^\infty \mu(A_n),$$

donde se han utilizado *iv*), *vi*) y el concepto de serie. ∎

Nota 10.3. Obsérvese que ahora la σ-subaditividad asegura que, en cualquier espacio de medida, *la unión numerable de conjuntos de medida cero es de medida cero.*

Definición 10.4 (Propiedad c.p.d.)**.** Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida. Se dice que una propiedad $P(\omega)$ relativa a un punto genérico $\omega \in \Omega$ se verifica *casi por doquier* con respecto a μ (μ-c.p.d.), o bien, para casi todo punto de Ω con respecto a μ (μ-ct $\omega \in \Omega$), si el conjunto de puntos donde dicha propiedad no se verifica es un conjunto de medida cero. Se omite la referencia a la medida μ si no hay lugar a confusión.

Por ejemplo, la función parte entera $E : \mathbb{R} \to \mathbb{R}$ es continua c.p.d. ya que el conjunto $\mathbb{Z}$ de puntos de discontinuidad tiene medida cero (según la Nota 10.3).

Si $(\Omega, \mathscr{A}, \mu)$ es un espacio de medida y $\{P_n\}$ es una sucesión de propiedades relativas a un punto genérico $\omega \in \Omega$ tal que cada una de las P_n se verifica c.p.d., entonces se verifican todas simultáneamente c.p.d., es decir, si se considera la propiedad P determinada por: "P se verifica en ω si, y sólo si, todas las P_n se verifican en ω", entonces P se verifica c.p.d. En efecto, si para cada n natural $E_n \subset \Omega$ es el conjunto de medida cero en el que no se verifica P_n, entonces el conjunto

$$E := \{\omega \in \Omega : P \text{ no se verifica en } \omega\}$$

es de medida cero, pues $E = \cup_{n \in \mathbb{N}} E_n$.

Ejemplos 10.5 (σ-álgebras y medidas)**.**

a) La σ-álgebra generada por una familia de subconjuntos de Ω.

La familia $\mathscr{P}(\Omega)$ constituida por todos los subconjuntos de Ω es una σ-álgebra en Ω. Si $\{\mathscr{A}_i : i \in I\}$ es una familia de σ-álgebras en Ω, entonces $\mathscr{A} = \cap_{i\in I}\mathscr{A}_i$ es una σ-álgebra en Ω. En consecuencia si $\mathscr{D}$ es una familia de subconjuntos de Ω, entonces la intersección de todas las σ-álgebras en Ω que contienen a $\mathscr{D}$ (entre las cuales hay al menos una: $\mathscr{P}(\Omega)$) es la menor σ-álgebra en Ω que contiene a $\mathscr{D}$ y recibe el nombre de *σ-álgebra generada por $\mathscr{D}$*.

b) La σ-álgebra de Borel y la medida de Borel-Lebesgue.

Todo espacio topológico $(X,\mathfrak{I})$ se considerará un espacio medible con la σ-álgebra generada por la familia $\mathfrak{I}$ de los abiertos de X. A esta σ-álgebra se la llama σ-álgebra de Borel.

La σ-álgebra de Borel en $\mathbb{R}^N$ es la σ-álgebra generada por la familia $\mathfrak{I}$ de los abiertos de $\mathbb{R}^N$, se nota por $\mathscr{B}$ y sus elementos se denominan conjuntos *borelianos*.

Es claro que los conjuntos cerrados son borelianos, así como que $\mathscr{B}$ está generada también por la familia $\mathscr{F}$ de los cerrados de $\mathbb{R}^N$. De hecho $\mathscr{B}$ contiene las siguientes familias de subconjuntos topológicamente relevantes:

i) Conjuntos del tipo G_δ que son aquellos conjuntos que se pueden expresar como intersección numerable (y decreciente, si se quiere) de abiertos.

ii) Conjuntos del tipo F_σ que son aquellos conjuntos que se pueden expresar como unión numerable de cerrados. Es inmediato comprobar que un conjunto B del tipo F_σ se puede expresar como unión numerable y creciente de conjuntos compactos. En efecto si $B = \cup_{n=1}^{\infty} F_n$, es una expresión de B como unión de cerrados, también B se puede expresar de la forma $B = \cup_{n=1}^{\infty} K_n$, donde para cada n

$$K_n = \overline{B}(0,n) \cap (F_1 \cup \ldots \cup F_n).$$

Se verá más adelante que $\mathscr{B}$ es la σ-álgebra generada por los intervalos acotados de $\mathbb{R}^N$ (ver Corolario 10.9). La relevancia de esta σ-álgebra radica en el hecho de que *existe una única medida sobre $\mathscr{B}$ que extiende el volumen de los intervalos acotados* (ver Teorema 10.15). Esta medida recibe el nombre de *medida de Borel-Lebesgue*, se nota por λ y se puede describir de la siguiente manera:

$$\lambda(B) = \inf\left\{ \sum_{n=}^{\infty} v(I_n) : B \subset \cup_{n\in\mathbb{N}} I_n, I_n \text{ intervalos acotados} \right\}, \ \forall B \in \mathscr{B}$$

(ver de nuevo 10.15).

La expresión que define la medida de Borel-Lebesgue tiene perfecto sentido para cualquier subconjunto de $\mathbb{R}^N$ dando lugar a la llamada *medida exterior de Lebesgue*, λ^*. Desgraciadamente λ^* no es una medida sobre la σ-álgebra $\mathscr{P}(\mathbb{R}^N)$ (ver ejercicios 10.8 y 10.9).

c) La σ-álgebra y la medida de Lebesgue.

La σ-álgebra de Lebesgue en $\mathbb{R}^N$ (que presentamos en el Teorema 10.5), se nota por $\mathcal{M}$ y es la mayor σ-álgebra en $\mathbb{R}^N$ que contiene a los intervalos acotados y sobre la cual la medida exterior de Lebesgue es aditiva (ver ejercicio 10.4). De hecho la restricción de λ^* a $\mathcal{M}$ es la única medida sobre $\mathcal{M}$ que extiende el volumen de los intervalos acotados (ver Teorema 10.15). Esta medida recibe el nombre de medida de Lebesgue y la notaremos también por λ.

Curiosamente $\mathcal{M}$ es sólo un ligero agrandamiento de $\mathcal{B}$, de hecho es la "completación" de $\mathcal{B}$, es decir:

$$\mathcal{M} = \{B \cup Z : B \in \mathcal{B}, Z \subset A \in \mathcal{B} \text{ con } \lambda(A) = 0\}$$

y para $B \cup Z$ con $B \in \mathcal{B}$ y $Z \subset A \in \mathcal{B}$ con $\lambda(A) = 0$ se define $\lambda(B \cup Z) = \lambda(B)$ (ver una vez más 10.15 junto con 10.13). Es conocido sin embargo que existen extensiones de la medida de Lebesgue, aunque no hay ninguna forma razonable de medir en todo $\mathcal{P}(\mathbb{R}^N)$.

d) La σ-álgebra y la medida inducida.

Si $(\Omega, \mathcal{A})$ es un espacio medible y E es un subconjunto medible de Ω no vacío, es inmediato comprobar que

$$\mathcal{A}_E := \{E \cap A : A \in \mathcal{A}\} \quad (= \{A \in \mathcal{A} : A \subset E\})$$

es una σ-álgebra en E, que se denomina σ-álgebra inducida.

Si $(\Omega, \mathcal{A}, \mu)$ es un espacio de medida y E es un subconjunto medible no vacío de Ω, es inmediato probar que la restricción, μ_E, de μ a la σ-álgebra $\mathcal{A}_E$ es una medida sobre $\mathcal{A}_E$ que se denomina *medida inducida*. En lo sucesivo cualquier subconjunto medible no vacío de un espacio medible (resp. de medida) se considerará como un espacio medible (resp. de medida) con la σ-álgebra (resp. la medida) inducida. La terna $(E, \mathcal{A}_E, \mu_E)$ se denomina *espacio de medida inducido*.

10.2. Construcción de la medida de Lebesgue en $\mathbb{R}^N$.

Comenzamos ahora la construcción de la medida de Lebesgue en $\mathbb{R}^N$. Como anunciábamos, el germen de dicha construcción es el volumen de los intervalos de dimensión N, concepto que precisamos a continuación.

Definición 10.6 (Volumen de un intervalo de dimensión N)**.** Un intervalo en $\mathbb{R}^N$ o *intervalo de dimensión N* es un conjunto de la forma

$$I = I_1 \times I_2 \times \cdots \times I_N$$

donde $I_1, I_2, \dots, I_N$ son intervalos (cada uno de ellos de cualquier tipo, abierto, cerrado o semiabierto, acotado o no) en $\mathbb{R}$. Cuando el intervalo I es acotado y no vacío se define su *volumen* por:

$$v(I) := (\sup I_1 - \inf I_1) \cdots (\sup I_N - \inf I_N).$$

Por coherencia convenimos que $v(\emptyset) = 0$.

Si $N = 1$ el volumen 1-dimensional de un intervalo acotado $I \subset \mathbb{R}$ se llama la *longitud* de I y se representa por $\ell(I)$. Para $N = 2$ ($N = 3$) el volumen 2-dimensional (3-dimensional) de un intervalo acotado en $\mathbb{R}^2$ ($\mathbb{R}^3$), que no es otra cosa que un rectángulo (ortoedro) de lados paralelos a los ejes, es el *área* (*volumen*) de dicho rectángulo (ortoedro).

Notaremos por $\mathscr{J}$ a la familia de los *intervalos acotados* de $\mathbb{R}^N$.

Para algunas demostraciones es más cómodo restringirse a una familia más pequeña de intervalos acotados.

Definición 10.7 (cubo diádico de dimensión N)**.** Dado $a = (a_1, \dots, a_N) \in \mathbb{R}^N$ y $\delta > 0$, llamamos *N-cubo* de lado δ y vértice a al intervalo acotado

$$Q(a, \delta) = [a_1, a_1 + \delta[\times [a_2, a_2 + \delta[\times \ldots \times [a_N, a_N + \delta[.$$

Para cada $n \in \mathbb{N}$ notaremos por P_n el conjunto de todos los N-cubos de lado $\frac{1}{2^n}$ con vértice en un punto x de $\mathbb{R}^N$ cuyas coordenadas sean múltiplos enteros de $\frac{1}{2^n}$. Es decir:

$$P_n = \left\{ Q\left(x, \frac{1}{2^n}\right) : x \in \mathbb{R}^N, 2^n x \in \mathbb{Z}^N \right\}.$$

Tales intervalos acotados se denominan *cubos diádicos* de dimensión N.

No es difícil probar que para cada natural n la familia P_n constituye una partición (numerable) de $\mathbb{R}^N$.

En $\mathbb{R}^N$ tenemos el siguiente resultado acerca de la expresión de un conjunto abierto como unión de cubos diádicos disjuntos, que será frecuentemente usado.

Proposición 10.8 (Descomposición canónica en cubos diádicos)**.** *Todo conjunto abierto no vacío G en $\mathbb{R}^N$ es unión de una sucesión $\{Q_n\}$ de N-cubos diádicos dos a dos disjuntos y cuya adherencia está contenida en G. Además si $K \subset G$ y K es compacto entonces se verifica que $K \subset \cup_{n=1}^m Q_n$ para algún $m \in \mathbb{N}$.*

Demostración:

Para $n = 1$ sea S_1 el conjunto de los N-cubos diádicos de lado $\frac{1}{2}$ cuya adherencia está contenida en G. Para $n \in \mathbb{N}$, $n > 1$ sea S_n el conjunto de los N-cubos diádicos de lado $\frac{1}{2^n}$ cuya adherencia está contenida en G y no están contenidos en ningún N-cubo diádico de lado mayor que $\frac{1}{2^n}$ cuya adherencia esté contenida a su vez en G. Como $\cup_{n=1}^\infty S_n$ es una familia numerable de N-cubos diádicos, podemos considerar una enumeración $\{Q_n : n \in \mathbb{N}\}$ de ésta. Probamos ahora que la sucesión $\{Q_n\}$ satisface las propiedades deseadas.

Para cada natural n definimos

$$F_n := \begin{cases} \cup\{Q : Q \in S_n\} & \text{si } S_n \neq \emptyset \\ \emptyset & \text{si } S_n = \emptyset \end{cases}$$

Es claro que $\cup_{p=1}^{\infty} Q_p = \cup_{n=1}^{\infty} F_n$ y que $G \supset \cup_{n=1}^{\infty} F_n$. Veamos que $G \subset \cup_{n=1}^{\infty} F_n$. Para ello sea $x \in G$; por ser G abierto en $\mathbb{R}^N$ existe $\rho > 0$ tal que $\overline{B}_\infty(x,\rho) \subset G$. Sea $n \in \mathbb{N}$ tal que $\frac{1}{2^n} < \rho$ y sea $Q \in P_n$ tal que $x \in Q$ (la familia P_n es una partición de $\mathbb{R}^N$). Para todo $y \in \overline{Q}$ se tiene que $\|y-x\|_\infty \leq \frac{1}{2^n} < \rho$, así resulta que $\overline{Q} \subset G$; por tanto, o bien $Q \in S_n$, o bien $Q \subset Q'$ con $Q' \in S_m$ y $m < n$. En cualquier caso se tiene que $Q \subset \cup_{n=1}^{\infty} F_n$ por lo que $x \in \cup_{n=1}^{\infty} F_n$. Hemos demostrado así que $G = \cup_{p=1}^{\infty} Q_p$.

Veamos que $Q_p \cap Q_q = \emptyset$ para $p \neq q$. Sea $Q_p \in S_m$ y $Q_q \in S_n$. Si $m = n$ entonces Q_p y Q_q son N-cubos diádicos distintos de igual lado y por tanto se tiene que $Q_p \cap Q_q = \emptyset$. Si $m \neq n$, Q_p y Q_q son N-cubos diádicos de distinto lado y, por definición de los conjuntos S_n, el de menor lado no está contenido en el de mayor lado, luego se tiene que $Q_p \cap Q_q = \emptyset$.

Sea ahora $K \subset G$ compacto. Si $G = \mathbb{R}^N$ entonces $Q_p \in S_1$, $\forall p \in \mathbb{N}$ y la existencia de m es inmediata. En otro caso, sea:

$$\alpha = \inf\{\|y-x\|_\infty : x \in K, y \in \mathbb{R}^N \setminus G\}.$$

Por ser K compacto y $\mathbb{R}^N \setminus G$ cerrado y ambos disjuntos, ha de ser $\alpha > 0$. Sea n_0 el más pequeño número natural tal que $\frac{1}{2^{n_0}} < \alpha$. Si ahora $x \in K$ y $Q \in P_{n_0}$ es tal que $x \in Q$, se tiene que $\overline{Q} \subset G$ por lo cual, o bien $Q \in S_{n_0}$, o bien $Q \subset Q'$ con $Q' \in S_m$ y $m < n_0$, así resulta que

$$K \subset \cup_{n=1}^{n_0} F_n.$$

Como K es un conjunto acotado es claro que para todo $n \in \mathbb{N}$ el conjunto

$$\{Q : Q \in S_n, Q \cap K \neq \emptyset\}$$

es finito, en consecuencia la familia

$$\{Q : Q \in S_n, Q \cap K \neq \emptyset, 1 \leq n \leq n_0\}$$

es finita y, según acabamos de ver, K está contenido en la unión de dicha familia, luego $\exists m \in \mathbb{N}$ tal que $K \subset \cup_{n=1}^{m} Q_n$. ∎

Corolario 10.9. *La σ-álgebra generada por $\mathscr{J}$ coincide con la σ-álgebra de Borel $\mathscr{B}$.*

Demostración:

Sea $\mathscr{A}$ la σ-álgebra generada por $\mathscr{J}$. En vista del resultado anterior se tiene que $\Im \subset \mathscr{A}$ y por tanto $\mathscr{B} \subset \mathscr{A}$. Veamos que $\mathscr{J} \subset \mathscr{B}$ y en consecuencia $\mathscr{A} \subset \mathscr{B}$. En efecto, basta pensar que un intervalo acotado de dimensión N es la intersección de $2N$ semiespacios que obviamente son borelianos por ser abiertos o cerrados. ∎

Si A es un subconjunto de $\mathbb{R}^N$, la manera más natural en la que cabe pensar medir A es sin duda la que da la siguiente definición.

Definición 10.10 (Medida exterior de Lebesgue)**.** La medida exterior de Lebesgue es por definición la aplicación $\lambda^* : \mathscr{P}(\mathbb{R}^N) \to [0,\infty]$ dada por

$$\lambda^*(A) := \inf\left\{\sum_{n=1}^{\infty} v(I_n) : A \subset \cup_{n\in\mathbb{N}} I_n,\, I_n \in \mathscr{J},\ \forall n \in \mathbb{N}\right\},\ \forall A \subset \mathbb{R}^N$$

(obviamente, existen sucesiones $\{I_n\}$ que verifican las condiciones exigidas).

Probemos que también

$$\lambda^*(A) = \inf\left\{\sum_{n=1}^{\infty} v(I_n) : A \subset \cup_{n\in\mathbb{N}} I_n,\ I_n \in \mathscr{J},\, I_n \text{ abierto}, \forall n \in \mathbb{N}\right\},\ \forall A \subset \mathbb{R}^N$$

Para abreviar notamos por $\alpha(A)$ al segundo miembro de la igualdad a demostrar. Es claro que $\lambda^*(A) \le \alpha(A)$. Si $\lambda^*(A) = \infty$, entonces $\lambda^*(A) = \alpha(A)$. Si $\lambda^*(A) < \infty$, para probar que $\alpha(A) \le \lambda^*(A)$ observemos que si $I = I_1 \times \ldots \times I_N$ es un intervalo acotado, y si para $0 < \delta$ definimos

$$I(\delta) :=]\inf I_1 - \delta, \sup I_1 + \delta[\times \cdots \times] \inf I_N - \delta, \sup I_N + \delta[.$$

se tiene que $I(\delta)$ es un intervalo abierto acotado con $I \subset I(\delta)$ y claramente se verifica que

$$v(I(\delta)) \to v(I) \text{ cuando } \delta \to 0.$$

Ahora, dado $\varepsilon > 0$, tomemos una sucesión $\{I_n\}$ en $\mathscr{J}$ tal que

$$A \subset \cup_{n=1}^{\infty} I_n \quad \text{y} \quad \sum_{n=1}^{\infty} v(I_n) \le \lambda^*(A) + \frac{\varepsilon}{2}.$$

Como acabamos de ver, existe para cada $n \in \mathbb{N}$ un intervalo abierto acotado J_n tal que

$$I_n \subset J_n \quad \text{y} \quad v(J_n) \le v(I_n) + \frac{\varepsilon}{2^{n+1}}.$$

Concluimos que

$$\alpha(A) \le \sum_{n=1}^{\infty} v(J_n) \le \sum_{n=1}^{\infty}\left(v(I_n) + \frac{\varepsilon}{2^{n+1}}\right) = \frac{\varepsilon}{2} + \sum_{n=1}^{\infty} v(I_n) \le \lambda^*(A) + \varepsilon$$

y por tanto $\alpha(A) \leq \lambda^*(A)$. ■

Seguidamente exponemos las propiedades básicas de la medida exterior de Lebesgue.

Proposición 10.11 (Propiedades de la medida exterior de Lebesgue)**.** *La aplicación* $\lambda^* : \mathscr{P}(\mathbb{R}^N) \to [0,\infty]$ *definida por*

$$\lambda^*(A) := \inf\left\{\sum_{n=1}^{\infty} v(I_n) : A \subset \cup_{n\in\mathbb{N}} I_n, I_n \in \mathscr{J}, \forall n \in \mathbb{N}\right\}, \forall A \subset \mathbb{R}^N$$

satisface las siguientes propiedades:

i) $\lambda^*(\emptyset) = 0$.

ii) $[A,B \subset \mathbb{R}^N, A \subset B] \Rightarrow \lambda^*(A) \leq \lambda^*(B)$ *(monotonía).*

iii) $[A_n \subset \mathbb{R}^N, \forall n \in \mathbb{N}] \Rightarrow \lambda^*\left(\cup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \lambda^*(A_n)$ *(σ-subaditividad).*

Demostración:

i) Es inmediata ya que el vacío es un intervalo acotado de volumen cero.

ii) Es consecuencia de que si una sucesión $\{I_n\}$ de elementos de $\mathscr{J}$ cubre B también cubre A.

iii) Si $\sum_{n=1}^{\infty} \lambda^*(A_n) = \infty$, no hay nada que probar. En otro caso sea $\varepsilon > 0$. Para cada $n \in \mathbb{N}$ tomemos una sucesión $\{I_{m,n}\}_{m\in\mathbb{N}}$ de elementos de $\mathscr{J}$ que cubre A_n y tal que

$$\sum_{m=1}^{\infty} v(I_{m,n}) \leq \lambda^*(A_n) + \frac{\varepsilon}{2^n}.$$

Denotemos $A = \cup_{n=1}^{\infty} A_n$. Si σ es cualquier biyección de $\mathbb{N}$ sobre $\mathbb{N}\times\mathbb{N}$, de $A \subset \cup_{k=1}^{\infty} I_{\sigma(k)}$ deducimos, teniendo en cuenta el Apéndice A sobre $[0,\infty]$, que

$$\lambda^*(A) \leq \sum_{k=1}^{\infty} v(I_{\sigma(k)}) = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} v(I_{m,n}) \leq \sum_{n=1}^{\infty} \lambda^*(A_n) + \varepsilon$$

y la arbitrariedad de ε demuestra la σ-subaditividad. ■

Definición 10.12 (medida exterior)**.** Una <u>*medida exterior*</u> en un conjunto no vacío Ω es por definición una aplicación $\mu^* : \mathscr{P}(\Omega) \to [0,\infty]$ verificando:

i) $\mu^*(\emptyset) = 0$.

ii) $[A,B \subset \Omega, A \subset B] \Rightarrow \mu^*(A) \leq \mu^*(B)$ (monotonía).

iii) $[A_n \subset \Omega, \forall n \in \mathbb{N}] \Rightarrow \mu^*\left(\cup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$ (σ-subaditividad).

Proposición 10.13 (Regularidad de la medida exterior de Lebesgue)**.** *Dado* $A \subset \mathbb{R}^N$*, existe* B *boreliano tal que* $A \subset B$ *y* $\lambda^*(A) = \lambda^*(B)$.

Demostración:

Si $\lambda^*(A) = \infty$ podemos tomar $B = \mathbb{R}^N$. En otro caso para cada $n \in \mathbb{N}$ existe una sucesión $\{I_{m,n}\}_{m\in\mathbb{N}}$ de intervalos abiertos acotados que cubre A tal que

$$\sum_{m=1}^{\infty} v(I_{m,n}) \leq \lambda^*(A) + \frac{1}{n}.$$

En consecuencia el conjunto $G_n = \cup_{m=1}^{\infty} I_{m,n}$ es un abierto que contiene a A y verifica $\lambda^*(G_n) \leq \lambda^*(A) + \frac{1}{n}$. Finalmente el conjunto $B = \cap_{n=1}^{\infty} G_n$ es un boreliano (de tipo G_δ) que verifica $A \subset B \subset G_n,\ \forall n \in \mathbb{N}$, luego

$$\lambda^*(A) \leq \lambda^*(B) \leq \lambda^*(G_n) \leq \lambda^*(A) + \frac{1}{n},$$

con lo que $\lambda^*(A) = \lambda^*(B)$. ∎

Del resultado anterior y de las propiedades de la medida de Borel-Lebesgue se deduce el siguiente resultado, cuya demostración tenemos que posponer hasta haber probado efectivamente que $\lambda^*_{|\mathscr{B}}$ es una medida.

Proposición 10.14 (Crecimiento cont. de la medida exterior de Lebesgue)**.** *Si* $\{A_n\}$ *es una sucesión creciente de subconjuntos de* $\mathbb{R}^N$*, entonces*

$$\lambda^*\left(\cup_{n=1}^{\infty} A_n\right) = \lim \lambda^*(A_n).$$

10.3. Existencia y unicidad de la medida de Lebesgue

Disponemos ya del bagaje necesario para presentar el teorema fundamental de la lección en el que se describe el espacio de medida $(\mathbb{R}^N, \mathscr{M}, \lambda)$.

Teorema 10.15 (Existencia y unicidad de la medida de Lebesgue)**.**

i) La familia de subconjuntos de $\mathbb{R}^N$

$$\mathscr{M} = \{B \cup Z : B \in \mathscr{B},\ \lambda^*(Z) = 0\}$$

es una σ*-álgebra en* $\mathbb{R}^N$ *que contiene a la familia de los intervalos acotados.*

ii) La restricción a $\mathscr{M}$ *(resp.* $\mathscr{B}$*) de la medida exterior de Lebesgue, denotada por* λ*, es la única medida sobre* $\mathscr{M}$ *(resp.* $\mathscr{B}$*) que extiende el volumen de los intervalos acotados.*

Obsérvese que los conjuntos de medida cero de $\mathscr{M}$ son simplemente aquellos subconjuntos Z de $\mathbb{R}^N$ tales que $\lambda^*(Z) = 0$. En consecuencia el espacio de medida $(\mathbb{R}^N, \mathscr{M}, \lambda)$ es completo. Es interesante observar que los conjuntos de medida cero tienen interior vacío (Hágase!).

Demostración:

$\boxed{\textbf{i}}$) Es claro que $\mathbb{R}^N \in \mathscr{M}$. Si $A \cup Z \in \mathscr{M}$, con $A \in \mathscr{B}$ y $\lambda^*(Z) = 0$, entonces, en virtud de la regularidad de la medida exterior de Lebesgue, existe $B \in \mathscr{B}$ tal que $Z \subset B$ y $\lambda^*(B) = 0$. En esta situación podemos escribir

$$(A \cup Z)^C = A^C \cap Z^C = A^C \cap [B^C \cup (Z^C \cap B)] = [A^C \cap B^C] \cup [A^C \cap (Z^C \cap B)],$$

y por tanto $(A \cup Z)^C = A' \cup Z'$, siendo $A' = A^C \cap B^C \in \mathscr{B}$ y $Z' = A^C \cap (Z^C \cap B)$ que satisface $\lambda^*(Z') \leq \lambda^*(B) = 0$, por lo que $(A \cup Z)^C \in \mathscr{M}$.

Si ahora $\{A_n \cup Z_n\}$ es una sucesión de elementos de $\mathscr{M}$, entonces

$$\cup_{n=1}^{\infty}(A_n \cup Z_n) = (\cup_{n=1}^{\infty} A_n) \ \cup \ (\cup_{n=1}^{\infty} Z_n) \in \mathscr{M}$$

puesto que $\cup_{n=1}^{\infty} A_n \in \mathscr{B}$ y $\lambda^* \left(\cup_{n=1}^{\infty} Z_n\right) \leq \sum_{n=1}^{\infty} \lambda^*(Z_n) = 0$.

Finalmente el Corolario 10.9 nos asegura que $\mathscr{J} \subset \mathscr{M}$ (si se quiere una demostración alternativa de este hecho, sin utilizar el citado corolario, basta pensar que un intervalo acotado es unión de un intervalo abierto y el conjunto Z unión de sus eventuales caras que claramente verifica $\lambda^*(Z) = 0$).

$\boxed{\textbf{ii}}$) La prueba de este apartado es laboriosa. Comprobaremos en primer lugar que la restricción de λ^* a $\mathscr{M}$ es una medida. Supuesto que λ^* sea aditiva en la σ-álgebra $\mathscr{M}$ obsérvese que para cualquier $E \in \mathscr{M}$ necesariamente ha de satisfacerse que

$$\lambda^*(A) = \lambda^*(A \cap E) + \lambda^*(A \backslash E), \ \forall A \subset \mathbb{R}^N.$$

En efecto, dado $A \subset \mathbb{R}^N$ sea $B \in \mathscr{B}$ como en 10.13. Se tiene

$$\lambda^*(A) = \lambda^*(B) = \lambda^*(B \cap E) + \lambda^*(B \backslash E) \geq \lambda^*(A \cap E) + \lambda^*(A \backslash E)$$

y la subaditividad de λ^* (es inmediato que una medida exterior es subaditiva) nos permite concluir que la anterior desigualdad es de hecho una igualdad. Es natural entonces considerar la familia

$$\mathscr{C} = \{E \subset \mathbb{R}^N : \lambda^*(A) = \lambda^*(A \cap E) + \lambda^*(A \setminus E), \ \forall A \subset \mathbb{R}^N\},$$

es decir, la familia de los conjuntos que "parten bien respecto a la medida exterior" cualquier conjunto.

Probaremos seguidamente las siguientes afirmaciones:

a) $\mathscr{C}$ es una σ-álgebra y λ^* es una medida sobre $\mathscr{C}$.

b) La σ-álgebra $\mathscr{M}$ está contenida en la σ-álgebra $\mathscr{C}$.

Como consecuencia de estos apartados se concluye que λ es una medida.

$\boxed{\textbf{a}}$) Es inmediato comprobar que $\mathbb{R}^N \in \mathscr{C}$ y que para cualquier $A \in \mathscr{C}$ también $A^C \in \mathscr{C}$. Dados $E, F \in \mathscr{C}$ y $A \subset \mathbb{R}^N$, tenemos:

$$\lambda^*(A) = \lambda^*(A \cap E) + \lambda^*(A \cap E^C) =$$

$$\lambda^*(A\cap E\cap F)+\lambda^*(A\cap E\cap F^C)+\lambda^*(A\cap E^C\cap F)+\lambda^*(A\cap E^C\cap F^C).$$

Al sustituir A por $A\cap(E\cup F)$ queda

$$\lambda^*(A\cap(E\cup F))=\lambda^*(A\cap E\cap F)+\lambda^*(A\cap E\cap F^C)+\lambda^*(A\cap E^C\cap F), \qquad (1)$$

y por tanto

$$\lambda^*(A)=\lambda^*(A\cap(E\cup F))+\lambda^*(A\cap E^C\cap F^C)=\lambda^*(A\cap(E\cup F))+\lambda^*(A\cap(E\cup F)^C).$$

Obtenemos así que $E\cup F\in\mathscr{C}$. Por inducción se sigue que la unión finita de conjuntos de $\mathscr{C}$ pertenece a $\mathscr{C}$.

Sean ahora $E,F\in\mathscr{C}$ disjuntos, por (1) se tiene:

$$\lambda^*(A\cap(E\cup F))=\lambda^*(A\cap E)+\lambda^*(A\cap F),\ \forall A\subset\mathbb{R}^N.$$

Si ahora $E_1,E_2,\ldots,E_n\in\mathscr{C}$ son disjuntos dos a dos, obtenemos por inducción que:

$$\lambda^*(A\cap(E_1\cup\ldots\cup E_n))=\lambda^*(A\cap E_1)+\ldots+\lambda^*(A\cap E_n),\ \forall A\subset\mathbb{R}^N. \qquad (2)$$

Sea finalmente una sucesión $\{E_n\}$ de elementos de $\mathscr{C}$ disjuntos dos a dos y pongamos $F_n=E_1\cup\ldots\cup E_n$, $\forall n\in\mathbb{N}$, y $E=\cup_{n=1}^{\infty}E_n$. Entonces

$$\lambda^*(A)=\lambda^*(A\cap F_n)+\lambda^*(A\backslash F_n)=\sum_{k=1}^{n}\lambda^*(A\cap E_k)+\lambda^*(A\backslash F_n)\geq$$

$$\geq\sum_{k=1}^{n}\lambda^*(A\cap E_k)+\lambda^*(A\backslash E),$$

donde se ha usado (2) y la monotonía de λ^*. Haciendo que $n\to\infty$ y usando la σ-subaditividad y la subaditividad de λ^*, obtenemos

$$\lambda^*(A)\geq\sum_{n=1}^{\infty}\lambda^*(A\cap E_n)+\lambda^*(A\backslash E)\geq\lambda^*\left(\cup_{n=1}^{\infty}(A\cap E_n)\right)+\lambda^*(A\backslash E)=$$

$$=\lambda^*(A\cap E)+\lambda^*(A\backslash E)\geq\lambda^*(A),$$

lo que demuestra que $E\in\mathscr{C}$ y que:

$$\lambda^*(A)=\sum_{n=1}^{\infty}\lambda^*(A\cap E_n)+\lambda^*(A\backslash E),\ \forall A\subset\mathbb{R}^N. \qquad (3)$$

Tomando $A=E$ en (3) obtenemos la σ-aditividad de λ^*, es decir

$$\left[E_n\in\mathscr{C},\forall n\in\mathbb{N},E_i\cap E_j=\emptyset(i\neq j)\right]\Rightarrow\lambda^*\left(\cup_{n=1}^{\infty}E_n\right)=\sum_{n=1}^{\infty}\lambda^*(E_n).$$

Probemos finalmente que la unión numerable de elementos de $\mathscr{C}$ pertenece a $\mathscr{C}$. En primer lugar si $E,F\in\mathscr{C}$, se tiene que $E\cap F=(E^C\cup F^C)^C\in\mathscr{C}$ y en consecuencia $E\backslash F=E\cap F^C\in\mathscr{C}$. Sea ahora una sucesión cualquiera $\{E_n\}$ de elementos de $\mathscr{C}$. Definimos por recurrencia

$$F_1:=E_1 \text{ y } F_{n+1}:=E_{n+1}\backslash\cup_{k=1}^{n}E_k.$$

De lo ya demostrado $F_n \in \mathscr{C}$, $\forall n \in \mathbb{N}$. Claramente dichos conjuntos son disjuntos dos a dos por lo que concluimos que

$$\cup_{n=1}^{\infty} E_n = \cup_{n=1}^{\infty} F_n \in \mathscr{C}.$$

$\boxed{\mathbf{b})}$ Supongamos que S es un semiespacio de $\mathbb{R}^N$ del tipo:

$$\{(x_1,\ldots,x_N) \in \mathbb{R}^N : x_i < \alpha\} \ \ (\text{resp. } \ \le, >, \ge),$$

para algún i con $1 \le i \le N$ y algún $\alpha \in \mathbb{R}$. Veamos que $S \in \mathscr{C}$. Si $A \subset \mathbb{R}^N$ y $A \subset \cup_{n=1}^{\infty} I_n$ con $I_n \in \mathscr{J}$, $\forall n \in \mathbb{N}$, entonces los intervalos $I_n \cap S$ recubren $A \cap S$ y los intervalos $I_n \backslash S$ recubren $A \backslash S$ y por tanto se verifica

$$\lambda^*(A \cap S) + \lambda^*(A \backslash S) \le$$
$$\sum_{n=1}^{\infty} v(I_n \cap S) + \sum_{n=1}^{\infty} v(I_n \backslash S) = \sum_{n=1}^{\infty} (v(I_n \cap S) + v(I_n \backslash S)) = \sum_{n=1}^{\infty} v(I_n).$$

Consecuentemente $\lambda^*(A \cap S) + \lambda^*(A \backslash S) \le \lambda^*(A)$ y la subaditividad de la medida exterior nos permite concluir que la anterior desigualdad es de hecho una igualdad y que por tanto $S \in \mathscr{C}$. Como cualquier intervalo acotado de $\mathbb{R}^N$ se puede expresar como la intersección de $2N$ semiespacios del tipo anterior y éstos pertenecen a $\mathscr{C}$ podemos asegurar que también los intervalos acotados pertenecen a $\mathscr{C}$. Como $\mathscr{B}$ es la σ-álgebra que genera $\mathscr{J}$ podemos ahora concluir que $\mathscr{B} \subset \mathscr{C}$. Supongamos finalmente que $Z \subset \mathbb{R}^N$ verifica que $\lambda^*(Z) = 0$. Para cualquier $A \subset \mathbb{R}^N$ se tiene entonces

$$\lambda^*(A \cap Z) + \lambda^*(A \backslash Z) = \lambda^*(A \backslash Z) \le \lambda^*(A),$$

donde se ha utilizado la monotonía de λ^*. La subaditividad de λ^* nos permite concluir que la anterior desigualdad es de hecho una igualdad y por tanto $Z \in \mathscr{C}$. Hemos probado que $\mathscr{M} \subset \mathscr{C}$ (de hecho probaremos enseguida que ambas σ-álgebras coinciden).

Una vez probado que λ es una medida, demostraremos ahora que extiende el volumen de los intervalos acotados, esto es,

$$\lambda(I) = v(I), \ \forall I \in \mathscr{J}.$$

Esta es sin duda la parte más difícil del teorema que de hecho requiere la utilización del Teorema de Heine-Borel-Lebesgue (Apéndice A del Tema 2). Supongamos que I es un intervalo acotado de $\mathbb{R}^N$. Es claro que $\lambda(I) \le v(I)$, pues I es un recubrimiento de sí mismo. En consecuencia si $v(I) = 0$, también $\lambda(I) = 0$. Para probar la igualdad en el caso $v(I) > 0$, observemos que si $I = I_1 \times \ldots \times I_N$, y para cada δ tal que

$$0 < \delta < \frac{1}{2} \min\{\sup I_1 - \inf I_1, \ldots, \sup I_N - \inf I_N\}$$

definimos

$$I(\delta) = [\inf I_1 + \delta, \sup I_1 - \delta] \times \ldots \times [\inf I_N + \delta, \sup I_N - \delta],$$

se tiene que $I(\delta)$ es un intervalo cerrado tal que $I(\delta) \subset I$ y

$$v(I(\delta)) \to v(I) \text{ cuando } \delta \to 0.$$

Fijado $\varepsilon > 0$, podemos en consecuencia tomar un intervalo cerrado K contenido en I tal que

$$v(I) < v(K) + \varepsilon.$$

Sea $\{I_n\}$ una sucesión de intervalos abiertos y acotados de $\mathbb{R}^N$ que recubran I y tal que

$$\sum_{n=1}^{\infty} v(I_n) \leq \lambda(I) + \varepsilon.$$

Puesto que K es compacto ha de ocurrir que $K \subset I_1 \cup \ldots \cup I_m$ para conveniente $m \in \mathbb{N}$ (Teorema de Heine-Borel-Lebesgue). Se verifica entonces que

$$v(K) \leq v(I_1) + \ldots + v(I_m) \leq \sum_{n=1}^{\infty} v(I_n) \leq \lambda(I) + \varepsilon$$

(para comprobar la primera desigualdad véase el Apéndice B sobre "Subaditividad del volumen"), y por tanto $v(I) < \lambda(I) + 2\varepsilon$. Como quiera que la anterior desigualdad es válida para cualquier positivo ε podemos concluir que $v(I) \leq \lambda(I)$ y en consecuencia $\lambda(I) = v(I)$.

Por último, probaremos la unicidad anunciada en el teorema haciendo uso de la magnífica relación existente entre la medida de Lebesgue y la topología de $\mathbb{R}^N$ que recogemos en el siguiente resultado.

Teorema 10.16. *Para $E \subset \mathbb{R}^N$ equivalen:*

i) $E \in \mathscr{M}$.

ii) $E \in \mathscr{C}$, *esto es,* $\lambda^*(A) = \lambda^*(A \cap E) + \lambda^*(A \backslash E)$, $\forall A \subset \mathbb{R}^N$.

iii) $\forall \varepsilon > 0, \exists G$ *abierto* $: E \subset G$ *y* $\lambda^*(G \backslash E) < \varepsilon$.

iv) $\exists A$ *de tipo* $G_\delta : E \subset A$ *y* $\lambda^*(A \backslash E) = 0$.

v) $\forall \varepsilon > 0, \exists F$ *cerrado* $: F \subset E$ *y* $\lambda^*(E \backslash F) < \varepsilon$.

vi) $\exists B$ *de tipo* $F_\sigma : B \subset E$ *y* $\lambda^*(E \backslash B) = 0$.

Para todo conjunto medible E se verifican:

$$\lambda(E) = \left\{ \begin{array}{ll} \inf\{\lambda(G) : G \textit{ abierto}, E \subset G\} & (\underline{\textit{regularidad exterior de }}\lambda) \\ \sup\{\lambda(K) : K \textit{ compacto}, K \subset E\} & (\underline{\textit{regularidad interior de }}\lambda) \end{array} \right\}$$

La caracterización de los conjuntos medibles dada por *ii*) se debe a Carathéodory y no requiere usar la topología.

Demostración:

Comenzamos probando que para $E \subset \mathbb{R}^N$ se verifica

$$\lambda^*(E) = \inf\{\lambda(G) : G \text{ abierto}, E \subset G\},$$

lo que es una generalización de la regularidad exterior. Notemos por comodidad

$$\alpha = \inf\{\lambda(G) : G \text{ abierto}, E \subset G\}.$$

Por monotonía es $\lambda^*(E) \leq \alpha$ por lo que si $\lambda^*(E) = \infty$ entonces también $\alpha = \infty$. En otro caso, dado $\varepsilon > 0$ tomemos una sucesión $\{I_n\}$ de intervalos abiertos acotados cuya unión contenga a E, y tal que

$$\sum_{n=1}^{\infty} v(I_n) < \lambda^*(E) + \varepsilon.$$

Entonces $G = \cup_{n\in\mathbb{N}} I_n$ es un abierto que contiene a E, con lo que usando la definición de medida exterior de Lebesgue, tenemos

$$\alpha \leq \lambda(G) = \lambda^*(G) \leq \sum_{n=1}^{\infty} v(I_n) < \lambda^*(E) + \varepsilon.$$

Siendo $\varepsilon > 0$ arbitrario, se sigue que $\alpha \leq \lambda^*(E)$.

$\boxed{\textbf{i)} \Rightarrow \textbf{ii)}}$ En el teorema anterior se ha probado que $\mathscr{M} \subset \mathscr{C}$, donde

$$\mathscr{C} = \{E \subset \mathbb{R}^N : \lambda^*(A) = \lambda^*(A \cap E) + \lambda^*(A \backslash E),\ \forall A \subset \mathbb{R}^N\}.$$

$\boxed{\textbf{ii)} \Rightarrow \textbf{iii)}}$ Si $\lambda^*(E) < \infty$, fijemos $\varepsilon > 0$ y tomemos un abierto G con $E \subset G$ y $\lambda(G) < \lambda^*(E) + \varepsilon$. Por ii)

$$\lambda(G) = \lambda^*(G \cap E) + \lambda^*(G \backslash E) = \lambda^*(E) + \lambda^*(G \backslash E).$$

En consecuencia

$$\lambda^*(G \backslash E) < \varepsilon.$$

En el caso en que $\lambda^*(E) = \infty$, definamos $E_n = E \cap B_\infty(0,n), \forall n \in \mathbb{N}$. Como para cada natural n se verifican

$$E_n \in \mathscr{C} \quad \text{y} \quad \lambda^*(E_n) \leq \lambda(B_\infty(0,n)) = (2n)^N < \infty,$$

por lo ya demostrado, existe un abierto G_n conteniendo a E_n y tal que

$$\lambda^*(G_n \backslash E_n) < \frac{\varepsilon}{2^n}.$$

Pongamos $G = \cup_{n\in\mathbb{N}} G_n$; G es abierto, y como $E = \cup_{n\in\mathbb{N}} E_n$, G cubre a E. Además como $G \backslash E \subset \cup_{n\in\mathbb{N}} (G_n \backslash E_n)$ se tiene

$$\lambda^*(G \backslash E) \leq \lambda^*\left(\cup_{n\in\mathbb{N}} (G_n \backslash E_n)\right) \leq \sum_{n=1}^{\infty} \lambda^*(G_n \backslash E_n) < \varepsilon$$

donde se ha empleado la monotonía y la σ-subaditividad.

$\boxed{\textbf{iii)} \Rightarrow \textbf{iv)}}$ Para cada natural n sea G_n un abierto tal que

$$E \subset G_n \quad \text{y} \quad \lambda^*(G_n \backslash E) < \frac{1}{n}.$$

Definiendo $A = \cap_{n \in \mathbb{N}} G_n$ se obtiene un conjunto de tipo G_δ con $E \subset A$ y, puesto que $A \backslash E \subset G_n \backslash E$, $\forall n \in \mathbb{N}$, se tiene

$$\lambda^*(A \backslash E) \leq \lambda^*(G_n \backslash E) < \frac{1}{n}, \ \forall n \in \mathbb{N},$$

y por tanto $\lambda^*(A \backslash E) = 0$.

$\boxed{\textbf{iv)} \Rightarrow \textbf{i)}}$ La hipótesis nos dice que $A, A \backslash E \in \mathscr{M}$ (A es boreliano y $\lambda^*(A \backslash E) = 0$ por lo que ambos pertenecen a $\mathscr{M}$), luego $E = A \backslash (A \backslash E) \in \mathscr{M}$.

Hemos probado que $i) \Leftrightarrow ii) \Leftrightarrow iii) \Leftrightarrow iv)$.

Para comprobar que también $i) \Leftrightarrow v) \Leftrightarrow vi)$ basta tener en cuenta que

$$E \in \mathscr{M} \Leftrightarrow E^C \in \mathscr{M}$$

y que $v)$ (resp. $vi)$) son las escrituras de $iii)$ (resp. $iv)$) para el conjunto E^C.

Probemos finalmente la regularidad interior. Sea E medible y notemos

$$\alpha = \sup\{\lambda(K) : K \text{ compacto}, K \subset E\}.$$

Es claro que $\alpha \leq \lambda(E)$. Para probar la otra desigualdad consideremos para cada natural n el conjunto $E_n = E \cap B_\infty(0, n)$. Puesto que $\{E_n\}$ es una sucesión creciente de conjuntos medibles con unión E, por el crecimiento continuo

$$\lambda(E) = \lim \lambda(E_n).$$

Por $v)$, para cada natural n existe F_n compacto contenido en E_n tal que

$$\lambda(E_n \backslash F_n) < \frac{1}{n}.$$

En consecuencia, $\{F_n\}$ es una sucesión de compactos contenidos en E tales que

$$\lambda(E_n) = \lambda(F_n) + \lambda(E_n \backslash F_n), \ \forall n \in \mathbb{N},$$

y por tanto, tomando limite obtenemos

$$\lambda(E) = \lim \lambda(F_n).$$

Hemos probado que $\lambda(E) \leq \alpha$. ∎

Nótese que $vi)$ es un refinamiento de la definición de conjunto medible puesto que asegura que un tal conjunto medible E se puede expresar de la forma $E = B \cup Z$ con B del tipo F_σ, $\lambda^*(Z) = 0$ y $B \cap Z = \emptyset$.

Demostración de la unicidad de la medida de Lebesgue.

Sea μ una medida definida sobre la σ-álgebra $\mathscr{M}$ (resp. $\mathscr{B}$) que extiende el volumen de los intervalos acotados. Probemos que para cualquier conjunto medible E se tiene que $\mu(E) \leq \lambda(E)$. En efecto sea $\{I_n\}$ una sucesión de intervalos acotados tales que

$$E \subset \cup_{n=1}^{\infty} I_n.$$

De la monotonía y la σ-subaditividad de la medida μ se deduce que

$$\mu(E) \leq \mu\left(\cup_{n=1}^{\infty} I_n\right) \leq \sum_{n=1}^{\infty} \mu(I_n) = \sum_{n=1}^{\infty} v(I_n)$$

y en consecuencia

$$\mu(E) \leq \lambda^*(E) = \lambda(E).$$

Conviene resaltar que hemos probado que cualquier medida que extienda el volumen de los intervalos (definida sobre cualquier σ-álgebra que contenga a $\mathscr{J}$) es más pequeña que la medida exterior de Lebesgue.

Probemos ahora la desigualdad contraria. Sea G un abierto de $\mathbb{R}^N$. Si $\{Q_n\}$ es la descomposición canónica de G en cubos diádicos de dimensión N (Proposición 10.8) se tiene que

$$\mu(G) = \mu\left(\cup_{n=1}^{\infty} Q_n\right) = \sum_{n=1}^{\infty} \mu(Q_n) = \sum_{n=1}^{\infty} v(Q_n) = \sum_{n=1}^{\infty} \lambda(Q_n) = \lambda\left(\cup_{n=1}^{\infty} Q_n\right) = \lambda(G),$$

donde se ha utilizado la σ-aditividad de las medidas.

Sea K un compacto de $\mathbb{R}^N$. Si I es un intervalo abierto acotado que contenga a K se tiene de lo ya demostrado que

$$\mu(K) = \mu(I) - \mu(I \setminus K) = \lambda(I) - \lambda(I \setminus K) = \lambda(K).$$

Ahora la regularidad interior de la medida de Lebesgue nos asegura que para cada conjunto medible Lebesgue E se tiene

$$\lambda(E) = \sup\{\lambda(K) : K \text{ compacto}, K \subset E\},$$

y por tanto

$$\lambda(E) = \sup\{\mu(K) : K \text{ compacto}, K \subset E\} \leq \mu(E),$$

donde se ha utilizado la monotonía de la medida μ. ∎

Ahora ya podemos probar la Proposición 10.14.

Demostración del crecimiento continuo de la medida exterior de Lebesgue.

Como para cada n natural se tiene que $A_n \subset \cup_{n=1}^{\infty} A_n$, deducimos de la monotonía de la medida exterior que $\lambda^*(A_n) \leq \lambda^*\left(\cup_{n=1}^{\infty} A_n\right)$, y en consecuencia

$$\lim \lambda^*(A_n) \leq \lambda^*\left(\cup_{n=1}^{\infty} A_n\right).$$

Para probar la desigualdad contraria podemos suponer que $\lim \lambda^*(A_n) < \infty$. Para cada natural n se puede elegir un boreliano B_n tal que $A_n \subset B_n$ y $\lambda^*(A_n) = \lambda(B_n)$. Notemos $C_n = \cup_{k=1}^{n} B_k$. Es claro que $\{C_n\}$ es una sucesión creciente de borelianos tal que

$$\lambda^*(A_n) \leq \lambda(C_n), \forall n \in \mathbb{N}.$$

Probemos por inducción que

$$\lambda^*(A_n) = \lambda(C_n), \forall n \in \mathbb{N}.$$

Se tiene que $\lambda^*(A_1) = \lambda(C_1)$ y para cada n natural que

$$\lambda(C_{n+1}) - \lambda(C_n) = \lambda(C_{n+1} \backslash C_n) = \lambda(B_{n+1} \backslash (B_{n+1} \cap C_n)) =$$

$$\lambda(B_{n+1}) - \lambda(B_{n+1} \cap C_n) \leq \lambda(B_{n+1}) - \lambda^*(A_n) = \lambda^*(A_{n+1}) - \lambda^*(A_n),$$

donde se ha utilizado la hipótesis de inducción. Del desarrollo anterior y teniendo en cuenta de nuevo la hipótesis de inducción

$$\lambda(C_{n+1}) \leq \lambda^*(A_{n+1}),$$

y por tanto

$$\lambda(C_{n+1}) = \lambda^*(A_{n+1}).$$

Por último, utilizando el crecimiento continuo de la medida de Borel-Lebesgue, concluimos

$$\lambda^*\left(\cup_{n=1}^{\infty} A_n\right) \leq \lambda\left(\cup_{n=1}^{\infty} C_n\right) = \lim \lambda(C_n) = \lim \lambda^*(A_n).$$

■

A continuación estudiamos el magnífico comportamiento de la medida de Lebesgue frente a las transformaciones afines. Empezamos probando que *si* $\varphi : \mathbb{R}^N \to \mathbb{R}^N$ *es una aplicación continua, entonces* $\varphi^{-1}(B) \in \mathscr{B}$, $\forall B \in \mathscr{B}$. En efecto, la familia

$$\{B \in \mathscr{B} : \varphi^{-1}(B) \in \mathscr{B}\}$$

es una σ-álgebra que contiene los abiertos, luego coincide con la σ-álgebra de Borel y por tanto

$$\varphi^{-1}(B) \in \mathscr{B}, \ \forall B \in \mathscr{B}.$$

En particular *los homeomorfismos de* $\mathbb{R}^N$ *sobre* $\mathbb{R}^N$ *conservan los borelianos.* Conviene llamar la atención sobre el hecho de que, a diferencia de lo que ocurre con los borelianos, la imagen inversa por una función continua de un medible puede no serlo. De hecho *existen homeomorfismos de* $\mathbb{R}$ *sobre* $\mathbb{R}$ *que no conservan los conjuntos medibles* (ver ejercicio 10.9). Sin embargo si es cierto que *el trasladado de un medible es medible.* En efecto, como las traslaciones conservan los conjuntos borelianos, basta probar que si $a \in \mathbb{R}^N$ y $Z \subset \mathbb{R}^N$ con $\lambda(Z) = 0$, entonces $\lambda(a + Z) = 0$, lo cual se deduce de que $v(I) = v(a + I)$, $\forall I \in \mathscr{J}$.

10.4. Caracterización de la medida de Lebesgue

Teorema 10.17 (Caracterización de la medida de Lebesgue)**.**

i) Si μ es una medida sobre $\mathscr{M}$ (resp. $\mathscr{B}$) invariante por traslaciones y tal que $\alpha := \mu([0,1[^N) < \infty$, entonces $\mu = \alpha\lambda$.

ii) La medida de Lebesgue es la única medida sobre $\mathscr{M}$ invariante por traslaciones para la que la medida del intervalo $[0,1[^N$ es 1.

Demostración:

$\boxed{\textbf{i)}}$ Notemos

$$Q_0 = [0,1[^N \quad y \quad Q_n = \frac{1}{2^n} Q_0, \ \ \forall n \in \mathbb{N}.$$

Para cada natural n, Q_0 es unión de 2^{nN} intervalos trasladados de Q_n dos a dos disjuntos, por lo que usando que μ es invariante por traslaciones, se tiene que

$$\alpha = \mu(Q_0) = \sum_{j=1}^{2^{nN}} \mu(Q_n) = 2^{nN}\,\mu(Q_n),$$

esto es,

$$\mu(Q_n) = \frac{\alpha}{2^{nN}} = \alpha v(Q_n).$$

Utilizando de nuevo que μ es invariante por traslaciones deducimos que

$$\mu(Q) = \alpha\, v(Q), \text{ para cualquier cubo diádico.}$$

La descomposición de un abierto en unión de diádicos disjuntos (Proposición 10.8) nos asegura que

$$\mu(G) = \alpha\lambda(G) \text{ para cualquier conjunto abierto G.}$$

Sea ahora K un compacto y tomemos I intervalo abierto acotado conteniendo a K. Entonces como $\mu(I) < \infty$, tenemos

$$\mu(K) = \mu(I) - \mu(I \backslash K) = \alpha\lambda(I) - \alpha\lambda(I \backslash K) = \alpha[\lambda(I) - \lambda(I \setminus K)] = \alpha\lambda(K).$$

Para $E \in \mathscr{M}$, en virtud de las regularidades de la medida de Lebesgue y de la monotonía de μ, se tiene que

$$\alpha\lambda(E) = \alpha \sup\{\lambda(K) : K \subset E \text{ compacto}\} = \sup\{\alpha\lambda(K) : K \subset E \text{ compacto}\} =$$

$$\sup\{\mu(K) : K \subset E \text{ compacto}\} \leq \mu(E).$$

$$\alpha\lambda(E) = \alpha \inf\{\lambda(G) : E \subset G \text{ abierto}\} = \inf\{\alpha\lambda(G) : E \subset G \text{ abierto}\} =$$

$$\inf\{\mu(G) : E \subset G \text{ abierto}\} \geq \mu(E).$$

$\boxed{\textbf{ii}}$) Veamos que λ es invariante por traslaciones. Para cada $a \in \mathbb{R}^N$, definimos $\mu : \mathscr{M} \to [0,\infty]$ por

$$\mu(E) = \lambda(a+E),\ \forall E \in \mathscr{M}$$

y basta probar, en virtud del teorema de existencia y unicidad de la medida de Lebesgue, que μ es una medida sobre $\mathscr{M}$ que extiende el volumen de los intervalos. Sea $\{E_n\}$ una sucesión de conjuntos medibles disjuntos. Se tiene que

$$\mu\left(\cup_{n=1}^{\infty} E_n\right) = \lambda\left(a + \cup_{n=1}^{\infty} E_n\right) = \lambda\left(\cup_{n=1}^{\infty}(a+E_n)\right) = \sum_{n=1}^{\infty} \lambda(a+E_n) = \sum_{n=1}^{\infty} \mu(E_n),$$

donde se utilizado que los conjuntos $a+E_n$ son disjuntos dos a dos. Finalmente

$$\mu(I) = \lambda(a+I) = v(a+I) = v(I),\ \forall I \in \mathscr{J}.$$

La unicidad es consecuencia inmediata de i). ∎

10.5. Comportamiento de la medida de Lebesgue frente a aplicaciones

Proposición 10.18. *Sea $T : \mathbb{R}^N \to \mathbb{R}^N$ una aplicación lineal. Para todo conjunto medible $E \subset \mathbb{R}^N$ se verifica que el conjunto $T(E)$ es medible y*

$$\lambda(T(E)) = |\det T|\lambda(E),$$

donde notamos $\det T$ al determinante de la matriz asociada a T. En particular si T es una isometría euclídea, entonces

$$\lambda(T(E)) = \lambda(E),\ \forall E \in \mathscr{M}.$$

Demostración:

Supongamos que $\det T \neq 0$. Entonces T es un isomorfismo y por lo tanto aplica borelianos en borelianos. Esto nos permite definir $\mu : \mathscr{B} \to [0,\infty]$ por

$$\mu(B) = \lambda(T(B)),\ \forall B \in \mathscr{B}$$

que es claramente una medida invariante por traslaciones y tal que $\mu([0,1[^N) < \infty$ (por ser $T([0,1[^N)$ acotado). Así, en virtud del Teorema 10.17 i), se sigue que existe $\alpha_T \geq 0$ tal que

$$\mu(B) = \alpha_T \lambda(B),\ \forall B \in \mathscr{B},$$

y por tanto

$$\lambda(T(B)) = \alpha_T \lambda(B),\ \forall B \in \mathscr{B}.$$

Si ahora Z es de medida cero, entonces, en virtud de la regularidad de la medida exterior de Lebesgue, existe B boreliano con $Z \subset B$ tal que $\lambda(B) = 0$. Se tiene pues que

$$T(Z) \subset T(B) \;\Rightarrow\; \lambda^*(T(Z)) \leq \lambda(T(B)) = \alpha_T \lambda(B) = 0 \;\Rightarrow\; \lambda^*(T(Z)) = 0.$$

En consecuencia,

$$\lambda(T(E)) = \alpha_T \lambda(E),\ \forall E \in \mathcal{M}.$$

Seguidamente probamos que $\alpha_T = 1$ en el caso particular de que T sea una isometría. En efecto, tomando como E la bola unidad euclídea B, al ser $T(B) = B$, obtenemos

$$\lambda(B) = \lambda(T(B)) = \alpha_T \lambda(B).$$

AsÝ $\alpha_T = 1$ y vale el enunciado pues al ser T una isometría, su matriz asociada es ortogonal y en consecuencia tiene determinante ± 1(véase la Proposición 7.25). Hemos probado que si T es una isometría, entonces

$$\lambda(T(E)) = \lambda(E),\ \forall E \in \mathcal{M}.$$

En el caso general en que T sea una aplicación lineal con $\det T \neq 0$, es sabido (véase Apéndice C) que existen isometrías lineales Q_1 y Q_2 y una aplicación lineal D, tal que $D(e_i) = \alpha_i e_i\ \ (1 \leq i \leq N)$, donde $\alpha_1, \ldots, \alpha_N$ son reales positivos, tales que

$$T = Q_1 D Q_2.$$

Es inmediato comprobar que

$$\lambda(D([0,1]^N)) = \lambda([0,\alpha_1] \times \ldots \times [0,\alpha_N]) = \prod_{j=1}^{N} \alpha_j = \det D = |\det T|,$$

y en consecuencia, $\alpha_D = \det D$, como queríamos demostrar. Hemos probado el enunciado para el operador diagonal D.

Por el resultado ya probado para isometrías y para operadores diagonales, concluimos que para cada $E \in \mathcal{M}$ se verifica

$$\lambda(T(E)) = \lambda((Q_1 D Q_2)(E)) = \lambda((D Q_2)(E)) = \det D\, \lambda(Q_2(E)) =$$

$$= |\det T|\, \lambda(Q_2(E)) = |\det T|\, \lambda(E).$$

Supongamos finalmente que $\det T = 0$. En esta situación $T(\mathbb{R}^N)$ está incluido en un hiperplano de $\mathbb{R}^N$ y por tanto existe una isometría lineal Q verificando que $Q(T(\mathbb{R}^N)) \subset \{0\} \times \mathbb{R}^{N-1}$. En consecuencia, por lo ya demostrado,

$$\lambda(T(\mathbb{R}^N)) = \lambda(Q(T(\mathbb{R}^N))) \leq \lambda(\{0\} \times \mathbb{R}^{N-1}) = 0$$

y de ello deducimos que

$$\lambda(T(E)) = 0,\ \forall E \in \mathcal{M}.$$

■

Finalizamos la lección estudiando el comportamiento de la medida de Lebesgue frente a las aplicaciones de clase $\mathscr{C}^1$.

Lema 10.19. *Sea $G \subset \mathbb{R}^N$ un abierto no vacío. Existe entonces una sucesión $\{B_n\}$ de bolas abiertas euclídeas disjuntas entre sí de manera que*

$$\lambda\left(G \setminus \bigcup_{n=1}^{\infty} B_n\right) = 0.$$

Demostración. Supongamos en primer lugar que $\lambda(G) < \infty$. Pongamos

$$G = \bigcup_{k=1}^{\infty} Q_k$$

(descomposición canónica en cubos diádicos (Proposición 10.8)). Como

$$\lambda(G) = \sum_{k=1}^{\infty} \lambda(Q_k)$$

existe $p_1 \in \mathbb{N}$ tal que

$$\sum_{k=1}^{p_1} \lambda(Q_k) > \frac{\lambda(G)}{2}.$$

Para $k = 1, 2, \ldots, p_1$ sean a_k el centro de Q_k y l_k el lado. Consideremos

$$B_2\left(a_k, \frac{l_k}{2}\right) \subset Q_k \quad (k = 1, 2, \ldots, p_1).$$

Se tiene que

$$\lambda\left(B_2\left(a_k, \frac{l_k}{2}\right)\right) = \lambda\left(B_2\left(0, \frac{l_k}{2}\right)\right) = \lambda\left(\frac{l_k}{2}\, B_2(0,1)\right) = \left(\frac{l_k}{2}\right)^N \lambda(B_2(0,1)) =$$

$$(l_k)^N \frac{\lambda(B_2(0,1))}{2^N} = \lambda(Q_k)\, C,$$

donde se han utilizado el Teorema 10.17 y la Proposición 10.18. Así existe una constante C con $0 < C < 1$ tal que

$$\lambda\left(B_2\left(a_k, \frac{l_k}{2}\right)\right) = C\, \lambda(Q_k),\ \forall k = 1, 2, \ldots, p_1 .$$

Nos interesa una buena mayoración de la medida del conjunto $G \setminus \bigcup_{k=1}^{p_1} B_k$, se tiene que

$$\lambda\left(G \setminus \bigcup_{k=1}^{p_1} B_k\right) = \lambda(G) - \sum_{k=1}^{p_1} \lambda(B_k) = \lambda(G) - C \sum_{k=1}^{p_1} \lambda(Q_k) \leq$$

$$\lambda(G) - C\, \frac{\lambda(G)}{2} = \left(1 - \frac{C}{2}\right) \lambda(G) .$$

Consideramos ahora el abierto $G \setminus \bigcup_{k=1}^{p_1} \overline{B_k}$ y se aplica el proceso anterior para obtener

$$B_{p_1+1}, \ldots, B_{p_2}$$

bolas euclídeas disjuntas entre sí contenidas en $G\setminus\bigcup_{k=1}^{p_1}\overline{B_k}$ (y por tanto disjuntas con las $B_k,\ k=1,2,\ldots,p_1$) de manera que

$$\lambda\left(G\setminus\bigcup_{k=1}^{p_2}B_k\right)=\lambda\left(\left(G\setminus\bigcup_{k=1}^{p_1}\overline{B_k})\right)\setminus\bigcup_{p_1+1}^{p_2}B_k)\right)\leq$$

$$\left(1-\frac{C}{2}\right)\lambda\left(\left(G\setminus\bigcup_{k=1}^{p_1}\overline{B_k})\right)\right)\leq\left(1-\frac{C}{2}\right)^2\lambda(G),$$

donde se han utilizado propiedades elementales de la medida de Lebesgue.

Por un proceso de inducción, en el paso n-ésimo, existe un $p_n\in\mathbb{N}$ y bolas euclídeas disjuntas entre sí

$$B_1,\ldots,B_{p_n}$$

tales que

$$\lambda\left(G\setminus\bigcup_{k=1}^{p_n}B_k\right)\leq\left(1-\frac{C}{2}\right)^n\lambda(G)\ .$$

Así

$$\lambda\left(G\setminus\bigcup_{k=1}^{p_n}B_k\right)\leq\left(1-\frac{C}{2}\right)^n\lambda(G),\ \forall n\in\mathbb{N}$$

y por tanto

$$\lambda\left(G\setminus\bigcup_{k=1}^{\infty}B_k\right)=0\ .$$

Supongamos ahora el caso en que $\lambda(G)=\infty$. La descomposición canónica de un abierto en cubos diádicos (Proposición 10.8), nos permite escribir

$$G=\bigcup_{n=1}^{\infty}Q_n.$$

Aplicamos lo ya probado a cada $\overset{\circ}{Q}_n$, y basta ordenar las bolas euclídeas en una sucesión. ∎

Proposición 10.20. *Sean $G\subset\mathbb{R}^N$ abierto, $f:G\to\mathbb{R}^N$ y $K\geq 0$ tales que*

$$\|f(x)-f(y)\|_2\leq K\|x-y\|_2,\ \forall x,y\in G.$$

Entonces

$$\lambda^*(f(E))\leq K^N\lambda^*(E),\ \forall E\subset G.$$

Demostración. Si $\lambda^*(E)=\infty$ no hay nada que probar. En otro caso, fijado $\varepsilon>0$ existe H abierto con $E\subset H\subset G$ tal que

$$\lambda(H)<\lambda^*(E)+\varepsilon$$

(regularidad de la medida exterior de Lebesgue (Proposición 10.13)).

Pongamos

$$H=\bigcup_{n=1}^{\infty}B_2(a_n,r_n)\cup Z,$$

donde $\{B_2(a_n, r_n)\}$ es una sucesión de bolas euclídeas abiertas disjuntas dos a dos y Z es un conjunto de medida cero (Lema 10.19). Se tiene entonces que

$$f(E) \subset f(H) = \bigcup_{n=1}^{\infty} f(B_2(a_n, r_n)) \cup f(Z) \subset B_2(f(a_n), Kr_n) \cup f(Z) ,$$

donde se ha utilizado la hipótesis de lipschitzianidad.

Por un lado se tiene que

$$\lambda\left(\bigcup_{n=1}^{\infty} B_2(f(a_n), Kr_n)\right) \leq \sum_{n=1}^{\infty} \lambda\left(B_2(f(a_n), Kr_n)\right) =$$

$$\sum_{n=1}^{\infty} K^N \lambda\left(B_2(f(a_n), r_n)\right) = K^N \lambda(H) .$$

Por otra parte sea $L \subset G$ es un abierto tal que $Z \subset L \subset G$ con $\lambda(L) < \varepsilon$ (regularidad exterior de la medida de Lebesgue (Teorema 10.16)).

La descomposición canónica de un abierto en cubos diádicos nos da

$$L = \bigcup_{n=1}^{\infty} Q_n .$$

Para cada n notemos b_n al centro y s_n el semilado del cubo diádico Q_n; así

$$\overline{Q}_n = \overline{B}_\infty(b_n, s_n), \ \forall n \in \mathbb{N}$$

y por tanto $L = \bigcup_{n=1}^{\infty} \overline{B}_\infty(b_n, s_n)$. Tenemos pues que

$$f(Z) \subset f(L) \subset f\left(\bigcup_{n=1}^{\infty} \overline{B}_\infty(b_n, s_n)\right) \subset \bigcup_{n=1}^{\infty} \overline{B}_\infty(f(b_n), \sqrt{N} s_n)$$

y también

$$\lambda^*(f(Z)) \leq \sum_{n=1}^{\infty} \left(\sqrt{N}\right)^N s_N^N = \sum_{n=1}^{\infty} \left(\sqrt{N}\right)^N \frac{(2s_n)^N}{2^N} = \left(\frac{\sqrt{N}}{2}\right)^N \sum_{n=1}^{\infty} (2s_n)^N =$$

$$\left(\frac{\sqrt{N}}{2}\right)^N \lambda(L) < \left(\frac{\sqrt{N}}{2}\right)^N \varepsilon .$$

Resumiendo, fijado $\varepsilon > 0$, hemos probado que

$$\lambda^*(f(E)) \leq K^N \lambda(H) + \left(\frac{\sqrt{N}}{2}\right)^N \varepsilon \leq K^N(\lambda^*(E) + \varepsilon) + \left(\frac{\sqrt{N}}{2}\right)^N \varepsilon$$

y la arbitrariedad de ε prueba la proposición. ■

Lema 10.21. *Sea $A \subset \mathbb{R}^N$ cualquiera. Cada recubrimiento abierto $\mathscr{U}$ de A admite un subrecubrimiento numerable.*

Demostración. Pongamos $\mathscr{U} = \{G_i : \ i \in I\}$ Se tiene que

$$A \subset \bigcup_{i \in I} G_i.$$

Para cada $x \in A$ existe $i \in I$ tal que $x \in G_i$. Por ser G_i abierto existen $b_x \in G_i \cap \mathbb{Q}^N$ y $r_x \in \mathbb{Q}^+$ tales que

$$x \in B_2(b_x, r_x) \subset G_i.$$

Es claro que

$$A \subset \cup_{x \in A} B_2(b_x, r_x)$$

y que la familia $\{B_2(b_x, r_x) : x \in A\}$ es numerable (¡Obsérvese que estas bolas se repiten muchas veces!) Sea $\{B_n\}$ una enumeración de dichas bolas y notemos U_n a un abierto tal que $B_n \subset U_n \in \mathscr{U}$. Es Claro que

$$A \subset \bigcup_{n=1}^{\infty} U_n.$$

∎

Proposición 10.22. *Sean $G \subset \mathbb{R}^N$ abierto y $f : G \to \mathbb{R}^N$ una función de clase $\mathscr{C}^1$. Se verifican las siguientes propiedades:*

i) $f(Z)$ es de medida cero para todo $Z \subset G$ de medida cero.

ii) $f(E)$ es medible para todo $E \subset G$ medible.

Demostración. i) Para cada natural n sea $G_n = \{x \in G : \|Df(x)\| < n\}$. Es claro que G_n es abierto y que basta probar que

$$\lambda(f(Z \cap G_n)) = 0, \ \forall n \in \mathbb{N}.$$

Es sabido que G_n se puede expresar como una unión numerable de bolas abiertas (no necesariamente disjuntas entre sí (véase el Lema 10.21)) . Pongamos

$$\{B_2(b_x, r_x) : x \in G_n\} = \{B_2(b_k, r_k) : k \in \mathbb{N}\}.$$

El teorema del valor medio nos asegura que para cada k natural la función f verifica

$$\|f(x) - f(y)\|_2 \leq n\|x - y\|_2, \ \forall x, y \in B_2(b_k, r_k),$$

y de la Proposición 10.20 se deduce que

$$\lambda^*(f(Z \cap G_n)) \leq \sum_{k=1}^{\infty} \lambda^*(f(Z \cap B_2(b_k, r_k)) \leq n^N \sum_{k=1}^{\infty} \lambda^*(Z \cap B_2(b_k, r_k)) = 0.$$

ii) Sea $E \subset G$ medible. El apartado *v*) del Teorema 10.16 nos proporciona una sucesión $\{F_n\}$ de cerrados contenidos en E tal que $\lambda\left(E \setminus \cup_{n=1}^{\infty} F_n\right) = 0$. Consideremos para cada $n \in \mathbb{N}$ el compacto

$$K_n = (F_1 \cup \ldots \cup F_n) \cap \overline{B}(0,n).$$

Se tiene entonces que

$$K_n \subset E,\ \forall n \in \mathbb{N} \quad \text{y} \quad \lambda\left(E \setminus \cup_{n=1}^{\infty} K_n\right) = 0.$$

Como

$$f(E) = f\left(\cup_{n=1}^{\infty} K_n \cup \left(E \setminus \cup_{n=1}^{\infty} K_n\right)\right) = \cup_{n=1}^{\infty} f(K_n) \cup f\left(E \setminus \cup_{n=1}^{\infty} K_n\right),$$

concluimos que $f(E)$ es medible pues la continuidad conserva la compacidad y el apartado *i*) nos asegura que $f\left(E \setminus \cup_{n=1}^{\infty} K_n\right)$ es medible (de hecho es un conjunto de medida cero). ■

10.6. Apéndice A: Orden, topología y aritmética en $[0,\infty]$.

El conjunto
$$[0,\infty] := \{x \in \mathbb{R} : 0 \le x\} \cup \{\infty\}$$
se considera como un conjunto totalmente ordenado, extendiendo el orden usual de $\mathbb{R}_0^+$ mediante el convenio
$$x \le \infty, \ \forall x \in [0,\infty].$$
Es inmediato que todo subconjunto de $[0,\infty]$ tiene supremo e ínfimo. En $[0,\infty]$ se considera la topología del orden, para la cual los conjuntos de la forma $[0,\alpha[$ y $]\beta,\infty]$ con $\alpha,\beta \in \mathbb{Q}^+$, forman una subbase numerable. Dicha topología coincide con la topología de la compactación por un punto de $\mathbb{R}_0^+$. De esta forma $[0,\infty]$ se convierte en un espacio métrico compacto (homeomorfo al intervalo $[0,1]$). Es inmediato que toda sucesión monótona creciente de elementos de $[0,\infty]$ es convergente (al supremo del conjunto de sus términos).

Si extendemos la operación suma en $[0,\infty]$ mediante
$$x+\infty = \infty + x = \infty, \ \forall x \in [0,\infty]$$
tenemos claramente que la suma es asociativa y conmutativa y que 0 es el elemento neutro. Además la suma es compatible con el orden:
$$[x,y \in [0,\infty], x \le y] \Rightarrow x+z \le y+z, \ \forall z \in [0,\infty].$$
Si $\sum_{n\ge 1} a_n$ es una serie de elementos de $[0,\infty]$, entonces, al ser creciente la sucesión de sumas parciales, dicha serie es convergente (aunque su término general no tienda a cero).

Notemos que la suma es continua, es decir:
$$\left[a,b,a_n,b_n \in [0,\infty], \ \forall n \in \mathbb{N}, \left\{ \begin{array}{c} \{a_n\} \to a \\ \{b_n\} \to b \end{array} \right\} \right] \Rightarrow \{a_n+b_n\} \to a+b.$$
Como consecuencia se tiene, por ejemplo, que
$$[a_n,b_n \in [0,\infty], \ \forall n \in \mathbb{N}] \Rightarrow \sum_{n=1}^{\infty}(a_n+b_n) = \sum_{n=1}^{\infty} a_n + \sum_{n=1}^{\infty} b_n. \qquad (*)$$

Más aún, si $a : \mathbb{N}\times\mathbb{N} \to [0,\infty]$ es una sucesión doble de $[0,\infty]$ y σ es una biyección de $\mathbb{N}$ sobre $\mathbb{N}\times\mathbb{N}$, se tiene:
$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{m,n} = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{m,n} = \sum_{k=1}^{\infty} a_{\sigma(k)}.$$

En efecto. Veamos, por ejemplo, la igualdad
$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{m,n} = \sum_{k=1}^{\infty} a_{\sigma(k)}.$$

Claramente para $N \in \mathbb{N}$ fijo, se tiene

$$\sum_{m=1}^{M} \sum_{n=1}^{N} a_{m,n} \leq \sum_{k=1}^{\infty} a_{\sigma(k)}, \ \forall M \in \mathbb{N},$$

luego

$$\sum_{n=1}^{N} \sum_{m=1}^{\infty} a_{m,n} = \sum_{m=1}^{\infty} \sum_{n=1}^{N} a_{m,n} \leq \sum_{k=1}^{\infty} a_{\sigma(k)},$$

(donde se ha utilizado $(*)$) y por tanto

$$\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{m,n} \leq \sum_{k=1}^{\infty} a_{\sigma(k)}.$$

Recíprocamente, para $K \in \mathbb{N}$ fijo, existen $M, N \in \mathbb{N}$ tales que

$$\sum_{k=1}^{K} a_{\sigma(k)} \leq \sum_{m=1}^{M} \sum_{n=1}^{N} a_{m,n},$$

luego

$$\sum_{k=1}^{K} a_{\sigma(k)} \leq \sum_{m=1}^{\infty} \sum_{n=1}^{N} a_{m,n} = \sum_{n=1}^{N} \sum_{m=1}^{\infty} a_{m,n} \leq \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{m,n}$$

(donde se vuelve a utilizar $(*)$), y por tanto

$$\sum_{k=1}^{\infty} a_{\sigma(k)} \leq \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{m,n}.$$

Aplicando lo ya demostrado a $b_{m,n} = a_{n,m}$ y a la biyección $\tau \circ \sigma : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$, donde $\tau(m,n) = (n,m)$, se obtiene la otra igualdad.

La definición de producto es algo más problemática. Extendemos el producto usual en $\mathbb{R}_0^+$ mediante el convenio:

$$x\,\infty = \infty\, x = \infty, \ \forall x \in]0,\infty], \ \ 0\,\infty = \infty\, 0 = 0\ .$$

Es inmediato comprobar que este producto es asociativo, conmutativo y distributivo respecto de la suma, así como que 1 es el elemento neutro. El haber definido

$$0\,\infty = \infty\, 0 = 0$$

hace evidente que el producto no sea continuo. No obstante es claro que si $\{a_n\}$ y $\{b_n\}$ son sucesiones crecientes de elementos de $[0,\infty]$, $\{a_n\} \to a$ y $\{b_n\} \to b$, entonces $\{a_n b_n\} \to ab$. En particular,

$$[a, a_n \in [0,\infty], \forall n \in \mathbb{N}] \Rightarrow a \sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} a\, a_n.$$

10.7. Apéndice B: "Subaditividad del volumen".

Si un intervalo acotado está incluido en la unión de un número finito de intervalos acotados, esto es $I \subset I_1 \cup I_2 \cup \ldots \cup I_n$, entonces

$$v(I) \leq v(I_1) + \ldots + v(I_n).$$

Demostración:

i) Empezaremos probando que si un intervalo acotado I es unión finita de intervalos $I_1, \ldots, I_n$ disjuntos dos a dos, entonces

$$v(I) = v(I_1) + \ldots + v(I_n) \quad (\textit{aditividad del volumen en } \mathscr{J}).$$

Fijemos $j \in \{1, \ldots, N\}$, sean S, T semirrectas disjuntas cuya unión sea $\mathbb{R}$ y pongamos

$$\hat{S} := \{x = (x_1, \ldots, x_N) \in I : x_j \in S\},$$
$$\hat{T} := \{x = (x_1, \ldots, x_N) \in I : x_j \in T\}. \qquad (*)$$

Se deduce directamente de la definición de volumen que

$$v(I) = v(\hat{S}) + v(\hat{T})$$

puesto que $\ell(\pi_j(I)) = \ell(\pi_j(I \cap \hat{S})) + \ell(\pi_j(I \cap \hat{T}))$, mientras que para $k \neq j$ es $\ell(\pi_k(I)) = \ell(\pi_k(I \cap \hat{S})) = \ell(\pi_k(I \cap \hat{T}))$.

Probaremos la proposición por inducción sobre el número de intervalos que intervienen en la descomposición. Acabamos de demostrar que el enunciado es cierto siempre que en la partición intervienen sólo dos intervalos. Sea $n > 2$ un natural tal que el enunciado es cierto para todos los números naturales precedentes. Sea I un intervalo acotado de $\mathbb{R}^N$ tal que $I = I_1 \cup \ldots \cup I_n$ para intervalos disjuntos. Sean $I_1 = A_1 \times \ldots \times A_N$, $I_n = B_1 \times \ldots \times B_N$ las expresiones de I_1 e I_n como producto de intervalos acotados de $\mathbb{R}$, sea $j \in \{1, \ldots, N\}$ tal que $A_j \cap B_j = \emptyset$ (por ser $I_1 \cap I_n = \emptyset$ tal j existe), sean S y T semirrectas disjuntas de $\mathbb{R}$ tales que

$$A_j \subset S, B_j \subset T, S \cup T = \mathbb{R}$$

y definamos $\hat{S}$ y $\hat{T}$ como en $(*)$. Por ser $I_1 \subset \hat{S}$, $I_n \subset \hat{T}$ tenemos:

$$\hat{S} = (\hat{S} \cap I_1) \cup \ldots \cup (\hat{S} \cap I_{n-1}), \hat{T} = (\hat{T} \cap I_2) \cup \ldots \cup (\hat{T} \cap I_n)$$

expresiones de $\hat{S}$ y $\hat{T}$ como uniones de $n-1$ intervalos acotados dos a dos disjuntos. Aplicando la hipótesis de inducción y que $v(\emptyset) = 0$, tenemos:

$$v(I) = v(\hat{S}) + v(\hat{T}) =$$

$$\big(v(\hat{S} \cap I_1) + \cdots + v(\hat{S} \cap I_{n-1})\big) + \big(v(\hat{T} \cap I_2) + \cdots + v(\hat{T} \cap I_n)\big) =$$

$$\big(v(\hat{S}\cap I_1)+v(\hat{T}\cap I_1)\big)+\cdots+\big(v(\hat{S}\cap I_n)+v(\hat{T}\cap I_n)\big)=v(I_1)+\cdots+v(I_n),$$

ya que para $k=1,\ldots,n$

$$I_k=\big(\hat{S}\cap I_k\big)\cup\big(\hat{T}\cap I_k\big)$$

es una partición disjunta de I_k en dos intervalos.

ii) Probemos que la diferencia de dos intervalos acotados se puede expresar como unión finita disjunta de intervalos acotados.

Como la intersección de dos intervalos es un intervalo, es suficiente probar que para cada intervalo I de $\mathbb{R}^N$, el conjunto $\mathbb{R}^N\backslash I$ se puede expresar como unión finita de intervalos disjuntos dos a dos. Para evitar casos triviales, podemos suponer que $\emptyset\neq I\neq\mathbb{R}^N$.

Probaremos que es cierta la afirmación anterior por inducción sobre N. Para $N=1$, es claro que el complemento de un intervalo es una semirrecta o bien es unión disjunta de dos semirrectas, por tanto, $\mathbb{R}\backslash I$ se expresa como unión de, como máximo, dos intervalos disjuntos. Supuesto que es cierta la afirmación para N, y para cualquier intervalo de $\mathbb{R}^N$, sea I un intervalo de $\mathbb{R}^{N+1}$, luego I coincide con el producto de $N+1$ intervalos de $\mathbb{R}$, que llamaremos I_j $(1\leq j\leq N+1)$. Si notamos $J:=I_1\times\ldots\times I_N$, entonces J es un intervalo de $\mathbb{R}^N$. Por hipótesis de inducción, existe una familia finita $\{C_1,\ldots,C_n\}$ de intervalos de $\mathbb{R}^N$, dos a dos disjuntos, tales que

$$\mathbb{R}^N\backslash J=\cup_{i=1}^n C_i;$$

también $\mathbb{R}\backslash I_{N+1}$ se puede expresar como unión de dos intervalos disjuntos, que llamamos J_1,J_2. Entonces, es claro que

$$\mathbb{R}^{N+1}\backslash I=\Big((\mathbb{R}^N\backslash J)\times\mathbb{R}\big)\cup\big(J\times(\mathbb{R}\backslash I_{N+1})\Big)=\cup_{i=1}^n(C_i\times\mathbb{R})\cup(J\times J_1)\cup(J\times J_2).$$

Es fácil comprobar que los conjuntos que aparecen en la descomposición anterior son intervalos de $\mathbb{R}^{N+1}$ disjuntos dos a dos.

La subaditividad del volumen es consecuencia de *i*) y de la monotonía del volumen. En efecto, si notamos

$$J=I_1\cup(I_2\backslash I_1)\cup(I_3\backslash(I_2\cup I_1))\ldots=I_1\cup(I_2\backslash I_1)\cup((I_3\backslash I_2)\backslash I_1)\ldots.$$

se tiene que

$$v(I)\leq v(J)\leq v(I_1)+\cdots+v(I_n)$$

donde se han utilizado la monotonía del volumen y el apartado *ii*) ya que, por ejemplo,

$$I_2=(I_2\setminus I_1)\cup(I_2\cap I_1).$$

∎

10.8. Apéndice C: "Descomposición de un isomorfismo lineal."

Si $T \in \text{Iso}\,(\mathbb{R}^N)$, probaremos que existen dos isometrías $Q_1, Q_2 \in \mathscr{L}(\mathbb{R}^N)$ (norma euclídea) y un operador diagonal D en $\mathbb{R}^N$ con valores propios positivos tales que $Q_1 D Q_2 = T$.

Demostración:
La idea de la demostración está tomada de [RaWe, ∮ 4.5, p. 98]. Antes de iniciar la demostración, es inmediato comprobar la siguiente identidad algebraica para una matriz B cuadrada de orden N:

$$(xB^t|z) = (x|zB), \quad \forall x, z \in \mathbb{R}^N$$

Para comprobar la igualdad, como ambas expresiones definen aplicaciones bilineales, basta comprobar que coinciden sobre el producto de vectores de una base. En efecto, si $1 \leq i, j \leq N$ tenemos

$$(e_i B^t|e_j) = ((b_{1i}, \ldots, b_{Ni})|e_j) = b_{ji} = (e_i|(b_{j1}, \ldots, b_{jN})) = (e_i|e_j B) \qquad (*)$$

Sea A la matriz asociada a T (en términos de la base canónica), esto es,

$$T(x) = xA^t, \; \forall x \in \mathbb{R}^N.$$

Entonces, la matriz $S = A^t A$ es simétrica y no singular por ser T un isomorfismo.

Además, la forma cuadrática asociada a S es definida positiva, ya que, si $x \in \mathbb{R}^N \backslash \{0\}$, usando $(*)$ y que A es una matriz inversible, tenemos

$$(S(x)|x) = (xA^t A|x) = (xA^t|xA^t) = \|xA^t\|_2^2 > 0.$$

Por tanto el operador S tiene N valores propios reales y positivos, que llamamos $\lambda_1, \ldots, \lambda_N$ (algunos pueden tener multiplicidad mayor que 1 y en tal caso aparecen en la lista repetidos tantas veces como indique su orden de multiplicidad). En general, ocurre que vectores propios asociados a distintos valores propios son ortogonales (es muy fácil de comprobar). Si un valor propio tiene multiplicidad $k > 1$, en el subespacio propio asociado, que tiene dimensión k, se puede obtener un sistema ortonormal con k elementos. Así podemos conseguir una base ortonormal $\{y_1, \ldots, y_N\}$ de $\mathbb{R}^N$ tal que

$$S(y_j) = \lambda_j y_j \quad (1 \leq j \leq N).$$

Llamamos $\mu_j = \sqrt{\lambda_j} > 0$. Los operadores lineales que intervendrán en la descomposición de T verifican (por definición)

$$Q_2(y_j) = e_j, \quad D(e_j) = \mu_j e_j, \quad Q_1(e_j) = \frac{1}{\mu_j} T(y_j) \quad (1 \leq j \leq N).$$

Es claro que Q_2 es una isometría , ya que transforma la base ortonormal $\{y_1, \ldots, y_n\}$ en otra ortonormal (la base canónica). También es claro que D es diagonal, con valores propios positivos y que se verifica que $T = Q_1DQ_2$, ya que es ambos operadores lineales coinciden sobre la base $\{e_1,\ldots,e_N\}$.

Sólo resta probar que Q_1 es una isometría y para ello, en vista de la Proposición 7.25, basta comprobar la imagen de la base canónica es una base ortonormal. En efecto, se tiene que

$$(Q_1(e_k)|Q_1(e_j)) = \frac{1}{\mu_k\mu_j}(T(y_k)|T(y_j)) = \frac{1}{\mu_k\mu_j}(y_kA^t|y_jA^t) =$$

$$\frac{1}{\mu_k\mu_j}(y_k|y_jA^tA) = \frac{1}{\mu_k\mu_j}(y_k|S(y_j)) = \frac{\lambda_j}{\mu_k\mu_j}(y_k|y_j) = \delta_{k,j},$$

donde se ha usado $(*)$, que la base es ortonormal y que $\lambda_j = \mu_j^2$ $(1 \leq j \leq N)$. ■

10.9. Apéndice D: "Conjuntos ternarios de Cantor y función singular de Lebesgue"

Probamos que para cada $\alpha \in [0,1[$ existe un conjunto ternario de Cantor C_α de medida α (en particular el conjunto de Cantor original es de medida cero y no numerable). Los conjuntos de Cantor son diseminados (despreciables topológicamente, esto es: compactos de interior vacío) y tienen medida tan cerca a la unidad como se quiera.

Definición 10.23 (Conjuntos ternarios de Cantor)**.** Sea $I_0 = [0,1]$. Tomemos $\rho \in]0, \frac{1}{2}[$, en principio sin restricción adicional.

El primer paso de la construcción de los conjuntos ternarios de Cantor consiste en suprimir del intervalo I_0 el intervalo abierto centrado en el punto medio de I_0 con longitud ρ, intervalo que notaremos $I_{1,1}$.

En el segundo paso se suprimen de cada uno de los dos intervalos que quedan, los intervalos abiertos centrados en el punto medio de cada uno de ellos con longitud ρ^2, intervalos que notaremos $I_{2,1}, I_{2,2}$.

En el tercero se suprimen de cada uno de los cuatro intervalos que quedan, los intervalos abiertos centrados en el punto medio de cada uno de ellos con longitud ρ^3, intervalos que notaremos $I_{3,1}, I_{3,2}, I_{3,3}, I_{3,4}$.

Y así sucesivamente.

Es claro que la suma de las longitudes de los intervalos que hemos suprimido es

$$\rho\left(1 + 2\rho + (2\rho)^2 + \cdots + (2\rho)^n + \cdots\right) .$$

Estaría feo que la suma de las longitudes de los intervalos que quitamos fuese mayor que 1, es decir, hemos de imponer que

$$\frac{\rho}{1-2\rho} \leq 1 \quad \Longleftrightarrow \quad \rho \leq \frac{1}{3} .$$

Resumiendo,

$$\rho \in]0, \frac{1}{3}] .$$

El conjunto ternario de Cantor asociado a ρ y que notaremos C_ρ es el siguiente

$$C_\rho = I_0 \setminus \left(I_{1,1} \cup (I_{2,1} \cup I_{2,2}) \cup (I_{3,1} \cup I_{3,2} \cup I_{3,3} \cup I_{3,4}) \cup \cdots\right) ,$$

cuya medida es claramente

$$\lambda(C_\rho) = \frac{1-3\rho}{1-2\rho} .$$

Habida cuenta que la función

$$\rho \longmapsto g(\rho) = \frac{1-3\rho}{1-2\rho}$$

es continua, estrictamente decreciente y verifica que

$$\lim_{\rho\to 0} g(\rho) = 1 \quad \text{y que} \quad g\big(\frac{1}{3}\big) = 0\ ,$$

deducimos que

$$g\big(]0,\frac{1}{3}]\big) = [0,1[\ .$$

Hemos probado que si $\alpha \in [0,1[$, existe un conjunto ternario de Cantor con medida α . En particular, el conjunto de Cantor original, $C_{\frac{1}{3}}$, que en lo sucesivo lo notaremos simplemente por C es de medida cero.

Probamos ahora que existe una función $f : [0,1] \to [0,1]$ creciente, continua y tal que $f(C) = [0,1]$ (de lo que se deduce que el conjunto de Cantor original C no es numerable). Dicha función se suele utilizar para construir un homeomorfismo de $\mathbb{R}$ sobre $\mathbb{R}$ que no conserva los conjuntos de medida cero (véase el Ejercicio 10.9). Más adelante veremos también que esta función uniformemente continua no es absolutamente continua.

Definición 10.24 (Función singular de Lebesgue)**.** Sean $I_{n,k}$ $(k = 1,2,\ldots,2^{n-1})$ los intervalos que se suprimen en el paso n-ésimo de la construcción del conjunto de Cantor original. Definimos la función singular de Lebesgue $f : [0,1] \to [0,1]$ como

$$f(x) = \begin{cases} 0 & \text{, si } x = 0 \\ \frac{2k-1}{2^n} & \text{, si } x \in I_{n,k} \text{ para } n \in \mathbb{N},\ k \in \{1,\ldots,2^{n-1}\} \end{cases}$$

y

$$f(x) = \sup\{f(t) :\ t \in [0,1] \setminus C\ ,\ t < x\}\ ,\ \forall x \in C \setminus \{0\}\ .$$

El proceso para obtener la expresión de f en el intervalo $I_{n,k}$ se puede describir como sigue: una vez definida f en 0 como 0 y en 1 como 1, en cada uno de los intervalos que se suprimen en la construcción de C se define f como el valor medio de los valores que toma en los puntos más próximos a dicho intervalo para los cuales ya está definida.

Nótese que para $x \in I_{n,k}$ se tiene que

$$f(x) = \frac{2k-1}{2^n} = \frac{\frac{k-1}{2^{n-1}} + \frac{k}{2^{n-1}}}{2}$$

y que por tanto, si k es par el intervalo anterior en el que está ya definida f es $I_{n-1,\frac{k}{2}}$, y si k es impar el intervalo posterior en el que está ya definida f es $I_{n-1,\frac{k+1}{2}}$. En ambos casos el otro intervalo colindante es o bien de la forma $I_{p,q}$ con $1 \leq p \leq n-2$ o bien $\{0\},\{1\}$ si $k = 1,\ k = 2^{n-1}$. Es claro que f es creciente y que

$$f([0,1]) \supset \left\{\frac{2k-1}{2^n} :\ n \in \mathbb{N}\ ,\ k \in \{i,\ldots,2^{n-1}\}\right\}$$

que es un conjunto denso en $[0,1]$. Por otra parte, puesto que f es creciente, si no fuese continua en un punto $a \in [0,1]$, alguno de los intervalos $]f(a-),f(a)[\ ,\ ]f(a),f(a+)[$ no sería vacío y no estaría contenido en $f([0,1])$, contradiciendo la inclusión antes mencionada. Por continuidad f toma en C todos los valores del conjunto anterior, y al ser $f(C)$ compacto, concluimos que $f(C) = [0,1]$. Además es evidente que f es derivable en $[0,1] \setminus C$ con derivada nula en dichos puntos.

Conviene saber que existen funciones de $[0,1]$ en $[0,1]$ continuas y estrictamente crecientes con derivada nula casi por doquier.

10.10. Referencias recomendadas.

[Jur], [Ber], [Guz], [LuMa], [RaWe] y [Ru].

10.11. Resumen del resultados del Tema 10.

En $\mathbb{R}^N$ se define la medida exterior de Lebesgue λ^* por:

$$\lambda^*(A) := \inf\left\{\sum_{n=1}^{\infty} v(I_n) : A \subseteq \cup_{n\in\mathbb{N}} I_n, I_n \in \mathscr{J}, \forall n \in \mathbb{N}\right\}, \ \forall A \subset \mathbb{R}^N$$

donde $\mathscr{J}$ es la familia de los intervalos acotados de $\mathbb{R}^N$. En la anterior definición se pueden tomar los intervalos abiertos.

En $\mathbb{R}^N$ se destacan las siguientes σ-álgebras:

1. La σ-álgebra de Borel, $\mathscr{B}$, generada por $\mathfrak{I}$ (abiertos). También está generada por $\mathscr{F}$ (cerrados), así como por $\mathscr{J}$.

2. La σ-álgebra de Lebesgue, $\mathscr{M}$, definida por

$$\begin{aligned}\mathscr{M} &:= \{B \cup Z : B \text{ boreliano y } \lambda^*(Z) = 0\}\\ &= \{E \subset \mathbb{R}^N : \lambda^*(A) = \lambda^*(A \cap E) + \lambda^*(A \backslash E), \ \forall A \subset \mathbb{R}^N\} =: \mathscr{C}\end{aligned}$$

(la segunda descripción debida a Caratheodory la hemos usado como un instrumento de demostración).

Se verifican las siguientes afirmaciones:

a) $\mathscr{B} \subsetneq \mathscr{M} \subsetneq \mathscr{P}(\mathbb{R}^N)$ (Ejercicios 10.6 y 10.8).

b) $\mathscr{M}$ es la mayor σ-álgebra que contiene $\mathscr{J}$ y sobre la que λ^* es aditiva (Ejercicio 10.4).

c) λ^* no es aditiva en $\mathscr{P}(\mathbb{R}^N) \Leftrightarrow \mathscr{M} \neq \mathscr{P}(\mathbb{R}^N)$ (Ejercicio 10.8).

Sea $E \subset \mathbb{R}^N$. Equivalen las siguientes afirmaciones:

i) $E \in \mathscr{M}$.

ii) $\forall \varepsilon > 0, \exists G \in \mathfrak{I} : E \subset G$ y $\lambda^*(G \backslash E) < \varepsilon$.

iii) $\exists A \in G_\delta$ (intersec. numerable de abiertos) $: E \subset A$ y $\lambda^*(A \backslash E) = 0$.

iv) $\forall \varepsilon > 0, \exists F \in \mathscr{F} : F \subset E$ y $\lambda^*(E \backslash F) < \varepsilon$.

v) $\exists B \in F_\sigma$ (unión numerable de cerrados) y $\lambda^*(E \backslash B) = 0$.

Nótese que v) es un refinamiento de la definición de conjunto medible y que los cerrados se pueden tomar acotados (es decir: compactos).

Teorema de existencia y unicidad de la medida de Lebesgue.
La medida de Lebesgue λ es la restricción de λ^* a $\mathscr{M}$ y es la única medida sobre $\mathscr{M}$ que extiende el volumen de los intervalos acotados.

Sea $E \in \mathscr{M}$, se tiene

$$\lambda(E) = \left\{ \begin{array}{l} \inf\{\lambda(G) : G \text{ abierto}, E \subset G\} \\ \sup\{\lambda(K) : K \text{ compacto}, K \subset E\} \end{array} \right\} \text{regularidad} \left\{ \begin{array}{l} \text{exterior} \\ \text{interior} \end{array} \right\} \text{de } \lambda.$$

Teorema de caracterización de la medida de Lebesgue.

i) *Si μ es una medida sobre $\mathscr{M}$ (resp. $\mathscr{B}$) invariante por traslaciones y tal que $\alpha := \mu([0,1[^N) < \infty$, entonces $\mu = \alpha\lambda$.*

ii) *La medida de Lebesgue es la única medida sobre $\mathscr{M}$ invariante por traslaciones para la que la medida del intervalo $[0,1[^N$ es 1.*

Medida de Lebesgue y homeomorfismos.
Todo homeomorfismo de $\mathbb{R}^N$ sobre $\mathbb{R}^N$ conserva los conjuntos borelianos. Sin embargo existen homeomorfismos de $\mathbb{R}$ sobre $\mathbb{R}$ que no conservan los conjuntos medibles.

Medida de Lebesgue y aplicaciones lineales.
Sea $T : \mathbb{R}^N \to \mathbb{R}^N$ una aplicación lineal. Para todo conjunto medible $E \subset \mathbb{R}^N$ se verifica que el conjunto $T(E)$ es medible y

$$\lambda(T(E)) = |\det T|\lambda(E),$$

donde notamos $\det T$ *al determinante de la matriz asociada a T. En particular si T es una isometría euclídea, entonces*

$$\lambda(T(E)) = \lambda(E), \ \forall E \in \mathscr{M}.$$

Medida de Lebesgue y aplicaciones de clase $\mathscr{C}^1$.
Sean $G \subset \mathbb{R}^N$ abierto y $f : G \to \mathbb{R}^N$ una función de clase $\mathscr{C}^1$. Se verifican las siguientes propiedades:

i) *$f(Z)$ es de medida cero para todo $Z \subset G$ de medida cero.*

ii) *$f(E)$ es medible para todo $E \subset G$ medible.*

Conjunto ternario de Cantor
Es un subconjunto C de $[0,1]$ no numerable, de medida cero y tal que todos sus puntos son de acumulación.

Función singular de Lebesgue
Es una función $f : [0,1] \to [0,1]$ continua y creciente con derivada nula casi por doquier. Además $f(C) = [0,1]$ (de lo que se deduce que C es no numerable).

10.12. Ejercicios del Tema 10

10.1 **(*)** Probar que

i) Dos cubos diádicos o bien son disjuntos o uno de ellos contiene al otro.

ii) Para todo $n \in \mathbb{N}$ se verifica que la familia P_n de los N-cubos diádicos de lado $\frac{1}{2^n}$ constituye una partición de $\mathbb{R}^N$.

10.2 **Teorema de Carathéodory.**

Sea μ^* una medida exterior en un conjunto no vacío Ω. Probar que la familia

$$\mathscr{A} := \{E \subset \Omega : \mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^C),\ \forall A \subset \Omega\}$$

es σ-álgebra y la restricción de μ^* a $\mathscr{A}$ es una medida completa .

10.3 Probar que $\mathscr{M}$ es la mayor σ-álgebra que contiene los intervalos acotados y sobre la que λ^* es aditiva.

10.4 **(*) Expresión analítica del Conjunto de Cantor**

Probar que C es el conjunto de los números reales x que se pueden expresar de la forma

$$x = \sum_{n=1}^{\infty} \frac{\alpha_n}{3^n}, \text{ con } \alpha_n \in \{0,2\}, \forall n \in \mathbb{N}.$$

Probar también que C es compacto y que todos sus puntos son de acumulación.

<u>Indicación</u>: Si para cada natural n notamos $J_{n,k}$ $(k = 1,2,\ldots,2^n)$ los intervalos que quedan en la construcción del conjunto de Cantor, esto es,

$$C = \cap_{n=1}^{\infty} C_n,$$

donde

$$C_n = \bigcup_{k=1}^{2^n} J_{n,k},\ \forall n \in \mathbb{N}.$$

Entonces cada intervalo $J_{n,k}$ tiene la forma

$$\left[\sum_{j=1}^{n} \frac{\alpha_j}{3^j}, \sum_{j=1}^{n} \frac{\alpha_j}{3^j} + \frac{1}{3^n}\right], \text{ con } \alpha_1,\ldots,\alpha_n \in \{0,2\}.$$

Probar también que:

1. $\frac{1}{4}$ está en C y no es extremo de ningún intervalo de los que se suprimen en la construcción de C.
2. C contiene irracionales. Dar un ejemplo.

3. $C+C=[0,2]$

Indicación: Sea $a\in[0,2]$ y sea $\frac{a}{2}=\sum_{n=1}^{\infty}\frac{\alpha_n}{3^n}$, con $\alpha_n\in\{0,1,2\}$. Considerar

$$x=\sum_{n=1}^{\infty}\frac{\beta_n}{3^n}\text{ con }\beta_n=\begin{cases}\alpha_n & \text{si }\alpha_n\in\{0,2\}\\ 2\alpha_n & \text{si }\alpha_n=1\end{cases}\quad;$$

$$y=\sum_{n=1}^{\infty}\frac{\gamma_n}{3^n}\text{ con }\gamma_n=\begin{cases}\alpha_n & \text{si }\alpha_n\in\{0,2\}\\ 0 & \text{si }\alpha_n=1\end{cases}$$

10.5 **(*) Expresión analítica de la función singular de Lebesgue**

Probar que

$$f\left(\sum_{n=1}^{\infty}\frac{\alpha_n}{3^n}\right)=\sum_{n=1}^{\infty}\frac{\alpha_n}{2^{n+1}}.$$

10.6 **Existencia de conjuntos no medibles**.

i) Probar que la familia $\{x+\mathbb{Q}:x\in\mathbb{R}\}$ es una partición de $\mathbb{R}$.
Indicación: La relación $x\simeq y$ si $y-x\in\mathbb{Q}$ es una relación de equivalencia en $\mathbb{R}$.

ii) Pongamos $\{x+\mathbb{Q}:x\in\mathbb{R}\}=\{A_i:i\in I\}$ ($A_i\neq A_j$, para $i\neq j$) y para cada $i\in I$ sea $x_i\in A_i\cap[0,1]$. Probar que el conjunto $E=\{x_i:i\in I\}$ no es medible.
Indicación: Si $\{q_n:n\in\mathbb{N}\}$ es una numeración de $[-1,1]\cap\mathbb{Q}$, entonces $[0,1]\subseteq\cup_{n=1}^{\infty}(q_n+E)\subset[-1,2]$ y los conjuntos q_n+E son disjuntos entre sí.

iii) Probar que cualquier subconjunto medible de E tiene medida cero.
Indicación: Si $A\subset E$, entonces $\cup_{n=1}^{\infty}(q_n+A)\subset[-1,2]$ y los conjuntos q_n+A son disjuntos entre sí.

iv) Sea $M\subset\mathbb{R}$ con $\lambda^*(M)>0$. Probar que M contiene un subconjunto no medible.
Indicación: $M=\cup_{q\in\mathbb{Q}}M\cap(q+E)$.

10.7 Probar que la existencia de conjuntos no medibles equivale a la no aditividad de λ^*.

10.8 Probar que existe un conjunto $Z\subset\mathbb{R}$ de medida cero que no es boreliano y existe un homeomorfismo φ de $\mathbb{R}$ sobre $\mathbb{R}$ tal que $\varphi(Z)$ no es medible.

Indicación: Prolongar la función singular de Lebesgue, f, por

$$f(x)=\begin{cases}0 & \text{si }x\le 0\\ 1 & \text{si }x\ge 1\end{cases}$$

y considerar la función $\varphi:\mathbb{R}\to\mathbb{R}$ definida por

$$\varphi(x)=x+f(x),\ \forall x\in\mathbb{R}.$$

Probar que $\lambda(\varphi([0,1]\backslash C))=1$ y deducir de ello que $\lambda(\varphi(C))=1$. Sea $E\subset\varphi(C)$ no medible. Tomar $Z:=\varphi^{-1}(E)$.

10.9 Sean A un abierto de $\mathbb{R}^N$ y $f : A \to \mathbb{R}^M$ una función de clase $\mathscr{C}^1$ con $N < M$. Probar que $f(A)$ es de medida cero.

10.10 Toda variedad diferenciable en $\mathbb{R}^N$ de dimensión $k < N$ es de medida cero.

10.11 Sean $x_1, x_2, \ldots, x_N, a \in \mathbb{R}^N$ y

$$P := \left\{ x \in \mathbb{R}^N : x = a + \sum_{n=1}^{N} t_n x_n, t_n \in [0,1], 1 \leq n \leq N \right\}$$

(N-paralelogramo en $\mathbb{R}^N$, generalización del paralelogramo en $\mathbb{R}^2$ y del paralelepípedo en $\mathbb{R}^3$). Probar que

$$\lambda(P) = |\det\,(x_1, x_2, \ldots, x_N)|.$$

Deducir de lo anterior el área del triángulo.

10.12 **(*)** Probar que toda aplicación $f : \mathbb{R}^N \to \mathbb{R}^N$ lipschitziana para cualquier norma conserva los conjuntos de medida cero y los conjuntos medibles.

Indicación: Aplíquense el Teorema de Haussdorf, la Proposición 10.20 y la prueba del apartado *ii*) de la Proposición 10.22.

10.13 **(*) Medida interior. Medibilidad original de Lebesgue.**

Recordamos que para cada $A \subset \mathbb{R}^N$ la medida exterior de A, $\lambda^*(A)$, verifica

$$\lambda^*(A) = \inf\{\lambda(G) : G \text{ abierto } \; A \subset G\}.$$

Definimos ahora la medida interior de A, $\lambda_*(A)$, como sigue

$$\lambda_*(A) := \sup\{\lambda(K) : K \text{ compacto } \; K \subset A\}.$$

Sabemos también que si A es un conjunto medible, entonces

$$\lambda_*(A) = \lambda^*(A)(= \lambda(A)).$$

Probar

1. $\lambda_*(A) \leq \lambda^*(A),\ \forall A \subset \mathbb{R}^N$.
2. Que existe $A \subset \mathbb{R}$ no medible tal que $\lambda_*(A) = \lambda^*(A)$.
   Indicación: considerar un conjunto no medible del intervalo $[0,1]$.
3. Si $\lambda^*(A) < \infty$, Entonces A es medible si, y sólo si, $\lambda_*(A) = \lambda^*(A)$.
4. A es medible si, y sólo si,

$$\lambda_*(A \cap B(0,n)) = \lambda^*(A \cap B(0,n)),\ \forall n \in \mathbb{N}.$$

5. Si A es medible y $B \subset A$, entonces

$$\lambda(A) = \lambda_*(B) + \lambda^*(A \setminus B).$$

6. Definición de Lebesgue de medibilidad. Sea A acotado y sea I un intervalo acotado que contiene a A

   a) Comprobar que
   $$\lambda_*(A) = v(I) - \lambda^*(I \setminus A).$$
   El número $v(I) - \lambda^*(I \setminus A)$ es la definición de medida interior de A dada por Lebesgue (obsérvese la independencia del intervalo I elegido).

   b) Lebesgue definió la medibilidad de un conjunto A acotado por la propiedad
   $$v(I) - \lambda^*(I \setminus A) = \lambda^*(A),$$
   donde I es un intervalo acotado que contiene a A. Comprobar que un conjunto acotado es medible en el sentido anterior si, y sólo si, es medible.

   c) Lebesgue definió también la medibilidad de un conjunto A cualquiera por la propiedad de que $A \cap I$ fuese medible para cualquier intervalo acotado I. Comprobar que esta medibilidad coincide con la medibilidad.

10.14 **(*)** Probar que el producto cartesiano de conjuntos medibles es medible.

Indicación: Sean p, q naturales fijos. Probar que

i) Si $Z \subset \mathbb{R}^p$ es de medida cero, y k es un natural, entonces $Z \times [-k, k]^q$ es un conjunto de medida cero en $\mathbb{R}^{p+q}$. Deducir que también es de medida cero el conjunto $Z \times \mathbb{R}^q$.

ii) Si $E \subset \mathbb{R}^p$ es medible y $C \subset \mathbb{R}^q$ es cerrado, entonces $E \times C$ es medible.

iii) Si $E \subset \mathbb{R}^p$ es medible y $B \subset \mathbb{R}^q$ es un F_σ, entonces $E \times B$ es medible.

10.13. Soluciones a los ejercicios del Tema 10.

10.1 *i)* Sean $Q_1 = Q\left(a, \frac{1}{2^n}\right) \in P_n, Q_2 = Q\left(b, \frac{1}{2^m}\right) \in P_m$ con intersección no vacía. Supongamos por ejemplo que $n \le m$ y sea $z \in Q_1 \cap Q_2$. Notando a_j, b_j, z_j la coordenada j-ésima de a, b, z respectivamente, tenemos que por ser $z \in Q_1 \cap Q_2$ se verifica para $1 \le j \le d$

$$\max\{a_j, b_j\} \le z_j < \min\left\{a_j + \frac{1}{2^n}, b_j + \frac{1}{2^m}\right\} \qquad (*)$$

en particular $a_j \le z_j < b_j + \frac{1}{2^m}$ por lo que $2^m a_j < 2^m b_j + 1$ y, puesto que $2^m a_j$ y $2^m b_j$ son números enteros, se tiene que $2^m a_j \le 2^m b_j$, es decir, $a_j \le b_j$, para $1 \le j \le d$, lo que junto con (*) implica que $b \in Q_1$.

Si es $n = m$ razonando igual con la desigualdad

$$b_j \le z_j < a_j + \frac{1}{2^n}$$

se obtiene que $b_j \le a_j, 1 \le j \le d$, por lo que $a = b$ y, en este caso $Q_1 = Q_2$.

Si $n < m$, como $b \in Q_1$ se tiene $0 \le b_j - a_j < \frac{1}{2^n}$ por lo que $0 \le 2^m(b_j - a_j) < \frac{2^m}{2^n}$ y, como ésta es una desigualdad entre números enteros, se deduce que $0 \le 2^m(b_j - a_j) \le \frac{2^m}{2^n} - 1$ esto es, $0 \le b_j - a_j \le \frac{1}{2^n} - \frac{1}{2^m}$ para $1 \le j \le d$. Si ahora $x \in Q_2$ se tiene $0 \le x_j - b_j < \frac{1}{2^m}$ para $1 \le j \le d$, y sumando esta desigualdad con la anterior obtenemos que $0 \le x_j - a_j < \frac{1}{2^n}$, para $1 \le j \le d$, es decir $x \in Q_1$, por lo que $Q_2 \subset Q_1$.

ii) Después de lo visto en el apartado $i)$ basta tener en cuenta que dado $x \in \mathbb{R}^N$ es $x \in Q\left(\frac{\mathscr{E}(2^n x)}{2^n}, \frac{1}{2^n}\right) \in P_n$, donde para $x = (x_1, x_2, \ldots, x_N) \in \mathbb{R}^N$ hemos notado

$$\mathscr{E}(x) = (E(x_1), E(x_2), \ldots, E(x_N))$$

siendo E la función "parte entera".

10.2 El teorema de Carathéodory está hecho en teoría salvo que la medida μ^* es completa, esto es

$$[Z \in \mathscr{A}, \mu(Z) = 0, B \subset Z] \Leftrightarrow B \in \mathscr{A}.$$

En efecto sea $A \subset \Omega$, como $\mu^*(B) = 0$ se tiene

$$\mu^*(A \cap B) + \mu^*(A \backslash B) \le \mu^*(B) + \mu^*(A) = \mu^*(A),$$

donde se utilizado la monotonía de la medida exterior. La subaditividad de μ^* nos permite concluir que la anterior desigualdad es de hecho una igualdad.

10.3 Sea $\mathscr{A}$ una σ-álgebra que contiene a los intervalos acotados y sobre la que la medida exterior es aditiva. Por contener los intervalos contiene a la σ-álgebra $\mathscr{B}$ de los borelianos. Como la σ-álgebra $\mathscr{M}$ la hemos caracterizado por $E \in \mathscr{M}$ si, y sólo si

$$\lambda^*(A) \geq \lambda^*(A \cap E) + \lambda^*(A \cap E^C), \;\; \forall A \subset \mathbb{R}^N,$$

bastará probar que todo conjunto $E \in \mathscr{A}$ cumple tal desigualdad para concluir que $\mathscr{A} \subset \mathscr{M}$, es decir $\mathscr{M}$ es la mayor σ-álgebra que contiene los intervalos acotados y sobre la que λ^* es aditiva. Sea $E \in \mathscr{A}$. La regularidad de la medida exterior de Lebesgue nos asegura que para cada $A \subset \mathbb{R}^N$ existe B boreliano tal que $A \subset B$ y $\lambda^*(A) = \lambda^*(B)$. Así para cada $A \subset \mathbb{R}^N$, se tiene

$$\lambda^*(A) = \lambda^*(B) = \lambda^*\Big((B \cap E) \cup (B \cap E^C)\Big) =$$

$$\lambda^*(B \cap E) + \lambda^*(B \cap E^C) \geq \lambda^*(A \cap E) + \lambda^*(A \cap E^C),$$

donde se ha utilizado que $B \cap E, B \cap E^C \in \mathscr{A}$.

10.4 Fijados $\alpha_j \in \{0,2\}, \;\; \forall j \in \mathbb{N}$, es claro que la serie $\sum_{j \geq 1} \frac{\alpha_j}{3^j}$ es convergente (es de términos positivos y sumas parciales acotadas por 1). Así

$$\sum_{j=1}^{\infty} \frac{\alpha_j}{3^j} \in [0,1].$$

Veamos que para cada $n \in \mathbb{N}$ y para cada $k = 1, \ldots, 2^n$ los intervalos $J_{n,k}$ tienen la forma

$$\left[\sum_{j=1}^{n} \frac{\alpha_j}{3^j}, \sum_{j=1}^{n} \frac{\alpha_j}{3^j} + \frac{1}{3^n}\right], \quad \text{con } \alpha_1, \ldots, \alpha_n \in \{0,2\}.$$

Para $n = 1$ se verifican las expresiones. En efecto

$$J_{1,1} = \left[0, \frac{1}{3}\right] = \left[\frac{0}{3}, \frac{1}{3}\right] \text{ y } J_{1,2} = \left[\frac{2}{3}, 1\right] = \left[\frac{2}{3}, \frac{2}{3} + \frac{1}{3}\right]. \tag{1}$$

Supongamos que las expresiones son ciertas para n y las probaremos para $n+1$. Sea

$$J_{n,k} = \left[a, a + \frac{1}{3^n}\right[. \tag{2}$$

Dicho intervalo da lugar a dos: su tercio izquierda

$$J_{n+1,2k-1} = \left[a, a + \frac{1}{3^{n+1}}\right] = \left[a + \frac{0}{3^{n+1}}, a + \frac{1}{3^{n+1}}\right]$$

y su tercio derecha

$$J_{n+1,2k} = \left[a + \frac{2}{3^{n+1}}, a + \frac{1}{3^n}\right] = \left[a + \frac{2}{3^{n+1}}, a + \frac{2}{3^{n+1}} + \frac{1}{3^{n+1}}\right],$$

ya que

$$\frac{1}{3^n} = \frac{2}{3^{n+1}} + \frac{1}{3^{n+1}}.$$

En virtud de las expresiones anteriores y teniendo en cuenta que, para cada $n \in \mathbb{N}$,

$$\frac{1}{3^n} = \frac{2}{3^{n+1}} + \frac{2}{3^{n+2}} + \dots$$

concluimos que

$$x = \sum_{j=1}^{\infty} \frac{\alpha_j}{3^j} \text{ con } \alpha_j \in \{0,2\} \Leftrightarrow x \in C.$$

En efecto. Sea $\{\alpha_j\}$ con $\alpha_j \in \{0,2\}, \forall j \in \mathbb{N}$ y sea $x = \sum_{j=1}^{\infty} \frac{\alpha_j}{3^j}$. Para cada $n \in \mathbb{N}$ se tiene

$$\sum_{j=1}^{n} \frac{\alpha_j}{3^j} \leq x = \sum_{j=1}^{n} \frac{\alpha_j}{3^j} + \sum_{j=n+1}^{\infty} \frac{\alpha_j}{3^j} \leq \sum_{j=1}^{n} \frac{\alpha_j}{3^j} + \sum_{j=n+1}^{\infty} \frac{2}{3^j} = \sum_{j=1}^{n} \frac{\alpha_j}{3^j} + \frac{1}{3^n},$$

lo que implica que $x \in C_n, \forall n \in \mathbb{N}$, esto es $x \in C$.

En cuanto a la otra implicación, fijado $x \in C$, se construye inductivamente una sucesión $\{\alpha_j\}$ con $\alpha_j \in \{0,2\}, \forall j \in \mathbb{N}$ tal que

$$\sum_{j=1}^{n} \frac{\alpha_j}{3^j} \leq x \leq \sum_{j=1}^{n} \frac{\alpha_j}{3^j} + \frac{1}{3^n}, \forall n \in \mathbb{N}.$$

Definimos

$$\alpha_1 = \begin{cases} 0 & \text{si } x \in J_{1,1} \\ 2 & \text{si } x \in J_{1,2} \end{cases} \text{ (ver (1))}.$$

Supuesto definidos $\alpha_1, \dots, \alpha_n$, si tomamos en (2) $a = \sum_{j=1}^{n} \frac{\alpha_j}{3^j}$, entonces $x \in J_{n,k}$ y definimos

$$\alpha_{n+1} = \begin{cases} 0 & \text{si } x \in J_{n+1,2k-1} \\ 2 & \text{si } x \in J_{n+1,2k} \end{cases}.$$

Es claro que C es cerrado y acotado. Veamos que todos sus puntos son de acumulación. Sea $x \in C$. Para cada $n \in \mathbb{N}$ existe $k \in \mathbb{N}$ con $1 \leq k \leq 2^n$ tal que $x \in J_{n,k} = [a_n, b_n]$. Sea c_n uno de los extremos de $J_{n,k}$ tal que $x \neq c_n$. Entonces $\{c_n\} \to x$.

Demostremos ahora los últimos tres puntos:

1. $$\frac{1}{4} = \frac{2}{3^2} + \frac{2}{3^4} + \dots$$

2. $$a = \frac{2}{3} + \frac{0}{3^2} + \frac{2}{3^3} + \frac{0}{3^4} + \frac{0}{3^5} + \frac{2}{3^6} + \dots$$

3. Está hecho en la indicación, pues es claro que $x, y \in C$ y $\frac{a}{2} = \frac{x}{2} + \frac{y}{2}$.

10.5 Veamos ahora que

$$f\left(\sum_{n=1}^{\infty}\frac{\alpha_n}{3^n}\right)=\sum_{n=1}^{\infty}\frac{\alpha_n}{2^{n+1}}.$$

Para cada sucesión $\{\alpha_n\}$ con $\alpha_n\in\{0,2\},\forall n\in\mathbb{N}$, notaremos

$$\sum_{n=1}^{\infty}\frac{\alpha_n}{2^{n+1}}:=g\left(\sum_{n=1}^{\infty}\frac{\alpha_n}{3^n}\right).$$

Empezaremos probando que la expresión es cierta para los extremos de los intervalos $J_{n,k}$ con n natural y $k=1,2,\ldots,2^n$. Para $n=1$ es cierta. En efecto:

$$J_{1,1}\begin{cases} f(0)=0 \text{ y } g(0)=0\\ f\left(\frac{1}{3}\right)=\frac{1}{2} \text{ y } g\left(\frac{1}{3}\right)=g\left(\frac{2}{3^2}+\frac{2}{3^3}+\ldots\right)=\frac{1}{2^2}+\frac{1}{2^3}+\ldots=\frac{1}{2}\end{cases}$$

y

$$J_{1,2}\begin{cases} f\left(\frac{2}{3}\right)=\frac{1}{2} \text{ y } g\left(\frac{2}{3}\right)=\frac{1}{2}\\ f(1)=1 \text{ y } g(1)=g\left(\frac{2}{3}+\frac{2}{3^2}+\ldots\right)=\frac{1}{2}+\frac{2}{2^2}+\ldots=1\end{cases}$$

$$J_{1,2}\begin{cases} f\left(\frac{2}{3}\right)=\frac{1}{2} \text{ y } g\left(\frac{2}{3}\right)=\frac{1}{2}\\ f(1)=1 \text{ y } g(1)=g\left(\frac{2}{3}+\frac{2}{3^2}+\ldots\right)=\frac{1}{2}+\frac{2}{2^2}+\ldots=1\end{cases}$$

Supongamos que las expresiones son ciertas para n y las probaremos para $n+1$. Sea

$$J_{n,k}=\left[a,a+\frac{1}{3^n}\right].$$

Tan sólo hemos de comprobar que

$$g\left(a+\frac{1}{3^{n+1}}\right)=f\left(a+\frac{1}{3^{n+1}}\right)\left[:=\frac{f(a)+f\left(a+\frac{1}{3^n}\right)}{2}=:\right]$$

$$=f\left(a+\frac{2}{3^{n+1}}\right)=g\left(a+\frac{2}{3^{n+1}}\right)$$

Teniendo en cuenta la hipótesis de inducción obtenemos que

$$f\left(a+\frac{1}{3^n}\right)=g\left(a+\frac{1}{3^n}\right)=g\left(a+\frac{2}{2^{n+1}}+\frac{2}{2^{n+2}}+\ldots\right)=$$

$$g(a)+\frac{1}{2^{n+1}}+\frac{1}{2^{n+2}}+\ldots=g(a)+\frac{1}{2^n}=f(a)+\frac{1}{2^n},$$

y en consecuencia

$$f\left(a+\frac{1}{3^{n+1}}\right)=f\left(a+\frac{2}{3^{n+1}}\right)=$$

$$\frac{f(a)+f\left(a+\frac{1}{3^n}\right)}{2}=\frac{f(a)+f(a)+\frac{1}{2^n}}{2}=f(a)+\frac{1}{2^{n+1}}.$$

Calculemos ahora, utilizando nuevamente la hipótesis de inducción, los valores de g en dichos puntos:

$$g\left(a+\frac{1}{3^{n+1}}\right)=g\left(a+\frac{2}{3^{n+2}}+\frac{2}{3^{n+3}}+\dots\right)=$$

$$g(a)+\frac{1}{2^{n+2}}+\frac{1}{2^{n+3}}+\dots=g(a)+\frac{1}{2^{n+1}}=f(a)+\frac{1}{2^{n+1}}$$

y

$$g\left(a+\frac{2}{3^{n+1}}\right)=g(a)+\frac{1}{2^{n+1}}=f(a)+\frac{1}{2^{n+1}}.$$

Por último sea x un elemento del conjunto de Cantor C, y sea

$$x=\sum_{n=1}^{\infty}\frac{\alpha_n}{3^n}\text{ con }\alpha_n\in\{0,2\}$$

Si para cada natural n notamos $x_n=\sum_{k=1}^{n}\frac{\alpha_k}{3^k}$, es claro que $\{x_n\}\to x$ y que cada x_n es un extremo izquierdo de un intervalo $J_{n,k}$, luego para dichos puntos vale la fórmula, esto es

$$f(x_n)=\sum_{k=1}^{n}\frac{\alpha_k}{2^{k+1}}.$$

Por continuidad

$$f(x)=\lim f(x_n)=\sum_{n=1}^{\infty}\frac{\alpha_n}{2^{n+1}}.$$

10.6 *i)* La relación $x\simeq y\Leftrightarrow y-x\in\mathbb{Q}$ es una relación de equivalencia y la clase de equivalencia de x es $x+\mathbb{Q}$.

ii) Como para cada $i\in I$ se tiene que

$$A_i\cap[0,1]\neq\emptyset$$

(de hecho $\overline{A_i}=\overline{x+\mathbb{Q}}=\mathbb{R}$), podemos escribir

$$A_i=x_i+\mathbb{Q},\text{ con }x_i\in[0,1].$$

Al ser $E=\{x_i : i\in I\}$, se tiene que $E\subset[0,1]$.
Veamos que $[0,1]\subset\cup_{n=1}^{\infty}(q_n+E)$. En efecto: $x\in[0,1]\Rightarrow x+Q=x_i+Q$, para conveniente $i\in I$. En consecuencia $x-x_i\in\mathbb{Q}\cap[-1,1]$ y por tanto $x-x_i=q_n$, para conveniente $n\in\mathbb{N}$, luego $x\in q_n+E$.
Veamos ahora que los conjuntos q_n+E son disjuntos:

$$(q_n+E)\cap(q_m+E)\neq\emptyset\Rightarrow q_n+x_i=q_m+x_j\Rightarrow i=j\text{ (partición)}\Rightarrow n=m.$$

La inclusión $\cup_{n=1}^{\infty}(q_n+E)\subset[-1,2]$ es obvia.

Finalmente, si E fuese medible, entonces sería

$$1 = \lambda([0,1]) \leq \sum_{n=1}^{\infty} \lambda(q_n + E) \leq \lambda([-1,2]) = 3$$

y la invarianza por traslaciones nos da el absurdo pues $\lambda(E)$ no puede ser 0 ni mayor que 0.

iii) Los $q_n + A$ son disjuntos entre sí, pues lo son los $q_n + E$.

iv) $\{q + E : q \in \mathbb{Q}\}$ es una partición numerable de $\mathbb{R}$ (se prueba como en *ii*)). E es no medible, luego los $q + E$ son también no medibles y los únicos subconjuntos de $q + E$ que son medibles son los de medida cero. Se tiene

$$0 < \lambda^*(M) = \lambda^*\left(\bigcup_{q\in\mathbb{Q}} M \cap (q + E)\right) \leq \sum_{n=1}^{\infty} \lambda^*(M \cap (q_n + E)).$$

Si los $M \cap (q_n + E)$ fuesen todos medibles, serían de medida cero y en consecuencia

$$\sum_{n=1}^{\infty} \lambda^*(M \cap (q_n + E)) = 0$$

lo que es absurdo. Hemos probado que M tiene subconjuntos no medibles.

10.7 Sabemos que $\mathscr{M} = \mathscr{C}$, esto es

$$E \in \mathscr{M} \Leftrightarrow \lambda^*(A) \geq \lambda^*(A \cap E) + \lambda^*(A \cap E^C), \ \ \forall A \subset \mathbb{R}^N.$$

Si $E \subset \mathbb{R}^N$ es no medible, entonces existe un subconjunto A de $\mathbb{R}^N$ tal que

$$\lambda^*((A \cap E) \cup (A \cap E^C) = \lambda^*(A) < \lambda^*(A \cap E) + \lambda^*(A \cap E^C)$$

y λ^* no es aditiva.

Supongamos que λ^* no es aditiva. Sean A, B subconjuntos disjuntos de $\mathbb{R}^N$ tales que

$$\lambda^*(A \cup B) < \lambda^*(A) + \lambda^*(B).$$

Se tiene entonces que

$$\lambda^*(A \cup B) < \lambda^*((A \cup B) \cap A) + \lambda^*((A \cup B) \cap A^C)$$

y el conjunto A no es medible.

10.8 Es claro que φ es un homeomorfismo estrictamente creciente de $\mathbb{R}$ sobre $\mathbb{R}$, así como que

$$\varphi([0,1]) = [0,2]$$

y en consecuencia

$$\varphi([0,1]\backslash C) = [0,2]\backslash \varphi(C).$$

Se tiene que

$$[0,1]\backslash C = \cup_{n=1}^{\infty} \cup_{k=1}^{2^{n-1}} I_{n,k}$$

y por tanto

$$\varphi([0,1]\backslash C) = \cup_{n=1}^{\infty} \cup_{k=1}^{2^{n-1}} \varphi(I_{n,k})$$

de donde se deduce

$$2 - \lambda(\varphi(C)) = \lambda([0,2]\backslash \varphi(C)) = \sum_{n=1}^{\infty} \sum_{k=1}^{2^{n-1}} \lambda(\varphi(I_{n,k})).$$

Teniendo en cuenta ahora que para cada $n \in \mathbb{N}$ y $k \in \{1, \dots, 2^{n-1}\}$ es

$$f(x) = \frac{2k-1}{2^n}, \forall x \in I_{n,k}$$

obtenemos que

$$2 - \lambda(\varphi(C)) = \sum_{n=1}^{\infty} \sum_{k=1}^{2^{n-1}} \lambda\left(\frac{2k-1}{2^n} + I_{n,k}\right) = \sum_{n=1}^{\infty} \sum_{k=1}^{2^{n-1}} \lambda(I_{n,k}),$$

y en consecuencia

$$2 - \lambda(\varphi(C)) = \sum_{n=1}^{\infty} \sum_{k=1}^{2^{n-1}} \frac{1}{3^n} = \sum_{n=1}^{\infty} \frac{2^{n-1}}{3^n} = \frac{1}{2} \sum_{n=1}^{\infty} \left(\frac{2}{3}\right)^n = 1.$$

Hemos probado que

$$\lambda(\varphi(C)) = 1.$$

Sea $E \subset \varphi(C)$ no medible (ejercicio 10.6) y definamos $Z = \varphi^{-1}(E)$. Z es de medida cero (es un subconjunto de C) y no es boreliano (si lo fuese, también lo sería $\varphi(Z)$). Hemos probado que los homeomorfismos no conservan los conjuntos medibles.

10.9 Definamos $B := A \times \mathbb{R}^{M-N} \subseteq \mathbb{R}^M$ y $g : B \to \mathbb{R}^M$ por

$$g(x,y) = f(x), \ \forall (x,y) \in B.$$

El conjunto B es abierto de $\mathbb{R}^M$ y g es de clase $\mathscr{C}^1$ en B. El conjunto $A \times \{0\} \subset B$ es de medida cero en $\mathbb{R}^M$ por estar contenido en un hiperplano, en consecuencia, al aplicar las funciones de clase $\mathscr{C}^1$ conjuntos de medida cero en conjuntos de medida cero, se sigue que $g(A \times \{0\}) = f(A)$ es de medida cero.

10.10 Sea M una variedad diferenciable en $\mathbb{R}^N$ de dimensión $k < N$. Sabemos que dado $x \in M$, existe $\rho > 0$, $U \subset \mathbb{R}^k$ abierto y $p : U \to \mathbb{R}^N$ de clase $\mathscr{C}^1$ tal que $B(x,\rho) \cap M \subset p(U)$, y como $p(U)$ es de medida cero (ejercicio anterior) se sigue que $B(x,\rho) \cap M$ es de medida cero en $\mathbb{R}^N$. La bola $B(x,\rho)$ contiene una bola B con centro y radio racionales tal que $x \in B$; la familia $\mathscr{B}$ de tales bolas es numerable. Ya que

$$M = \cup_{B \in \mathscr{B}} (B \cap M) \text{ y } B \cap M \text{ tiene medida cero,}$$

se sigue que M tiene medida cero.

10.11 Consideremos la aplicación lineal $T : \mathbb{R}^N \to \mathbb{R}^N$ dada por

$$T(e_k) = x_k, \quad \forall k \in \{1, 2 \ldots, N\}.$$

Es claro que si notamos $I = [0,1]^N$, entonces $T(I) = P - a$. En consecuencia

$$\lambda(P) = \lambda(P - a) = \lambda(T(I)) = |\det T| \lambda(I) = |\det T| = |\det (x_1, x_2, \ldots, x_N)|,$$

donde se ha utilizado que la medida de Lebesgue es invariante por traslaciones y el comportamiento de la medida de Lebesgue frente a las aplicaciones lineales.

Probemos finalmente la fórmula del área del triángulo. Es claro que el área de un triángulo A es la mitad del área del paralelogramo con lo que

$$A = \frac{1}{2} |\det (x,y)|.$$

Como podemos trasladar y girar, no quita generalidad suponer que $x = (b,0)$ e $y = (a,h)$ con $b, h > 0$. Así

$$A = \frac{1}{2} \begin{vmatrix} b & 0 \\ a & h \end{vmatrix} = \frac{bh}{2}.$$

10.12 Del Teorema de Hausdorff se deduce que f es también lipschitziana para $\|\cdot\|_2$. La Proposición 10.20 nos asegura que f conserva los conjuntos de medida cero. Ahora, la demostración del apartado *ii*) de la Proposición 10.22 nos dice que f conserva los conjuntos medibles.

10.13

1. Sean $K \subset A \subset G$, K compacto y G abierto. La monotonía de la medida de Lebesgue nos asegura que $\lambda(K) \leq \lambda(G)$. Así $\lambda(K) \leq \lambda^*(A)$ y también $\lambda_*(A) \leq \lambda^*(A)$.
2. Sea E un subconjunto no medible del intervalo $[0,1]$ (véase ejercicio 10.6). El conjunto $A = E \cup]1, +\infty[$ es no medible y $\lambda_*(A) = \infty$.
3. Basta probar que si $\lambda_*(A) = \lambda^*(A) < \infty$, entonces A es medible. Sea $\varepsilon > 0$, tomamos $K \subset A \subset G$ con

$$\lambda_*(A) - \frac{\varepsilon}{2} < \lambda(K) \quad \text{y} \quad \lambda(G) < \lambda^*(A) + \frac{\varepsilon}{2}.$$

Se tiene que

$$\lambda_*(G \setminus A) \leq \lambda_*(G \setminus K) \leq \lambda(G \setminus K) = \lambda(G) - \lambda(K) < \varepsilon,$$

y esto es equivalente a la medibilidad de A.

4. Acabamos de probar que para cada $n \in \mathbb{N}$ el conjunto $A \cap B(0,n)$ es medible y $A = \bigcup_{n=1}^{\infty}(A \cap B(0,n))$.
5. Sea C medible con $A \setminus B \subset C \subset A$ con $\lambda^*(A \setminus B) = \lambda(C)$. Se tiene
$$\lambda(A) = \lambda(C) + \lambda(A \setminus C) = \lambda^*(A \setminus B) + \lambda(A \setminus C) \le \lambda^*(A \setminus B) + \lambda_*(B),$$
donde se ha utilizado que $A \setminus C \subset B$.
Sea ahora $\varepsilon > 0$ y sea $K \subset B$ compacto tal que $\lambda_*(B) - \varepsilon < \lambda(K)$. Se tiene que
$$\lambda(A) = \lambda(K) + \lambda(A \setminus K) > \lambda_*(B) - \varepsilon + \lambda^*(A \setminus B).$$
Así
$$\lambda(A) \ge \lambda_*(B) + \lambda^*(A \setminus B).$$
6. *a*) Es consecuencia del apartado 5. sin más que cambiar A por I y B por A.
   b) Como $\lambda^*(A) < \infty$, sabemos (apartado 3.), que es medible si, y sólo si, $\lambda_*(A) = \lambda^*(A)$, lo que coincide con la definición de Lebesgue.
   c) Es consecuencia del apartado 4.

10.14 Sean A, B conjuntos medibles acotados de $\mathbb{R}^p, \mathbb{R}^q$ respectivamente. Queremos probar que $A \times B$ es un conjunto medible de $\mathbb{R}^{p+q}$. En efecto, sean I, J intervalos acotados de $\mathbb{R}^p, \mathbb{R}^q$ respectivamente tales que $A \subset I, B \subset J$. Se tiene que $A \times B \subset I \times J$

i) Basta observar que si $Z \subset \cup_{n=1}^{\infty} I_n$, entonces
$$Z \times [-k,k]^q \subset \cup_{n=1}^{\infty}(I_n \times [-k,k]^q) := \cup_{n=1}^{\infty} J_n \qquad \text{y} \qquad v(J_n) = (2k)^q v(I_n),$$
y que la unión numerable de conjuntos de medida cero lo es.

ii) Al ser E medible se puede expresar como
$$E = \cup_{n=1}^{\infty} C_n \cup Z$$
donde para cada n, C_n es cerrado de $\mathbb{R}^p$, y Z es de medida cero. Se sigue que
$$E \times C = \cup_{n=1}^{\infty}(C_n \times C) \cup (Z \times C).$$
Para cada n natural $C_n \times C$ es cerrado de $\mathbb{R}^{p+q}$, y, en virtud de *i*), $Z \times C$ es de medida cero como subconjunto de $Z \times \mathbb{R}^q$, por tanto $E \times C$ es medible.

iii) Es consecuencia de *ii*), de la distributividad del producto cartesiano y de que la unión numerable de medibles es medible.
Sea finalmente F medible de $\mathbb{R}^q$. Existen B un F_σ de $\mathbb{R}^q$ y Z un conjunto de medida cero de $\mathbb{R}^q$ tales que $F = B \cup Z$. Se tiene
$$E \times F = (E \times B) \cup (E \times Z)$$
y basta aplicar *i*), *iii*) y que la unión de dos medibles es medible.

Tema 11

Integral asociada a una medida

Hacia finales del siglo XIX resultó claro para muchos matemáticos que la integral de Riemann debía cambiarse por algún tipo de integral capaz de integrar más funciones y sobre todo mejor avenida con los procesos de convergencia. Se hacía necesaria una integral que tuviese teoremas de convergencia que, en condiciones muy generales, permitiesen intercambiar el límite y la integral. Entre los intentos hechos en esta dirección, los más notables fueron debidos a Jordan, Borel, Young y Lebesgue. La construcción de Lebesgue resultó ser la más exitosa.

Esquemáticamente, he aquí la idea principal:

La teoría de la integral de Riemann tiene como punto de partida la consideración para cada función f acotada sobre un intervalo $[a,b]$ de las llamadas sumas de Riemann

$$\sum_{j=1}^{m} f(x_j)\lambda(I_j),$$

donde $I_1,\ldots,I_m$ son intervalos disjuntos cuya unión es $[a,b]$; $x_j \in I_j$, y $\lambda(I_j)$ denota, la longitud de I_j $(1 \leq j \leq m)$.

Lebesgue descubrió que se obtenía una teoría de la integración completamente satisfactoria si se permitía que los conjuntos I_j de la suma anterior pertenecieran a una clase más amplia de subconjuntos de la recta, los llamados “conjuntos medibles", y si la clase de las funciones que se consideran se amplía a la que él llamó “funciones medibles” apareciendo así lo que se denominan *sumas de Lebesgue* de f correspondientes a particiones en la imagen de dicha función:

$$\sum_{k=1}^{n} y_k\lambda\left(f^{-1}(J_k)\right),$$

donde $J_1,\ldots,J_n$ son intervalos disjuntos cuya unión es $\mathbb{R}$; $y_k \in J_k$, y $\lambda\left(f^{-1}(J_k)\right)$ denota la medida del conjunto medible $f^{-1}(J_k)$ para $k = 1,\ldots,n$. Es de destacar que esta idea, sencilla pero brillante, de considerar particiones de la imagen de f en vez de en el dominio de f dio lugar a la teoría de la medida y la integral de Lebesgue.

La divulgación de dicha idea se suele expresar con el siguiente símil: “Si Riemann y Lebesgue tuviesen que contar un montón de monedas sobre una mesa éstos procederían de distinta forma. Riemann cuadricularía la mesa, contaría las monedas de cada cuadrícula y sumaría; sin embargo, Lebesgue clasificaría las monedas, contaría las que hay de cada clase y sumaría".

El paso de la teoría de integración de Riemann a la de Lebesgue es un proceso de completación similar al paso de $\mathbb{Q}$ a $\mathbb{R}$, por lo que no es de extrañar que la integral de Lebesgue disponga de mejores resultados de convergencia.

Sabemos que la medida de Lebesgue λ está muy relacionada con la topología. En esta lección presentamos una versión abstracta de la integral de Lebesgue, relativa a una medida arbitraria μ sobre cualquier espacio medible $(X,\mathscr{A})$. Esta teoría abstracta, que no es en forma alguna más difícil que el caso particular de la recta real, muestra que una gran parte de la teoría de la integración es independiente de cualquier topología en el conjunto subyacente y, por supuesto, nos proporciona un instrumento de mucha mayor aplicabilidad.

Como hemos comentado, la clase de las funciones medibles juega un papel fundamental en la teoría de la integración. Dicha clase tiene algunas propiedades básicas en común con otra clase muy importante de funciones, la clase de las funciones continuas. Es útil tener en consideración estas analogías: por una parte los conceptos de espacio topológico, conjunto abierto y función continua y por otra los de espacio medible, conjunto medible y función medible. Esta relación es más clara si el marco en que se tratan es abstracto, y esto (más que el mero deseo de generalidad) es lo que motiva el enfoque elegido.

11.1. Función medible.

Definición 11.1 (Función medible)**.** Sean $(X,\mathscr{A})$ e $(Y,\mathscr{B})$ espacios medibles. Se dice que $f : X \longrightarrow Y$ es una *función medibles* si la imagen inversa por f de cada conjunto medible de Y es un conjunto medible de X, es decir,

$$f^{-1}(B) \in \mathscr{A}, \quad \forall B \in \mathscr{B}. \tag{11.1.1}$$

Es inmediato que la composición de funciones medibles es medible.

Recordemos que todo espacio topológico, salvo mención expresa, se considerará un espacio medible con la σ-álgebra de Borel.

Teniendo en cuenta que

$$f^{-1}(Y) = X,\ \ f^{-1}(Y\backslash B) = X\backslash f^{-1}(B)\ \text{ y }\ f^{-1}(\cup_{n\in\mathbb{N}}B_n) = \cup_{n\in\mathbb{N}}f^{-1}(B_n),$$

podemos afirmar que la familia

$$\left\{B \subset Y : f^{-1}(B) \in \mathscr{A}\right\}$$

es una σ-álgebra en Y. En consecuencia, para comprobar que f es medible basta verificar la condición 11.1.1 para una familia de subconjuntos de Y que genere la σ-álgebra $\mathscr{B}$. En particular, las funciones continuas de $\mathbb{R}^N$ en $\mathbb{R}^M$ son medibles si se considera en $\mathbb{R}^N$ cualquier σ-álgebra que contenga a los borelianos (por ejemplo $\mathscr{M}$) y en $\mathbb{R}^M$

la σ-álgebra de Borel. En cuyo caso, basta comprobar la condición 11.1.1 para los intervalos abiertos acotados pues sabemos que todo abierto es unión numerable de intervalos abiertos acotados (véase la tantas veces citada Proposición 10.8).

Ejemplos 11.2 (funciones medibles)**.**

1. Toda función constante f de X en Y es medible.

2. La restricción a un subconjunto medible (con la σ-álgebra inducida) de una función medible es medible.

3. Recordemos que la función característica de un subconjunto A de X, χ_A, es la función de X en $\mathbb{R}$ definida por
$$\chi_A(x) = \begin{cases} 1 & \text{si} \quad x \in A \\ 0 & \text{si} \quad x \notin A \end{cases}.$$
Se tiene que χ_A es medible en $(X, \mathscr{A})$ si, y sólo si, $A \in \mathscr{A}$. En efecto, basta observar que para cualquier subconjunto B de $\mathbb{R}$ se verifica que
$$\chi_A^{-1}(B) = \begin{cases} X & \text{si} \quad 0,1 \in B \\ A & \text{si} \quad 0 \notin B, 1 \in B \\ X \backslash A & \text{si} \quad 0 \in B, 1 \notin B \\ \emptyset & \text{si} \quad 0,1 \notin B \end{cases}.$$

4. Sea $(X, \mathscr{A})$ un espacio medible. Si $E \in \mathscr{A}$ y $f : E \longrightarrow \mathbb{R}$, entonces a f le asociamos la función $f\chi_E : X \longrightarrow \mathbb{R}$ (prolongación por cero fuera de E) dada por
$$f\chi_E(x) = \begin{cases} f(x) & \text{si} \quad x \in E \\ 0 & \text{si} \quad x \in X \backslash E. \end{cases}$$
Entonces las siguientes afirmaciones son equivalentes

   i) f es medible en $(E, \mathscr{A}_E)$.

   ii) $f\chi_E$ es medible en $(X, \mathscr{A})$.

   En efecto, para B abierto de $\mathbb{R}$ se tiene que
   i) $\Rightarrow$ ii)
$$(f\chi_E)^{-1}(B) = \{x \in X : (f\chi_E)(x) \in B\} =$$
$$\begin{cases} E^C \cup \{x \in E : f(x) \in B\} = E^C \cup f^{-1}(B) & \text{si} \quad 0 \in B \\ \{x \in E : f(x) \in B\} = f^{-1}(B) & \text{si} \quad 0 \notin B \end{cases}$$
   ii) $\Rightarrow$ i)
   La inclusión de E en X es medible, luego la restricción de $f\chi_E$ a E es medible, por ser composición de medibles ($f = f\chi_E \circ i_E$). ∎

5. Sean $(X,\mathscr{A},\mu)$ un espacio de medida completo e Y un espacio topológico. Si $f,g : X \longrightarrow Y$ son funciones iguales c.p.d., entonces f es medible si, y sólo si, lo es g.

   En efecto, supongamos por ejemplo que la función f es medible y que el conjunto $Z = \{x \in X : f(x) \neq g(x)\}$ es de medida cero. Para B medible de Y se tiene que

   $$g^{-1}(B) = (g^{-1}(B) \cap Z) \cup (g^{-1}(B) \cap Z^C) = (g^{-1}(B) \cap Z) \cup (f^{-1}(B) \cap Z^C)$$

   es medible ya que $g^{-1}(B) \cap Z$ es medible por ser μ completa, y $f^{-1}(B) \cap Z^C$ es medible por ser f medible. ■

   Obsérvese que si la medida no es completa dos funciones iguales c.p.d. no tienen que ser simultáneamente medibles. En efecto, si es A un subconjunto no medible de un conjunto Z de medida cero, entonces las funciones $f,g : X \longrightarrow \mathbb{R}$ definidas por

   $$f(x) = 0,\ \ \forall x \in X, \quad g(x) = \begin{cases} 0 & \text{si} \quad x \in X \backslash Z \\ 1 & \text{si} \quad x \in A \\ 2 & \text{si} \quad x \in Z \backslash A \end{cases}$$

   son iguales c.p.d. pues

   $$\{x \in X : f(x) \neq g(x)\} = Z.$$

   f es medible y sin embargo, como

   $$g^{-1}(\{1\}) = \{x \in X : g(x) = 1\} = A,$$

   y $\{1\}$ es un boreliano, concluimos que g no es medible.

6. Sean $(\mathbb{R}^N,\mathscr{M},\lambda)$ el espacio de medida de Lebesgue y $E \in \mathscr{M}$. Son medibles las siguientes funciones reales definidas en E:

   i) Las continuas.

   ii) Las continuas c.p.d.

   iii) Las iguales c.p.d. a una medible.

   iv) Las composiciones de una medible con una continua.

   En efecto,
   i) y iv) se han probado antes.
   ii) Sea $f : E \longrightarrow \mathbb{R}$ continua c.p.d. Sea Z el subconjunto de E de medida cero en el que f no es continua. $f_{|E\backslash Z}$ es medible por ser continua y $f_{|Z}$ es claramente medible. En consecuencia, si $B \subset \mathbb{R}$ es abierto, entonces

   $$f^{-1}(B) = \left(f_{|E\backslash Z}\right)^{-1}(B) \cup \left(f_{|Z}\right)^{-1}(B) \in \mathscr{M},$$

   y, por tanto, f es medible.
   iii) Consecuencia del apartado 5. ■

Nota: Es conveniente advertir que en el espacio medible $(\mathbb{R}^N,\mathscr{M})$ cualquier conjunto susceptible de ser construido con las técnicas básicas de la Teoría de conjuntos es medible y, en consecuencia, cualquier función definida mediante el uso de las técnicas habituales del Análisis es medible.

11.2. Propiedades de las funciones medibles.

Proposición 11.3 (Caracterización de las funciones medibles)**.** *Sean* $(\Omega, \mathscr{A})$ *un espacio medible y* $f : \Omega \longrightarrow \mathbb{R}$ *una función. Equivalen:*

i) f *es medible.*

ii) $\{\omega \in \Omega : f(\omega) < \alpha\} \in \mathscr{A}, \ \forall \alpha \in \mathbb{R}.$

iii) $\{\omega \in \Omega : f(\omega) \geq \alpha\} \in \mathscr{A}, \ \forall \alpha \in \mathbb{R}.$

iv) $\{\omega \in \Omega : f(\omega) \leq \alpha\} \in \mathscr{A}, \ \forall \alpha \in \mathbb{R}.$

v) $\{\omega \in \Omega : f(\omega) > \alpha\} \in \mathscr{A}, \ \forall \alpha \in \mathbb{R}.$

Demostración:

Es obvio que la primera afirmación implica cualquiera de las restantes. Para probar el recíproco es suficiente demostrar que cualquiera de las familias de intervalos siguientes

a) $\{]-\infty, \alpha[: \alpha \in \mathbb{R}\}$

b) $\{[\alpha, +\infty[: \alpha \in \mathbb{R}\}$

c) $\{]-\infty, \alpha] : \alpha \in \mathbb{R}\}$

d) $\{]\alpha, +\infty[: \alpha \in \mathbb{R}\}$

generan la σ-álgebra de Borel de $\mathbb{R}$, para lo que basta comprobar que la σ-álgebra generada contiene a los intervalos abiertos acotados.

Por ejemplo, en el caso b) basta observar que

$$[\alpha, +\infty[= \mathbb{R} \setminus]-\infty, \alpha[, \ \ ]\alpha, +\infty[= \bigcup_{n=1}^{\infty} \left[\alpha + \frac{1}{n}, +\infty\right[\quad \text{y} \quad]\alpha, \beta[=]-\infty, \beta[\cap]\alpha, +\infty[.$$

Los otros casos se dejan como ejercicio (¡Hágase!). ∎

Dadas $f, g : \Omega \to \mathbb{R}$ se definen las funciones $f \vee g, f \wedge g, f^+, f^- : \Omega \to \mathbb{R}$ por

$$(f \vee g)(\omega) := \max\{f(\omega), g(\omega)\}, \ \ (f \wedge g)(\omega) := \min\{f(\omega), g(\omega)\},$$
$$f^+ := 0 \vee f, \ \ f^- := 0 \vee (-f).$$

Es inmediato comprobar

$$f = f^+ - f^-, \ \ |f| = f^+ + f^-, \ \ f \vee g = \frac{1}{2}(f + g + |f - g|), \ \ f \wedge g = \frac{1}{2}(f + g - |f - g|).$$

Proposición 11.4 (Estabilidad algebraica de las funciones medibles)**.** *Sean* $(\Omega, \mathscr{A})$ *un espacio medible,* $f, g : \Omega \longrightarrow \mathbb{R}$ *funciones medibles,* α *un número real y* p *un positivo, entonces las funciones*

$$f + g, \ \alpha f, \ fg, \ |f|^p, \ f^+ \ \text{y} \ f^-$$

son también medibles. En consecuencia, el conjunto de las funciones medibles reales es una subálgebra del álgebra de las funciones reales definidas en Ω*. Si además suponemos que* f *no se anula, entonces la función* $\frac{1}{f}$ *también es medible.*

Demostración:

Empezaremos probando que la función $h:\Omega\longrightarrow\mathbb{R}^2$ definida por

$$h(\omega)=(f(\omega),g(\omega)),\quad \forall\omega\in\Omega$$

es también medible. Si G es un abierto de $\mathbb{R}^2$, entonces, de nuevo la Proposición 10.8 nos asegura la existencia de dos sucesiones $\{I_n\},\{J_n\}$ de intervalos acotados de $\mathbb{R}$ tales que $G=\bigcup_{n=1}^{\infty}I_n\times J_n$ con lo que concluimos que

$$h^{-1}(G)=\bigcup_{n=1}^{\infty}h^{-1}(I_n\times J_n)=\bigcup_{n=1}^{\infty}\left[f^{-1}(I_n)\cap g^{-1}(J_n)\right]\in\mathscr{A}.$$

Como las funciones de $\mathbb{R}^2$ en $\mathbb{R}$

$$(x,y)\mapsto x+y,\quad (x,y)\mapsto xy$$

son continuas, se deduce que las funciones obtenidas por composiciones con h de éstas son medibles, esto es $f+g$ y fg son medibles.

Como las funciones de $\mathbb{R}$ en $\mathbb{R}$

$$x\mapsto\alpha x,\ \ x\mapsto|x|^p,\quad x\mapsto\max\{x,0\},\quad \text{y}\ \ x\mapsto\max\{-x,0\}$$

son continuas, deducimos que αf, $|f|^p$, f^+ y f^- son medibles.

Por último, si f no se anula, basta componer la función medible f con la función continua $x\mapsto\frac{1}{x}$ de $\mathbb{R}^*$ en $\mathbb{R}$ para obtener la medibilidad de $\frac{1}{f}$. ∎

Proposición 11.5 (Estabilidad analítica de las func. medibles)**.** *Sea $(\Omega,\mathscr{A})$ un espacio medible. Si $\{f_n\}$ es una sucesión de funciones medibles definidas en Ω, entonces se verifican las siguientes propiedades:*

i) El conjunto $E=\{\omega\in\Omega:\{f_n(\omega)\}$ es mayorada $\}$ es medible y si dicho conjunto no es vacío, la función $f:E\longrightarrow\mathbb{R}$ definida por $f(\omega)=\sup\{f_n(\omega):n\in\mathbb{N}\}$, $\forall\omega\in E$, es medible.

ii) El conjunto $E=\{\omega\in\Omega:\{f_n(\omega)\}$ es minorada $\}$ es medible y si dicho conjunto no es vacío, la función $f:E\longrightarrow\mathbb{R}$ definida por $f(\omega)=\inf\{f_n(\omega):n\in\mathbb{N}\}$, $\forall\omega\in E$, es medible.

iii) El conjunto $E=\{\omega\in\Omega:\limsup f_n(\omega)\in\mathbb{R}\}$ es medible y si dicho conjunto no es vacío, la función $f:E\longrightarrow\mathbb{R}$ definida por $f(\omega)=\limsup f_n(\omega)$, $\forall\omega\in E$, es medible.

iv) El conjunto $E=\{\omega\in\Omega:\liminf f_n(\omega)\in\mathbb{R}\}$ es medible y en caso de ser no es vacío, la función $f:E\longrightarrow\mathbb{R}$ definida por $f(\omega)=\liminf f_n(\omega)$, $\forall\omega\in E$, es medible.

v) El conjunto $E=\{\omega\in\Omega:\{f_n(\omega)\}$ es convergente $\}$ es medible y si dicho conjunto no es vacío la función $f:E\longrightarrow\mathbb{R}$ definida por $f(\omega)=\lim f_n(\omega)$, $\forall\omega\in E$, es medible.

Demostración:

i) Es inmediato que

$$E = \bigcup_{m=1}^{\infty} \left[\bigcap_{n=1}^{\infty} \{\omega \in \Omega : f_n(\omega) \leq m\} \right],$$

y, por tanto, $E \in \mathscr{A}$. Por otra parte, supuesto $E \neq \emptyset$, para cada $\alpha \in \mathbb{R}$ se tiene que

$$\{\omega \in E : f(\omega) \leq \alpha\} = \bigcap_{n=1}^{\infty} \{\omega \in E : f_n(\omega) \leq \alpha\} =$$

$$= E \bigcap \cap_{n=1}^{\infty} \{\omega \in \Omega : f_n(\omega) \leq \alpha\}$$

es medible, de donde se deduce que f es medible.

ii) Aplíquese i) a la sucesión $\{-f_n\}$.

iii) El conjunto

$$F = \{\omega \in \Omega : \{f_n(\omega)\} \text{ es mayorada }\}$$

es medible por i). Es claro que si $F = \emptyset$, entonces también $E = \emptyset$. Supuesto $F \neq \emptyset$, consideremos, para cada natural n, la función $g_n : F \longrightarrow \mathbb{R}$ definida por

$$g_n(\omega) = \sup\{f_k(\omega) : k \geq n\}, \ \ \forall \omega \in F,$$

que es medible por i). Es claro que

$$E = \{\omega \in F : \{g_n(\omega)\} \text{ es minorada }\}$$

que es medible por ii). Supuesto $E \neq \emptyset$ se tiene que $f(\omega) = \inf\{g_n(\omega) : n \in \mathbb{N}\}$ también es medible por ii).

iv) Aplíquese iii) a la sucesión $\{-f_n\}$.

v) Sean

$$F_1 = \{\omega \in \Omega : \limsup f_n(\omega) \in \mathbb{R}\}$$
$$F_2 = \{\omega \in \Omega : \liminf f_n(\omega) \in \mathbb{R}\},$$

conjuntos que son medibles (por iii) y iv)). Es claro que si $F_1 \cap F_2 = \emptyset$, entonces $E = \emptyset$. Supuesto $F_1 \cap F_2 \neq \emptyset$, las funciones $g : F_1 \longrightarrow \mathbb{R}$ definida por $g(\omega) = \limsup f_n(\omega)$, $\forall \omega \in F_1$ y $h : F_2 \longrightarrow \mathbb{R}$ definida por $h(\omega) = \liminf f_n(\omega)$, $\forall \omega \in F_2$ son medibles por iii) y iv), respectivamente. Entonces

$$E = \{\omega \in F_1 \cap F_2 : g(\omega) = h(\omega)\} = F_1 \cap F_2 \cap (g-h)^{-1}(\{0\})\}$$

es medible y finalmente, supuesto $E \neq \emptyset$, como $f(\omega) = g(\omega)(= h(\omega)), \forall \omega \in E$, concluimos que la función f también es medible.

■

11.3. Funciones simples. Teorema de aproximación de Lebesgue.

Consideramos ahora las funciones medibles más sencillas: las funciones simples. Será clave en la teoría el hecho de que cualquier función medible positiva es límite puntual de una sucesión creciente de funciones simples positivas.

Definición 11.6 (Función simple)**.** Sea $(\Omega, \mathscr{A})$ un espacio medible. Una función $s : \Omega \longrightarrow \mathbb{R}$ es una *función simple* si es medible y su imagen es un conjunto finito, esto es, $s(\Omega) = \{\alpha_1, \ldots, \alpha_m\}$ para convenientes $m \in \mathbb{N}$ y $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$. Claramente los conjuntos

$$A_i := \{\omega \in \Omega : s(\omega) = \alpha_i\} \quad (i = 1, \ldots, m),$$

son medibles y forman una partición de Ω. La función s se expresa entonces en la forma

$$s = \sum_{i=1}^{m} \alpha_i \chi_{A_i}.$$

Proposición 11.7 (Estabilidad algebraica de las func. simples)**.** *Sea $(\Omega, \mathscr{A})$ un espacio medible. En el espacio vectorial de las funciones reales definidas en Ω, se verifica que el conjunto de las funciones simples es el subespacio vectorial generado por las funciones características de los conjuntos medibles. Además si s y t son simples, entonces st es simple.*

Demostración:
Inmediata de la definición, de la estabilidad algebraica de las funciones medibles y del Ejemplo 11.2.2. ∎

El siguiente resultado es uno de los máximos exponentes del "saber hacer"de Lebesgue. Como ya hemos comentado, a diferencia de lo que se hace en la integral de Riemann, en la que se subdivide el dominio de la función para aproximarla por funciones escalonadas, lo que se hará para la integral que pretendemos definir es subdividir la imagen de la función para conseguir aproximarla por funciones simples positivas.

Teorema 11.8 (de aproximación de Lebesgue)**.** *Sean $(\Omega, \mathscr{A})$ un espacio medible y $f : \Omega \longrightarrow [0, \infty[$ una función medible. Entonces, existe una sucesión creciente $\{s_n\}$ de funciones simples positivas que converge puntualmente a f en Ω.*

Si además f está mayorada, entonces existe una sucesión creciente de funciones simples positivas $\{s_n\}$ que converge uniformemente a f en Ω.

Demostración:
Fijado n natural, definimos

$$E_n = \Big\{\omega \in \Omega : n \leq f(\omega)\Big\}$$

y, para $k = 1, 2, \ldots, n2^n$,

$$E_{n,k} := \Big\{\omega \in \Omega : \frac{k-1}{2^n} \leq f(\omega) < \frac{k}{2^n}\Big\}.$$

Como f es medible, los conjuntos E_n y $E_{n,k}$ también lo son. Definimos entonces la función simple

$$s_n := \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \chi_{E_{n,k}} + n\chi_{E_n}.$$

(obsérvese que los conjuntos cuyas funciones características aparecen en la definición anterior forman una partición de Ω).

Probemos que la sucesión $\{s_n\}$ así obtenida es puntualmente creciente, es decir,

$$s_n(\omega) \leq s_{n+1}(\omega), \quad \forall n \in \mathbb{N},\ \forall \omega \in \Omega.$$

Dado un natural n y $\omega \in \Omega$, entonces si $\omega \in E_n$, se tiene

$$s_n(\omega) = n = \frac{n2^{n+1}}{2^{n+1}} \leq f(\omega),$$

y, por tanto, $s_n(\omega) \leq s_{n+1}(\omega)$. En otro caso, si $\omega \in E_{n,k}$, entonces al ser

$$\left[\frac{k-1}{2^n}, \frac{k}{2^n}\right[= \left[\frac{2k-2}{2^{n+1}}, \frac{2k-1}{2^{n+1}}\right[\cup \left[\frac{2k-1}{2^{n+1}}, \frac{2k}{2^{n+1}}\right[.$$

o bien $\omega \in E_{n+1,2k-1}$ o bien $\omega \in E_{n+1,2k}$. En el primer caso

$$s_n(\omega) = \frac{k-1}{2^n} = s_{n+1}(\omega),$$

mientras que en el segundo

$$s_n(\omega) = \frac{k-1}{2^n} < \frac{2k-1}{2^{n+1}} = s_{n+1}(\omega).$$

Veamos ahora que para $\omega \in \Omega$ se verifica

$$\{s_n(\omega)\} \to f(\omega).$$

Sea $\omega \in \Omega$ y sea n un natural tal que $f(\omega) < n$. Se tiene entonces que

$$s_n(\omega) \leq f(\omega) < s_n(\omega) + \frac{1}{2^n}.$$

Así, para n suficientemente grande, se verifica que

$$0 \leq f(\omega) - s_n(\omega) < \frac{1}{2^n}, \tag{11.3.1}$$

de donde se obtiene la convergencia anunciada.

Finalmente, si además f está mayorada, tenemos que

$$\exists m \in \mathbb{N} \ : \ E_n = \emptyset,\ \forall n \geq m,$$

con lo cual de 11.3.1 deducimos que

$$0 \leq f(\omega) - s_n(\omega) < \frac{1}{2^n}, \qquad \forall \omega \in \Omega,\ \forall n \geq m.$$

Hemos probado que en este caso la convergencia es uniforme. ∎

11.4. Integral de funciones simples positivas.

En el caso en que dispongamos de una medida μ sobre $\mathscr{A}$, hay una forma natural de definir la integral de una función simple positiva. Comprobamos en primer lugar que la definición que daremos de integral es independiente de la expresión de la función simple que elijamos. Supongamos que se da la igualdad

$$\sum_{i=1}^{m} \alpha_i \chi_{A_i} = \sum_{j=1}^{n} \beta_j \chi_{B_j},$$

donde $m,n \in \mathbb{N},\ 0 \leq \alpha_i, \beta_j < +\infty\ \ \forall i,j$ y $\{A_1,\dots,A_m\},\{B_1,\dots,B_n\}$ son dos particiones de Ω en conjuntos medibles, entonces comprobaremos que se verifica la igualdad

$$\sum_{i=1}^{m} \alpha_i \mu(A_i) = \sum_{j=1}^{n} \beta_j \mu(B_j).$$

En efecto, usando que la medida es aditiva y que si para ciertos índices i,j se verifica que $A_i \cap B_j \neq \emptyset$, entonces evaluando la función simple en esta intersección se tiene $\alpha_i = \beta_j$, obtenemos

$$\sum_{i=1}^{m} \alpha_i \mu(A_i) = \sum_{i=1}^{m} \alpha_i \mu\Big(A_i \cap \Big(\cup_{j=1}^{n} B_j\Big)\Big) = \sum_{i=1}^{m} \alpha_i \mu\Big(\cup_{j=1}^{n}(A_i \cap B_j)\Big) =$$

$$= \sum_{i=1}^{m} \alpha_i \sum_{j=1}^{n} \mu(A_i \cap B_j) = \sum_{i=1}^{m}\sum_{j=1}^{n} \alpha_i \mu(A_i \cap B_j) = \sum_{j=1}^{n}\sum_{i=1}^{m} \alpha_i \mu(A_i \cap B_j) =$$

$$= \sum_{j=1}^{n}\sum_{i=1}^{m} \beta_j \mu(A_i \cap B_j) = \sum_{j=1}^{n} \beta_j \sum_{i=1}^{m} \mu(A_i \cap B_j) = \sum_{j=1}^{n} \beta_j \mu\Big(\cup_{i=1}^{m}(B_j \cap A_i)\Big) = \sum_{j=1}^{n} \beta_j \mu(B_j).$$

Definición 11.9 (Integral de una función simple positiva)**.** Sea $(\Omega,\mathscr{A},\mu)$ un espacio de medida. Si $s = \sum_{i=1}^{m} \alpha_i \chi_{A_i} : \Omega \longrightarrow [0,\infty[$, donde $\{A_1,\dots,A_m\}$ es una partición de Ω en conjuntos medibles, entonces se define la integral de la función simple s como el elemento de $[0,\infty]$ dado por:

$$\int_\Omega s\, d\mu := \sum_{i=1}^{m} \alpha_i \mu(A_i).$$

Proposición 11.10 (Prop. de la integral de las funciones simples positivas)**.** *Sea* $(\Omega,\mathscr{A},\mu)$ *un espacio de medida. Si* $s,t : \Omega \longrightarrow [0,\infty[$ *son funciones simples y* $\alpha \in [0,\infty[$*, entonces se verifican las siguientes afirmaciones:*

i) $\int_\Omega (s+t)\, d\mu = \int_\Omega s\, d\mu + \int_\Omega t\, d\mu$.

ii) $\int_\Omega (\alpha s)\, d\mu = \alpha \int_\Omega s\, d\mu$.

iii) $\Big[s \leq t\Big] \Rightarrow \int_\Omega s\, d\mu \leq \int_\Omega t\, d\mu$.

iv) $\int_\Omega s\, d\mu = 0$ *si, y sólo si,* $s = 0$ *c.p.d.*

Demostración:

Pongamos

$$s = \sum_{i=1}^{m} \alpha_i \chi_{A_i} \quad \text{y} \quad t = \sum_{j=1}^{n} \beta_j \chi_{B_j},$$

con m, n naturales, $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_n \in [0, \infty[$ y $\{A_1, \ldots, A_m\}, \{B_1, \ldots B_n\}$ particiones de Ω en conjuntos medibles.

i) Es claro que

$$\{A_i \cap B_j : 1 \leq i \leq m, 1 \leq j \leq n\}$$

es una partición de Ω en conjuntos medibles (algunos de los cuales pueden ser el vacío y por tanto se pueden suprimir). Se tiene que

$$s + t = \sum_{i=1}^{m} \sum_{j=1}^{n} (\alpha_i + \beta_j) \chi_{A_i \cap B_j}.$$

Por tanto,

$$\int_{\Omega} (s+t)\, d\mu = \sum_{i=1}^{m} \sum_{j=1}^{n} (\alpha_i + \beta_j) \mu(A_i \cap B_j) =$$

$$\sum_{i=1}^{m} \alpha_i \sum_{j=1}^{n} \mu(A_i \cap B_j) + \sum_{j=1}^{n} \beta_j \sum_{i=1}^{m} \mu(A_i \cap B_j).$$

Usando que los conjuntos $\{B_1, \ldots, B_n\}, \{A_1, \ldots, A_m\}$ son particiones de Ω y la aditividad de la medida se tiene que

$$\sum_{j=1}^{n} \mu(A_i \cap B_j) = \mu(A_i), \qquad \sum_{i=1}^{m} \mu(A_i \cap B_j) = \mu(B_j).$$

Usando estas igualdades válidas para $1 \leq i \leq m, 1 \leq j \leq n$, se obtiene que

$$\int_{\Omega} (s+t)\, d\mu = \int_{\Omega} s\, d\mu + \int_{\Omega} t\, d\mu.$$

ii) Se tiene que

$$\int_{\Omega} (\alpha s)\, d\mu = \sum_{i=1}^{m} \alpha \alpha_i \mu(A_i) = \alpha \sum_{i=1}^{m} \alpha_i \mu(A_i) = \alpha \int_{\Omega} s\, d\mu.$$

iii) Como $s \leq t$ por hipótesis, entonces si $\omega \in A_i \cap B_j \neq \emptyset$ tenemos que $s(\omega) = \alpha_i \leq \beta_j = t(\omega)$. Por tanto

$$\int_{\Omega} s\, d\mu = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_i \mu(A_i \cap B_j) \leq \sum_{i=1}^{m} \sum_{j=1}^{n} \beta_j \mu(A_i \cap B_j) = \int_{\Omega} t\, d\mu.$$

iv) Se tiene que

$$\int_\Omega s\,d\mu = 0 \Leftrightarrow \sum_{i=1}^{m} \alpha_i\mu(A_i) = 0 \Leftrightarrow \alpha_i\mu(A_i) = 0, i = 1,\dots,m \Leftrightarrow$$

$$\left\{\begin{array}{l} \alpha_i = 0 \\ \text{ó} \\ \mu(A_i) = 0 \end{array}\right. \quad 1 \le i \le m \quad \Leftrightarrow s = 0 \ \text{ c.p.d..}$$

∎

De la proposición anterior se deduce que

$$\int_\Omega s\,d\mu = \max\left\{\int_\Omega t\,d\mu : t \ \text{ simple positiva, } t \le s\right\}.$$

Podemos, por tanto, extender la definición de integral de la siguiente forma:

11.5. Integral de una función medible positiva.

Definición 11.11 (Integral de una función medible positiva)**.** Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida. Si $f : \Omega \longrightarrow [0,\infty[$ es una función medible, se define la integral de f como el elemento de $[0,\infty]$ dado por

$$\int_\Omega f\,d\mu := \sup\left\{\int_\Omega s\,d\mu : \ s \ \text{ simple },0 \le s \le f\right\}.$$

El siguiente resultado junto con el Teorema de aproximación de Lebesgue permiten obtener cómodamente las propiedades de la integral de las funciones medibles positivas.

Proposición 11.12. *Sean $(\Omega, \mathscr{A}, \mu)$ un espacio de medida y $f : \Omega \longrightarrow [0,\infty[$ una función medible. Si $\{s_n\}$ es una sucesión creciente de funciones simples positivas que converge puntualmente a f en Ω, entonces*

$$\int_\Omega f\,d\mu = \lim \int_\Omega s_n\,d\mu.$$

Demostración:

De la definición de integral para funciones medibles positivas se deduce que

$$\int_\Omega s_n \, d\mu \leq \int_\Omega f \, d\mu, \qquad \forall n \in \mathbb{N},$$

y, por tanto,

$$\lim \int_\Omega s_n \, d\mu \leq \int_\Omega f \, d\mu.$$

Para probar la otra desigualdad, sea s una función simple positiva con $s \leq f$ y sea $0 < \rho < 1$. Para cada natural n definimos

$$F_n := \Big\{ \omega \in \Omega : s_n(\omega) \geq \rho s(\omega) \Big\}.$$

Obtenemos así una sucesión creciente de conjuntos medibles tal que $\bigcup_{n=1}^\infty F_n = \Omega$. En efecto, para cada $\omega \in \Omega$, se tiene

$$\begin{cases} f(\omega) = 0 \;\; \Rightarrow & \omega \in F_1 \\ \\ f(\omega) > 0 \;\; \Rightarrow & \rho s(\omega) < f(\omega) \; (\rho < 1) \;\; \Rightarrow \exists n \in \mathbb{N} : \omega \in F_n \end{cases}.$$

Supuesto que $s = \sum_{k=1}^m \alpha_k \chi_{A_k}$, para cada natural n se verifica que

$$\begin{aligned} \rho \sum_{k=1}^m \alpha_k \, \mu(A_k \cap F_n) &= \sum_{k=1}^m \rho \, \alpha_k \, \mu(A_k \cap F_n) = \\ &\int_\Omega \rho \, s \, \chi_{F_n} \, d\mu \leq \int_\Omega s_n \, \chi_{F_n} \, d\mu \leq \int_\Omega s_n \, d\mu, \end{aligned} \tag{11.5.1}$$

donde se ha utilizado la definición de F_n y la monotonía de la integral 11.10.iii). Para cada $k \in \{1, \ldots, m\}$, $\big\{A_k \cap F_n\big\}_{n \in \mathbb{N}}$ es una sucesión creciente de conjuntos medibles tal que

$$\bigcup_{n=1}^\infty \big(A_k \cap F_n\big) = A_k,$$

y por tanto se sigue de la propiedad de crecimiento continuo de μ que

$$\mu(A_k) = \lim_n \mu\big(A_k \cap F_n\big).$$

Tomando límite cuando $n \to \infty$ en 11.5.1 obtenemos que

$$\rho \sum_{k=1}^m \alpha_k \mu(A_k) \leq \lim \int_\Omega s_n \, d\mu,$$

es decir,

$$\rho \int_\Omega s \, d\mu \leq \lim \int_\Omega s_n \, d\mu.$$

Tomando ahora en la desigualdad anterior límite cuando $\rho \to 1$ obtenemos que

$$\int_\Omega s \, d\mu \leq \lim \int_\Omega s_n \, d\mu.$$

Considerando por último la arbitrariedad de s, concluimos que

$$\int_{\Omega} f \, d\mu \leq \lim \int_{\Omega} s_n \, d\mu.$$

∎

Proposición 11.13 (Prop. de la integral de las funciones medibles posit.)**.** *Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida. Si $f, g : \Omega \longrightarrow [0, \infty[$ son funciones medibles y $\alpha \in [0, \infty[$, entonces se verifican las siguientes afirmaciones:*

i) $\int_{\Omega}(f+g) \, d\mu = \int_{\Omega} f \, d\mu + \int_{\Omega} g \, d\mu$.

ii) $\int_{\Omega}(\alpha f) \, d\mu = \alpha \int_{\Omega} f \, d\mu$.

iii) $\left[f \leq g\right] \Rightarrow \int_{\Omega} f \, d\mu \leq \int_{\Omega} g \, d\mu$.

iv) $\int_{\Omega} f \, d\mu = 0$ *si, y sólo si,* $f = 0$ *c.p.d.*

Además si $f = g$ c.p.d., entonces $\int_{\Omega} f \, d\mu = \int_{\Omega} g \, d\mu$.

Demostración:

i) El Teorema de aproximación de Lebesgue nos asegura la existencia de dos sucesiones crecientes $\{s_n\}$ y $\{t_n\}$ de funciones simples positivas con $\{s_n\} \nearrow f$ y $\{t_n\} \nearrow g$. En consecuencia $\{s_n + t_n\}$ es también una sucesión creciente de funciones simples positivas con $\{s_n + t_n\} \nearrow f + g$. En virtud de la proposición anterior se tiene que

$$\int_{\Omega}(f+g) \, d\mu = \lim \int_{\Omega}(s_n + t_n) \, d\mu =$$

$$= \lim \int_{\Omega} s_n \, d\mu + \lim \int_{\Omega} t_n \, d\mu = \int_{\Omega} f \, d\mu + \int_{\Omega} g \, d\mu.$$

ii) Es inmediata de la definición.

iii) Si $s : \Omega \longrightarrow [0, \infty[$ es una función simple con $s \leq f$, entonces $s \leq g$, y por tanto

$$\int_{\Omega} s \, d\mu \leq \int_{\Omega} g \, d\mu.$$

En consecuencia

$$\int_{\Omega} f \, d\mu \leq \int_{\Omega} g \, d\mu.$$

En el argumento hemos utilizado dos veces la definición de integral de una función medible positiva.

iv) Sea $\{s_n\}$ una sucesión creciente de funciones simples positivas convergente a f.

Supuesto que $\int_\Omega f\, d\mu = 0$, entonces $\int_\Omega s_n\, d\mu = 0$, $\forall n \in \mathbb{N}$, y en virtud del apartado iv) de la Proposición 11.10 es $s_n = 0$ c.p.d., $\forall n \in \mathbb{N}$, de donde concluimos que $f = 0$ c.p.d. (pues $\{\omega \in \Omega : f(\omega) \neq 0\} \subset \cup_{n\in\mathbb{N}}\{\omega \in \Omega : s_n(\omega) \neq 0\}$).

Supongamos ahora que $f = 0$ c.p.d. Entonces, como $0 \le s_n \le f$, se tiene que

$$s_n = 0 \qquad \text{c.p.d.,} \qquad \forall n \in \mathbb{N},$$

luego

$$\int_\Omega s_n\, d\mu = 0, \quad \forall n \in \mathbb{N}$$

y por tanto $\int_\Omega f\, d\mu = 0$.

Finalmente, como para cualesquiera dos funciones medibles positivas f y g se verifica

$$f \le |f - g| + g,$$

de los apartados iii) y i) anteriores se deduce que

$$\int_\Omega f\, d\mu \le \int_\Omega |f - g|\, d\mu + \int_\Omega g\, d\mu.$$

Si suponemos que $f = g$ c.p.d., entonces, teniendo en cuenta iv) deducimos que

$$\int_\Omega f\, d\mu \le \int_\Omega g\, d\mu.$$

De igual forma

$$\int_\Omega g\, d\mu \le \int_\Omega f\, d\mu,$$

por lo que

$$\int_\Omega f\, d\mu = \int_\Omega g\, d\mu.$$

■

Nota: De la última afirmación se deduce que para calcular la integral de una función medible positiva pueden ignorarse, si así se desea, los valores que toma dicha función en un conjunto de medida cero. Esto es, el conocimiento de una función medible positiva en algún subconjunto medible E de Ω tal que $\mu(\Omega \backslash E) = 0$, es suficiente para calcular la integral de dicha función.

11.6. Función integrable e integral.

Definición 11.14 (Función integrable e integral de una función integrable)**.** Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida. Diremos que una función medible $f : \Omega \longrightarrow \mathbb{R}$ es integrable si la función medible positiva $|f|$ tiene integral finita. Si f es integrable, entonces las funciones f^+ y f^- son medibles positivas y tienen integral finita, puesto que $f^+ \leq |f|$ y $f^- \leq |f|$, podemos definir

$$\int_\Omega f \, d\mu := \int_\Omega f^+ \, d\mu - \int_\Omega f^- \, d\mu.$$

Notaremos $\mathscr{L}(\mu)$ al conjunto de las funciones $f : \Omega \longrightarrow \mathbb{R}$ integrables, es decir,

$$\mathscr{L}(\mu) := \left\{ f : \Omega \longrightarrow \mathbb{R} : f \text{ es medible con } \int_\Omega |f| \, d\mu < \infty \right\}.$$

Claramente si una función medible positiva tiene integral finita, entonces ésta es integrable y su integral coincide con la integral como función medible positiva (lo cual legitima la notación introducida). Más generalmente, conviene observar que si g y h son funciones medibles positivas tales que

$$\int_\Omega g \, d\mu < \infty, \quad \int_\Omega h \, d\mu < \infty \quad \text{y} \quad f = g - h,$$

entonces

$$f \in \mathscr{L}(\mu) \quad \text{y} \quad \int_\Omega f \, d\mu = \int_\Omega g \, d\mu - \int_\Omega h \, d\mu.$$

En efecto, es claro que $|f|$ es medible y las propiedades de la integral de una función medible positiva nos aseguran que

$$\int_\Omega |f| \, d\mu = \int_\Omega |g - h| \, d\mu \leq \int_\Omega (g + h) \, d\mu = \int_\Omega g \, d\mu + \int_\Omega h \, d\mu < \infty.$$

Por otra parte, al ser $f^+ + h = f^- + g$, se sigue que

$$\int_\Omega f^+ \, d\mu + \int_\Omega h \, d\mu = \int_\Omega f^- \, d\mu + \int_\Omega g \, d\mu,$$

y en consecuencia

$$\int_\Omega f \, d\mu = \int_\Omega g \, d\mu - \int_\Omega h \, d\mu.$$

■

Proposición 11.15 (Propiedades de $\mathscr{L}(\mu)$ y de la integral)**.** *Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida. Si $f, g : \Omega \longrightarrow \mathbb{R}$ son funciones integrables y α es un número real, se verifican las siguientes afirmaciones:*

i) $f + g$ es integrable con $\int_\Omega (f + g) \, d\mu = \int_\Omega f \, d\mu + \int_\Omega g \, d\mu$.

ii) αf es integrable con $\int_\Omega (\alpha f) \, d\mu = \alpha \int_\Omega f \, d\mu$.

iii) $\left[f \leq g\right] \Rightarrow \int_\Omega f\, d\mu \leq \int_\Omega g\, d\mu$.

iv) $|\int_\Omega f\, d\mu| \leq \int_\Omega |f|\, d\mu$.

Además si f y g son medibles e iguales c.p.d., entonces

$$f \text{ es integrable } \Leftrightarrow g \text{ es integrable,}$$

en cuyo caso

$$\int_\Omega f\, d\mu = \int_\Omega g\, d\mu.$$

Demostración:

i) Puesto que $f+g = (f^+ + g^+) - (f^- + g^-)$, el resultado se deduce de la aditividad de la integral para funciones medibles positivas y de la observación hecha antes.

ii) Sabemos que αf es medible y que

$$\int_\Omega |\alpha f|\, d\mu = \int_\Omega |\alpha|\, |f|\, d\mu = |\alpha| \int_\Omega |f|\, d\mu < \infty,$$

y por tanto αf es integrable. Por otra parte $\alpha f = \alpha f^+ - \alpha f^-$, por lo que si $\alpha \geq 0$ se tiene que

$$\int_\Omega (\alpha f)\, d\mu = \int_\Omega (\alpha f^+)\, d\mu - \int_\Omega (\alpha f^-)\, d\mu =$$

$$= \alpha \int_\Omega f^+\, d\mu - \alpha \int_\Omega f^-\, d\mu = \alpha \int_\Omega f\, d\mu,$$

mientras que si $\alpha < 0$, entonces

$$\int_\Omega (\alpha f)\, d\mu = \int_\Omega (-\alpha f^-)\, d\mu - \int_\Omega (-\alpha f^+)\, d\mu =$$

$$= -\alpha \int_\Omega f^-\, d\mu + \alpha \int_\Omega f^+\, d\mu = \alpha \int_\Omega f\, d\mu.$$

iii) Se tiene que

$$f^+ - f^- \leq g^+ - g^- \Rightarrow f^+ + g^- \leq g^+ + f^- \Rightarrow$$

$$\Rightarrow \int_\Omega \left(f^+ + g^-\right) d\mu \leq \int_\Omega \left(g^+ + f^-\right) d\mu \Rightarrow$$

$$\int_\Omega f^+\, d\mu + \int_\Omega g^-\, d\mu \leq \int_\Omega g^+\, d\mu + \int_\Omega f^-\, d\mu \quad \Rightarrow$$

$$\Rightarrow \quad \int_\Omega f^+\, d\mu - \int_\Omega f^-\, d\mu \leq \int_\Omega g^+\, d\mu - \int_\Omega g^-\, d\mu,$$

es decir,

$$\int_\Omega f\, d\mu \leq \int_\Omega g\, d\mu.$$

iv) Como $|f|$ es claramente integrable y $-|f| \leq f \leq |f|$, los apartados anteriores nos dan

$$-\int_\Omega |f|\, d\mu \leq \int_\Omega f\, d\mu \leq \int_\Omega |f|\, d\mu,$$

es decir,

$$\left|\int_\Omega f\, d\mu\right| \leq \int_\Omega |f|\, d\mu.$$

Finalmente si f, g son medibles e iguales c.p.d., entonces $|f| = |g|$ c.p.d. y la Proposición 11.13 garantiza que

$$\int_\Omega |f|\, d\mu = \int_\Omega |g|\, d\mu,$$

lo que prueba que f es integrable si, y sólo si, lo es g. Por otra parte, como $|f-g| = 0$ c.p.d., si f y g son integrables, entonces concluimos que

$$\left|\int_\Omega f\, d\mu - \int_\Omega g\, d\mu\right| = \left|\int_\Omega (f-g)\, d\mu\right| \leq \int_\Omega |f-g|\, d\mu = 0,$$

donde se ha aplicado nuevamente la Proposición 11.13. ∎

Nota: De la última afirmación de la proposición anterior se deduce que para estudiar la integrabilidad, y el valor de la integral cuando proceda, de una función medible pueden ignorarse los valores que dicha función toma en un conjunto cualquiera de medida cero.

Proposición 11.16. *Sean* $(\Omega, \mathscr{A}, \mu)$ *un espacio de medida,* $E \in \mathscr{A}$*,* $E \subset D \subset \Omega$ *y* $f : D \longrightarrow \mathbb{R}$ *una función. Entonces* $f_{|E} \in \mathscr{L}(\mu_E)$ *si, y sólo si,* $f\chi_E \in \mathscr{L}(\mu)$*, en cuyo caso*

$$\int_E f_{|E}\, d\mu_E = \int_\Omega f\chi_E\, d\mu.$$

Demostración:

Sabemos que $f_{|E}$ es medible en $(E, \mathscr{A}_E)$ si, y sólo si, $f\chi_E$ es medible en $(\Omega, \mathscr{A})$, por lo que para la demostración daremos por supuesta la medibilidad de ambas funciones.

Supongamos en primer lugar que $0 \leq f$. Tomemos una sucesión creciente de funciones simples positivas $\{s_n\}$ que converge puntualmente a $f\chi_E$. Es claro entonces que $\left\{s_{n|E}\right\}$ es una sucesión creciente de funciones simples positivas definidas en E que converge puntualmente a $f_{|E}$. Por la Proposición 11.12 aplicada a las funciones $f_{|E}$ y $f\chi_E$ tenemos que

$$\int_E f_{|E}\, d\mu_E = \lim \int_E s_{n|E}\, d\mu_E \quad \text{y} \quad \int_\Omega f\chi_E\, d\mu = \lim \int_\Omega s_n\chi_E\, d\mu.$$

Observemos que para cada $n \in \mathbb{N}$ se verifica que

$$s_n(\omega) = 0, \qquad \forall \omega \in E^C,$$

luego si

$$s_n = \sum_{k=1}^m \alpha_k \chi_{A_k}$$

para convenientes $m \in \mathbb{N}, \alpha_1, \dots, \alpha_m \in [0,\infty[$ y $\{A_1, \dots, A_m\}$ partición de Ω en conjuntos medibles, necesariamente ha de ocurrir que

$$\alpha_k = 0, \quad \text{si } 1 \leq k \leq m \ \text{ y } \ A_k \cap E^C \neq \emptyset,$$

luego

$$\sum_{k=1}^{m} \alpha_k \mu_E(A_k \cap E) = \sum_{k=1}^{m} \alpha_k \mu(A_k),$$

y, por tanto

$$\int_E s_{n|E}\, d\mu_E = \int_\Omega s_n \chi_E\, d\mu.$$

Ahora, de las anteriores expresiones se sigue que

$$\int_E f_{|E}\, d\mu_E = \int_\Omega f\chi_E\, d\mu,$$

de donde se obtiene el enunciado. Finalmente, para f arbitraria, aplicando lo anterior a $|f|$ obtenemos que

$$\int_E |f_{|E}|\, d\mu_E = \int_E |f|_{|E}\, d\mu_E = \int_\Omega |f|\chi_E\, d\mu = \int_\Omega |f\chi_E|\, d\mu,$$

por lo que

$$f_{|E} \in \mathscr{L}(\mu_E) \ \Leftrightarrow \ f\chi_E \in \mathscr{L}(\mu).$$

Además, es ese supuesto, dado que

$$f_{|E} = f^+_{|E} - f^-_{|E} \qquad \text{y} \qquad f\chi_E = f^+\chi_E - f^-\chi_E,$$

de nuevo aplicando lo anterior a f^+ y f^- obtenemos que

$$\int_E f_{|E}\, d\mu_E = \int_E f^+_{|E}\, d\mu_E - \int_E f^-_{|E}\, d\mu_E = \int_\Omega f^+\chi_E\, d\mu - \int_\Omega f^-\chi_E\, d\mu = \int_\Omega f\chi_E\, d\mu.$$

∎

Definición 11.17 (Integral de func. en conjunto medible y definida c.p.d.)**.** Sean $(\Omega, \mathscr{A}, \mu)$ un espacio de medida, $D \subset \Omega$ y $f : D \longrightarrow \mathbb{R}$ una función. Si $E \in \mathscr{A}$ con $E \subset D$, entonces se dice que f es *integrable en E* si se verifica que $f_{|E}$ es una función integrable en el espacio de medida $(E, \mathscr{A}_E, \mu_E)$, y en tal caso definimos la *integral de f en E* por

$$\int_E f\, d\mu := \int_E f_{|E}\, d\mu_E.$$

En virtud del resultado anterior este concepto se puede también expresar en términos del espacio de medida inicial. En efecto, f es integrable en E si se verifica que $f\chi_E$ es integrable en el espacio de medida $(\Omega, \mathscr{A}, \mu)$, y en tal caso la integral de f en E viene dada por

$$\int_E f\, d\mu = \int_\Omega f\chi_E\, d\mu.$$

Una función f definida c.p.d. en Ω es integrable si lo es en un conjunto medible E con $\mu(E^C) = 0$ en el que esté definida, en cuyo caso definimos

$$\int_\Omega f\, d\mu := \int_E f\, d\mu.$$

Esta definición no depende del conjunto medible E. Obsérvese que si F es cualquier otro conjunto medible con $\mu(F^C) = 0$ en el que f esté definida y sea integrable, entonces $f\chi_E = f\chi_F$ c.p.d. pues coinciden en $E \cap F$ (con f) y

$$\mu\left((E\cap F)^C\right) = \mu\left(E^C \cup F^C\right) \leq \mu\left(E^C\right) + \mu\left(F^C\right) = 0,$$

con lo que

$$\int_F f\, d\mu = \int_\Omega f\chi_F\, d\mu = \int_\Omega f\chi_E\, d\mu = \int_E f\, d\mu.$$

Proposición 11.18 (Aditividad respecto del conjunto de integración)**.** *Sean $(\Omega, \mathscr{A}, \mu)$ un espacio de medida, $D \subset \Omega$ y $f : D \longrightarrow \mathbb{R}$ una función. Si $E, F \in \mathscr{A}$ son disjuntos con $E \cup F \subset D$, entonces f es integrable en $E \cup F$ si, y sólo si, f es integrable en E y en F, y en tal caso*

$$\int_{E\cup F} f\, d\mu = \int_E f\, d\mu + \int_F f\, d\mu.$$

Demostración:
Puesto que

$$f\chi_{E\cup F} = f\chi_E + f\chi_F,$$

de las propiedades de las funciones medibles se sigue inmediatamente que

$f\chi_{E\cup F}$ es medible si, y sólo si, $f\chi_E$ y $f\chi_F$ son medibles.

En tal caso, de las propiedades de la integral de las funciones medibles positivas se tiene que

$$\int_{E\cup F} |f|\, d\mu = \int_E |f|\, d\mu + \int_F |f|\, d\mu,$$

de donde se deduce que

f es integrable en $E \cup F$ si, y sólo si, lo es en E y en F.

En ese supuesto, se verifica la fórmula del enunciado. ■

Ejemplos 11.19.

1. Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida. Entonces el espacio vectorial $\mathbf{F}$ generado por las funciones características de los conjuntos de medida finita es un subespacio de $\mathscr{L}(\mu)$. Además
$$\int_\Omega \sum_{i=1}^m \alpha_i \chi_{E_i}\, d\mu = \sum_{i=1}^m \alpha_i \mu(E_i), \qquad \forall \sum_{i=1}^m \alpha_i \chi_{E_i} \in \mathbf{F}.$$

En efecto, es claro que si E es un conjunto medible con $\mu(E) < \infty$, entonces la función χ_E, al ser medible positiva con integral acotada, es integrable y su integral es $\int_\Omega \chi_E \, d\mu = \mu(E)$. El resto es consecuencia de la linealidad de la integral.

De hecho **F** es el espacio vectorial de las funciones simples integrables (¡Compruébese!). En el próximo teorema comprobaremos que cualquier función integrable es "aproximable" por funciones de **F**.

2. Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida:

   i) Si $f : \Omega \longrightarrow \mathbb{R}$ es una función medible y existe $g : \Omega \longrightarrow \mathbb{R}$ integrable tal que $|f| \leq g$, entonces f es integrable.

   ii) Si $E \in \mathscr{A}$ con $\mu(E) < \infty$ y $f : E \longrightarrow \mathbb{R}$ es medible y acotada, entonces f es integrable. En particular, toda función continua definida en un subconjunto compacto de $\mathbb{R}^N$ es integrable (respecto de la medida de Lebesgue).

   En efecto, para probar i) basta tener en cuenta la monotonía de la integral para funciones medibles positivas y la definición de función integrable. Ahora ii) es consecuencia de que $|f| \leq M\chi_E$, donde M es una cota de f.

   Es obligado advertir, sin embargo, que el cálculo de la integral de una tal función incluso en las situaciones más sencillas es todavía inabordable. La función

$$f(x,y) = x + y, \qquad \forall (x,y) \in [0,1] \times [0,1]$$

   es un ejemplo de ello.

3. Si el lector conoce la integral de Riemann puede deducir inmediatamente que toda función integrable en el sentido de Riemann es integrable y que ambas integrales coinciden. En efecto, sean $[a,b]$ un intervalo compacto de $\mathbb{R}$ y $f : [a,b] \longrightarrow \mathbb{R}$ una función integrable Riemann. No quita generalidad suponer que $0 \leq f$, pues en otro caso basta tener presente que $f = f^+ - f^-$.

   Para cada natural n se considera la partición de $[a,b]$

$$P_n = \left\{ x_k = a + k\frac{b-a}{2^n} : \;\; k = 0, 1, \ldots, 2^n \right\}$$

   y las funciones simples $e_n, E_n : [a,b] \longrightarrow [0, +\infty[$ definidas por

$$e_n = \sum_{k=1}^{2^n} m_k \chi_{[x_{k-1}, x_k[} + f(b)\chi_{\{b\}}, \qquad E_n = \sum_{k=1}^{2^n} M_k \chi_{[x_{k-1}, x_k[} + f(b)\chi_{\{b\}}$$

   donde para $k = 1, \ldots, 2^n$,

$$m_k = \inf\Big(f([x_{k-1}, x_k])\Big) \quad \text{y} \quad M_k = \sup\Big(f([x_{k-1}, x_k])\Big).$$

Es inmediato comprobar que para cada natural n se verifica que

$$\int_{[a,b]} e_n d\lambda = I(f,P_n) := \sum_{k=1}^{2^n} m_k(x_k - x_{k-1}),$$

$$\int_{[a,b]} E_n d\lambda = S(f,P_n) := \sum_{k=1}^{2^n} M_k(x_k - x_{k-1}),$$

y

$$e_n \leq e_{n+1} \leq f \leq E_{n+1} \leq E_n.$$

Si notamos

$$e = \lim e_n \quad \text{y} \quad E = \lim E_n,$$

se tiene que e y E son medibles y acotadas y el ejemplo anterior garantiza que son integrables. Además la Proposición 11.12 nos asegura que

$$\int_{[a,b]} e\, d\lambda = \lim \int_{[a,b]} e_n\, d\lambda = \int_a^b f(x)\, dx = \lim \int_{[a,b]} E_n\, d\lambda = \int_{[a,b]} E\, d\lambda.$$

Consecuentemente la función medible positiva $E - e$ tiene integral cero y por tanto

$$e = E = f \ \text{ c.p.d.}$$

La Proposición 11.15 garantiza que f es integrable y que

$$\int_{[a,b]} f\, d\lambda = \int_{[a,b]} e\, d\lambda = \int_{[a,b]} E\, d\lambda = \int_a^b f(x)\, dx.$$

El ejemplo anterior legitima la introducción de la siguiente notación. Sean $I \subset \mathbb{R}$ un intervalo de extremos α y β con $-\infty \leq \alpha < \beta \leq +\infty$ y $f : I \longrightarrow \mathbb{R}$ integrable (respecto de la medida de Lebesgue). Entonces se usa la notación

$$\int_\alpha^\beta f(x)\, dx$$

para designar la integral $\int_I f\, d\lambda$.

Análogamente cuando se consideran funciones integrables en $\mathbb{R}^N$ su integral se nota por

$$\int_{\mathbb{R}^N} f(x_1,\ldots,x_N)\, d(x_1,\ldots,x_N).$$

4. El siguiente ejemplo pone de manifiesto que el conjunto de las funciones integrables (Lebesgue) es estrictamente mayor que el de funciones integrables según Riemann. La función $f : \mathbb{R} \to \mathbb{R}$ definida por

$$f(x) = \begin{cases} 0 & \text{si } x \notin \mathbb{Q} \\ x & \text{si } x \in \mathbb{Q} \end{cases},$$

es integrable con integral nula (pues $f = 0$ c.p.d.) y, sin embargo, no está definida en un intervalo compacto, ni es acotada y sólo es continua en un punto.

11.7. Densidad de las funciones simples en $\mathscr{L}(\mu)$.

Teorema 11.20 (densidad)**.** *Sea $(\Omega,\mathscr{A},\mu)$ un espacio de medida. Entonces el espacio vectorial* $\mathbf{F}$ *de las funciones simples integrables es denso en $\mathscr{L}(\mu)$, es decir: dada una función integrable f existe una sucesión $\{f_n\}$ de funciones de* $\mathbf{F}$ *tal que*

$$\lim \int_\Omega |f-f_n| \ d\mu = 0.$$

Además se puede conseguir que

$$\{f_n\} \to f \text{ puntualmente en } \Omega.$$

Demostración:

Sea f una función integrable, entonces f^+ y f^- son también funciones integrables. Sean $\{s_n\}$ y $\{t_n\}$ dos sucesiones crecientes de funciones simples positivas tales que

$$\{s_n\} \to f^+, \quad \{t_n\} \to f^-$$

(por ejemplo las dadas por el Teorema de aproximación de Lebesgue 11.8). Sabemos que

$$\int_\Omega f^+ \, d\mu = \lim \int_\Omega s_n \, d\mu \quad \text{y} \quad \int_\Omega f^- \, d\mu = \lim \int_\Omega t_n \, d\mu$$

(Proposición 11.12). Pongamos para cada natural n, $f_n := s_n - t_n$. Es inmediato que $\{f_n\} \to f$. Al ser

$$0 \le s_n \le f^+, \qquad 0 \le t_n \le f^- \qquad (n \in \mathbb{N}),$$

concluimos que

$$s_n, t_n \in \mathbf{F}, \ \ \forall n \in \mathbb{N},$$

y, por tanto, $f_n \in \mathbf{F}$, $\forall n \in \mathbb{N}$.

Veamos que

$$\lim \int_\Omega |f-f_n| \, d\mu = 0.$$

En efecto, para cada natural n se tiene que

$$\int_\Omega |f_n - f| \, d\mu = \int_\Omega |s_n - t_n - (f^+ - f^-)| \, d\mu \le \int_\Omega |s_n - f^+| \, d\mu + \int_\Omega |t_n - f^-| \, d\mu =$$

$$\int_\Omega f^+ \, d\mu - \int_\Omega s_n \, d\mu + \int_\Omega f^- \, d\mu - \int_\Omega t_n \, d\mu,$$

y basta tener en cuenta que la última sucesión tiene límite cero. ■

11.8. Referencias recomendadas.

[Jur], [Ber], [Guz] y [Ru].

11.9. Resumen del resultados del Tema 11.

Función medible.
Sean $(X,\mathscr{A})$ y $(Y,\mathscr{B})$ espacios medibles. Una función $f : X \longrightarrow Y$ se dice medible si

$$f^{-1}(B) \in \mathscr{A}, \quad \forall B \in \mathscr{B}$$

equivalentemente, si la condición anterior se verifica para una familia de conjuntos que genere la σ-álgebra $\mathscr{B}$.

En el espacio medible $(\mathbb{R}^N,\mathscr{M})$ cualquier función real definida mediante el uso de las técnicas habituales del Análisis es medible.

Función simple.
Sea $(\Omega,\mathscr{A})$ un espacio medible. Una función $s : \Omega \longrightarrow \mathbb{R}$ es una función simple si es medible y su imagen es finita, esto es, $s(\Omega) = \{\alpha_1,\ldots,\alpha_m\}$. Claramente los conjuntos

$$A_i := \{\omega \in \Omega : s(\omega) = \alpha_i\} \qquad (i = 1,\ldots,m)$$

son medibles y forman una partición de Ω. La función s se expresa entonces en la forma

$$s = \sum_{i=1}^{m} \alpha_i \chi_{A_i}.$$

Teorema de aproximación de Lebesgue.
Sean $(\Omega,\mathscr{A})$ un espacio medible y $f : \Omega \longrightarrow [0,\infty[$ una función medile. Entonces existe una sucesión creciente $\{s_n\}$ de funciones simples positivas que converge puntualmente a f en Ω.

Si además f está mayorada, entonces existe una sucesión creciente de funciones simples positivas $\{s_n\}$ que converge uniformemente a f en Ω .

Integral de una función simple positiva.
Sea $(\Omega,\mathscr{A},\mu)$ un espacio de medida. Si $s = \sum_{i=1}^{m} \alpha_i \chi_{A_i} : \Omega \longrightarrow [0,\infty[$ es una función simple, se define la integral de s como el elemento de $[0,\infty]$ dado por:

$$\int_\Omega s \, d\mu := \sum_{i=1}^{m} \alpha_i \mu(A_i).$$

Integral de una función medible positiva.
Sea $(\Omega,\mathscr{A},\mu)$ un espacio de medida. Si $f : \Omega \longrightarrow [0,\infty[$ es una función medible, se define la integral de f como el elemento de $[0,\infty]$ dado por:

$$\int_\Omega f \, d\mu := \sup\left\{ \int_\Omega s \, d\mu : \;\; s \;\; \text{simple}, \; 0 \le s \le f \right\}.$$

Proposición

Sean $(\Omega, \mathscr{A}, \mu)$ *un espacio de medida y* $f : \Omega \longrightarrow [0, \infty[$ *una función medible. Si* $\{s_n\}$ *es una sucesión creciente de funciones simples positivas que converge puntualmente a* f *en* Ω*, entonces*

$$\int_\Omega f \, d\mu = \lim \int_\Omega s_n \, d\mu.$$

Función integrable e integral de una función.

Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida. Diremos que una función medible $f : \Omega \longrightarrow \mathbb{R}$ es integrable si la función medible positiva $|f|$ tiene integral finita. Si f es integrable definimos la integral de f por

$$\int_\Omega f \, d\mu := \int_\Omega f^+ \, d\mu - \int_\Omega f^- \, d\mu.$$

Propiedades de la integral.

Sea $(\Omega, \mathscr{A}, \mu)$ *un espacio de medida. Si* $f, g : \Omega \longrightarrow \mathbb{R}$ *son funciones integrables y* α *es un número real, se verifican las siguientes afirmaciones:*

i) $f + g$ *es integrable y* $\int_\Omega (f+g) \, d\mu = \int_\Omega f \, d\mu + \int_\Omega g \, d\mu$.

ii) αf *es integrable y* $\int_\Omega (\alpha f) \, d\mu = \alpha \int_\Omega f \, d\mu$.

iii) $\left[f \leq g\right] \Rightarrow \int_\Omega f \, d\mu \leq \int_\Omega g \, d\mu$.

iv) $|\int_\Omega f \, d\mu| \leq \int_\Omega |f| \, d\mu$.

Además si f *y* g *son medibles e iguales c.p.d., entonces*

$$f \text{ es integrable } \Leftrightarrow g \text{ es integrable},$$

en cuyo caso $\int_\Omega f \, d\mu = \int_\Omega g \, d\mu$.

Integral de una func. en un conjunto y de una func. definida c.p.d.

Una función f definida en un conjunto que contenga a un conjunto medible E se dice integrable en E si $f_{|E}$ es integrable en $(E, \mathscr{A}_E, \mu_E)$, en cuyo caso se define la integral de f en E por

$$\int_E f d\mu := \int_E f_{|E} d\mu_E = \int_\Omega f \chi_E d\mu.$$

Una función f definida c.p.d. en Ω es integrable si lo es en un conjunto medible E con $\mu(E^C) = 0$ en el que esté definida, en cuyo caso definimos

$$\int_\Omega f \, d\mu := \int_E f \, d\mu.$$

Aditividad de la integral respecto del conjunto de integración.
Sean $(\Omega, \mathscr{A}, \mu)$ un espacio de medida, $D \subset \Omega$ y $f : D \longrightarrow \mathbb{R}$ una función. Si $E, F \in \mathscr{A}$ con $E \cap F = \emptyset$ y $E \cup F \subset D$, entonces f es integrable en $E \cup F$ si, y sólo si, f es integrable en E y en F, en cuyo caso $\int_{E\cup F} f\, d\mu = \int_E f\, d\mu + \int_F f\, d\mu$.

Teorema de densidad.
Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida. Entonces el espacio vectorial $\mathbf{F}$ de las funciones simples integrables es denso en $\mathscr{L}(\mu)$, es decir, dada una función integrable f existe una sucesión $\{f_n\}$ de funciones de $\mathbf{F}$ tal que

$$\lim \int_\Omega |f - f_n|\, d\mu = 0.$$

Además se puede conseguir que $\{f_n\} \to f$ puntualmente en Ω.

11.10. Ejercicios del Tema 11.

11.1 Sean $(\Omega, \mathscr{A}, \mu)$ un espacio de medida y $f : \Omega \longrightarrow [0, \infty[$ una función medible. Probar, sin utilizar la Proposición 11.12, que

$$\int_{\Omega} f\, d\mu = 0 \Leftrightarrow f = 0 \quad \text{c.p.d.}$$

Indicación: Basta considerar, para cada natural n, el conjunto

$$E_n := \left\{ \omega \in \Omega : f(\omega) > \frac{1}{n} \right\}.$$

Justificar que $\mu(E_n) = 0, \forall n \in \mathbb{N}$ y deducir, utilizando el crecimiento continuo de la medida μ, que $\mu\left(\left\{ \omega \in \Omega : f(\omega) > 0 \right\}\right) = 0$.

11.2 Sean $(\Omega, \mathscr{A}, \mu)$ un espacio de medida y $f : \Omega \longrightarrow \mathbb{R}$ una función integrable. Probar que $f = 0$ c.p.d. si, y sólo si,

$$\int_E f\, d\mu = 0 \qquad \text{para cada } E \in \mathscr{A}.$$

11.3 Sea $(\Omega, \mathscr{A})$ un espacio medible. Probar que si $\{f_n\}$ es una sucesión de funciones medibles en Ω, entonces se verifican las siguientes propiedades:

i) El conjunto $E = \left\{ \omega \in \Omega : \sum_{n \geq 1} f_n(\omega) \ \text{ es convergente} \right\}$ es medible y si dicho conjunto no es vacío la función $f : E \longrightarrow \mathbb{R}$ definida por $f(\omega) = \sum_{n=1}^{\infty} f_n(\omega), \ \forall \omega \in E$, es medible.

ii) El conjunto $E = \left\{ \omega \in \Omega : \sum_{n \geq 1} f_n(\omega) \quad \text{es absolutamente convergente} \right\}$ es medible.

11.4 Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida. Dadas $f, g, h : \Omega \longrightarrow \mathbb{R}$, probar que

i) Si f es integrable y g es medible acotada, entonces fg es integrable.

ii) Si f es medible y g, h son integrables con $g \leq f \leq h$, entonces f es integrable (toda función medible comprendida entre dos integrables es integrable).

iii) Si f es medible y g, h son integrables, entonces $h \vee (f \wedge g)$ y $h \wedge (f \vee g)$ son integrables, donde las funciones $f \vee g$ y $f \wedge g$, se definen por:

$$(f \vee g)(\omega) = \max\left\{ f(\omega), g(\omega) \right\} \quad \text{y} \quad (f \wedge g)(\omega) = \min\left\{ f(\omega), g(\omega) \right\}$$

para todo ω en Ω.

Nota: El producto de dos funciones integrables no tiene por qué ser integrable. Más tarde daremos un ejemplo.

11.5 Sean $(\Omega, \mathscr{A}, \mu)$ un espacio de medida y $f : \Omega \longrightarrow \mathbb{R}$ medible con $0 \leq f < 1$. Probar que existe una sucesión $\{A_n\}$ de conjuntos medibles tales que

$$f = \sum_{n=1}^{\infty} \frac{1}{2^n} \chi_{A_n}$$

uniformemente en Ω. Probar también que

$$\int_{\Omega} f \, d\mu = \sum_{n=1}^{\infty} \frac{1}{2^n} \mu(A_n).$$

Indicación: Para cada $n \geq 2$ considerar la función $2^n(s_n - s_{n-1})$, donde $\{s_n\}$ es la sucesión que se construye en la demostración del Teorema de aproximación de Lebesgue. Para probar la última igualdad utilizar la Proposición 11.12.

11.6 Sean $(\Omega, \mathscr{A}, \mu)$ un espacio de medida y $f : \Omega \longrightarrow \mathbb{R}$ medible. Supóngase que $\{f_n\}$ es una sucesión de funciones medibles de Ω en $\mathbb{R}$ tal que para cada natural n

$$\mu\Big(\Big\{\omega \in \Omega : f(\omega) \neq f_n(\omega)\Big\}\Big) < \frac{1}{2^n}.$$

Probar que $\{f_n\}$ converge a f c.p.d.

Indicación: Para cada natural n, considérese

$$A_n = \Big\{\omega \in \Omega : f(\omega) \neq f_n(\omega)\Big\} \qquad \text{y} \qquad B_n = \bigcup_{k=n}^{\infty} A_k.$$

Sea $B = \bigcap_{n=1}^{\infty} B_n$. Probar que $\mu(B) = 0$ y que $\{f_n\} \to f$ en B^C.

11.7 (*) **Teorema de Luzin.**
Sea $f : \mathbb{R}^N \longrightarrow \mathbb{R}$ una función medible. Probar que para cada positivo ε, existe $F \subset \mathbb{R}^N$ cerrado tal que

$$\lambda(F^C) < \varepsilon \qquad \text{y} \qquad f_{|F} \qquad \text{es continua.}$$

Indicación: Probar el teorema para

i) Funciones simples (Si A es medible, tómese C cerrado y G abierto tales que $C \subset A \subset G$ y $\lambda(G \backslash C) < \varepsilon$ y considérese el cerrado $F = C \cup (G^C)$ que cumple que $\lambda(F^C) < \varepsilon$ y $(\chi_A)_{|F}$ es continua.

ii) Funciones medibles acotadas (considerar una sucesión de funciones simples que converja uniformemente a f (Teorema de aproximación de Lebesgue)).

iii) Funciones medibles (componer f con un homeomorfismo de $\mathbb{R}$ sobre $]-1,1[$).

11.8 Sea $\mathbf{C}_{00}(\mathbb{R}^N) = \{f : \mathbb{R}^N \longrightarrow \mathbb{R} : f \text{ es continua de soporte compacto}\}$, donde el soporte de f es el conjunto

$$\text{sop } (f) = \overline{\{x \in \mathbb{R}^N : f(x) \neq 0\}}.$$

Probar que

i) Si $f \in \mathbf{C}_{00}(\mathbb{R}^N)$, entonces f es integrable.

ii) (*) Una función $f : \mathbb{R}^N \longrightarrow \mathbb{R}$ es medible si, y sólo, existe una sucesión $\{f_n\}$ en $\mathbf{C}_{00}(\mathbb{R}^N)$ tal que $\{f_n\} \to f$ c.p.d.

Indicación: Para cada $n \in \mathbb{N}$ aplicar el Teorema de Luzin para obtener $F_n \subset \mathbb{R}^N$ cerrado con $\lambda(F_n^C) < \frac{1}{2^n}$ y $f_{|F_n}$ continua. Aplicar ahora el Teorema de Tietze (una función real continua definida en un cerrado se puede extender de manera continua a todo $\mathbb{R}^N$). Además si la función es acotada, entonces la extensión puede ser elegida con la misma cota a la función continua $f_{|F_n}$ para obtener $g_n : \mathbb{R}^N \longrightarrow \mathbb{R}$ continua tal que

$$\lambda\left(\left\{x \in \mathbb{R}^N : f(x) \neq g_n(x)\right\}\right) < \frac{1}{2^n}.$$

Para cada n, sea φ_n una función continua con $0 \leq \varphi_n \leq 1$ verificando

$$\varphi_n(x) = \begin{cases} 1 & \text{si} \quad x \in \overline{B}(0,n) \\ \\ 0 & \text{si} \quad x \notin B(0,n+1) \end{cases} \qquad \text{(Lema de Uryshon, Ejercicio 2.16)}$$

Por último para cada n sea $f_n := g_n \varphi_n$.

11.9 (*) **Teorema de Egorov.**
Sean $(\Omega, \mathscr{A}, \mu)$ un espacio de medida y $\{f_n\}$ una sucesión de funciones medibles en Ω que converge c.p.d. a f. Dados un conjunto E de medida finita y un positivo ε, probar que existe $A \subset E$ medible tal que $\lambda(A) < \varepsilon$ y $\{f_n\}$ converge uniformemente a f en $E \backslash A$.

Indicación: Para $k, m \in \mathbb{N}$ definir

$$E_{k,m} := \bigcup_{n=m}^{\infty} \left\{\omega \in E : |f_n(\omega) - f(\omega)| \geq \frac{1}{k}\right\}$$

y probar que $\lim_{m\to\infty} \mu\left(E_{k,m}\right) = 0$. Si $\varepsilon > 0$, para cada $k \in \mathbb{N}$ existe $m_k \in \mathbb{N}$ tal que $\mu\left(E_{k,m_k}\right) < \frac{\varepsilon}{2^k}$. Considerar $A = \bigcup_{k=1}^{\infty} E_{k,m_k}$ y probar que

$$\mu(A) < \varepsilon, \qquad |f_n(\omega) - f(\omega)| < \frac{1}{k}, \qquad \forall \omega \in E \backslash A, \quad n \geq m_k.$$

11.11. Soluciones a los ejercicios del Tema 11.

11.1 Supongamos que $\int_\Omega f\, d\mu = 0$. Para cada natural n, definimos

$$E_n := \left\{ \omega \in \Omega : f(\omega) > \frac{1}{n} \right\}.$$

Obtenemos así una sucesión creciente de conjuntos medibles verificando que

$$\bigcup_{n=1}^{\infty} E_n = \left\{ \omega \in \Omega : f(\omega) > 0 \right\}.$$

Es claro que

$$\frac{1}{n}\chi_{E_n} \leq f, \qquad \forall n \in \mathbb{N}$$

y por tanto

$$\frac{1}{n}\mu(E_n) = \int_\Omega \frac{1}{n}\chi_{E_n}\, d\mu \leq \int_\Omega f\, d\mu, \qquad \forall n \in \mathbb{N}.$$

Deducimos de esto que

$$\mu(E_n) = 0, \qquad \forall n \in \mathbb{N},$$

y del crecimiento continuo de la medida μ concluimos que

$$\mu\left(\left\{ \omega \in \Omega : f(\omega) > 0 \right\} \right) = 0,$$

por tanto $f = 0$ c.p.d.

Nota: Aunque no se utiliza la Proposición 11.12, se utiliza la propiedad de crecimiento continuo de una medida que es la base de su demostración.

11.2 Es claro que si $f = 0$ c.p.d., entonces $\int_E f\, d\mu = 0$, $\forall E \in \mathscr{A}$. Supongamos ahora que $\int_E f\, d\mu = 0$, $\forall E \in \mathscr{A}$. Consideremos el conjunto medible

$$E := \left\{ \omega \in \Omega : f(\omega) \geq 0 \right\}.$$

Se tiene entonces

$$\int_\Omega f^+\, d\mu = \int_E f^+\, d\mu = \int_E f\, d\mu = 0,$$

y el ejercicio anterior (Proposición 11.13.iv) nos asegura que $f^+ = 0$ c.p.d. Análogamente $f^- = 0$ c.p.d., luego $f = 0$ c.p.d.

11.3 *i)* Basta aplicar el apartado v) de la Proposición 11.5 a la sucesión de sumas parciales de la serie: $\left\{ \sum_{k=1}^{n} f_k(\omega) \right\}$.

ii) Basta aplicar el apartado v) de la Proposición 11.5 a la sucesión

$$\left\{\sum_{k=1}^{n} |f_k(\omega)|\right\}.$$

11.4 *i)* $|fg|$ es medible (Proposición 11.4) y si $|g| < K$ se sigue de las propiedades de la integral de las funciones medibles positivas (Proposición 11.13) que

$$\int_{\Omega} |f(\omega)|\,|g(\omega)|\,d\mu \leq \int_{\Omega} K|f(\omega)|\,d\mu = K\int_{\Omega} |f(\omega)|\,d\mu < \infty.$$

ii)

$$g \leq f \leq h \Rightarrow |f| \leq |g| \vee |h| = \frac{1}{2}\Big(|g| + |h| + \big||g| - |h|\big|\Big),$$

y f es integrable, ya que es medible y se verifica la desigualdad anterior.

iii) Teniendo en cuenta las propiedades de las funciones medibles y de las funciones integrables, las siguientes expresiones

$$f \vee g = \frac{1}{2}\Big(f + g + |f - g|\Big), \qquad f \wedge g = \frac{1}{2}\Big(f + g - |f - g|\Big)$$

nos aseguran que $h \vee (f \wedge g)$ y $h \wedge (f \vee g)$ son medibles y $h \vee g$ y $h \wedge g$ son integrables. Basta ahora tener en cuenta el apartado ii) y las desigualdades

$$h \leq h \vee (f \wedge g) \leq h \vee g, \qquad h \wedge g \leq h \wedge (f \vee g) \leq h$$

para concluir que $h \vee (f \wedge g)$ y $h \wedge (f \vee g)$ son integrables.

11.5 Sea $\{s_n\}$ la sucesión creciente de funciones simples positivas que converge uniformemente hacia f que se construye en la demostración del Teorema de aproximación de Lebesgue (11.8). La función medible positiva $2s_1$ sólo toma los valores $\{0, 1\}$, luego es la función característica de un conjunto medible. Llamamos

$$A_1 := \Big\{\omega \in \Omega : 2s_1(\omega) = 1\Big\} = \Big\{\omega \in \Omega : \frac{1}{2} \leq f(\omega) < 1\Big\}.$$

Para $n \geq 2$ se tiene que

$$\text{ó bien} \quad s_n(\omega) = s_{n-1}(\omega), \quad \text{ó bien } s_n(\omega) = s_{n-1}(\omega) + \frac{1}{2^n},$$

de donde se deduce que la función medible positiva

$$2^n\big(s_n - s_{n-1}\big)$$

sólo toma también los valores $\{0, 1\}$, luego es la función característica de un conjunto medible A_n. Es fácil comprobar que

$$A_n = \bigcup_{k=1}^{2^{n-1}} \left\{\omega \in \Omega : \frac{2k-1}{2^n} \leq f(\omega) < \frac{2k}{2^n}\right\}.$$

La sucesión de sumas parciales de la serie

$$\sum_{n\geq 1} \frac{1}{2^n}\chi_{A_n},$$

equivalentemente

$$s_1 + (s_2 - s_1) + \cdots + (s_n - s_{n-1}) + \cdots,$$

es la sucesión $\{s_n\}$ que converge uniformemente a f en Ω.

Finalmente la Proposición 11.12 nos asegura que

$$\int_\Omega f d\mu = \lim \int_\Omega s_n d\mu = \lim \sum_{k=1}^{n} \frac{1}{2^k}\mu(A_k) = \sum_{n=1}^{\infty} \frac{1}{2^n}\mu(A_n).$$

11.6 Para cada natural n, consideremos

$$A_n = \Big\{\omega \in \Omega : f(\omega) \neq f_n(\omega)\Big\}, \qquad B_n = \bigcup_{k=n}^{\infty} A_k.$$

La sucesión $\{B_n\}$ es decreciente y

$$\mu(B_n) = \mu\Big(\bigcup_{k=n}^{\infty} A_k\Big) \leq \sum_{k=n}^{\infty} \mu(A_k) < \sum_{k=n}^{\infty} \frac{1}{2^k} = \frac{1}{2^{n-1}}.$$

La propiedad de decrecimiento continuo de μ, nos asegura que

$$\mu(B) = \mu\Big(\bigcap_{n=1}^{\infty} B_n\Big) = \lim \mu(B_n) \leq \lim \frac{1}{2^{n-1}} = 0.$$

Se tiene que

$$B^C = \Big(\bigcap_{n=1}^{\infty} B_n\Big)^C = \bigcup_{n=1}^{\infty}\left(\bigcup_{k=n}^{\infty} A_k\right)^C = \bigcup_{n=1}^{\infty}\Big(\bigcap_{k=n}^{\infty} A_k^C\Big) =$$

$$= \bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}(A_k^C) = \bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\Big\{\omega \in \Omega : f(\omega) = f_k(\omega)\Big\}.$$

Así

$$\omega \in B^C \Rightarrow \exists m \in \mathbb{N} : k \geq m, \quad \text{entonces} \quad f_k(\omega) = f(\omega),$$

y en consecuencia hemos probado que

$$\{f_n(\omega)\} \to f(\omega),\ \forall \omega \in B^C.$$

11.7 **Teorema de Luzin.**

i) Si A es medible, tomamos C cerrado y G abierto tales que $C \subset A \subset G$ y $\lambda(G \backslash C) < \varepsilon$ y consideramos el cerrado $F = C \cup (G^C)$ que cumple que $\lambda(F^C) = \lambda(G \backslash C) < \varepsilon$ y $(\chi_A)_{|F}$ es continua (en C vale 1 y en $\mathbb{R}^{\backslash}G$ vale 0 y ambos son cerrados, de donde se deduce que la imagen inversa de cualquier cerrado de $\mathbb{R}$ es cerrada)

Sea ahora $s = \sum_{k=1}^{m} \alpha_k \chi_{A_k}$. Sean $F_1, \ldots, F_m$ cerrados tales que $\left(\chi_{A_k}\right)_{|F_k}$ sea continua y $\lambda(F_k^C) < \frac{\varepsilon}{m}$. El conjunto $F = F_1 \cap \ldots \cap F_m$ es cerrado, $s_{|F}$ es continua y

$$\lambda(F^C) = \lambda\left((F_1^C) \cup \ldots \cup (F_m^C)\right) \leq \lambda(F_1^C) + \ldots + \lambda(F_m^C) < \varepsilon.$$

ii) Sea f medible acotada, entonces existe una sucesión $\{s_n\}$ de funciones simples que converge uniformemente a f (basta aplicar el Teorema de aproximación de Lebesgue a f^+ y f^-). Por i)

$$\forall n \in \mathbb{N},\ \exists F_n \subset \mathbb{R}^N \ \text{ cerrado : } \ \lambda(F_n^C) < \frac{\varepsilon}{2^n} \ \text{ y } \ s_{n|F_n} \ \text{ continua.}$$

Sea $F = \bigcap_{n=1}^{\infty} F_n$, conjunto que es cerrado. Se tiene que

$$\lambda(F^C) = \lambda\left(\bigcup_{n=1}^{\infty} F_n^C\right) \leq \sum_{n=1}^{\infty} \lambda(F_n^C) < \varepsilon.$$

$$\left.\begin{array}{l} s_{n|F} \ \text{ es continua, para todo } n \\ s_{n|F} \longrightarrow f_{|F} \ \text{ uniformemente} \end{array}\right\} \Rightarrow f_{|F} \ \text{ es continua.}$$

iii) Sea $\varphi : \mathbb{R} \longrightarrow]-1,1[$ un homeomorfismo, por ejemplo

$$\varphi(x) = \frac{x}{1+|x|}, \quad \forall x \in \mathbb{R}.$$

La función $\varphi \circ f$ es medible acotada, luego por ii) existe F cerrado tal que $\lambda(F^C) < \varepsilon$ y $\varphi \circ f_{|F} = (\varphi \circ f)_{|F}$ es continua. Pero entonces

$$f_{|F} = \varphi^{-1} \circ (\varphi \circ f_{|F})$$

también es continua.

11.8 Sea

$$\mathbf{C}_{00}(\mathbb{R}^N) = \left\{ f : \mathbb{R}^N \to \mathbb{R} : f \text{ es continua de soporte compacto} \right\},$$

donde el soporte de f es el conjunto

$$\text{sop } (f) = \overline{\{x \in \mathbb{R}^N : f(x) \neq 0\}}.$$

i) Si $f \in \mathbf{C}_{00}(\mathbb{R}^N)$, probaremos que f es integrable.

La propiedad de compacidad nos asegura la existencia de $M > 0$ tal $|f| \leq M$, con lo que si sop $(f) = K$ se tiene que

$$|f| \leq M \chi_K \in \mathscr{L}(\lambda).$$

ii) Ahora probaremos que f es medible si, y sólo si, existe una sucesión $\{f_n\}$ en $\mathbf{C}_{00}(\mathbb{R}^N)$ tal que $\{f_n\} \to f$ c.p.d.

Es claro que si existe $\{f_n\}$ en $\mathbf{C}_{00}(\mathbb{R}^N)$ tal que $\{f_n\} \to f$ c.p.d., entonces f es medible (límite c.p.d. de medibles).

Recíprocamente, para cada $n \in \mathbb{N}$, si aplicamos el Teorema de Luzin, obtenemos un cerrado $F_n \subset \mathbb{R}^N$ con $\lambda(F_n^C) < \frac{1}{2^n}$ y $f_{|F_n}$ continua. Usamos ahora el Teorema de Tietze (una función real continua definida en un cerrado se puede extender de manera continua a todo $\mathbb{R}^N$). Además si la función es acotada, entonces la extensión puede ser elegida con la misma cota de la función continua $f_{|F_n}$. Obtenemos entonces una función $g_n : \mathbb{R}^N \longrightarrow \mathbb{R}$ continua tal que

$$\lambda\Big(\Big\{x \in \mathbb{R}^N : f(x) \neq g_n(x)\Big\}\Big) < \frac{1}{2^n}.$$

Para cada n, sea φ_n una función continua en $\mathbb{R}^N$ con $0 \leq \varphi_n \leq 1$ verificando

$$\varphi_n(x) = \begin{cases} 1 \quad \text{si} \quad x \in \overline{B}(0,n) \\ \\ 0 \quad \text{si} \quad x \notin B(0,n+1) \end{cases} \quad \text{(Ejercicio 2.16)} \quad .$$

Por último, para cada natural n, definimos $f_n := g_n\varphi_n$. Comprobemos que $f_n \in \mathbf{C}_{00}(\mathbb{R}^N)$ y que $\{f_n\} \to f$ c.p.d. En efecto, el Ejercicio 11.6 asegura que $\{g_n\} \to g$ c.p.d. Si $x \in \mathbb{R}^N$ es un punto de convergencia de esta sucesión, también lo es de $\{f_n\}$ sin más que observar que si $x \in \overline{B}(0,m)$, entonces

$$n \geq m \quad \Rightarrow \quad f_n(x) = g_n(x).$$

11.9 **Teorema de Egorov.**

Para $k,m \in \mathbb{N}$, definimos

$$E_{k,m} := \bigcup_{n=m}^{\infty} \Big\{\omega \in E : |f_n(\omega) - f(\omega)| \geq \frac{1}{k}\Big\}.$$

Es inmediato que para cada natural k, la sucesión $\{E_{k,m}\}_{m\in\mathbb{N}}$ es decreciente y como

$$E \backslash \bigcap_{m=1}^{\infty} E_{k,m} = \bigcup_{m=1}^{\infty} (E\backslash E_{k,m}) =$$

$$= \bigcup_{m=1}^{\infty}\bigcap_{n=m}^{\infty}\Big\{\omega \in E : |f_n(\omega) - f(\omega)| < \frac{1}{k} \ \text{ si } \ n \geq m\Big\}$$

que es casi todo E (contiene los puntos donde $\{f_n\}$ converge a f, en consecuencia

$$\mu\Big(E \backslash \bigcap_{m=1}^{\infty} E_{k,m}\Big) = \mu(E),$$

y al ser $\mu(E) < \infty$, se sigue que

$$\mu\Big(\bigcap_{m=1}^{\infty} E_{k,m}\Big) = 0.$$

Como $\mu(E) < \infty$, la propiedad de decrecimiento continuo nos asegura que

$$\lim_{m\to\infty} \mu\left(E_{k,m}\right) = 0.$$

Si $\varepsilon > 0$, es claro que para cada $k \in \mathbb{N}$ existe $m_k \in \mathbb{N}$ tal que

$$\mu\left(E_{k,m_k}\right) < \frac{\varepsilon}{2^k}.$$

Si notamos $A = \bigcup_{k=1}^{\infty} E_{k,m_k}$, es inmediato comprobar que

$$\mu(A) < \varepsilon \quad \text{y} \quad |f_n(\omega) - f(\omega)| < \frac{1}{k}, \qquad \forall \omega \in E \backslash A, n \geq m_k,$$

pues

$$\mu(A) \leq \sum_{k=1}^{\infty} \mu\left(E_{k,m_k}\right) < \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} = \varepsilon,$$

y

$$E\backslash A = E\backslash \bigcup_{k=1}^{\infty} E_{k,m_k} = \bigcap_{k=1}^{\infty}\Big(E\backslash E_{k,m_k}\Big) =$$

$$= \bigcap_{k=1}^{\infty}\Big\{\omega \in E : |f_n(\omega) - f(\omega)| < \frac{1}{k} \quad \text{si} \quad n \geq m_k\Big\}.$$

Tema 12

Teoremas de convergencia

Es usual en Análisis Matemático trabajar con funciones que están definidas como límite (en algún sentido) de una sucesión de funciones. Es fundamental, por tanto, ser capaces de deducir la integrabilidad de una tal función a partir del conocimiento de la integrabilidad de las funciones que la aproximan y, en su caso, relacionar el valor de la integral de la función límite con el valor de las integrales de las funciones aproximantes.

En esta lección obtendremos los teoremas de convergencia para la integral de Lebesgue. El Teorema de la convergencia monótona (TCM 12.1) y su primera consecuencia, el Teorema de la convergencia dominada (TCD 12.5), nos permitirán, en condiciones razonables, intercambiar el límite y la integral. El Lema de Fatou (12.7), otra consecuencia del T.C.M., nos proporcionará un resultado aceptable incluso careciendo de convergencia. Probaremos también la complitud del espacio $\mathscr{L}(\mu)$ (12.9) lo cual nos permitirá mostrar como $\mathscr{L}(\mu)$ se puede obtener a partir de $\mathbf{F}$ (el espacio de las funciones simples integrables) de igual modo que los números reales se obtienen a partir de los números racionales.

Para la integral de Lebesgue en $\mathbb{R}^N$ disponemos de un subespacio de $\mathbf{F}(\mathbb{R}^N)$ especialmente tangible: las funciones escalonadas. También disponemos de una clase especialmente destacada de funciones: las funciones continuas de soporte compacto. Terminaremos la lección probando el Teorema de densidad 12.12 que afirma que ambas clases de funciones son densas en el espacio $\mathscr{L}(\lambda)$ de las funciones integrables en $\mathbb{R}^N$.

Recordemos que si $(\Omega, \mathscr{A}, \mu)$ es un espacio de medida y $\{f_n\}$ es una sucesión de funciones de Ω en $\mathbb{R}$, se dice que $\{f_n\}$ converge casi por doquier, o bien que $\{f_n(\omega)\}$ converge para casi todo punto ω de Ω, si el conjunto de puntos $\omega \in \Omega$ en los que la sucesión $\{f_n(\omega)\}$ no es convergente es de medida cero. En ese caso, dada una función f de Ω en $\mathbb{R}$, se dice que la sucesión $\{f_n\}$ converge casi por doquier a f, o bien, que la sucesión $\{f_n(\omega)\}$ converge a $f(\omega)$ para casi todo $\omega \in \Omega$ (abreviadamente $\{f_n\} \to f$ c.p.d. o $f(\omega) = \lim f_n(\omega)$, ct $\omega \in \Omega$) si el conjunto

$$\{\omega \in \Omega : \{f_n(\omega)\} \text{ no converge a } f(\omega)\}$$

es de medida cero.

12.1. Teorema de la convergencia monótona .

En 1906 el matemático italiano Beppo Levi publicó un interesante resultado sobre la integración término a término de series de funciones positivas, una versión equivalente de dicho resultado se refiere a la integración término a término de sucesiones crecientes de funciones y es conocida como "Teorema de la convergencia monótona".

Teorema 12.1 (convergencia monótona (TCM))**.** *Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida. Si $\{f_n\}$ es una sucesión monótona de funciones integrables en Ω verificando que la sucesión de sus integrales $\left\{\int_\Omega f_n \, d\mu\right\}$ está acotada, entonces $\{f_n\}$ converge* c.p.d. *a una función f integrable en Ω y*

$$\int_\Omega f \, d\mu = \lim \int_\Omega f_n \, d\mu.$$

Demostración:

Sea $\{f_n\}$ una sucesión monótona de funciones integrables en Ω verificando que la sucesión de sus integrales $\left\{\int_\Omega f_n \, d\mu\right\}$ está acotada. Supongamos en primer lugar que $\{f_n\}$ es una sucesión creciente de funciones positivas y sea $M > 0$ tal que

$$\int_\Omega f_n \, d\mu \leq M, \ \forall n \in \mathbb{N}.$$

Como las funciones f_n son medibles, el conjunto $\{\omega \in \Omega : \ \lim f_n(\omega) \in \mathbb{R}\}$ es medible (Proposición 11.5.v), y por tanto el conjunto

$$E := \{\omega \in \Omega : \ \lim f_n(\omega) = +\infty\}$$

también es medible. Probaremos que $\mu(E) = 0$. Fijemos $K > 0$ y para cada n natural consideremos el conjunto medible

$$E_n = \{\omega \in \Omega : \ f_n(\omega) \geq K\}.$$

Se tiene que

$$K \ \mu(E_n) = \int_{E_n} K \, d\mu \leq \int_{E_n} f_n \, d\mu \leq \int_\Omega f_n \, d\mu \leq M,$$

y por tanto $\mu(E_n) \leq \dfrac{M}{K}$. Como la sucesión de funciones es creciente, entonces la sucesión de conjuntos $\{E_n\}$ también lo es, por tanto, por el crecimiento continuo de la medida sabemos que

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim \mu(E_n) \leq \frac{M}{K}.$$

Como $E \subset \bigcup_{n=1}^{\infty} E_n$, entonces $\mu(E) \leq \dfrac{M}{K}$. De la arbitrariedad de K, obtenemos que $\mu(E) = 0$.

Sea $f : \Omega \longrightarrow \mathbb{R}$ la función definida por

$$f(\omega) = \begin{cases} 0 \text{ si } \ \omega \in E \\ \lim f_n(\omega) \text{ si } \ \omega \in E^C \end{cases}.$$

Como $f = \lim f_n \chi_{E^C}$, f es medible y claramente $\{f_n\} \longrightarrow f$ c.p.d. Probaremos que

$$\int_\Omega f \, d\mu = \lim \int_\Omega f_n \, d\mu$$

(y en consecuencia la función $f = |f|$ será integrable). Fijado $n \in \mathbb{N}$ el conjunto

$$\{\omega \in \Omega : \; f_n(\omega) \neq \big(f_n \chi_{E^C}\big)(\omega)\} = \{\omega \in \Omega : \; f_n(\omega)\big(1 - \chi_{E^C}(\omega)\big) \neq 0\} =$$

$$E \cap \{\omega \in \Omega : f_n(\omega) \neq 0\}$$

es de medida cero y $f_n \chi_{E^C} \leq f$. En consecuencia

$$\int_\Omega f_n \, d\mu = \int_\Omega f_n \chi_{E^C} \, d\mu \leq \int_\Omega f \, d\mu,$$

y por tanto

$$\lim \int_\Omega f_n \, d\mu \leq \int_\Omega f \, d\mu.$$

Para probar la otra desigualdad, sea s una función simple positiva con $s \leq f$ y sea $0 < \rho < 1$. Para cada natural n definimos el conjunto

$$F_n := \{\omega \in \Omega : \; \rho s(\omega) \leq f_n(\omega)\}.$$

Obtenemos así una sucesión creciente de conjuntos medibles tal que $\bigcup_{n=1}^\infty F_n = \Omega$. En efecto, para cada $\omega \in \Omega \begin{cases} \text{si } f(\omega) = 0 \Rightarrow \omega \in F_1 \\ \text{si } f(\omega) > 0 \Rightarrow \rho s(\omega) < f(\omega) \; (\rho < 1) \; \Rightarrow \exists n \in \mathbb{N} : \; \omega \in F_n \end{cases}$ Supuesto que $s = \sum_{k=1}^m \alpha_k \chi_{A_k}$, para nada n natural se verifica que

$$\rho \sum_{k=1}^m \alpha_k \, \mu(A_k \cap F_n) = \sum_{k=1}^m \rho \; \alpha_k \, \mu(A_k \cap F_n) = \int_\Omega \rho \; s \; \chi_{F_n} \, d\mu \leq \int_\Omega f_n \; \chi_{F_n} \, d\mu \leq \int_\Omega f_n \, d\mu, \tag{12.1.1}$$

donde se han utilizado la definición de F_n y la monotonía de la integral. Para cada $k \in \{1, \ldots, m\}$, $\{A_k \cap F_n\}_{n \in \mathbb{N}}$ es una sucesión creciente de conjuntos medibles tal que $\bigcup_{n=1}^\infty \big(A_k \cap F_n\big) = A_k$, y por tanto se deduce del crecimiento continuo de la medida μ que $\mu(A_k) = \lim_n \mu(A_k \cap F_n)$. Tomando límite cuando $n \to \infty$ en 12.1.1 obtenemos en consecuencia que

$$\rho \sum_{k=1}^m \alpha_k \; \mu(A_k) \leq \lim \int_\Omega f_n \, d\mu,$$

es decir

$$\rho \int_\Omega s \, d\mu \leq \lim \int_\Omega f_n \, d\mu.$$

Tomando ahora en la desigualdad anterior límite cuando $\rho \to 1$ obtenemos que

$$\int_\Omega s \, d\mu \leq \lim \int_\Omega f_n \, d\mu.$$

Considerando por último la arbitrariedad de s, concluimos que

$$\int_\Omega f\,d\mu \le \lim \int_\Omega f_n\,d\mu.$$

Supongamos ahora que $\{f_n\}$ es una sucesión creciente sin ninguna otra limitación. La sucesión $\{f_n - f_1\}$ es también creciente y además $f_n - f_1 \ge 0$, $\forall n \in \mathbb{N}$. Como claramente la sucesión $\left\{\int_\Omega (f_n - f_1)\,d\mu\right\}$ está mayorada, podemos asegurar que

$$\{f_n - f_1\} \longrightarrow g \text{ c.p.d.}$$

para una cierta función integrable g y que

$$\int_\Omega g\,d\mu = \lim \int_\Omega (f_n - f_1)\,d\mu.$$

Definimos entonces $f = g + f_1$ y observamos que $\{f_n\} \longrightarrow f$ c.p.d. y que

$$\int_\Omega f\,d\mu = \int_\Omega (g + f_1)\,d\mu = \lim \int_\Omega (f_n - f_1)\,d\mu + \int_\Omega f_1\,d\mu = \lim \int_\Omega f_n\,d\mu.$$

Por último, si la sucesión $\{f_n\}$ es decreciente notemos que basta aplicar lo ya demostrado a la sucesión $\{-f_n\}$. En efecto, si $\{-f_n\} \longrightarrow g$ c.p.d. con g integrable y

$$\int_\Omega g\,d\mu = \lim \int_\Omega (-f_n)\,d\mu,$$

entonces $\{f_n\} \longrightarrow f = -g$ c.p.d. y

$$\int_\Omega f\,d\mu = \int_\Omega (-g)\,d\mu = -\int_\Omega g\,d\mu =$$
$$-\lim \int_\Omega (-f_n)\,d\mu = -\lim\left(-\int_\Omega f_n\,d\mu\right) = \lim \int_\Omega f_n\,d\mu.$$

■

Es conveniente codificar el siguiente resultado práctico que será utilizado en la prueba del Teorema de la convergencia creciente para funciones medibles positivas 12.3.

Corolario 12.2 (versión práctica del TCM)**.** *Sean $(\Omega, \mathscr{A}, \mu)$ un espacio de medida y $\{f_n\}$ es una sucesión monótona de funciones integrables en Ω verificando que la sucesión de sus integrales $\left\{\int_\Omega f_n\,d\mu\right\}$ está acotada. Si $f : \Omega :\to \mathbb{R}$ es una función medible tal que $\{f_n\} \longrightarrow f$* c.p.d.*, entonces f es integrable en Ω y*

$$\int_\Omega f\,d\mu = \lim \int_\Omega f_n\,d\mu.$$

Demostración:

Por el Teorema 12.1 existe $g : \Omega \to \mathbb{R}$ integrable tal que

$$\{f_n\} \longrightarrow g \text{ c.p.d. y } \int_\Omega \mathrm{g}\, \mathrm{d}\mu = \lim \int_\Omega \mathrm{f_n}\, \mathrm{d}\mu.$$

Comprobemos que $f = g$ c.p.d. Si

$$A = \{\omega \in \Omega : \; g(\omega) = \lim f_n(\omega)\} \text{ y } B = \{\omega \in \Omega : \; f(\omega) = \lim f_n(\omega)\} \, ,$$

entonces $A, B \in \mathscr{A}$ y $\mu(A^C) = 0 = \mu(B^C)$. De la inclusión

$$A \cap B \subset \{\omega \in \Omega : f(\omega) = g(\omega)\},$$

se obtiene que $\{\omega \in \Omega : f(\omega) \neq g(\omega)\} \subset (A \cap B)^c = A^c \cup B^c$, y el primer conjunto, que es medible, tiene medida cero. Como f y g son medibles, y hemos obtenido que coinciden c.p.d., se sigue que f es integrable y su integral coincide con la de g. ∎

Corolario 12.3 (teorema de la conver. crec. para func. medibles positivas TCC))**.** *Sea* $(\Omega, \mathscr{A}, \mu)$ *un espacio de medida. Si* $\{f_n\}$ *es una sucesión creciente de funciones medibles positivas en* Ω *que converge* c.p.d. *a una función medible positiva* f*, entonces*

$$\int_\Omega f \, d\mu = \lim \int_\Omega f_n \, d\mu.$$

Demostración:

La sucesión $\left\{ \int_\Omega f_n \, d\mu \right\}$ es creciente. Si es acotada, el resultado se sigue del Corolario anterior. Si no es acotada, entonces $\lim \int_\Omega f_n \, d\mu = \infty$ y al verificarse que

$$\int_\Omega f_n \, d\mu \leq \int_\Omega f \, d\mu, \; \forall n \in \mathbb{N},$$

deducimos que también $\int_\Omega f \, d\mu = \infty$. ∎

La última consecuencia inmediata del Teorema 12.1 que resaltamos nos asegura la continuidad absoluta de una medida definida por integración de una función medible positiva.

Corolario 12.4. *Sean* $(\Omega, \mathscr{A}, \mu)$ *un espacio de medida y* f *una función medible positiva en* Ω*. Entonces* $\nu : \mathscr{A} \longrightarrow [0, \infty]$ *definida por*

$$\nu(E) = \int_E f \, d\mu, \; \forall E \in \mathscr{A}$$

es una medida <u>*absolutamente continua*</u> *respecto de* μ*, es decir:*

$$[E \in \mathscr{A}, \mu(E) = 0] \Longrightarrow \nu(E) = 0.$$

Además si f *es integrable, entonces: dado* $\varepsilon > 0$ *existe* $\delta > 0$ *de manera que* $\nu(E) < \varepsilon$ *para todo conjunto medible* E *con* $\mu(E) < \delta$ *.*

Demostración:

Sea $E = \bigcup_{n\in\mathbb{N}} E_n$, donde $\{E_n\}$ es una sucesión de elementos disjuntos de $\mathscr{A}$. Obsérvese que

$$f\,\chi_E = f\sum_{n=1}^{\infty}\chi_{E_n} = \sum_{n=1}^{\infty} f\,\chi_{E_n},$$

$$\nu(E) = \int_E f\,d\mu = \int_\Omega f\,\chi_E\,d\mu \quad \text{y}$$

$$\nu(E_n) = \int_{E_n} f\,d\mu = \int_\Omega f\,\chi_{E_n}\,d\mu$$

donde se ha utilizado que la integral extendida a un conjunto medible de una función f medible positiva coincide con la integral de la prolongación por cero de f fuera de dicho conjunto. Se tiene que

$$\nu(E) = \int_\Omega f\,\chi_E\,d\mu = \int_\Omega \lim\Big(\sum_{k=1}^{n} f\,\chi_{E_k}\Big)d\mu =$$

$$\lim\int_\Omega\Big(\sum_{k=1}^{n} f\,\chi_{E_k}\Big)d\mu = \lim\sum_{k=1}^{n}\int_\Omega f\,\chi_{E_k}\,d\mu = \sum_{k=1}^{\infty}\int_\Omega f\,\chi_{E_k}\,d\mu = \sum_{k=1}^{\infty}\nu(E_k)$$

donde se ha utilizado el Corolario anterior. Así ν es una medida. Por último si $\mu(E) = 0$, entonces el apartado iv) de la Proposición 11.13 nos asegura que $\int_\Omega f\,\chi_E\,d\mu = 0$, esto es, $\nu(E) = 0$. Hemos probado que ν es absolutamente continua respecto μ.

Supongamos ahora que f es una función integrable. Sea $\varepsilon > 0$. Para cada natural n sea $A_n = \{\omega\in\Omega:\ f(\omega)\le n\}$. Se tiene que $A_n\in\mathscr{A}$, $\forall n\in\mathbb{N}$ y que $f\,\chi_{A_n}\nearrow f$. El Corolario 12.2 garantiza que

$$\int_\Omega f\,\chi_{A_n}\,d\mu \nearrow \int_\Omega f\,d\mu.$$

Existe entonces $n\in\mathbb{N}$ tal que $\int_\Omega\big(f - f\,\chi_{A_n}\big)\,d\mu < \dfrac{\varepsilon}{2}$. Tomamos $\delta = \dfrac{\varepsilon}{2n}$, si $E\in\mathscr{A}$ con $\mu(E) < \delta$, entonces

$$\begin{aligned}\nu(E) &= \int_E f\,d\mu = \int_E\big(f - f\,\chi_{A_n}\big)d\mu + \int_E f\,\chi_{A_n}\,d\mu\\ &\le \int_\Omega\big(f - f\,\chi_{A_n}\big)d\mu + \int_E n\,d\mu < \frac{\varepsilon}{2} + n\frac{\varepsilon}{2n} = \varepsilon.\end{aligned}$$

■

En la lección siguiente se verá un ejemplo que muestra que es esencial exigir que f sea integrable para que ν verifique la última propiedad del corolario anterior.

12.2. Teorema de la convergencia dominada y Lema de Fatou

En la práctica no siempre nos encontramos con sucesiones monótonas de funciones, por lo que el Teorema 12.1 no puede ser utilizado en tales ocasiones. Afortunadamente, la limitación que impone la monotonía se suprime en el siguiente teorema de convergencia, que a cambio exige la condición de dominación de la sucesión. La dominación suele acaecer con frecuencia y normalmente no es difícil su comprobación.

Teorema 12.5 (convergencia dominada (TCD))**.** *Sea* $(\Omega, \mathscr{A}, \mu)$ *un espacio de medida. Sea* $\{f_n\}$ *una sucesión de funciones integrables que converge* c.p.d. *en* Ω *a una función* _medible_ f *y supongamos que existe una función integrable* g *tal que*

$$|f_n| \leq g, \text{ c.p.d. }, \forall \text{n} \in \mathbb{N}.$$

Entonces f *es integrable y*

$$\lim \int_\Omega |f - f_n| \, d\mu = 0.$$

En consecuencia

$$\int_\Omega f \, d\mu = \lim \int_\Omega f_n \, d\mu.$$

Demostración:

Sea $E \in \mathscr{A}$ con $\mu(E^C) = 0$ tal que para cada $\omega \in E$ se verifica

$$\{f_n(\omega)\} \longrightarrow f(\omega) \text{ y } |f_n(\omega)| \leq g(\omega), \, \forall n \in \mathbb{N}.$$

Para cada natural n es $|f_n \, \chi_E| \leq g \, \chi_E$ y en consecuencia $|f \, \chi_E| \leq g \, \chi_E$ con lo que $f \, \chi_E$ es integrable y, por tanto, f es integrable.

Para cada natural n, sea $g_n : E \longrightarrow \mathbb{R}$ la función definida por

$$g_n(\omega) = \sup\{|f(\omega) - f_k(\omega)| : \, k \geq n\}.$$

Se tiene que $0 \leq g_n \, \chi_E \leq 2g \, \chi_E$, con lo que $g_n \, \chi_E$, que es medible, también es integrable. Como $\{g_n \, \chi_E\}$ es una sucesión decreciente de funciones integrables positivas que converge a cero, el Corolario 12.2 garantiza que

$$\lim \int_\Omega g_n \, \chi_E \, d\mu = 0.$$

Como $|f - f_n| \chi_E \leq g_n \, \chi_E$, $\forall n \in \mathbb{N}$, se tiene que

$$0 \leq \int_\Omega |f - f_n| d\mu = \int_\Omega |f - f_n| \chi_E \, d\mu \leq \int_\Omega g_n \, \chi_E \, d\mu, \, \forall n \in \mathbb{N}$$

y por tanto concluimos que

$$\lim \int_\Omega |f - f_n| \, d\mu = 0.$$

Finalmente, al verificarse

$$\left|\int_\Omega f\,d\mu - \int_\Omega f_n\,d\mu\right| = \left|\int_\Omega (f-f_n)\,d\mu\right| \le \int_\Omega |f-f_n|\,d\mu,$$

se tiene en consecuencia que

$$\int_\Omega f\,d\mu = \lim \int_\Omega f_n\,d\mu.$$

■

Nota 12.6. Si $(\Omega, \mathscr{A}, \mu)$ es un espacio de medida completo no hay que exigir la medibilidad de f en el teorema (véase el apartado v) de la Proposición 11.5 y el Ejemplo 11.2.5).

Mostraremos seguidamente que, incluso careciendo de convergencia puntual de la sucesión de funciones, podemos conseguir un resultado del mismo tipo.

Teorema 12.7 (Lema de Fatou)**.** *Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida. Sea $\{f_n\}$ una sucesión de funciones medibles y positivas en Ω tal que*

$$\liminf \int_\Omega f_n\,d\mu < \infty.$$

Entonces, existe una función f integrable en Ω tal que

$$\liminf\, f_n = f \text{ c.p.d. } \textit{y} \quad \int_\Omega f\,d\mu \le \liminf \int_\Omega f_n\,d\mu.$$

Demostración:

Para cada natural n, consideremos la función $g_n : \Omega \longrightarrow \mathbb{R}$ definida por

$$g_n(\omega) = \inf\{f_k(\omega) : \; k \ge n\}.$$

Se tiene que $0 \le g_n \le f_k$, para $k \ge n$, luego $0 \le \int_\Omega g_n\,d\mu \le \int_\Omega f_k\,d\mu$, para $k \ge n$, y en consecuencia $0 \le \int_\Omega g_n\,d\mu \le \inf\left\{\int_\Omega f_k\,d\mu : \; k \ge n\right\}$ y, por tanto,

$$0 \le \int_\Omega g_n\,d\mu \le \liminf \int_\Omega f_n\,d\mu < \infty.$$

Como la sucesión $\{g_n\}$ es creciente, el Teorema 12.1 nos asegura que $\{g_n\}$ converge c.p.d. hacia una función integrable f y que $\int_\Omega f\,d\mu = \lim \int_\Omega g_n\,d\mu$. Hemos probado que

$$\int_\Omega f\,d\mu \le \liminf \int_\Omega f_n\,d\mu.$$

■

En el Ejercicio 12.2.i) se da un ejemplo de una sucesión de funciones que cumple la hipótesis del Lema de Fatou para la cual la desigualdad del enunciado es estricta.

12.3. Teorema de la convergencia absoluta

Con frecuencia nos encontramos con el problema de integrar funciones definidas como la suma de una serie de funciones (piénsese por ejemplo en las funciones definidas por una serie de potencias). La conjunción de los Teoremas 12.1 y 12.5 nos permiten a continuación establecer un procedimiento muy natural para resolver este problema de integración.

Teorema 12.8 (convergencia absoluta (TCA))**.** *Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida. Sea $\sum_{n\geq 1} f_n$ una serie de funciones integrables en Ω y supongamos que*

$$\sum_{n=1}^{\infty} \int_{\Omega} |f_n| \, d\mu < \infty.$$

Entonces la serie $\sum_{n\geq 1} f_n$ converge absolutamente c.p.d.*, existe una función f integrable en Ω tal que $\sum_{n=1}^{\infty} f_n = f$* c.p.d. *y además*

$$\lim \int_{\Omega} \Big| f - \sum_{k=1}^{n} f_k \Big| \, d\mu = 0.$$

En consecuencia

$$\int_{\Omega} f \, d\mu = \sum_{n=1}^{\infty} \int_{\Omega} f_n \, d\mu.$$

Demostración:

Para cada $n \in \mathbb{N}$, sea $G_n = \sum_{k=1}^{n} |f_k|$. $\{G_n\}$ es una sucesión creciente de funciones integrables tal que

$$\int_{\Omega} G_n \, d\mu = \sum_{k=1}^{n} \int_{\Omega} |f_k| \, d\mu \leq \sum_{n=1}^{\infty} \int_{\Omega} |f_n| \, d\mu < \infty.$$

El Teorema (12.1) garantiza, en consecuencia, la existencia de una función integrable G tal que $\{G_n\}$ converge a G c.p.d. Este hecho nos permite afirmar que la serie $\sum_{n\geq 1} f_n$ converge absolutamente c.p.d., esto es el conjunto medible

$$E = \Big\{ \omega \in \Omega : \sum_{n=1}^{\infty} |f_n(\omega)| < \infty \Big\}$$

verifica que $\mu(E^C) = 0$. Como toda serie absolutamente convergente es convergente podemos definir la función medible $f : \Omega \longrightarrow \mathbb{R}$ por

$$f(\omega) = \begin{cases} \sum_{n=1}^{\infty} f_n(\omega) \text{ si } \omega \in E \\ 0 \qquad\qquad\quad \text{si } \omega \in E^C \end{cases},$$

es decir,

$$f = \sum_{n=1}^{\infty} f_n \, \chi_E.$$

Como el conjunto medible $\left\{\omega \in \Omega : \sum_{n=1}^{\infty} f_n(\omega) \neq f(\omega)\right\}$ es un subconjunto de E^C, se sigue que es de medida cero, y en consecuencia

$$\sum_{n=1}^{\infty} f_n = f \text{ c.p.d.}$$

Veamos que la función f cumple las condiciones del teorema. Para ello consideremos la sucesión $\{F_n\}$ de funciones integrables en Ω definida por

$$F_n = \sum_{k=1}^{n} f_k, \ \forall n \in \mathbb{N}.$$

Obsérvese que $\{F_n\} \longrightarrow f$ c.p.d. y que es fácil comprobar que $|F_n| \leq G$ c.p.d. , $\forall n \in \mathbb{N}$. El Teorema (12.5) nos permite concluir que f es integrable y que

$$\lim \int_{\Omega} |f - F_n| \, d\mu = 0.$$

Por tanto,

$$\int_{\Omega} f \, d\mu = \lim \int_{\Omega} F_n \, d\mu = \lim \int_{\Omega} \Big(\sum_{k=1}^{n} f_k\Big) d\mu = \lim \sum_{k=1}^{n} \int_{\Omega} f_k \, d\mu = \sum_{n=1}^{\infty} \int_{\Omega} f_n \, d\mu$$

■

12.4. Teorema de Riesz.

Los teoremas de convergencia precedentes culminan ahora con el siguiente resultado.

Teorema 12.9 (Riesz)**.** *Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida. Sea $\{f_n\}$ una sucesión de funciones integrables en Ω tal que*

$$\lim_{p,q \longrightarrow \infty} \int_{\Omega} |f_p - f_q| \, d\mu = 0,$$

es decir:

$$\forall \varepsilon > 0, \ \exists m \in \mathbb{N} : \ p, q \geq m \Longrightarrow \int_{\Omega} |f_p - f_q| \, d\mu < \varepsilon.$$

Existe entonces una función f integrable en Ω tal que

$$\lim \int_{\Omega} |f - f_n| \, d\mu = 0$$

y existe, además, una sucesión parcial $\{f_{\sigma(n)}\}$ que converge c.p.d. *a f.*

Demostración:

La condición satisfecha por $\{f_n\}$ permite construir una sucesión parcial $\{f_{\sigma(n)}\}$ de $\{f_n\}$ verificando

$$\int_\Omega |f_{\sigma(n+1)} - f_{\sigma(n)}|\, d\mu < \frac{1}{2^n},\ \forall n \in \mathbb{N}. \tag{12.4.1}$$

En efecto, definamos por inducción una aplicación $\sigma : \mathbb{N} \longrightarrow \mathbb{N}$ de la siguiente forma: $\sigma(1)$ es el menor de los números naturales k para los cuales se verifica que:

$$\int_\Omega |f_p - f_q|\, d\mu < \frac{1}{2} \text{ siempre que } p, q \geq k.$$

Dado $n \in \mathbb{N}$ supongamos definidos $\sigma(1) < \sigma(2) < \cdots < \sigma(n)$ números naturales cumpliendo que para $j = 1, 2, \ldots, n$ se verifica

$$\int_\Omega |f_p - f_q|\, d\mu < \frac{1}{2^j} \text{ siempre que } p, q \geq \sigma(j).$$

La condición satisfecha por $\{f_n\}$ nos permite afirmar que el conjunto

$$\left\{ k \in \mathbb{N} :\ k > \sigma(n) \text{ y } \int_\Omega |f_p - f_q|\, d\mu < \frac{1}{2^{n+1}} \text{ para } q, p \geq k \right\}$$

no es vacío. Definimos $\sigma(n+1)$ como el mínimo de dicho conjunto. Es claro que la aplicación σ es estrictamente creciente y además verifica las desigualdades 12.4.1.

El Teorema 12.8 garantiza entonces que la serie de funciones $\sum_{n\geq 1} F_n$ definida por

$$F_1 = f_{\sigma(1)} \quad , \quad F_{n+1} = f_{\sigma(n+1)} - f_{\sigma(n)},\ \forall n \in \mathbb{N}$$

converge c.p.d. a una función f integrable en Ω y además

$$\lim \int_\Omega \left| f - \sum_{k=1}^{n} F_k \right| d\mu = 0.$$

Obsérvese que $\sum_{k=1}^{n} F_k = f_{\sigma(n)},\ \forall n \in \mathbb{N}$. Por tanto $\{f_{\sigma(n)}\}$ converge c.p.d. a f y

$$\lim \int_\Omega |f - f_{\sigma(n)}|\, d\mu = 0.$$

De esto último deducimos que

$$\lim \int_\Omega |f - f_n|\, d\mu = 0.$$

En efecto, sea $\varepsilon > 0$ y sea $m \in \mathbb{N}$ tal que

$$\left[p, q \geq m \Longrightarrow \int_\Omega |f_p - f_q|\, d\mu < \frac{\varepsilon}{2}\right] \text{ y } \left[n \geq m \Longrightarrow \int_\Omega |f - f_{\sigma(n)}|\, d\mu < \frac{\varepsilon}{2}\right].$$

Entonces para $n \geq m$ tenemos que

$$\int_\Omega |f - f_n|\, d\mu \leq \int_\Omega |f - f_{\sigma(n)}|\, d\mu + \int_\Omega |f_{\sigma(n)} - f_n|\, d\mu < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

donde hemos utilizado que $n \leq \sigma(n)$, $\forall n \in \mathbb{N}$. ■

La aplicación $\|\cdot\|_1 : \mathscr{L}(\mu) \longrightarrow \mathbb{R}_0^+$ definida por

$$\|f\|_1 = \int_\Omega |f| \, d\mu$$

satisface claramente las propiedades

1. $\|\alpha f\|_1 = |\alpha| \|f\|_1$, $\forall \alpha \in \mathbb{R}$, $\forall f \in \mathscr{L}(\mu)$ *(homogeneidad).*
2. $\|f + g\|_1 \leq \|f\|_1 + \|g\|_1$, $\forall f, g \in \mathscr{L}(\mu)$ *(desigualdad triangular)* .

La única deficiencia de $\|\cdot\|_1$ (para ser una norma) es que puede ser $\|f\|_1 = 0$ sin que la función f sea cero. Este problema desaparece al considerar el espacio $\mathbf{L}_1(\mu)$ definido como sigue

$$\mathbf{L}_1(\mu) = \mathscr{L}(\mu)/\mathbf{N}(\mu)$$

donde $\mathbf{N}(\mu)$ es el subespacio vectorial de $\mathscr{L}(\mu)$ definido por

$$\mathbf{N}(\mu) = \{f : \Omega \longrightarrow \mathbb{R} : \; f \text{ es medible y } f = 0 \text{ c.p.d.}\}.$$

En consecuencia,

$$f + \mathbf{N}(\mu) = g + \mathbf{N}(\mu)$$

si, y sólo si,

$$f = g \text{ c.p.d., esto es, } \; \mu(\{\omega \in \Omega : \; f(\omega) \neq g(\omega)\}) = 0.$$

En el caso de que $(\Omega, \mathscr{A}, \mu)$ sea un espacio de medida completo la descripción de $\mathbf{N}(\mu)$ es más simple:

$$\mathbf{N}(\mu) = \{f : \Omega \longrightarrow \mathbb{R} : \; f = 0 \; \text{ c.p.d.}\}.$$

Esencialmente $\mathbf{L}_1(\mu)$ no es otra cosa que que el mismo espacio $\mathscr{L}(\mu)$ en el que se considera la igualdad c.p.d. en lugar de la igualdad ordinaria de funciones. De manera natural $\|\cdot\|_1$ se convierte en una norma en $\mathbf{L}_1(\mu)$ sin más que definir $\|f + \mathbf{N}(\mu)\|_1 = \|f\|_1$ (¡Compruébese!).

La convergencia respecto de la norma $\|\cdot\|_1$ se denomina <u>*convergencia en media*</u>. El Teorema de Riesz proporciona la complitud de $\mathbf{L}_1(\mu)$.

Teorema 12.10 (complitud de $\mathbf{L}_1(\mu)$)**.** *Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida. Entonces $\big(\mathbf{L}_1(\mu), \|\cdot\|_1\big)$ es un espacio de Banach.*

Ejemplo: El espacio de Banach $(\ell_1, \|\cdot\|_1)$.

Consideremos el caso particular en que $\Omega = \mathbb{N}$, $\mathscr{A} = \mathscr{P}(\mathbb{N})$, y μ es la medida definida por

$$\mu(A) = \begin{cases} \text{número de elementos de } A & \text{si } A \text{ es finito} \\ \infty & \text{si } A \text{ no es finito} \end{cases}$$

para A subconjunto de $\mathbb{N}$. Entonces toda función (sucesión) $x:\mathbb{N}\to\mathbb{R}$ es medible. Calculemos la integral de $|x|$ para $x:\mathbb{N}\to\mathbb{R}$. Como

$$|x_1|\chi_{\{1\}}+\cdots+|x_n|\chi_{\{n\}}\nearrow|x|,$$

el Teorema de la convergencia creciente para funciones medibles positivas (12.3) nos dice que

$$\int_{\mathbb{N}}[|x_1|\chi_{\{1\}}+\cdots+|x_n|\chi_{\{n\}}]\,d\mu=\sum_{k=1}^{n}|x_k|\nearrow\int_{\mathbb{N}}|x|\,d\mu,$$

esto es,

$$\int_{\mathbb{N}}|x|\,d\mu=\sum_{n=1}^{\infty}|x_n|.$$

Así, el espacio $\mathscr{L}(\mu)$ es el conjunto de las series absolutamente convergentes. Como sólo el conjunto vacío tiene medida cero, en este caso no hay diferencia entre $\mathbf{L}_1(\mu)$ y $\mathscr{L}(\mu)$ pues $\mathbf{N}(\mu)=\{0\}$. Al espacio $\mathbf{L}_1(\mu)$ resultante lo notaremos ℓ_1. Una sucesión $x:\mathbb{N}\to\mathbb{R}$ está en ℓ_1 si, y sólo si, la serie $\sum_{n\geq 1}|x_n|$ es convergente, esto es,

$$\ell_1=\left\{x:\mathbb{N}\to\mathbb{R}:\sum_{n=1}^{\infty}|x_n|<\infty\right\}.$$

Ahora el Teorema de la convergencia dominada (12.5) nos permite calcular la integral de $x\in\ell_1$. En efecto, como

$$x_1\chi_{\{1\}}+\cdots+x_n\chi_{\{n\}}\to x,$$

y

$$|x_1\chi_{\{1\}}+\cdots+x_n\chi_{\{n\}}|\leq|x|,\ \forall n\in\mathbb{N}$$

concluimos que

$$\int_{\mathbb{N}}[x_1\chi_{\{1\}}+\cdots+x_n\chi_{\{n\}}]\,d\mu=\sum_{k=1}^{n}x_k\to\int_{\mathbb{N}}x\,d\mu,$$

esto es,

$$\int_{\mathbb{N}}x\,d\mu=\sum_{n=1}^{\infty}x_n.$$

Finalmente el Teorema de Riesz (12.9) nos asegura que la norma

$$\|x\|_1:=\sum_{n=1}^{\infty}|x_n|\ \ \forall x\in\ell_1$$

es completa.

12.5. Subespacios densos en $\mathscr{L}(\mathbb{R}^N)$.

Cuando tratamos la integral de Lebesgue en $\mathbb{R}^N$ podemos considerar dos clases especialmente destacadas de funciones: las funciones escalonadas y las funciones continuas de soporte compacto. Como anunciamos en la introducción probaremos, como primera consecuencia de los teoremas de convergencia, la densidad en $\big(\mathscr{L}(\lambda), \|\cdot\|_1\big)$ de ambas clases de funciones.

Definición 12.11. $h : \mathbb{R}^N \longrightarrow \mathbb{R}$ es una función escalonada si existen m natural, $\alpha_1, \ldots, \alpha_m$ reales e $I_1, \ldots, I_m$ intervalos acotados tales que h se puede escribir de la forma

$$h = \sum_{k=1}^{m} \alpha_k \, \chi_{I_k}.$$

Es claro que el conjunto $\mathbf{E}(\mathbb{R}^N)$ de todas las funciones escalonadas en $\mathbb{R}^N$ es un subespacio del espacio $\mathbf{F}(\mathbb{R}^N)$ de las funciones simples integrables y que

$$\int h \, d\lambda = \sum_{k=1}^{m} \alpha_k \, v(I_k).$$

Teorema 12.12 (densidad)**.**

i) *El espacio vectorial $\mathbf{C}_{00}(\mathbb{R}^N)$ de las funciones continuas de soporte compacto es denso en $\big(\mathscr{L}(\lambda), \|\cdot\|_1\big)$, es decir, dada una función integrable f existe una sucesión $\{f_n\}$ en $\mathbf{C}_{00}(\mathbb{R}^N)$ tal que*

$$\lim \|f - f_n\|_1 = 0.$$

ii) *El espacio vectorial $\mathbf{E}(\mathbb{R}^N)$ de las funciones escalonadas es denso en $\big(\mathscr{L}(\lambda), \|\cdot\|_1\big)$, es decir, dada una función integrable f existe una sucesión $\{f_n\}$ en $\mathbf{E}(\mathbb{R}^N)$ tal que*

$$\lim \|f - f_n\|_1 = 0.$$

Además se puede conseguir en ambos casos que la sucesión converja también c.p.d.

Demostración:
Teniendo en cuenta el Teorema 11.20 bastará probar que dados $E \in \mathscr{M}$ con $\lambda(E) < \infty$ y $\varepsilon > 0$, existe $g \in \mathbf{C}_{00}(\mathbb{R}^N)$(resp. $h \in \mathbf{E}(\mathbb{R}^N)$) tal que

$$\|\chi_E - g\|_1 < \varepsilon \quad (\text{resp.} \quad \|\chi_E - h\|_1 < \varepsilon).$$

$\boxed{\textbf{i}}$) La regularidad interior de la medida de Lebesgue nos asegura que existe un compacto $K \subset E$ tal que $\lambda(E \setminus K) < \dfrac{\varepsilon}{2}$. Sea G un abierto acotado tal que $K \subset G$. La función $\varphi : \mathbb{R}^N \longrightarrow \mathbb{R}$ definida por

$$\varphi(x) = \frac{\operatorname{dist}(x, \mathbb{R}^N \setminus G)}{\operatorname{dist}(x, \mathbb{R}^N \setminus G) + \operatorname{dist}(x, K)}, \ \forall x \in \mathbb{R}^N,$$

es continua con $0 \le \varphi \le 1$ y además $\varphi(x) = \begin{cases} 1 & \text{si } x \in K \\ 0 & \text{si } x \notin G \end{cases}$, $\forall x \in \mathbb{R}^N$. Es claro que $\{\varphi^n\}$ es una sucesión decreciente en $\mathbf{C}_{00}(\mathbb{R}^N)$ que converge a χ_K. El Corolario 12.2 nos asegura que existe $m \in \mathbb{N}$ tal que $\|\chi_K - \varphi^m\|_1 < \dfrac{\varepsilon}{2}$. Tomemos $g = \varphi^m$. Se verifica entonces

$$\|\chi_E - g\|_1 \le \|\chi_E - \chi_K\|_1 + \|\chi_K - g\|_1 = \lambda(E \setminus K) + \|\chi_K - g\|_1 < \varepsilon.$$

$\boxed{\textbf{ii}}$) La regularidad exterior de la medida de Lebesgue nos asegura que existe un abierto G tal que $E \subset G$ y $\lambda(G \setminus E) < \dfrac{\varepsilon}{2}$. La Proposición 10.8 (de descomposición canónica de un abierto en cubos diádicos) nos asegura que existe una sucesión $\{I_n\}$ de intervalos acotados disjuntos dos a dos tal que $G = \bigcup_{n=1}^{\infty} I_n$. Es claro que $\left\{\sum_{k=1}^{n} \chi_{I_k}\right\}$ es una sucesión creciente en $\mathbf{E}(\mathbb{R}^N)$ que converge a χ_G . Como $\lambda(G) = \lambda(G \setminus E) + \lambda(E) < \infty$ el Corolario 12.2 nos asegura ahora que existe $m \in \mathbb{N}$ tal que $\left\|\chi_G - \sum_{k=1}^{m} \chi_{I_k}\right\|_1 < \dfrac{\varepsilon}{2}$. Tomemos $h = \sum_{k=1}^{m} \chi_{I_k}$. Se verifica entonces

$$\|\chi_E - h\|_1 \le \|\chi_E - \chi_G\|_1 + \|\chi_G - h\|_1 = \lambda(G \setminus E) + \|\chi_G - h\|_1 < \varepsilon.$$

Ahora es clara la densidad de $\mathbf{C}_{00}(\mathbb{R}^N)$ y $\mathbf{E}(\mathbb{R}^N)$ en $\mathcal{L}(\lambda)$ así como que, para una función integrable f dada, por el procedimiento habitual podemos construir una sucesión $\{f_n\}$ en dichos espacios tal que

$$\lim \|f - f_n\|_1 = 0.$$

Aplicando a dicha sucesión el Teorema de Riesz (12.9) se obtiene una función integrable f_0 y una sucesión parcial $\{f_{\sigma(n)}\}$ de $\{f_n\}$ tales que

$$\|f_0 - f_n\|_1 \longrightarrow 0 \quad \text{y} \quad \{f_{\sigma(n)}\} \longrightarrow f_0 \text{ c.p.d.}$$

Como

$$\|f_0 - f\| \le \|f_0 - f_n\|_1 + \|f_n - f\|_1, \ \forall n \in \mathbb{N},$$

deducimos que $\|f - f_0\|_1 = 0$ y en consecuencia $f = f_0$ c.p.d. Así la sucesión de funciones $\{f_{\sigma(n)}\}$ de $\mathbf{C}_{00}(\mathbb{R}^N)$ (resp. $\mathbf{E}(\mathbb{R}^N)$) verifica

$$\|f - f_{\sigma(n)}\|_1 \longrightarrow 0 \quad \text{y} \quad \{f_{\sigma(n)}\} \longrightarrow f \text{ c.p.d.}$$

∎

12.6. Resumen del resultados del Tema 12.

Teorema de la convergencia monótona (TCM). *Sea* $(\Omega, \mathscr{A}, \mu)$ *un espacio de medida. Si* $\{f_n\}$ *es una sucesión monótona de funciones integrables en* Ω *verificando que la sucesión de sus integrales* $\left\{ \int_\Omega f_n \, d\mu \right\}$ *está acotada, entonces* $\{f_n\}$ *converge* c.p.d. *a una función* f *integrable en* Ω *y*

$$\int_\Omega f \, d\mu = \lim \int_\Omega f_n \, d\mu.$$

Corolario (versión práctica del TCM). *Sea* $(\Omega, \mathscr{A}, \mu)$ *un espacio de medida. Sea* $\{f_n\}$ *una sucesión monótona de funciones integrables en* Ω *verificando que la sucesión de sus integrales* $\left\{ \int_\Omega f_n \, d\mu \right\}$ *está acotada. Si* $f : \Omega \to \mathbb{R}$ *es una función medible tal que* $\{f_n\} \longrightarrow f$ c.p.d., *entonces* f *es integrable en* Ω *y*

$$\int_\Omega f \, d\mu = \lim \int_\Omega f_n \, d\mu.$$

Corolario (teorema de la convergencia creciente para funciones medibles positivas (TCC). *Sea* $(\Omega, \mathscr{A}, \mu)$ *un espacio de medida. Si* $\{f_n\}$ *es una sucesión creciente de funciones medibles positivas en* Ω *que converge* c.p.d. *a una función medible positiva* f, *entonces*

$$\int_\Omega f \, d\mu = \lim \int_\Omega f_n \, d\mu.$$

Corolario. *Sean* $(\Omega, \mathscr{A}, \mu)$ *un espacio de medida y* f *una función medible positiva en* Ω. *Entonces* $\nu : \mathscr{A} \longrightarrow [0, \infty]$ *definida por*

$$\nu(E) = \int_E f \, d\mu, \ \forall E \in \mathscr{A}$$

es una medida <u>*absolutamente continua*</u> *respecto de* μ *, es decir:*

$$[E \in \mathscr{A}, \mu(E) = 0] \Longrightarrow \nu(E) = 0.$$

Además si f *es integrable, entonces, dado* $\varepsilon > 0$ *existe* $\delta > 0$ *de manera que* $\nu(E) < \varepsilon$ *para todo conjunto medible* E *con* $\mu(E) < \delta$ *.*

Teorema de la convergencia dominada (TCD). *Sea* $(\Omega, \mathscr{A}, \mu)$ *un espacio de medida. Sea* $\{f_n\}$ *una sucesión de funciones integrables que converge* c.p.d. *en* Ω *a una función* <u>*medible*</u> f *y supongamos que existe una función integrable* g *tal que*

$$|f_n| \leq g, \text{ c.p.d. }, \forall \text{n} \in \mathbb{N}.$$

Entonces f *es integrable y*

$$\lim \int_\Omega |f - f_n| \, d\mu = 0.$$

En consecuencia

$$\int_\Omega f\, d\mu = \lim \int_\Omega f_n\, d\mu.$$

Nota. Si $(\Omega, \mathscr{A}, \mu)$ es un espacio de medida completo no hay que exigir la medibilidad de f en el T.C.D.

Lema de Fatou. *Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida. Si $\{f_n\}$ es una sucesión de funciones medibles y positivas en Ω tal que*

$$\liminf \int_\Omega f_n\, d\mu < \infty,$$

entonces existe una función f integrable en Ω tal que

$$\liminf\, f_n = f \text{ c.p.d.} \;\; y \quad \int_\Omega f\, d\mu \le \liminf \int_\Omega f_n\, d\mu.$$

Teorema de la convergencia absoluta (TCA). *Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida. Sea $\sum_{n\ge 1} f_n$ una serie de funciones integrables en Ω y supongamos que*

$$\sum_{n=1}^{\infty} \int_\Omega |f_n|\, d\mu < \infty.$$

Entonces la serie $\sum_{n\ge 1} f_n$ converge absolutamente c.p.d.*, existe una función f integrable en Ω tal que $\sum_{n=1}^{\infty} f_n = f$* c.p.d. *y además*

$$\lim \int_\Omega \left| f - \sum_{k=1}^{n} f_k \right| d\mu = 0.$$

En consecuencia

$$\int_\Omega f\, d\mu = \sum_{n=1}^{\infty} \int_\Omega f_n\, d\mu.$$

Teorema de Riesz. *Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida. Sea $\{f_n\}$ una sucesión de funciones integrables en Ω tal que*

$$\lim_{p,q \longrightarrow \infty} \int_\Omega |f_p - f_q|\, d\mu = 0,$$

es decir:

$$\forall \varepsilon > 0,\ \exists m \in \mathbb{N}:\ p,q \ge m \Longrightarrow \int_\Omega |f_p - f_q|\, d\mu < \varepsilon.$$

Existe entonces una función f integrable en Ω tal que

$$\lim \int_\Omega |f - f_n|\, d\mu = 0$$

y existe, además, una sucesión parcial $\{f_{\sigma(n)}\}$ que converge c.p.d. *a f.*

El espacio $\mathbf{L}_1(\mu)$.

$\mathbf{L}_1(\mu) = \mathscr{L}(\mu)/\mathbf{N}(\mu)$, donde $\mathbf{N}(\mu) = \{f : \Omega \longrightarrow \mathbb{R} : \ f \ \text{ es medible y } \ f = 0 \text{ c.p.d.}\}$ Esencialmente $\mathbf{L}_1(\mu)$ no es otra cosa que que el mismo espacio $\mathscr{L}(\mu)$ en el que se considera la igualdad c.p.d. en lugar de la igualdad ordinaria de funciones. De manera natural $\|\cdot\|_1$ se convierte en una norma en $\mathbf{L}_1(\mu)$ sin más que definir $\|f+\mathbf{N}(\mu)\|_1 = \|f\|_1$. La convergencia respecto de la norma $\|\cdot\|_1$ se denomina *convergencia en media*.

Teorema de complitud de $\mathbf{L}_1(\mu)$. *Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida. Entonces $\big(\mathbf{L}_1(\mu), \|\cdot\|_1\big)$ es un espacio de Banach.*

Teorema de densidad.

i) El espacio vectorial $\mathbf{C}_{00}(\mathbb{R}^N)$ de las funciones continuas de soporte compacto es denso en $\big(\mathscr{L}(\lambda), \|\cdot\|_1\big)$, es decir, dada una función integrable f existe una sucesión $\{f_n\}$ en $\mathbf{C}_{00}(\mathbb{R}^N)$ tal que

$$\lim \|f - f_n\|_1 = 0.$$

ii) El espacio vectorial $\mathbf{E}(\mathbb{R}^N)$ de las funciones escalonadas es denso en $\big(\mathscr{L}(\lambda), \|\cdot\|_1\big)$, es decir, dada una función integrable f existe una sucesión $\{f_n\}$ en $\mathbf{E}(\mathbb{R}^N)$ tal que

$$\lim \|f - f_n\|_1 = 0.$$

Además se puede conseguir en ambos casos que la sucesión converja también c.p.d. a f.

12.7. Ejercicios del Tema 12.

12.1 Sean $(\Omega, \mathscr{A}, \mu)$ un espacio de medida, $\{f_n\}$ una sucesión de funciones medibles y f, g dos funciones medibles. Probar las siguientes afirmaciones:

i) Si $\{f_n\} \longrightarrow f$ c.p.d. y $\{f_n\} \longrightarrow g$ c.p.d. Entonces $f = g$ c.p.d.

ii) Si $\{f_n\} \longrightarrow f$ c.p.d. y $f = g$ c.p.d. Entonces $\{f_n\} \longrightarrow g$ c.p.d.

12.2 Considerar las siguientes sucesiones $\{f_n\}$ de funciones reales de variable real

$$i)\ \ f_n = \frac{1}{2n}\chi_{[-n,n]}\ \ ;\ \ ii)\ \ f_n = n^2\chi_{[\frac{1}{n+1},\frac{1}{n}]}.$$

Estudiar en cada caso la convergencia puntual y comparar $\int \lim f_n$ con $\lim \int f_n$.

12.3 Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida con $\mu(\Omega) < \infty$ y sea $\{f_n\}$ una sucesión de funciones medibles en Ω que converge puntualmente a una función f. Supongamos que existe una constante $M \geq 0$ tal que $|f_n| \leq M$, $\forall n \in \mathbb{N}$. Probar que f es integrable y que

$$\lim \int_\Omega |f - f_n|\, d\mu = 0$$

Dar un ejemplo mostrando que la hipótesis $\mu(\Omega) < \infty$ no puede ser suprimida.

12.4 Calcular $\lim \int f_n$ para cada una de las siguientes sucesiones $\{f_n\}$ de funciones de $]0,1[$ en $\mathbb{R}$:

$$i)\ \ \frac{nx}{1+n^2x^2}\ \ ;\ \ ii)\ \ \frac{1}{n}\log(x+n)\mathrm{e}^{-x}\cos(x)\ \ ;\ \ iii)\ \ \frac{1+nx}{(1+x)^n}\ .$$

12.5 Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida con $\mu(\Omega) < \infty$ y sea $\{f_n\}$ una sucesión de funciones integrables en Ω que converge uniformemente a una función f. Probar que f es integrable y que

$$\lim \int_\Omega |f - f_n|\, d\mu = 0$$

Dar un ejemplo mostrando que la hipótesis $\mu(\Omega) < \infty$ no puede ser suprimida.

12.6 Dar un ejemplo de una sucesión de funciones integrables $\{f_n\}$ tal que

$$\|f_n\|_1 \leq 1,\ \forall n \in \mathbb{N}$$

y que no tiene ninguna sucesión parcial convergente en media.

12.7 Considerar la siguiente sucesión de intervalos de $\mathbb{R}$:

$$I_1 = [0,1] \;,$$
$$I_2 = \left[0, \frac{1}{2}\right] \;,\; I_3 = \left[\frac{1}{2}, 1\right] \;,$$
$$I_4 = \left[0, \frac{1}{4}\right] \;,\; I_5 = \left[\frac{1}{4}, \frac{2}{4}\right] \;,\; I_6 = \left[\frac{2}{4}, \frac{3}{4}\right] \;,\; I_7 = \left[\frac{3}{4}, 1\right]$$
$$\ldots\ldots\ldots\ldots\ldots$$

Probar que la sucesión $\{\chi_{I_n}\}$ converge en media a cero y no converge en ningún punto de $[0,1]$. Obtener una sucesión parcial de $\{\chi_{I_n}\}$ que converja a cero c.p.d.

12.8 Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida y $\mathbf{A}$ un subconjunto denso en $\mathscr{L}(\mu)$ (por ejemplo $\mathbf{F}, \mathbf{C}_{00}(\mathbb{R}^N), \mathbf{E}(\mathbb{R}^N)$ cuando se consideran las funciones integrables en $\mathbb{R}^N$). Probar que para $f : \Omega \longrightarrow \mathbb{R}$ medible equivalen las siguientes afirmaciones.

i) f es integrable.

ii) Existe una sucesión $\{f_n\}$ en $\mathbf{A}$ de Cauchy en media que converge c.p.d. a la función f .

En cuyo caso se tiene además

$$\int_\Omega f \, d\mu = \lim \int f_n \, d\mu.$$

Se resalta la analogía con el método de Cantor de construcción de un modelo de números reales a partir de los números racionales.
Indicación: Utilizar el Teorema de Riesz y el Ejercicio 12.1.

12.9 Para cada natural n sea

$$f_n(x) = a_n \; \mathrm{sen}(nx) + b_n \; \cos(nx), \; \forall x \in \mathbb{R},$$

donde $\{a_n\}, \{b_n\}$ son dos sucesiones de números reales. Se supone que $\{f_n\} \longrightarrow 1$ c.p.d. en $[0, 2\pi]$. Deducir que la sucesión $\{|a_n| + |b_n|\}$ no está acotada.
Indicación: Utilizar el Ejercicio 12.3.

12.10 Probar que la sucesión de funciones $\{f_n\}$ definidas por

$$f_n(x) = \mathrm{sen}(nx), \; \forall x \in \mathbb{R}, \; \forall n \in \mathbb{N}$$

no tiene ninguna parcial que que converge c.p.d. en $\mathbb{R}$.
Indicación: Si existiese una tal sucesión parcial $\{f_{\sigma(n)}\}$ considerar la sucesión $\{g_n\}$ definida para cada n natural

$$g_n = \left(f_{\sigma(n+1)} - f_{\sigma(n)}\right)^2 \chi_{[0,2\pi]}.$$

12.11 Sea $(\Omega, \mathscr{A}, \mu)$ un espacio de medida y $f \in \mathscr{L}(\mu)$. Probar que

$$n\,\mu\big(\{\omega \in \Omega : \; |f(\omega)| \geq n\}\big) \longrightarrow 0.$$

Si f no es integrable, ¿es cierto la anterior afirmación para funciones medibles positivas? En caso negativo, dése un contraejemplo.

12.12 Dada $f \in \mathscr{L}(\mathbb{R}^N)$ definamos para cada $h \in \mathbb{R}^N$ la función $f_h : \mathbb{R}^N \longrightarrow \mathbb{R}$ dada por $f_h(x) = f(x+h)$, $\forall x \in \mathbb{R}^N$. Probar que $\lim_{h\to 0} \int |f - f_h| = 0$.
Indicación: Probarlo en primer lugar para funciones continuas de soporte compacto.

12.13 ***Teorema de Lebesgue de caracterización las funciones Riemann integrables.** Sea $f : [a,b] \longrightarrow \mathbb{R}$ acotada. Entonces f es Riemann integrable si, y sólo si, f es continua c.p.d.
Indicación: Para cada n natural considerar la partición de $[a,b]$

$$P_n = \{x_k = a + k\,\frac{b-a}{2^n} : \; k = 0, 1, \ldots, 2^n\}$$

y las funciones simples $e_n, E_n : [a,b] \longrightarrow \mathbb{R}$ definidas por

$$e_n = \sum_{k=1}^{2^n} m_k\, \chi_{[x_{k-1},x_k[} + f(b)\chi_{\{b\}} \;\; , \;\; E_n = \sum_{k=1}^{2^n} M_k\, \chi_{[x_{k-1},x_k[} + f(b)\chi_{\{b\}}$$

donde para $k = 1, \ldots, 2^n$,

$$m_k = \inf(f([x_{k-1},x_k[)) \quad \text{y} \quad M_k = \sup(f([x_{k-1},x_k[)).$$

Probar que

i) $\int_{[a,b]} e_n\, d\lambda = I(f,P_n)$ y $\int_{[a,b]} E_n\, d\lambda = S(f,P_n)$.

ii) $e_n \leq e_{n+1} \leq f \leq E_{n+1} \leq E_n$, $\forall n \in \mathbb{N}$.

iii) Si $e = \lim e_n$ y $E = \lim E_n$, entonces

a) f es continua c.p.d. si, y sólo si, $e = E$ c.p.d.

b) Las funciones e y E son integrables y además

$$\int_{[a,b]} e\, d\lambda = \underline{\int_a^b} f(x)dx \quad \text{y} \quad \int_{[a,b]} E\, d\lambda = \overline{\int_a^b} f(x)dx \; \int_{[a,b]} e_n\, \lambda$$

iv) f es integrable Riemann si, y sólo si, $E = e$ c.p.d.

12.14 * Sea G un abierto de $\mathbb{R}^N$. Dada una función f integrable en G y un número positivo ε existe una función φ de soporte compacto que admite derivadas parciales continuas de todos los órdenes ($\varphi \in \mathbf{C}_{00}^{\infty}(\mathbb{R}^N)$), tal que sop $\varphi \subset G$ y $\int_G |f - \varphi| < \varepsilon$.
Indicación: En virtud del Teorema de densidad (12.12 apartado ii)) basta con aproximar la función característica de un intervalo acotado contenido en G por una función de $\mathbf{C}_{00}^{\infty}(\mathbb{R}^N)$. Estudiar esta aproximación en primer lugar en el caso $N = 1$. En esta situación observar que si $a, b, \delta \in \mathbb{R}$ con $a < b$ y $0 < \delta < \dfrac{b-a}{2}$, entonces la función $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$ definida por

$$\varphi(x) = \frac{\gamma(x-a)\gamma(b-x}{\gamma(x-(b-\delta)) + \gamma((a+\delta)-x) + \gamma(x-a)\gamma(b-x)}, \ \forall x \in \mathbb{R}$$

donde $\gamma(x) = \begin{cases} e^{-\frac{1}{x}} \text{ si } x > 0 \\ 0 \quad \text{si } x \leq 0 \end{cases}$ verifica las siguientes propiedades:

i) $\varphi \in \mathbf{C}_{00}^{\infty}(\mathbb{R})$.

ii) sop $(\varphi) \subset [a, b]$.

iii) $0 \leq \varphi(x) \leq 1, \ \forall x \in \mathbb{R}$.

iv) $\varphi(x) = 1, \ \forall x \in [a+\delta, b-\delta]$.

v) $\varphi(x) = 0, \ \forall x \notin [a, b]$.

12.8. Soluciones a los ejercicios del Tema 12.

12.1 Consideremos los siguientes conjuntos medibles

$$A = \{\omega \in \Omega : \{f_n(\omega)\} \to f(\omega)\}\ ,\ B = \{\omega \in \Omega : \{f_n(\omega)\} \to g(\omega)\}$$

$$\text{y} \quad D = \{\omega \in \Omega : f(\omega) = g(\omega)\}.$$

i) $\mu(A^C) = \mu(B^C) = 0 \Longrightarrow \mu\big((A\cap B)^C\big) = 0$. El resultado se sigue de que $A\cap B \subset D$.

ii) $\mu(A^C) = \mu(D^C) = 0 \Longrightarrow \mu\big((A\cap D)^C\big) = 0$. El resultado se sigue de que $A\cap D \subset B$.

12.2 Todas las funciones que aparecen es este ejercicio son integrables, de hecho son funciones simples integrables.

i) $f_n = \frac{1}{2n}\chi_{[-n,n]}$. $\{f_n\} \longrightarrow 0$ en $\mathbb{R}$ (de hecho converge uniformemente).

$$\int f_n = 1,\ \forall n \in \mathbb{N}\ \ ;\ \ \int 0 = 0 \neq 1 = \lim \int f_n.$$

Es fácil comprobar que no hay monotonía ni acotación. Se puede hacer directamente o bien pensando que en el caso de que existiera monotonía el T.C.M. daría la igualdad y en el caso de que existiese acotación el T.C.D. daría también la igualdad. Este ejercicio muestra asimismo que el Lema de Fatou no se puede aspirar a probar la igualdad.

ii) $f_n = n^2\chi_{[\frac{1}{n+1},\frac{1}{n}]}$. $\{f_n\} \longrightarrow 0$ en $\mathbb{R}$.

$$\int f_n = \frac{n}{n+1},\ \forall n \in \mathbb{N}\ \ ;\ \ \int 0 = 0 \neq 1 = \lim \int f_n.$$

12.3 Las funciones constantes son integrables. En efecto la función constantemente igual M , $M\ \chi_\Omega$, es simple integrable con integral igual a $M\ \mu(\Omega)$. El ejercicio es consecuencia del T.C.D.

El apartado i) del ejercicio anterior muestra que la hipótesis $\mu(\Omega) < \infty$ no puede ser suprimida.

12.4 En todos los ejercicios se verifica que $\{f_n\} \longrightarrow 0$ c.p.d. Como no hay monotonía hay que utilizar el ejercicio anterior para concluir que $\lim \int f_n = 0$. El problema estriba en la mayorar las sucesiones.

i) $\dfrac{nx}{1+n^2x^2}$

$$0 \leq (1-nx)^2 \Longrightarrow 0 \leq nx \leq \frac{1+n^2x^2}{2},\ \forall n \in \mathbb{N} \Longrightarrow M = \frac{1}{2}\ .$$

ii) $\dfrac{1}{n}\log(x+n)\mathrm{e}^{-x}\cos(x)$

$$|f_n(x)| \leq \frac{\log(1+n)}{n} = \frac{\log(1+n)-\log(1)}{n} \leq$$
$$\leq \frac{1}{1+\vartheta} \leq 1,\ (0<\vartheta<n)\ \forall n\in\mathbb{N}$$

donde se ha utilizado el Teorema del valor medio. Así $M=1$.

iii) $\dfrac{1+nx}{(1+x)^n}$
$1+nx \leq (1+x)^n,\ \forall n\in\mathbb{N} \Longrightarrow M=1$.

12.5 Para todo natural n la función $|f-f_n|$ es medible. La monotonía de la integral para funciones medibles positivas nos asegura que

$$\int_\Omega |f-f_n|\,d\mu \leq \mu(\Omega)\,\|f-f_n\|_\infty.$$

La sucesión $\left\{\|f-f_n\|_\infty\right\}$ en $[0,\infty]$ converge a cero por hipótesis. Sea $m\in\mathbb{N}$ tal que $\left\{\|f-f_m\|_\infty\right\}\leq 1$. Como $\int_\Omega |f-f_m|\,d\mu \leq \mu(\Omega)$ concluimos que $f-f_m$ es integrable y en consecuencia $f=f_m+(f-f_m)$ también es integrable. Como $\left\{\|f-f_n\|_\infty\right\}\longrightarrow 0$, concluimos que

$$\lim \int_\Omega |f-f_n|\,d\mu = 0.$$

El apartado i) del Ejercicio 12.2 muestra que la hipótesis $\mu(\Omega)<\infty$ no puede ser suprimida.

12.6 Se trata de dar un ejemplo de una sucesión de funciones integrables que sea acotada en media y no tenga ninguna parcial convergente ($\mathscr{L}(\lambda)$ es de dimensión infinita).

$$f_n = \chi_{[n,n+1[^N},\ \forall n\in\mathbb{N}.$$

Las funciones son simples integrables con $\|f_n\|_1=1,\ \forall n\in\mathbb{N}$ y

$$p\neq q \Longrightarrow |f_p-f_q| = f_p+f_q \quad \text{y en consecuencia} \quad \|f_p-f_q\|_1 = 2\ .$$

Así $\{f_n\}$ no tiene ninguna parcial de Cauchy, luego no tiene ninguna parcial convergente.

12.7 Se trata de dar un ejemplo que ponga de manifiesto que en el Teorema de Riesz (12.9) la sucesión que converge c.p.d. tiene que ser una "sucesión parcial" de la dada. La siguiente sucesión de intervalos de $\mathbb{R}$:

$$I_1=[0,1]\ ,$$
$$I_2=\left[0,\frac{1}{2}\right]\ ,\ I_3=\left[\frac{1}{2},1\right]\ ,$$
$$I_4=\left[0,\frac{1}{4}\right]\ ,\ I_5=\left[\frac{1}{4},\frac{2}{4}\right]\ ,\ I_6=\left[\frac{2}{4},\frac{3}{4}\right]\ ,\ I_7=\left[\frac{3}{4},1\right]$$
$$\cdots\cdots\cdots\cdots\cdots$$

es tal que $\{l(I_n)\} \longrightarrow 0$, y en consecuencia $\{\chi_{I_n}\}$ converge en $\left(\mathscr{L}([0,1]), \|\cdot\|_1\right)$ a cero. Sea $x \in [0,1]$. Es obvio que x pertenece a infinitos I_n y no pertenece a infinitos I_n, así la sucesión $\{\chi_{I_n}\}$ toma infinitas veces el valor cero e infinitas veces el valor 1 y por tanto no es convergente. Por último la sucesión $\{\chi_{I_{2^n}}\}$ converge a cero c.p.d. (de hecho en $\mathbb{R} \setminus \{0\}$).

12.8 $i) \Rightarrow ii)$ Sea $\{g_n\}$ una sucesión en $\mathbf{A}$ tal que $\{g_n\} \longrightarrow f$ en $\left(\mathscr{L}(\mu), \|\cdot\|_1\right)$. El Teorema de Riesz (12.9) nos asegura la existencia de una función integrable g tal que

$$\{g_n\} \longrightarrow g \ \text{ en } \ \left(\mathscr{L}(\mu), \|\cdot\|_1\right) \ \text{ y } \ \exists\{g_{\sigma(n)}\} \longrightarrow g \text{ c.p.d.}$$

Es claro que $f = g$ c.p.d. La sucesión $\{f_n\} = \{g_{\sigma(n)}\}$ cumple las condiciones de la afirmación $ii)$ (véase el Ejercicio 12.1)

$ii) \Rightarrow i)$ Sea $\{f_n\}$ una sucesión verificando $ii)$. El Teorema de Riesz nos asegura que existe una función integrable g tal que

$$\{f_n\} \longrightarrow g \ \text{ en } \ \left(\mathscr{L}(\mu), \|\cdot\|_1\right) \ \text{ y } \ \exists\{f_{\sigma(n)}\} \longrightarrow g \text{ c.p.d.}$$

de donde se deduce del Ejercicio 12.1 que $f = g$ c.p.d. Como f es medible concluimos que también es integrable. Por último de

$$\|f - f_n\|_1 \leq \|g - f_n\|_1 + \|f - g\|_1, \ \forall n \in \mathbb{N}$$

se deduce que $\{f_n\}$ converge en media a f y en particular

$$\int_\Omega f \, d\mu = \lim \int_\Omega f_n \, d\mu.$$

12.9 Un fácil cálculo nos da que $\int_0^{2\pi} f_n(x)\, dx = 0, \ \forall n \in \mathbb{N}$. Como

$$|f_n(x)| \leq |a_n| + |b_n|, \ \forall x \in [0, 2\pi], \ \forall n \in \mathbb{N},$$

si, razonando por reducción al absurdo, la sucesión $\{|a_n| + |b_n|\}$ estuviese acotada se podría aplicar el Ejercicio 12.3 y se llegaría a la siguiente contradicción:

$$2\pi = \int_0^{2\pi} dx = \int_0^{2\pi} \lim f_n(x)\, dx = \lim \int_0^{2\pi} f_n(x)\, dx = 0.$$

12.10 Razonando por reducción al absurdo, si una sucesión parcial $\{f_{\sigma(n)}\}$ convergiese c.p.d., considerando la sucesión de funciones (Riemann) integrables

$$\{g_n\} = \left\{ \left(f_{\sigma(n+1)} - f_{\sigma(n)}\right)^2 \chi_{[0,2\pi]} \right\}$$

se tendría con un sencillo cálculo que

$$\int g_n(x)\,dx = \int_0^{2\pi} \left(f_{\sigma(n+1)}(x) - f_{\sigma(n)(x)}\right)^2\,dx = \pi + \pi - 0 = 2\pi,\ \forall n \in \mathbb{N},$$

$$\left[\operatorname{sen}^2 x = \frac{1-\cos 2x}{2},\ \ \operatorname{sen} px\ \operatorname{sen} qx = \frac{1}{2}(\cos(p-q)x - \cos(p+q)x)\right]$$

$$|g_n| \leq 4\,\chi_{[0,2\pi]},\ \forall n \in \mathbb{N}$$

y el Ejercicio 12.3 nos daría el siguiente absurdo:

$$2\pi = \lim \int g_n(x)\,dx = \int \lim g_n(x)\,dx = \int 0\,dx = 0.$$

12.11 Sea $f \in \mathscr{L}(\mu)$. Para cada natural n sea $A_n := \{\omega \in \Omega : |f(\omega)| \geq n\}$. Se tiene que

$$0 \leq n\,\chi_{A_n} \leq |f|,\ \forall n \in \mathbb{N}.$$

Como $\{n\,\chi_{A_n}\} \longrightarrow 0$, el T.C.D. nos asegura que

$$\{n\,\mu(A_n)\} = \left\{\int_\Omega n\,\chi_{A_n}\,d\mu\right\} \longrightarrow 0.$$

La función $f :]0,1[\to \mathbb{R}$ dada por $f(x) = \frac{1}{x}, \forall x \in]0,1[$ es medible positiva y sirve de contraejemplo.

12.12 Dada $f \in \mathscr{L}(\mu)$ definimos para cada $h \in \mathbb{R}^N$ la función $f_h : \mathbb{R}^N \longrightarrow \mathbb{R}$ dada por

$$f_h(x) = f(x+h),\ \forall x \in \mathbb{R}^N.$$

Probemos en primer lugar que $f_h \in \mathscr{L}(\mu)$. Es inmediato que si $\{s_n\}$ es una sucesión creciente de funciones simples integrables que converge a f^+ y definimos

$$s_n^h(x) = s_n(x+h),\ \forall x \in \mathbb{R}^N,$$

entonces la sucesión $\{s_n^h\}$ es una sucesión creciente de funciones simples integrables que converge a f_h^+ y claramente

$$\int s_n^h(x)\,dx = \int s_n(x)dx,\ \forall n \in \mathbb{N}.$$

En consecuencia del T.C.M. se deduce que f_h es integrable y que

$$\int f_h(x)\,dx = \int f(x)\,dx.$$

Sea $\varepsilon > 0$. El Teorema 12.12 nos asegura que existe $\varphi \in \mathbf{C}_{00}(\mathbb{R}^N)$ tal que $\|f-\varphi\|_1 < \dfrac{\varepsilon}{3}$. Sea $h \in \mathbb{R}^N$ y notemos $\varphi_h(x) = \varphi(x+h),\ \forall x \in \mathbb{R}^N$. Se tiene que

$$\varphi_h \in \mathbf{C}_{00}(\mathbb{R}^N) \ \text{ y que } \ \|f_h - \varphi_h\|_1 < \frac{\varepsilon}{3}\ .$$

Así

$$\|f-f_h\|_1 \leq \|f-\varphi\|_1 + \|f_h-\varphi_h\|_1 + \|\varphi-\varphi_h\|_1 \leq \frac{2\varepsilon}{3} + \|\varphi-\varphi_h\|_1.$$

Veamos finalmente que el último sumando se puede controlar. En efecto, se puede suponer que $\|h\| < 1$ para conveniente norma de $\mathbb{R}^N$. Se tiene que

$$K = \left\{x \in \mathbb{R}^N : \ \operatorname{dist}(x, \operatorname{sop}\ (\varphi)) \leq 1\right\}$$

es un compacto de $\mathbb{R}^N$ verificando sop (φ), sop $(\varphi_h) =$ sop $(\varphi) - h \subset K$. La continuidad uniforme de $\varphi - \varphi_h$ (Teorema de Heine) nos asegura la existencia de $\delta > 0$ tal que

$$\|h\| < \delta \Longrightarrow \|\varphi - \varphi_h\|_\infty \ < \ \frac{\varepsilon}{3\,\lambda(K)+1}\ .$$

El resto es consecuencia de la desigualdad

$$\int |\varphi(x) - \varphi_h(x)|\, dx \leq \|\varphi-\varphi_h\|_\infty\, \lambda(K).$$

12.13 i) y ii) son conocidas (véase el Ejemplo 10.19.3). Probamos iii):

Apartado **a**), esto es

$$f \ \text{ es continua c.p.d.} \Longleftrightarrow e = E \ \text{c.p.d.}$$

Sea $Z_1 = \bigcup_{n=1}^\infty P_n$.

$\boxed{\Rightarrow}$ Sean $Z_2 = \{x \in [a,b] : \ f \text{ no es continua en } x\}$ y $Z = Z_1 \cup Z_2$. Probamos que para cada $x \in [a,b] \setminus Z$ se verifica que $e(x) = E(x)$. En efecto, sea $\varepsilon > 0$, existe $\delta > 0$ tal

$$t \in]x-\delta, x+\delta[\cap [a,b] \Rightarrow |f(x) - f(t)| \leq \varepsilon$$

Es claro que existe un natural n y un intervalo I de P_n tal que $I \subset]x-\delta, x+\delta[$, con lo que

$$\begin{gathered} 0 \leq E_n(x) - e_n(x) = M_k - m_k \leq \\ f(t_1) - f(t_2) + \varepsilon \leq |f(x) - f(t_1)| + |f(x) - f(t_2)| + \varepsilon \leq 3\,\varepsilon, \end{gathered}$$

para convenientes $t_1, t_2 \in I$. Hemos probado que $e(x) = E(x)$.

$\boxed{\Leftarrow}$ Sea $Z_3 = \{x \in [a,b] : \ e(x) < E(x)\}$ y $Z? = Z_1 \cup Z_3$. Sea $x \in [a,b] \setminus Z'$. Veamos que f es continua en x. Dado $\varepsilon > 0$ existe un natural n tal que $E_n(x) - e_n(x) \leq \varepsilon$. Sea I un intervalo de P_n tal que $x \in \overset{\circ}{I}$, y sea $\delta > 0$ tal que $]x-\delta, x+\delta[\subset I$. Se tiene entonces que

$$t \in]x-\delta, x+\delta[\Longrightarrow |f(x) - f(t)| \leq E_n(x) - e_n(x) \leq \varepsilon \ .$$

Apartado **b**). Sea M una cota de f, esto es, $-M \leq f(x) \leq M$, $\forall x \in [a,b]$. Es claro que para cada natural n es $-M \leq e_n \leq E_n \leq M$. Por otra parte $\{e_n\}$ y $\{E_n\}$ son sucesiones monótonas de funciones integrables que convergen puntualmente a e y E respectivamente, y además

$$-M(b-a) \leq \int_{[a,b]} e_n \, d\lambda \leq \int_{[a,b]} e_n \, d\lambda \leq M(b-a).$$

El T.C.M. nos da que

$$\int_{[a,b]} e \, d\lambda = \lim \int_{[a,b]} e_n \, d\lambda$$

o lo que es lo mismo

$$\int_{[a,b]} e \, d\lambda = \lim I(f, P_n)$$

y en consecuencia el Teorema de Darboux nos asegura que

$$\int_{[a,b]} e \, d\lambda = \underline{\int_a^b} f(x) dx \ .$$

De manera análoga

$$\int_{[a,b]} E \, d\lambda = \overline{\int_a^b} f(x) dx \ .$$

Probemos finalmente iv). Acabamos de probar que

$$f \ \text{ es integrable Riemann } \iff \int_{[a,b]} (E - e) \, d\lambda = 0$$

lo que equivale, en virtud del apartado iv) de la Proposición 10.13, a que $E = e$ c.p.d. que a su vez es equivalente, según hemos probado en el apartado **a**), a que f sea continua c.p.d.

12.14 <u>Caso $N = 1$</u>:

$$\int_G \left| \chi_{[a,b]} - \varphi \right| = \int_a^b (\mathbf{1} - \varphi) = \int_a^{a+\delta} (\mathbf{1} - \varphi) + \int_{b-\delta}^b (\mathbf{1} - \varphi) \leq \delta + \delta = 2 \ \delta \ .$$

<u>Caso general</u>: Sea $I = I_1 \times \cdots \times I_N$. Se toma $\varphi : \mathbb{R}^N \longrightarrow \mathbb{R}$ la función definida por

$$\varphi(x_1, \ldots, x_N) = \varphi_1(x_1) \cdots \varphi_N(x_N) \ ,$$

donde $\varphi_1, \ldots, \varphi_N$ son las funciones asociadas a $I_1, \ldots, I_N$ por el procedimiento dado en la indicación. Se tiene que

$$\int_G |\chi_I - \varphi| = \int_{I_1 \times \cdots \times I_N} (\mathbf{1} - \varphi) \leq \lambda(I_1 \times \cdots \times I_N) - \lambda(I'_1 \times \cdots \times I'_N) \ ,$$

donde

$$I'_k \subset I_K \ \text{ y } \ l(I_k) = l(I'_k) + 2\ \delta_k \quad (k = 1, \ldots, N).$$

Así

$$\lambda(I_1 \times \cdots \times I_N) - \lambda(I'_1 \times \cdots \times I'_N) \longrightarrow 0 \ \text{ cuando } \ \delta = \max\{\delta_1, \ldots, \delta_N\} \longrightarrow 0.$$

Tema 13

Técnicas de integración en una variable.

El objetivo de este tema es el estudio de la integral de Lebesgue en $\mathbb{R}$, esto es, la integral asociada al espacio de medida $(R, \mathscr{M}, \lambda)$. Tras el asentamiento en las lecciones precedentes de los pilares básicos que sostienen la teoría de la integración, podemos ahora abordar la cuestión fundamental para nosotros: proporcionar métodos que permitan reconocer cuando una función $f : I \longrightarrow \mathbb{R}$, donde I es un intervalo de $\mathbb{R}$, es integrable y, en caso afirmativo, calcular su integral. Es usual notar $\mathscr{L}(I)$ al conjunto de las funciones integrables en I.

Es importante tener presente que el problema de integración planteado contiene dos cuestiones distintas: una es la integrabilidad de la función (hasta ahora sabemos que son integrables las funciones nulas c.p.d., las combinaciones lineales de funciones características de conjuntos de medida finita y las funciones continuas de soporte compacto) y otra es calcular, supuesto que la función sea integrable, el valor de su integral. En este sentido debemos advertir que presentamos tres tipos de técnicas de integración:

- Métodos que proporcionan la integrabilidad de f aunque no el valor de la integral.

- Métodos que proporcionan la integrabilidad de f y el valor de su integral.

- Métodos que proporcionan el valor de la integral, supuesto que la función sea integrable.

El lector tendrá cuidado en lo sucesivo de advertir con precisión la información que cada resultado proporciona al respecto.

13.1. Notación y aditividad de la integral.

Notación 13.1. Sean I un intervalo de $\mathbb{R}$ y $f : I \to \mathbb{R}$ una función real. Al ser

$$\chi_I = \chi_{\overset{\circ}{I}} \text{ c.p.d.}$$

la función f es integrable en I si, y sólo si, lo es en $\overset{\circ}{I}$; en cuyo caso

$$\int_I f\, d\lambda = \int_{\overset{\circ}{I}} f\, d\lambda.$$

Este hecho nos permite considerar, a efectos de integración, sólo intervalos de $\mathbb{R}$ de la forma $]\alpha,\beta[$ con

$$-\infty \leq \alpha < \beta \leq +\infty.$$

Si $f \in \mathscr{L}(]\alpha,\beta[)$ notaremos *su integral* por

$$\int_\alpha^\beta f(x)\, dx.$$

Es usual definir

$$\int_\beta^\alpha f(x)\, dx := -\int_\alpha^\beta f(x)\, dx.$$

Conviene resaltar que la notación es coherente con la ya utilizada para la integral de Riemann

El siguiente resultado es una herramienta indispensable en muchas demostraciones.

Proposición 13.2 (Aditividad de la integral respecto del intervalo de integración)**.**

i) Sean $f :]\alpha,\beta[\to \mathbb{R}$ *una función y* $\gamma \in]\alpha,\beta[$*. Entonces*

$$f \in \mathscr{L}^1(]\alpha,\beta[) \;\Leftrightarrow f \in \mathscr{L}^1(]\alpha,\gamma[) \cap \mathscr{L}^1(]\gamma,\beta[),$$

en cuyo caso

$$\int_\alpha^\beta f(x)\, dx = \int_\alpha^\gamma f(x) dx + \int_\gamma^\beta f(x)\, dx.$$

ii) Sean $\alpha,\beta,\gamma \in [-\infty,+\infty]$ *y* $f :]a,b[\longrightarrow \mathbb{R}$ *una función integrable, donde* $a = \min\{\alpha,\beta,\gamma\}$ *y* $b = \max\{\alpha,\beta,\gamma\}$*. Entonces*

$$\int_\alpha^\beta f(x)\, dx = \int_\alpha^\gamma f(x)\, dx + \int_\gamma^\beta f(x)\, dx.$$

Demostración:

i) Es un caso particular de la Proposición 11.18.

ii) La fórmula que se pretende demostrar es equivalente a

$$\int_\alpha^\beta f(x)\, dx + \int_\beta^\gamma f(x)\, dx + \int_\gamma^\alpha f(x)\, dx = 0, \qquad (*)$$

fórmula que es trivial si dos de los tres puntos que aparecen en ella coinciden, pues en tal caso se reduce a

$$\int_a^b f(x)\, dx + \int_b^a f(x)\, dx = 0.$$

Además, si se permutan de cualquier forma los puntos α,β y γ, el primer miembro de la fórmula $(*)$ permanece ó queda multiplicado por -1. Ello permite suponer que $\alpha < \gamma < \beta$, en cuyo caso la fórmula es cierta por el apartado i). ∎

13.2. Teorema fundamental del cálculo.

El teorema fundamental del cálculo pone de manifiesto que la derivación y la integración son operaciones inversas, esto es

$$\left(\int_a^x f(t)\,dt\right)' = f(x) \text{ para casi todo punto } x \in]\alpha,\beta[,$$

$$f(x) = f(a) + \int_a^x f'(t)\,dt \text{ para todo punto } x \in]\alpha,\beta[.$$

Antes de concretar estas afirmaciones conviene introducir un concepto más débil que el de integrabilidad.

Definición 13.3 (integrabilidad local e integral indefenida)**.** Una función $f :]\alpha,\beta[\to \mathbb{R}$ es *localmente integrable* si es integrable en cada intervalo compacto contenido en $]\alpha,\beta[$. Fijado un punto $a \in]\alpha,\beta[$, se define la *integral indefenida de* f *de origen* a por

$$F(x) := \int_a^x f(t)\,dt,\ \forall x \in]\alpha,\beta[.$$

Es inmediato de la proposición anterior que toda función integrable es localmente integrable. Probaremos enseguida que la función

$$f(x) = \frac{1}{x},\ \forall x \in]0,1[$$

es localmente integrable pero no integrable. También es claro que dos integrales indefinidas se diferencian en una costante.

Es fácil comprobar que toda función localmente integrable es medible. En efecto, sean $\{a_n\}$ y $\{b_n\}$ sucesiones de números reales del intervalo $]\alpha,\beta[$ con $a_n \le b_n,\ \forall n \in \mathbb{N}$ tales que $\{a_n\} \searrow \alpha$ y $\{b_n\} \nearrow \beta$, entonces $f\chi_{[a_n,b_n]} \to f$ c.p.d. y es claro que f es medible, ya que cada función $f\chi_{[a_n,b_n]}$ es medible por ser integrable.

Resaltamos también que en virtud de la propiedad de compacidad las funciones continuas son localmente integrables.

Teorema 13.4 (fundamental del cálculo)**.** *Sea $f :]\alpha,\beta[\to \mathbb{R}$ una función localmente integrable y sea $a \in]\alpha,\beta[$ un punto fijo. La integral indefinida de f de origen a, esto es*

$$F(x) = \int_a^x f(t)\,dt,\ \forall x \in]\alpha,\beta[$$

verifica las siguientes afirmaciones:

i) Es continua.

ii) Es derivable c.p.d. con $F' = f$ c.p.d. Además F es derivable en cada punto x del intervalo $]\alpha,\beta[$ en el que f es continua con $F'(x) = f(x)$.

Demostración. **i)** Sean $x, x_n \in]\alpha, \beta[$ con $\{x_n\} \to x$. Vamos a probar que $\{F(x_n)\} \to F(x)$. Fijamos $[c,d] \subset]\alpha, \beta[$ tal que

$$x, x_n \in [c,d], \ \forall n \in \mathbb{N}.$$

Notamos por último I_n el intervalo de extremos x, x_n. Se tiene que

$$f\,\chi_{I_n} \longrightarrow 0 \quad , \quad |f\,\chi_{I_n}| \leq |f| \in \mathscr{L}([c,d]).$$

Como

$$|F(x_n) - F(x)| = \left| \int_x^{x_n} f(t)\ dt \right| \leq \int_c^d f(t)\ |\chi_{I_n}(t)|\ dt,$$

concluimos, en virtud del teorema de la convergencia dominada (Teorema 12.5), que

$$\{F(x_n)\} \longrightarrow F(x).$$

La demostración de este apartado para funciones integrables Riemann es una consecuencia inmediata de la acotación de la función f y de las propiedades de la integral.

ii) Supongamos que f es continua en un punto $x \in]\alpha, \beta[$. Dado $\varepsilon > 0$, la continuidad de f en x nos asegura que existe $\delta > 0$ tal que

$$(y \in [a,b],\ \ 0 < |y - x| < \delta) \Rightarrow \ \ |f(y) - f(x)| < \varepsilon.$$

Para un tal y se verifica

$$\left| \frac{F(y) - F(x)}{y - x} - f(x) \right| = \frac{1}{|y-x|} \left| \int_a^y f(t)\ dt - \int_a^x f(t)\ dt - (y-x)f(x) \right| =$$

$$\frac{1}{|y-x|} \left| \int_x^y f(t)dt - (y-x)f(x) \right| = \frac{1}{|y-x|} \left| \int_x^y f(t)\ dt - \int_x^y f(x)\ dt \right| =$$

$$\frac{1}{|y-x|} \left| \int_x^y \Big(f(t) - f(x)\Big)\ dt \right| \leq \frac{1}{|y-x|} \left| \int_x^y |f(t) - f(x)|\ dt \right| \leq \frac{1}{|y-x|} \left| \int_x^y \varepsilon\ dt \right| = \varepsilon,$$

donde se ha utilizado la aditividad respecto del intervalo de integración (Proposición 13.2). Hemos probado que la función F es derivable en x y que $F'(x) = f(x)$.

Hemos demostrado este apartado para funciones continuas.

También es cierto para funciones Riemann integrables, ya que la función f es continua c.p.d. en virtud del teorema de Lebesgue de caracterización de las funciones Riemann integrables (Ejercicio 12.13).

Para funciones integrables (Lebesgue) la demostración de este apartado es muy laboriosa y requiere la presentación de técnicas complicadas impropias de un primer curso de Análisis Matemático.

∎

Teorema 13.5 (regla de Barrow)**.** *Sea $f :]\alpha,\beta[\to \mathbb{R}$ una función que admite primitiva, esto es existe $G :]\alpha,\beta[\to \mathbb{R}$ derivable tal que*

$$G'(x) = f(x),\ \forall x \in]\alpha,\beta[.$$

Se verifican las siguientes afirmaciones

i) f *es medible.*

ii) Si $f :]\alpha,\beta[\to [0,+\infty[$ *, entonces*

$$\int_\alpha^\beta f(x)\,dx = \lim_{x\to\beta} G(x) - \lim_{x\to\alpha} G(x)\,.$$

iii) Si f *es integrable, entonces la primitiva* G *tiene límites (reales) en* α,β *, y*

$$\int_\alpha^\beta f(x)\,dx = \lim_{x\to\beta} G(x) - \lim_{x\to\alpha} G(x).$$

Demostración. Sea $[a,b] \subset]\alpha,\beta[$ fijo. Consideramos la sucesión $\{G_n\}$ de funciones reales definidas en $[a,b]$ por:

(13.2.1) $$G_n(x) := \frac{G\left(x+\frac{1}{n}\right) - G(x)}{\frac{1}{n}},\ \forall x \in [a,b]$$

(si m es un natural tal que $b+\frac{1}{m} < \beta$ y $n \geq m$, entonces $x+\frac{1}{n} \in]\alpha,\beta[,\ \forall x \in [a,b]$).

Al ser G una primitiva de f, es claro que

(13.2.2) $$\{G_n(x)\} \longrightarrow f(x),\ \forall x \in [a,b]\ .$$

Consideramos ahora dos sucesiones $\{a_n\}$ y $\{b_n\}$ en $]\alpha,\beta[$ con

$$a_n < b_n,\ \forall n \in \mathbb{N},\ \{a_n\} \searrow \alpha\ ,\ \{b_n\} \nearrow \beta\ .$$

Se verifica que

(13.2.3) $$f\chi_{[a_n,b_n]} \to f\ .$$

i) En virtud de 13.2.2 la función $f_{[a,b]}$ es medible como limite de una sucesión de funciones continuas, y en virtud de 13.2.3 la función f también es medible por ser límite de una sucesión de funciones medibles.

Lo parte complicada de la demostración de los apartados **ii)** y **iii)** es la prueba de la que nos atrevemos a bautizar como *regla de Barrow para intervalos compactos*, esto es

(13.2.4) $$\int_a^b f(x)\,dx = G(b) - G(a),\ \forall\ [a,b] \subset]\ \alpha,\beta[.$$

ii) Supuesto que 13.2.4 es válida para funciones medibles positivas, de 13.2.3 y del teorema de la convergencia creciente para funciones medibles positivas (Corolario 12.3), se sigue que

$$G(b_n) - G(a_n) = \int_{a_n}^{b_n} f(x)\, dx \nearrow \int_{\alpha}^{\beta} f(x)\, dx,$$

esto es

$$\int_{\alpha}^{\beta} f(x)\, dx = \lim_{x\to\beta} G(x) - \lim_{x\to\alpha} G(x).$$

Nótese que al ser G creciente existen los límites

$$\lim_{x\to\alpha} G(x) \in [-\infty, +\infty[\ \ , \ \ \lim_{x\to\alpha} G(x) \in]-\infty, +\infty]$$

y por consiguiente f es integrable si, y sólo si, ambos límites (que existen) son reales.

Finalmente la prueba de la igualdad 13.2.4 es una fácil consecuencia de la parte sencilla del teorema fundamental del cálculo (Teorema 13.4) si la función f es continua, en efecto

$$(F(x) - G(x))' = 0, \forall x \in [a, b],$$

y por consiguiente ambas funciones se diferencian en una constante, que se calcula y es igual a $G(b)$).

También es una fácil consecuencia del teorema del valor medio si f es integrable Riemann.

En el caso de considerar funciones positivas generales empezamos probando que la función f es localmente integrable. Para ello consideramos las funciones definidas en 13.2.1. Se tiene que

$$\begin{aligned}
\int_a^b G_n(x)\, dx =& n\left(\int_a^b G\left(x+\frac{1}{n}\right)\, dx - \int_a^b G(x)\, dx\right)\\
=& n\left(\int_{a+\frac{1}{n}}^{b+\frac{1}{n}} G(x)\, dx - \int_a^b G(x)\, dx\right)\\
=& n\left(\int_b^{b+\frac{1}{n}} G(x)\, dx - \int_a^{a+\frac{1}{n}} G(x)\, dx\right)\\
\leq& n\left(\int_b^{b+\frac{1}{n}} G(b)\, dx - \int_a^{a+\frac{1}{n}} G(a)\, dx\right) = G(b) - G(a),
\end{aligned}$$

donde se han utilizado la invarianza por traslaciones de la medida de Lebesgue y el teorema fundamental del cálculo (Teorema 13.4).

Ahora el lema de Fatou (Teorema 12.7) nos asegura que

$$\int_a^b f(x) dx \leq \liminf \int_a^b G_n(x)\, dx \leq G(b) - G(a).$$

Finalmente vemos que la igualdad 13.2.4 es ya una consecuencia del teorema de la convergencia dominada (Teorema 12.5). La función *positiva* $f_{[a,b+\frac{1}{m}]}$ es integrable y en virtud de 13.2.2

$$\int_a^b n\left(G(1+\frac{1}{n})-G(x)\right)\,dx \longrightarrow \int_a^b f(x)\,dx,$$

y el resultado es ahora consecuencia de la siguiente cuenta ya conocida

$$n\left(\int_a^b G\left(x+\frac{1}{n}\right)\,dx-\int_a^b G(x)\,dx\right)=n\left(\int_{a+\frac{1}{n}}^{b+\frac{1}{n}} G(x)\,dx-\int_a^b G(x)\,dx\right)=$$

$$n\left(\int_b^{b+\frac{1}{n}} G(x)\,dx-\int_a^{a+\frac{1}{n}} G(x)\,dx\right)\longrightarrow G(b)-G(a),$$

donde se han vuelto a utilizar la invarianza por traslaciones de la medida de Lebesgue y el teorema fundamental del cálculo (Teorema 13.4). Queda así probada lña validez de 13.2.4 para funciones positivas.

iii) Supuesto probada 13.2.4 para funciones reales, sea $\gamma\in]\alpha,\beta[$. Existe $m\in\mathbb{N}$ tal que

$$n\geq m \;\Rightarrow\; a_n<\gamma.$$

Para $n\geq m$ se tiene que

$$\int_{a_n}^{\gamma} f(x)\,dx=G(\gamma)-G(a_n).$$

Habida cuenta que

$$\{f\chi_{[a_n,\gamma]}\}\to f\chi_{]\alpha,\gamma[}\quad \text{y}\quad |f\chi_{[a_n,\gamma]}|\leq|f|,\ \forall n\in\mathbb{N}\,,$$

la integrabilidad de $|f|$ nos permite utilizar el teorema de la convergencia dominada (Teorema 12.5) para concluir que

$$\left\{\int_{a_n}^{\gamma} f(x)\,dx\right\}\to\int_{\alpha}^{\gamma} f(x)\,dx$$

En consecuencia la sucesión $\{G(a_n)\}$ es convergente y

$$\int_{\alpha}^{\gamma} f(x)\,dx=G(\gamma)-\lim_{x\to\alpha}G(x).$$

Análogamente existe el límite de G en β y

$$\int_{\gamma}^{\beta} f(x)\,dx=\lim_{x\to\beta}G(x)-G(\gamma)\,.$$

Finalmente, la aditividad de la integral respecto del intervalo de integración nos permite concluir

$$\int_{\alpha}^{\beta} f(x)\,dx=\lim_{x\to\beta}G(x)-\lim_{x\to\alpha}G(x).$$

Como en el apartado manterior la prueba de la igualdad 13.2.4 es también una fácil si la función f es integrable Riemann. Para funciones integrables (Lebesgue) es muy laboriosa y requiere la presentación de técnicas complicadas impropias de un primer curso de Análisis Matemático. Por ello nos limitamos a considerar la siguiente clase de funciones

Definición 13.6 (Acotación local)**.** $f:]\alpha,\beta[\to\mathbb{R}$ se dice localmente acotada si es acotada en cada intervalo compacto contenido en $]\alpha,\beta[$.

Es claro que toda función medible localmente acotada es localmente integrable. Obsérvese que toda función continua es localmente acotada.

Probamos finalmente, para estan funciones (no tan particulares), la regla de Barrow para intervalos compactos . El teorema del valor medio nos asegura la existencia de $0<\theta_n<1$ tal que $G_n(x)=f(x+\frac{\theta_n}{n})$, con lo que para cada $x\in[a,b]$ se verifica que

$$|G_n(x)|\leq\|f_{[a,b+\frac{1}{m}]}\|_\infty\in\mathscr{L}([a,b]),$$

y se puede utilizar nuevamente el teorema de la convergencia dominada para concluir que

$$\int_a^b f(x)dx=G(b)-Ga).$$

■

Es bueno resaltar que la última afirmación es inmejorable en el sentido de que es cierta siempre que tenga sentido formularse.

Por otra parte es claro que cualquier función que se obtenga sumando a G una función constante es también primitiva de f. Probamos a continuación que de esta forma se obtienen todas las primitivas de f. En efecto, si H es otra primitiva de f, entonces la función $G-H$ es derivable con derivada cero luego constante (teorema del valor medio).

Nota 13.7 (criterio de integrabilidad)**.** Una función

$$f:]\alpha,\beta[\to[0,+\infty[$$

que admite primitiva es integrable, si, y sólo si, su primitiva G tiene límites (reales) en α,β.

En el caso de funciones integrables $f:]\alpha,\beta[\to\mathbb{R}$, la condición anterior es necesaria. Se podría pensar aventuradamente que toda función

$$f:]\alpha,\beta[\to\mathbb{R}$$

con primitiva G que tenga tiene límites (reales) en α,β es integrable. El siguiente ejemplo pone de manifiesto que esto no es así.

Ejemplo 13.8. Consideremos la función $f:\mathbb{R}^+\to\mathbb{R}$ definida por

$$f(t)=\frac{\operatorname{sen} t}{t}.$$

El teorema fundamental del cálculo (Teorema13.4) nos asegura que la función $G:\mathbb{R}^+\to\mathbb{R}$ dada por

$$G(x)=\int_1^x f(t)\,dt$$

es una primitiva de la función f en $\mathbb{R}^+$.

Como $f(x) \to 1$ cuando $x \to 0$ y el intervalo $]0,1[$ es acotado, podemos deducir que la función f es integrable en $]0,1[$, esto es, existe $\lim_{x\to 0} G(x)$. Para probar que también existe $\lim_{x\to +\infty} G(x) \in \mathbb{R}$, basta con comprobar, en virtud del teorema de complitud de $\mathbb{R}$, que si $\{a_n\}$ es una sucesión de $]1,+\infty[$, con $\{a_n\} \nearrow +\infty$, entonces $\big\{G(a_n)\big\}$ es una sucesión de Cauchy. Este hecho se deduce inmediatamente de la siguiente desigualdad: Para cualesquiera $a, b \in \mathbb{R}$ con $1 < a < b$ se verifica que

$$\left|\int_a^b \frac{\operatorname{sen} t}{t}\, dt\right| = \left|\Big[\frac{-\cos t}{t}\Big]_a^b - \int_a^b \frac{\cos t}{t^2}\, dt\right| \le \frac{1}{a} + \frac{1}{b} + \int_a^b \frac{dt}{t^2} = \frac{2}{a}.$$

Si f fuese integrable, también lo sería $|f|$, lo que es imposible ya que

$$\int_0^{n\pi} \left|\frac{\operatorname{sen} t}{t}\right| dt = \sum_{k=1}^{n} \int_{(k-1)\pi}^{k\pi} \frac{|\operatorname{sen} t|}{t}\, dt \ge \sum_{k=1}^{n} \int_{(k-1)\pi}^{k\pi} \frac{|\operatorname{sen} t|}{k\pi}\, dt = \frac{2}{\pi}\big(1 + \frac{1}{2} + \cdots + \frac{1}{n}\big).$$

Es digno de resaltar que, utilizando técnicas de la integral de Lebesgue (teoremas de Fubini, Tonelli, de la convergencia dominada y la integración por partes) se prueba que

$$\lim_{x\to+\infty} \int_0^x \frac{\operatorname{sen} t}{t}\, dt = \int_0^{+\infty} \frac{\operatorname{sen} t}{t}\, dt = \frac{\pi}{2}.$$

(ver Ejercicio 14.10).

□

Ahora la regla de Barrow (Teorema 13.5) nos permite discutir la integrabilidad y calcular la integral en su caso de las funciones potenciales y de la exponencial.

Ejemplo 13.9 (Función potencial). Para cada ρ real sea $f_\rho : \mathbb{R}^+ \longrightarrow \mathbb{R}$ la función definida por

$$f_\rho(x) = x^\rho, \quad \forall x \in \mathbb{R}^+.$$

Entonces se verifica:

i) $f_\rho \notin \mathscr{L}(\mathbb{R}^+)$, $\forall \rho \in \mathbb{R}$.

ii) Si $\beta \in \mathbb{R}^+$, entonces $f_\rho \in \mathscr{L}(]0,\beta[)$ si y sólo si, $\rho > -1$, siendo

$$\int_0^\beta x^\rho\, dx = \frac{\beta^{\rho+1}}{\rho+1}$$

y, en particular,

$$\int_0^1 \frac{dx}{\sqrt{x}} = 2.$$

iii) Si $\alpha \in \mathbb{R}^+$, entonces $f_\rho \in \mathscr{L}(]\alpha,+\infty[)$ si, y sólo si, $\rho < -1$, y en tal caso

$$\int_\alpha^{+\infty} x^\rho\, dx = -\frac{\alpha^{\rho+1}}{\rho+1}$$

y en particular

$$\int_1^{+\infty} \frac{dx}{x^2} = 1.$$

iv) Si $0 < \alpha < \beta < +\infty$, entonces $f_\rho \in \mathscr{L}(]\alpha,\beta[)$, $\forall \rho \in \mathbb{R}$, siendo

$$\int_\alpha^\beta x^\rho \, dx = \begin{cases} \frac{\beta^{\rho+1}-\alpha^{\rho+1}}{\rho+1} & \text{si} \quad \rho \neq -1 \\ \\ \log\left(\frac{\beta}{\alpha}\right) & \text{si} \quad \rho = -1 \end{cases} .$$

f_ρ es una función positiva y medible que, para cualquier real ρ, admite primitiva en $\mathbb{R}^+$ dada por

$$G(x) = \begin{cases} \frac{x^{\rho+1}}{\rho+1} & \text{si} \quad \rho \neq -1 \\ \\ \log(x) & \text{si} \quad \rho = -1 \end{cases} .$$

Puesto que

$$\lim_{x\to 0} G(x) = \begin{cases} -\infty & \text{si} \quad \rho+1 \leq 0 \\ 0 & \text{si} \quad \rho+1 > 0 \end{cases}$$

y

$$\lim_{x\to +\infty} G(x) = \begin{cases} +\infty & \text{si} \quad \rho+1 \geq 0 \\ 0 & \text{si} \quad \rho+1 < 0 \end{cases}$$

el enunciado se sigue de la regla de Barrow. □

Ejemplo 13.10 (Función exponencial). La función $f : \mathbb{R}^+ \to \mathbb{R}$ definida por

$$f(x) = e^{-x}, \ \forall x \in \mathbb{R}^+$$

es integrable y

$$\int_0^{+\infty} e^{-x} dx = 1.$$

Es medible positiva y admite primitiva en $\mathbb{R}^+$ dada por

$$G(x) = -e^{-x}.$$

13.3. Integración por partes.

Otro método de cálculo de integrales es la fórmula de integración por partes.

Proposición 13.11 (Fórmula de integración por partes)**.** *Sean $u, v :]\alpha,\beta[\to \mathbb{R}$ funciones derivables tales que $u'v$ y uv' son integrables. Entonces existen*

$$\lim_{x\to\alpha} u(x)v(x) \ \text{ y } \ \lim_{x\to\beta} u(x)v(x)$$

y se verifica

$$\int_\alpha^\beta u(t)v'(t)\, dt = \lim_{x\to\beta} u(x)v(x) - \lim_{x\to\alpha} u(x)v(x) - \int_\alpha^\beta u'(t)v(t)\, dt.$$

Demostración:

uv es una primitiva de la función $u'v + uv'$ que es integrable y en consecuencia el resultado se deduce directamente de la regla de Barrow (Teorema (13.5). ■

Hay que observar que la utilidad de la anterior fórmula depende de la habilidad que se tenga para expresar el integrando en la forma adecuada.

13.4. Cambio de variable.

Los dos siguientes resultados teóricos de este apartado son particularizaciones para integrales simples de los correspondientes para integrales múltiples (Teoremas 14.9 y 14.10).

Teorema 13.12 (**Cambio de variable para funciones medibles positivas).** *Sean I y J intervalos abiertos de $\mathbb{R}$, ϕ un difeomorfismo de clase $\mathscr{C}^1$ de I sobre J y $f : J \longrightarrow [0, +\infty[$ una función medible. Entonces*

$$\int_J f(x)\, dx = \int_I f(\phi(t))|\phi'(t)|\, dt.$$

Teorema 13.13 (**Cambio de variable).** *Sean I y J intervalos abiertos de $\mathbb{R}$, ϕ un difeomorfismo de clase $\mathscr{C}^1$ de I sobre J y $f : J \to \mathbb{R}$ una función medible. Entonces*

$$f \in \mathscr{L}(J) \;\Leftrightarrow (f \circ \phi)|\phi'| \in \mathscr{L}(I),$$

en cuyo caso

$$\int_J f(x)\, dx = \int_I f(\phi(t))|\phi'(t)|\, dt.$$

La práctica del cambio de variable para funciones reales de variable real se resume en el siguiente:

Corolario 13.14. (Fórmula del cambio de variable). *Sean $-\infty \leq \alpha < \beta \leq +\infty$, f una función real medible definida en $]\alpha, \beta[$ y $\phi :]a, b[\to]\alpha, \beta[$ un difeomorfismo de clase $\mathscr{C}^1$. Entonces*

$$f \in \mathscr{L}(]\alpha, \beta]) \Longleftrightarrow (f \circ \phi)\phi' \in \mathscr{L}(]a, b[)$$

y en caso afirmativo vale la siguiente fórmula del cambio de variable:

$$\int_\alpha^\beta f(x)\, dx = \int_{\phi^{-1}(\alpha^+)}^{\phi^{-1}(\beta^-)} f(\phi(t))\phi'(t)\, dt.$$

Demostración. Es sabido que una función derivable $\phi :]a,b[\to]\alpha,\beta]$ es un difeomorfismo si, y sólo si,

$$\phi'(t) \neq 0, \quad \forall t \in]a,b[\;,$$

en cuyo caso es bien conocido (teorema del valor medio) que ϕ es estrictamente monótona y lo mismo le ocurre a ϕ^{-1}. Sólo se puede dar una de las siguiente posibilidades excluyentes entre sí:

1. $\phi'(t) > 0$, $\forall t \in]a,b[$.
2. $\phi'(t) < 0$, $\forall t \in]a,b[$.

Que se traducen en

1. $|\phi'| = \phi'$, $\phi^{-1}(\alpha^+) = a$ y $\phi^{-1}(\beta^-) = b$.
2. $|\phi'| = -\phi'$, $\phi^{-1}(\alpha^+) = b$ y $\phi^{-1}(\beta^-) = a$.

La fórmula del cambio de variable es ahora consecuencia del teorema de cambio de variable (Teorema 13.13), sin más que recordar que

$$(f \circ \phi)|\phi'| \in \mathscr{L}(I) \;\Leftrightarrow\; (f \circ \phi)\phi' \in \mathscr{L}(I).$$

■

El Ejercicio 13.0 contiene abundantes ejemplos en los que se puede aplicar esta fórmula.

Nota 13.15. Utilizando el teorema fundamental del cálculo (Teorema 13.4) se prueba una versión más general del teorema del cambio de variable donde sólo se le exige a la función ϕ que sea un difeomorfismo, es decir $\phi'(t) \neq 0$, $\forall t \in I$.

13.5. Criterio de comparación.

En el estudio de la integrabilidad de una función localmente integrable $f : I \to \mathbb{R}$ definida en un intervalo I de $\mathbb{R}$ es decisivo el comportamiento de la función en los extremos del intervalo. Puesto que sabemos que para cada $c \in I$:

$$f \text{ es integrable en } I \text{ si, y sólo si, } f \text{ es integrable en } I \cap]-\infty, c] \text{ y en } I \cap [c, +\infty[,$$

se sigue que podemos centrar nuestro estudio en intervalos semiabiertos y, por una simple cuestión de simetría, nos bastará con enunciar los resultados para intervalos de la forma:

$$[\alpha, \beta[\quad \text{con } -\infty < \alpha < \beta \leq +\infty.$$

Es importante destacar que la idea que subyace en lo que sigue es la siguiente propiedad elemental:

Si $f : I \longrightarrow \mathbb{R}$ es una función medible para la que existe una función $g : I \to \mathbb{R}$ integrable tal que $|f| \leq g$, entonces f es integrable.

El primer resultado muestra que si $\beta < +\infty$ y la función es acotada en las cercanías de β, entonces f es integrable.

Proposición 13.16. *Sean $-\infty < \alpha < \beta < +\infty$ y $f : [\alpha, \beta[\longrightarrow \mathbb{R}$ una función localmente integrable. Supongamos que existen $c \in [\alpha, \beta[$ y $M > 0$ tales que*

$$|f| \leq M, \qquad \forall x \in [c, \beta[$$

(lo cual ocurre en particular si f tiene límite finito en β). Entonces f es integrable.

Demostración:
Como f es integrable en $[\alpha, c]$ por hipótesis y es también integrable en $[c, \beta[$ por ser acotada, concluimos que f es integrable. ■

Conviene resaltar que en esta proposición es esencial $\beta < +\infty$, recuérdese que

$$\int_1^{+\infty} \frac{dx}{x} = +\infty \,.$$

Ejemplos 13.17.

a) La función de $]0,1[$ en $\mathbb{R}$ definida por $x \longmapsto \operatorname{sen} \dfrac{1}{x}$ es medible (por ser continua) y al ser acotada (pese a que no tiene límite en 0), es integrable.

b) La condición no es necesaria. La función de $]0,1[$ en $\mathbb{R}$ definida por $x \longmapsto \dfrac{1}{\sqrt{x}}$ no está acotada y, sin embargo, es integrable de hecho

$$\int_0^1 \frac{dx}{\sqrt{x}} = 2.$$

□

Proposición 13.18 (Criterio de comparación). *Sean $-\infty < \alpha < \beta \leq +\infty$ y $f, g : [\alpha, \beta[\longrightarrow \mathbb{R}$ funciones localmente integrables con $g(x) \neq 0,\ \forall x \in [\alpha, \beta[$. Supongamos que*

$$\lim_{x \to \beta} \frac{f(x)}{g(x)} = L.$$

Entonces

i) Si $L \in \mathbb{R}^*$, f es integrable si, y sólo si, g es integrable.

ii) Si $L = 0$ y g es integrable, entonces f es integrable.

iii) Si $L = \pm\infty$ y f es integrable, entonces g es integrable.

Demostración:

i) Existe $c \in]\alpha,\beta[$ tal que

$$\frac{|L|}{2} < \left|\frac{f(x)}{g(x)}\right| < 2|L|, \quad \forall x \in [c,\beta[$$

y, en consecuencia,

$$\frac{|L|}{2}\,|g(x)| < |f(x)| < 2|L|\,|g(x)|, \qquad \forall x \in [c,\beta[$$

de donde se deduce que f es integrable en $[c,\beta[$ si, y sólo si, g es integrable en $[c,\beta[$. Como f y g son integrables en $[\alpha,c]$, de lo anterior deducimos i).

ii) Existe $c \in]\alpha,\beta[$ tal que

$$\left|\frac{f(x)}{g(x)}\right| < 1, \qquad \forall x \in [c,\beta[$$

y en consecuencia

$$|f(x)| < |g(x)|, \qquad \forall x \in [c,\beta[$$

de donde, supuesto g integrable, deducimos que f es integrable en $[c,\beta[$. Como f es integrable en $[\alpha,c]$ de lo anterior concluimos que f es integrable en $[\alpha,\beta[$.

iii) Existe $c \in]\alpha,\beta[$ tal que $\left|\frac{f(x)}{g(x)}\right| > 1,\ \forall x \in [c,\beta[$ y en consecuencia

$$|f(x)| > |g(x)|, \forall x \in [c,\beta[$$

de donde, supuesto f integrable, deducimos que g es integrable en $[c,\beta[$. Como g es integrable en $[\alpha,c]$ de lo anterior concluimos que g es integrable en $[\alpha,\beta[$. ∎

Ejemplo 13.19. La función $f:\mathbb{R}\to\mathbb{R}$ definida por

$$f(x) = e^{-x^2},\ \forall x \in \mathbb{R},$$

es localmente integrable por ser continua. Por simetría, nos podemos limitar a estudiar su integrabilidad en $[0,+\infty[$ y basta comparar con e^{-x} que es integrable $[0,+\infty[$ (Ejemplo 13.10) ya que

$$\frac{e^{-x^2}}{e^{-x}} = e^{x-x^2} \to 0 \quad \text{cuando} \quad x \to +\infty.$$

□

13.6. Funciones definidas por integrales

En esta sección $(\Omega, \mathscr{A}, \mu)$ es un espacio de medida, C un conjunto y $f : C \times \Omega \to \mathbb{R}$ una función tales que

$$\forall t \in C, \quad \text{la función } \ \omega \longmapsto f(t,\omega) \ \text{ es integrable en } \Omega.$$

Se trata de estudiar propiedades de la función $F : C \to \mathbb{R}$ definida por integrales de la siguiente forma

$$F(t) := \int_\Omega f(t,\omega)\, d\mu \ , \forall t \in C,$$

a partir de propiedades de la función f. Concretamente estamos interesados en dar condiciones suficientes que permitan asegurar la continuidad, derivabilidad y el cálculo de su derivada.

Teorema 13.20 (de continuidad de funciones definidas por integrales). *Sean (E,d) un espacio métrico, $(\Omega, \mathscr{A}, \mu)$ un espacio de medida y $f : E \times \Omega \to \mathbb{R}$ una función satisfaciendo las siguientes condiciones:*

i) $\forall t \in E$, *la función* $\omega \longmapsto f(t,\omega)$ *es integrable en* Ω.

ii) $\forall \omega \in \Omega$, *la función* $t \longmapsto f(t,\omega)$ *es continua en* E.

iii) Existe $g : \Omega \to \mathbb{R}$ *integrable tal que*

$$|f(t,\omega)| \leq g(\omega),\ \forall t \in E,\ \forall \omega \in \Omega\ .$$

Entonces la función $F : E \to \mathbb{R}$, *definida por*

$$F(t) = \int_\Omega f(t,\omega)\, d\mu\ ,$$

es continua en E.

Demostración. Sea $t_0 \in E$ y sea $\{t_n\}$ una sucesión en E convergente a t_0. Para cada natural n sea $\Phi_n : \Omega \to \mathbb{R}$ la función definida por

$$\Phi_n(\omega) = f(t_n,\omega),\ \forall \omega \in \Omega.$$

Por la hipótesis i) cada Φ_n es integrable en Ω, y la aplicación $\omega \longmapsto f(t_0,\omega)$ es medible (de hecho también integrable). Por la hipótesis ii) se tiene que

$$\{\Phi_n(\omega)\} \longrightarrow f(t_0,\omega),\ \forall \omega \in \Omega.$$

Por último la hipótesis iii) nos asegura la mayoración:

$$|\Phi_n(\omega)| \leq g(\omega),\ \forall \omega \in \Omega.$$

Así el teorema de la convergencia dominada (Teorema 12.1) nos permite concluir que

$$\lim \int_\Omega \Phi_n(\omega)\, d\mu = \int_\Omega f(t_0,\omega)\, d\mu,$$

es decir,

$$\left\{\int_\Omega f(t_n,\omega)\, d\mu\right\} \longrightarrow \int_\Omega f(t_0,\omega)\, d\mu,$$

o lo que es igual,

$$\{F(t_n)\} \longrightarrow F(t_0).$$

∎

Teorema 13.21 (de derivación de funciones definidas por integrales). *Sean I un intervalo de $\mathbb{R}$, $(\Omega,\mathscr{A},\mu)$ un espacio de medida y $f : I\times\Omega \to \mathbb{R}$ una función satisfaciendo las siguientes condiciones:*

i) $\forall t \in I$, la función $\omega \longmapsto f(t,\omega)$ es integrable en Ω.

ii) $\forall \omega \in \Omega$, la función $t \longmapsto f(t,\omega)$ es derivable en I.

iii) Existe $g : \Omega \to \mathbb{R}$ integrable tal que

$$\left|\frac{\partial f}{\partial t}(t,\omega)\right| \leq g(\omega),\ \forall t \in I,\ \forall \omega \in \Omega\ .$$

Entonces la función $F : I \to \mathbb{R}$, definida por

$$F(t) = \int_\Omega f(t,\omega)\, d\mu\ ,$$

es derivable en I con

$$F'(t) = \int_\Omega \frac{\partial f}{\partial t}(t,\omega)\, d\mu,\ \forall t \in I.$$

Demostración. Sea $t_0 \in I$ y sea $\{t_n\}$ una sucesión de puntos de I distintos de t_0 y convergente a t_0. Para cada natural n y para cada $\omega \in \Omega$, la función $f(\cdot,\omega)$, en virtud de la hipótesis ii), es derivable en el intervalo determinado por los puntos t_0 y t_n. El Teorema del valor medio nos garantiza, en consecuencia, la existencia de un punto c_n del intervalo abierto de extremos t_0 y t_n tal que

$$\left|\frac{f(t_n,\omega)-f(t_0,\omega)}{t_n-t_0}\right| = \left|\frac{\partial f}{\partial t}(c_n,\omega)\right|.$$

La hipótesis iii) nos asegura ahora que

$$\left|\frac{f(t_n,\omega)-f(t_0,\omega)}{t_n-t_0}\right| \leq g(\omega),\ \forall \omega \in \Omega.$$

Si para cada natural n se define $\Phi_n : \Omega \to \mathbb{R}$ por

$$\Phi_n(\omega) = \frac{f(t_n,\omega) - f(t_0,\omega)}{t_n - t_0},$$

obtenemos una sucesión de funciones integrables en Ω (hipótesis i)), tal que

$$|\Phi_n(\omega)| \leq g(\omega),\ \forall \omega \in \Omega.$$

Es claro que para cada $\omega \in \Omega$ se verifica que

$$\{\Phi_n(\omega)\} \longrightarrow \frac{\partial f}{\partial t}(t_0,\omega),$$

y como como la función $\omega \longrightarrow \dfrac{\partial f}{\partial t}\ (t_0,\omega)$ es medible (para cada $t \in I$ la función $\omega \longmapsto \dfrac{\partial f}{\partial t}(t,\omega)$ es medible por ser límite de funciones medibles), el teorema de la convergencia dominada (Teorema 12.5) nos asegura que dicha función es integrable y que

$$\lim \int_\Omega \frac{f(t_n,\omega) - f(t_0,\omega)}{t_n - t_0}\, d\mu = \int_\Omega \frac{\partial f}{\partial t}\ (t_0,\omega),$$

es decir,

$$\frac{F(t_n) - F(t_0)}{t_n - t_0} \longmapsto \int_\Omega \frac{\partial f}{\partial t}(t_0,\omega),$$

o lo que es igual que la función F es derivable en t_0 con

$$F'(t_0) = \int_\Omega \frac{\partial f}{\partial t}\ (t_0,\omega).$$

■

Nota 13.22. El apartado *iii*) de ambos teoremas, la mayoración por una función integrable, es la condición más difícil de verificar. A este respecto conviene resaltar que tanto la continuidad como la derivabilidad son propiedades locales, lo que facilita en ocasiones la tarea de encontrar tal mayoración. También un apropiado cambio de variable puede facilitar la tarea de estudiar la continuidad y derivabilidad.

El siguiente ejemplo estudia una de las funciones importantes del Análisis.

Ejemplo 13.23 (La función Γ). Sea $\Gamma : \mathbb{R}^+ \to \mathbb{R}$ la función definida por

$$\Gamma(t) = \int_0^{+\infty} x^{t-1} e^{-x}\, dx.$$

Probar que

i) Γ es derivable y dar una expresión de su derivada.

ii) $\lim_{t\to 0} \Gamma(t) = +\infty$.

En efecto: consideremos la función $f : \mathbb{R}^+ \times \mathbb{R}^+ :\longrightarrow \mathbb{R}$ definida por

$$f(t,x) = x^{t-1}e^{-x}.$$

Es fácil comprobar (ver Ejercicio 13.3) que

$$t \in \mathbb{R}^+ \Longrightarrow x^{t-1}e^{-x} \in \mathscr{L}^1(\mathbb{R}^+),$$

es decir, está definida la función

$$\Gamma(t) = \int_0^{+\infty} x^{t-1}e^{-x}\,dx,\ \forall t \in \mathbb{R}^+.$$

Derivando respecto a la variable t, se tiene que

$$\frac{\partial f}{\partial t}(t,x) = x^{t-1}\ln x\, e^{-x},\ \forall t \in \mathbb{R}^+.$$

La dificultad, como ya hemos comentado, está en comprobar la hipótesis *iii*) del Teorema 13.21 : la existencia de una función $g : \Omega :\longrightarrow \mathbb{R}$ integrable tal que

$$\left|\frac{\partial f}{\partial t}(t,\omega)\right| \leq g(\omega),\ \forall t \in \mathbb{R}^+,\ \forall \omega \in \Omega,$$

viene en nuestra ayuda el carácter local. Sea $0 < t_0$. Tomemos $\alpha, \beta \in \mathbb{R}$ tales que

$$0 < \alpha < t_0 < \beta.$$

Comprobemos que dicha hipótesis se verifica para $I \times \Omega =]\alpha, \beta[\times \mathbb{R}^+$. En efecto, como

$$\alpha - 1 < t - 1 < \beta - 1 \Longrightarrow x^{t-1} \leq max\left\{x^{\alpha-1}, x^{\beta-1}\right\} \leq x^{\alpha-1} + x^{\beta-1},$$

se tiene que

$$\left|\frac{\partial f}{\partial t}(t,x)\right| = x^{t-1}e^{-x}|\ln x| \leq \left(x^{\alpha-1} + x^{\beta-1}\right)e^{-x}|\ln x|,\ \forall (t,x) \in]\alpha,\beta[\times \mathbb{R}^+,$$

con lo cual basta probar que

$$\left(x^{\alpha-1} + x^{\beta-1}\right)e^{-x}|\ln x| \in \mathscr{L}^1(R^+).$$

En $]0,1[$ esta función se comporta como $x^\rho \ln x$ que es fácil comprobar (véase el Ejercicio 13.3) es integrable si, y sólo si, $\rho > -1$. En $[1,+\infty[$ se comporta como x^{-2} que es integrable. En efecto

$$\frac{x^\rho e^{-x}\ \ln x}{x^{-2}} = \frac{x^{\rho+2}\ \ln x}{e^x} = \frac{x^{\rho+3}}{e^x}\ \frac{\ln x}{x} \longrightarrow 0 \ \text{ cuando } \ x \longrightarrow +\infty.$$

Por último el carácter local de la derivabilidad nos asegura que la función Γ es derivable con

$$\Gamma'(t) = \int_0^{+\infty} x^{t-1}e^{-x}\ \ln x\,dx, \forall t \in \mathbb{R}^+.$$

Veamos ahora el apartado $ii)$. Sea $\{t_n\}$ una sucesión de positivos decreciente a cero. Para cada natural n consideremos la función $f_n :]0,1[\to \mathbb{R}$ definida por

$$f_n(x) = x^{t_n - 1} e^{-x}.$$

$\{f_n\}$ es una sucesión creciente de funciones medibles positivas que converge a la función

$$f(x) = x^{-1} e^{-x}, \ \forall x \in]0,1[$$

que es medible positiva pero no integrable (véase el Ejercicio 13.3). El Teorema de la convergencia creciente para funciones medibles positivas (Corolario 12.3) nos asegura que

$$\int_0^1 x^{t_n - 1} e^{-x} \, dx \nearrow \int_0^1 x^{-1} e^{-x} \, dx = +\infty.$$

La afirmación ii) es ahora consecuencia de la desigualdad

$$\Gamma(t) \geq \int_0^1 x^{t-1} e^{-x} \, dx, \ \forall t \in \mathbb{R}^+.$$

□

13.7. Continuidad absoluta

Proposición 13.24. *Sea $f : [a,b] \to \mathbb{R}$ una función integrable, y $F : [a,b] \to \mathbb{R}$ la función definida por*

$$F(x) = \int_a^x f(t) \, dt, \ \forall x \in [a,b]$$

verifica que para cada $\varepsilon > 0$ existe $\delta > 0$ tal que si $\{]a_i, b_i[\}_{i=1,2,\ldots,n}$ es un conjunto finito de subintervalos de $[a,b]$ disjuntos dos a dos con $\sum_{i=1}^{n} (b_i - a_i) < \delta$, entonces

$$\sum_{i=1}^{n} |F(b_i) - F(a_i)| < \varepsilon \, .$$

Demostración. Sea $\varepsilon > 0$. Sabemos (ver Corolario 12.4) que existe $\delta > 0$ tal que

$$\big[E \subset [a,b] \ \text{ medible con } \ \lambda(E) < \delta \big] \quad \Longrightarrow \quad \int_E |f(x)| dx < \varepsilon \, .$$

Sea finalmente $\{]a_i, b_i[\}_{i=1,2,\ldots,n}$ un conjunto finito de subintervalos de $[a,b]$ disjuntos con $\sum_{i=1}^{n} (b_i - a_i) < \delta$. El conjunto $E = \bigcup_{i=1}^{n}]a_i, b_i[$ verifica que $\lambda(E) < \delta$ y, en consecuencia,

$$\sum_{i=1}^{n} |F(b_i) - F(a_i)| = \sum_{i=1}^{n} \left| \int_{a_i}^{b_i} f(t) \, dt \right| \leq \sum_{i=1}^{n} \int_{a_i}^{b_i} |f(t)| \, dt = \int_E |f(t)| \, dt < \varepsilon \, .$$

■

Definición 13.25. Función absolutamente continua. Una función $f:[a,b]\to\mathbb{R}$ se dice absolutamente continua si para cada $\varepsilon>0$ existe $\delta>0$ tal que si $\{]a_i,b_i[\}_{i=1,2,\dots,n}$ es un conjunto finito de subintervalos de $[a,b]$ disjuntos dos a dos con $\sum_{i=1}^{n}(b_i-a_i)<\delta$, entonces

$$\sum_{i=1}^{n}|f(b_i)-f(a_i)|<\varepsilon\,.$$

Nótese que la condición anterior proporciona para $n=1$ la continuidad uniforme de la función f. Sin embargo no toda función uniformemente continua es absolutamente continua. La función singular de Lebesgue (10.24) es uniformemente continua y no es absolutamente continua como consecuencia del siguiente resultado.

Teorema 13.26. *Sea $f:[a,b]\longrightarrow\mathbb{R}$ una función absolutamente continua. Entonces f es derivable* c.p.d. *en $[a,b]$, la función f' es integrable en $[a,b]$ y se verifica que*

$$f(x)-f(a)=\int_a^x f'(t)\,dt,\ \forall x\in[a,b]\,.$$

Es decir, las funciones absolutamente continuas son precisamente aquellas que se pueden reconstruir a partir de su derivada.

La demostración de este teorema requiere también la utilización de técnicas complejas impropias de un primer curso de Análisis

13.8. Resumen de resultados del Tema 13.

Teorema fundamental del cálculo (TFC) *Sea $f:]\alpha,\beta[\to\mathbb{R}$ una función localmente integrable y sea $a\in]\alpha,\beta[$ un punto fijo. La integral indefinida de f de orige a , esto es*

$$F(x)=\int_a^x f(t)\,dt,\ \forall x\in]\alpha,\beta[$$

verifica las siguientes afirmaciones:

i) Es continua.

ii) Es derivable c.p.d. con $F'=f$ c.p.d. Además F es derivable en cada punto x del intervalo $]\alpha,\beta[$ en el que f es continua con $F'(x)=f(x)$.

Es claro que dos integrales indefinidas se diferencian en una constante.

Teorema (Regla de Barrow) *Sea $f:]\alpha,\beta[\to\mathbb{R}$ una función que admite primitiva, esto es existe $G:]\alpha,\beta[\to\mathbb{R}$ derivable tal que*

$$G'(x)=f(x),\ \forall x\in]\alpha,\beta[.$$

Se verifican las siguientes afirmaciones

i) f es medible.

ii) Si $f:]\alpha,\beta[\to[0,+\infty[$, entonces

$$\int_\alpha^\beta f(x)\,dx=\lim_{x\to\beta}G(x)-\lim_{x\to\alpha}G(x)\,.$$

iii) Si f es integrable, entonces la primitiva G tiene límites (reales) en α,β , y

$$\int_\alpha^\beta f(x)\,dx=\lim_{x\to\beta}G(x)-\lim_{x\to\alpha}G(x).$$

Cualquier función que se obtenga sumando a G una función constante es también primitiva de f y de esta forma se obtienen todas las primitivas de f.

Criterio de integrabilidad *Una función $f:]\alpha,\beta[\to[0,+\infty[$*
que admite primitiva es integrable, si, y sólo si, su primitiva G tiene límites (reales) en α,β.
En el caso de funciones integrables $f:]\alpha,\beta[\to\mathbb{R}$, la condición anterior es sólo necesaria.

Fórmula de integración por partes *Sean $u,v:]\alpha,\beta[\to\mathbb{R}$ funciones derivables tales que $u'v$ y uv' son integrables. Entonces existen*

$$\lim_{x\to\alpha}u(x)v(x)\ \text{ y }\ \lim_{x\to\beta}u(x)v(x)$$

y se verifica

$$\int_\alpha^\beta u(t)v'(t)\,dt = \lim_{x\to\beta} u(x)v(x) - \lim_{x\to\alpha} u(x)v(x) - \int_\alpha^\beta u'(t)v(t)\,dt.$$

Teorema de cambio de variable para funciones medibles positivas *Sean I y J intervalos abiertos de $\mathbb{R}$, ϕ un difeomorfismo de clase $\mathscr{C}^1$ de I sobre J y $f : J \to [0,+\infty[$ una función medible. Entonces*

$$\int_J f(x)\,dx = \int_I f(\phi(t))|\phi'(t)|\,dt.$$

Teorema de cambio de variable *Sean I y J intervalos abiertos de $\mathbb{R}$, ϕ un difeomorfismo de clase $\mathscr{C}^1$ de clase $\mathscr{C}^1$ I sobre J y $f : J \to \mathbb{R}$ una función medible. Entonces*

$$f \in \mathscr{L}(J) \;\Leftrightarrow (f\circ\phi)|\phi'| \in \mathscr{L}^1(I),$$

en cuyo caso

$$\int_J f(x)\,dx = \int_I f(\phi(t))|\phi'(t)|\,dt.$$

Fórmula del cambio de variable *Sean $-\infty \le \alpha < \beta \le +\infty$, f una función real medible definida en $]\alpha,\beta[$ y $\phi :]a,b[\to]\alpha,\beta[$ un difeomorfismo de clase $\mathscr{C}^1$. Entonces*

$$f \in \mathscr{L}^1(]\alpha,\beta]) \Longleftrightarrow (f\circ\phi)\phi' \in \mathscr{L}^1(]a,b[)$$

y en caso afirmativo vale la siguiente fórmula del cambio de variable:

$$\int_\alpha^\beta f(x)\,dx = \int_{\phi^{-1}(\alpha^+)}^{\phi^{-1}(\beta^-)} f(\phi(t))\phi'(t)\,dt.$$

Criterio de comparación *Sean $-\infty < \alpha < \beta \le +\infty$ y $f,g : [\alpha,\beta[\to \mathbb{R}$ funciones localmente integrables con $g(x) \neq 0$, $\forall x \in [\alpha,\beta[$. Supongamos que*

$$\lim_{x\to\beta} \frac{f(x)}{g(x)} = L.$$

Entonces

i) Si $L \in \mathbb{R}^*$, f es integrable si, y sólo si, g es integrable.

ii) Si $L = 0$ y g es integrable, entonces f es integrable.

iii) Si $L = \pm\infty$ y f es integrable, entonces g es integrable.

Teorema de continuidad de funciones definidas por integrales *Sean (E,d) un espacio métrico, $(\Omega,\mathscr{A},\mu)$ un espacio de medida y $f : E\times\Omega \to \mathbb{R}$ una función satisfaciendo las siguientes condiciones:*

i) $\forall t \in E$, *la función* $\omega \longmapsto f(t,\omega)$ *es integrable en* Ω.

ii) $\forall \omega \in \Omega$, *la función* $t \longmapsto f(t,\omega)$ *es continua en* E.

iii) Existe $g : \Omega \to \mathbb{R}$ *integrable tal que*

$$|f(t,\omega)| \leq g(\omega),\ \forall t \in E,\ \forall \omega \in \Omega\ .$$

Entonces la función $F : E \to \mathbb{R}$, *definida por*

$$F(t) = \int_\Omega f(t,\omega)\, d\mu\ ,$$

es continua en E.

Teorema de derivación de funciones definidas por integrales *Sean I un intervalo de $\mathbb{R}$, $(\Omega,\mathscr{A},\mu)$ un espacio de medida y $f : I\times\Omega \to \mathbb{R}$ una función satisfaciendo las siguientes condiciones:*

i) $\forall t \in I$, *la función* $\omega \longmapsto f(t,\omega)$ *es integrable en* Ω.

ii) $\forall \omega \in \Omega$, *la función* $t \longmapsto f(t,\omega)$ *es derivable en* I.

iii) Existe $g : \Omega \to \mathbb{R}$ *integrable tal que*

$$\left|\frac{\partial f}{\partial t}(t,\omega)\right| \leq g(\omega),\ \forall t \in I,\ \forall \omega \in \Omega\ .$$

Entonces la función $F : I \to \mathbb{R}$, *definida por*

$$F(t) = \int_\Omega f(t,\omega)\, d\mu\ ,$$

es derivable en I *con*

$$F'(t) = \int_\Omega \frac{\partial f}{\partial t}(t,\omega)\, d\mu,\ \forall t \in I.$$

Definición de función absolutamente continua *Una función $f : [a,b] \to \mathbb{R}$ se dice absolutamente continua si para cada $\varepsilon > 0$ existe $\delta > 0$ tal que si $\{]a_i,b_i[\}_{i=1,2,\dots,n}$ es un conjunto finito de subintervalos de $[a,b]$ disjuntos dos a dos con $\sum_{i=1}^n (b_i - a_i) < \delta$, entonces*

$$\sum_{i=1}^n |f(b_i) - f(a_i)| < \varepsilon\, .$$

Teorema de caracterización de las integrales indefinidas *Sea $f : [a,b] \to \mathbb{R}$ una función. Equivalen las siguiente afirmaciones*

1. *f es absolutamente continua.*

2. *f es derivable* c.p.d. *en $[a,b]$, la función f' es integrable en $[a,b]$ y se verifica que*

$$f(x) - f(a) = \int_a^x f'(t)\, dt,\ \forall x \in [a,b]\,.$$

Es decir, las funciones absolutamente continuas son precisamente aquellas que se pueden reconstruir a partir de su derivada.

13.9. Ejercicios del Tema 13.

13.0. Probar que existen las siguientes integrales y que tienen el valor que se indica en cada caso:

$$\int_0^1 \frac{dx}{1+e^x} = 1+\ln\left(\frac{2}{1+e}\right) \ , \ \int_0^{\frac{1}{2}} \frac{dx}{\sqrt{20+8x-x^2}} = \operatorname{arc\,sen}\left(\frac{2}{3}\right) - \operatorname{arc\,sen}\left(\frac{7}{12}\right) \ ,$$

$$\int_0^3 \frac{dx}{\sqrt{9-x^2}} = \frac{\pi}{2} \ , \ \int_0^1 \frac{x^2}{\sqrt{1-x^6}}\,dx = \frac{\pi}{6} \ ,$$

$$\int_1^{+\infty} \frac{x-1}{x^3-3x^2+x+5}\,dx = \frac{3\pi+\ln 2}{10} \ , \ \int_0^{+\infty} \frac{x}{3+x^4}\,dx = \frac{\sqrt{3}\pi}{12}$$

$$\int_{-\infty}^{+\infty} \frac{dx}{e^x+e^{-x}} = \frac{\pi}{2} \ , \ \int_{-\pi}^{\pi} \cos^2 x\,dx = \int_{-\pi}^{\pi} \operatorname{sen}^2 x\,dx = \pi \ ,$$

$$\int_0^{+\infty} e^{-\alpha x}\cos(\beta x)\,dx = \frac{\alpha}{\alpha^2+\beta^2} \ , \int_0^{+\infty} e^{-\alpha x}\operatorname{sen}(\beta x)\,dx = \frac{\beta}{\alpha^2+\beta^2} \ (\alpha>0, \beta\in\mathbb{R}),$$

$$\int_{-1}^1 \sqrt{1-x^2}\,dx = \frac{\pi}{2} \ , \ \int_{-\pi}^{\pi} (1+\cos x)^2\,dx = 3\pi \ , \ \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} |\operatorname{sen} x|^3\,dx = \frac{4}{3} \ ,$$

$$\int_0^{\frac{\pi}{2}} \operatorname{sen}^2\vartheta\ \cos^2\vartheta\ d\vartheta = \frac{\pi}{16} \ , \ \int_0^1 \left(1-\rho^{\frac{2}{3}}\right)^{\frac{3}{2}} \rho\,d\rho = \frac{8}{35}$$

$$\int_0^{+\infty} \frac{dx}{1+x^2+y^2} = \frac{\pi}{2\sqrt{1+y^2}} \ , \ \int_0^1 \frac{dy}{\sqrt{1+y^2}} = \ln(1+\sqrt{2}) \ ,$$

$$\int_0^{+\infty} \frac{dx}{(1+y)(1+yx^2)} = \frac{\pi}{2(1+y)\sqrt{y}} \ (y>0), \int_0^{+\infty} \frac{dy}{(1+y)\sqrt{y}} = \pi \ , \ \int_0^1 \ln x\,dx = -1,$$

$$\int_0^{+\infty} \frac{dy}{(1+y)(1+yx^2)} = \int_0^{+\infty} \frac{1}{1-x^2}\left(\frac{1}{1+y} - \frac{x^2}{1+yx^2}\right)dy = \frac{2\ln x}{x^2-1} \ \ (x\neq\pm 1).$$

13.1 Para cualquier subconjunto E medible de $]0,1[$ se define

$$\nu(E) = \int_E \frac{dx}{x}.$$

Probar que ν es una medida en $]0,1[$ absolutamente continua respecto de la medida de Lebesgue λ en $]0,1[$ y sin embargo

$$0<\delta<\frac{1}{2} \Longrightarrow \nu(]\delta,2\delta[) = \ln 2 \quad \text{y} \quad \lambda(]\delta,2\delta[) = \delta \ .$$

¿Contradice este hecho la segunda parte del Corolario 12.4?

13.2 Dar un ejemplo de una función integrable en $]0,1[$ cuyo cuadrado no lo es.

13.3 Estudiar la integrabilidad de las siguientes funciones:

a) $x^{\rho}\ e^{-\gamma x}\ \ (\rho \in \mathbb{R},\ \gamma > 0)$ en $\mathbb{R}^{+}$.

b) $\dfrac{x}{e^{x}-1}$ en $\mathbb{R}^{+}$.

c) $x^{\rho}\ \ln x\ \ (\rho \in \mathbb{R})$ en $\mathbb{R}^{+}$.

d) $\dfrac{1}{\sqrt{x}}\ \mathrm{sen}\,\dfrac{1}{x}$ en $]0,1[$.

e) $x^{\alpha-1}(1-x)^{\beta-1}\ (\alpha,\beta \in \mathbb{R})$ en $]0,1[$.

f) $\ln x\ \ln(1+x)$ en $]0,1[$.

g) $x^{\rho}\ \mathrm{sen}\,x\ \ (\rho \in \mathbb{R})$ en $]1,+\infty[$.

Indicación: Usar en cada caso el criterio de comparación.

a): En 0, comparar con x^{ρ}; en $+\infty$, comparar con x^{-2}.

b): En 0 tiene límite; comparar con $\dfrac{x}{e^{x}}$ en $+\infty$ y usar el apartado anterior.

c): Comparar con x^{α} para conveniente α.

d): Comparar con $\dfrac{1}{\sqrt{x}}$.

e): En cero, comparar con $x^{\alpha-1}$; en 1, comparar con $(1-x)^{\beta-1}$.

f): En 0 y en 1, la función tiene límite.

g): Si $\rho < -1$, comparar con x^{ρ}; si $\rho = -1$, ver Ejemplo 13.9; por último si $-1 < \rho$, usar el caso anterior.

13.4 Justificar, haciendo uso en cada caso de un conveniente teorema de convergencia, las siguientes igualdades:

i)
$$\lim \int_0^1 \frac{nx\ln x}{1+n^2x^2}\,dx = 0.$$

ii)
$$\lim \int_0^1 \frac{n\sqrt{x}}{1+n^2x^2}\,dx = 0.$$

iii)
$$\lim \int_0^1 \frac{n^{\frac{3}{2}}x}{1+n^2x^2}\,dx = 0.$$

iv)
$$\lim \int_0^{+\infty} \left(1+\frac{x^2}{n}\right)^{-n} \operatorname{sen}\left(\frac{x}{n}\right) dx = 0$$

v)
$$\lim \int_0^{+\infty} \frac{dx}{\left(1+\frac{x}{n}\right)^n x^{\frac{1}{n}}} = 1.$$

vi)
$$\lim \int_0^{n} \left(1-\frac{x}{n}\right)^n x^{\alpha-1} dx = \int_0^{+\infty} e^{-x} x^{\alpha-1} dx \ \ (\alpha>0).$$

vii)
$$\int_0^{+\infty} \frac{x}{1+e^x}\, dx = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2}\ .$$

viii)
$$\int_0^{+\infty} \frac{e^{-ax}}{1+e^{-bx}} dx = \sum_{n=0}^{\infty} \frac{(-1)^n}{a+nb} \ \ (a,b>0) \ \text{ y deducir que}$$
$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = \frac{\pi}{4} \ , \ \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} = \ln 2.$$

ix)
$$\int_0^{+\infty} \frac{xe^{-ax}}{1-e^{-bx}} dx = \sum_{n=0}^{\infty} \frac{1}{(a+nb)^2} \ \ (a,b>0).$$

x)
$$\int_0^{1} \frac{x^p}{1-x} \ln\left(\frac{1}{x}\right) dx = \sum_{n=1}^{\infty} \frac{1}{(p+n)^2} \ (p>-1).$$

Indicación: Se emplea el teorema de la convergencia dominada (Teorema 12.5) en todos los apartados salvo en vii), ix) y x) en los que se utiliza el teorema de la convergencia absoluta (Teorema 12.8). La dificultad de comprobar las hipótesis del teorema de la convergemcia dominada es siempre la mayoración de la sucesión de funciones. Conviene tener presenta la desigualdad entre las medias geométrica y aritmética así como la siguiente identidad

$$0<|r|<1 \Longrightarrow \frac{1}{1+r} = \sum_{n=0}^{\infty} (-1)^n\, r^n,$$

que es consecuencia de la expresión que da la suma de los infinitas términos de una progresión geométrica de razón menor que uno y de la igualdad $\dfrac{1}{1+r} = \dfrac{1}{1-(-r)}$.

13.5 Estudiar la continuidad y derivabilidad de las siguientes funciones:

i)

$$F(t)=\int_0^{+\infty}\frac{e^{-x}-e^{-tx}}{x}\,dx,\ \forall t\in\mathbb{R}^+.$$

ii)

$$F(t)=\int_0^{\pi}\ln(1+\,t\cos x)\,dx,\ \forall t\in]-1,1[.$$

13.6 Probar que la función $F:\mathbb{R}\to\mathbb{R}$ definida por

$$F(t)=\int_{-\infty}^{+\infty}\cos(tx)\,e^{\frac{-x^2}{2}}\,dx,$$

es derivable. Utilizar el método de integración por partes en la expresión de F' para obtener que $F'(t)=-tF(t),\ \forall t\in\mathbb{R}$. Deducir de ello que la función de R en R dada por

$$t\longmapsto F(t)\,e^{\frac{t^2}{2}}$$

tiene derivada nula. Por último concluir, utilizando el Teorema del valor medio, que

$$F(t)=C\,e^{\frac{-t^2}{2}},\ \forall t\in\mathbb{R},$$

donde

$$C=\int_{-\infty}^{+\infty}e^{\frac{-x^2}{2}}\,dx.$$

Esta constante se calcula en el Ejercicio 14.7, concretamente $C=\sqrt{2\pi}$.

13.7 Probar que

1. La función
$$f(x)=\sqrt{x},\ \forall x\in[0,1]$$
es absolutamente continua.
2. La función
$$g(x)=\begin{cases}0 & \text{si } x=0\\ x\,\mathrm{sen}\left(\frac{\pi}{2x}\right) & \text{si } x\in]0,1]\end{cases}$$
no es absolutamente continua.

13.8 Sea $f:]\alpha,\beta[\to\mathbb{R}$ una función localmente integrable. Probar que f es integrable si, y sólo si, el conjunto

$$\left\{\int_s^t|f(x)|dx:\ \alpha<s<t<\beta\right\}$$

está acotado. En caso afirmativo

$$\int_\alpha^\beta f(x)\,dx=\lim\int_{s_n}^{t_n}f(x)\,dx$$

para cualesquiera sucesiones $\{s_n\}$ y $\{t_n\}$ en $]\alpha,\beta[$ con

$$s_n < t_n,\ \forall n \in \mathbb{N},\ \{s_n\} \searrow \alpha,\ \{t_n\} \nearrow \beta.$$

13.9 Estudiar la derivabilidad de la función

$$F(t) := \int_0^{+\infty} \frac{\operatorname{sen}(tx)}{1+x^2}\, dx\ ,\forall t \in \mathbb{R}^+$$

Indicación: Realizar el cambio de variable $y = tx$ y aplicar el teorema de derivación (Teorema13.21)

13.10 Sean I,J intervalos abiertos de $\mathbb{R}$ y $f : I \times J \to \mathbb{R}$ una función continua verificando

1. Para cada t de I la función $x \longmapsto f(t,x)$ está en $\mathscr{L}^1(J)$.
2. Para cada x en J la función $t \longmapsto f(t,x)$ es de clase $\mathscr{C}^1(J)$.
3. La función $\left|\dfrac{\partial f(t,x)}{\partial t}\right|$ está dominada por una función integrable.

Sea $g : I \to J$ una función de clase $\mathscr{C}^1$. Probar que para $a \in J$,la función

$$F(t) := \int_a^{g(t)} f(t,x)\, dx,\ \forall t \in I$$

es derivable con derivada

$$F'(t) = \int_a^{g(t)} \frac{\partial f(t,x)}{\partial t}\, dx + f(t,g(t))\, g'(t)\ ,\forall t \in I.$$

Indicación: Considerar la función

$$\phi : I \times J \to \mathbb{R}$$

dada por

$$\phi(u,v) = \int_a^v f(u,x)\, dx\ .$$

13.10. Soluciones a los ejercicios del Tema 13.

13.1 La función f es medible (positiva) por ser continua. La primera parte del Corolario 12.4 nos asegura que ν es una medida en $]0,1[$ absolutamente continua respecto de λ. No se contradice la segunda parte del citado teorema pues f no es integrable (véase el Ejemplo 13.9).

13.2 La función $x \longmapsto \dfrac{1}{\sqrt{x}}$ es integrable y su cuadrado, $x \longmapsto \dfrac{1}{x}$, no lo es (véase el Ejemplo 13.9).

13.3 **a)** $\boxed{x^\rho\ e^{-\gamma x} \in \mathscr{L}^1(\mathbb{R}^+) \Longleftrightarrow \rho > -1}$, para cada $\gamma \in \mathbb{R}^+$.

En efecto, si $x \in]0,1[$ se tiene que $x^\rho\ e^{-\gamma} \leq x^\rho\ e^{-\gamma x} \leq x^\rho$, esto es, la función se comporta como x^ρ. Así (véase Ejemplo 13.9) la función es integrable en $]0,1[$ si, y sólo si, $\rho > -1$. En $[1,+\infty[$ se compara con x^{-2} (ver Criterio de comparación (Proposición 13.18)) y la función es integrable:

$$\lim_{x\to+\infty} \frac{x^\rho\ e^{-\gamma x}}{x^{-2}} = \lim_{x\to+\infty} \frac{x^{\rho+2}}{e^{\gamma x}} = 0.$$

b) $\boxed{\dfrac{x}{e^x - 1} \in \mathscr{L}^1(\mathbb{R}^+)}$.

En efecto, en $]0,1[$ la función es integrable pues es continua y tiene límite en 0 y en 1. En $[1,+\infty[$ se comporta como $x\, e^{-x}$, que acabamos de ver que es integrable. En efecto,

$$\lim_{x\to+\infty} \frac{e^x}{e^x - 1} = 1.$$

c) $\boxed{x^\rho\ \ln x \notin \mathscr{L}^1(\mathbb{R}^+)}$.

En $[1,+\infty[$ la función es integrable si, y sólo si $\rho < -1$. En el caso $\rho < -1$, tomemos $\alpha \in \mathbb{R}$ con $\rho < \alpha < -1$, y basta comparar con x^α, que sabemos es integrable (véase Ejemplo 13.9):

$$\lim_{x\to+\infty} \frac{x^\rho \ln x}{x^\alpha} = \lim_{x\to+\infty} \frac{\ln x}{x^{\alpha-\rho}} = 0.$$

Si por el contrario $\rho \geq -1$, entonces basta compara con $\dfrac{1}{x}$ para concluir que la función no es integrable. En efecto:

$$\lim_{x\to+\infty} \frac{\frac{1}{x}}{x^\rho\ \ln x} = \lim_{x\to+\infty} \frac{1}{x^{\rho+1}\ \ln x} = 0.$$

En $]0,1[$ la función es integrable si, y sólo si $\rho > -1$. En el caso $\rho > -1$, tomemos $\alpha \in \mathbb{R}$ con $\rho > \alpha > -1$, y basta comparar con x^{α}, que sabemos es integrable (véase Ejemplo 13.9):

$$\lim_{x\to 0} \frac{x^{\rho} \ln x}{x^{\alpha}} = \lim_{x\to 0} x^{\rho-\alpha} \ \ln x = 0.$$

Si por el contrario $\rho \leq -1$, entonces basta comparar con $\frac{1}{x}$ para concluir que la función no es integrable. En efecto:

$$\lim_{x\to 0} \frac{\frac{1}{x}}{x^{\rho} \ \ln x} \lim_{x\to 0} \frac{x^{-(\rho+1)}}{\ln x} = 0.$$

d) $\boxed{\frac{1}{\sqrt{x}} \ \mathrm{sen} \ \frac{1}{x} \in \mathscr{L}^1(]0,1[)}$.

En efecto,

$$\left| \frac{1}{\sqrt{x}} \ \mathrm{sen} \frac{1}{x} \right| \leq \frac{1}{\sqrt{x}} \in \mathscr{L}^1(]0,1[).$$

También se puede razonar que la función es continua y tiene límite en 0 y en 1.

e) $\boxed{x^{\alpha-1}(1-x)^{\beta-1} \in \mathscr{L}^1(]0,1[) \Longleftrightarrow \alpha,\beta > 0}$.

En efecto,

$$\left.\begin{array}{l} x^{\alpha-1}(1-x)^{\beta-1} \in \mathscr{L}^1(]0,\frac{1}{2}[) \Longleftrightarrow \alpha-1 > -1 \\ x^{\alpha-1}(1-x)^{\beta-1} \in \mathscr{L}^1(]\frac{1}{2},1[) \Longleftrightarrow \beta-1 > -1 \end{array}\right\} \Longrightarrow$$
$$x^{\alpha-1}(1-x)^{\beta-1} \in \mathscr{L}^1(]0,1[) \Longleftrightarrow \alpha,\beta > 0.$$

f) $\boxed{\ln x \ \ln(1+x) \in \mathscr{L}^1(]0,1[)}$.

En efecto, es una función continua y tiene límite en 0 y en 1:

$$\lim_{x\to 0} \ \ln x \ \ln(1+x) = \lim_{x\to 1} \ \ln x \ \ln(1+x) = 0.$$

Para calcular el primer límite téngase en cuenta que

$$\ln x \ \ln(1+x) = \ln(1+x)^{\ln x} = \ln \left((1+x)^{\frac{1}{x}} \right)^{x \ \ln x}$$
$$\lim_{x\to 0} (1+x)^{x} = e \quad \text{y} \quad \lim_{x\to 0} x \ \ln x = 0.$$

g) $\boxed{x^\rho \ \operatorname{sen} x \in \mathscr{L}^1(]1,+\infty[) \Longleftrightarrow \rho < -1}$.
En efecto,

$$|x^\rho \ \operatorname{sen} x| \leq x^\rho \Longrightarrow \left[\rho < -1 \Longrightarrow x^\rho \ \operatorname{sen} x \in \mathscr{L}^1(]1,+\infty[)\right].$$

Si $\rho = -1$, entonces $\dfrac{\operatorname{sen} x}{x} \notin \mathscr{L}(]1,+\infty[)$ (Ejemplo13.8). Por último si $-1 < \rho$, al ser $|x^\rho \ \operatorname{sen} x| \geq \left|\dfrac{\operatorname{sen} x}{x}\right|$, concluimos que $x^\rho \ \operatorname{sen} x \notin \mathscr{L}^1(]1,+\infty[)$.

13.4 Es suficiente usar el teorema de la convergencia dominada (Teorema 12.5) para las sucesiones de funciones $\{f_n\}$ en la mayoría de los apartados. En los restantes se usa el teorema de la convergencia absoluta (Teorema12.8).

i) $f_n(x) = \dfrac{nx\log x}{1+n^2x^2}$. Es claro que la sucesión de funciones $\{f_n\}$ converge puntualmente a cero y además si $x \in]0,1[$, entonces se verifica que

$$|f_n(x)| \leq |\log x|.$$

Aplicando el método de integración por partes y la Regla de Barrow (Teorema 13.5), se obtiene que la función $x \mapsto \log x$ es integrable en $]0,1[$ y basta aplicar el criterio de comparación y la medibilidad de las funciones de la sucesión y el hecho de que cada una de las funciones racionales $x \mapsto \dfrac{nx}{1+n^2x^2}$ está dominada por $\dfrac{1}{2}$ $\big($para esto basta derivar la anterior función para cada n y evaluar en el único punto crítico, o bien, desarrollar la desigualdad $0 \leq (1-nx)^2\big)$. Por último, como consecuencia del Teorema de la convergencia dominada, la sucesión de integrales anterior converge a cero.

ii) Esta vez es claro que la sucesión de funciones también converge puntualmente a cero y se aplica el mismo argumento del apartado anterior a las funciones

$$f_n(x) = \frac{nx}{1+n^2x^2}\frac{1}{\sqrt{x}} \leq \frac{1}{\sqrt{x}},$$

que son claramente integrables en $]0,1[$.

iii) De nuevo se trata de la sucesión de las integrales de una sucesión de funciones que converge puntualmente a cero en $]0,1[$. Para usar el teorema de la convergencia dominada, basta considerar la cadena de desigualdades siguiente

$$x \in]0,1[\Rightarrow |f_n(x)| = \frac{(nx)^{\frac{3}{2}}}{1+n^2x^2}\frac{1}{\sqrt{x}} \leq \frac{1}{\sqrt{x}} \in \mathscr{L}^1(]0,1[),$$

donde hemos usado que la función $t \mapsto \dfrac{t^{\frac{3}{2}}}{1+t^2}$ está acotada por 1 $\left(\text{hacer el cambio } x = t^{\frac{1}{2}} \text{ y derivar la función } x \mapsto \dfrac{x^3}{1+x^4}\right)$.

iv) Por la continuidad de la función seno en cero, y por ser la sucesión $\left\{\left(1+\frac{x^2}{n}\right)^{-n}\right\}$ convergente a $\exp(-x^2)$, la sucesión de funciones $\{f_n\}$ dada por

$$f_n(x) = \left(1+\frac{x^2}{n}\right)^{-n} \operatorname{sen}\frac{x}{n}$$

converge puntualmente a cero.

Usando además que la sucesión $\left(1+\frac{x^2}{n}\right)^{-n}$ claramente verifica al desigualdad

$$\left(1+\frac{x^2}{n}\right)^n = 1+n\frac{x^2}{n}+\ldots \geq 1+x^2,$$

de donde se obtiene la desigualdad

$$|f_n(x)| \leq \frac{1}{1+x^2}$$

Por último, la integrabilidad en $\mathbb{R}^+$ de la última función (basta usar el criterio de comparación) y el teorema de la convergencia dominada nos asegura que la sucesión de integrales del enunciado converge a cero.

v) Es inmediato comprobar que en este caso la sucesión de funciones

$$f_n(x) = \frac{1}{\left(1+\frac{x}{n}\right)^n x^{\frac{1}{n}}}$$

converge puntualmente a la función $x \mapsto e^{-x}$. Para $n \geq 2$ se tiene que

$$|f_n(x)| \leq \frac{1}{\sqrt{x}} \qquad \forall x \in]0,1[, \qquad\qquad |f_n(x)| \leq \frac{1}{\left(1+\frac{x}{2}\right)^2} \qquad \forall x \geq 1$$

Las funciones que se han usado para mayorar son integrables, por tanto, aplicando el teorema de la convergencia dominada se obtiene que la sucesión de integrales converge a $\int_0^{+\infty} e^{-x}\,dx = 1$ (Ejemplo13.10).

vi) Para cada natural n, consideremos la función integrable dada por

$$f_n(x) = \left(1-\frac{x}{n}\right)^n x^{\alpha-1}\chi_{]0,n[}.$$

Aplicando la desigualdad de las medias para los números

$$a_1 = \ldots = a_n = 1-\frac{x}{n}, \qquad a_{n+1} = 1,$$

queda

$$\left(\left(1-\frac{x}{n}\right)^n\right)^{\frac{1}{n+1}} \leq \frac{n-x+1}{n+1} = 1-\frac{x}{n+1}$$

de donde se deduce inmediatamente la desigualdad

$$\left(1-\frac{x}{n}\right)^n \leq \left(1-\frac{x}{n+1}\right)^{n+1}.$$

Se tiene entonces

$$|f_n(x)| \leq e^{-x}x^{\alpha-1} \in \mathscr{L}^1(\mathbb{R}^+),$$

donde se ha usado lo probado en el problema anterior. Por último, el teorema de la convergencia monótona o de la convergencia dominada nos da la igualdad.

vii) Para $x>0$ se tiene

$$\frac{x}{1+e^x} = \frac{x}{e^x}\frac{1}{1+e^{-x}} = \frac{x}{e^x}\sum_{n=0}^{\infty}(-1)^n e^{-nx} =$$

$$= \sum_{n=0}^{\infty}(-1)^n x e^{-(n+1)x} = \sum_{n=1}^{\infty}(-1)^{n+1} x e^{-nx}.$$

Aplicaremos el teorema de la convergencia absoluta a la sucesión de funciones integrables

$$f_n(x) = (-1)^{n+1} x e^{-nx}, \qquad \forall x \in \mathbb{R}_0^+, \quad \forall n \in \mathbb{N}.$$

Integrando por partes, calculamos

$$\int_0^{+\infty} |f_n| = \int_0^{+\infty} x e^{-nx}\, dx = \frac{1}{n^2}.$$

Por tanto, al ser la serie $\sum_{n\geq 1}\frac{1}{n^2}$ convergente, se puede aplicar el teorema de la convergencia absoluta, obteniéndose así el enunciado del apartado vii).

viii) Primero desarrollamos en serie de la misma forma que en el apartado anterior, obteniéndose

$$\frac{e^{-ax}}{1+e^{-bx}} = e^{-ax}\sum_{n=0}^{\infty}(-1)^n e^{-bnx} = \sum_{n=0}^{\infty}(-1)^n e^{-(a+bn)x}.$$

En este caso, al integrar término a término se obtiene la serie armónica y, por tanto, no podemos usar el teorema de la convergencia absoluta como en el apartado anterior. En cambio, sí podremos usar el teorema de la convergencia dominada para la sucesión de funciones

$$f_n(x) = \sum_{k=0}^{n}(-1)^k e^{-(a+bk)x} = \sum_{k=0}^{n} e^{-ax}\left(-e^{-bx}\right)^k, \qquad \forall x \in \mathbb{R}_0^+, \forall n \in \mathbb{N}.$$

Calculando las sumas parciales de la serie anterior, tenemos

$$|f_n(x)| = \frac{|(-1)^{n+1}e^{-(a+bn)x}e^{-bx} - e^{-ax}|}{1+e^{-bx}} \leq$$

$$\leq \frac{2e^{-ax}}{1+e^{-bx}} \leq 2e^{-ax} \in \mathscr{L}^1(\mathbb{R}^+).$$

Por último, tenemos que para $a=1, b=2$

$$\int_0^{+\infty} \frac{e^{-x}}{1+e^{-2x}}\, dx = -\arctan(e^{-x})\Big]_{x=0}^{x=+\infty} = \frac{\pi}{4},$$

Para $a=1, b=1$, obtenemos

$$\int_0^{+\infty} \frac{e^{-x}}{1+e^{-x}}\, dx = -\log(1+e^{-x})\Big]_{x=0}^{x=+\infty} = \log 2.$$

ix) De nuevo podemos aplicar el teorema de la convergencia absoluta. En este caso, desarrollamos en serie de la siguiente forma:

$$\frac{xe^{-ax}}{1-e^{-bx}} = xe^{-ax}\sum_{n=0}^{\infty} e^{-nbx} = \sum_{n=0}^{\infty} xe^{-(a+bn)x}.$$

Tomamos la sucesión de funciones

$$f_n(x) = xe^{-(a+bn)x}, \quad \forall x \in \mathbb{R}_0^+, \forall n \in \mathbb{N}.$$

Aplicando el método de integración por partes, obtenemos que

$$\int_0^{+\infty} |f_n| = \int_0^{+\infty} xe^{-(a+bn)x}\, dx = \Big[\text{tomando } u=x, v'=e^{-(a+nb)x}\Big] = \frac{1}{(a+nb)^2}$$

y teniendo en cuenta que la serie $\displaystyle\sum_{n\geq 1}\int_0^{+\infty} |f_n|$ es convergente, el teorema de la convergencia absoluta nos da el resultado que se enunció.

x) Se tiene que

$$\frac{x^p}{1-x}\log\left(\frac{1}{x}\right) = x^p\log\left(\frac{1}{x}\right)\sum_{b=0}^{\infty} x^n = \sum_{n=0}^{\infty} x^{n+p}\log\left(\frac{1}{x}\right) \quad (0<x<1).$$

Ahora aplicaremos de nuevo el teorema de la convergencia absoluta para la sucesión de funciones integrables

$$f_n(x) = x^{n+p}\log\left(\frac{1}{x}\right),\ \forall x \in]0,1],\ \forall n \in \mathbb{N}\cup\{0\},$$

lo que es posible, al ser

$$\int_0^1 |f_n| = \int_0^1 x^{n+p}\log\left(\frac{1}{x}\right)\, dx =$$

$$\Big[\text{int. partes } u=-\log x, v'=x^{n+p}\Big] = \frac{1}{(n+p+1)^2},$$

tenemos que la serie

$$\sum_{n\geq 1}\int_0^1 |f_n|$$

es convergente. Obtenemos entonces que

$$\int_0^1 \frac{x^p}{1-x}\log\left(\frac{1}{x}\right)\, dx = \sum_{n=0}^{\infty}\frac{1}{(n+p+1)^2} = \sum_{n=1}^{\infty}\frac{1}{(p+n)^2}.$$

13.5 i) Sea $f:\mathbb{R}^+\times\mathbb{R}^+ :\longrightarrow \mathbb{R}$ la función definida por

$$f(t,x)=\frac{e^{-x}-e^{-tx}}{x}.$$

Sea $t\in\mathbb{R}^+$. La función de $\mathbb{R}^+$ en $\mathbb{R}$ dada por $x\longmapsto \dfrac{e^{-x}-e^{-tx}}{x}$ es integrable. En efecto, es localmente integrable por ser continua, y tiene límite en 0 igual a $t-1$, y en $+\infty$ se compara con x^{-2}. Así está definida la función

$$F(t)=\int_0^{+\infty}\frac{e^{-x}-e^{-tx}}{x}\,dx,\ \forall t\in\mathbb{R}^+.$$

Derivando respecto a la variable t, se tiene que

$$\frac{\partial f}{\partial t}(t,x)=e^{-tx},\ \forall t\in\mathbb{R}^+.$$

Sea $0<t_0$. Tomemos $\alpha\in\mathbb{R}$ tal que $0<\alpha<t_0$. Para $t\in]\alpha,+\infty[$ es

$$\left|\frac{\partial f}{\partial t}(t,x)\right|\le e^{-\alpha x}\in\mathscr{L}^1(R^+),$$

el teorema de derivación (Teorema13.21) para $I\times\Omega=]\alpha,+\infty[\times\mathbb{R}^+$ y el carácter local de la derivabilidad nos aseguran que la función F es derivable (en particular continua) con

$$F'(t)=\int_0^{+\infty}e^{-tx}\,dx,\ \forall t\in\mathbb{R}^+,$$

es decir

$$F'(t)=\frac{1}{t}\Rightarrow\ F(t)=\ln(t)+C\ .$$

Evaluando la función F en el punto 1, obtenemos que $C=0$. Así

$$F(t)=\ln(t),\ \forall t\in\mathbb{R}^+\ .$$

ii) Sea $f:]-1,1[\times]0,\pi[:\longrightarrow\mathbb{R}$ la función definida por

$$f(t,x)=\ln(1+t\ \cos x).$$

Sea $t\in]-1,1[$. La función de $]0,\pi[$ en $\mathbb{R}$ dada por $x\longmapsto\ln(1+t\ \cos x)$ es integrable (continua y tiene límite en los extremos del intervalo de integración). Así está definida la función

$$F(t)=\int_0^{\pi}\ln(1+t\ \cos x)\,dx,\ \forall t\in]-1,1[.$$

Derivando respecto a la variable t, se tiene que

$$\frac{\partial f}{\partial t}(t,x)=\frac{\cos x}{1+t\ \cos x},\ \forall t\in]-1,1[.$$

Sea $t_0 \in]-1,1[$. Tomemos $\alpha \in \mathbb{R}$ tal que $|t_0| < \alpha < 1$. Para $t \in]-\alpha,+\alpha[$ es

$$\left|\frac{\partial f}{\partial t}(t,x)\right| = \left|\frac{\cos x}{1+t\ \cos x}\right| \leq \frac{|\cos x|}{1-|t||\cos x|} \leq \frac{1}{1-\alpha|\cos x|} \in \mathscr{L}^1(]0,\pi[),$$

el teorema de derivación (Teorema 13.21) para $I \times \Omega =]-\alpha,+\alpha[\times]0,\pi[$ y el carácter local de la derivabilidad nos aseguran que la función F es derivable (en particular continua) con

$$F'(t) = \int_0^\pi \frac{\cos x}{1+t\ \cos x}\, dx,\ \forall t \in]-1,1[.$$

13.6 Sea $f : \mathbb{R}\times\mathbb{R} \longrightarrow \mathbb{R}$ la función definida por

$$f(t,x) = \cos(tx)\ e^{\frac{-x^2}{2}}.$$

Sea $t \in \mathbb{R}$. La función de $\mathbb{R}$ en $\mathbb{R}$ dada por $x \longmapsto \cos(tx)\ e^{\frac{-x^2}{2}}$ es integrable. En efecto $|f(t,x)| \leq e^{\frac{-x^2}{2}} \in \mathscr{L}^1(\mathbb{R})$ (véase el Ejemplo 13.19, de hecho en el Ejercicio 14.7 se calcula dicha integral). Así está definida la función

$$F(t) = \int_{-\infty}^{+\infty} \cos(tx)\ e^{\frac{-x^2}{2}}\, dx,\ \forall t \in \mathbb{R}.$$

Derivando respecto a la variable t, se tiene que

$$\frac{\partial f}{\partial t}(t,x) = -x\,\mathrm{sen}(tx)\ e^{\frac{-x^2}{2}},\ \forall t \in \mathbb{R}.$$

Para $t \in \mathbb{R}$ es

$$\left|\frac{\partial f}{\partial t}(t,x)\right| \leq |x|e^{\frac{-x^2}{2}} \in \mathscr{L}^1(\mathbb{R}).$$

En efecto, dicha función es medible positiva y por simetría se tiene

$$\int_{-\infty}^{+\infty} |x|e^{\frac{-x^2}{2}}\, dx = 2\int_0^{+\infty} x\ e^{\frac{-x^2}{2}}\, dx = 2\left[-e^{\frac{-x^2}{2}}\right]_{x=0}^{x=+\infty} = 2(-0+1) = 2,$$

luego es integrable (por tener integral finita). El teorema de derivación (Teorema 13.21) nos asegura que la función F es derivable (en particular continua) con

$$F'(t) = \int_{-\infty}^{+\infty} -x\ \mathrm{sen}(tx)\ e^{\frac{-x^2}{2}}\, dx = 0 - \int_{-\infty}^{+\infty} t\ \cos(tx)\ e^{\frac{-x^2}{2}}\, dx = -tF(t)\ ,$$

donde se ha utilizado el método de integración por partes. Así la función de $\mathbb{R}$ en $\mathbb{R}$ definida por

$$t \longmapsto F(t)\ e^{\frac{t^2}{2}}$$

tiene derivada nula. En consecuencia, en virtud del teorema del valor medio, existe $C \in \mathbb{R}$, tal que

$$F(t) = C\ e^{\frac{-t^2}{2}},\ \forall t \in \mathbb{R},$$

y basta evaluar la constante en cero

$$C = F(0) = \int_{-\infty}^{+\infty} e^{\frac{-x^2}{2}}.$$

Esta constante se calcula en el Ejercicio 14.7, concretamente $C = \sqrt{2\pi}$.

Hemos probado que

$$F(t) = \sqrt{2\pi}\, e^{\frac{-t^2}{2}},\ \forall t \in \mathbb{R}.$$

13.7 1. Para probar que la función f es absolutamente continua basta tener en cuenta que

$$f(x) = \int_0^x \frac{1}{2\sqrt{x}},\ \forall x \in [0,1]\ ,$$

y el resultado se sigue de la Proposición 13.24 .

2. La función g no es absolutamente continua, en efecto sea $\delta > 0$. Tomemos $n_0 \in \mathbb{N}$ tal que $\dfrac{1}{4n_0} < \delta$. Consideremos los subintervalos disjuntos del intervalo $[0,1]$

$$I_k = [a_k, b_k] = \left[\frac{1}{4k+1}, \frac{1}{4k}\right],\ k = n_0\ ,\ n_0+1\ ,\ \ldots\ ,n_0+n\ .$$

Es claro que

$$\sum_{k=n_0}^{n_0+n} \ell(I_k) < \delta\ ,$$

mientras que

$$\sum_{k=n_0}^{n_0+n} |g(b_k) - g(a_k)| = \sum_{k=n_0}^{n_0+n} \frac{1}{4k+1} > \frac{1}{4} \sum_{k=n_0}^{n_0+n} \frac{1}{k+1} > 1$$

si n es suficientemente grande pues la serie

$$\sum_{k \geq n_0} \frac{1}{k+1}$$

es divergente.

13.8 Sea

$$M := \sup \left\{ \int_s^t |f(x)|dx :\ \alpha < s < t < \beta \right\} \in [0,\infty].$$

Si f es integrable, entonces

$$\int_s^t |f(x)|dx \leq \int_\alpha^\beta |f(x)|\, dx\ (\alpha < s < t < \beta)$$

y en consecuencia $M \in \mathbb{R}$.

Recíprocamente supongamos $M \in \mathbb{R}$, consideremos $\{s_n\}$ y $\{t_n\}$ sucesiones en $]\alpha,\beta[$ con

$$s_n < t_n \text{ para cada natural } n \text{ y } \{s_n\} \searrow \alpha,\ \{t_n\} \nearrow \beta\ .$$

Para cada natural n definimos la función $f_n = |f|\ \chi_{[s_n,t_n]}$. Como $\{f_n\} \nearrow |f|$ y la sucesión $\{\int f_n\}$ esta acotada, la versión práctica del teorema de la convergencia monótona (Corolario 12.2) nos asegura que $|f| \in \mathscr{L}^1(]\alpha,\beta[)$, y en consecuencia, al ser f medible, también $f \in \mathscr{L}^1(]\alpha,\beta[)$. Finalmente sean $\{s_n\},\{t_n\}$ como en el enunciado, consideremos para cada natural n la función $g_n = f\chi_{]s_n,t_n[}$, es claro que la sucesión $\{g_n\} \to f$, así, en virtud del teorema de la convergencia dominada (Teorema 12.5), concluimos que

$$\int_\alpha^\beta f(x)\,dx = \lim \int_{s_n}^{t_n} f(x)\,dx\ .$$

Obsérvese que la fórmula anterior es válida aunque las sucesiones no sean monótonas, esto es, basta que

$$s_n < t_n \text{ para cada natural } n \text{ y } \{s_n\} \to \alpha,\ \{t_n\} \to \beta\ .$$

13.9 Aplicaremos el Teorema de derivación. Para ello, empezamos por comprobar que la función $x \mapsto \frac{\operatorname{sen} tx}{1+x^2}$ es integrable en $\mathbb{R}^+$ para cada $t > 0$ fijo. Basta usar que la función anterior es medible por ser continua, el criterio de comparación, la desigualdad

$$\left|\frac{\operatorname{sen} tx}{1+x^2}\right| \le \frac{1}{1+x^2},\quad \forall x > 0$$

y la regla de Barrow para la función positiva $x \mapsto \frac{1}{1+x^2}$, cuya primitiva (arctan) tiene límites en 0 y en $+\infty$.

Ahora podemos aplicar el Teorema del cambio de variable. Para cada $t > 0$ fijo, hacemos el cambio $y = tx$, con lo que obtenemos

$$F(t) = \int_0^{+\infty} \frac{\operatorname{sen} y}{1+\frac{y^2}{t^2}}\frac{1}{t}\,dy = \int_0^{+\infty} \frac{t\operatorname{sen} y}{t^2+y^2}\,dy\ .$$

Ahora aplicaremos el Teorema de derivación para la función

$$f(t,x) = \frac{t\operatorname{sen} x}{t^2+x^2}\quad (t,x>0)\ .$$

a) El Teorema del cambio de variable (por el argumento anterior) nos asegura que para cada $t > 0$ fijo la función, $x \mapsto f(t,x)$ es integrable en $\mathbb{R}^+$.

b) Es claro que para cada $x > 0$ fijo, la función $t \mapsto f(t,x)$ es derivable en $\mathbb{R}^+$ con derivada

$$\frac{\partial f}{\partial t}(t,x) = \frac{\operatorname{sen} x(t^2+x^2) - t\operatorname{sen} x(2t)}{(t^2+x^2)^2}\ .$$

c) Para comprobar la última hipótesis del Teorema de derivación, basta acotar, para cada $0 < \alpha < \beta$ y $t \in [\alpha, \beta]$, como sigue

$$\left|\frac{\partial f}{\partial t}(t,x)\right| \leq$$

$$\frac{|\operatorname{sen} x|}{t^2+x^2} + \frac{2t^2}{(t^2+x^2)^2} \leq$$

$$\frac{|\operatorname{sen} x|}{\alpha^2+x^2} + \frac{2\beta^2}{(\alpha^2+x^2)^2}\ .$$

La primera función de la desigualdad anterior es integrable. Por ser continua en $[0,1]$, es integrable en $[0,1]$ y, por ser comparable en $+\infty$ con $\frac{1}{1+x^2}$, que es integrable. La segunda función obviamente es localmente integrable y comparable en $+\infty$ con $\frac{1}{x^4}$, que es integrable en $[1,+\infty[$.

13.10 Veamos que la función ϕ tiene derivadas parciales continuidad y, por tanto, es una función de clase $\mathscr{C}^1$.

Derivabilidad respecto de u. Fijamos $u \in I$. La función

$$u \longmapsto \int_a^v f(v,x\,du$$

es derivable por el teorema fundamental del cálculo (para funciones continuas) con derivada $f(u,v)$, $\forall v \in J$.

Derivabilidad respecto de u. Fijamos $v \in I$. El teorema de derivación para funciones definidas por integrales a la función

$$u \longmapsto \int_a^v f(u,x)\ dx$$

(se cumplen las condiciones del enunciado) y nos sale derivable con deriva

$$u \longmapsto \int_a^v \frac{\partial f(u,x)}{\partial u}\ dx\ ,$$

función que es continua por el teorema de continuidad para funciones definidas por integrales.

Hemos probado que la función ϕ es de clase $\mathscr{C}^1$. Observemos ahora que

$$F(t) = \phi)t, g(t)),\ \forall t \in I\ .$$

Por la regla de la cadana F es derivable y su derivada es

$$F'(t) = \frac{\partial\phi(t,g(t))}{\partial u}\ 1 + \frac{\partial\phi(t,g(t))}{\partial v}\ g'(t)\ ,$$

esto es

$$F'(t) = \int_a^{g(t)} \frac{\partial f(t,x)}{\partial t}\ dx + f(t,g(t))\ g'(t)\ .$$

Tema 14

Técnicas de integración en varias variables.

El objetivo de este tema es el estudio de la integral de Lebesgue en $\mathbb{R}^N$, esto es, la integral asociada al espacio de medida $(R^N, \mathscr{M}, \lambda)$. Tras el asentamiento en las lecciones precedentes de los pilares básicos que sostienen la teoría de la integración, podemos ahora abordar la cuestión fundamental para nosotros: proporcionar métodos que permitan reconocer cuando una función $f : E \longrightarrow \mathbb{R}$, donde E es un subconjunto medible de $\mathbb{R}^N$, es integrable y, en caso afirmativo, calcular su integral. Es usual notar $\mathscr{L}^1(E)$ al conjunto de las funciones integrables en E.

Es importante tener presente que el problema de integración planteado contiene dos cuestiones distintas: una es la integrabilidad de la función (hasta ahora sabemos que son integrables las funciones nulas c.p.d., las combinaciones lineales de funciones características de conjuntos de medida finita y las funciones continuas de soporte compacto) y otra es calcular, supuesto que la función sea integrable, el valor de su integral. En este sentido debemos advertir que presentamos tres tipos de técnicas de integración:

- Métodos que proporcionan la integrabilidad de f aunque no el valor de la integral.

- Métodos que proporcionan la integrabilidad de f y el valor de su integral.

- Métodos que proporcionan el valor de la integral, supuesto que la función sea integrable.

El lector tendrá cuidado en lo sucesivo de advertir con precisión la información que cada resultado proporciona al respecto.

Para calcular integrales múltiples no se dispone de ningún procedimiento elemental comparable a la regla de Barrow (Teorema 13.5), pero se dispone de un resultado fundamental que relaciona la integral en $\mathbb{R}^N$ con integraciones sucesivas en espacios de menor dimensión, lo que en última instancia permite evaluar integrales múltiples realizando integraciones iteradas en cada una de las variables. Este resultado es el Teorema de Fubini.

14.1. Teorema de Fubini.

Teorema 14.1 (Fubini)**.** *Sean p y q naturales y $f : \mathbb{R}^p \times \mathbb{R}^q \longrightarrow \mathbb{R}$ una función integrable. Entonces se verifican las siguientes propiedades:*

i) Para casi todo $y \in \mathbb{R}^q$, la función definida en $\mathbb{R}^p$ por

$$x \longmapsto f(x,y)$$

es integrable en $\mathbb{R}^p$ y la función definida casi por doquier en $\mathbb{R}^q$ por

$$y \longmapsto \int_{\mathbb{R}^p} f(x,y)\, dx$$

es integrable en $\mathbb{R}^q$.

ii) Para casi todo $x \in \mathbb{R}^p$ la función definida en $\mathbb{R}^q$ por

$$y \longmapsto f(x,y)$$

es integrable en $\mathbb{R}^q$ y la función definida casi por doquier en $\mathbb{R}^p$ por

$$x \longmapsto \int_{\mathbb{R}^q} (x,y)\, dy$$

es integrable en $\mathbb{R}^p$.

iii)

$$\int_{\mathbb{R}^p \times \mathbb{R}^q} f(x,y)\, d(x,y) = \int_{\mathbb{R}^q} \left[\int_{\mathbb{R}^p} f(x,y)\, dx \right] dy = \int_{\mathbb{R}^p} \left[\int_{\mathbb{R}^q} f(x,y)\, dy \right] dx.$$

Demostración. Designemos por $\mathscr{F}$ el conjunto de todas las funciones $f : \mathbb{R}^{p+q} \to \mathbb{R}$ integrables que satisfacen i), ii) y iii). La demostración consiste en probar que $\mathscr{F} = \mathscr{L}^1(\mathbb{R}^{p+q})$, lo que haremos en varias etapas. Puesto que los papeles de p y q son intercambiables, es suficiente comprobar i) y la primera igualdad de iii).

a) $\mathscr{F}$ es un espacio vectorial: La demostración es consecuencia inmediata de que la unión de dos conjuntos de medida cero es de medida cero y de las propiedades de la integral de Lebesgue.

b) $\mathscr{F}$ es estable por paso al límite de sucesiones monótonas: Si $\{f_n\}$ es una sucesión monótona de funciones de $\mathscr{F}$ que converge puntualmente hacia una función integrable f, entonces $f \in \mathscr{F}$.

En efecto, podemos suponer que $\{f_n\}$ es creciente (en otro caso consideraríamos $\{-f_n\}$). En primer lugar, como para todo $(x,y) \in \mathbb{R}^p \times \mathbb{R}^q$ se verifica que

$$f_n(x,y) \leq f(x,y),$$

la monotonía de la integral nos asegura que

$$\int_{\mathbb{R}^p\times\mathbb{R}^q} f_n(x,y)\, d(x,y) \leq \int_{\mathbb{R}^p\times\mathbb{R}^q} f(x,y)\, d(x,y), \qquad \forall n\in\mathbb{N} \tag{1}$$

y se deduce de la versión práctica del teorema de la convergencia monótona (Corolario 12.2) para funciones de $\mathscr{L}^1(\mathbb{R}^{p+q})$), que

$$\int_{\mathbb{R}^p\times\mathbb{R}^q} f_n(x,y)\, d(x,y) \;\longrightarrow\; \int_{\mathbb{R}^p\times\mathbb{R}^q} f(x,y)\, d(x,y). \tag{2}$$

Utilizando que la unión numerable de conjuntos de medida cero es de medida cero, se deduce la existencia de $Z\subset\mathbb{R}^q$ de medida (q-dimensional) cero tal que para todo $y\in\mathbb{R}^q\backslash Z$ y para todo $n\in\mathbb{N}$, la aplicación

$$x \longmapsto f_n(x,y)$$

es integrable en $\mathbb{R}^p$ y la función definida casi por doquier en $\mathbb{R}^q$ por

$$F_n(y) = \int_{\mathbb{R}^p} f_n(x,y)\, dx$$

es integrable en $\mathbb{R}^q$, siendo además

$$\int_{\mathbb{R}^q} F_n(y)\, dy = \int_{\mathbb{R}^q} \left[\int_{\mathbb{R}^p} f_n(x,y)\, dx\right] dy = \int_{\mathbb{R}^p\times\mathbb{R}^q} f_n(x,y)\, d(x,y).$$

La sucesión $\{F_n\}$ es claramente creciente (podemos definir $F_n(y)=0,\ \forall y\in Z$). La desigualdad (1) nos permite, en consecuencia, aplicar el Corolario 12.2 para funciones de $\mathscr{L}^1(\mathbb{R}^q)$), para deducir que para casi todo $y\in\mathbb{R}^q$, la sucesión $\{F_n(y)\}$ es convergente, luego acotada. Ahora, como para cada $y\in\mathbb{R}^q$ es

$$\big\{f_n(x,y)\big\} \leq f(x,y),$$

de nuevo el Corolario 12.2 para funciones de $\mathscr{L}^1(\mathbb{R}^p)$), nos asegura que para casi todo $y\in\mathbb{R}^q$ la función

$$x \longmapsto f(x,y)$$

es integrable en $\mathbb{R}^p$ y que

$$\int_{\mathbb{R}^p} f_n(x,y)dx \to \int_{\mathbb{R}^p} f(x,y)\, dx.$$

Puesto que por hipótesis

$$\int_{\mathbb{R}^q} \left[\int_{\mathbb{R}^q} f_n(x,y)\, dx\right] dy = \int_{\mathbb{R}^p\times\mathbb{R}^q} f_n(x,y)\, d(x,y), \qquad \forall n\in\mathbb{N} \tag{3}$$

de nuevo la desigualdad (1)nos permite aplicar el Corolario 12.2 para funciones de $\mathscr{L}^1(\mathbb{R}^q)$), para afirmar que la función definida casi por doquier en $\mathbb{R}^q$ por

$$y \longmapsto \int_{\mathbb{R}^p} f(x,y)\, dx$$

es integrable en $\mathbb{R}^q$ y además

$$\int_{\mathbb{R}^q}\left[\int_{\mathbb{R}^p} f_n(x,y)\,dx\right]\,dy \longrightarrow \int_{\mathbb{R}^q}\left[\int_{\mathbb{R}^p} f(x,y)\,dx\right]\,dy.$$

Por último, teniendo ahora en cuenta (2) y (3), concluimos que

$$\int_{\mathbb{R}^p\times\mathbb{R}^q} f(x,y)\,d(x,y) = \int_{\mathbb{R}^q}\left[\int_{\mathbb{R}^p} f(x,y)\,dx\right]\,dy.$$

c) Si $E\subset\mathbb{R}^{p+q}\equiv\mathbb{R}^p\times\mathbb{R}^q$ es medible con $\lambda(E)<\infty$, entonces $\chi_E\in\mathscr{F}$.

La demostración de este apartado la dividiremos en varios subapartados:

c.1) Si I es un intervalo acotado de $\mathbb{R}^p$ y J es un intervalo acotado de $\mathbb{R}^q$, entonces $\chi_{I\times J}\in\mathscr{F}$.
En efecto,

$$\chi_{I\times J}(x,y)=\chi_I(x)\chi_J(y),\quad \forall(x,y)\in\mathbb{R}^p\times\mathbb{R}^q$$

y, consecuentemente, fijado $y\in\mathbb{R}^q$ la función definida en $\mathbb{R}^p$ por

$$x\longrightarrow\chi_{I\times J}(x,y)$$

es la función

$$\left\{\begin{array}{lll}\chi_I & \text{si} & y\in J\\ 0 & \text{si} & y\notin J\end{array}\right.$$

que es integrable con integral

$$\int_{\mathbb{R}^p}\chi_{I\times J}(x,y)\,dx=\left\{\begin{array}{lll}v(I) & \text{si} & y\in J\\ 0 & \text{si} & y\notin J\end{array}\right\}=v(I)\chi_J(y),$$

lo que prueba i).
Probemos la primera igualdad de iii)

$$\int_{\mathbb{R}^p\times\mathbb{R}^q}\chi_{I\times J}(x,y)\,d(x,y)=\lambda(I\times J)=v(I)v(J)=$$

$$=\int_{\mathbb{R}^q}v(I)\chi_J(y)\,dy=\int_{\mathbb{R}^q}\left[\int_{\mathbb{R}^p}\chi_{I\times J}(x,y)\,dx\right]\,dy.$$

c.2) Si $G\subset\mathbb{R}^{p+q}$ es un conjunto abierto con $\lambda(G)<\infty$, entonces $\chi_G\in\mathscr{F}$.
Pongamos $G=\cup_{n=1}^{\infty}I_n$ para conveniente sucesión $\{I_n\}$ de intervalos acotados de $\mathbb{R}^{p+q}$ disjuntos dos a dos (Proposición 10.8). Entonces

$$\chi_G=\sum_{n=1}^{\infty}\chi_{I_n}.$$

Definamos

$$f_n = \sum_{k=1}^{n} \chi_{I_k}, \quad \forall n \in \mathbb{N}.$$

Por los apartados a) y c.1), $f_n \in \mathscr{F}$, $\forall n \in \mathbb{N}$. Por b), concluimos que $\chi_G \in \mathscr{F}$.

c.3) Si $A \subset \mathbb{R}^{p+q}$ es un conjunto G_δ con $\lambda(A) < \infty$, entonces $\chi_A \in \mathscr{F}$.
En efecto, sea $\{G_n\}$ una sucesión decreciente de conjuntos abiertos de $\mathbb{R}^{p+q}$ tal que

$$\lambda(G_n) < \infty, \quad \forall n \in \mathbb{N} \quad \text{y} \quad A = \bigcap_{n=1}^{\infty} G_n.$$

Entonces $\left\{\chi_{G_n}\right\} \searrow \chi_A$. El apartado b) nos permite asegurar que $\chi_A \in \mathscr{F}$.

c.4) Si $Z \subset \mathbb{R}^{p+q}$ es un conjunto de medida cero, entonces $\chi_Z \in \mathscr{F}$.
En efecto, existe una sucesión $\{G_n\}$ de conjuntos abiertos de $\mathbb{R}^{p+q}$ tal que

$$Z \subset G_{n+1} \subset G_n, \ \forall n \in \mathbb{N} \quad \text{y} \quad \{\lambda(G_n)\} \searrow 0.$$

Sea A= $\bigcap_{n=1}^{\infty} G_n$. Este conjunto satisface

$$Z \subset A, \ \lambda(A) = 0 \text{ y } \ \chi_A \in \mathscr{F} \quad \text{(subapartado anterior)}.$$

Se tiene, por tanto, que

$$0 = \lambda(A) = \int_{\mathbb{R}^p \times \mathbb{R}^q} \chi_A(x,y)\, d(x,y) = \int_{\mathbb{R}^p} \left[\int_{\mathbb{R}^p} \chi_A(x,y)\, dx\right] dy,$$

y en consecuencia, en virtud de la Proposición 11.13 iv),

$$\int_{\mathbb{R}^p} \chi_A(x,y)\, dx = 0, \quad (c.p.d.\ y \in \mathbb{R}^q)$$

y, fijado uno de estos y, también

$$\chi_A(x,y) = 0, (c.t.\ x \in \mathbb{R}^p).$$

Al ser

$$0 \le \chi_Z(x,y) \le \chi_A(x,y),$$

concluimos que para casi todo $y \in \mathbb{R}^q$, la aplicación $x \to \chi_Z(x,y)$ es cero c.p.d., y por lo tanto integrable. De nuevo la Proposición 11.13 iv) nos asegura que para casi todo $y \in \mathbb{R}^q$ es

$$\int_{\mathbb{R}^p} \chi_Z(x,y)\, dx = 0,$$

y, por tanto, la aplicación definida casi por doquier en $\mathbb{R}^q$ por

$$y \to \int_{\mathbb{R}^p} \chi_Z(x,y)\, dx$$

es nula c.p.d. y por tanto, integrable. Hemos probado i) para χ_Z. Otra vez la Proposición 11.13 iv) nos dice que

$$\int_{\mathbb{R}^q}\left[\int_{\mathbb{R}^p}\chi_Z(x,y)\,dx\right]\,dy = 0$$

y como

$$0 = \lambda(Z) = \int_{\mathbb{R}^p\times\mathbb{R}^q}\chi_Z(x,y)\,d(x,y)$$

la primera igualdad de iii) queda probada.

Estamos ya en condiciones de verificar el apartado **c)**. Sea $E\subset\mathbb{R}^{p+q}$ un conjunto medible con medida finita. El Teorema 10.16 nos asegura la existencia de un conjunto G_δ, A, y un conjunto de medida cero Z tales que

$$A = E\cup Z \quad \text{y} \quad E\cap Z = \emptyset.$$

Se tiene que

$$\chi_E = \chi_A - \chi_Z.$$

Como $\lambda(A) < \infty$, en virtud de los apartados c.3), c.4) y a), concluimos que $\chi_E\in\mathscr{F}$.

d) Si $f:\mathbb{R}^{p+q}\longrightarrow\mathbb{R}$ es una función integrable, entonces $f\in\mathscr{F}$.

En efecto, en virtud del apartado a), bastará probar que f^+ y f^- están en $\mathscr{F}$. Veamos que $f^+\in\mathscr{F}$ (la comprobación de que $f^-\in\mathscr{F}$ es idéntica). El Teorema de aproximación de Lebesgue 11.8 nos asegura que existe una sucesión creciente $\{s_n\}$ de funciones simples positivas que converge a f^+. Es claro que las funciones s_n son integrables y, en virtud de los apartados a) y c), cada $s_n\in\mathscr{F}$. Por último, el apartado b) nos permite concluir que $f^+\in\mathscr{F}$. ∎

Nota 14.2. i) y ii) se resumen diciendo que la función f tiene integrales iteradas y iii) que ambas son iguales e iguales a la integral de la función f.

Corolario 14.3. *Supongamos que $E\subset\mathbb{R}^{p+q}$ es medible, entonces:*

i) Para cada casi todo $x\in\mathbb{R}^p$, el conjunto

$$E(x) := \{y\in\mathbb{R}^q : (x,y)\in E\}$$

es medible.

ii) Para cada casi todo $y\in\mathbb{R}^q$, el conjunto

$$E(y) := \{x\in\mathbb{R}^p : (x,y)\in E\}$$

es medible.

Demostración. Si $\lambda(E) < \infty$, basta aplicar el Teorema de Fubini a χ_E. En otro caso, la propiedad es cierta para cada conjunto $E_n = E\cap B_\infty(0,n)$, y basta pensar que

$$E = \bigcup_{n=1}^{\infty} E_n.$$

∎

Fórmula 14.4 (**del teorema de Fubini en conjuntos medibles**).

a) **En el plano.** *Si* $E \subset \mathbb{R}^2$ *es un conjunto medible y* $f \in \mathscr{L}^1(E)$ *, entonces*

$$\int_E f(x,y)\ d(x,y) = \int_{\alpha_1}^{\beta_1} \left[\int_{E(x)} f(x,y)\ dy\right]\ dx = \int_{\alpha_2}^{\beta_2} \left[\int_{E(y)} f(x,y)\ dx\right]\ dy,$$

siendo

$$\alpha_1 = \inf E_1,\ \ \beta_1 = \sup E_1,\ \ \alpha_2 = \inf E_2,\ \ \beta_2 = \sup E_2$$

donde

$$E_1 = \pi_1(E), \quad y \quad E_2 = \pi_2(E)$$

(π_1, π_2 *son las proyecciones sobre los ejes coordenados) y para cada* $x \in]\alpha_1, \beta_1[$ *(resp.* $y \in]\alpha_2, \beta_2[$*) es*

$$E(x) = \{y \in \mathbb{R} : (x,y) \in E\} \quad (\textit{resp. } E(y) = \{x \in \mathbb{R} : (x,y) \in E\}).$$

En particular, cuando $E = I \times J$*, siendo* I, J *intervalos de* $\mathbb{R}$*, entonces*

$$\int_E f(x,y)\ d(x,y) = \int_I \left[\int_J f(x,y)\ dy\right]\ dx = \int_J \left[\int_I f(x,y)\ dx\right]\ dy.$$

Demostración. Teniendo en cuenta el teorema de Fubini y el hecho de que

$$\chi_E(x,y) = \chi_{]\alpha_1,\beta_1[}(x)\chi_{E(x)}(y), \ \ ct\ x \in \mathbb{R}, \forall y \in \mathbb{R}$$

se tiene que:

$$\int_E f(x,y)\ d(x,y) = \int_{\mathbb{R}^2} f(x,y)\chi_E(x,y)\ d(x,y) = \int_{\mathbb{R}} \left[\int_{\mathbb{R}} f(x,y)\chi_E(x,y)\ dy\right]\ dx =$$

$$= \int_{\mathbb{R}} \left[\int_{\mathbb{R}} f(x,y)\chi_{]\alpha_1,\beta_1[}(x)\chi_{E(x)}(y)\ dy\right]\ dx = \int_{\mathbb{R}} \left[\int_{\mathbb{R}} f(x,y)\ \chi_{E(x)}(y)\ dy\right] \chi_{]\alpha_1,\beta_1[}(x)\ dx =$$

$$= \int_{]\alpha_1,\beta_1[} \left[\int_{E(x)} f(x,y)\ dy\right]\ dx = \int_{\alpha_1}^{\beta_1} \left[\int_{E(x)} f(x,y)\ dy\right]\ dx.$$

Finalmente, un intercambio de papeles, da la otra igualdad. ∎

Ejemplo: Área del círculo.

Sea $E = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq r^2\}$ *(*$r > 0$*). Calcularemos* $\lambda(E)$*. Se tiene*

$$\lambda(E) = \int_E d(x,y) = \int_{-r}^{r} \left[\int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}}\ dy\right]\ dx = \int_{-r}^{r} 2\sqrt{r^2-x^2}\ dx =$$

$$\left[\text{c. v. } \ x = r \operatorname{sen} t\ \right] = 2r^2 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2 t\ dt \left[\text{utilizar que } \ \cos^2 t = \frac{1+\cos 2t}{2}\right] = \pi r^2.$$

b) En el espacio $\mathbb{R}^3$. *Sea $E \subset \mathbb{R}^3$ un conjunto medible y $f \in \mathscr{L}^1(E)$. La integral de la función f se puede calcular mediante una integración simple y una integración doble o mediante tres integraciones simples (¿Cuántas posibilidades hay? $P_3 + 2V_3^2 = 18$). Por ejemplo, podemos calcular la integral de f como sigue:*

$$\int_E f(x,y,z)\, d(x,y,z) = \int_{\alpha_3}^{\beta_3} \left[\int_{E(z)} f(x,y,z)\, d(x,y) \right] dz,$$

siendo

$$\alpha_3 = \inf E_3, \quad \beta_3 = \sup E_3$$

donde $E_3 = \pi_3(E)$, y para cada $z \in]\alpha_3, \beta_3[$ es

$$E(z) = \{(x,y) \in \mathbb{R}^2 : (x,y,z) \in E\}.$$

Ejemplo: Volumen de la esfera.
Sea $E = \{(x,y,z) \in \mathbb{R}^3 : x^2+y^2+z^2 \leq r^2\}$ ($r > 0$). Calculamos $\lambda(E)$. Se tiene

$$\lambda(E) = \int_E d(x,y,z) = \int_{-r}^{r} \left[\int_{E(z)} d(x,y) \right] dz =$$

$$= \int_{-r}^{r} \left[\int_{\{(x,y)\in\mathbb{R}^2 : x^2+y^2 \leq r^2 - z^2\}} d(x,y) \right] dz =$$

$$= \int_{-r}^{r} \pi(r^2 - z^2)\, dz = \pi \left[r^2 z - \frac{z^3}{3} \right]_{z=-r}^{z=r} = \frac{4}{3}\pi r^3.$$

donde se han usado los cálculos del ejemplo anterior.

Obsérvese que el Teorema de Fubini da una condición necesaria para que una función $f : \mathbb{R}^{p+q} \equiv \mathbb{R}^p \times \mathbb{R}^q \longrightarrow \mathbb{R}$ sea integrable; a saber: ambas integrales iteradas deben existir y ser iguales. Tal condición no es, sin embargo, suficiente.

Ejemplos 14.5.

a) La función $f :]0,1[\times]0,1[\longrightarrow \mathbb{R}$ definida por

$$f(x,y) = \frac{x^2 - y^2}{(x^2+y^2)^2},$$

admite las dos integrales iteradas pero no son iguales y en consecuencia f no es integrable. En efecto, como

$$\frac{\partial}{\partial x}\left(\frac{-x}{x^2+y^2} \right) = \frac{x^2-y^2}{(x^2+y^2)^2}$$

se tiene, utilizando la regla de Barrow que

$$\int_0^1 \frac{x^2-y^2}{(x^2+y^2)^2}\, dx = \frac{-x}{x^2+y^2}\bigg]_{x=0}^{x=1} = \frac{-1}{1+y^2},$$

y, por tanto

$$\int_0^1 \left[\int_0^1 f(x,y)\,dx\right]\,dy = \int_0^1 \frac{-1}{1+y^2}\,dy = -\arctan y\Big]_{y=0}^{y=1} = \frac{-\pi}{4}.$$

Análogamente se prueba que

$$\int_0^1 \left[\int_0^1 f(x,y)\,dy\right]\,dx = \frac{\pi}{4}.$$

b) La función $f:\mathbb{R}\times\mathbb{R}\longrightarrow\mathbb{R}$ definida por

$$f(x,y) = \frac{xy}{(x^2+y^2)^2},$$

admite las dos integrales iteradas y son iguales, pero, sin embargo, no es integrable. En efecto, si lo fuese

$$\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty} f(x,y)\,dx\right]\,dy =$$

$$= \int_{-\infty}^{+\infty}\left[\int_{-\infty}^{0} f(x,y)\,dx + \int_0^{+\infty} f(x,y)\,dx\right]\,dy =$$

$$= \int_{-\infty}^{+\infty}\left[\int_0^{+\infty} (-f(x,y)+f(x,y))\,dx\right]\,dy = \int_{-\infty}^{+\infty} 0\,dy = 0,$$

donde se ha utilizado que $f(-x,y) = -f(x,y),\ \forall (x,y)\in\mathbb{R}$.

Análogamente se prueba que

$$\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty} f(x,y)\,dy\right]\,dx = 0.$$

Veamos que no es integrable. Si fuese f integrable también lo sería $|f|$ con lo que, en virtud del Teorema de Fubini, se tendría que

$$\int_{\mathbb{R}^2} |f(x,y)|\,d(x,y) = \int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty} |f(x,y)|\,dy\right]\,dx,$$

pero

$$\int_{-\infty}^{+\infty} |f(x,y)|\,dx = \int_{-\infty}^{+\infty} \frac{|x|\,|y|}{(x^2+y^2)^2}\,dx =$$

$$|y|\int_0^{+\infty} \frac{2x}{(x^2+y^2)^2}\,dx = |y|\left[\frac{-1}{x^2+y^2}\right]_{x=0}^{x=+\infty} = \frac{1}{|y|},$$

y, en consecuencia,

$$\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty} |f(x,y)|\,dx\right]\,dy = +\infty$$

pues la función de $\mathbb{R}^+$ en $\mathbb{R}$ definida por $y\rightarrow\frac{1}{y}$ no es integrable (véase el Ejemplo 13.9).

■

Obsérvese que se ha usado el hecho de que la existencia de las integrales iteradas de $|f|$ es una condición necesaria para la integrabilidad de f. El siguiente resultado afirma que dicha condición es también suficiente para que una función medible sea integrable.

14.2. Teorema de Tonelli.

Teorema 14.6 (Tonelli)**.** *Sean p y q naturales cualesquiera y sea $f:\mathbb{R}^{p+q}\equiv\mathbb{R}^p\times\mathbb{R}^q\longrightarrow\mathbb{R}$ una función medible y supongamos que es finita alguna de las siguientes integrales iteradas:*

$$\int_{\mathbb{R}^p}\left[\int_{\mathbb{R}^q}|f(x,y)|\,dy\right]\,dx,\quad \int_{\mathbb{R}^q}\left[\int_{\mathbb{R}^p}|f(x,y)|\,dx\right]\,dy.$$

Entonces f es integrable.

Demostración. Para cada $n\in\mathbb{N}$, la función $f_n:\mathbb{R}^{p+q}\longrightarrow\mathbb{R}$ definida por

$$f_n=(|f|\wedge n)\chi_{B(0,n)}$$

es integrable (puesto que está dominada por $n\chi_{B(0,n)}$). El Teorema de Fubini nos dice ahora que

$$\int_{\mathbb{R}^p\times\mathbb{R}^q}f_n(x,y)\,d(x,y)=\int_{\mathbb{R}^q}\left[\int_{\mathbb{R}^p}f_n(x,y)\,dx\right]\,dy=\int_{\mathbb{R}^p}\left[\int_{\mathbb{R}^q}f_n(x,y)\,dy\right]\,dx\ ,$$

y, en consecuencia, las integrales iteradas de f_n están mayoradas por la integral iteradas de $|f|$ que hemos supuesto finita. Así la sucesión

$$\left\{\int f_n\right\}$$

está mayorada. Puesto que la sucesión $\{f_n\}$ es creciente y converge puntualmente a $|f|$, la versión práctica del teorema de la convergencia monótona (Corolario 12.2) nos asegura que la función $|f|$ es integrable y en consecuencia f también lo es. ∎

Combinando los teoremas de Fubini y Tonelli, codificamos el siguiente importante criterio de integrabilidad.

Corolario 14.7 (**Criterio de integrabilidad**)**.** *Sean p y q naturales. Una función medible $f:\mathbb{R}^{p+q}\longrightarrow\mathbb{R}$ es integrable si, y sólo si, alguna de las siguientes integrales iteradas es finita:*

$$\int_{\mathbb{R}^p}\left[\int_{\mathbb{R}^q}|f(x,y)|\,dy\right]\,dx,\quad \int_{\mathbb{R}^q}\left[\int_{\mathbb{R}^p}|f(x,y)|\,dx\right]\,dy.$$

En tal caso ambas coinciden, existen las integrales iteradas de f y son iguales a la integral de f.

Para $f:\mathbb{R}^N\longrightarrow[0,+\infty[$ medible, el conjunto ordenado de f es por definición

$$P_f:=\left\{(x,y)\in\mathbb{R}^{N+1}:0<y<f(x)\right\}.$$

No es difícil comprobar que la función $\phi:\mathbb{R}^{N+1}\longrightarrow\mathbb{R}$ dada por

$$\phi(x,y)=f(x)-y$$

es medible. En efecto, $\Phi = f \circ \pi_1 - \pi_2$ y todo se reduce a probar que $f \circ \pi_1$ es medible. Si $G \subset \mathbb{R}$ es abierto, entonces

$$(f \circ \pi_1)^{-1}(G) = f^{-1}(G) \times \mathbb{R}$$

que es medible (véase Ejercicio 10.14). Finalmente, el conjunto

$$P_f = \left\{(x,y) \in \mathbb{R}^{N+1} : 0 < \phi(x,y)\right\} \cap \left(\mathbb{R}^N \times \mathbb{R}^+\right)$$

es medible.

A continuación haciendo uso del corolario anterior daremos una muy importante interpretación de la integral de una función en términos de la medida de Lebesgue.

Proposición 14.8 (Medida del conjunto ordenado). *Sea $f : \mathbb{R}^N \to [0,+\infty[$ una función medible. Entonces f es integrable si, y sólo si, el conjunto ordenado de f tiene medida finita, en cuyo caso*

$$\int_{\mathbb{R}^N} f(x)\, dx = \lambda\left(P_f\right).$$

Demostración. Como la función $\chi_{P_f} : \mathbb{R}^{N+1} \to [0,+\infty[$ es medible podemos aplicarle el corolario anterior para deducir que χ_{P_f} es integrable (es decir $\lambda\left(P_f\right) < +\infty$) si, y sólo si, la integral iterada

$$\int_{\mathbb{R}^N} \left[\int_{\mathbb{R}} \chi_{P_f}(x,y)\, dy\right] dx$$

es finita. Puesto que

$$\chi_{P_f}(x,y) = \chi_A(x)\chi_{]0,f(x)[}(y)$$

donde

$$A = \left\{x \in \mathbb{R}^N : f(x) > 0\right\},$$

se tiene para cada $x \in \mathbb{R}^N$ que

$$\int_{\mathbb{R}} \chi_{P_f}(x,y)\, dy = \chi_A(x) f(x) = f(x).$$

Concluimos en consecuencia que la finitud de la integral iterada de χ_{P_f} antes considerada equivale a la integrabilidad de f y que en tal caso se tiene:

$$\lambda\left(P_f\right) = \int_{\mathbb{R}^{N+1}} \chi_{P_f}(x,y)\, d(x,y) = \int_{\mathbb{R}^N} f(x)\, dx.$$

■

Ejemplo: área del círculo.

Sea $E = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$. Se trata de calcular $\lambda(E)$. Por la invarianza de λ por isometrías es $\lambda(E) = 2\lambda\left(P_f\right)$, siendo $f :]-1,1[\to \mathbb{R}$ la función definida por

$$f(x) = \sqrt{1-x^2}, \forall x \in]-1,1[.$$

Así

$$\lambda(E) = 2\int_{-1}^{1} \sqrt{1-x^2}\, dx = \pi.$$

■

14.3. Teorema del cambio de variable.

Teorema 14.9 (cambio de variable para funciones medibles positivas)**.**
Sean Ω y G abiertos de $\mathbb{R}^N$, ϕ un difeomorfismo de clase $\mathscr{C}^1$ de Ω sobre G y $f : G \to [0,+\infty[$ una función medible. Entonces

$$\int_{\phi(E)} f(x)\, dx = \int_E f(\phi(t))\, |\det J_\phi(t)|\, dt, \ \forall E \subset \Omega \text{ medible.}$$

Demostración. La Proposición 10.22 nos asegura que ϕ y ϕ^{-1} aplican conjuntos medibles en conjuntos medibles, esto es,

$$E \subset \Omega \text{ es medible} \Leftrightarrow \phi(E) \text{ es medible.}$$

Observemos ahora que la integral del segundo miembro de la fórmula del enunciado tiene sentido por ser $(f \circ \phi)\, |\det J_\phi|$ medible positiva. En efecto, $|\det J_\phi|$ es medible por ser continua. Veamos que $f \circ \phi$ es medible. Si U es un abierto de $[0,+\infty[$, entonces

$$(f \circ \phi)^{-1}(U) = \phi^{-1}\left(f^{-1}(U)\right)$$

que es medible por el comentario anterior. Si la igualdad es cierta, en particular lo será para $f = \mathbf{1}$, es decir

$$\int_{\phi(E)} dx = \int_E |\det J_\phi(t)|\, dt, \ \forall E \subset \Omega \text{ medible,}$$

equivalentemente

$$\lambda(\phi(E)) = \int_E |\det J_\phi(t)|\, dt, \ \forall E \subset \Omega \text{ medible ,}$$

lo que es cierto si ϕ es lineal (ver Proposición10.18) y en otro caso se aproxima por una aplicación lineal.

Probaremos la anterior igualdad en dos etapas:

a) *Fijados $a \in \Omega$ y $\rho > 1$ existe $\delta > 0$ tal que $B(a,\delta) \subset \Omega$ y se verifica que*

$$\rho^{-1} \int_E |\det J_\phi(t)|\, dt \le \lambda(\phi(E)) \le \int_E |\det J_\phi(t)|\, dt, \ \ \forall E \subset B(a,\delta) \text{ medible.}$$

En efecto, si notamos $T = D\phi(a)$, obsérvese que existe $\varepsilon \in]0,1[$ tal que

$$\rho^{-1} \le \frac{\left(1-\varepsilon \left\|T^{-1}\right\|\right)^N}{1+\varepsilon} < \frac{\left(1+\varepsilon \left\|T^{-1}\right\|\right)^N}{1-\varepsilon} \le \rho.$$

Fijemos un tal ε y denotemos

$$\alpha := 1-\varepsilon \left\|T^{-1}\right\| \quad \text{y} \quad \beta := 1+\varepsilon \left\|T^{-1}\right\| .$$

Así la anterior cadena de desigualdades puede escribirse:

$$\rho^{-1} \le \frac{\alpha^N}{1+\varepsilon} < \frac{\beta^N}{1-\varepsilon} \le \rho.$$

Ahora, tomemos $\delta > 0$ tal que $B(a,\delta) \subset \Omega$ y

$$\|x-a\| < \delta \Rightarrow \left\{ \begin{array}{ll} \|D\phi(x) - T\| < \varepsilon & (1) \\ 1-\varepsilon < \dfrac{\det J_\phi(x)}{\det J_\phi(a)} < 1+\varepsilon & (2) \end{array} \right. .$$

del teorema del valor medio (Teorema 4.5) y de (1) se deduce que

$$x,y \in B(a,\delta) \Rightarrow \left\{ \begin{array}{l} \|\phi(x)-\phi(y)\| \le \beta \|T(x)-T(y)\| \\ \alpha\|T(x)-T(y)\| \le \|\phi(x)-\phi(y)\| \end{array} \right. .$$

En efecto, si $x = y$ no hay nada que probar. En otro caso

$$\begin{gathered} \|\phi(x)-\phi(y)\| \le \|\phi(x)-\phi(y)-T(x-y)+T(x-y)\| = \\ \|(\phi(x)-T(x))-(\phi(y)-T(y))\| + \|T(x-y)\| \le \\ \|x-y\| \sup\{\|D\phi(z)-T\| : z \in]x,y[\} + \|T(x-y)\| \le \\ \varepsilon\|x-y\| + \|T(x-y)\| = \varepsilon \left\|T^{-1}(T(x-y))\right\| + \|T(x-y)\| \le \\ \varepsilon \left\|T^{-1}\right\| \; \|T(x-y)\| + \|T(x-y)\| = \beta \|T(x)-T(y)\|, \end{gathered}$$

y así mismo

$$\begin{gathered} \|T(x)-T(y)\| \le \|\phi(x)-\phi(y)-T(x-y)\| + \|\phi(x)-\phi(y)\| = \\ \|(\phi(x)-T(x))-(\phi(y)-T(y))\| + \|\phi(x)-\phi(y)\| \le \\ \|x-y\| \sup\{\|D\phi(z)-T\| : z \in]x,y[\} + \|\phi(x)-\phi(y)\| \le \\ \varepsilon\|x-y\| + \|\phi(x)-\phi(y)\| = \varepsilon\|T^{-1}(T(x-y))\| + \|\phi(x)-\phi(y)\| \le \\ \varepsilon\|T^{-1}\| \; \|T(x-y)\| + \|\phi(x)-\phi(y)\|, \end{gathered}$$

luego

$$\alpha\|T(x)-T(y)\| \le \|\phi(x)-\phi(y)\|.$$

Las desigualdades anteriores se escriben también

$$x,y \in B(a,\delta) \Rightarrow \left\{ \begin{array}{l} \left\|\left(\phi\circ T^{-1}\right)(T(x)) - \left(\phi\circ T^{-1}\right)(T(y))\right\| \le \beta\|T(x)-T(y)\| \\ \left\|\left(T\circ\phi^{-1}\right)(\phi(x)) - \left(T\circ\phi^{-1}\right)(\phi(y))\right\| \le \frac{1}{\alpha}\|\phi(x)-\phi(y)\|. \end{array} \right.$$

Equivalentemente

$$u,v \in T(B(a,\delta)) \Rightarrow \left\|\left(\phi\circ T^{-1}\right)(u) - \left(\phi\circ T^{-1}\right)(v)\right\| \le \beta\|u-v\|$$

$$u,v \in \phi(B(a,\delta)) \Rightarrow \left\|\left(T\circ\phi^{-1}\right)(u) - \left(T\circ\phi^{-1}\right)(v)\right\| \le \frac{1}{\alpha}\|u-v\|.$$

Al verificar $\phi\circ T^{-1}$ y $T\circ\phi^{-1}$ las desigualdades anteriores en los abiertos $T(B(a,\delta))$ y $\phi(B(a,\delta))$ respectivamente, las Proposiciones 10.20 y 10.22, nos aseguran ahora que si $E \subset B(a,\delta)$ es medible, entonces

$$\lambda(\phi(E)) = \lambda\left(\left(\phi\circ T^{-1}\right)(T(E))\right) \le \beta^N \lambda(T(E)) = \beta^N |\det J_\phi(a)| \lambda(E)$$

$$| \det J_\phi(a) | \lambda(E) = \lambda(T(E)) = \lambda\left(\left(T \circ \phi^{-1}\right)(\phi(E))\right) \leq \alpha^{-N} \lambda(\phi(E)).$$

Hemos probado que si $E \subset B(a, \delta)$ es medible, entonces se verifica que

$$\alpha^N \mid \det J_\phi(a) \mid \ \lambda(E) \leq \lambda(\phi(E)) \leq \beta^N \mid \det J_\phi(a) \mid \ \lambda(E). \qquad (*)$$

Por otra parte (2) se escribe también

$$(1-\varepsilon) \mid \det J_\phi(a) \mid < \mid \det J_\phi(t) \mid < (1+\varepsilon) \mid \det J_\phi(a) \mid, \ \forall t \in B(a, \delta)\ ,$$

y en consecuencia para cada $E \subset B(a, \delta)$ medible se verifica también

$$\int_E (1-\varepsilon) \mid \det J_\phi(a) \mid \ dt \leq \int_E \mid \det J_\phi(t) \mid \ dt \leq \int_E (1+\varepsilon) \mid \det J_\phi(a) \mid \ dt\ ,$$

o equivalentemente

$$(1-\varepsilon) \mid \det J_\phi(a) \mid \lambda(E) \leq \int_E \mid \det J_\phi(t) \mid \ dt \leq (1+\varepsilon) \mid \det J_\phi(a) \mid \ \lambda(E)\ . \qquad (**)$$

La conjunción de (*) y (**) prueba el apartado **a)**.

b) *Para todo conjunto medible E incluido en Ω se verifica:*

$$\lambda(\phi(E)) = \int_E \mid \det J_\phi(t) \mid \ dt.$$

Fijemos $\rho > 1$. Comprobemos que Ω se puede expresar como unión numerable de bolas abiertas, en cada una de las cuales se cumplen las desigualdades del apartado **a)**. En efecto, para cada $x \in \Omega$ existe $\delta(x) > 0$ tal que se cumplen las condiciones del apartado a). Entonces, el recubrimiento

$$\mathscr{U} = \{B(x, , \delta(x)) : \ x \in \Omega\}$$

admite, en virtud del Lema 10.21 , un subrecubrimiento numerable.

Sea $E \subset \Omega$ medible. Es inmediato que E se puede expresar como una unión numerable de conjuntos medibles disjuntos entre sí

$$E = \bigcup_{n=1}^{\infty} E_n \ \text{ con } \ E_n \subset B(b_n, r_n), \quad \forall n \in \mathbb{N}.$$

El apartado **a)** nos garantiza que para cada natural n se tiene

$$\rho^{-1} \int_{E_n} \mid \det J_\phi(t) \mid \ dt \leq \lambda(\phi(E_n)) \leq \rho \int_{E_n} \mid \det J_\phi(t) \mid \ dt.$$

Como la aplicación

$$E \to \int_E \mid \det J_\phi(t) \mid \ dt$$

es una medida sobre la σ-álgebra $\mathscr{M}(\Omega)$ (Corolario 12.4), la suma de las anteriores desigualdades proporciona la siguiente:

$$\rho^{-1}\int_E |\det J_\phi(t)|\ dt \le \lambda(\phi(E)) \le \rho\int_E |\det J_\phi(t)|\ dt.$$

Por último de la arbitrariedad de ρ se deduce el apartado **b)**.

c) *Para toda función simple positiva s definida en G se verifica que*

$$\int_{\phi(E)} s(x)\ dx = \int_E s(\phi(t))\,|\det J_\phi(t)|\ dt,\ \forall E\subset\Omega \text{ medible.}$$

Por linealidad de la integral basta probar que si $E\subset\Omega$ es medible, entonces

$$\int_{\phi(E)} \chi_A(x)dx = \int_E \chi_A(\phi(t))\,|\det J_\phi(t)|\ dt,\ \forall A\subset G \text{ medible},$$

o lo que es lo mismo

$$\int_{\phi(E)\cap A} dx = \int_{E\cap\phi^{-1}(A)} |\det J_\phi(t)|\ dt,\ \forall A\subset G \text{ medible}$$

es decir

$$\lambda(\phi(E)\cap A) = \int_{E\cap\phi^{-1}(A)} |\det J_\phi(t)|\ dt,\ \forall A\subset G \text{ medible}$$

lo que se deduce fácilmente de **b)** por ser ϕ una aplicación biyectiva.

d) *Para toda función medible positiva f definida en G se verifica que*

$$\int_{\phi(E)} f(x)\ dx = \int_E f(\phi(t))\,|\det J_\phi(t)|\ dt,\ \forall E\subset\Omega \text{ medible.}$$

El teorema de aproximación de Lebesgue (Teorema 11.8) nos asegura que existe una sucesión creciente $\{s_n\}$ de funciones simples positivas que converge a f y basta aplicar **c)** y el teorema de convergencia creciente para funciones medibles positivas (Teorema 12.3). ∎

Teorema 14.10 (cambio de variable). *Sean Ω y G abiertos de $\mathbb{R}^N$, ϕ un difeomorfismo de clase $\mathscr{C}^1$ de Ω sobre G, $E\subset\Omega$ medible y $f:\phi(E)\to\mathbb{R}$ una función medible. Entonces*

$$f\in\mathscr{L}^1(\phi(E)) \Leftrightarrow (f\circ\phi)\,|\det J_\phi| \in\mathscr{L}^1(E),$$

en cuyo caso

$$\int_{\phi(E)} f(x)\ dx = \int_E f(\phi(t))\,|\det J_\phi(t)|\ dt.$$

Demostración. Por definición, f es integrable en $\phi(E)$ si $\int_{\phi(E)} |f(x)| \; dx < \infty$, lo que es equivalente, en virtud del teorema anterior, a que

$$\int_E |f|(\phi(t)) \,|\det J_\phi(t)| \; dt = \int_E |(f(\phi(t))\,| |\det J_\phi(t)| \, dt < \infty,$$

es decir $(f \circ \phi) \,|\det J_\phi\,|$ es integrable en E.

En el caso de que f sea integrable en $\phi(E)$ aplicamos nuevamente el teorema anterior para obtener

$$\int_{\phi(E)} f(x)\, dx = \int_{\phi(E)} f^+(x)\, dx - \int_{\phi(E)} f^-(x)\, dx =$$

$$\int_E f^+(\phi(t))\,|\det J_\phi(t)| \; dt - \int_E f^-(\phi(t))\,|\det J_\phi(t)| \; dt =$$

$$\int_E \left[f(\phi(t))\,|\det J_\phi(t)\,|\right]^+ \; dt - \int_E \left[f(\phi(t))\,|\det J_\phi(t)\,|\right]^- \; dt =$$

$$\int_E f(\phi(t))\,|\det J_\phi(t)| \; dt.$$

∎

Corolario 14.11. (Fórmula de cambio de variable). *En las condiciones del teorema del cambio de variable (Teorema 14.10), obsérvese que si además $\lambda\left(G^C\right) = 0$, entonces: Para cualesquiera $E \subset \mathbb{R}^N$ medible y $f : E \to \mathbb{R}$ medible se verifica*

$$f \in \mathscr{L}^1(E) \Leftrightarrow (f \circ \phi)|\det J_\phi| \in \mathscr{L}^1\left(\phi^{-1}(E)\right).$$

en cuyo caso se tiene la siguiente fórmula del cambio de variable

$$\int_E f(x)\, dx = \int_{\phi^{-1}(E)} f(\phi(t))|\det J_\phi(t)|\, dt \;.$$

Demostración. En efecto, como $\lambda(E \backslash G) = 0$, es claro que

$$f \in \mathscr{L}^1(E) \Leftrightarrow f \in \mathscr{L}^1(E \cap G)$$

en cuyo caso

$$\int_E f(x)\, dx = \int_{E\cap G} f(x)\, dx.$$

El teorema del cambio de variable (Teorema 14.10) nos asegura ahora que

$$f \in \mathscr{L}^1(E \cap G) \Leftrightarrow (f \circ \phi)\,|\det J_\phi\,| \in \mathscr{L}^1\left(\phi^{-1}(E \cap G)\right),$$

en cuyo caso

$$\int_{E\cap G} f(x)\, dx = \int_{\phi^{-1}(E\cap G)} f(\phi(t))\,|\det J_\phi(t)| \; dt.$$

Puesto que $\phi^{-1}(E \cap G) = \phi^{-1}(E)$ se sigue de ambos resultados que

$$f \in \mathscr{L}^1(E) \Leftrightarrow (f \circ \phi)\,|\det J_\phi\,| \in \mathscr{L}^1\left(\phi^{-1}(E)\right),$$

en cuyo caso

$$\int_E f(x)\,dx = \int_{\phi^{-1}(E)} f(\phi(t))\,|\det J_\phi(t)|\ \,dt,$$

lo que prueba la fórmula del cambio de variable que es la manera habitual de utilizar el Teorema 14.10 para calcular integrales. ∎

Nota 14.12. Para reconocer que ϕ es un difeomorfismo de clase $\mathscr{C}^1$ de Ω sobre $\phi(\Omega)$ es usual utilizar el teorema "global" de la función inversa (Corolario 8.8). En la práctica el conjunto de integración $\phi(E)$ es conocido y el problema radica en encontrar un cambio de variable adecuado y en conocer E, esto es, el conjunto de $\mathbb{R}^N$ que se aplica mediante el difeomorfismo ϕ en el conjunto de partida. Es usual también, una vez efectuado el cambio de variable, utilizar el teorema de Fubini (Teorema 14.1) para calcular la integral.

14.4. Coordenadas polares, cilíndricas y esféricas.

a) **Coordenadas polares en el plano.**

Dadas por el difeomorfismo ϕ de clase $\mathscr{C}^\infty$, del abierto $\Omega = \mathbb{R}^+ \times]-\pi,\pi[$ sobre el abierto $G = \mathbb{R}^2 \backslash \{(x,0) : x \leq 0\}$, definido por

$$\phi(\rho,\vartheta) = (\rho\cos\vartheta, \rho\,\mathrm{sen}\,\vartheta)$$

con

$$\det J_\phi(\rho,\vartheta) = \rho > 0, \forall(\rho,\vartheta) \in \Omega.$$

La fórmula del cambio de variable deviene para $E \subset \mathbb{R}^2$ medible y $f : E \to \mathbb{R}$ medible en

$$f \in \mathscr{L}^1(E) \Leftrightarrow \rho f(\rho\cos\vartheta, \rho\,\mathrm{sen}\,\vartheta) \in \mathscr{L}^1\left(\phi^{-1}(E)\right),$$

en cuyo caso

$$\int_E f(x,y)\,d(x,y) = \int_{\phi^{-1}(E)} f(\rho\cos\vartheta, \rho\,\mathrm{sen}\,\vartheta)\rho\,d(\rho,\vartheta).$$

Ejemplo: área del círculo.

Sea $E = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq r^2\}$ con $r > 0$. Se trata de calcular $\lambda(E)$. Se tiene

$$\lambda(E) = \int_E d(x,y) = \int_{]0,r[\times]-\pi,\pi[} \rho\,d(\rho,\vartheta) = \int_0^r \left[\int_{-\pi}^{\pi} \rho\,d\vartheta\right] d\rho = \pi r^2,$$

donde se han utilizado los teoremas de Fubini, del cambio de variable y la regla de Barrow.

b) **Coordenadas cilíndricas en el espacio.**

Dadas por el difeomorfismo ϕ de clase $\mathscr{C}^\infty$, del abierto $\Omega = \mathbb{R}^+ \times]-\pi,\pi[\times\mathbb{R}$ sobre el abierto $G = \mathbb{R}^3 \backslash \{(x,0,z) : x \leq 0\}$, definido por

$$\phi(\rho,\vartheta,z) = (\rho\cos\vartheta, \rho\operatorname{sen}\vartheta, z)$$

con

$$\det J_\phi(\rho,\vartheta,z) = \rho > 0, \forall(\rho,\vartheta,z) \in \Omega.$$

La fórmula del cambio de variable viene dada ahora para $E \subset \mathbb{R}^3$ medible y $f : E \longrightarrow \mathbb{R}$ medible, por

$$f \in \mathscr{L}^1(E) \Leftrightarrow \rho f(\rho\cos\vartheta, \rho\operatorname{sen}\vartheta, z) \in \mathscr{L}^1\left(\phi^{-1}(E)\right),$$

en cuyo caso

$$\int_E f(x,y,z)\, d(x,y,z) = \int_{\phi^{-1}(E)} f(\rho\cos\vartheta, \rho\operatorname{sen}\vartheta, z)\rho\, d(\rho,\vartheta,z)$$

Ejemplo: volumen del cilindro.

Sea $E = \{(x,y,z) \in \mathbb{R}^3 : x^2+y^2 \leq r^2, 0 \leq z \leq h\}$ con $r,h > 0$. Se trata de calcular $\lambda(E)$. Se tiene

$$\lambda(E) = \int_E d(x,y,z) = \int_{]0,r[\times]-\pi,\pi[\times]0,h[} \rho\, d(\rho,\vartheta,z) =$$

$$\int_0^r \left[\int_{-\pi}^{\pi} \left[\int_0^h \rho\, dz\right] d\vartheta\right] d\rho = \pi r^2 h.$$

c) **Coordenadas esféricas en el espacio.**

Dadas por el difeomorfismo ϕ de clase $\mathscr{C}^\infty$, del abierto $\Omega = \mathbb{R}^+ \times]-\pi,\pi[\times\left]\frac{-\pi}{2},\frac{\pi}{2}\right[$ sobre el abierto $G = \mathbb{R}^3 \backslash \{(x,0,z) : x \leq 0\}$, definido por

$$\phi(\rho,\vartheta,\varphi) = (\rho\cos\vartheta\cos\varphi, \rho\operatorname{sen}\vartheta\cos\varphi, \rho\operatorname{sen}\varphi)$$

con

$$\det J_\phi(\rho,\vartheta,\varphi) = \rho^2\cos\varphi > 0, \forall(\rho,\vartheta,\varphi) \in \Omega.$$

La fórmula del cambio de variable viene dada ahora para $E \subset \mathbb{R}^3$ medible y $f : E \longrightarrow \mathbb{R}$ medible por

$$f \in \mathscr{L}^1(E) \Leftrightarrow \rho^2\cos\varphi f(\rho\cos\vartheta\cos\varphi, \rho\operatorname{sen}\vartheta\cos\varphi, \rho\operatorname{sen}\varphi) \in \mathscr{L}^1\left(\phi^{-1}(E)\right),$$

en cuyo caso

$$\int_E f(x,y,z)d(x,y,z) = \int_{\phi^{-1}(E)} f(\rho\cos\vartheta\cos\varphi, \rho\operatorname{sen}\vartheta\cos\varphi, \rho\operatorname{sen}\varphi)\rho^2\cos\varphi d(\rho,\vartheta,\varphi).$$

Ejemplo: volumen de la esfera.

Sea $E=\{(x,y,z)\in\mathbb{R}^3 : x^2+y^2+z^2\leq r^2\}$ con $r>0$. Se trata de calcular $\lambda(E)$. Se tiene

$$\lambda(E)=\int_E d(x,y,z)=\int_{]0,r[\times]-\pi,\pi[\times]\frac{-\pi}{2},\frac{\pi}{2}[}\rho^2\cos\varphi\, d(\rho,\vartheta,\varphi)=$$

$$\int_0^r\left[\int_{-\pi}^{\pi}\left[\int_{\frac{-\pi}{2}}^{\frac{\pi}{2}}\rho^2\cos\varphi d\varphi\right]d\vartheta\right]d\rho=\int_0^r\left[\int_{-\pi}^{\pi}2\rho^2 d\vartheta\right]d\rho=\int_0^r 4\pi\rho^2 d\rho=\frac{4}{3}\pi r^3.$$

14.5. Resumen de resultados del Tema 14.

Teorema de Fubini *Sean p y q naturales y $f : \mathbb{R}^p \times \mathbb{R}^q \longrightarrow \mathbb{R}$ una función integrable. Entonces se verifican las siguientes propiedades:*

i) Para casi todo $y \in \mathbb{R}^q$, la función definida en $\mathbb{R}^p$ por

$$x \to f(x,y)$$

es integrable en $\mathbb{R}^p$ y la función definida casi por doquier en $\mathbb{R}^q$ por

$$y \longrightarrow \int_{\mathbb{R}^p} f(x,y)\, dx$$

es integrable en $\mathbb{R}^q$.

ii) Para casi todo $x \in \mathbb{R}^p$ la función definida en $\mathbb{R}^q$ por

$$y \longrightarrow f(x,y)$$

es integrable en $\mathbb{R}^q$ y la función definida casi por doquier en $\mathbb{R}^p$ por

$$x \to \int_{\mathbb{R}^q} (x,y)\, dy$$

es integrable en $\mathbb{R}^p$.

iii)

$$\int_{\mathbb{R}^p \times \mathbb{R}^q} f(x,y)\, d(x,y) = \int_{\mathbb{R}^q} \left[\int_{\mathbb{R}^p} f(x,y)\, dx \right] dy = \int_{\mathbb{R}^p} \left[\int_{\mathbb{R}^q} f(x,y)\, dy \right] dx.$$

Fórmula del teorema de Fubini en conjuntos medibles

a) En el plano. *Si $E \subset \mathbb{R}^2$ es un conjunto medible y $f \in \mathscr{L}^1(E)$, entonces*

$$\int_E f(x,y)\, d(x,y) = \int_{\alpha_1}^{\beta_1} \left[\int_{E(x)} f(x,y)\, dy \right] dx = \int_{\alpha_2}^{\beta_2} \left[\int_{E(y)} f(x,y)\, dx \right] dy,$$

siendo

$$\alpha_1 = \inf E_1, \;\; \beta_1 = \sup E_1, \;\; \alpha_2 = \inf E_2, \;\; \beta_2 = \sup E_2$$

donde

$$E_1 = \pi_1(E), \quad y \quad E_2 = \pi_2(E)$$

(π_1, π_2 son las proyecciones sobre los ejes coordenados) y para cada $x \in]\alpha_1, \beta_1[$ (resp. $y \in]\alpha_2, \beta_2[$) es

$$E(x) = \{y \in \mathbb{R} : (x,y) \in E\} \quad (\text{resp. } E(y) = \{x \in \mathbb{R} : (x,y) \in E\}).$$

En particular, cuando $E = I \times J$, siendo I, J intervalos de $\mathbb{R}$, entonces

$$\int_E f(x,y)\, d(x,y) = \int_I \left[\int_J f(x,y)\, dy \right] dx = \int_J \left[\int_I f(x,y)\, dx \right] dy.$$

b) En el espacio $\mathbb{R}^3$. *Sea $E \subset \mathbb{R}^3$ un conjunto medible y $f \in \mathscr{L}^1(E)$. La integral de la función f se puede calcular mediante una integración simple y una integración doble o mediante tres integraciones simples (¿Cuántas posibilidades hay?). Por ejemplo, podemos calcular la integral de f como sigue:*

$$\int_E f(x,y,z)\, d(x,y,z) = \int_{\alpha_3}^{\beta_3} \left[\int_{E(z)} f(x,y,z)\, d(x,y) \right] dz,$$

siendo

$$\alpha_3 = \inf E_3, \quad \beta_3 = \sup E_3$$

donde $E_3 = \pi_3(E)$, y para cada $z \in]\alpha_3, \beta_3[$ es

$$E(z) = \{(x,y) \in \mathbb{R}^2 : (x,y,z) \in E\}.$$

Teorema de Tonelli *Sean p y q naturales cualesquiera y sea $f : \mathbb{R}^{p+q} \equiv \mathbb{R}^p \times \mathbb{R}^q \longrightarrow \mathbb{R}$ una función medible y supongamos que es finita alguna de las siguientes integrales iteradas:*

$$\int_{\mathbb{R}^p} \left[\int_{\mathbb{R}^q} |f(x,y)|\, dy \right] dx, \quad \int_{\mathbb{R}^q} \left[\int_{\mathbb{R}^p} |f(x,y)|\, dx \right] dy.$$

Entonces f es integrable.

Criterio de integrabilidad *Sean p y q naturales. Una función medible $f : \mathbb{R}^{p+q} \longrightarrow \mathbb{R}$ es integrable si, y sólo si, alguna de las siguientes integrales iteradas es finita:*

$$\int_{\mathbb{R}^p} \left[\int_{\mathbb{R}^q} |f(x,y)|\, dy \right] dx, \quad \int_{\mathbb{R}^q} \left[\int_{\mathbb{R}^p} |f(x,y)|\, dx \right] dy.$$

En tal caso existen las dos y coinciden. Entonces también existen las integrales iteradas de f y son iguales a la integral de f.

Medida del conjunto ordenado *Sea $f : \mathbb{R}^N \to [0, +\infty[$ una función medible. Entonces f es integrable si, y sólo si, el conjunto ordenado de f*

$$P_f := \left\{ (x,y) \in \mathbb{R}^{N+1} : 0 < y < f(x) \right\}$$

tiene medida finita, en cuyo caso

$$\int_{\mathbb{R}^N} f(x)\, dx = \lambda\left(P_f\right).$$

Teorema de cambio de variable para funciones medibles positivas *Sean Ω y G abiertos de $\mathbb{R}^N$, ϕ un difeomorfismo de clase $\mathscr{C}^1$ de Ω sobre G y $f : G \to [0, +\infty[$ una función medible. Entonces*

$$\int_{\phi(E)} f(x)\, dx = \int_E f(\phi(t))\, |\det J_\phi(t)|\, dt, \ \forall E \subset \Omega \text{ medible.}$$

Teorema de cambio de variable *Sean Ω y G abiertos de $\mathbb{R}^N$, ϕ un difeomorfismo de clase $\mathscr{C}^1$ de Ω sobre G, $E \subset \Omega$ medible y $f : \phi(E) \to \mathbb{R}$ una función medible. Entonces*

$$f \in \mathscr{L}^1(\phi(E)) \Leftrightarrow (f \circ \phi) \mid \det J_\phi \mid \in \mathscr{L}^1(E),$$

en cuyo caso

$$\int_{\phi(E)} f(x)\, dx = \int_E f(\phi(t)) \mid \det J_\phi(t) \mid\, dt.$$

Fórmula de cambio de variable *En las condiciones del teorema del cambio de variable (Teorema 14.10), obsérvese que si además $\lambda\left(G^C\right) = 0$, entonces: Para cualesquiera $E \subset \mathbb{R}^N$ medible y $f : E \to \mathbb{R}$ medible se verifica*

$$f \in \mathscr{L}^1(E) \Leftrightarrow (f \circ \phi)|\det J_\phi| \in \mathscr{L}^1\left(\phi^{-1}(E)\right).$$

en cuyo caso se tiene la siguiente fórmula de cambio de variable

$$\int_E f(x)\, dx = \int_{\phi^{-1}(E)} f(\phi(t))|\det J_\phi(t)|\, dt\ .$$

Coordenadas polares en el plano. *Dadas por el difeomorfismo ϕ de clase $\mathscr{C}^\infty$ del abierto $\Omega = \mathbb{R}^+ \times]-\pi, \pi[$ sobre el abierto $G = \mathbb{R}^2 \backslash \{(x,0) : x \leq 0\}$, definido por*

$$\phi(\rho, \vartheta) = (\rho \cos \vartheta, \rho \operatorname{sen} \vartheta)$$

con

$$\det J_\phi(\rho, \vartheta) = \rho > 0, \forall (\rho, \vartheta) \in \Omega.$$

La fórmula del cambio de variable deviene para $E \subset \mathbb{R}^2$ medible y $f : E \to \mathbb{R}$ medible en

$$f \in \mathscr{L}^1(E) \Leftrightarrow \rho f(\rho \cos \vartheta, \rho \operatorname{sen} \vartheta) \in \mathscr{L}^1\left(\phi^{-1}(E)\right),$$

en cuyo caso

$$\int_E f(x,y)\, d(x,y) = \int_{\phi^{-1}(E)} f(\rho \cos \vartheta, \rho \operatorname{sen} \vartheta) \rho\, d(\rho, \vartheta).$$

Coordenadas cilíndricas en el espacio. *Dadas por el difeomorfismo ϕ de clase $\mathscr{C}^\infty$ del abierto $\Omega = \mathbb{R}^+ \times]-\pi, \pi[\times \mathbb{R}$ sobre el abierto $G = \mathbb{R}^3 \backslash \{(x,0,z) : x \leq 0\}$, definido por*

$$\phi(\rho, \vartheta, z) = (\rho \cos \vartheta, \rho \operatorname{sen} \vartheta, z)$$

con

$$\det J_\phi(\rho, \vartheta, z) = \rho > 0, \forall (\rho, \vartheta, z) \in \Omega.$$

La fórmula de cambio de variable viene dada ahora para $E \subset \mathbb{R}^3$ medible y $f : E \to \mathbb{R}$ medible, por

$$f \in \mathscr{L}^1(E) \Leftrightarrow \rho f(\rho \cos \vartheta, \rho \operatorname{sen} \vartheta, z) \in \mathscr{L}^1\left(\phi^{-1}(E)\right),$$

en cuyo caso

$$\int_E f(x,y,z)\, d(x,y,z) = \int_{\phi^{-1}(E)} f(\rho\cos\vartheta, \rho\,\mathrm{sen}\,\vartheta, z)\rho\, d(\rho,\vartheta,z)$$

Coordenadas esféricas en el espacio. *Dadas por el difeomorfismo ϕ de clase $\mathscr{C}^\infty$ del abierto $\Omega = \mathbb{R}^+ \times]-\pi,\pi[\times]\frac{-\pi}{2},\frac{\pi}{2}[$ sobre el abierto $G = \mathbb{R}^3 \backslash \{(x,0,z) : x \leq 0\}$, definido por*

$$\phi(\rho,\vartheta,\varphi) = (\rho\cos\vartheta\cos\varphi, \rho\,\mathrm{sen}\,\vartheta\cos\varphi, \rho\,\mathrm{sen}\,\varphi)$$

con

$$\det J_\phi(\rho,\vartheta,\varphi) = \rho^2\cos\varphi > 0, \forall(\rho,\vartheta,\varphi) \in \Omega.$$

La fórmula del cambio de variable viene dada ahora para $E \subset \mathbb{R}^3$ medible y $f : E \to \mathbb{R}$ medible por

$$f \in \mathscr{L}^1(E) \Leftrightarrow \rho^2\cos\varphi f(\rho\cos\vartheta\cos\varphi, \rho\,\mathrm{sen}\,\vartheta\cos\varphi, \rho\,\mathrm{sen}\,\varphi) \in \mathscr{L}^1\left(\phi^{-1}(E)\right),$$

en cuyo caso

$$\int_E f(x,y,z) d(x,y,z) = \int_{\phi^{-1}(E)} f(\rho\cos\vartheta\cos\varphi, \rho\,\mathrm{sen}\,\vartheta\cos\varphi, \rho\,\mathrm{sen}\,\varphi)\rho^2\cos\varphi d(\rho,\vartheta,\varphi).$$

14.6. Ejercicios del Tema 14.

14.1 Estudiar la integrabilidad de las siguientes funciones:

i)
$$f(x,y) = e^{-xy} - 2e^{-2xy} \text{ en }]0,1[\times]1,+\infty[$$

ii)
$$f(x,y) = (x-y)e^{-(x-y)^2} \text{ en } \mathbb{R}^+ \times \mathbb{R}^+.$$

iii)
$$f(x,y) = \operatorname{sen}\left(\frac{1}{x-y}\right) \text{ en } \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}.$$

iv)
$$f(x,y) = e^{-(x+y)} \text{ en } \{(x,y) \in \mathbb{R}^2 : 0 < x, \ 0 < y < x\}.$$

v)
$$f(x,y) = e^{-xy} \text{ en } \left\{(x,y) \in \mathbb{R}^2 : 1 < x, \ 0 < y < \frac{1}{x}\right\}.$$

vi)
$$f(x,y) = e^{-|x-y|} \text{ en }]-1,1[\times\mathbb{R}.$$

vii)
$$f(x,y) = \frac{x}{(x^2+y^2)\sqrt{1-x}} \text{ en } \{(x,y) \in \mathbb{R}^2 : 0 < x < 1, \ 0 < y < x\}.$$

viii)
$$f(x,y) = \frac{\cos(xy)}{(1+y^2)\sqrt{\operatorname{sen} x}} \text{ en }]0,1[\times\mathbb{R}^+.$$

ix)
$$f(x,y,z) = \frac{1}{(x^2+y^2+z^2)^\alpha} \text{ en } \mathbb{R}^3 \ (\alpha \in \mathbb{R}).$$

14.2 Estudiar la integrabilidad de las siguientes funciones:

i) $(x,y) \to \frac{xy}{(x^2+y^2)^\alpha}$ ($\alpha > 0$ es un parámetro).

ii) $1 < \alpha < 2 \Longrightarrow (x,y) \to \frac{\operatorname{sen} x \ \operatorname{sen} y}{(x^2+y^2)^\alpha}$.

14.3 Sean $x_1,\dots,x_N,a\in\mathbb{R}^N$ y sea

$$T:=\left\{x\in\mathbb{R}^N:\ x=a+\sum_{k=1}^{N}t_kx_k,\ \sum_{k=1}^{N}t_k\le 1,\ 0\le t_k,\ 1\le k\le N\right\}$$

el $(N+1)$-edro en $\mathbb{R}^N$ (generalización del triángulo en $\mathbb{R}^2$ y del tetraedro en $\mathbb{R}^3$). Probar que

$$\lambda(T)=\frac{1}{n!}\ |\det\,(x_1,\dots,x_N)|.$$

Indicación: Tomar como variables las t_k y proceder por inducción sobre N, para calcular la medida del conjunto

$$\left\{t\in\mathbb{R}^N:\ \sum_{k=1}^{N}t_k\le 1,\ 0\le t_k,\ 1\le k\le N\right\}.$$

14.4 Calcular el volumen de los siguientes conjuntos:

i) Bóveda de Viviani:

$$E=\left\{(x,y,z)\in\mathbb{R}^3:\left(x-\frac{1}{2}\right)^2+y^2\le\frac{1}{4}\right\}\bigcap$$

$$\left\{(x,y,z)\in\mathbb{R}^3:\ x^2+y^2+z^2\le 1,\ 0\le z\right\}.$$

ii)

$$E=\left\{(x,y,z)\in\mathbb{R}^+\times\mathbb{R}^+\times\mathbb{R}^+:\ x^{\frac{2}{3}}+y^{\frac{2}{3}}+z^{\frac{2}{3}}\le 1\right\}.$$

Indicación: Hacer uso de la transformación

$$\begin{cases}x=\rho\ \cos^3\vartheta\\ y=\rho\ \operatorname{sen}^3\vartheta\end{cases}.$$

iii)

$$E=\left\{(x,y,z)\in\mathbb{R}^3:\ x^2+y^2\le(1-z)^2,\ x^2+y^2\le\frac{z}{2},\ z\le 1\right\}.$$

Indicación: E es la intersección de un cono y un paraboloide (un cucurucho de helado invertido).

iv)

$$E=\left\{(x,y,z)\in\mathbb{R}^3:\ x^2+y^2+z^2\le 3,\ x^2+y^2-z^2\le 1\right\}.$$

Indicación: E es la intersección de un hiperboloide (diábolo) y una esfera.

14.5 **Área de la cardioide.**
La curva en $\mathbb{R}^2$, cuya ecuación en coordenadas polares viene dada por

$$\rho = 2a(1+\cos\vartheta) \quad (a \in \mathbb{R}^+,\ -\pi < \vartheta \le \pi)$$

se llama una <u>*cardioide*</u>. Sea

$$E = \{(\rho\ \cos\vartheta, \rho\ \operatorname{sen}\vartheta) :\ -\pi < \vartheta \le \pi,\ 0 \le \rho \le 2a(1+\cos\vartheta)\}.$$

Calcular $\lambda(E)$.

14.6 **Sólidos de revolución.**
Sea $E \subset \mathbb{R}_0^+ \times \mathbb{R}$ un conjunto medible. Consideremos el conjunto

$$S = \left\{(x,y,z) \in R^3 :\ \left(\sqrt{x^2+y^2}, z\right) \in E\right\}$$

(sólido de revolución obtenido al girar el conjunto E contenido en el plano XZ en torno al eje OZ). Probar que S es medible y que

$$\lambda(S) = 2\pi \int_E x\, d(x,z).$$

Aplicación: probar que

a) El volumen del toro engendrado por la rotación en torno al eje OZ del círculo $(x-1)^2 + (z-1)^2 \le \frac{1}{4}$ (salvavidas) es $\frac{\pi^2}{2}$.

b) El conjunto

$$E = \left\{(x,z) \in \mathbb{R}^2 :\ 1 < x,\ 0 < z < \frac{1}{x^2}\right\}$$

tiene área uno y que el sólido engendrado al rotar dicho conjunto en torno al eje OZ tiene volumen infinito.

14.7 **Integral de Gauss.**
Probar que

$$\int_0^{+\infty} e^{-ax^2}\, dx = \frac{1}{2}\sqrt{\frac{\pi}{a}} \quad (a > 0).$$

<u>Indicación:</u> Probar, utilizando el Teorema del cambio de variable, que la función $f : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}$ definida por

$$f(x,y) = e^{-a(x^2+y^2)} \quad (x,y \in \mathbb{R}^+)$$

es integrable y calcular su integral.

14.8 Probar que la función
$$f(x,y)=\frac{1}{1+x^2+y^2}$$
es integrable en $]0,+\infty[\times]0,1[$ y deducir que:
$$\int_0^{\frac{\pi}{2}}\frac{\operatorname{arc\,tg}(\cos x)}{\cos x}\,dx=\frac{\pi}{2}\,\ln(1+\sqrt{2}).$$

14.9 Probar que
$$\int_0^{+\infty}\frac{\ln x}{x^2-1}\,dx=\frac{1}{2}\int_{\mathbb{R}^+\times\mathbb{R}^+}\frac{d(x,y)}{(1+x^2y)(1+y)}=\frac{\pi^2}{4}.$$
Deducir que
$$\int_0^1\frac{\ln x}{x-1}\,dx=\frac{\pi^2}{6}$$
y obtener a partir de ello que
$$\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{6}.$$

14.10 Probar que
$$\lim_{t\to+\infty}\int_0^t\frac{\operatorname{sen}x}{x}\,dx=\frac{\pi}{2}\,.$$
En el Ejemplo 13.8 se probó la existencia de dicho límite, ahora se trata de calcularlo

Indicaciones:

a) Probar, usando los Teoremas de Fubini y Tonelli, que la función
$$F(x,y)=e^{-xy}\operatorname{sen}x$$
es integrable en $]0,n[\times]0,\infty[$ y que
$$\int_0^n\frac{\operatorname{sen}x}{x}\,dx=\int_0^{+\infty}\left[\int_0^n e^{-xy}\operatorname{sen}x\,dx\right]dy.$$

b) Para cada natural n, sea $f_n:]0,+\infty[\to\mathbb{R}$ la función definida por
$$f_n(y)=\int_0^n e^{-xy}\operatorname{sen}x\,dx.$$
Probar, integrando por partes, que
$$\{f_n(y)\}\longrightarrow\frac{1}{1+y^2},$$
Probar además que
$$|f_n(y)|\leq\frac{2}{1+y^2},\ \forall n\in\mathbb{N}.$$

c) Deducir finalmente, utilizando el teorema de la convergencia dominada (Teorema 12.5) que
$$\lim_{t\to+\infty}\int_0^t \frac{\operatorname{sen} x}{x}\,dx=\frac{\pi}{2}.$$

14.11 Calcular las siguientes integrales:

i)
$$\int_E \frac{x}{\sqrt{1+x^2+y^2}}\,d(x,y)\ \text{ con }\ E=\{(x,y)\in\mathbb{R}^2:\ y^2\le 2x,\ 0\le x\le 2\}.$$

ii)
$$\int_E \sqrt{xy}\,d(x,y)\ \text{ con }\ E=\left\{(x,y)\in\mathbb{R}^2:\ 0\le y,\left(\frac{2x+y}{4}\right)^2\le\frac{x}{6}\right\}.$$

iii)
$$\int_E \frac{2xy(2-3x)}{x^2+2y^2}\,d(x,y)\ \text{ con }\ E=\{(x,y)\in\mathbb{R}^2:\ 0<x,\ 0<y,\ y^2\le\frac{1}{2}-x\}.$$

iv)
$$\int_E (x+y)\,d(x,y)\ \text{ con }\ E=\{(x,y)\in\mathbb{R}^2:\ y^2\le x+2,\ x^2\le y+2\}.$$

v)
$$\int_E y^2\,d(x,y)\ \text{ con }\ E=\{(x,y)\in\mathbb{R}^2:\ x^2+y^2\le R^2\}.$$

vi)
$$\int_E \frac{x^2+y^2}{\sqrt{1+x^2+y^2}}\,d(x,y)\ \text{ con }\ E=\left\{(x,y)\in\mathbb{R}^2:\ x^2+y^2<1,\ 0<y\right\}.$$

vii)
$$\int_E \left(\frac{x^2}{a^2}+\frac{y^2}{b^2}\right)d(x,y)\ \text{ con }\ E=\left\{(x,y)\in\mathbb{R}^2:\ \frac{x^2}{a^2}+\frac{y^2}{b^2}<1\right\}.$$

viii)
$$\int_E \frac{d(x,y)}{(1+x^2+y^2)^{\frac{3}{2}}}\ \text{ con }\ E=\left\{(x,y)\in\mathbb{R}^2:\ 0<x<1,\ 0<y<\frac{1}{\sqrt{3}}\right\}.$$

ix)
$$\int_E \frac{d(x,y,z)}{(1+x+y+z)^3}$$
con
$$E=\left\{(x,y,z)\in\mathbb{R}^3:\ 0\le x,\ 0\le y,\ 0\le z,\ x+y+z\le 1\right\}.$$

x)
$$\int_E (x^2+y^2+z^2)\,d(x,y,z)$$
con
$$E=\left\{(x,y,z)\in\mathbb{R}^3:0\le x,\ 0\le y,\ 0\le z,\ \frac{x}{a}+\frac{y}{b}+\frac{x}{c}\le 1\right\}\ (a,b,c>0).$$

14.7. Soluciones a los ejercicios del Tema 14.

14.1 Estudiar la integrabilidad de las siguientes funciones:

i) $\boxed{f(x,y) = e^{-xy} - 2e^{-2xy} \notin \mathscr{L}^1(]0,1[\times]1,+\infty[)}$
Este ejercicio se puede hacer hallando las integrales iteradas de la función f que son distintas y por el teorema de Fubini (Teorema14.1) la función no es integrable. El método es peligroso pues si después de hacer las cuentas fuesen iguales de ahí no se deduce nada y el trabajo se habría perdido.
Como
$$\int_0^1 \left(e^{-xy} - 2e^{-2xy}\right)\, dy = \frac{e^{-2x} - e^{-x}}{x}$$
y
$$\int_1^{+\infty} \left(e^{-xy} - 2e^{-2xy}\right)\, dx = \frac{e^{-y} - e^{-2y}}{y}$$
de ser las integrales iteradas iguales concluiríamos que
$$\int_1^{+\infty} \frac{e^{-x} - e^{-2x}}{x}\, dx = 0\ ,$$
lo que es absurdo pues el integrando es negativo.

ii) $\boxed{f(x,y) = (x-y)e^{-(x-y)^2} \notin \mathscr{L}^1(\mathbb{R}^+ \times \mathbb{R}^+)}$

iii) $\boxed{f(x,y) = \operatorname{sen}\left(\frac{1}{x-y}\right) \in \mathscr{L}^1(\{(x,y) \in \mathbb{R}^2 :\ x^2 + y^2 \leq 1\})}$
En efecto, f es continua, está acotada y el conjunto de integración es de medida finita.

iv) $\boxed{f(x,y) = e^{-(x+y)} \in \mathscr{L}^1(\{(x,y) \in \mathbb{R}^2 :\ 0 < x,\ 0 < y < x\})}$
En efecto, f es medible positiva y
$$\int_0^{+\infty} \left[\int_0^x e^{-x-y} dy\right] dx = \int_0^{+\infty} e^{-x} \left[\int_0^x e^{-y} dy\right] dx = \int_0^{+\infty} e^{-x}\left(1 - e^{-x}\right) dx = \frac{1}{2}\ .$$

Si se quiere integrar en el otro orden hay que tener en cuenta que
$$\{(x,y) \in \mathbb{R}^2 : 0 < x,\ 0 < y < x\} = \{(x,y) \in \mathbb{R}^2 : 0 < y,\ y < x\}\ .$$
$$\int_0^{+\infty} \left[\int_y^{+\infty} e^{-x-y} dx\right] dy = \frac{1}{2}\ .$$

v) $\boxed{f(x,y) = e^{-xy} \notin \mathscr{L}^1\big(\{(x,y) \in \mathbb{R}^2 :\ 1 < x,\ 0 < y < \tfrac{1}{x}\}\big)}$
En efecto la función f es medible positiva. Como

$$\int_1^{+\infty} \Big[\int_0^{\frac{1}{x}} e^{-xy}\, dy\Big] dx = \int_1^{+\infty} \Big[-\frac{1}{x}\, e^{-xy}\Big]_0^{\frac{1}{x}} dx = \int_1^{+\infty} (1-\frac{1}{e})\frac{1}{x} = +\infty\,,$$

f no es integrable en virtud del teorema de Fubini.

vi) $\boxed{f(x,y) = e^{-|x-y|} \in \mathscr{L}^1(]-1,1[\times\mathbb{R})}$
En efecto la función f es medible positiva. Como

$$\int_{-1}^{1} \Big[\int_{-\infty}^{+\infty} e^{-|x-y|} dy\Big]\, dx = \int_{-1}^{1} \Big[\int_{-\infty}^{x} e^{y-x}\, dy - \int_{x}^{+\infty} e^{x-y} dy\Big]\, dx =$$

$$\int_{-1}^{1} \Big(\big[e^{y-x}\big]_{-\infty}^{x} + \big[-e^{x-y}\big]_{x}^{+\infty}\Big)\, dx = \int_{-1}^{1} 2\, dx = 4,$$

Luego por el teorema de Tonelli (Teorema 14.6) la función f es integrable y por el teorema de Fubini

$$\int_{]-1,1[\times\mathbb{R}} f(x,y)\, d(x,y) = 4.$$

vii) $\boxed{f(x,y) = \frac{x}{(x^2+y^2)\sqrt{1-x}} \in \mathscr{L}^1(\{(x,y) \in \mathbb{R}^2 :\ 0 < x,\ 0 < y < x\})}$ En efecto la función f es medible positiva. Como

$$\int_0^1 \Big[\int_0^x \frac{x}{(x^2+y^2)\sqrt{1-x}}\, dy\Big]\, dx = \frac{\pi}{2}$$

Luego por el teorema de Tonelli la función f es integrable y por el teorema de Fubini

$$\int_{\{(x,y)\in\mathbb{R}^2:\ 0<x<1,\ 0<y<\frac{1}{x}\}} f(x,y)\, d(x,y) = \frac{\pi}{2}$$

viii) $\boxed{f(x,y) = \frac{\cos(xy)}{(1+y^2)\sqrt{\operatorname{sen} x}} \in \mathscr{L}^1(]0,1[\times\mathbb{R}^+)}$

$$|f(x,y)| \leq g(x,y) = \frac{1}{1+y^2)\sqrt{\operatorname{sen} x}}$$

La función g es medible positiva y

$$\int_0^{+\infty} g(x,y)\, dy = \frac{\pi}{4}\, \frac{1}{\sqrt{\operatorname{sen} x}}$$

que es integrable en $]0,1[$ $\big($comparar con $\frac{1}{\sqrt{x}}\big)$, luego por por el teorema de Tonelli la función g es integrable e igual le ocurre a la función f (medible y acotada por una integrable).

ix) $\boxed{f(x,y,z)=\frac{1}{(x^2+y^2+z^2)^\alpha}\notin\mathscr{L}^1(\mathbb{R}^3)(\alpha\in\mathbb{R})}$

Veamos que la función f no es integrable en $\mathbb{R}^3$. Usando el teorema de cambio de variable (Teorema 14.10) para coordenadas esféricas, y el criterio de integrabilidad (Corolario 14.7), la integrabilidad de f en la bola unidad equivale a la integrabilidad de la función $\rho\mapsto\rho^{2-2\alpha}$ en $]0,1[$ (respect. en $]1,+\infty[$), lo cual ocurre si, sólo si, $\alpha<\frac{3}{2}$ (respectivamente $\alpha>\frac{3}{2}$).

En consecuencia, f no es integrable para ningún valor de α.

Para los valores de α en los que f es integrable en la bola unidad euclídea, teniendo en cuenta que en este caso, el determinante del jacobiano del cambio de variable en (ρ,θ,φ) es $\rho^2\cos\varphi$, usando los teoremas del cambio de variable y de Fubini, se verifica

$$\int_{B(0,1)}f(x,y,z)\,d(x,y,z)=\int_0^1\Big[\int_{-\pi}^{\pi}\Big[\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\rho^{2-2\alpha}\cos\varphi\,d\varphi\Big]\,d\theta\Big]\,d\rho=$$

$$4\pi\int_0^1\rho^{2-2\alpha}\,d\varphi=\frac{4\pi}{3-2\alpha}\,.$$

Análogamente

$$\int_{\mathbb{R}^3\setminus B(0,1)}f(x,y,z)\,d(x,y,z)=\frac{4\pi}{2\alpha-3}\,.$$

14.2 i) $\boxed{(x,y)\to\frac{xy}{(x^2+y^2)^\alpha}\in\mathscr{L}^1(\{(x,y)\in\mathbb{R}^2:\ x^2+y^2>1\})\Longleftrightarrow\alpha>2}$.

En efecto: la fórmula de cambio de variable a coordenadas polares nos dice que

$$f\in\mathscr{L}^1\left(\{(x,y)\in\mathbb{R}^2:\ x^2+y^2>1\}\right)\Longleftrightarrow\rho^{3-2\alpha}\,\cos\vartheta\ \operatorname{sen}\vartheta\mathscr{L}^1(]1,+\infty[\times]-\pi,\pi[)\,,$$

lo que equivale, en virtud del criterio de integrabilidad (Corolario14.7), a la finitud de la integral

$$\int_1^{+\infty}\left[\int_{-\pi}^{\pi}\rho^{3-2\alpha}|\cos\vartheta\ \operatorname{sen}\vartheta|\,d\vartheta\right]\,d\rho,$$

lo que a su vez equivale a la finitud de

$$\int_1^{+\infty}\rho^{3-2\alpha}\,d\rho\,,$$

esto es $\alpha>2$.

ii) $\boxed{1<\alpha<2\Longrightarrow(x,y)\to\frac{\operatorname{sen}x\ \operatorname{sen}y}{(x^2+y^2)^\alpha}\in\mathscr{L}^1(\mathbb{R}^2)}$.

Notemos

$$f(x,y)=\frac{\operatorname{sen}x\ \operatorname{sen}y}{(x^2+y^2)^\alpha},\ \forall(x,y)\in\mathbb{R}^2\backslash\{(0,0)\}.$$

La aditividad de la integral respecto el recinto de integración nos asegura que

$$f\in\mathscr{L}^1(\mathbb{R}^2)\Longleftrightarrow f\in\mathscr{L}^1(B(0,1))\cap\mathscr{L}^1(\mathbb{R}^2\setminus B(0,1))\,.$$

Si hacemos la mayoración

$$|f(x,y)| \le \frac{1}{(x^2+y^2)^\alpha}$$

no llegamos a nada pues

$$\frac{1}{(x^2+y^2)^\alpha} \notin \mathscr{L}^1(\mathbb{R}^2).$$

Sin embargo esta mayoración es buena para trabajar en $\mathbb{R}^2 \setminus B(0,1)$. En efecto,

$$1 < \alpha \Longleftrightarrow \rho^{1-2\alpha} \in \mathscr{L}^1(]1,+\infty[) \Longleftrightarrow \frac{1}{(x^2+y^2)^\alpha} \in \mathscr{L}^1(\mathbb{R}^2 \setminus B(0,1))$$

$$\Longrightarrow f \in \mathscr{L}^1(\mathbb{R}^2 \setminus B(0,1)).$$

Para trabajar en $B(0,1)$ hay que afinar más y utilizar la acotación

$$|\operatorname{sen} x| \le |x| \quad , \quad |\operatorname{sen} y| \le |y|.$$

En efecto,

$$\alpha < 2 \Longleftrightarrow \rho^{3-2\alpha} \in \mathscr{L}^1(]0,1[) \Longleftrightarrow \frac{|x|\ |y|}{(x^2+y^2)^\alpha} \in \mathscr{L}^1(B(0,1))$$

$$\Longrightarrow f \in \mathscr{L}^1(B(0,1)).$$

En ambos casos se han utilizado la fórmula de cambio de variable para coordenadas polares, la integrabilidad de las funciones potenciales, el criterio de integrabilidad y que toda función medible acotada por una integrable es integrable.

14.3 Sean $x_1,\ldots,x_N,a \in \mathbb{R}^N$ y sea

$$T := \left\{ x \in \mathbb{R}^N :\ x = a + \sum_{k=1}^N t_k x_k,\ \sum_{k=1}^N t_k \le 1,\ 0 \le t_k,\ 1 \le k \le N \right\}$$

el $(N+1)$-edro en $\mathbb{R}^N$ (generalización del triángulo en $\mathbb{R}^2$ y del tetraedro en $\mathbb{R}^3$). Probar que

$$\lambda(T) = \frac{1}{N!}\ |\det\,(x_1,\ldots,x_N)|.$$

Para cada natural N, consideramos el conjunto

$$A_N := \left\{ t \in \mathbb{R}^N : \sum_{k=1}^N t_k \le 1, 0 \le t_k\ (1 \le k \le N) \right\}$$

y la aplicación lineal $S : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ dada por

$$S(e_k) = x_k \qquad (1 \le k \le N)\ .$$

Por ser S lineal y llevar la base canónica en los vectores $\{x_1,\ldots,x_N\}$ tenemos que

$$S(t)=S(t_1e_1+\ldots+t_Ne_N)=t_1S(e_1)+\ldots+t_NS(e_N)=t_1x_1+\ldots+t_Nx_N,$$

y en consecuencia

$$S(A_N)=T-a\,.$$

Obtenemos entonces que

$$\lambda(T)=\lambda(T-a)=\lambda(S(A_N))=|\det S|\;\lambda(A_N)=|\det\,(x_1,\ldots,x_N)|\;\lambda(A_N)$$

y probaremos por inducción que

$$\lambda(A_N)=\frac{1}{N!}\,.$$

Para $N=1$ es inmediato. Supongamos que es cierto para un natural N y lo probaremos para $N+1$. Así

$$\lambda(A_{N+1})=\int_{\mathbb{R}^{N+1}}\chi_{A_{N+1}}\,d(t_1,\ldots,t_{N+1})=$$

$$=\int_0^1\left[\int_{\{t\in\mathbb{R}^N:\ \sum_{k=1}^N t_k\leq 1-t_{N+1},\ 0\leq t_k\ (1\leq k\leq N\}}d(t_1,\ldots,t_N)\right]\,dt_{N+1}.$$

Como el conjunto

$$\left\{t\in\mathbb{R}^N:\ \sum_{k=1}^N t_k\leq 1-t_{N+1},\ 0\leq t_k, 1\leq k\leq N\right\}$$

es la imagen de A_N mediante la homotecia de razón $1-t_{N+1}$, usando las propiedades de la medida de Lebesgue y la hipótesis de inducción, concluimos que

$$\lambda(A_{N+1})=\int_0^1\Big[(1-t_{N+1})^N\lambda(A_N)\Big]\,dt_{N+1}=\frac{1}{N+1}\lambda(A_N)=\frac{1}{N+1}\frac{1}{N!}=\frac{1}{(N+1)!}\,.$$

14.4 Calcular el volumen de los siguientes conjuntos:

i) Bóveda de Viviani: Intersección de un cilindro circular con la semiesfera, esto es,

$$E=\left\{(x,y,z)\in\mathbb{R}^3:\left(x-\frac{1}{2}\right)^2+y^2\leq\frac{1}{4}\right\}\bigcap$$

$$\left\{(x,y,z)\in\mathbb{R}^3:\ x^2+y^2+z^2\leq 1,\ 0\leq z\right\}.$$

El conjunto E es compacto, luego no hay problema de integración. La bóveda de Viviani se puede ver como el conjunto ordenado limitado por la función

$$(x,y)\longmapsto\sqrt{1-x^2-y^2}$$

sobre el recinto de integración

$$F = B_2\left(\left(\frac{1}{2},0\right),\frac{1}{2}\right) = \left\{(x,y)\in\mathbb{R}^2:\ \left(x-\frac{1}{2}\right)^2 + y^2) \le \frac{1}{4}\right\}.$$

La Proposición 14.8 nos asegura que

$$\lambda(E) = \int_F \sqrt{1-x^2-y^2}\, d(x,y).$$

El integrando nos aconseja utilizar coordenadas polares. Haciendo en F el cambio $\begin{cases} x = \rho\ \cos\vartheta \\ y = \rho\ \operatorname{sen}\vartheta \end{cases}$, queda

$$\rho(\rho - \cos\vartheta) \le 0 \Longleftrightarrow \rho \le \cos\vartheta \ \text{ c.p.d. } \ \left(\Longrightarrow -\frac{\pi}{2} \le \vartheta \le \frac{\pi}{2}\right),$$

esto es, para cada valor de la variable ϑ entre $-\frac{\pi}{2}$ y $\frac{\pi}{2}$ la variable ρ varía en $]0,\cos\vartheta]$. Así

$$\phi^{-1}(F) = \left\{(\rho,\vartheta):\ -\frac{\pi}{2} \le \vartheta \le \frac{\pi}{2},\ 0 < \rho \le \cos\vartheta\right\}.$$

La fórmula del cambio de variable (Corolario 14.11) nos da que

$$\lambda(E) = \int_{\phi^{-1}(F)} \sqrt{1-\rho^2}\,\rho\ d(\rho,\vartheta).$$

Ahora el teorema de Fubini (Teorema 14.1) nos hace concluir que

$$\lambda(E) = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left[\int_0^{\cos\vartheta} \sqrt{1-\rho^2}\ \rho\ d\rho\right] d\vartheta = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left[\left[-\frac{(1-\rho^2)^{\frac{3}{2}}}{3}\right]_{\rho=0}^{\rho=\cos\vartheta}\right] d\vartheta =$$
$$\frac{1}{3}\left(\pi - \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} |\operatorname{sen}\vartheta|^3 d\vartheta\right) = \frac{1}{3}\left(\pi - 2\int_0^{\frac{\pi}{2}} \operatorname{sen}^3\vartheta\ d\vartheta\right) = \frac{3\pi-4}{9}.$$

ii) El conjunto E es compacto, luego no hay problema de integración. Dicho conjunto se puede ver como el conjunto ordenado limitado por la función

$$(x,y) \longmapsto \left(1 - x^{\frac{2}{3}} - y^{\frac{2}{3}}\right)^{\frac{3}{2}}$$

sobre el recinto de integración

$$F = \left\{(x,y)\in\mathbb{R}^+\times\mathbb{R}^+:\ x^{\frac{2}{3}} + y^{\frac{2}{3}} \le 1\right\}.$$

La Proposición 14.8 nos asegura que

$$\lambda(E) = \int_F \left(1 - x^{\frac{2}{3}} - y^{\frac{2}{3}}\right)^{\frac{3}{2}}\ d(x,y).$$

Haciendo uso de la transformación indicada, esto es, el difeomorfismo ϕ de clase $\mathbf{C}^1$ del abierto $\mathbb{R}^+ \times \left]0, \frac{\pi}{2}\right[$ sobre el abierto $\mathbb{R}^+ \times \mathbb{R}^+$ dado por

$$\left\{\begin{matrix} x = \rho \ \cos^3 \vartheta \\ y = \rho \ \operatorname{sen}^3 \vartheta \end{matrix}\right\} \Longrightarrow \left|\det \ J_\phi(\rho, \vartheta)\right| = 3 \ \rho \ \operatorname{sen}^2 \theta \ \cos^2 \vartheta > 0.$$

Utilizando en F el citado cambio de variable nos queda

$$\rho^{\frac{2}{3}}(\cos^2 \vartheta + \operatorname{sen}^2 \vartheta) \leq 1 \Longleftrightarrow \rho^{\frac{2}{3}} \leq 1 \Longleftrightarrow 0 < \rho \leq 1,$$

es decir, para cada valor de la variable ϑ entre 0 y $\dfrac{\pi}{2}$, la variable ρ varía en $]0, 1]$. Así

$$\phi^{-1}(F) = \left\{(\rho, \vartheta) : \ 0 \leq \vartheta \leq \frac{\pi}{2}, \ 0 < \rho \leq 1\right\} \ \text{ c.p.d. .}$$

La fórmula del cambio de variable (Corolario 14.10) nos da que

$$\lambda(E) = \int_{\phi^{-1}(F)} \left(1 - \rho^{\frac{2}{3}}\right)^{\frac{3}{2}} 3 \ \rho \ \operatorname{sen}^2 \theta \ \cos^2 \vartheta \ d(\rho, \vartheta).$$

Ahora el Teorema de Fubini 14.1 nos hace concluir que

$$\lambda(E) = \int_0^{\frac{\pi}{2}} \left[\int_0^1 \left(1 - \rho^{\frac{2}{3}}\right)^{\frac{3}{2}} 3 \ \rho \ \operatorname{sen}^2 \theta \ \cos^2 \vartheta \ d\rho\right] d\vartheta =$$

$$\left[\int_0^{\frac{\pi}{2}} \operatorname{sen}^2 \vartheta \ \cos^2 \vartheta \ d\vartheta\right] \left[\int_0^1 \left(1 - \rho^{\frac{2}{3}}\right)^{\frac{3}{2}} 3 \ \rho \ d\rho\right] = \frac{\pi}{16} \times \frac{8}{35} = \frac{\pi}{70},$$

donde se han usado dos resultados del Ejercicio 13.0 $\big($en el segundo se aconseja el cambio de variable $\rho = \left(1 - x^2\right)^{\frac{3}{2}}\big)$.

iii) El conjunto E es compacto, luego no hay problema de integración. Se trata de calcular el volumen del sólido interior al cono

$$\left\{(x, y, z) \in \mathbb{R}^3 : \ x^2 + y^2 \leq (1 - z)^2\right\}$$

y al paraboloide

$$\left\{(x, y, z) \in \mathbb{R}^3 : \ x^2 + y^2 \leq \frac{z}{2}, \ z \leq 1\right\}$$

por debajo del vértice del cono

$$V = (0, 0, 1)$$

(un cucurucho de helado invertido). Se tiene que

$$\lambda(E) = \int_{\mathbb{R}^3} \chi_E \ d(x, y, z) = \int_E \ d(x, y, z).$$

Lo primero es resolver el sistema

$$\left\{\begin{matrix} x^2 + y^2 = (1 - z)^2 \\ x^2 + y^2 = \dfrac{z}{2} \end{matrix}\right. \Longrightarrow 2z^2 - 5z + 2 = 0 \Longleftrightarrow \left\{\begin{matrix} z = \frac{1}{2} \\ z = 2 \end{matrix}\right. .$$

La solución $z = 2$ es superior al vértice V del cono.

Primer método: El Teorema de Fubini nos da

$$\lambda(E) = \int_0^1 \left[\int_{E(z)} d(x,y) \right] dz =$$

$$\int_0^{\frac{1}{2}} \left[\int_{E(z)} d(x,y) \right] dz + \int_{\frac{1}{2}}^1 \left[\int_{E(z)} d(x,y) \right] dz =$$

$$\int_0^{\frac{1}{2}} \left[\int_{\{(x,y)\in\mathbb{R}^2 :\ x^2+y^2\le \frac{z}{2}\}} d(x,y) \right] dz + \int_{\frac{1}{2}}^1 \left[\int_{\{(x,y)\in\mathbb{R}^2 :\ x^2+y^2\le (1-z)^2\}} d(x,y) \right] dz =$$

$$\pi \int_0^{\frac{1}{2}} \frac{z}{2}\, dz + \pi \int_{\frac{1}{2}}^1 (1-z)^2\, dz = \frac{\pi}{16} + \frac{\pi}{24} = \frac{5\pi}{48},$$

donde se ha utilizado la fórmula del área de los círculos de radios $\sqrt{\frac{z}{2}}$ y $1 - z$ respectivamente.

Segundo método: Si proyectamos el círculo de intersección de ambas superficies, obtenemos el recinto de integración

$$F = \left\{ (x,y) \in \mathbb{R}^2 :\ x^2 + y^2 \le \frac{1}{4} \right\}.$$

Para cada $(x,y) \in F$ la variable z varía entre el paraboloide y el cono. Así el Teorema de Fubini nos da también

$$\lambda(E) = \int_F \left[\int_{2(x^2+y^2)}^{1-\sqrt{x^2+y^2}} dz \right] d(x,y) =$$

$$\int_F \left(1 - \sqrt{x^2+y^2} - 2(x^2+y^2) \right) d(x,y) =$$

$$\int_{]0,\frac{1}{2}[\times]-\pi,\pi[} \left(1 - \rho - 2\rho^2\right) \rho\, d(\rho,\vartheta) =$$

$$\int_0^{\frac{1}{2}} \left[\int_{-\pi}^{\pi} \left(1-\rho-2\rho^2\right) \rho\, d\vartheta \right] d\rho = 2\pi \int_0^{\frac{1}{2}} \left(\rho - \rho^2 - 2\rho^3\right) d\rho = \frac{5\pi}{48},$$

donde se han utilizado la fórmula del cambio de variable a coordenadas polares y otra vez el Teorema de Fubini para la integral doble.

iv) El conjunto E es compacto, luego no hay problema de integración. Se trata de calcular el volumen del sólido que el hiperboloide (diábolo)

$$\{(x,y,z) \in \mathbb{R}^3 :\ x^2 + y^2 - z^2 \le 1\}$$

delimita con la esfera

$$\{(x,y,z) \in \mathbb{R}^3 :\ x^2 + y^2 + z^2 \le 3\}.$$

Se tiene que

$$\lambda(E)=\int_{\mathbb{R}^3}\chi_E\, d(x,y,z)=\int_E\, d(x,y,z).$$

Lo primero es resolver el sistema

$$\begin{cases} x^2+y^2+z^2=3 \\ x^2+y^2-z^2=1 \end{cases} \Longrightarrow 3-z^2=1+z^2 \Longleftrightarrow z=\pm 1.$$

Por razones de simetría nos limitamos al conjunto

$$\{(x,y,z)\in\mathbb{R}^3:\ x^2+y^2+z^2\leq 3,\ x^2+y^2-z^2\leq 1,\ 0\leq z\},$$

con lo que la única solución es $z=1$.

<u>Primer método:</u> El Teorema de Fubini nos da

$$\lambda(E)=2\int_0^{\sqrt{3}}\left[\int_{E(z)}d(x,y)\right]dz=$$

$$2\left(\int_0^1\left[\int_{E(z)}d(x,y)\right]dz+\int_1^{\sqrt{3}}\left[\int_{E(z)}d(x,y)\right]dz\right)=$$

$$2\int_0^1\left[\int_{\{(x,y)\in\mathbb{R}^2:\ x^2+y^2\leq 1+z^2\}}d(x,y)\right]dz+$$

$$2\int_1^{\sqrt{3}}\left[\int_{\{(x,y)\in\mathbb{R}^2:\ x^2+y^2\leq 3-z^2\}}d(x,y)\right]dz=$$

$$2\pi\left(\int_0^1(1+z^2)\,dz+\int_1^{\sqrt{3}}3-z^2\,dz\right)=$$

$$2\pi\left(\left(1+\frac{1}{3}\right)+\left(3\sqrt{3}-3-\sqrt{3}+\frac{1}{3}\right)\right)=\frac{4(3\sqrt{3}-2)\pi}{3},$$

donde se ha utilizado la fórmula del área de los círculos de radios $\sqrt{1+z^2}$ y $\sqrt{3-z^2}$ respectivamente.

<u>Segundo método:</u> Si proyectamos el círculo de intersección de ambas superficies, obtenemos el recinto de integración

$$F=\left\{(x,y)\in\mathbb{R}^2:\ x^2+y^2\leq 2\right\},$$

que interesa descomponer de la forma

$$F=F_1\cup F_2=\left\{(x,y)\in\mathbb{R}^2:\ x^2+y^2\leq 1\right\}\cup\left\{(x,y)\in\mathbb{R}^2:\ 1<x^2+y^2\leq 2\right\}.$$

Para cada $(x,y)\in F_1$ la variable z varía entre el plano OXY y la esfera y para cada $(x,y)\in F_2$ la variable z varía entre el hiperboloide y la esfera. Así el Teorema de

Fubini nos da también

$$\lambda(E) = 2\left(\int_{F_1}\left[\int_0^{\sqrt{3-x^2-y^2}} dz\right] d(x,y) + \int_{F_2}\left[\int_{\sqrt{x^2+y^2-1}}^{\sqrt{3-x^2-y^2}} dz\right] d(x,y)\right) =$$

$$2\left(\int_{F_1}\sqrt{3-x^2-y^2}\, d(x,y) + \int_{F_2}\left(\sqrt{3-x^2-y^2} - \sqrt{x^2+y^2-1}\right) d(x,y)\right) =$$

$$2\left(\int_{]0,1]\times]-\pi,\pi[}\sqrt{3-\rho^2}\,\rho\, d(\rho,\vartheta) + \int_{]1,\sqrt{2}[\times]-\pi,\pi[}\left(\sqrt{3-\rho^2} - \sqrt{\rho^2-1}\right)\rho\, d(\rho,\vartheta)\right) =$$

$$2\left(\int_0^1\left[\int_{-\pi}^{\pi}\sqrt{3-\rho^2}\,\rho\, d\vartheta\right] d\rho + \int_1^{\sqrt{2}}\left[\int_{-\pi}^{\pi}\left(\sqrt{3-\rho^2} - \sqrt{\rho^2-1}\right)\rho\, d\vartheta\right] d\rho\right) =$$

$$-4\pi\times\frac{2}{3}\times\frac{1}{2}\left(\left[(3-\rho^2)^{\frac{3}{2}}\right]_0^1 + \left[(3-\rho^2)^{\frac{3}{2}}\right]_1^{\sqrt{2}} + \left[(\rho^2-1)^{\frac{3}{2}}\right]_1^{\sqrt{2}}\right) =$$

$$\frac{-4\pi}{3}\left(\left[(3-\rho^2)^{\frac{3}{2}}\right]_0^{\sqrt{2}} + \left[(\rho^2-1)^{\frac{3}{2}}\right]_1^{\sqrt{2}}\right) = \frac{4(3\sqrt{3}-2)\pi}{3},$$

donde se han utilizado la fórmula del cambio de variable a coordenadas polares y otra vez el Teorema de Fubini para la integral doble.

14.5 Área de la cardioide.

La curva en $\mathbb{R}^2$, cuya ecuación en coordenadas polares viene dada por

$$\rho = 2a(1+\cos\vartheta) \quad (a\in\mathbb{R}^+,\ -\pi<\vartheta\le\pi)$$

se llama una <u>*cardioide*</u> . Sea

$$E = \{(\rho\ \cos\vartheta, \rho\ \operatorname{sen}\vartheta):\ -\pi<\vartheta\le\pi,\ 0\le\rho\le 2a(1+\cos\vartheta)\}.$$

Calcular $\lambda(E)$.

Conviene <u>dibujar</u> la cardioide lo que justifica su nombre. Se tiene que E es un compacto luego medible con medida finita y

$$\lambda(E) = \int_E d(x,y)\ .$$

Es fácil ver que si ϕ es el cambio a polares, entonces

$$\phi^{-1}(E) = \{(\rho,\vartheta): -\pi<\vartheta<\pi, 0<\rho<2a(1+\cos\vartheta\}\ c.p.d.$$

La fórmula de cambio de variable (Corolario 14.11) a coordenadas polares nos da

$$\lambda(E) = \int_{\phi^{-1}(E)} \rho\ d(\rho,\vartheta)\ .$$

El teorema de Fubini (Teorema14.1) nos hace concluir

$$\lambda(E) = \int_{-\pi}^{\pi}\left[\int_0^{2a(1+\cos\vartheta)}\rho\ d\rho\right] d\vartheta = 2a^2\int_{-\pi}^{\pi}(1+\cos\vartheta)^2 = 6a^2\pi\ .$$

14.6 **Sólidos de revolución.**
Sea $E \subset \mathbb{R}_0^+ \times \mathbb{R}$ un conjunto medible. Consideremos el conjunto

$$S = \left\{ (x,y,z) \in \mathbb{R}^3 : \left(\sqrt{x^2+y^2}, z \right) \in E \right\}$$

(sólido de revolución obtenido al girar el conjunto E contenido en el plano XZ en torno al eje OZ). Probar que S es medible y que

$$\lambda(S) = 2\pi \int_E x \, d(x,z).$$

En efecto consideremos el difeomorfismo de clase $\mathscr{C}^\infty$

$$\phi : \mathbb{R}^+ \times]-\pi,\pi[\times \mathbb{R} \longrightarrow \mathbb{R}^3 \setminus \{(x,0,z) : \ x \le 0\}$$

que da el cambio a coordenadas cilíndricas. Utilizando dichas coordenadas cuando $\vartheta = 0$ el sólido de revolución se confunde con E, y para otro valor de la variable ϑ las variables (ρ,z) se mueven en el girado de E, que se comporta como E a todos los efectos, así

$$\chi_{\phi^{-1}(S)}(\rho,\vartheta,z) = \chi_E(\rho,z)\ \chi_{]-\pi,\pi[}(\vartheta), \text{ c.t. } (\rho,\vartheta,z)$$

es decir

$$\phi^{-1}(S) = \{(\rho,\vartheta,z) : \ (\rho,z) \in E, \ -\pi < \vartheta < \pi\} \equiv E \times]-\pi,\pi[$$

de donde se deduce que $\phi^{-1}(S)$ es medible y por tanto S es medible. Se tiene que

$$\lambda(S) = \int_S d(x,y,z) = \int_{\phi^{-1}(S)} \rho \ d(\rho,\vartheta,z) = \int_E \left[\int_{-\pi}^{\pi} \rho \ d\vartheta \right] d(\rho,z) =$$

$$2\pi \int_E \rho \ d(\rho,z) = 2\pi \int_E x \ d(x,z),$$

donde se han utilizado la fórmula de cambio de variable a coordenadas cilíndricas y el teorema de Fubini.

Aplicación: probar que

a) El volumen del toro engendrado por la rotación en torno al eje OZ del círculo $(x-1)^2 + (z-1)^2 \le \frac{1}{4}$ (salvavidas) es $\frac{\pi^2}{2}$.
En efecto, haciendo el cambio

$$\begin{cases} x = 1 + \rho \operatorname{sen} \ \vartheta \\ y = 1 + \rho \cos \ \vartheta \end{cases}$$

se tiene que

$$\int_E x \ d(x,z) = \int_0^{\frac{1}{2}} \left[\int_{-\pi}^{\pi} (1 + \rho \cos \ \vartheta) \rho \ d\vartheta \right] d\rho = \int_0^{\frac{1}{2}} 2\pi \ \rho \ d\rho = \frac{\pi}{4}$$

b) El conjunto
$$E = \left\{ (x,z) \in \mathbb{R}^2 : \ 1 < x, \ 0 < z < \frac{1}{x^2} \right\}$$
tiene área uno y que el sólido engendrado al rotar dicho conjunto en torno al eje OZ tiene volumen infinito.

En efecto
$$\lambda(E) = \int_1^{+\infty} \left[\int_0^{\frac{1}{x^2}} dz \right] dx = \int_1^{+\infty} \frac{dx}{x^2} = \left[-\frac{1}{x} \right]_1^{\infty} = 1$$
$$\lambda(S) = 2\pi \int_E x d(x,z) = 2\pi \int_1^{+\infty} \left[\int_0^{\frac{1}{x^2}} x\, dz \right] dx = 2\pi \int_1^{+\infty} \frac{dx}{x} = +\infty \, .$$

14.7 **Integral de Gauss.**
Sea $g : \mathbb{R}^+ \times]0, \frac{\pi}{2}[\to \mathbb{R}$ la función definida por
$$g(\rho,\vartheta) = f(\rho \ \cos\vartheta, \rho \ \operatorname{sen}\vartheta)\rho = \rho \ e^{-a\rho^2}.$$
El teorema del cambio de variable (Teorema 14.10) nos asegura que
$$f \in \mathscr{L}(\mathbb{R}^+ \times \mathbb{R}^+) \Longleftrightarrow g \in \mathscr{L}\left(\mathbb{R}^+ \times \left]0, \frac{\pi}{2}\right[\right),$$
en cuyo caso sus integrales coincide. Se tiene que
$$\int_0^{+\infty} \left[\int_0^{\frac{\pi}{2}} \rho \ e^{-a\rho^2} \ d\vartheta \right] d\rho = \frac{\pi}{2} \int_o^{+\infty} \rho \ e^{-a\rho^2} \ d\rho = \frac{\pi}{2} \left[\frac{-e^{-a\rho}}{2a} \right]_{\rho=0}^{\rho=+\infty} = \frac{\pi}{4a},$$
donde se ha utilizado la regla de Barrow (Teorema13.5). Los teoremas de Tonelli y Fubini (Teoremas14.6 y 14.1) nos aseguran que g es integrable y que
$$\int_{\mathbb{R}^+ \times]0, \frac{\pi}{2}[} \rho \ e^{-a\rho^2} \ d(\rho,\theta) = \frac{\pi}{4a}.$$
Ahora el teorema del cambio de variable nos asegura la integrabilidad de f y que
$$\int_{\mathbb{R}^+ \times \mathbb{R}^+} e^{-a(x^2+y^2)} \ d(x,y) = \frac{\pi}{4a}.$$
Una nueva aplicación del teorema de Fubini nos dice que la función $x \longmapsto e^{-ax^2}$ es integrable en $\mathbb{R}^+$ y que
$$\frac{\pi}{4a} = \int_{\mathbb{R}^+} \left[\int_{\mathbb{R}^+} e^{-a(x^2+y^2)} \ dy \right] dx = \int_{\mathbb{R}^+} e^{-ax^2} \left[\int_{\mathbb{R}^+} e^{-ay^2} \ dy \right] dx =$$
$$\left[\int_{\mathbb{R}^+} e^{-ax^2} \ dx \right] \left[\int_{\mathbb{R}^+} e^{-ay^2} \ dy \right] = \left[\int_{\mathbb{R}^+} e^{-ax^2} \ dx \right]^2,$$
esto es,
$$\int_0^{+\infty} e^{-ax^2} \ dx = \frac{1}{2}\sqrt{\frac{\pi}{a}} \ \ (a > 0).$$

14.8 Probar que la función

$$f(x,y) = \frac{1}{1+x^2+y^2}$$

es integrable en $]0,+\infty[\times]0,1[$ y deducir que:

$$\int_0^{\frac{\pi}{2}} \frac{\operatorname{arc\,tg}(\cos x)}{\cos x}\, dx = \frac{\pi}{2}\ \ln(1+\sqrt{2}).$$

En efecto, la función f es medible positiva, con lo que en virtud del criterio de integrabilidad (Corolario 14.7) es integrable si, y sólo si, la siguinte integral es finita

$$\int_0^1 \left[\int_0^{+\infty} \frac{dx}{1+x^2+y^2}\right] dy = \int_0^1 \left[\int_0^{+\infty} \frac{1}{1+y^2} \frac{dx}{1+\left(\frac{x}{\sqrt{1+y^2}}\right)^2}\right] dy =$$

$$\int_0^1 \left[\frac{1}{\sqrt{1+y^2}} \int_0^{+\infty} \frac{\frac{dx}{\sqrt{1+y^2}}}{1+\left(\frac{x}{\sqrt{1+y^2}}\right)^2}\right] dy =$$

$$\int_0^1 \left[\left[\arctan\left(\frac{x}{\sqrt{1+y^2}}\right)\right]_{x=0}^{x=+\infty} \frac{1}{\sqrt{1+y^2}}\right] dy =$$

$$\frac{\pi}{2}\int_0^1 \frac{dy}{\sqrt{1+y^2}} = \frac{\pi}{2}\left[\log(y+\sqrt{1+y^2}) = \operatorname{arg\,senh} y\right]_{y=0}^{y=1} = \frac{\pi\log(1+\sqrt{2})}{2}\ .$$

El teorema de Fubini (Teorema 14.1) nos asegura ahora que

$$\frac{\pi\log(1+\sqrt{2}}{2} = \int_0^{+\infty}\left[\int_0^1 \frac{dx}{1+x^2+y^2}\right] dx =$$

$$= \int_0^{+\infty} \arctan\left(\frac{1}{\sqrt{1+x^2}}\right)\frac{dx}{\sqrt{1+x^2}} = \left(x = \tan t\right) = \int_0^{\frac{\pi}{2}} \frac{\arctan(\cos t)}{\cos t}\, dt\ .$$

Hemos probado que

$$\int_0^{\pi/2} \frac{\arctan(\cos x)}{\cos x}\, dx = \frac{\pi\log(1+\sqrt{2})}{2}$$

14.9 La función $f:\mathbb{R}^+\times\mathbb{R}^+ \longrightarrow \mathbb{R}$ definida por

$$f(x,y) = \frac{1}{2(1+y)(1+x^2y)}$$

es medible positiva, el criterio de integrabilidad (Corolario 14.7) nos dice que f es integrable si, y sólo si, la siguiente integral es finita

$$\frac{1}{2}\int_0^{+\infty}\Big[\int_0^{+\infty}\frac{1}{(1+y)(1+yx^2)}\,dx\Big]\,dy =$$

$$= \frac{1}{2}\int_0^{+\infty}\Big[\frac{1}{(1+y)\sqrt{y}}\Big[\arctan\sqrt{y}x\Big]_{x=0}^{x=\infty}\Big]\,dy = \frac{\pi}{2}\int_0^{+\infty}\frac{1}{2(1+y)\sqrt{y}}\,dy =$$

$$= \frac{\pi}{2}\Big[\arctan\sqrt{y}\Big]_{y=0}^{y=+\infty} = \frac{\pi^2}{4}\ .$$

El teorema de Fubini (Teorema 14.1) nos asegura ahora que

$$\frac{\pi^2}{4} = \frac{1}{2}\int_0^{+\infty}\Big[\int_0^{+\infty}\frac{dy}{(1+y)(1+yx^2)}\Big]\,dx\ .$$

Como se verifica c.p.d. la identidad

$$f(x,y) = \frac{1}{2(1-x^2)}\left(\frac{1}{1+y} - \frac{x^2}{1+yx^2}\right),$$

(salvo en la semirrecta $x=1, y>0$), entonces se tiene que

$$\int_0^{+\infty}\frac{dy}{(1+y)(1+yx^2)} = \int_0^{+\infty}\frac{1}{1-x^2}\left(\frac{1}{1+y} - \frac{x^2}{1+yx^2}\right)\,dy =$$

$$= \frac{1}{1-x^2}\left[\log\left(\frac{1+y}{1+yx^2}\right)\right]_{y=0}^{y=+\infty} = \frac{1}{1-x^2}\left[\log\left(\frac{1}{x^2}\right) - \log 1\right] = \frac{2\log x}{x^2-1}$$

y, en consecuencia,

$$\int_0^{+\infty}\frac{\log x}{x^2-1}\,dx = \frac{\pi^2}{4}\ .$$

Como

$$\int_0^{+\infty}\frac{\log x}{x^2-1}\,dx = \int_0^1\frac{\log x}{x^2-1}\,dx + \int_1^{+\infty}\frac{\log x}{x^2-1}\,dx = \left(x = \frac{1}{t}\right) =$$

$$= \int_0^1\frac{\log x}{x^2-1}\,dx + \int_1^0\frac{\log t}{1-t^2}\,dt = 2\int_0^1\frac{\log x}{x^2-1}\,dx\ ,$$

concluimos que

$$\int_0^1\frac{\log x}{x^2-1} = \frac{\pi^2}{8}\ .$$

Llamamos ahora

$$A = \int_0^1\frac{\log x}{x-1}\,dx, \quad B = \int_0^1\frac{\log x}{x+1}\,dx\ .$$

Obsérvese que no hay problema de integrabilidad, ya que en el punto 1 tienen límite y en el punto 0, al ser

$$\lim_{x\to 0}\frac{\frac{\log x}{x-1}}{\log x} = -1 \neq 0, \qquad \lim_{x\to 0}\frac{\frac{\log x}{x+1}}{\log x} = 1 \neq 0\ ,$$

basta usar el criterio de comparación para comprobar que ambas funciones son integrables, ya que la función log es integrable en $]0,1[$.

Se sigue pues

$$\left.\begin{array}{r} A+B=\displaystyle\int_0^1 \frac{2x\log x}{x^2-1}\,dx = \left(x^2=t\right) = \frac{A}{2} \\ A-B = 2\displaystyle\int_0^1 \frac{\log x}{x^2-1}\,dx = \frac{\pi^2}{4} \end{array}\right\} \quad \Rightarrow \quad A = \int_0^1 \frac{\log x}{x-1}\,dx = \frac{\pi^2}{6} \quad .$$

Por último, como

$$0<x<1 \Rightarrow \frac{1}{1-x} = \sum_{n=0}^{\infty} x^n \ ,$$

se tiene que

$$\int_0^1 \frac{\log x}{x-1}\,dx = \int_0^1 \Big(\sum_{n=0}^{\infty}(-\log x)x^n\Big)\,dx \ ,$$

y como para $n \in \mathbb{N}\cup\{0\}$ se verifica la igualdad

$$\int_0^1 -(\log x)x^n\,dx = \left[\begin{array}{l} u=-\log x,\ u'=\frac{-1}{x} \\ v'=x^n,\ v=\frac{x^{n+1}}{n+1} \end{array}\right] = \int_0^1 \frac{x^n}{n+1}\,dx = \frac{1}{(n+1)^2}$$

y la serie $\sum_{n\geq 1}\frac{1}{n^2}$ es convergente, en virtud del teorema de la convergencia absoluta (Teorema 12.8) tenemos que

$$\int_0^1 \frac{\log x}{x-1}\,dx = \int_0^1 \left(\sum_{n=0}^{\infty}(-\log x)x^n\right)\,dx \ =$$

$$= \sum_{n=0}^{\infty}\left(\int_0^1 (-\log x)x^n\,dx\right) = \sum_{n=1}^{\infty}\frac{1}{n^2} \ .$$

Hemos probado, por tanto que

$$\sum_{n=1}^{\infty}\frac{1}{n^2} = \frac{\pi^2}{6} \ .$$

14.10 **a)** F es medible, por ser continua, y para cada n natural es

$$\text{(14.7.1)} \qquad \int_0^{+\infty} |F(x,y)|dy = \left[-\frac{|\operatorname{sen} x|}{x}\,e^{-xy}\right]_{y=0}^{y=+\infty} = \frac{|\operatorname{sen} x|}{x} \in \mathscr{L}^1(]1,n[),$$

pues esta función es continua y tiene límite en 0 y en n. El criterio de integrabilidad (Corolario 14.7) nos usegura que F es integrable en $]0,n[\times]0,+\infty[$. El teorema de Fubini nos dice ahora que

$$\int_0^n \left[\int_0^{+\infty} F(x,y)dy\right]dx = \int_0^{+\infty}\left[\int_0^n F(x,y)dx\right]dy.$$

Procediendo como en la igualdad 14.7.1 calculamos la primera integral del primer miembro, obteniendo que

$$\int_0^n \frac{\operatorname{sen} x}{x}\,dx = \int_0^{+\infty}\left[\int_0^n e^{-xy}\ \operatorname{sen} x\,dx\right]dy.$$

b) El teorema de Fubini asegura que cada f_n está definida c.p.d. (de hecho esta definida en todo punto) y que es integrable. Integrando por partes dos veces se tiene para cada n natural que

$$\begin{gathered} f_n(y) = \Big[-e^{-xy}\ \cos x\Big]_{x=0}^{x=n} - y\int_0^n e^{-xy}\ \cos x\,dx = \\ 1-e^{-ny}\ \cos n - y\int_0^n e^{-xy}\ \cos x\,dx = \\ 1-e^{ny}\ \cos n - y\left(\Big[e^{-xy}\ \operatorname{sen} x\Big]_{x=0}^{x=n} + y\int_0^b e^{-xy}\ \operatorname{sen} x\,ax\right) = \\ 1-e^{ny}\ \cos n - y\left(e^{-ny}\ \operatorname{sen} n + y f_n(y)\right), \end{gathered}$$

luego

$$f_n(y) = \frac{1-e^{-ny}\ (\cos n + y\ \operatorname{sen} n)}{1+y^2},\ \forall n\in\mathbb{N},$$

y en consecuencia

$$\{f_n(y)\} \longrightarrow \frac{1}{1+y^2}\ .$$

Además, para cada natural n se verifica también que

$$\left|\frac{\cos n + y\operatorname{sen} n}{e^{ny}}\right| \le \frac{1+y}{e^y} \le 1,$$

de donde se sigue para cada natural n que

$$|f_n(y)| \le \frac{2}{1+y^2} \in \mathscr{L}^1(]0,+\infty[).$$

c) Como el límite existe, basta calcular

$$\lim\int_0^n \frac{\operatorname{sen} x}{x}\,dx\ .$$

El apartado **b)** nos permite aplicar el teorema de la convergencia dominada (Teorema 12.5) a la sucesión $\{f_n\}$ para obtener, teniendo en cuenta el apartado **a)**, que

$$\lim\int_0^n \frac{\operatorname{sen} x}{x}\,dx = \lim\int_0^{+\infty} f_n(y)\,dy = \int_0^{+\infty}\frac{dy}{1+y^2} = \Big[\operatorname{arc\,tg} y\Big]_{y=0}^{y=+\infty} = \frac{\pi}{2}\ .$$

Hemos probado que

$$\lim_{t\to+\infty}\int_0^t \frac{\operatorname{sen} x}{x}\,dx = \frac{\pi}{2}.$$

14.11 i)

$$\int_E \frac{x}{\sqrt{1+x^2+y^2}}\, d(x,y) \text{ con } E = \{(x,y) \in \mathbb{R}^2 :\ y^2 \leq\ 2x,\ 0 \leq x \leq 2\}.$$

La función anterior es continua en $\mathbb{R}^2$, por tanto, es integrable sobre compactos y E es compacto. Usando el Teorema de Fubini, tenemos

$$\int_E \frac{x}{\sqrt{1+x^2+y^2}}\, d(x,y) = \int_{-2}^{2}\Big[\int_{\frac{y^2}{2}}^{2} \frac{x}{\sqrt{1+x^2+y^2}}\, dx\Big]\, dy =$$

$$D = \int_{-2}^{2}\Big[\sqrt{1+x^2+y^2}\Big]_{x=\frac{y^2}{2}}^{x=2}\, dy = \int_{-2}^{2}\sqrt{5+y^2}\, dy - \int_{-2}^{2}(\frac{y^2}{2}+1)\, dy =$$

$$= \Big(6+\frac{5}{2}\log 5\Big) - \frac{20}{3} = \frac{-2}{3} + \frac{5}{2}\log 5\ .$$

ii)

$$\int_E \sqrt{xy}\, d(x,y) \text{ con } E = \left\{(x,y) \in \mathbb{R}^2 :\ 0 \leq y, \left(\frac{2x+y}{4}\right)^2 \leq \frac{x}{6}\right\}.$$

De la segunda condición se deduce que $x \geq 0$ y entonces ha de verificarse que $\dfrac{x}{2}+\dfrac{y}{4} \leq \sqrt{\frac{x}{6}}$, es decir, $2x+y \leq \frac{4}{\sqrt{6}}\sqrt{x}$; como además, $y \geq 0$, entonces E es la región de $\mathbb{R}^2$ del primer cuadrante limitada por la función $f : \mathbb{R}^+ \longrightarrow \mathbb{R}$ dada por

$$f(x) = \frac{4}{\sqrt{6}}\sqrt{x} - 2x,$$

Como además, $\left(\frac{x}{2}\right)^2 \leq \frac{x}{6} \Rightarrow x \leq \frac{2}{3}$, entonces

$$\int_E \sqrt{xy}\, d(x,y) = \int_0^{2/3}\Big[\int_0^{4\sqrt{\frac{x}{6}}-2x}\sqrt{xy}\, dy\Big]\, dx =$$

$$= \int_0^{2/3} \frac{3}{2}\sqrt{xy}^{\frac{3}{2}}\Big]_{y=0}^{y=\frac{4}{\sqrt{6}}\sqrt{x}-2x} dx = \int_0^{2/3}\frac{3}{2}\sqrt{x}\Big(\frac{4}{\sqrt{6}}\sqrt{x}-2x\Big)^{\frac{3}{2}}\, dx = \ldots$$

iii)

$$\int_E \frac{2xy(2-3x)}{x^2+2y^2}\, d(x,y) \text{ con } E = \{(x,y) \in \mathbb{R}^2 :\ 0 < x,\ 0 < y,\ y^2 \leq \frac{1}{2} - x\}.$$

$$\int_E \frac{2xy(2-3x)}{x^2+2y^2}\, d(x,y) = \int_0^{1/2}\Big[\int_0^{\sqrt{1/2-x}} \frac{2xy(2-3x)}{x^2+2y^2}\, dy\Big]\, dx =$$

$$= \int_0^{1/2}\Big[\frac{x(2-3x)}{2}\log(x^2+2y^2)\Big]_{y=0}^{y=\sqrt{1/2-x}}\, dx =$$

$$= \int_0^{1/2}\frac{x(2-3x)}{2}\log\frac{x^2-2x+1}{x^2}\, dx = \ldots$$

iv)
$$\int_E (x+y)\, d(x,y) \ \text{ con } \ E = \{(x,y) \in \mathbb{R}^2 : \ y^2 \leq x+2, \ x^2 \leq y+2\}.$$

El dominio es la región comprendida en el interior de dos parábolas

$$\int_E (x+y)\, d(x,y) = \int_{\frac{-1-\sqrt{5}}{2}}^{-1} \left[\int_{x^2-2}^{\sqrt{x+2}} (xy+y)\, dy\right] dx +$$
$$+\int_{-1}^{\frac{-1+\sqrt{5}}{2}} \left[\int_{-\sqrt{x+2}}^{\sqrt{x+2}} (x+y)\, dy\right] dx + \int_{\frac{-1+\sqrt{5}}{2}}^{2} \left[\int_{x^2-2}^{\sqrt{x+2}} (x+y)\, dy\right] dx = \cdots$$

v)
$$\int_E y^2\, d(x,y) \ \text{ con } \ E = \{(x,y) \in \mathbb{R}^2 : \ x^2+y^2 \leq R^2\}.$$
$$\int_E y^2\, d(x,y) = \int_{-\pi}^{\pi} \left[\int_0^R \rho^3 \operatorname{sen}^2\theta\, d\rho\right] d\theta = \frac{R^4}{4}\int_{-\pi}^{\pi} \operatorname{sen}^2\theta\, d\theta = \frac{\pi R^4}{4}\ .$$

vi)
$$\int_E \frac{x^2+y^2}{\sqrt{1+x^2+y^2}}\, d(x,y) \ \text{ con } \ E = \left\{(x,y) \in \mathbb{R}^2 : \ x^2+y^2 < 1, \ 0 < y\right\}.$$

E es medible con medida finita y $f\chi_E$ está acotada por 1, luego $f\chi_E$ es integrable.

$$\int_E \frac{x^2+y^2}{\sqrt{1+x^2+y^2}}\, d(x,y) = \int_0^{\pi} \left[\int_0^1 \frac{\rho^3}{\sqrt{1+\rho^2}}\, d\rho\right] d\theta = \begin{cases} u = \rho^2 \\ v' = \rho(1+\rho^2)^{\frac{1}{2}} \end{cases}$$
$$\int_0^{\pi} \left[\rho^2(1+\rho^2)\frac{1}{2} - \frac{2}{3}(1+\rho^2)^{\frac{3}{2}}\right]_{\rho=0}^{\rho=1} = \frac{(2-\sqrt{2})\pi}{3}\ .$$

vii)
$$\int_E \left(\frac{x^2}{a^2}+\frac{y^2}{b^2}\right) d(x,y) \ \text{ con } \ E = \left\{(x,y) \in \mathbb{R}^2 : \ \frac{x^2}{a^2}+\frac{y^2}{b^2} < 1\right\}.$$

E es medible y tiene medida finta, por ser compacto y además $f\chi_E$ está da por 1, por tanto $f\chi_E$ es integrable. Haciendo el cambio de variable $\begin{cases} x = a\rho\cos\theta \\ y = b\rho\operatorname{sen}\theta \end{cases}$, se tiene que el valor absoluto del determinante del jacobiano del cambio de variable es $ab\rho$ (estamos suponiendo que $a, b > 0$). La función que da el cambio de variable es

$$\Phi : \mathbb{R}^+ \times \left]-\pi, \pi\right[\longrightarrow \mathbb{R}^2$$

la dada por

$$\Phi(\rho,\theta) = (a\rho\cos\theta, b\rho\operatorname{sen}\theta) \qquad (\rho > 0, \theta \in]-\pi,\pi[)$$

es de clase infinito, inyectiva y su imagen es el conjunto $\mathbb{R}^2 \backslash \{(x,0) : x \le 0\}$. Así el Teorema global de la función inversa nos asegura que Φ es un difeomorfismo de clase infinito y podemos aplicar el teorema del cambio de variable (Corolario 8.8), obteniéndose

$$\int_E \left(\frac{x^2}{a^2}+\frac{y^2}{b^2}\right)\, d(x,y) = \int_{-\pi}^{\pi}\left[\int_0^1 ab\rho^3\, d\rho\right]\, d\theta = \frac{ab\pi}{2}\ .$$

viii)

$$\int_E \frac{d(x,y)}{(1+x^2+y^2)^{\frac{3}{2}}}\ \text{ con }\ E=\left\{(x,y)\in\mathbb{R}^2 :\ 0<x<1,\ 0<y<\frac{1}{\sqrt{3}}\right\}.$$

El recinto de integración es un rectángulo de lados 1 y $\dfrac{1}{\sqrt{3}}$. En general, si

$$\alpha = \arctan\left(\frac{b}{a}\right),$$

en el cambio a coordenadas polares de tiene:

$$0<\theta<\alpha \Rightarrow 0<\rho<\frac{a}{\cos\theta};\qquad \alpha<\theta<\frac{\pi}{2}\Rightarrow 0<\rho<\frac{b}{\operatorname{sen}\theta}\ .$$

En nuestro caso $\alpha = \arctan\dfrac{1}{\sqrt{3}} = \dfrac{\pi}{6}$

$$I := \int_E \frac{1}{(1+x^2+y^2)^{\frac{3}{2}}}\, d(x,y) =$$

$$= \int_0^{\frac{\pi}{6}}\left[\int_0^{\frac{1}{\cos\theta}} \frac{\rho}{(1+\rho^2)^{\frac{3}{2}}}\right]\, d\theta + \int_{\frac{\pi}{6}}^{\frac{\pi}{2}}\left[\int_0^{\frac{1}{\sqrt{3}\cos\theta}} \frac{\rho}{(1+\rho^2)^{\frac{3}{2}}}\right]\, d\theta =$$

$$= \left(\frac{\pi}{6}-\arctan\frac{\sqrt{7}}{7}\right)+\left(\arctan\frac{\sqrt{7}}{3}-\frac{\pi}{6}\right)=$$

$$= \arctan\frac{\sqrt{7}}{3}-\arctan\frac{\sqrt{7}}{7}\ .$$

ix)

$$\int_E \frac{d(x,y,z)}{(1+x+y+z)^3}$$

con

$$E=\left\{(x,y,z)\in\mathbb{R}^3 :\ 0\le x,\ 0\le y,\ 0\le z,\ x+y+z\le 1\right\}.$$

$$\int_E \frac{1}{(1+x+y+z)^3}\, d(x,y,z) = \int_0^1\left[\int_{\{(x,y)\in\mathbb{R}^2 : x+y\le -z\}} \frac{1}{(1+x+y+z)^3}\, d(x,y)\right]\, dz =$$

$$\int_0^1\left[\int_0^{1-z}\left[\int_0^{1-y-z} frac1(1+x+y+z)^3\, dx\right]\, dy\right]\, dz =$$

$$\int_0^1 \left[\int_0^{1-z} \left[\frac{-1}{2(1+x+y+z)^2} \right]_{x=0}^{x=1} dy \right] dz = \cdots$$

x)

$$\int_E (x^2+y^2+z^2)\, d(x,y,z)$$

con

$$E = \left\{ (x,y,z) \in \mathbb{R}^3 : 0 \le x,\ 0 \le y,\ 0 \le z,\ \frac{x}{a} + \frac{y}{b} + \frac{x}{c} \le 1 \right\}\ (a,b,c>0).$$

$$\int_E (x^2+y^2+z^2)\, d(x,y,z) =$$

$$\int_0^c \left[\int_{\{(x,y)\in\mathbb{R}^2:\ 0\le x,\ 0\le y,\ \frac{x}{a}+\frac{y}{b}\le 1-\frac{z}{c}\}} (x^2+y^2+z^2)\, d(x,y) \right] dz =$$

$$\int_0^c \left[\int_0^{b\left(1-\frac{z}{c}\right)} \left[\int_0^{a\left(1-\frac{z}{c}-\frac{y}{b}\right)} (x^2+y^2+z^2)\, dx \right] dy \right] dz = \cdots$$

Capítulo IV

Referencias

Bibliografía

[ApPa] C. Aparicio y R. Payá, *Análisis Matemático I*. Servicio de Publicaciones de la Universidad de Granada, 1999 (Séptima edición). 1.24, 1.4

[BaGa] R. M. Barbolla, M. García y otros, *Introducción al Análisis Real*. Alhambra, Madrid, 1981. 1.24, 1.4

[Ber] Berberian, S. K., *Fundamentals of Real Analysis*, Springer, Nueva York, 1998. 2.7, 10.10, 11.8

[BRV] Bombal, F., Rodríguez, L y Vera, G., *Problemas de Análisis Matemático (2. Cálculo diferencial)*, Editorial AC, Madrid, 1988 3.10

[Bra] Brabenec, R. L., *Introduction to Real Analysis*, Thomson Information Publishing Group, Boston, 1990. 2.7

[Bri] Bridges, D. S., *Foundations of Real and Abstract Analysis*, Graduate Texts in Mathematics, Springer, Nueva York, 1998. 2.7

[Cra] Craven, B.D., *Functions of several variables*, Chapman and Hall, Nueva York, 1981. 2.7, 3.10, 8.4, 9.6

[Fe] Fernández Viña, J.A., *Análisis Matemático II, Topología y Cálculo Diferencial*, Tecnos, Madrid, 1984. 2.7, 3.10

[FeSa] Fernández Viña, J.A.y Sánchez Mañes, E., *Ejercicios y complementos de Análisis Matemático II*, Tecnos, Madrid, 1986. 3.10

[Gau] E. Gaughan, *Introdución al Análisis*. Alhambra, Madrid, 1972. 1.24, 1.4

[Guz] M. Guzmán y B. Rubio, *Integración: Teoría y técnicas*, Alhambra, Madrid, 1979. 4.3, 4.4, 10.10, 11.8

[HoYo] J.G. Hocking y G.S. Young, *Topología*. Reverté, Barcelona, 1966. 4.2, 4.4

[Jur] Jürgen, J., *Postmodern Analysis*, Universitext, Springer, Berlin, 1998. 2.7, 3.10, 8.4, 10.10, 11.8

[Lin] E. Linés, *Principios de Análisis Matemático*. Reverté, Barcelona, 1983. 1.24, 1.4

[LuMa] J. Lukes y J. Maly, *Measure and integral*. Matfyzpress, Praga, 1995. 10.10

[MaHo] Marsden, J.E. y Hoffman, M.J., *Análisis Clásico Elemental*, Segunda Edición, Addison-Wesley Iberoamericana, Argentina, 1998. 2.7, 3.10, 4.4, 5.7, 6.3, 7.3, 8.4

[McBo] E. J. McShane y T. A. Botts, *Análisis Real*. Aguilar, Madrid, 1971. 1.24, 1.4

[RaWe] L. Rade y B. Westergren, *Mathematics Handbook for Science and Engineering*. Studentlitteratur, Lund (Suecia), 1998 (Cuarta edición). 10.8, 10.10

[Rey] J. Rey Pastor, *Elementos de Análisis Algebraico*. Herederos de J. Rey Pastor, Madrid, 1966. 1.24, 1.4

[Ru] W. Rudin, *Real and complex Analysis*. McGraw-Hill, 1966. 10.10, 11.8

[SoSi] P. N. de Souza y J-N Silva, *Berkeley Problems in Mathematics*. Springer-Verlag, New York, 1999. 1.4, 1.6, 2.7

[Stra] G. Strang, *Linear algebra and its applications*, Academic Press. 7.3

[Stro] Stromberg, K.E., *An introduction to classical real analysis*, Wadsworth International Group, Belmont, California, 1981. 2.7, 7.2

[Thi] S. Thio de Pol, *Primos o algunos dislates sobre números*. Alhambra, Madrid, 1977. 1.2, 1.4

Índice alfabético